



ÁREA 7:

MÉTODOS CUANTITATIVOS



Universidad de Valladolid
Departamento de Economía Aplicada

ESTUDIO DE LOS CAMBIOS EN LOS HÁBITOS DE CONSUMO ESPAÑOL A NIVEL REGIONAL MEDIANTE EL NUEVO SISTEMA CON EL QUE SE ELABORA EL ÍNDICE DE PRECIOS DE CONSUMO

SANTIAGO RODRIGUEZ FEIJOÓ

e-mail: srfeijoo@dmc.ulpgc.es

CARLOS GONZÁLEZ CORREA

e-mail: cgoncor@gmail.com

ALEJANDRO RODRÍGUEZ CARO

e-mail: arcaro@dmc.ulpgc.es

Departamento de MÉTODOS CUANTITATIVOS EN ECONOMÍA Y GESTIÓN
UNIVERSIDAD DE LAS PALMAS DE GRAN CANARIA

Área temática: Métodos Cuantitativos

Resumen

El Instituto Nacional de Estadística español introdujo en el año 2001 una nueva fórmula para calcular el Índice de Precios de Consumo (IPC), la Fórmula de Laspeyres Encadenado. Esta modificación responde a la necesidad de ajustarse al Reglamento (CE) 2454/1997 de 10 de diciembre de 1997 donde se concretan las normas mínimas para las ponderaciones presentes en los índices de precios. El nuevo IPC, a partir de cualquier nivel de agregación, se calcula usando los índices elementales junto con unas ponderaciones que se actualizan anualmente, a diferencia del IPC que se elaboraba anteriormente, cuyas ponderaciones permanecían fijas mientras no variase la base del índice.

En el trabajo se usan los índices de precios, las ponderaciones y el gasto regional per cápita para calcular un índice de cantidades anual para el período 2001-2006. A partir de los índices de precios y cantidades anuales se calculan sus correlaciones y distintas elasticidades medias, con el objeto de caracterizar los hábitos de consumo de la población española a nivel autonómico, así como sus cambios.

Palabras clave: Hábitos de consumo, IPC, Índice de cantidades.

Abstract

The Spanish National Statistics Institute introduced, in 2001, a new formula for calculate the Prices Consumer Index (CPI), the Laspeyres linked. This modification responds to the necessity to adjust to Commission Regulation (EC) 2454/97 of 10 December 1997. The new CPI is calculated using the elementary prices index and the weight, that are updated annually.

In this work, the consumer prices index, the weights and the regional expenditure are used to calculate an annual quantities index for period 2001-2006. We use the consumer price and the quantities index to estimate your correlations and several types of elasticity, with the intention of characterizing the habits of consumer of the Spanish population at autonomic level, as well as its changes.

Key words: Habits of Consumer, CPI, Quantities Index.

1. Introducción.

El encadenamiento de los índices ha llevado a considerar la variación de las ponderaciones, en la elaboración del IPC, como uno de los elementos fundamentales para que representen la relación real entre los gastos de las familias y los cambios de precios.

Es por ello que siguiendo las directrices de Eurostat [UE, (1997)] y debido a la continua disponibilidad de datos sobre gastos provenientes de la Encuesta Continua de Presupuestos Familiares (ECPF), una de las modificaciones más importantes que ha realizado el Instituto Nacional de Estadística [INE, (2002)], a partir del cambio de sistema mediante el cual se elabora el IPC en el año 2001, es la actualización continua de sus ponderaciones.

En este trabajo se utilizan estas ponderaciones, junto con información procedente de la ECPF y el propio IPC, con el objeto de estudiar los cambios en los hábitos de consumo de la población española. Para ello, en el epígrafe 2 se desarrolla el marco teórico que incorpora la información de las ponderaciones junto con el gasto que se obtiene de la ECPF, con el objeto de estimar un índice de cantidades de periodicidad anual. Este índice, junto con el índice de precios, IPC, nos permite calcular distintas elasticidades que, combinándolas con la correlación entre la evolución de los precios y de las cantidades, identifican distintos mercados, que reflejan como están evolucionando los hábitos de consumo de una determinada población.

En el epígrafe tercero, se aplica parte del esquema teórico desarrollado en el punto 2 al caso español, desagregando a nivel de rúbricas, un total de 57, y a nivel de las comunidades autónomas. Además, únicamente para el conjunto del Estado español, se analizan las elasticidades para el conjunto de 57 rúbricas. En el último epígrafe se enumeran las principales conclusiones obtenidas.

2. Determinación de los cambios en los hábitos de Consumo. Planteamiento teórico

El cálculo del IPC se realiza combinando los índices de precios elementales con unas ponderaciones que reflejan la importancia en el consumo total del gasto en el bien o servicio considerado. Tradicionalmente las ponderaciones que se utilizaban eran únicas

durante todos los periodos para los cuales se mantenía la base fija. A partir de las críticas realizadas por diferentes teóricos sobre los sesgos producidos en el IPC como consecuencia de la permanencia durante un periodo de tiempo relativamente largo de las citadas ponderaciones, especialmente las contenidas en el Informe Boskin, algunas agencias estadísticas aceptan la recomendación de encadenar los índices. En este sentido, Eurostat ha decidido que los Índices de Precios de Consumo se elaboren mediante esta técnica. Ésta consiste básicamente en referir los precios del periodo corriente a los precios de diciembre del año anterior. El encadenamiento obliga a actualizar anualmente las ponderaciones, que se calculan a partir de las proporciones de gasto correspondientes al año anterior, con lo que se consigue obtener una mayor concordancia entre los diversos gastos de las familias y el índice de cambio de precio elaborado con las fórmulas actualmente utilizadas y debidas a Laspeyres.

La ponderación en cada periodo de tiempo t del artículo i es W_t^i y se calculan como:

$$W_t^i = \frac{\text{gasto realizado en el artículo } i \text{ en el año anterior}}{\text{gasto total realizado en el año anterior}} = \frac{G_t^i}{G_t^T} \quad (1)$$

Los cambios de esta ponderación señalan la evolución del gasto en el bien o servicio i y, por tanto, los cambios que se producen en los hábitos de consumo de las familias, en gran medida debidos a los cambios de precios. En conjunto, señalan la sustitución que se ha producido en el consumo de cada uno de los bienes como consecuencia de la evolución de todos los precios de los bienes y servicios que componen la cesta de la compra.

Una forma de analizar esta evolución es a través de la razón de estas ponderaciones pertenecientes a los dos periodos que se utilizan para el cálculo de estos índices:

$$\frac{W_1^i}{W_0^i} = \frac{\frac{G_1^i}{G_1^T}}{\frac{G_0^i}{G_0^T}} = \frac{G_1^i \times G_0^T}{G_0^i \times G_1^T} = \frac{p_1^i \times q_1^i \times G_0^T}{p_0^i \times q_0^i \times G_1^T} \quad (2)$$

Donde p_t^i y q_t^i representan el precio y la cantidad correspondiente al bien o servicio i en el instante $t = \{0,1\}$.

Por otra parte, sabemos que los índices de precios y cantidades del bien o servicio i se obtienen como:

$$I_p^i = \frac{p_1^i}{p_0^i}; \quad I_q^i = \frac{q_1^i}{q_0^i} \quad (3)$$

Combinando estas expresiones con la expresada en (2) se obtiene la expresión del índice de cantidades reflejado en (4):

$$I_q^i = \frac{G_1^T}{G_0^T} \times \frac{W_1^i}{W_0^i} \times \frac{1}{I_p^i} \quad (4)$$

A partir de estos índices se puede establecer una relación entre los precios y las cantidades consumidas, que permita analizar el comportamiento de los consumidores ante las variaciones de precios.

Sin embargo, hay que tener en cuenta que los índices presentes en (4) tienen características distintas. Mientras que el índice de precios I_p^i es un cociente de variables stock, el índice de cantidades definido I_q^i es un cociente de variables flujo, en la medida que sus componentes se corresponden con las cantidades consumidas en el periodo, que no necesariamente se producen en una sola transacción. Además, dado que normalmente la información que sustenta tanto las cantidades totales gastadas como su proporción se obtiene a partir de encuestas dirigidas a la población residente en el ámbito territorial del índice de precios, ambas cantidades están determinadas por la población existente en el periodo que se considere, con lo que también lo está el índice de cantidades. Por tanto, para el análisis conjunto de ambos índices, el de cantidades debe corregirse en función de la población.

Si N_t es la población en el año t , el índice de cantidades corregido se puede definir tal como se muestra en (5)

$$I_q^{hi} = \frac{\frac{q_1^i}{N_1}}{\frac{q_0^i}{N_0}} = \frac{N_0}{N_1} \times I_q^i \Rightarrow I_q^i = \frac{N_1}{N_0} \times I_q^{hi} \quad (5)$$

Con lo que la expresión (4) se transforma en

$$I_q^i = \frac{N_1}{N_0} \times I_q^{hi} = \frac{G_1^T}{G_0^T} \times \frac{W_1^i}{W_0^i} \times \frac{1}{I_p^i} \quad (6)$$

De donde es inmediato obtener (7)

$$I_q^{hi} = \frac{N_0}{N_1} \times \frac{G_1^T}{G_0^T} \times \frac{W_1^i}{W_0^i} \times \frac{1}{I_p^i} = \frac{\frac{G_1^T}{N_1}}{\frac{G_0^T}{N_0}} \times \frac{W_1^i}{W_0^i} \times \frac{1}{I_p^i} \quad (7)$$

Con lo que se consigue establecer una relación entre el índice de precios I_p^i y el índice de cantidades homogeneizado en términos de población I_q^{hi} a través de las ponderaciones y del gasto per cápita de las familias. Su expresión genérica será:

$${}^{t/0}I_q^{hi} = \frac{N_0}{N_t} \times \frac{G_t^T}{G_0^T} \times \frac{W_t^i}{W_0^i} \times \frac{1}{{}^{t/0}I_p^i} \quad (8)$$

donde:

${}^{t/0}I_q^{hi}$ es un índice de cantidades homogeneizado en términos de población del bien o servicio i entre el periodo base y el corriente.

${}^{t/0}I_p^i$ es el índice de precios del bien o servicio i entre el periodo base y el corriente.

N_t , G_t^T y W_t^i son la población total, el gasto total realizado y la participación del gasto en el bien o servicio i en cada uno de los periodos. Además, $\frac{G_t^T}{N_t}$ no es más que el gasto per cápita de las familias destinadas al consumo en cada periodo.

Para bienes y servicios normales, los crecimientos de los precios y las cantidades deben tener signo contrario, por lo que sus índices deben tener una correlación entre ellos con signo negativo. Cuando la correlación es positiva, es decir, ante crecimiento de los precios se produce un incremento de las cantidades, señala que el poder de “resistencia” que tienen los consumidores para oponerse a las subidas de precios es menor. Asimismo, si algún territorio se distancia del comportamiento general, significa un menor grado de “monopolio” de los oferentes en ese mercado territorial.

Por el contrario, cuanto mayor sea el valor negativo de la correlación, más posibilidades tienen los consumidores de reaccionar ante las subidas de precios. Esta reacción depende de la variación del índice de precios, puesto que, si éste crece, la reacción se concreta en variar el consumo hacia bienes y servicios sustitutivos o en renunciar a su consumo. Si el índice de precios decrece, la reacción que refleja la correlación negativa es un incremento del índice de cantidades.

Ahora bien, en cada instante de tiempo la cesta de productos es fija y el consumidor reacciona más en términos relativos de precios entre productos que con el precio del propio producto. Es decir, responde a su percepción de si un precio sube mucho o poco, en virtud de la media de crecimiento de todos los precios, el IPC general. Por tanto, consideran que un bien o servicio ha crecido mucho o poco si su crecimiento es mayor o menor que el citado índice general.

Considerando conjuntamente la relación entre los índices de precios y cantidades per cápita y el crecimiento relativo de los precios se obtiene la tipología de mercado que se muestra en el cuadro 1

Cuadro 1. Tipología de los mercados

	Índices de precios de un bien o servicio con crecimiento superior al IPC general	Índices de precios de un bien o servicio con crecimiento inferior al IPC general o con caída de precios
Correlación entre índices de precios y cantidades positiva	M1: Los consumidores no pueden evitar consumir este bien o servicio. Predominio de los oferentes en el mercado.	M2: Los bienes y servicios representados tienen alguna característica, que puede ser la moderación de los precios, que los hace deseables para el consumidor.
Correlación entre índices de precios y cantidades negativa	M3: Los consumidores, ante crecimientos de precios, pueden evitar consumir este bien o servicio, bien sea reduciendo el consumo o sustituyéndolo por otro.	M4: Los bienes y servicios representados son en alguna forma indeseables para los consumidores, bien sea porque se están sustituyendo por otros o porque han dejado de representar los hábitos de consumo.

Para completar el estudio de los cambios en los hábitos de consumo es necesario cuantificar de alguna manera el grado de sustitución que se produce entre bienes y servicios como consecuencia de las variaciones producidas por los cambios en los precios. Para ello, en este trabajo se utiliza la elasticidad media definida por Allen (1934). Esta se define como se muestra en (9)

$$\varepsilon_{Q^i}^{P^i} = \frac{\Delta q^i}{\left(\frac{q_1^i + q_t^i}{2}\right)} \times \frac{\left(\frac{p_1^i + p_t^i}{2}\right)}{\Delta p^i} = \frac{(q_t^i - q_1^i)}{(q_1^i + q_t^i)} \times \frac{(p_1^i + p_t^i)}{(p_t^i - p_1^i)} \quad (9)$$

El hecho de que las cantidades sean unas magnitudes flujo, obliga a trabajar en términos per cápita. Ello nos lleva a definir una elasticidad media per cápita de la forma que se muestra en (10).

$$\varepsilon_{q^i}^{P^i} = \frac{\left(\frac{q_t^i}{N_t} - \frac{q_1^i}{N_1}\right)}{\left(\frac{q_t^i}{N_t} + \frac{q_1^i}{N_1}\right)} \times \frac{(p_1^i + p_t^i)}{(p_t^i - p_1^i)} \quad (10)$$

Multiplicando y dividiendo (10) por $\frac{q_1^i}{N_1} \times p_1^i$ se tiene (11)

$$\varepsilon_{q^i}^{P^i} = \frac{\left(\frac{q_t^i}{N_t} - \frac{q_1^i}{N_1}\right)}{\frac{q_1^i}{N_1}} \times \frac{(p_1^i + p_t^i)}{\frac{p_1^i}{p_t^i - p_1^i}} \quad (11)$$

Sustituyendo (7) en (11) y reordenando términos se obtiene (12)

$$\varepsilon_{Q^i}^{P^i} = \frac{(I_q^{hi} - 1)}{(I_q^{hi} + 1)} \times \frac{(I_p^i + 1)}{(I_p^i - 1)} = \frac{1 - I_q^{hi} \times I_p^i + (I_p^i - I_q^{hi})}{1 - I_q^{hi} \times I_p^i - (I_p^i - I_q^{hi})} \quad (12)$$

Es de señalar que estas elasticidades se refieren a la variación que se produce entre los precios y cantidades consumidas, y se obtienen a partir de los índices de precios y cantidades per cápita. El valor de estas elasticidades nos señala la reacción de los consumidores ante variaciones de los precios. Si es negativa y los precios han crecido, los consumidores han disminuido el consumo, mientras que si los precios han disminuido los consumidores han aumentado el consumo. Si es positiva señala que a pesar de todo, los consumidores lo han aumentado en el primer caso y lo han disminuido en el segundo. Estas variaciones también se pueden distinguir entre los que son mayores o menores que +1 o -1, señalando un crecimiento/decrecimiento más o menos que proporcional al cambio de precio producido.

También es posible obtener la elasticidad cruzada en la medida que en (9) se sustituya P_t^i por P_t^j . Por tanto, la elasticidad cruzada se obtiene a partir de la siguiente expresión:

$$\mathcal{E}_{Q^j}^{P^j} = \frac{1 - I_q^{hi} \times I_p^j + (I_p^j - I_q^{hi})}{1 - I_q^{hi} \times I_p^j - (I_p^j - I_q^{hi})} \quad (13)$$

Estas elasticidades cruzadas reflejan los cambios en las cantidades de cada uno de los bienes y servicios como consecuencia de cada cambio de precio del resto de los bienes y servicios presentes en la cesta de la compra. Por lo que, en conjunto reflejan la sustitución que se produce en las cantidades consumidas ante las variaciones de precios de todos y cada uno de los bienes y servicios presentes en la cesta de la compra.

Este conjunto de elasticidades cruzadas se puede analizar desde una doble perspectiva. Por un lado, los cambios, en promedio, de las cantidades consumidas en un bien o servicio k como consecuencia de los cambios de precios del resto de los bienes y servicios (\mathcal{E}_1^k). Por otro, los que, en promedio, se producen en las cantidades del resto de los bienes presentes en el conjunto de los bienes y servicios como consecuencia del cambio del precio de un bien o servicio k (\mathcal{E}_2^k). Es decir, podemos definir los siguientes estadísticos:

$$\mathcal{E}_1^k = \frac{1}{n-1} \times \left(\sum_{\substack{j=1 \\ j \neq k}}^n \mathcal{E}_{Q^j}^{P^j} \right) \quad (14)$$

$$\varepsilon_2^k = \frac{1}{n-1} \times \left(\sum_{\substack{j=1 \\ j \neq k}}^n \varepsilon_{Q^j}^{P^k} \right) \quad (15)$$

Para este análisis conjunto, debemos tener en cuenta que estas elasticidades cruzadas, independiente de la fórmula de obtención, esencialmente representan un cociente entre variaciones relativas de cantidades y de precios, es decir:

$$\varepsilon_{Q^j}^{P^i} = \frac{\frac{\Delta q_i}{q_i}}{\frac{\Delta p_j}{p_j}} \quad (16)$$

Si el cambio relativo en precios fuese el mismo para todos los bienes y servicios, la expresión (14) junto con la (16) la podríamos escribir como (17)

$$\varepsilon_1^k = \frac{1}{n-1} \times \beta \times \sum_{\substack{i=1 \\ i \neq k}}^n \left(\frac{\Delta q_k}{q_k} \right)^i \quad (17)$$

Donde $\left(\frac{\Delta q_k}{q_k} \right)^i$ representa el cambio relativo de las cantidades del producto k debido al cambio de precios del producto i . Si además hacemos $\beta = n-1$, (17) se puede expresar como (18).

$$\begin{aligned} \varepsilon_1^k &= \left[\left(\frac{q_k^1}{q_k^0} \right)^1 + \left(\frac{q_k^1}{q_k^0} \right)^2 + \dots + \left(\frac{q_k^1}{q_k^0} \right)^{k-1} + \left(\frac{q_k^1}{q_k^0} \right)^{k+1} + \dots + \left(\frac{q_k^1}{q_k^0} \right)^n - (n-1) \right] = \\ &= \frac{1}{q_k^0} \times \left[\sum_{\substack{i=1 \\ i \neq k}}^n (q_k^1)^i - (n-1) \times q_k^0 \right] \quad (18) \end{aligned}$$

La expresión (18) es fácilmente interpretable. Si $\varepsilon_1^k = 0$, ello implica que, en términos medios, el incremento de precios relativos generalizado e igual en los (n-1) productos restantes de la cesta de la compra no tiene efectos sobre las cantidades consumidas de la

rúbrica k. Si $\varepsilon_1^k > 0$, implicaría que la respuesta ante un incremento generalizado de los precios relativos en el resto de productos de la cesta de la compra produce un incremento en las cantidades consumidas en el producto k. Es decir, en términos medios, el producto k es sustitutivo con respecto al resto de productos que componen la cesta de la compra. Por el contrario, si $\varepsilon_1^k < 0$, ello implicaría cierto carácter complementario, en el sentido de que el incremento de precios generalizado reduciría el consumo del productos k.

El estadístico ε_2^k lo podemos escribir como (19), indicando $\left(\frac{\Delta q_j}{q_j}\right)^k$ el cambio relativo que muestra la cantidad consumida en el producto j al incrementarse el precio relativo del producto k.

$$\varepsilon_2^k = \frac{1}{n-1} \frac{p_k}{\Delta p_k} \left[\sum_{\substack{j=1 \\ j \neq k}}^n \left(\frac{\Delta q_j}{q_j}\right)^k \right] \quad (19)$$

De forma similar a como se hizo con ε_1^k , ε_2^k se puede escribir en función de los índices de cantidades.

$$\varepsilon_2^k = \frac{1}{n-1} \frac{p_k}{\Delta p_k} \left[\sum_{\substack{j=1 \\ j \neq k}}^n \left(\frac{q_j^1}{q_j^0}\right)^k - (n-1) \right] \quad (20)$$

En (20), $\left(\frac{q_j^1}{q_j^0}\right)^k$ denota el índice de cantidades del producto j en el instante temporal 1 con respecto al 0, debido al cambio relativo en el precio del producto k. Si $\varepsilon_2^k = 0$, implica que un incremento relativo en el precio del producto k no produce cambios, en términos medios, y sobre los índices, en las cantidades consumidas en el resto de los productos presentes en la cesta de la compra. Si $\varepsilon_2^k > 0$, el encarecimiento del producto k, provoca un mayor consumo medio del resto de productos que forman la cesta de la compra, mientras que, si $\varepsilon_2^k < 0$, lo que se produce es una reducción. Es decir, de alguna

manera, ε_2^k vuelve a medir el carácter sustitutivo-complementario del producto k frente al resto de los productos. Este carácter se determina por la elasticidad precio del producto k en la medida que es la que señala el crecimiento/decrecimiento de su consumo. Así, si $\varepsilon_{Q^i}^{P^i} > 0$, lo que señala que ante crecimientos de precios aumenta la cantidad consumida, y $\varepsilon_2^k > 0$, lo que refleja que el consumo del resto de productos también aumenta, expresa su carácter complementario. Si $\varepsilon_2^k < 0$ será un producto sustitutivo. Por el contrario, si $\varepsilon_{Q^i}^{P^i} < 0$ la caracterización es la contraria.

Por tanto, el estudio de los cambios en los hábitos de consumo no puede realizarse únicamente con ε_1^k y ε_2^k , y obviar ni la elasticidad precio directa, ni la estructura de mercados que se define en el cuadro 1. Por ello, para alcanzar el objetivo perseguido, es necesario analizar conjuntamente todas las medidas descritas en las páginas previas. Para ello partiremos de la estructura de mercados definidos en el Cuadro 1.

Los mercados M1 y M2 contienen a los productos que presentan una correlación positiva entre el índice de precios y el de cantidades, lo que conlleva una elasticidad precio también positiva. Además, el mercado M1 muestra un cambio fuerte en los precios. Por tanto, son bienes muy atractivos o necesarios para el consumidor. En este sentido, se esperaría que si el resto de los productos se hacen más caros, el consumidor reaccionaría consumiendo más productos que se sitúan en este mercado. Es decir, ε_1^k se espera que sea mayor que cero. Ahora bien, si sus precios siguen subiendo y $\varepsilon_2^k > 0$ indicaría que el consumidor considera que ya no le es rentable seguir pagando dicho precio y prefiere pasar a consumir otros productos. Sin embargo, si $\varepsilon_2^k < 0$, ello significaría que el consumidor sigue sin poder renunciar a consumir productos que se encuentran en el mercado M1.

El mercado M2 lo forman productos cuyos precios y cantidades han evolucionado con el mismo signo, pero siguen siendo productos con un crecimiento de precios moderados con respecto al conjunto de todas las rúbricas. Son productos en los cuales la propia moderación de sus precios puede ser un acicate para su consumo. Para estos bienes se

esperaría que ε_1^k fuese mayor que cero, puesto que ello implicaría que los productos pertenecientes al mercado M2 se hacen aún más baratos. Nuevamente el valor de ε_2^k nos marcaría un punto de inflexión en el comportamiento del consumidor. Si $\varepsilon_2^k > 0$, un incremento del precio en el producto k produciría un incremento en el consumo en el resto de productos, lo cual confirmaría que el bajo precio relativo es el que incentiva el consumo de estos productos, de tal forma que si se hace relativamente más caro se pasa a consumir otro tipo de producto. Por el contrario, si $\varepsilon_2^k < 0$ indicaría que estos bienes son, en algún sentido, fundamentales para el consumidor.

Los productos pertenecientes a los mercados M3 y M4 se caracterizan por presentar una correlación negativa entre los índices de precios y cantidades y, por tanto, presentarán una elasticidad precio directa negativa. El mercado M3 se caracteriza, además, por contener productos con crecimiento de precios altos. Por ello, si ε_1^k fuese mayor que cero, implicaría que, al hacerse relativamente más baratos los productos del mercado M3 se consumirían más, mientras que si $\varepsilon_1^k < 0$, mostraría que el consumo de los productos del mercado M3 es complementario del consumo del resto de los productos de la cesta de la compra. Esta complementariedad es compatible con $\varepsilon_2^k < 0$. Ahora bien, si ε_1^k fuese menor que cero y $\varepsilon_2^k > 0$ estaría mostrando a la vez cierta complementariedad y cierta sustituibilidad. Es decir, serían bienes deseables para el consumidor porque, aunque tienen un precio muy elevado, se consumen en unas cantidades lo suficientemente elevadas como para que una unidad más de su consumo tenga un coste de oportunidad muy elevado. Esto hace que si el resto de productos incrementa su precio no se consuma más de estos productos y, sin embargo, si su precio crece se hace demasiado caro y se pasa a consumir otros productos.

El mercado M4 contiene productos que, además de presentar una elasticidad precio directa negativa, sus precios han crecido de forma moderada. Son bienes poco atractivos para el consumidor. Esta conclusión se vería apoyada si ε_1^k fuese menor que cero, ya que ello significaría que al incrementar el resto de precios, se reduce la cantidad consumida en este producto. En la misma línea, se esperaría que ε_2^k fuese mayor que

cero, dado que ello mostraría que la razón de consumir productos que se encuentran en M4 es su bajo precio. Ahora bien, si estos crecen, se pasa a consumir otros bienes.

3. Determinación de los cambios en los hábitos de Consumo. Aplicación al estado español y sus comunidades autónomas

Para aplicar la metodología propuesta en el epígrafe anterior, el primer paso consiste en estimar G^T correspondiente a la expresión (8). Desgraciadamente este dato no está disponible, pero lo podemos estimar siguiendo el siguiente razonamiento. Conociendo el gasto ECPF [INE (2006b)] en un grupo se puede conocer el gasto total de la cesta que usa el IPC si se conoce el peso del citado grupo en el gasto total de la citada cesta. En general, la composición de los grupos en que se divide el Gasto de la ECPF es similar a la de los del IPC, excepto aquellos en que los gastos de autoconsumo, autosuministro, de salario en especie, comidas gratuitas o bonificadas y alquiler imputado a la vivienda en la que reside el hogar (cuando es propietario de la misma o la tiene cedida gratuita o semigratuitamente por otros hogares o instituciones) son importantes. Los grupos donde estas cantidades son más significativas son el Grupo 1: Alimentación y Bebidas no alcohólicas, el Grupo 4: Vivienda, agua, electricidad y otros combustibles, el Grupo 9: Ocio, espectáculos y cultura y el Grupo 12: Otros bienes y servicios. Por tanto, a partir de los gastos ECPF del resto de los grupos se puede estimar el gasto total IPC. Obviamente, en cada cálculo intervienen gastos residuales no adaptables al IPC que evitan que la estimación de este gasto total en la cesta del IPC sea idéntica en todos los grupos. Dado que no tenemos ningún criterio por el cual decidir que cifra en la mejor, nuestra propuesta es usar como estimación el valor de la media geométrica de los resultados obtenidos en cada uno de ellos.

No obstante, no hay que olvidar que en el cálculo del gasto total de las familias también interviene el crecimiento de la población, por lo que se hace necesario relativizar dicho gasto en función de la evolución de la misma.

Por tanto, se puede considerar que este cálculo es válido, tanto para las rúbricas nacionales como para cada una de las comunidades ya que, aunque introduce un evidente sesgo, como influye por igual en los cálculos de todos los índices de cantidades de cada Comunidad, se convierte en un sesgo aceptable.

A partir de la fórmula reflejada en (6) y con los datos de precios y ponderaciones referidos al periodo 2001-2006 [INE, (2006a)], se han obtenido los índices de cantidades y su correlación con los índices de precios de las rúbricas nacionales y de todas las Comunidades, donde las G^T se han obtenido como se ha señalado anteriormente.

3.1 Análisis de los mercados en la España de las Autonomías 2001-2006

En el cuadro 2 se muestran, para España y para las comunidades autónomas, la asignación de las 57 rúbricas a cada uno de los cuatro mercados que se definen en el cuadro 1. En el cuadro 2, la primera fila se corresponde con los códigos de los distintos territorios, cuya identificación se muestra en el anexo 1, mientras que en la primera columna se encuentran los códigos de las distintas rúbricas, cuya identificación se puede consultar en el anexo 2.

Al analizar el cuadro 2, lo primero que destaca es la heterogeneidad en los hábitos de consumo de los españoles. Únicamente las rúbricas R13, aceites y grasas, la R47, comunicaciones, y la R48, objetos recreativos, tienen un comportamiento homogéneo para todo el conjunto de comunidades autónomas. Los aceites y las grasas, para el conjunto del estado español, se ha caracterizado, a lo largo del período 2001-2006, por presentar un fuerte incremento de precios, por encima del crecimiento medio de la cesta de la compra, y por reducir el consumo de este tipo de bienes.

Por el contrario, los bienes y servicios integrados en las rúbricas de comunicaciones y objetos recreativos han visto crecer sus precios de forma moderada y esa moderación en los precios ha favorecido en crecimiento de las cantidades consumidas. Hay también otras rúbricas que muestran un elevado nivel de homogeneidad a nivel del conjunto del estado español. Por ejemplo, la R32, reparación de calzado, que pertenece al mercado M3 en todos los territorios, con la excepción de la Rioja, que forma parte del mercado M1. Dentro de la heterogeneidad que muestra el cuadro 2, el mercado más homogéneo territorialmente es el M3. Éste, junto con el M2 son los mercados que representan, en principio y sin entrar en consideraciones de elasticidades cruzadas, una conducta del consumidor más acorde con lo esperado, en el sentido de que a mayor crecimiento de precios menos consumo y a menor crecimiento de precios más consumo.

Cuadro 2. Identificación de mercados en España y sus comunidades autónomas

	C00	C01	C02	C03	C04	C05	C06	C07	C08	C09	C10	C11	C12	C13	C14	C15	C16	C17	C18
R01	M2	M2	M2	M4	M2	M4	M2	M2	M2	M2	M2	M2	M4						
R02	M3	M3	M1	M3	M3	M1	M3	M3	M3	M3	M1	M3	M1	M1	M1	M3	M1	M1	M3
R03	M3	M3	M1	M3	M3	M4	M3	M3	M3	M3	M3	M3	M2	M3	M3	M3	M1	M1	M3
R04	M3	M3	M3	M3	M3	M1	M4	M3	M1	M3									
R05	M2	M2	M1	M4	M3	M2	M3	M4	M4	M3	M2	M4	M2	M2	M3	M4	M1	M2	M4
R06	M1	M3	M1	M3	M2	M1	M3	M3	M3	M3	M3	M4	M2	M1	M3	M2	M1	M2	M3
R07	M2	M2	M2	M4	M2	M2	M2	M2	M4	M4	M2	M4	M2	M2	M2	M2	M2	M2	M4
R08	M3	M3	M2	M3	M3	M2	M3	M4	M4	M3	M3	M4	M2	M1	M1	M3	M1	M1	M4
R09	M2	M2	M2	M4	M1	M2	M4	M1	M4	M4	M2	M4	M2	M1	M2	M2	M1	M2	M3
R10	M3	M3	M3	M3	M3	M1	M3	M3	M3	M3	M1	M3	M3	M3	M3	M3	M3	M1	M3
R11	M2	M2	M2	M4	M2	M2	M4	M1	M4	M4	M2	M4	M2	M1	M2	M1	M1	M2	M3
R12	M2	M4																	
R13	M3																		
R14	M3	M3	M1	M3	M1	M1	M3	M3	M3	M3	M1	M3	M3	M1	M3	M1	M1	M1	M3
R15	M3	M3	M3	M3	M3	M1	M3	M1	M1	M4									
R16	M1	M1	M2	M1	M1	M1	M1	M1	M3	M1	M1	M3	M1	M2	M1	M1	M1	M2	M3
R17	M2	M2	M2	M4	M2	M1	M2	M2	M4	M2	M2	M4	M2	M2	M2	M2	M2	M2	M4
R18	M3	M3	M3	M3	M1	M1	M3												
R19	M4	M4	M4	M3	M4	M2	M4	M2	M4	M4	M1	M4							
R20	M2	M2	M2	M2	M2	M2	M4	M4	M2	M4	M4	M2							
R21	M2	M2	M1	M3	M1	M1	M3	M1	M2	M2	M1	M3	M2	M1	M2	M2	M1	M2	M4
R22	M2	M2	M2	M3	M1	M1	M2	M2	M3	M2	M2	M4	M2						
R23	M2	M2	M2	M4	M2	M2	M4	M2	M4	M4	M2	M4	M2	M2	M2	M2	M2	M2	M4
R24	M3	M1	M1	M4															
R25	M4	M4	M2	M4	M4	M2	M4	M4	M4	M4	M3	M4	M2	M2	M2	M4	M2	M2	M4
R26	M2	M2	M2	M4	M3	M2	M4	M2	M4	M4	M2	M4	M2	M2	M1	M4	M2	M2	M2
R27	M2	M2	M2	M4	M2	M2	M3	M2	M4	M4	M1	M4	M4	M2	M2	M1	M2	M1	M2
R28	M4	M4	M2	M3	M4	M2	M3	M4	M4	M4	M3	M4	M2	M2	M2	M4	M2	M2	M3
R29	M2	M4	M2	M4	M2	M2	M4	M4	M4	M4	M2	M4	M2	M2	M2	M4	M2	M2	M2
R30	M2	M2	M1	M3	M2	M2	M2	M1	M4	M3	M2	M4	M1	M2	M1	M3	M2	M2	M2
R31	M2	M2	M2	M4	M2	M2	M4	M2	M4	M3	M4	M3	M2	M2	M2	M4	M2	M1	M2
R32	M3	M1	M3																
R33	M1	M1	M1	M3	M1	M1	M3	M1	M3	M3	M1	M3	M1	M1	M1	M1	M1	M1	M3
R34	M3	M3	M3	M3	M4	M2	M4	M3	M3	M4	M3	M3	M3	M2	M1	M3	M2	M1	M4
R35	M3	M1	M1	M4															
R36	M4	M4	M3	M4	M4	M2	M4	M3	M4	M3	M3	M4	M3	M3	M4	M4	M4	M1	M4
R37	M2	M2	M2	M4	M2	M2	M2	M2	M4	M4	M2	M4	M2						
R38	M2	M2	M2	M4	M2	M4	M2												
R39	M4	M2	M2	M4	M2	M2	M3	M4	M4	M4	M4	M4	M2	M2	M2	M4	M2	M2	M2
R40	M2	M2	M2	M4	M2	M2	M4	M2	M4	M4	M2	M4	M2						
R41	M1	M1	M1	M3	M1	M3	M1	M1	M1	M1	M1	M1	M2						
R42	M1	M1	M1	M4	M1	M1	M3	M1	M4	M3	M1	M3	M2	M1	M1	M1	M1	M1	M4
R43	M2	M2	M2	M4	M2	M2	M2	M4	M2	M2	M2	M4	M4	M2	M2	M2	M2	M2	M2
R44	M3	M3	M1	M3	M3	M1	M3	M4	M1	M1	M3								
R45	M3	M3	M3	M3	M2	M1	M3	M1	M3	M3	M3	M1	M3						
R46	M3	M3	M1	M3	M3	M1	M3	M3	M3	M3	M3	M3	M1	M1	M1	M3	M1	M1	M3
R47	M2																		
R48	M2																		
R49	M2	M2	M1	M4	M2	M2	M3	M2	M4	M4	M2	M4	M2	M2	M2	M2	M1	M2	M2
R50	M2	M2	M2	M2	M4	M2	M2	M2	M4	M3	M1	M2	M2	M2	M2	M1	M1	M2	M4
R51	M3	M3	M1	M3	M3	M1	M4	M3	M3	M3	M3	M3	M1	M1	M2	M3	M1	M1	M3
R52	M3	M3	M3	M3	M3	M1	M3	M3	M4	M3	M3	M3	M3	M1	M3	M3	M1	M1	M3
R53	M3	M3	M3	M3	M3	M1	M3	M1	M3										
R54	M3	M1	M1	M4	M1	M1	M3	M1	M4	M3	M1	M3	M2	M1	M1	M1	M1	M1	M2
R55	M2	M2	M2	M2	M1	M2	M2	M2	M2	M1	M2	M2	M2	M2	M1	M1	M2	M1	M2
R56	M1	M3	M1	M3	M1	M1	M3	M3	M3	M3	M3	M3	M1	M1	M1	M1	M1	M1	M3
R57	M1	M1	M1	M4	M1	M1	M3	M1	M3	M3	M1	M3	M1	M1	M1	M2	M1	M1	M3

Esto justifica la elaboración del Cuadro 3, que persigue simplificar la descripción general de los hábitos de consumo a nivel territorial. En este nuevo cuadro, por filas se identifican las comunidades autónomas y por columnas se recogen para distintos mercados y agrupaciones de los mismos la siguiente información:

- N°, es el número de rúbricas que se integran en cada mercado o agrupación de los mismos
- %N, es el porcentaje (sobre 57) de rúbricas que se integran en cada mercado o agrupación de los mismos
- %G, es la suma de las ponderaciones en porcentaje de las rúbricas que se integran en cada uno de los mercados o agrupación de los mismos

Por ejemplo, para el conjunto del mercado español (C00), 25 rúbricas integran el mercado M1, lo cual supone el 43.9% de las rúbricas y el 34.2% del gasto, que por término medio, realiza un español.

Como se observa en el cuadro 3, la cesta de la compra de España se divide entre los mercados de forma desigual. Así, el 72,5 % del gasto se realiza en rúbricas en las que los agentes se comportan como resulta habitual, es decir, ante crecimientos medios menores aumentan su consumo (M2) y ante crecimientos mayores disminuyen el consumo (M3), y que, además, significan el 78,9% de las rúbricas.

Estos gastos se reparten de forma simétrica entre las que tienen un incremento menor de los precios (M2), el 43,9 % de las rúbricas y el 34,2 % del gasto y las que lo tienen con incrementos mayores de precios (M3), el 35,1 % de las rúbricas y el 38,4 % de los gastos. Estos datos parecen señalar que, en el primer caso, las rúbricas tienen poco peso (el promedio de las ponderaciones de las 25 rúbricas es 1,37 %) y, en el segundo, lo tienen mayor (el promedio de las ponderaciones de las 20 rúbricas es 1,92 %, aunque este promedio está muy influido por la presencia de la R44, Transporte personal, con un peso del 15,81 %, con lo que el promedio, sin considerar esta rúbrica se reduce a 1,19%).

Por su parte, el gasto producido en mercados donde los agentes no actúan de la forma esperada es bastante diferente, puesto que mientras que entre las que, a pesar de subir los precios aumenta el consumo (M1) son el 12,3% de las rúbricas y que significan el 21,8 % del gasto, en las que, a pesar de disminuir los precios disminuyen las cantidades (M4) son el 8,8 % de las rúbricas y solo el 5,7 % del gasto. Estos datos parecen señalar que en el primer caso las rúbricas tienen mucho peso (el promedio de las ponderaciones de las 7 rúbricas es 3,11 %) y en el segundo lo tienen menor (el promedio de las ponderaciones de las 5 rúbricas es 1,14 %).

Cuadro 3. Descripción del consumo en España y sus comunidades autónomas

	M2			M3			M2+M3			M1			M4			M1+M4		
	Nº	% N	% G	Nº	% N	% G	Nº	% N	% G	Nº	% N	% G	Nº	% N	% G	Nº	% N	% G
C00	25	43,9	34,2	20	35,1	38,4	45	78,9	72,5	7	12,3	21,8	5	8,8	5,7	12	21,1	27,5
C01	25	43,9	35,7	21	36,8	50,7	46	80,7	86,3	6	10,5	7,6	5	8,8	6,0	11	19,3	13,7
C02	26	45,6	37,8	13	22,8	18,0	39	68,4	55,7	17	29,8	43,9	1	1,8	0,4	18	31,6	44,3
C03	6	10,5	10,7	29	50,9	58,8	35	61,4	69,5	0	0,0	0,0	22	38,6	30,5	22	38,6	30,5
C04	21	36,8	24,3	17	29,8	35,4	38	66,7	59,7	13	22,8	30,0	6	10,5	10,3	19	33,3	40,3
C05	28	49,1	43,8	4	7,0	6,0	32	56,1	49,8	23	40,4	49,2	2	3,5	1,0	25	43,9	50,2
C06	13	22,8	19,9	28	49,1	55,1	41	71,9	75,1	2	3,5	1,6	14	24,6	23,3	16	28,1	24,9
C07	18	31,6	26,4	21	36,8	51,3	39	68,4	77,7	10	17,5	12,8	8	14,0	9,5	18	31,6	22,3
C08	8	14,0	13,5	23	40,4	54,6	31	54,4	68,1	1	1,8	0,9	25	43,9	31,0	26	45,6	31,9
C09	9	15,8	13,6	29	50,9	58,7	38	66,7	72,3	3	5,3	4,7	16	28,1	23,0	19	33,3	27,7
C10	21	36,8	31,0	21	36,8	51,9	42	73,7	82,9	12	21,1	16,1	3	5,3	1,1	15	26,3	17,1
C11	5	8,8	8,5	27	47,4	55,1	32	56,1	63,7	0	0,0	0,0	25	43,9	36,3	25	43,9	36,3
C12	30	52,6	42,9	15	26,3	33,5	45	78,9	76,5	9	15,8	20,6	3	5,3	2,9	12	21,1	23,5
C13	27	47,4	36,0	12	21,1	27,8	39	68,4	63,8	17	29,8	35,9	1	1,8	0,3	18	31,6	36,2
C14	26	45,6	31,8	16	28,1	31,3	42	73,7	63,1	14	24,6	34,8	1	1,8	2,1	15	26,3	36,9
C15	18	31,6	22,0	18	31,6	20,5	36	63,2	42,5	11	19,3	26,8	10	17,5	30,7	21	36,8	57,5
C16	23	40,4	33,8	8	14,0	7,1	31	54,4	40,9	24	42,1	56,4	2	3,5	2,7	26	45,6	59,1
C17	27	47,4	34,5	2	3,5	1,1	29	50,9	35,6	28	49,1	64,4	0	0,0	0,0	28	49,1	64,4
C18	18	31,6	26,4	22	38,6	45,0	40	70,2	71,3	0	0,0	0,0	17	29,8	28,7	17	29,8	28,7

Cuando se realiza este análisis en las Comunidades autónomas se observa, por un lado, que existen cuatro comunidades en que la división ente mercados lleva a que algunas no tienen rúbricas en alguno de los mercados. Así, en Aragón, Extremadura y Ceuta y Melilla, todas las rúbricas de los mercados que no se comportan de forma esperada se agrupan en el M4 sin que ninguna de ellas se comporte como el M1. Por su parte, en la Rioja, estos mercados se agrupan en el M1, sin que existan gastos en el M4. Por otro lado, en cuatro comunidades, Canarias, Navarra, País Vasco y la Rioja, el gasto que se

produce en rúbricas donde los agentes se comportan de forma esperada es inferior al 50 %. Además, en todos estos casos el porcentaje de rúbricas supera el 50%. También es de destacar que en estas últimas, excepto Navarra, existe una evidente desproporción entre las rúbricas que se comportan como del M2 respecto a las que se comportan del M3.

3.2 Los hábitos de consumo en España. Un análisis en función de las elasticidades medias

Las elasticidades complementan el análisis de la evolución de los diferentes mercados que intervienen en la cesta de la compra, señalando tanto la evolución del consumo como sus características.

Mercado M1: Las rúbricas que se encuadran en este mercado se caracterizan por el alto crecimiento de sus precios en relación al cambio general y que, a pesar de todo, se han incrementado sus consumos.

Cuadro 4. Rúbricas españolas incluidas en el Mercado 1

		Prom. IPC	Corr.	$\varepsilon_{q^i}^{p^i}$	ε_1^k	ε_2^k	Pond /100
	General	1,065	0,29				100
R06	Carne de ave	1,094	0,08	-0,06	-0,08	0,10	0,91
R16	Legumbres y hortalizas frescas	1,085	0,63	0,55	0,56	0,11	1,01
R33	Vivienda en alquiler	1,088	0,44	0,16	0,20	0,09	2,24
R41	Servicios para el hogar	1,098	0,74	0,72	1,02	0,07	1,39
R42	Servicios médicos y similares	1,078	0,52	0,28	0,33	0,10	1,90
R56	Turismo y hostelería	1,085	0,02	0,02	0,03	0,10	12,44
R57	Bienes y servicios no contemplados en otra parte	1,074	0,63	0,48	0,57	0,09	1,89
	PROMEDIO de los valores absolutos			0,40	0,10	0,32	3,11

Como se observa en el Cuadro 4 todas las rúbricas, excepto la R06: Carne de ave, tienen las elasticidades precios positivas señalando su carácter de necesarios o muy deseados por el consumidor. Asimismo, las ε_1^k son positivas, señalando su carácter de productos sustitutivos del resto de las rúbricas, en el sentido señalado de que aumentan

su consumo si el resto de los precios crece en la medida que se abaratan relativamente estas rúbricas. El hecho de que ε_2^k también sean positivos refleja un cierto carácter de complementariedad en el sentido señalado de que si sus precios siguen creciendo, el incremento del coste de oportunidad los llevaría a consumir otros productos. Estas características se pueden aplicar también a la rúbrica R06 en la medida que su elasticidad precio y ε_1^k son negativas y que ε_2^k sea positivo lo que señala que es un producto sustitutivo de otras rúbricas que no pertenecen a este mercado y que se elige como alternativa a estos productos por la poca importancia que tiene su consumo. Esto nos lleva a concluir que el mercado M1 lo forman bienes y servicios muy relacionados con la calidad de vida de los consumidores, y éstos estarían dispuestos a consumir más si su crecimiento de precios se moderase.

Mercado M2: Las rúbricas que se encuadran en este mercado se caracterizan porque el crecimiento de sus precios, en relación al cambio general, es menor y que, en consecuencia, se han incrementado sus consumos.

Hay que señalar que en este mercado se han incluido las rúbricas que, al descender los precios se han incrementado las cantidades, como lo reflejan tanto el signo de la correlación de los índices de precios y cantidades como la elasticidad precio. En general, excepto la rúbrica R26: Prendas de vestir de mujer, el hecho de que el signo de la elasticidad precio y que el promedio de sus valores absolutos sea muy alto, los identifican como bienes deseables para los consumidores. Asimismo, las ε_1^k son positivas, lo que indica el carácter sustitutivo del resto de las rúbricas, en el sentido señalado de que aumentan su consumo si el resto de los precios crece en la medida que se abaratan relativamente estas rúbricas.

El hecho de que ε_2^k también sean positivos confirma que son los precios bajos los que determinan el deseo de consumirlos. Por tanto, si sus precios suben, el incremento del coste de oportunidad los llevaría a consumir otros productos.

Cuadro 5. Rúbricas españolas incluidas en el Mercado 2

		Prom. IPC	Corr.	$\varepsilon_{q^j}^{p^j}$	ε_1^k	ε_2^k	Pond. /100
	General	1,065	0,29				100
R01	Cereales y derivados	1,038	0,92	1,60	0,91	0,19	1,55
R05	Carne de porcino	1,060	0,09	0,25	0,25	0,09	0,79
R07	Otras carnes	1,038	0,59	0,92	0,56	0,18	2,31
R09	Crustáceos, moluscos y preparados de pescado	1,056	0,34	0,30	0,26	0,13	1,43
R11	Leche	1,056	0,36	0,32	0,29	0,13	1,31
R12	Productos lácteos	1,037	0,82	2,07	1,00	0,21	1,62
R17	Preparados de legumbres y hortalizas	1,034	0,58	1,36	0,67	0,22	0,50
R20	Azúcar	1,002	0,34	12,58	0,75	1,39	0,13
R21	Otros preparados alimenticios	1,064	0,92	0,92	0,78	0,13	0,95
R22	Agua mineral, refrescos y zumos	1,056	0,90	0,87	0,63	0,15	0,83
R23	Bebidas alcohólicas	1,034	0,46	0,42	0,23	0,22	0,89
R26	Prendas de vestir de mujer	1,050	0,18	-0,17	-0,10	0,20	3,54
R27	Prendas de vestir de niño y bebé	1,045	0,60	0,66	0,34	0,22	1,08
R29	Calzado de Hombre	1,047	0,18	0,06	0,04	0,18	0,72
R30	Calzado de mujer	1,062	0,56	0,27	0,22	0,14	0,88
R31	Calzado de niño	1,045	0,57	0,23	0,13	0,21	0,36
R37	Textil y accesorios para el hogar	1,042	0,71	0,86	0,57	0,17	0,60
R38	Electrodomésticos y reparaciones	0,984	-0,98	-4,24	0,98	-0,51	1,06
R40	Artículos no duraderos para el hogar	1,015	0,35	2,61	0,62	0,44	1,74
R43	Medicamentos y material terapéutico	0,994	-0,90	-4,87	0,88	-0,69	1,55
R47	Comunicaciones	0,965	-0,97	-6,76	3,69	-0,12	2,91
R48	Objetos recreativos	0,912	-0,99	-1,67	2,48	-0,06	2,36
R49	Publicaciones	1,053	0,90	0,49	0,32	0,18	1,19
R50	Esparcimiento	1,060	0,69	0,71	0,61	0,12	1,52
R55	Artículos de uso personal	1,059	0,78	1,49	1,35	0,11	2,34
	PROMEDIO de los valores absolutos			0,75	0,26	1,87	1,37

Mercado M3: Las rúbricas que se encuadran en este mercado se caracterizan porque el crecimiento de sus precios en relación al cambio general es mayor y que, en consecuencia se han disminuido sus consumos.

Cuadro 6. Rúbricas españolas incluidas en el Mercado 3

		Prom. IPC	Corr.	$\varepsilon_{q^i}^{p^i}$	ε_1^k	ε_2^k	Pond /100.
	General	1,065	0,29				100
R02	Pan	1,118	-0,35	-0,08	-0,13	0,08	1,91
R03	Carne de vacuno	1,084	-0,37	-0,40	-0,59	0,09	1,22
R04	Carne de ovino	1,115	-0,87	-0,82	-1,33	0,09	0,44
R08	Pescado fresco y congelado	1,077	-0,07	-0,06	-0,07	0,10	1,74
R10	Huevos	1,134	-0,99	-0,40	-0,58	0,09	0,24
R13	Aceites y grasas	1,230	-0,94	-0,61	-2,28	0,04	0,79
R14	Frutas frescas	1,148	-0,32	-0,04	-0,06	0,08	1,55
R15	Fretas en conserva y frutos secos	1,181	-0,78	-0,24	-0,62	0,05	0,29
R18	Patatas y sus preparados	1,119	-0,55	-0,36	-0,74	0,07	0,35
R24	Tabaco	1,098	-0,85	-0,42	-0,57	0,10	2,28
R32	Reparación de calzado	1,118	-0,94	-0,85	-1,48	0,09	0,02
R34	Calefacción, alumbrado y distribución de agua	1,071	-0,56	-0,25	-0,35	0,09	4,94
R35	Conservación de la vivienda	1,093	-0,97	-0,76	-1,08	0,10	3,58
R44	Transporte personal	1,075	-0,85	-0,27	-0,33	0,11	15,81
R45	Transporte público urbano	1,101	-0,67	-0,26	-0,38	0,09	0,58
R46	Transporte público interurbano	1,109	-0,48	-0,07	-0,12	0,07	0,53
R51	Educación infantil y primaria	1,099	-0,66	-0,18	-0,26	0,09	0,46
R52	Educación secundaria	1,092	-0,77	-0,37	-0,49	0,10	0,46
R53	Educación universitaria	1,104	-0,89	-0,50	-0,76	0,09	0,62
R54	Otras enseñanzas	1,076	-0,07	-0,11	-0,12	0,11	0,57
	PROMEDIO de los valores absolutos			0,62	0,09	0,35	1,92

En general, el signo de la elasticidad precio negativo señala su carácter de bienes sustituibles para el consumidor. Asimismo, las ε_1^k son negativas, señalando su carácter de productos complementarios del resto de las rúbricas, en el sentido señalado de que

aumentan su consumo si el resto de los precios crece en la medida que se abaratan relativamente estas rúbricas.

El hecho de que ε_2^k también sean positivos refleja un cierto carácter de complementariedad en el sentido señalado de que si sus precios siguen creciendo, el incremento del coste de oportunidad los llevaría a consumir otros productos.

Mercado M4: Las rúbricas que se encuadran en este mercado se caracterizan porque el crecimiento de sus precios en relación al cambio general es menor y que, a pesar de todo, han disminuido sus consumos.

Cuadro 7. Rúbricas españolas incluidas en el Mercado 4

		Prom. IPC	Corr.	$\varepsilon_{q^j}^{p^j}$	ε_1^k	ε_2^k	Pond /100.
	General	1,065	0,29				100
R19	Café, cacao e infusiones	1,045	-0,64	-0,58	-0,48	0,16	0,40
R25	Prendas de vestir de hombres	1,048	-0,27	-0,34	-0,21	0,20	2,83
R28	Complementos y reparación de prendas de vestir	1,054	-0,53	-0,36	-0,28	0,16	0,28
R36	Muebles y revestimientos de suelo	1,064	-0,91	-0,74	-0,74	0,14	1,83
R39	Utensilios y herramientas para el hogar	1,044	-0,37	-0,18	-0,12	0,18	0,37
	PROMEDIO de los valores absolutos			0,37	0,17	0,44	1,14

En general, el signo de la elasticidad precio negativo señala su carácter de bienes normales para el consumidor. Además, dado que han sufrido un cambio moderado en sus precios, indica el carácter poco necesario de estos bienes y servicios. Asimismo, las ε_1^k son negativas, señalando su carácter de productos complementarios del resto de las rúbricas, en el sentido señalado de que aumentan su consumo si el resto de los precios crece en la medida que relativamente se abaratan mucho más estas rúbricas.

El hecho de que ε_2^k sean positivos confirma el carácter de complementariedad en el sentido señalado de que si sus precios siguen creciendo, el incremento del coste de

oportunidad los llevaría a consumir otros productos. Este mercado aglutina los productos menos deseados para los consumidores.

4.- Conclusiones

En este trabajo se diseña un método de análisis que permite estudiar los hábitos de consumo y su evolución temporal para un determinado territorio. Dicho método se basa en la información procedente de la elaboración de un IPC encadenado, lo cual supone que las ponderaciones se van modificando periódicamente. Los cambios en las ponderaciones nos muestran los cambios en los hábitos de consumo y nos permiten estimar unos índices de cantidades percapita que, junto con el IPC anual, son usados en este trabajo como base para el cálculo de distintas medidas que pretenden describir el comportamiento del consumidor en el mercado.

Esta metodología se aplicó al mercado español en su conjunto y por comunidades autónomas, permitiendo identificar cuatro mercados. El denominado M1 contiene bienes muy deseables para los consumidores, pero cuyos precios han crecido por encima de la media, lo que hace que el consumidor sólo esté dispuesto a incrementar su consumo percapita si sus precios se hacen relativamente más baratos. Son bienes y servicios que reflejan calidad de vida. El mercado M2 lo constituyen un conjunto de productos para los cuales la propia moderación de precios se identifica como el elemento fundamental que determina su incremento de consumo. Son productos importantes pero no imprescindibles para el consumidor. El mercado M3 lo forman productos para los cuales el consumidor está dispuesto a renunciar a parte de su consumo si sus precios siguen subiendo. Por último, el mercado M4 lo forman los bienes y servicios poco atractivos para el consumidor. Son bienes obsoletos.

Desde el punto de vista territorial, la conclusión más relevante es la gran heterogeneidad detectada en los hábitos de consumo a nivel de comunidades autónomas. Son muy pocas las rúbricas que se sitúan, para todas las comunidades autónomas, siempre en el mismo mercado.

A nivel español destaca la importancia que se le da a la rúbrica R56: Turismo y hostelería que se sitúa en el mercado M1, a pesar de que las elasticidades entre precios y cantidades y la correlación entre los índices de precios y cantidades tienen un valor muy

próximo a cero. Las legumbres y hortalizas frescas, la vivienda en alquiler, los servicios para el hogar y los servicios médicos también forman parte del mercado M1. El mercado M2 lo forman productos alimenticios, vestido, calzado, elementos para el hogar, comunicaciones y ocio. En el mercado M3 se sitúan, además algunas rúbricas de alimentación (carnes, pescados, aceites, huevos), el transporte y la educación junto con los gastos relacionados con el uso de la vivienda (calefacción, alumbrado, agua, conservación de la vivienda). Por último, el mercado M4 sólo contiene 5 rúbricas que en su conjunto únicamente representan el 5.7% del gasto total medio de un español.

Bibliografía

Allen, R.G.D. (1934): "The Concept of Arc Elasticity of Demand", *The Review of Economic Studies*, vol. 1, nº 3, pp. 226-230.

Boskin, M. et al. (1996): "Toward a more accurate measure of cost of living", Final Report to the Senate Finance Committee.

INE, (2002): IPC. Documento Metodológico, Documento Electrónico, http://www.ine.es/ipc01/metodologia_ipc2001.pdf.

INE, (2006a): **Nivel, Calidad y Condiciones de Vida. Índices de Precios de Consumo**, <http://www.ine.es/inebase>. Página electrónica. Datos de IPC: precios y ponderaciones.

INE, (2006b): **Nivel, Calidad y Condiciones de Vida. Encuesta Continua de Presupuestos Familiares**, <http://www.ine.es/inebase>. Página electrónica. Datos de gastos de los hogares por grupos y comunidades autónomas.

UE, (1997): Reglamento (CE) N° 2454/1997 del Consejo de 10 de diciembre de 1997 Relativo a los Índices Armonizados de Precios de Consumo.

Anexo 1. Códigos territoriales

Territorio	España	Andalucía	Aragón	Asturias	Baleares	Canarias
Código	C00	C01	C02	C03	C04	C05
Territorio	Cantabria	Castilla León	Castilla Mancha	Cataluña	C Valenciana	Extremadura
Código	C06	C07	C08	C09	C10	C11
Territorio	Galicia	Madrid	Murcia	Navarra	País Vasco	Rioja
Código	C12	C13	C14	C15	C16	C17
Territorio	Ceuta Melilla					
Código	C18					

Anexo 2. Códigos de rúbricas

Rúbrica	Código	Rúbrica	Código
Cereales y derivados	R01	Calzado de mujer	R30
Pan	R02	Calzado de niño	R31
Carne de vacuno	R03	Reparación de calzado	R32
Carne de ovino	R04	Vivienda en alquiler	R33
Carne de porcino	R05	Calefacción, alumbrado y distribución de agua	R34
Carne de ave	R06	Conservación de la vivienda	R35
Otras carnes	R07	Muebles y revestimientos de suelo	R36
Pescado fresco y congelado	R08	Textil y accesorios para el hogar	R37
Crustáceos, moluscos y preparados de pescado	R09	Electrodomésticos y reparaciones	R38
Huevos	R10	Utensilios y herramientas para el hogar	R39
Leche	R11	Artículos no duraderos para el hogar	R40
Productos lácteos	R12	Servicios para el hogar	R41
Aceites y grasas	R13	Servicios médicos y similares	R42
Frutas frescas	R14	Medicamentos y material terapéutico	R43
Fretas en conserva y frutos secos	R15	Transporte personal	R44
Legumbres y hortalizas frescas	R16	Transporte público urbano	R45
Preparados de legumbres y hortalizas	R17	Transporte público interurbano	R46
Patatas y sus preparados	R18	Comunicaciones	R47
Café, cacao e infusiones	R19	Objetos recreativos	R48
Azúcar	R20	Publicaciones	R49
Otros preparados alimenticios	R21	Esparcimiento	R50
Agua mineral, refrescos y zumos	R22	Educación infantil y primaria	R51
Bebidas alcohólicas	R23	Educación secundaria	R52
Tabaco	R24	Educación universitaria	R53
Prendas de vestir de hombres	R25	Otras enseñanzas	R54
Prendas de vestir de mujer	R26	Artículos de uso personal	R55
Prendas de vestir de niño y bebé	R27	Turismo y hostelería	R56
Complementos y reparación de prendas de vestir	R28	Bienes y servicios no contemplados en otra parte	R57
Calzado de Hombre	R29		

UN ESQUEMA DE ELABORACIÓN DEL IPC MÁS CERCANO AL ICV

SANTIAGO RODRÍGUEZ FEIJOÓ
e-mail:srfeijoo@dmc.ulpgc.es

CARLOS GONZÁLEZ CORREA
e-mail:cgoncor@gmail.com

ALEJANDRO RODRÍGUEZ CARO
e-mail:arcaro@dmc.ulpgc.es

Departamento de MÉTODOS CUANTITATIVOS EN ECONOMÍA Y GESTIÓN
UNIVERSIDAD DE LAS PALMAS DE GRAN CANARIA

Área temática: Métodos Cuantitativos.

Resumen

En el presente trabajo se hace una propuesta para el cálculo de un índice del Coste de la Vida (ICV) a partir de la estructura de trabajo y de la información estadística con la que se calcula el Índice de Precios de Consumo (IPC). Para ello, en primer lugar, se analizan los objetivos para los cuales las agencias de estadística (Bureau of Labor Statistic, Eurostat, etc...) calculan el IPC. En segundo lugar, y, dado que el objetivo no es único y se suele relacionar con el nivel de vida, se estudian estos dos conceptos, llegando a la conclusión de que cuando el objetivo del IPC es calcular el Índice del Coste de la Vida, el nivel elemental del cálculo del IPC se convierte en la principal fuente de sesgo, debido a la rigidez que implican las fórmulas que actualmente se usan para el cálculo del índice elemental, en términos de sustitución de productos y establecimientos. En tercer lugar, y bajo el supuesto de que se desea medir los cambios en el nivel de vida debido a los cambios en los precios, se deduce la interpretación desde el punto de vista estadístico de los índices de precios teóricos obtenidos a partir de una función de utilidad con elasticidad de sustitución constante. Por último, se propone el cálculo de un indicador de Índice del Coste de la Vida que combine distintas fórmulas para la obtención de los índices de los agregados elementales con el objeto de que éste sea capaz de captar el distinto nivel de sustitución que se puede producir dentro de cada uno de esos agregados elementales.

Palabras clave: Inflación, IPC, Índice del Coste de la Vida, CES.

Abstract

In the present work, we make a proposal for the construction of the Cost-of-Living Index (COLI), using the same structure of work and the same statistical information with which the Consumer Price Index (CPI) is calculated. Our first purpose in the present paper is to analyze the reasons for which the Statistic Agencies (Bureau of Labor Statistic, Eurostat, ...) calculate the CPI. Secondly, and, since the objective is not unique and usually it is related to the cost of living, we studied these two concepts and we concluded that, when the objective of the CPI is to calculate the COLI, the lower-level of the calculation of the CPI is the main source of bias, due to the rigidity that implies the formulas that at the moment are used for the calculation of the elementary index, in terms of substitution of products and establishments. Thirdly, and under the assumption that we want to measure the changes in the cost-of-living due to the changes in the prices, we deduce the statistical interpretation of the prices index that are obtained from a CES utility function. Finally, we propose to calculate a COLI that combines different formulas for the elementary Index. The objective of this index is to catch the different level from substitution that can be produced within each one of elementary aggregates.

Key words: Inflation, CPI, Cost-of-Living Index , CES.

1. Introducción

El control de la inflación es un elemento fundamental en la política económica de cualquier país. La razón se encuentra, por un lado, en su capacidad para modificar los flujos de capital entre regiones, en un mundo en donde el capital cada vez tiene más libertad para moverse geográficamente y, por otro lado, en que la inflación es el factor básico a la hora de fijar los cambios en el coste de la vida y, en consecuencia, en los cambios en la calidad de vida de los ciudadanos. Por tanto, la falta de control en los precios a nivel regional trae como consecuencia una ineficiente asignación territorial del capital y una divergencia en los niveles de vida de los ciudadanos.

El indicador más completo de la inflación es el Índice de Precios de Consumo (IPC), lo cual le convierte en el mejor indicador de los cambios en el nivel de vida de los ciudadanos. Ahora bien, el IPC no mide la inflación general de una economía, puesto que sólo tiene en cuenta el consumo de los hogares, ni tampoco mide el nivel de vida, dado que teóricamente sólo considera precios. Esta parcialidad del IPC ha provocado cierta divergencia entre los encargados de elaborarlo y los académicos que se centran en su estudio. Los primeros afirman que el IPC lo que pretende medir es la evolución pura de los precios de consumo [De Castro y González (2000), pág. 59], los segundos, cuando analizan el IPC lo hacen identificando sesgos que, en muchas de las ocasiones, están relacionados más con los cambios en el nivel de vida que con los precios puros. En esta última línea se encuentran, entre otros, el informe Boskin, que afirma que el Bureau of Labor Statistic (BLS) de Estados Unidos debiera fijar como objetivo de la medición de los precios de consumo un índice del coste de la vida [Boskin (1996), pág. 2]. En el mismo sentido se manifiesta Schultze y Mackie editores cuando afirman que el BLS debe acercar el IPC a un índice del coste de la vida [Schultze and Mackie (2002), Executive Summary, pág. 1]. Diewert (1998), Schultze (2003), Hausman, J. (2003), etc... se manifiestan de forma similar.

En gran parte, el distanciamiento entre los que elaboran el IPC y los académicos que lo estudian nace de la forma en la que se define esta medida y en su uso. Por ejemplo, el BLS define el IPC como un programa que produce mensualmente datos acerca de los cambios en los precios pagados por los consumidores urbanos para una cesta

representativa de bienes y servicios.¹ La misma fuente indica que el IPC combina teoría económica, técnicas estadísticas y datos para obtener una medida del cambio de los precios para el sector del consumo de la economía americana [BLS (2006), pág. 1] y reconoce indirectamente que existen problemas en su construcción, muchos de los cuales podrían resolverse si se abordara su cálculo a través de un esquema de trabajo como el que define el concepto de Índice del Coste de la Vida (ICV) [BLS (2006), pág. 2]. El enfoque del ICV no está exento de dificultades, con lo cual el BLS concluye que el objetivo del IPC es ofrecer una aproximación a un ICV condicional que incluye solamente los precios de mercado de bienes y servicios [BLS (2006), pág. 3].

La posición de la Unión Europea (UE) es distinta. En este caso, el IPC es el instrumento para medir la inflación que permita cuantificar el nivel de estabilidad de los precios de los distintos países que aspiran a formar parte de la Unión Monetaria. En el Reglamento (CE) N° 2494/95 del Consejo relativo a los índices armonizados de precios de consumo (IPCA), se reconoce la parcialidad de esta medida como indicador de la inflación, pero se defiende como un elemento fundamental de la misma. El objetivo de la Unión Europea es armonizar los distintos IPC de cada uno de los países, para que midan lo mismo. Sin embargo, dentro de la UE cada país calcula su propio IPC y, en algunos casos, se vuelven a mezclar los conceptos de inflación y calidad de vida. Por ejemplo, en España, el Instituto Nacional de Estadística clasifica el IPC en un área que denomina "Nivel, Calidad y Condiciones de Vida". La propia clasificación del indicador nos hace pensar en un uso que aborda aspectos que van más allá de la inflación y están más relacionados con un ICV. De hecho, INE (2006) define el IPC a través del objetivo que se persigue con su cálculo diciendo que "mide la evolución del nivel de precios de los bienes y servicios de consumo adquiridos por los hogares residentes en España" [INE (2006), pág. 11].

En este trabajo nos planteamos cómo conciliar el IPC y el ICV desde el punto de vista de la estructura de trabajo que aporta el concepto de ICV, pero con la información que utiliza en la actualidad el IPC, con el objeto de disponer de medidas más precisas, tanto de la inflación, como de los cambios que provocan los precios en el nivel de vida de los ciudadanos. Para ello, en el epígrafe siguiente se analizan algunos aspectos del IPC, en

¹ [Http://www.bls.gov/cpi/#overview](http://www.bls.gov/cpi/#overview), "The **Consumer Price Indexes (CPI)** program produces monthly data on changes in the prices paid by urban consumers for a representative basket of goods and services"

el punto tercero se aborda, desde el punto de vista estadístico, el cálculo de los índices de precios teóricos obtenidos a partir de las funciones de utilidad con elasticidad de sustitución constante (CES) y, en apartado cuatro, se plantea una forma de elaborar un ICV con los mismos datos con los que en la actualidad se calcula el IPC. En el último epígrafe se enumeran las principales conclusiones.

2.- El IPC, algunos aspectos de su elaboración y su relación con el Nivel de Vida

En la elaboración del IPC hay, al menos, dos elementos que son comunes, tanto para el BLS como para la UE. El primero de estos elementos es la información de base que, a su vez, determina el segundo elemento común, y que se refiere a que el IPC se elabora en dos fases. La fase elemental es aquella en la cual únicamente se utilizan datos de precios y da lugar al cálculo de los índices elementales. La/s fase/s agregada/s parte/n de los citados índices elementales y se combinan con la estructura del gasto para obtener distintos IPC agregados, bien por territorios, por necesidades o por ambos simultáneamente.

En la primera de las fases, la elemental, únicamente se utilizan datos de precios recogidos en los puntos de venta. Entre los diferentes países puede diferir mucho la forma en la que se seleccionan los establecimientos en donde se recogen los precios, la frecuencia con la que se observan, la propia selección de los productos para los cuales se miden los precios, etc. Ahora bien, existe un aspecto común en todos los casos: siempre se observan precios en los puntos de venta, nunca en los consumidores. Es decir, lo que se observa son los precios de la oferta.

Esta forma de recoger los datos tiene consecuencias muy importantes, tanto si el IPC se entiende como una medida de la inflación, como si se ve como una medida del nivel de vida. Ello se debe a que los precios medidos en los establecimientos constituyen la única información de la que se dispone para el instante actual y, además, sólo se utilizan en la fase elemental, puesto que, en las siguientes fases, ya se usan los índices de precios elementales y no los precios. Cualquier error o inconsistencia que se produzca en la transformación de los precios en sus índices elementales no puede ser subsanado en ninguna fase posterior de la elaboración del IPC.

En estas circunstancias, la fórmula de cálculo del índice elemental se convierte en un elemento clave a la hora de incorporar sesgos en el resultado final del IPC. De hecho, la mayoría de los sesgos que se identifican en la literatura se producen dentro del estrato elemental. Por ejemplo, los cuatro sesgos que Schultze (2003) identifica como particularmente importantes, a saber, el sesgo por sustitución de productos, el sesgo por cambios en la calidad, el sesgo por la aparición de nuevos productos y el sesgo a la hora de definir los bienes y servicios que deben formar parte del IPC, se producen en el nivel elemental. Lo mismo sucede con los que se analizan en Hausman (2003).

Numéricamente el sesgo no es trivial. En la Tabla 1 se muestran las diferencias relativas entre las inflaciones calculadas para cinco productos pertenecientes al grupo de alimentación entre dos instantes de tiempo separados seis meses, en función de qué fórmula de cálculo se ha utilizado para obtener el índice elemental, con respecto a la fórmula de Jevons. Los datos son del año 2005 y pertenecen a la Comunidad Autónoma de Canarias²

Tabla 1. Diferencias relativas en los índices elementales medida tomando como referencia la fórmula de Jevons					
	Producto 1	Producto 2	Producto 3	Producto 4	Producto 5
Carli	204.08	71.50	400.22	914.61	142.78
Dutot	4.40	-12.38	43.12	11.47	4.68
CES (2)	-6.14	-11.21	-41.07	-5.18	-8.15
Carli: Media aritmética de los índices de precios					
Dutot: Cociente entre las medias de los precios en cada instante de tiempo que se compara					
Jevons: Media geométrica de los índices de precios					
CES (2): Índice de precios teórico de una función de elasticidad de sustitución constante igual a 2					

Es decir, si en vez de utilizar la fórmula de Jevons se utilizase la de Carli, la inflación medida para el producto 1 sería un 204% superior, para el producto 2 un 71.5% superior, etc...La fórmula de Carli es especialmente llamativa, pero las diferencias son

² Los datos fueron suministrados por el Instituto de Estadística de Canarias.

lo suficientemente importantes entre todas las fórmulas³ como para que se deban de tener en cuenta. El BLS utiliza la fórmula de Jevons y Eurostat permite tanto ésta como la de Dutot y, sólo en casos excepcionales, se puede usar la de Carli.

En cualquier caso, todos los índices elementales, al no incorporar ninguna ponderación explícita del volumen de productos realmente demandados, son indicativos de la inflación que ofertan los puntos de venta, que denominaremos inflación ofertada. Ahora bien, los consumidores, ante una inflación ofertada en el instante t con respecto a un instante cero, pueden aceptarla en su totalidad, sólo parcialmente o no aceptarla. Es decir, existe otra inflación, que es la que realmente soportan los consumidores, que denominaremos inflación demandada. Esta inflación es la que mide la inflación de los consumidores y es la que debiera usarse para estimar un ICV.

Caben, al menos, tres posibilidades para calcular la inflación demandada. Primera, obtener datos de cantidades también a nivel elemental. Esta es la propuesta, por ejemplo, de Hausman (2003), que se base en que en la actualidad muchos establecimientos ya disponen de estos datos, puesto que el sistema de venta que utilizan se realiza mediante el escaneo electrónico de los actos de consumo. Sin embargo, reconoce que, en la actualidad, existen lagunas de información y, en cualquier caso, esto sería una ruptura total de la forma de trabajar de las agencias oficiales de encargadas de elaborar el IPC. Segunda, obtener los precios, no de los puntos de venta, sino de los consumidores. Esta alternativa presenta problemas de representatividad, sobre todo cuando se refiere a productos duraderos. Tercero, incorporar formulas de cálculo de los índices elementales que se correspondan con comportamientos teóricos de los consumidores. Las dos primeras soluciones obligan a modificar el conjunto de datos actualmente utilizados en la elaboración del IPC, la tercera no obligatoriamente.

En cualquier caso, todo depende de lo que se quiera medir. Si el IPC se utiliza como medida de la inflación de la economía, el uso de los precios de la oferta o de la demanda son igualmente válidos. Ahora bien, para mantener la coherencia, si se decide usar precios de oferta, éstos deben usarse en todos los niveles y no solo en la elemental. Sin embargo, en las fases de agregación superiores al nivel elemental, se incorporan pesos de la demanda, que sí se obtienen a partir de encuestas realizadas a los consumidores.

³ Para la justificación estadística de estas diferencias véase Rodríguez, González y Rodríguez (2005).

La conclusión, llegado este punto, es que, tanto en los objetivos para los cuales se diseña el IPC, como en la forma de elaborarlo, se están mezclando componentes de distinta naturaleza que dificulta su uso y su capacidad para representar aquello para lo que fue diseñado. Así, el BLS se define en términos de ICV pero a nivel elemental únicamente utiliza datos de precios actuales obtenidos desde la oferta, no desde la demanda. Por otra parte, la UE se define en términos de inflación de la economía y en la fase elemental utiliza precios de la oferta para, en las siguientes fases, mezclarlos con datos de la demanda de consumo.

3.- El ICV en un contexto de Utilidad CES. Su interpretación estadística.

Por cuestiones operativas, las agencias de estadística necesitan una fórmula que se puede expresar como una media ponderada de los índices elementales. Es decir, una expresión de la forma de [1].

$$I_{t/0}^h = \sum_{j=1}^{n_h} \omega_j^h I_{t/0}^j \quad [1]$$

Siendo t y 0 dos instantes de tiempo, con t posterior a 0, h es un determinado nivel agregado que está formado por n_h estratos elementales, $I_{t/0}^j$ es el índice del agregado elemental j-ésimo y ω_j^h es la ponderación del estrato elemental j-ésimo dentro del agregado h, que cumple $\sum_{j=1}^{n_h} \omega_j^h = 1$. Estas ponderaciones se corresponden con la demanda y en la actualidad, dado que la fórmula de cálculo es la de Laspeyres, se refieren a un período anterior al instante actual, denotado por t. Dependiendo de que se utilice un Laspeyres o un Laspeyres encadenado, la ponderación se corresponde con el período base, denotado por 0 o con un período t' comprendido entre 0 y t, respectivamente.

Si las ponderaciones con las que se elaboran los niveles agregados del IPC se actualizan de forma continua, es evidente que la estructura de la inflación desde el lado de la oferta tiende a converger con la de la demanda. Es decir, a medio plazo los consumidores se ajustan a la oferta y los productores a la demanda a la hora de satisfacer/ofrecer necesidades/productos.

Esta forma de calcular el IPC tiene varias implicaciones: primero, si se usa la fórmula de Laspeyres (lo habitual), para calcular el índice en el instante t el único dato disponible para ese instante de tiempo es el precio de los productos obtenido en los establecimientos que lo ofertan (obsérvese que ponerlo a la venta no implica que los consumidores lo adquieran). Segundo, estos precios sólo se utilizan para calcular los índices elementales. Tercero, la hipótesis con la que se trabaja cuando se usa una ponderación fija para dos instantes de tiempo distintos es que la estructura de la demanda de cada nivel de agregación, que definen las propias ponderaciones, no se modifica. Es decir, todos los cambios de precios se ajustan dentro de cada uno de los niveles agregados que se calculan y no afectan al resto de agregados en términos de modificación de las pautas del gasto de consumidor. Por tanto, si un índice elemental crece mucho, se reduce la cantidad consumida de los productos que contiene, para mantener constante el gasto realizado en el conjunto de productos que forman el agregado elemental, pero no afecta a la demanda del resto de productos a los que puede acceder el consumidor para satisfacer la misma u otra necesidad (suponiendo constante todos los demás precios).

Si pensamos en términos de necesidades del consumidor, si las ponderaciones se corresponden con un instante de tiempo próximo a t y si la inflación es moderada, la hipótesis de mantenimiento del gasto es más realista cuanto más heterogéneas sean las necesidades de los agregados que se combinan. Es decir, si el aceite sube de precio, el consumidor, antes que modificar la cantidad de carne que consume, actuará reduciendo la cantidad de aceite para mantener el gasto en esta partida y, por tanto, no afectará al resto de su consumo. Ahora bien, en el estrato elemental estamos hablando de bienes y servicios que todos ellos satisfacen una misma necesidad del consumidor. El ejemplo anterior no considera la posibilidad de que se produzcan cambios de las pautas de consumo dentro del propio estrato elemental. Por ejemplo, buscando un establecimiento más barato o buscando un aceite alternativo que sea más barato y que le permita consumir las mismas cantidades.

Ahora bien, las necesidades del consumidor pueden tener una naturaleza muy diversa. No es lo mismo una necesidad de ocio, que una de salud o que una de alimentación. En principio, sería razonable pensar que si un producto de ocio se encarece, el consumidor lo sustituya por otro de precio relativo menor, incluso puede llegar a dejar de

consumirlo. Sin embargo, cuando hablamos de una necesidad de salud el consumidor puede ver el precio como un indicativo de calidad. De hecho, en este tipo de necesidades la función de demanda del consumidor viene determinada, fundamentalmente, por la función de producción del ofertante. Si la necesidad del consumidor es del grupo de alimentación, existe una amplia variedad de alternativas de consumo, que hace que dentro del estrato elemental se produzca una alta sustitución, aunque es más difícil que se deje de satisfacer la necesidad que motiva el acto de consumo.

Como ya se ha comentado, el índice del estrato elemental j-ésimo, $I_{t/0}^j$, se calcula como una función g de los precios observados en los establecimientos en los instantes de tiempo 0 y t. Si denotamos por k a las unidades de observación de los precios en el estrato elemental j, el $I_{t/0}^j = g(p_j^0, p_j^t)$. Las funciones que se usan en la actualidad son el cociente de las medias de los precios, fórmula de Dutot, y el cociente del producto de los precios, fórmula de Jevons, para cada uno de los instantes de tiempo que se comparan. Ambas fórmulas se corresponden con los índices de precios verdaderos correspondientes a una determinada función de utilidad CES.

Para una Economía que tiene un conjunto de n bienes de consumo, la función de utilidad CES homogénea de grado 1 se define como:

$$U^t = A [\sum_{i=1}^n a_i (q_i^t)^\rho]^{1/\rho} \quad [2]$$

Siendo U la utilidad, t el tiempo, A es un parámetro de escala, a_i es la ponderación del producto i y q_i^t son las cantidades consumidas del producto i en el instante t. Las fórmulas que se utilizan para calcular los índices elementales se refieren a los casos límite de una función U en la cual $a_i = \frac{1}{n}, \forall i$. Si $\rho \rightarrow 0$, se obtiene una función de utilidad Cobb-Douglas, cuyo índice de precios verdadero es la fórmula de Jevons, mientras que si $\rho \rightarrow \infty$, da lugar a una función Leontief, cuyo índice de precios verdadero es la fórmula de Dutot.

Volviendo al caso general definido en [2], es inmediato demostrar que la elasticidad de sustitución es igual a $\varepsilon = \frac{1}{1-\rho}$. Además, con esta nomenclatura, el índice de precios teórico es [3].

$$I_{t/Q} = \frac{\left[\sum_{i=1}^n \left(\frac{a_i}{p_i^\rho} \right)^\varepsilon \right]^{\frac{1}{\rho\varepsilon}}}{\left[\sum_{i=1}^n \left(\frac{a_i}{p_i^\rho} \right)^\varepsilon \right]^{\frac{1}{\rho\varepsilon}}} \quad [3]$$

Desde el punto de vista axiomático, [3] cumple todas las propiedades que ILO(2003) enumera como deseables para cualquier índice elemental, con la excepción de la propiedad de proporcionalidad, propiedad que no se encuentra dentro de las consideradas fundamentales.

La expresión [3] no facilita ni la interpretación del índice ni se ajusta a la forma con la que trabajan las agencias de estadística. Ahora, bien, si definimos una función f para la unidad de observación muestral k perteneciente al estrato elemental j en el instante t como [4]

$$f_k^t = \frac{a_k^\varepsilon \sum_k ({}^k p_j^t)^{\rho\varepsilon}}{({}^k p_j^t)^{\rho\varepsilon}}, \quad [4]$$

Es inmediato comprobar que f_k^t toma valores más grandes cuando a_k crece y cuando, en su conjunto, los precios son más altos, y se hace más pequeño cuando el precio en la observación muestral k -ésima es más grande. Es decir, ante una situación dada definida por un conjunto de preferencias, $\{a_k\}$, por un conjunto de precios dentro del agregado elemental j , $\{{}^k p_j^t\}$, y con una utilidad del tipo CES, la función f_k^t será más pequeña cuanto mayor sea el precio en la unidad muestral k -ésima. En consecuencia, si se define una ponderación del tipo [5]

$$\omega_k^t = \frac{f_k^t}{\sum_k f_k^t}, \quad [5]$$

ésta se caracteriza por mantener una relación inversa con los precios. Es evidente que la relación inversa depende también de los parámetros que definen la función de utilidad, las constantes α_k y ε .

Si definimos un índice de precios para el agregado elemental j como [6]

$$I_{t/0}^j = \left[\sum_k \omega_k^t \left[I_{t/0}^k \right]^{\rho \varepsilon} \right]^{\frac{1}{\rho \varepsilon}} = \left[\sum_k \omega_k^t \left[\frac{p_j^t}{p_j^0} \right]^{\rho \varepsilon} \right]^{\frac{1}{\rho \varepsilon}}, \quad [6]$$

sustituyendo [3] en [4] y el resultado en [5], es inmediato comprobar que las expresiones [3] y [6] son iguales.

Es decir, los índices teóricos que se obtienen de una función de utilidad tipo CES se pueden interpretar como promedios de los índices de precios, en donde la ponderación es inversamente proporcional a los precios del instante actual y depende de los parámetros que definen la función de utilidad, en concreto de la participación de cada uno de los índices y de la elasticidad de sustitución. A este mismo resultado se llega si se parte de las ponderaciones referidas al instante base, usando una media armónica definida como [7].

$$I_{t/0}^j = \left[\sum_k \omega_k^0 \left[I_{t/0}^k \right]^{-\rho \varepsilon} \right]^{-\frac{1}{\rho \varepsilon}} = \left[\sum_k \omega_k^0 \left[\frac{p_j^t}{p_j^0} \right]^{-\rho \varepsilon} \right]^{-\frac{1}{\rho \varepsilon}} \quad [7]$$

En la actualidad las agencias de estadística no usan ponderaciones explícitas en el cálculo de los índices elementales. Además, las constantes α_k son iguales para todos los precios observados, usándose las fórmulas de Dutot y Jevons, que, como hemos dicho, se corresponden con funciones de utilidad Leontief y Cobb-Douglas respectivamente. Al no usar ponderaciones para cada índice de precios de una unidad muestral, la fórmula de Dutot pondera de manera directa cada uno de estos índices por el precio en el instante cero. Es decir, la importancia de un cambio en precios es mayor cuanto mayor sea su precio en el instante de referencia. Este resultado es compatible con una función de utilidad del tipo Leontief, pero no es adecuado para representar el comportamiento del consumidor en la mayoría de los bienes, sobre todo cuando se está trabajando con un conjunto de productos que todos ellos tienen capacidad para satisfacer una única necesidad. Recordemos que una de las propiedades que tiene la función tipo Leontief es la complementariedad.

En los últimos años, la fórmula más utilizada para el índice del agregado elemental es la de Jevons. Como ya hemos dicho, se corresponde con una utilidad Cobb-Douglas que, a su vez, presenta una elasticidad de sustitución igual a 1. Dado que no se usa ningún tipo de ponderación explícita, todos los índices de precios de las unidades muestrales tienen el mismo peso, sea cual sea la posición que dichos precios presentan en la distribución de los mismos. Al trabajar con una distribución de utilidad de Cobb-Douglas sabemos que la elasticidad precio directa es -1 y la cruzada es 0. Por ello, una conclusión del uso de esta fórmula es que, cuando un precio en una unidad muestral cambia, siempre se produce una reducción equivalente de la cantidad demandada en dicha unidad muestral, no afectando al resto de productos del estrato elemental. Nuevamente este resultado puede no representar al comportamiento del consumidor, puesto que en algunos de los estratos elementales los productos son lo suficientemente homogéneos como para que el consumidor sustituya el producto dentro del agregado con el fin de mantener la máxima cantidad de producto consumido. Sin embargo, para algún tipo de necesidades, sí puede ser adecuado, por ejemplo, en necesidades de ocio. Los productos que satisfacen esta necesidad son, por definición, heterogéneos, con lo cual si a un individuo le gusta ir al teatro, el incremento de su precio puede provocar una reducción de la demanda de teatro, sin que ello afecta a la parte de gasto del resto de productos de consumo dentro del propio estrato elemental.

4.- Propuestas de elaboración de IPC e ICV

El uso de una única fórmula, sea la de Dutot o la de Jevons, para calcular todos los índices de precios elementales de un IPC no parece la opción más adecuada, debido a cada uno de ellos restringe el comportamiento del consumidor a un escenario muy concreto y en absoluto general. El elemento fundamental en la elección de la fórmula del estrato elemental, cuando no se dispone de ponderaciones, se encuentra en el nivel de sustitución que presentan las unidades muestrales que forman cada agregado elemental, que, a su vez, por el tipo de necesidad que el consumidor pretende satisfacer.

Para muchos estratos elementales la fórmula debe recoger la sustitución que se produce entre los distintos bienes y establecimientos que ofrecen productos englobados en dicho estrato elemental. Dado que a nivel elemental únicamente se dispone de información de

precios y que estos cambian a lo largo del tiempo, la fórmula que se utilice debe recoger los cambios en la demanda que se produce debido a los cambios en los precios.

Como esta respuesta es distinta en función de la necesidad que se pretende satisfacer, la fórmula debe permitir ponderar de forma distinta por los precios en el período actual. No olvidemos que el índice que se desea calcular se refiere al instante t y los únicos datos disponibles para ese instante de tiempo son los precios observados en los puntos de venta.

Por ello, nuestra propuesta es que los estratos elementales se deben clasificar en varios grupos dependiendo del nivel de homogeneidad de las unidades muestrales mediante las cuales se representa. Uno de estos grupos estaría formado por aquellos estratos elementales que satisfacen necesidades fundamentales y para las cuales el ofertante tiene una posición de control sobre el mercado (por cuestiones geográficas, monopolios, etc.). En estos estratos, el consumidor necesita mantener las cantidades que consume, para lo cual incluso puede llegar a reducir el consumo de otros bienes con el fin de incrementar el gasto en este estrato elemental. En este caso, la fórmula a utilizar podría ser la de Dutot.

Un segundo grupo de estratos lo formarían aquellos que satisfacen necesidades no básicas mediante unos productos poco homogéneos. Un ejemplo de este tipo de estrato son algunos productos de ocio. En ellos el consumidor puede determinar la proporción de gasto que le asigna a la satisfacción de esta necesidad y, luego, fijar la cantidad en función de los precios. En este caso la fórmula a utilizar es la de Jevons.

Sin embargo, en una Economía desarrollada, el grupo más amplio lo formarían estratos elementales en los cuales el consumidor tiene cierta capacidad para sustituir sus patrones de consumo, en función de cómo los ofertantes definan la distribución de precios en cada instante de tiempo. Dentro de cada uno de estos estratos, el consumidor cambia de establecimiento y/o modifica la cantidad de cada uno de los productos que consume, como respuesta a la inflación que los productores ofertan en cada instante de tiempo. Es decir, existe una alta capacidad de sustitución dentro del estrato elemental. En este caso, dado que no se dispone más que de precios dentro del estrato, la única solución coherente es usar como fórmula elemental alguna expresión que incorpore una

relación inversa con la distribución de precios en el instante actual. Esta propiedad la tienen los índices de precios teóricos correspondientes a una función CES.

Por ejemplo, si se supone que un determinado estrato presenta una elasticidad de sustitución igual a 2, la fórmula para el índice de precios del agregado elemental, bajo las condiciones actuales de información utilizadas para calcular el IPC, sería [8].

$$I_{t/0}^j = \frac{\sum_k \frac{k}{k_{p_j}^0}}{\sum_k \frac{k}{k_{p_j}^t}} \quad [8]$$

Este resultado coincide con el que se llegaría usando [9], con lo cual, el cambio en los precios simples está ponderado de forma inversa por el precio de cada uno de los bienes del estrato en el período actual.

$$I_{t/0}^j = \left[\sum_k \omega_k^t \left[I_{t/0}^k \right] \right], \text{ con } \omega_k^t = \frac{f_k^t}{\sum_k f_k^t} \text{ y } f_k^t = \frac{\sum_k (k_{p_j}^t)}{(k_{p_j}^t)} \quad [9]$$

Las fórmulas [8] y [9] se corresponden con el índice de precios teórico de una función de utilidad $U^t = A \left[\sum_{i=1}^n a_i (\sqrt{q_i^t}) \right]^2$ con valores iguales para todo a_i . Es inmediato comprobar que para esta función U^t la elasticidad de las cantidades con respecto a su propio precio se encuentra comprendida entre 0 y -2 obteniéndose mediante la expresión [10]

$$\varepsilon_{q_i}^{p_i} = \frac{\frac{a_i^2}{p_i}}{\sum_i \frac{a_i^2}{p_i}} - 2, \quad [10]$$

mientras que la elasticidad precio cruzada toma valores entre 0 y 1, siendo su expresión la que se muestra en [11].

$$\epsilon_{q_i}^{p_j} = \frac{\frac{a_j}{p_j}}{\sum_i \frac{a_i}{p_i}} \quad [11]$$

Si comparamos estos resultados con los que se derivan de utilizar como fórmula para los índices de precios la de Dutot o la de Jevons, las diferencias son claras. En primer lugar, las elasticidades precio, tanto propia como cruzada, son distintas para cada uno de los productos. En segundo lugar, los signos son los esperados por la teoría del consumidor. Y, tercero, la sustitución se realiza en función de la posición que ocupa el precio que cambia dentro de la distribución del conjunto de todos los precios.

Volviendo al caso general, es evidente que cuanto más grande sea la elasticidad de sustitución, ϵ , el cambio en el precio del producto más caro tiene menos importancia en el cálculo del índice del agregado elemental. Esto es más aceptable desde el punto de vista teórico cuanto más homogéneos sean los productos que forman el agregado elemental.

El uso de las fórmulas [3], [6] o [7] como fórmula para el agregado elemental, supone usar un índice elemental que se corresponde con un comportamiento esperado del consumidor. A partir de estos índices se obtiene un ICV combinándolos con las ponderaciones obtenidas mediante una encuesta de consumo. Es evidente que el ICV así calculado es un índice teórico. Su capacidad para representar el verdadero cambio en el nivel de vida, debido al cambio de los precios, pasa por la representatividad de las elasticidades estimadas dentro de cada estrato elemental y el carácter constante al que obliga la función CES.

La forma de medir estas elasticidades es múltiple. Una de las opciones es la que defiende, entre otros, Hausman (2003) y que consiste en obtener cantidades, también, dentro del agregado elemental. Su opinión es que actualmente existen ya datos de precios y cantidades, que aquellos que los establecimientos obtienen mediante el proceso de lectura electrónica que utilizan en el momento de la venta. Esto significa que ya existen datos muestrales de precios y cantidades a nivel elemental para muchos tipos de gastos de las familias, que podrían ser usados para estimar un ICV. Sin embargo, en nuestra opinión, esta opción perpetúa un viejo problema, y es que los datos que se

toman vuelven a proceder de los puntos de venta y no del comportamiento del consumidor, con lo cual, los movimientos de los consumidores dentro del marco que define la oferta pueden no observarse. Por otro lado, esto supondría un cambio radical en el trabajo que realizan las agencias de estadística para elaborar el IPC-ICV.

Nuestra propuesta pasa por modificar, mediante la incorporación de nuevas preguntas, las actuales encuestas sobre consumo que se realizan a las familias y que permiten obtener las ponderaciones de gasto con las que se agregan los índices de precios elementales. Las preguntas que se incorporarían en dicha encuesta tendrían como objetivo obtener información directamente del consumidor, que permitiera estimar el nivel de sustitución dentro de sus hábitos de consumo y para cada necesidad que el consumidor pretende satisfacer. Sería preguntas como las que se muestran en el Cuadro 1.

Con esta información se clasificarían los estratos elementales en términos de elasticidad y se usaría ésta, junto con la fórmula de los índices de precios teóricos de una función CES, para calcular cada índice de precios elemental con la expresión que más se ajuste a las características de la necesidad que se pretende satisfacer con el consumo de los productos que contiene el estrato elemental.

Cuadro1. Propuestas de preguntas a incluir en la ECPF

- a) Si se incrementase el precio del conjunto de productos/establecimientos que usa Usted actualmente para satisfacer la necesidad X, ordene las siguientes alternativas con un 1 si es la primera opción que Usted usaría y un 5 si es la última, partiendo del supuesto que su renta total no cambia.
- o Incrementar el Gasto para mantener el patrón de consumo(cantidades y establecimientos)sabiendo que ello no me permitirá satisfacer alguna otra necesidad
 - o Mantener el Gasto constante buscando establecimientos más baratos
 - o Mantener el Gasto constante buscando productos alternativos más baratos
 - o Mantener el Gasto constante buscando productos y establecimientos más baratos
 - o Reducir el gasto quedando mis necesidades X insatisfechas
- b) Partiendo de su actual patrón de consumo (definido por los productos que consume y los establecimientos donde los adquiere) que le permite satisfacer la necesidad X, indique numéricamente cual sería su reacción ante los siguientes incrementos de precios en su patrón de consumo, suponiendo que su renta total no cambia.

Si el incremento de precios en mi patrón de consumo fuese del:	Estaría dispuesto a Reducir las cantidades que actualmente consumo en mis establecimientos actuales hasta un:	Y estaría dispuesto a cambiar mi patrón de consumo siempre y cuando los nuevos precios fuesen más baratos al menos en un:
1%%%
2%%%
5%%%
10%%%
15%%%
25%%%

5.- Conclusiones

Los principales problemas de sesgos que se producen en la elaboración del IPC, cuando se desea usar como medida de cambio en el coste de la vida, se incorporan en la primera fase de su cálculo, el nivel elemental. En gran medida, ello es debido a que los únicos datos que se utilizan para calcular los índices elementales no proceden de los consumidores y las fórmulas que actualmente se usan se corresponden con comportamientos de los consumidores que se explican mediante funciones de utilidad del tipo Leontief y Cobb-Douglas, cuya capacidad de representación es, cuanto menos, incompleta. Si se quiere calcular un ICV mediante un esquema de elaboración similar al del IPC es necesario que los índices elementales se obtengan con una mayor variabilidad en el comportamiento que sus fórmulas de cálculo actuales representan.

La propuesta que se hace en este trabajo es utilizar fórmulas de cálculo para los índices elementales procedentes de índices de precios teóricos correspondientes a funciones de utilidad CES, para lo cual se propone modificar las encuestas de consumo con las que se obtienen las ponderaciones que se usan a nivel superior al elemental, con el fin de que estas recojan la información que permita estimar la elasticidad de sustitución para cada uno de los estratos elementales en los cuales se dividen los bienes y servicios de consumo. Esto supondría el no utilizar una única fórmula de cálculo para los índices de precios elementales, ajustándose, en cada caso, a los niveles de sustitución dentro de cada estrato elemental que manifiestan los consumidores.

La familia de fórmulas propuesta se caracteriza por mantener una relación inversa con los precios, tanto en el período actual como en el de referencia, por tener en cuenta la distribución de los mismos y por tener un comportamiento, desde el punto de vista axiomático, similar al de las fórmulas que se usan en la actualidad.

Con el esquema de trabajo propuesto y, a pesar de que los datos de precios que se observan se refieren a los precios que ofertan los establecimientos, todas las fases de elaboración del IPC-ICV se ponderan desde la perspectiva del consumidor, obteniéndose una medida más cercana al verdadero cambio en el coste de la vida de los consumidores.

Bibliografía

BLS (2006). "Chapter 17, The Consumer Price Index", Handbook of Methods, <http://www.bls.gov/opub/hom/pdf/homch17.pdf>, pág. 1

Boskin, M. et al. 1996. "Toward a More Accurate Measure of Cost of Living", Final Report to the Senate Finance Committee.

De Castro, M.A. and Ignacio González (2000). "Índices de Precios de Consumo. Índices del Coste de la Vida. Fundamentos y Aspectos Metodológicos Más Destacados", *Estadística Española*, Vol. 42, Núm. 146, pp. 59-82.

Diewert, W. E. (1998). "Index Number Issues in the Consumer Price Index", *Journal of Economic Perspectives*, Vol. 12, nº 1, pp. 47-58.

Hausman, J. (2003). "Sources of Bias and Solutions to Bias in the Consumer Price Index", *Journal of Economic Perspectives*. Vol. 17, Nº 1, pp. 23-44.

INE, (2006). IPC. Documento Metodológico, Documento Electrónico, http://www.ine.es/ipc01/metodologia_ipc2001.pdf.

ILO, (2003). CPI Manual, Documento Electrónico: <http://www.ilo.org/public/english/bureau/stat/guides/cpi>, International Labour Organization.

Rodríguez, S., Carlos González y Alejandro Rodríguez, (2005). "Inconsistencia Económica de las Fórmulas de Carli, Dutot y Jevons para Calcular el Índice Elemental del IPC con Estratos Básicos Homogéneos", *Estadística Española*, Vol. 47, nº. 160, pp. 423-466.

Schultze, C. and Christopher Mackie, eds. 2002. *At What Price* Washington, D.C.: National Academy Press.

Schultze, C. (2003). "The Consumer Price Index: Conceptual Issues and Practical Suggestions", *Journal of Economic Perspectives*, Vol. 17, nº 1, pp. 3-32.

UE, (1995). Reglamento (CE) Nº 2494/95 del Consejo de 23 de Octubre de 1995 Relativo a los Índices Armonizados de Precios de Consumo.

CORRUPCIÓN PÚBLICA: ACTUACIÓN DEL GOBIERNO Y CUANTÍAS DE SOBORNO

M. D. SOTO TORRES
e-mail: lolasoto@eco.uva.es

R. FERNÁNDEZ LECHÓN
e-mail: ramonfer@eco.uva.es

Departamento de ECONOMÍA APLICADA
UNIVERSIDAD DE VALLADOLID

Área temática: Métodos Cuantitativos

Resumen

Considerando un mercado en el que participan un número finito de empresas, sin barreras de entrada y salida, aunque la permanencia en el mercado requiera el pago de sobornos como consecuencia de procesos de corrupción, se determina la evolución dinámica del tamaño del mercado, si las empresas siguen un equilibrio de Nash-Cournot, y las cantidades pagadas por sobornos ante tres posibles actuaciones del gobierno: no se lucha contra la corrupción, se adoptan medidas moderadas o se adopta una clara política para combatirla. El trabajo comprueba, después de analizar el problema de un agente corrupto representativo, que las cantidades pagadas por cada empresa al agente corrupto podrían disminuir a medida que el gobierno intensifica su lucha contra la corrupción.

Palabras clave: Corrupción, medidas gubernamentales, cuantías de soborno.

Abstract

This paper considers a model about a market, without barriers of entry and exit, in which a finite number of firms take part. Nevertheless, as consequence of public corruption activities, the permanence in the market of any firm implies a bribe payment. If the firms follow a Nash-Cournot equilibrium, the paper determines the evolution dynamics of the size market and the bribe payments under three possible performances of the government. Firstly, it is assumed that the government does not combat the corruption; secondly, the government adopts a moderate policy for combating it and, finally, the government adopts a clear policy of fight against it. After of analysing the problem associated to a representative corrupt agent, the paper checks as the bribe payment of each firm could diminish if the government steps up the fight against the corruption.

Key words: Corruption, government policies, bribe payment.

1. Introducción

En las últimas décadas, el problema de la corrupción ha llegado a ser uno de los temas más ampliamente debatido en los foros políticos, sociales y económicos como consecuencia de su incidencia negativa sobre diversos aspectos entre los que pueden citarse: el entorno, los derechos humanos, las instituciones democráticas, el desarrollo económico, la pobreza, las libertades o la justicia. Distintas organizaciones internacionales como el Banco Mundial, el Fondo Monetario Internacional, las Naciones Unidas o la Organización para la Cooperación y el Desarrollo Económico, han desarrollado y puesto en práctica numerosas estrategias para luchar contra ella. No obstante, la corrupción es persistente y omnipresente en todo el mundo. En mayor o menor medida, afecta a todas las sociedades independientemente de su estado de desarrollo y del tipo de régimen político-económico dominante en ella.

Aunque la corrupción es claramente clandestina y, por tanto, difícil de cuantificar, su incidencia y grado de penetración en ciertos países puede ser comprobada utilizando datos de corrupción publicados por diferentes organizaciones. Por ejemplo, Political Risk Services Inc.¹, es una empresa privada que anualmente publica el International Country Risk Guide en el que se analiza el grado de corrupción en ciertos países y está construido teniendo en cuenta los resultados de encuestas realizadas a inversores extranjeros en aquellos países; Transparencia Internacional² publica el índice de corrupción percibido, el primero fue publicado en 1995, y se construye sobre la opinión de expertos e inversores en el país correspondiente. Este índice abarca más de 150 países en el periodo 2004-2006. El IMD³ es otro índice de corrupción publicado por el Institute for Management Development in the World Competitiveness Yearbook y está basado en investigaciones de expertos abarcando a más de 50 países.

El significado de corrupción puede ser distinto entre un país y otro debido a diferencias culturales, legales e incluso, religiosas. No obstante, para normalizar el concepto Transparencia Internacional define corrupción como el abuso de una posición pública para una ganancia privada. Alguna razón por la que la corrupción está presente en todas las sociedades es aportada por Blackburn et al. (2006) al afirmar que la corrupción

¹ www.prsgroup.com

² www.transparencia.org.es

³ www.imd.ch

pública es un aspecto inevitable en las sociedades como consecuencia de las transferencias de responsabilidades del gobierno a los burócratas. Pope (2000) afirma que la corrupción puede tener lugar donde exista una clara combinación de oportunidad e inclinación; la corrupción puede ser iniciada desde una cualquiera de dos posibles posiciones: un soborno es ofrecido a un funcionario o el funcionario requiere un pago ilícito. Los que ofrecen sobornos lo hacen con objeto de conseguir alguna cosa que de otro modo sería difícil de obtener, o que para conseguirlo sería necesario saltarse las reglas establecidas o también, sobornan pensando que no podrían alcanzar sus objetivos a no ser que induzcan a ello a un funcionario público. Los funcionarios, por su parte, podrían rehusar servir a sus clientes a menos que se ofrezca una compensación por un trabajo que entra dentro de sus obligaciones.

En este último caso, sería posible diferenciar entre un pequeño soborno y la gran corrupción llevada a cabo por altos funcionarios públicos que son sobornados con grandes cantidades monetarias que se ocultan en paraísos fiscales. Mauro (1997), quizá el autor más ampliamente citado en las investigaciones científicas que abordan el problema de la corrupción, en su análisis sobre las causas y las consecuencias de la corrupción, justifica, en parte, los pequeños sobornos a los funcionarios públicos, es aquello que él denomina corrupción de baja intensidad, para que determinados procesos administrativos legales no se dilaten innecesariamente.

El trabajo de Ades et al. (1999) señala que dependiendo de la actividad desarrollada por los profesionales éstos proponen distintas medidas para luchar contra la corrupción. Los legisladores consideran que uno de los caminos a seguir para luchar contra la corrupción es la reforma de los sistemas legales así como el incremento de las penalizaciones a los corruptos; los inversores estarían de acuerdo en que el problema de la corrupción pública se encuentra muy ligado a los salarios de los funcionarios, de modo que el mantenimiento de los salarios de los burócratas iguales a los salarios del sector privado, bajo semejantes responsabilidades, debería ser considerado por los gobiernos como una política adecuada para luchar contra la corrupción. Por último, los economistas propondrían medidas encaminadas a promover y fomentar la competencia, asegurando que un sistema claramente competitivo no podría convivir con el pago de sobornos. No obstante, los resultados teóricos sobre los efectos de la competencia en la corrupción son ambiguos. Bliss et al. (1997) comprueban que países que progresivamente han ido

incrementando su nivel de competencia, han experimentado, en ocasiones, un paulatino recrudescimiento de la corrupción. Estos autores sugieren que la competencia es más un instrumento para controlar la corrupción, que para provocar una efectiva reducción de la misma.

El objetivo de este trabajo, es realizar un desarrollo teórico para comprobar como políticas activas de persecución de la corrupción provocarían una disminución de su intensidad. Para alcanzar este objetivo, se construye un modelo sobre un mercado de libre entrada y salida, en el que participan un número finito de empresas aunque potencialmente podrían participar un número suficientemente amplio de productores. Se supone que las empresas del mercado compiten en cantidades de producción y no en precio, siendo la producción total del mercado la determinada por ese factor. La permanencia en el mercado de cada empresa supondrá el abono de una cantidad a ciertos burócratas corruptos cuya cuantía será determinada en función del grado de persecución del gobierno contra la corrupción y de los beneficios de las empresas.

En una primera etapa se supone que no se lucha contra la corrupción, posteriormente se considera que el gobierno sí adopta alguna medida para combatirla, entonces el burócrata corrupto tendrá alguna posibilidad de ser capturado. En este caso, considerando un programa no lineal con restricciones en desigualdad y cuya función objetivo es la ganancia esperada del funcionario corrupto, se determina bajo que condiciones todas las empresas del mercado pagarán sobornos y la cuantía del pago del soborno de cada una de ellas. En una tercera etapa, el modelo supondrá que el gobierno endurece las medidas para luchar contra la corrupción provocando que el corrupto tenga más posibilidades de ser capturado. En este último caso, el problema asociado al corrupto proporciona una solución óptima cuyo pago de soborno nunca supera al pago de soborno de las dos situaciones previas.

Algunos aspectos de la formulación del modelo que presentamos fueron previamente estudiados por Emerson (2006); no obstante, existen claras diferencias respecto al enfoque y tratamiento dado a los modelos. En nuestra aproximación se determina la evolución del tamaño de mercado desde la evolución de los precios, presentando uno de los estados de equilibrio dinámicos del tamaño del mercado las mismas características de estabilidad que el precio de equilibrio dinámico. En nuestro enfoque, es el mercado

el que determina su tamaño de acuerdo con la oferta y demanda de sus productos y las cuantías del soborno se adaptan a los beneficios que tienen las empresas del mercado. Emerson, por su parte, considera que el pago por soborno de cada empresa limita el número de empresas en el mercado. Esta diferencia se debe a que en nuestra aproximación, siempre se considera un problema para el corrupto cuya formulación garantiza la existencia de solución óptima como consecuencia de la introducción de una restricción que tiene en cuenta los beneficios de cada empresa, ya que si el corrupto solicita una cantidad superior, la empresa optaría por abandonar el mercado.

El trabajo está dividido en secciones. En la segunda se enumeran las hipótesis de comportamiento de los productores y se obtiene la ecuación en diferencias que guía la evolución del tamaño del mercado. En la tercera, se determina el pago de soborno de cada empresa dependiendo de las medidas adoptadas por el gobierno en la lucha contra la corrupción. En esta sección se proponen y resuelven distintos ejemplos para mostrar el alcance de los resultados obtenidos. El trabajo termina con unas conclusiones.

2. Características generales del modelo

Consideremos un mercado en el que participan un número finito de productores de bienes o servicios idénticos o en el caso de ser diferentes, se supone son fuertemente sustitutivos. En el mercado podrían participar potencialmente un número ilimitado de empresas, aunque todas las empresas que en él participan se ven obligadas a pagar a una determinada cantidad a ciertos burócratas corruptos.

Supondremos que ninguno de los productores tiene capacidad para determinar el precio de su producción y todos ellos conocen las funciones de oferta y demanda idénticas para sus productos.

Dividamos el horizonte temporal en intervalos de la misma amplitud, meses, trimestres, semestres, años, etc.

Si la demanda de los productos en cualquier intervalo $(t-1, t]$ puede expresarse como una función lineal del precio $D(t) = a - bp(t)$, siendo a y b parámetros positivos y la oferta también puede ser expresada como una función lineal del precio, pero del periodo anterior, $S(t) = s + dp(t-1)$, siendo s y d positivos con $a > s$, tendremos que cuando el

mercado esté en equilibrio, oferta igual a demanda, el precio de equilibrio verificará la ecuación en diferencias lineal de primer orden:

$$p(t) = \frac{a-s}{b} - \frac{d}{b} p(t-1),$$

que admite un único estado de equilibrio dinámico $p^* = \frac{a-s}{b+d}$ positivo, ya que a precios nullos la función de demanda supera a la función de oferta.

Como la pendiente de la ecuación en diferencias que siguen los precios, en valor absoluto, coincide con el cociente de la pendiente de la función de demanda y de la función de oferta, dependiendo de si ese cociente es menor, igual o mayor a la unidad, el estado de equilibrio es asintóticamente estable ($b > d$), estable ($b = d$) o inestable ($d > b$). Por tanto, de acuerdo con la relación entre los parámetros b y d , la evolución de los precios de equilibrio del mercado hacia el estado de equilibrio dinámico, que también es un precio de equilibrio de mercado, sigue tres comportamientos diferentes como muestra el gráfico 1.

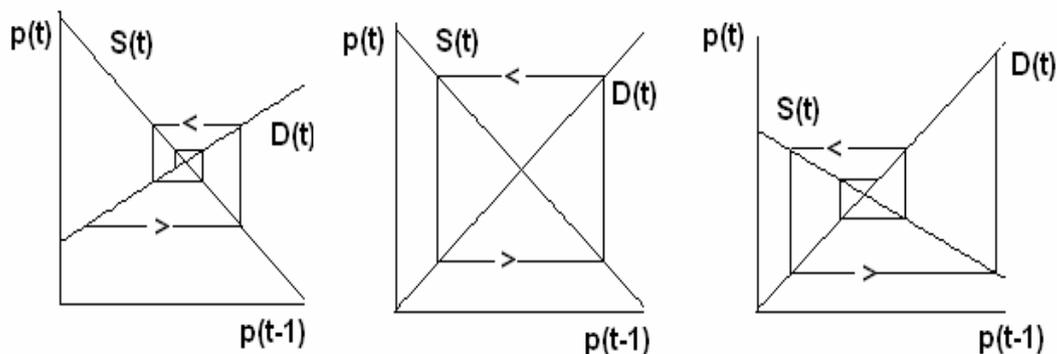


Gráfico 1: Evolución de los precios

Para comprobar de manera intuitiva la evolución de los precios de equilibrio que muestra el gráfico 1, rescribimos la ecuación en diferencias que siguen los precios de equilibrio en función del precio de equilibrio dinámico

$$p(t) - p^* = \frac{d}{b} (p^* - p(t-1)).$$

Por tanto, cuando el estado de equilibrio dinámico es asintóticamente estable⁴, independientemente del precio de equilibrio del mercado en el momento inicial, los precios de equilibrios convergen hacia el precio de equilibrio dinámico tomando valores superiores e inferiores al del precio de equilibrio dinámico; si el estado de equilibrio es estable, los precios oscilarán alrededor del estado de equilibrio dinámico repitiéndose el precio inicial cada dos periodos, mientras que si el estado de equilibrio es inestable, los precios de equilibrio de mercado se alejan del precio de equilibrio dinámico adoptando posiciones cada vez más alejadas del precio dinámico.

En cada intervalo $(t-1, t]$ en el que se divide el horizonte temporal del modelo, el precio de la producción puede obtenerse considerando la función inversa de la función de demanda $p(t) = \frac{a - D(t)}{b}$ y cada empresa que participa en el mercado, conociendo ese precio, determina su producción en el horizonte temporal del modelo tratando de maximizar los resultados en cada intervalo $\{\pi_i(t)\}_{i=1}^{\infty}$, donde el resultado de cada empresa se expresa

$$\pi_i(t) = X_i(t)p(t) - cX_i(t) - B(t) = X_i(t) \frac{a - \sum_{i=1}^{N(t)} X_i(t)}{b} - cX_i(t) - B(t),$$

siendo, $c > 0$ el coste unitario de producción, supuesto constante e idéntico para todas las empresas, $X_i(t)$ la producción de la empresa i en el intervalo, $N(t)$ el número de empresas que participan en el mercado en el intervalo y $B(t)$ el pago de cada empresa en el intervalo a los burócratas corruptos.

Si las empresas del mercado determinan su producción, en cada intervalo, siguiendo un equilibrio de Nash no cooperativo, donde el pago es su función de beneficios, las curvas de reacción de cada empresa permiten afirmar que la producción de cada empresa es idéntica y verifica

$$X_i(t) = b[p(t) - c] \quad i = 1, 2, \dots, N(t),$$

⁴ En particular, cuando el estado de equilibrio es asintóticamente estable la trayectoria que siguen los precios en su tendencia hacia el estado de equilibrio dinámico es conocida como el proceso de formación de la tela de araña.

donde sustituyendo el precio en función de la demanda y operando se encuentra que la producción de cada empresa satisface la relación

$$X_i(t) = \frac{a - cb}{1 + N(t)}, \quad i = 1, 2, \dots, N(t),$$

positiva si $a > cb$, condición que será impuesta de ahora en adelante.

La producción de cada empresa, por tanto, es independiente del pago realizado a los corruptos y coincide con el obtenido en el modelo del oligopolio de Cournot que establece la disminución progresiva de la producción de cada empresa cuando aumenta el tamaño del mercado, siendo la producción máxima la del monopolio, que también provoca el mayor precio. Por el contrario cuando el tamaño del mercado es suficientemente grande, los beneficios de las empresas tienden hacia cero.

Conociendo la producción de cada empresa, obtenemos la producción de todas las empresas que participan en el mercado durante el intervalo

$$\sum_{i=1}^{N(t)} X_i(t) = (a - cb) \frac{N(t)}{1 + N(t)} = D(t),$$

luego el porcentaje de participación de cada empresa en la producción del mercado en cualquier intervalo es:

$$\frac{X_i(t)}{\sum_{i=1}^{N(t)} X_i(t)} = \frac{1}{N(t)},$$

que es inversamente proporcional al tamaño del mercado en ese intervalo.

Como el precio puede ser determinado en función de la función inversa de la función de demanda, conociendo la expresión de la función de demanda, determinamos el precio en función del tamaño del mercado

$$p(t) = \frac{a}{b} - \frac{a - cb}{b} \frac{N(t)}{1 + N(t)}$$

positivo y función decreciente del número de empresas y creciente respecto al coste unitario de producción.

2.1 La evolución del tamaño del mercado

Como en cada intervalo, el equilibrio del mercado requiere que la oferta sea igual a la demanda, sustituyendo la última expresión encontrada en la ecuación en diferencias que verifican los precios de equilibrio, obtendremos que el número de empresas satisface una ecuación en diferencias, no lineal,

$$(a - cb) \frac{N(t+1)}{1 + N(t+1)} = s + \frac{da}{b} - \frac{d(a - cb)}{b} \frac{N(t)}{1 + N(t)},$$

que alternativamente puede expresarse

$$N(t+1) = \frac{s + \frac{da}{b} + (s + dc)N(t)}{a - cb - s - \frac{da}{b} + (a - s - c(b + d))N(t)} = f(N(t)),$$

que nos indica que el tamaño del mercado podría estar indeterminado si en un intervalo

$$N(t) = \frac{cb + s + \frac{da}{b} - a}{a - s - c(b + d)}.$$

Por el contrario, si en ningún intervalo, el tamaño del mercado toma el valor expresado previamente, la ecuación en diferencias no lineal para el tamaño del mercado determina que el número de empresas admite dos estados de equilibrio dinámicos $N_* = -1$, que carece de significado y

$$N^* = \frac{(sb + da)}{b(a - s - c(b + d))}$$

positivo si el coste unitario de producción verifica $0 < c < \frac{a - s}{b + d} < \frac{a}{b}$.

Observamos que la cuantía que determina el estado de equilibrio dinámico N^* crece si crece el coste unitario de producción, la oferta a precio nulo o el parámetro d . Decece con la demanda a precio nulo y su variación no está totalmente definida si varía el parámetro b .

Como $|f'(N^*)| = \frac{d}{b}$, el estado de equilibrio dinámico es asintóticamente estable si $d < b$, estable si $d = b$ e inestable si $b < d$. Por tanto, las características de estabilidad

asociadas al estado de equilibrio dinámico N^* coinciden con las obtenidas para el precio de equilibrio p^* .

El gráfico 2, muestra los dos estados de equilibrio que presenta la ecuación en diferencias $N(t+1) = f(N(t))$ y la asíntota de la función f .

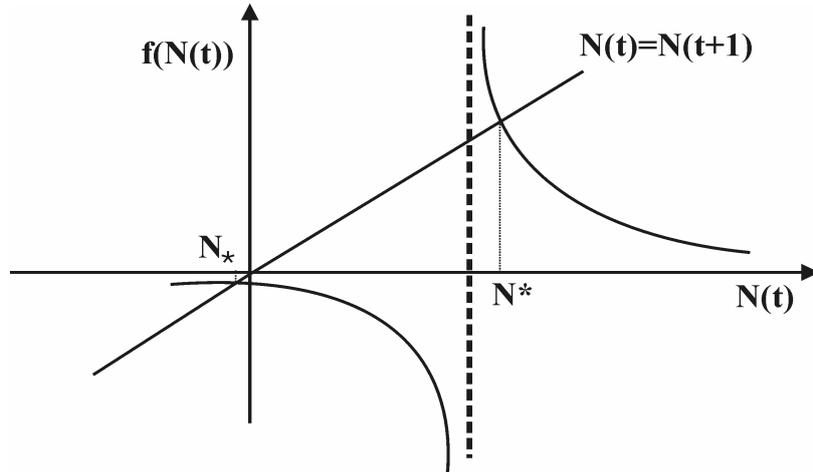


Gráfico 2: Estados de equilibrio para el tamaño del mercado

Observemos que la condición $0 < c < \frac{a-s}{b+d} = p^*$ que garantiza la existencia del estado de equilibrio dinámico N^* siempre se verifica, ya que desde los resultados alcanzados tenemos que los beneficios de cada empresa, eliminando el pago de soborno, alcanzan el valor

$$(p(t) - c)X_i(t) = (p(t) - c) \frac{a - cb}{1 + N(t)} = \frac{1}{b} \left(\frac{a - cb}{1 + N(t)} \right)^2 > 0,$$

y como la evolución de los precios puede expresarse en función de su precio de equilibrio, $p(t) - p^* = \frac{d}{b}(p^* - p(t-1))$, para cualquier intervalo tenemos que si $c = p^*$, y si el beneficio de cada empresa, eliminado el pago de soborno, $(p(t) - c)X_i(t)$ es positivo en un intervalo, en el siguiente es negativo, que no es posible; tampoco puede ser $c > p^*$, ya que entonces $(p(t) - p^*)X_i(t) > (p(t) - c)X_i(t) > 0 \Rightarrow p(t) > p^*$, en cualquier intervalo que no es cierto.

3. Análisis de la actuación del gobierno

En un mercado sin barreras de entrada ni de salida, donde se llevan a cabo idénticas producciones o producciones fuertemente sustitutivas, los resultados de las empresas deben tender a anularse en el largo plazo, ya que si una empresa presenta resultados positivos cualquier otro potencial competidor podría estar dispuesto a entrar en el mercado para producir un bien sustitutivo de aquel que produce la empresa con resultados positivos. En ese caso, podría ocurrir que se modificase la demanda del producto de la empresa con resultados positivos y, como consecuencia, la empresa podría verse obligada a abandonar el mercado.

Los resultados de las empresas en cada intervalo constituyen, por tanto, un elemento para decidir la permanencia o no en el mercado de las empresas, luego condicionarán la cuantía del soborno, ya que cualquier empresa abandonaría el mercado si sus resultados son negativos. Distinguimos tres casos para determinar la cuantía del pago a los corruptos dependiendo de que el gobierno no actúe contra la corrupción, el caso de que la corrupción sea perseguida y un régimen que podría denominarse más democrático, donde la actuación contra la corrupción es más intensiva.

3.1 El gobierno no lucha contra la corrupción

El pago de soborno de cada empresa, en el caso de que se permita actuar libremente a los corruptos, puede obtenerse en cada intervalo $(t-1, t]$ como aquel que anule sus resultados⁵, ya que será la cantidad máxima que puede asumir cada empresa para no tener que abandonar el mercado, pues si los corruptos requieren una cantidad superior, la empresa optaría por abandonar el mercado en vez de pagar el soborno y tampoco los corruptos requerirán una cantidad inferior al no temer ningún tipo de represalia por solicitarla.

Teniendo en cuenta los desarrollos de la sección anterior, encontramos

$$\pi_i(t) = 0 \Rightarrow B(t) = \frac{1}{b} \left(\frac{a - cb}{1 + N(t)} \right)^2$$

⁵ Peretto considera esta misma opción para determinar los gastos en investigación y desarrollo de las empresas de un mercado sin barreras de entrada y salida.

que es una función decreciente del número de empresas en el mercado y del coste unitario de producción, luego si el tamaño de mercado o el coste unitario aumentan, el pago de soborno de cada empresa disminuye; en particular, si el tamaño del mercado crece o el coste de producción tiende a $\frac{a}{b}$, el pago de soborno tiende a cero.

Desde un punto de vista dinámico, la expresión anterior muestra que la evolución de la cuantía del pago por soborno se adapta a la evolución del tamaño del mercado y si el estado de equilibrio del número de empresas es asintóticamente estable, el pago de soborno tenderá hacia un valor estacionario

$$B^* = \frac{1}{b} \left(\frac{a - cb}{1 + N^*} \right)^2 = b(p^* - c)^2,$$

y la renta alcanzada por los corruptos será $N^* B^* = (p^* - c) \frac{sb + da}{b + d}$.

Si el estado de equilibrio dinámico del número de empresas es estable, los corruptos sólo obtendrían dos posibles pagos de sobornos, mientras que la evolución de los pagos por soborno se iría alejando progresivamente de B^* si el estado de equilibrio dinámico del número de empresas es inestable.

3.2 El gobierno lucha contra la corrupción

Si se adopta una clara política de lucha contra la corrupción, habrá tareas de vigilancia por parte del gobierno para impedir el cobro de sobornos y los corruptos tendrán una cierta probabilidad de ser capturados por esa actividad lo que supondrá para ellos, al menos, la pérdida de las cantidades sobornadas.

Supongamos que los corruptos, si hay varios, actúan conjuntamente tomando las mismas decisiones ante cualquier situación, por lo que podemos considerar que sólo actúa un corrupto. Consideremos, además, que la probabilidad de que el corrupto sea capturado depende de la intensidad y frecuencia del cobro de sobornos, luego esta probabilidad será función del tamaño del mercado y de la cantidad sobornada a cada empresa. Denotemos por $1 - \Phi(B(t), r(t)N(t))$ la probabilidad de que el corrupto sea capturado en sus actividades de soborno durante el intervalo $(t-1, t]$, donde $B(t)$ es la cantidad que se soborna a cada empresa supuesta idéntica para cada empresa sobornada,

$r(t)$, con $0 \leq r(t) \leq 1$, es el porcentaje de productores que pagan soborno en el intervalo y $N(t)$ el tamaño del mercado.

Sea Φ de clase C^r , $r > 1$, y que mientras no halla lugar a confusión la dependencia respecto al tiempo de las variables será eliminada en la notación. La función de probabilidad Φ verificará condiciones adicionales al tener en cuenta que si no se soborna a ninguna empresa o el corrupto no recibe ninguna cantidad, la probabilidad de ser capturado es nula, luego $\Phi(B, rN) = 1$ si $B r = 0$. También es realista suponer que, la probabilidad de que el corrupto sea capturado crece si la cantidad sobornada a cada empresa crece, luego $\Phi'_B(B, rN) < 0$ y si el porcentaje de empresas sobornadas crece, también crece la probabilidad de capturar al corrupto, luego $\Phi'_2(B, rN) < 0$. Asimismo, por ser una función de probabilidad, $0 < \Phi(B, rN) \leq 1$, para cualquier $B > 0$ y para todo $r \in (0, 1]$.

Supondremos que si el corrupto es capturado, el gobierno rescata el soborno y él no recibe nada; mientras que, en el caso de no ser capturado recibe el pago del soborno, luego su ganancia esperada en la actividad de sobornar en cada intervalo y para todo el mercado es $\Phi(B, rN) B r N$.

Si el corrupto desea determinar la cantidad a sobornar a cada empresa y el porcentaje de empresas a sobornar en cualquier intervalo de modo que su ganancia esperada sea máxima, ha de resolver el problema no lineal con restricciones en desigualdad siguiente:

$$\max_{B, r} \Phi(B, rN) B r N, \quad s. a : 0 \leq B \leq \frac{1}{b} \frac{(a - cb)^2}{(1 + N)^2}, \quad 0 \leq r \leq 1,$$

donde la primera restricción indica que el pago de soborno de cada empresa será nulo o positivo. En este último caso, la cuantía de soborno será a lo sumo una cantidad que proporcione unos resultados nulos a cada empresa, ya que, en caso contrario, como ya ha sido mencionado las empresas abandonarían el mercado como consecuencia de obtener resultados negativos. La segunda restricción del problema indica que el porcentaje de empresas sobornadas será nulo o positivo, y en este último caso, no podrá superar a la unidad.

El problema no lineal formulado para el corrupto admite solución óptima global al verificarse el teorema de Weierstrass y corresponderá a una solución factible interior o bien pertenecerá a la frontera del conjunto de soluciones posibles. Considerando la función lagrangiana

$$L(B, r, \lambda, \mu) = \Phi(B, rN)B r N + \lambda \left[\frac{1}{b} \frac{(a - cb)^2}{(1 + N)^2} - B \right] + \mu[1 - r],$$

donde λ y μ son multiplicadores de Lagrange, las condiciones de primer orden de Kuhn-Tucker, determinan que el máximo global verificará:

$$L'_B(B, r, \lambda, \mu) = \Phi'_B(B, rN)B r N + \Phi(B, rN)r N - \lambda \leq 0, \quad BL'_B(B, r, \lambda, \mu) = 0, \quad B \geq 0,$$

$$L'_r(B, r, \lambda, \mu) = \Phi'_2(B, rN)B r N^2 + \Phi(B, rN)B N - \mu \leq 0, \quad r L'_r(B, r, \lambda, \mu) = 0, \quad r \geq 0,$$

$$\lambda \left[\frac{1}{b} \frac{(a - cb)^2}{(1 + N)^2} - B \right] = 0, \quad B \leq \frac{1}{b} \frac{(a - cb)^2}{(1 + N)^2}, \quad \lambda \geq 0,$$

$$\mu[1 - r] = 0, \quad r \leq 1, \quad \mu \geq 0.$$

Como para aquellas soluciones factibles que incluyan $B=0$ o $r=0$, el valor de la función objetivo se anula, ninguna solución posible que tenga al menos una componente nula puede corresponder a un máximo del problema, luego la solución óptima verificará

$$B > 0 \Rightarrow \Phi'_B(B, rN)B r N + \Phi(B, rN)r N = \lambda,$$

$$r > 0 \Rightarrow \Phi'_2(B, rN)B r N^2 + \Phi(B, rN)B N = \mu.$$

Como los multiplicadores de Lagrange pueden ser positivos o nulos, tenemos las siguientes posibilidades respecto a la solución óptima $(\bar{B}, N\bar{r})$ del problema en cada intervalo de tiempo:

$$1. \quad \lambda = 0, \mu = 0 \Rightarrow \bar{B} = -\frac{\Phi(\bar{B}, N\bar{r})}{\Phi'_B(\bar{B}, N\bar{r})} \leq \frac{1}{b} \frac{(a - cb)^2}{(1 + N)^2}, \quad N\bar{r} = -\frac{\Phi(\bar{B}, N\bar{r})}{\Phi'_2(\bar{B}, N\bar{r})} \leq N,$$

$$2. \quad \lambda > 0, \mu = 0 \Rightarrow \bar{B} = \frac{1}{b} \frac{(a - cb)^2}{(1 + N)^2} < -\frac{\Phi(\bar{B}, N\bar{r})}{\Phi'_B(\bar{B}, N\bar{r})}, \quad N\bar{r} = -\frac{\Phi(\bar{B}, N\bar{r})}{\Phi'_2(\bar{B}, N\bar{r})} \leq N,$$

$$3. \quad \lambda = 0, \mu > 0 \Rightarrow \bar{r} = 1, \quad \bar{B} = -\frac{\Phi(\bar{B}, N)}{\Phi'_B(\bar{B}, N)} \leq \frac{1}{b} \frac{(a - cb)^2}{(1 + N)^2}, \quad N < -\frac{\Phi(\bar{B}, N)}{\Phi'_2(\bar{B}, N)},$$

$$4. \quad \lambda > 0, \mu > 0 \Rightarrow \bar{r} = 1, \quad \bar{B} = \frac{1}{b} \frac{(a - cb)^2}{(1 + N)^2} < -\frac{\Phi(\bar{B}, N)}{\Phi'_B(\bar{B}, N)}, \quad N < -\frac{\Phi(\bar{B}, N)}{\Phi'_2(\bar{B}, N)}.$$

La siguiente proposición establece una condición suficiente que garantiza que no es óptimo imponer sobornos sólo a una parte del mercado. Esta posible solución óptima se elimina del problema del corrupto al considerar que el modelo propuesto no aprecia características específicas en alguna de las empresas participantes del mercado que permita distinguirla de las demás. Entonces, el considerar que unas empresas pagan un soborno y las demás no, no tendría cabida dentro del contexto en el que está planteado el modelo.

Para eliminar esta posible solución óptima observamos, desde las condiciones de primer orden, que si fuera óptimo imponer sobornos sólo a una parte del mercado, el multiplicador asociado al tamaño del mercado en el problema del corrupto sería nulo y la fracción del mercado pagando los sobornos vendría determinada por medio de un punto crítico de la función objetivo considerada exclusivamente como una función de la variable r . Por tanto, si la función objetivo del problema carece de puntos críticos respecto a la proporción r , no es óptimo imponer sobornos sólo a una parte del mercado. Basándonos en este razonamiento, se enuncia la siguiente proposición.

Proposición 1 Si para cualquier cuantía de soborno y tamaño de mercado, la función objetivo del problema

$$\max_{B,r} \Phi(B,rN)B r N, \quad s. a: 0 \leq B \leq \frac{1}{b} \frac{(a-cb)^2}{(1+N)^2}, \quad 0 \leq r \leq 1,$$

es estrictamente creciente en $r \in (0,1)$, $\bar{r} = 1$ forma parte de las componentes de la solución óptima del problema.

También si la función objetivo del problema, considerada como una función de la variable r , es cuasiconvexa⁶ en $(0,1]$, $\bar{r} = 1$ forma parte de las componentes de la solución óptima del problema, ya que un máximo global de una función cuasiconvexa no constante se encuentra en la frontera del conjunto convexo de soluciones posibles.

Si se soborna a todo el mercado, las condiciones necesarias de primer orden para el problema del corrupto determinan que la solución óptima respecto a la cuantía óptima del soborno para cada tamaño del mercado N , corresponde a la solución factible

⁶ Una función $f : C \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$, siendo C convexo y no vacío, es cuasiconvexa si verifica $f(\alpha \bar{x}_1 + (1-\alpha)\bar{x}_2) \leq \max\{f(\bar{x}_1), f(\bar{x}_2)\}$ para todo $x_1, x_2 \in C$ y $\alpha \in (0,1)$.

$B = -\frac{\Phi(B, N)}{\Phi'_B(B, N)}$, que puede ser interior o frontera del conjunto factible, pero siempre

perteneciendo a la curva de nivel cero de la función g definida por la expresión

$g(B) = \Phi(B, N) + B\Phi'_B(B, N)$ o al punto frontera $B = \frac{1}{b} \frac{(a - cb)^2}{(1 + N)^2}$, perteneciente a una

curva de nivel positivo de g . Observemos que este último valor coincide con el pago por soborno efectuado por las empresas cuando no se toma ninguna medida para luchar contra la corrupción.

La siguiente proposición determina condiciones suficientes de óptimo para las cuantías de soborno.

Proposición 2: Si dado un tamaño de mercado N , la función objetivo del problema

$$\max_B \Phi(B, N)B N, \quad s. a : 0 \leq B \leq \frac{1}{b} \frac{(a - cb)^2}{(1 + N)^2},$$

es pseudocóncava⁷ y verifica

1. $-\frac{\Phi(\bar{B}, N)}{\Phi'_B(\bar{B}, N)} = \bar{B} \leq \frac{1}{b} \frac{(a - cb)^2}{(1 + N)^2}$, entonces \bar{B} es la cuantía de soborno óptima.
2. $\bar{B} = \frac{1}{b} \frac{(a - cb)^2}{(1 + N)^2} < \min \left\{ B \in (0, \bar{B}) : B = -\frac{\Phi(B, N)}{\Phi'_B(B, N)} \right\}$, entonces \bar{B} es la cuantía

de soborno óptima.

Para comprobar la proposición, observemos para el primer apartado que \bar{B} es máximo global del programa irrestricto, $\max \Phi(B, N)B N$, por ser un punto crítico de la función objetivo que es una función pseudocóncava. Entonces, también es máximo global del programa del agente corrupto al ser solución factible de ese problema. Con respecto al segundo apartado, si la solución factible en la frontera \bar{B} no es máximo global del programa del agente corrupto, que admite solución óptima por verificarse el teorema de Weierstrass, otra solución factible $B_1 \neq \bar{B}$ será la solución óptima. En este caso, el

⁷ Una función diferenciable $f : C \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$, siendo C abierto y no vacío, es pseudocóncava si verifica $f(\bar{x}_1) < f(\bar{x}_2) \Rightarrow (\bar{x}_2 - \bar{x}_1)^t \nabla f(\bar{x}_1) > 0$ para todo $x_1, x_2 \in C$.

máximo global B_1 verificará las condiciones de Kuhn-Tucker, luego será un punto crítico de la función objetivo en contra de la hipótesis que mantiene este apartado.

En el siguiente apartado se consideran varios ejemplos que presentan distintas soluciones óptimas de acuerdo con la proposición 2.

3.2.1 Tipos de soluciones para el problema del corrupto

El problema formulado para el corrupto siempre admite solución óptima aunque si corresponde a un punto en la frontera, entonces las actuaciones encaminadas a luchar contra la corrupción no provocarían una variación de las cuantías de soborno respecto al caso de que no se tomen medidas. Esta situación ocurre si la probabilidad de no ser capturado viene determinada por la función fraccionada $\Phi(B, N) = \frac{1}{1 + BN}$, ya que

$$\Phi'_B(B, N) = -\frac{N}{(1 + BN)^2} \Rightarrow \Phi(B, N) + B\Phi'_B(B, N) = \frac{1}{(1 + BN)^2} \neq 0,$$

por tanto, la solución óptima del problema del corrupto, cuya función objetivo es pseudocóncava respecto a la cuantía de soborno y estrictamente creciente respecto a la proporción del mercado r , es sobornar una cuantía $\bar{B} = \frac{1}{b} \frac{(a - cb)^2}{(1 + N)^2}$ como en el caso de que el gobierno no tome medidas contra la corrupción.

Por otro lado, si podemos aplicar el teorema de las funciones implícitas, sobre la curva de nivel cero de la función que se obtiene al derivar la función objetivo del problema del corrupto respecto a la cuantía de soborno, resulta más sencillo caracterizar la solución óptima del problema teniendo en cuenta la proposición 2. En tal situación, si el tamaño del mercado es N_1 y existe $B_1 = -\frac{\Phi(B_1, N_1)}{\Phi'_B(B_1, N_1)}$, esto es, el par (B_1, N_1) pertenece a la curva de nivel cero de la función $\Phi(B, N) + B\Phi'_B(B, N)$, tal que adicionalmente se verifica la condición $B_1\Phi''_{BB}(B_1, N_1) + 2\Phi'_B(B_1, N_1) \neq 0$, por el teorema de las funciones implícitas, existirá una única función $B = f(N)$ definida en un entorno $B(N_1, \delta)$ tal que si $N \in B(N_1, \delta)$, tendremos $\Phi(f(N), N) + f(N)\Phi'_B(f(N), N) = 0$.

Luego si la probabilidad de no ser capturado permite que se verifica el teorema de las funciones implícitas sobre la curva de nivel cero de la función $\Phi(B, N) + B\Phi'_B(B, N)$ y

si $f(N) \leq \frac{1}{b} \frac{(a-cb)^2}{(1+N)^2}$, entonces $B = f(N)$ determinará la cuantía de soborno correspondiente a la solución óptima para el tamaño de mercado N y si se verifica la condición $\frac{1}{b} \frac{(a-cb)^2}{(1+N)^2} < f(N)$, el problema del corrupto admite como máximo global el punto frontera del conjunto de soluciones posibles.

Un ejemplo donde existe solución óptima interior para la cuantía de soborno se obtiene considerando que la probabilidad de no ser capturado está dada por la función

$$\Phi(B, N) = \frac{1}{1 + B^2 N^{\frac{1}{2}}}, \text{ que garantiza que la función objetivo del problema del agente}$$

corrupto es pseudocóncava respecto a la cuantía del soborno y estrictamente creciente respecto a la proporción r . Tenemos

$$\Phi(B, N) = \frac{1}{1 + B^2 N^{\frac{1}{2}}} \Rightarrow \Phi(B, N) + B\Phi'_B(B, N) = \frac{1 - B^2 N^{\frac{1}{2}}}{(1 + B^2 N^{\frac{1}{2}})^2}$$

y la última expresión se anula para $B = N^{-\frac{1}{4}}$. Si superponemos las representaciones gráficas (gráfico 3) de esta función y la frontera del conjunto de soluciones posibles,

$$B = \frac{k}{(1+N)^2}, \text{ donde para simplificar la notación } k = \frac{(a-cb)^2}{b}, \text{ encontramos que si el}$$

tamaño del mercado es inferior o igual a N_1 , la solución óptima está determinada por la curva de nivel cero de la función derivada de la función objetivo y, por tanto, en estos casos las medidas del gobierno provocan la disminución de las cuantías de los pagos de soborno. Cuando el tamaño de mercado supere a N_1 , las cuantías de los pagos no se ven modificadas por las medidas gubernamentales, ya que la solución óptima se encuentra en curvas de nivel positivas y en la frontera del conjunto de oportunidades. En particular, y como consecuencia de los valores proporcionados a los parámetros del modelo para la representación gráfica de la frontera del conjunto de soluciones posibles que aseguran que el estado de equilibrio dinámico del número de empresas es asintóticamente estable, cuando el mercado tienda hacia ese valor, las cuantías no se ven modificadas por las políticas gubernamentales contra la corrupción.

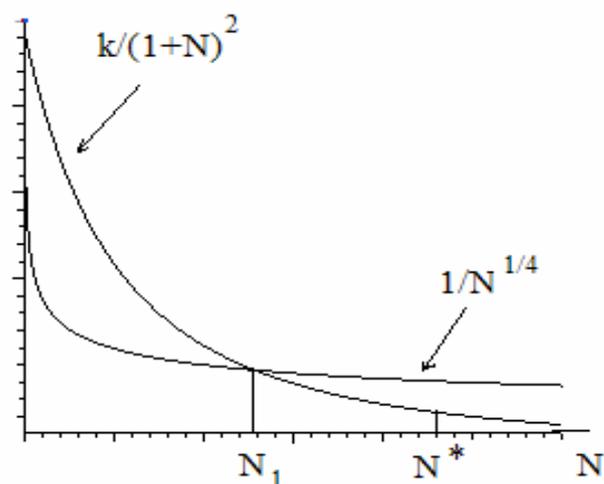


Gráfico 3: Variaciones de las cuantías de soborno

3.3 El gobierno endurece la lucha contra la corrupción

La intensificación de las medidas para luchar contra la corrupción significará que la probabilidad de que el burócrata corrupto sea atrapado crece ante el mismo soborno y el mismo tamaño del mercado. Por tanto si Φ^2 denota la probabilidad de no ser atrapado manteniendo un régimen moderado de lucha contra la corrupción y Φ^1 es esa misma probabilidad pero con medidas adicionales, se verificará

$$1 - \Phi^2(B, r N) < 1 - \Phi^1(B, r N) \Rightarrow \Phi^1(B, r N) < \Phi^2(B, r N).$$

En este supuesto, el corrupto puede plantearse el mismo problema que fue analizado en la sección anterior pero ahora con la función Φ^1 . Si mantenemos para las funciones las hipótesis que fueron consideradas en las proposiciones 2 y 3, el corrupto encontrará la misma variedad de soluciones óptimas. No obstante, en la nueva situación, la solución óptima podría implicar un valor inferior de pago de soborno, en el caso de que el óptimo no se alcance en el contorno.

Suponiendo que no es óptimo sobornar a una parte del mercado, la siguiente proposición establece una condición suficiente para que el pago de soborno sea inferior cuando la probabilidad de ser capturado aumenta.

Proposición 3 Si para un tamaño de mercado N existen $B^1 = -\frac{\Phi^1(B^1, N)}{\Phi^1_B(B^1, N)}$ y

$B^2 = -\frac{\Phi^2(B^2, N)}{\Phi^2_B(B^2, N)}$, con $\Phi^1(B, rN) < \Phi^2(B, rN)$ y $\Phi^1_B(B^1, N) \leq \Phi^2_B(B^2, N)$, entonces

$B^1 < B^2$.

Para probar la condición, observemos que si $\Phi^1_B(B^1, N) \leq \Phi^2_B(B^2, N)$, entonces

$$\frac{-\Phi^2_B(B^2, N)}{-\Phi^1_B(B^1, N)} \leq 1 \Rightarrow \frac{\Phi^2(B^2, N)}{B^2} \frac{B^1}{\Phi^1(B^1, N)} \leq 1 \Rightarrow \frac{\Phi^2(B^2, N)}{\Phi^1(B^1, N)} \leq \frac{B^2}{B^1}$$

y si $B^2 \leq B^1$, desde la última relación tendremos $\Phi^2(B^2, N) \leq \Phi^1(B^1, N)$. Por ser Φ^1 decreciente respecto a su variable B , también $\Phi^1(B^1, N) \leq \Phi^1(B^2, N)$ y como $\Phi^1(B, N) < \Phi^2(B, N)$ para cualquier pago de soborno, también $\Phi^1(B^2, N) < \Phi^2(B^2, N)$ y, por tanto, $\Phi^1(B^1, N) \leq \Phi^1(B^2, N) < \Phi^2(B^2, N)$ que contradice el resultado anterior.

Luego, si se verifica la proposición 3, y la solución óptima del problema del corrupto se encuentra sobre la curva de nivel cero, la cuantía óptima de soborno disminuye al aumentar la probabilidad de ser capturado. No obstante, puede ocurrir que para unas medidas de lucha contra la corrupción existan, para el problema del corrupto, soluciones óptimas en la frontera y a medida que se endurece la lucha contra la corrupción las soluciones óptimas pasen a ser interiores, como ocurre en el siguiente ejemplo.

Si para cualquier tamaño de mercado y cuantía de soborno, las probabilidades de que el burócrata corrupto no sea capturado ante dos políticas diferentes verifican:

$$\Phi^1(B, N) = \frac{1}{1 + 2B^2N^{\frac{1}{2}}} < \frac{1}{1 + B^2N^{\frac{1}{2}}} = \Phi^2(B, N),$$

la función implícita, definida desde la relación $\Phi^1(B, N) + B\Phi^1_B(B, N) = 0$, es

$B = \frac{1}{2}N^{\frac{1}{4}}$ y la curva de nivel cero para la derivada de la función $B N \Phi^2(B, N)$

determina la función implícita $B = N^{\frac{1}{4}}$. El gráfico 4 muestra las representaciones de las funciones Φ^1 y Φ^2 como funciones de B y las funciones implícitas anteriores para

cada función de probabilidad junto con función que determina la frontera del conjunto admisible.

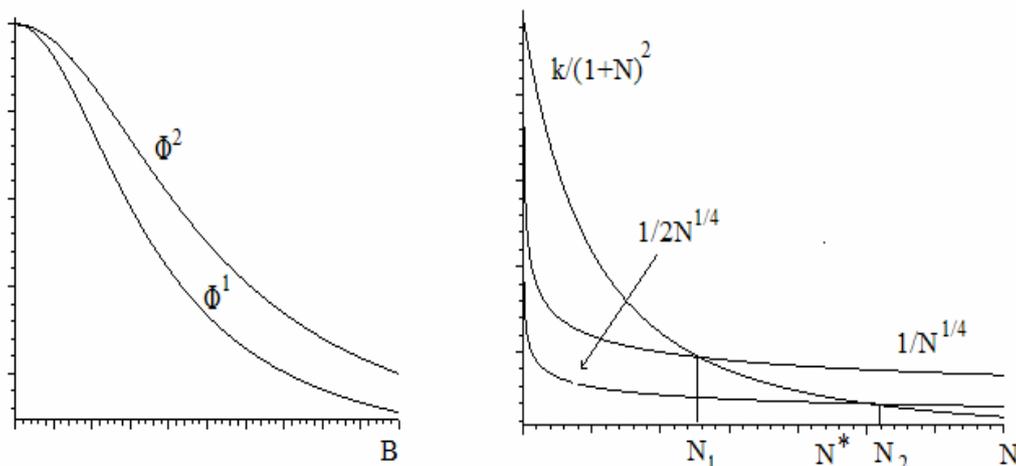


Gráfico 4: La lucha contra la corrupción disminuye la cuantía de soborno

Observamos en el gráfico de la derecha que manteniendo una política contra la corrupción representada por la función Φ^2 , la cuantía de soborno es menor que la que determina la frontera si el tamaño de mercado es inferior a N_1 como también recogía el gráfico 3; si la política contra la corrupción está representada por la función Φ^1 , la solución óptima del problema del corrupto determina una cuantía de soborno inferior a la recogida por Φ^2 siempre que el tamaño de mercado sea inferior o igual a N_2 , a partir de ese punto la política de no tomar medidas en la lucha contra la corrupción proporciona la menor cuantía de soborno. Observemos también que ahora el estado de equilibrio dinámico del número de empresas pertenece al intervalo $[N_1, N_2]$, que implica menores pagos de sobornos cuando el tamaño del mercado tienda hacia él si se adoptan medidas intensivas.

4. Conclusiones

En este trabajo se ha desarrollado un modelo para comprobar que la lucha contra la intensidad de la corrupción en los mercados de bienes puede ser efectiva si se adoptan medidas para combatirla. Con este objetivo se han estudiado tres casos que modelizan tres políticas gubernamentales en la lucha contra la corrupción. Si la política es no combatirla todas las empresas del mercado pagarán una cantidad cuya evolución dinámica sigue el comportamiento del tamaño del mercado. Si el gobierno adopta

alguna medida para combatir la corrupción, la cuantía de soborno puede disminuir, en cuyo caso, la cuantía de soborno depende de las medidas públicas adoptadas más que de la evolución dinámica del mercado. Similar situación ocurre, si se intensifican las medidas, un régimen más democrático en palabras de Emerson, donde bajo ciertas condiciones la cuantía de soborno disminuirá respecto a un control más suave de la corrupción.

Como la lucha contra la corrupción no es una tarea sencilla, las medidas para combatirla requieren cierto esfuerzo y este hecho es patente en la modelización teórica que presentamos, expresados a través de los requerimientos para las funciones que recogen las distintas políticas públicas estudiadas.

El trabajo podría proporcionar nuevas aportaciones si se considera en el mismo planteamiento que los costes de producción de las empresas pueden ser distintos. Esta característica permitiría analizar si además de disminuir la intensidad de la corrupción, también disminuye su frecuencia.

Bibliografía

Ades, A. y Di Tella, R. (1999): “Rents, Competition, and Corruption”, *The American Economic Review*, 89, pp. 982-993.

Blackburn, K., Bose N., y Emranul M. (2006): “The incidence and persistence of corruption in economic development”, *Journal of Economic Dynamics and Control*, 30 (12), pp. 2447-2467.

Bliss, C. y Di Tella, R. (1997): “Does Competition Kill Corruption?”, *The Journal of Political Economy*, 105, pp. 1001-1023.

Emerson, P.M. (2006): “Corruption, competition and democracy”, *Journal of Development Economics*, 81, pp. 193-212.

Mauro, P. (1997): “Why worry about corruption?”, *International Monetary Fund*, Series: Economic Issues (2), pages 15.

Peretto, P.F. (1999): “Industrial development, technological change and long-run growth”, *Journal of Development Economics*, 59, pp. 389-417.

Pope, J. (2000): *Confronting corruption: The elements of a national integrity system*, Transparency International (TI), Malaysia.

Varian, H.R. (1998): *Análisis Microeconómico*, 3ª Edición Antoni Bosch Editor, Barcelona.

DIFERENCIA MEDIA DE GINI, DESIGUALDAD, SOLAPAMIENTO Y PROXIMIDAD

CALLEALTA BARROSO, Fco. Javier
e-mail: franciscoj.callealta@uah.es

Departamento de Estadística, Estructura Económica y O.E.I.
UNIVERSIDAD DE ALCALÁ

Área temática: Métodos Cuantitativos.

Resumen

La proximidad entre distribuciones de tamaños de magnitudes económicas resultantes de observaciones realizadas sobre elementos poblacionales puede ser estudiada, tanto desde una perspectiva estática (semejanzas y diferencias en un determinado corte temporal) como desde una perspectiva dinámica (convergencia o divergencia a lo largo del tiempo), a partir del análisis de las proyecciones de dichas distribuciones sobre un espacio factorial euclídeo óptimamente reducido y obtenido mediante técnicas de escalamiento multidimensional (Callealta, 2005 y 2006). Para ello es esencial partir de una medida de proximidad que refleje adecuadamente las diferencias existentes entre aquéllas, no existiendo una única posibilidad y siendo la elección de ésta determinante en la obtención de los resultados y para la interpretación de sus significados. Partiendo de la Diferencia Media de Gini, ampliamente empleada para el análisis comparativo de poblaciones desde la óptica de la desigualdad, puede derivarse razonablemente también una medida de la proximidad entre las distribuciones comparadas. Consecuentemente con la hipótesis de que la Diferencia Media de Gini incorpora información sobre diversos aspectos de las distribuciones comparadas, este trabajo trata de profundizar sobre el contenido informativo de esta medida y de descomponerla convenientemente de acuerdo con los distintos aspectos de que informa.

Palabras clave: Diferencia Media de Gini, Desigualdad, Disparidad, Solapamiento, Proximidad.

Abstract

The increasing number of economic surveys is causing a great interest about size distributions of a large number of economic magnitudes. Analysis of proximity between such kinds of size distributions can be conducted from both static point of view (i.e. distance analysis) and dynamic point of view (i.e. convergence analysis). In both cases, multidimensional scaling has shown to be a very useful technique that provides the possibility of projecting those size distributions into certain points of an optimally reduced Euclidean space, reducing the problem of comparing size distributions to that of comparing just points. However, correct application of multidimensional scaling techniques requires computing a dissimilarity matrix according to an adequate proximity measure that reflects precisely the proximity concept adapted to the problem. Election of such a proximity measure is not a trivial problem and analysis results could be highly affected by it. A quite adequate proximity measure can be deduced using the well-known Gini mean difference, when applied to a particular decomposition of compared populations (Callealta, 2005 y 2006). Consequently with the hypothesis of Gini mean difference integrates information not only about inequality but also about other different aspects of compared distributions, this work deals with decomposition of this measure and interpretation of its components.

Key words: (max. 6). Gini mean difference, inequality, disparity, proximity, overlapping

1 Introducción.

El análisis dinámico del grado de acercamiento o convergencia entre distribuciones de los tamaños de una cierta magnitud económica en diversas poblaciones puede realizarse con la ayuda de técnicas de escalamiento multidimensional (Callealta, 2005 y 2006). De esta forma, el análisis se basaría fundamentalmente sobre la forma y evolución de las trayectorias temporales visualizadas en un cierto espacio factorial de dimensiones óptimamente reducido, y construidas de acuerdo con la información contenida en una matriz de las proximidades entre aquéllas, previamente calculadas de forma adecuada.

En aquéllos trabajos nos centrábamos sobre el objetivo de medir las disparidades o diferencias que pueden presentar diferentes distribuciones personales de ingreso, con la finalidad de emplear tal medida como indicador de convergencia entre éstas. Así, proponíamos una nueva medida de disimilaridad que se manifestaba útil para ello.

Ahora, consecuentemente con estos planteamientos, en este trabajo nos proponemos profundizar sobre la relación existente entre la citada medida de disimilaridad propuesta y la diferencia media de Gini calculada entre las poblaciones comparadas, pensando que en ésta se mezclan, presumiblemente, informaciones que tiene que ver no sólo con el grado de desigualdad relativas a éstas, sino también sobre el grado de proximidad y sobre el grado de solapamiento existente entre las mismas.

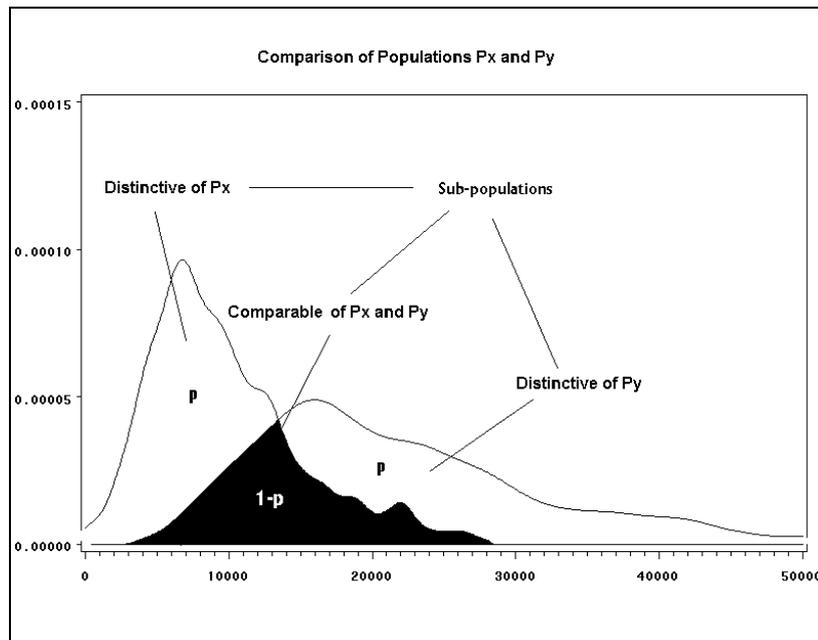
2 Propuestas de descomposición de las poblaciones comparadas y de una medida de disimilaridad entre éstas.

2.1 Descomposición de las poblaciones comparadas

Siguiendo lo expuesto en Callealta (2005 y 2006), supongamos que son X e Y un par de variables aleatorias independientes y absolutamente continuas con funciones de densidad respectivas $f_X(x)$ y $f_Y(y)$ definidas sobre \mathfrak{R} , representativas de los comportamientos probabilísticos de sendas poblaciones P_X y P_Y , para una magnitud observada.

Al comparar estas poblaciones, representadas por sus respectivas distribuciones de probabilidad (densidades de probabilidad), podemos observar que se pueden extraer dos subpoblaciones disjuntas de cada una de ellas (P_{X^*} y P_C de P_X ; P_{Y^*} y $P_{C'}$ de P_Y) de forma que dos de ellas (P_C y $P_{C'}$) son perfectamente comparables y equiparables por sus valores, pudiéndose diferenciar claramente de las otras dos subpoblaciones (P_{X^*} y P_{Y^*} , una de cada población de partida), perfectamente separadas por estar compuestas por aquellos elementos de cada población para los que, una vez apartadas las subpoblaciones P_C y $P_{C'}$, no existe ningún elemento más de la otra población que tome su mismo valor (subpoblaciones de elementos “diferenciales” o “no equiparables”), tal y como se refleja de forma intuitiva en el gráfico 1.

Gráfico 1



Así,

$$P_X = P_{X^*} \cup P_C \quad , \quad P_{X^*} \cap P_C = \phi \quad y \quad P_Y = P_{Y^*} \cup P_{C'} \quad , \quad P_{Y^*} \cap P_{C'} = \phi$$

siendo las variables C y C' , representantes respectivamente de los comportamientos de las poblaciones equiparables P_C y $P_{C'}$, variables idénticamente distribuidas.

De acuerdo con esta apreciación, para cualquier par de variables absolutamente continuas X e Y objetos de comparación, podemos definir las siguientes variables auxiliares:

- a) Variable C representativa del comportamiento de las subpoblaciones “equiparables”, P_C y P_{C^*} ; o subpoblaciones de P_X y P_Y para las que se pueden encontrar sendas subpoblaciones de la otra población con características (valores de la variable) similares (comportamiento común para ambas variables X e Y relacionado con la parte sombreada en el gráfico 1), y cuya función de densidad vendría dada por:

$$f_C(t) = \frac{\text{Min}(f_X(t), f_Y(t))}{1-p} = \begin{cases} \frac{f_X(t)}{1-p} & , \text{ si } f_X(t) \leq f_Y(t) \\ \frac{f_Y(t)}{1-p} & , \text{ si } f_Y(t) \leq f_X(t) \end{cases}$$

siendo $1-p$ la proporción de cada una de las poblaciones de P_X y P_Y que es equiparable con una misma proporción de la otra:

$$1-p = \int_{\{y|f_X(y) > f_Y(y)\}} f_Y(y) dy + \int_{\{x|f_Y(x) > f_X(x)\}} f_X(x) dx$$

- b) Variable X^* representativa del comportamiento de la subpoblación diferenciada para X, $P_{X^*} = P_X - P_C$; o subpoblación de P_X complementaria a la seleccionada como “equiparable” a otra de P_Y , con características específicas de X y para la que, una vez extraídas las poblaciones “equiparables”, no es posible encontrar otros elementos de P_Y equiparables con los propios (relacionado con la parte izquierda no sombreada en el gráfico 1). Su función de densidad vendría dada por:

$$f_{X^*}(x) = \frac{f_X(x) - f_Y(x)}{p} \cdot I\{f_X(x) > f_Y(x)\} = \begin{cases} \frac{f_X(x) - f_Y(x)}{p} & , \text{ si } f_X(x) > f_Y(x) \\ 0 & , \text{ si } f_X(x) \leq f_Y(x) \end{cases}$$

donde ahora p representa la proporción de la población X que no es equiparable a ninguna de la Y , y donde $I\{\cdot\}$ es la función indicadora de la proposición entre llaves¹.

- c) Variable Y^* representativa del comportamiento de la Subpoblación diferenciada para Y , $P_{Y^*} = P_Y - P_C$; o subpoblación de P_Y complementaria a la seleccionada como “equiparable” a otra de P_X , con características específicas de Y y para la que, una vez extraídas las poblaciones “equiparables”, no es posible encontrar otros elementos de P_X equiparables con los propios (relacionado con la parte derecha no sombreada en el gráfico 1). Su función de densidad vendría dada por:

$$f_{Y^*}(y) = \frac{f_Y(y) - f_X(y)}{p} \cdot I\{f_Y(y) > f_X(y)\} = \begin{cases} \frac{f_Y(y) - f_X(y)}{p} & , si \quad f_Y(y) > f_X(y) \\ 0 & , si \quad f_Y(y) \leq f_X(y) \end{cases}$$

donde ahora p representa la proporción de la población Y que no es equiparable a ninguna de la X .

Con estas definiciones, la variable original X puede expresarse como una mixtura de las variables X^* y C con pesos p y $1-p$ respectivamente y, de forma análoga, la variable original Y puede expresarse como una mixtura de las variables Y^* y C' con pesos p y $1-p$ respectivamente. Consecuentemente, pueden expresarse sus funciones de densidad y de distribución como sigue:

$$f_X(x) = (1-p) \cdot f_C(x) + p \cdot f_{X^*}(x) \quad ; \quad f_Y(y) = (1-p) \cdot f_C(y) + p \cdot f_{Y^*}(y)$$

$$F_X(x) = (1-p) \cdot F_C(x) + p \cdot F_{X^*}(x) \quad ; \quad F_Y(y) = (1-p) \cdot F_C(y) + p \cdot F_{Y^*}(y)$$

donde la variable C informa de las características de las subpoblaciones P_C y P_C' seleccionadas como equiparables en ambas poblaciones X e Y , de proporción $1-p$, mientras que las variables X^* e Y^* informan de las subpoblaciones P_{X^*} y P_{Y^*} , de

¹ La función indicadora para una proposición A vale $I\{A\} = \begin{cases} 1 & si \quad A \text{ es cierto} \\ 0 & si \quad A \text{ es falso} \end{cases}$

proporciones p , específicas, diferenciadas o no equiparables de sendas poblaciones comparadas X e Y .

Algunas propiedades de estas distribuciones, asociadas a las descomposiciones poblacionales propuestas, son:

- a) La proporción de población diferenciada es: $p = \frac{1}{2} \cdot \int_{-\infty}^{+\infty} |f_X(t) - f_Y(t)| \cdot dt$
- b) Las medias de estas distribuciones auxiliares descomponen las medias de las distribuciones originales, informando de las contribuciones a éstas de cada subpoblación equiparable y específica, en función de sus pesos en las correspondiente mixturas:

$$E[X] = pE[X^*] + (1-p)E[C] \quad , \quad E[Y] = pE[Y^*] + (1-p)E[C]$$

Consecuentemente, otras propiedades derivadas son las siguientes:

- c) $(1-p)E[C] = E[X] - pE[X^*] = E[Y] - pE[Y^*]$
- d) $E[X] - E[Y] = p(E[X^*] - E[Y^*])$
- e) $E[X] + E[Y] = p(E[X^*] + E[Y^*]) + 2(1-p) \cdot E[C]$

2.2 La medida de disimilaridad asociada

A partir de esta propuesta de descomposición de las distribuciones comparadas, en Callealta (2005 y 2006) se propone, como medida de la disimilaridad entre dos distribuciones X e Y , la diferencia media de Gini, $\Delta_{X^*Y^*}$, calculado entre las distribuciones asociadas X^* e Y^* de sus respectivas subpoblaciones diferenciadas, o no equiparables, según las definiciones precedentes, ponderada por el producto de las proporciones que estas representan de las poblaciones originales X e Y^2 .

² Obsérvese que la introducción de la ponderación responde, en primer lugar, a nuestro interés de hacer que la medida sea lo más intuitiva posible (conduce a la evaluación directa de diferencias entre las subpoblaciones diferenciadas), a la vez que, en segundo lugar, a introducir en la expresión el efecto de la magnitud relativa de las poblaciones que son realmente específicas o no equiparables (peso de las proporciones de población no equiparables).

$$d(X, Y) = p^2 \cdot E\left[|Y^* - X^*|\right] = p^2 \cdot \Delta_{X^*Y^*} =$$

$$= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} |y - x| \cdot (f_X(x) - f_Y(x)) I\{f_X(x) > f_Y(x)\} \cdot (f_Y(y) - f_X(y)) I\{f_Y(y) > f_X(y)\} dx \cdot dy$$

Entre algunas de sus propiedades, destacamos las siguientes:

- a) $d(X, Y) = 0 \Leftrightarrow X = Y$ c.s.
- b) La medida $d(X, Y)$ crece con la diferencia entre X e Y; es decir, crece tanto con el aumento de la proporción que representan las subpoblaciones diferenciadas o no equiparables de X e Y, como con el aumento de la separación entre éstas.
- c) La medida $d(X, Y)$ es simétrica: $d(X, Y) = p^2 \cdot E\left[|X^* - Y^*|\right] = d(Y, X)$
- d) $0 \leq d(X, Y) \leq \Delta_{XY}$, siendo Δ_{XY} la diferencia media de Gini entre las variables X e Y, $\Delta_{XY} = E[|X - Y|]$
- e) Si $X > Y$ cs ó $X < Y$ cs $\Rightarrow d(X, Y) = \Delta_{XY} = |\mu_Y - \mu_X|$
- f) esta medida de disimilaridad es invariante ante traslaciones comunes para las variables comparadas; y se ve afectada proporcionalmente por el factor de escala ante cambios de escalas comunes para las variables comparadas.

Sin embargo, es de notar que esta medida de disimilaridad, que mide proximidad entre distribuciones en el sentido planteado, no cumple estrictamente la propiedad triangular; por lo que no llega a ser estrictamente una medida de distancia (Callealta, 2005 y 2006).

3 Diferencia Media de Gini y mixtura de variables.

3.1 Caso General: descomposición de las diferencias medias de Gini cuando cada una de las poblaciones comparadas son mixturas de 2 subpoblaciones.

En este caso, sea X una v.a. con función de distribución

$$F_X(t) = p_U \cdot F_U(t) + p_V \cdot F_V(t) \quad , p_U \geq 0, p_V \geq 0, p_U + p_V = 1$$

y sea Y una v.a. con función de distribución

$$F_Y(t) = p_W \cdot F_W(t) + p_S \cdot F_S(t) \quad , p_W \geq 0, p_S \geq 0, p_W + p_S = 1$$

Bajo la hipótesis de ser X e Y independientes, entonces:

$$\begin{aligned} \Delta_{XY} &= E[|X - Y|] = \\ &= p_U \cdot p_W \cdot E[|U - W|] + p_V \cdot p_W \cdot E[|V - W|] + p_U \cdot p_S \cdot E[|U - S|] + p_V \cdot p_S \cdot E[|V - S|] = \quad [1] \\ &= p_U \cdot p_W \cdot \Delta_{UW} + p_V \cdot p_W \cdot \Delta_{VW} + p_U \cdot p_S \cdot \Delta_{US} + p_V \cdot p_S \cdot \Delta_{VS} \end{aligned}$$

Y si X' es independiente y está idénticamente distribuida que X tal que:

$$F_{X'}(t) = p_U \cdot F_U(t) + p_V \cdot F_V(t) \quad , p_U \geq 0, p_V \geq 0, p_U + p_V = 1, U' \text{ iid } U, V' \text{ iid } V$$

entonces:

$$\begin{aligned} \Delta_{XX} &= E[|X - X'|] = p_U^2 \cdot E[|U - U'|] + 2 \cdot p_U \cdot p_V \cdot E[|U - V'|] + p_V^2 \cdot E[|V - V'|] = \quad [2] \\ &= p_U^2 \cdot \Delta_{UU} + 2 \cdot p_U \cdot p_V \cdot \Delta_{UV} + p_V^2 \cdot \Delta_{VV} \end{aligned}$$

3.2 Descomposiciones de las diferencias medias de Gini entre variables X e Y , como mixturas de las distribuciones X^*, Y^* y C

Según la descomposición de X e Y como mixturas de las subpoblaciones X^*, Y^* y C se verifica que:

$$F_X(t) = p \cdot F_{X^*}(t) + (1-p) \cdot F_C(t) \quad , \quad F_Y(t) = p \cdot F_{Y^*}(t) + (1-p) \cdot F_C(t)$$

Entonces, aplicando la expresión [1] a este caso particular, obtendremos:

$$\Delta_{XX} = p^2 \cdot \Delta_{X^*X^*} + 2 \cdot p \cdot (1-p) \cdot \Delta_{X^*C} + (1-p)^2 \cdot \Delta_{CC} \quad [3]$$

$$\Delta_{YY} = p^2 \cdot \Delta_{Y^*Y^*} + 2 \cdot p \cdot (1-p) \cdot \Delta_{Y^*C} + (1-p)^2 \cdot \Delta_{CC} \quad [4]$$

y aplicando la expresión [2], análogamente obtendremos:

$$\Delta_{XY} = p^2 \cdot \Delta_{X^*Y^*} + p \cdot (1-p) \cdot \Delta_{X^*C} + p \cdot (1-p) \cdot \Delta_{Y^*C} + (1-p)^2 \cdot \Delta_{CC} \quad [5]$$

Nótese que haciendo uso de [3] y [4] podemos escribir finalmente que:

$$\Delta_{XY} = p^2 \cdot \Delta_{X^*Y^*} + p \cdot (1-p) \cdot [\Delta_{X^*C} + \Delta_{Y^*C}] + (1-p)^2 \cdot \Delta_{CC} \quad [6]$$

o, equivalentemente:

$$\Delta_{XY} = p^2 \cdot \Delta_{X^*Y^*} + \frac{1}{2} \{ \Delta_{XX} - p^2 \cdot \Delta_{X^*X^*} + \Delta_{YY} - p^2 \cdot \Delta_{Y^*Y^*} \} \quad [7]$$

3.3 Descomposición de la diferencia media de Gini para la Población Agregada

$$P_T = P_X \cup P_Y$$

Si las poblaciones P_X y P_Y tienen pesos respectivos p_X y $p_Y=1-p_X$ en la población conjunta P_T , entonces ésta es una mixtura de las anteriores, y la variable T que rige su comportamiento posee una función de distribución

$$F_T(t) = p_X \cdot F_X(t) + (1-p_X) \cdot F_Y(t)$$

En este caso, aplicando [1], obtenemos:

$$\Delta_{TT} = p_X^2 \cdot \Delta_{XX} + 2 \cdot p_X \cdot (1-p_X) \cdot \Delta_{XY} + (1-p_X)^2 \cdot \Delta_{YY} \quad [8]$$

y teniendo en cuenta la expresión [6], se deduce que:

$$\Delta_{TT} = p_X^2 \cdot \Delta_{XX} + (1-p_X)^2 \cdot \Delta_{YY} + 2 \cdot p_X \cdot (1-p_X) \cdot [p^2 \cdot \Delta_{X^*Y^*} + p(1-p) \{ \Delta_{X^*C} + \Delta_{Y^*C} \} + (1-p)^2 \cdot \Delta_{CC}] \quad [9]$$

o, alternativamente, teniendo en cuenta [7], que:

$$\begin{aligned} \Delta_{TT} &= p_X^2 \cdot \Delta_{XX} + (1-p_X)^2 \cdot \Delta_{YY} + \\ &+ 2 \cdot p_X \cdot (1-p_X) \cdot \left[p^2 \cdot \Delta_{X^*Y^*} + \frac{1}{2} \{ \Delta_{XX} - p^2 \cdot \Delta_{X^*X^*} + \Delta_{YY} - p^2 \cdot \Delta_{Y^*Y^*} \} \right] = \\ &= p_X \cdot \Delta_{XX} + (1-p_X) \cdot \Delta_{YY} + p_X \cdot (1-p_X) \cdot p^2 \cdot [2 \cdot \Delta_{X^*Y^*} - \{ \Delta_{X^*X^*} + \Delta_{Y^*Y^*} \}] \end{aligned} \quad [10]$$

Ahora bien, sustituyendo en [10] los valores de Δ_{XX} y Δ_{YY} dados en [3] y [4], se obtiene también que:

$$\begin{aligned}
\Delta_{TT} &= p_X \cdot (p^2 \cdot \Delta_{X^*X^*} + 2 \cdot p \cdot (1-p) \cdot \Delta_{X^*C} + (1-p)^2 \cdot \Delta_{CC}) + \\
&+ (1-p_X) \cdot (p^2 \cdot \Delta_{Y^*Y^*} + 2 \cdot p \cdot (1-p) \cdot \Delta_{Y^*C} + (1-p)^2 \cdot \Delta_{CC}) + p_X \cdot (1-p_X) \cdot p^2 \cdot [2 \cdot \Delta_{X^*Y^*} - \{\Delta_{X^*X^*} + \Delta_{Y^*Y^*}\}] = \\
&= (1-p)^2 \cdot \Delta_{CC} \cdot (p_X + (1-p_X)) + p^2 \Delta_{X^*X^*} \cdot (p_X - p_X \cdot (1-p_X)) + p^2 \Delta_{Y^*Y^*} \cdot ((1-p_X) - p_X \cdot (1-p_X)) + \\
&+ 2p_X \cdot p(1-p) \Delta_{X^*C} + 2(1-p_X) \cdot p(1-p) \Delta_{Y^*C} + 2p_X \cdot (1-p_X) \cdot p^2 \cdot \Delta_{X^*Y^*}
\end{aligned}$$

de donde:

$$\begin{aligned}
\Delta_{TT} &= (1-p)^2 \Delta_{CC} + p_X^2 \cdot p^2 \Delta_{X^*X^*} + (1-p_X)^2 \cdot p^2 \Delta_{Y^*Y^*} + \\
&+ 2p_X \cdot p(1-p) \Delta_{X^*C} + 2(1-p_X) \cdot p(1-p) \Delta_{Y^*C} + 2p_X \cdot (1-p_X) \cdot p^2 \cdot \Delta_{X^*Y^*}
\end{aligned} \tag{11}$$

4 Interpretación de las componentes de la diferencia media de Gini:

La expresión [11] nos permite descomponer la diferencia media de Gini de una población compuesta por dos grupos a partir de las poblaciones auxiliares “comparables” y “específicas” definidas previamente.

Obsérvese que en dicha fórmula aparecen las siguientes componentes a las que podemos dar, en principio, los siguientes significados:

- Proporción de población en el núcleo de solapamiento (población comparable)

$$P = 1 - p = 1 - \frac{1}{2} \cdot \int_{-\infty}^{+\infty} |f_X(t) - f_Y(t)| dt$$

- Indicador del núcleo de solapamiento

$$I_{CC} = (1-p)^2 \cdot \Delta_{CC} = 2 \cdot (1-p)^2 \mu_C \cdot G_{CC}$$

- Indicador de integración X^* con el núcleo del solapamiento

$$I_{X^*C} = p \cdot (1-p) \cdot \Delta_{X^*C} = \frac{1}{2} \{ \Delta_{XX} - p^2 \cdot \Delta_{X^*X^*} - (1-p)^2 \cdot \Delta_{CC} \}$$

- Indicador de integración Y^* con el núcleo del solapamiento

$$I_{Y^*C} = p \cdot (1-p) \cdot \Delta_{Y^*C} = \frac{1}{2} \{ \Delta_{YY} - p^2 \cdot \Delta_{Y^*Y^*} - (1-p)^2 \cdot \Delta_{CC} \}$$

- Indicador de la diversidad singular de X (de X^*)

$$I_{X^*X^*} = p^2 \cdot \Delta_{X^*X^*} = 2p^2 \cdot \mu_{X^*} \cdot G_{X^*X^*}$$

- Indicador de la diversidad singular de Y (de Y*)

$$I_{Y^*Y^*} = p^2 \cdot \Delta_{Y^*Y^*} = 2p^2 \cdot \mu_{Y^*} \cdot G_{Y^*Y^*}$$

- Indicador de disimilaridad diferencial entre X e Y

$$\delta_{XY} = p^2 \cdot \Delta_{X^*Y^*}$$

Consecuentemente, de [3], [4], [5] y [11], obtenemos las siguientes descomposiciones de las respectivas diferencias medias de Gini, en función de éstas componentes:

$$\Delta_{XX} = I_{X^*X^*} + 2I_{X^*C} + I_{CC} \quad [12]$$

$$\Delta_{YY} = I_{Y^*Y^*} + 2I_{Y^*C} + I_{CC} \quad [13]$$

$$\Delta_{XY} = I_{CC} + I_{X^*C} + I_{Y^*C} + \delta_{XY} \quad [14]$$

$$\Delta_{TT} = p_X^2 I_{X^*X^*} + (1-p_X)^2 I_{Y^*Y^*} + I_{CC} + 2p_X I_{X^*C} + 2(1-p_X) I_{Y^*C} + 2p_X(1-p_X) \delta_{XY} \quad [15]$$

5 Índice de Gini y mixtura de Variables

5.1 Descomposición de los Índices de Gini relativos a X e Y, como mixtura de las subpoblaciones X*, Y* y C.

De [3], para X, obtenemos que:

$$\begin{aligned} \mu_X \cdot G_X &= p^2 \cdot \mu_{X^*} \cdot G_{X^*} + p \cdot (1-p) \cdot (\mu_{X^*} + \mu_C) G_{X^*C} + (1-p)^2 \cdot \mu_C \cdot G_C = \\ &= \frac{1}{2} I_{X^*X^*} + I_{X^*C} + \frac{1}{2} I_{CC} \end{aligned} \quad [16]$$

de [4], para Y,

$$\begin{aligned} \mu_Y \cdot G_Y &= p^2 \cdot \mu_{Y^*} \cdot G_{Y^*} + p \cdot (1-p) \cdot (\mu_{Y^*} + \mu_C) G_{Y^*C} + (1-p)^2 \cdot \mu_C \cdot G_C = \\ &= \frac{1}{2} I_{Y^*Y^*} + I_{Y^*C} + \frac{1}{2} I_{CC} \end{aligned} \quad [17]$$

y de [7],

$$\begin{aligned} (\mu_X + \mu_Y) \cdot G_{XY} &= p^2 \cdot \Delta_{X^*Y^*} + \mu_X \cdot G_X - p^2 \cdot \mu_{X^*} \cdot G_{X^*} + \mu_Y \cdot G_Y - p^2 \cdot \mu_{Y^*} \cdot G_{Y^*} = \\ &= I_{CC} + I_{X^*C} + I_{Y^*C} + \delta_{XY} \end{aligned} \quad [18]$$

5.2 Descomposición del Índice de Gini para la Población Agregada $P_T = P_X \cup P_Y$

De [8], para T,

$$\mu_T \cdot G_T = p_X^2 \cdot \mu_X \cdot G_X + p_X \cdot (1 - p_X) \cdot (\mu_X + \mu_Y) G_{XY} + (1 - p_X)^2 \cdot \mu_Y \cdot G_Y \quad [19]$$

y de [10],

$$\mu_T \cdot G_T = p_X \cdot \mu_X \cdot G_X + (1 - p_X) \cdot \mu_Y \cdot G_Y + p_X \cdot (1 - p_X) \cdot p^2 \cdot \left\{ \Delta_{X^*Y^*} - \mu_{X^*} \cdot G_{X^*} - \mu_{Y^*} \cdot G_{Y^*} \right\} \quad [20]$$

Sin embargo, obsérvese que de [10] puede obtenerse la descomposición:

$$\begin{aligned} \Delta_{TT} &= p_X \cdot \Delta_{XX} + (1 - p_X) \cdot \Delta_{YY} + p_X \cdot (1 - p_X) \cdot \left[2p^2 \Delta_{X^*Y^*} - \Delta_{XX} p^2 \cdot \frac{\Delta_{X^*X^*}}{\Delta_{XX}} - \Delta_{YY} p^2 \cdot \frac{\Delta_{Y^*Y^*}}{\Delta_{YY}} \right] = \\ &= p_X \cdot \Delta_{XX} \cdot \left(1 - (1 - p_X) p^2 \cdot \frac{\Delta_{X^*X^*}}{\Delta_{XX}} \right) + (1 - p_X) \cdot \Delta_{YY} \cdot \left(1 - p_X p^2 \cdot \frac{\Delta_{Y^*Y^*}}{\Delta_{YY}} \right) + 2p_X \cdot (1 - p_X) \cdot p^2 \Delta_{X^*Y^*} \end{aligned} \quad [21]$$

de donde

$$\begin{aligned} \mu_T G_T &= p_X \cdot \mu_X G_X \cdot \left(1 - (1 - p_X) p^2 \cdot \frac{\Delta_{X^*X^*}}{\Delta_{XX}} \right) + (1 - p_X) \cdot \mu_Y G_Y \cdot \left(1 - p_X p^2 \cdot \frac{\Delta_{Y^*Y^*}}{\Delta_{YY}} \right) + \\ &+ p_X \cdot (1 - p_X) \cdot \delta_{XY} \end{aligned} \quad [22]$$

en la que podemos ver cierta analogía con la descomposición que hace Yitzhaki (1994).³

Aquí la aproximación al índice de Gini entre grupos viene representada proporcionalmente por la medida de la disimilaridad entre poblaciones propuesta, y las expresiones entre paréntesis parecen informar de un cierto grado de solapamiento, tomando valores tanto menores cuanto más se alejan una de otras, si bien no se encuentran normalizados.

³ En Yitzhaki (1994), la fórmula referenciada allí como (21) presenta la descomposición por grupos del Índice de Gini de la población agregada, como

$$G_T = \frac{p_X \mu_X}{\mu_T} \cdot O_X \cdot G_X + \frac{p_Y \mu_Y}{\mu_T} \cdot O_Y \cdot G_Y + GB$$

siendo GB una aproximación al índice de Gini entre grupos, y O_X y O_Y los indicadores de solapamiento de X e Y sobre la población agregada, respectivamente.

6 Conclusiones

En este trabajo se han desarrollado diversas descomposiciones de la Diferencia Media de Gini y del índice de concentración de Gini para una población constituida por dos grupos, de acuerdo con la propuesta realizada para la descomposición de éstos grupos constituyentes, con el objetivo de tratar de conectar en dichas descomposiciones de la diferencia media y del índice de concentración de Gini los conceptos de desigualdad, proximidad y solapamiento entre aquéllos. A partir de aquí, los próximos pasos deberán encaminarse hacia la identificación de un índice normalizado de solapamiento compatible con estas descomposiciones, y hacia la aplicación de éstas sobre distribuciones empíricamente observadas, para evaluar empíricamente su interés descriptivo, así como precisar los presumibles significados de sus componentes.

Bibliografía

- Callealta, F.J. (2006) “A new measure of dissimilarity between distributions: application to the analysis of income distributions convergence in the European Union” en *Distribution Models Theory*. World Scientific Publishing. Chapter 8, pp. 125-160.
- Callealta, F.J. (2005) “Convergencia entre Distribuciones Personales de Renta de Países de la Unión Europea” en *Actas de la XXXI Reunión de Estudios Regionales Asociación Española de Ciencia Regional* en CD-ROM.
- Dagum, C. (1980) “Inequality Measures Between income Distributions with Applications”. *Econometrica*, Vol. 48, nº 7, pp. 1971-1803.
- Yitzhaki, S. (1994) “Economic distance and overlapping of distributions”. *Journal of Econometrics*, nº 61. pp. 147-150.

SELECCIÓN DE NUEVOS PRODUCTOS TURÍSTICOS EN PINAR DEL RÍO, CUBA.

PÉREZ LEÓN, VÍCTOR¹

e-mail: VP_LEON@MAT.UPR.EDU.CU

CAMARGO TORIBIO, ISIS²

e-mail: ISIS@MAT.UPR.EDU.CU

CABALLERO FERNÁNDEZ, RAFAEL³

e-mail: RAFAEL.CABALLERO@UMA.ES

GONZÁLEZ LOZANO MERCEDES³

e-mail: M_GONZALEZ@UMA.ES

Departamento de MATEMÁTICAS¹

UNIVERSIDAD DE PINAR DEL RÍO. CUBA

CENTRO DE ESTUDIOS DE GERENCIA, DESARROLLO LOCAL Y TURISMO (GEDELTUR)²

UNIVERSIDAD DE PINAR DEL RÍO. CUBA

Departamento de ECONOMÍA APLICADA (MATEMÁTICAS)³

UNIVERSIDAD DE MÁLAGA

Área temática: Métodos Cuantitativos.

Resumen

La introducción de nuevos productos turísticos es una de las opciones del Ministerio de Turismo en Cuba (MINTUR) para elevar el desarrollo del sector turístico, lo cual incide satisfactoriamente en beneficio del país. El trabajo que se presenta pone en práctica la aplicación de las técnicas de decisión multicriterio en el proceso de toma de decisiones en la Delegación del Ministerio de Turismo en Pinar del Río, Cuba, con el objetivo de seleccionar un nuevo producto turístico que debe ser implementado inicialmente en la provincia para que, su empleo de forma sostenible, genere los ingresos necesarios que permitan la puesta en marcha de los demás productos propuestos.

Se tienen en consideración cuatro productos turísticos para su posible implantación y desarrollo y, para cada uno de ellos, diversos criterios, que representan indicadores de sostenibilidad, de carácter económico y ecológico, de los cuales, varios se contraponen. Para la elección de la mejor alternativa se utilizan métodos multicriterios discretos.

La información sobre la cual se toman las decisiones fue brindada principalmente por la Delegación del MINTUR en Pinar del Río, el Grupo Extrahotelero Palmares y otras organizaciones implicadas.

Palabras clave: Producto Turístico, Métodos Multicriterio.

Abstract

The introduction of new tourist products is one of the options of the Ministry of Tourism in Cuba (MINTUR) to increase the tourist sector development and to obtain more benefits for the country. This work applies Multicriteria Decision Making techniques to aid the Tourism Ministry Delegation in Pinar del Rio, Cuba, in the process of decision making to select a new tourist product in the area. This new product must be implemented so that its use, in a sustainable way, generates the income needed to implement some other proposed products.

We will consider four new possible products for implementation, each of them being evaluated for several different criteria measuring sustainability, benefits and ecological aspects. Some of these criteria are in strong conflict, and this is the main reason to use a Multicriteria Decision Making tool to choose the most suitable alternative.

The information and data used for this work was provided by the MINTUR Delegation of Pinar del Rio, the Grupo Extrahotelero Palmares and some other organizations implied in the process.

Key words: Tourist Product, Multicriteria Decision Making.

1. Introducción

A partir de 1990 el turismo se ha presentado como uno de los pilares fundamentales de la economía cubana. Desde entonces el número de turistas que visitan la Isla aumenta cada año. Ello evidencia la preferencia de Cuba para la industria del ocio. En 1990 se recibieron 340 mil turistas y, ya en el año 2005 la cifra de visitantes fue de 2,3 millones, según datos de la Oficina Nacional de Estadística (ONE), situándose en el tercer lugar entre los receptores de turismo en el Caribe, con el 12% del total de visitas al área (Organización Mundial del Turismo, (OMT), 2006).

Entre los atractivos de la oferta turística predomina el producto “Sol y Playa”, siendo el principal motivo de las visitas al país. Algo similar ocurre en la mayoría de los destinos del Caribe Insular, motivo por el cual se debe considerar la diversificación de la oferta turística en Cuba, uno de los principales objetivos del Ministerio de Turismo de Cuba (MINTUR).

Cuba cuenta con un área de 109886 km², divididos entre la isla de Cuba, Isla de la Juventud y los cayos adyacentes, de ellos, 5353 km² son de áreas protegidas: 7 parques nacionales, 3 reservas ecológicas, 7 reservas florísticas manejadas, 12 refugios de fauna, 1 elemento natural destacado¹, 2 paisajes naturales protegidos y 3 áreas de recursos de manejo. Entre esta diversidad de áreas, 15 cuentan con reconocimiento internacional en el año 2005.

La provincia que posee la mayor cantidad de áreas protegidas es Pinar del Río, con una superficie total de 1540 km² divididos entre el Parque Nacional Viñales, la Sierra del Rosario y el Parque Nacional Guanahacabibes; además es significativo destacar que dentro de las seis Reservas del la Biosfera que tiene Cuba, dos de ellas se ubican en esta provincia (Península de Guanahacabibes y La Sierra del Rosario) (Camargo, 2003). Con estos datos es posible afirmar que la provincia de Pinar del Río cuenta con muy buenas condiciones para el aprovechamiento de los recursos naturales con fines turísticos.

¹ Es un área que contiene una o más características naturales de valor destacado o excepcional, por su rareza implícita y sus cualidades representativas o estéticas y que puede contener valores histórico-culturales asociados, siendo manejadas con el fin de conservar dichas características.

En el año 2005 por concepto de turismo en Pinar del Río se ingresaron \$ 40271300 y se empleó el 1,93% de la población económicamente activa. Se recibieron 339419 visitantes, lo que significó solo el 14,6% del número de visitantes al País. Con una estancia promedio de 1,8 días (Oficina Territorial de Estadística, (OTE), 2006), por debajo de la media nacional en ese año, que fue de 6,6 días (ONE 2006). La Tabla 1 muestra la cantidad de turistas, por país de origen, llegados a Pinar de Río en el 2005, entendiendo por turista aquella persona que permanece, por lo menos, una noche en un medio de alojamiento colectivo o privado en el lugar visitado.

País	Turistas	% respecto al total
Alemania	26704	21
Reino Unido	22309	11
Francia	21528	20
Holanda	15650	41
España	8733	4,5
Italia	6803	4
Canada	5171	0,8
Bélgica	4818	23
Suiza	4157	19
EE.UU.	3207	9
Dinamarca	2951	32
Mexico	2380	2,7
Austria	1633	10
Brasil	536	3
Otros	34229	6
Total	160809	6,93

Fuente: Elaboración propia a partir de ONE y OTE

Tabla 1: Turistas llegados a Pinar del Río en el año 2005.

Los principales turistas que acuden a Pinar del Río son los que provienen de Alemania, Reino Unido, Francia, Holanda y España, los tres primeros con cifras por encima de los 20000. Con respecto al total de turistas emitidos por cada país, el 41% de los holandeses que vienen a Cuba, visitan Pinar del Río, de los provenientes de Dinamarca lo hace el 32% y más del 20% de los que viajan desde Bélgica, Alemania y Francia seleccionan a Pinar del Río como destino principal.

Más del 50% de las visitas a Pinar del Río son realizadas en el día (OTE, 2006), pues son ofrecidas como parte del paquete de productos. Es esta la causa del bajo índice de pernoctaciones que existe actualmente en el territorio, lo que trae consigo que los ingresos por concepto de turismo no sean los mejores.

La Delegación del Ministerio de Turismo (MINTUR) en Pinar del Río tiene como principal objetivo amentar los ingresos generados por la actividad turística y para ello, se ha propuesto elevar la estancia promedio de los visitantes a la provincia, variable esta que es de gran interés tanto político como social, en los destinos turísticos (Palmer, et al., 2006). Para alcanzar este propósito se está analizando la elaboración de nuevos productos turísticos, que contribuyan a diversificar la oferta en la zona, otro de los objetivos del MINTUR. El diseño, elaboración y puesta en marcha de los nuevos productos correrá a cargo del “Grupo Empresarial Extrahotelero Sucursal Palmares, Pinar del Río”, perteneciente a la Delegación del MINTUR.

A continuación se mencionan los 4 nuevos productos turísticos que se pretenden poner en marcha y el municipio pinareño en que tendrá lugar, cada uno orientado a distintos segmentos de mercado dentro del turismo. Esta información fue ofrecida por los especialistas de campo del Grupo Palmares.

- ✓ Centro de Buceo en Cayo Jutías. Minas de Matahambre.
- ✓ Pesca recreativa de Black Bass. Complejo Cuyaguaje, Guane y Sandino.
- ✓ Excursión a Cayo San Felipe. La Coloma, Pinar del Río.
- ✓ Sendero de observación de Aves. Finca San Vicente, Viñales.

Actualmente la empresa no cuenta con el capital suficiente para desarrollarlos todos de forma conjunta, por lo cual se precisa seleccionar uno que será puesto en marcha inicialmente y, manteniendo su desarrollo y utilización de forma sostenible, lograr la generación de los ingresos necesarios para la implementación de los restantes productos. Por consiguiente, el objetivo de este trabajo es la selección del producto turístico (alternativa) a desarrollar de forma inicial, para lo cual se tendrán en cuenta varios criterios, principalmente de corte económico y ecológico, que son de interés para el centro decisor, en este caso, el Grupo Extrahotelero Palmares.

Los criterios que se tendrán en cuenta en el problema son criterios de medida de indicadores de sostenibilidad (Hardi & Barg, 1997; World Bank, 1997; Farsari y Prastacos, 2002) puesto que en la ejecución de determinadas acciones y políticas, estos constituyen herramientas de evaluación que permiten chequear cuán exitosas han sido las acciones y políticas seleccionadas y si los humanos están actuando adecuadamente en su camino hacia el desarrollo sostenible². Además, los indicadores de sostenibilidad son herramientas de planificación que ayudan a elegir entre alternativas (Hardi & Barg, 1997).

2. Planteamiento del Problema

El problema planteado es la selección de un producto turístico, entre una serie de nuevos productos, considerando distintos aspectos. En este problema existen varias alternativas posibles y criterios a tener en cuenta, algunos de ellos contrapuestos, lo que hace que sea complejo encontrar una solución adecuada que equilibre todos los aspectos. Por lo tanto, este problema puede abordarse desde la teoría de la Toma de Decisiones Multicriterio.

Para poder determinar cuál es el producto que se va a elegir de entre los cuatro que se consideran, se deben maximizar los ingresos que serán generados por cada producto. De acuerdo con la OMT (2005), los ingresos se utilizan ampliamente como indicador de importancia para el turismo como sector económico y como instrumento de vigilancia de los cambios producidos en el ámbito turístico a escala mundial. Además, constituyen un indicador fundamental para evaluar los beneficios del turismo en una comunidad o destino.

Como se pretende aumentar las pernoctaciones en los destinos turísticos, la oferta turística de la comunidad influye sobremanera en ello, siempre y cuando esté diseñada de acuerdo a las expectativas de los clientes. Mientras mayor sea la oferta turística, mayor cantidad de días permanecerán los visitantes en el destino. Por consiguiente, en el modelo que se formulará se tratará de maximizar el número de atracciones que componen la oferta turística de la comunidad de destino.

² Indicadores de desarrollo sostenible para los destinos turísticos. Guía práctica. OMT, 2005. pp. 9-10.

La capacidad de alojamiento influye también de gran manera en el problema, por lo que será maximizada, para garantizar de esa manera que exista capacidad de acogida para toda la demanda y lograr que se abarque todo el turismo posible. La capacidad de alojamiento turístico es uno de los mecanismos más importantes para controlar eficazmente la presión humana sobre los recursos, sobre el medio local (OMT, 2005; Ivars, 2001). La falta de este control produce insostenibilidad (OMT, 2005). Además, expresada en un plano muy local, permite evaluar la presión turística.

Hasta el momento se han presentado criterios meramente económicos, correspondientes a una de las aristas de la sostenibilidad (Choi y Sirakaya, 2005, OMT, 2005) pero, no debemos descuidar el plano ecológico, por ello se proponen dos criterios adicionales que evidenciarán esta dimensión de la sostenibilidad.

En el modelo se pretende maximizar la capacidad de carga del sitio donde estará enclavado el nuevo producto. Este criterio se emplea para indicar a cuántos turistas se pueden dar cabida en un determinado lugar o zona sin dañarlo o sin menoscabar el grado de satisfacción de los visitantes como manifestaron Mathienson y Wal (1992) y OMT (2005). Permite, además, estructurar una estrategia de desarrollo que debe trasladarse necesariamente a una programación del crecimiento urbano-turístico y a una asignación coherente del uso del suelo (Ivars, 2001).

Otro criterio a maximizar será el número de proyectos de investigación sobre el medio ambiente en las áreas o localidades donde serán implementados los nuevos productos. Es un indicador que evidencia el interés por el medioambiente en el área (Farsari y Prastacos, 2000; OMT, 2005).

El carácter multiobjetivo del problema descrito viene dado por la existencia de criterios en conflicto pues, a simple vista no se puede determinar cuál será la mejor alternativa porque no existe una que domine a las demás completamente. Sobre problemas de este tipo se puede encontrar en Current et al., (1990) una extensa recopilación, en la cual se revisan aspectos multiobjetivo de problemas de localización.

Veamos ahora cuál va a ser la formulación del problema con el que nos enfrentamos.

El problema en cuestión consiste en la selección del primer producto turístico que entrará en vigor. Se seleccionará de entre las s alternativas candidatas (productos que conforman el conjunto J).

Las variables que conformarán el modelo serán las variables binarias y_j , que se encuentran asociadas a los productos (tomarán valor 1 si es seleccionado el producto j con $j \in J$, y 0 en caso contrario).

Teniendo en cuenta la notación establecida, a continuación definiremos los criterios considerados.

✓ Ingresos esperados:

Los ingresos esperados anualmente para cada nuevo producto turístico son el resultado del producto del precio por la cantidad de turistas que demandan dicho producto al año.

- Precio:

Para cada producto se determinó el precio a que será ofertado: el sendero de observación de aves tendrá el mismo precio que otro producto similar que estuvo en explotación. Este dato fue solicitado a la oficina comercial de la Delegación del MINTUR en Pinar del Río. Algo igual sucede con el precio de la excursión al cayo San Felipe. Se estableció contacto con el Club de actividades Náuticas en la provincia de Cienfuegos, la zona central de Cuba, donde se brinda este tipo de producto y el centro decisor determinó que el precio a que se ofertará es el mismo.

El precio del producto pesca recreativa de Black Bass fue fijado por el Grupo Extrahotelero Palmares. Con ayuda de los analistas, obtuvieron los precios a que se ofertan estos productos en Estados Unidos y México, fuertes competidores en la rama en el área.

Para hallar el precio a que serán ofrecidos los servicios en el centro de buceo, se solicitó al Centro Internacional de Buceo “María la Gorda”, situado en Pinar del Río, los datos de las últimas 1000 ventas realizadas, independientemente del tipo de actividad que haya sido demandada, con ello, se determinó el precio medio ponderado del número de actividades que se ofertan en él. El valor que se obtuvo fue el que seleccionó el centro decisor como precio del producto centro de buceo en Cayo Jutías.

- Demanda:

La demanda de este tipo de productos (anual) fue obtenida en los estudios que han sido realizados con anterioridad por el Grupo Extrahotelero Palmares, necesarios para

justificar ante el organismo superior (MINTUR) la creación de los nuevos productos turísticos.

Con los datos que han sido especificados, la expresión que recoge los ingresos esperados viene dada como:

$$\text{Max } \sum_{j=1}^s p_j d_j y_j, \text{ donde}$$

p_j son los precios correspondientes a cada producto j ($j=1, \dots, s$).

d_j representa la demanda que tiene este producto j ($j=1, \dots, s$).

Producto	Precio unitario por sesión	Demanda	Ingresos
Centro de Buceo:	52	150	7800
Observación de aves	5	90	450
Pesca de Black Bass	300	100	30000
Excursión a San Felipe	40	500	20000

Fuente: Elaboración propia a partir de datos del Grupo Extrahotelero Palmares.

Tabla 2: Ingresos anuales (CUC³).

✓ Oferta turística:

Este criterio se tendrá en cuenta porque permite la diversificación de la oferta en la localidad (Farsari y Prastacos, 2001) y, apoyado en la eliminación de los costes de traslación para los turistas, permitirá el aumento de las pernoctaciones.

Para determinar el valor de este criterio para cada alternativa de selección se realizó un recuento en la Oficina Territorial de Estadística, de los sitios de interés turístico como cines, teatros, museos, casas de cultura, etc., que existen en cada uno de los municipios en que tendrán lugar las alternativas de selección. Se incluyen en la relación, los productos que son ofertados en el catálogo de ventas de la Agencia de Viajes Cubanacán, la encargada de comercializar todos los productos que se brindan para el turismo internacional en el país.

La expresión que será empleada para la representación de estos datos es:

³ Peso Cubano Convertible.

$$\text{Max} \sum_{j=1}^s o_j y_j, \text{ donde}$$

o_j es la cantidad de atracciones que componen la oferta turística de la alternativa j ($j=1, \dots, s$).

Municipio	Cines	Casa de cultura	Teatros	Museos	Campismo	Oferta de Cubanacán	Total
Minas de Matahambre	5			1		1	7
Viñales	3	1		1	1	20	26
Guane y Sandino	6	2	1	2	1	5	17
Pinar del Río	3	1	1	3	1	6	15

Fuente: Elaboración propia a partir de OTE, Cubanacán y Campismo Popular.

Tabla 3: Número de Atracciones turísticas en los municipios.

✓ Capacidad de alojamiento:

Se tomará en cuenta para la capacidad de alojamiento en los municipios el número de habitaciones que se ofertan para el hospedaje de turistas internacionales, ya sea en hoteles, campismos y casas particulares que rentan habitaciones en MLC⁴

Los datos con respecto a la capacidad hotelera fueron solicitados a la oficina comercial de la Delegación Provincial del MINTUR en Pinar del Río. Para determinar la oferta de alojamiento en casas particulares se visitó la Oficina Nacional de Administración Tributaria (ONAT), que registra el pago de los impuestos por la tenencia de viviendas que sirven de alojamiento para el turismo internacional, entre otras.

La capacidad de alojamiento en el problema viene dada por la expresión:

$$\text{Max} \sum_{j=1}^s c_j y_j, \text{ donde}$$

c_j es la cantidad de habitaciones disponibles en el municipio donde se desarrollará el producto j ($j=1, \dots, s$).

Municipio	Hoteles	Particulares	Total
Minas de Matahambre	-	1	1
Viñales	192	262	454
Guane y Sandino	83	4	87
Pinar del Río	223	57	280

Fuente: Elaboración propia a partir de datos de ONAT, MINTUR.

Tabla 4: Capacidad de alojamiento (Habitaciones).

⁴ Moneda Libremente Convertible.

✓ Capacidad de carga:

La “capacidad de carga” como instrumento de gestión es, en sí mismo, un indicador “sintético”, ya que incluye numerosos elementos y factores que influyen en la evolución del territorio, pero su análisis se centra en la supervisión cuantitativa de los usuarios, lo que precisamente le diferencia de los indicadores habituales (indicadores de categoría) que supervisan los recursos (OMT, 2005).

Este concepto tiene interés para el turismo, en particular porque se refiere a límites y umbrales que no queremos sobrepasar. Mide el grado de utilización sostenible (OMT, 2005). Es también, uno de los indicadores que se demandan para supervisar la sostenibilidad del entorno natural (Choi y Sirakaya, 2005).

Los valores para este criterio en cada uno de los nuevos productos los proporcionó el centro decisor. Para ello se apoyó en los estudios realizados por los especialistas de campo como parte de la planificación de los nuevos productos.

La expresión que nos permite determinar este objetivo es la que sigue:

$$Max \sum_{j=1}^s C_j y_j, \text{ donde}$$

C_j representa la capacidad de carga del sitio en que tendrá lugar el producto j ($j=1, \dots, s$).

Producto	Capacidad de carga del sitio
Centro de Buceo:	16
Observación de aves	10
Pesca de Black Bass	10
Excursión a San Felipe	30

Fuente: Grupo extrahotelero Palmares

Tabla 5: Capacidad de carga (turistas por sesión).

✓ Número de proyectos de investigación sobre el medio ambiente en el área:

De acuerdo con (Stuart, 1997), la importancia que las autoridades locales le brindan al turismo está directamente relacionada con el número de investigaciones sobre el medio ambiente desarrolladas en el área. Tal es el caso que para determinar los valores correspondientes a este objetivo el centro decisor brindó información acerca de las entidades que resultan de su interés para este criterio.

Para determinar el número de proyectos de investigación que se han llevado a cabo en el municipio donde tendrá lugar el producto se estableció comunicación telefónica o mediante e-mail con cada una de las entidades que aparecen en la primera columna de la tabla 6 y se recibió información acerca de los trabajos que han realizado en cada uno de los municipios de interés en la presente investigación.

La expresión que nos permite desarrollar este objetivo es:

$$\text{Max} \sum_{j=1}^s I_j y_j, \text{ donde}$$

I_j es el número de proyectos de investigación sobre el medio ambiente que se han llevado a cabo en el municipio donde tendrá lugar el producto j ($j=1, \dots, s$).

Número de trabajos de investigación	Minas de Matahambre	Viñales.	Guane y Sandino.	La Coloma, Pinar del Río
ENIA	2	2	3	-
Planif. F.	4	2	1	-
CEMARNA	-	3	-	-
Emp. Nac. Flora y Fauna	-	-	-	18
Eco Vida	-	39	42	-
MINTUR	1	1	1	1
Comité Prov. Espeleología	14	24	219	8
Total	21	71	266	27

Fuente: Elaboración propia a partir de informes de los organismos relacionados y el CITMA.

Tabla 6: Número de proyectos de investigación.

En la *Tabla 7* finalmente se exponen todos los datos del problema que se aborda. En ella aparecen señalados para cada criterio los valores máximos (en verde) y los valores mínimos (en rojo) y la última fila brinda la máxima diferencia existente entre las distintas alternativas en cada uno de los criterios.

Alternativas	Ingresos	Oferta Turística	Capacidad de Alojamiento	Capacidad de Carga	Proyectos de Investigación
Centro de Buceo	7800.00	7.00	1.00	16.00	21.00
Obsevación de Aves	450.00	26.00	454.00	10.00	71.00
Pesca de Black Bass	30000.00	17.00	87.00	10.00	266.00
Excursión a San Felipe	20000.00	15.00	280.00	30.00	27.00
Máxima Diferencia	29550.00	19.00	453.00	20.00	245.00

Tabla 7: Valores de los criterios para cada alternativa.

3. Resolución del Problema

El problema multicriterio discreto será abordado mediante el PROMETHEE (Preference Ranking Organisation Method for Enrichment Evaluations), propuesto por Brans et al. (1986), como uno de los métodos que sobresalen entre los que emplean las Relaciones de Superación que permiten ofrecer una ordenación parcial o total de las alternativas y presentado por Fernández Barberis, G. (2002) como de gran atractivo en el mundo de la decisión multicriterio.

Existen numerosas aplicaciones, entre las que se pueden destacar los problemas de selección de la óptima fuente de energía renovable para las instalaciones turísticas (Mladineo y Grabovac, 1988), la ubicación de instalaciones comerciales en ambiente competitivo (Karkazis, 1989), depósitos de desechos (Brigs et al., 1990, Vuk et al., 1991, Hokkanen y Salminen, 1997), emplazamiento de un parque temático (Fernández y Martín, 1999), localización de empresas de reciclaje de RAEE (Queiruga et al., 2004), localización de incineradoras de Residuos MER (Caballero et al., 2004), entre otras.

El software que se utiliza para resolver el problema mediante PROMETHEE es el Decisión Lab. Además, con este programa realizaremos un análisis de sensibilidad en los pesos de los criterios, y seguidamente un análisis de robustez de estos.

El método PROMETHEE fue diseñado con el fin de ayudar al decisor en los problemas de selección de ordenamiento de alternativas posibles sometidas a una evaluación multicriterio. Admite la existencia de alternativas incomparables y conduce al decisor a investigar bajo qué criterios las alternativas evidencian un buen comportamiento y bajo cuáles su desempeño es ineficiente, para proceder luego a efectuar la mejor elección de acuerdo a su propio esquema de preferencias.

La ventaja de los métodos PROMETHEE frente al resto de sus competidores es que requiere información adicional muy clara y precisa que puede ser fácilmente obtenida por el decisor con la permanente y activa ayuda del analista. Ello justifica la variedad de aplicaciones de estas técnicas, entre las que podemos encontrar aplicaciones a problemas de administración de desperdicios radioactivo (Briggs et al., 1990), problemas de diagnóstico médico de patologías (Du Bois et al., 1989) y con especial atención, a problemas de localización, concretamente, de instalaciones comerciales en ambiente competitivo (Karkazis, J., 1989), depósitos de desechos (Briggs et al., 1990;

Vuk et al., 1991), incineradores de materiales específicos de riesgo en Andalucía (Caballero et al., 2004), etc.

El análisis con el método PROMETHEE se complementa con el proceso GAIA (Geometrical Análisis for Intercative Assistance), un módulo de interacción visual (Mareschal y Brans, 1988) que ofrece al decisor una descripción gráfica clara de su problema de decisión, enfatizando los conflictos existentes entre los criterios y el impacto de los pesos en la decisión final. Este análisis es más descriptivo que los métodos PROMETHEE y está orientado gráficamente.

Para la localización de un único producto turístico se aplica el método PROMETHEE mediante el software Decisión Lab. La solución va a depender de las características técnicas del problema y de las preferencias del centro decisor. Estas vienen dadas por la información adicional requerida por el método PROMETHEE: Información entre los distintos criterios (intercriterios) e información propia de cada criterio (intracriterio).

La información entre los distintos criterios consiste en el establecimiento de pesos o ponderaciones que reflejan la importancia relativa que el decisor asigna a cada uno de ellos. Para ello hemos empleado el programa Expert Choice, que se basa en la metodología AHP (Analytic Hierarchy Process), propuesta por Saaty (1977). Con esta herramienta se determinan los pesos de cada criterio en dependencia de las preferencias del centro decisor. Además, esta metodología permite ofrecer la ordenación de las alternativas; esto se lleva a cabo estableciendo una comparación para cada par de alternativas con respecto a cada criterio. Luego, es posible visualizar el peso total asignado a cada una de las alternativas, resultando como mejor solución, aquella que mayor ponderación obtenga.

Las comparaciones dieron como resultado iguales ponderaciones para los criterios capacidad de carga y el número de proyectos de investigación, que fueron a su vez los de mayor fortaleza, con un valor de 0.283. Algo similar pasa con los criterios ingresos y oferta turística, que presentan igual ponderación, aunque con un valor menor que los dos anteriores, en este caso 0.163. Luego se observa la capacidad de alojamiento, con una ponderación de 0.109. Estos resultados demuestran que el centro decisor le confiere más importancia a los criterios de orden ecológico que a los económicos. El

ratio de inconsistencia es de 0.02, inferior a 0.1, por lo cual se puede considerar el resultado obtenido como aceptable para realizar el análisis posterior.

A su vez PROMETHEE modeliza las preferencias dentro de cada criterio asignando a cada uno de ellos un criterio generalizado que tiene en cuenta las desviaciones y las escalas de los criterios. Estos criterios generalizados se obtienen determinando una función de preferencia para cada uno de ellos. Las funciones que se han seleccionado son todas del tipo lineal para todos los criterios, pues son de fácil comprensión para el centro decisor, además las más utilizadas y las que mejor se adecuan al problema en cuestión. Los umbrales de preferencia e indiferencia elegidos para los distintos criterios aparecen en la tabla posterior.

Umbrales	Ingresos	Oferta Turística	Capacidad de Alojamiento	Capacidad de Carga	Proyectos de Investigación
Indiferencia	6000	8	100	6	60
Preferencia	20000	11	200	10	100

Tabla 8: Umbrales de preferencia e indiferencia.

Con toda la información que se ha obtenido se establece la ordenación de las alternativas mediante el PROMETHEE. Dicha ordenación aparece en el *Gráfico 1* donde podemos ver que la mejor alternativa es el producto “Pesca recreativa de Black Bass”, de acuerdo con la distribución de los pesos considerados para cada uno de los criterios, pero seguida, muy de cerca, por la alternativa “Excursión a Cayo San Felipe”. Solo se ha presentado la ordenación total debido a que no existen alternativas incomparables en nuestro problema.

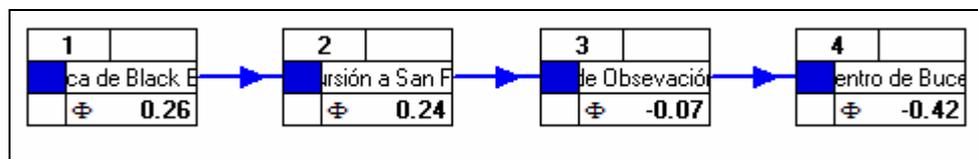


Gráfico 1: Ordenación total.

Mediante el análisis de los perfiles de las alternativas se pueden establecer comparaciones que dan mayor fiabilidad a la selección que se desea realizar, pues permiten determinar la potencialidad de los atributos en cada una de estas. Para ello se establecerá la comparación de las alternativas Pesca de Black Bass y Excursión a San Felipe, que son las que ocupan el primero y segundo puesto. En la tabla que sigue, se

recogen los valores de los criterios para los productos que serán comparados, además de los valores medios que toman los criterios en las cuatro alternativas

Alternativas	Ingresos	Oferta Turística	Capacidad de Alojamiento	Capacidad de Carga	Proyectos de Investigación
Pesca de Black Bass	30000	17	87	10	266
Excursión a San Felipe	20000	15	280	30	27
Valor Medio	14562.5	16	205.5	16.5	96.25

Tabla 9: Valores de los criterios de las alternativas que se comparan.

Para la alternativa pesca de Black Bass los valores de los atributos ingresos, oferta turística y proyectos de investigación se encuentran por encima de la media de los valores de estos criterios y la capacidad de alojamiento y la capacidad de carga se encuentran por debajo. Como se maximizan todos los criterios, en el perfil de la pesca de Black Bass (*Gráfico 2*) se representan por encima de 0 los criterios con valores mayores que la media.

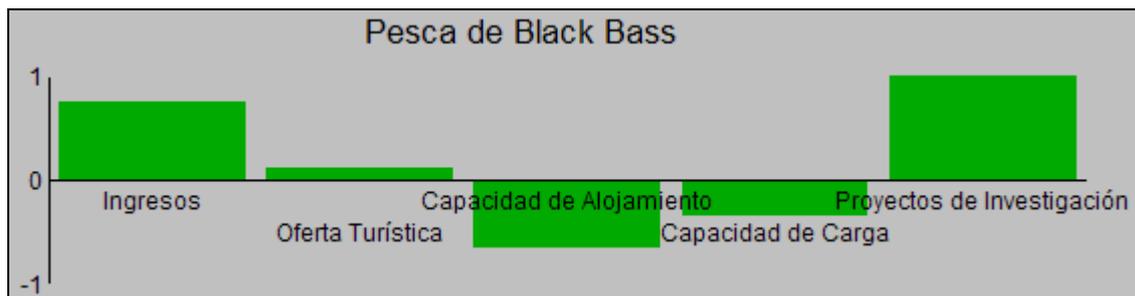


Gráfico 2: Perfil de Pesca de Black Bass.

Para la alternativa excursión a San Felipe, los valores que se corresponden con los criterios oferta turística y proyectos de investigación son los únicos que están por debajo de la media, o sea, salvo estos dos criterios, todos los demás se representan por encima de 0 como se muestra en el *Gráfico 3*.

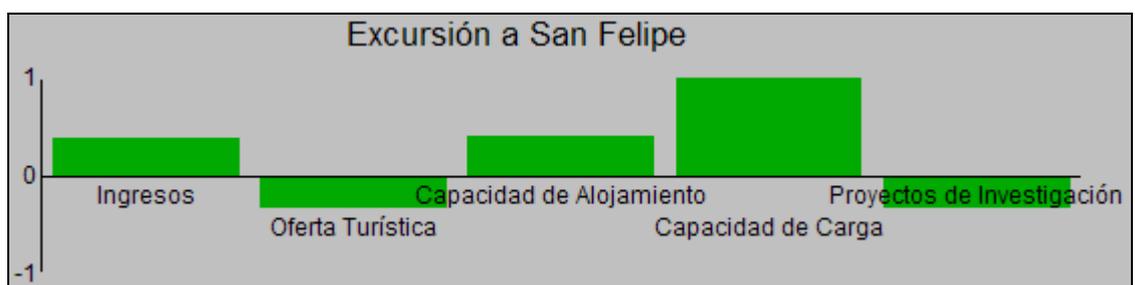


Gráfico 3: Perfil de Excursión a San Felipe.

Con ánimo de fortalecer el análisis, se emplea el proceso GAIA, que consiste en un módulo de interacción visual complementario de la metodología PROMETHEE (Mareschal and Brans, 1988). Este ofrece al decisor una descripción gráfica de su problema de decisión, enfatizando los conflictos existentes entre los criterios y el impacto de los pesos en la decisión final. Hace el análisis más descriptivo y está orientado de forma gráfica (Fernández, 2002).

El análisis GAIA completo se fundamenta en el análisis de los flujos netos obtenidos a partir de la descomposición del flujo neto global. En comparación con las evaluaciones de los criterios, los flujos unicriterio contienen mayor cantidad de información acerca de la estructura de preferencias del decisor debido al uso de las funciones de preferencia. Estos flujos están expresados en escalas similares y son independientes de las escalas originales de los criterios.

Así pues, se proyectan sobre el plano GAIA los puntos que representarán las alternativas y los vectores unitarios de los ejes de coordenadas que representarán a los criterios. Este análisis permite distinguir qué alternativas son buenas bajo cada criterio, dado que estas se localizarán en la misma dirección del eje correspondiente en el plano.

Se tendrá en cuenta que los criterios representados por ejes en direcciones similares expresan preferencias afines, mientras que los que se encuentran en direcciones opuestas corresponden a criterios en conflicto entre sí. Se debe destacar además, que la longitud de los ejes que representan a cada uno de los criterios es en sí una medida del poder de discriminación relativo de los criterios respecto al conjunto de alternativas. El *Gráfico 4* permitirá sacar las conclusiones con respecto al problema que se trata.

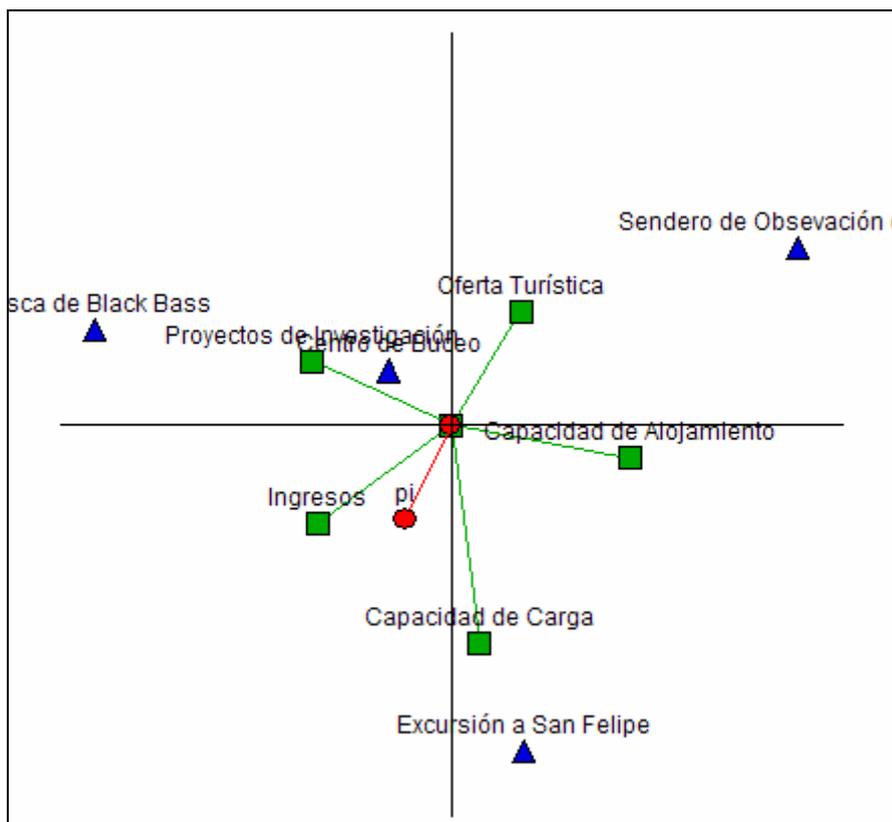


Gráfico 4: Plano GAIA.

Los criterios ingresos y oferta turística se encuentran en conflicto, pues sus vectores unitarios tienen direcciones opuestas. Otro par de criterios que están en conflicto también son la capacidad de alojamiento y el criterio número de proyectos de investigación.

La calidad de la información que se brinda está directamente relacionada con el porcentaje Λ , que explica la cantidad de información que conserva el plano. En este caso el plano explica el 78,16% de la variabilidad. Como Λ es superior al 70%, podemos concluir con la afirmación de que el plano ofrece una representación bastante fiable del problema de decisión al que se enfrenta el Grupo Extrahotelero Palmares.

Luego de realizar el análisis gráfico se lleva a cabo un análisis de sensibilidad de los pesos. Esto da la posibilidad de determinar en qué intervalo se pueden mover los pesos sin que varíe la solución del problema. Se debe destacar que cuanto mayor sean estos intervalos, más estables serán las ponderaciones.

Criterios	Absolute values			Relative values (%)		
	Weight	Min	Max	Weight	Min	Max
Ingresos	0.1630	0.1058	11.742	16.28%	11.21%	66.02%
Oferta Turística	0.1630	0.1132	0.4449	16.28%	11.90%	34.68%
Capacidad de Alojamiento	0.1090	0.0000	0.1303	10.89%	0.00%	12.75%
Capacidad de Carga	0.2830	0.0481	0.2996	28.27%	6.28%	29.44%
Proyectos de Investigación	0.2830	0.2664	Infinity	28.27%	27.06%	100.00%

Tabla 10: Intervalos de estabilidad para las ponderaciones.

La Tabla 10 indica que el peso correspondiente al criterio capacidad de alojamiento puede disminuir hasta cero y ello no provocaría cambios en la solución, sin embargo, un aumento de tan solo 0.0213 unidades del valor original de su peso, provocaría variación en la solución. En cambio, el criterio proyectos de investigación solo puede disminuir el valor de su peso en 0,0166 sin que ello varíe la solución; pero, puede aumentar de forma indefinida y el resultado sería el mismo.

Como se ha comentado anteriormente, la metodología AHP permite ordenar las alternativas de acuerdo a las preferencias del centro decisor, por ello, decidimos aplicar este método a nuestros datos con el fin de establecer comparaciones en cuanto a la solución obtenida. Como se aprecia en el Gráfico 5, la ordenación de las alternativas es similar a la anterior, pues el producto “Pesca recreativa de Black Bass” es el de mayor valoración y el resto de las alternativas se muestra en el mismo orden de preferencia que se obtuvo al aplicar el método PROMETHEE alcanzándose un índice de inconsistencia de 0.04, lo que nos permite calificar de aceptable el resultado obtenido.

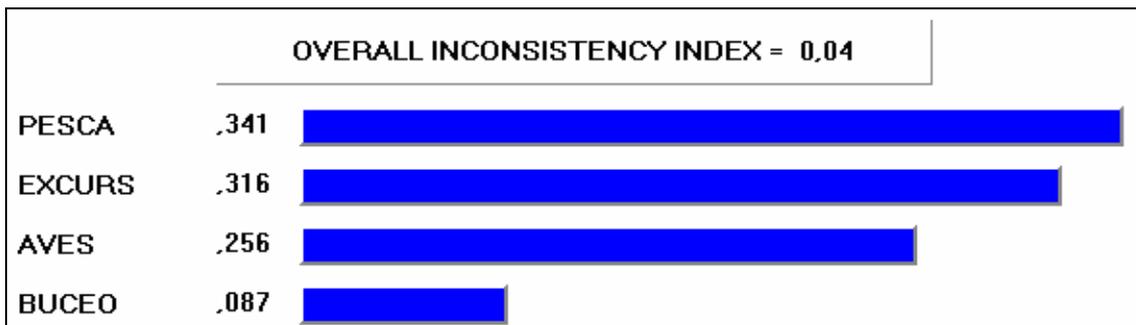


Gráfico 5: Ordenación de las alternativas mediante AHP.

El problema fue abordado además desde la Teoría de la Utilidad Multiatributo (MAUT), con el fin de realizar un análisis más consistente. Esta teoría ha sido desarrollada especialmente por Keeney y Raiffa (1976, 1993), a partir de la teoría de utilidad unidimensional de Von Neumann y Morgenstern (1944), con el fin de expresar las

preferencias del centro decisor en términos de la utilidad que le reporta. En particular, empleamos la función de utilidad aditiva. La valoración de las distintas alternativas (función de utilidad), en este caso, resulta de sumar las contribuciones de cada uno de los atributos considerados, adecuadamente ponderados en función de su importancia.

Al convertir los datos del problema en utilidades se debe tener en cuenta que los objetivos pretenden ser maximizados, por lo que los valores más altos que se puedan obtener en cada uno de los criterios representan una mayor utilidad para el centro decisor, de igual modo, la peor utilidad la reportan aquellos criterios con valores más bajos. De este modo, consideramos que las funciones de utilidad del decisor, para todos los criterios, eran lineales, con valores comprendidos en el intervalo $[0,10]$, teniendo un mínimo en el valor 0 y un máximo en el valor 10.

El *Gráfico 6* proporciona la ordenación mediante los valores de las sumas ponderadas para cada criterio, en el cual la suma ponderada de mayor valor corresponde a la alternativa “Pesca de Black Bass”, que nuevamente sobresale como el mejor producto turístico a implementar. El orden de las restantes soluciones se comporta de igual forma que en los resultados hallados con anterioridad.

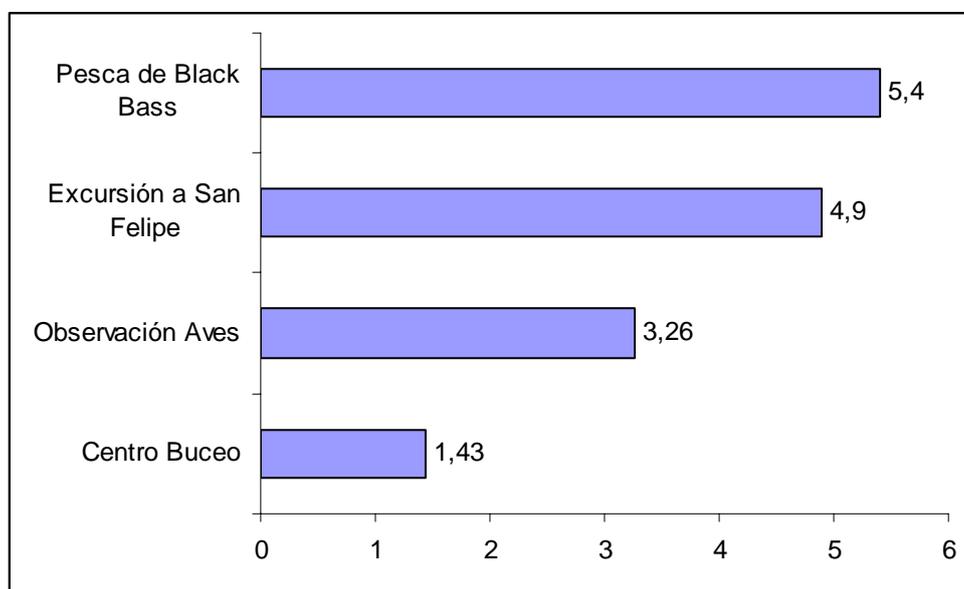


Gráfico 6: Ordenación mediante MAUT.

4. Conclusiones

El trabajo que se presenta fue realizado con el objetivo de solucionar un problema surgido en la Delegación del Ministerio de Turismo en Pinar del Río, a la hora de seleccionar, de un grupo de productos turísticos, aquel que será el primero en ponerse en marcha para que sea explotado de forma sostenible y garantice el aumento de la oferta turística, genere los ingresos necesarios para desarrollar los restantes y contribuya a incrementar las pernoctaciones de los visitantes a Pinar del Río. Para este proceso de selección se han utilizado técnicas multicriterio.

Tras presentar los productos turísticos y los indicadores de tipo económico y ecológico de importancia para el centro decisor se empleó el programa Expert Choice para encontrar las ponderaciones o los pesos que tienen cada criterio para el decisor y, se aplicaron varios métodos para encontrar la mejor alternativa, en concreto, el método PROMETHEE, la metodología AHP, que también posibilita ordenar las alternativas y, por último, le Teoría de la Utilidad Multiatributo (MAUT). Con todos ellos, la solución a que se llegó fue similar.

La ordenación obtenida fue presentada al Grupo extrahotelero Palmares, el encargado de los nuevos productos en el Ministerio del Turismo en Pinar del Río y es evidencia de la gran ayuda que proporciona la aplicación de las técnicas empleadas en la búsqueda de soluciones a los problemas reales que se presentan en las organizaciones.

Bibliografía

Brans, J.P., Mareschal, B. y Vincke, Ph. (1984): "PROMETHEE: A new family of outranking method for MCMD" en Brans, J.P. Edi. *Operational Research '84*. North Holland, pp. 447-490.

Brans, J.P., Mareschal, B. y Vincke, Ph. (1986): "How to select and how to rank projects: The PROMETHEE method", *European Journal of Operational Research*, 24, pp. 228-238.

Brans, J.P. y Vincke, Ph. (1985): "A preference ranking organization method: The PROMETHEE method for MCDM", *Management Science*, 31, 6, pp. 647-656.

Briggs, T., Kunsh, P.L. y Mareschal, B. (1990): "Nuclear waste management: An application of the multicriteria PROMETHEE method", *European Journal of Operational Research*, 4, 1, pp.1-10.

Caballero, R., González, M., Guerrero, F. M., Molina, J. y Paralera, C. (2004): "Problema de localización de incineradoras de residuos MER bajo un enfoque

multicriterio: Una aplicación al caso andaluz”, *Anales de Economía Aplicada: XVIII Reunión Anual de León 2004*.

Camargo, I. A. (2003): *Aportaciones Metodológicas para la Implementación del Turismo Rural en la República de Cuba. Estudio de Caso: Zona Turística de Viñales*. Tesis (en opción al grado científico de Doctor en Ciencias Geográficas). Universidad de Alicante.

Choi, H. C. y Sirakaya, E. (2005): “Sustainability indicators for managing community tourism”, *Tourism Management* (In Press).

CITMA (2006): *Plan de Ciencia y Técnica del Ministerio de Ciencia Tecnología y Medio Ambiente*. Delegación Territorial Pinar del Río, Cuba.

Comité Espeleológico de Pinar del Río (2006): *Informe Anual del Comité Espeleológico Provincial a la Junta Directiva Nacional*. Pinar del Río, Cuba.

Cubanacán (2006): *Catálogo de Cubanacán Viajes*. Agencia de viajes Cubanacán, Pinar del Río. Cuba.

Current, J., Min, H. y Echilling, D. (1990): “Multiobjective analysis of facility locations decisions”, *European Journal of Operational Research*, 49, pp.295-307.

Daymond, J, S. (1997): “Indicators of sustainable tourism in New Zeland: A goverment perspective”, *Journal of Sustainable Tourism*, 5, pp. 279-293.

Du Bois, P., Brans, J.P., Cantraine, F. y Mareschal, B. (1989): “MEDICIS: An expert system for computer-aided diagnosis using PROMETHEE multicriteria method”, *European Journal of Operatinal Research*, 39, pp. 284-292.

Empresa Nacional de Flora y Fauna. Informe Anual de Investigación (2006).

Empresa Nacional de Investigaciones Aplicadas (MICONS). (2006): Archivo Técnico.

Farsari, Y. y Prastacos, P. (2000): "Sustainable tourism indicators. Case-study for the municipality of Hersonissos", *Proceedings of The International Scientific Conference on "Tourism on Islands and Specific Destinations"*, University of the Aegean, Chios.

Farsari, Y. y Prastacos, P. (2001): “Sustainable tourism indicators for Mediterranean established destinations”, *Tourism Today*, 1, 1, pp.103-121.

Farsari, Y. y Prastacos, P. (2002): Sustainable development indicators: An overview. <http://www.iac.mforth.gr/regional/papers/Asteras-English.pdf> Acceso 3/2/07.

Fernández, G. y Martín, M. A. (1999): “La decisión multicriterio en la elección del emplazamiento de un parque temático. Una aplicación práctica”. *VII Jornadas ASEPUMA*.

Fernández Barberis, G. M. (2002): “Los métodos PROMEHEE: Una metodología de ayuda a la toma de decisiones multicriterio discreta”. *Toma de Decisiones con Criterios Múltiples*, 1, pp. 5-20.

Hardi, P. y Barg, S. (1997): *Measuring Sustainable Development: Review of Current Practice*, Occasional Paper N. 17, Industry Canada, Ontario.

Hokkanen, J. y Salminen, P. (1997): “Locating a waste treatment facility by multicriteria analysis”, *Journal of Multicriteria Decision Analysis*, 6, 3, pp. 175-184.

- Ivars, J. A. (2001): *Planificación y Gestión del Desarrollo Turístico Sostenible: Propuestas para la Creación de un Sistema de Indicadores*. Universidad de Alicante, España. 1. Documento de trabajo.
- Karkaziz, J. (1989): "Facilities location in a competitive environment: A PROMETHEE based multiple criteria analysis", *European Journal of Operational Research*, 42, 3, pp. 294-304.
- Keeney, L. y Raiffa, H. (1976, 1993): *Decisions with Multiple Objectives: Preferences and Value Tradeoffs*. 1ª edn. Wiley, New York; 2ª edn. Cambridge University Press, Cambridge, GB.
- Mareschal, B. y Brans, J.P. (1988): "Geometrical representation for MCDM, The GAIA procedure", *European Journal of Operational Research*, 34, pp. 69-77.
- Mathienson, A. y Wall, G. (1992): *Tourism. Economical, Physical and Social Impacts*. Essex.
- Mladineo, N. y Grabovac, J. (1988): "The application of multicriteria analysis in the selection of the optimal renewable energy sources for tourist facilities". *Proceedingszbornik Radova, Yugoslavia*, pp. 110-121.
- OMT (2005): *Indicadores de Sostenibilidad para los Destinos Turísticos. Guía Práctica*, edn. Madrid, España: Organización Mundial del Turismo.
- OMT (2006): Barómetro OMT del Turismo Mundial, 4, 2. Madrid, España.
- ONAT (2006): *Informe Anual sobre Arrendadores*. Oficina Nacional de Administración Tributaria. Pinar del Río. Cuba.
- ONE (2006): *Anuario Estadístico Cuba 2005* edn. 2006. La Habana, Cuba.
- OTE (2006): *Anuario Estadístico. Pinar del Río 2005* edn. Pinar del Río, Cuba.
- Palmer, A., Beltrán, M. y Cortiñas, P. (2006): "Robust estimators end bootstrap confidence intervals applied to tourism spending", *Tourism Management*, 27, pp. 42-50.
- Queiruga, D., González -Benito, J., Walter, G. y Spengler, T. (2004): *Propuesta de un Método para la Elección de Buenas Alternativas para la Localización de Empresas de Reciclaje de RAEE en España*. Documento de Trabajo. "Nuevas tendencias en dirección de empresas".
- Quintana, R., Figuerola M., Chirivella, M., Lima, D., Figueras, M. A. y García, A. (2005): *Efectos y Futuro del Turismo en la Economía Cubana*, edn. Montevideo, Uruguay: Tradinco S.A.
- Saaty, T.L. (1977): "A scaling method for priorities in hierarchical structures", *Journal of Mathematical Psychology*, 15, pp. 234-281.
- Stuart, J. D. (1997): "Indicators of sustainable tourism in New Zealand: A local government perspective", *Journal of Sustainable Tourism*, 5, 4, pp.279-293.
- Von Neumann, J. y Morgenstern, O. (1944): *Theory of Games and Economic Behaviour*. Princeton University Press, New Jersey.
- Vuk, D., Kozelj, B. y Mladineo, N. (1991): "Application of multicriterial analysis on the selection of the location for disposal of communal waste", *European Journal of Operational Research*, 55, 2, pp. 211-217.

World Bank (1997): *Expanding the Measure of Wealth. Indicators of Environmentally Sustainable Development*, Washington D.C., The World Bank.

DIFERENCIAS DE GÉNERO EN LA TEORÍA NEOCLÁSICA DE LA DEMANDA: UN ANÁLISIS EXPERIMENTAL

José Carlos Rodríguez Alcantud
Facultad de Economía y Empresa
Universidad de Salamanca
e-mail: jcr@usal.es

Luis Daniel López Matos
Facultad de Ciencias Sociales y Jurídicas
Universidad Católica de Ávila
e-mail: daniel.lopez@ucavila.es

Carlos Rodríguez Palmero
Facultad de Ciencias Económicas y Empresariales
Universidad de Valladolid
e-mail: cpalmero@eco.uva.es

ÁREA TEMÁTICA: Métodos cuantitativos

Resumen

Este trabajo examina la existencia de diferencias de género en el comportamiento de un consumidor. Los resultados revelan que las mujeres presentan un mayor grado de eficiencia en sus decisiones de consumo que los hombres. Sin embargo, no existen diferencias significativas de género entre aquellos individuos que conocen los postulados de la teoría neoclásica de la demanda.

Palabras clave: Economía Experimental, diferencias de género, preferencia revelada, GARP, Afriat, comportamiento del consumidor, experimento de Montecarlo.

Abstract

This paper examines the existence of gender differences concerning the consumer's behaviour. The results reveal that women show a higher degree of coherence in their consumer decisions than men. Nevertheless, there are no significant gender differences regarding the individuals that know the postulates of the neoclassic demand theory.

Keywords: Experimental economics, gender differences, revealed preference, GARP, Afriat, consumer's behaviour, Montecarlo Experiment.

1. Introducción

La investigación en ciencias sociales con excepción de la economía indica diferencias¹ substanciales de comportamiento entre los hombres y las mujeres. La literatura general de la psicología contiene muchos estudios primarios y meta-analíticos que presentan evidencias de género en comportamiento social, sexual, motor, actitudes, capacidad cognoscitiva, toma de decisiones y rasgos de la personalidad. En una revisión extensa de la literatura, Eagly (1995, p. 148) concluye que los psicólogos “están en general de acuerdo que sus resultados meta-analíticos encuentran evidencias de diferencias de género”. Los reexámenes feministas recientes de la teoría económica realizados por Ferber (1995), Nelson (1995) y Strober (1994) han conducido a los economistas experimentales a estudiar la existencia de diferencias de género en el comportamiento económico de un individuo y plantearse la siguiente pregunta: ¿existen diferencias significativas de género en la toma de decisiones económicas?

En la literatura económica existen varias evidencias que sugieren la existencia de diferencias de género en el comportamiento económico de un individuo. Por ejemplo, en experimentos de financiación de bienes públicos se aprecia una clara diferencia de comportamiento entre los hombres y las mujeres (v. Brown-Kruse y Hummels (1993), Nowell y Tinkler (1994) y Seguíno et al. (1996)), siendo los resultados contradictorios en cuanto a qué género exhibe un mayor comportamiento free-riding. Resultados similares se exhiben en diversos estudios empíricos que analizan la existencia de diferencias de género en las inversiones (v. Jianakoplos y Bernasek (1998)), en la toma de decisiones bajo incertidumbre (v. Schubert et al. (1999) y Powell y Ansic (1999)) experimentos del juego del ultimátum o del dilema del prisionero (v. Eckel y Grossman (2001) y Frank et al. (1993)). Sin embargo, existen otros estudios experimentales que no encuentran ninguna diferencia significativa de género en el comportamiento altruista de un individuo (v. Isaac et al. (1985), Stockard et al. (1988); Cadsby y Maynes (1998,

¹ En psicología existen diversos estudios que muestran una clara diferencia de género en el comportamiento moral o en el comportamiento altruista (Eagly y Crowley (1986), Gilligan (1982), Uesugi y Vinacke (1963) y Vinacke (1959)). En sociología, los estudios encuentran diferencias de género en cuestiones de criminalidad y uso ilícito y prescrito de la droga (Cooperstock y Pamell (1982), Gottfredson e Hirschi (1990) y Kandel y Logan (1984) entre otros). En teoría política diversos estudios muestran una clara diferencia de género en el comportamiento político de hombres y mujeres. El comportamiento de la votación de las mujeres está más influenciado por cuestiones sociales que el de hombres (Baxter y Lansing (1983) y Goertzel (1983) entre otros).

2005)) o en el grado de aversión al riesgo (v. Schubert et al. (1999) y Powell y Ansic (1999)). Eckel y Grossmann (1988) y Powell y Ansic (1997) afirman que estos resultados contradictorios pueden ser consecuencia de la existencia de fallos metodológicos en el diseño experimental. La existencia de diferentes fenómenos cognitivos como el anclaje o el marco de referencia focal y la experiencia pueden influir en la aparición de diferencias de género en el comportamiento económico de un individuo (v. Bromiley y Curley (1992), Eagly (1995) y Unger (1990) entre otros).

En este trabajo proporcionamos los resultados de dos experimentos económicos de la teoría del consumidor que han sido diseñados para responder a las siguientes preguntas: ¿existen diferencias significativas de género en el grado de coherencia de las decisiones económicas de un individuo? En caso afirmativo, ¿son las mujeres consumidores más eficientes? La teoría pura del consumidor tiene por objeto estudiar la estructura de las decisiones de consumo realizadas por un consumidor racional cuando éste se enfrenta con diversos precios y rentas monetarias. El enfoque de la preferencia revelada de Samuelson (1938, 1948) y Houthakker (1950) postula la racionalidad de la conducta de un consumidor directamente en términos de su función de demanda. La economía experimental permite reproducir las condiciones ideales que propugna la teoría neoclásica de la demanda, dado que en un laboratorio se pueden crear unas condiciones muy próximas a aquellas condiciones ideales del modelo teórico del comportamiento optimizador del consumidor, y tener una elevada fiabilidad de los diferentes tests no paramétricos utilizados para analizar tanto la estructura de las decisiones de consumo de un individuo como la posible existencia de diferencias de género en el proceso de elección de un consumidor individual. Como consecuencia de ello, podemos afirmar que una interpretación experimental de la teoría de la preferencia revelada es consistente con la metodología positivista de las contribuciones formuladas por Samuelson en sus artículos seminales y que la economía experimental proporciona el marco idóneo para analizar la posible existencia de diferencias de comportamiento entre las mujeres y los hombres en sus decisiones de consumo.

La organización del trabajo es la siguiente. En la sección segunda proporcionamos un breve resumen del marco teórico de la teoría de la preferencia revelada asociada a un conjunto finito de datos de demanda en un contexto competitivo. En la sección tercera presentamos el marco teórico del análisis de la sensibilidad de los diferentes tests no paramétricos de la preferencia revelada utilizados en el estudio empírico del comportamiento de un consumidor. En la sección cuarta presentamos el diseño experimental del experimento realizado. En la sección quinta presentamos los resultados de nuestro experimento y realizamos una evaluación de los mismos.

2. Axiomas de la preferencia revelada

La teoría de la preferencia revelada postula la racionalidad de la conducta de un consumidor directamente en términos de ciertas condiciones algebraicas, conocidas como condiciones de preferencia revelada. En su versión reducida, la teoría de la preferencia revelada se limita a un conjunto de datos observables de precios y cantidades demandadas en un contexto competitivo, intentando descubrir todas las implicaciones empíricas de la hipótesis de consistencia del individuo. Supongamos que hay k bienes y que el consumidor puede elegir eventualmente cualquier cantidad no negativa de cada bien. Un vector² de demanda $\bar{x} = (x_1, \dots, x_k)$ es un vector de k componentes cuya componente i -ésima especifica la cantidad del bien i que se ha demandado. Designaremos los precios de mercado de los bienes mediante un vector de precios $\bar{p} = (p_1, \dots, p_k)$ tal que $\bar{p} \in \mathfrak{R}_{++}^k$. Nuestro punto de partida será la observación de cualquier serie finita de datos de demanda (\bar{p}_t, \bar{x}_t) , $t = 1, \dots, n$, donde cada par (\bar{p}_t, \bar{x}_t) expresa la cesta \bar{x}_t demandada a precios normalizados \bar{p}_t . La racionalidad de esta serie de observaciones se medirá por el grado de cumplimiento de los siguientes postulados de racionalidad.

² Para todo $k \in \mathbb{N}$ y para cada $\bar{x}, \bar{y} \in \mathfrak{R}^k$, “ $\bar{x} \geq \bar{y}$ ” significa que $x_i \geq y_i$ para todo i , “ $\bar{x} > \bar{y}$ ” significa que $x_i \geq y_i$ para todo i y $x_{i_0} > y_{i_0}$ para al menos un i_0 y “ $\bar{x} \gg \bar{y}$ ” significa que $x_i > y_i$ para todo i . Denotamos por $\mathfrak{R}_+^k = \{\bar{x} \in \mathfrak{R}^k : \bar{x} \geq \bar{0}\}$ y $\mathfrak{R}_{++}^k = \{\bar{x} \in \mathfrak{R}^k : \bar{x} > \bar{0}\}$

Definición 1 Sean \bar{x}_t y \bar{x}_s cestas demandadas a precios (normalizados) \bar{p}_t y \bar{p}_s respectivamente. Se dice que el consumidor:

- a) revela directamente que prefiere \bar{x}_t a \bar{x}_s , y se denota por $\bar{x}_t R^0 \bar{x}_s$, si $\bar{p}_t \bar{x}_t \geq \bar{p}_s \bar{x}_s$.
- b) revela que prefiere \bar{x}_t a \bar{x}_s , y se denota por $\bar{x}_t R \bar{x}_s$, si existen $\bar{x}_{t_1}, \dots, \bar{x}_{t_k}$ satisfaciendo $\bar{x}_t R^0 \bar{x}_{t_1}, \bar{x}_{t_1} R^0 \bar{x}_{t_2}, \dots, \bar{x}_{t_k} R^0 \bar{x}_s$.

Samuelson (1948) propone una condición necesaria para que una serie finita de observaciones de consumo sea racionalizable por una función de utilidad, el axioma débil de la preferencia revelada. En el contexto en que nos hallamos este postulado se traduce en lo siguiente:

Definición 2 (Axioma Débil de la Preferencia Revelada) El conjunto de observaciones de consumo $\{(\bar{p}_t, \bar{x}_t)\}_{t=1}^n$ satisface el Axioma Débil de la Preferencia revelada (WARP) si para todo par de observaciones i, j se cumple:: $\bar{x}_i R^0 \bar{x}_j$, y además, $\bar{x}_i \neq \bar{x}_j$, entonces $\bar{p}_j \bar{x}_j < \bar{p}_i \bar{x}_i$.

Houthakker (1950) demuestra que una condición necesaria y suficiente para que la función de demanda se derive de la relación de preferencia de un consumidor es que dicha función satisfaga el llamado axioma fuerte de la conducta del consumidor, también denominado axioma fuerte de la preferencia revelada.

Definición 3 (Axioma Fuerte de la Preferencia Revelada) El conjunto de observaciones de consumo $\{(\bar{p}_t, \bar{x}_t)\}_{t=1}^n$ satisface el Axioma Fuerte de la Preferencia revelada (SARP) si para todo par de observaciones i, j se cumple:: $\bar{x}_i R \bar{x}_j$, y además, $\bar{x}_i \neq \bar{x}_j$, entonces $\bar{p}_j \bar{x}_j < \bar{p}_i \bar{x}_i$.

El axioma fuerte de la preferencia revelada es, claramente, una generalización del axioma débil. Rose (1958) y Afriat (1965) demuestran la equivalencia entre el axioma débil y el axioma fuerte de la preferencia revelada en un espacio de consumo formado por dos bienes. Sin embargo, esta equivalencia no es cierta en espacios de consumo con más de dos bienes. Gale (1960) y Kamiya (1963) presentan dos contraejemplos que muestran la no equivalencia entre WARP y SARP. Varian (1982) define una nueva condición algebraica de preferencia revelada, el axioma generalizado de la preferencia

revelada³ (GARP), necesaria y suficiente para que un conjunto finito de observaciones de consumo sean consistentes con la hipótesis de la maximización de la utilidad.

Definición 3 *El conjunto de observaciones de consumo $\{(\bar{p}_t, \bar{x}_t)\}_{t=1}^n$ satisface el Axioma Generalizado de la Preferencia revelada (GARP) si para todo par de observaciones i, j se cumple: $\bar{x}_i R \bar{x}_j$, entonces $\bar{p}_j \bar{x}_j < \bar{p}_i \bar{x}_i$.*

Si los datos satisfacen GARP se puede construir una función de utilidad $u(\cdot)$ no saciada, continua, monótona y cóncava que racionaliza los datos en el sentido que para toda observación t , $u(\bar{x}_t) \geq u(\bar{x})$ para todo \bar{x} tal que $\bar{p}_t \bar{x}_t \geq \bar{p}_t \bar{x}$. La principal diferencia entre GARP y SARP radica en que GARP permite la multivaluación de la función de demanda y no exige la convexidad estricta de la relación de preferencia del consumidor que racionaliza la conducta observada. Varian (1982) presenta un procedimiento, basado en el algoritmo de la clausura transitiva de Warshall, que permite contrastar empíricamente en un tiempo de ejecución proporcional a n^3 tanto la consistencia de un conjunto finito de datos con la teoría neoclásica de la demanda como las observaciones de consumo inconsistentes⁴ con GARP y, además, diseña un procedimiento constructivo⁵ que permite obtener explícitamente esta función de utilidad.

3. Fiabilidad de los Tests No-paramétricos GARP

Diversos autores han contrastado empíricamente la consistencia de diferentes conjuntos de datos con la hipótesis de la maximización de la utilidad desde una perspectiva no paramétrica (v. Landsburg (1981), Sippel (1997) y Mattei (1994) entre otros). Utilizando series temporales de datos de consumo agregado, Landsburg (1981) y Varian (1982) encuentran que los datos satisfacen GARP. Varian (1982) afirmó que los datos de consumo agregados muy probablemente satisfacen las condiciones de preferencia revelada por una cuestión trivial: “the changes expenditure from year to year were large relative to the changes in relative prices. Hence budget sets rarely intercepted in ways

³ Este axioma es equivalente a la propiedad de la consistencia cíclica propuesta por Afriat en 1967.

⁴ Sea $\{(\bar{p}_t, \bar{x}_t)\}_{t=1}^n$ un conjunto finito de observaciones de demanda. Se dice que \bar{x}_t es inconsistente con GARP si y sólo existe \bar{x}_s satisfaciendo $\bar{x}_s R \bar{x}_t$ y $\bar{x}_t P^0 \bar{x}_s$ (v. Varian; 1982).

⁵ Demostraciones alternativas de este resultado pueden verse en Afriat (1967), Diewert (1973) y Scarf, Tood y Fostel (2004) entre otros.

that would generate a GARP violation”. Bronars ofreció en 1987 una contribución novedosa en la investigación de la sensibilidad de los diferentes tests no paramétricos de la preferencia revelada. Bronars elige la hipótesis de Becker (1962) de elección aleatoria sobre un conjunto presupuestario para realizar un análisis ex-ante de las diferentes axiomas de preferencia revelada. Este autor, utilizando los datos de Varian (1982), encontró que el 67% de los consumidores con comportamiento aleatorio violaron GARP, contrariamente a lo planteado por Varian.

Becker (1962) y Chant (1963) proponen el primer modelo de comportamiento aleatorio de un consumidor. En este modelo, denominado *random money choice*, el consumidor agota toda su renta monetaria en la compra de los distintos bienes que forman el conjunto de consumo. El consumidor se gasta cada unidad de dinero en un cierto bien con una cierta probabilidad⁶. De este modo, si β_{ij} es la probabilidad de elegir el bien i en la situación j , m_j es el nivel de renta monetaria disponible por el consumidor y x_{ij} la cantidad demandada del bien i en la situación j se tiene que:

$$x_{ij} = \beta_{ij} \cdot \frac{m_j}{p_{ij}} \quad \beta_{ij} > 0, \quad \sum_{i=1}^K \beta_{ij} = 1, \quad j = 1, \dots, n \quad (1)$$

Por su parte, Battalio, Dwyer y Kagel (1987) plantean un modelo de comportamiento aleatorio del consumidor, *random goods choice*, donde el consumidor en cada situación presupuestaria selecciona una cantidad aleatoria, x_{ij} , de cada uno de los bienes del conjunto de consumo. Si β_{ij} es la proporción aleatoria asociada al bien i en la situación j y m_j el nivel de renta monetaria en la situación j -ésima se tiene

$$x_{ij} = \beta_{ij} \cdot \frac{m_j}{\sum_i \beta_{ij} p_{ij}} \quad \left(\beta_{ij} > 0, \quad \sum_i \beta_{ij} = 1, \quad i = 1, \dots, n \right) \quad (2)$$

como ecuación para la función de demanda.

Battalio, Dwyer y Kagel (1987) plantean que los consumidores pueden necesitar consumir una cantidad mínima de cada uno de los bienes del conjunto de consumo,

⁶ La simulación de cualquier modelo de comportamiento aleatorio de un consumidor requiere de una gran cantidad de números aleatorios y, en general, de sucesiones de realizaciones de variables aleatorias. En este trabajo hemos utilizado el generador de números aleatorios Mersenne-Twister, propuesto por Matsumoto y Nishimura en 1998, que posee un período de $2^{19937} - 1$ y que permite que la sucesión de pseudo números aleatorios generada se comporte como una sucesión de variables aleatorias continuas e independientes con distribución uniforme en el intervalo real (0,1).

donde este consumo mínimo puede estar determinado psicológicamente o ser necesario para la supervivencia del individuo. El *Generalized Minimum Needs Hypothesis* o modelo generalizado de necesidad mínima caracteriza este comportamiento aleatorio. En este modelo, la cantidad demandada del bien i en la situación presupuestaria j es:

$$x_{ij} = \gamma_i + \frac{\beta_{ij} p_{ij}^{-\delta}}{\sum_i \beta_{ij} p_{ij}^{1-\sigma}} \left(m_j - \sum_i \gamma_i p_{ij} \right) \quad \left(\beta_{ij} > 0, \sum_i \beta_{ij} = 1, i = 1, \dots, n \right) \quad (3)$$

Bronars (1987) propone un modelo de comportamiento aleatorio del consumidor donde el consumidor, en cada situación presupuestaria, se gasta aleatoriamente un porcentaje de su renta monetaria en cada uno de los bienes del conjunto de consumo. Si $\left(\beta_{ij} / \sum_i \beta_{ij} \right) \cdot m_j$ es el porcentaje aleatorio de dinero que se gasta el consumidor en el bien i en la situación j se tiene:

$$x_{ij} = \frac{\beta_{ij}}{\sum_i \beta_{ij}} \cdot \frac{m_j}{p_{ij}} \quad \beta_{ij} > 0, i = 1, \dots, n \quad (4)$$

como función de demanda. El test de Bronar's se ha utilizado en diferentes estudios experimentales de la teoría neoclásica de la demanda para analizar la fiabilidad de los diferentes test no paramétricos de la preferencia revelada desde una perspectiva ex-ante (v. Sippel (1997), Mattei (2000), Harbaugh (2001), Andreoni y Miller (2002) y Février (2004) entre otros).

En la siguiente tabla se resumen los anteriores modelos de comportamiento aleatorio del consumidor:

Modelo	Restricciones paramétricas			(x_{ij}) Función de demanda
	σ	γ	$m_j - \sum_i \gamma_i p_{ij}$	
Random money choice	1	0	–	$\beta_{ij} \cdot \frac{m_j}{p_{ij}}$
Random goods choice	0	0	–	$\beta_{ij} \cdot \frac{m_j}{\sum_i \beta_{ij} p_{ij}}$
Generalized minimum-needs	> 0	Ver texto	≥ 0	Ecuación (3)
Bronars' Test	1	0	–	$\frac{\beta_{ij}}{\sum_i \beta_{ij}} \cdot \frac{m_j}{p_{ij}}$

4. Diseño experimental

Los experimentos se realizaron los días 10, 11, 17 y 18 de mayo de 2005 en horas de clase, sin previo aviso, y en presencia del profesor de la asignatura. En el primer experimento participaron un total de 339 alumnos de la Universidad Católica de Ávila,

mientras que el segundo experimento lo realizaron 67 alumnos que cursan Economía y Administración y Dirección de Empresas en esta universidad. Nos presentamos en las clases y el profesor de la asignatura correspondiente nos presentó a los alumnos. Tras la conclusión de esta presentación, se informó a los alumnos que en este experimento se iba a simular el comportamiento racional de un consumidor, y se realizó una breve exposición del experimento. Posteriormente, se les preguntó si deseaban participar. Aquellos alumnos que mostraron un especial interés por participar se desplazaron a una sala informática. Una vez allí, se les proporcionó una hoja donde se describía el experimento y éstos podían analizar, de forma visual, los distintos atributos y características de los diferentes bienes de consumo con el objetivo de que cada participante se creara una imagen comprensible de cada uno de los bienes utilizados en el experimento. Los estudiantes realizaron sus elecciones de consumo de forma voluntaria y anónima, sin ninguna compensación económica y en un ambiente de seriedad y silencio con la única restricción de agotar toda la renta monetaria disponible en cada uno de las 20 y 15 situaciones presupuestarias correspondientes al primer y segundo experimento respectivamente. El participante, a diferencia de los experimentos realizados por Sippel (1997), Mattei (2000) y Février (2004), entre otros, no percibió ningún tipo de incentivo por participar en los experimentos y, además, no podía demandar una cantidad que no se correspondiese con las distintas unidades físicas de cada uno de los bienes.

En varios estudios empíricos se aprecia un comportamiento inconsistente con la teoría neoclásica de la demanda (v. Sippel (1997), Mattei (2000) y Février (2004) entre otros). Dicha inconsistencia⁷ puede haber estado ocasionada por la existencia de una variación temporal en las preferencias de los participantes, por la existencia de algún error de medida o por algún fenómeno cognitivo que puede influir sobre el comportamiento de un consumidor. Para efectuar sus elecciones de consumo los participantes dispusieron de una aplicación informática implementada en Visual C++, la cual no permitía al alumno gastarse una cantidad mayor o menor de la renta monetaria disponible en cada

⁷ Kahneman y Amos Tversky en numerosos trabajos han presentado pruebas sustanciales de que en un entorno de laboratorio los individuos, en ocasiones, se desvían extraordinaria y sistemáticamente de las predicciones del modelo estándar. Para ver un resumen de los problemas, debates y resultados en economía experimental véanse, entre otros, los trabajos de Kagel y Roth (1995), Binmore (1999), Loomes (1999) y Starmer (1999).

situación presupuestaria. Esta aplicación reproduce todos los requerimientos teóricos de la teoría del consumidor, eliminando, por tanto, cualquier error de medida o aleatorio y cualquier variación temporal de las preferencias de los participantes. Al finalizar las elecciones de consumo en cada situación presupuestaria, el alumno pulsaba una tecla >> para que aparecieran los nuevos precios y la nueva renta monetaria disponible correspondientes a la siguiente situación presupuestaria. Esta aplicación impide que se almacenen las decisiones de consumo hasta que el participante no haya acabado de completar todos los escenarios posibles. El participante en todo momento puede revisar y rehacer, si lo considera oportuno, sus decisiones. Una vez que el participante ha realizado sus elecciones de consumo en todos los escenarios posibles y revisado, si lo hubiese considerado oportuno, sus decisiones, éste pulsa “Aceptar” y la aplicación almacena al mismo tiempo sus elecciones de consumo. Además, nuestra aplicación ha sido implementada con el objetivo de intentar eliminar, en la medida de lo posible, diferentes fenómenos cognitivos como el anclaje⁸ o el marco de referencia focal⁹, dado que la aplicación ordena aleatoriamente los diferentes bienes de consumo y, además, en cada situación presupuestaria, el participante podía demandar en el orden que deseara los diferentes bienes del conjunto de consumo.

El conjunto de consumo estaba formado por 8 bienes: paquetes de cigarrillos (Marlboro, Fortuna, Ducados,...), copas (whisky, ron, vodka,...), entradas de cine, cenas, CD's R-W 700 MB (10X), DVD+R 4.7GB (8X) y frutos secos (pipas, maíces, cacahuetes, etc). Estos bienes de consumo intentan incorporar un atractivo emocional en los participantes en nuestro experimento, despertar su interés, así como cumplir una función de identificación con el grupo de referencia. Es decir, son productos que conllevan un mayor nivel de implicación y, como consecuencia de ello, un mayor esfuerzo cognitivo en la toma de decisiones.

⁸ Las respuestas se ven influidas por indicaciones en la pregunta.

⁹ El formato de las preguntas cambia la prioridad de diferentes aspectos de la teoría cognitiva.

Los precios de la primera situación presupuestaria fueron extraídos de la realidad con el objetivo de que éstos sirviesen como precios de referencia, tanto externos como interno¹⁰, a los participantes del experimento. Para cada una de las diferentes situaciones presupuestarias, con excepción de la última, hemos realizado una simulación de los precios de los distintos bienes con el objeto de aumentar el número de intersecciones de los diferentes conjuntos presupuestarios y, en consecuencia, incrementar la potencia de los diferentes tests no paramétricos utilizados para analizar la consistencia de las decisiones de consumo de un consumidor. Para contrastar la homogeneidad de grado cero de la función de demanda hemos considerado, de manera similar a lo realizado por Sippel (1997) los mismos precios y la misma renta monetaria en la situación presupuestaria inicial y final. Los precios simulados para cada uno de los bienes estaban limitados por unos precios reales de referencia mínimo y máximo¹¹, siendo la terminación de cada uno de ellos 0, 5 y 9, dada la evidencia del uso de estas terminaciones en los precios reales de las distintas economías. La renta monetaria varió entre los 24 y los 36 euros en cada escenario.

En el segundo experimento intentamos dificultar lo máximo posible el proceso de elección racional haciendo que tanto los precios de cada uno de los bienes como la renta monetaria variaran muy poco de un escenario a otro. El conjunto de consumo también estaba formado por 8 bienes: pipas, maíz gigante, almendras saladas, cacahuets con miel, revuelto de frutos secos, gominolas, avellanas y pistachos. La elección de estos productos estuvo fundamentada por dos razones principalmente: En primer lugar, por ser todos los bienes perfectamente divisibles. En segundo lugar, por la similitud en los precios reales de estos bienes entre los diferentes establecimientos. Los precios fueron simulados a partir de una matriz de compensación tipo Slutsky y estaban limitados por unos precios de referencia reales mínimo y máximo resultantes de una muestra realizada en la ciudad de Ávila a 15 establecimientos diferentes. Los precios de la primera situación presupuestaria, de manera análoga a lo realizado en el primer experimento, se

¹⁰ Los precios de referencia externos son los precios que recibe el consumidor a través de distintos canales de información externos y que le sirven para efectuar comparaciones, refiriéndose al precio de la marca líder o el más económico.

¹¹ En la mente de los consumidores existe generalmente una banda de precios aceptables para el nivel de calidad del bien o servicio al que enfrenta su decisión de compra. Los precios que no pertenecen a dicha banda por encontrarse fuera de ella, ya sea por ser demasiado bajos o demasiado altos, son considerados insatisfactorios y consecuentemente se rechazan.

correspondieron con los precios de estos bienes en uno de estos establecimientos de la muestra que fue seleccionado aleatoriamente. La terminación de cada uno de estos precios fue 0 y 5, dada la evidencia del uso de estas terminaciones en los precios reales de estos productos. La renta monetaria varió entre los 9.58 y los 10.36 euros en cada escenario, siendo 10 euros la renta monetaria de la primera situación presupuestaria.

5. Resultados empíricos

A continuación, realizamos un análisis descriptivo de los resultados obtenidos en nuestros experimentos. En el primer experimento observamos que el 67.2% de los participantes violaron GARP, mientras que el 46.2% de los participantes del segundo experimento también exhibió un comportamiento inconsistente con la hipótesis de la maximización de la utilidad. En numerosos estudios empíricos, como hemos señalado con anterioridad, se ha comprobado la inconsistencia de las decisiones de consumo de un consumidor con los diferentes axiomas de la preferencia revelada. Al hablar de inconsistencia nos referimos a que el comportamiento observado del consumidor viola el axioma generalizado de la preferencia revelada y, por consiguiente, sus observaciones de consumo no proceden del correspondiente comportamiento optimizador (v. Battalio (1973), Sippel (1997), Mattei (2000) y Février (2004) entre otros).

Desde el trabajo seminal de Afriat en 1973 sobre la eficiencia en el consumo, han aparecido en la literatura económica multitud de índices que tratan de determinar el nivel de coherencia de las decisiones de consumo realizadas por un consumidor (v. Afriat (1973), Varian (1990), Swofford & Whitney (1987), Famulari (1995) y Gross (1995) entre otros). Swofford & Whitney (1987) y McMillan & Amoako-Tuffour (1988) sugieren, como medida de eficiencia, comparar el número de observaciones de consumo que violan el axioma generalizado de la preferencia revelada con el número máximo posible. Famulari (1996), inspirándose en los trabajos de Swofford & Whitney (1987) y McMillan & Amoako-Tuffour (1988), proporciona un nuevo índice de eficiencia al comparar el número de violaciones del axioma generalizado con el número de lotes de bienes comparables por la relación de preferencia revelada indirecta asociada al conjunto de observaciones de consumo. Para incrementar la potencia del test anterior Famulari realiza una restricción del test anterior considerando sólo aquellas

observaciones de consumo cuyo gasto¹² sea similar. Resulta obvio que cuanto menor sea tanto el índice de Swofford & Withney y como el índice de Famulari, más eficiente será el consumidor en sus decisiones de consumo, y viceversa. Afriat (1973) y Varian (1990) definen una medida de “eficiencia parcial” que se puede interpretar como una medida estadística de la bondad del ajuste del modelo maximizador de la utilidad. Para todo conjunto $\{(\bar{p}_t, \bar{x}_t)\}_{t=1}^n$ de observaciones de consumo se considera un conjunto $e = (e_t)_{t=1}^n$ de índices de eficiencia tal que $0 < e_t < 1, \forall t = 1, \dots, n$. La relación de preferencia revelada directa o indirecta asociada a $(e_t)_{t=1}^n$ se define de la siguiente manera:

Definición 3 (Niveles de eficiencia) *Sea $(\bar{p}_t, \bar{x}_t), t = 1, \dots, n$, una serie finita de datos de demanda. Se dice que la cesta demandada \bar{x}_s se revela directamente preferida, al nivel de eficiencia e_t , a \bar{x}_t y se denota por $\bar{x}_t R^0(e_t) \bar{x}_s$ si $e_t \bar{p}_t \bar{x}_t \geq \bar{p}_t \bar{x}_s$.*

Se dice que \bar{x}_t se revela directa o indirectamente preferida a \bar{x}_s bajo el vector de “niveles de eficiencia” \bar{e} y se denota por $\bar{x}_t R(\bar{e}) \bar{x}_s$ si existen si existen $\bar{x}_{t_1}, \dots, \bar{x}_{t_k}$ satisfaciendo $\bar{x}_t R^0(e_{t_1}) \bar{x}_{t_1}, \bar{x}_{t_1} R^0(e_{t_2}) \bar{x}_{t_2}, \dots, \bar{x}_{t_k} R^0(e_{t_k}) \bar{x}_s$

Con estas definiciones podemos describir el siguiente axioma de racionalidad del comportamiento ante la demanda:

Definición 3 (Varian; 1990) *La serie finita de datos de demanda $(\bar{p}_t, \bar{x}_t), t = 1, \dots, n$, satisface GARP(\bar{e}) si para todo \bar{x}_t y \bar{x}_s tal que $\bar{x}_t R(\bar{e}) \bar{x}_s$ se tiene $e_s \bar{p}_s \bar{x}_s \geq \bar{p}_s \bar{x}_t$.*

Afriat (1973), a propuesta de H. Rubin, propuso usar una misma cota uniforme $0 < e^* < 1$ en todas las observaciones de consumo. El índice crítico e^* de coste de eficiencia¹³ (CCEI) se ha convertido tanto por su simplicidad en el cálculo como por su interpretación económica, en la medida de eficiencia más utilizada en la literatura económica para analizar la bondad del ajuste del modelo optimizador del consumidor. Varian (1990) propone utilizar un e_t distinto en cada observación de consumo y

¹² Sea $\{(\bar{p}_t, \bar{x}_t)\}_{t=1}^n$ un conjunto de observaciones de consumo. Se dice que el gasto asociado a \bar{x}_i y \bar{x}_j es similar si y sólo si $2 \cdot \left(\frac{m_i - m_j}{m_i + m_j} \right) \leq K$, donde $m_i = \bar{p}_i \bar{x}_i$ y $m_j = \bar{p}_j \bar{x}_j$.

¹³ Houtmann y Maks (1987) proponen un procedimiento basado en el método de la bisección que permite encontrar el mayor e^* donde los datos satisfacen GARP.

encontrar los $(e_t)_{t=1}^n$ que se aproximen más a uno (en alguna norma). La función de sobrecompensación de Samuelson permite construir esta medida de eficiencia:

Definición 7 Sea $\{(\bar{p}_t, \bar{x}_t)\}_{t=1}^n$ un conjunto de observaciones de consumo y R la relación de preferencia revelada indirecta sobre este conjunto. Se considera la función m de $\{(\bar{p}_t, \bar{x}_t)\}_{t=1}^n$ en \mathfrak{R}_+ definida por:

$$m(\bar{p}, \bar{x}, R) = \inf_{\bar{y}, \bar{x}} \bar{p}\bar{x}$$

A m se le denomina función generalizada de compensación.

El siguiente conjunto de “índices de eficiencia” permite determinar el grado de coherencia de las decisiones de consumo de un individuo, además de precisar qué observaciones son las causantes de esta falta de coherencia.

Definición 8 (Varian, 1990) Sea $\{(\bar{p}_t, \bar{x}_t)\}_{t=1}^n$ un conjunto de observaciones de consumo.

Llamaremos índice de eficiencia de Varian al conjunto de índices, $i = (i_t)_{t=1}^n$, tal que

$$i_t = \frac{m(\bar{p}, \bar{x}, R)}{\bar{p}_t \bar{x}_t}, \quad \forall t = 1, \dots, n \quad (5)$$

A partir del conjunto de índices de eficiencia $i = (i_t)_{t=1}^n$ se puede construir de manera natural la siguiente medida de la bondad del ajuste del modelo maximizador de la utilidad definida por: $i^* = \frac{1}{\sqrt{n}} |i|$. El conjunto de observaciones de consumo $\{(\bar{p}_t, \bar{x}_t)\}_{t=1}^n$ satisface GARP si y sólo si $i^* = 1$. Además, $0 < i^* < 1$ para todo conjunto de observaciones de consumo inconsistente con GARP. El consumidor será más eficiente cuanto mayor sea el valor de i^* .

La siguiente tabla muestra el número y porcentaje de los participantes inconsistentes del primer experimento para diferentes índices de Afriat. Además, muestra los resultados en promedio de las diferentes medidas de eficiencia consideradas.

EXPERIMENTO I.1						
PARTICIPANTES VIOLANDO GARP						
	$e^* = 1$	$e^* = 0,99$	$e^* = 0,98$	$e^* = 0,97$	$e^* = 0,96$	$e^* = 0,95$
Hombres	156 (73.58 %)	121 (57.07 %)	91 (42.92 %)	62 (29.24 %)	44 (20.75 %)	39 (18.40 %)
Mujeres	80 (62.49 %)	60 (47.24 %)	46 (36.22 %)	26 (20.47 %)	17 (13.38 %)	10 (07.87 %)
ÍNDICES DE EFICIENCIA						
	Varian's index	Famulari's index	Swofford's index	Afriat's index	obs. incons.	num. viol.
Hombres	0.9829	0.0613	0.0416	0.9773	5.2667	15.1801
Mujeres	0.9862	0.0505	0.0336	0.9834	4.3203	13.7969
PROMEDIO DE VIOLACIONES GARP						
	$e^* = 1$	$e^* = 0,99$	$e^* = 0,98$	$e^* = 0,97$	$e^* = 0,96$	$e^* = 0,95$
Hombres	15.1801	8.2275	5.2654	3.1801	1.7820	1.1754
Mujeres	13.7969	8.3359	3.8984	1.7500	1.1328	0.6719
PROMEDIO DE VIOLACIONES SARP						
	$e^* = 1$	$e^* = 0,99$	$e^* = 0,98$	$e^* = 0,97$	$e^* = 0,96$	$e^* = 0,95$
Hombres	16.9384	8.2275	5.2654	3.1801	1.7820	1.1754
Mujeres	15.2734	8.3359	3.8984	1.7500	1.1328	0.6719
PROMEDIO DE VIOLACIONES WARP						
	$e^* = 1$	$e^* = 0,99$	$e^* = 0,98$	$e^* = 0,97$	$e^* = 0,96$	$e^* = 0,95$
Hombres	7.7725	3.7251	2.4360	1.3555	0.9005	0.5782
Mujeres	6.7187	3.1562	1.8750	1.0781	0.7187	0.3438

La tabla anterior exhibe un porcentaje¹⁴ de participantes inconsistentes (73.5% hombres, 63.5% mujeres y 69.8% total) superior al exhibido en otros experimentos. Sippel (1997) obtuvo un 41.7% y un 63.3% de participantes inconsistentes en su primer y segundo experimento respectivamente. Mattei (2000) en tres experimentos con 8 bienes y 20 situaciones presupuestarias diferentes obtuvo un porcentaje de “participantes inconsistentes” igual al 25%, 44% y 32% en su primer, segundo y tercer experimento respectivamente. Février (2004) obtuvo sólo un 29% de participantes inconsistentes, aunque este porcentaje no es muy significativo debido al reducido número de situaciones presupuestarias que considera en su estudio. Los resultados del primer experimento muestran cómo las mujeres presentan en promedio un nivel de eficiencia superior al exhibido por los hombres en sus decisiones de consumo.

¹⁴ Entre los participantes inconsistentes incluimos a los 13 participantes (8 hombres y 5 mujeres) que, habiendo mostrado un comportamiento inconsistente con la teoría neoclásica de la demanda, presentan un índice de Afriat superior a 0.99999; los cuáles podrían ser considerados como individuos “consistentes” con GARP.

En la literatura económica existen varias evidencias que sugieren que la diferencia de género en el comportamiento económico de un individuo disminuye entre aquellos participantes que posean la misma formación económica y que conozcan los postulados de los diferentes modelos económicos (v. Cadsby y Maynes (1994) y Powell y Ansic (1997) entre otros). Para analizar esta hipótesis limitamos nuestro estudio a los 67 estudiantes de Economía y Administración y Dirección de Empresas que han realizado el experimento. En la siguiente tabla presentamos un análisis descriptivo de los resultados obtenidos en el segundo experimento.

EXPERIMENTO I.2 (ADE, ECONOMÍA)						
PARTICIPANTES VIOLANDO GARP						
	$e^* = 1$	$e^* = 0,99$	$e^* = 0,98$	$e^* = 0,97$	$e^* = 0,96$	$e^* = 0,95$
Hombres	17 (50.00 %)	10 (29.41 %)	8 (23.53 %)	5 (14.70 %)	3 (08.82 %)	3 (08.82 %)
Mujeres	14 (42.42 %)	9 (27.27 %)	8 (24.24 %)	3 (09.09 %)	2 (06.06 %)	0 (00.00 %)
ÍNDICES DE EFICIENCIA						
	Varian's index	Famulari's index	Swofford's index	Afriat's index	obs. cons.	num. viol.
Hombres	0.9955	0.0154	0.0105	0.9887	2.1429	3.9429
Mujeres	0.9956	0.0223	0.0142	0.9918	2.3438	5.0313
PROMEDIO DE VIOLACIONES GARP						
	$e^* = 1$	$e^* = 0,99$	$e^* = 0,98$	$e^* = 0,97$	$e^* = 0,96$	$e^* = 0,95$
Hombres	3.9429	2.0286	1.0857	0.5714	0.3429	0.1714
Mujeres	5.0313	1.1875	0.6875	0.1875	0.1250	0.0000
PROMEDIO DE VIOLACIONES SARP						
	$e^* = 1$	$e^* = 0,99$	$e^* = 0,98$	$e^* = 0,97$	$e^* = 0,96$	$e^* = 0,95$
Hombres	4.9714	2.0286	1.0857	0.5714	0.3429	0.1714
Mujeres	5.9688	1.1875	0.6875	0.1875	0.1250	0.0000
PROMEDIO DE VIOLACIONES WARP						
	$e^* = 1$	$e^* = 0,99$	$e^* = 0,98$	$e^* = 0,97$	$e^* = 0,96$	$e^* = 0,95$
Hombres	3.6000	1.3714	0.8000	0.4000	0.2857	0.1714
Mujeres	3.3750	0.9375	0.6250	0.1875	0.1250	0.0000

En la tabla anterior observamos cómo el porcentaje¹⁵ de “participantes inconsistentes” es similar entre los participantes varones (48.5%) y entre las mujeres (43.7%). Además, observamos cómo el grado de coherencia de las decisiones de consumo efectuadas tanto por los participantes varones como por las mujeres es bastante similar, lo que hace pensar en la no existencia de diferencias de género entre aquellos individuos que posean una formación económica similar y que estén familiarizados con los diferentes modelos

¹⁵ Entre los participantes inconsistentes incluimos a los 4 participantes (1 hombre y 3 mujeres) que cursan Economía o ADE que, habiendo mostrado un comportamiento inconsistente con la hipótesis de la maximización de la utilidad, presentan un índice de Afriat superior a 0.9999.

económicos. Con el objeto de corroborar o rechazar esta suposición realizamos el segundo experimento descrito en la sección anterior a estos mismos estudiantes de Economía y Administración de Empresas.

Los resultados de este experimento se muestran en la siguiente tabla.

EXPERIMENTO II						
PARTICIPANTES VIOLANDO GARP						
	$e^* = 1$	$e^* = 0,99$	$e^* = 0,98$	$e^* = 0,97$	$e^* = 0,96$	$e^* = 0,95$
Hombres	23 (67.65 %)	13 (38.23 %)	10 (29.41 %)	3 (08.82 %)	0 (00.00 %)	0 (00.00 %)
Mujeres	18 (54.54 %)	10 (27.78 %)	6 (18.18 %)	2 (06.06 %)	0 (00.00 %)	0 (00.00 %)
ÍNDICES DE EFICIENCIA						
	Varian's index	Famulari's index	Swofford's index	Afriat's index	obs. incons.	num. viol.
Hombres	0.9901	0.0901	0.0886	0.9905	4.6765	19.3539
Mujeres	0.9867	0.1332	0.1314	0.9916	5.2424	23.3939
PROMEDIO DE VIOLACIONES GARP						
	$e^* = 1$	$e^* = 0,99$	$e^* = 0,98$	$e^* = 0,97$	$e^* = 0,96$	$e^* = 0,95$
Hombres	19.3529	7.2647	1.8529	0.3529	0.0000	0.0000
Mujeres	23.3939	6.6667	2.0909	0.4242	0.0000	0.0000
PROMEDIO DE VIOLACIONES SARP						
	$e^* = 1$	$e^* = 0,99$	$e^* = 0,98$	$e^* = 0,97$	$e^* = 0,96$	$e^* = 0,95$
Hombres	19.3529	7.2647	1.8529	0.3529	0.0000	0.0000
Mujeres	23.3939	6.6667	2.0909	0.4242	0.0000	0.0000
PROMEDIO DE VIOLACIONES WARP						
	$e^* = 1$	$e^* = 0,99$	$e^* = 0,98$	$e^* = 0,97$	$e^* = 0,96$	$e^* = 0,95$
Hombres	6.2941	2.2353	0.7059	0.2353	0.0000	0.0000
Mujeres	8.4848	2.6061	0.9697	0.3030	0.0000	0.0000

Los resultados del segundo experimento muestran un porcentaje de participantes inconsistentes superior en los hombres (67.6%) que en las mujeres (54.5%), aunque éstos parecen presentar un mayor grado de eficiencia en sus decisiones de consumo que las mujeres al presentar, en promedio mayores índices de Afriat-Varian y menores índices de Swofford, Famulari y un menor número tanto de observaciones inconsistentes como de violaciones con el axioma generalizado de la preferencia revelada.

El porcentaje de “participantes inconsistentes” es superior al encontrado en anteriores experimentos de la teoría del consumidor¹⁶. En el primer experimento observamos cómo el 69% de los participantes muestra un comportamiento inconsistente. Sin

¹⁶ véanse los trabajos realizados por Mattei (2000), Sippel (1997) y Février (2004) entre otros.

embargo, en este experimento el porcentaje de “participantes inconsistentes” entre los estudiantes de Economía y Administración y Dirección de Empresas es similar al exhibido en otros experimentos previos, a pesar de no proporcionar ningún tipo de incentivo -monetario o no- a nuestros participantes. En el segundo experimento, realizado sobre los mismos 67 alumnos, apreciamos un aumento en el porcentaje de participantes inconsistentes entre los hombres (48.5% vs 67.6%) y entre las mujeres (43.7% vs 54.5%). Estos participantes tenían en ambos experimentos la misma motivación intrínseca¹⁷, por lo que es posible que en el primer experimento los contrastes no paramétricos de la preferencia revelada no recojan los posibles comportamientos inconsistentes.

La siguiente tabla muestra el análisis de la sensibilidad efectuado donde podemos apreciar una elevadísima fiabilidad -muy próxima al 100%- de los tests no paramétricos utilizados en ambos experimentos. Como consecuencia de ello, podemos afirmar que muy probablemente este aumento en el porcentaje de “participantes inconsistentes” esté ocasionado por la existencia de umbrales en la percepción¹⁸ de individuo.

Tests comportamiento aleatorio	Número de simulaciones: 50000 Individuos que violan GARP (Experimento I)					
	$e^* = 1$	$e^* = 0,99$	$e^* = 0,98$	$e^* = 0,97$	$e^* = 0,96$	$e^* = 0,95$
Random money choice	97.99 %	86.70 %	64.83 %	40.90 %	22.23 %	11.43 %
Random goods choice	99.74 %	96.83 %	87.11 %	68.60 %	46.46 %	28.00 %
G.M.N.H.	99.48 %	95.22 %	82.88 %	62.73 %	41.14 %	24.11 %
Bronars' Test	98.46 %	86.94 %	66.25 %	42.99 %	24.54 %	12.68 %
	Individuos que violan GARP (Experimento II)					
	$e^* = 1$	$e^* = 0,99$	$e^* = 0,98$	$e^* = 0,97$	$e^* = 0,96$	$e^* = 0,95$
Random money choice	99.90 %	84.46 %	31.91 %	7.37 %	1.57 %	0.27 %
Random goods choice	99.98 %	86.09 %	29.58 %	6.08 %	1.13 %	0.19 %
G.M.N.H.	99.98 %	91.72 %	45.01 %	12.99 %	3.14 %	0.77 %
Bronars' Test	99.93 %	84.76 %	32.35 %	8.12 %	1.83 %	0.38 %

En una revisión bibliográfica sobre la influencia de los incentivos en las decisiones económicas de un individuo, Camerer (1999) afirma que “The data show that incentives sometimes improve performance, but often don't. This unsurprising conclusion implies that we should immediately push beyond debating the caricatured positions that incentives always help or never help”. A pesar de la existencia de evidencias contradictorias sobre la influencia de los incentivos en los resultados experimentales, el

¹⁷ En ambos estudios no se les proporcionó ningún tipo de incentivo y los experimentos se realizaron uno a continuación del otro.

¹⁸ Véanse entre otros los trabajos de McFadden (1999, 2001) sobre diferentes aspectos psicológicos en las decisiones de elección de un individuo.

elevado porcentaje de “participantes inconsistentes” ha podido estar ocasionado por la inexistencia de incentivos en nuestro experimento económico. Para responder a esta hipótesis formulada, analizaremos si los participantes de nuestros experimentos han presentado un comportamiento aleatorio en sus elecciones de consumo. Para ello, realizamos 50000 simulaciones de Montecarlo para cada uno de los “modelos de comportamiento aleatorio del consumidor” considerados en la tercera sección y realizamos, a partir de los datos obtenidos en las simulaciones, un contraste no paramétrico de independencia de dos muestras para contrastar la hipótesis nula de la no existencia de diferencias de comportamiento entre los participantes de nuestros experimentos y los resultados obtenidos a partir de las diferentes simulaciones de Montecarlo efectuadas.

La siguiente tabla exhibe los resultados de los contrastes no paramétricos de independencia de dos muestras de Kolmogorov-Smirnov y de Mann-Whitney para cada uno de los diferentes modelos de comportamiento aleatorios de un consumidor considerados. Nuestros resultados permiten rechazar la hipótesis nula de la no existencia de diferencia de comportamiento entre los 50000 individuos con comportamiento aleatorio y los participantes de nuestros experimentos.

Experimento I										
	Num.		Statistic M-W		Prob. M-W		Statistic K-S		Prob. K-S	
	Part.	Sim.	Afriat	Varian	Afriat	Varian	Afriat	Varian	Afriat	Varian
Participantes	—	—	—	—	—	—	—	—	—	—
Random money choice	339	50000	-11.234	-12.039	0.000	0.000	6.622	7.228	0.000	0.000
Random goods choice	339	50000	-17.813	-19.984	0.000	0.000	9.198	10.693	0.000	0.000
G.M.N.H.	339	50000	-16.897	-17.508	0.000	0.000	8.765	9.863	0.000	0.000
Bronars' Test	339	50000	-11.559	-12.269	0.000	0.000	6.654	7.037	0.000	0.000
Participantes	—	—	—	—	—	—	—	—	—	—
Random money choice	339	50000	-13.852	-13.681	0.000	0.000	7.484	7.546	0.000	0.000
Random goods choice	339	50000	-21.037	-21.419	0.000	0.000	10.734	10.893	0.000	0.000
G.M.N.H.	339	50000	-19.703	-19.615	0.000	0.000	10.206	10.021	0.000	0.000
Bronars' Test	339	50000	-13.965	-13.781	0.000	0.000	7.484	7.536	0.000	0.000
Experimento II										
	Num.		Statistic M-W		Prob. M-W		Statistic K-S		Prob. K-S	
	Part.	Sim.	Afriat	Varian	Afriat	Varian	Afriat	Varian	Afriat	Varian
Participantes	—	—	—	—	—	—	—	—	—	—
Random money choice	67	50000	-7.436	-8.40	0.000	0.000	4.488	4.932	0.000	0.000
Random goods choice	67	50000	-7.430	-8.444	0.000	0.000	4.641	5.137	0.000	0.000
G.M.N.H.	67	50000	-8.443	-9.301	0.000	0.000	4.941	5.377	0.000	0.000
Bronars' Test	67	50000	-7.580	-8.453	0.000	0.000	4.536	4.936	0.000	0.000
Participantes	—	—	—	—	—	—	—	—	—	—
Random money choice	67	50000	-10.990	-10.440	0.000	0.000	5.565	5.450	0.000	0.000
Random goods choice	67	50000	-11.683	-11.390	0.000	0.000	5.981	5.976	0.000	0.000
G.M.N.H.	67	50000	-11.780	-11.384	0.000	0.000	6.017	6.037	0.000	0.000
Bronars' Test	67	50000	-10.950	-10.405	0.000	0.000	5.514	5.416	0.000	0.000

Los resultados del primer experimento sugieren la existencia entre la población general de diferencias de comportamiento en el consumo. En efecto, el porcentaje de “participantes inconsistentes” en el primer experimento es sensiblemente superior entre los estudiantes varones (73.5%) que entre las mujeres (63.5%). Además, las mujeres presentan un mayor grado de coherencia en sus decisiones de consumo que los hombres, dado que éstos últimos presentan en promedio un menor índice de Afriat, un menor índice de Varian, un mayor índice de Famulari, índice de Swofford, de observaciones inconsistentes y de violaciones y, por tanto, un menor grado de eficiencia en sus decisiones de consumo. Sin embargo, no está tan clara esta diferencia de comportamiento entre aquellos individuos que poseen una formación económica similar y conocen los postulados de la teoría económica. Entonces, resulta natural plantearse las siguientes preguntas: ¿existen diferencias de género en el proceso de elección de un consumidor? En caso afirmativo, ¿son estadísticamente significativas estas diferencias de género? Para responder a estas preguntas realizamos dos contrastes de hipótesis no paramétricos de independencia de dos muestras, el test de Kolmogorov-Smirnov y el test de Mann-Whitney para cada uno de los índices de eficiencia considerados. Los resultados obtenidos se muestran en la siguiente tabla:

Experimento I.1						
	Participantes		Stat. M-W	Prob.	Stat. K-S	Prob.
	H	M				
Afriat's index	212	127	-2.334	0.020	1.551	0.016
Varian's index	212	127	-1.797	0.072	1.060	0.211
Famulari's index	212	127	-1.993	0.053	1.216	0.104
Swofford's index	212	127	-2.092	0.036	1.284	0.074
Inconsistent number	212	127	-1.947	0.052	1.122	0.161
Experimento I.2						
	Participantes		Stat. M-W	Prob.	Stat. K-S	Prob.
	H	M				
Afriat's index	34	33	-0.896	0.370	0.565	0.906
Varian's index	34	33	-0.293	0.769	0.507	0.959
Famulari's index	34	33	-0.325	0.745	0.397	0.997
Swofford's index	34	33	-0.177	0.860	0.437	0.991
Inconsistent number	34	33	-0.232	0.816	0.390	0.959
Experimento II						
	Participantes		Stat. M-W	Prob.	Stat. K-S	Prob.
	H	M				
Afriat's index	34	33	-0.575	0.565	0.580	0.890
Varian's index	34	33	-0.258	0.796	0.9118	0.377
Famulari's index	34	33	-0.303	0.762	0.667	0.767
Swofford's index	34	33	-0.341	0.733	0.667	0.767
Inconsistent number	34	33	-0.129	0.897	0.667	0.767

La tabla anterior muestra cómo entre la población general las mujeres son más eficientes -en promedio- en sus decisiones de consumo que los hombres y existen diferencias de género en el grado de racionalidad de las elecciones de consumo realizadas por un individuo. Sin embargo, en el primer experimento podemos colegir aproximadamente con una probabilidad de 0.98 la no existencia de diferencias de género entre aquellos individuos que poseen una formación económica similar y que conocen los postulados de los diferentes modelos económicos, mientras que en el segundo experimento lo podemos afirmar con una probabilidad de 0.8 aproximadamente. En relación con la existencia de diferencias en el grado de eficiencia en el consumo, podemos asegurar con una probabilidad de 0.75 aproximadamente la no existencia de diferencias de género en el grado de coherencia de las decisiones de consumo efectuadas por los estudiantes de Economía o Administración de Empresas.

6. Conclusión

En este trabajo hemos obtenido resultados contradictorios en relación con la existencia de diferencias de género en el grado de coherencia de las elecciones de consumo de un individuo. Considerando la población en general hemos encontrado que las mujeres son consumidores más eficientes que los hombres. Sin embargo, no existen diferencias de género entre individuos que poseen una formación similar y están familiarizados con los postulados de la teoría pura del comportamiento del consumidor. Los resultados obtenidos, al igual que sucede en anteriores experimentos de la teoría del consumidor (v. Sippel (1997), Mattei (2000) y Février (2004) entre otros) no son muy favorables con la teoría del comportamiento del consumidor.

A pesar de la inexistencia de incentivos, hemos demostrado que los participantes en los experimentos han tenido el suficiente grado de motivación intrínseca que les ha llevado a realizar elecciones de consumo “casi consistentes” con la teoría neoclásica de la demanda. Camerer (1999) en una revisión bibliográfica sobre la influencia de los incentivos en los resultados experimentales afirma que “los incentivos a veces ayudan a obtener mejores resultados pero en otras ocasiones no”. No podemos afirmar rotundamente que el número de participantes inconsistentes, y el número de

violaciones, con los diferentes axiomas de la preferencia revelada, hubiera sido menor si hubiésemos utilizado incentivos monetarios en la realización de nuestros experimentos, ya que siempre es posible sugerir que el comportamiento de un consumidor se aproximaría más al modelo optimizador de la utilidad si se aumentaran los incentivos. Pero, ¿cuál debería ser el pago con el que debemos de incentivar a los participantes para eliminar todas estas anomalías? Vernon Smith (1982) propone para que los experimentos “funcionen” que exista un procedimiento de valor inducido en el diseño de los experimentos. Así, plantea una metodología de realización de experimentos donde los participantes muestren deseos de insaciabilidad, obtengan recompensas suficientemente altas y obvias, exista una restricción a la comunicación entre sujetos, etc. Sin embargo, admite que los preceptos anteriores no pueden garantizar ninguna correspondencia entre recompensas monetarias observables y preferencias que, en principio, no son “directamente observables”. Tampoco podemos afirmar la existencia de una correlación (positiva o negativa) entre la presencia de incentivos y las diferencias de género en el comportamiento racional de un consumidor en un contexto determinista y en un ambiente sin riesgo como en el que nos encontramos.

Referencias

- [1] Afriat, S.N. (1965) The equivalence in two dimensions of the strong and weak axioms of revealed preference. *Metroeconomica*, Vol. 17, 24-28.
- [2] Afriat, S.N. (1967) The construction of utility functions from expenditure data. *International Economic Review* 8(1), 67-77.
- [3] Afriat, S.N. (1973) On a system of inequalities in demand analysis: An extension of the classical method. *International Economic Review*, Vol. 14(2), 460-472.
- [4] Andreoni, J. & Miller, J.H. (2002) Giving According to GARP: An Experimental Test of the Consistency of Preferences for Altruism. *Econometrica*, Vol. 70(2), 737-753.
- [5] Battalio, R.C. et al. (1973) A Test of consumer demand theory using observations of individual consumer purchases. *Western Economic Journal* Vol. XI, 411-428.
- [6] Battalio, R.C., Dwyer, G.P. & Kagel, J.H. (1987) Tests of competing theories of consumer choice and the representative consumer hypothesis. *The Economic Journal*, Vol. 97, 842-856.
- [7] Baxter, S., Lansing, M. (1983) *Women and Politics: The visible majority*. Ann Arbor, MI: University of Michigan Press.
- [8] Becker, G. (1962) Irrational Behavior and Economic Theory. *Journal of Political Economy*, Vol. 70, 1-13.
- [9] Binmore, K. (1999) Why experiment in economics?. *Economic Journal*, Vol. 109(2), F16-24.

- [10] Bromiley, P., Curley, S. (1992) Individual difference in risk taking. In: Yates, J.F. (Ed.), *Risk Taking Behaviour*. Wiley, Chichester, UK.
- [11] Bronars, S. (1987) The Power of Nonparametric Tests of Preference Maximization. *Econometrica*, Vol. 55 (3), 693-698.
- [12] Brown-Kruse, J., Hummels, D. (1993) Gender effects in laboratory public goods contribution: do individuals put their money where their mouth is? *Journal of Economic Behavior and Organization*, Vol. 22, 255-267.
- [13] Cadsby, C.B., Maynes, E. (1994) Threshold public goods games: New experimental evidence, University of Guelph. Working Paper 1994-4.
- [14] Cadsby, C.B., Maynes, E. (1998a) Gender and free riding in a threshold public goods game: experimental evidence. *Journal of Economic Behavior and Organization*, Vol. 34, 603-620.
- [15] Camerer, C.; Hogarth, H. (1999) The effects of financial incentives in experiments: A Review and capital-labor-production framework. *Journal of Risk and Uncertainty*, Vol. 19(1), 7-42.
- [16] Chant, J.F. (1963) Irrational behavior and economic theory: a comment. *Journal of Political Economy*, Vol. 71, 505-510.
- [17] Cooperstock, R., Parnell, P. (1982) Research on psychotropic drug use: a review of findings and methods. *Social Science and Medicine*, Vol. 16, 1179-1196.
- [18] Diewert, E. (1973) Afriat and revealed preference theory. *Review of Economic Studies*, Vol. 40, 419-426.
- [19] Eagly, A., Crowley, M. (1986) Gender and helping behavior: a meta-analytic review of the social psychological literature. *Psychological Bulletin*, Vol. 100, 283-308.
- [20] Eagly, A., Karau, S., Makhijani, M. (1995) Gender and the effectiveness of leaders: A meta-Analysis. *Psychological Bulletin*, Vol. 117(1), 125-145.
- [21] Eagly, A. (1995) The science and politics of comparing women and men. *American Psychologist*, Vol. 50(3), 145-158.
- [22] Eckel, C., Grossman, P. (1996) The relative price of fairness: gender differences in a punishment game. *Journal of Economic Behavior and Organization*, Vol. 30, 143-158.
- [23] Eckel, C., Grossman, P. (1998) Are women less selfish than men? Evidence from dictator experiments. *The Economic Journal*, Vol. 108, 726-735.
- [24] Eckel, C., Grossman, P. (2001) Chivalry and solidarity in ultimatum games. *Economic Inquiry*, Vol. 39, 171-188.
- [25] Eckel, C., Grossman, P. (2004) Differences in the economic decisions of men and women: experimental evidence. Unpublished manuscript (forthcoming in Plott, C., Smith, V. (Eds.), *Handbook of Results in Experimental Economics*. North-Holland, New York).
- [26] Famulari, M. (1995) A household-based, nonparametric test of demand theory. *The Review of Economics and Statistics* 77(4), 372-382.
- [27] Fengler, M. & Winter, J. (2001) Psychological pricing points and price adjustment in German retail markets, mimeo, Mannheim Universität.
- [28] Ferber, M. (1995) The study of economics: a feminist critique. *American Economic Review*, Vol. 85 (2), 357-361.
- [29] Février P., Visser M. (2004) A Study of Consumer Behavior Using Laboratory Data. *Experimental Economics*, Vol. 7(1), 93-114.
- [30] Frank, R., Gilovich, T., Regan, D. (1993) Does studying economics inhibit cooperation?. *Journal of Economic Perspectives*, Vol. 7(2), 159-171.

- [31] Gale, D. (1960) A Note on Revealed Preference. *Economica*, Vol. 27(108), 348-354.
- [32] Gilligan, C. (1982) *In a Different Voice: Psychological Theory and Women's Development*. Harvard University Press, Cambridge, MA.
- [33] Goertzel, T. (1983) The gender gap: sex, famil income, and political opinions in the early 1980's. *Journal of Political and Military Sociology*, Vol. 11, 209-222.
- [34] Gottfredson, M., Hirschi (1990) *A general theory of crime*. Standford, CA: Standford University Press.
- [35] Gross, J. (1995) Testing data for consistency with revealed preference. *The Review of Economics and Statistics* 77(4), 701-710.
- [36] Harbaugh, W., Krause K. & Berry, T. (2001) GARP for kids: On the development of rational choice behavior. *American Economic Review*, Vol. 91(5), 1539-1545.
- [37] Harrison, G. (1989) Theory and misbehavior of first-price auctions. *American Economic Review*, Vol. 79(4), 749-762.
- [38] Holdershaw, J., Gendall, P. & Garland, R. (1997) The Widespread Use Of Odd Pricing In The Retail Sector. *Marketing Bulletin*, Vol. 7, 53-58.
- [39] Houthakker, H. (1950) Revealed Preference and the Utility Function. *Economica*, Vol. 17, 159-174.
- [40] Houtman, M. & Maks, J. (1987) The existence of homothetic utility functions generating dutch consumer data. Working Paper. University of Groningen.
- [41] Isaac, R., McCue, K., Plott, C. (1985) Public goods provision in an experimental market. *Journal of Public Economics*, Vol. 26, 51-74.
- [42] Jianakoplos, N., Bernasek, A. (1998) Are women more risk averse? *Economic Inquiry*, Vol. 36, 620-630.
- [43] Kagel, J.H., Rachlin, H. & Green, L. (1981) Demand curves for animals consumers. *Quarterly Journal of Economics*, Vol. 66, 1-15.
- [44] Kagel, J. & Roth, A. (1995) *The Handbook of Experimental Economics*. Princeton, Princeton University Press.
- [45] Kandel, D., Logan, J. (1984) Patterns of drug use from adolescence to young adulthood: periods of risk for initiation, continued use, and discontinuation. *American Journal of Public Health*, Vol. 74, 660-666.
- [46] Kamiya, D. (1963) A Note on the Strong Axiom of Revealed Preference, Vol. 30(117), 83-84.
- [47] Landsburg, S. (1981) Taste change in the United Kingdom, 1900-1955. *Journal of Political Economy*, Vol. 89(11), 92-104.
- [48] Loomes, G. (1999) Some lessons from past experiments and some challenges for the future. *Economic Journal*, Vol. 109(2), F35-F45.
- [49] Mattei, A. (1994) La Coherence des Choix des Cosommateurs Suisses. *Revue Suisse d'Economie Politique et de Statistique*, Vol. 30, 3-20.
- [50] Mattei, A. (2000) Full-scale real tests of consumer behavior using experimental data. *Journal of Economic Behavior and Organization*, Vol. 43, 487-497.
- [51] Mason, C., Phillips, O., Redington, D. (1991) The role of gender in a non-cooperative game. *Journal of Economic Behavior and Organization*, Vol. 15, 215-235.
- [52] McFadden, D. (1999) Rationality for Economists? *Journal of Risk and Uncertainly*, Vol. 19 (1-3), 73-105.
- [53] McFadden, D. (2001) Economic choices. *American Economic Review*, Vol. 91(3),351-378.

- [54] Nelson, J. (1995) Feminism and economics. *Journal of Economic Perspectives*, Vol. 9(2), 131-148.
- [55] Nowell, C., Tinkler, S. (1994) The influence of gender in a non-cooperative game. *Journal of Economic Behavior and Organization*, Vol. 25, 25-36.
- [56] Powell, M., Ansic, D. (1997) Gender differences in risk behaviour in financial decision-making: An experimental analysis. *Journal of Economic Psychology*, Vol. 18(6), 605-628.
- [57] Powell, M., Ansic, D. (1999) Gender differences in financial decision-making: a new approach for experimental economic analysis. Unpublished manuscript (presented at the AEA Conference, New York, January 1999).
- [58] Rose, H. (1958) Consistency of Preference: The Two-Commodity Case. *Review of Economic Studies*, Vol. 24, 124-125.
- [59] Samuelson, P. (1948) Consumption Theory in Terms of Revealed Preference. *Economica*, Vol. 15, 243-253.
- [60] Scarf, H.E., Fostel, A., & Todd, M.J. (2004) Two New Proofs of Afriat's Theorem. *Economic Theory*, Vol. 24(1), 211-219.
- [61] Schubert, R., Brown, M., Gysler, M., Brachinger, H. (1999) Financial decision making: are women really more risk averse? *American Economic Review (Papers and Proceedings)*, Vol. 89, 381-385.
- [62] Seguino, S., Stevens, T., Lutz, M. (1996) Gender and cooperative behavior: economic man rides alone. *Feminist Economics*, Vol. 2, 1-21.
- [63] Sippel, R. (1997) An Experiment On The Pure Theory Of Consumer's Behaviour. *The Economic Journal*, Vol. 107, 1431-44.
- [64] Smith, V.L. (1982) Microeconomic Systems as an Experimental Science. *American Economic Review*, Vol. 72, 923-955.
- [65] Starmer, C. (1999) Experimental Economics: Hard Science or Wasteful Tinkering?. *Economic Journal*, Vol. 109(29), F5-F15.
- [66] Stockard, J., Van de Kragt, A., Dodge, P. (1988) Gender roles and behavior in social dilemmas: are there sex differences in cooperation and in its justification? *Social Psychology Quarterly*, Vol. 51(2), 154-163.
- [67] Strober, M. (1994) Rethinking economics through a feminist lens. *American Economic Review*, Vol. 84(2), 143-147.
- [68] Swofford, J. & Whitney, G. (1987) Nonparametric Tests of Utility Maximization and Weak Separability for Consumption, Leisure and Money. *Review of Economics and Statistics*, Vol. 69(3), 458-64
- [69] Unger, R. (1990) Imperfect reflections of reality: Psychology constructs gender. In: Hare-Mustin, R., Marecek, J. (Eds), *Making a Difference; Psychology and the Construction of Gender*. Yale University Press, New Haven, USA.
- [70] Usuegi, T., Vinacke, E. (1963) Strategy in a feminine game. *Sociometry*, Vol. 26, 75-88.
- [71] Varian, H. (1982) The nonparametric approach to demand analysis. *Econometrica*, Vol. 50(4), 945-971.
- [72] Varian, H. (1990) Goodness-of-fit in optimizing models. *Journal of Econometrics*, Vol. 46, 125-140.
- [73] Vinacke, E. (1959) Sex roles in a three-person game. *Sociometry*, Vol. 22, 343-360.

UNA EXPLICACIÓN DEL COMPORTAMIENTO ERRÁTICO DEL INTERVALO DE WALD EN EL MODELO BINOMIAL

JESÚS BASULTO SANTOS
e-mail: basulto@us.es

FCO. JAVIER ORTEGA IRIZO
e-mail: fjortega@us.es

JOSÉ ANTONIO CAMÚÑEZ RUÍZ
e-mail: camunez@us.es

Departamento de ECONOMÍA APLICADA I
UNIVERSIDAD DE SEVILLA

Área temática: Métodos cuantitativos

Resumen

Es conocido que el intervalo de Wald para estimar la proporción de éxito en una distribución binomial presenta un comportamiento muy errático con respecto a su probabilidad de cubrimiento al hacer variar los parámetros n y p . Habitualmente se ha considerado que este comportamiento es “esencialmente impredecible”. Sin embargo, el análisis de la causa de estos descensos permite obtener una fórmula que proporciona, fijado p , todos los valores de n en los que se produce un brusco descenso de la probabilidad de cubrimiento. Dicho análisis también permite comprender por qué el intervalo de Wilson presenta un comportamiento menos errático que el de Wald.

Palabras clave: Distribución Binomial, Probabilidad de Cubrimiento, Valor Nominal, Intervalo de Wald, Intervalo de Wilson.

Abstract

The erratic behaviour of the coverage probability of the Wald interval of a binomial proportion has previously been remarked on the literature. In addition, “unlucky” values of n again arise in the same “unpredictable” way. In section 4, we present an explication on the “unlucky” of n and, also, we propose a method to calculate, fixed p , the values of n where significant change in coverage probability occurs. From the analysis of the Wald interval, we show that the performance the Wilson interval is less erratic.

Key words: Binomial Distribution, Coverage Probability, Nominal value, Wald interval, Wilson Interval.

1. Introducción

En muchos libros de estadística se suele proponer el siguiente intervalo de Wald, $\hat{p} \pm z_{\frac{\alpha}{2}} n^{-1/2} [\hat{p}(1-\hat{p})]^{1/2}$, como un intervalo de confianza para estimar el parámetro p de la distribución binomial, y donde $\hat{p} = \frac{x}{n}$ es la proporción en una muestra de tamaño n , $z_{\frac{\alpha}{2}}$ es el percentil $100\left(1-\frac{\alpha}{2}\right)$ de una distribución normal $(0,1)$ y $1-\alpha$ es el coeficiente de confianza o valor nominal. Este intervalo es evaluado por medio de muestras aleatorias, de tamaño n , generadas por una distribución binomial con parámetro p . Un criterio para evaluar este intervalo de Wald es calcular la probabilidad con que el intervalo aleatorio de Wald cubre el valor del parámetro p . A medida que la probabilidad de cubrimiento se aleja del valor nominal, el intervalo de Wald será poco útil, caso de que el valor de cubrimiento sea mucho más pequeño que el valor nominal o de amplitud no óptima, caso de que la probabilidad de cubrimiento sea mucho mayor que el valor nominal.

Si la probabilidad de cubrimiento del intervalo de Wald lo representamos por la expresión $c(p,n)$, entonces afirmar que el intervalo de Wald, con valor nominal $1-\alpha$, es eficaz es equivalente a que $c(p,n) = 1-\alpha$, para los valores del parámetro $p \in [0,1]$ y $n \geq 1$.

En esta y la siguiente parte del trabajo vamos a calcular las probabilidades de cubrimiento del intervalo de Wald, con un valor nominal de 0.95, cuando variamos los tamaños de la muestra, n , y el valor del parámetro p . Seguiremos el trabajo Brown, Cal y Dasgupta (2001), donde los autores señalan los hechos siguientes: (a) Hay pares de valores (p,n) , “con mucha suerte”, es decir, con probabilidades de cubrimiento próximas al valor nominal 0.95 y, otros pares con “mala suerte”, es decir, con probabilidades de cubrimiento que se apartan totalmente del valor nominal 0.95; (b) para un valor fijo de p , la función $c(p,n)$ no es creciente, es decir, que al aumentar el tamaño de la muestra no siempre logramos que la probabilidad de cubrimiento; (c) para un valor fijo de p , la función $c(p,n)$ calculada para el intervalo de Wald tiende a situarse por debajo del valor nominal 0.95 y (d) para un valor fijo de n , la función $c(p,n)$ es una curva en dientes de sierra y además presenta grandes oscilaciones para valores del parámetro próximos a cero o a la unidad.

En nuestro trabajo buscamos explicar los comportamientos caóticos de los cubrimientos del intervalo de Wald, aportando un procedimiento para obtener los valores del tamaño n para los que la probabilidad de cubrimiento sufre la mayor caída local.

Un intervalo alternativo al de Wald es el usado por Wilson en 1927. En el libro de Spiegel (1991) se afirma en la página 211 que “un método más exacto para obtener los límites de confianza se presenta en el problema 9.12”. En dicho problema se obtiene precisamente el intervalo de

Wilson. La expresión del intervalo de Wilson es la siguiente: $\frac{x + \frac{k^2}{2}}{n + k^2} \pm \frac{k n^{1/2}}{n + k^2} \left[\hat{p}\hat{q} + \frac{k^2}{4n} \right]^{1/2}$, donde $\hat{q} = 1 - \hat{p}$ y $k = z_{\frac{\alpha}{2}}$. Veremos que el intervalo de Wilson mejora los comportamientos de las probabilidades de cubrimientos frente al intervalo de Wald.

2. Intervalo de Wald

En la Tabla I recogemos las probabilidades de cubrimiento, entre corchetes, para distintos valores de n cuando $p = 0,5$ y $\alpha = 0,05$.

Tabla I
Probabilidades de Cubrimiento para distintos n y $p = 0,5$ con $\alpha = 0,05$, intervalo de Wald.

Valores de n con Buena suerte	17[0,951]	20[0,958]	25[0,956]	49 [0,955]
Valores de n con Mala suerte	10[0,891]	12[0,854]	13[0,907]	40 [0,919]

Se observa que hay valores de n con buena suerte, así para $n=17$ la probabilidad de cubrimiento, 0.951, está muy próxima a su valor nominal 0,95; mientras que para $n=12$ la probabilidad de cubrimiento, 0.854 se aleja totalmente de su valor nominal 0,95. Estas suertes siguen produciéndose cuando n es igual a 49 y 40 respectivamente.

En la Tabla II recogemos las probabilidades de cubrimiento, entre corchetes, para distintos valores de n cuando $p = 0,2$ y $\alpha = 0,05$. De nuevo se observa la buena y mala suerte en la elección de n . Además se observa que cuando aumenta el tamaño de n las probabilidades de cubrimiento disminuyen alejándose de su valor nominal. En efecto, para $n=31$ la probabilidad de

cubrimiento es 0,949, mientras que para $n=92$ vale 0,922. Este ultimo es contrario al comportamiento razonable que debería tener un intervalo de confianza cuando aumenta suficientemente el tamaño de n .

Tabla II

Probabilidades de Cubrimiento para distintos n y $p = 0,2$ con $\alpha = 0,05$, intervalo de Wald.

Valores de n con Buena suerte	31[0,949]	38[0,948]	72[0,954]	97 [0,951]
Valores de n con Mala suerte	25[0,884]	32[0,889]	66[0,916]	92[0,922]

En la Figura I hemos calculado las probabilidades de cubrimiento para $n=100$, $\alpha = 0,05$ y distintos valores de p del intervalo $[0; 0,5]$. Para valores de p del intervalo $[0,5; 1]$ las probabilidades de cubrimiento correspondientes son idénticas.

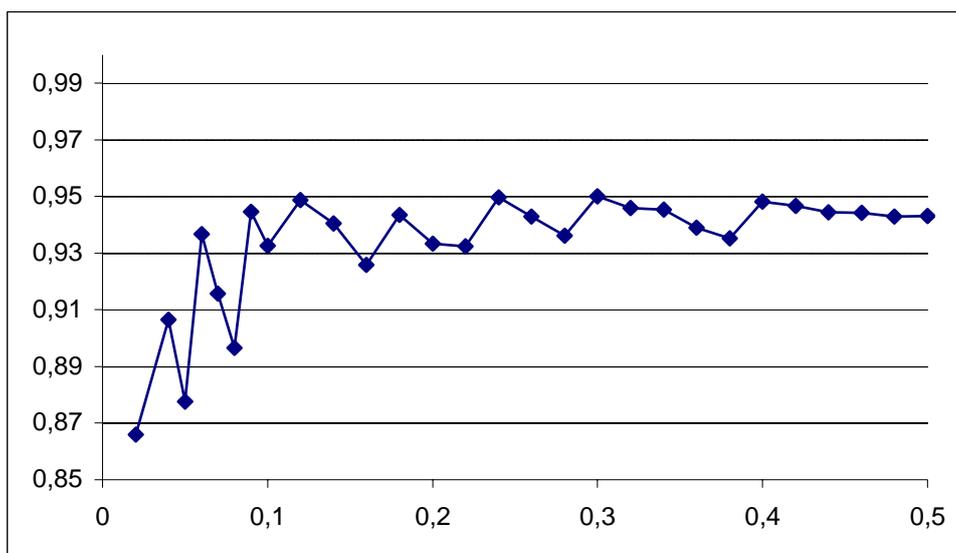


Figura I. Probabilidades de Cubrimiento para distintos valores de p del intervalo $[0, 0,5]$ cuando $n = 100$ y $\alpha = 0,05$. Intervalo de Wald.

En la Figura I se observa que la curva de las probabilidades de cubrimiento presenta grandes oscilaciones en el entorno de $p = 0$, que por simetría también ocurre para $p = 1$. Pero lo más sorprendente es observa que la curva de las probabilidades de cubrimiento está situada por debajo de la línea que marca el valor nominal 0.95, excepto para algunos valores.

Las grandes oscilaciones de las probabilidades de cubrimiento para valores de p cercanos a cero que observamos en la Figura I para $n = 100$ se mantienen para grandes valores aunque ahora

las oscilaciones son menores. En la Tabla III se observa que para $p = 0.005$, $\alpha = 0,05$, al variar los tamaños de la muestras para n grande, siguen observándose oscilaciones significativas en las probabilidades de cubrimiento.

Tabla III

Probabilidades de Cubrimiento para distintos n , cuando $p = 0.005$ y $\alpha = 0,05$, intervalo de Wald.

Valores de n con mala suerte	592	954	1583	1875
$c(0,005, n)$	0,792	0,852	0,889	0,898

Por último, veamos algunas reglas que suelen proponer los libros de enseñanza de la estadística para decidir sobre el uso del intervalo de confianza de Wald.

1. $np, np(1-p)$ sean ≥ 5 (o 10);
2. $np(1-p) \geq 5$ (o 10);
3. $n \geq 100$.

Con estas reglas los autores de los libros de texto pretenden ayudar a los alumnos para aproximar la distribución binomial por la normal. Estas reglas no son muy adecuadas cuando se aplican a un problema de inferencia como el que pretende dar respuesta el intervalo de Wald. En efecto, la dos primeras reglas no son adecuadas, por ejemplo, para $n = 40$ y $p = 0,5$, siendo $n \cdot p = n \cdot (1-p) = 20$ y también $n \cdot p \cdot (1-p) = 10$, pero la probabilidad de cubrimiento vale 0,919, que se aleja totalmente de su valor nominal 0,95.

En cuanto la tercera regla, que también busca la aproximación de la distribución binomial por la normal, puede ser equívoca. En efecto, de la Figura I, si

$p = 0,5$, la regla es adecuada, pero si el parámetro se aleja del 0,5, por ejemplo $p = 0,2$, la probabilidad de cubrimiento vale 0,933, o si $p = 0,28$, entonces su probabilidad de cubrimiento vale 0,936, etc.

3. Intervalo de Wilson

Este intervalo fue usado por Wilson en una notas publicadas en 1927 sobre “Probable inference, the law of sucesión, and statistical inference”. Wilson aproximó la distribución binomial por una normal y consideró la cantidad pivotal siguiente: $\frac{x - np}{\sqrt{np(1-p)}}$ que se aproxima a la distribución normal (0,1). Al invertir esta cantidad pivotal obtuvo el intervalo que hemos presentado en nuestra introducción con el nombre de intervalo de Wilson. En el libro de Spiegel (1991) se obtiene este intervalo en el problema 9.12. Esta justificación es más pedagógica en el intervalo de Wilson que en el de Wald, ya que en el de Wald la varianza \hat{p} es estimado por $\hat{p}(1-\hat{p})/n$, que es consistente, y la cantidad pivotal que se invierte es ahora $\frac{x - np}{\sqrt{n\hat{p}(1-\hat{p})}}$, que también tiende en distribución a una distribución normal (0,1).

Si el valor central del intervalo de Wald es el estimador es \hat{p} , es decir, la proporción de la

muestra, el valor central del intervalo de Wilson es el estimador $\frac{\hat{p} + \frac{t}{2}}{1+t}$, donde $t = \frac{z_{\alpha/2}^2}{n}$.

En la Figura II, hemos calculado las probabilidades de cubrimiento $n=100$, $\alpha = 0,05$ y distintos valores de p del intervalo $[0, 0,5]$. Para valores de p del intervalo $[0,5 ; 1]$ las probabilidad de cubrimiento son idénticos a las anteriores.

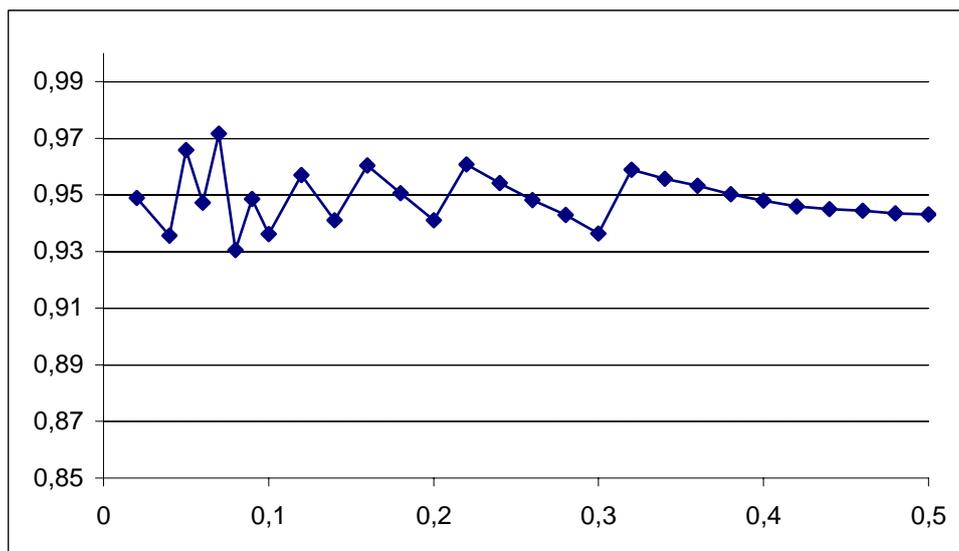


Figura II. Probabilidades de Cubrimiento para distintos valores de p del intervalo $[0, 0,5]$ cuando $n=100$ y $\alpha = 0,05$, intervalo de Wilson.

Si comparamos con la Figura I del intervalo de Wald, vemos el cambio tan espectacular de esta nueva curva, la Figura II, frente a la de Wald. Wilson en su trabajo de 1927 habla de que la “chance” de su intervalo de que contenga a p puede ser mejorada por su nuevo intervalo. Recordemos que la palabra chance se aplicaba a probabilidades objetivas, como en el caso de repetición en las mismas condiciones de un cierto experimento, mientras la palabra probabilidad se usaba en otros casos.

En la Tabla IV recogemos las probabilidades de cubrimiento, entre corchetes, para distintitos valores de n cuando $p = 0,5$ y $\alpha = 0,05$. Vemos ahora que los problemas del intervalo de Wald no se presentan en el intervalo de Wilson. Las probabilidades de cubrimiento oscilan alrededor del coeficiente de confianza del 0,95.

Tabla IV

Probabilidades de Cubrimiento para distintos n y $p = 0,5$ con $\alpha = 0,05$, intervalo de Wilson

Valores de n	17[0,951]	20[0,959]	25[0,957]	49 [0,955]
Valores de n	10[0,978]	12[0,961]	13[0,977]	40 [0,961]

En la Tabla V recogemos las probabilidad de cubrimiento, entre corchetes, para distintos valores de n cuando $p = 0,2$ y $\alpha = 0,05$.

Tabla V

Probabilidades de Cubrimiento para distintos n y $p = 0,2$ con $\alpha = 0,05$, intervalo de Wilson.

Valores de n	31[0,958]	38[0,960]	72[0,962]	97 [0,958]
Valores de n	25[0,925]	32[0,952]	66[0,955]	92[0,949]

Ahora, para $n=31$ la probabilidad de cubrimiento es 0,958, y para $n=92$ vale 0,949. Es decir, al aumentar el tamaño de la muestra, la probabilidad de cubrimiento se aproxima más a su valor nominal. Se trata de un comportamiento razonable. En otros casos cuando aumentamos el tamaño de la muestra, se observa que la probabilidad de cubrimiento también aumenta.

4. Análisis empírico del comportamiento irregular del intervalo de Wald.

En esta sección, analizaremos con más detalle algunos ejemplos concretos, al objeto de buscar una explicación del comportamiento aparentemente errático de la probabilidad de cubrimiento del intervalo de Wald. Nos centraremos en el caso de mantener un valor de p fijo haciendo variar n .

Fijado un α , a partir de la expresión del intervalo de Wald, podemos obtener la probabilidad de cubrimiento $C(n, p)$ dados unos valores de p y n como la suma de las probabilidades de ciertos valores en la distribución $Bi(n, p)$. Concretamente,

$$C(n, p) = \sum_{j \in (L_1(n, p), L_2(n, p))} P[X(n, p) = j],$$

donde $X(n, p) \sim Bi(n, p)$ y $L_1(n, p)$ y $L_2(n, p)$ son las soluciones de las ecuaciones

$$\frac{1}{n} \left[x + z_{\alpha/2} \sqrt{\frac{x(n-x)}{n}} \right] = p \quad \text{y} \quad \frac{1}{n} \left[x - z_{\alpha/2} \sqrt{\frac{x(n-x)}{n}} \right] = p, \quad \text{respectivamente.}$$

Por ejemplo, para $n=31$ y $p=0.2$ al resolver las ecuaciones anteriores obtenemos $L_1(31, 0.2) = 2.9822$ y $L_2(31, 0.2) = 11.4685$, por lo que la probabilidad de cubrimiento del intervalo de Wald será

$$C(31, 0.2) = \sum_{j=3}^{11} P[X(31, 0.2) = j] = 0.9498.$$

Cuando se analiza la probabilidad de cubrimiento fijado p y haciendo aumentar n , la característica más llamativa es que en ciertos valores de n se producen bruscos descensos de dicha probabilidad, como puede apreciarse en la figura III (en la que se ha tomado $p = 0.25$). Además, podemos observar cómo dicha probabilidad se mantiene casi siempre por debajo del valor nominal $1 - \alpha$.

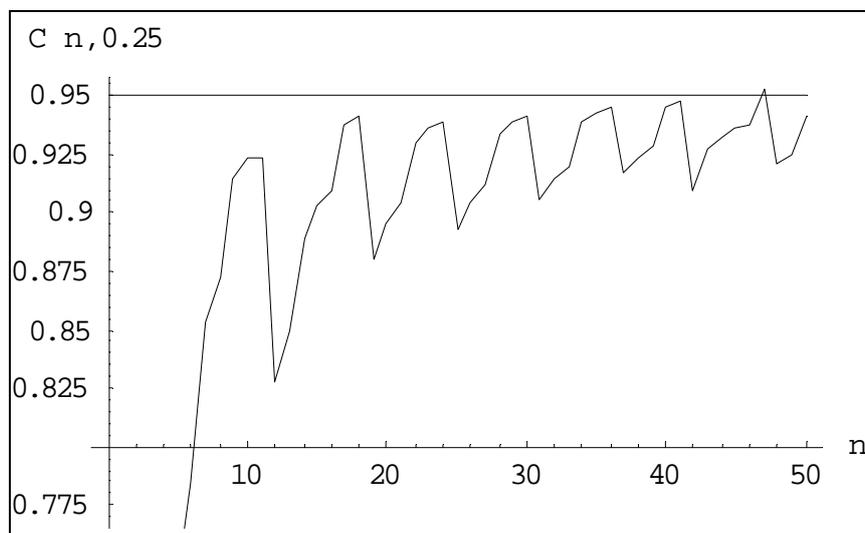


Figura III. Probabilidades de Cubrimiento para distintos valores de n y $p = 0.25$. Intervalo de Wald.

En Brown, Cal y Dasgupta (2001), acerca de estos valores de n , los autores señalan: “Subsequent “unlucky” values of n again arise in the same **unpredictable** way, as one...”¹. Veamos en algunos ejemplos, cuál es la causa de estos bruscos descensos en la probabilidad de cubrimiento.

Ejemplo 1: $p = 0'4$, $n = 32$.

En este caso, podemos comprobar que $C(32,0'4) = 0'9543$ mientras que para $n = 33$ la probabilidad de cubrimiento es $C(33,0'4) = 0'9245$. En este caso, se obtiene $L_1(32,0'4) = 7'9992$, $L_2(32,0'4) = 18'2868$ mientras que para $n = 33$ los límites son $L_1(33,0'4) = 8'3123$, $L_2(33,0'4) = 18'7759$ por lo que las probabilidades de cubrimiento vienen

dadas por $C(32,0'4) = \sum_{j=8}^{18} P[X(32,0'4) = j]$ y $C(33,0'4) = \sum_{j=9}^{18} P[X(33,0'4) = j]$. En la tabla VI y

en la figura IV podemos apreciar que las diferencias de probabilidad de los valores 8 a 18 cuando pasamos de $n = 32$ a $n = 33$ son aproximadamente simétricas y de suma cercana a 0. Así, **el hecho que resulta decisivo en la diferencia de cubrimiento es la inclusión del caso $j = 8$ en la suma de probabilidades para $n = 32$ y la exclusión del mismo cuando pasamos a $n = 33$** , ya que la probabilidad del valor $j = 8$ (que es igual a 0'0327) es muy superior (del orden de 10 veces) a la suma de diferencias de probabilidades para el resto de valores que sí aparecen en ambos sumatorios (que es igual a 0'0029).

Tabla VI
Probabilidades de cubrimiento para $n=32$, $n=33$, $p=0'4$, Intervalo de Wald

j	$P[X(32,0'4) = j]$	$P[X(33,0'4) = j]$	Diferenc.
8	0,0327	***	***
9	0,0581	0,0479	-0,0102
10	0,0890	0,0766	-0,0124
11	0,1187	0,1068	-0,0119
12	0,1385	0,1306	-0,0079
13	0,1421	0,1406	-0,0014
14	0,1285	0,1339	0,0054
15	0,1028	0,1131	0,0103
16	0,0728	0,0848	0,0120
17	0,0457	0,0566	0,0109
18	0,0254	0,0335	0,0081
Sumas	0,9543	0,9245	0,0029

¹ Brown, L. D., Cal, T., y Dasgupta, A. (2001). “Interval Estimation for a Binomial Proportion”. Statistical Science, vol. 16, No. 2, 101-133., Página 104.

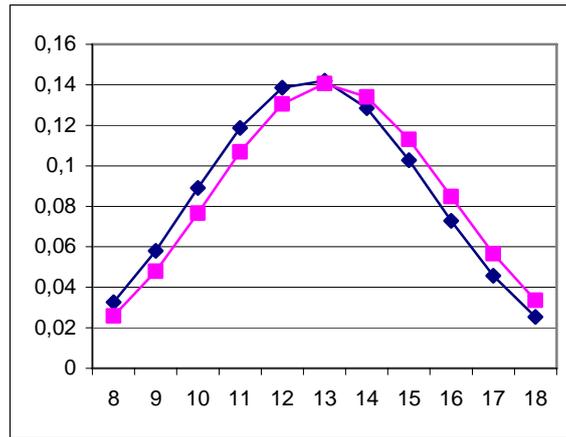


Figura IV. Probabilidades de Binomiales $n=32$ y $n=33$ ($p=0.4$)

Ejemplo 2: $p = 0.2$, $n = 31$.

En este caso, podemos comprobar que $C(31,0.2) = 0.9498$ mientras que para $n = 32$ la probabilidad de cubrimiento es $C(32,0.2) = 0.8902$. En este caso, se obtiene $L_1(31,0.2) = 2.9822$, $L_2(31,0.2) = 11.4685$ mientras que para $n = 32$ los límites son $L_1(32,0.2) = 3.1140$, $L_2(32,0.2) = 11.7439$ por lo que las probabilidades de cubrimiento vienen

dadas por $C(31,0.2) = \sum_{j=3}^{11} P[X(31,0.2) = j]$ y $C(32,0.2) = \sum_{j=9}^{11} P[X(32,0.2) = j]$. En la tabla VII y

en la figura V podemos apreciar que las diferencias de probabilidad de los valores 3 a 11 cuando pasamos de $n = 31$ a $n = 32$ ya no son aproximadamente simétricas puesto que consideramos más valores a la derecha de la moda que a la izquierda. No obstante, estas diferencias en valor absoluto disminuyen con respecto al caso del ejemplo 1, resultando de nuevo que la probabilidad del valor $j = 3$ (que es igual a 0.0696) es muy superior a la suma de diferencias de probabilidades para el resto de valores que sí aparecen en ambos sumatorios (que es igual a 0.0099). Así, volvemos a observar en este ejemplo que **el hecho que resulta decisivo en la diferencia de cubrimiento es la inclusión del caso $j = 3$ en la suma de probabilidades para $n = 31$ y la exclusión del mismo cuando pasamos a $n = 32$.**

Tabla VII

Probabilidades de cubrimiento para $n=31, n=32, p=0'2$, Intervalo de Wald

j	$P[X(31,0'2) = j]$	$P[X(32,0'2) = j]$	Diferenc.
3	0,0696	****	****
4	0,1217	0,1113	-0,0104
5	0,1643	0,1558	-0,0085
6	0,1780	0,1753	-0,0027
7	0,1589	0,1628	0,0038
8	0,1192	0,1272	0,0079
9	0,0762	0,0848	0,0086
10	0,0419	0,0487	0,0069
11	0,0200	0,0244	0,0044
Sumas	0,9498	0,8902	0,0099

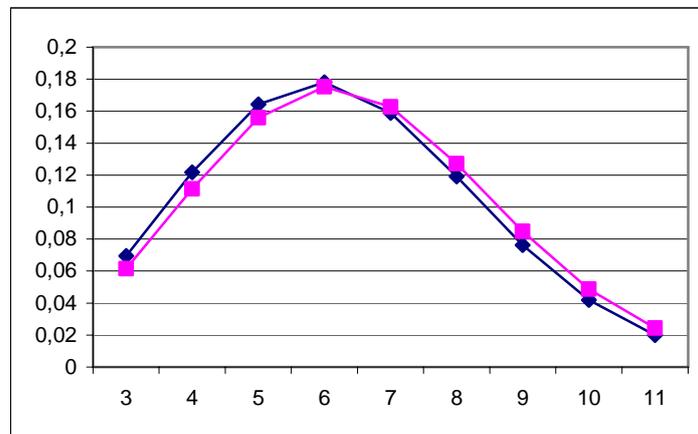


Figura V. Probabilidades de Binomiales $n=31$ y $n=32$ ($p=0'2$)

Ejemplo 3: $p = 0'005, n = 591$.

En este caso, podemos comprobar que $C(591,0'005) = 0'9449$ mientras que para $n = 592$ la probabilidad de cubrimiento es $C(592,0'005) = 0'7921$. En este caso, se obtiene $L_1(591,0'005) = 0'9983, L_2(591,0'005) = 8'6902$ mientras que para $n = 592$ los límites son $L_1(592,0'005) = 1'0009, L_2(592,0'005) = 8'6977$ por lo que las probabilidades de cubrimiento son

$$C(591,0'005) = \sum_{j=1}^8 P[X(591,0'005) = j] \quad \text{y} \quad C(592,0'005) = \sum_{j=2}^8 P[X(592,0'005) = j].$$

En la tabla VIII y en la figura VI podemos apreciar que las diferencias de probabilidad de los valores 1 a 8 cuando pasamos de $n = 591$ a $n = 592$ ya son totalmente asimétricas pero, no obstante, estas diferencias en valor absoluto son muy cercanas a cero, resultando de nuevo que la probabilidad del

valor $j=1$ (que es igual a 0'1535) es muy superior (del orden de 200 veces) a la suma de diferencias de probabilidades para el resto de valores que sí aparecen en ambos sumatorios (que es igual a 0'0007). Así, volvemos a observar en este ejemplo que **el hecho que resulta decisivo en la diferencia de cubrimiento es la inclusión del caso $j=1$ en la suma de probabilidades para $n=591$ y la exclusión del mismo cuando pasamos a $n=592$.**

Tabla VIII

Probabilidades de cubrimiento para $n=591$, $n=592$, $p=0'005$, Intervalo de Wald

j	$P[X(591,0'005) = j]$	$P[X(592,0'005) = j]$	Diferenc.
1	0,1535	***	***
2	0,2276	0,2272	-0,0004
3	0,2245	0,2246	0,0000
4	0,1659	0,1662	0,0003
5	0,0979	0,0982	0,0003
6	0,0480	0,0483	0,0002
7	0,0202	0,0203	0,0001
8	0,0074	0,0075	0,0001
Sumas	0,9449	0,7922	0,0007

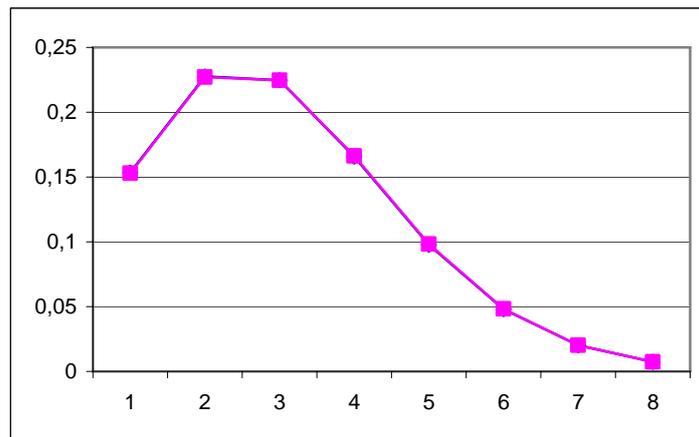


Figura VI. Probabilidades de Binomiales $n=591$ y $n=592$ ($p=0'005$)

El análisis de algunos ejemplos, nos llevó a calcular las probabilidades de cubrimiento y de los valores $L_1(n, p)$ y $L_2(n, p)$ para diversos valores de n y p (todos ellos con $p \leq 0'5$). A modo de ejemplo, reproducimos en la tabla IX los cálculos efectuados para $p = 0'5$. Hemos resaltado los casos en los que se produce disminución en la probabilidad de cubrimiento al pasar de n a $n+1$, que, como puede apreciarse, siempre coinciden con los casos en los que cambia la parte entera de la función $L_1(n, p)$.

Tabla IX

Probabilidades de cubrimiento y extremos del sumatorio, $p=0.5$. Intervalo de Wald

n	$L_1(n, 0.5)$	$L_2(n, 0.5)$	$C(n, 0.5)$
10	2.3659	7.6341	0.8906
11	2.7018	8.2982	0.9346
12	3.0454	8.9546	0.8540
13	3.3956	9.6044	0.9077
14	3.7519	10.2481	0.9427
15	4.1135	10.8865	0.8815
16	4.4799	11.5201	0.9232
17	4.8508	12.1492	0.9510
18	5.2256	12.7744	0.9037
19	5.6041	13.3959	0.9364
20	5.9860	14.0140	0.9586
21	6.3710	14.6290	0.9216
22	6.7589	15.2411	0.9475
23	7.1495	15.8505	0.9069

En todos los casos que hemos considerado, hemos observado, sin excepción alguna, que se verifica la siguiente regla empírica (para $p \leq 0.5$):

$$C(n+1, p) < C(n, p) \Leftrightarrow E[L_1(n+1, p)] > E[L_1(n, p)],$$

donde $E[\square]$ representa la función parte entera.

Fijado un valor de α , esta regla nos va a permitir para cada valor de p , “predecir” todos los valores de n en los que se va a producir un descenso en la probabilidad de cubrimiento. Por tanto, la aparición de estos valores no es imprevisible, como afirman Brown, Cal y Dasgupta (2001).

En efecto, podemos obtener la función $L_1(n, p)$ explícitamente, ya que la ecuación implicada (al elevarla al cuadrado para eliminar la raíz) es de segundo grado. De las dos soluciones, la única que verifica la ecuación original es:

$$L_1(n, p) = \frac{n(2np - z_{\alpha/2}^2) - \sqrt{4n^3 p z_{\alpha/2}^2 - 4n^3 p^2 z_{\alpha/2}^2 + n^2 z_{\alpha/2}^4}}{2(n + z_{\alpha/2}^2)}$$

Esta función es estrictamente creciente (como función de n). Resolviendo las ecuaciones $L_1(n, p) = k$, $k \in \mathbb{N}$, y siendo n_k^* , $k \in \mathbb{N}$ las soluciones de dichas ecuaciones, obtenemos que los valores de n en los que $E[L_1(n+1, p)] > E[L_1(n, p)]$ vienen dados por la sucesión $\{E[n_k^*]\}_{k \in \mathbb{N}}$, por lo que los valores de n en los que se producirá un descenso en la probabilidad de cubrimiento

vienen dados por $n_k = E[n_k^*] + 1$, $k \in \mathbb{N}$. Además, las ecuaciones $L_1(n, p) = k$, $k \in \mathbb{N}$ pueden resolverse explícitamente, por lo que finalmente podemos obtener la fórmula que nos permite calcular todos los valores de n en los que descenderá la probabilidad de cubrimiento. En efecto, la ecuación $L_1(n, p) = k$ finalmente conducen a una ecuación de tercer grado, con una única solución válida que viene dada por:

$$n_k^* = \frac{2k}{p} - \frac{2^{1/3}(-k^2 p^2 - 3kp^2 z_{\alpha/2}^2)}{h(p, z_{\alpha/2}, k)} + \frac{h(p, z_{\alpha/2}, k)}{2^{1/3} 3p^2},$$

donde

$$h(p, z_{\alpha/2}, k) = \left(-2k^3 p^3 + 18k^2 p^3 z_{\alpha/2}^2 - 27k^2 p^4 z_{\alpha/2}^2 + \sqrt{4(-kp^2 - 3kp^2 z_{\alpha/2}^2)^3 + (-2k^3 p^3 + 18k^2 p^3 z_{\alpha/2}^2 - 27k^2 p^4 z_{\alpha/2}^2)^2} \right)^{1/3}$$

Por ejemplo, para $p = 0.5$, los primeros términos de la sucesión vienen dados por $\{6, 9, 12, 15, 18, 21, 23, 26, 28, 31, \dots\}$. En la tabla IX hemos observado que se produce disminución de cubrimiento en los valores 12, 15, 18, 21, y 23. A partir de la sucesión podemos saber también sin necesidad de calcular las probabilidades que se produce disminución en los valores 6, 9, 26, 28, 31, ...

En Brown, Cal y Dasgupta (2001), se representa gráficamente la probabilidad de cubrimiento para $p = 0.2$ y n variando desde 25 a 100. En el caso $p = 0.2$ (tomando, como siempre $\alpha = 0.05$), los primeros términos de nuestra sucesión son: $\{15, 24, 32, 39, 46, 53, 60, 66, 73, 79, 86, 92, 98, 105, 111, 117, 123, \dots\}$, valores que concuerdan totalmente con los apreciados a través del cálculo directo de las probabilidades de cubrimiento. En este mismo artículo, también se considera el caso $p = 0.005$, apreciándose descenso en la probabilidad de cubrimiento para los valores de n 592, 954, 1279, 1583 y 1876. Los primeros valores de nuestra sucesión en este caso son $\{592, 954, 1279, 1583, 1876, 2159, 2436, 2708, 2975, 3238, \dots\}$, lo que nos permitiría señalar los próximos valores de n en los que ocurrirá un descenso de la probabilidad de cubrimiento. En efecto, las probabilidades de cubrimiento en el entorno de $n = 2159$ se ofrecen en la tabla X, donde podemos apreciar que el descenso se produce efectivamente en el valor predicho.

Tabla X

Probabilidades de cubrimiento y extremos del sumatorio, $p=0'005$. Intervalo de Wald

n	$L_1(n, 0'005)$	$L_2(n, 0'005)$	$C(n, 0'005)$
2155	5.9863	19.3600	0.950078
2156	5.9898	19.3665	0.950169
2157	5.9934	19.3729	0.950259
2158	5.9970	19.3793	0.950349
2159	6.0006	19.3857	0.905555
2160	6.0041	19.3922	0.905744
2161	6.0077	19.3986	0.905932
2162	6.0113	19.4050	0.906120

5. Análisis empírico del comportamiento del intervalo de Wilson. Comparación con el de Wald.

En esta sección consideraremos algunos ejemplos concretos para el caso del intervalos de Wilson, al objeto de buscar una explicación empírica al hecho de que este intervalo tenga en general un comportamiento menos errático que el de Wald debido a que los descensos apreciados en las probabilidades son mucho menos bruscos. Seguiremos centrándonos en considerar valores fijos de p mientras que hacemos variar n . En la figura VII, hemos representado las probabilidades de cubrimiento del intervalo de Wilson para $p = 0'25$ y distintos valores de n . Si la comparamos con la figura III, en la que representamos las probabilidades de cubrimiento en el intervalo de Wald para los mismos valores de n y p , podemos apreciar como el caso de Wilson presenta fluctuaciones menos marcadas. Observemos además que el cubrimiento del intervalo de Wilson se acerca mucho más al valor nominal (que en este caso es $0'95$).

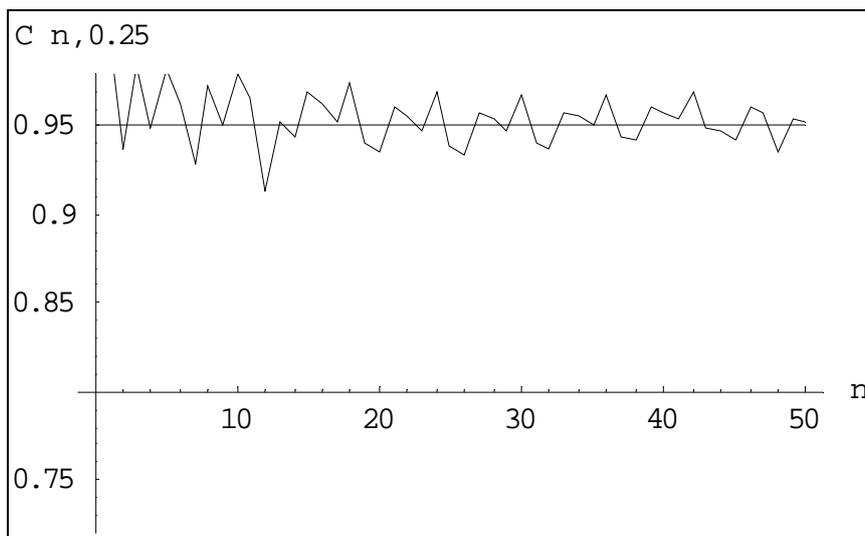


Figura VII. Probabilidades de Cubrimiento para distintos valores de n y $p=0'25$. Intervalo de Wilson.

Fijado un α , a partir de la expresión del intervalo de Wilson, podemos obtener la probabilidad de cubrimiento $C^*(n, p)$ dados unos valores de p y n como la suma de las probabilidades de ciertos valores en la distribución $Bi(n, p)$. Concretamente,

$$C(n, p) = \sum_{j \in (L_1^*(n, p), L_2^*(n, p))} P[X(n, p) = j],$$

donde $X(n, p) \sim Bi(n, p)$ y $L_1^*(n, p)$ y $L_2^*(n, p)$ son las soluciones de las ecuaciones

$$\frac{x + \frac{z_{\alpha/2}^2}{2}}{n + z_{\alpha/2}^2} + \frac{z_{\alpha/2} n^{1/2}}{n + z_{\alpha/2}^2} \left[\frac{x}{n} \left(1 - \frac{x}{n} \right) + \frac{z_{\alpha/2}^2}{4n} \right]^{1/2} = p \quad \text{y} \quad \frac{x + \frac{z_{\alpha/2}^2}{2}}{n + z_{\alpha/2}^2} - \frac{z_{\alpha/2} n^{1/2}}{n + z_{\alpha/2}^2} \left[\frac{x}{n} \left(1 - \frac{x}{n} \right) + \frac{z_{\alpha/2}^2}{4n} \right]^{1/2} = p,$$

respectivamente. Por ejemplo, para $n=31$ y $p=0'2$ al resolver las ecuaciones anteriores obtenemos $L_1^*(31, 0'2) = 1.8350$ y $L_2^*(31, 0'2) = 10.5650$, por lo que la probabilidad de cubrimiento

$$\text{del intervalo de Wilson será } C^*(31, 0'2) = \sum_{j=2}^{10} P[X(31, 0'2) = j] = 0'9586.$$

Ejemplo 4: $p = 0'2$, $n = 32$.

En este caso, se obtiene $L_1^*(32, 0'2) = 1.9651$ y $L_1^*(33, 0'2) = 2.0964$, es decir, se produce un salto en la parte entera de la función $L_1(n, 0'2)$ al pasar de $n = 32$ a $n = 33$. No obstante, podemos comprobar que las probabilidades de cubrimiento de los respectivos intervalos de Wilson vienen dadas por $C^*(32, 0'2) = 0.9518$ y $C^*(33, 0'2) = 0.9516$. La ausencia de un descenso brusco de probabilidad se debe a que el salto que se produce en L_1^* (lo que supone perder un término en el sumando que nos proporciona la probabilidad de cubrimiento) se compensa porque L_2^* también “salta” (en concreto, $L_2^*(32, 0'2) = 10.8350$ y $L_2^*(33, 0'2) = 11.1037$). Como puede apreciarse en la tabla XI y en la figura VIII, sigue ocurriendo que las diferencias de probabilidad en los términos comunes a ambos sumatorios en total tienden a sumar un valor cercano a cero (notar que en el caso de Wilson la figura es más simétrica que en el de Wald); pero ahora cada sumatorio tiene un término “extra” que no aparece en el otro y debido a la simetría aproximada, ambos términos tienen probabilidad muy similar con lo que en conjunto, las probabilidades de cubrimiento son muy similares.

Tabla XI

Probabilidades de cubrimiento para $n=32, n=33, p=0'2$, Intervalo de Wilson

j	$P[X(32,0'2) = j]$	$P[X(33,0'2) = j]$	Diferenc.
2	0,0246	***	***
3	0,0614	0,0540	-0,0074
4	0,1113	0,1013	-0,0100
5	0,1558	0,1469	-0,0089
6	0,1753	0,1714	-0,0039
7	0,1628	0,1653	0,0025
8	0,1272	0,1343	0,0071
9	0,0848	0,0932	0,0085
10	0,0487	0,0559	0,0072
11	***	0,0292	***
Sumas	0,9518	0,9516	-0,0048

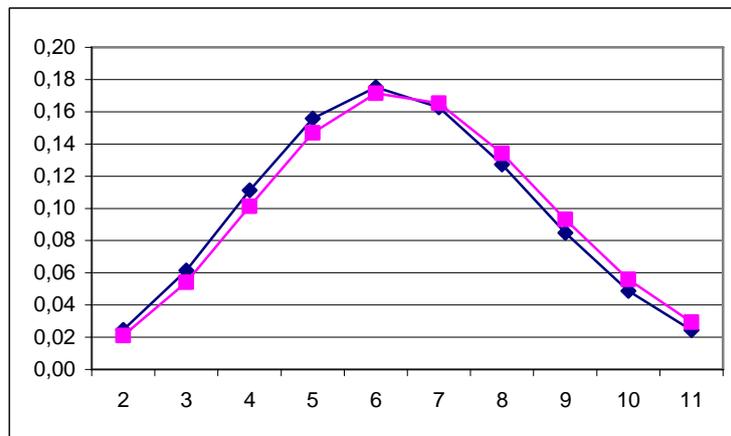


Figura VIII. Probabilidades de Binomiales $n=32$ y $n=33$ ($p=0'2$). Int. Wilson

Tras haber realizado los cálculos correspondientes en multitud de situaciones distintas, comprobamos empíricamente que en el intervalo de Wald (para $p \leq 0'5$) cuando aumenta la parte entera de L_1 **nunca** cambia la parte entera de L_2 (lo que conlleva descenso brusco de probabilidad). Por su parte en el caso del intervalo de Wilson, cuando aumenta la parte entera de L_1^* en muchas ocasiones aumenta también la parte entera de L_2^* y esta es la causa de que las variaciones sean más suaves en el intervalo de Wilson. De hecho, los descensos más importantes en el intervalo de Wilson también se dan cuando aumenta la parte entera de L_1^* mientras que la parte entera de L_2^* se mantiene, pero esta situación es poco frecuente (esto ocurre, por ejemplo, para $p = 0'4$, $n = 30$ donde podemos observar que $L_1^*(30,0'4) = 6'7409$, $L_1^*(31,0'4) = 7'0539$, $L_2^*(30,0'4) = 17'2591$ y $L_2^*(31,0'4) = 17'7461$, siendo las probabilidades de cubrimiento $C^*(30,0'4) = 0.9616$ y $C^*(31,0'4) = 0.9350$).

5. Conclusiones y cuestiones abiertas.

Como hemos indicado en la sección 4, para calcular la probabilidad de cubrimiento del intervalo de Wald para unos valores n y p , debemos sumar las probabilidades de todos los enteros contenidos en un cierto intervalo $(L_1(n, p), L_2(n, p))$. La primera conclusión que queremos destacar es que el comportamiento errático del intervalo de Wald se debe a que en ciertos casos, al pasar de n a $n+1$ (para un $p \leq 0.5$ fijo) dicho intervalo pierde un entero (por el extremo inferior). Empíricamente, hemos visto que: i) Cuando aumenta la parte entera de L_1 , **nunca** aumenta la parte entera de L_2 y ii) El aumento de la parte entera de L_1 al pasar de n a $n+1$ queda biunívocamente asociada al descenso en la probabilidad de cubrimiento al pasar de n a $n+1$.

A partir de las reglas anteriores, podemos obtener una fórmula que nos proporciona todos los valores de n en los que se produce un descenso en la probabilidad de cubrimiento (para un $p \leq 0.5$ fijo), por lo que dichos valores no tienen un comportamiento impredecible, como indican Brown, Cal y Dasgupta (2001).

Al comparar el intervalo de Wald con el de Wilson, observamos en este último un comportamiento menos errático, y esto se debe a que en muchas de las ocasiones en las que se produce un aumento en la parte entera de L_1^* , lo que provoca la pérdida de un entero por el extremo inferior, éste se compensa con un aumento en la parte entera de L_2^* , lo que provoca la ganancia de un entero por el extremo superior (recordemos que la probabilidad de cubrimiento del intervalo de Wilson se obtenía sumando las probabilidades de todos los enteros contenidos en un cierto intervalo $(L_1^*(n, p), L_2^*(n, p))$, como vimos en la sección 5).

Destaquemos que en el caso del intervalo de Wald no hemos encontrado ninguna excepción a las reglas empíricas expuestas, mientras que en el intervalo de Wilson la regla que hemos enunciado se da con excepciones.

Como cuestiones abiertas para continuar con esta investigación quedaría en primer lugar buscar demostraciones formales del resultado empírico obtenido y por otra parte sería interesante buscar alguna explicación al hecho de que el intervalo de Wilson habitualmente tenga una probabilidad de cubrimiento más cercana al valor nominal que el de Wald.

Referencias Bibliográficas

Brown, L. D., Cal, T., y Dasgupta, A. (2001). "Interval Estimation for a Binomial Proportion". Statistical Science. vol. 16, No. 2, 101-133.

Spiegel, M. R., (1991), Estadística. Madrid. McGraw-Hill. 1993. 217-218.

Wilson, E. B. (1927). "Probable inference, the law of succession, and statistical inference". J.Amer. Statist. Assoc. 22, 209-212.

MODELIZACIÓN Y ANÁLISIS DE DATOS SOBRE PROPORCIONES

JAVIER PALAREA ALBALADEJO

e-mail: jpalarea@pdi.ucam.edu

Departamento de Informática de Sistemas
UNIVERSIDAD CATÓLICA SAN ANTONIO

JOSEP ANTONI MARTIN FERNÁNDEZ

e-mail: josepantoni.martin@udg.edu

Departamento de Informática y Matemática Aplicada
UNIVERSIDAD DE GIRONA

JUAN GÓMEZ GARCÍA

e-mail: jgomezg@um.es

Departamento de Métodos Cuantitativos para la Economía
UNIVERSIDAD DE MURCIA

Área temática: Métodos Cuantitativos.

Resumen

En determinados ámbitos es frecuente trabajar con datos que representan partes de un total: proporciones, porcentajes, partes por millón, o similar. Es el caso de, por ejemplo, la distribución del presupuesto familiar, la composición de una cartera de inversión, el empleo del tiempo diario en distintas actividades, la distribución de ventas en distintas regiones, y otros. Las restricciones de no negatividad y de suma constante que caracterizan este tipo de datos implican que las técnicas multivariantes habitualmente utilizadas no son adecuadas para su análisis y modelización. La cuestión clave es que la geometría del espacio muestral sobre el que se definen, el simplex, es diferente de la clásica geometría Euclídea del espacio real. En el presente trabajo revisaremos los fundamentos teóricos, la geometría del simplex, la metodología log-cociente, e importantes aspectos prácticos como el problema de las componentes nulas. El apartado teórico se acompaña de ejemplos que motivan la exposición y ponen de relieve la naturaleza del problema que se está tratando. Finalmente, se ilustra la aplicación práctica de las ideas y métodos estudiados a través de ejemplos del campo económico.

Palabras clave: Proporciones, Datos Composicionales, Simplex, Análisis Log-cociente.

Abstract

In some scopes analysts frequently work with data that represent parts of a whole: proportions, percentages, parts per million, or similar. This is the case of household budgets, investment portfolios, time budgets, sales by region, and others. The no-negativity and constant-sum constraints characterizing this kind of data imply that common multivariate techniques are not suitable for modelling purposes. The key feature is that the geometry of the sample space on which data are defined, the simplex, differs from the classical Euclidean geometry of real space. In this paper, the theoretical foundations, the geometry of the simplex, the log-ratio methodology and some other relevant practical issues will be reviewed. The theoretical aspects are illustrated by examples which motivate the exposition. Finally, some applications from the economic field are presented.

Key words: Proportions, Compositional Data, Simplex, Log-ratio Analysis.

1. Introducción

Ya en 1897 Karl Pearson (Pearson, 1897) advertía de las dificultades para interpretar la relación entre variables que representaran partes de un total. Habitualmente, dichas variables vienen expresadas en proporciones, porcentajes, partes por millón, o similar. Pongamos como ejemplos en el campo económico la distribución del presupuesto familiar en distintas partidas de gasto, la composición relativa de una cartera de inversión, el patrón de actividad de una cadena de producción, la composición étnica de una población, el empleo del tiempo diario en distintas actividades, la distribución de las ventas de un producto en distintas regiones, la distribución de trabajadores en distintos sectores de actividad, y otros. La característica fundamental de un vector de datos de este tipo es que la suma de sus componentes, siempre positivas, es una constante. Así, un cambio en una componente conlleva un cambio en, al menos, una de las demás componentes. Por lo que hay que replantearse, entre otros, el concepto de independencia. Además, una fila cualquiera de la matriz de covarianzas de una muestra de vectores de proporciones siempre tiene al menos un elemento negativo y su suma es igual a 0. Esto implica que la matriz de covarianzas es singular y que las correlaciones no varían libremente en el habitual intervalo $[-1,1]$.

La cuestión clave es que la geometría del espacio muestral sobre el que se define un vector de proporciones es diferente de la clásica geometría Euclídea de \mathbb{R}^D , por ello las técnicas multivariantes habitualmente utilizadas, y fundamentadas en esta geometría, no son directamente aplicables. Para llamar la atención del lector proponemos el siguiente ejemplo. Consideremos los patrones de actividad de dos operarios (A, B) entre tres actividades mutuamente excluyentes (X, Y, Z) en dos periodos (1, 2) de 120 minutos. La asignación de tiempo a cada actividad en el periodo 1 es [10, 80, 30] para el operario A y [20, 80, 20] para el B. En el periodo 2 estos tiempos son, respectivamente, [20, 70, 30] y [30, 70, 20]. La dedicación del operario A a la actividad X se dobla de un periodo a otro, mientras que para la Y se reduce en un 12.5% y no cambia para la Z. El operario B incrementa el tiempo dedicado a X en un 50%, mientras que se reduce en un 12.5% el dedicado a Y y tampoco se modifica el dedicado a Z. Una medida de diferencia adecuada debería detectar un mayor cambio en el patrón de actividad de A que en el de B. Sin embargo, si medimos esta diferencia entre periodos utilizando la distancia Euclídea habitual obtenemos un valor de 14.1421, el mismo para los dos operarios.

Pensemos en las consecuencias de este hecho en, por ejemplo, un análisis cluster. De manera que cualquier medida basada de forma más o menos explícita en la distancia Euclídea no se comportará bien en este contexto.

Aún así, durante el pasado siglo se han aplicado, y aún se aplican, técnicas estadísticas multivariantes para datos reales sobre este tipo de variables, sin tener en cuenta su especial naturaleza. Cómo trabajar con la restricción de suma constante ha sido durante largo tiempo un problema aparentemente irresoluble, y el hecho de no tenerla en cuenta, o de interpretarla de forma inadecuada, ha llevado a conclusiones erróneas en muchos análisis de datos. Algunas actuaciones particulares en este línea se han orientado hacia intentar salvar los problemas técnicos y no hacia un estudio ni del tipo de relación relevante entre las partes de un total ni de su peculiar espacio muestral (véase p. ej. Bohling y otros, 1998). No es hasta los años 80 con la publicación de la monografía de Aitchison (1986) cuando se dispone de una obra de referencia sobre los fundamentos teóricos y sobre una metodología específica para el análisis estadístico de tales vectores de proporciones, que desde entonces suelen denominarse *datos composicionales*.

En la sección 2 resumiremos los fundamentos sobre los que se asienta el análisis de datos composicionales. A continuación, en la sección 3, destacaremos los aspectos más relevantes de la metodología basada en transformaciones para trasladar los datos al espacio real donde poder aplicar las técnicas habituales. Al mismo tiempo, pondremos de relieve algunos de los problemas prácticos que pueden surgir. Finalmente, dedicaremos la sección 4 a destacar la aplicación de las herramientas del análisis de datos composicionales en el campo de la economía.

2. Fundamentos matemáticos

Formalmente, un dato composicional es un vector $\mathbf{x} = [x_1, \dots, x_D]$ constituido por D partes o componentes positivas cuyo espacio muestral es el simplex S^D definido como $S^D = \{\mathbf{x} = [x_1, \dots, x_D] : x_1 > 0, \dots, x_D > 0; \sum_j x_j = c\}$. En realidad, el valor de c es irrelevante desde un punto de vista matemático, por lo que en adelante consideraremos $c = 1$, esto es, las componentes x_j están medidas en proporciones. Este espacio muestral tiene una estructura de espacio vectorial inducida por las operaciones *perturbación* y *potenciación*. Para cualquier par de composiciones \mathbf{x} y \mathbf{x}' de S^D la operación

perturbación se define como $\mathbf{x} \oplus \mathbf{x}' = C(x_1 x'_1, \dots, x_D x'_D)$ y la operación de potenciación como $\alpha \otimes \mathbf{x} = C(x_1^\alpha, \dots, x_D^\alpha)$, donde α es un valor real. La perturbación es una operación fundamental encargada de describir el *cambio* composicional en el simplex. El operador C se denomina *clausura* y normaliza un vector de componentes positivas dividiendo cada una por el valor de la suma de todas ellas. Para tener una idea intuitiva del significado de estas operaciones básicas en el simplex diremos que la perturbación es el equivalente a la traslación o suma en espacios reales y la potenciación es el equivalente al producto por escalar. A partir de ellas podemos definir la *diferencia* composicional como $\mathbf{x} \ominus \mathbf{x}' = \mathbf{x} \oplus (-1) \otimes \mathbf{x}' = C(x_1 / x'_1, \dots, x_D / x'_D)$. El papel de marginales en el simplex lo hacen las *subcomposiciones*, que son las proyecciones del simplex S^D , el espacio de las composiciones con D partes, sobre un sub-simplex de dimensión menor, digamos S^d , obtenidas mediante la clausura de un subvector formado por d de las partes de una composición en S^D . Por ejemplo, a partir de la composición $\mathbf{x} = [0.2, 0.6, 0.1, 0.1]$ de S^4 podemos obtener la subcomposición $\mathbf{s} = [0.2222, 0.6667, 0.1111]$ de S^3 clausurando el subvector formado por las 3 primeras componentes. Notar que las operaciones elementales en espacios reales no se comportan consistentemente cuando se aplican sobre composiciones. Así, la suma o resta de composiciones no da como resultado otra composición, esto es, pueden producir un elemento que no pertenece al simplex.

Adicionalmente, es posible definir un producto interior en S^D

$$\langle \mathbf{x}, \mathbf{x}' \rangle = \left[\frac{1}{D} \sum_{i < j} \ln \frac{x_i}{x_j} \ln \frac{x'_i}{x'_j} \right], \quad (1)$$

que confiere al simplex estructura de espacio Euclídeo (Billheimer, Guttorp y Fagan, 2001; Pawlowsky-Glahn y Egozcue, 2001, 2002). Desde un punto de vista estadístico esta estructura no es irrelevante, ya que las medidas básicas de tendencia central, variabilidad y distancia deben ser coherentes con la naturaleza de los datos. El producto interior (1) induce una distancia en S^D definida como

$$d_a(\mathbf{x}, \mathbf{x}') = \left[\frac{1}{D} \sum_{i < j} \left(\ln \frac{x_i}{x_j} - \ln \frac{x'_i}{x'_j} \right)^2 \right]^{1/2} = \left[\sum_{i=1}^D \left(\ln \frac{x_i}{g(\mathbf{x})} - \ln \frac{x'_i}{g(\mathbf{x}')} \right)^2 \right]^{1/2}, \quad (2)$$

donde $g(\mathbf{x}) = (x_1 \cdots x_D)^{1/D}$ es la media geométrica de las componentes de la composición \mathbf{x} . La distancia (2) se conoce como *distancia de Aitchison* y es totalmente compatible con las operaciones básicas en el simplex y con la particular naturaleza de los datos composicionales (Aitchison y otros, 2000). La estructura de espacio Euclídeo implica que sobre el simplex pueden definirse todos los entes y conceptos geométricos usuales tales como líneas, ángulos, ortogonalidad, paralelismo, etc. Sin embargo, aunque matemáticamente equivalentes, las características, y el aspecto, de estos elementos en la geometría del simplex difieren de aquellas a las que estamos habituados en la geometría Euclídea clásica del espacio real. Para evitar confusiones nos referiremos a la geometría del simplex como *geometría de Aitchison*.

3. Análisis log-cociente

Una metodología adecuada para el análisis de datos composicionales debe tener en cuenta algunos principios lógicos necesarios y las características del simplex como espacio muestral sobre el que se definen. La idea principal es que las composiciones sólo proporcionan información sobre la magnitud relativa de sus partes, y no pueden justificarse interpretaciones que involucren a las magnitudes absolutas. Se asume que el valor de la suma de las partes es irrelevante. Por lo tanto, cualquier aseveración sobre una composición debe hacerse en términos de los cocientes entre las partes, los cuales medirán dicha relación relativa. Así, una función aplicable sobre composiciones deberá ser invariante por cambios de escala y expresable en términos de cocientes entre las partes.

Trabajar con cocientes asegura además un principio lógico básico: la *coherencia subcomposicional*. Consideremos dos analistas financieros, A y B, interesados en estudiar la relación entre las proporciones de capital invertido por los clientes en distintos productos financieros. Supongamos que existen cuatro productos, siendo x_1 , x_2 , x_3 y x_4 las proporciones de capital invertido en cada uno de ellos, y que el analista A tiene acceso a toda la información sobre las cantidades invertidas, trabaja con la composición $[x_1, x_2, x_3, x_4]$, mientras que el analista B sólo tiene acceso a la información sobre los tres primeros, trabaja con la subcomposición $[s_1, s_2, s_3]$. Aún así, lo lógico sería que si B calcula el coeficiente de correlación lineal habitual entre s_1 y s_2 obtenga el mismo resultado que si A lo calcula entre x_1 y x_2 , las proporciones se refieren

a los mismos productos financieros. Pero esto no es así cuando la medida se calcula sobre las proporciones absolutas. Recogemos a continuación los datos que maneja cada analista sobre los mismos tres clientes:

Datos analista A: [x_1 x_2 x_3 x_4]	Datos analista B: [s_1 s_2 s_3]
[0.2, 0.1, 0.2, 0.5]	[0.4, 0.2, 0.4]
[0.1, 0.1, 0.1, 0.7]	[0.3333, 0.3333, 0.3333]
[0.6, 0.2, 0.1, 0.1]	[0.6667, 0.2222, 0.1111]

El coeficiente de correlación lineal entre las proporciones de capital invertido en los productos 1 y 2 es 0.982 para el analista A y -0.5291 para el B. Por lo tanto, queda claro que esta medida de dependencia no es coherente cuando se aplica directamente sobre proporciones. Sin embargo, es fácil comprobar que los cocientes entre las partes x_i / x_j no varían al pasar de la composición a la subcomposición. Así, para el cliente 1, $x_1 / x_2 = s_1 / s_2 = 2$. Por lo que si trabajamos con los cocientes evitaremos el problema de la incoherencia subcomposicional.

Para destacar un poco más la conveniencia de trabajar con cocientes entre las partes, retomemos el ejemplo de la introducción sobre el patrón de actividad de los dos operarios. Vimos que la distancia Euclídea no se comportaba según lo esperado, proporcionaba la misma diferencia entre periodos para los dos operarios. En la sección 2 se define la distancia de Aitchison, la cual se deduce de la propia geometría del simplex. Si nos fijamos en la expresión (2) vemos que se expresa en términos de cocientes entre las partes. Utilizando (2) como medida de diferencia entre los patrones de actividad de cada operario en los dos periodos obtenemos un valor de 0.6276 para el operario A, y un valor de 0.3970 para el operario B. Esto es, un mayor cambio en el patrón de actividad de A que en el de B, que es lo que efectivamente ocurre.

3.1 Transformaciones log-cociente

Una vez que se ha puesto de relieve la necesidad de centrar la atención en los cocientes entre las partes surge la pregunta sobre qué tipo de cocientes utilizar. La piedra angular de la metodología propuesta por Aitchison (1986) es la transformación de una composición definida sobre S^D en un vector que involucre los cocientes entre las partes y que esté definido sobre el espacio real. Si esa transformación es biyectiva se establece

una correspondencia uno a uno entre las composiciones en el simplex y los correspondientes vectores transformados reales. De esta manera, cualquier problema que afecte a composiciones queda expresado en términos de tales vectores transformados, con lo que se tiene la posibilidad de resolverlo utilizando las técnicas multivariantes habituales en espacios reales. Aitchison propone fundamentalmente dos tipos de transformaciones: la *transformación log-cociente aditiva* (alr) sobre \mathbb{R}^{D-1} , definida como

$$\text{alr}(\mathbf{x}) = \left[\ln \frac{x_1}{x_D}, \dots, \ln \frac{x_{D-1}}{x_D} \right], \quad (3)$$

y la *transformación log-cociente centrada* (clr) sobre \mathbb{R}^D , definida como

$$\text{clr}(\mathbf{x}) = \left[\ln \frac{x_1}{g(\mathbf{x})}, \dots, \ln \frac{x_D}{g(\mathbf{x})} \right]. \quad (4)$$

El hecho de tomar logaritmos de los cocientes sólo responde a una conveniencia matemática. Los log-cocientes son más manejables y además permiten que se cumplan algunas propiedades sencillas. Por ejemplo, no existe una relación exacta entre las varianzas $\text{var}(x_i/x_j)$ y $\text{var}(x_j/x_i)$ y, sin embargo, sí se cumple que $\text{var}[\ln(x_i/x_j)] = \text{var}[\ln(x_j/x_i)]$.

La transformación alr está más vinculada a modelizaciones paramétricas ya que a partir de ella es posible definir una clase de distribuciones, la *normal logística aditiva*, (Aitchison y Shen, 1980), con densidad

$$f(\mathbf{x}) = \frac{\det(2\pi\Sigma)^{-1/2}}{x_1 \cdots x_D} \exp\left\{-\frac{1}{2}(\text{alr}(\mathbf{x}) - \mu)\Sigma^{-1}(\text{alr}(\mathbf{x}) - \mu)^T\right\},$$

donde μ y Σ son, respectivamente, el vector de medias y la matriz de covarianzas de los datos alr-transformados. Esta distribución, que denotaremos $\text{aln}(\mu, \Sigma)$, juega un papel central en el estudio de distribuciones en S^D , análogo al de la distribución normal multivariante en \mathbb{R}^D . Además resulta que el vector alr-transformado sigue una distribución normal multivariante en \mathbb{R}^{D-1} . En Mateu-Figueras (2003) puede encontrarse un estudio detallado sobre la definición y caracterización de modelos de probabilidad en el simplex. Un inconveniente de la transformación alr es su asimetría respecto a las partes de la composición. Aquella utilizada como denominador cobra un protagonismo

especial. Por ello, siempre que se trabaje con datos alr-transformados, debe comprobarse que la técnica estadística utilizada es invariante frente a permutaciones de las partes. Sin embargo, el principal inconveniente es que no es una transformación isométrica. En consecuencia, los ángulos y distancias en el simplex, con la métrica de Aitchison, no pueden asociarse con ángulos y distancias en el espacio real, con la métrica Euclídea. A pesar de esto, la consistencia de los resultados está garantizada cuando la inferencia se basa en la función de verosimilitud de la distribución aln (Aitchison, 1986). La transformación clr está más vinculada a contextos no paramétricos, ya que a partir de ella es posible especificar la distancia de Aitchison (2) en términos de la distancia Euclídea entre los vectores clr-transformados, esto es,

$$d_a(\mathbf{x}, \mathbf{x}') = d_e(\text{clr}(\mathbf{x}), \text{clr}(\mathbf{x}')).$$

La transformación clr es simétrica e isométrica, pero la imagen de S^D queda realmente restringida a un subespacio de R^D y la matriz de covarianzas de los datos clr-transformados es singular, del mismo modo que ocurría con la matriz de covarianzas aplicada directamente sobre las composiciones.

La caracterización del simplex como espacio Euclídeo sugiere a Egozcue y otros (2003) proponer la transformación *log-cociente isométrica* (ilr), que salva los principales inconvenientes de las dos anteriores. La transformación ilr de una composición \mathbf{x} se define como $\text{ilr}(\mathbf{x}) = [\langle \mathbf{x}, \mathbf{e}_1 \rangle, \dots, \langle \mathbf{x}, \mathbf{e}_{D-1} \rangle]$, donde $\mathbf{e}_1, \dots, \mathbf{e}_{D-1}$ es una base ortonormal obtenida a partir del producto interior (1). Esto es, las componentes del vector ilr-transformado son las coordenadas de la composición \mathbf{x} respecto a la base $\mathbf{e}_1, \dots, \mathbf{e}_{D-1}$. Esta transformación es isométrica y los datos ilr-transformados se representan en los habituales ejes ortogonales. En Egozcue y otros (2003) los autores obtienen una expresión explícita para la transformación ilr dada una base ortonormal particular:

$$\text{ilr}(\mathbf{x}) = \mathbf{y} = [y_1, \dots, y_{D-1}] \in R^{D-1}, \text{ donde } y_i = \frac{1}{\sqrt{i(i+1)}} \ln \left(\frac{\prod_{j=1}^i x_j}{(x_{i+1})^i} \right).$$

El principal problema aquí es determinar cuál es la base ortonormal más apropiada para un problema concreto, la que proporciona las expresiones que hacen más fácil la interpretación de los resultados, ya que las coordenadas ilr no suelen ser fáciles de

interpretar. Una posibilidad es utilizar las coordenadas y expresar los resultados en la base canónica de \mathbb{R}^D sin abandonar el simplex. Trabajar en coordenadas permite definir distribuciones de probabilidad en el simplex, estimando los parámetros con los procedimientos habituales a partir de la expresión en coordenadas de los datos originales (Pawlowsky-Glahn, 2003; Mateu-Figueras y Pawlowsky-Glahn, 2007). Notar que realmente esta estrategia para *escapar* de un espacio muestral restringido expresando los datos en coordenadas respecto a una base ortonormal es aplicable a cualquier problema definido sobre un espacio Euclídeo (véase Pawlowsky-Glahn, 2003).

3.2. Medidas de posición y variabilidad en un conjunto de datos composicionales

Como sabemos, la medida de posición central más utilizada sobre conjuntos de datos en un espacio real es la media aritmética o centroide del conjunto. Sabemos además que es un buen estimador de la media de una variable aleatoria \mathbf{x} . Desde un punto de vista geométrico, la media $\mu = E[\mathbf{x}]$ de una variable aleatoria \mathbf{x} en el espacio real minimiza la distancia Euclídea al cuadrado esperada $E[d_e(\mathbf{x}, \xi)^2]$ y la varianza queda expresada como $\text{Var}(\mathbf{x}) = E[d_e(\mathbf{x}, \mu)^2]$. Trasladando esta caracterización a la geometría del simplex, Aitchison (2001) define el centro $\text{cen}(\mathbf{x})$ de una composición aleatoria \mathbf{x} en S^D como el valor que minimiza $E[d_a(\mathbf{x}, \xi)^2]$. Dada una muestra $\mathbf{X} = (\mathbf{x}_1, \dots, \mathbf{x}_n)$ de datos composicionales, la *media geométrica composicional* definida como

$$g(\mathbf{X}) = C(g_1, \dots, g_D), \quad (5)$$

donde $g_j = (x_{1j} \cdots x_{nj})^{1/n}$, es un estimador lineal insesgado óptimo del centro $\text{cen}(\mathbf{x})$ (Pawlowsky-Glahn y Egozcue, 2002). Destacar que en este contexto la linealidad se expresa como $\alpha_1 \otimes \mathbf{x}_1 \oplus \alpha_2 \otimes \mathbf{x}_2 \oplus \dots \oplus \alpha_n \otimes \mathbf{x}_n$, donde los α_i son coeficientes reales. En la Figura 1 representamos una muestra de puntos del simplex S^3 en un *diagrama ternario*. Puede apreciarse cómo la media geométrica composicional (o), además de ser compatible con las operaciones básicas en el simplex, representa el centro de gravedad de la nube de puntos composicionales de forma más adecuada que la habitual media aritmética (*). En general, esto ocurrirá siempre que la nube de puntos en S^3 tenga una forma aproximadamente cóncava.

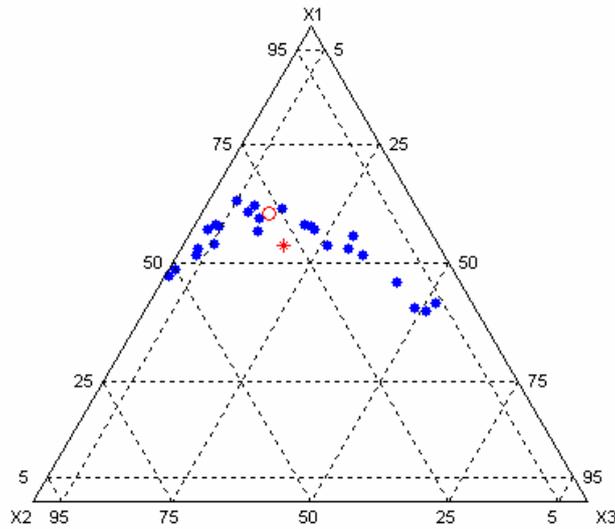


Figura 1. Diagrama ternario. Media geométrica composicional (o) y media aritmética (*) de un conjunto de datos composicionales.

Siguiendo la misma línea, la varianza entorno a $\text{cen}(\mathbf{x})$ viene dada por $E[d_a(\mathbf{x}, \text{cen}(\mathbf{x}))^2]$. Pawlowsky-Glahn y Egozcue (2002) establecen que esta varianza coincide con la medida de *variabilidad total* $\text{tot var}(\mathbf{x})$ definida en Aitchison (1997) como la suma de las *varianzas log-cociente* $\text{var}(\ln x_i / x_j)$ dividida por el número de partes. De manera que su equivalente muestral

$$\text{totvar}(\mathbf{X}) = \frac{1}{n} \sum_{i=1}^n d_a^2(\mathbf{x}_i, \mathbf{g}(\mathbf{X})) \quad (6)$$

es una medida de variabilidad adecuada y coherente con la métrica del símplex.

Las medidas (5) y (6) están relacionadas con las medidas muestrales habituales aplicadas sobre los datos log-cociente transformados. Así, se tiene que:

1. $\text{alr}(\mathbf{g}(\mathbf{X})) = \overline{\text{alr}(\mathbf{X})}$, $\text{clr}(\mathbf{g}(\mathbf{X})) = \overline{\text{clr}(\mathbf{X})}$ y $\text{ilr}(\mathbf{g}(\mathbf{X})) = \overline{\text{ilr}(\mathbf{X})}$, donde $\overline{\text{alr}(\mathbf{X})}$, $\overline{\text{clr}(\mathbf{X})}$ y $\overline{\text{ilr}(\mathbf{X})}$ representan las medias aritméticas de los respectivos conjuntos de datos transformados.

2. $\text{totvar}(\mathbf{X}) = \text{tr}(\text{cov}(\text{clr}(\mathbf{X}))) = \text{tr}(\text{cov}(\text{ilr}(\mathbf{X})))$. Destacar que tal relación no puede establecerse con la transformación alr debido a que ésta no es isométrica.

El patrón de variabilidad de una composición \mathbf{x} en S^D queda completamente determinada por la *matriz de variación* $\mathbf{T} = [\tau_{ij}]$ formada por las varianzas log-cociente $\tau_{ij} = \text{var}[\ln(x_i / x_j)]$, con $i = 1, \dots, D-1$ y $j = i+1, \dots, D$. Esta matriz es simétrica y tiene ceros en la diagonal principal. Aunque no puede expresarse como la matriz de covarianzas estándar de un vector, sí que está relacionada con las matrices de covarianzas de los vectores alr-, clr-, y ilr-transformados mediante simples operaciones matriciales (Aitchison, 1986; Egozcue y otros, 2003). Además, tiene la ventaja de que la matriz de variación de cualquier subcomposición se obtiene simplemente extrayendo las varianzas log-cociente asociadas a las partes que forma la subcomposición. En consonancia con la metodología log-cociente, las varianzas log-cociente miden la variabilidad relativa de una componente x_i respecto a otra x_j . Así, $\text{var}[\ln(x_i / x_j)] = 0$ implica que el cociente x_i / x_j es constante, por lo que existe una relación de proporcionalidad perfecta entre x_i y x_j , mientras que $\text{var}[\ln(x_i / x_j)] = \infty$ implica que una relación de proporcionalidad inexistente. Esta relación de proporcionalidad sustituye a la habitual correlación entre variables. Es sencillo obtener una medida que esté acotada en el intervalo $[0,1]$ considerando $v_{ij} = \exp(-\sqrt{\tau_{ij}})$. De esta manera, un valor $v_{ij} = 0$ indica ausencia de relación de proporcionalidad y un valor $v_{ij} = 1$ indica relación de proporcionalidad perfecta. Destacar que si el interés es la independencia, ésta debe expresarse en términos de independencia de subcomposiciones. Por ejemplo, dada una composición de 5 partes, la independencia entre las subcomposiciones $[s_1, s_2, s_3]$ y $[s_4, s_5]$ se expresa como $\text{cov}[\ln(x_1 / x_3); \ln(x_4 / x_5)] = 0$ y $\text{cov}[\ln(x_2 / x_3); \ln(x_4 / x_5)] = 0$. Estas medidas son simplemente estimadas a partir de sus análogas muestrales aplicadas sobre los correspondientes log-cocientes observados.

3.3 El problema de los ceros y su tratamiento

En los puntos anteriores hemos puesto de manifiesto que el correcto análisis estadístico de vectores de proporciones gira entorno a los cocientes entre sus partes. Sin embargo, si un vector presenta alguna componente nula resulta imposible considerar todos los cocientes de la forma x_i / x_j . De manera que ni las transformaciones log-cociente, ni la distancia de Aitchison, ni las medidas descriptivas son aplicables en este caso. En la

literatura sobre datos composicionales se han distinguido dos tipos de ceros: los *ceros esenciales* y los *ceros por redondeo*. Los ceros esenciales se corresponden con componentes que realmente toman un valor nulo. A veces surgen por una excesiva desagregación o bien como un indicador de la existencia de diferentes submuestras dentro del conjunto de datos. Así, en la distribución de un presupuesto familiar en distintas partidas de gasto pueden distinguirse las familias que consumen tabaco o bebidas alcohólicas de las que no lo hacen. La fusión, *amalgamamiento*, de algunas de las partes o el análisis de submuestras de forma separada puede salvar el problema en algunos casos. En Aitchison y Kay (2003) y Bacon-shone (2003) pueden encontrarse algunas propuestas para tratar este tipo de ceros.

Un problema distinto es el de los ceros por redondeo, ceros que se corresponden con valores que no han podido observarse porque limitaciones en los instrumentos de medida o en el procedimiento de recogida y tratamiento de datos, o incluso políticas o culturales, impiden que se registren cuantías pequeñas que no superan cierto umbral de detección. El dato composicional con ceros por redondeo es un tipo especial de dato incompleto y, por lo tanto, es susceptible de ser tratado mediante técnicas de imputación o reemplazamiento. El procedimiento elegido debe preservar la estructura de covarianzas y ser coherente con las propiedades métricas específicas de los datos composicionales. En Martín-Fernández, Barceló-Vidal y Pawlowsky-Glahn (2003) se propone un procedimiento no paramétrico de reemplazamiento que satisface estas condiciones. Consiste en imputar de forma multiplicativa los valores nulos con un valor pequeño prefijado proporcional al umbral de detección. Con este procedimiento se obtienen buenos resultados cuando el número de ceros es reducido o en conjuntos de datos pequeños donde no es posible, o es arriesgado, establecer hipótesis distribucionales sobre los datos. El inconveniente lo encontramos en la necesidad medir la sensibilidad de los resultados al valor elegido para imputar, además de que tiende a subestimar la variabilidad. Como alternativa paramétrica cuando se dispone de conjuntos de datos relativamente grandes, Palarea-Albaladejo, Martín-Fernández y Gómez-García (2007) proponen una versión del algoritmo EM basada en la distribución de probabilidad normal logística aditiva. Este procedimiento imputa los ceros teniendo en cuenta la información suministrada por el resto de componentes y corrige sustancialmente la tendencia a la subestimación de la variabilidad del reemplazamiento

multiplicativo. Para agilizar la exposición de los ejemplos de la sección 4 asumiremos que los datos no contienen ceros o que éstos han sido previamente tratados con alguno de los métodos anteriores.

4. Algunos ejemplos seleccionados en economía

Hasta la fecha, el volumen de trabajos que hacen uso de la metodología log-cociente en el campo de la economía y las ciencias sociales es relativamente pequeño. Destacamos los trabajos de Fry, Fry y McLaren (2000), donde se aplica a la especificación de un modelo microeconómico para el análisis de la distribución del presupuesto familiar; Katz y King (1999), donde se propone un modelo para estudiar cómo la distribución geográfica de los resultados electorales en un sistema multipartito dependen de una serie de variables socioeconómicas; Larrosa (2003), que estudia los patrones de la composición del stock de capital en distintos países; Anyadike-Danes (2003), que analiza la composición de la población no trabajadora en el Reino Unido y su relación con el nivel de desempleo. En las subsecciones siguientes ilustraremos la aplicación práctica del análisis log-cociente a través de unos ejemplos del ámbito económico.

4.1. Distribución de trabajadores inmigrantes por sectores de actividad.

En los últimos años, el dinamismo de nuestra economía ha provocado una importante entrada de mano de obra inmigrante para cubrir las necesidades del mercado laboral español. Es apreciable que el peso de las distintas nacionalidades no ha sido el mismo, destacando las americanas y norteafricanas, en concreto Ecuador y Marruecos, aunque se detecta una importancia creciente de otros grupos como es el caso de los ciudadanos europeos fuera del ámbito de la UE (informe 2/2004 del CES). Cabe preguntarse cómo varía la demanda de trabajadores por sectores de actividad según sea la nacionalidad de los mismos. Por ejemplo, contrastaremos si hay diferencias significativas entre la distribución por sectores de actividad de los trabajadores procedentes de países latinoamericanos y africanos (grupo 1, 27 países) y la de los trabajadores de otras nacionalidades (grupo 2, 31 países). Utilizaremos datos elaborados a partir de la información suministrada por el Boletín Estadístico de Extranjería e Inmigración, nº 6, de Julio de 2005, editado por el Ministerio de Trabajo y Asuntos Sociales.

Se consideran cuatro sectores: agricultura, industria, construcción y servicios. La Figura 2a muestra los datos de cada país en S^4 en un *diagrama cuaternario* distinguiendo los

continentes de origen. Se observa que el sector servicios es el que más trabajadores extranjeros atrae, especialmente los procedentes del continente americano. El sector industria es el que menos trabajadores inmigrantes contrata. Si nos quedamos con la subcomposición formada por los tres sectores principales, los datos se proyectan sobre el símplex S^3 . Esto nos permite visualizar más claramente la distribución sectorial de los trabajadores de los distintos países en la Figura 2b.

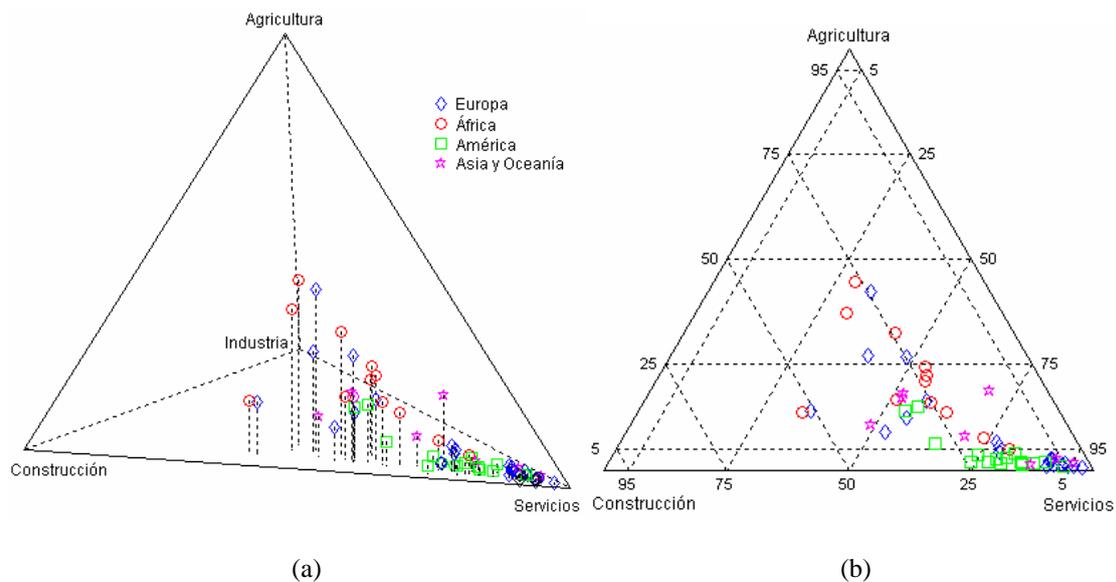


Figura 2. Representación gráfica de la distribución porcentual de los trabajadores extranjeros en los distintos sectores de actividad distinguiendo continente de procedencia.

En la Tabla 1 tenemos un resumen descriptivo de los datos (véase subsección 3.2.). Observamos que la gran mayoría del trabajo inmigrante se concentra en el sector servicios, tanto de forma global como cuando se distingue entre inmigrantes latinoamericanos y africanos y el resto.

		[agricul, industr, constru, servici]
Global	$g(\mathbf{X})$	[5.3027, 6.6366, 16.3326, 71.7281]
	$\text{totvar}(\mathbf{X})$	1.5072
Países Latinoamericanos y africanos (grupo 1)	$g(\mathbf{X}_1)$	[6.4736, 6.9263, 21.1360, 65.4641]
	$\text{totvar}(\mathbf{X}_1)$	1.3548
Otras nacionalidades	$g(\mathbf{X}_2)$	[4.3880, 6.2955, 12.8461, 76.4704]

(grupo 2)	totvar(\mathbf{X}_2)	1.5637
-----------	--------------------------	--------

Tabla 1. Medidas descriptivas de la distribución de trabajadores inmigrantes según sector de actividad, a nivel global y distinguiendo nacionalidad de origen.

Asumiendo, respectivamente, modelos $\text{aln}(\mu_1, \Sigma)$ y $\text{aln}(\mu_2, \Sigma)$ para los datos de los grupos 1 y 2, contrastamos la hipótesis $\mu_1 = \mu_2$ mediante el estadístico Traza de Hotelling para muestras independientes aplicado sobre las muestras alr-transformadas. Se obtiene un valor del estadístico de 0.193 con un p-valor asociado igual a 0.022. Destacar que estos valores son independientes de la componente utilizada como denominar en la transformación alr. Al habitual 95% de confianza rechazaríamos la hipótesis de igualdad de medias, esto es, existe una diferencia significativa en la distribución sectorial de los dos grupos de trabajadores inmigrantes considerados. Un buen indicador de la naturaleza de esta diferencia puede obtenerse como $g(\mathbf{X}_1) \ominus g(\mathbf{X}_2) = [29.0591, 21.6707, 32.4081, 16.8621]$. Este vector diferencia composicional muestra que el grupo 1 se caracteriza frente al grupo 2 por un mayor peso relativo en los sectores construcción y agricultura, respecto a los sectores industria y servicios.

4.2. Empleo del tiempo según la edad.

En las últimas décadas, con el diseño e implementación de políticas de bienestar en los países desarrollados, hay un creciente interés por el estudio del uso que los ciudadanos hacen de su tiempo. Esta información sirve de apoyo a la formulación de políticas familiares, de igualdad de género, etc.

En este ejemplo utilizaremos datos elaborados a partir de la Encuesta de Empleo del Tiempo realizada por el INE en 2004 con el fin de construir un modelo de regresión que explique la relación entre el uso del tiempo y la edad de los individuos. Para simplificar la exposición, se agrupan las actividades en 3 grupos: 1. estudios o trabajo; 2. cuidados personales, hogar y familia; 3. ocio y vida social. Hemos considerado una submuestra aleatoria de 442 individuos sin proporciones nulas en los grupos de actividades considerados. Destacar que la distribución del tiempo es una variable composicional en origen. El total, 24 horas, es el mismo para todos. De hecho, podríamos trabajar directamente en unidades de tiempo y no con proporciones.

El modelo a estimar puede expresarse en términos de los datos alr-transformados como

$$\begin{cases} \log(\text{act}_1 / \text{act}_3) = \beta_{01} + \beta_{11} \log(\text{edad}) + \varepsilon_1 \\ \log(\text{act}_2 / \text{act}_3) = \beta_{02} + \beta_{12} \log(\text{edad}) + \varepsilon_2, \end{cases} \quad (7)$$

donde act_1 , act_2 y act_3 denotan, respectivamente, las proporciones de tiempo dedicadas a las actividades de los grupos 1, 2 y 3. Destacar que este modelo de regresión es invariante frente a permutaciones de las partes de la composición $\mathbf{act} = [\text{act}_1, \text{act}_2, \text{act}_3]$. Esto es, si se toma como divisor de la transformación alr una parte distinta a act_3 se obtienen resultados compatibles. A continuación, se estiman los parámetros del modelo mediante el habitual procedimiento de mínimos cuadrados ordinarios y se llevan al simplex utilizando la inversa de la transformación alr (3). Tenemos entonces la *recta* de regresión composicional dada por

$$[\text{act}_1, \text{act}_2, \text{act}_3] = [0.1129, 0.6036, 0.2836] \oplus [0.4354, 0.3270, 0.2376] \otimes \log(\text{edad}).$$

Para facilitar la interpretación de los resultados, calculamos las distribuciones de tiempo estimadas por el modelo y las ordenamos de menor a mayor edad. Denotemos mediante $\hat{\mathbf{g}}_{\max}$ y $\hat{\mathbf{g}}_{\min}$ las estimaciones de las distribuciones de tiempo asociadas, respectivamente, a la mayor y menor edad de la muestra. Si calculamos la diferencia composicional $\hat{\mathbf{g}}_{\max} \ominus \hat{\mathbf{g}}_{\min}$ se obtiene $[0.5484, 0.2992, 0.1524]$, mientras que $\hat{\mathbf{g}}_{\min} \ominus \hat{\mathbf{g}}_{\max}$ es igual a $[0.1555, 0.2849, 0.5596]$. Estos resultados ponen claramente de manifiesto el cambio en la distribución del tiempo según la edad. Conforme se incrementa ésta aumenta el tiempo dedicado a trabajo o estudios y disminuye el dedicado a ocio y vida social. Podemos decir que el intercambio de tiempo se produce básicamente entre estos dos grupos. El valor de la media geométrica composicional, $[0.2996, 0.6046, 0.0957]$, revela que el grupo cuidados personales, hogar y familia es al que más tiempo se dedica. Puede comprobarse también que esto es así independientemente de la edad. Destaca que la relación de este grupo de actividad con el segundo y tercer grupo en importancia, que como hemos visto cambian según la edad, se mantiene casi constante en términos relativos. Así, $0.5484/0.2992 \cong 0.5596/0.2849$ y $0.1524/0.2992 \cong 0.1555/0.2849$.

En la Figura 3 representamos los datos observados y la *recta* de regresión composicional en el simplex S^3 correspondiente a la composición \mathbf{act} . Las flechas indican el sentido del desplazamiento conforme se incrementa la edad.

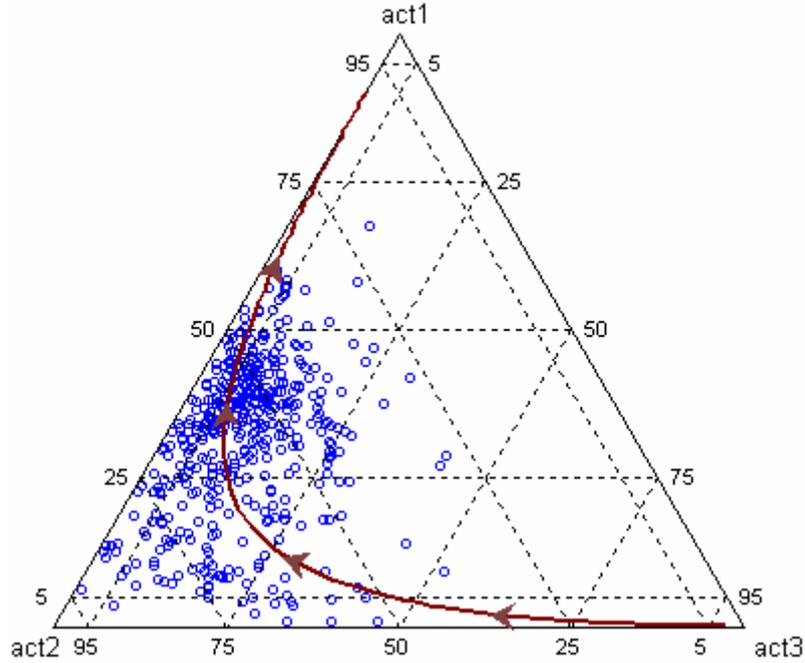


Figura 3. Distribución del tiempo diario entre tres grupos de actividades: valores observados y recta de regresión composicional estimada en función de la edad.

Finalmente, destacaremos que también es posible realizar un análisis semejante en el plano alr-transformado. Así, las rectas de regresión estimadas a partir de (7) pueden interpretarse como las ecuaciones paramétricas de la recta que relaciona $\log(\text{act}_2 / \text{act}_3)$ y $\log(\text{act}_1 / \text{act}_3)$, siendo $\log(\text{edad})$ la variable paramétrica. La ecuación de dicha recta vendrá dada entonces por

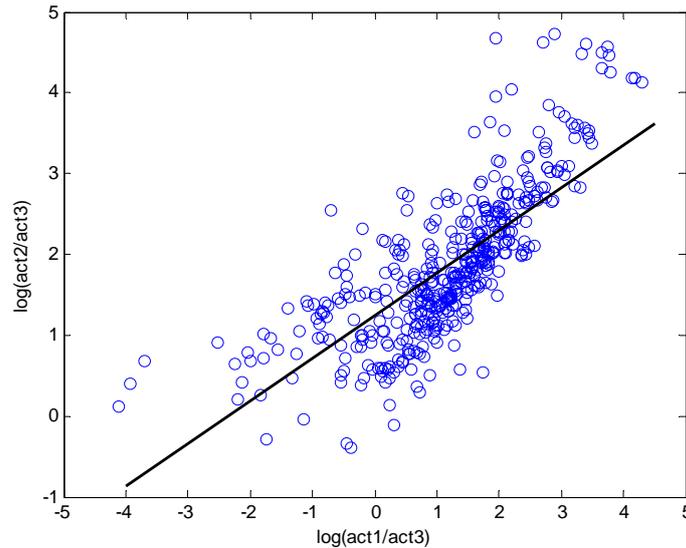
$$\log(\text{act}_2 / \text{act}_3) = \frac{\hat{\beta}_{02}\hat{\beta}_{11} - \hat{\beta}_{12}\hat{\beta}_{01}}{\hat{\beta}_{11}} + \frac{\hat{\beta}_{12}}{\hat{\beta}_{11}} \log(\text{act}_1 / \text{act}_3),$$

que en este caso resulta ser igual a $\log(\text{act}_2 / \text{act}_3) = 1.2412 + 0.5274 \log(\text{act}_1 / \text{act}_3)$. En la Figura 4 representamos los puntos en el plano alr junto a la recta estimada. La pendiente de la recta es aproximadamente 1/2, por lo que un incremento de 2 unidades en $\log(\text{act}_1 / \text{act}_3)$ se traduce aproximadamente en un incremento de 1 unidad en $\log(\text{act}_2 / \text{act}_3)$. Realizando algunas operaciones llegamos a la expresión equivalente:

$$\text{act}_1 / \text{act}_3 \propto (\text{act}_2 / \text{act}_3)^{1.8961}. \quad (8)$$

De esta manera se observa de forma más evidente una relación cuadrática entre los incrementos en $\text{act}_1 / \text{act}_3$ y los incrementos en $\text{act}_2 / \text{act}_3$. Sabemos que la proporción

act_2 se mantiene aproximadamente constante con la edad, por lo que el incremento *cuadrático* de act_1/act_3 en relación al incremento de act_2/act_3 se explica necesariamente por el efecto conjunto de la disminución en act_3 junto al incremento en act_1 . De modo que llegamos a la misma conclusión que cuando realizamos el análisis sobre el símplex.



5. Comentarios finales

Al iniciar un proceso de modelización y análisis de datos es necesario prestar especial atención a las características del espacio muestral sobre el que están definidos. Esto es imprescindible para la elaboración de modelos estadísticos apropiados. Cuando se trabaja con datos numéricos, habitualmente, se asume que se trata de datos reales. Sin embargo, los datos que se refieren a partes de un total, entre otros, no se corresponden con valores continuos sobre todo el espacio real, sino que se definen sobre un subconjunto de este último sujeto a las restricciones de no negatividad y de suma constante. Sin embargo, la mayoría de las veces se analizan recurriendo a las técnicas multivariantes estándar para datos reales no restringidos, quizás por su apariencia de vectores de números reales. Aunque en algunos casos prácticos puede que se llegue a conclusiones similares, lo importante es que, de base, no se está actuando correctamente. Las medidas de posición, variabilidad, similaridad, etc. utilizadas en el espacio real no son coherentes con la estructura geométrico-algebraica del símplex, por

lo que nada garantiza que las inferencias tengan algún significado y se correspondan en cierta medida con la realidad subyacente.

A lo largo de este trabajo se han revisado los fundamentos que determinan el marco formal dentro del cual desarrollar técnicas y medidas coherentes con las características del simplex como espacio muestral. Reconociendo que los datos contienen sólo información sobre las magnitudes relativas, la metodología log-cociente permite trasladar el problema al más familiar espacio real, y aplicar entonces las herramientas de análisis usuales. La implementación de esta estrategia no está exenta de problemas prácticos. En concreto, han sido comentados los problemas derivados de la existencia de observaciones con partes nulas y algunas de las soluciones propuestas. Como se ha destacado, la caracterización del simplex como espacio Euclídeo abre la posibilidad de circunscribir la modelización estadística al simplex, sin necesidad de recurrir a transformaciones de los datos, redefiniendo los conceptos fundamentales de forma adecuada. Esta idea marca una de las principales líneas de investigación actuales en esta área.

Finalmente, para ilustrar la aplicación práctica del análisis log-cociente, se han considerado dos ejemplos en el campo económico empleando técnicas de uso habitual entre los investigadores. Estos ejemplos ponen de manifiesto las diferencias en el modo de actuar y de interpretar los resultados cuando se trabaja con composiciones.

Bibliografía

Aitchison, J. (1986): *The statistical analysis of compositional data*, Chapman & Hall, London. Reimpresión por Blackburn Press en 2003.

Aitchison, J. (1997): "The one-hour course in compositional data analysis or compositional data analysis is easy", en Pawlowsky-Glahn, V. (ed.), *Proceedings of the Third Annual Conference of the IAMG*, Barcelona, pp. 3-35.

Aitchison, J., Shen, S. M. (1980): "Logistic normal distributions: some properties and uses", *Biometrika*, 67, pp. 261-272.

Aitchison, J., Barceló-Vidal, C., Martín-Fernández, J. A., Pawlowsky-Glahn, V. (2000): "Logratio analysis and compositional distance", *Mathematical Geology*, 32, pp. 271-275.

Aitchison, J., Kay, J. W. (2003): "Possible solutions of some essential zero problems in compositional data analysis", en Thió-Henestrosa, S., Martín-Fernández, J. A. (eds.), *First Compositional Data Analysis Workshop - CoDaWork'03*. Universitat de Girona.

- Anyadike-Danes, M. (2003): "The allometry of- non-employment", en Thió-Henestrosa, S., Martín-Fernández, J. A. (eds.), *First Compositional Data Analysis Workshop - CoDaWork'03*. Universitat de Girona.
- Bacon-Shone, J. (2003): "Modelling structural zeros in compositional data", en Thió-Henestrosa, S., Martín-Fernández, J. A. (eds.), *First Compositional Data Analysis Workshop - CoDaWork'03*. Universitat de Girona.
- Billheimer, D., Guttorp, P., Fagan, W. F. (2001): "Statistical interpretation of species composition", *J. Am. Stat. Assoc.*, 96, pp. 1205-1214.
- Bohling, G. C., Davis, J. C., Olea, R. A. Harff, J. (1998): "Singularity and nonnormality in the classification of compositional data", *Mathematical Geology*, 30, pp. 5-20.
- Egozcue, J. J., Pawlowsky-Glahn, V., Mateu-Figueras, G., Barceló-Vidal, C. (2003): "Isometric logratio transformation for compositional data analysis", *Mathematical Geology*, 35, pp. 279-300.
- Fry, J. M., Fry, T. R. L., McLaren, K. R. (2000): "Compositional data analysis and zeros in micro data", *Applied Economics*, 32, pp. 953-959.
- Katz, J. N., King, G. (1999): "A statistical model for multiparty electoral data", *American Political Science Review*, 93, pp. 15-32.
- Larrosa, J. M. (2003): "A compositional statistical analysis of capital stock", en Thió-Henestrosa, S., Martín-Fernández, J. A. (eds.), *First Compositional Data Analysis Workshop - CoDaWork'03*. Universitat de Girona.
- Martín-Fernández, J. A., Barceló-Vidal, C., Pawlowsky-Glahn, V. (2003): "Dealing with zeros and missing values in compositional data sets", *Mathematical Geology*, 35, pp. 253-278.
- Mateu-Figueras, G. (2003): *Modelos de Distribución sobre el Simplex*, tesis doctoral, Universitat Politècnica Catalunya.
- Mateu-Figueras, G. Pawlowsky-Glahn, V. (2007): "The skew-normal distribution on SD", *Comm. Statist. Theory Methods*, 36, (en prensa).
- Palarea-Albaladejo, J., Martín-Fernández, J. A., Gómez-García, J. (2007): "A parametric approach for dealing with compositional rounded zeros", *Mathematical Geology*, (en prensa).
- Pawlowsky-Glahn, V. (2003): "Statistical modelling on coordinates", en Thió-Henestrosa, S., J. A. Martín-Fernández (eds.), *First Compositional Data Analysis Workshop - CoDaWork'03*. Universitat de Girona.
- Pawlowsky-Glahn, V., Egozcue, J. J. (2001): "Geometric approach to statistical analysis on the simplex", *Stoch. Environ. Res. Risk Assess.*, 15, pp. 384-398.
- Pawlowsky-Glahn, V., Egozcue, J. J. (2002): "BLU estimators and compositional data", *Mathematical Geology*, 34, pp. 259-274.
- Pearson, K. (1897): "Mathematical contributions to the theory of evolution. On a form of spurious correlation which may arise when indices are used in the measurements of organs", *Proc. R. Soc.*, 60, pp. 489-498.

LA DISTRIBUCIÓN TRONCOPIRAMIDAL Y LOS SISTEMAS DE DISTRIBUCIONES TIPO PEARSON

RAFAEL HERRERÍAS PLEGUEZUELO

e-mail: rherrer@ugr.es

JOSÉ MANUEL HERRERÍAS VELASCO

e-mail: jmherrer@ugr.es

Departamento de Métodos Cuantitativos para la Economía y la Empresa
UNIVERSIDAD DE GRANADA

Área temática: Métodos Cuantitativos .

Resumen

En este trabajo se realiza, en primer lugar, un estudio probabilístico completo de la distribución troncopiramidal como extensión de la distribución trapezoidal univariante, esta última se utiliza como modelo probabilístico en el método PERT, el Análisis de inversiones y el método de valoración de las dos funciones de distribución.

En segundo lugar, se obtienen las distribuciones marginales de la distribución troncopiramidal, comprobándose, que a su vez son análogas a las distribuciones del sistema de Pearson continuo univariante.

Finalmente se simplifican las expresiones de la distribución troncopiramidal estandarizando los recorridos de las variables originales.

Palabras clave: distribución troncopiramidal, distribución trapezoidal, sistema de Pearson.

Abstract

This paper describes, firstly, a complete probabilistic study of pyramidal trunk distribution as an extension of univariate trapezoidal distribution, which is used as a probabilistic model in the PERT model, in investment analysis and in the evaluation method of the two distribution functions.

Secondly, we obtain the marginal distributions of pyramidal trunk distribution, which are analogous to the distributions of the Pearson univariate continuous system.

Finally, we simplify the expressions of the pyramidal trunk distribution, standardising the ranges of the original variables.

Key words: Pyramidal trunk Distribution, Trapezoidal Distribution, Pearson system

1. Introducción

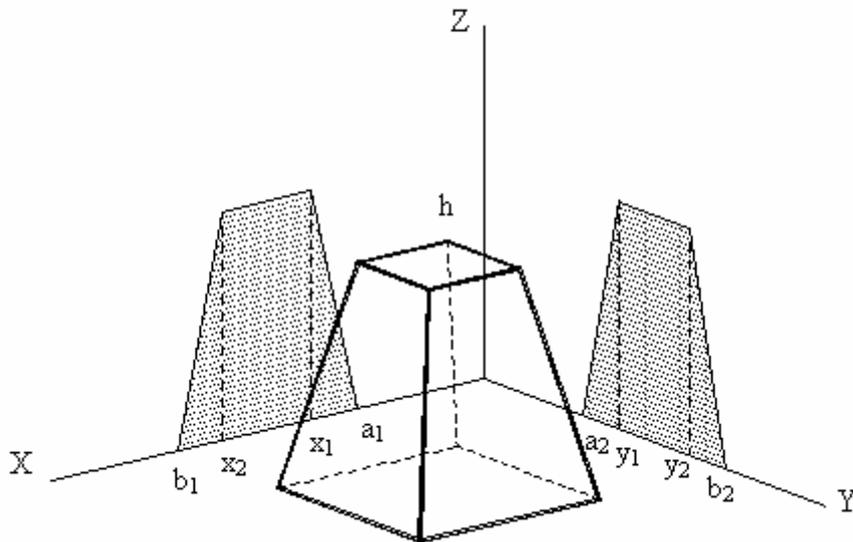
La distribución troncopiramidal surge, de manera natural, como una extensión al dominio bidimensional de la distribución trapezoidal univariante (especificada mediante los valores mínimo, a , máximo, b , e intervalo modal, m_1 y m_2 , del recorrido de la variable) con el objetivo final de utilizarla como modelo probabilístico en los campos donde esta última ha sido usada con éxito, a saber, en la Teoría General de Valoración, en particular, en el Método de Valoración de las Dos Funciones de Distribución (MVDFD), introducido por Ballesteros, E.(1973) como método de las dos betas, en el que se sugiere la posibilidad de usar la distribución beta como modelo para la valoración sintética de tierras, con posterioridad fue ampliado por Romero, C.(1977) a las distribuciones rectangulares y triangulares, Lozano, J.J. (1996) utiliza una clase especial de distribuciones trapezoidales y Herrerías, García, Cruz y Herrerías, (2001) extienden el método al uso de distribuciones trapezoidales de cualquier tipo.

En este trabajo se obtienen las distribuciones marginales de la distribución troncopiramidal, comprobándose que satisfacen la ecuación diferencial lineal, de primer orden homogénea, clásica del sistema de Pearson de distribuciones continuas univariantes, Elderton y Johnson (1969), por lo que las distribuciones marginales de la distribución troncopiramidal son distribuciones de Pearson.

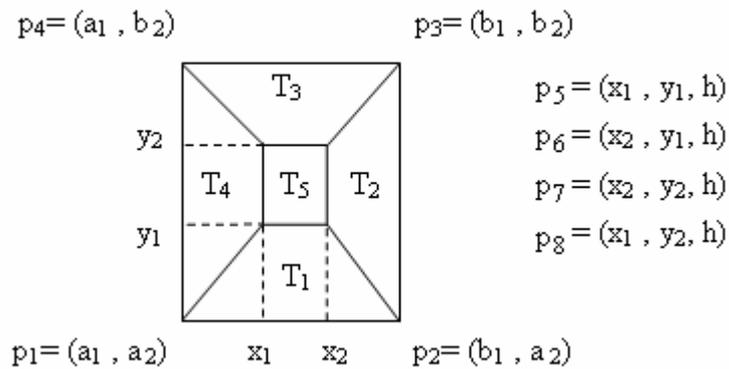
La inclusión de las distribuciones marginales en el sistema de Pearson, trae como consecuencia el conocimiento de una gran parte de las propiedades de la distribución, a pesar de ello se han obtenido las características estocásticas más importantes de la distribución troncopiramidal.

2. Función de densidad de la distribución troncopiramidal

Las caras del tronco de pirámide pueden determinarse fácilmente por planos determinados por tres puntos: dos en la base y el tercero correspondiente a uno de los dos vértices del tronco de pirámide, donde la cota de estos puntos o altura del tronco de pirámide, servirá de constante normalizadora para la distribución bivariante. Su representación gráfica es:



Proyectando el tronco de pirámide en el plano $Z = 0$. Se denotan por T_i ($i = 1,2,3,4,5$) las diferentes regiones que conforman los recorridos de (X, Y) .



Se determinan los planos que conforman las cinco caras de la troncopiramidal, a partir de la ecuación del plano que pasa por tres puntos:

- $$\Pi(p_1, p_5, p_2) \equiv \begin{vmatrix} x & y & z & 1 \\ a_1 & a_2 & 0 & 1 \\ x_1 & y_1 & h & 1 \\ b_1 & a_2 & 0 & 1 \end{vmatrix} = 0 \quad \Rightarrow$$

$$\Pi(p_1, p_5, p_2) \equiv -h(a_1 - b_1)y + [a_1(y_1 - a_2) + b_1(a_2 - y_1)]z + a_2(a_1 - b_1)h = 0$$

Dividiendo por $(a_1 - b_1)$:

$$-hy + (y_1 - a_2)z + a_2h = 0 \Rightarrow$$

$$\Rightarrow (a_2 - y)h + (y_1 - a_2)z = 0 \Rightarrow z = \frac{y - a_2}{y_1 - a_2} h \quad \text{si } (x, y) \in T_1 \quad (1)$$

- $\Pi(p_2, p_6, p_3) \equiv h(a_2 - b_2)x + [b_1(a_2 - b_2) + x_2(b_2 - a_2)]z + b_1(b_2 - a_2)h = 0$

Dividiendo por $(a_2 - b_2)$:
$$z = \frac{b_1 - x}{b_1 - x_2} h \quad \text{si } (x, y) \in T_2 \quad (2)$$

- $\Pi(p_3, p_7, p_4) \equiv -h(b_1 - a_1)y + [b_1(y_2 - b_2) + a_1(b_2 - y_2)]z + b_2(b_1 - a_1)h = 0$

Dividiendo por $(b_1 - a_1)$:
$$z = \frac{b_2 - y}{b_2 - y_2} h \quad \text{si } (x, y) \in T_3 \quad (3)$$

- $\Pi(p_1, p_8, p_4) \equiv h(a_2 - b_2)x + [a_1(a_2 - b_2) + x_1(b_2 - a_2)]z + a_1(b_2 - a_2)h = 0$

Dividiendo por $(a_2 - b_2)$:
$$z = \frac{x - a_1}{x_1 - a_1} h \quad \text{si } (x, y) \in T_4 \quad (4)$$

- $\Pi(p_5, p_6, p_7) \equiv z(x_1y_2 + x_2y_1 - x_2y_2 - x_1y_1) - h(x_1y_2 + x_2y_1 - x_2y_2 - x_1y_1) = 0$

Dividiendo por $(a_2 - b_2)$:
$$z = h \quad \text{si } (x, y) \in T_5 \quad (5)$$

A partir de la ecuación de una recta que pasa por dos puntos, se obtienen las proyecciones de las aristas de la troncopiramidal en el plano $Z = 0$:

$$r_{12} \text{ es } \frac{x - b_1}{x_2 - b_1} = \frac{y - a_2}{y_1 - a_2} \Rightarrow \begin{cases} x_{12} = b_1 + \frac{x_2 - b_1}{y_1 - a_2} (y - a_2) \\ y_{12} = a_2 + \frac{y_1 - a_2}{x_2 - b_1} (x - b_1) \end{cases} \quad (6)$$

$$r_{23} \text{ es } \frac{x - b_1}{x_2 - b_1} = \frac{y - b_2}{y_2 - b_2} \Rightarrow \begin{cases} x_{23} = b_1 + \frac{x_2 - b_1}{y_2 - b_2}(y - b_2) \\ y_{23} = b_2 + \frac{y_2 - b_2}{x_2 - b_1}(x - b_1) \end{cases} \quad (7)$$

$$r_{34} \text{ es } \frac{x - a_1}{x_1 - a_1} = \frac{y - b_2}{y_2 - b_2} \Rightarrow \begin{cases} x_{34} = a_1 + \frac{x_1 - a_1}{y_2 - b_2}(y - b_2) \\ y_{34} = b_2 + \frac{y_2 - b_2}{x_1 - a_1}(x - a_1) \end{cases} \quad (8)$$

$$r_{41} \text{ es } \frac{x - a_1}{x_1 - a_1} = \frac{y - a_2}{y_1 - a_2} \Rightarrow \begin{cases} x_{41} = a_1 + \frac{x_1 - a_1}{y_1 - a_2}(y - a_2) \\ y_{41} = a_2 + \frac{y_1 - a_2}{x_1 - a_1}(x - a_1) \end{cases} \quad (9)$$

Donde se ha denotado por r_{ij} ($i, j = 1,2,3,4$) la recta frontera de las regiones T_i y T_j . Análogamente se denota por x_{ij} e y_{ij} las correspondientes abscisas u ordenadas de los puntos de la recta r_{ij} . Finalmente, se denotan por $x_{ij}(y_0)$ e $y_{ij}(x_0)$ las abscisas u ordenadas de los puntos de la recta r_{ij} particularizada al valor y_0 ó x_0 respectivamente. Esta notación facilitará el cálculo integral que se precisa en el estudio de la distribución de probabilidad.

De (1), (2), (3), (4) y (5) se obtiene la función de densidad troncopiramidal, especificando el valor de h .

Para determinar h o constante normalizadora se recurre a que la integral de la función de densidad debe cumplir:

$$\int_{a_1}^{b_1} \int_{a_2}^{b_2} z(x, y) dx dy = 1$$

Antes de calcular la integral, se comprueba que:

$$z(x, y) \geq 0 \quad \forall (x, y) \in T_i \quad (i = 1,2,3,4,5) \Leftrightarrow \begin{cases} a_1 < x < b_1 \\ a_2 < y < b_2 \end{cases}$$

para poder asegurar que $z(x, y)$ es una verdadera función de densidad.

Se denota por z_i ($i = 1, 2, 3, 4, 5$) la función z de las variables X e Y en las regiones T_i ($i = 1, 2, 3, 4, 5$), estas son respectivamente (1), (2), (3), (4) y (5).

Evidentemente:

$$\begin{aligned} \int_{a_1}^{b_1} \int_{a_2}^{b_2} z(x, y) dx dy &= \iint_{T_1} z(x, y) dx dy + \iint_{T_2} z(x, y) dx dy + \iint_{T_3} z(x, y) dx dy + \\ &+ \iint_{T_4} z(x, y) dx dy + \iint_{T_5} z(x, y) dx dy = \int_{a_2}^{y_1} \int_{x_{41}}^{x_{12}} z_1 dx dy + \\ &+ \int_{x_2}^{b_1} \int_{y_{12}}^{y_{23}} z_2 dy dx + \int_{y_2}^{b_2} \int_{x_{34}}^{x_{23}} z_3 dx dy + \int_{a_1}^{x_1} \int_{y_{41}}^{y_{34}} z_4 dy dx + \int_{x_1}^{x_2} \int_{y_1}^{y_2} z_5 dy dx = \\ &= \frac{h}{6} ((b_2 - a_2 + y_2 - y_1)(b_1 - a_1 + x_2 - x_1) + (b_1 - a_1)(b_2 - a_2) + (x_2 - x_1)(y_2 - y_1)) \end{aligned}$$

Igualando a 1 se tiene:

$$h = \frac{6}{(b_2 - a_2 + y_2 - y_1)(b_1 - a_1 + x_2 - x_1) + (b_1 - a_1)(b_2 - a_2) + (x_2 - x_1)(y_2 - y_1)} \quad (10)$$

La expresión de la función de densidad de la distribución troncopiramidal es la siguiente:

$$z(x, y) = \begin{cases} \frac{y - a_2}{y_1 - a_2} h & \text{si } (x, y) \in T_1 \\ \frac{b_1 - x}{b_1 - x_2} h & \text{si } (x, y) \in T_2 \\ \frac{b_2 - y}{b_2 - y_2} h & \text{si } (x, y) \in T_3 \\ \frac{x - a_1}{x_1 - a_1} h & \text{si } (x, y) \in T_4 \\ h & \text{si } (x, y) \in T_5 \\ 0 & \text{en otro caso} \end{cases} \quad (11)$$

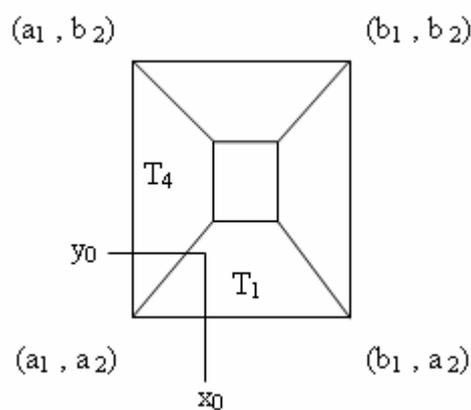
O lo que es lo mismo:

$$z(x, y) = \begin{cases} \frac{y - a_2}{y_1 - a_2} h & \text{si } x_{41} \leq x \leq x_{12} \wedge a_2 < y < y_1 \\ \frac{b_1 - x}{b_1 - x_2} h & \text{si } x_2 < x < b_1 \wedge y_{12} \leq y \leq y_{23} \\ \frac{b_2 - y}{b_2 - y_2} h & \text{si } x_{34} \leq x \leq x_{23} \wedge y_2 < y < b_2 \\ \frac{x - a_1}{x_1 - a_1} h & \text{si } a_1 < x < x_1 \wedge y_{41} \leq y \leq y_{34} \\ h & \text{si } x_1 < x < x_2 \wedge y_1 \leq y \leq y_2 \\ 0 & \text{En otro caso} \end{cases} \quad (12)$$

3. Función de distribución de la distribución troncopiramidal

En el cálculo de la función de distribución hay que distinguir varios casos:

1. Si $(x_0, y_0) \in T_1$, se tiene que:

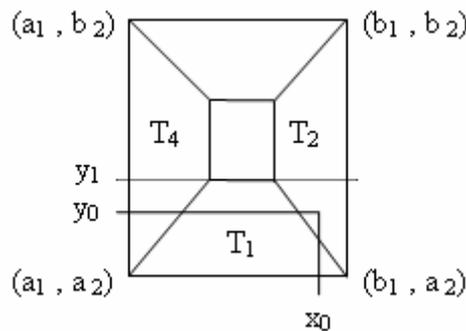


$$\begin{aligned} F(x_0, y_0) &= \int_{a_2}^{y_0} \int_{a_1}^{x_{41}} z_4 dx dy + \int_{a_2}^{y_0} \int_{x_{41}}^{x_0} z_1 dx dy = \\ &= \frac{h}{2} \frac{(x_0 - a_1)(y_0 - a_2)^2}{y_1 - a_2} - \frac{h}{6} \frac{x_1 - a_1}{(y_1 - a_2)^2} (y_0 - a_2)^3 \end{aligned} \quad (13)$$

Puede comprobarse que $F(a_1, a_2) = 0$ y $\frac{\partial^2 F(x_0, y_0)}{\partial x_0 \partial y_0} = h \frac{y_0 - a_2}{y_1 - a_2}$, que es la forma funcional de la función de densidad en la región T_1 .

2. Si $(x_0, y_0) \in T_2$, hay que distinguir los tres casos siguientes:

i. $(x_0, y_0) \in T_2$; con $y_0 \leq y_1$, se tiene que:

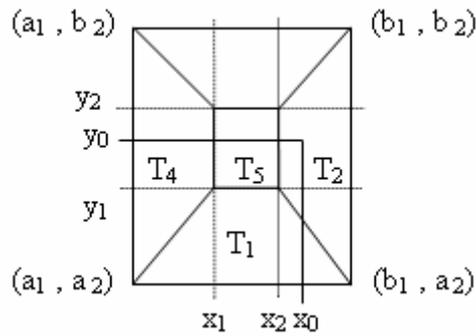


$$\begin{aligned}
 F(x_0, y_0) &= \int_{a_2}^{y_0} \int_{a_1}^{x_{41}} z_4 dx dy + \int_{y_{12}(x_0)}^{y_0} \int_{x_{41}}^{x_{12}} z_1 dx dy + \int_{y_{12}(x_0)}^{y_0} \int_{x_{12}}^{x_0} z_2 dx dy + \\
 &+ \int_{a_2}^{y_{12}(x_0)} \int_{x_{41}}^{x_0} z_1 dx dy = \frac{h}{2} \frac{b_1 - a_1}{y_1 - a_2} (y_0 - a_2)^2 - \frac{h}{6} \frac{b_1 - a_1 - (x_2 - x_1)}{(y_1 - a_2)^2} (y_0 - a_2)^3 + \quad (14) \\
 &+ \frac{h}{6} \frac{y_1 - a_2}{(b_1 - x_2)^2} (b_1 - x_0)^3 - \frac{h}{2} \frac{(y_0 - a_2)(b_1 - x_0)^2}{b_1 - x_2}
 \end{aligned}$$

Puede comprobarse que $\frac{\partial^2 F(x_0, y_0)}{\partial x_0 \partial y_0} = h \frac{b_1 - x_0}{b_1 - x_2}$, que es la forma funcional de

la función de densidad en la región T_2 .

ii. $(x_0, y_0) \in T_2$; con $y_1 < y_0 < y_2$, se tiene que:

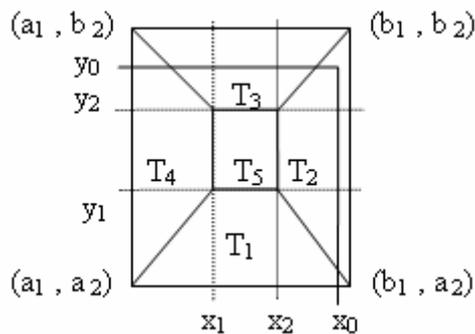


$$\begin{aligned}
 F(x_0, y_0) &= \int_{a_1}^{x_1} \int_{y_1}^{y_0} z_4 dx dy + \int_{a_2}^{y_1} \int_{x_1}^{x_2} z_1 dx dy + \int_{x_2}^{x_0} \int_{a_2}^{y_1} z_1 dx dy + \\
 &+ \int_{x_1}^{x_2} \int_{y_1}^{y_0} z_5 dy dx + \int_{x_2}^{x_0} \int_{y_1}^{y_0} z_2 dx dy = \\
 &= \frac{h}{2} (b_1 - x_1)(y_0 - a_2) - \frac{h}{2} \frac{(y_0 - a_2)(b_1 - x_0)^2}{b_1 - x_2} + \frac{h}{2} (x_2 - a_1)(y_0 - y_1) + \\
 &+ \frac{h}{3} (x_1 - a_1)(y_1 - a_2) - \frac{h}{6} (b_1 - x_2)(y_1 - a_2) + \frac{h}{6} \frac{y_1 - a_2}{(b_1 - x_2)^2} (b_1 - x_0)^3
 \end{aligned} \tag{15}$$

Puede comprobarse que $\frac{\partial^2 F(x_0, y_0)}{\partial x_0 \partial y_0} = h \frac{b_1 - x_0}{b_1 - x_2}$, que es la forma funcional de

la función de densidad en la región T_2 .

iii. $(x_0, y_0) \in T_2$; con $y_0 > y_2$, se tiene que:



$$\begin{aligned}
F(x_0, y_0) &= 1 - \left[\int_{y_0}^{b_2} \int_{a_1}^{x_{34}} z_4 dx dy + \int_{y_0}^{b_2} \int_{x_{34}}^{x_{23}} z_3 dx dy + \int_{x_0}^{b_1} \int_{y_{12}}^{y_{23}} z_2 dy dx + \right. \\
&\quad \left. + \int_{y_0}^{y_{23}(x_0)} \int_{x_{23}}^{x_0} z_2 dx dy + \int_{x_0}^{b_1} \int_{a_2}^{y_{12}} z_1 dy dx \right] = \\
&= 1 - \left[\frac{h}{2} \frac{b_1 - a_1}{b_2 - y_2} (b_2 - y_0)^2 - \frac{h}{6} \frac{b_1 - a_1 - (x_2 - x_1)}{(b_2 - y_2)^2} (b_2 - y_0)^3 - \right. \\
&\quad \left. - \frac{h}{6} \frac{y_1 - a_2}{(b_1 - x_2)^2} (b_1 - x_0)^3 + \frac{h}{2} \frac{(y_0 - a_2)(b_1 - x_0)^2}{b_1 - x_2} \right] \quad (16)
\end{aligned}$$

Puede comprobarse que $F(b_1, b_2) = 1$ y $\frac{\partial^2 F(x_0, y_0)}{\partial x_0 \partial y_0} = h \frac{b_1 - x_0}{b_1 - x_2}$, que es la forma funcional de la función de densidad en la región T_2 .

3. Si $(x_0, y_0) \in T_3$, hay que distinguir tres casos:

i. $(x_0, y_0) \in T_3$; con $x_0 \leq x_1$, se tiene que:

$$\begin{aligned}
F(x_0, y_0) &= \int_{x_{34}(y_0)}^{x_0} \int_{y_{34}}^{y_0} z_3 dy dx + \int_{x_{34}(y_0)}^{x_0} \int_{y_{14}}^{y_{34}} z_4 dy dx + \\
&\quad + \int_{a_1}^{x_{34}(y_0)} \int_{y_{14}}^{y_0} z_4 dy dx + \int_{a_1}^{x_0} \int_{a_2}^{y_{14}} z_1 dy dx = \\
&= \frac{h}{2} \frac{b_2 - a_2}{x_1 - a_1} (x_0 - a_1)^2 - \frac{h}{6} \frac{b_2 - a_2 - (y_2 - y_1)}{(x_1 - a_1)^2} (x_0 - a_1)^3 + \\
&\quad + \frac{h}{6} \frac{x_1 - a_1}{(b_2 - y_2)^2} (b_2 - y_0)^3 - \frac{h}{2} \frac{(x_0 - a_1)(b_2 - y_0)^2}{b_2 - y_2} \quad (17)
\end{aligned}$$

Puede comprobarse que $\frac{\partial^2 F(x_0, y_0)}{\partial x_0 \partial y_0} = h \frac{b_2 - y_0}{b_2 - y_2}$, que es la forma funcional de la función de densidad en la región T_3 .

ii. $(x_0, y_0) \in T_3$; con $x_1 < x_0 < x_2$, se tiene que:

$$\begin{aligned}
F(x_0, y_0) &= \int_{a_1}^{x_1} \int_{y_{41}}^{y_2} z_4 dx dy + \int_{y_2}^{y_0} \int_{a_1}^{x_{34}} z_4 dy dx + \int_{a_2}^{y_1} \int_{x_{41}}^{x_0} z_1 dx dy + \\
&+ \int_{x_1}^{x_0} \int_{y_1}^{y_2} z_5 dy dx + \int_{y_2}^{y_0} \int_{x_{34}}^{x_0} z_3 dx dy = \\
&= \frac{h}{2} (x_0 - a_1)(b_2 - y_1) - \frac{h}{2} \frac{(x_0 - a_1)(b_2 - y_0)^2}{b_2 - y_2} + \frac{h}{2} (x_0 - x_1)(y_2 - a_2) + \\
&+ \frac{h}{3} (x_1 - a_1)(y_1 - a_2) - \frac{h}{6} (x_1 - a_1)(b_2 - y_2) + \frac{h}{6} \frac{x_1 - a_1}{(b_2 - y_2)^2} (b_2 - y_0)^3
\end{aligned} \tag{18}$$

Puede comprobarse que $\frac{\partial^2 F(x_0, y_0)}{\partial x_0 \partial y_0} = h \frac{b_2 - y_0}{b_2 - y_2}$, que es la forma funcional de la función de densidad en la región T_3 .

iii. $(x_0, y_0) \in T_3$; con $x_0 > x_2$, se tiene que:

$$\begin{aligned}
F(x_0, y_0) &= 1 - \left[\int_{y_0}^{b_2} \int_{a_1}^{x_{34}} z_4 dx dy + \int_{y_0}^{b_2} \int_{x_{34}}^{x_{23}} z_3 dx dy + \int_{y_{23}(x_0)}^{y_0} \int_{x_0}^{x_{23}} z_3 dx dy + \right. \\
&\left. + \int_{x_0}^{b_1} \int_{y_{12}}^{y_{23}} z_2 dy dx + \int_{x_0}^{b_1} \int_{a_2}^{y_{12}} z_1 dy dx \right] = \\
&= 1 - \left[\frac{h}{2} \frac{(x_0 - a_1)(b_2 - y_0)^2}{b_2 - y_2} - \frac{h}{6} \frac{b_2 - a_2 - (y_2 - y_1)}{(b_1 - x_2)^2} (b_1 - x_0)^3 - \right. \\
&\left. - \frac{h}{6} \frac{x_1 - a_1}{(b_2 - y_2)^2} (b_2 - y_0)^3 + \frac{h}{2} \frac{b_2 - a_2}{b_1 - x_2} (b_1 - x_0)^2 \right]
\end{aligned} \tag{19}$$

Puede comprobarse que $F(b_1, b_2) = 1$ y $\frac{\partial^2 F(x_0, y_0)}{\partial x_0 \partial y_0} = h \frac{b_2 - y_0}{b_2 - y_2}$, que es la forma funcional de la función de densidad en la región T_3 .

4. Si $(x_0, y_0) \in T_4$:

$$\begin{aligned}
F(x_0, y_0) &= \int_{a_1}^{x_0} \int_{y_{41}}^{y_0} z_4 dy dx + \int_{a_1}^{x_0} \int_{a_2}^{y_{41}} z_1 dy dx = \\
&= \frac{h}{2} \frac{(y_0 - a_2)(x_0 - a_1)^2}{x_1 - a_1} - \frac{h}{6} \frac{y_1 - a_2}{(x_1 - a_1)^2} (x_0 - a_1)^3
\end{aligned} \tag{20}$$

Puede comprobarse que $F(a_1, a_2) = 0$ y $\frac{\partial^2 F(x_0, y_0)}{\partial x_0 \partial y_0} = h \frac{x_0 - a_1}{x_1 - a_1}$, que es la forma funcional de la función de densidad en la región T_4 .

5. Si $(x_0, y_0) \in T_5$:

$$\begin{aligned} F(x_0, y_0) &= \int_{a_1}^{x_1} \int_{y_{41}}^{y_0} z_4 dy dx + \int_{a_2}^{y_1} \int_{x_{41}}^{x_0} z_1 dx dy + \int_{x_1}^{x_0} \int_{y_1}^{y_0} z_5 dy dx = \\ &= \frac{h}{2}(x_0 - x_1)(y_0 - a_2) + \frac{h}{2}(x_0 - a_1)(y_0 - y_1) + \frac{h}{3}(x_1 - a_1)(y_1 - a_2) \end{aligned} \quad (21)$$

Puede comprobarse que $\frac{\partial^2 F(x_0, y_0)}{\partial x_0 \partial y_0} = h$, que es la forma funcional de la función de densidad en la región T_5 .

4. Funciones de densidad marginales

4.1. Densidad marginal de X

Se distinguen tres casos para obtener $f_1(x)$:

i. Si $a_1 < x \leq x_1$, se tiene:

$$\int_{a_2}^{y_{41}} z_1 dy + \int_{y_{41}}^{y_{34}} z_4 dy + \int_{y_{34}}^{b_2} z_3 dy = h \frac{b_2 - a_2}{x_1 - a_1} (x - a_1) - \frac{h}{2} \frac{(b_2 - a_2) - (y_2 - y_1)}{(x_1 - a_1)^2} (x - a_1)^2 \quad (22)$$

ii. Si $x_1 < x \leq x_2$, se tiene:

$$\int_{a_2}^{y_1} z_1 dy + \int_{y_1}^{y_2} z_5 dy + \int_{y_2}^{b_2} z_3 dy = \frac{h}{2} (b_2 - a_2 + y_2 - y_1) \quad (23)$$

iii. Si $x_2 \leq x < b_1$, se tiene:

$$\int_{a_2}^{y_{12}} z_1 dy + \int_{y_{12}}^{y_{23}} z_2 dy + \int_{y_{23}}^{b_2} z_3 dy = h \frac{b_2 - a_2}{b_1 - x_2} (b_1 - x) - \frac{h}{2} \frac{(b_2 - a_2) - (y_2 - y_1)}{(b_1 - x_2)^2} (b_1 - x)^2 \quad (24)$$

Fácilmente se demuestra que $\int_{a_1}^{b_1} f_1(x) dx = 1$, teniendo en cuenta (10).

El momento no centrado de orden 1 de la variable X tiene la siguiente expresión:

$$\alpha_{10} = E[x] = \frac{h}{12} \left[\begin{aligned} &(b_1^2 - a_1^2)(b_2 - a_2) + \frac{h}{12}(x_2^2 - x_1^2)(y_2 - y_1) + \\ &+ \frac{h}{24}((b_1 + x_2)^2 - (x_1 + a_1)^2)(b_2 - a_2 + y_2 - y_1) \end{aligned} \right] \quad (25)$$

El momento no centrado de orden 2 de la variable X tiene la siguiente expresión:

$$\alpha_{20} = \frac{h}{60} \left[\begin{aligned} &((b_1 + x_2)^3 - (x_1 + a_1)^3)(b_2 - a_2 + y_2 - y_1) + 3(b_2 - a_2)(b_1^3 - a_1^3) + \\ &+ (b_2 - a_2)(a_1 x_1^2 - b_1 x_2^2) + (y_2 - y_1)(3(x_2^3 - x_1^3) + x_1 a_1^2 - x_2 b_1^2) \end{aligned} \right] \quad (26)$$

De donde la varianza de la variable X, puede expresarse de la siguiente forma:

$$\text{var}_x = \frac{h}{2880} \left\{ \begin{aligned} &48 \left[\begin{aligned} &((b_1 + x_2)^3 - (x_1 + a_1)^3)(b_2 - a_2 + y_2 - y_1) + 3(b_2 - a_2)(b_1^3 - a_1^3) + \\ &+ (b_2 - a_2)(a_1 x_1^2 - b_1 x_2^2) + (y_2 - y_1)(3(x_2^3 - x_1^3) + x_1 a_1^2 - x_2 b_1^2) \end{aligned} \right] - \\ &- 5h \left[\begin{aligned} &2(b_1^2 - a_1^2)(b_2 - a_2) + 2(x_2^2 - x_1^2)(y_2 - y_1) + \\ &+ ((b_1 + x_2)^2 - (x_1 + a_1)^2)(b_2 - a_2 + y_2 - y_1) \end{aligned} \right]^2 \end{aligned} \right\} \quad (27)$$

4.2. Densidad marginal de Y

Se distinguen tres casos para obtener $f_2(y)$:

i. Si $a_2 < y \leq y_1$, se tiene:

$$\int_{a_1}^{x_{41}} z_4 dx + \int_{x_{41}}^{x_{12}} z_1 dx + \int_{x_{12}}^{b_1} z_2 dx = h \frac{b_1 - a_1}{y_1 - a_2} (y - a_2) - \frac{h}{2} \frac{(b_1 - a_1) - (x_2 - x_1)}{(y_1 - a_2)^2} (y - a_2)^2 \quad (28)$$

ii. Si $y_1 \leq y \leq y_2$, se tiene:

$$\int_{a_1}^{x_1} z_4 dx + \int_{x_1}^{x_2} z_5 dx + \int_{x_2}^{b_1} z_2 dx = \frac{h}{2} (b_1 - a_1 + x_2 - x_1) \quad (29)$$

iii. Si $y_2 \leq y < b_2$, se tiene:

$$\int_{a_1}^{x_{34}} z_4 dx + \int_{x_{34}}^{x_{23}} z_3 dx + \int_{x_{23}}^{b_1} z_2 dx = h \frac{b_1 - a_1}{b_2 - y_2} (b_2 - y) - \frac{h}{2} \frac{(b_1 - a_1) - (x_2 - x_1)}{(b_2 - y_2)^2} (b_2 - y)^2 \quad (30)$$

Fácilmente se demuestra que $\int_{a_2}^{b_2} f_2(y) dy = 1$, teniendo en cuenta (10).

El momento no centrado de orden 1 de la variable Y tiene la siguiente expresión:

$$\alpha_{01} = E[y] = \left[\begin{aligned} & \frac{h}{12} (b_2^2 - a_2^2) (b_1 - a_1) + \frac{h}{12} (y_2^2 - y_1^2) (x_2 - x_1) + \\ & + \frac{h}{24} ((b_2 + y_2)^2 - (y_1 + a_2)^2) (b_1 - a_1 + x_2 - x_1) \end{aligned} \right] \quad (31)$$

El momento no centrado de orden 2 de la variable Y tiene la siguiente expresión:

$$\alpha_{02} = \frac{h}{60} \left[\begin{aligned} & ((b_2 + y_2)^3 - (y_1 + a_2)^3) (b_1 - a_1 + x_2 - x_1) + 3(b_1 - a_1) (b_2^3 - a_2^3) + \\ & + (b_1 - a_1) (a_2 y_1^2 - b_2 y_2^2) + (x_2 - x_1) (3(y_2^3 - y_1^3) + y_1 a_2^2 - y_2 b_2^2) \end{aligned} \right] \quad (32)$$

De donde la varianza de Y, puede expresarse de la siguiente forma:

$$\text{var}_y = \frac{h}{2880} \left\{ \begin{aligned} & 48 \left[\begin{aligned} & ((b_2 + y_2)^3 - (y_1 + a_2)^3) (b_1 - a_1 + x_2 - x_1) + 3(b_1 - a_1) (b_2^3 - a_2^3) + \\ & + (b_1 - a_1) (a_2 y_1^2 - b_2 y_2^2) + (x_2 - x_1) (3(y_2^3 - y_1^3) + y_1 a_2^2 - y_2 b_2^2) \end{aligned} \right] - \\ & - 5h \left[\begin{aligned} & 2(b_2^2 - a_2^2) (b_1 - a_1) + 2(y_2^2 - y_1^2) (x_2 - x_1) + \\ & + ((b_2 + y_2)^2 - (y_1 + a_2)^2) (b_1 - a_1 + x_2 - x_1) \end{aligned} \right]^2 \end{aligned} \right\} \quad (33)$$

Proposición

Las funciones de densidad marginales de la distribución troncopiramidal satisfacen la ecuación diferencial del sistema de Pearson, correspondiente a las distribuciones univariantes continuas.

$$f_1(x) = \begin{cases} h \frac{b_2 - a_2}{x_1 - a_1} (x - a_1) - \frac{h}{2} \frac{(b_2 - a_2) - (y_2 - y_1)}{(x_1 - a_1)^2} (x - a_1)^2 & \text{si } a_1 < x \leq x_1 \\ \frac{h}{2} (b_2 - a_2 + y_2 - y_1) & \text{si } x_1 \leq x \leq x_2 \\ h \frac{b_2 - a_2}{b_1 - x_2} (b_1 - x) - \frac{h}{2} \frac{(b_2 - a_2) - (y_2 - y_1)}{(b_1 - x_2)^2} (b_1 - x)^2 & \text{si } x_2 \leq x < b_1 \end{cases} \quad (34)$$

Como $f_1(x_1) = f_1(x_2) = \frac{h}{2} (b_2 - a_2 + y_2 - y_1)$ y además:

$$f_1'(x) = \begin{cases} h \frac{b_2 - a_2}{x_1 - a_1} - h \frac{(b_2 - a_2) - (y_2 - y_1)}{(x_1 - a_1)^2} (x - a_1) & \text{si } a_1 < x \leq x_1 \\ 0 & \text{si } x_1 \leq x \leq x_2 \\ -h \frac{b_2 - a_2}{b_1 - x_2} + h \frac{(b_2 - a_2) - (y_2 - y_1)}{(b_1 - x_2)^2} (b_1 - x) & \text{si } x_2 \leq x < b_1 \end{cases} \quad (35)$$

Obsérvese que $f_1'(x_1^-) = h \frac{y_2 - y_1}{x_1 - a_1}$ y $f_1'(x_2^+) = -h \frac{y_2 - y_1}{b_1 - x_2}$.

Si se nota por:

$$p = \frac{(b_2 - a_2)(x_1 - a_1)}{b_2 - a_2 - (y_2 - y_1)} \quad \text{y} \quad q = \frac{(b_2 - a_2)(b_1 - x_2)}{b_2 - a_2 - (y_2 - y_1)}$$

Se tiene:

$$\frac{f_1'(x)}{f_1(x)} = \begin{cases} \frac{x - (a_1 + p)}{\frac{x^2}{2} - (a_1 + p)x + a_1 p + \frac{a_1^2}{2}} & \text{si } a_1 < x < x_1 \\ 0 & \text{si } x_1 < x < x_2 \\ \frac{x - (b_1 - q)}{\frac{x^2}{2} - (b_1 - q)x - b_1 q + \frac{b_1^2}{2}} & \text{si } x_2 < x < b_1 \end{cases} \quad (36)$$

Obsérvese que las raíces de los denominadores son respectivamente $(a_1 y 2p + a_1)$ y $(b_1 y b_1 - 2q)$.

5. La distribución troncopiramidal estandarizada

Como es habitual, el recorrido finito de una variable aleatoria puede transformarse en un recorrido sobre $(0,1)$. Este recorrido estandarizado se obtiene restando al valor original de la variable el valor más pequeño y el resultado se divide por el recorrido de la variable original, esto es, la diferencia entre el mayor y el menor de los valores de la variable original, por tanto, si el recorrido de la variable X es (a,b) , la variable estandarizada $X^* = \frac{X - a}{b - a}$ evidentemente toma sus valores en $(0,1)$.

En el caso de la distribución piramidal pueden transformarse los recorridos de las variables X e Y , $R_X(a_1, b_1)$ y $R_Y(a_2, b_2)$ en recorridos estandarizados $R_{X^*}(0,1)$ y $R_{Y^*}(0,1)$ con lo que las fórmulas obtenidas anteriormente para el caso de una distribución piramidal general, se convierten en fórmulas mucho más sencillas y simples en su aplicación empírica.

A continuación se detalla, a manera de ejemplo, como resultan las fórmulas generales, cuando se realiza la estandarización en los recorridos de las variables:

Aplicando que:

$$x^* = \frac{x - a_1}{b_1 - a_1} \quad e \quad y^* = \frac{y - a_2}{b_2 - a_2} \quad (37)$$

Se tiene:

$$a_1^* = 0 = a_2^*, \quad b_1^* = 1 = b_2^*, \quad (38)$$

$$x_1^* = \frac{x_1 - a_1}{b_1 - a_1}, \quad x_2^* = \frac{x_2 - a_1}{b_1 - a_1}, \quad y_1^* = \frac{y_1 - a_2}{b_2 - a_2}, \quad y_2^* = \frac{y_2 - a_2}{b_2 - a_2}$$

La expresión de la constante normalizadora h queda reducida a :

$$h^* = \frac{6}{1 + (x_2^* - x_1^*)(y_2^* - y_1^*) + (1 + x_2^* - x_1^*)(1 + y_2^* - y_1^*)} \quad (39)$$

La función de densidad queda reducida a:

$$f(x^*, y^*) = \begin{cases} \frac{h^*}{y_1^*} y^* & \text{si } (x^*, y^*) \in T_1 \\ \frac{h^*}{(1 - x_2^*)} (1 - x^*) & \text{si } (x^*, y^*) \in T_2 \\ \frac{h^*}{(1 - y_2^*)} (1 - y^*) & \text{si } (x^*, y^*) \in T_3 \\ \frac{h^*}{x_1^*} x^* & \text{si } (x^*, y^*) \in T_4 \\ h^* & \text{si } (x^*, y^*) \in T_5 \\ 0 & \text{en otro caso} \end{cases} \quad (40)$$

Los momentos no centrados de orden 1 tienen las expresiones siguientes:

$$E[x^*] = \frac{h^*}{24} \left(2 + (1 + 2x_2^*)(1 + y_2^* - y_1^*) + (1 + 3y_2^* - 3y_1^*)(x_2^{*2} - x_1^{*2}) \right) \quad (41)$$

$$E[y^*] = \frac{h^*}{24} \left(2 + (1 + 2y_2^*)(1 + x_2^* - x_1^*) + (1 + 3x_2^* - 3x_1^*)(y_2^{*2} - y_1^{*2}) \right)$$

Los momentos no centrados de orden 2 tienen las siguientes expresiones:

$$\alpha_{20}^* = \frac{h^*}{60} \left(3 + x_2^* - x_2^{*2} + (1 + 2x_2^* + 3x_2^{*2})(1 + y_2^* - y_1^*) + (1 + 4y_2^* - 4y_1^*)(x_2^{*3} - x_1^{*3}) \right) \quad (42)$$

$$\alpha_{02}^* = \frac{h^*}{60} \left(3 + y_2^* - y_2^{*2} + (1 + 2y_2^* + 3y_2^{*2})(1 + x_2^* - x_1^*) + (1 + 4x_2^* - 4x_1^*)(y_2^{*3} - y_1^{*3}) \right)$$

Las varianzas de las variables estandarizadas tienen las expresiones siguientes:

$$\sigma_{x^*}^2 = \frac{h^*}{2880} \left[48 \left(3 + x_2^* - x_2^{*2} + (1 + 2x_2^* + 3x_2^{*2})(1 + y_2^* - y_1^*) + (1 + 4y_2^* - 4y_1^*)(x_2^{*3} - x_1^{*3}) \right) - 5h^* \left(2 + (1 + 2x_2^*)(1 + y_2^* - y_1^*) + (1 + 3y_2^* - 3y_1^*)(x_2^{*2} - x_1^{*2}) \right)^2 \right] \quad (43)$$

$$\sigma_{y^*}^2 = \frac{h^*}{2880} \left[48 \left(3 + y_2^* - y_2^{*2} + (1 + 2y_2^* + 3y_2^{*2})(1 + x_2^* - x_1^*) + (1 + 4x_2^* - 4x_1^*)(y_2^{*3} - y_1^{*3}) \right) - 5h^* \left(2 + (1 + 2y_2^*)(1 + x_2^* - x_1^*) + (1 + 3x_2^* - 3x_1^*)(y_2^{*2} - y_1^{*2}) \right)^2 \right]$$

La covarianza de X^* e Y^* puede expresarse de la siguiente forma:

$$\text{cov}_{x^* y^*} = \frac{h^*}{2880} \left[24 \left(5 + (1 + 3x_2^* + 6(x_2^{*2} - x_1^{*2}))(y_2^{*2} - y_1^{*2}) + (1 + x_2^* + x_2^{*2} - x_1^{*2})(1 + 3y_2^*) + x_2^*(2 + y_2^*) \right) - 5h^* \left[\left(2 + (1 + 2x_2^*)(1 + y_2^* - y_1^*) + (1 + 3y_2^* - 3y_1^*)(x_2^{*2} - x_1^{*2}) \right) \right] \left[\left(2 + (1 + 2y_2^*)(1 + x_2^* - x_1^*) + (1 + 3x_2^* - 3x_1^*)(y_2^{*2} - y_1^{*2}) \right) \right] \right] \quad (44)$$

Bibliografía

BALLESTERO, E. (1973) "Nota sobre un nuevo método rápido de valoración". Revista de Estudios Agrosociales, 85, 75-78.

ELDERTON, W. P. and JOHNSON, N. L. (1969). "Systems of frequency curves". Cambridge University Press.

HERRERÍAS, R.; GARCÍA, J.; CRUZ, S. y HERRERÍAS, J. M. (2001) "Il modello probabilistico trapezoidale, nel metodo delle due distribuzioni della teoria generale di valutazione". Genio Rurale. Anno LXIV Abril 2001, n° 4, 3-9

LOZANO, J. J. (1996) *“Tasación urbana: una metodología para informes de tasación masiva”*. Tesis Doctoral. Universidad Politécnica Madrid.

ROMERO, C. (1977) *“Valoración por el método de las dos distribuciones Beta: una extensión”*. Revista de Economía Política, 75, 47-62. Madrid.

CONCILIACIÓN ESPACIAL-SECTORIAL DE TRIMESTRALIZACIÓN DE SERIES. UNA APLICACIÓN A LAS CONTABILIDADES PROVINCIALES¹.

JOSÉ LUIS ROJO GARCÍA
e-mail: (rojo_s@eco.uva.es)

JOSÉ ANTONIO SANZ GÓMEZ
e-mail: (josea.sanz@eco.uva.es)

Departamento de Economía Aplicada)
UNIVERSIDAD DE VALLADOLID

Área temática: Métodos cuantitativos.

Resumen

En un trabajo anterior, los autores desarrollaron un modelo bayesiano jerárquico, con especificaciones normal-gamma, para la desagregación temporal de series económicas, en presencia de indicadores. La comparación con otros métodos avalaba la calidad del procedimiento desarrollado, comparándolo con procedimientos clásicos (Chow-lin, con especificación AR(1) o de ruido blanco para las perturbaciones). Se ilustra el método con una desagregación de los valores añadidos regionales en la Contabilidad Regional del INE.

Cuando se desagregan temporalmente varias series sujetas a identidades contables (por ejemplo sectoriales y/o geográficas), se plantea la conciliación entre desagregaciones temporales individuales e identidades transversales. Existen aportaciones de interés en este campo, alguna de ellas recientes, desde distintas metodologías, pero no se han realizado, que conozcamos, propuestas desde la óptica bayesiana.

Existen referencias adecuadas para los procedimientos clásicos, mientras que trabajos recientes han abordado la conciliación de series desestacionalizadas provenientes de series brutas ya armonizadas.

En este trabajo se realiza la conciliación simultánea entre la desagregación temporal y la espacial-sectorial (siguiendo una tabla de doble entrada) mediante un modelo conjunto bayesiano jerárquico normal-gamma, que resultaría de la yuxtaposición de los modelos individuales y del agregado, con una estructura de covarianzas adecuada, y suponiendo conocidos indicadores de comportamiento. Se obtiene la solución conjuntamente óptima y se ilustra su funcionamiento mediante una aplicación que obtiene estimaciones trimestrales provinciales con desagregación sectorial a partir de la Contabilidad Regional de España y de la Contabilidad Regional Trimestral de Castilla y León

Palabras clave: Métodos bayesianos, Análisis regional, Desagregación temporal, Cuentas regionales.

SPATIAL-SECTORAL BENCHMARKING FOR A TWO-WAY CLASSIFIED TABLE OF TIME SERIES. IMPLEMENTATION FOR THE REGIONAL AND PROVINCIAL ACCOUNTS

Abstract

In a previous paper, we present a bayesian hierarchical normal-gamma model, allowing for benchmarking of time series, using indicators. The model has been compared with classical ones (AR(1)-Chow-Lin, among others), and its practical interest showed by benchmarking annual data from the Spanish Regional Accounts (I.N.E.)

The disaggregation of time series linked by accounting constraints (contemporaneous and/or temporal ones) involves the conciliation of individual disaggregations. In recent years there have been relevant contributions to this topic but, as far as we know, the bayesian approach has not been used.

Classical approaches are well described in the literature, and recent proposals point to the conciliation of SA series from conciliated original ones.

In this paper, we deal with the simultaneous benchmarking of a two-way classified table of time series, by using a bayesian hierarchical normal-gamma model that superimposes the individual and the global indicators models and by settling the appropriate random structure. Authors derive the joint optimal solution and, finally, the method is used to benchmark the province's Spanish Regional Accounts (Spanish Statistics Institute) with the Quarterly Regional Accounts (Regional Statistics Institute of Castilla y León)

Key words: Bayesian methods, Benchmarking, Regional Accounts.

¹Trabajo OTP/06/06 realizado bajo subvención de la Consejería de Hacienda de la Junta de Castilla y León (Dirección General de Estadística), concedida por Resolución de 28-09-06 (Convocatoria de 10-04-06, de concurso público para la realización de trabajos y proyectos de investigación en materia de Estadística aplicada, BOCYL nº 78 de 21-04-06).

1. Introducción

En un trabajo anterior (Rojo y Sanz (2005)), los autores desarrollaron un procedimiento bayesiano para estimar una serie trimestral, dada la correspondiente agregación anual, y de acuerdo con la información proporcionada por un conjunto de indicadores trimestrales. El objetivo era la distribución de magnitudes anuales, en el sentido en que Friedman (1962) da a este término.

Desde el punto de vista distribucional, se trataba de un modelo bayesiano jerárquico, con una especificación de las denominadas normal-gamma (véase, por ejemplo, Spiegelhalter et. al. (1999) o Box y Tiao (1973)), por establecer distribuciones normales para las variables de posición y distribuciones gamma para las de precisión. Ello garantiza que el problema se mueve en las coordenadas de una de las familias conjugadas de distribuciones, con marginales gamma o de Student, según los casos.

Una de las aportaciones más novedosas de dicho trabajo residió en el sistema elegido para el establecimiento de las distribuciones marginales a priori. La idea (véase asimismo Rojo y Sanz (2005)) es que la utilización de distribuciones neutrales (o aproximadamente neutrales) a priori proporciona distribuciones a posteriori sobre-dispersas, por el escaso peso de la verosimilitud sobre el conjunto del modelo.

La determinación de las distribuciones marginales a priori obtenidas en este apartado permitió la asignación de valores razonables para los parámetros implicados en las mismas, de manera que las distribuciones utilizadas eran propias. Ello se conseguía a partir de una pre-trimestralización de los valores anuales de la serie, pre-trimestralización que utiliza un procedimiento sin indicadores, en concreto, el método de Boot-Feibes y Lisman (Boot y Feibes, (1967), BFL a partir de ahora). Las conclusiones que afloraron en este apartado indican que es posible realizar asignaciones a priori de los parámetros e hiperparámetros de las distribuciones a priori, exceptuando la matriz P , implicada en la distribución a priori de la serie trimestral (debido al excesivo número de parámetros implicados), que se tomó igual a la matriz identidad en las aplicaciones. Ello no implica que, a posteriori, sea igual a la identidad, obteniéndose correlaciones no nulas a posteriori.

Finalmente se obtuvieron las distribuciones marginales a posteriori de los parámetros y variables implicados, restringidas por la necesaria conciliación entre datos trimestrales y anuales, y el estimador óptimo (de Bayes), esperanza de la distribución marginal a posteriori. Las distribuciones son gammas o de Student multivariantes, con lo cual la expresión de su esperanza es conocida, no siendo necesarios procedimientos de simulación para su obtención. Los autores elaboraron programas de aplicación en TSP y en SPLUS para su desarrollo práctico, que permitieron evidenciar la calidad del método propuesto mediante un ejemplo tomado de la Contabilidad Nacional Trimestral en varios escenarios para las perturbaciones del modelo de verosimilitud, comparando la solución obtenida, tanto con los “verdaderos” valores, como con la solución que proporciona el procedimiento clásico más idóneo en cada caso (Denton, Chow-Lin o Chow-Lin con esquema de Fernández, véanse Denton (1971), Chow y Lin (1971) o Fernández (1981)).

Aunque es necesario realizar comparaciones sistemáticas mediante procedimientos de simulación (véase Rodríguez Feijóo y otros (2003)) las conclusiones de la anterior aplicación establecieron la mayor calidad del procedimiento bayesiano obtenido frente a los procedimientos clásicos, en los tres escenarios. Es posible que la regularidad implicada por el procedimiento bayesiano, junto con el hecho de que la serie original de la Contabilidad Nacional Trimestral fuera una ciclo-tendencia, y por ello suave, haya realizado más la calidad de nuestro método.

Desde entonces, las técnicas de trimestralización han alcanzado nuevos desarrollos (véase, por ejemplo, Di Fonzo (2003) o Santos Silva y Cardoso (2001) sobre procedimientos dinámicos, Jerez et. al. (2005), Proietti (2005) o Palate (2005) sobre uso del Filtro de Kalman). Sin embargo, no se ha avanzado grandemente en el terreno de la conciliación entre la desagregación temporal y la espacial-sectorial. Laniel y Fyfe (1989) propusieron en su día una solución iterativa, que discrimina entre la importancia relativa de las distintas restricciones, utilizando el algoritmo denominado “raking ratio” y, básicamente, los resultados se siguen encontrando al nivel de los obtenidos y referenciados en Pavía (1997) salvo el reciente trabajo de Di Fonzo y Marini (2005) o el de Queneville y Rancourt (2005). Y por otro lado, no existen, que conozcamos, referencias de metodología bayesiana en este terreno. Sin embargo, los procedimientos

bayesianos pueden presentar ventajas adicionales cuando la base estadística informativa se refiere a pequeñas áreas (véase Rojo y Sanz (2002)), lo que es importante en las aplicaciones provinciales, debido a su menor tamaño y a la ausencia de datos o a la falta de calidad estadística de los mismos.

En este trabajo se realiza dicha conciliación mediante un modelo conjunto bayesiano normal-gamma, con una estructura de covarianzas adecuada. Aunque en un principio se pensó en que los indicadores del procedimiento fueran trimestralizaciones no armonizadas, en la línea argumental del trabajo de Pavía (1997), finalmente la opción elegida utiliza en la verosimilitud indicadores de comportamiento, para poder cubrir trabajos como el ya citado de Di Fonzo y Marini (2005), que se aplica a la conciliación de series mensuales desestacionalizadas provenientes de series brutas armonizadas, entre otros puntos de interés.

En la línea de los trabajos más recientes de los autores, las distribuciones a priori son informativas (no neutrales, en consecuencia) y el modelo global utilizará las distribuciones a posteriori, restringidas, para las series trimestrales, en la línea de Rojo y Sanz (2005).

El citado trabajo de Di Fonzo y Marini (DiFonzo y Marini, 2005) se plantea un objetivo relacionado con el nuestro, la armonización de series desestacionalizadas provenientes de series brutas armonizadas. Adicionalmente, plantean la armonización de series con restricciones siguiendo una tabla de doble entrada, que es el problema que nosotros abordamos, si bien la solución utiliza distinta metodología y no permite el uso de indicadores, sino sólo de aproximaciones. Dichos autores señalan que el trabajo se puede aplicar a objetivos más generales relacionados con identidades contables transversales (como por ejemplo los abordados en Chen y Dagum (1997), Eurostat (1999) o Di Fonzo (2002)).

2. Planteamiento del problema

Supongamos que se dispone de un marco contable para una Contabilidad provincial, en el que la región resulta del agregado de R provincias, estando a su vez la actividad económica desagregada en S sectores o mejor, ramas de actividad.

Se suponen conocidos los agregados anuales para cada provincia y rama, en un conjunto homogéneo, metodológicamente hablando, de N años

$$y_{ij} = (y_{ij1}, \dots, y_{ijN})', \quad i = 1, \dots, R, \quad j = 1, \dots, S$$

y se suponen asimismo conocidos los datos trimestrales para el conjunto regional correspondientes a cada una de las S ramas, para los n trimestres correspondientes a los N años (suponemos años completos, $n = 4N$), esto es, conocemos $b_j, j = 1, \dots, S$, siendo $b_j = (b_{j1}, \dots, b_{jN})', j = 1, \dots, S$.

El objetivo es la obtención de una estimación para las cifras trimestrales desagregadas por sectores y provincias, $x_{ij} = (x_{ij1}, \dots, x_{ijn})', i = 1, \dots, R, j = 1, \dots, S$, de manera que se verifican las restricciones que llamaremos “anuales”,

$$y_{ija} = \sum_{t=4(a-1)+1}^{4a} x_{ijt}, \quad a = 1, \dots, N, \quad i = 1, \dots, R, \quad j = 1, \dots, S \quad [1]$$

y las que denominaremos “regionales”,

$$b_j = \sum_{i=1}^R x_{ij}, \quad j = 1, \dots, S \quad [2]$$

Este sistema de series trimestrales con una doble clasificación lo podemos visualizar de forma gráfica como

Provincia	Rama de actividad					Total
	1	...	j	...	S	
1	x_{11}	...	x_{1j}	...	x_{1S}	a_1
\vdots	\vdots	\ddots	\vdots	\ddots	\vdots	\vdots
i	x_{i1}	...	x_{ij}	...	x_{iS}	a_i
\vdots	\vdots	\ddots	\vdots	\ddots	\vdots	\vdots
R	x_{R1}	...	x_{Rj}	...	x_{RS}	a_R
Total	b_1	...	b_j	...	b_S	z

Por ejemplo, en la Contabilidad Regional del INE, se dispone de valores anuales desde el lado de la oferta para las 17 Comunidades autónomas, y de la Contabilidad Nacional Trimestral que proporciona el conjunto nacional con frecuencia trimestral en el

entendimiento de que las magnitudes anuales y los datos de la Contabilidad Regional Trimestral de la D.G. de Estadística de la Junta de Castilla y León estuvieran armonizadas.

Di Fonzo y Marini (2003) establecen distintas situaciones a estudiar, dependiendo de cuáles entre las restricciones (anuales, regionales, sectoriales) sean conocidas. Nuestro trabajo podría transformarse para resolver dichas situaciones, si bien nos centramos en la aquí propuesta por adaptarse a la realidad estadística regional.

Obsérvese que estamos tomando las variables como “flujos”, de manera que la agregación se produce por suma. Si las variables fueran “índices”, la agregación se obtendría como media. Los cambios son anecdóticos, y no se reflejan en la profundidad del trabajo. Otro tanto ocurre si la cifra anual fuera la del último trimestre (saldo) o la del primero. Señalemos también que este procedimiento está diseñado para trimestralizar datos anuales, que es la aplicación propuesta. Con otras frecuencias (mensualización de datos trimestrales o anuales) las modificaciones son asimismo de menor entidad.

3. Formulación bayesiana. Distribuciones a priori

Sea x la matriz columna $x = (x'_{11}, \dots, x'_{R1}, \dots, x'_{1S}, \dots, x'_{RS})$ de dimensión $nRS \times 1$ con los valores trimestrales a estimar. Como en Rojo y Sanz (2005) supondremos una distribución normal a priori, dada la precisión τ ,

$$\pi(x | \tau, D) \propto \tau^{m/2} \exp\left\{-\frac{\tau}{2}(x - \mu)' P (x - \mu)\right\}$$

siendo $m = nRS$, P una matriz $m \times m$ no singular y $D = \{y_{ij}, b_j, i = 1, \dots, R, j = 1, \dots, S\}$

los datos. En la forma cuadrática normal se introduce una forma cuadrática adicional que proporciona mayores probabilidades a priori a las series más regulares, para evitar la volatilidad habitual de los métodos de trimestralización. Se añade, por tanto, un factor

$$\exp\left\{-\frac{\tau}{2} \sum_{i=1}^R \sum_{j=1}^S \sum_{t=2}^n (x_{ijt} - x_{ij,t-1})^2\right\} = \exp\left\{-\frac{\tau}{2} x' [I_{RS} \otimes D' D] x\right\}$$

donde D es la matriz $(n-1) \times n$ de primeras diferencias ($d_{ii} = -1, d_{i,i+1} = 1, i = 1, \dots, n-1$, y el resto ceros). Un mayor grado de alisado se obtiene introduciendo un factor $\exp\left\{-\frac{\tau}{2}x' [I_{RS} \otimes D_2' D_2] x\right\}$ donde D_2 es la matriz $(n-2) \times n$ que produce segundas diferencias, ($d_{ii} = 1, d_{i,i+1} = -2, d_{i,i+2} = 1, i = 1, \dots, n-2$). En este trabajo utilizaremos la corrección mediante las primeras diferencias.

No utilizaremos la matriz de primeras diferencias de Denton (Denton, 1971) que añade una fila inicial con un uno en su primera columna, propuesta para simplificar el cálculo de inversas al convertir a D en una matriz cuadrada (hay una versión similar para las segundas diferencias) ya que es innecesario en nuestro procedimiento y atrae el primer valor de la serie hacia el cero, haciendo muy dependiente el resultado de las escalas empleadas. Pueden verse las críticas de Cholette (1984) o Sanz (1981) a la aproximación de Denton.

Finalmente imponemos las restricciones, esto es, asignamos densidades a priori nulas a los valores que no respeten las restricciones anuales [1] o las regionales [2].

En definitiva, la distribución a priori de x condicionada por τ resulta

$$\pi(x | \tau, D) \propto \tau^{m/2} \cdot \exp\left\{-\frac{\tau}{2}[(x - \mu)' P(x - \mu) + x'(I_{RS} \otimes D' D)x]\right\}$$

sujeta a [1] y [2].

Para τ suponemos una distribución marginal a priori gamma,

$$\pi(\tau | D) \propto \tau^{\alpha-1} \cdot \exp\{-\tau\beta\}, \quad \tau > 0 \quad [3]$$

lo que garantiza que trabajaremos en un marco determinado por familias conjugadas de distribuciones. En definitiva, la distribución a priori conjunta de x y τ resulta

$$\pi(x, \tau | D) \propto \tau^{m/2+\alpha-1} \cdot \exp\left\{-\frac{\tau}{2}[2\beta + (x - \mu)' P(x - \mu) + x'(I_{RS} \otimes D' D)x]\right\}$$

En cuanto a la verosimilitud, para cada una de las RS series trimestrales se supone un modelo lineal de indicadores con perturbación aditiva,

$$x_{ijt} = \sum_{k=1}^{K_{ij}} \delta_{ijk} z_{ijt}^k + \varepsilon_{ijt}, \quad t = 1, \dots, n$$

El número de indicadores no tiene por qué ser el mismo para todas las series, puesto que la cantidad de información estadística para cada provincia y/o rama de actividad puede ser diferente. Podemos emplear un único indicador por serie (cuando el indicador es una aproximación o se trabaja con series de diferencias), pudiendo cada modelo tener o no un término constante. Aunque en la aplicación que se presenta más adelante se supone un único indicador y un término constante, las modificaciones para otras especificaciones no exigen cambios ni en los resultados ni el programa implementado puesto que sólo afectan a la dimensión de las matrices Z y δ .

En notación matricial, el modelo de verosimilitud se escribe como $x = Z\delta + \varepsilon$, siendo Z una matriz diagonal por bloques, $Z = \text{diag}(Z_{11}, Z_{21}, \dots, Z_{RS})$, de dimensión $nRS \times K$, con $K = \sum_{i=1}^R \sum_{j=1}^S K_{ij}$ y Z_{ij} la matriz cuyas columnas son los indicadores correspondientes al modelo para la serie x_{ij} .

La matriz $\varepsilon = (\varepsilon'_{11}, \dots, \varepsilon'_{R1}, \dots, \varepsilon'_{1S}, \dots, \varepsilon'_{RS})'$ es la matriz general de perturbaciones del modelo. Supondremos que el término de error de cada una de las series es normal en su distribución condicionada por τ (el hecho de que este parámetro de precisión sea el mismo que el anterior, no supone restricción alguna, porque puede actuarse sobre cada matriz $P_{\varepsilon_{ij}}$ posterior), y que es independiente a priori de los términos de error correspondientes a otras series, esto es, a priori $(\varepsilon | \tau, D) \rightarrow N_m(0, \tau^{-1} P_{\varepsilon}^{-1})$ donde $P_{\varepsilon}^{-1} = \text{diag}(P_{\varepsilon_{11}}^{-1}, P_{\varepsilon_{21}}^{-1}, \dots, P_{\varepsilon_{RS}}^{-1})$ es diagonal por bloques.

La estructura aleatoria de los errores puede fijarse en cada caso, siendo las asignaciones para los parámetros implicados diferentes para cada serie, aunque tomaremos el mismo modelo para cada una de dichas series.

En definitiva, la verosimilitud resulta

$$L(x, \delta, \tau | D) \propto \tau^{m/2} \cdot \exp \left\{ -\frac{\tau}{2} (x - Z\delta)' P_{\varepsilon} (x - Z\delta) \right\}$$

Para δ supondremos una distribución a priori, dada τ , normal K -dimensional $N_K(\bar{\delta}, \tau^{-1} P_{\delta}^{-1})$. Combinando la distribución a priori con la verosimilitud, obtenemos la distribución a posteriori,

$$p(x, \delta, \tau | D) \propto \tau^{m+\alpha-1+K/2} \cdot \exp \left\{ -\frac{\tau}{2} \left[2\beta + (x - \mu)' P (x - \mu) + x' (I_{RS} \otimes D' D) x + (\delta - \bar{\delta})' P_{\delta} (\delta - \bar{\delta}) + (x - Z\delta)' P_{\varepsilon} (x - Z\delta) \right] \right\} \quad [4]$$

con x sujeta a las restricciones anuales y regionales, [1] y [2].

Otra expresión para las restricciones

La existencia de restricciones lineales en una distribución continua conduce a una distribución degenerada. Su efecto aparente sería, por tanto, hacernos abandonar la familia de distribuciones normal-gamma, lo que dificultaría notablemente la obtención de las soluciones de Bayes al problema de trimestralización planteado. No obstante, es conocido (véase, por ejemplo, Rojo y Sanz (2005)) que basta con introducir la restricción en la densidad conjunta a posteriori para obtener el modelo restringido, esto es, verificando la restricción, y esta restricción nos mantendrá, por su linealidad y por afectar a un parámetro de posición, dentro de la misma familia conjugada de distribuciones.

Las restricciones [1] y [2] expresan relaciones lineales entre los elementos de x . Cada restricción lineal independiente de las demás elimina una variable, esto es, aunque disponemos de nRS variables (los elementos de x) sólo son “libres” $3N(R-1)S$ de ellas. Nuestra propuesta consiste en eliminar los cuartos trimestres de cada año (que se determinan por la restricción anual [1]) y los valores de una de las provincias (la R -ésima, que se determinará por la restricción regional [2]).

En concreto, si llamamos $x_r = (x_{11}^r, \dots, x_{R-1,1}^r, \dots, x_{1S}^r, \dots, x_{R-1,S}^r)'$ a la matriz columna con los valores libres (no afectados por las restricciones), donde para

$i = 1, \dots, R-1, j = 1, \dots, S$ x_{ij}^r es la matriz columna con sólo los tres primeros trimestres de cada año, se comprueba que

$$x = I_S \otimes \left\{ \begin{pmatrix} I_{R-1} \\ -I_{R-1}' \end{pmatrix} \otimes [I_N \otimes A] \right\} \cdot x_r + I_S \otimes \left\{ \begin{pmatrix} I_{R-1} | 0 \\ -I_{R-1}' | 1 \end{pmatrix} \otimes I_n \right\} \cdot b, \quad [5]$$

donde $b = (x_{11}^{4a}', \dots, x_{R-1,1}^{4a}', b_1', x_{12}^{4a}', \dots, x_{R-1,2}^{4a}', b_2', \dots, x_{1,S}^{4a}', \dots, x_{R-1,S}^{4a}', b_S')'$,

siendo $x_{ij}^{4a} = (0, 0, 0, y_{ij1}, 0, 0, 0, y_{ij2}, \dots, 0, 0, 0, y_{ijN})'$, $i = 1, \dots, R-1, j = 1, \dots, S$ la matriz columna trimestral $n \times 1$ que contiene el correspondiente valor de la serie anual en la posición del cuarto trimestre de cada año, y siendo $A = (I_3, -I_3)'$.

Nótese que la expresión [5] es una relación lineal que proporciona x en función de x_r . En primer lugar la utilizaremos para obtener la distribución restringida, y una vez obtenida la estimación óptima de Bayes para x_r , esta misma expresión nos proporcionará la solución final para x y su dispersión a posteriori.

4. Distribuciones marginales a posteriori. Solución de Bayes

La obtención de las distribuciones a posteriori restringidas se deduce con pequeñas variaciones de las construcciones que realizamos en Rojo y Sanz (2005) ya que, salvo los diferentes significados de las matrices implicadas, en parte debidas a que nuestro problema no es la trimestralización de una única serie con restricciones anuales, sino de RS series con restricciones anuales y regionales, las expresiones analíticas obtenidas son completamente equivalentes. Sólo detallaremos, por tanto, los resultados, describiendo la obtención tanto de la distribución marginal a posteriori restringida para x_r , como de las distribuciones marginales a posteriori restringidas para δ y τ .

Integrando la posteriori conjunta [4] con respecto a τ y a δ , obtenemos la distribución no restringida a posteriori para x ,

$$p(x | D) \propto \left[1 - (x - M^{-1}N)' \frac{M}{B} (x - M^{-1}N) \right]^{-(2m+2\alpha)/2}$$

donde $M = P + [I_{RS} \otimes D'D] + P_\varepsilon - P_\varepsilon Z P_{\delta\varepsilon}^{-1} Z' P_\varepsilon$, $N = P_\varepsilon Z P_{\delta\varepsilon}^{-1} P_\delta \bar{\delta} + P\mu$ y $B = T - N'M^{-1}N$, siendo $P_{\delta\varepsilon} = P_\delta + Z' P_\varepsilon Z$ y $T = 2\beta + \mu' P\mu + \bar{\delta}' P_\delta \bar{\delta} - \bar{\delta}' P_\delta P_{\delta\varepsilon}^{-1} P_\delta \bar{\delta}$, que es una distribución Multivariante de Student² (MS-t, desde ahora) con $m + 2\alpha$ grados de libertad, matriz de posición $M^{-1}N$ y matriz de escala $(m + 2\alpha)M / B$. Por tanto, la matriz de covarianzas es $V_x = \frac{B}{m + 2\alpha - 2} M^{-1}$.

La distribución restringida (que lo será de x_r) se obtiene sustituyendo x por el resultado de la expresión [5]. Análogamente a la expresión [8] de Rojo y Sanz (2005), se trata de una función lineal, por lo que el resultado será el equivalente a la expresión [26] de dicho trabajo de los autores, con las necesarias adaptaciones notacionales. En concreto, la distribución a posteriori de x_r (o dicho de otra forma, de x con las restricciones anuales y regionales) resulta

$$p(x_r | D) \propto \left[1 - (x_r - A_r)' \frac{W' MW}{B_r} (x_r - A_r) \right]^{-(m+\alpha)}$$

siendo $A_r = (W' MW)^{-1} W' M (M^{-1} N - y_v)$ y

$$B_r = B + (M^{-1} N - y_v)' M \left[M^{-1} - W (W' MW)^{-1} W' \right] M (M^{-1} N - y_v)$$

donde $y_v = I_S \otimes \left\{ \left(\begin{array}{c|c} I_{R-1} & 0 \\ \hline -I_{R-1} & 1 \end{array} \right) \otimes I_n \right\} \cdot b$ y $W = I_S \otimes \left\{ \left(\begin{array}{c} I_{R-1} \\ \hline -I_{R-1} \end{array} \right) \otimes [I_N \otimes A] \right\}$.

Se trata de una distribución MS-t de dimensión $3N(R-1)S$ (la dimensión de x_r) y, por tanto, de grados de libertad $(5R+3)NS + 2\alpha$, matriz de posición A_r y escala

$((5R+3)NS + 2\alpha) \frac{W' MW}{B_r}$, con lo que su matriz de covarianzas es,

$$V_{x_r} = \frac{B_r}{(5R+3)NS + 2\alpha - 2} (W' MW)^{-1}.$$

² Véase Zellner (1971) para la versión concreta de las distribuciones de Student utilizadas en este trabajo.

En cuanto a la obtención de las distribuciones marginales a posteriori restringidas para δ y τ , el punto de partida no puede ser su distribución sin restringir, porque entonces no podríamos insertar las restricciones [5] en el modelo. Partiendo de la distribución a posteriori de $(x, \delta, \tau | D)$, expresión [4], obtenemos la distribución de $(x_r, \delta, \tau | D)$, esto es, introducimos las restricciones en el modelo. Integrando dicha distribución en x_r , obtenemos la posteriori restringida marginal de δ y τ , dado D , de la que se obtienen las dos distribuciones a posteriori marginales restringidas de δ y de τ . Los resultados son los siguientes:

La distribución marginal a posteriori restringida de τ es

$$p(\tau | D) \propto \tau^{nRS + \alpha - 1 - 3N(R-1)S/2} \cdot \exp\left\{-\frac{\tau}{2} B_\delta\right\}$$

donde

$$B_\delta = 2\beta + y_v' [I_{RS} \otimes D' D] y_v + \bar{\delta}' P_\delta \bar{\delta} + (\mu - y_v)' P (\mu - y_v) + y_v' P_\varepsilon y_v - (P\mu - P_r y_v)' Q_r (P\mu - P_r y_v)$$

siendo $P_r = P + [I_{RS} \otimes D' D] + P_\varepsilon$ y $Q_r = W(W' P_r W)^{-1} W'$. Se trata, por tanto, de una distribución gamma³ $\gamma\left(\frac{B_\delta}{2}, nRS + \alpha - \frac{3N(R-1)S}{2}\right)$.

Para δ , su distribución marginal a posteriori restringida es

$$p(\delta | D) \propto [B_\delta + (\delta - A_\delta)' M_\delta (\delta - A_\delta)]^{-\frac{K \cdot (5R+3)NS + 2\alpha}{2}}$$

siendo $M_\delta = P_\delta + Z' P_\varepsilon (P_\varepsilon^{-1} - Q_r) P_\varepsilon Z$ y $A_\delta = M_\delta^{-1} [P_\delta \bar{\delta} + Z' P_\varepsilon (y_v + Q_r (P\mu - P_r y_v))]$.

En definitiva, se trata de una MS-t, con $(5R+3)NS + 2\alpha$ grados de libertad, matriz de posición A_δ y escala $[(5R+3)NS + 2\alpha] M_\delta / B_\delta$, con lo que su matriz de covarianzas

$$\text{resulta } V_\delta = \frac{B_\delta}{(5R+3)NS + 2\alpha - 2} M_\delta^{-1}$$

Solución de Bayes

Sabemos que el estimador óptimo de Bayes para una función de pérdida cuadrática es la esperanza de la distribución a posteriori. Es inmediato, entonces, que la solución, \hat{x}_r , para el vector restringido, es

$$\hat{x}_r = E_{post} [x_r] = (W' MW)^{-1} W' M (M^{-1} N - y_v)$$

y una vez obtenida la solución \hat{x}_r obtenemos la solución para x , $\hat{x} = W \hat{x}_r + y_v$

Para el resto de parámetros, las soluciones siguen el mismo camino. Así, la solución para δ será $\hat{\delta} = E_{post} [\delta] = A_\delta$ y, para τ , su esperanza a posteriori,

$$\hat{\tau} = E_{post} [\tau] = \frac{5NRS + 2\alpha + 3NS}{B_\delta}$$

5. Asignación de valores de los parámetros de las distribuciones a priori

Hemos señalado en la introducción que nos alejaremos de las propuestas que sugieren la utilización de distribuciones no informativas, o neutrales en el sentido de Jeffreys. Se trata de distribuciones sobre-dispersas que únicamente son válidas cuando ninguna otra asignación es posible o, cuando la verosimilitud domina a las distribuciones a priori (por ejemplo, en procesos ligados a la extracción de muestras numerosas). En las formulaciones bayesianas ligadas a modelos macroeconómicos, es frecuente que la dificultad (o imposibilidad) de obtener muestras de tamaño importante conduzca a que la sobredispersión asociada a las distribuciones neutrales a priori se traslade a las posteriori y, en consecuencia, a todo el proceso. Ello es más dramático si cabe cuando se simula la distribución a posteriori mediante muestreo de Gibbs para estimar la esperanza de dicha posteriori, ya que la dispersión de las simulaciones crea efectos extremos y la simulación con frecuencia colapsa. Los autores hemos tenido ocasión de observar estas consecuencias en ensayos previos, en los que suponíamos distribuciones de Laplace en lugar de normales.

³Véase el significado de los parámetros de la distribución gamma utilizado por los autores, y que pusimos de manifiesto al asignar a τ la distribución marginal a priori.

La utilización de distribuciones no neutrales nos traslada, no obstante, a otro problema, a saber, la asignación de valores razonables para los parámetros o hiperparámetros implicados en las mismas. Como ya sugerimos en Rojo y Sanz (2005) y empleado anteriormente en otro contexto en Rojo y Sanz (2004), nuestra propuesta se relaciona con las denominadas *g-priors* propuestas por Zellner (véase Zellner (1986)), al menos desde la idea de que el pasado nos proporciona una idea a priori que puede estudiarse en la estimación del comportamiento presente.

Previamente, es necesario disponer de las distribuciones marginales a priori, puesto que la utilización de las distribuciones condicionadas no proporciona soluciones aceptables. Siguiendo Rojo y Sanz (2005) reseñamos los resultados con la notación actual:

Distribución marginal a priori de τ : (expresión [3]) era una distribución gamma, $\gamma(\beta, \alpha)$. Los momentos de τ^{-1} , varianza de alguna de las distribuciones propuestas,

resultan $E[\tau^{-1}] = \frac{\beta}{\alpha - 1}$, ($\alpha > 1$) y $Var[\tau^{-1}] = \frac{\beta^2}{(\alpha - 1)^2(\alpha - 2)}$, ($\alpha > 2$), de donde

obtenemos

$$\alpha = 2 + \frac{(E[\tau^{-1}])^2}{Var[\tau^{-1}]} \quad \text{y} \quad \beta = (\alpha - 1)E[\tau^{-1}].$$

Distribución marginal a priori de δ : Integrando en τ la priori de δ y τ , obtenemos

$$\pi(\delta | D) \propto \left[2\beta + (\delta - \bar{\delta})' P_{\delta} (\delta - \bar{\delta}) \right]^{-\frac{K+2\alpha}{2}},$$

una MS-t de dimensión K con 2α grados de libertad, matriz de posición $\bar{\delta}$ y matriz de precisión $\frac{\alpha}{\beta} P_{\delta}$, por lo que su matriz de covarianzas resulta $Cov_{\pi}(\delta) = \frac{\beta}{\alpha - 1} P_{\delta}^{-1}$.

Distribución marginal a priori de ε : Integrando en τ la priori conjunta de ε y τ (obtenida a partir de la marginal de τ y de la de ε condicionada por τ), obtenemos

$$\pi(\varepsilon | D) \propto \left[2\beta + \varepsilon' P_{\varepsilon} \varepsilon \right]^{-\frac{m+2\alpha}{2}}$$

Es, obviamente, una MS-t con 2α grados de libertad, matriz de posición igual a la matriz nula de dimensión m y matriz de precisión $\frac{\alpha}{\beta}P_\varepsilon$, de donde su matriz de varianzas y covarianzas resulta $\text{Cov}_\pi(\varepsilon) = \frac{\beta}{\alpha-1}P_\varepsilon^{-1}$

De la lectura de las distribuciones anteriores podemos concluir que los parámetros implicados son: α y β (que intervienen en la distribución marginal a priori de τ), el vector $\bar{\delta}$ y la matriz P_δ que permiten concretar la distribución a priori de δ , el vector μ y la matriz P implicados en la distribución de x y finalmente P_ε , implicada en la verosimilitud.

Supondremos que $P = I_{RS}$, esto es, que los valores de las series trimestrales son independientes a priori. El sistema no dispone de los suficientes grados de libertad como para asignar valores a los elementos de dicha matriz, salvo que supusiéramos un modelo de comportamiento que redujera drásticamente el número de sus parámetros.

El procedimiento empleado para la asignación se parece al empleado en Rojo y Sanz (2005) si bien presenta particularidades notables en algunos episodios.

Inicialmente obtenemos una estimación sin indicadores de las series x , estimación que almacenamos en una matriz columna \tilde{x} . En concreto, se trimestralizan las RS series mediante el procedimiento de Denton sin indicadores o mediante el de Boot-Feibes y Lisman. Se asigna entonces $\mu = \tilde{x}$.

Se divide entonces el periodo temporal en estudio en H segmentos o *paquetes*, y llamamos $\tilde{x}_1, \dots, \tilde{x}_H$ a la división de la serie trimestral provisional \tilde{x} obtenida. Obsérvese que, a diferencia del trabajo anterior de los autores, no se trata de dividir \tilde{x} en segmentos, ya que se dividen los periodos temporales y, en consecuencia, cada serie.

Tomamos entonces Z_h , $h=1, \dots, H$ la misma división temporal de la matriz de indicadores en segmentos. Obsérvese que el número de filas de Z es nRS y el de Z_h es aproximadamente nRS/H por lo que no son previsible problemas de rango en dicha

matriz, aunque el número de años, N , no sea grande (véase en este sentido la aplicación práctica propuesta, en la que $N = 5$ y no existen problemas de rango en las regresiones).

Se obtienen entonces H observaciones de la matriz δ , $\hat{\delta}_h = (Z_h' Z_h)^{-1} Z_h' x_h$, $h = 1, \dots, H$ y otras tantas de τ^{-1} , $\tau_h^{-1} = \frac{1}{T-K} [x_h' x_h - \hat{\delta}_h' (Z_h' Z_h) \hat{\delta}_h]$, $h = 1, \dots, H$ donde T es el tamaño de cada lote y K el número total de indicadores (incluyendo, en su caso, los términos constantes).

De aquí, obtenemos los valores de $\alpha = 2 + \frac{(\overline{\tau^{-1}})^2}{S_{\tau^{-1}}^2}$ y de $\beta = (\alpha - 1) \overline{\tau^{-1}}$, donde $\overline{\tau^{-1}}$ y

$S_{\tau^{-1}}^2$ son la media y cuasi-varianza de las observaciones τ_h^{-1} , $h = 1, \dots, H$.

Por otro lado, se asignará $\bar{\delta}$ como la media aritmética de las asignaciones $\hat{\delta}_1, \dots, \hat{\delta}_H$. De estas estimaciones, se obtiene adicionalmente una estimación de $\text{Cov}_\pi(\delta)$ y de ésta, la de P_δ . En efecto,

$$P_\delta = \frac{\beta}{\alpha - 1} \text{Cov}_\pi^{-1}(\delta)$$

Se supone que, a priori, están incorrelacionados los parámetros de la matriz δ correspondientes a distintos modelos, con lo que la matriz $\text{Cov}_\pi^{-1}(\delta)$ será diagonal por bloques, disponiendo de RS bloques y siendo el tamaño de cada bloque igual al número de parámetros implicados en el modelo referente a la serie correspondiente.

Faltaría por decidir la asignación de valores a priori a la matriz P_ε . Como

$\text{Cov}_\pi(\varepsilon) = \frac{\beta}{\alpha - 1} P_\varepsilon^{-1}$, entonces $P_\varepsilon = \frac{\beta}{\alpha - 1} \text{Cov}_\pi^{-1}(\varepsilon)$. En consecuencia, P_ε dependerá de

los valores que especifiquemos sobre $\text{Cov}_\pi(\varepsilon)$. Aunque resulta tradicional suponer una estructura AR(1) o de ruido blanco en el comportamiento de las perturbaciones del modelo de indicadores, nuestra matriz $\text{Cov}_\pi(\varepsilon)$ contiene varianzas y covarianzas para las perturbaciones de los RS modelos. Se supondrá a priori incorrelación entre las

perturbaciones de distintos modelos, de manera que se realizarán asignaciones en las matrices de varianzas y covarianzas de las perturbaciones de los modelos individuales. Dicho de otra forma, $Cov_{\pi}(\varepsilon)$ se toma diagonal por bloques, existiendo RS bloques cuadrados de dimensión n , que denominaremos $Cov_{\pi}(\varepsilon_{ij})$, $i=1,\dots,R$, $j=1,\dots,S$. La utilización de modelos diferenciados sólo obliga a analizar individualmente cada modelo, lo que dificulta el trabajo cuando R o S son grandes, si bien no lo complica formalmente. En este trabajo supondremos el mismo tipo de modelo para cada una de las series. Así,

- Si suponemos que, a priori, las perturbaciones de cada modelo siguen un ruido blanco⁴ tomaremos $Cov_{\pi}(\varepsilon_{ij}) = \sigma_{ij}^{-2} I_n$, $i=1,\dots,R$, $j=1,\dots,S$.
- Si suponemos un esquema AR(1), debe tomarse⁵

$$Cov_{\pi}^{-1}(\varepsilon_{ij}) = \sigma_{ij}^{-2} \begin{pmatrix} 1 & -\rho_{ij} & 0 & \cdots & 0 & 0 & 0 \\ -\rho_{ij} & 1 + \rho_{ij}^2 & -\rho_{ij} & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & -\rho_{ij} & 1 + \rho_{ij}^2 & -\rho_{ij} \\ 0 & 0 & 0 & \cdots & 0 & -\rho_{ij} & 1 \end{pmatrix}$$

donde ρ_{ij} se asignará a priori como el coeficiente de autocorrelación de los residuos MCO de la regresión de \tilde{x}_{ij} (recuérdese, la estimación BFL de la serie trimestral) sobre los indicadores Z_{ij} , y σ_{ij}^2 como la varianza de los $m-1$ transformados de los residuos $\varepsilon_{ij} - \rho_{ij} \cdot \varepsilon_{i,j-1}$ que siguen un ruido blanco.

- Si se supone un esquema de camino aleatorio, se realiza la asignación construyendo los residuos MCO de la regresión de \tilde{x}_{ij} sobre los indicadores Z_{ij} como para el esquema AR(1) y asignando ahora σ_{ij}^2 como la varianza de los $m-1$ transformados de los residuos $\varepsilon_{ij} - \varepsilon_{i,j-1}$ que siguen un ruido blanco. Se toma entonces

4 Esta hipótesis es insostenible para el método de Chow-Lin ya que, recuérdese, los residuos de la regresión anual del modelo de indicadores se repartían en partes anuales entre los cuatro trimestres, produciendo saltos entre el cuarto trimestre y el primero. Para nosotros no es inaceptable, porque no produce dichos efectos, véase Rojo y Sanz (2005).

5 Véase, p. ej., Dhrymes (1984), pág. 113.

$$Cov_{\pi}^{-1}(\varepsilon_{ij}) = \sigma_{ij}^{-2} \begin{pmatrix} 2 & -1 & 0 & \cdots & 0 & 0 & 0 \\ -1 & 2 & -1 & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & -1 & 2 & -1 \\ 0 & 0 & 0 & \cdots & 0 & -1 & 1 \end{pmatrix}$$

Otros esquemas son posibles (por ejemplo, el de Litterman, véase Litterman (1983)) y se desarrollan con procedimientos análogos.

6. Una propuesta de aplicación

En este epígrafe utilizaremos el procedimiento propuesto para constatar la viabilidad de su implementación práctica así como para mostrar cómo puede obtenerse una Contabilidad provincial congruente con la Contabilidad Regional.

El primero de los datos básicos es el agregado regional trimestral. Muchos Institutos estadísticos regionales proporcionan dicho agregado, desagregado por ramas de actividad. Para Castilla y León, podría obtenerse encadenando crecimientos en volumen provenientes de la D.G. de Estadística regional, aunque a efectos de ilustración hemos tomado una trimestralización sin indicadores de la Contabilidad Regional (anual) del INE en volumen, obteniéndose a través de los índices de volumen encadenados al valor del año 2000.

Para que esta información sea congruente con la anual provincial, la desagregación por ramas debe ser asimismo congruente con la existente en el ámbito provincial. Tomamos, en consecuencia, cuatro ramas, Agricultura, Industria, Construcción y Servicios Totales, y añadimos una quinta “rama”, Impuestos Netos sobre los productos, de forma que la suma coincida con el PIBpm provincial o regional según los ámbitos.

En el ámbito provincial se dispondrá de los valores añadidos en euros de 2000 a precios básicos anuales, tomados de la Contabilidad Regional del I.N.E. Han sido deflacionados bajo la hipótesis de que los deflatores regionales son aproximaciones suficientemente adecuadas de los provinciales (una hipótesis, obviamente, simplificadora, debido al elevado nivel de agregación de las ramas (se ha realizado a 7 ramas de actividad) pero insustituible mientras no se disponga de deflatores provinciales.)

Nuestro procedimiento se alimenta de una solución inicial, la trimestralización (no armonizada) de los valores anuales provinciales por ramas a precios básicos (en euros de 2000). Aunque los autores disponemos de un procedimiento propio de trimestralización individual de series (el desarrollado en Rojo y Sanz (2005)) no lo hemos utilizado para que no existieran sospechas de acuerdo metodológico entre ambos procedimientos y se juzgase sólo la viabilidad del ahora presentado. Hemos utilizado, en consecuencia, ECOTRIM (versión 1.01), implementado para la Comisión Europea por Roberto Barcellan y Tommaso Di Fonzo, y dentro de él, el procedimiento de Denton sin indicadores, y con diferencias de primer orden.

Por otro lado, nuestro procedimiento utiliza indicadores. En el ejercicio que planteamos tendremos un único indicador para cada provincia, rama y trimestre, si bien este planteamiento es provisional. Como ya hemos señalado, el procedimiento puede incorporar más indicadores, e incluso distinto número de ellos para cada provincia y/o rama. En esta ilustración, el indicador que utilizaremos será (por homogeneidad) el empleo (en miles de ocupados) en cada una de las nueve provincias y cada una de los cuatro sectores: Agricultura, Industria, Construcción y Servicios, proporcionados por el I.N.E. en serie homogénea para el periodo de análisis (obsérvese que, al no emplear el año 2005, no se nos plantea la necesidad de corregir el cambio metodológico). Para los Impuestos Netos sobre los productos, el indicador utilizado es el valor añadido a precios básicos total provincial (sin impuestos, obviamente).

Los resultados de la ilustración se proporcionan en el Anexo. Si se observan las variaciones de las series se constatará que el grado de “regularidad” alcanzado a veces no es muy alto. Ello ocurre porque se está estimando el valor añadido provincial por ramas sin desestacionalizar, esto es, en serie bruta. Se han utilizado, por tanto, indicadores en serie bruta. Los autores hemos realizado una trimestralización utilizando la serie desestacionalizada mediante el procedimiento X12-ARIMA en su versión implementada por Eviews, constatándose la mayor regularidad (menor variabilidad) de los resultados. El grado de regularidad de ambas soluciones (desestacionalizada o en serie bruta) se ha comparado utilizando como medida de regularidad

$FD_{ij} = \frac{1}{n-1} \sum_{t=2}^n |\tilde{x}_{ij,t} - \tilde{x}_{ij,t-1}|$ para cada serie con cada procedimiento. Se constata que

efectivamente, el uso de series desestacionalizadas proporciona trimestralizaciones más regulares, como era de esperar. Obviamente, se alcanzaría una mayor regularidad acudiendo al procedimiento de segundas diferencias en el establecimiento de la distribución a priori de la serie trimestral.

Para la obtención de los hiperparámetros implicados en las distribuciones a priori, elegimos $H = 5$ lotes, correspondientes a cada uno de los años 2000 a 2004. Por tanto, cada lote tiene 4 trimestres.

El número de indicadores (columnas de la matriz Z) es $K = 2 \cdot 45 = 90$ indicadores, puesto que los modelos tienen término constante.

Se opta por una hipótesis de ruido blanco a priori para los modelos de indicadores. Obviamente, la distribución a posteriori no es un ruido blanco. Como señalamos en la formulación teórica, la matriz P se considera igual a la identidad (de orden 720).

La inversión de las matrices P_ε , P_δ y $Z'Z$ no se realiza sobre su dimensión total (la primera es de dimensión 720 y las dos restantes de dimensión 90) sino que, como son diagonales por bloques, se invierte cada bloque separadamente. Ello tiene como consecuencia que ciertas casi-singularidades derivadas de multicolinealidades imperfectas no se presenten. Adicionalmente, la matriz $P_{\delta\varepsilon}$, que se invierte en el procedimiento, es también diagonal por bloques, por lo que se actúa de forma análoga.

7. Conclusiones

En este trabajo se desarrolla un procedimiento original para conciliar espacial y sectorialmente series trimestrales. Dicha conciliación se realiza mediante un modelo conjunto bayesiano normal-gamma, que yuxtapone de los modelos provinciales y regional, con una estructura de covarianzas adecuada. Los indicadores utilizados no son aproximaciones, sino que se toman indicadores de comportamiento.

Las distribuciones a priori empleadas son informativas (no neutrales, en consecuencia) y el modelo global utilizará las distribuciones a posteriori, restringidas, para las series trimestrales, en la línea de Rojo y Sanz (2005).

El modelo estimado obtiene la solución de Bayes para las series trimestrales para cada desagregación espacial y sectorial, así como medidas de dispersión de las series estimadas.

La solución se obtiene en un paso (no es iterativa) y permite utilizar distintos esquemas para la función de verosimilitud. Supone conocidos, para cada una de las ramas de actividad, los valores anuales de las series desagregadas geográficamente (provinciales) y los valores trimestrales de las series regionales, de manera que la solución armoniza con todas esas restricciones. Nuestro trabajo demuestra que la solución puede reducirse a la propuesta para una única serie de Rojo y Sanz (2005) facilitando su desarrollo conceptual y suministrando la solución óptima para una función de pérdida cuadrática.

La ventaja de la propuesta bayesiana es que, en general, resulta más adecuada para resolver situaciones en las que la información disponible es corta, lo que ocurre con frecuencia en las Contabilidades para áreas pequeñas como son nuestras provincias.

El procedimiento se amplía sin problemas a otras frecuencias, aunque a lo largo del trabajo se habla de trimestralización de datos anuales por simplificar la descripción. Basta cambiar el número de periodos de cada periodo agregado (el 4 utilizado por un 12, si se trata de mensualizar) sin que sea necesaria ninguna otra modificación.

Tampoco altera el procedimiento la diversidad de indicadores por serie (que en la aplicación se ha tomado como 2 para cualquiera de las 45 series, al asumir un modelo de verosimilitud con término constante)

Obsérvese que nuestro procedimiento permite suponer un esquema de ruido blanco para la verosimilitud, posibilidad que está vetada a procedimientos clásicos, por presentarse el problema del escalón entre cuartos y primeros trimestres.

Ciertamente, puede suponerse desconocido el agregado trimestral regional. Sin embargo, ese problema ya ha sido resuelto por los autores debido a que entonces sólo se aplica la restricción “anual”, por lo que la solución se reduce a 45 trimestralizaciones separadas.

Los autores se plantean hacia el futuro la comparación de su método con procedimientos más clásicos, y el estudio de las dispersiones (bayesianas y clásicas) del estimador obtenido. Otro objeto de interés reside en la automatización del procedimiento implementado, de manera que se pueda utilizar automáticamente para cualquier opción temporal y cualquier esquema para los errores del modelo de verosimilitud.

El trabajo se ilustra con una aplicación a la elaboración de Contabilidades Trimestrales provinciales para las 9 provincias castellano-leonesas. Esta propuesta utiliza la Contabilidad Regional (anual) del I.N.E. y supone conocida la Contabilidad Trimestral de Castilla y León (CTCYL) Base 2000 (adaptada a los nuevos cambios metodológicos contables). El propósito de esta ilustración es, por un lado, demostrar que el procedimiento es factible y, por otro, desarrollar la implementación del procedimiento con la herramienta estadística adecuada, en este caso SPLUS.

Bibliografía

Boot, J.C.G. y Feibes, W. (1967): "On Glejser's derivation of monthly figures from yearly data". *Cahiers Économiques de Bruxelles*, 36, 539-546.

Box, G.E.P. y Tiao, G.C. (1973): *Bayesian Inference in Statistical Analysis* Addison Wesley, Reading, Mass.

Chen, Z.G. y Dagum, E.B. (1997): "A recursive method for predicting variables with temporal and contemporaneous constraints" *American Statistical Association*, Proceedings of the Business and Economic Statistics Section.

Cholette, P.A. (1984): "Adjusting sub-annual series to yearly benchmarks". *Survey Methodology* 10(1) 35-49.

Chow, G. y Lin, A.L. (1971): "Best Linear Unbiased Interpolation, Distribution and Extrapolation of Time Series by related series", *The Review of Economics and Statistics*, 53(4), 372-375.

Denton, F.T. (1971): "Adjustment of monthly or quarterly series to annual totals: an approach based on quadratic minimization", *Journal of the American Statistical Association*, 66(333), 99-102.

Dhrymes, Ph.J. (1984): *Econometría*. Ed. AC, Madrid.

Di Fonzo T. (2002): "Temporal disaggregation of a system of time series when the aggregate is known. Optimal vs. adjustment methods" *Workshop on Quarterly National Accounts*, Eurostat, Theme 2 Economy and Finance, pp. 63-77

Di Fonzo T. (2003): "Temporal Disaggregation of Economic Time Series: towards a dynamic extension". *European Commission, Working Papers & Studies*.

Di Fonzo, T. y Marini, M. (2003): “Benchmarking systems of seasonally adjusted time series according to Denton's movement preservation principle”. Università di Padova, Dipartimento di Scienze Statistiche, Working paper 2003.9 (disponible en http://homes.stat.unipd.it/difonzo/public_file/wp_2003_09.pdf).

Di Fonzo, T. y Marini, M. (2005): “Benchmarking a system of time series: Denton's movement preservation principle vs. a data based procedure”. *Working Papers and Studies*. European Commission, Luxembourg.

Eurostat (1999): *Handbook of Quarterly National Accounts*. Luxembourg: European Commission.

Fernández, R.B. (1981): “A Methodological note on the estimation of Time Series”, *The Review of Economics and Statistics*, 63(3), 471-478.

Friedman, M. (1962): “The Interpolation of Time Series by related Series”. *Journal of the American Statistical Association*, 57 (diciembre), 729-757.

Jerez, M., Casals, J. y Sotoca, S. (2005): “Empirical modeling of time series sampled at different frequencies”. *Working Papers and Studies*. European Commission, Luxembourg.

Laniel, N. y Fyfe, K. (1989): “Benchmarking of economic time series” *American Statistical Association Proceedings of the Business and Economic Statistics Section*, pp. 462-466.

Litterman, R.B. (1983): “A Random Walk, Markov Model for the Distribution of Time Series” *Journal of Business and Economic Statistics*, Vol. 1, nº 2, pp. 169-173.

Palate, J. (2005): “Reusable components for benchmarking using Kalman filters”. *Working Papers and Studies*. European Commission, Luxembourg.

Pavía, J.M. (1997): “Trimestralización de series anuales”. *Universidad de Valencia, Tesis doctoral*.

Proietti, T. (2005): “Temporal disaggregation by state space methods: dynamic regression methods revisited”. *Working Papers and Studies*. European Commission, Luxembourg.

Quenneville, B. y Rancourt, E. (2005): “Simple methods to restore the additivity of a system of time series”. *Working Papers and Studies*. European Commission, Luxembourg.

Rodríguez Feijóo, S., Rodríguez Caro, A., Dávila Quintana, D. (2003): “Methods for quarterly disaggregation without indicators: A comparative study using simulation” *Computational Statistics & Data Analysis* 43, pp.63-78.

Rojo, J.L. y Sanz, J.A. (1999): “Una propuesta bayesiana para la distribución de Contabilidades regionales por procedimientos indirectos”, en el libro *Cambios Regionales en la U.E. y Nuevos Retos Territoriales*, AEER, ISBN: 84-607-3322-X, Madrid, 2001, pp. 1-19.

Rojo, J.L. y Sanz, J.A. (2000): “Estimaciones de magnitudes en pequeñas areas” Trabajo subvencionado por la Dirección General de Estadística de la Junta de Castilla y León (no publicado).

- Rojo, J.L. y Sanz, J.A. (2002): “Estimaciones para pequeñas áreas: un enfoque bayesiano a un problema de distribución”. *Estudios de Economía aplicada*. 20-I, pp. 217-240.
- Rojo, J.L. y Sanz, J.A. (2004): “Desagregación espacial para pequeñas áreas: un modelo bayesiano normal-gamma”. *Estudios de Economía aplicada*. 22-III, pp. 487-516.
- Rojo, J.L. y Sanz, J.A. (2005): “A Bayesian benchmarking method with applications to the Quarterly National Accounts”. Luxembourg: Office for Official Publications of the European Communities (ISSN 1725-4825; Cat. No. K3-DT-05-013-EN-N). *European Communities*, 2005.
- Santos Silva, J.M.C. y Cardoso, F.N. (2001): “The Chow-Lin method using dynamic models”. *Economic Modelling*, 18: 269-280.
- Sanz, R. (1981): “Métodos de desagregación temporal de series económicas” *Banco España. Servicio Estudios. Estudios económicos* nº 22.
- Spiegelhalter, D., Thomas, A. y Best, N. (1999): *WinBUGS, V. 1.2: User Manual* MRC Biostatistics Unit, Cambridge, U.K.
- Zellner, A. (1971): *An Introduction to Bayesian Inference in Econometrics*, J. Wiley & Sons, Inc. N. York.
- Zellner, A. (1986): “On assessing prior distributions and bayesian regression with g-prior distributions”, en *Bayesian Inference and Decision Techniques* Elsevier Science Publishers, B.V., N. York.

Anexo: Estimaciones finales (armonizadas)

ÁVILA

	2000-I	2000-II	2000-III	2000-IV	2001-I	2001-II	2001-III	2001-IV	2002-I	2002-II	2002-III	2002-IV	2003-I	2003-II	2003-III	2003-IV	2004-I	2004-II	2004-III	2004-IV	
Unidad: miles de euros																					
Agricultura, ganadería y pesca	44.410,1	45.275,9	43.872,0	43.902,1	41.285,7	39.972,4	39.822,0	39.996,8	47.038,1	44.694,2	42.216,5	41.974,2	41.650,5	39.887,8	40.970,6	44.579,4	40.538,6	40.058,9	40.357,2	37.251,2	
Industria	89.294,1	76.531,2	76.726,3	71.887,3	69.970,1	71.175,9	79.544,2	69.396,4	84.310,5	72.534,8	72.763,8	78.125,6	77.165,6	65.373,3	61.903,4	68.463,5	61.649,0	63.993,9	62.023,5	62.766,1	
Construcción	51.865,4	47.052,6	49.418,8	48.756,2	47.463,6	48.974,6	44.862,6	38.474,6	38.232,3	38.071,5	50.228,8	62.831,9	61.632,6	42.216,8	61.318,0	72.007,1	76.202,4	68.389,7	70.027,4	100.223,9	
Servicios	288.979,7	290.070,9	289.973,3	291.259,0	297.794,0	306.697,4	286.277,9	294.581,7	314.873,1	314.089,9	280.677,1	314.446,6	308.296,7	331.981,0	327.799,1	341.258,4	378.370,8	380.730,0	385.816,7	374.796,5	
Impuestos netos sobre los productos	48.517,7	48.371,1	48.077,7	47.837,6	48.403,0	47.960,3	47.661,7	47.504,7	48.925,5	49.378,6	50.294,9	51.674,7	53.088,2	55.311,7	57.915,5	60.899,8	62.219,2	66.530,3	71.787,9	79.991,6	
Píbm	523.037	507.702	501.068	501.442	504.916	514.681	498.168	510.437	533.389	518.715	496.181	549.063	542.125	562.937	569.907	607.698	648.180	660.336	658.375	702.990	

BURGOS

	2000-I	2000-II	2000-III	2000-IV	2001-I	2001-II	2001-III	2001-IV	2002-I	2002-II	2002-III	2002-IV	2003-I	2003-II	2003-III	2003-IV	2004-I	2004-II	2004-III	2004-IV	
Unidad: miles de euros																					
Agricultura, ganadería y pesca	116.025,7	114.450,0	114.694,6	112.293,7	107.382,6	100.009,1	89.682,2	84.103,4	71.693,0	75.569,6	82.962,2	84.721,1	80.167,6	80.503,4	77.485,7	80.675,1	80.382,2	88.730,7	91.190,8	91.500,8	
Industria	455.976,9	594.477,9	471.510,4	469.805,0	447.145,2	478.719,3	494.234,6	464.293,7	467.210,8	491.697,9	492.665,1	464.348,4	519.437,1	488.542,1	492.294,1	545.062,9	558.994,9	545.732,6	562.826,4	581.386,3	
Construcción	88.885,1	68.494,0	91.289,9	97.909,0	98.300,1	82.259,0	102.483,8	116.525,1	102.694,1	86.664,2	131.477,1	139.292,1	128.544,6	176.030,0	177.596,7	132.247,4	163.423,0	197.874,7	187.853,2	190.747,0	
Servicios	690.292,2	690.704,6	674.615,5	713.261,5	764.727,7	728.388,7	637.220,3	725.794,0	757.405,7	798.372,3	798.372,3	772.867,8	796.006,8	864.158,0	876.625,0	905.876,0	941.326,9	956.811,1	1.038.345,2	992.347,1	
Impuestos netos sobre los productos	143.535,0	143.409,8	143.159,4	142.783,9	142.303,6	142.439,3	143.211,4	144.619,0	146.386,2	149.484,9	153.738,0	159.065,3	164.528,1	171.664,4	179.475,0	187.960,2	198.464,1	207.126,3	215.276,3	222.391,4	
Píbm	1.494.715	1.521.537	1.495.470	1.530.653	1.559.869	1.531.814	1.455.832	1.535.336	1.545.250	1.580.244	1.659.215	1.620.325	1.686.704	1.780.698	1.803.446	1.851.812	1.951.601	1.977.231	2.093.492	2.079.910	

LEÓN

	2000-I	2000-II	2000-III	2000-IV	2001-I	2001-II	2001-III	2001-IV	2002-I	2002-II	2002-III	2002-IV	2003-I	2003-II	2003-III	2003-IV	2004-I	2004-II	2004-III	2004-IV	
Unidad: miles de euros																					
Agricultura, ganadería y pesca	108.871,0	109.639,6	111.092,9	106.209,4	111.182,4	109.708,4	111.067,6	107.742,5	97.716,2	96.753,5	99.090,9	99.714,2	90.476,3	92.729,2	95.385,0	90.707,4	92.955,7	88.944,6	88.246,6	81.963,4	
Industria	333.670,6	320.458,5	326.194,1	329.323,8	337.240,3	333.207,9	340.796,4	344.883,0	350.203,0	348.372,3	348.721,7	346.703,8	360.381,5	375.326,0	376.484,9	375.620,0	387.579,6	404.853,1	402.317,3	408.554,5	
Construcción	124.335,4	100.146,0	106.935,2	161.538,6	141.133,7	144.218,4	143.325,6	162.158,6	146.230,1	151.325,2	163.974,8	173.837,1	169.839,3	210.833,5	215.966,6	229.162,0	226.785,6	240.867,7	259.856,4	289.736,0	
Servicios	836.893,8	824.152,0	835.238,4	862.273,8	818.115,9	850.552,7	936.219,9	927.313,6	936.863,5	896.430,0	991.544,0	944.796,0	1.036.051,3	901.727,7	1.050.401,1	1.057.537,1	1.047.485,1	1.065.213,4	1.166.407,7	1.310.772,6	
Impuestos netos sobre los productos	146.738,1	147.312,0	146.459,7	150.181,2	150.267,8	152.366,2	154.579,6	160.102,8	161.558,6	163.186,9	165.249,4	167.309,7	171.655,3	173.532,4	178.149,4	187.960,2	194.242,0	204.523,0	213.382,2	221.671,7	
Píbm	1.549.509	1.549.872	1.557.921	1.599.527	1.557.940	1.590.084	1.685.767	1.670.874	1.704.512	1.635.866	1.751.966	1.722.559	1.834.032	1.718.944	1.819.694	1.934.680	1.961.705	1.999.179	2.119.511	2.290.336	

PALENCIA

	2000-I	2000-II	2000-III	2000-IV	2001-I	2001-II	2001-III	2001-IV	2002-I	2002-II	2002-III	2002-IV	2003-I	2003-II	2003-III	2003-IV	2004-I	2004-II	2004-III	2004-IV	
Unidad: miles de euros																					
Agricultura, ganadería y pesca	82.286,3	82.147,1	80.446,4	79.063,2	77.071,0	75.572,5	76.040,3	75.877,3	79.737,0	78.204,6	74.746,5	73.066,4	74.555,9	75.331,4	76.335,3	76.561,7	80.214,4	80.603,3	79.517,9	77.939,8	
Industria	145.661,0	145.570,4	141.536,6	148.790,1	149.235,6	153.128,1	148.403,5	158.503,5	163.437,1	162.661,1	167.439,0	160.670,7	181.546,8	193.147,8	184.603,8	176.198,2	187.740,4	192.883,7	186.809,0	186.809,0	
Construcción	40.006,0	35.193,2	32.777,1	33.451,7	31.928,7	42.262,5	42.468,2	46.898,2	40.716,1	49.306,5	60.912,9	62.442,5	56.869,4	50.600,3	63.304,4	61.955,5	61.454,6	68.959,6	77.481,8	89.736,1	
Servicios	279.748,6	315.721,5	319.988,5	297.545,3	325.081,1	324.318,9	308.361,0	266.886,5	299.257,1	329.508,7	328.003,2	358.571,4	332.851,7	360.082,0	364.509,4	384.272,1	384.483,1	389.429,2	385.734,2	457.851,4	
Impuestos netos sobre los productos	59.351,5	59.258,7	59.073,1	58.794,6	59.078,5	59.256,2	60.102,8	61.558,4	63.186,9	65.249,4	67.309,7	69.367,9	71.655,3	73.532,4	78.149,4	87.960,2	94.242,0	104.523,0	113.382,2	121.671,7	
Píbm	607.053	637.891	633.822	617.645	642.335	654.536	635.376	689.434	652.334	684.530	698.411	733.978	698.693	741.503	773.727	786.543	784.483	822.849	836.779	909.597	

SALAMANCA

	2000-I	2000-II	2000-III	2000-IV	2001-I	2001-II	2001-III	2001-IV	2002-I	2002-II	2002-III	2002-IV	2003-I	2003-II	2003-III	2003-IV	2004-I	2004-II	2004-III	2004-IV	
Unidad: miles de euros																					
Agricultura, ganadería y pesca	78.319,9	78.880,5	73.486,5	75.582,0	79.237,3	77.094,5	78.924,5	78.771,4	75.192,7	73.484,9	74.056,3	72.869,6	74.216,0	68.001,2	69.626,6	72.836,6	70.172,3	68.321,9	68.338,1	69.019,2	
Industria	115.507,1	106.385,8	108.493,9	107.303,2	119.982,8	114.399,5	119.915,7	129.000,6	118.927,7	117.425,1	118.819,3	121.286,3	128.628,4	133.359,4	138.091,1	137.520,3	136.437,3	146.239,5	143.244,2	147.859,6	
Construcción	75.397,4	84.373,2	111.424,0	116.118,4	96.196,6	108.179,2	129.972,5	129.107,1	145.242,3	163.280,3	113.903,1	125.121,3	157.007,1	173.591,4	121.211,6	170.453,9	165.365,6	201.676,0	212.087,8	193.953,2	
Servicios	685.484,7	679.093,0	697.456,5	703.983,8	699.076,9	719.439,5	702.585,3	724.289,4	742.271,2	750.034,5	752.549,5	768.740,1	805.339,2	833.101,7	855.137,9	889.566,6	891.630,2	916.327,7	927.638,1	975.421,7	
Impuestos netos sobre los productos	99.865,0	100.471,4	101.644,1	103.403,4	104.474,0	106.425,7	107.983,4	109.147,1	110.002,5	111.912,2	114.961,9	119.151,7	124.375,5	129.615,7	134.796,5	139.827,8	146.022,2	151.538,7	157.590,7	164.205,1	
Píbm	1.034.594	1.046.214	1.020.505	1.106.391	1.097.968	1.124.538	1.139.381	1.163.316	1.191.538	1.218.105	1.174.290	1.207.149	1.289.564	1.338.266	1.319.652	1.390.215	1.429.655	1.483.194	1.508.906	1.550.356	

SEGOVIA

	2000-I	2000-II	2000-III	2000-IV	2001-I	2001-II	2001-III	2001-IV	2002-I	2002-II	2002-III	2002-IV	2003-I	2003-II	2003-III	2003-IV	2004-I	2004-II	2004-III	2004-IV	
Unidad: miles de euros																					
Agricultura, ganadería y pesca	82.082,8	77.904,6	81.279,3	76.293,3	73.981,7	71.997,7	66.890,1	70.303,5	72.800,7	79.420,0	78.018,3	78.879,9	67.731,4	67.009,6	63.341,7	62.556,4	58.554,0	67.857,3	64.320,2	62.150,3	
Industria	75.629,4	69.200,1	71.588,6	67.847,9	71.480,0	66.747,4	70.584,3	78.124,6	74.419,2	76.057,5	77.395,4	81.157,6	68.293,6	72.531,6	79.476,3	77.803,8	69.212,4	81.556,6	83.386,1	89.596,3	
Construcción																					

AGREGACIÓN DE ESTRUCTURAS DE PREFERENCIAS EN LA E-COGNOCRACIA *

MARÍA TERESA ESCOBAR URMENETA
mescobar@unizar.es

JOSÉ MARÍA MORENO JIMÉNEZ
moreno@unizar.es

Grupo Decisión Multicriterio Zaragoza. Facultad de Económicas
UNIVERSIDAD DE ZARAGOZA

Área temática: Métodos cuantitativos.

Resumen

El trabajo presenta una aplicación del método ‘*Agregación de Estructuras de Preferencia (AIPS)*’, procedimiento propuesto para la toma de decisiones en grupo con el Proceso Analítico Jerárquico (AHP), a un caso real de e-participación abordado desde la perspectiva de la e-cognocracia. En dicho experimento los participantes han emitido sus juicios en dos rondas de opinión diferentes separadas por un proceso de discusión realizado a través de un foro. Esta forma de incorporar las opiniones ha permitido analizar la evolución temporal de las estructuras de preferencia individuales y colectivas. En este sentido, el trabajo extiende a un número elevado de actores, la propuesta anteriormente efectuada por los autores para un número reducido de participantes.

Palabras clave: Decisión en Grupo, Proceso Analítico Jerárquico (AHP), e-participación, e-democracia, e-cognocracia.

Abstract

This paper presents an application of the procedure ‘*Aggregation of Individual Preference Structures (AIPS)*’ which was proposed for group decision making with the Analytic Hierarchy Process (AHP) to a real case of e-participation coped with from the perspective of e-cognocracy. In this experiment, the participants have expressed their opinions in two rounds separated by a discussion process carried in a forum. This way to incorporate the opinions of decision makers has allowed us to analyse the temporal evolution of the individual and collective preference structures. In this sense, the paper extends to a higher number of actors, the proposal previously made by the authors for a small number of participants.

Key words: Group Decision Making, Analytic Hierarchy Process (AHP), e-participation, e-democracy, e-cognocracy.

* Este trabajo ha sido parcialmente financiado por el proyecto de investigación “*Herramientas decisionales para la e-cognocracia*” del Ministerio de Educación y Ciencia (Ref. TSI2005-02511)

1. Introducción

En los últimos años se están llevando a cabo diversas experiencias de participación ciudadana a través de la red (Ayuntamiento de Zaragoza, Madrid,...). Uno de los problemas relacionados con dichas experiencias es la necesidad de disponer de procedimientos que permitan identificar las posiciones de los diferentes individuos y agregar sus opiniones en una función representativa del conjunto.

Uno de estos procedimientos es el *Método Contextualizado de Borda* (MCB) propuesto por Moreno-Jiménez y Escobar (2005). En este trabajo se presenta una aplicación de dicho procedimiento a una experiencia realizada dentro del ámbito de la participación electrónica, concretamente de la nueva propuesta de sistema democrático denominada e-cognocracia (Moreno-Jiménez, 2003, 2004, 2006; Moreno-Jiménez y Polasek, 2003, 2005).

Este procedimiento utiliza el *Proceso Analítico Jerárquico* (AHP) en: (i) la modelización del problema; (ii) la incorporación de las interdependencias entre las alternativas, las intensidades de las preferencias y la incertidumbre del decisor al emitir sus juicios y (iii) la *Agregación de Estructuras de Preferencia Individuales* (AIPS) como procedimiento de síntesis a la hora de ordenar las alternativas (Escobar y Moreno-Jiménez, 2006a). Este último procedimiento, combina el uso de las marcas empleadas en los métodos de Borda (Ludwin, 1978; Saari, 1990) con ideas de los procedimientos de agregación seguidos en AHP.

Para alcanzar estos objetivos el trabajo se ha estructurado como sigue. En la Sección 2 se introducen las ideas propuestas por el nuevo sistema democrático denominado e-cognocracia. La Sección 3 revisa la toma de decisiones en grupo con el Proceso Analítico Jerárquico y el procedimiento de Agregación de Estructuras de Preferencia Individuales. La Sección 4 presenta la aplicación práctica en la que se ha aplicado el procedimiento AIPS y, finalmente, en la Sección 5 se destacan las conclusiones más importantes del estudio.

2. e-participación, e-democracia y e-cognocracia

Durante los últimos años, favorecido por las directrices marcadas por la política europea (véase el VI Programa Marco de Investigación), se está produciendo un notable desarrollo de herramientas analíticas e informáticas que potencien la participación

electrónica (e-participación) de los ciudadanos en la toma de decisiones públicas relativas al gobierno de la sociedad (Moreno-Jiménez y otros, 2007).

Muchos son los dominios de la actividad pública en los que se están implementando estas ideas (e-administración, e-discusión, e-votación, e-democracia, e-cognocracia,...) y muchas las herramientas colaborativas (blogs, foros, wikis,...) empleadas con tal fin. Una de las aplicaciones más novedosas e innovadoras es su uso en el contexto del modelo de representación democrático conocido como e-cognocracia (Moreno-Jiménez, 2003, 2006; Moreno-Jiménez y Polasek, 2003).

La *e-cognocracia* es un nuevo modelo de representación democrática que, basado en el evolucionismo de los sistemas vivos, combina la democracia representativa (“nueva derecha”) con la democracia participativa (“nueva izquierda”) para desarrollar una democracia cognitiva que persigue: (i) dar una nueva orientación a la democracia dirigida a la creación y difusión social del conocimiento derivado de la resolución científica de los problemas complejos planteados en el ámbito de las decisiones públicas relativas al gobierno de la sociedad (objetivo estratégico); (ii) solventar algunas de las limitaciones de la democracia tradicional, dando nuevo sentido a conceptos como participación, transparencia, control, comunicación y seguridad (objetivos tácticos) y (iii) desarrollar nuevas herramientas decisionales que permitan una implementación de la e-cognocracia que sea rigurosa, amigable y segura (metodológica y tecnológicamente), aplicando estas herramientas a diferentes situaciones reales (objetivos operativos).

La esencia de este nuevo sistema democrático, la e-cognocracia, viene reflejada en la frase: “*No hay democracia sin libertad, ni libertad sin conocimiento*”.

La *e-cognocracia* se puede esquematizar en los siguientes puntos:

1. Los ciudadanos podrán elegir entre participar en el sistema democrático depositando su confianza en un partido político como tradicionalmente se viene haciendo (*e-voting*), o implicarse directamente en la resolución de los problemas, aportando opiniones e ideas que permitan mejorar el conocimiento del proceso decisional seguido por el sistema (e-cognocracia).
2. Los escaños del parlamento se distribuirán en dos partes: (i) una ocupada por los partidos tradicionales como habitualmente se hace (parte pública) y (ii) la otra

reservada para los ciudadanos que desean implicarse directamente en el proceso democrático (parte privada).

3. No todos los temas tienen que ser obligatoriamente objeto de resolución mediante implicación directa de los ciudadanos.
4. La implicación directa de los ciudadanos en la toma de decisiones va orientada a la mejora y difusión del conocimiento social sobre el tópico en discusión, y a la expansión del conocimiento relativo al método científico seguido en la resolución del problema.
5. El conocimiento relevante del problema se refiere a patrones de comportamiento, estructuras de preferencia, hechos estilizados y tendencias.
6. Habitualmente la resolución del problema consistiría en la ordenación de un conjunto de alternativas discretas entre las que se seleccionará la mejor o un conjunto de las mismas. Para tratar este tipo de situaciones es necesario un marco multicriterio que permita, la resolución de problemas con múltiples escenarios, criterios y actores, en los que la incorporación de los aspectos intangibles y subjetivos es fundamental para la resolución del problema.

3. Decisión en Grupo con AHP: el procedimiento de Agregación de Estructuras de Preferencia (AIPS)

El Proceso Analítico Jerárquico (Saaty, 1977, 1980, 1996; Moreno-Jiménez, 2002) es una técnica multicriterio que permite la resolución de problemas complejos en los que intervienen múltiples escenarios, actores y criterios. AHP permite capturar la visión de la realidad de los actores implicados en la resolución de un problema y proporcionar las prioridades de las alternativas comparadas en una escala abstracta común (escala absoluta), válida para la ordenación y selección de los candidatos.

La metodología de AHP consta de cuatro etapas: (i) *Modelización*; (ii) *Valoración*; (iii) *Priorización* y (iv) *Síntesis*. En la primera (Modelización) se construye un modelo jerárquico en el que se incluyen escenarios, actores, criterios, subcriterios, atributos (subcriterios del último nivel de desagregación) y, por último, las alternativas. En la segunda (Valoración) se incorporan mediante comparaciones pareadas las preferencias de los individuos según la escala fundamental de Saaty (Saaty, 1980). En la tercera (Priorización) se calculan, mediante alguno de los procedimientos de priorización existentes, las *prioridades locales* (de cada nodo respecto del que cuelga) y las

prioridades globales (de cada nodo respecto al nodo raíz o meta del problema). Por último, se calculan las *prioridades totales* de las alternativas mediante la síntesis de las prioridades globales de las mismas. Una característica de esta técnica es que permite evaluar la *inconsistencia* del decisor a la hora de emitir los juicios mediante comparaciones pareadas.

En el caso de decisión en grupo con AHP (Saaty, 1989; Iz y Gardiner, 1993; Bryson, 1996; Condon et al. 2003; Moreno-Jiménez et al. 2005), donde todos los actores persiguen un mismo fin de manera coordinada, se supone una única y común jerarquía para modelizar el problema. A la hora de obtener las prioridades para el grupo o colectivo, se suelen seguir diferentes procedimientos (Saaty, 1980; Aczel y Saaty, 1983; Dyer y Forman, 1992; Ramanathan y Ganesh 1994; Van den Honert y Lootsma, 1997; Forman y Peniwati, 1998; Van den Honert, 2001; Altuzarra y otros, 2007; Gargallo y otros, 2006), siendo los dos más extendidos en la práctica los siguientes: (i) la *agregación de juicios individuales* (AIJ) y (ii) la *agregación de prioridades individuales* (AIP).

Suponiendo un contexto local (un único criterio en la jerarquía) con n alternativas (A_1, \dots, A_n) y r decisores (D_1, \dots, D_r) , y denotando por $A^{[k]} = (a_{ij}^{[k]})$ la matriz de comparaciones pareadas del k -ésimo decisor ($k=1, \dots, r$; $i, j = 1, \dots, n$) y por β_k su importancia relativa en el grupo ($\beta_k \geq 0$, $\sum_{k=1}^r \beta_k = 1$), las prioridades de las alternativas comparadas según los dos procedimientos AIJ y AIP se obtienen como:

(i) *Aggregación de Juicios Individuales* (AIJ): A partir de las matrices de juicios individuales, $A^{[k]}$ $k=1, \dots, r$, se construye una matriz de juicios para el grupo $A^{[G]} = (a_{ij}^{[G]})$ utilizando cualquier procedimiento de agregación de juicios individuales y, a partir de ella, se obtienen las prioridades de las alternativas, $w^{[G/J]} = (w_i^{[G/J]})$ mediante cualquier procedimiento de priorización.

(ii) *Aggregación de Prioridades Individuales* (AIP): A partir de las matrices de juicios individuales, $A^{[k]}$ $k=1, \dots, r$, se obtienen las prioridades individuales mediante cualquier procedimiento de priorización, $w^{[k]} = (w_i^{[k]})$, $k = 1, \dots, r$ y, a partir de ellas, utilizando

cualquier procedimiento de agregación de prioridades individuales se obtienen las prioridades de las alternativas, $w^{[G/P]} = (w_i^{[G/P]})$.

Los dos métodos de agregación anteriores trabajan con valores precisos para los juicios, restricción ésta muy fuerte en la práctica, sobre todo cuando se trabaja con atributos intangibles. Por estos motivos se presentó un nuevo procedimiento de votación (Moreno-Jiménez y Escobar, 2005; Escobar y Moreno-Jiménez, 2006a) en el que se agregan las estructuras de preferencia individuales obtenidas al incorporar, mediante intervalos de juicio, la incertidumbre del decisor.

Definición 1: Dadas n alternativas A_1, \dots, A_n , se define

(i) Un *Ranking* de las n alternativas, $R_j = (r_{j1}, \dots, r_{jn})$, como una de las posibles ordenaciones de las mismas. En total hay $n!$ posibles ordenaciones o rankings, $\mathbf{R} = \{R_1, \dots, R_j, \dots, R_s, s=n!, j=1, \dots, s\}$, donde $r_{jp} = i$ significa que la alternativa i está situada en la posición p -ésima ($p=1, \dots, n$) del ranking j -ésimo ($j=1, \dots, s$).

(ii) Una *Estructura de Preferencia* (EP) para n alternativas es la distribución de probabilidad de los rankings que se presentan para las mismas.

Conforme a la situación (contexto local) y notación introducida más arriba, así como la Definición 1, el esquema del nuevo procedimiento de agregación de las estructuras de preferencia individuales es el siguiente.

Procedimiento:

El método propuesto para obtener la *importancia global de una alternativa para el grupo* ($W_i^G = W^G(A_i), i = 1, \dots, n$) consiste en los siguientes pasos:

1. Cada decisor compara las alternativas permitiendo que sus comparaciones sean imprecisas e incompletas.
2. Se obtienen las estructuras de preferencia asociadas a cada decisor.
3. Se calcula la distancia entre cada par de decisores.
4. Posteriormente, se calcula la intensidad relativa o importancia holística (global) de cada posible ordenación en el conjunto de posibles estructuras de preferencia.
5. Finalmente se sintetiza la importancia holística de cada alternativa para el grupo.

Veamos cada uno de estos cinco pasos con más detalle.

Paso 1: Cada decisor D_k ($k=1,\dots,r$), utilizando alguna de la técnicas de decisión multicriterio -AHP en este caso-, compara las n alternativas A_i ($i=1,\dots,n$) proporcionando una matriz de juicios $A^{[k]} = (a_{ij}^{[k]})$ que puede ser incompleta¹ e imprecisa². De esta manera, la intensidad de las preferencias de los decisores es incluida en el modelo. Además, la medida de la inconsistencia permite evaluar comportamientos poco honrados de los actores.

Paso 2: Para cada decisor D_k se calcula la estructura de preferencia resultante $(R_j^{[k]}; p_j^{[k]})$, $k=1,\dots,r$, $j=1,\dots,s$. De esta forma se incorpora la incertidumbre del decisor a la hora de emitir sus juicios. El resultado viene dado en una tabla como la mostrada a continuación (Tabla 1).

Tabla 1. Estructuras de preferencia para los decisores

Dec.\Rank.	R_1	...	R_j	...	R_s	Suma
D_1	$p_1^{[1]}$...	$p_j^{[1]}$...	$p_s^{[1]}$	1
...	1
D_k	$p_1^{[k]}$...	$p_j^{[k]}$...	$p_s^{[k]}$	1
...	1
D_r	$p_1^{[r]}$...	$p_j^{[r]}$...	$p_s^{[r]}$	1

Paso 3: Se calcula la *distancia* $d_{kk'} = d(D_k, D_{k'})$ entre las estructuras de preferencia asociadas a dos decisores cualesquiera $(D_k, D_{k'})$, $k, k' \in \{1,\dots,r\}$ utilizando la siguiente expresión (Altuzarra et al, 2005):

$$d_{kk'} = d(D_k, D_{k'}) = \left[\left(p^{[k]} - p^{[k']} \right)^T S \left(p^{[k]} - p^{[k']} \right) \right]^{1/2}$$

donde $p^{[k]} = (p_1^{[k]}, \dots, p_s^{[k]})^T$ y S es una matriz de similitudes cuyos elementos s_{jl} miden la *similitud* entre cada par de rankings utilizando una medida de correlación ordinal, en este caso el coeficiente por rangos de Spearman (ρ):

$$s_{jl} = \rho(R_j, R_l) = 1 - \frac{6 \sum_{i=1}^n d_i^2}{n^3 - n}$$

¹ Los juicios emitidos deben, al menos, permitir la conexión de todos los nodos.

² En cada entrada de la matriz se tiene un intervalo de juicios.

con $d_i = p - q$, donde $r_{jp} = r_{iq} = i$, esto es, d_i es la diferencia entre las posiciones que ocupa la alternativa A_i en las dos ordenaciones. De esta manera se capturan las interdependencias entre las alternativas comparadas en el problema.

Paso 4: Se calcula la *importancia relativa para el grupo*, $\Pi^G(R_j)$, de cada *ranking* R_j $j=1, \dots, s$ para aquellos rankings mostrados por algún decisor (con probabilidades no nulas). Habitualmente, la importancia relativa de un ranking, en general de cualquier elemento de una tabla de contingencia, se obtiene marginalizando, esto es, tomando un valor promedio de la respectiva fila o columna, como puede ser la media aritmética

$$(\Pi_a^G(R_j) = \sum_{k=1}^r \beta_k p_j^{[k]}) \text{ o la geométrica } (\Pi_g^G(R_j) = \prod_{k=1}^r (p_j^{[k]})^{\beta_k}).$$

Desgraciadamente esta forma de proceder no captura las interdependencias entre las alternativas. Para recoger este aspecto, así como las interrelaciones entre los actores implicados, se evalúa la distancia del grupo al ranking considerado:

$$D^G(R_j) = \sum_{k=1}^r \beta_k d(D_k, R_j)$$

A partir de esta distancia se define la *importancia relativa para el grupo del ranking* R_j como:

$$\Pi_h^G(R_j) = \frac{1/D^G(R_j)}{\sum_j 1/D^G(R_j)}, \quad j = 1, \dots, s$$

Conocida la importancia relativa para el grupo de cada ranking se pueden determinar el *ranking más preferido* por el grupo (mayor $\Pi_h^G(R_j)$, $j=1, \dots, s$), así como la *estructura de preferencia representativa* del mismo, que se define como:

$$(R_j^{[G]}; p_j^{[G]} = \Pi_h^G(R_j), j=1, \dots, s).$$

Con la información anterior se puede establecer el nuevo procedimiento de votación, conocido como votación de Borda contextualizada de la siguiente forma:

Paso 5: Utilizando una idea similar a los métodos de Borda, se define la *importancia global para el grupo* de cada una de las *alternativas* consideradas, $W_i^G = W^G(A_i)$, $i = 1, \dots, n$, como:

$$W^G(A_i) = \sum_{p=1}^n (n-p+1) \sum_{j \in S_p(A_i)} \prod^G(R_j), \quad i = 1, \dots, n \quad [1]$$

siendo $S_p(A_i) = \{j \mid A_i \text{ ocupa la posición } p\text{-ésima en el ranking } R_j, j=1, \dots, s\}$

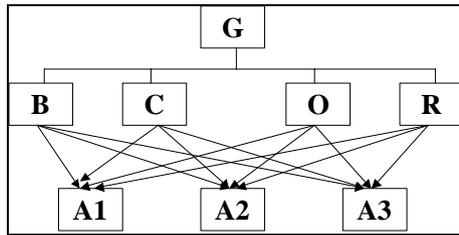
Los valores $W_i^G = W^G(A_i), i = 1, \dots, n$ permiten resolver los problemas tipo α ($P.\alpha$), esto es seleccionar la mejor alternativa (Roy, 1993), pues establecen un orden total entre los candidatos comparados.

Escobar y Moreno-Jiménez (2006b) analizan algunas cuestiones abiertas de los métodos empleados en el procedimiento AIPS, encontrando que dicho procedimiento permite discriminar las intensidades entre las diferentes posiciones, al asociar a cada marca la importancia holística de cada ordenación. En cuanto a los comportamientos deshonestos, éstos se ven limitados por la redundancia en los juicios y por la utilización de medidas de inconsistencia. Por último, la escala derivada al utilizar AHP proporciona una escala común para todos los individuos. Además, en dicho trabajo se analizaron algunas propiedades del procedimiento de agregación de estructuras de preferencia desde la óptica de la Teoría de la Elección Social.

4. Aplicación en e-cognocracia

En lo que sigue se presenta la aplicación a un caso real de e-participación abordado desde la perspectiva de la e-cognocracia en el que se utiliza el procedimiento de Agregación de Estructuras de Preferencia (AIPS) para agregar las opiniones de los diferentes individuos participantes. El experimento tuvo lugar durante las II Jornadas de Comercio Electrónico y Administración Electrónica (JCEL) llevadas a cabo en Zaragoza del 28 de febrero al 2 de marzo de 2007 (<http://jcel.unizar.es/index.php>). En dichas jornadas se realizó un taller sobre voto electrónico en el que los asistentes pudieron participar con sus opiniones y votaciones. El tema sobre el que se solicitó su opinión fue el de instalación de una base de inteligencia de la OTAN en la ciudad de Zaragoza, cuestión de actualidad en ese momento en la ciudad. Para abordar dicho problema se consideró una estructura jerárquica (ver Figura 1) con cuatro criterios (Beneficios, Costes, Oportunidades y Riesgos) y tres alternativas: (A1) ‘Instalar la base en Zaragoza’; (A2) ‘Instalar la base a más de 50 km’; y (A3) ‘No instalar la base’.

Figura 1. Jerarquía del problema



Un total de 21 participantes emitieron sus juicios a través de la red en dos rondas de opinión diferentes, separadas por un proceso de discusión realizado a través de un foro (Moreno-Jiménez y otros, 2007). Las prioridades finales alcanzadas para cada uno de los participantes en las dos rondas, aplicando el procedimiento de priorización de la media geométrica por filas son los que se muestran en la Tabla 2. En dicha tabla se muestran también los resultados que se obtienen aplicando el procedimiento de agregación de prioridades individuales (AIP).

Tabla 2. Prioridades finales individuales y agregadas (AIP)

Ronda 1				Ronda 2			
Individuo	Prioridades finales			Individuo	Prioridades finales		
	A1	A2	A3		A1	A2	A3
1	0,3333	0,3333	0,3333	1	0,5850	0,2045	0,2105
2	0,4036	0,2884	0,3080	2	0,4536	0,2337	0,3127
3	0,1564	0,5950	0,2487	3	0,4790	0,1822	0,3388
4	0,5735	0,3726	0,0538	4	0,6966	0,2599	0,0435
5	0,0795	0,3440	0,5765	5	0,2097	0,2180	0,5722
6	0,6422	0,1176	0,2402	6	0,8182	0,0909	0,0909
7	0,4934	0,1300	0,3767	7	0,0416	0,1799	0,7785
8	0,2412	0,2431	0,5157	8	0,3182	0,4741	0,2076
9	0,1057	0,2096	0,6847	9	0,0606	0,1886	0,7508
10	0,1547	0,2786	0,5667	10	0,2898	0,5352	0,1750
11	0,6472	0,2878	0,0650	11	0,7987	0,1480	0,0532
12	0,1998	0,1273	0,6729	12	0,2282	0,1677	0,6041
13	0,5783	0,3097	0,1120	13	0,5467	0,3657	0,0875
14	0,7337	0,1985	0,0678	14	0,7536	0,1913	0,0551
15	0,3401	0,3843	0,2756	15	0,4523	0,2911	0,2566
16	0,1210	0,5407	0,3383	16	0,0893	0,3350	0,5758
17	0,1474	0,6324	0,2202	17	0,7785	0,1799	0,0416
18	0,7509	0,1991	0,0500	18	0,4728	0,4728	0,0544
19	0,1198	0,4274	0,4528	19	0,2911	0,6284	0,0805
20	0,4454	0,2751	0,2795	20	0,5563	0,3498	0,0940
21	0,1830	0,5717	0,2453	21	0,3038	0,5312	0,1651
AIP	0,3441	0,3570	0,2989	AIP	0,4434	0,3380	0,2187

Se puede observar que en cuanto a las prioridades agregadas en la primera ronda recibe una mayor prioridad la alternativa A2, seguida de la alternativa A1 con una prioridad ligeramente inferior, quedando posicionada en último lugar la alternativa A3. En la

segunda ronda la alternativa preferida pasa a ser la A1, seguida de la A2 (a mayor distancia que en la Ronda 1) y en último lugar la alternativa A3.

Para aplicar el procedimiento de agregación de estructuras de preferencia individuales se toma como punto de partida las ordenaciones que proponen cada uno de los individuos participantes en el proceso. En la Tabla 3 se resumen dichas ordenaciones en las dos rondas de opinión.

Tabla 3. Ordenaciones o rankings propuestos por los individuos en las dos rondas de opinión

Ronda 1		Ronda 2	
Individuo	Rankings	Individuo	Rankings
1	Todas	1	132
2	132	2	132
3	231	3	132
4	123	4	123
5	321	5	321
6	132	6	123; 132
7	132	7	321
8	321	8	213
9	321	9	321
10	321	10	213
11	123	11	123
12	312	12	312
13	123	13	123
14	123	14	123
15	213	15	123
16	231	16	321
17	231	17	123
18	123	18	123; 213
19	321	19	213
20	132	20	123
21	231	21	213
AIP	213	AIP	123

Del total de individuos (21), ocho (38,10%) mantienen la misma ordenación de las alternativas de la primera a la segunda ronda; dos individuos (9,52%) modifican la ordenación de las dos primeras alternativas y otros dos modifican la ordenación de las dos últimas; cuatro individuos (19,05%) pasan a considerar la alternativa preferida en la primera ronda a la última en la segunda mientras uno de los individuos pasa a ordenar en primer lugar la alternativa que ocupaba el último lugar en la primera ronda.

En cuanto a la ordenación que se obtiene tras aplicar el procedimiento de agregación de prioridades individuales (AIP) como ya se ha señalado previamente, se modifica la ordenación de las dos primeras alternativas. En cambio, cuando se consideran las

probabilidades que cada individuo da a cada posible ordenación y se agregan dichas probabilidades, los resultados que se obtienen son diferentes como se puede ver en la Tabla 4. Concretamente, en la primera ronda, el ranking 213 que es el resultante de aplicar AIP recibe una de las probabilidades más pequeñas, mientras que en la segunda ronda el ranking 123, obtenido al utilizar el procedimiento AIP, sí que recibe la mayor probabilidad.

La diferencia de resultados para la ronda 1 proviene del hecho de que aunque un número pequeño de individuos propone la ordenación 213, las prioridades promedio sí se ordenan de acuerdo con dicho ranking.

Tabla 4. Estructuras de preferencia con juicios precisos

Ronda 1	Probabilidades de los rankings					
Individuo	123	132	213	231	312	321
1	0,167	0,167	0,167	0,167	0,167	0,167
2	0	1	0	0	0	0
3	0	0	0	1	0	0
4	1	0	0	0	0	0
5	0	0	0	0	0	1
6	0	1	0	0	0	0
7	0	1	0	0	0	0
8	0	0	0	0	0	1
9	0	0	0	0	0	1
10	0	0	0	0	0	1
11	1	0	0	0	0	0
12	0	0	0	0	1	0
13	1	0	0	0	0	0
14	1	0	0	0	0	0
15	0	0	1	0	0	0
16	0	0	0	1	0	0
17	0	0	0	1	0	0
18	1	0	0	0	0	0
19	0	0	0	0	0	1
20	0	1	0	0	0	0
21	0	0	0	1	0	0
AIPS I	0,246	0,198	0,056	0,198	0,056	0,246

Ronda 2	Probabilidades de los rankings					
Individuo	123	132	213	231	312	321
1	0	1	0	0	0	0
2	0	1	0	0	0	0
3	0	1	0	0	0	0
4	1	0	0	0	0	0
5	0	0	0	0	0	1
6	0,5	0,5	0	0	0	0
7	0	0	0	0	0	1
8	0	0	1	0	0	0
9	0	0	0	0	0	1
10	0	0	1	0	0	0
11	1	0	0	0	0	0
12	0	0	0	0	1	0
13	1	0	0	0	0	0
14	1	0	0	0	0	0
15	1	0	0	0	0	0
16	0	0	0	0	0	1
17	1	0	0	0	0	0
18	0,5	0	0,5	0	0	0
19	0	0	1	0	0	0
20	1	0	0	0	0	0
21	0	0	1	0	0	0
AIPS I	0,381	0,167	0,214	0,000	0,048	0,190

Utilizando la información proporcionada en la Tabla 4 se ha aplicado el procedimiento AIPS descrito en la sección anterior. Al emplear este procedimiento se tiene en cuenta la relación o distancia existente entre los diferentes rankings. Los resultados obtenidos se muestran en el Gráfico 1.

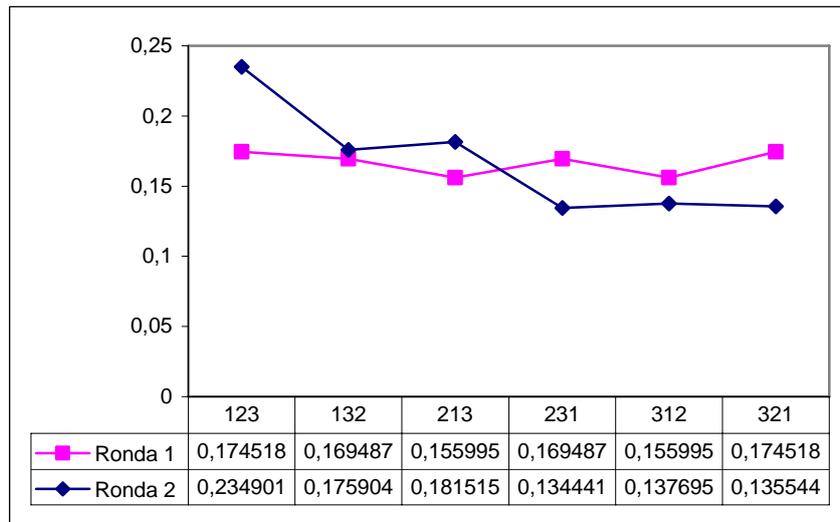


Gráfico 1. Resultados del procedimiento AIPS con juicios precisos

Se observa que en la ronda 1, los rankings que reciben una mayor importancia son el 123 y el 321 seguidos a poca distancia del resto de ordenaciones. Sin embargo, en la ronda 2 las posiciones se clarifican y el ranking de mayor importancia holística resulta ser el 123 seguido del 213.

El procedimiento AIPS permite también obtener la importancia asociada a cada alternativa, siguiendo un procedimiento similar al del método de Borda. Los resultados que se alcanzan para este experimento se muestran en el Gráfico 2, observándose que en la primera ronda la importancia holística que reciben las tres alternativas es exactamente la misma, mientras que en la segunda ronda hay una pequeña ventaja de la alternativa A1 sobre A2 y a su vez, de ésta sobre A3.

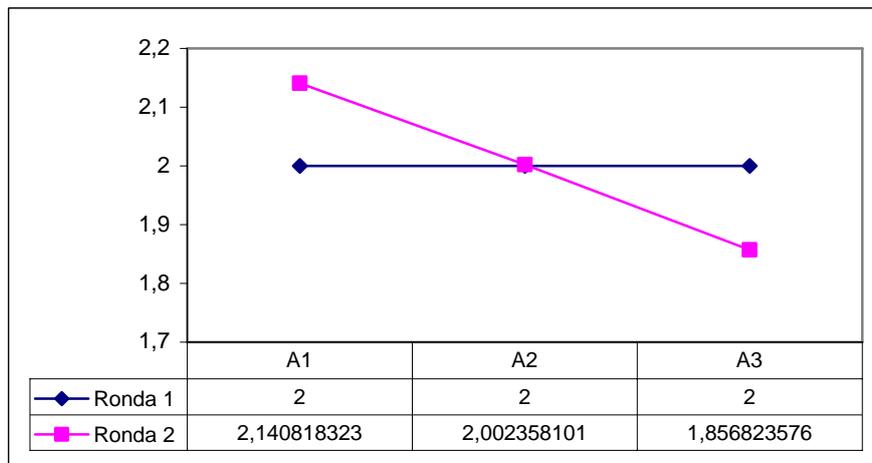


Gráfico 2. Importancia de las alternativas al aplicar AIPS con juicios precisos

Estos resultados se han obtenido al aplicar el procedimiento AIPS a los juicios que los individuos emitieron de forma precisa. Aunque el procedimiento tiene en cuenta las interdependencias entre las diferentes ordenaciones, la consideración del ranking de forma precisa no permite diferenciar entre dos individuos que ordenen las alternativas exactamente igual, uno diferenciándolas ampliamente y otro que apenas considere diferencias entre las mismas. Para poder capturar esta información que el decisor proporciona al emitir sus juicios se ha aprovechado una de las ventajas del procedimiento AIPS, que es la posibilidad de considerar juicios imprecisos.

Para ello, se ha vuelto a aplicar dicho procedimiento partiendo de juicios imprecisos, concretamente, en la matriz de juicios del primer nivel de la jerarquía, en la que los individuos comparan la importancia de los criterios con respecto a la meta. Los juicios originales se han sustituido por intervalos de juicio centrados en dicho valor original y con una amplitud/variación en la escala de juicios de AHP ($\{1/9, \dots, 1/2, 1, 2, \dots, 9\}$) de +/- 1 unidades.

Tabla 5. Estructuras de preferencia con juicios imprecisos

Ronda 1	Probabilidades de los rankings					
Individuo	123	132	213	231	312	321
1	0,167	0,167	0,167	0,167	0,167	0,167
2	0,56	0,44	0	0	0	0
3	0	0	0	1	0	0
4	1	0	0	0	0	0
5	0	0	0	0	0	1
6	0	1	0	0	0	0
7	0	1	0	0	0	0
8	0	0	0	0	0,978	0,022
9	0	0	0	0	0	1
10	0	0	0	0	0	1
11	1	0	0	0	0	0
12	0	0	0	0	1	0
13	1	0	0	0	0	0
14	1	0	0	0	0	0
15	0,008	0	0,992	0	0	0
16	0	0	0	1	0	0
17	0	0	0	1	0	0
18	1	0	0	0	0	0
19	0	0	0	0,095	0	0,905
20	0	1	0	0	0	0
21	0	0	0,009	0,991	0	0

Ronda 2	Probabilidades de los rankings					
Individuo	123	132	213	231	312	321
1	0,168	0,832	0	0	0	0
2	0	1	0	0	0	0
3	0	1	0	0	0	0
4	1	0	0	0	0	0
5	0	0	0	0	0,087	0,913
6	0,5	0,5	0	0	0	0
7	0	0	0	0	0	1
8	0	0	1	0	0	0
9	0	0	0	0	0	1
10	0	0	1	0	0	0
11	1	0	0	0	0	0
12	0	0	0	0	1	0
13	1	0	0	0	0	0
14	1	0	0	0	0	0
15	1	0	0	0	0	0
16	0	0	0	0	0	1
17	1	0	0	0	0	0
18	0,5	0	0,5	0	0	0
19	0	0	1	0	0	0
20	1	0	0	0	0	0
21	0	0	1	0	0	0

En la Tabla 5 se muestran las estructuras de preferencia obtenidas para cada individuo tras realizar un experimento de simulación de tamaño 1000 generando valores dentro de

los intervalos de juicio fijados y calculando posteriormente para cada combinación de valores simulados las respectivas prioridades. Los resultados alcanzados tras aplicar de nuevo el procedimiento AIPS a las estructuras de preferencia de la Tabla 5 son los mostrados en el Gráfico 3.

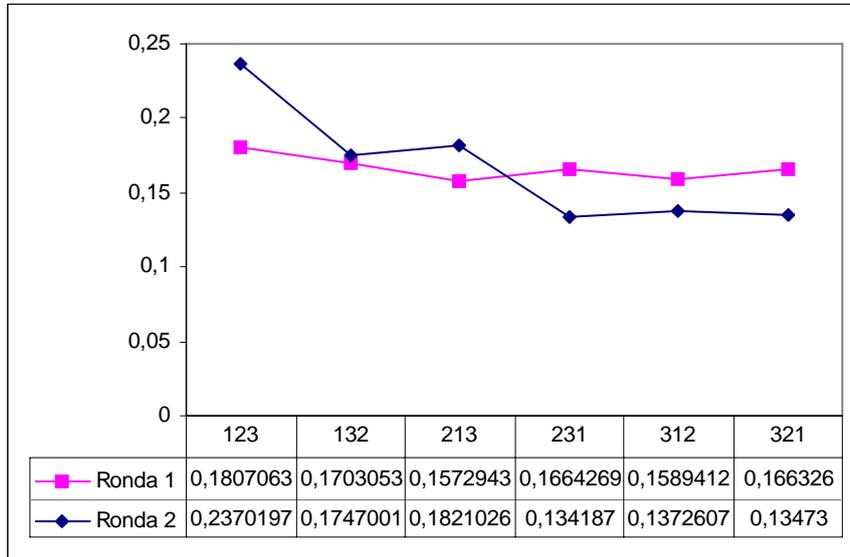


Gráfico 3. Resultados del procedimiento AIPS con juicios imprecisos

En la ronda 1 el ranking que recibe una mayor importancia es el 123 seguido del 132. En la ronda 2 las diferencias son algo mayores, y el ranking con una mayor importancia sigue siendo el 123 seguido en este caso del 213.

Al aplicar AIPS para calcular la importancia asociada a cada alternativa, se obtienen los resultados que se muestran en el Gráfico 4, en el que se vuelve a visualizar que las opiniones tras la discusión en el foro se clarificaron a favor de la alternativa A1.

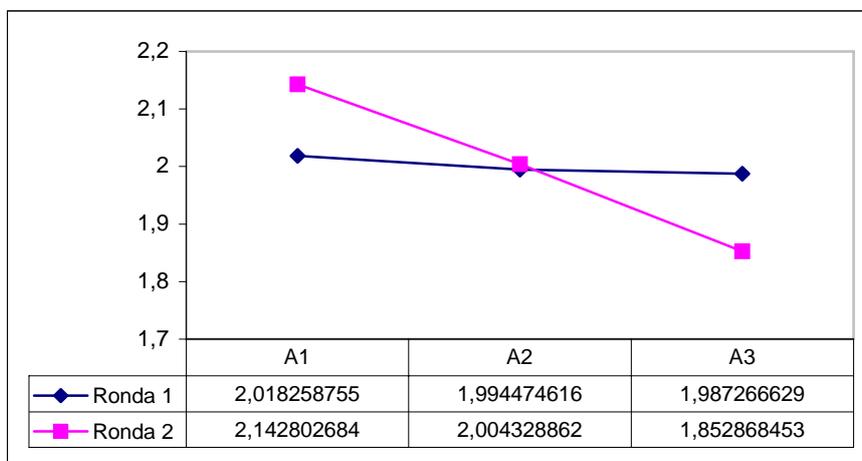


Gráfico 4. Importancia de las alternativas al aplicar AIPS con juicios imprecisos

5. Conclusiones

En este trabajo se aplica el método de decisión con múltiples actores conocido como *Agregación de Estructuras de Preferencia Individuales* (AIPS) a la resolución de un problema de localización planteado en el taller sobre “Voto electrónico y e-cognocracia” realizado dentro de las II Jornadas sobre Comercio Electrónico celebradas en Zaragoza (<http://jcel.unizar.es/index.php>).

La aplicación del método AIPS permite capturar la interdependencia existente entre las alternativas e incorporar la incertidumbre inherente en el proceso de emisión de juicios. Por otro lado, el uso de AIPS pone de manifiesto el cambio acaecido en las estructuras de preferencias individuales durante el proceso de discusión llevado a cabo entre los participantes en el taller. También ha permitido analizar la evolución temporal de las estructuras de preferencia colectivas. La tendencia de este cambio coincide con la detectada con otras herramientas decisionales usadas en este tipo de estudios (Moreno-Jiménez y otros, 2007; Turón y Moreno-Jiménez, 2007). Por último, destacar que con el experimento realizado se ha extendido la aplicación del procedimiento AIPS a un número de decidores más amplio que el considerado hasta el momento en trabajos previos (Moreno-Jiménez y Escobar, 2005; Escobar y Moreno-Jiménez, 2006a).

Bibliografía

Aczél, J.; Saaty, T.L. (1983): “Procedures for Synthesizing Ratio Judgements”, *Journal of Mathematical Psychology* 27 (1), 93-102.

Altuzarra, A.; Moreno-Jiménez, J.M.; Salvador, M. (2005): “Searching for consensus in AHP-group decision making. A Bayesian approach”, *Proceedings CODAWORK'05*, Gerona.

Altuzarra, A.; Moreno-Jiménez, J.M.; Salvador, M. (2007): “A Bayesian Priorization Procedure for AHP-Group Decision Making”, *European Journal of Operational Research*, 182, 367-382.

Bryson, N. (1996): “Group Decision Making and the analytic hierarchy process: exploring the consensus-relevant information content”, *Computers and Operations Research*, 23, 27-35.

Condon, E.; Golden, B.; Wasil, E. (2003): “Visualizing group decisions in the analytic hierarchy process”, *Computers and Operations Research*, 30 (10), 1435-1445.

Dyer, R.F.; Forman, E.H. (1992): “Group Decision Support with the Analytic Hierarchy Process”, *Decision Support Systems* 8, 99-124.

Escobar, M.T.; Moreno-Jiménez, J.M. (2006a): “Aggregation of Individual Preference Structures in AHP-Group Decision Making”, *Group Decision and Negotiation*, DOI: 10.1007/s10726-006-9050-x.

- Escobar, M.T.; Moreno-Jiménez, J.M. (2006b): “Notas sobre el método ‘Agregación de Estructuras de Preferencia’ en Decisión en Grupo con AHP”, *Anales de Economía Aplicada* 2006. CD.
- Forman, E.; Peniwati, K. (1998): “Aggregating individual judgements and priorities with the Analytic Hierarchy Process”, *European Journal of Operational Research*, 108, 165-169.
- Gargallo, M.P.; Moreno-Jiménez, J.M.; Salvador, M. (2006): “AHP-Group Decision Making. A Bayesian approach based on mixtures for group pattern identification”, *Group Decision and Negotiation*, DOI: 10.1007/s10726-006-9068-0.
- Iz, P.H.; Gardiner, L.R. (1993): “Analysis of Multiple Criteria Decision Support Systems for Cooperative Groups”, *Group Decision and Negotiation*, 2, 61-79.
- Ludwin, W.B. (1978): “Strategic Voting and the Borda Method”, *Public Choice*, 33, 85-90.
- Moreno-Jiménez, J.M. (2002): “El Proceso Analítico Jerárquico. Fundamentos. Metodología y Aplicaciones”. En Caballero, R. y Fernández, G.M. Toma de decisiones con criterios múltiples. *RECT@ Revista Electrónica de Comunicaciones y Trabajos de ASEPUMA*, Serie Monografías nº 1, 21-53.
- Moreno-Jiménez, J.M. (2003): “Las Nuevas Tecnologías y la Representación Democrática del Inmigrante”, *IV Jornadas Jurídicas de Albarracín. Consejo General del Poder Judicial (TSJA)*.
- Moreno-Jiménez, J.M. (2004): “E-cognocracia y Representación Democrática del Inmigrante”, *Anales de Economía Aplicada* (CD).
- Moreno-Jiménez, J.M. (2006): “E-cognocracia: Nueva Sociedad, Nueva Democracia”, *Estudios de Economía Aplicada*, 24(1-2), 559-581.
- Moreno-Jiménez, J.M.; Aguarón, J.; Raluy, A.; Turón, A. (2005): “A Spreadsheet Module for consistent consensus building in AHP-Group Decision Making”, *Group Decision and Negotiation*, 14 (2), 89-108.
- Moreno-Jiménez, J.M.; Escobar, M.T. (2005): “Votación e intensidad de las preferencias. Una aproximación basada en AHP”, *Anales de Economía Aplicada* 2005.
- Moreno-Jiménez, J.M.; Piles, J.; Ruiz, J.; Salazar, J.L.; Sanz, A. (2007): “E-cognocracia y voto electrónico. Aplicación a un problema de localización”, *XXII Anales de Economía Aplicada*. Valladolid.
- Moreno-Jiménez, J.M.; Polasek, W. (2003): “E-democracy and Knowledge. A Multicriteria Framework for the New Democratic Era”, *Journal Multicriteria Decision Analysis*, 12, 163-176.
- Moreno-Jiménez, J.M.; Polasek, W. (2005): “E-cognocracy and the participation of immigrants in e-governance”. En Böhlen y otros (eds.): *TED Conference on e-government 2005. Electronic democracy: The challenge ahead*. University Rudolf Trauner-Verlag, *Schriftenreihe Informatik*, 13, 18-26.
- Ramanathan, R.; Ganesh, L.S. (1994): “Group preference aggregation methods employed in AHP: An evaluation and intrinsic process for deriving members' weightages”, *European Journal of Operational Research*, 79, 249-265.

- Roy, B. (1993): "Decision science or decision-aid science?". *European Journal of Operational Research*, 66, 184-203.
- Saari, D. (1990): "The Borda Dictionary". *Social Choice and Welfare*, 7, 279-317.
- Saaty, T.L. (1977): "A Scaling Method for Priorities in Hierarchical Structures", *Journal of Mathematical Psychology*, 15 (3), 234-281.
- Saaty, T.L. (1980): *Multicriteria Decision Making: The Analytic Hierarchy Process*, Mc Graw-Hill, New York. (2nd print 1990, RSW Pub. Pittsburgh).
- Saaty, T.L. (1989): "Group Decision Making and the AHP". En: Golden, B.L., Wasil, E.A., Harker, P.T. (eds) *The Analytic Hierarchy Process: Application and Studies*, 59-67.
- Saaty, T.L. (1996): *The Analytic Network Process*, RSW Publications.
- Turón, A.; Moreno-Jiménez, J.M. (2007): "Herramientas de visualización gráfica para el análisis de la evolución dinámica en la e-cognocracia", *XXII Anales de Economía Aplicada*. Valladolid.
- Van den Honert, R. C. (2001): "Decisional Power in Group Decision Making: A Note on the Allocation of Group Members' Weights in the Multiplicative AHP and SMART", *Group Decision and Negotiation*, 10(3), 275-286.
- Van den Honert, R.C.; Lootsma, F. A. (1997): "Group preference aggregation in the multiplicative AHP The model of the group decision process and Pareto optimality", *European Journal of Operational Research*, 96(2), 363-370.

ST INDEX. Propuesta de un índice sintético global para la medición del desarrollo turístico sostenible

MARCELINO SÁNCHEZ RIVERO

e-mail: sanriver@unex.es

Departamento de Economía Aplicada I
UNIVERSIDAD DE EXTREMADURA

JUAN IGNACIO PULIDO FERNÁNDEZ

e-mail: jipulido@ujaen.es

Departamento de Economía
UNIVERSIDAD DE JAÉN

Área temática: Métodos cuantitativos

Resumen

Este trabajo establece las bases metodológicas para el diseño de un índice sintético global para la medición del desarrollo turístico sostenible, que se denominará ST INDEX. El cálculo de este índice responderá a la cuádruple dimensión de la sostenibilidad y facilita el seguimiento y control de resultados de las políticas de fomento del turismo sostenible, a la vez que sirve de elemento de medición de la convergencia entre distintos destinos turísticos, que dispondrán, además, de un conocimiento más profundo sobre los elementos que condicionan su capacidad competitiva en los mercados. La elaboración de índices sintéticos de desarrollo sostenible es una propuesta aún muy reciente, por lo que las aportaciones teóricas son todavía muy escasas. En el caso del turismo sostenible, por el momento no se ha producido ningún avance, por lo que la investigación que se plantea en esta propuesta tiene un carácter de exclusivamente novedosa, teniendo como referente el Monitor de Competitividad Turística del WTTC, la aportación al World Economic Forum (WEF) de los grupos de trabajo de las Universidades de Yale y de Columbia, que han diseñado el Environmental Sustainability Index (ESI) y, para el caso de España, el trabajo realizado por un equipo de investigadores de la Universidad de A Coruña para el diseño de un Índice Sintético de Desarrollo Sostenible (Synthetic Index of Sustainable Development, o SISD), que han aplicado al análisis del desarrollo sostenible de las Comunidades Autónomas con buen resultado.

Palabras clave: desarrollo económico, sostenibilidad, turismo sostenible, índice sintético, ST INDEX, gestión territorial.

Abstract

This paper establishes the methodological bases for the design of a global synthetic index for the measurement of the sustainable tourism development. This index will be denominated ST INDEX. The calculation of this index will respond to the quadruple dimension of the sustainability and it facilitates the monitoring and control of results of the development policies of the sustainable tourism, at the same time it also serves as element of measurement of the convergence among different tourism destinations that will prepare of a deeper knowledge on the elements that condition its competitiveness in the markets.

The elaboration of synthetic indexes of sustainable development is a very recent proposal, for what the theoretical contributions are still very scarce. In the case of the sustainable tourism, at the moment any advance has not taken place, for what the issue that we propose in this paper has a character of exclusively novel, having as relating the Tourism Competitiveness Monitor to the WTTC, the contribution to the World Economic Forum (WEF) of the groups of work of the Universities of Yale and of Columbia that the Environmental Sustainability Index has designed (ESI) and, for the case of Spain, the work carried out by a team of researchers of the University of A Coruña for the design of a Synthetic Index of Sustainable Development (SISD) that have applied to the analysis of the sustainable development of the Autonomous Communities with good result.

Key words: economic development, sustainability, sustainable tourism, synthetic index, ST INDEX, territorial management.

1. Introducción

El de la sostenibilidad es uno de los debates que más interés ha generado entre los investigadores de la última década. El turismo, una de las principales actividades económicas a nivel mundial, no ha quedado al margen de este debate, que ha tenido una parte sustancial en la necesidad de definir herramientas con las que medir la evolución de los destinos turísticos en el proceso de transición que implica el desarrollo sostenible.

Afrontar la medida de la sostenibilidad no es, ni mucho menos, una tarea sencilla. Son varias las razones que dificultan su cuantificación. En primer lugar, se trata de un concepto todavía excesivamente abstracto y, si se quiere, latente, es decir, no directamente observable, lo que obliga a ser cuantificado mediante indicadores indirectos que se relacionan, en mayor o menor medida, con la sostenibilidad. En segundo lugar, la sostenibilidad tiene un carácter claramente multivariante, puesto que son varios los aspectos con los que guarda relación. Su cuádruple naturaleza social, económica, medioambiental e institucional es una muestra evidente de hasta qué punto es una realidad multivariante. En tercer lugar, la medida de la sostenibilidad es relativa, en el sentido de que, dependiendo de la unidad de análisis que se tome como referencia, las unidades analizadas pueden ser más o menos sostenibles. Por último, su cuantificación viene dada, en la mayor parte de los casos, en diversas unidades de medida. Si fuera posible utilizar una unidad de medida común para la totalidad de indicadores que la cuantifican, expresada, por ejemplo, en términos monetarios, es evidente que su valoración cuantitativa sería considerablemente más sencilla de lo que es en realidad como consecuencia de la existencia de indicadores, especialmente los medioambientales, expresados en unidades de medida de lo más variado (emisiones de CO₂, tratamiento de residuos, producción de energías renovables, etc.).

El objeto de este trabajo es, precisamente, el de establecer una metodología con la que abordar la creación de un índice sintético de sostenibilidad turística, que se ha denominado ST INDEX, y que viene a cubrir las necesidades de agregación de información que plantea el propio concepto de sostenibilidad, facilitando, a la vez, una herramienta también muy útil para evaluar la gestión de los destinos turísticos y para comparar los esfuerzos realizados en materia de sostenibilidad por cada uno de ellos.

La primera parte se dedica al análisis del concepto de turismo sostenible y de los elementos que lo configuran, así como a justificar la necesidad de medir la sostenibilidad del turismo y de valorar el papel de los indicadores en esta tarea, poniendo especial énfasis en la necesidad de un índice sintético de turismo sostenible. Seguidamente, se repasan las principales aportaciones metodológicas para el cálculo de índices sintéticos, centrando el análisis en los tres métodos que servirán, posteriormente, para establecer la propuesta del ST INDEX. Esta propuesta queda definida en la tercera y última parte de este trabajo.

2. La sostenibilidad de la actividad turística

El interés científico por un desarrollo económico equilibrado y respetuoso con los recursos obtenidos de la Madre Naturaleza no es nuevo. Ya, desde finales del siglo XVII, autores como Petty, Cantillon, Hume o Quesnay se planteaban la relación entre el ser humano y la naturaleza, y las posibilidades de éste para mejorar las capacidades productivas de la tierra, que hasta entonces había sido considerada como la única fuente generadora de riqueza. Sin embargo, hay que esperar a la década de los sesenta del siglo XX para ver el comienzo de una nueva etapa, que tiene su origen en la preocupación por los, cada vez más importantes, problemas medioambientales que se empiezan a constatar en las sociedades industrializadas y por el agotamiento de los recursos naturales como consecuencia del impresionante consumo de recursos a que había dado lugar la revolución neoclásica. La década de los setenta es especialmente importante en este cambio de actitud de los economistas respecto a las cuestiones ambientales. El *Informe Meadows* (1972) y, sobre todo, la crisis del petróleo (1973) ponen de manifiesto los serios problemas a los que se enfrenta el modelo neoclásico de crecimiento económico ilimitado y sirven de punto de inflexión en el desinterés que el discurso económico imperante había tenido hasta entonces por estas cuestiones.

Sin embargo, no será hasta finales de los 80, incluso ya entrados los 90, cuando el debate sobre la sostenibilidad adquiera carta de naturaleza propia, convirtiéndose desde entonces en uno de los principales temas de investigación por parte de la comunidad científica y de interés político y social a nivel mundial. Inicialmente, este debate tuvo una marcada influencia medioambientalista, aunque, con el paso del tiempo, se han producido importantes avances en el estudio de la sostenibilidad y de los factores que la

condicionan, pasando a considerarse otros aspectos (económico, sociocultural e institucional) que actualmente son clave en el estudio de este paradigma, sin obviar, por supuesto, el ámbito medioambiental¹.

2.1. El debate en torno al paradigma del turismo sostenible

La traslación del debate científico de la sostenibilidad al ámbito turístico es aún más reciente, aunque, como señalan Ivars et al (2001, p. 6), la discusión en torno a la creciente importancia del turismo en los procesos de desarrollo económico y sus implicaciones ambientales, culturales y antropológicas se inicia mucho antes en el mundo académico, con trabajos tan importantes como los de Budowski (1976), Cohen (1978), Krippendorf (1978) o De Kadt (1979), por citar algunos; y, para el caso español, los análisis de Cals (1974), Gaviria (1976) o Jurdao (1979).

El concepto de turismo sostenible se ha consolidado en la literatura, científica y no científica, de la última década como pieza básica en el debate sobre la gestión de los recursos turísticos. Sin embargo, el pretendido desarrollo sostenible de la actividad turística, que ha sido objeto de innumerables propuestas dialécticas, en buena parte de los casos no ha llegado a concretarse en un inventario de políticas eficaces, capaces de garantizar que el aprovechamiento productivo del turismo sea compatible con la conservación del patrimonio natural y cultural del territorio en el que se desarrolla y con un modelo de gestión abierto y participativo.

La sostenibilidad no es una estrategia *per sé*, sino que implica un cambio del paradigma en el aprovechamiento turístico de los recursos, que no supone ausencia de crecimiento ni pérdida de eficacia, y garantiza en cambio una mejora en la calidad de vida de los residentes en los destinos y de los visitantes, reconociendo que existen límites reales a la disponibilidad de recursos y, por tanto, deben tomarse medidas para reducir el consumo material que eleva los costes ambientales, culturales y sociales.

Se requiere, en consecuencia, un equilibrio entre los objetivos a corto y a largo plazo, bajo el convencimiento de que pueden derivarse todo tipo de beneficios (incluso económicos) con un uso equilibrado de los recursos. Pero, lograr ese equilibrio, requiere la implementación en lo concreto de un proceso de toma de decisiones y una serie de

¹ Para un análisis más detallado de la evolución del concepto de sostenibilidad aplicado al turismo y de los factores que condicionan su consecución, pueden verse, entre otros, Dourojeanni (1997), Hall (2000),

actuaciones relativas al uso de los recursos necesarias para alcanzar el escenario deseado.

Como señalan Fullana y Ayuso (2002, p. 29), “el concepto de desarrollo sostenible no contiene directrices sobre la relación deseable entre las acciones de la sociedad humana y el estado del mundo natural”. Es evidente, además, como ya se ha puesto de manifiesto, que “no puede asociarse con un proyecto acabado ni supone una meta” (Sierra, 2002, p. 22), y tampoco “puede identificarse con un estado ideal, prácticamente utópico” (Ivars et al, 2001, p. 9). Se trata, por tanto, de un proceso de cambio cualitativo, abierto y en continua evolución; de un proceso de transición hacia un nuevo marco de relaciones entre los seres humanos y el ecosistema global, que exige nuevos fundamentos teóricos, soportes analíticos y planteamientos operativos, en la medida en que el marco conceptual tradicional presenta importantes limitaciones para afrontar con garantías este proceso.

De hecho, los avances en el conocimiento de la sostenibilidad han permitido ampliar el concepto inicial, incorporando nuevos planteamientos vinculados con el equilibrio entre los aspectos ambientales, socioculturales y económicos del turismo; la necesidad de liderazgo político para afrontar el proceso de cambio hacia la sostenibilidad del turismo; la participación informada de todos los agentes participantes; la aplicación de los principios de la sostenibilidad en todos los sectores de la actividad turística, en todas las formas de turismo y en todos los destinos turísticos; o con los objetivos mundiales de eliminación de la pobreza. En definitiva, así lo reconoce la Organización Mundial del Turismo (OMT) en su nueva definición de turismo sostenible, realizada por el Comité de Desarrollo Sostenible del Turismo (OMT, 2004: 7): “Las directrices para el desarrollo sostenible del turismo y las prácticas de gestión sostenible son aplicables a todas las formas de turismo en todos los tipos de destinos, incluidos el turismo de masas y los diversos segmentos turísticos. Los principios de sostenibilidad se refieren a los aspectos ambiental, económico y sociocultural del desarrollo turístico, habiéndose de establecer un equilibrio adecuado entre esas tres dimensiones para garantizar su sostenibilidad a largo plazo.

Por tanto, el turismo sostenible debe:

Swarbrooke (2000), Ivars et al (2001), Fullana y Ayuso (2002), Weaver (2005) o Pulido (2006a).

1. Dar un paso óptimo a los recursos ambientales que son un elemento fundamental del desarrollo turístico, manteniendo los procesos ecológicos esenciales y ayudando a conservar los recursos naturales y la diversidad biológica.
2. Respetar la autenticidad sociocultural de las comunidades anfitrionas, conservar sus activos culturales arquitectónicos y sus valores tradicionales, y contribuir al entendimiento y a la tolerancia interculturales.
3. Asegurar unas actividades económicas viables a largo plazo, que reporten a todos los agentes unos beneficios socioeconómicos bien distribuidos, entre los que se cuenten oportunidades de empleo estable y de obtención de ingresos y servicios sociales para las comunidades anfitrionas, y que contribuyan a la reducción de la pobreza.

El desarrollo sostenible del turismo exige la participación informada de todos los agentes relevantes, así como un liderazgo político firme para lograr una colaboración amplia y establecer un consenso. El logro de un turismo sostenible es un proceso continuo y requiere un seguimiento constante de los impactos para introducir las medidas preventivas o correctivas que resulten necesarias.

El turismo sostenible debe reportar también un alto grado de satisfacción a los turistas y representar para ellos una experiencia significativa, que los haga más conscientes de los problemas de la sostenibilidad y fomente en ellos unas prácticas turísticas sostenibles”.

Una definición amplia que sintetiza los aspectos básicos del turismo sostenible y que consolida la idea de la cuádruple dimensión de la sostenibilidad como eje fundamental del análisis del proceso de transición, quedando patente que la consecución del desarrollo sostenible en cualquier destino turístico requiere de un planteamiento integral desde la óptica de estas cuatro dimensiones y, en consecuencia, no se conseguirá abordándola desde la visión parcial de cualquiera de ellas.

2.2. ¿Por qué medir el turismo sostenible?

El análisis de la sostenibilidad se ha convertido, pues, en uno de los aspectos más atractivos e interesantes de la investigación en turismo, sobre todo porque su aplicabilidad real al desarrollo turístico de los territorios deja de ser una cuestión más o

menos etérea para convertirse en un aspecto fundamental del futuro de los destinos. Si a ello se une el hecho de que, según todas las previsiones e indicadores, el turismo se consolidará en el siglo XXI como la primera industria mundial, parece evidente la importancia y oportunidad de profundizar en el estudio de los factores clave de la sostenibilidad de los destinos turísticos.

La competitividad de los destinos turísticos dependerá, entre otras cuestiones, de su capacidad para definir y aplicar un modelo de desarrollo turístico acorde con los criterios de sostenibilidad económica, ambiental, sociocultural e institucional, y que, en consecuencia, les permita mantener sus recursos y atractivos en permanente buen estado para ser consumidos por el turista, presente y futuro. Este planteamiento conduce a un importante supuesto: la sostenibilidad de un destino turístico es una variable clave en el posicionamiento competitivo dentro de su mercado de referencia.

La eficiente respuesta a estos retos requiere, por un lado, un profundo compromiso social, involucrando a todos los sectores de la sociedad con un espíritu de responsabilidad compartida, aunque diferenciada según se trate de instituciones, empresas, organizaciones no gubernamentales, individuos, etc.; y, por otro, “repensar” el conjunto de instrumentos y mecanismos de gestión para incentivar cambios significativos en las actuales prácticas turísticas, teniendo en cuenta que es necesario reestructurar los modos de producción, distribución y consumo.

Hace falta una nueva política turística que disponga de una variada gama de instrumentos jurídicos, normativos, administrativos, sociales, institucionales, educativos, participativos, etc., y que supere los principios neoliberales en la concepción de los esquemas de gestión turística sostenible para garantizar el reforzamiento simultáneo de los procesos turísticos y territoriales-medioambientales en beneficio de ambos y del desarrollo económico global.

A pesar de todas las iniciativas puestas en marcha en la última década, tanto por el sector público como por las empresas y organizaciones no gubernamentales, la OMT presentó un informe en la *Cumbre Mundial sobre Desarrollo Sostenible* (Johanesburgo, 2002) sobre los avances conseguidos a nivel internacional en materia de desarrollo sostenible del turismo. La conclusión era desilusionante: “si bien, en términos generales, la necesidad de aplicar en el turismo una planificación sistemática se ha

convertido en un enfoque ampliamente aceptado, y las políticas y estrategias de turismo recomendadas por las organizaciones internacionales incorporan la cuestión medioambiental y la sostenibilidad social y económica, su aplicación efectiva por gobiernos centrales, regionales y locales y por los empresarios turísticos ha sido lenta y parcial” (Yunis, 2003: 19).

Según el propio informe, el principal reto a que se enfrenta la consecución de un turismo sostenible es el de salvar las distancias que actualmente existe entre la generación de orientaciones metodológicas, sugerencias de política turística y conocimientos técnicos y su aplicación práctica en la planificación real y en el desarrollo de destinos turísticos por parte de las administraciones públicas, así como en la ejecución de proyectos turísticos y el funcionamiento cotidiano de las empresas turísticas.

Para la superación este reto, la OMT propone, entre otras medidas, realizar un mayor esfuerzo para adaptar los conocimientos metodológicos y tecnológicos existentes a las situaciones específicas de cada territorio y dotar a las administraciones públicas, especialmente a las locales, a las empresas turísticas y, en general, a todas las partes interesadas en cualquier proceso de desarrollo turístico, de mecanismos adecuados para supervisar el cumplimiento de los criterios de sostenibilidad establecidos. Es necesario, por tanto, disponer de un marco operativo que ayude a los gestores a conocer en todo momento y de forma sistemática y objetiva la consecución de los objetivos de sostenibilidad respecto al desarrollo turístico de dicho territorio; objetivos que, por otra parte, habrán de ser previamente establecidos y que se deberán modificar en la medida en que cambie la situación del territorio, el comportamiento turístico o las preferencias locales.

En consecuencia, se convierte en objetivo principal de cualquier destino turístico la medición del nivel de sostenibilidad alcanzado por su proceso de desarrollo turístico, a fin de verificar el comportamiento favorable o desfavorable de su modelo de desarrollo según estos principios y adoptar los cambios necesarios en caso de no alcanzar los objetivos propuestos. Como señala Ko (2005: 432), “si el desarrollo sostenible es uno de los objetivos actuales más importantes de la industria del turismo, entonces la industria necesita poder medir su rendimiento e impactos en este área”.

2.3. El uso de indicadores para medir el turismo sostenible

Desde los años 90 del siglo pasado, se han ido discutiendo propuestas muy diversas de indicadores cuyo objeto era, precisamente, contribuir al diseño de marcos operativos que garanticen una evaluación tangible de la sostenibilidad de diferentes proyectos o modelos de desarrollo turístico².

El concepto de indicador ha sido establecido por no pocos autores. Así, Castell (2000: 35) lo define como “variable o relación entre variables (índices) de cuya medición se pueden obtener referencias ciertas sobre la evolución del sistema en que está inmersa. Las variables indicadoras son aquellas sensibles a cambios y tendencias de origen natural o humano, incluyendo el manejo”. Por otro lado, Sancho et al (2002) incluyen en la definición de su modelo AUDITSOS dos conceptos de indicador, uno de la OMT (1995), como “aquel instrumento cuantitativo y sintético que facilitaría el análisis y evaluación de la información de manera que, junto cualquier otro tipo de instrumentos, los actores decidores reduzcan la probabilidad de adoptar inadvertidamente decisiones desafortunadas”, y otro de James (2000)³, según el cual “un indicador será el instrumento que permitirá a los gestores de los destinos turísticos evaluar la sostenibilidad del mismo, conociendo cuales son los atributos clave, cual es su evolución, y cuales son los riesgos que les afectan o podrían llegar a afectar”.

Por su parte, la OCDE (1993) define los indicadores como “variables que ofrecen información sintética sobre determinados fenómenos o espacios gracias a la concesión de un valor suplementario que amplía la significación del parámetro individualmente considerado”. Y el Ministerio de Medio Ambiente (2001: 16), refiriéndose al concepto de indicador ambiental, señala que se trata de “una variable que ha sido socialmente dotado de un significado añadido al derivado de su propia configuración científica, con el fin de reflejar de forma sintética una preocupación social con respecto al medio ambiente e insertarla coherentemente en el proceso de toma de decisiones”.

En definitiva, como señalan los autores del proyecto METASIG, “la creación de indicadores implica una construcción científico-técnica, amparada en una valoración

² Para la revisión de algunas de estas propuestas, pueden consultarse, entre otros, Manning (1999), Miller (2001), Sirakaya et al (2001), Ivars et al (2001), Sancho et al (2002), Bloyer et al (2004), White et al (2006), Sancho y García (2006).

³ Publicado posteriormente en España como James (2004).

social previa, que otorga a determinados parámetros valor de representación sobre procesos relacionados con el desarrollo sostenible o, desde el punto de vista ambiental, con la presión sobre el medio, las condiciones ambientales objetivas y las políticas de respuesta” (Ivars et al, 2001: 37). Los indicadores permiten, pues, comprender cual es la situación actual, definir el estado ideal y pronosticar si, con las actuaciones ejecutadas, se está más cerca o más lejos de dicho estado, así como prevenir, anticipar y solucionar problemas (Bermejo y Nebreda, 1998).

Las tres funciones básicas de los indicadores (OCDE, 1997) son: simplificación, cuantificación y comunicación. Los indicadores han de ser representaciones empíricas de la realidad en las que se reduzcan el número de componentes. Además, han de medir cuantitativamente (al menos establecer una escala) el fenómeno a representar. Y, por último, el indicador ha de utilizarse para transmitir la información referente al objeto de estudio. Para el caso de los indicadores sociales, Fricker (1998) amplía estas tres funciones a cinco, pudiendo tener una utilidad informativa, predictiva, orientada hacia la resolución de problemas, evaluadora de programas y, por último, definitoria de objetivos; mientras que Miller (2001) se refiere a las siguientes funciones que pueden esperarse de los indicadores de sostenibilidad: identificación de problemas, mejor conocimiento y mayor concienciación pública acerca de esos problemas y definición, desarrollo y evaluación de políticas.

Los principales avances se han producido en el ámbito de la definición de indicadores de tipo ambiental (Ministerio de Medio Ambiente, 2003), es decir, aquellos que facilitan la evaluación de la incidencia del turismo en el medio ambiente, mientras que la incorporación de las demás dimensiones del turismo sostenible (sociocultural, económica e institucional) se ha planteado, en la mayoría de los casos, a través de una adaptación de los modelos utilizados para la definición de indicadores ambientales.

En efecto, se han definido, básicamente, tres modelos para abordar el proceso de construcción de indicadores ambientales, que son los que se han aplicado posteriormente para la obtención de indicadores de turismo sostenible. En España, el modelo adoptado por los autores de AUDITSOS (Sancho et al, 2002) es el aplicado por la OCDE (1993, 2000). Conocido internacionalmente como *modelo PSR* (Presión-Estado-Respuesta), se basa en el concepto de causalidad, es decir, las actividades

humanas ejercen Presiones sobre el medio ambiente y modifican la cualidad y calidad (Estado) de los recursos naturales. La sociedad responde a estos cambios a través de políticas ambientales, macroeconómicas y sectoriales (Respuestas). Éstas últimas producen una retroalimentación dirigida a modificar las presiones a través de las actividades humanas. A partir del modelo PSR, la Agencia Europea de Medio Ambiente (1998) creó un nuevo marco analítico, conocido como *modelo DPSIR* (correspondiente a los términos ingleses *driving forces-pressure-state-impact-response*), que incorpora las causas de la presión y los impactos producidos como consecuencia de las modificaciones que se han producido en el estado de las condiciones ambientales. Este es el modelo adoptado por Ivars et al (2001) para el desarrollo de METASIG. Finalmente, la WCPA ha propuesto un modelo especialmente pensado para valorar el estado de los espacios naturales protegidos y evaluar su gestión. El *modelo WCPA* (Hockings et al, 2000), que ya ha sido adoptado en España para la evaluación del sistema de espacios protegidos de Cataluña, propone las siguientes categorías de indicadores: de contexto, de planificación y legislación, de medios, de procesos, de actividades y servicios, y de resultados.

El problema está en que no existe una lista de indicadores universales y unánimemente aceptados. Es más, en la medida en que un indicador describe un proceso específico de control (y no una información exclusivamente numérica), los indicadores están fuertemente vinculados a los procesos de los que forman parte. Y, consecuentemente, las aportaciones realizadas hasta el momento sólo admiten una comparación muy parcial –variable a variable, o indicador a indicador-, ya que no establecen una forma homogénea, sintética y global para llevarla a cabo entre los distintos territorios o economías (González et al, 2004).

2.3. La necesidad de un índice sintético de turismo sostenible

Efectivamente, la mayor parte de las propuestas realizadas hasta ahora se basan en la construcción de indicadores para evaluar de forma independiente los diferentes componentes de la sostenibilidad, pero se ha avanzado muy poco en el diseño de indicadores que integren esa cuádruple dimensión. Es necesario, por tanto, encontrar indicadores adaptados al carácter integral y multidimensional de los procesos de desarrollo, aún a sabiendas de que este objetivo implica mayor complejidad en su

diseño, obtención e interpretación (Bermejo y Nebreda, 1998), lo que obliga a la creación de sistemas de indicadores adaptados a la especificidad de cada territorio, distinguiendo entre los distintos tipos de espacios turísticos, su grado de desarrollo y otras circunstancias que influyen tanto en la evolución de la actividad turística como en su orientación hacia el desarrollo sostenible (Ivars et al, 2001).

A ello puede contribuir el establecimiento de bases metodológicas para el diseño de un índice sintético global de desarrollo turístico sostenible que permita, por un lado, sistematizar la información acerca de los factores que condicionan la sostenibilidad de cualquier destino turístico, y, por otro, la comparación entre todos estos destinos, a fin de delimitar su esfuerzo, en términos comparados, respecto a los objetivos de la sostenibilidad.

Los indicadores sintéticos o índices son medidas adimensionales, resultado de combinar varios indicadores simples mediante un sistema de ponderación que jerarquiza los componentes. La información que se obtiene de estos indicadores es mayor, si bien la interpretación de la misma es, en muchos casos, más complicada y con ciertas restricciones.

3. Aportaciones metodológicas para el cálculo de índices sintéticos. Su aplicación al análisis del desarrollo sostenible o del turismo

Durante la última década han avanzado algunas iniciativas de elaboración de índices sintéticos para las diferentes dimensiones del desarrollo sostenible (United Nations, 2001b). También, en este caso, es la dimensión medioambiental de la sostenibilidad la que acumula un mayor número de experiencias (Living Planet Index, The Ecological Footprint,...), si bien, hasta el momento no existe un acuerdo internacional acerca de la validez general de ninguna de ellas. Respecto a la dimensión económica, se suele cuestionar la inconsistencia de las metodologías utilizadas para el cálculo de índices como el Genuine Progress Indicator (GPI) o el Index of Sustainable Economic Welfare (ISEW), de los que también se censura la subjetividad de sus resultados. Los índices sintéticos para medir la dimensión social están aún menos desarrollados. El más conocido, aunque también muy cuestionado, es el Human Development Index. Y, respecto a la dimensión institucional, aún no se ha desarrollado índices sintéticos de relevancia.

El gran problema de la mayoría de estos índices está en el hecho de que, por lo general, no existe información pública sobre las metodologías utilizadas, ni sobre los procesos de agregación seguidos, lo que dificulta su uso por otros organismos o instituciones, así como la comparación de resultados.

Seguidamente, se analizan tres de las propuestas metodológicas más recientes para la construcción de índices sintéticos. Las dos primeras (Monitor de Competitividad Turística y ESI) han tenido un nivel de acogida aceptable en el ámbito internacional, mientras que la tercera (SISD), que es una variación del ESI, supone la primera propuesta de agregación de indicadores realizada en España para la obtención de un índice sintético de desarrollo sostenible aplicada al análisis regional.

3.1. El Monitor de Competitividad Turística

El World Travel & Tourism Council (WTTC) es un organismo internacional que agrupa a las 100 empresas turísticas de titularidad privada más importantes del mundo, y que tiene por objeto concienciar a la población mundial del impacto que los viajes y el turismo tienen en la generación de riqueza y en la creación del empleo, apostando de forma decidida por un desarrollo turístico sostenible. En 2001 este organismo puso en funcionamiento un Monitor de Competitividad Turística que intenta medir hasta qué punto cada país ofrece un entorno competitivo para el desarrollo turístico. Este Monitor considera ocho grandes dimensiones de la competitividad turística (competitividad en precios, turismo humano, infraestructuras, medio ambiente, tecnología, apertura turística, desarrollo social y recursos humanos) que incluyen, a su vez, diversos indicadores de competitividad turística (65 en total) para casi 200 países de todo el mundo. Cada una de estas 65 variables es transformada en un índice que refleja la actuación relativa de cada país en relación con los demás. Esta normalización se lleva a cabo la metodología propuesta por el Programa de las Naciones Unidas para el Desarrollo (PNUD), es decir:

$$I_{ij} = \frac{x_{ij} - \min(x_{ij})}{\max(x_{ij}) - \min(x_{ij})}$$

siendo x_{ij} el valor alcanzado por la variable X_j en el país i . De esta forma, I_{ij} valdrá 1 para el país en el que se registra el máximo valor de la variable, y valdrá 0 para el país en el que observe el mínimo valor de la variable.

Dado que algunas variables turísticas están definidas de forma inversa (esto es, mayor competitividad cuanto menor sea el valor de la variable, como sucede, entre otros, con el índice de precios hoteleros o con los impuestos sobre bienes y servicios), es preciso utilizar la siguiente normalización inversa:

$$I_{ij} = \frac{\max(x_{ij}) - x_{ij}}{\max(x_{ij}) - \min(x_{ij})}$$

lo que garantiza un valor del índice igual a 1 para el país más competitivo y de 0 para el menos competitivo.

Una vez normalizadas todas las variables que sirven de base para el Monitor de Competitividad, el WTTC obtiene un índice agregado para cada una de las ocho dimensiones de la competitividad turística. Dicho índice se calcula simplemente de la siguiente forma:

$$S_i^{(k)} = \sum_{j=1}^m I_{ij}^{(k)}$$

siendo m el número de variables que integran cada dimensión de la competitividad turística, y donde (k) hace referencia a la k -ésima dimensión (para $k = 1$ a 8). Finalmente, el anterior índice agregado se normaliza de la siguiente forma:

$$C_i^{(k)} = \frac{S_i^{(k)} - \min(S_i^{(k)})}{\max(S_i^{(k)}) - \min(S_i^{(k)})}$$

De esta forma, se obtienen un total de ocho índices agregados de competitividad turística. Los principales inconvenientes de esta metodología son los siguientes:

- a) No se utilizan todas las variables originales para la obtención de estos índices agregados, debido a la falta de información estadística en un elevado porcentaje de países para algunas de dichas variables.

- b) El índice agregado se calcula como una suma simple de las variables estandarizadas, lo que significa que no se establecen ponderaciones para cada una de las variables que forman el índice agregado.
- c) A pesar de la obtención de ocho índices agregados (uno por cada dimensión de la competitividad turística), no se llega a proponer un único índice sintético que resuma la información aportada por los ocho índices agregados parciales.

Para salvar este último inconveniente, Gooroochurn & Sugiyarto (2005) proponen agregar los ocho índices parciales en un único índice ponderado de competitividad turística. Para ello, una vez obtenidos los valores estandarizados de las variables I_{ij} , obtienen el siguiente índice agregado para cada una de las ocho dimensiones de la competitividad:

$$y_i^{(k)} = \frac{\sum_{j=1}^{n_k} I_{ij}}{n_k}$$

siendo n_k el número de variables que componen cada dimensión. Una vez realizada esta agregación, el índice ponderado de competitividad turística que proponen es el siguiente:

$$z_i = \sum_{k=1}^8 \omega_k y_i^{(k)}$$

donde ω_k es el coeficiente de ponderación asignado a cada dimensión de la competitividad turística. Para determinar este coeficiente, los autores utilizan un análisis factorial confirmatorio, para lo cual consideran como variable (o dimensión) latente el índice agregado de competitividad turística y como variables observadas los ocho índices agregados $y_i^{(k)}$. Una vez obtenidos los coeficientes estimados asociados a las dimensiones consideradas, $\hat{\beta}_k$, los coeficientes de ponderación se calculan a través de la siguiente expresión:

$$\omega_k = \frac{|\hat{\beta}_k|}{\sum_{k=1}^8 |\hat{\beta}_k|}$$

Esta metodología para la obtención de índices sintéticos tiene la ventaja de que cada dimensión del fenómeno turístico analizado es ponderada mediante una metodología estadística robusta, como el análisis factorial confirmatorio. Sin embargo, los índices agregados parciales $y_i^{(k)}$ se obtienen sin diferenciar el diferente papel que cada variable tiene en la construcción de cada dimensión, al obtenerse dichos índices parciales como una media simple de variables normalizadas.

3.2. El Environmental Sustainability Index (ESI)

El Índice de Sostenibilidad Medioambiental (Environmental Sustainability Index, ESI) es elaborado, desde 2002, por el World Economic Forum (WEF) a partir de las bases metodológicas fijadas por el Yale Center for Environmental Law and Policy de la Universidad de Yale y por el Center for International Herat Science Information Network de la Universidad de Columbia. El último índice calculado hasta el momento, el correspondiente a 2005, está construido a partir de 76 variables, integradas a su vez en 21 indicadores de sostenibilidad medioambiental, de forma que cada indicador se construye a partir de entre 2 y 12 variables, dependiendo del caso y el valor final del ESI es una media igualmente ponderada⁴ de los valores alcanzados por los 21 indicadores calculados. Se calcula para un total de 146 países y considera 5 grandes dimensiones (sistemas medioambientales, reducción de las presiones sobre el medio ambiente, reducción de la vulnerabilidad humana a las presiones medioambientales, capacidad de respuesta social e institucional a los desafíos medioambientales y administración global). Además, el ESI debe ser interpretado en términos de probabilidad, puesto que el valor final de este índice para un país determinado “cuantifica la probabilidad de que dicho país sea capaz de preservar sus recursos

⁴ En opinión de los dos centros de investigación que elaboran este índice, ponderar de la misma forma a todos los indicadores calculados responde a varias razones. En primer lugar, a que una agregación simple es transparente y fácil de entender. En segundo lugar, a que, tras consultar a expertos medioambientales, ninguno de ellos se atrevió a conceder a unos indicadores una importancia sustancialmente mayor o menor que a los demás. Finalmente, a que la utilización de ciertos métodos estadísticos para obtener coeficientes de ponderación otorgan valores prácticamente iguales a los 21 indicadores. En cualquier caso, se reconoce que es probable que, cuando el análisis se realiza a nivel de un país, existan diferentes prioridades para los distintos componentes de la sostenibilidad.

medioambientales de forma efectiva a lo largo de varias décadas” y “evalúa la capacidad del país para evitar un mayor deterioro medioambiental” (ESI, 2005: 23).

Dado que las diferentes variables utilizadas en la construcción del ESI están expresadas en diferentes unidades de medida (especialmente, las variables medioambientales), el primer paso del proceso consiste en normalizar la totalidad de variables, para evitar las distorsiones que provocarían aquellas variables con varianzas elevadas. Este proceso de normalización de las variables se realiza mediante la siguiente expresión cuando exista una relación directa entre el valor de la variable y el nivel de sostenibilidad:

$$z_{ij} = \frac{x_{ij} - \bar{x}}{S_j}$$

siendo \bar{x} la media y S_j la desviación estándar de la variable, x_{ij} el valor de la variable X_j para la unidad de análisis i , y z_{ij} es el valor normalizado de x_{ij} .

Si, por el contrario, valores más elevados de la variable implican menor nivel de sostenibilidad (relación inversa), la normalización se lleva a cabo de la siguiente forma:

$$z_{ij} = \frac{\bar{x} - x_{ij}}{S_j}$$

Dado que el proceso de imputación y agregación de variables se basa en la asunción de una distribución normal multivariante, se comprueba hasta qué punto cada variable puede ser considerada normal, empleando para ello el coeficiente de asimetría. Así, todas aquellas variables cuyo coeficiente de asimetría sea superior a 2 es transformada a través de una función logarítmica o de una función potencial.

Esta transformación no sólo conseguirá que las variables altamente asimétricas corrijan su grado de asimetría, sino también que se reduzca el impacto de los valores extremos (*outliers*) de las variables. En cualquier caso, y para reducir totalmente dichos valores extremos (con las consecuencias negativas que los mismos tienen en el proceso de agregación de variables), la totalidad de las variables son sometidas a un proceso de “recorte” de las colas de sus distribuciones respectivas. Este “recorte” consiste en calcular los percentiles 2,5 y 97,5 y en sustituir todos los valores de la variable que superen el percentil 97,5 por el valor de dicho percentil y todos los valores que se sitúen

por debajo del percentil 2,5 por el valor de este último percentil. Es evidente que este recorte cambia la distribución de la variable y perjudica (o beneficia, según el caso) a aquellas unidades de análisis que tengan varios valores extremos⁵.

Una vez obtenidos los valores normalizados “recortados” (\tilde{z}_{ij}), el valor de cada indicador se obtiene como una suma ponderada de los valores \tilde{z}_{ij} de las variables que lo componen, es decir:

$$I_{ir} = \sum_{j=1}^p w_j \tilde{z}_{ij} \text{ para } r = 1, 2, \dots, 21$$

siendo p el número de variables que componen el r -ésimo indicador. El coeficiente de ponderación w_j es el mismo para la totalidad de las variables que forman el indicador.

Finalmente, el ESI es, simplemente, la suma igualmente ponderada de los 21 indicadores:

$$ESI_i = \sum_{r=1}^{21} \omega_r I_{ir}$$

Al objeto de facilitar la comparación entre los diferentes países, los valores ESI_i se convierten en el percentil normal estándar, que toma valores comprendidos entre 0 y 100.

Como se puede apreciar, la metodología utilizada en la construcción del ESI considera la misma ponderación para todas las variables que componen cada indicador y para todos los indicadores que forman el índice sintético final. En este sentido, es importante señalar que los autores del índice reconocen que “cuando el objetivo es diseñar el mejor índice posible, es importante considerar las técnicas estadísticas más avanzadas”. Pero, cuando la transparencia y la facilidad de interpretación por no expertos es lo prioritario, la metodología del ESI es una alternativa válida y útil (ESI, 2005: 65-66).

3.3. El Índice Sintético de Desarrollo Sostenible (Synthetic Index of Sustainable Development, SISD)

⁵ En cualquier caso, los autores del ESI justifican el uso de este “recorte” en el hecho de que el mismo afecta a un porcentaje mínimo de los datos y que, en consecuencia, sus beneficios superan ampliamente a sus inconvenientes.

A partir de la base metodológica del ESI, González et al (2004) proponen un índice sintético de desarrollo sostenible (SISD), que modifica ligeramente el proceso de cálculo del índice final. Tras normalizar y recortar las 78 variables que sirven de base para la construcción del SISD, el valor de cada indicador (se consideran 2 indicadores para la dimensión económica de la sostenibilidad; 2 para la dimensión institucional; 3 para la dimensión medioambiental; y 6 para la dimensión social) se calcula mediante la media simple de los valores \tilde{z}_{ij} de las variables que componen cada indicador⁶.

Finalmente, la última fase de obtención del SISD se lleva a cabo de la siguiente forma (González et al, 2004: 97):

- “a) Transformar los valores de la variable tipificada (...), de forma que puedan comprenderse y compararse. Para ello el z-score de cada indicador se convierte en el percentil normal estándar, con valor teórico comprendido entre 0 y 100.*
- b) Calcular los valores de los percentiles de cada componente o dimensión del SISD (económica, institucional, medioambiental y social) media la media ponderada de los percentiles calculados para los indicadores.*
- c) Por último, el valor del Índice SISD se obtiene a través de la ponderación de la media de los percentiles calculados para cada una de las dimensiones o componentes del índice.”*

Los resultados obtenidos tras su aplicación al análisis de la sostenibilidad en las Comunidades Autónomas españolas permiten a los autores del SISD asegurar que se trata de un instrumento válido para analizar y comparar la sostenibilidad entre diversos territorios, tanto de forma estática como dinámica, a través del análisis de la evaluación de los valores de la misma y de sus componentes para cada territorio.

4. Propuesta de un índice sintético de sostenibilidad turística mediante modelos factoriales: el ST INDEX

Recogiendo los aspectos metodológicos más interesantes de las anteriores propuestas para el desarrollo de un índice sintético, se presenta seguidamente una propuesta de

⁶ Cada indicador está compuesta, a su vez, por varios subindicadores, los cuales están formados por varias variables. Para conocer cómo se articulan las variables, los subindicadores, los indicadores y las

diseño de un índice sintético de sostenibilidad turística, que hemos bautizado con el nombre de ST INDEX (acrónimo de Sustainable Tourism Index).

Para ello, el punto de partida serán las cuatro grandes dimensiones de la sostenibilidad, es decir, la dimensión social, la dimensión medioambiental, la dimensión económica y, finalmente, la dimensión institucional. Para cada una de estas cuatro dimensiones, se considerarán un total de n_k (para $k = 1$ a 4) indicadores iniciales. Dada la diferente naturaleza de estos indicadores y las diferentes escalas de medida en que los mismos pueden venir expresados, el primer paso será normalizar todos estos índices de forma similar a como se realiza en el Monitor de Competitividad del WTTC, es decir: dado el indicador de sostenibilidad X_j , el valor que el mismo alcanza para el destino turístico i , x_{ij} se transformará en el valor normalizado y_{ij} , si el sentido de la sostenibilidad es directo (es decir, a mayor sostenibilidad, mayor valor del indicador), a través de la siguiente expresión

$$y_{ij} = \frac{x_{ij} - \min(x_{ij})}{\max(x_{ij}) - \min(x_{ij})}$$

empleándose, por el contrario, la siguiente expresión si el sentido de la sostenibilidad es inverso (menor valor del indicador implica grado mayor de sostenibilidad):

$$y_{ij} = \frac{\max(x_{ij}) - x_{ij}}{\max(x_{ij}) - \min(x_{ij})}$$

De esta forma, se garantiza un recorrido de los valores estandarizados y_{ij} entre 0 y 1, de forma que, cuanto más se aproxime a 1 dicho valor, mayor grado de sostenibilidad turística poseerá el destino i .

Seguidamente, para cada dimensión de la sostenibilidad, se utilizará un modelo de análisis factorial común para determinar la ponderación de cada indicador parcial en la obtención de un índice agregado. De esta forma, y admitiendo el carácter latente de cada dimensión, los coeficientes de ponderación φ_j vendrán dados por la correlación existente entre cada indicador parcial normalizado y la dimensión latente que explica el

dimensiones de la sostenibilidad, puede consultarse González et al (2004).

mayor porcentaje posible de varianza de dichos indicadores (esto es, $\phi_j = \alpha_{j1}$; siendo α_{j1} el coeficiente de correlación entre el indicador normalizado Y_j y el primer factor común extraído). Así, por ejemplo, el índice agregado de la dimensión social de la sostenibilidad turística vendría dado por la siguiente expresión:

$$I_S = \alpha_{11}Y_1 + \alpha_{21}Y_2 + \dots + \alpha_{n_k1}Y_{n_k}$$

Con este sistema de ponderaciones, se garantiza que los indicadores más correlacionados con la dimensión social de la sostenibilidad turística tengan mayor peso en el cálculo del correspondiente índice agregado.

En el caso de que el primer factor explique un escaso porcentaje de las varianzas de los indicadores que formarán el índice agregado será necesario considerar factores adicionales. La consideración de más de un factor común introduce la necesidad, en la mayor parte de los casos, de una rotación factorial⁷ que garantice que cada indicador esté altamente correlacionado únicamente con uno de los factores extraídos. Cuando se lleve a cabo una rotación factorial, los coeficientes de ponderación ϕ_j se obtendrán de la matriz de cargas factoriales rotadas (matriz de coeficientes β_{jr} , para $j=1,2,\dots,n_k$ y para $r=1,2,\dots,m$), por lo que, si se considera un total de m factores comunes, se obtendrían las siguientes expresiones intermedias:

$$G_1 = \beta_{11}Y_1 + \beta_{21}Y_2 + \dots + \beta_{n_k1}Y_{n_k}$$

$$G_2 = \beta_{12}Y_1 + \beta_{22}Y_2 + \dots + \beta_{n_k2}Y_{n_k}$$

...

$$G_m = \beta_{1m}Y_1 + \beta_{2m}Y_2 + \dots + \beta_{n_km}Y_{n_k}$$

A partir de las expresiones anteriores, el índice agregado para la dimensión social de la sostenibilidad turística se obtendrá de la siguiente forma:

$$I_S = \omega_1G_1 + \omega_2G_2 + \omega_3G_3 + \dots + \omega_mG_m$$

⁷ El método de rotación más utilizado en la práctica es el método varimax normalizado, que garantiza que, una vez rotados, los factores extraídos seguirán siendo mutuamente ortogonales, como lo son los factores antes de la rotación.

donde $\omega_i = \frac{\lambda_i^*}{\sum_{i=1}^m \lambda_i^*}$; para $i = 1, 2, \dots, m$, siendo λ_i^* ($i = 1, 2, \dots, m$) el autovalor asociado al

factor común rotado G_i .

Puesto que los valores de estos índices agregados pueden ser tanto positivos como negativos, al objeto de facilitar la interpretación de los mismos y, sobre todo, de permitir un análisis comparativo entre destinos turísticos, el índice agregado obtenido será transformado mediante la función de Casalmiglia (1990), que viene dada por la siguiente expresión:

$$S_s = \begin{cases} 1 + \frac{k-1}{2} \exp(I_s) & \text{si } I_s < 0 \\ k - \frac{k-1}{2} \exp(-I_s) & \text{si } I_s \geq 0 \end{cases}$$

De esta forma, el nuevo índice transformado, S_s , tomará siempre valores comprendidos entre 0 y k (un valor de k especialmente interesante a efectos de interpretación es el valor 100).

Una vez obtenidos los índices agregados transformados para las cuatro dimensiones de la sostenibilidad turística, S_s , S_m , S_e e S_l se puede recurrir a un análisis factorial confirmatorio (como proponen Gooroochurn y Sugiyarto, 2005) para cuantificar el peso relativo de cada dimensión en la medida de la sostenibilidad turística, siempre y cuando el número de destinos turísticos analizado sea suficientemente grande. De esta forma, y una vez estimados los correspondientes coeficientes, el índice sintético de sostenibilidad turística que proponemos (el ST INDEX) vendría dado por la siguiente expresión:

$$ST_i = \omega_1 S_{Si} + \omega_2 S_{Mi} + \omega_3 S_{Ei} + \omega_4 S_{Li}$$

donde $\omega_k = \frac{|\hat{\beta}_k|}{\sum_{k=1}^4 |\hat{\beta}_k|}$, para $k = 1, 2, 3, 4$ y verificándose que $\sum_{k=1}^4 \omega_k = 1$.

Como complemento al diseño propuesto, es interesante determinar hasta qué punto los resultados obtenidos con cada uno de los índices sintéticos transformados y, a partir de ellos, con el índice sintético ST INDEX son congruentes con los indicadores

normalizados de sostenibilidad turística Y_j obtenidos en el primer paso del diseño propuesto. Para ello, se calculará primero la matriz \mathbf{M}_k , que computará las distancias euclídeas cuadradas entre cada par de unidades turísticas analizadas considerando únicamente los n_k indicadores de sostenibilidad utilizados para cuantificar esa dimensión concreta de la sostenibilidad turística, empleando para ello la siguiente expresión:

$$m_{rs} = \sum_{j=1}^{n_k} (y_{rj} - y_{sj})^2 \text{ para } r, s = 1, 2, \dots, w$$

Si se consideran un total de w destinos turísticos, la matriz \mathbf{M}_k será una matriz cuadrada y simétrica de orden w . Seguidamente, se calculará la matriz \mathbf{S}_k , que cuantificará la distancia entre unidades turísticas respecto al índice agregado de la dimensión de la sostenibilidad que se esté considerando. Así, por ejemplo, los elementos de la matriz \mathbf{S}_k , que es también una matriz cuadrada y simétrica de orden w , se calcularán de la siguiente forma si se está considerando la dimensión social de la sostenibilidad turística:

$$s_{rs} = (S_{Sr} - S_{Ss})^2 \text{ para } r, s = 1, 2, \dots, w$$

Pues bien, la correlación existente entre los elementos de ambas matrices se puede utilizar como indicador del grado de coherencia de los indicadores calculados. De esta forma, si esta correlación $r_{\mathbf{M},\mathbf{S}}$ es elevada se podrá concluir que el índice agregado de la dimensión social de la sostenibilidad S_s está correctamente construido puesto que está altamente relacionado con los indicadores utilizados para cuantificar dicha dimensión.

5. Conclusiones

La construcción de índices sintéticos se percibe como una solución ideal, aunque no exenta de crítica, para el análisis de la sostenibilidad, toda vez que la realidad multidimensional que contempla este paradigma dificulta su percepción desde una sola óptica. Así, el ST INDEX, cuyo diseño se ha presentado en este trabajo, viene a responder a esta necesidad de análisis multidimensional, permitiendo normalizar los diferentes indicadores utilizados hasta ahora y las distintas escalas de medida en que están expresados.

La metodología presentada contempla la consideración del carácter latente de la cuádruple dimensión de la sostenibilidad utilizando un modelo de análisis factorial común para determinar la ponderación de cada indicador parcial en la obtención de un índice agregado. De esta manera, se garantiza que los indicadores más correlacionados con cada una de las dimensiones tengan mayor peso en el posterior cálculo del índice agregado.

Además, se ha propuesto el cálculo de un indicador del grado de coherencia de los índices calculados que permitirá determinar si éstos están bien contruidos, lo asegura la correcta aplicación del ST INDEX.

El próximo paso, en el que trabajan actualmente los autores de este trabajo, será la aplicación del ST INDEX para el análisis de diferentes tipologías de destinos turísticos, a fin de validar la metodología propuesta y la coherencia de los resultados obtenidos. Esta tarea permitirá, además, mejorar en un futuro el índice propuesto. De hecho, ya pueden adelantarse algunos de los retos que se están afrontando:

a) Tratamiento de los datos perdidos (*missing data*): la metodología expuesta presupone la existencia de datos para la totalidad de indicadores y de destinos turísticos analizados. Sin embargo, en muchas ocasiones los datos no están disponibles, lo que obliga a desarrollar una metodología para el tratamiento de los datos perdidos, puesto que, de lo contrario, habría que prescindir tanto del indicador como del destino turístico al que se refiera ese dato perdido.

b) Dinamizar la metodología: el ST INDEX es un índice sintético de carácter estático, en el sentido de que ofrece una visión de la sostenibilidad turística referida a un único período de tiempo (normalmente, a un año). Pero es interesante arbitrar mecanismos que permitan establecer comparaciones intertemporales. Otra cuestión preocupante es la relativa a los cambios en las ponderaciones en los distintos años, es decir, hasta qué punto las ponderaciones son estables en el tiempo o, por el contrario, son relativamente volátiles.

c) Posibilidad de combinar este método, en las diferentes dimensiones de la sostenibilidad, con los planteamientos de los modelos PSR o DPSIR, que parecen ser los generalmente más aceptados para explicar las relaciones de causalidad en la sostenibilidad. Así, por ejemplo, el “Sistema Español de Indicadores Ambientales de

Turismo” incluye 27 indicadores que se ajustan a cada componente de los modelos anteriores, pero no especifica a qué dimensión concreta de la sostenibilidad (social, medioambiental, económica o institucional) se refieren.

Bibliografía

Bermejo, R. y Nebreda, A. (1998): “Conceptos e instrumentos para la sostenibilidad local”. *Cuadernos Bakeaz*, 26, pp. 1-18.

Blackstock K.L., McCrum G., Scott A. y White V. (2006): *A Framework For Developing Indicators of Sustainable Tourism*. Macaulay Institute.

Blackstock K.L., Scott, A.J., White V. y McCrum G. (2006): *Indicators and Sustainable Tourism: Summary of Interview Findings*. Macaulay Institute.

Bloyer, J., Gustke, L. y Leung, Y. (2004): “Indicators for sustainable tourism development: Crossing the divide from definition to actions”. En Pineda, F.; Brebbia, C. y Múgica, M (eds.): *Sustainable Tourism*. Southampton: WiT, pp. 109-15.

Bossel, H. (1999): “Indicators for Sustainable Development: Theory, Method, Applications”. A report to the Balaton Group. International Institute for Sustainable Development.

Briassoulis, H. (2001): “Sustainable Development and its Indicators: Through a (Planner’s) Glass Darkly”. *Journal of Environmental Planning and Management*, 44 (3), 409-427.

Calsamiglia, X. (1990): “La financiación de las Comunidades Autónomas y el principio de solidaridad”. *Economía Pública*, 6, pp. 3-43.

Castell, C. (2000): “Los programas de seguimiento ecológico en espacios naturales protegidos”. En AA. VV.: *Actas del 5º Congreso EUROPARC España*, Fundación Fernando González Bernáldez, Madrid, pp. 30-36.

Dourojeanni, A. (1997): *Procedimientos de gestión para un desarrollo sostenible*, ONU. Comisión Económica para América Latina y el Caribe (CEPAL), Santiago de Chile.

Fricke, A. (1998): “Measuring up to sustainability”. *Futures*, 39 (4), pp. 367-375.

Fullana, P. y Ayuso, S. (2002): *Turismo sostenible*, Rubes Editorial, Barcelona.

González, F., Martín, F. y Fernández, M. (2004): “Medición del desarrollo sostenible y análisis regional: diseño y aplicación de un índice sintético global a las comunidades autónomas españolas”. *Investigaciones Regionales*, 5, pp. 91-112.

Gooroochurn, N. y Sugiyarto, G. (2005): “Competitiveness indicators in the travel and tourism industry”. *Tourism Economics*, 11 (1), pp. 25-43.

Guy, G.B. and Kibert, C.J. (1998): “Developing Indicators of Sustainability: US Experience”. *Building Research and Information*, 26 (1), pp. 39-45.

Hall, C. M. (2000): *Tourism planning. Policies, processes and relationships*, Prentice Hall, Essex.

- Hass, J.L.; Brunvoll, F. y Hoie, H. (2003): *Overview of Sustainable Development Indicators Used by National and International Agencies*. OECD Statistics Working Papers, 2002/2, OECD Publishing.
- Hockings, M. et al (2000): *Evaluating effectiveness a framework for assessing the management of protected areas*, UICN, Gland y Cambridge.
- Hughes, G. (2002): "Environmental indicators". *Annals of Tourism Research*, 29 (2), pp. 457-477.
- Ivars, J.A. (coord.) et al (2001): *Planificación y gestión del desarrollo turístico sostenible: propuestas para la creación de un sistema de indicadores (proyecto METASIG)*, Universidad de Alicante, Alicante.
- James, D. (2000): "Local sustainable tourism indicator". En *5th International Forum on Tourism Statistics*, Caledonian University Conference Proceeding, Glasgow.
- James, D. (2004): "Local sustainable tourism indicador". *Estudios Turísticos*, nº 161-162, pp. 219-232.
- Jobson, J.D. (1992): *Applied Multivariate Data Analysis. Volume II: Categorical and Multivariate Methods*. Editorial Springer-Verlag, New York, pp. 388-426.
- Ko, J.T.G. (2001): "Assessing progress of tourism sustainability". *Annals of Tourism Research*, 28 (3), pp. 817-820.
- Liu, Z. (2003): "Sustainable Tourism Development: A Critique". *Journal of Sustainable Tourism*, 11 (6), pp. 459-475.
- Manning, T. (1999): "Indicators of Tourism Sustainability". *Tourism Management*, 20 (1), pp. 179-181.
- Miller, G. (2001): "The development of indicators for sustainable tourism: results of a Delphi survey of tourism researchers". *Tourism Management*, 22 (4), 351-362.
- Ministerio de Medio Ambiente (1996): *Indicadores ambientales. Una propuesta para España*. Madrid.
- Ministerio de Medio Ambiente (2001): *Sistema español de indicadores ambientales: área de medio urbano*. Serie Monografías, Madrid.
- Ministerio de Medio Ambiente (2003): *Sistema español de indicadores ambientales de turismo*. Madrid.
- OCDE (1991): *Environmental Indicators. A preliminary Set*. París.
- OCDE (1993): *OECD core set of indicators for environmental performance reviews*. Environment Monographs n. 83. París.
- OCDE (1997): *Better understanding our cities. The rule of urban indicators*. OCDE, París.
- OCDE (2000): *Frameworks to measure sustainable development*. OECD Expert Workshop, París.
- OCDE (2003): *Environmental indicators: development, measurement and use*. Paris.

- Organización Mundial del Turismo (1995): *Lo que todo gestor turístico debe saber. Guía práctica para el desarrollo y uso de indicadores de turismo sostenible*, OMT, Madrid.
- Organización Mundial del Turismo (2004): *Indicators of sustainable development for tourism destinations: a guidebook*. OMT, Madrid.
- Pintér, L.; Hardi, P. y Bartelmus, P. (2005): *Indicators of Sustainable Development: Proposals for a Way Forward*. Discussion Paper Prepared under a Consulting Agreement on behalf of the UN Division for Sustainable Development. *mimeo*
- Pulido, J. I. (2004): “Turismo y medio ambiente. Planteamientos para la reorientación medioambiental de la política turística española”. En Torres, E. y Pulido, J. I. (coord.): *Quaderns de Política Econòmica*, núm. 7. Monografía “La política turística en España”. Valencia, pp. 93-114.
- Pulido, J. I. (2006a): “Turismo y sostenibilidad”. En Sáez, A., Martín, P. y Pulido, J. I. (coord.): *Estructura económica del turismo*. Editorial SÍNTESIS. Madrid, pp. 327-366.
- Pulido, J. I. (2006b): “Sostenibilidad de los destinos turísticos”. En López, D. y Pulido, J. I. (dir.): *La actividad turística española en 2005 (Edición 2006)*, Centro de Estudios Ramón Areces, S. A. y Asociación Española de Expertos Científicos en Turismo (AECIT). Jaén, pp. 279-295.
- Sánchez, M. (2007): “Construcción de índices sintéticos de sostenibilidad turística mediante modelos factoriales: aplicación al Sistema Español de Indicadores Ambientales de Turismo”. Comunicación XI Congreso de AECIT. *mimeo*
- Sancho, A. et al (2002): *Auditoría de sostenibilidad en los destinos turísticos*. Instituto de Economía Internacional. Universidad de Valencia, Valencia.
- Sancho, A. y García, G. (2006): “¿Qué indica un indicador? Análisis comparativo en los destinos turísticos”. *Revista de Análisis Turístico*, 2, pp. 69-85.
- Segnestam, L. (2002): *Indicators of Environment and Sustainable Development. Theories and Practical Experience*. The World Bank. Environmental Economics Series. Paper N° 69.
- Sierra, V. (2002): “Desarrollo sostenible: acotaciones conceptuales y revisiones estratégicas”. *Boletín Económico ICE*, nº 2.749, pp. 13-23.
- Sirakaya, E.; Jamal, T.B. y Choy, H.-S (2001): “Developing Indicators for Destination Sustainability”. En Weaver, D.B. (ed.): *The Encyclopedia of Ecotourism*. CAB International, pp. 411-432.
- Swarbrooke, J. (2000): *Sustainable tourism management*, CABI Publishing, Oxon.
- United Nations (2001a): *Indicators of sustainable development. Framework and Methodologies*. Department of Economic and Social Affairs. Division For Sustainable Development. Background Paper for the Ninth Session of the Commission on Sustainable Development. New York.
- United Nations (2001b): *Report on the Aggregation of Indicators of Sustainable Development*. U. N. Department of Economic and Social Affairs. Division for Sustainable Development. Background Paper for the Ninth Session of the Commission on Sustainable Development. New York.

United Nations (2006): Global Trends and Status of Indicators of Sustainable Development. Department of Economic and Social Affairs. Division For Sustainable Development. Background Paper for the Fourteenth Session of the Commission on Sustainable Development. New York.

Uriel, E. (1995): *Análisis de datos: series temporales y análisis multivariante*, pp. 343-379. Editorial AC. Madrid.

Vera, J.F. y Ivars, J.A. (2003): “Measuring sustainability in a mass tourist destination: pressures, perceptions and policy responses in Torre Vieja, Spain”. *Journal of Sustainable Tourism*, 11(2-3), pp. 181-203.

Weaver, D. (2005): *Sustainable tourism*, Butterworth-Heinemann, Oxford.

White V., McCrum G., Blackstock K.L. y Scott A. (2006): *Indicators and Sustainable Tourism: Literature Review*. Macaulay Institute.

World Bank (2002): *World Development Report 2003: Sustainable development in a dynamic world: transforming institutions, growth, and quality of life*. Oxford University Press-World Bank, New York.

World Economic Forum (2005): 2005 Environmental Sustainability Index. <http://sedac.ciesin.columbia.edu/es/esi/ESI2005.pdf>

Yunis, E. (2003): “El marco internacional de la gestión sostenible del turismo: la visión de la Organización Mundial del Turismo”. A *DISTANCIA. Monográfico “Turismo y sostenibilidad”*, 21 (1), pp. 16-21.

El Índice de Consistencia Geométrico. Umbrales exactos y empíricos

Juan Aguarón Joven
e-mail: aguaron@unizar.es

José María Moreno Jiménez
e-mail: moreno@unizar.es

Departamento de Estructura e Historia Económica y Economía Pública
UNIVERSIDAD DE ZARAGOZA

Área temática: Métodos Cuantitativos

Resumen

Este trabajo, presenta dos formas para obtener los umbrales del Índice de Consistencia Geométrico (Geometric Consistency Index, GCI) asociados con los de la correspondiente Razón de Consistencia (Consistency Ratio, CR). El primero proporciona para el caso $n=3$ los umbrales del GCI de una manera exacta, utilizando para ello las relaciones analíticas existentes entre las dos medidas de inconsistencia (umbrales exactos). El segundo proporciona en un caso general ($n = 3...16$) los umbrales para el GCI mediante un estudio de simulación (umbrales empíricos).

Palabras clave: Multicriterio, AHP, Priorización, Consistencia

Abstract

This paper, dedicated to obtaining the thresholds for the Geometric Consistency Index (GCI) associated with the corresponding of the Consistency Ratio (CR), presents two ways to derive these values. The first, Exact Method, provides in an exact way for the case of $n=3$, the thresholds for the GCI by using the derived analytical relation between the GCI and the CR (exact thresholds). The second, Empirical Method, provides in a general situation ($n = 3, \dots, 16$) the thresholds for the GCI by way of a simulation study (empirical thresholds).

Key words: Multicriteria, AHP, Priorization, Consistency

1. Introducción

Uno de los aspectos o propiedades más destacados a la hora de comparar las diferentes Técnicas de Decisión Multicriterio, es la posibilidad de evaluar el grado de consistencia que tiene el decisor, en general los actores participantes en el proceso de toma de decisiones, a la hora de incorporar sus preferencias en el modelo utilizado para la resolución del problema.

En este sentido, el Proceso Analítico Jerárquico (AHP) propuesto por Thomas Saaty (Saaty, 1977, 1980) es una técnica de decisión multicriterio que permite evaluar la consistencia del decisor al emitir los juicios correspondientes a los elementos de las matrices recíprocas de comparaciones pareadas a través de las cuales incorpora al modelo su estructura de preferencias.

Dada una matriz recíproca de comparaciones pareadas $A_{n \times n} = (a_{ij})$, entendemos por *consistencia* la transitividad cardinal de los juicios de la esto es:

$$a_{ij}a_{jk} = a_{ik} \quad i, j, k = 1 \dots n$$

Saaty sugirió que, cuando se usa el método del vector propio principal (EVM) como procedimiento de priorización, una medida de la inconsistencia podía ser capturada por el Índice de Consistencia:

$$CI = \frac{\lambda_{\max} - n}{n - 1}$$

Finalmente, para estandarizar dicho valor, se construye la Razón de Consistencia (CR), obtenida como el cociente entre CI y el Random Index (RI) que no es sino el valor esperado de la Razón de Consistencia para cada valor de n ¹:

$$RC = \frac{CI}{RI(n)} = \frac{\lambda_{\max} - n}{RI(n)(n - 1)}$$

Las prioridades obtenidas a partir de la matriz de juicios son admitidas si el valor de RC es inferior al 10% (5% y 8% para matrices de orden 3 y 4 respectivamente). En caso contrario, es preciso realizar una revisión de los juicios (Saaty, 1994).

¹ El Random Index se define como la esperanza de la Razón de Consistencia cuando los juicios de la matriz siguen distribuciones discretas uniformes en los valores $\{1/9, 1/8, \dots, 1/2, 1, 2, \dots, 8, 9\}$

Aunque el EVM es el procedimiento de priorización empleado habitualmente en AHP, en los últimos años, debido a sus propiedades analíticas y psicológicas (Barzilai y Lootsma, 1997; Bruga, 2000; Aguarón y Moreno, 1999; Escobar y Moreno-Jiménez, 2000), está empezando a extenderse la utilización del Método de la Media Geométrica por Filas (RGMM) que proporciona (salvo factor de normalización) las prioridades de las alternativas como

$$\omega_i = \left(\prod_{j=1}^n a_{ij} \right)^{1/n}$$

Aguarón y Moreno (2003) proponen el Índice de Consistencia Geométrico (GCI) para medir la inconsistencia cuando se emplea el RGGM como método de priorización:

$$GCI = \frac{2}{(n-1)(n-2)} \sum_{i < j} \log^2 e_{ij} = \frac{2}{(n-1)(n-2)} \sum_{i < j} \log^2 \frac{a_{ij} \omega_j}{\omega_i}$$

Además, Aguarón y Moreno (2003) demuestran que, cuando la inconsistencia es baja, se puede utilizar una aproximación lineal entre la Razón de consistencia de Saaty y el Índice de Consistencia Geométrica. A partir de esta relación lineal se obtiene unos umbrales aproximados para el CGI equivalentes a los del CR:

Tabla 1. Umbrales aproximados del GCI

CR	0.01	0.05	0.1	0.15
GCI (n=3)	0.0314	0.1573	0.3147	0.4720
GCI (n=4)	0.0352	0.1763	0.3526	0.5289
GCI (n>4)	~0.037	~0.185	~0.370	~0.555

Cuando la inconsistencia aumenta, dicha relación no es bastante precisa y proporciona ligeras sobreestimaciones de los verdaderos umbrales, debido a la concavidad de la relación real entre ambas medidas. Para establecer estos umbrales de una forma más precisa, en este trabajo presentamos dos aproximaciones: la primera es válida exclusivamente para el caso $n = 3$, y proporciona el umbral exacto para el CGI asociado con cualquier valor del CR de Saaty empleando una expresión cerrada; en la segunda proporcionamos los umbrales por medio de un estudio de simulación.

La expresión exacta se ha obtenido exclusivamente para $n=3$. Sin embargo, hemos incluido este resultado por la importancia del caso, debido a sus características prácticas y matemáticas. Muchas de las aplicaciones reales incluyen en su nivel más general de la jerarquía tres elementos o de nodos (por ejemplo, criterios económicos, sociales y

ambientales). Además, para $n = 3$, las prioridades obtenidas con el método de EGV y el método RGMM coinciden.

Para $n > 3$, no hemos podido obtener expresiones cerradas para la relación entre el GCI y el CR. Incluso, en muchos casos no está garantizada la relación biunívoca entre ambas, como probaremos en este trabajo. Por ello, hemos desarrollado un estudio de simulación generando 1.200.000 matrices recíprocas, calculando los correspondientes GCI y CR. A partir de estos pares de valores (GCI, CR), hemos estimado para $n=3, \dots, 16$, los umbrales del GCI asociados a ciertos valores del CR.

La estructura del es la siguiente: la sección 2 presenta el método exacto para el caso de $n=3$; la sección 3 incluye el método empírico y, finalmente, la sección 4 recoge las conclusiones más importantes del trabajo.

2. Umbrales exactos para el GCI

Para obtener de manera precisa los umbrales del GCI asociados con cualquier valor del CR, comenzamos obteniendo, para el caso $n=3$, la relación exacta entre ellos.

Lema 1 Para $n=3$ se cumple

$$GCI = 3[\operatorname{acosh}(CR \times RI(3) + 1)]^2$$

DEMOSTRACIÓN. Si $n=3$, los dos procedimientos de priorización proporcionan el mismo vector de prioridades y, por lo tanto, la misma matriz de residuos. Además, esta matriz tiene la forma:

$$\begin{pmatrix} 1 & a & 1/a \\ 1/a & 1 & a \\ a & 1/a & 1 \end{pmatrix}$$

Las dos medidas de inconsistencia son: $GCI = 3 \log^2 a$ y $CR = \frac{a + 1/a - 2}{2RI(3)}$

y el CR se puede expresar como

$$CR = \frac{1}{RI(3)} \left[\frac{1}{2}(a + 1/a) - 1 \right] = \frac{1}{RI(3)} [\operatorname{cosh}(\log a) - 1]$$

de donde $\log a = \operatorname{arcosh}(CR \times RI(3) + 1)$ y, despejando $\log a$ de la expresión del GCI, obtenemos $GCI = 3[\operatorname{acosh}(CR \times RI(3) + 1)]^2$ #

A partir de esta relación, los valores exactos del GCI asociados a los correspondientes valores del CR vienen indicados en la Tabla 2. Como se puede observar en la Figura 1, esta relación es aproximadamente lineal.

Tabla 2. Umbrales exactos del GCI para n=3

CR	0.01	0.03	0.05	0.08	0.10	0.15	0.20	0.25
GCI	0.0314	0.0942	0.1566	0.2500	0.3199	0.4660	0.6187	0.7701

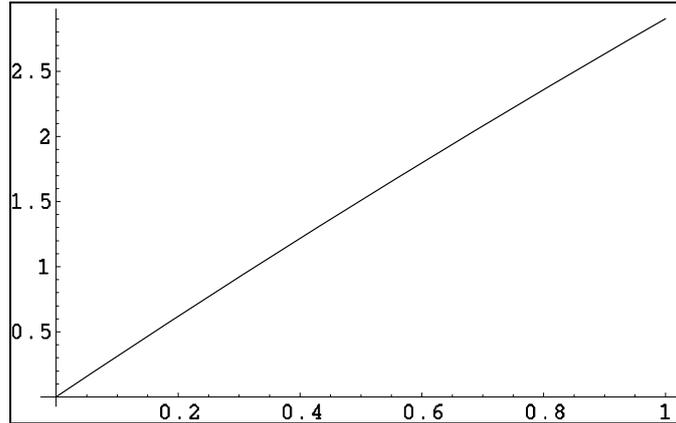


Figura 1. Valores del CR y GCI para n=3

A pesar de la limitación de este resultado, válido solamente para $n = 3$, su importancia práctica y el hecho de que las prioridades obtenidas por con el método EVM coinciden con las obtenidas con el RGMM, resulta interesante su consideración.

Lema 2 Para $n = 4$ y $a > 0$, las dos matrices A y B siguientes

$$A = \begin{pmatrix} 1 & a & 1 & 1/a \\ 1/a & 1 & a & 1 \\ 1 & 1/a & 1 & a \\ a & 1 & 1/a & 1 \end{pmatrix} \quad B = \begin{pmatrix} 1 & b & b & b \\ 1/b & 1 & c & 1/c \\ 1/b & 1/c & 1 & c \\ 1/b & c & 1/c & 1 \end{pmatrix}$$

donde b y c están definidas por

$$b = \frac{1}{3} \left(1 + a + \frac{1}{a} \right) \quad (1)$$

$$3b = \frac{1}{b} + c + \frac{1}{c} \quad (2)$$

tienen el mismo CR y diferente GCI.

DEMOSTRACIÓN. Ver Aguarón y Moreno 2007.

Lema 3 Para $n > 4$ es posible construir dos matrices recíprocas de orden n con el mismo CR y diferentes GCI

DEMOSTRACIÓN. Ver Aguarón y Moreno 2007.

3. Umbrales empíricos para el GCI

Para obtener los umbrales para el GCI con mayor precisión, hemos desarrollado tres experimentos de simulación diferentes, basados en 1.200.000 matrices simuladas y los correspondientes pares (GCI_i, CR_i) calculados a partir de ellas.

- El primero proporciona los valores tabulados para el GCI correspondientes a un valor específico del CR, una vez que éstos han sido ordenados. En este caso los valores obtenidos se recogen en la Tabla 3.

Tabla 3. Valores tabulados del GCI para diferentes CR

n	CR				
	0.01	0.05	0.1	0.15	0.2
3	0.0314	0.1567	0.3120	0.4660	0.6187
4	0.0351	0.1729	0.3406	0.5071	0.6486
5	0.0369	0.1825	0.3457	0.5267	0.6659
6	0.0373	0.1815	0.3407	0.5224	0.6818
7	0.0373	0.1828	0.3586	0.5187	0.6919
8	0.0372	0.1818	0.3533	0.5107	0.6904
9	0.0371	0.1814	0.3494	0.5222	0.6734
10	0.0368	0.1780	0.3447	0.5157	0.6682
11	0.0367	0.1796	0.3469	0.5126	0.6684
12	0.0366	0.1784	0.3504	0.5070	0.6743
13	0.0365	0.1757	0.3513	0.5113	0.6649
14	0.0364	0.1771	0.3501	0.5115	0.6611
15	0.0363	0.1779	0.3470	0.4991	0.6578
16	0.0362	0.1772	0.3459	0.5067	0.6592

- El segundo proporciona los umbrales del GCI asociados con un valor dado para el CR, a , como un promedio de las proyecciones del GCI sobre la recta $CR = a$ correspondientes a los valores del CR en el intervalo $[a - \delta, a + \delta]$. Las

proyecciones vienen dadas por $G\hat{C}I_i = \frac{GCI_i}{CR_i} a$ (ver Figura 2)

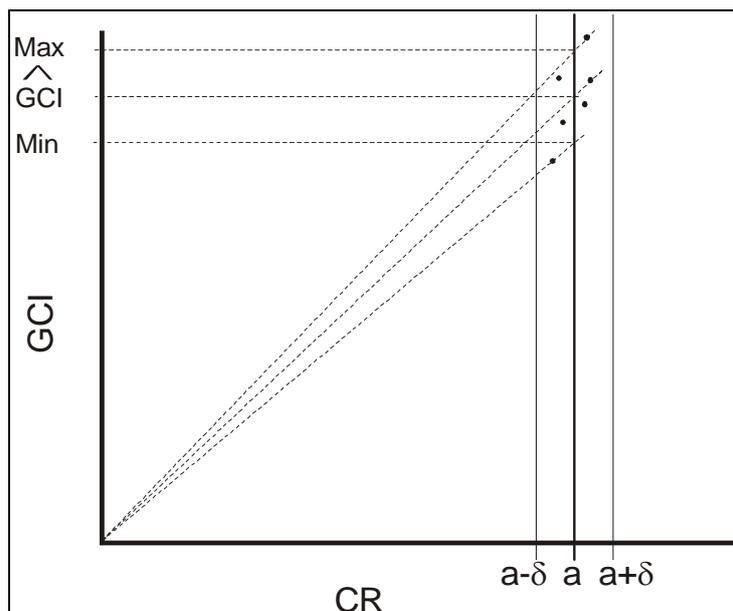


Figura 2 Proyecciones del GCI sobre la línea $CR=a$

A partir de estas proyecciones calculamos el mínimo, máximo y promedio de los valores $G\hat{C}I_i$ asociados con los umbrales $CR = a$. Estos valores, para $a = 0.10$ y $\delta = 0.001$ se muestran en la Tabla 4.

Tabla 4. Rangos del GCI para $CR = 0.10$

n	OBS	MIN	AVG	MAX
4	535	0,3300	0,3396	0,3508
5	616	0,3341	0,3525	0,3673
6	689	0,3265	0,3527	0,3692
7	702	0,3250	0,3532	0,3680
8	694	0,3244	0,3519	0,3664
9	711	0,3259	0,3517	0,3648
10	680	0,3320	0,3491	0,3626
11	613	0,3299	0,3489	0,3613
12	594	0,3334	0,3478	0,3587
13	535	0,3325	0,3471	0,3577
14	539	0,3322	0,3463	0,3572
15	411	0,3341	0,3458	0,3557
16	384	0,3352	0,3456	0,3555

- Finalmente, en el tercer experimento calculamos los umbrales a partir de los coeficientes de regresión estimados para los pares (GCI_i, CR_i) con $CR_i \in [a - \delta, a + \delta]$. Como se puede ver, por ejemplo en el caso $CR=0.10$, los coeficientes de regresión son muy cercanos (ver) para pequeñas modificaciones de δ (Tabla 5).

Tabla 5. Regresiones entre GCI y CR entorno a CR=0.1

n	0.09-0.11			0.08-0.12		
	N	m	r^2	N	m	r^2
3	3309	3.120	1.000	6589	3.120	1.000
4	5039	3.401	0.983	10177	3.400	0.995
5	6211	3.522	0.911	12357	3.520	0.974
6	6875	3.531	0.898	13745	3.530	0.969
7	6854	3.526	0.918	13770	3.522	0.974
8	7044	3.518	0.932	13876	3.515	0.979
9	6844	3.513	0.949	13622	3.509	0.985
10	6888	3.495	0.960	13643	3.491	0.988
11	6674	3.489	0.969	13624	3.484	0.990
12	6498	3.479	0.973	13718	3.475	0.992
13	6132	3.473	0.978	13740	3.469	0.993
14	6138	3.465	0.981	14168	3.462	0.994
15	5633	3.459	0.983	14040	3.456	0.995
16	5498	3.454	0.985	14557	3.450	0.996

De forma similar hemos obtenido los umbrales empíricos para el GCI asociados con diferentes valores del CR. Estos resultados se pueden ver en la Tabla 6

Tabla 6. Umbrales empíricos para el GCI

CR	0.01	0.05	0.1	0.15	0.2
<i>GCI</i> (n=3)	0.0314	0.1567	0.3120	0.4660	0.6187
<i>GCI</i> (n=4)	0.0351	0.1729	0.3406	0.5071	0.6486
<i>GCI</i> (n>4)	~0.037	~0.176	~0.330	~0.480	~0.637

4. Conclusiones

El empleo del método de la Media Geométrica por Filas (RGMM) presenta ventajas con respecto al clásico método del vector propio (EVM), especialmente desde el punto de vista de sus propiedades analíticas y operativas. Sin embargo, su uso estaba limitado por la falta de una medida de la inconsistencia acorde con la naturaleza del método.

La propuesta del Índice de Consistencia Geométrico ha permitido evaluar dicha inconsistencia, pero el establecimiento de umbrales adecuados para esta medida se limitaba al empleo de una relación aproximada entre las dos medidas, CR y GCI, (Aguaron y Moreno 2003).

En el presente trabajo hemos obtenido los umbrales exactos para el caso n=3 y hemos demostrado que para n>3 no existe una relación exacta (ni siquiera única) entre ambas medidas. No obstante, para estas situaciones se ha determinado empíricamente los umbrales promedios del GCI equivalentes al CR. Además, se ha observado que el

intervalo de variación del GCI es muy pequeño, por lo que se pueden dar por buenos los umbrales obtenidos y, en consecuencia, es posible medir la inconsistencia de una forma análoga y equivalente a la realizada con el CR.

Agradecimientos

Este trabajo ha sido parcialmente sufragado por el proyecto del Ministerio de Educación y Ciencia “Herramientas Decisionales para la e-cognocracia” (ref. TSI2005-02511).

Bibliografía

Aguarón, J.; Moreno-Jiménez, J.M. (2000): Local Stability Intervals in the Analytic Hierarchy Process, *European Journal of Operational Research*, 125(1), 114-133.

Aguarón, J.; Moreno-Jiménez, J.M. (2003): The Geometric Consistency Index: Approximated Thresholds, *European Journal of Operational Research*, 147, 137-145.

Aguarón, J.; Moreno-Jiménez, J.M. (2007): The Geometric Consistency Index: Approximated Thresholds, (*en evaluación*).

Barzilai, J.; Lootsma, F.A. (1997): Power relation and group aggregation in the multiplicative Ahp and SMART, *Journal of Multi-Criteria Decision Analysis*, 6, 155-165. (creo 6).

Brugha, C.M. (2000): Relative Measurement and the Power Function, *European Journal of Operational Research* 121, 627-640.

Escobar, M.T.; Moreno-Jiménez, J.M. (2000): Reciprocal Distributions in the Analytic Hierarchy Process, *European Journal of Operational Research* 123(1), 154-174.

Saaty, T.L. (1977): A scaling method for priorities in hierarchical structures, *J. Math. Psychol.*, 15, pp. 234-281.

Saaty T. L. (1980): *Multicriteria Decision Making: The Analytic Hierarchy Process*. McGraw-Hill. New York.

Saaty, T.L. (1994): *Fundamentals of Decision Making*. RSW Publications.

JUSTIFICACIONES DE LA REGLA DE BORDA: UNA REVISIÓN CRÍTICA

MIGUEL MARTÍNEZ PANERO
e-mail: panero@eco.uva.es

Departamento de Economía Aplicada
UNIVERSIDAD DE VALLADOLID

Área temática: Métodos cuantitativos.

Resumen

Mediante la regla de votación de Borda cada elector asigna puntuaciones escalonadas y descendentes a los candidatos de manera correlativa a su mayor o menor mérito. De esta forma, si hay n alternativas en litigio, la mejor valorada obtiene $n-1$ puntos, la segunda opción $n-2$, puntos, y así sucesivamente hasta llegar al peor candidato, que recibe 0 puntos. Las puntuaciones individualmente otorgadas se suman para determinar una evaluación colectiva que decide el ganador (o ganadores, en caso de empate).

Sobre el método de Borda, que ha generado una extensa literatura en el marco de la Teoría de la Elección Social, siempre se ha debatido la siguiente pregunta: ¿Por qué tales ponderaciones (números naturales) y no otras? En el presente trabajo se realiza una valoración crítica de las respuestas que ha suscitado la cuestión de la posible arbitrariedad del espectro de puntuaciones empleado: desde los alegatos “filosóficos” del propio Borda (1770) y Laplace (1795) utilizando supuestos estadísticos de equiprobabilidad, o la apología que del método realiza Morales (1797, 1805) como manera idónea de “calcular la opinión”, pasando por argumentos más recientes, con método axiomático, como los realizados por Goodman – Markowitz (1952) y culminados por Young (1974, 1975), o la “justificación parcial” proporcionada por Black (1976), hasta los tratamientos métricos de Cook – Seiford (1982) (a partir de de Kendall (1962)) y Farkas – Nitzan (1979), o las aproximaciones con metodología DEA, para llegar por fin la defensa de Saari en numerosos trabajos con técnicas de tipo geométrico que, hoy por hoy, le han convertido en el máximo valedor de la regla de Borda como procedimiento óptimo de decisión colectiva.

Palabras clave: método (regla) de Borda, regla de puntuación, ponderación (peso), discretización.

Abstract

Under the Borda voting rule each voter assigns equidistant descending scores to candidates by order of merit. In this way, if there are n alternatives, the best of them obtains $n-1$ points, the second best $n-2$ points, and so on, up to the worst candidate, which is given 0 points. These individual values are added up in order to determine a collective outcome, the highest scored alternative(s) being the winner(s).

Concerning the Borda count, which has provided a wide literature in Social Choice Theory, a question has arisen, even from Borda's times: Why such range of scores (positive integer numbers) instead of other possible choices? This paper critically surveys the answers to the subject of arbitrariness of weights appearing in the Borda count. Firstly, the “philosophical” arguments by Borda (1770) himself and Laplace (1795), taking into account statistical hypotheses of equiprobability, or the apology by Morales (1797, 1805) considering the Borda count as the appropriate way to “calculate the opinion”; in second place, more recent approaches with axiomatic method such as those proposed by Goodman – Markowitz (1952) and more successfully by Young (1974, 1975), or the “partial justification” provided by Black (1976); then, metric treatments by Cook – Seiford (1982) (coming from Kendall (1962)) and Farkas – Nitzan (1979), or within the DEA context; finally, Saari's defence with geometric techniques considering the Borda rule as the optimal method in collective decision making.

Key words: Borda count (rule), scoring rule, weight index, discretization.

1. Introducción

En 1770 el *chevalier* Jean-Charles de Borda, un reputado ingeniero militar y oficial de marina con aportaciones a la dinámica de fluidos y al perfeccionamiento de instrumentos náuticos y cartográficos¹, expuso ante la Academia de Ciencias de París una memoria ajena a sus intereses científicos primordiales, pero que, en buena medida, le ha reportado su fama póstuma. En calidad de miembro de la citada institución, Borda había constatado que el método de pluralidad empleado para elegir a los nuevos miembros entre los aspirantes a las distintas secciones (Mecánica, Astronomía, etc.) podía tener efectos perversos². Por ello, redactó una propuesta justificada de enmienda de tal procedimiento de selección, propugnando en su lugar el que hoy lleva su nombre, que sería implantado en la Academia desde 1796 hasta 1804, fecha en que dejó de ser utilizado por indicación expresa de Napoleón³.

Condorcet parafraseó como sigue el método de Borda, para criticarlo a su vez (citamos por la traducción⁴ de López de Peñalver [1799] (1992, p. 57)):

“Por aquí puede verse el objeto que se propone el Señor Borda, que es graduar el mérito por el número de votos que tiene cada sujeto, dando cierto valor a los votos respecto de los lugares en que se dan. Así, suponiendo que el voto en último lugar valga cierta cantidad, y que el voto en penúltimo lugar valga dicha cantidad mas otra que llamaremos P , supone que el lugar precedente valdrá dicha primera cantidad mas 2 veces P ; que el otro lugar inmediato valdrá dicha cantidad mas 3 veces P , y así sucesivamente.

¹ Especial atención merece su círculo repetidor, que tuvo un papel primordial en la implantación del sistema métrico decimal.

² Sobre los defectos en que incurre la regla de pluralidad, véanse la propia memoria de Borda [1770](1784) o Morales (1797). Un análisis crítico más reciente (pues es aún usada hoy en día en numerosos contextos decisionales) aparece en García Lapresta – Martínez Panero (2000).

³ Curiosamente, en 1997 Napoleón había sido elegido miembro de la sección de Ciencias Físicas y Matemáticas, subsección de Artes Mecánicas, mediante el método de Borda.

⁴ Entre 1798 y 1799, se editaron en castellano las *Cartas a una Princesa de Alemania sobre diversos Temas de Física y Filosofía*, obra de divulgación de Leonhard Euler, en traducción de López de Peñalver. Esta versión incorpora la apostilla, que Condorcet incluyó en su edición de la obra de Euler, en la que el autor francés reseña la Memoria que Borda presentó ante la Real Academia de Ciencias de París, donde éste defendía su método de votación con puntuaciones.

De esta manera para hallar la cantidad de opinión, como los valores antes señalados a los lugares solo son relativos y pueden suponerse iguales a la unidad, se multiplicará el número de votos en último lugar por 1, el de los penúltimos por 2, el de los que siguen por 3, y así sucesivamente”.

Lo prolijo de la cita (muy en el estilo de la época) queda compensado por su innegable interés histórico y científico. Efectivamente, en la citada exposición se consideran distintos juegos de ponderaciones o pesos para implementar la regla de Borda (si bien equivalentes, en el sentido de que todos determinan la misma decisión colectiva), cuyo espectro, desde su misma propuesta hasta nuestros días, ha sido cuestionado por su presunta arbitrariedad⁵.

En el presente trabajo se hace un repaso crítico de las justificaciones de tal rango de puntuaciones empleadas en la regla de Borda. Así, en la Sección 2, se enuncian los razonamientos del propio Borda o de Laplace, quienes preestablecen equidistribuciones de valores a asignar con criterios probabilísticos o estadísticos. Así mismo, se expone un argumento alternativo de Borda (secundado independientemente por Morales) que basa las puntuaciones otorgadas en una forma compacta de tener en cuenta simultáneamente todas las victorias por mayoría simple en las comparaciones por parejas de las alternativas. A continuación, en la Sección 3, se exponen distintas justificaciones axiomáticas del método, en las que diversas combinaciones de propiedades deseables desde supuestos ético-democráticos determinan que la única regla de votación⁶ que las satisface es justamente la de Borda. En la Sección 4 se trata su redescubrimiento por Kendall y desarrollos posteriores basados en tratamientos métricos y de cercanía al consenso, así como con metodología DEA. La panorámica presentada culmina en la Sección 5, donde se esboza el razonamiento de tipo geométrico con el que Saari se ha manifestado recurrente e incontrovertiblemente a favor de dicha regla. Por fin, se concluye el trabajo con algunas consideraciones y posibles extensiones del marco de preferencias utilizado.

⁵ Es inmediato verificar que, para una misma situación preferencial, el ganador puede depender del juego de ponderaciones utilizado. En Stein – Mizzi – Pfaffenberger (1994) se analiza mediante una condición de dominancia la posibilidad de que una alternativa sea ganadora cualquiera que sea el sistema de puntuaciones utilizado.

⁶ A lo largo del texto se emplearán recurrentemente los conceptos de *regla de votación*, *función de bienestar social* y *función de elección social*, para cuyo significado y diferencias según autores remitimos a Taylor (2005, pp. 7–8).

2. Borda, Laplace, Morales: apologías⁷ del método

Aunque se suelen datar los orígenes de la regla de Borda a finales del siglo XVIII, estos deberían retrotraerse varios siglos más si tenemos en cuenta que Nicolás de Cusa, en el libro III de su obra *De Concordantia Catolica*, publicada en 1434, ya discute en un contexto electoral un método que es justamente el de Borda, con puntuaciones individuales escalonadas, desde la unidad hasta el número total de candidatos, asignadas a los mismos de forma secreta⁸. Ésta es la razón por la que McLean – Urken (1995, pp. 24) han reivindicado su prioridad y sugerido que debería hablarse de “regla de Cusa”, en vez de “regla de Borda”.

Sin embargo Cusa se preocupó más de aspectos meramente procedimentales que del trasfondo del método, al contrario que Borda. Y es que, si bien la memoria de este último, un texto breve y programático, tiene fines eminentemente pragmáticos (implementar un método idóneo para ser usado en el seno de la Academia de Ciencias de París), su incisión y aguda visión del problema electoral le ha hecho merecer un destacado puesto en la historia de la Teoría de la Elección Social.

Borda [1770] (1784, p. 659), como ya se ha indicado, postula una cardinalización equidistante del ordenamiento de alternativas utilizando lo que Black (1958, p. 182) ha denominado “teoría de probabilidad a nivel de sentido común”:

“Digo que el grado de superioridad que este elector concede a *A* sobre *B* debe considerarse el mismo que el que otorga a *B* sobre *C*. En verdad, ya que el segundo candidato *B* es *igualmente susceptible* de todos los grados de mérito entre los de los otros candidatos *A* y *C*, no hay razón para sostener que el elector que ha decidido este orden de los 3 candidatos haya deseado colocarle más cerca de *A* que de *C*.”

Las cursivas son nuestras, e indican una de las críticas de Condorcet al método de Borda (véase López de Peñalver [1799] (1992), p. 59):

⁷ Empleamos esta palabra en su significado de defensa o (auto)justificación, que procede de su etimología griega y se ha conservado más en el ámbito anglosajón que en el castellano (así, por ejemplo, en la célebre *Apología de Sócrates* de Platón o en *A Mathematician's Apology* de Hardy).

⁸ El capítulo en el que Cusa expone su propuesta de la regla de Borda aparece traducido al inglés en McLean – Urken (1995, pp. 77–78).

“En la resolución de esta cuestión se supone que los grados de mérito o de preferencia entre los candidatos son iguales y reputados por iguales en la opinión de los censores. Esto admitiría bastante contradicción, y las más de las veces no puede suponerse esta igualdad, a lo menos en la opinión de los electores”.

También Grazia (1953) hace recaer una de las debilidades del razonamiento de Borda en la cita anterior, toda vez que el argumento negativo de ausencia de información, aunque puede tener sentido en el marco de la memoria,

“[...] no le faculta para justificar lo contrario [esto es, que los pesos asignados deben ser exactamente proporcionales a las posiciones]. Y esto es exactamente lo que Borda supone hasta el punto de distorsionar la escala electoral.”

Laplace dió una justificación estadística más sofisticada (aunque también controvertida, como veremos) al requerimiento de Borda⁹. Es bien conocido el carácter elíptico de los razonamientos de Laplace. Tanto es así que, en ocasiones, sus palabras: “Es fácil comprobar...” requieren arduos cálculos por parte del lector para comprobar la veracidad de lo afirmado. Esto es exactamente lo que ocurre en este caso, en el que Laplace [1819] (1987, pp. 114–115) sostiene:

“Supongamos que a cada elector se le da una urna conteniendo una infinidad de bolas, por medio de las cuales poder matizar todos los grados de mérito de los candidatos; supongamos además que saca de la urna un número proporcional a los méritos de cada candidato [...]. Pero las papeletas [en las que sólo se ordenan a los aspirantes] lo único que indican es que el primero tiene más [bolas] que el segundo, el segundo que el tercero, y así sucesivamente. Suponiendo que [...] todas las combinaciones [...] son igualmente admisibles, [...] si estos números son muy elevados, como es de suponer que ocurra para que puedan expresar toda la gama de méritos, el análisis más elemental permite ver que los números que han de ser escritos

⁹ Aunque la cita de Laplace que sigue está tomada de su *Ensayo Filosófico sobre las Probabilidades*, editado en 1819, según Black (1958, pp. 181–182) los argumentos expuestos ya aparecían en 1795 en sus lecciones impartidas en la Escuela Normal. Por otra parte, la citada obra está compuesta sustancialmente a partir del la *Teoría Analítica de las Probabilidades*, publicada en 1812. De hecho, la parte del *Ensayo* sobre las votaciones en asambleas y comités reproduce las páginas prologales cx–xciv de esta última obra.

en cada papeleta al lado del último nombre, del penúltimo, etc., pueden ser representados por la progresión aritmética 0, 1, 2, etc¹⁰.”

Todhunter (1865, pp. 546–548) reconstruyó el “elemental” razonamiento sugerido por Laplace para obtener la discretización indicada como sigue:

“Sean t_1, t_2, \dots, t_n los méritos de los candidatos, empezando por el más meritorio [...]; el elector puede adscribir cualesquiera méritos a los aspirantes, ordenados de tal forma que ninguno sea mayor que el inmediatamente precedente, y a lo sumo con un valor a .

El valor medio del mérito del r -ésimo candidato será:

$$\frac{\int \int \dots \int t_r dt_1 dt_2 \dots dt_n}{\int \int \dots \int dt_1 dt_2 \dots dt_n}.$$

El registro de condiciones de integración efectuado por Laplace no es inteligible, y *establece* el resultado de integración sin explicación.”

Hasta aquí, todo está esencialmente en Laplace, quien en el Libro 2, pp. 277–278, de su *Teoría Analítica de las Probabilidades*, llega sin justificación al valor

$$\frac{(n-r+1)a}{n+1},$$

el mismo que obtiene Todhunter apoyándose en los teoremas de cambio de variable y de Lejeune-Dirichlet. De aquí se sigue que cada elector debería asignar el número n al candidato considerado el mejor, el número $n-1$ al siguiente, y así sucesivamente, resultando ganador el aspirante que obtenga mayor suma de puntuaciones. Es decir, se reproduce exactamente el método propugnado por Borda.

Cabe señalar que Laplace es consciente del sesgo que tomaría la elección idónea cuando los agentes rebajasen estratégicamente el mérito de unos candidatos en favor de otros¹¹.

¹⁰ En realidad Laplace escribe “1, 2, 3, etc.”. Sin embargo el matiz de traducción, no contemplado en el original, no incurre en error, pues ambas graduaciones, al diferir en una constante, proporcionan esencialmente el mismo método. De hecho, algo parecido ocurrió con Morales (1797), quien empleó la escala 1,2,3,... para cambiarla en Morales (1805) por la secuencia 0,1,2,... por las razones que se concretarán más adelante.

¹¹ También Condorcet (1785), en su *Discours Preliminaire*, sin citar a Borda (hace alusión veladamente a un “geómetra célebre”) pero refiriéndose sin género de duda a su método, reputa a éste de cabalístico. Tal vez se entienda mejor este comentario si se tiene en cuenta que en francés *cabale* tiene una acepción de “intriga o maniobra oculta”, de la que carece el término “cábala” en castellano. Cuando este punto débil le

Y también señala otra crítica recurrente en la literatura que le ha sido hecha al método de Borda: su susceptibilidad de seleccionar candidatos mediocres¹².

Daunou (1803) fue quizá el primero en cuestionar el razonamiento de Laplace, basado en intensidades de preferencia, sobre método de Borda, pues además de los aspectos estratégicos señalados, según su lectura se obtendría una escala de puntuaciones en progresión geométrica y no aritmética. Esta es interpretación la de Daunou¹³:

“¿Qué representan los números de Borda? Simplemente, según el ciudadano Laplace, que el primer candidato obtiene más bolas que el segundo, el segundo que el tercero, y así sucesivamente. [...] Laplace señala que cualquiera que sea el número de bolas que obtenga el primer candidato en una papeleta dada, todas las combinaciones de números menores que satisfagan las condiciones precedentes [i.e., los votantes actúan de buena fe y asignan las bolas de forma precisa] son igualmente admisibles. Y él añade que ‘el número de bolas obtenido por cada candidato se puede encontrar realizando el total de todos los números asignados a él y dividiendo ese total por el número total de combinaciones’. O sea, que a cada candidato se le asignará el término medio de los valores asignados respecto del candidato inmediatamente superior. De manera que si el primer candidato obtiene [$n=$] 20, el segundo tendrá 10. En verdad, ya que a este segundo candidato se le asignaría alguno de los términos de la serie del 1 al 19, que suman 190, y al dividirse este valor por el número de combinaciones, que son 19, obtenemos 10 como cociente. [...] De manera similar serían $n/4$ para el tercer candidato, $n/8$ para el cuarto, y así sucesivamente”.

Para Black (1958, pp. 182–183) la justificación estadística de Borda – Laplace es “insostenible” porque, al interpretar las puntuaciones otorgadas con un sentido utilitarista, se llega a que “un votante considere los méritos relativos de dos candidatos relacionados en alguna razón de tipo 5:4 ó 6:1, o alguna otra. Pero podemos sentirnos

fue señalado a Borda, éste contestó lacónicamente que su procedimiento sólo era válido para ser utilizado por hombres honorables (Black (1958, p.182).

¹² Aunque el defecto puede devenir virtud, denominada en este caso “respeto por la media”. Véase a este respecto Black (1958, p. 56), quien afirma que “el candidato que debe ser elegido es el mejor situado, por término medio, en las relaciones de preferencias de los votantes”.

bastante seguros de que la mente humana no opera de esta forma”. Además, señala Black, se incurre en el problema de la “comparación interpersonal de utilidades¹⁴”. Por su parte, Saari (1995, p. 20) tacha tal explicación de “filosófica” e “insatisfactoria”. De hecho, son varios los autores que coinciden en que el enfoque probabilístico dado a la Teoría de la Elección Social por Condorcet y Laplace en buena parte ha lastrado el desarrollo de la disciplina en Francia hasta su renacimiento, ya avanzado el siglo XIX, en el ámbito anglosajón.

No obstante, un tratamiento estadístico *a la Laplace*, que llena las lagunas explicativas de éste con herramientas formales modernas, ha sido retomado en Tanguiane (1991, pp. 80 y ss.) y Tangian (2000), quien obtiene las puntuaciones Borda normalizadas como los valores medios esperados (o esperanza matemática) de un vector n -dimensional aleatorio equidistribuido en el intervalo unidad¹⁵. Así mismo, Tangian (2000) sostiene que las puntuaciones enteras del método de Borda representan las utilidades cardinales de los agentes (sus “estimaciones latentes”) de manera suficientemente adecuada, de modo que para electorados grandes el modelo de Laplace (preciso, pero inmanejable) se aproxima al de Borda (grosero, pero manejable)¹⁶. Por otro lado, la regla de Borda fue “redescubierta” en un contexto estadístico por Kendall (1962), quien trató la agregación de rankings como un problema de estimación y propuso una solución de consenso que maximizaba el acuerdo entre los votantes¹⁷. Además, una conexión entre la regla de Borda y una “regla de la media” obtenida desde supuestos probabilísticos puede encontrarse en Intriligator (1973). Y recientemente Heckelman (2003) ha diseñado una “regla de Borda probabilística” que asigna los pesos mediante un mecanismo de lotería.

Cabe indicar que tal vez Borda [1770] (1784) previese críticas posteriores a la forma de equidistanciar las puntuaciones de su método debido a lo poco concluyente de su

¹³ Véase McLean – Urken (1993, pp. 262 y ss), de donde traducimos.

¹⁴ Sobre este controvertido problema, que cuestiona la validez misma de la Teoría de la Elección Social, véase Martínez Panero (2004, pp. 142–150) y las referencias allí indicadas.

¹⁵ Tangian (1991) se basa en el análisis orden-estadístico de Kendall – Stuart (1969, pp. 268 y ss.) y observa, muy atinadamente, que las puntuaciones Borda normalizadas obtenidas son justamente las coordenadas del vector de centros de gravedad de las utilidades cardinales de los agentes, algo que está implícito en los argumentos de Laplace.

¹⁶ Para Tangian (2000) “la anterior suposición [i.e., la presunción de Laplace de que las estimaciones de utilidad ordenadas de los agentes están equidistribuidas, una hipótesis tradicional en probabilidad y estadística en ausencia de información] implica en particular que la distribución de utilidades cardinales es independiente de las preferencias ordinales”.

¹⁷ Como veremos, el enfoque de Kendall está en la base de algunos de los tratamientos métricos de la regla de Borda que se expondrán en la siguiente sección.

razonamiento anterior, pues ya en su misma memoria proporcionaba una segunda vía justificativa para introducir los pesos utilizados. A continuación exponemos los argumentos con los que Borda defendió esta segunda forma de entender su regla siguiendo la exposición del matemático ilustrado José Isidoro Morales quien, al parecer, los redescubrió independientemente.

Es interesante la forma en que Morales tuvo conocimiento del método de Borda, que suscitó por parte del autor español una memoria (Morales (1797)) en la que también éste defendía el nuevo procedimiento frente a otros más al uso como, por ejemplo, el de pluralidad (véase Martínez Panero – García Lapresta (2003)). A través del periódico francés *La Décade Philosophique* Morales había sabido por una breve reseña que los miembros del Instituto Nacional de Francia (nombre de la Academia de Ciencias tras el periodo revolucionario) habían escogido cinco plazas vacantes utilizando un método de elección que él llama “de compensación y suma” (no menciona, por no conocerlo entonces, el nombre de Borda). Tal procedimiento le parece idóneo, justo y exacto para reflejar la opinión que los votantes, y en su memoria, también con fines prácticos de implementación efectiva, Morales lo describe así: cada elector ordena de mejor a peor los c candidatos. De cada ordenación se asignan c puntos al primer candidato, $c-1$ puntos al segundo candidato,..., 2 puntos al penúltimo, y 1 punto al último candidato. Se suman los puntos recibidos por cada candidato y se escoge al que ha recibido más puntos.

El autor español divulgó su obra en distintos ámbitos nacionales y extranjeros, de manera que llegó a ser conocida y valorada por el mismo Borda. No obstante, ocho años después, nuestro autor escribe un apéndice (Morales (1805)) a la obra anterior con el fin de contestar a reparos y objeciones recibidas. Morales detalla que se ha criticado su método de compensación y suma por la rigidez en la asignación de c , $c-1$,...,2, 1 puntos, ya que ésta no permite reflejar con exactitud la intensidad en la opinión, es decir, constriñe la libertad del votante y no refleja el mérito del votado. Argumenta entonces, para fundamentar estas puntuaciones, que en las elecciones binarias (con sólo dos candidatos) la mayoría simple es el método adecuado para elegir y en el hecho de que, cuando hay más de dos candidatos, se podrían hacer elecciones binarias entre cada uno de los posibles pares (en total, $c(c-1)/2$ elecciones binarias) y luego sumar los votos que cada candidato obtuvo. Ahora bien, tal información numérica queda de hecho capturada

en las puntuaciones obtenidas por el método de Borda: basta con que se emplee una escala $c-1, c-2, \dots, 2, 1, 0$, equivalente a la utilizada en Morales (1797). En resumen, Morales (1805, p. 23) deja así el estado de la cuestión:

“[...] Si se hace una elección por el orden de mérito, con la progresión de los números naturales, *empezada desde cero*, y de tantos términos cuantos sean los candidatos, hay una absoluta identidad entre este método de votar ó de elegir [la regla de Borda] y el de la elección binaria repetida tantas veces como combinaciones admite el número de candidatos [...] También será cierto que este método gozará de la misma exactitud y rigor que la elección binaria, que se toma por norma de toda de toda elección justa. Porque el tal método no es otra cosa que una elección binaria repetida, y el mérito de su invención es haber ahorrado esa repetición”.

Massó (2003) estima que ésta es

“una maravillosa justificación (ordinal) de su método basado en las puntuaciones [...]. La argumentación de Morales parece un antecedente a la justificación del axioma moderno de consistencia usado en las caracterizaciones axiomáticas de la mayoría de los conceptos de solución propuestos por la teoría de los juegos cooperativos (núcleo, valor de Shapley, nucleolo, etc.)”.

Y como también veremos, una noción de consistencia, si bien en el contexto de la Teoría de la Elección Social, ha resultado ser crucial en la más conocida de las caracterizaciones del método de Borda que se detallan a continuación.

3. Goodman – Markowitz, Gärdenfors, Young, Black: justificaciones axiomáticas

Arrow [1963] (1974, pp. 202–203), en un apartado histórico a modo de apéndice, se refiere de esta forma a los problemas que plantea la regla de Borda, ya señalados por Black (1958):

“[...] este método confiere ponderaciones iguales a las diferencias de rango de los candidatos y, así mismo, da la misma ponderación a los distintos

votantes. Lo primero plantea el problema de la mensurabilidad de la utilidad; lo segundo, el problema de las comparaciones interpersonales. Borda justifica el primer paso con un argumento esencialmente basado en la ignorancia, [y el segundo] sobre la base de la igualdad de los votantes. Estos temas han continuado repitiéndose. El razonamiento de Goodman – Markowitz (1952) puede considerarse, en efecto, como la [primera] justificación axiomática de la posición de Borda”.

El objetivo de Goodman – Markowitz (1952) fue modificar las condiciones del célebre teorema de imposibilidad de Arrow para enunciar resultados positivos. Así, demuestran que la regla de Borda es la única función de bienestar social que, además de los usuales y comúnmente admitidos principios de optimalidad fuerte de Pareto y simetría (anonimato), cumple una tercera condición que requiere que “la implicación [en el agregado] que supone el cambio de un nivel [en la consideración de un candidato] al siguiente sea la misma cualquiera que sea el nivel de partida”.

El carácter *ad hoc* de esta última condición (que de hecho se puede entender como una relectura del requerimiento de equidistancia de Borda) en cierta forma ha relegado este resultado frente a otras axiomatizaciones del método recurrentemente citadas en la literatura. Por ejemplo, la proporcionada por Gärdenfors (1973), quien establece que la regla de Borda es la única función de votación representable¹⁸ que satisface neutralidad (tratamiento simétrico de las alternativas), optimalidad débil de Pareto y una tercera condición denominada independencia posicional fuerte. Esta última propiedad se puede enunciar así: siempre que una alternativa no ocupe alguno de los niveles intermedios entre dos alternativas x e y dadas, y tampoco empate con ellas, la relación social entre x e y no se verá afectada por ella.

¹⁸ Una función de votación representable es básicamente una regla de puntuación (*scoring rule*), que se define mediante un vector de pesos con tantas componentes como alternativas de manera que cada agente asigna la puntuación s_1 a la alternativa que ocupe el primer puesto, s_2 a la segunda, y así sucesivamente ($s_1 \geq s_2 \geq \dots \geq s_n \geq 0$, $s_1 > s_n$). De este modo, la regla de Borda no es sino un caso particular de regla de puntuación con pesos $s_i = n - i$. A su vez, la regla de Borda así definida es equivalente a cualquier regla de puntuación con pesos en progresión aritmética de diferencia común positiva, esto es, tales que $s_i - s_{i+1} = d > 0$, $i = 1, \dots, n-1$. Nótese, así mismo, que la regla de pluralidad corresponde en este contexto a una regla con puntuaciones $s_1 = 1$, $s_2 = \dots = s_n = 0$.

El trabajo de Gärdenfors citado entiende la regla de Borda como un procedimiento esencialmente ordinal¹⁹, y la defiende de la “acusación injusta de ser un método (estúpido) de amalgamar intensidades de preferencias [de los agentes mediante números impuestos]”. Su importancia principal está en delimitar dos visiones irreconciliables del voto²⁰: la de Borda (posicionalista) y la de Condorcet (no posicionalista). Sin embargo, tampoco su caracterización de la regla de Borda es la más reconocida. Tal mérito le corresponde a Young (1974), quien probó que el método citado corresponde a la única función de elección social que satisface neutralidad, fidelidad, cancelación y consistencia. Young define la propiedad de fidelidad como el hecho de que “socialmente más preferido” e “individualmente más preferido” signifiquen lo mismo cuando la sociedad es unipersonal. Así mismo, la propiedad de cancelación establece que, si para cualquier par de alternativas, la preferencia de una frente a la otra por parte de un agente se puede compensar con la preferencia inversa por parte de otro agente, entonces todas las alternativas son ganadoras. Por fin, el principio de consistencia o refuerzo (*reinforcement*) se puede describir como el requerimiento de que si dos coaliciones por separado dan lugar a la misma decisión social, entonces cuando se unen en un solo bloque tal decisión se preserve.

En un trabajo posterior, Young (1975) caracterizó de nuevo la regla de Borda como la única de las reglas de puntuación (que, a su vez, vienen dadas por funciones de elección social que cumplen anonimato, neutralidad, consistencia y continuidad²¹) que además cumple la mencionada propiedad de cancelación. Desde entonces se han proporcionado numerosas axiomatizaciones (en ocasiones más accesibles): Fine – Fine (1974), Hansson – Sahlquist (1976), Fishburn – Geherlein (1976), Nitzan – Rubinstein (1981), Saari (1990), Debord (1992) y Van Newenhizen (1992), entre otros (véase Chebotarev – Shamis (1998)).

Aunque no se trata de una caracterización, queremos incorporar también en esta sección la “justificación parcial” de la regla de Borda proporcionada por Black (1976).

¹⁹ Ahora bien, un enfoque utilitarista-cardinal de las puntuaciones, que ya aparecía tácitamente en Borda [1770](1784), ha sido retomado por Black (1958 y 1976), Sugden (1981, pp. 140–145) y Marchant (2000), en este último caso con un tratamiento difuso (*fuzzy*) de las preferencias de los agentes.

²⁰ Sobre la polémica Borda – Condorcet y una vía de aproximación mediante sistemas de votación híbridos, véase Martínez Panero (en prensa).

Entendemos este calificativo en el sentido de que Black, al coincidir con las posiciones electorales de Condorcet y ser estas divergentes de las de Borda, nunca podría justificar totalmente la regla de propuesta por éste último, que puede proporcionar un ganador distinto del de Condorcet²² (aquel que vence una a una a todas las demás alternativas por mayoría simple).

El razonamiento de Black (1976) retoma un enfoque utilitarista ya apuntado en Black (1958) y se basa, además de en la definición clásica de la cuenta de Borda, en un segundo conteo individual que asigna a cada alternativa el número de alternativas a las que derrota menos el número de aquellas por las que es derrotada. Black (1976) demuestra entonces que existe una transformación afín entre ambos contadores (por lo que son equivalentes en el sentido de que ambos proporcionan la misma ordenación social) y que la suma de las puntuaciones emitidas por cada agente mediante el segundo contador es nula²³. Con tales prerequisites, Black se decanta por el uso del segundo contador (por motivos más bien estéticos, de elegancia en las demostraciones aportadas, si se tiene en cuenta a Coughlin (1980)) y justifica entonces el uso de este segundo contador interpretándolo en términos mayoritarios. Además, argumenta que permite compensación²⁴, tiene en cuenta toda la información contenida en las preferencias mostradas por los agentes, proporciona un orden social transitivo²⁵, y selecciona como ganadora la alternativa que, en media²⁶, ocupa mejores posiciones en las preferencias de los agentes (algo que, como ya hemos comentado, ya aparecía explícitamente en Black (1958)).

Señala también Black, como ya habían hecho en el pasado Condorcet, Laplace y Daunou, que el voto estratégico puede arruinar la implementación efectiva y utilidad de

²¹ Esta es una propiedad de tipo arquimediano que indica que si se replican las preferencias de una coalición de votantes un número suficiente de veces, entonces esta coalición (junto con sus clones) deviene ganadora.

²² Por tal razón se dice que la regla de Borda no es *Condorcet consistente*.

²³ Véase también Coughlin (1980). En Martínez Panero (2003) se demuestran ambas propiedades de forma inmediata (utilizando el axioma de reciprocidad) para reglas de Borda difusas que extienden los casos clásicos formulados en Black (1976).

²⁴ Como ya señalamos, método “de compensación y suma” fue el nombre dado por Morales (1797) a la regla de Borda.

²⁵ Estas dos últimas características diferencian radicalmente la bondad del método de Borda con las “patologías” que pueden presentar los métodos de pluralidad, que sólo toma en cuenta las mejores alternativas, y de mayoría simple, que puede producir ciclos en el agregado (paradoja de Condorcet).

²⁶ Sobre los argumentos de tipo media en conexión con la regla de Borda, véanse también Straffin Jr (1980) y Mueller (1979).

la regla de Borda. Ahora bien, a partir del teorema de Gibbard – Satterthwaite, se puede afirmar que tal susceptibilidad de manipulación no es exclusiva suya, sino común a todos los procedimientos de votación (véase Taylor (2005, pp. 60–69). Por este motivo, aunque la manipulabilidad, como también la inconsistencia de Condorcet del método de Borda (las dos principales críticas de Black) son incuestionables, éstas se deben analizar comparativamente con otros procedimientos, para lo que nos remitimos al análisis de Saari (Sección 4) que, como veremos, coloca a la regla de Borda en una posición de privilegio.

4. Cook *et al.*, Farkas – Nitzan: métricas sobre preferencias y metodología DEA

Cook – Seiford (1982) probaron, a partir de Kendall (1962), que la regla de Borda minimiza para la métrica 2 el desacuerdo total agregado (sobre las métricas p , $1 \leq p \leq \infty$, véanse, Cook – Kress – Seiford (1996) y González Pachón – Romero (1999)). En la misma línea, en la que late la idea de consenso (*distance-consensus approach*), Farkas – Nitzan (1979) también demostraron, a su vez, que el método de Borda coincide con la regla de cercanía a la unanimidad (*closeness to unanimity*) introducida por los citados autores en términos paretianos sugeridos por Sen (1977). Ulteriores desarrollos de esta línea de investigación deben recoger las aportaciones de González Pachón – Romero (1999) utilizando programación por metas, y de Marchant (2001) siguiendo los pasos de Farkas – Nitzan (1979) en un contexto que incluye torneos, órdenes débiles, semiórdenes, relaciones de preferencia difusa, etc.

Cabe así mismo señalar que la metodología DEA (Análisis Envoltente de Datos) ha proporcionado una atractiva manera de afrontar el problema de la asignación de pesos. A este respecto Cook – Kress (1990) han probado que un caso particular de lo que denominan *función de discriminación de intensidad* reproduce exactamente el modelos de Borda y Kendall. Por otro lado, también basándose en la metodología citada, Contreras – Hinojosa – Mármol (2005) han propuesto un método alternativo a las técnicas de consenso apuntadas anteriormente, evaluando las alternativas mediante pesos flexibles de manera que se obtenga una ponderación óptima para cada una de ellas. Y de nuevo, en su análisis, la regla de Borda aparece involucrada en la solución

que se obtiene al resolver el problema de programación lineal que supone tal optimización.

5. Saari: geometría del voto y “optimalidad” de la regla de Borda

Saari (1995, p.20) se arroga el mérito de haber dicho la última palabra sobre el tema que nos ocupa al afirmar rotundamente que la cuestión de la particular elección de pesos de la regla de Borda, que había permanecido abierta durante los dos últimos siglos, ha sido por fin contestada por él mediante su “geometría del voto”.

Desde tiempos de Borda [1770](1784) es bien conocido el hecho de que toda la problemática electoral surge cuando concurren 3 o más alternativas²⁷. Por esta razón, de modo paradigmático (para mantener el aparato matemático a un nivel razonable), Saari (1994, 1995) ha estudiado exhaustivamente el voto ponderado en el caso de 3 alternativas²⁸, situándolas como vértices de un triángulo de manera que las 6 zonas en las que éste queda descompuesto después de trazar sus medianas corresponden a una de las 6 posibles ordenaciones de preferencias según la lejanía de un punto interior genérico de las mismas a los vértices. (véase también Nurmi (1999)).

Mediante un análisis en el que se incide en la simetría intrínseca de la regla de Borda, Saari (1994, p.14) demuestra que tal método es el que minimiza la manipulabilidad de entre todos los sistemas de voto posicional, así como las posibles paradojas en las que pudieran incurrir tales procedimientos²⁹. Por otro lado, Van Newenhizen (1992) y Saari (2000) han probado que este método es la regla de puntuación menos susceptible de vulnerar el criterio de Condorcet, o lo que es lo mismo, con mayor *eficiencia de*

²⁷ De hecho, problemas de agenda (orden de presentación de alternativas para su voto por parejas) con más de tres opciones ya fueron tenidos en cuenta estratégicamente por Plinio el Joven en el contexto del senado romano (véanse Farquharson (1969), Riker (1986, pp. 78–88) y McLean – Urken (1995, pp. 14–16 y 67–70)).

²⁸ En el caso de 3 alternativas y una vez realizado un proceso de normalización, un procedimiento de voto ponderado (o si se quiere, una regla de puntuación) viene dado por un vector $w_s = (1, s, 0)$ donde 1 es el peso asignado a la mejor alternativa, $s \in [0, 1]$ es el que obtiene la segunda alternativa y 0 la tercera. Los valores extremos alcanzables por s , 0 y 1, corresponden a los procedimientos de pluralidad (cada agente selecciona la mejor alternativa) y antipluralidad (cada agente indica su peor alternativa). El caso intermedio $s = 1/2$ corresponde justamente a la regla de Borda.

²⁹ “[The Borda count] is the unique method to minimize the number and kind of [voting] paradoxes, to minimize the likelihood of a paradox, [and] to minimize the likelihood that a small group can successfully manipulate the outcome”.

*Condorcet*³⁰. Además, Saari – Barney (2003), incidiendo en un tema introducido en Saari (1995, p.153), también han probado que la simetría de la regla de Borda hace a este método el único entre los posicionales inmune a la paradoja de la reversión (*reversal bias*): para cualquier otra regla de puntuación existen situaciones no triviales³¹ en las que, a pesar de invertirse las preferencias individuales, no hay cambio en la elección social resultante respecto de las situaciones originales (sobre la excepcionalidad de la regla de Borda entre los métodos posicionales, véase también Nurmi (2004 y 2005)).

Ahora bien, la militancia de Saari en pro de la regla de Borda, que considera óptima (Saari (1995, p. 19), no se restringe únicamente a su *status* entre los sistemas de voto posicional. Saari ha sostenido (y sostiene) polémicas intelectuales (que nos hacen recordar el enfrentamiento Borda *versus* Condorcet) con Brams y Fishburn (introdutores y defensores del método de voto aprobatorio) y últimamente con Risse (que propugna métodos Condorcet consistentes, tales como la regla de Kemeny, en perjuicio de la regla de Borda). Sin embargo, a pesar de la profusión de trabajos en distintos frentes que avalan la implementabilidad de la regla de Borda por parte de Saari, está aún lejos de verificarse el desideratum de Young (1997) sobre su puesta en práctica³².

6. Conclusiones

En el presente trabajo se ha presentado una panorámica de las distintas justificaciones a la presunta arbitrariedad o imposición de las ponderaciones que aparecen en la regla de Borda. Como hemos comentado, ya Arrow señaló que éste es un tema que viene repitiéndose de tiempo en tiempo y, efectivamente, las distintas argumentaciones probabilístico-estadísticas, axiomáticas, métricas y basadas en la metodología DEA, así como las geométricas que hemos expuesto parecen avalar tal aserto. Ahora bien, a partir de la variada gama de técnicas y métodos empleados que hemos reseñado podemos

³⁰ La eficiencia de Condorcet se define como la probabilidad condicionada de que un procedimiento seleccione el ganador de Condorcet, supuesto que tal ganador exista.

³¹ Una situación trivial sería, por ejemplo, aquella en la que los órdenes lineales individuales fuesen todas las permutaciones circulares de uno dado, lo que provocaría un empate colectivo (o, desde un punto de vista no posicional, la aparición de un ciclo).

³² Young (1997, p. 200): “I predict that the time will come when [...] Borda’s rule will be considered a standard tool for legislative and committee decision making”.

hacer la siguiente valoración: si la importancia de un problema científico depende en buena medida de las disciplinas que se generan o desarrollan para resolverlo y que son, a su vez, susceptibles de ser empleadas en otros contextos, el problema tratado en el presente trabajo, lejos de ser una cuestión aislada, ha resultado sumamente fértil, ya que, si bien su tratamiento no ha generado *per se* las técnicas citadas, el análisis implicado en su resolución ha coadyuvado incuestionablemente en el desarrollo de las mismas.

Para concluir, queremos indicar una posible vía que, en cierta forma, elude el problema planteado, ya que no lo encara como tal, sino en un contexto más amplio. Tal posibilidad consiste en diseñar reglas de Borda más generales, en las que las que, aún manteniéndose su filosofía de agregación, se permita a los agentes manifestar sus preferencias de forma matizada (gradual o lingüística), y no meramente mediante las preferencias taxativas que son la base informacional de la regla de Borda “clásica”. De esta manera se generarían puntuaciones Borda que serían, según la forma de capturar lo más fielmente posible tales matices de información, números reales, intervalos, números difusos trapeciales o triangulares, dependiendo de la representación utilizada. Este es el tratamiento seguido en Marchant (1996a,1996b) García Lapresta – Martínez Panero (2000, 2002), García Lapresta – Llamazares – Martínez Panero (2005 y 2006) y García Lapresta – Martínez Panero – Meneses (en prensa).

Agradecimientos

El autor agradece a Jesús Basulto, de la Universidad de Sevilla, que le facilitara facsímiles de las obras de Condorcet, Laplace y Todhunter utilizadas en la primera sección del trabajo. Así mismo, reconoce el apoyo financiero proporcionado por la Junta de Castilla y León (Consejería de Educación y Cultura, Proyecto VA040A05), el Ministerio de Educación y Ciencia (Plan Nacional de Investigación Científica, Desarrollo e Innovación Tecnológica, Proyecto SEJ2006-024267/ECON) y FEDER.

Bibliografía

Arrow, K.J. (1974): *Elección Social y Valores Individuales*. Instituto de Estudios Fiscales, Madrid. [Traducción de Eusebio Aparicio Auñón de la segunda edición, corregida, en inglés, 1963: *Social Choice and Individual Values*. Introducción de Andreu Mas Colell. Primera edición, 1951].

Black, D. (1958): *The Theory of Committees and Elections*. Kluwer Academic Publishers, Boston.

Black, D. (1976): "Partial justification of the Borda count". *Public Choice* 28, pp. 1–16.

Borda, J.C. de (1784): *Mémoire sur les élections au scrutin*. Histoire de l'Académie Royale des Sciences, París. [Se reproduce, en inglés, en Grazia (1953) y en McLean – Urken (1995, pp. 81–89)].

Condorcet, J.A.M.N.C., marqués de (1785): *Essai sur l'Application de l'Analyse à la Probabilité des Décisions Rendues à la Pluralité des Voix*. L'Imprimerie Royale, París. [Se reproducen fragmentos escogidos, en inglés, en McLean – Urken (1995, pp. 91–112)].

Contreras, I. – Hinojosa, M.A. – Mármol, A.M. (2005): "A class of flexible weight indices for ranking alternatives". *IMA Journal of Management Mathematics* 16, pp. 71–85.

Cook, W.D. – Kress, M. (1990): "A data envelopment model for aggregating preference rankings". *Management Science* 36 (11), pp. 1302–1310.

Cook, W.D. – Kress, M. – Seiford, L.M. (1996): "A general framework for distanced-based consensus in ordinal ranking models". *European Journal of Operational Research* 96, pp. 392–397.

Cook, W.D. – Seiford, L.M. (1982): "On the Borda – Kendall consensus method for priority ranking problems". *Management Science* 28, pp. 621–637.

Coughlin (1979): "A direct characterization of Black's first Borda count". *Economics Letters* 4, pp. 131–133.

Chebotarev, P.Y. – Shamis, E. (1998): "Characterizations of scoring methods for preference aggregation". *Annals of Operations Research* 80, pp. 299–332.

Daunou, P.C.F. (1803): *Mémoire sur les Élections au Scrutin*. Baudouin, imprimeur de l'Institut National. París. [Se reproduce, en inglés, en McLean – Urken (1995, pp. 237–276)].

Debord, B. (1992): "An axiomatic characterization of Borda's k -choice function". *Social Choice and Welfare* 9, pp. 337–343.

Farkas, D. – Nitzan, S. (1979): "The Borda rule and Pareto stability: a comment". *Econometrica* 49, pp. 1305–1306.

Farquharson, R. (1969): *Theory of Voting*. Blackwell, Oxford.

Fine, B. – Fine, K. (1974): "Social choice and individual ranking I". *Review of Economic Studies* 41, pp. 459–475.

- Fishburn, P.C. – Gehrlein, W.V. (1976): “Borda’s rule, positional voting and Condorcet’s simple majority principle”. *Public Choice* 28, pp. 79–88.
- García Lapresta, J.L. – Llamazares, B. – Martínez Panero, M. (2005): “Reglas de Borda lingüísticas: análisis de sus propiedades”. En E. Herrera Viedma (ed.): *Procesos de Toma de Decisiones. Modelado y Agregación de Preferencias*. Universidad de Granada, Granada, pp. 69–78.
- García Lapresta, J.L. – Llamazares, B. – Martínez Panero, M. (2006): “Linguistic matrix aggregator operators: Extensions of the Borda rule”. En B. Reusch (ed.): *Computational Intelligence, Theory and Applications*. Springer Verlag, Berlín, pp. 561–576.
- García Lapresta, J.L. – Martínez Panero, M. (2000): “Análisis de algunos sistemas de votación a partir de la obra del ilustrado José Isidoro Morales”. *Hacienda Pública Española* 154, pp. 93–103.
- García Lapresta, J.L. – Martínez Panero, M. (2002): “Borda count versus approval voting: A fuzzy approach”. *Public Choice* 112, pp. 167–184.
- García Lapresta, J.L. – Martínez Panero, M. – Meneses, L.C. (en prensa): “Defining the Borda count in a linguistic decision making context. *Information Sciences*.”
- Gärdenfors, P. (1973): “Positionalist voting functions”. *Theory and Decision* 4, pp. 1–24.
- Goodman, L.A. – Markowitz, H. (1952): “Social welfare functions based on individual rankings”. *American Journal of Sociology* 58(3), pp. 257–262.
- González Pachón, J. – Romero, C. (1999): “Distance-based consensus methods: A goal programming approach”. *Omega* 27, pp. 341–347.
- Grazia, A. de (1953): “Mathematical derivation of an election system”. *Isis* 44, pp. 42–51.
- Hansson, B. – Sahlquist, H. (1976): “A proof technique for social choice with variable electorate”. *Journal of Economic Theory* 13, pp. 193–300.
- Heckelman, J.C. (2003): “Probabilistic Borda rule voting”. *Social Choice and Welfare* 21, pp. 455–468.
- Intriligator, M. (1979): “A probabilistic model of Social Choice”. *Review of Economic Studies* 40, pp. 553–560.
- Kendall, M. (1962): *Rank Correlation Methods*. Hafner, Nueva York.
- Kendall, M. – Stuart, A. (1977): *The Advanced Theory of Statistics. 1, Distribution Theory*. Charles Griffin, Londres.
- Laplace, P.S. de [1819] (1987): *Ensayo Filosófico sobre las Probabilidades*. Alianza Editorial, Madrid. [Traducción del francés de María Martínez Sierra].
- López de Peñalver, J. (1992): *Escritos*. Instituto de Cooperación Iberoamericana – Quinto Centenario – Antoni Bosch, editor – Instituto de Estudios Fiscales, Madrid. [Edición de Ernest Lluch de la obra original, 1799].

- Marchant, T (1996a): *Agrégation de relations valuées par la méthode de Borda en vue d'un rangement. Considérations axiomatiques*. Ph. D. Thesis, Université Libre de Bruxelles, Bruselas.
- Marchant, T. (1996b): "Valued relations aggregation with the Borda method". *Journal of Multi-Criteria Decision Analysis* 5, pp. 127–132.
- Marchant, T. (2000): "Does the Borda rule provide more than a ranking?" *Social Choice and Welfare* 17, pp. 381–391.
- Marchant, T. (2001): "The Borda rule and Pareto stability: A further comment". *Fuzzy Sets and Systems* 120, pp. 423–428.
- Martínez Panero, M. (2004): *Generalizaciones y Extensiones de la Regla de Votación de Borda*. Tesis doctoral, Universidad de Valladolid, Valladolid.
- Martínez Panero, M. (en prensa): "Métodos híbridos de votación bajo preferencias usuales y difusas". *Anales de Estudios Económicos y Empresariales*, XVII.
- Martínez Panero, M. – García Lapresta, J.L. (2003): *José Isidoro Morales, Precursor Ilustrado de la Teoría de la Elección Social: Edición Facsímil de la Memoria Matemática sobre el Cálculo de la Opinión en las Elecciones (1797) y Apéndice (1805)*. Secretariado de Publicaciones e Intercambio Editorial. Universidad de Valladolid, Valladolid. Prólogo de Salvador Barberà.
- Massó, J. (2003): Reseña de Martínez Panero – García Lapresta (2003). *La Gaceta de la Real Sociedad Matemática Española* 8, pp. 633–636.
- McLean, I. – Urken, A.B. (eds.) (1995): *Classics of Social Choice*. The University of Michigan Press, Ann Arbor.
- Morales, J.I. (1797): *Memoria Matemática sobre el Cálculo de la Opinión en las Elecciones*. Imprenta Real, Madrid.
- Morales, J.I. (1805): *Apéndice á la Memoria Matemática sobre el Cálculo de la Opinión en las Elecciones*. Imprenta de Sancha, Madrid.
- Mueller, D.C. (1979): *Public Choice*. Cambridge University Press, Londres. [Existe traducción al castellano de Juan Carlos Zapatero en Alianza Editorial, Madrid, 1984].
- Nitzan, S. – Rubinstein, A. (1981): "A further characterization of Borda ranking method". *Public Choice* 36, pp. 153–158.
- Nurmi, H. (1999): *Voting Paradoxes and How to Deal with Them*. Springer-Verlag, Berlín-Heidelberg.
- Nurmi H. (2004): "A comparison of some distance-based choice rules in ranking environments". *Theory and Decision* 57, pp. 5–24.
- Nurmi, H. (2005): "A responsive system". *Economics of Governance* 6, pp. 63–74.
- Riker, W.H. (1986): *The Art of Political Manipulation*. Yale University Press, New Haven.
- Saari, D.G. (1990): "The Borda dictionary". *Social Choice and Welfare* 7, pp. 279–317.
- Saari, D.G. (1994): *Geometry of Voting*. Springer-Verlag, Berlín.
- Saari, D.G. (1995): *Basic Geometry of Voting*. Springer-Verlag, Berlín.

- Saari, D.G. (2000): “Mathematical structure of voting paradoxes. II. Positional voting”. *Economic Theory* 15, pp. 55–102.
- Saari, D.G. – Barney, S. (2003): “Consequences of reversing preferences”. *Mathematical Intelligencer* 25(4), pp. 17–31.
- Sen, A.K. (1977): “Social Choice Theory: A re-examination”. *Econometrica* 45, pp. 53–89.
- Stein, W.E. – Mizzi, P.J. – Pfaffenberger, R.C. (1994): “A stochastic dominance analysis of ranked voting systems with scoring”. *European Journal of Operational Research* 74, pp. 78–85.
- Straffin Jr., P.D. (1980): *Topics in the Theory of Voting*. Birkhäuser, Boston.
- Sugden, R. (1981): *The Political Economy of Public Choice: An Introduction to Welfare Economics*. Martin Robertson, Oxford.
- Tangian, A.S. (2000): “Unlikelihood of Condorcet’s paradox in a large society”. *Social Choice and Welfare* 17, pp. 337–365.
- Tanguiane, A.S. (1991): *Aggregation and Representation of Preferences. Introduction to Mathematical Theory of Democracy*. Springer-Verlag, Berlín.
- Taylor, A.D. (2005): *Social Choice and the Mathematics of Manipulation*. Cambridge University Press, Nueva York.
- Todhunter, I. (1865): *A History of the Mathematical Theory of Probability from the Time of Pascal to that of Laplace*. Macmillan, Cambridge.
- Van Newenhizen, J. (1992): “The Borda method is most likely to respect the Condorcet principle”. *Economic Theory* 2, pp. 69–83.
- Young, H.P. (1974): “An axiomatization of Borda’s rule”. *Journal of Economic Theory* 9, pp. 43–52.
- Young, H.P. (1975): “Social Choice scoring functions”. *SIAM Journal on Applied Mathematics* 28, pp. 824–838.
- Young, H.P. (1997): “Group choice and individual judgements”, en Mueller, D. (ed.): *Perspectives on Public Choice: A Handbook*. Cambridge University Press, Cambridge, pp. 181–200.

COMPARACIÓN DE UN MODELO MULTIFACTORIAL DE LA ETTI: PARAMÉTRICO VERSUS NO PARAMÉTRICO

Lourdes Gómez del Valle
e-mail: lourdes@eco.uva.es

Julia Martínez Rodríguez
e-mail: julia@eco.uva.es

Departamento de Economía Aplicada
UNIVERSIDAD DE VALLADOLID

Área temática: Métodos Cuantitativos.

Resumen

En la literatura de los tipos de interés existe un debate continuo sobre cuántos factores son necesarios para explicar la estructura temporal de los tipos de interés (ETTI). Los modelos con un sólo factor ofrecen resultados muy interesantes debido a su sencillez. Sin embargo, son muchos los autores, por ejemplo Canabarro (1995), los que demuestran que son necesarios más factores para explicar la curva. Además los modelos multifactoriales son más adecuados si se quiere valorar activos complejos, como opciones sobre la curva de rendimientos, o planificar estrategias de inmunización. Estos modelos, fundamentalmente si se utilizan técnicas de estimación no paramétrica, presentan una mayor complejidad de estimación y valoración. Gómez-Valle y Martínez-Rodríguez (2006a) presentaron una nueva técnica para reducir el coste computacional de los modelos de dos factores.

En este trabajo consideramos un modelo de dos factores: el tipo de interés y la varianza de los cambios del tipo de interés, y mostramos que la técnica planteada por Gómez-Valle y Martínez-Rodríguez (2006a) para su estimación, además de reducir el coste computacional proporciona mejores resultados que modelos que utilizan técnicas paramétricas y modelos GARCH para su estimación, como Longstaff y Schwartz (1992).

Palabras clave: volatilidad estocástica, estimación no paramétrica, ETTI.

Abstract

In the financial literature there is no consensus about how many factors are necessary to explain the term structure of interest rates (TSSIR). The one-factor diffusion models are very attractive because their simplicity, however they have important unrealistic properties. In fact there are several authors, such as Canabarro (1995), who show that models with more stochastic factors describe more accurately the evolution of the term structure. Furthermore the multi-factor models are specially useful to price more complex interest-rate derivatives, such as the so-called yield options, or improve the efficiency of hedging strategies. However these models, specially if nonparametric methods are used, are more complex to estimate and to price. In order to reduce the higher computational cost of the two factor models, Gómez-Valle and Martínez-Rodríguez (2006a) showed a new approach to estimate these models.

In this paper, we consider a model with two factors: the interest rate and the instantaneous variance of the changes of the interest rate. Then, we show that the approach of Gómez-Valle and Martínez-Rodríguez (2006a) reduces the computational cost of estimation as well as provides better results than other two-factor parametric models based on the GARCH approach such as Longstaff and Schwartz (1992).

Key words: stochastic volatility, nonparametric approach, TSIR.

1. Introducción

El precio de los títulos de renta fija viene a menudo expresado en términos de varios tipos de interés y rendimientos, por tanto la comprensión de los modelos de este tipo equivale a entender el comportamiento de los tipos de interés. El concepto clave del análisis de los títulos de renta fija y del comportamiento de los tipos de interés es realmente la ETTI. Sin embargo, no es esta su única aplicación, dejando al margen todas sus aplicaciones macroeconómicas, desde un punto de vista financiero podemos afirmar que la ETTI es también una herramienta clave para la cobertura o la evaluación de estrategias de carteras de renta fija. La determinación de la dinámica de la ETTI permite definir medidas del riesgo, asociadas a variaciones de los tipos, que posibilita un mejor control de la eficacia de las estrategias de gestión de carteras, Nave (1988).

En la literatura se han considerado tradicionalmente modelos dinámicos de la ETTI que dependen únicamente de un factor o variable de estado que es el tipo de interés instantáneo, como por ejemplo Vasicek (1977) y Cox *et al* (1985). Estos modelos son muy atractivos para los investigadores y para los agentes financieros debido a su sencillez, ya que es posible conocer su solución exacta, se estiman fácilmente a partir de los datos del mercado y proporcionan curvas de rendimientos con formas muy realistas. Además son muchos los tests empíricos descritos en la literatura que aceptan ciertos modelos unifactoriales, Ferreira y Gil-Bazo (2004). Sin embargo, también presentan importantes desventajas. En primer lugar, la forma de la curva de rendimientos está muy restringida y estos modelos no son capaces de generar todos los tipos de variaciones de las curvas de rendimientos que se observan en los mercados. Finalmente y relacionado con el punto anterior, los cambios en períodos infinitesimales de tiempo de dos variables dependientes del tipo de interés están perfectamente correlacionados, Canabarro (1995) y Rebonato (1996).

Intuitivamente los modelos multifactoriales son, en general, más flexibles y son capaces de generar formas y movimientos adicionales de las curvas de rendimientos. Además los modelos multifactoriales permiten correlaciones no perfectas entre las diferentes variables que dependen de los tipos de interés.

Son muchos los estudios empíricos que han tratado de determinar cuántos factores son necesarios para explicar la ETTI, Litterman y Scheinkman (1991), Canabarro (1995), y los resultados parecen depender de los datos considerados, el período de observación y el procedimiento de estimación. Aunque sí que parece existir un consenso a la hora de determinar que son necesarios dos o tres factores para explicar el comportamiento de la ETTI de forma adecuada.

Sin embargo, también hay que tener en cuenta que los modelos multifactoriales son más difíciles de implementar y tienen un mayor coste computacional. Por lo tanto, es necesario encontrar un equilibrio entre una representación más realista del modelo y un menor coste computacional. De ahí que los modelos de dos factores, que son en los que nos centramos en este trabajo, puedan ser más atractivos para los agentes financieros.

Recientemente, debido a la imposición de restricciones arbitrarias sobre el comportamiento de las diferentes variables de estado, son muchos los autores que han comenzado a aplicar técnicas no paramétricas para la estimación de los modelos, Stanton (1997) en el caso de una variable y Boudoukh *et al* (1999) en el caso de dos variables.

En este trabajo, nosotros consideramos un modelo de dos factores: el tipo de interés instantáneo y la varianza instantánea de los cambios del tipo de interés, y comparamos su comportamiento a la hora de determinar las curvas de rendimiento cuando se utiliza el modelo paramétrico LS propuesto por Longstaff y Schwartz (1992) ó un modelo no paramétrico. Evidentemente la utilización de un modelo no paramétrico aumenta el coste computacional del modelo considerablemente, pero utilizando el enfoque propuesto por Gómez-Valle y Martínez-Rodríguez (2006a) obtenemos resultados muy favorables. Así observamos que para todos los vencimientos considerados, un enfoque no paramétrico proporciona curvas de rendimientos estimadas más próximas a las observadas que cuando se utiliza el modelo paramétrico LS.

El trabajo lo organizamos de la siguiente forma. En la Sección 2 describimos el funcionamiento de un modelo de ETTI con dos factores: el tipo de interés instantáneo y la varianza instantánea de las variaciones del tipo de interés. A continuación, en la Sección 3, mostramos los datos y las técnicas utilizadas para realizar el estudio empírico

objeto de este trabajo y, en la Sección 4 comparamos las curvas de rendimiento estimadas con los diferentes modelos con las observadas en el mercado de Estados Unidos. Finalmente, en la Sección 5 resumimos las principales conclusiones de este trabajo.

2. El modelo

En este trabajo nos centramos en modelos Markovianos homogéneos en el tiempo con dos variables de estado que son el tipo de interés instantáneo r_t y la varianza de los cambios del tipo de interés V_t . Para ello suponemos que estas variables siguen el siguiente proceso de difusión de Markov:

$$dr_t = \mu_r(r_t, V_t)dt + \sqrt{V_t}dW_1, \quad (1)$$

$$dV_t = \mu_V(r_t, V_t)dt + \sigma_V(r_t, V_t)dW_2, \quad (2)$$

$$E(dW_1, dW_2) = \rho(r_t, V_t),$$

donde las tendencias μ_r y μ_V , la difusión σ_V y el coeficiente de correlación ρ dependen de r_t y V_t .

Suponemos que $\mu_r(r_t, V_t)$, $\mu_V(r_t, V_t)$, $\sigma_V(r_t, V_t)$ y $\rho(r_t, V_t)$ están bien definidas de tal forma que exista la solución de (1) - (2) para un estado inicial arbitrario.

El precio de un bono cupón-cero en el instante actual t que vence en un instante T , $t \leq T$, y reporta 1 u.m. en dicho instante, lo consideramos como función del tipo de interés instantáneo y lo denotamos por $P(t, r, V; T)$, con

$$P(T, r, V; T) = 1. \quad (3)$$

Aplicando un razonamiento de arbitraje similar al propuesto por Black y Scholes (1973), construimos una cartera sin riesgo y utilizando la regla de diferenciación del Lema de Itô, como en Duffie (1996), el precio de un derivado del tipo de interés viene dado por la solución de la siguiente ecuación en derivadas parciales:

$$\begin{aligned} \frac{\partial P}{\partial t} + (\mu_r - \lambda_r \sqrt{V}) \frac{\partial P}{\partial r} + (\mu_V - \lambda_V \sigma_V) \frac{\partial P}{\partial V} + \frac{1}{2} V \frac{\partial^2 P}{\partial r^2} + \frac{1}{2} \sigma_V^2 \frac{\partial^2 P}{\partial V^2} \\ + \rho \sqrt{V} \sigma_V \frac{\partial^2 P}{\partial r \partial V} - rP = 0, \end{aligned} \quad (4)$$

donde λ_r y λ_V son los precios del riesgo de mercado asociados al tipo de interés y a la varianza. Estos precios del riesgo de mercado vienen determinados por el mercado y son los mismos para todos los derivados de los tipos de interés Björk (1998).

Para valorar un bono cupón-cero es necesario resolver la ecuación (4) sujeta a la condición final (3). El Teorema de Feynman-Kac bajo ciertas hipótesis, Öksendal (2003), nos permite expresar la solución de este problema de forma integral:

$$P(t, r, V; T) = E_t^* \left[\exp \left(- \int_t^T r_u du \right) \right], \quad (5)$$

donde E^* denota la esperanza condicionada bajo Q^* , conocida como medida de probabilidad neutral al riesgo.

El tanto de rendimiento al vencimiento, $R(t, r, V; T)$ es el rendimiento interno de un bono cupón cero en un instante de tiempo t y que vence en T , esto es,

$$R(t, r, V; T) = -\frac{1}{\tau} \log P(t, r, V; T) \quad t \leq T. \quad (6)$$

El conjunto de tantos al contado $R(t, r, V; T_j)$ ($t \leq T_j$, con $j = 1, 2, \dots, N$) recibe el nombre de estructura temporal de los tipos de interés al contado o curva de rendimientos (*yield curve*).

Para poder valorar cualquier derivado del tipo de interés es necesario conocer y estimar μ_r , μ_V , σ_V , λ_r , λ_V y ρ . A continuación se reemplazan en la ecuación (4) y se resuelve teniendo en cuenta la correspondiente condición final que depende del derivado a valorar. La ecuación (4) es la base para su resolución utilizando un método en diferencias finitas, y (5) si se resuelve utilizando el método de simulación de Monte Carlo.

LS es uno de los modelos de dos factores más conocido en la literatura. Longstaff y Schwartz (1992) presentan un modelo de equilibrio general de la ETTI, y parten de una descripción de la economía subyacente. Establecen las hipótesis acerca de la evolución estocástica de dos variables relevantes de la economía x e y , y de las preferencias de un inversor representativo. Suponen que la variable x afecta únicamente a los rendimientos esperados de las inversiones productivas y la variable y , afecta tanto a los rendimientos esperados como a la incertidumbre de los rendimientos de la inversión productiva. Ambas variables siguen los siguientes procesos independientes:

$$dx = (\gamma - \delta x)dt + \sqrt{x}dZ_1, \quad (7)$$

$$dy = (\eta - \zeta y)dt + \sqrt{y}dZ_2. \quad (8)$$

El hecho de que los dos factores anteriores no estén especificados no es relevante, ya que una de las implicaciones más importantes del modelo es que tanto el tipo de interés como la varianza de los cambios de los tipos de interés son funciones lineales de r y V :

$$r_t = \alpha x + \beta y,$$

$$V_t = \alpha^2 x + \beta y,$$

donde α y β provienen del proceso productivo que se impone en el modelo de equilibrio general, ver Longstaff y Schwartz (1992) para más información. Una ventaja fundamental de este modelo es que se conoce su solución exacta.

Estas dos variables estocásticas fueron inicialmente consideradas por Fong y Vasicek (1991), pero su modelo es de equilibrio parcial y no posee solución exacta. Adicionalmente los tipos de interés pueden tomar valores negativos. Posteriormente Andersen y Lund (1997) y Vetzal (1997) utilizaron también el tipo de interés y la varianza de los cambios del tipo de interés para explicar el comportamiento de la ETTI, pero imponiendo diferentes restricciones paramétricas a las funciones. Sin embargo, no existe ningún consenso sobre que funciones paramétricas son más adecuadas para explicar el comportamiento de la ETTI. Por tanto posteriormente, Boudoukh *et al*

(1999) y Downing (1999) para evitar imponer restricciones arbitrarias a las funciones de los procesos estocásticos, consideraron técnicas de estimación no paramétricas.

3. Estimación e implementación de los modelos

En esta sección comparamos los errores al estimar las curvas de rendimientos con el modelo paramétrico LS y un modelo no paramétrico.

Para poder estimar los coeficientes de la ecuación en derivadas parciales (4), necesitamos disponer de una serie de tiempo del tipo de interés instantáneo libre de riesgo y de su varianza. Nosotros utilizamos, concretamente, las observaciones semanales de los rendimientos de los Treasury Bills (o T-Bills) del mercado secundario de Estados Unidos, lo cual es una práctica habitual en la literatura, como por ejemplo Chan *et al* (1992) y Stanton (1997).

El período de estimación lo consideramos desde enero de 1983 hasta diciembre de 2005, y transformamos los rendimientos en tipos de interés anualizados sin realizar ningún ajuste específico por fines de semana o vacaciones. Los datos los obtenemos de la Reserva Federal de Estados Unidos (h.15) y tenemos un total de 1.200 observaciones.

El tipo de interés instantáneo no es directamente observable por lo que lo aproximamos con el rendimiento a 3 meses, ya que supone un compromiso entre la elección de un vencimiento lo suficientemente pequeño y que a su vez no incluya los efectos propios del mercado. Por ejemplo los rendimientos diarios pueden ser muy volátiles porque incluyen características puntuales propias del funcionamiento de los mercados. En la Figura 1a) representamos los rendimientos de los Treasury Bills a 3 meses y en la Figura 1b) sus primeras diferencias. La Tabla 1 recoge los estadísticos descriptivos de las series del tipo de interés y de sus diferencias.

	Media	Desv. típica	Máximo	Mínimo	ρ_1	ρ_2
r_t	0.051264	0.023166	0.107943	0.008409	0.998	0.996
$r_{t+1} - r_t$	3.85 e-6	5.05 e-5	9.57 e-5	-0.000169	0.195	-0.010

Tabla 1. Resumen de los estadísticos más importantes del tipo de interés observado desde enero de 1983 hasta diciembre de 2005 y sus primeras diferencias.

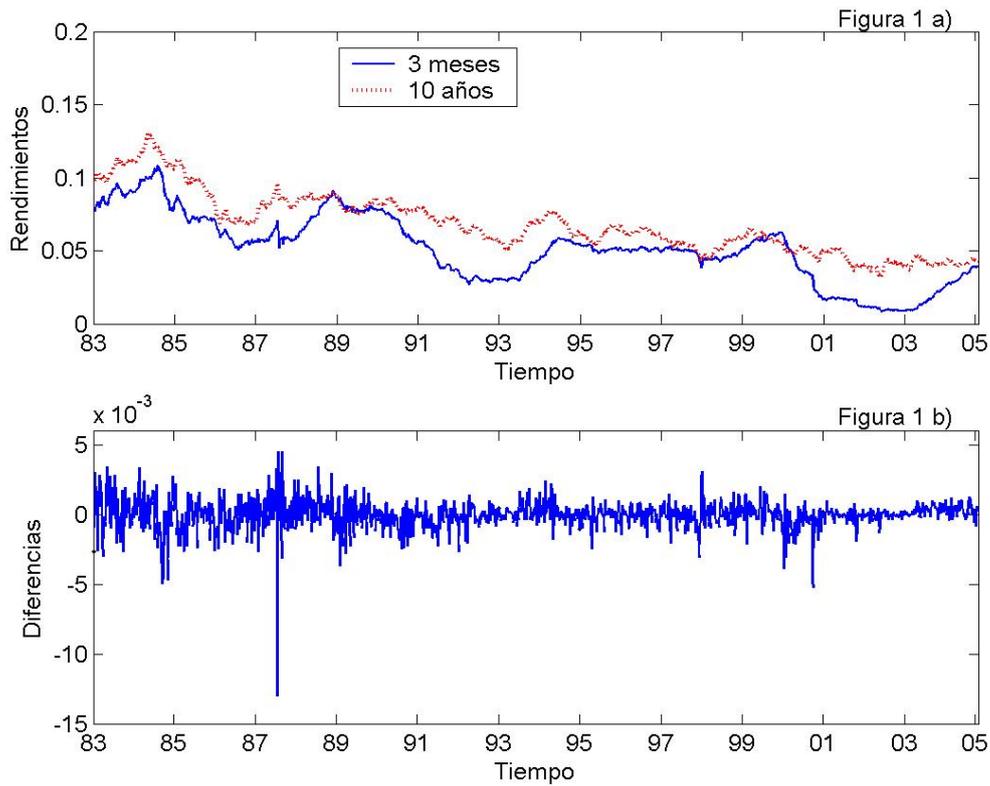


Figura 1: Fig 1a) Rendimientos de los Treasury Bills a 3 meses y de los bonos a vencimiento constante de 10 años de USA. Fig 1b) Primeras diferencias de los rendimientos de los Treasury Bills a 3 meses.

La varianza de los cambios del tipo de interés instantáneo no es tampoco directamente observable en el mercado, pero es posible estimarla. Algunos autores proponen su estimación mediante procesos GARCH, como Longstaff y Schwartz (1992). Concretamente, proponen un modelo GARCH(1,1) con varianza condicionada en la ecuación en media,

$$r_{t+1} - r_t = \alpha_0 + \alpha_1 r_t + \alpha_2 V_t + e_{t+1} \quad (9)$$

$$e_{t+1} \rightarrow N(0,1)$$

$$V_t = \beta_0 + \beta_1 r_t + \beta_2 V_{t-1} + \beta_3 e_t^2 \quad (10)$$

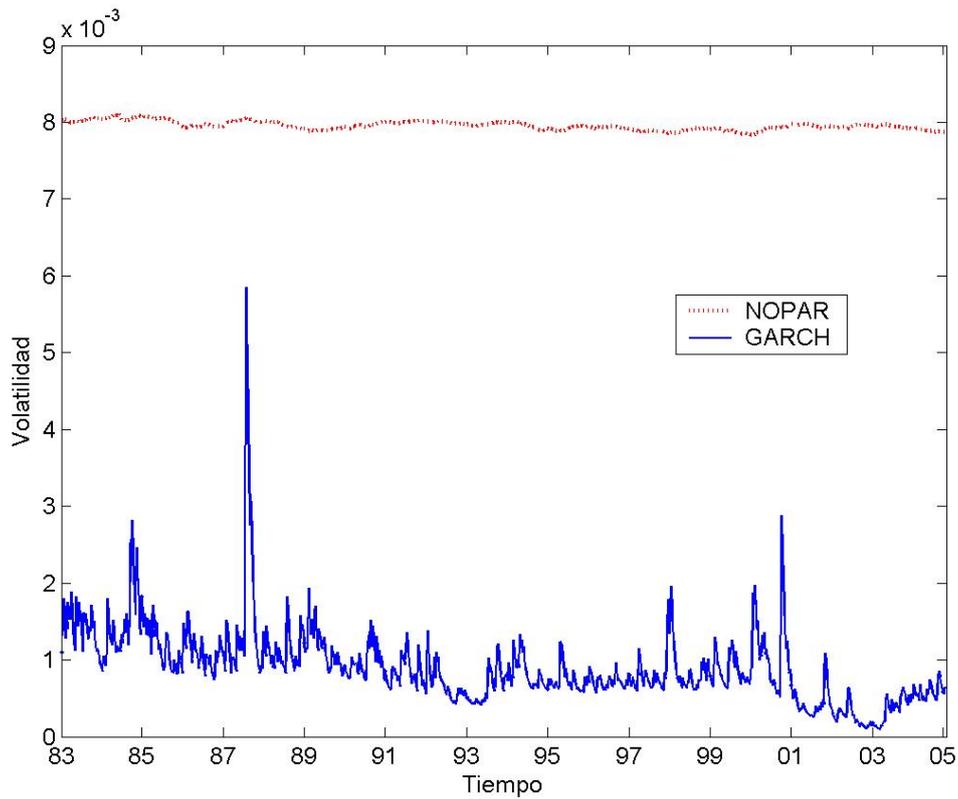


Figura 2: Volatilidad estimada utilizando un método no paramétrico y un método paramétrico (GARCH).

En la Tabla 2 mostramos los resultados de esta estimación y en la Figura 2 representamos gráficamente su comportamiento a lo largo del tiempo.

	α_0	α_1	α_2	β_0	β_1	β_2	β_3
Valor	1.21e-5	2.39e-4	-13.08	-1.85e-8	2.20e-6	6.85e-1	1.90e-1
Estadístico z	7.25e-1	2.31e-1	-2.49e-1	-5.84	18.39	57.67	15.61
p- Valor	4.62e-1	8.17e-1	8.03e-1	5.07e-9	0.00	0.00	0.00

Tabla 2: Parámetros estimados del modelo GARCH(1,1) para la volatilidad (9) - (10).

En primer lugar consideramos que la función $V(r, S)$ es inversible, es decir que podemos escribir el modelo (1) y (2) de forma equivalente

$$dr_t = \alpha_r(r_t, S_t)dt + \sqrt{V(r_t, S_t)}dZ_1^*,$$

$$dS_t = \alpha_s(r_t, S_t)dt + \beta_s(r_t, S_t)dZ_2^*,$$

donde Z_1 y Z_2 pueden estar correlacionados. La función $V(r, S)$ la estimamos utilizando la metodología descrita en Boudoukh *et al* (1999). En la Figura 2 representamos la volatilidad estimada utilizando un modelo GARCH y utilizando técnicas de estimación no paramétrica. En el caso de la estimación utilizando un modelo GARCH los valores oscilan entre 0 y $6e-3$ y en el caso de la estimación no paramétrica las variaciones son menores, pero los valores son mayores y próximos a $8e-3$.

Una vez analizados los factores del modelo, estimamos los parámetros del modelo LS y las funciones no paramétricas como paso previo a la obtención de la curva de rendimientos.

Los parámetros del modelo LS, (5)-(6) los estimamos utilizando únicamente la varianza y el tipo de interés observado mediante las relaciones obtenidas por Longstaff y Scwhartz (1993). En la Tabla 3 presentamos los valores de los seis parámetros estimados.

Finalmente, es necesario estimar el precio del riesgo de mercado y para ello, siguiendo a Longstaff y Scwhartz (1993), utilizamos la propiedad de que la solución exacta para un bono cupón-cero es conocida. Por tanto, este parámetro se obtiene minimizando la raíz cuadrada del error cuadrático medio,

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (P_i - \hat{P}_i)^2},$$

donde N es el número de observaciones, P_i el precio observado y \hat{P}_i el precio estimado para la observación i -ésima. Los precios observados que utilizamos son los correspondientes a los bonos con vencimiento constante a 6 meses, 1, 2, 3, 5, 7 y 10

años del mercado de Estados Unidos proporcionados por la Reserva Federal. El valor estimado de este parámetro es a partir de los datos anteriores -0.36821 .

En la implementación de las técnicas no paramétricas seguimos el enfoque propuesto por Gómez-Valle y Martínez-Rodríguez (2006a),

$$\frac{\partial R}{\partial T} \Big|_{T=t} = \frac{1}{2} (\mu_r(r, V) - \sigma_r(r, V) \lambda_r(r, V)), \quad (11)$$

$$\frac{\partial}{\partial T} (VP) \Big|_{T=t} = -r(t)V(t) + (\mu_v(r, V) - \sigma_v(r, V) \lambda_v(r, V)). \quad (12)$$

Este enfoque se basa en la estimación de la tendencia neutral al riesgo de ambas variables directamente a partir de la pendiente de la curva en el origen definida en Vasicek (1977).

Las funciones $\frac{\partial R}{\partial T}$ y $\frac{\partial (VP)}{\partial T}$ las aproximamos mediante integración numérica. En cuanto a la difusión de la varianza instantánea del tipo de interés, realizamos su estimación con una aproximación de primer orden como en Stanton (1997). En todos los casos utilizamos el método del núcleo para la estimación no paramétrica, ver Härdle (1990) para más información.

Una ventaja fundamental de este enfoque frente a otros, como el propuesto por Boudoukh *et al* (1999), consiste en que reducimos considerablemente el coste computacional del modelo ya que únicamente estimamos 4 funciones frente a las 6 que estiman Boudoukh *et al* (1999).

γ	δ	η	ξ	α	β
2.3310	4.074e-5	1.5295e-1	5.4986e-2	8.6710e-7	5.9716e-4

Tabla 3: Parámetros del modelo LS utilizando el sistema de ecuaciones propuesto por Longstaff y Schwartz (1993).

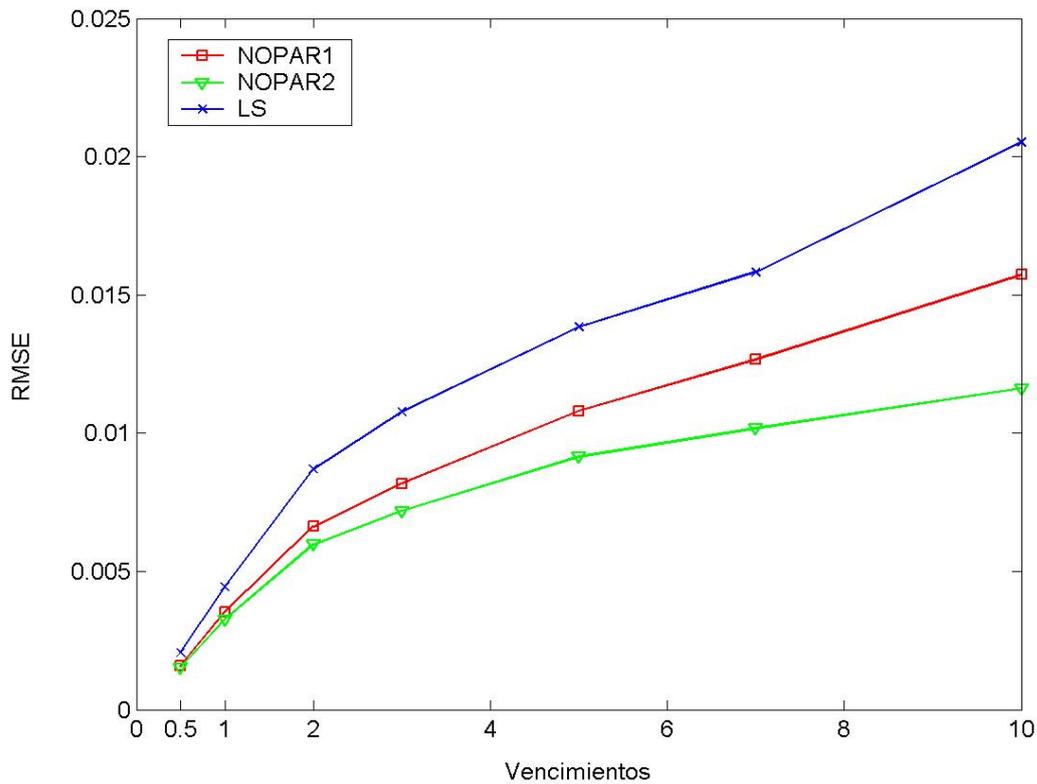


Figura 3: Error cuadrático medio de las curvas de rendimiento a diferentes vencimientos desde enero de 1983 hasta diciembre de 1995.

4. Valoración de bonos cupón cero

En esta sección obtenemos los precios de los bonos cupón-cero y posteriormente las curvas de rendimiento, utilizando el modelo paramétrico LS y un modelo no paramétrico basado en la técnica propuesta por Gómez-Valle y Martínez-Rodríguez (2006a). Posteriormente comparamos las curvas observadas frente a las estimadas con el modelo paramétrico frente al modelo no paramétrico.

La obtención de los precios en el modelo paramétrico LS es inmediata utilizando su solución exacta, Longstaff y Schwartz (1992) y es ampliamente conocida en la literatura. Posteriormente utilizamos la relación (6) para obtener la curva de rendimientos. Sin embargo, cuando empleamos un modelo no paramétrico la solución no es conocida por lo que es necesario aplicar un método numérico para la resolución de la ecuación (4) sujeta a la condición final (3). En la literatura se ha aplicado

habitualmente el Método de Simulación de Monte Carlo directamente sobre (5), sin embargo nosotros aplicamos un método en diferencias directamente sobre la ecuación en derivadas parciales que reduce su coste computacional, como en Gómez-Valle y Martínez-Rodríguez (2006a). Concretamente utilizamos una subrutina de la librería NAG para Fortran. De nuevo, obtenemos las curvas de rendimientos aplicando (6).

En la Figura 3 mostramos los resultados de esta estimación y, representamos el error cuadrático medio entre el rendimiento observado y el rendimiento estimado para cada uno de los modelos para períodos de vencimiento de 6 meses y 1, 2, 3, 5, 7 y 10 años. Así observamos que el modelo no paramétrico (NOPAR2) presenta menores errores para todos los vencimientos. Este resultado es análogo al obtenido en Gómez-Valle (2005) para modelos de un factor, donde se demostraba que el error cuadrático medio en modelos no paramétricos de un factor era menor que cuando se utilizaban modelos paramétricos como los de Cox *et al* (1985) y Vasicek (1977). A la hora de realizar comparaciones no tenemos en cuenta la técnica de estimación empleada por Boudoukh *et al* (1999) porque como ya hemos demostrado en trabajos anteriores Gómez-Valle y Martínez-Rodríguez (2006b), la técnica propuesta por Gómez-Valle y Martínez-Rodríguez (2006a) además de reducir el coste computacional, proporciona curvas de rendimientos más próximas a las observadas que la técnica propuesta por Boudoukh *et al* (1999).

Finalmente, en la Figura 3 también comparamos el comportamiento de los modelos anteriores con un modelo de un solo factor no paramétrico utilizando la técnica propuesta por Gómez-Valle y Martínez-Rodríguez (2006c), y que se basa en la estimación no paramétrica de la tendencia neutral al riesgo directamente a partir de los datos utilizando la aproximación,

$$\frac{\partial R}{\partial T} \Big|_{T=t} = \frac{1}{2} (\mu_r(r) - \sigma_r(r) \lambda_r(r)).$$

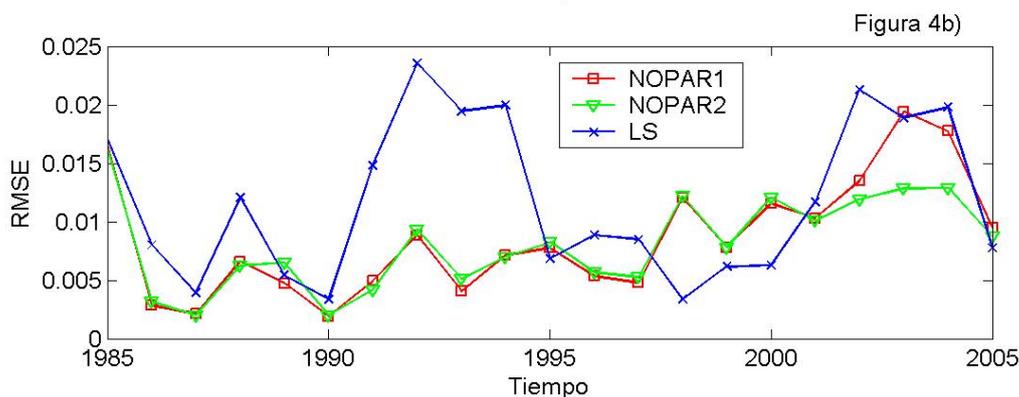
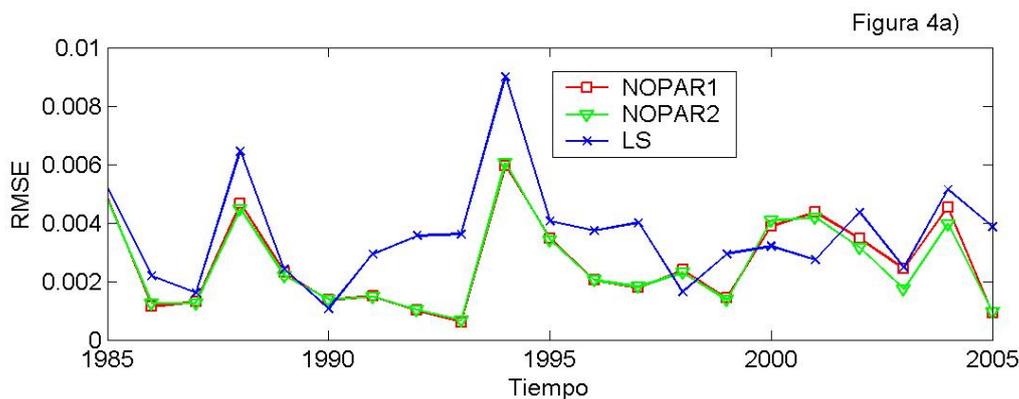


Figura 4: El error cuadrático medio anual de los diferentes modelos durante el periodo de estimación para los rendimientos a 1 año (Figura 4a) y a 5 años (Figura 4b).

La pendiente de la curva en el origen la aproximamos mediante integración numérica y las estimaciones las realizamos con el método no paramétrico de regresión local lineal. En este caso observamos, que este modelo no paramétrico unifactorial (NOPAR1) proporciona mayores errores que el modelo no paramétrico de dos factores (NOPAR2) para los diferentes vencimientos. Sin embargo, también observamos que proporciona menores errores que el modelo paramétrico de dos factores considerado. Finalmente, es importante destacar que en todos los casos el modelo NOPAR2 parece ser el más adecuado y exacto a la hora de determinar las curvas de rendimientos y, las diferencias aumentan con el período de vencimiento. Resultados análogos se obtienen en el trabajo de Corzo y Gómez (2005) .

En la Figura 4 mostramos el error cuadrático medio anual para los diferentes modelos y con vencimientos iguales a 1 año (Figura4a)) y 5 años (Figura 4b)). En ambas gráficas observamos que en general, los modelos no paramétricos presentan menores errores que

el modelo paramétrico LS, excepto algún año en concreto. El comportamiento de los modelos no paramétricos es muy similar y sus diferencias son pequeñas, pero en la mayoría de los casos el modelo NOPAR2 es más preciso a la hora de determinar los rendimientos de los bonos cupón-cero que el modelo NOPAR1. El comportamiento para el resto de vencimientos observados no varía.

5. Conclusiones

En este trabajo consideramos un modelo dinámico homogéneo de la ETTI, en el que suponemos dos variables de estado: el tipo de interés instantáneo y la varianza instantánea de las variaciones del tipo de interés. En la literatura son muchos los comentarios favorables al modelo paramétrico LS, de tal forma que puede considerarse como un referente a la hora de realizar comparaciones entre modelos.

Recientemente se han aplicado técnicas de estimación no paramétrica a los modelos existentes en la literatura, ya que no se conoce el comportamiento exacto de las variables de estado y por ello habitualmente, se imponen restricciones arbitrarias sobre el comportamiento de las funciones. Sin embargo, es interesante destacar que la implementación de un modelo no paramétrico aumenta considerablemente el coste computacional del modelo.

Por tanto, el objetivo de este trabajo consiste en analizar si un modelo de dos factores no paramétrico proporciona mejores resultados que uno paramétrico. Para ello utilizamos la técnica propuesta por Gómez-Valle y Martínez Rodríguez (2006a) y, observamos que las curvas estimadas con un modelo no paramétrico de dos factores se aproximan más a las observadas que las obtenidas con uno paramétrico. Para ello utilizamos datos observados del mercado de Estados Unidos desde enero de 1985 hasta diciembre de 2005.

Finalmente consideramos un modelo no paramétrico con un solo factor y observamos que al igual que sucede con otros estudios realizados con estimaciones paramétricas, un modelo de dos factores proporciona menores errores que uno con dos factores. Por tanto, podemos afirmar que a pesar de que un modelo no paramétrico de un factor ofrece curvas de rendimientos más próximas a las observadas que los modelos

paramétricos tradicionales de la ETTI (ver Gómez-Valle (2005)), la incorporación de una segunda variable de estado mejora los resultados, lo cual es fundamental para la posterior valoración de activos derivados del tipo de interés y la evaluación de estrategias de cobertura.

Bibliografía

Andersen, T. G. y J Lund (1997): “Estimating continuous-time stochastic volatility models of the short-term interest rate”, *Journal of Econometrics*, 77, pp. 343-377.

Björk, T. (1998): *Arbitrage Theory in Continuous Time*, Oxford University Press, Oxford, UK.

Boudoukh, J., M. Richardson, R. Stanton y R. F. Whitelaw (1999): “A multifactor, nonlinear, continuous-time model of interest rate volatility”, *Working Paper 99-042*, Stern School of Business.

Canabarro, E. (1995): “Where do one factor interest rate models fail?”, *The Journal of Fixed Income*, 5, 2, September, pp. 31-52.

Corzo T. y J. Gómez (2005): “Nonparametric estimation of convergence of interest rates: effects on bond pricing”, *Spanish Economic Review*, 7, pp. 167 – 190.

Cox, J. C., J. E. Ingersoll y S. A. Ross (1985): “A theory of the term structure of interest rates”, *Econometrica*, 53, pp. 385-408.

Chan, K. C., G. A. Karolyi, F. A. Longstaff y A. Sanders (1992): “An empirical comparison of alternative models of the short term interest rates”, *The journal of Finance*, 47, pp. 1209 – 1227.

Downing C. (1999): “Nonparametric estimation of multifactor continuous-time interest models”, *Board of Governors of the Federal Reserve System- Finance and Economics Discussion Series 99/62*.

Duffie, D. (1996): *Dynamic Asset Pricing Theory*, Princeton University Press, New Jersey.

Ferreira, E. y J. Gil-Bazo (2004): “Beyond Single-Factor affine term structure models”, *Journal of financial Econometrics*, 2, 4, pp. 565 – 591.

Fong H. G. y O. Vasiceck (1991): “Fixed-income volatility management”, *The Journal of Portfolio Management*, 17, 4, Summer, pp. 41 – 46.

Gómez-Valle, L. (2005): “Nuevos planteamientos en modelos unifactoriales de la estructura temporal de los tipos de interés”, Tesis doctoral nº 3156558, Proquest Information and Learning, España.

Gómez-Valle, L. y J. Martínez-Rodríguez (2006a): “Efficient estimation in models of the term structure of interest rates”, *ICNAAM 2006*, pp. 135 - 138.

Gómez-Valle, L. y J. Martínez-Rodríguez (2006b): “Direct nonparametric estimation of the risk-neutral dynamic: a two-factor model of the term structure”, *IX Congreso Hispano-Italiano de Matemática Financiera y Actuarial*.

Gómez-Valle, L. y J. Martínez-Rodríguez (2006c): “Modelling the term structure of interest rates: an efficient nonparametric approach”, aceptado en *Journal of Banking and Finance*.

Härdle, W. H., (1990): *Applied Nonparametric Regression*, *Econometric Society Monographs*, no. 19, Princeton University, New York.

Litterman, R. y J. Scheinkman (1991): “Common factors affecting bond returns”. *The Journal of Fixed Income*, June, pp. 54-61.

Longstaff, F. A. y E. Schwartz (1992): “Interest rate volatility and the term structure: a two factor general equilibrium model”, *The Journal of Finance*, XLVII, 4, September, pp. 1259 - 1282.

Longstaff, F. A. and Schwartz E. (1993): “Implementation of the Longstaff-Schwartz interest rate model”, *The Journal of Fixed Income*, 49, 4, July-August, pp. 70-74.

Nave, J. M. (1998): *Estructura temporal de los tipos de interés e inmunización financiera en el mercado español*. Tesis Doctoral. Universidad de Valencia. Departamento de Economía Financiera y Contabilidad.

Øksendal, B. (2003): *Stochastic differential Equations: An Introduction with Applications*, Springer-Verlag, Germany.

Rebonato, R. (1996): *Interest Rate Option Models. Understanding, Analysing and Using Models for Exotic Interest-Rate Options*. John Wiley & Sons, Ltd., England.

Stanton, R. (1997): “A nonparametric model of term structure dynamics and the market price of interest rate risk”, *Journal of Finance*, 52, 5, December, pp. 1973–2002.

Vasicek, O. (1977): “An equilibrium characterization of the term structure”, *Journal of Financial Economics*, Vol.5, pp. 177-188.

Vetzal, K. R. (1997): “Stochastic Volatility, Movements in the short term interest rates, and bond option values”, *Journal of Banking and Finance*, 21, pp. 169 – 196.

NUEVAS DISTRIBUCIONES EN LA METODOLOGÍA PERT

MARÍA DEL MAR LÓPEZ MARTÍN
e-mail: c.e.gamma@hotmail.com

CATALINA BEATRIZ GARCÍA GARCÍA
e-mail: cbgarcia@ugr.es

Departamento de Métodos Cuantitativos para la Economía y la Empresa
UNIVERSIDAD DE GRANADA

Área temática: Métodos Cuantitativos

Resumen

En el presente trabajo se presenta la distribución cúbica y la distribución bi-seno estudiándose sus características estocásticas principales y desarrollando su proceso de generalización. Desde el punto de vista práctico, se presentan como posibles alternativas a la distribución beta en la aplicación de la metodología PERT. Los criterios presentados por Taha (1981) y desarrollados por Herrerías (1989) consideran que la distribución aplicada, en la metodología PERT, debe tener una media moderada y una varianza lo mayor posible. Se han analizado las medias y las varianzas de las distribuciones presentadas y se han comparado con otras distribuciones aplicadas tradicionalmente en la metodología PERT, comprobándose que ambas son candidatas adecuadas como distribuciones subyacentes de la metodología PERT.

Palabras clave: Incertidumbre, metodología PERT, distribución bicúbica, distribución bi-seno.

Abstract

This work presents the cubic and bi-sine distributions, studies its principals stochastic characteristics and its generalization process. Considering practical aspect, these distributions are presented as alternative for the beta distribution in PERT methodology. The criteria presented by Taha (1981) and development by Herrerias (1989) considers that a appropriate candidate has to present a moderate average and a variance as large than possible. It has been analyzed the average and the variance of the presented distributions and has been compared with other distributions traditionally applied in methodology PERT, verifying that both are appropriate underlying distributions of methodology PERT.

Key words: uncertainty, PERT methodology, cubic distributions, bi-sine distributions.

1. INTRODUCCIÓN

Para controlar los tiempos de ejecución de los proyectos espaciales de la Armada de los Estados Unidos de América, se desarrolló *PROGRAM EVALUATION AND REVIEW TECHNIQUE*, denominado método PERT. Originalmente se utilizó para el control de tiempos del proyecto Polaris y actualmente se utiliza en todo el programa espacial, además de en otros ámbitos como, por ejemplo, la Investigación de Operaciones y, en general, el Análisis Económico. Como aplicaciones concretas, se destacan el estudio de la duración de un proyecto de fabricación en función de la duración de las diferentes tareas (Romero 1991) o el análisis de la bondad de un proyecto de inversión mediante sus diversos flujos de caja actualizados según su valor capital, (Suárez 1980).

En el presente trabajo se muestra la distribución bicúbica y la distribución bi-seno como posibles modelos alternativos aplicables en el ambiente de incertidumbre, estudiando sus principales características. A partir los criterios definidos por Taha (1981) y desarrollados por Herrerías (1989), concluimos que es una distribución adecuada para su uso en la metodología PERT.

En la sección 1, se presenta una recapitulación de los modelos probabilísticos univariantes utilizados en la metodología PERT.

En la sección 2, a partir de los estudios realizados por van Dorp y Kotz (2003), presentamos las expresiones de las funciones de la distribución bicúbica, estudiando sus principales características estocásticas y su forma.

En la sección 3, se presenta la distribución bi-seno, desarrollando las principales características de dicha distribución.

En la sección 4, se realiza un estudio de la distribución bicúbica y bi-seno aplicadas en el ámbito del PERT, comparando sus valores esperados y varianzas con las distribuciones más utilizadas en el ámbito de incertidumbre.

En la sección 5, a partir de los resultados obtenidos en la sección 2, se realiza una generalización de la distribución bicúbica a una rama a través del esquema multiplicativo, calculando sus momentos y realizando posteriormente un análisis sobre su forma.

1.1. MODELOS PROBABILÍSTICOS DE LA METODOLOGÍA PERT

La principal distribución utilizada en la metodología PERT para el estudio de la duración de una actividad asumía es la distribución beta con una media y desviación típica basadas en el valor modal y los valores extremos de la distribución subyacente. Luego, las principales críticas se debieron a que siempre se suponía que la duración de una actividad sigue una distribución beta, por ello Ben Yair (2000) realiza un trabajo en el que se comprueba que, en el caso de una distribución asimétrica a la derecha, las estimaciones de la moda y la desviación típica serán asintóticamente sobreestimadas.

Durante mucho tiempo varios autores han optado por presentar modificaciones sobre las fórmulas clásicas del PERT como se puede comprobar en Johnson (1998) y además se han aportado distintos modelos.

A continuación, se presenta de forma reducida, las principales distribuciones utilizadas en el ambiente de incertidumbre:

● **DISTRIBUCIÓN BETA** La distribución beta fue usada originalmente por los creadores de la metodología PERT, con el fin de superar los principales inconvenientes que presentaba la distribución normal. Dicha distribución presenta el inconveniente de que su varianza no incluye su valor modal, presentando Herrería (1995) una posible solución a dicho problema, pero ha sido utilizada con gran éxito en la metodología PERT por presentar unas buenas propiedades con respecto a su asimetría y curtosis. Dicha distribución tiene un recorrido limitado, no presentando colas infinitas y además de presentar una forma simétrica también puede adoptar asimetrías.

Con el fin de resaltar la rigidez de este modelo y salvaguardar la flexibilidad modeladora, Golenko-Ginzburg en 1988, mediante una reparametrización de la distribución beta, y posteriormente Herrerías en 1989, utilizando el sistema generador de Pearson, han desarrollado, por caminos distintos, unos modelos alternativos para estimar $E[X]$ y $Var[X]$.

● **DISTRIBUCIÓN TRIANGULAR** Es uno de los modelos más usados en el análisis de inversiones y se plantea como un modelo alternativo al modelos clásico, (Simpson; 1755,1757) . La distribución triangular puede ser simétrica o asimétrica siendo aplicable en el estudio de tiempo para completar ciertas actividades dentro de un proyecto global con cierta incertidumbre, Winston (1993). Recientemente ha adquirido popularidad por ser utilizada en el método de simulación de Monte Carlo, (Vose 1996).

● **DISTRIBUCIÓN TRAPEZOIDAL** Se presentó como una alternativa a la distribución beta, siendo utilizada en el estudio y detección del cáncer, Flehinger y Kimmel 1987 y Brown 1999 y en el análisis del riesgo por Pouliquen 1970, Powell y Wilson 1997 y Garvey 2000. Al ser dicha distribución tetraparamétrica no se puede ajustar perfectamente a partir de las tres estimaciones a , b y m , por ello Callejón, Pérez y Ramos en 1998 desarrollaron un procedimiento para disponer de una distribución trapezoidal en el caso de que solo se dispongan de los valores a , b y m , denominada **Trapezoidal CPR**.

● **DISTRIBUCIÓN TWO-SIDED POWER** Fue introducida por van Dorp y Kotz (2002) como una generalización de la distribución triangular. Esta distribución, pertenece al sistema de Pearson, Herrerías (2004). Dicha distribución fue introducida para extender las aplicaciones de la distribución triangular a los problemas asociados al riesgo y la incertidumbre, además se presenta como una alternativa versátil y flexible a la beta de dos y cuatro parámetros.

● **DISTRIBUCIÓN TOPP-LEONE** Es una distribución unimodal continua y acotada encontrándose en el artículo de Topp y Leone de 1955 los primeros estudios. En un principio recibió poca atención pero recientemente ha sido estudiada por Nadarajah y Kotz (2003), comprobando que dicha distribución es la más adecuada para ajustar fenómenos de tiempo por presentar una tasa de fallo en forma de bañera. Ghitany, Kotz y Xie (2005) dedican un artículo a las medidas de bondad y las características estocásticas de dicha distribución.

● **DISTRIBUCIÓN BIPARABÓLICA** Ha sido recientemente descrita por García (2007) como una generalización de la distribución parabólica. Dicha distribución está definida en un dominio acotado, determinada a partir de los tres valores clásicos y además su función de densidad es derivable en su valor modal, una característica que la distingue de las distribuciones más utilizadas en el ambiente de incertidumbre.

1.2. CONSTRUCCIÓN DE LA DISTRIBUCIÓN BICÚBICA

Según las investigaciones realizadas por van Dorp y Kotz (2003), si $p(.|\psi)$ es una función de densidad definida en el intervalo $[0,1]$ con parámetros, o vector de parámetros, ψ , entonces podemos definir la función de densidad unimodal en M de la siguiente forma:

$$g(t/M, p(./\psi)) = \begin{cases} p\left(\frac{t}{M}/\psi\right) & \text{si } 0 \leq t \leq M \\ p\left(\frac{1-t}{1-M}/\psi\right) & \text{si } M \leq t \leq 1 \end{cases} \quad (2.1)$$

donde $p(./\psi)$ es la función generadora de $g(t/M, p(./\psi))$.

Como la función de densidad asociada a una distribución cúbica viene dada por $p(y) = ay^3 + by^2 + cy + d$, con a, b, c y d valores reales, e imponiendo las condiciones que se observan en la Figura 2.1:

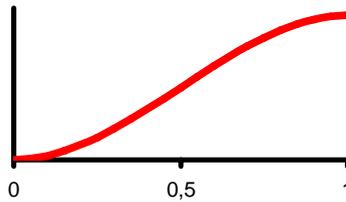


Figura 2.1: Representación de $p(y) = ay^3 + by^2 + cy + d$

1. Pasa por el origen de coordenadas, $p(0)=0$
2. Presenta un extremo relativo en $y=1$; $p'(1) = 0$
3. Presenta un extremo relativo en $y=0$; $p'(0) = 0$
4. $\int_0^1 p(y)dy = 1$

calculamos el valor de dichos parámetros, obteniendo pues, la expresión de la función de densidad cúbica generadora que viene dada por:

$$p(y) = 6\left(y^2 - \frac{2}{3}y^3\right) \quad (2.2)$$

A partir de la función de densidad obtenida podemos hallar la expresión de la función de densidad de la distribución bicúbica estandarizada descrita en (2.1):

$$g(t/M, p(y)) = \begin{cases} 6\left(\frac{t^2}{M^2} - \frac{2}{3}\frac{t^3}{M^3}\right) & \text{si } 0 \leq t \leq M \\ 6\left(\left(\frac{1-t}{1-M}\right)^2 - \frac{2}{3}\left(\frac{1-t}{1-M}\right)^3\right) & \text{si } M \leq t \leq 1 \end{cases} \quad (2.3)$$

Teniendo en cuenta la función de densidad descrita por van Dorp y Kotz (2003) definimos la función de distribución, G , utilizando la función de distribución $P(./\psi)$ asociada a la función de densidad generadora $p(./\psi)$:

$$G(t/M, P(./\psi)) = \begin{cases} M P\left(\frac{t}{M}/\psi\right) & \text{si } 0 \leq t \leq M \\ 1 - (1-M) P\left(\frac{1-t}{1-M}/\psi\right) & \text{si } M \leq t \leq 1 \end{cases} \quad (2.4)$$

Como la función de distribución asociada a $p(y)$ es $P(y) = 2y^3 - y^4$ definimos la función de distribución bicúbica estandarizada, utilizando el sistema generador de van Dorp y Kotz (2003), como:

$$G(t/M, P(y)) = \begin{cases} M \left(2 \frac{t^3}{M^3} - \frac{t^4}{M^4} \right) & \text{si } 0 \leq t \leq M \\ 1 - (1-M) \left(2 \left(\frac{1-t}{1-M} \right)^3 - \left(\frac{1-t}{1-M} \right)^4 \right) & \text{si } M \leq t \leq 1 \end{cases} \quad (2.5)$$

Para hallar las expresiones de la función de densidad y de distribución bicúbica para la variable x definida en el intervalo $[a, b]$, realizaremos la desestandarización mediante el

cambio $t = \frac{x-a}{b-a}$ y $M = \frac{m-a}{b-a}$ donde, a es el valor mínimo, m el valor más probable y

b el valor máximo, siendo estos valores aportados por el experto.

Para el caso de la función de densidad se obtiene:

$$f(x) = \begin{cases} \frac{-2}{(m-a)^3(b-a)} \left(2x^3 - 3(m+a)x^2 + 6amx - (3m-a)a^2 \right) & \text{si } a \leq x \leq m \\ \frac{-2}{(m-b)^3(b-a)} \left(2x^3 - 3(m+b)x^2 + 6bmx - (3m-b)b^2 \right) & \text{si } m \leq x \leq b \end{cases} \quad (2.6)$$

Dicha función es derivable en la moda, es decir, $f'_-(m) = f'_+(m) = 0$ y además $f(x) \geq 0$

ya que $\left(2x^3 - 3(m+a)x^2 + 6amx - (3m-a)a^2 \right) \leq 0$ y $\left(2x^3 - 3(m+b)x^2 + 6bmx - (3m-b)b^2 \right) \leq 0$.

La función de distribución asociada a la distribución bicúbica viene descrita por la expresión:

$$F(x) = \begin{cases} \frac{-(x-a)^3(a+x-2m)}{(m-a)^3(b-a)} & \text{si } a \leq x \leq m \\ 1 - \frac{(x-b)^3(b+x-2m)}{(m-b)^3(b-a)} & \text{si } m \leq x \leq b \end{cases} \quad (2.7)$$

cumpliendo una de las características más importantes de las funciones pertenecientes al sistema generador de van Dorp, $F(m) = \frac{m-a}{b-a} = M$, siendo F la función de distribución y M la moda estandarizada.

1.2.1 Principales Características Estocásticas

Partiendo de la función de densidad descrita en (2.1), van Dorp y Kotz describen la relación existente entre los momentos de la función generadora $p(y)$ y la función de densidad $g(t)$:

$$E[t^k/M, p] = M^{k+1} E[x^k/\psi] + \sum_{i=0}^k \binom{k}{i} (-1)^i (1-M)^{i+1} E[x^i/\psi] \quad (2.8)$$

A partir de (2.5), podemos describir los momentos de la distribución bicúbica, que vienen dados por:

$$E[t] = \frac{4M+3}{10} \quad (2.9)$$

$$Var[t] = \frac{12M^2 - 12M + 13}{300} \quad (2.10)$$

Utilizando la expresión de la función de densidad de la variable x con dominio de definición $[a, b]$, se obtiene que los momentos de dicha distribución vienen descritos por:

$$E[x] = \frac{3a + 4m + 3b}{10} \quad (2.11)$$

$$Var[x] = \frac{13a^2 + 13b^2 - 12bm + 12m^2 - 2a(7b + 6m)}{300} \quad (2.12)$$

Teniendo en cuenta la relación existente entre los momentos centrados y no centrados podemos hallar las expresiones del coeficiente de asimetría y curtosis que vienen dadas por:

$$\beta_1 = \frac{108(8M^3 - 12M^2 - 34M + 19)^2}{49(12M^2 - 12M + 13)^3} \quad (2.13)$$

$$\beta_2 = \frac{27(88M^4 - 176M^3 + 312M^2 - 224M + 113)}{7(12M^2 - 12M + 13)^2} \quad (2.14)$$

1.2.2. Estudio De La Forma De La Distribución Bicúbica

En esta sección realizaremos un estudio sobre la asimetría y el apuntamiento de la distribución bicúbica. Para ello, utilizaremos el coeficiente de asimetría y curtosis introducidos en la sección anterior. Para tener una idea más clara de la forma que presenta la distribución bicúbica, a continuación se muestra las representaciones gráficas del coeficiente de asimetría, β_1 , y del coeficiente de curtosis, β_2 .

En la Figura (2.2) se puede comprobar que cuanto más se aproxima M a 0.5, en el caso de la variable estandarizada, el valor del coeficiente de asimetría tiende a cero, indicando pues, que la distribución es simétrica para ese valor. Hay que destacar que el máximo valor que toma el coeficiente de asimetría es 0.362, por tanto dicha distribución presenta poca asimetría.

Sin embargo, en el estudio de la curtosis, se observa en la Figura (2.3), que el coeficiente de curtosis de la distribución bicúbica estandarizada toma valores en el intervalo [2.410,2.579]. Teniendo en cuenta que el coeficiente de curtosis de la distribución normal es 3, entonces obtenemos una distribución que tiene menor apuntamiento que la distribución normal. Por tanto, dicha distribución es platicúrtica.

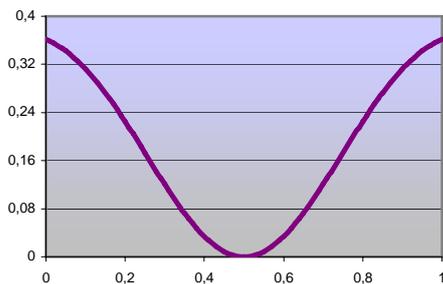


Figura 2.2: Representación del Coeficiente de Asimetría

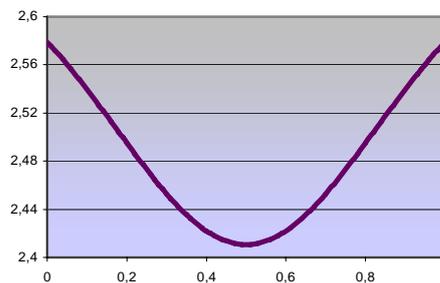


Figura 2.3: Representación del Coeficiente de Curtosis

1.3. CONSTRUCCIÓN DE LA DISTRIBUCIÓN BI-SENO

Para describir las expresiones de la función de densidad y distribución de la distribución bi-seno utilizaremos, al igual que para la distribución bicúbica, el sistema generador de van Dorp y Kotz (2203) dados por las expresiones (2.1) para el caso de la función de densidad y (2.4) para la función de distribución.

En primer lugar describiremos la expresión de la función de densidad generadora, $p(.|\psi)$. Teniendo en cuenta la Figura 3.1, la función de densidad asociada a la distribución seno es de la forma

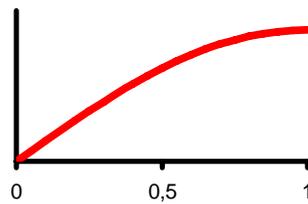


Figura 3.1: Representación de $p(y) = A \cdot \text{sen}(ay)$

luego, a partir de:

1. Pasa por el origen de coordenadas, $p(0)=0$
2. Presenta un extremo relativo en $y=1$; $p'(1) = 0$
3. Presenta un extremo relativo en $y=0$; $p'(0) = 0$
4. $\int_0^1 p(y)dy = 1$

se obtiene que la expresión de la función de densidad seno viene dada por:

$$p(y/\psi) = \frac{\pi}{2} \text{sen}\left(\frac{\pi}{2} y\right) \quad (3.1)$$

Por tanto, utilizando(2.1), obtenemos que la función de densidad estandarizada de la distribución bi-seno es:

$$g(t/M, p(y)) = \begin{cases} \frac{\pi}{2} \text{sen}\left(\frac{\pi}{2} \frac{t}{M}\right) & \text{si } 0 \leq t \leq M \\ \frac{\pi}{2} \text{sen}\left(\frac{\pi}{2} \frac{1-t}{1-M}\right) & \text{si } M \leq t \leq 1 \end{cases} \quad (3.2)$$

Realizando la desestandarización sobre $g(t/M, p(y))$ se obtiene la función de densidad definida a partir de la variable x , con dominio de definición $[a, b]$, que viene dada por:

$$f(x) = \begin{cases} \frac{\pi}{2} \frac{1}{b-a} \operatorname{sen}\left(\frac{\pi}{2} \frac{x-a}{m-a}\right) & \text{si } a \leq x \leq m \\ \frac{\pi}{2} \frac{1}{b-a} \operatorname{sen}\left(\frac{\pi}{2} \frac{x-b}{m-b}\right) & \text{si } m \leq x \leq b \end{cases} \quad (3.3)$$

donde a , m y b los valores aportados por el experto. Se comprueba fácilmente que dicha función es derivable en su valor modal, m .

A partir de (2.4) y sabiendo que la función de distribución asociada a la función de densidad generadora $p(y/\psi)$ viene descrita por $P(y/\psi) = 1 - \cos\left(\frac{\pi}{2}y\right)$, se obtiene la expresión de la función de distribución bi-seno estandarizada:

$$G(t/M, P(y)) = \begin{cases} M \left(1 - \cos\left(\frac{\pi}{2} \frac{t}{M}\right)\right) & \text{si } 0 \leq t \leq M \\ M + (1-M) \cos\left(\frac{\pi}{2} \frac{1-t}{1-M}\right) & \text{si } M \leq t \leq 1 \end{cases} \quad (3.4)$$

Y la función de distribución en x , con dominio $[a, b]$, sería:

$$F(x) = \begin{cases} \frac{m-a}{b-a} \left(1 - \cos\left(\frac{\pi}{2} \frac{x-a}{m-a}\right)\right) & \text{si } a \leq x \leq m \\ \frac{m-a}{b-a} + \frac{m-b}{b-a} \cos\left(\frac{\pi}{2} \frac{x-b}{m-b}\right) & \text{si } m \leq x \leq b \end{cases} \quad (3.5)$$

verificando que $F(m) = M$.

1.3.1 Principales Características Estocásticas

A partir de la expresión desarrollada por van Dorp y Kotz, (2.8), podemos hallar las expresiones de los momentos de la distribución bi-seno estandarizada, obteniendo que la esperanza matemática y la varianza viene dadas de la forma:

$$E[t] = \frac{\pi - 2 + (4 - \pi)M}{\pi} \quad (3.6)$$

$$\operatorname{Var}[t] = \frac{4(\pi - 3) - (M - 1)M(40 + (\pi - 16)\pi)}{\pi^2} \quad (3.7)$$

y para el caso de la distribución bi-seno en función de la variable x definida en el intervalo $[a, b]$:

$$E[x] = \frac{a(\pi - 2) - m(\pi - 4) + b(\pi - 2)}{\pi} \quad (3.8)$$

$$\text{Var}[x] = \frac{-4(3a^2 + 4ab + 3b^2) + 40(a+b)m - 40m^2 + 4(a+b-2m)^2 \pi - (m-a)(m-b)\pi^2}{\pi^2} \quad (3.9)$$

1.3.2 Estudio de la forma

Durante esta sección se realiza un estudio de la forma que presenta la distribución bi-seno, para ello utilizaremos tanto su coeficiente de asimetría como su coeficiente de curtosis. Haciendo uso de (2.8) y teniendo en cuenta la relación existente entre los momentos centrados y no centrados de una distribución, podemos hallar las expresiones del coeficiente de asimetría, β_1 , y del coeficiente de curtosis, β_2 , de la distribución bi-seno:

$$\beta_1 = -(2M - 1)^2 \frac{(A + BM - BM^2)^2}{12 - 4\pi - C(M - M^2)} \quad \text{donde} \quad \begin{cases} A = -16 + 24\pi - 6\pi^2 \\ B = 112 - 120\pi + 30\pi^2 - \pi^3 \\ C = 40 - 16\pi + \pi^2 \end{cases} \quad (3.10)$$

$$\beta_2 = \frac{A - BM + CM^2 - DM^3 + EM^4}{12 - 4\pi - F(M - FM^2)} \quad \text{donde} \quad \begin{cases} A = 528 - 96\pi - 48\pi^2 + 8\pi^3 \\ B = 2496 - 192\pi - 408\pi^2 + 72\pi^3 - \pi^4 \\ C = 4416 + 576\pi - 1320\pi^2 + 232\pi^3 - 4\pi^4 \\ D = 3840 + 1536\pi - 1824\pi^2 + 320\pi^3 - 6\pi^4 \\ E = 1920 + 768\pi - 912\pi^2 + 160\pi^3 - 3\pi^4 \\ F = 40 - 16\pi + \pi^2 \end{cases} \quad (3.11)$$

En la representación del coeficiente de asimetría, Figura 3.1, se observa que la distribución es simétrica cuando $M=0.5$, pero conforme nos alejamos de dicho valor el coeficiente va aumentando alcanzando como valor máximo, dentro del intervalo $[0,1]$, 0.179. Luego, dicha distribución presenta poca asimetría.

En la Figura (3.2), el recorrido del coeficiente de curtosis de la distribución bi-seno es $[2.193, 2.233]$, presenta una curtosis inferior a la distribución normal, 3.

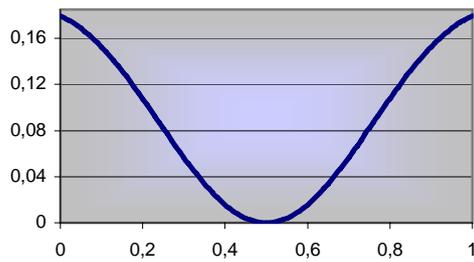


Figura 3.1: Coeficiente de Asimetría
distribución Bi-seno

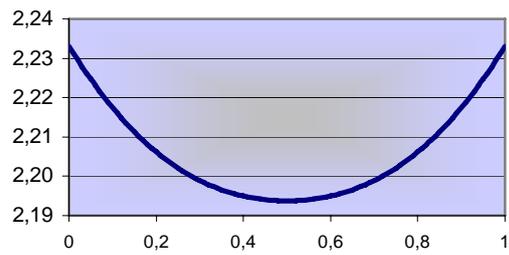


Figura 3.2: Coeficiente de Curtosis
distribución Bi-seno

1.4. LA DISTRIBUCIÓN BICÚBICA Y BI-SENO EN LA METODOLOGÍA PERT

Los principales criterios que deben cumplir los modelos probabilísticos para ser utilizados en análisis de inversiones en ambiente de riesgo, es que deben presentar una media 'moderada' y tener un valor de la varianza alto para maximizar el riesgo y no concluir con resultados finales optimistas. Por lo expuesto, si tenemos dos distribuciones que tienen la misma media, se elige la que tiene mayor varianza, pues refleja mayor incertidumbre y conduce, por tanto, a resultados más conservadores. Para ello, en esta sección estudiaremos el comportamiento de la esperanza matemática, la varianza y la altura modal de las principales distribuciones utilizadas en la metodología PERT.

Una vez obtenidos los valores pesimistas a , y optimista, b , aportados por el experto, los cuales determinan el rango donde fluctuará la variable x y m el valor modal, con $m \in [a, b]$ podemos describir las distribuciones.

Para realizar el estudio de la media, varianza y altura modal se estandariza el recorrido de la variable x al intervalo $[0,1]$ mediante el cambio de variable $t = \frac{x-a}{b-a}$ con valor

$$\text{modal } M = \frac{m-a}{b-a}.$$

En la tabla 4.1 se recoge los valores esperados y varianza de las distintas distribuciones utilizadas en la metodología PERT e incluyendo las distribuciones biparabólica, bicúbica y bi-seno. A partir de dicha información se representa los valores esperados, Figura 4.1, y las varianzas, Figura 4.2, de las distribuciones estandarizadas.

	Variable Aleatoria x	Variable Estandarizada t
Triangular	$E[x] = \frac{a+m+b}{3}$ $Var[x] = \frac{(b-a)^2 + (m-a)(m-b)}{18}$	$E[t] = \frac{M+1}{3}$ $Var[t] = \frac{1+M(M-1)}{18}$
Beta Clásica	$E[x] = \frac{a+4m+b}{6}$ $Var[x] = \frac{(b-a)^2}{36}$	$E[t] = \frac{4M+1}{6}$ $Var[t] = \frac{1}{36}$
Biparabólica	$E[x] = \frac{3a+2m+3b}{8}$ $Var[x] = \frac{12(m-a)^2 - 12(m-a)(m-b) + 19(b-a)^2}{320}$	$E[t] = \frac{2M+3}{8}$ $Var[t] = \frac{12M^2 - 12M + 19}{320}$
Bicúbica	$E[x] = \frac{3a+4m+3b}{10}$ $Var[x] = \frac{13a^2 + 13b^2 - 12bm + 12m^2 - 2a(7b+6m)}{300}$	$E[t] = \frac{4M+3}{10}$ $Var[t] = \frac{12M^2 - 12M + 13}{300}$
Bi-seno	$E[x] = \frac{a(\pi-2) - m(\pi-4) + b(\pi-2)}{\pi}$ $Var[x] = \frac{-4(3a^2 + 4ab + 3b^2) + 40(a+b)m - 40m^2}{\pi^2} + \frac{4(a+b-2m)^2 \pi - (m-a)(m-b)\pi^2}{\pi^2}$	$E[t] = \frac{\pi-2 + (4-\pi)M}{\pi}$ $Var[t] = \frac{4(\pi-3) - (M-1)M(40 + (\pi-16)\pi)}{\pi^2}$
Variable Estandarizada t		
Trapezoidal CPR	$E[t] = \begin{cases} \frac{1}{3} \left[M + \frac{3}{2} - \frac{1/2}{3/2 - M} \right] & si M < \frac{1}{2} \\ \frac{1}{3} \left[M + \frac{3}{2} - \frac{M}{1/2 - M} \right] & si M > \frac{1}{2} \end{cases}$ $Var[t] = \begin{cases} \frac{1}{18} \left[\frac{1}{4} + \frac{1}{2}M + (1-M)^2 - \frac{(1/2 - M)(1-M)}{(3/2 - M)^2} \right] & si M < \frac{1}{2} \\ \frac{1}{18} \left[\frac{1}{4} + \frac{1}{2}M + (1-M)^2 - \frac{(M - 1/2)M}{(M + 1/2)^2} \right] & si M > \frac{1}{2} \end{cases}$	

Tabla 4.1: Valores esperados y varianza de las distribuciones de la metodología PERT

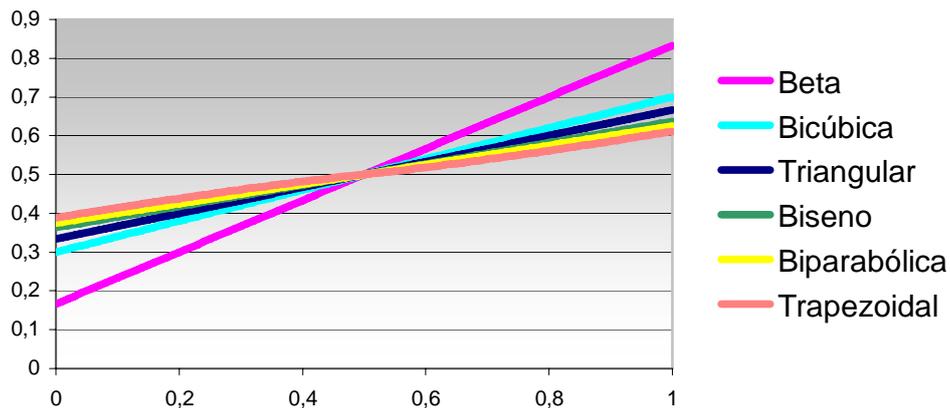


Figura 4.1: Comparación de los Valores Esperados

Se comprueba gráficamente que la distribución bicúbica y bi-seno tienen un valor esperado bastante moderado ya que el valor de sus medias está muy próximo al centro del intervalo, por tanto, proporcionan una esperanza 'centrada'.

Sin embargo, a la hora de estudiar la varianza de las distribuciones se puede observar que la distribución que presenta mayor varianza es la distribución bipolarabólica, pero hay que destacar que la distribución bi-seno tiene también un valor alto. Pero en el caso de la distribución bicúbica se comprueba que es una de las distribuciones que presenta menor varianza, estando únicamente la distribución beta por debajo de ella.

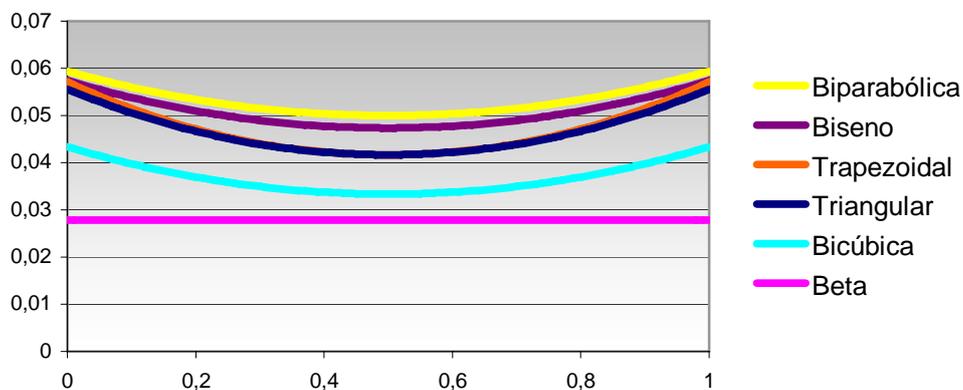


Figura 4.2: Comparación de la Varianza

Por último, estudiaremos las densidades de las distribuciones que se muestran en la Tabla 4.2 realizando su representación gráfica. En la Figura 4.3 se destaca que la distribución bicúbica alcanza la misma altura modal que la distribución triangular. Sin embargo, la distribución que tiene mayor ordenada es la distribución bi-seno.

Triangular	$h_T = \frac{2}{b-a}$	Bicúbica	$h_{bc} = \frac{2}{b-a}$
Uniforme	$h_U = \frac{1}{b-a}$	Bi-seno	$h_s = \frac{\pi}{2(b-a)}$
Biparabólica	$h_{bp} = \frac{3}{2(b-a)}$	Trapezoidal	$h_{CPR} = \begin{cases} \frac{2}{b-a-m+1/2} & \text{si } m < \frac{a+b}{2} \\ \frac{2}{b-a+m-1/2} & \text{si } m > \frac{a+b}{2} \end{cases}$

Tabla 4.2: Alturas Modales

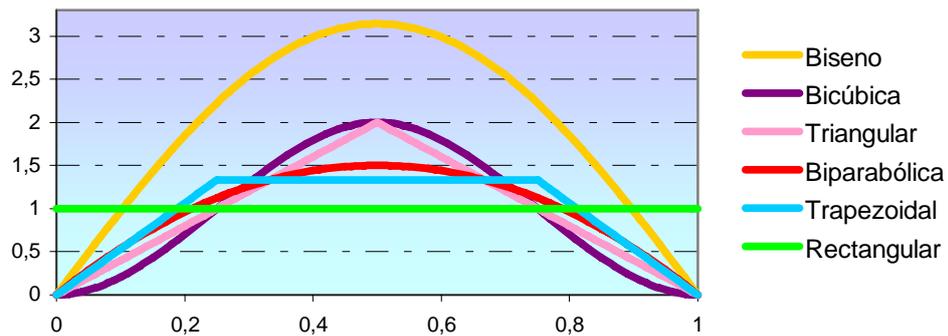


Figura 4.3: Comparación de densidades y alturas modales

1.5. GENERALIZACION DE LA DISTRIBUCION BICUBICA.

1.5.1. Aplicaciones Del Sistema Generador De Van Dorp Y Kotz Para La Generalización De La Distribución Bicúbica De Una Rama.

En esta sección se desarrolla la generalización de la distribución bicúbica mediante el esquema multiplicativo, estudiando sus principales características de forma y simetría. La generalización de la distribución bicúbica se obtiene introduciendo un nuevo parámetro n , permitiendo que la distribución presente una mayor flexibilidad.

Por tanto, la función de densidad generadora viene dada por:

$$p_1(y) = k \left(y^{2n} - \frac{2}{3} y^{3n} \right) \quad (5.1)$$

Sabiendo que, $\int_0^1 k \left(y^{2n} - \frac{2}{3} y^{3n} \right) dy = 1$, se obtiene que la función de densidad generadora viene dada por:

$$p_1(y) = \frac{3(2n+1)(3n+1)}{5n+1} \left(y^{2n} - \frac{2}{3} y^{3n} \right) \quad (5.2)$$

En el desarrollo de esta sección denotaremos a y como la variable generadora definida en el intervalo $[0,1]$, utilizada en la expresión (5.2). Para el caso de la distribución generada por el sistema generador de van Dorp y Kotz denotaremos a t como la variable generadora de dicho sistema con dominio $[0,1]$, definiendo así la distribución bicúbica generalizada estandarizada $SBCG_1$, y por último, denotaremos a la variable x a la variable desestandarizada cuyo dominio de definición es el intervalo $[a,b]$, dando lugar a la distribución BCG_1 .

A partir del sistema generador de van Dorp y Kotz, se describe la expresión de la función de densidad de $SBCG_1$:

$$g_1(t/M, p(. / n)) = \begin{cases} \frac{3(2n+1)(3n+1)}{5n+1} \left(\left(\frac{t}{M} \right)^{2n} - \frac{2}{3} \left(\frac{t}{M} \right)^{3n} \right) & \text{si } 0 \leq t \leq M \\ \frac{3(2n+1)(3n+1)}{5n+1} \left(\left(\frac{1-t}{1-M} \right)^{2n} - \frac{2}{3} \left(\frac{1-t}{1-M} \right)^{3n} \right) & \text{si } M \leq t \leq 1 \end{cases} \quad (5.3)$$

y la función de distribución de $SBCG_1$:

$$G_1(t/M, P(. / n)) = \begin{cases} \frac{3(2n+1)(3n+1)}{5n+1} M \left(\frac{1}{2n+1} \left(\frac{t}{M} \right)^{2n+1} - \frac{2}{3(3n+1)} \left(\frac{t}{M} \right)^{3n+1} \right) & \text{si } 0 \leq t \leq M \\ 1 + \frac{3(2n+1)(3n+1)(M-1)}{5n+1} \left(\frac{1}{2n+1} \left(\frac{1-t}{1-M} \right)^{2n+1} - \frac{2}{3(3n+1)} \left(\frac{1-t}{1-M} \right)^{3n+1} \right) & \text{si } M \leq t \leq 1 \end{cases} \quad (5.4)$$

Cuando $n=1$ en las expresiones (5.3) y (5.4) se obtienen la función de densidad (2.3) y de distribución (2.5).

Si realizamos la desestandarización de dichas funciones, realizando el cambio $t = \frac{x-a}{b-a}$ y

$M = \frac{m-a}{b-a}$, obtenemos las expresiones de la función de densidad y de distribución en función de la variable x , cuyo dominio de definición es $[a,b]$, de una distribución bicúbica generaliza de una rama BCG_1 :

Función de densidad de BCG_1

$$f_1(x) = \begin{cases} \frac{3(2n+1)(3n+1)}{5n+1} \left(\left(\frac{x-a}{m-a} \right)^{2n} - \frac{2}{3} \left(\frac{x-a}{m-a} \right)^{3n} \right) & \text{si } a \leq x \leq m \\ \frac{3(2n+1)(3n+1)}{5n+1} \left(\left(\frac{b-x}{b-m} \right)^{2n} - \frac{2}{3} \left(\frac{b-x}{b-m} \right)^{3n} \right) & \text{si } m \leq x \leq b \end{cases} \quad (5.5)$$

Función de distribución de BCG_1

$$F_1(x) = \begin{cases} \frac{3(2n+1)(3n+1)}{5n+1} \frac{m-a}{b-a} \left(\frac{1}{2n+1} \left(\frac{x-a}{m-a} \right)^{2n+1} - \frac{2}{3(3n+1)} \left(\frac{x-a}{m-a} \right)^{3n+1} \right) & \text{si } a \leq x \leq m \\ 1 + \frac{3(2n+1)(3n+1)}{5n+1} \left(\frac{m-a}{b-a} - 1 \right) \left(\frac{1}{2n+1} \left(\frac{b-x}{b-m} \right)^{2n+1} - \frac{2}{3(3n+1)} \left(\frac{b-x}{b-m} \right)^{3n+1} \right) & \text{si } m \leq x \leq b \end{cases} \quad (5.6)$$

Al igual que en el caso de las funciones estandarizadas, cuando sustituimos $n = 1$ en las funciones (5.5) y (5.6) obtenemos las expresiones (2.6) y (2.7).

1.5.2. Cálculo De Los Momentos

Para calcular los momentos de las variables t y x , debemos hallar primero los momentos de la variable generadora y . Para ello, utilizaremos la expresión de la función de densidad generadora

$$p_1(y) = \frac{3(2n+1)(3n+1)}{5n+1} \left(y^{2n} - \frac{2}{3} y^{3n} \right) \quad (5.7)$$

obteniendo que el momento de orden k , con $k \geq 1$, de la variable y , definida en el intervalo $[0,1]$, viene dada por la expresión

$$E[y^k] = \frac{(2n+1)(3n+1)(5n+k+1)}{(2n+k+1)(3n+k+1)(5n+1)} \quad (5.8)$$

Luego, utilizando la expresión (2.8) de van Dorp y Kotz (2003), podemos hallar la expresión del valor esperado y de la varianza de la distribución $SBCG_1$, obteniendo:

$$E[t/M, p] = \frac{30Mn^3 + (18M + 19)n^2 + 15n + 2}{30n^3 + 56n^2 + 30n + 4} \quad (5.9)$$

$$Var[t/M, p] = \frac{(289n^3 + 351n^2 + 128n + 12)(2n + 1)(3n + 1)}{12(n + 1)^2(2n + 3)(3n + 2)^2(5n + 1)^2} \quad (5.10)$$

Por tanto, el valor esperado de la distribución BCG_1 viene dado:

$$E[x] = \frac{a(19n^2 + 15n + 2) + (30n^3 + 18n^2)m + b(19n^2 + 15n + 2)}{30n^3 + 56n^2 + 30n + 4} \quad (5.11)$$

En (5.11), se observa que se pondera igual los extremos. Considerando $n = 1$ se obtiene que la ponderación de la moda es superior a la de los extremos

$$E[x] = \frac{3a + 4m + 3b}{10} \quad (5.12)$$

sin embargo, la ponderación de la moda es:

$$\frac{30n^3 + 18n^2}{30n^3 + 56n^2 + 30n + 4} \quad (5.13)$$

Se comprueba en la expresión (5.11), que en el valor esperado de la distribución bicúbica generalizada de una rama mediante el proceso multiplicativo, aparece el parámetro n ponderando los extremos y el valor más probable, al igual que ocurría con la distribución bipolarabólica, García (2007). Este hecho es destacable, ya que en el resto de distribuciones el parámetro solo pondera a m , como se observa en la Tabla 5.1.

Si hallamos el valor de n que pondera de igual forma a los extremos que al valor más probable se obtiene que, resolviendo la ecuación $19n^2 + 15n + 2 = 30n^3 + 18n^2$, $n = 0.781874$, pero como $n > 1$ entonces no podemos hallar un valor de n que pondere de igual forma a los extremos y a la moda. Hay que destacar que como $19n^2 + 15n + 2 < 30n^3 + 18n^2$ la distribución BCG_1 pondera más a la moda que a los extremos, $\forall n > 1$.

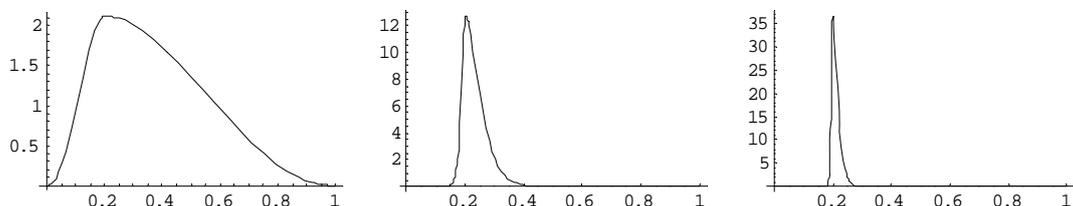
	Media
Beta Generalizada	$\mu = \frac{a + km + b}{k + 2}$
<i>TSP</i>	$\mu = \frac{a + (n-1)m + b}{n + 1}$
Pearson	$\mu = \frac{a + sm + b}{s + 2}; \quad b = \frac{1}{s}$
BPG_1	$\mu = \frac{a(7n + 2) + 6n^2m + b(7n + 2)}{6n^2 + 14n + 4}$
BCG_1	$\mu = \frac{a(19n^2 + 15n + 2) + (30n^3 + 18n^2)m + b(19n^2 + 15n + 2)}{30n^3 + 56n^2 + 30n + 4}$

Tabla 5.1: Valor Esperado de la variable aleatoria x
en distintas distribuciones subyacentes

Con el objeto de analizar el comportamiento de la distribución $SBCG_1$, se presenta en la Figura 5.1, la función de densidad de la distribución bicúbica generalizada para distintos valores de M y n . En dichas representaciones se observa que aumentando el tamaño de n , la función de densidad tiende a concentrarse en un punto.

Para valores de n más próximos a uno, y con $M \neq 0.5$ la función de densidad es asimétrica, sin embargo, cuando $M = 0.5$ la representación que adopta es casi simétrica, parecida a la función de densidad de una distribución normal. Por tanto, la distribución bicúbica reúne las condiciones para poder ser aplicable en el ámbito del PERT, ya que es una función de densidad asimétrica y simétrica dependiendo de los valores que consideremos de M y n y además el recorrido de la variable está acotado.

M=0.2



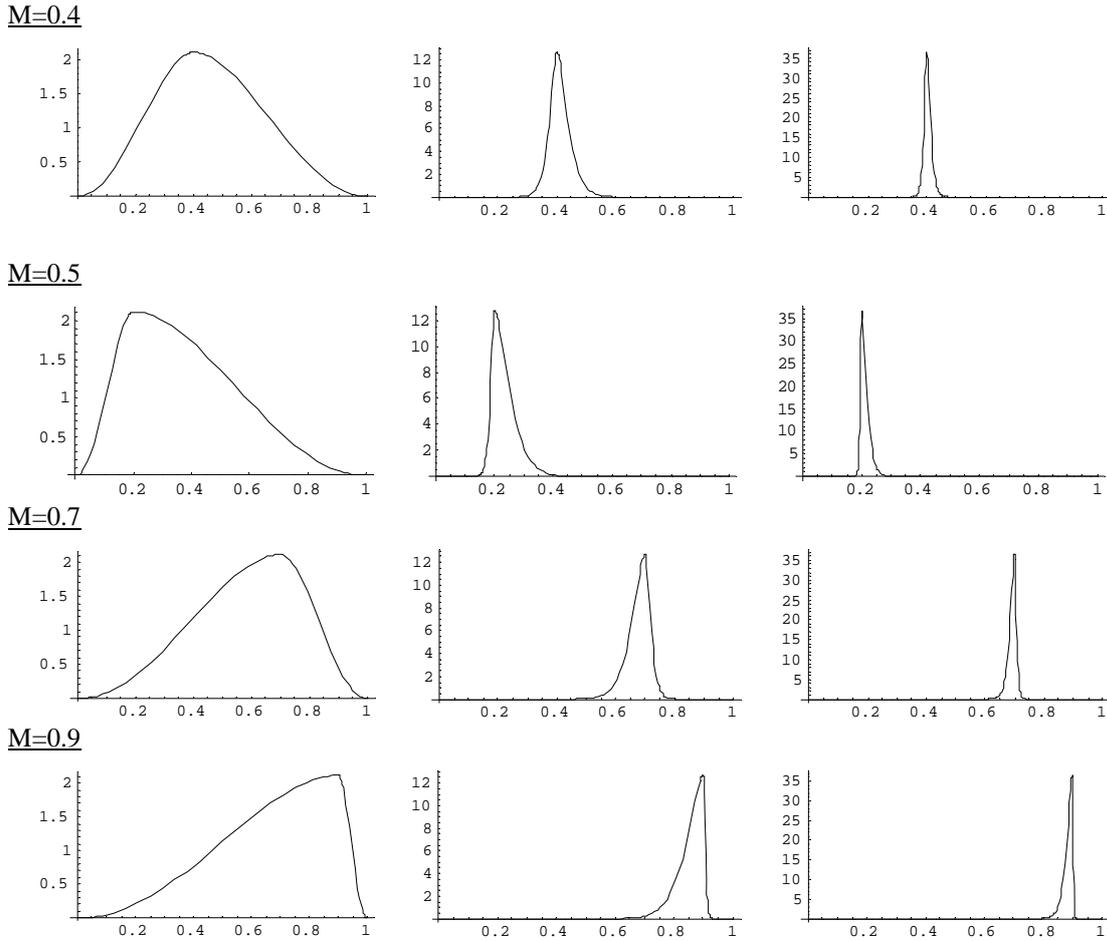


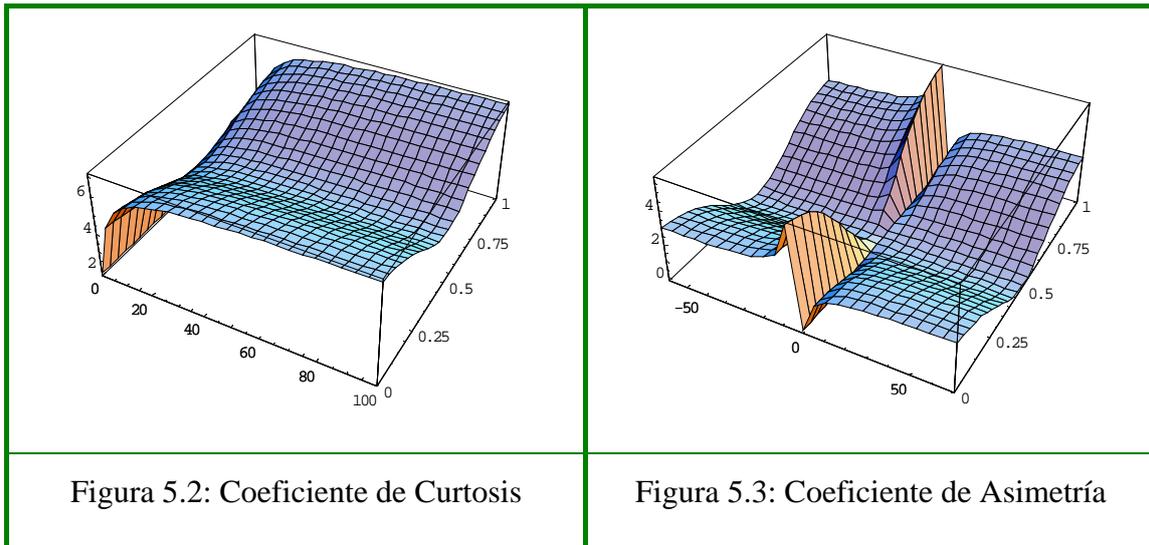
Figura 5.1: Función de densidad para distintos valores de M y $n = 1.1, n = 10, n = 30$

1.5.3. Estudio De La Asimetría Y Curtosis De La Distribución BCG_1

Haciendo uso de la relación de los momentos de la función generadora $p(y)$ y la función de densidad $g(t)$, desarrollados por van Dorp y Kotz (2003):

$$E\left[t^k/M, p\right] = M^{k+1}E\left[x^k/\psi\right] + \sum_{i=0}^k \binom{k}{i} (-1)^i (1-M)^{i+1} E\left[x^i/\psi\right] \quad (5.14)$$

podemos hallar los momentos centrados respecto a la media para obtener las expresiones del coeficiente de asimetría y de curtosis de la distribución bicúbica generalizada, cuyas representaciones gráficas vienen dadas en las Figuras 5.2 y 5.3



En la Figura 5.2 observamos que el coeficiente curtosis alcanza como máximo la distribución BCG_1 es de aproximadamente 6, presentando por tanto, un coeficiente de curtosis superior al de la distribución normal. Luego, podemos concluir que la distribución bicúbica generalizada de una rama es una distribución leptocúrtica.

En el caso del coeficiente de asimetría se observa en la Figura 5.3 que, cuando el tamaño de n va aumentando entonces el coeficiente se aproxima cada vez más a 4.

2. CONCLUSIÓN

En este documento se ha desarrollado dos distribuciones, bicúbica y bi-seno, como posibles alternativas para ser aplicadas en la metodología PERT.

Después de los estudios realizados sobre dichas distribuciones podemos concluir que dichas distribuciones son buenas candidatas para ser utilizadas en la metodología PERT, mejorando incluso algunas de las características de las distribuciones usadas habitualmente.

Bibliografía

- Ben Fair, A. *Upon implementing the beta distribution in Project management*. Department of industrial engineering and Management. Negev academic college of Engineering, Israel.
http://braude.ort.org.il/industrial/13thconf/html.%5Cfiles%2C113_p.pdf.
- Brown, S.L. (1999) An SAB report: *estimating uncertainties in Radiogenic Cancer Risk*, Washington, D.C.: Science and Advisory Board, United States Environmental Protection Agency, EPA-SAB-RAC-99-008.
- Callejón, J., Pérez, E. y Ramos, A. (1998). *La distribución trapezoidal como modelo probabilística para la metodología PERT*, X reunión ASEPELT-ESPAÑA. Universidad de Castilla La Mancha. CD-ROM-Fichero g26.
- Flehinger, B.J. y Kimmel, M. (1987). *The natural history of lung cancer in periodically screened population*. *Biométrica*, 43: 127-144.
- Carcía, C.B. (2007). Tesis Doctoral. *Generalizaciones de la distribución bipolar. Aplicaciones en el ámbito financiero y en el campo de la valoración*. Universidad de Granada, Febrero 2007.
- Garvey, P.R. (2000). *Probability methods for cost uncertainty analysis, a systems engineering perspective*. New York, NY: Marcel Dekker.
- Ghitany, M.E., Kotz, S. y Xie, M. (2005). On some reliability measures and their stochastic orderings for the Topp-Leone distribution. *Journal of applied Statistics*, vol.32, nº77, pp.715-722.
- Golenko-Ginzburg, D. (1998). *On the distribution of activity time in PERT*. *J.Opl. Soc.* Vol.39, nº8, pp.767-771.
- Herrerías, R. (1989). *Utilización de Modelos Probabilísticos Alternativos para el Método PERT. Aplicación al Análisis de Inversiones*. Estudios de Economía Aplicada. Secretariado de Publicaciones de la Universidad de Valladolid., pp. 89-112.
- Herrería, R. (1995). *Un nuevo uso de las tres estimaciones subjetivas del PERT*. IX Reunión Asepelt España. Vol. IV, pp. 411-416.
- Herrería, R. (2004). *La distribución $TSP(a,m,b,n)$ de van Dorp y Kotz como distribución de tipo Pearson univariante continua. Algunas aplicaciones estadísticas y económicas. Aspectos teóricos y aplicados en la generación de distribuciones de probabilidad*. Editorial: Entrelibros. EISBN: 84-931950-8-1.
- Johnson, D. (1997). *The triangular distribution as a Proxy for the beta distribution in risk analysis*. *Statistician*, 46, 387-398.
- Nadarafah, S. y Kotz, S.(2003). Moments of some J-shaped distributions. *Journal of Applied Statistics*, 30, pp.311-317.
- Pouliquen, L.Y. (1970). *Risk analysis in Project appraisal*. World Bank Staff Occasional Papers, 1. Baltimore, Md: John Hopkins University Press.
- Powell, M.R. y Wilson, J.D. (1997). *Risk assessment for national natural resource conservation programs*. Discussion Paper 97-149. Washington D.C.: Resources for the Future. pp. 370-418, Reproduced in *Biometrika*.
- Romero, C. (1991). *Técnicas de programación y control de proyectos*. Pirámide

Simpson, T. (1755) A letter to the Right Honourable George Earls of Macclesfield President of the Royal Society, on the advantage of taking the mean of a number of observations in practical astronomy. *Philosophical Transactions*, 49(1), pp.82-93.

Simpson, T. (1757). An attempt to show the advantage arising by taking the mean of a number of observations in practical astronomy. *Miscellaneous Tracts on some curious and very interesting Subjects in Mechanics, Physical Astronomy and Speculative Mathematics*, pp.64-75

Suárez, S. (1980). *Decisiones óptimas de inversión y financiación en la empresa*. Pirámide.

Taha, H.A. (1981) *Investigación de operaciones*. Representaciones y Servicios de Ingeniería S.A. México.

Topp, C.W. y Leone, F.C. (1955). A family of J-Shaped frequency functions, *Journal of the American Statistical Association*, 50 (269), pp. 209-219.

Van Dorp, J.R. y Kotz, S. (2002). *A Novel Extension of the Triangular Distribution and its Parameter Estimation*, *The Statistician*, 51, N° 1, pp. 63-79.

Van Dorp, J.R. y Kotz, S. (2003). *Generalizations of Two-Sided Power Distributions and their Convolution*. *Communications in Statistics: Theory and Methods*, Vol. 32, N° 9, pp. 1703-1723.

Vose, D. (1996). *Quantitative Risk Analysis: A guide to Monte Carlo simulation Modelling*. Wiley, New York.

Winston, W.L. (1993). *Operations Research, Applications and Algorithms*. Pacific Grove, CA: Duxbury Press.

ECONOMIC GROWTH AND EMISSION REDUCTION: A COMPARATIVE ANALYSIS WITH AND WITHOUT EMISSION RIGHTS BANKING BETWEEN PERIODS

MARÍA JESÚS JIMÉNEZ HERRERO
e-mail: chusjh@eco.uva.es

FRANCISCO JOSÉ CABO GARCÍA
e-mail: pcabo@eco.uva.es

GUIOMAR MARTÍN HERRÁN
e-mail: guiomar@eco.uva.es

Departamento de Economía Aplicada, Sección Matemáticas
UNIVERSIDAD DE VALLADOLID

Área temática: Métodos Cuantitativos.

Resumen

En este trabajo se propone un modelo de crecimiento económico endógeno con comercio de derechos de emisión. Una empresa localizada en un país firmante del Protocolo de Kyoto tiene que determinar la política óptima de inversión en capital físico y tecnológico. Ambos tipos de capital son esenciales para la producción, mientras que el capital tecnológico también reduce el ratio de emisiones por unidad de output. El país utiliza el comercio de derechos de emisión para cumplir con sus compromisos de reducción de emisiones con un menor coste en términos de crecimiento económico. En el mercado internacional de derechos de emisión bajo el auspicio del Protocolo de Kyoto, la transferencia de derechos de emisión, banking, de un periodo de compromiso a otro no está permitido, por lo que el valor de los derechos de emisión al final de cada periodo es nulo. El objetivo principal de este artículo es analizar cuál es el efecto sobre las principales variables económicas de la empresa de permitir el banking entre periodos. Para ese fin, el modelo planteado se resuelve mediante un algoritmo numérico que nos permite obtener el comportamiento óptimo bajo dos escenarios: con y sin banking entre periodos. Estos dos escenarios se comparan tanto en términos de crecimiento económico como en logros de reducción de emisiones. A su vez se lleva a cabo un análisis de sensibilidad de los diferentes parámetros del modelo bajo ambos escenarios para analizar la robustez de los resultados.

Palabras clave: Comercio de derechos de emisión, Control de la contaminación, Control óptimo, Acumulación de derechos.

Abstract

We set out a model of endogenous economic growth with emission trading. A firm located in a signatory country of the Kyoto Protocol has to determine optimal investment policies for physical and abatement capital. Both types of capital are essential for production, while the latter also reduces the ratio of emissions per unit of output. The country has access to emission rights trading so as to reach growth attaining its emissions reduction compromise. In the international emission trading market proposed by the Kyoto Protocol, the banking of emission rights from one period to another is not allowed, so the value emission rights at the end of each period is null. The main purpose of this paper is to analyze the effects of allowing emission rights banking between periods. To that end we develop a numerical algorithm to compute the optimal behaviour of the firms under two scenarios: with and without emission rights banking between periods. We compare these two scenarios in terms of economic growth and environmental achievements. A sensitivity analysis to variations of the different parameters of the model is carried out in both scenarios.

Key words: Emission Trading, Pollution Control, Optimal Control, Banking.

Financial support by MEC under project SEJ2005-03858/ECON and by JCYL under projects VA033B06 (second author) and VA045A06 (third author) is acknowledged. All projects are co-financed by FEDER funds.

1. Introduction

The environment has played an important role in the economy as a supplier of productive inputs which are essential to obtain a higher economic growth. However, the use of the environment as a store of industrial waste has carried out some environmental problems, being the Climate Change one of the most important. This problem not only has environmental implications, but also economic consequences, especially significant for the economic growth.

The Kyoto Protocol, adopted in 1997 and entered into force in 2005, constitutes one of the most important steps towards preventing Global Warming, which struggle started some years ago by the United Nations Climate Change Convention adopted in 1992 at the Rio Earth Summit. According to this agreement, the signatory developed countries have to reduce their aggregate emissions by at least a 5% below those in 1990 (that percentage is not the same for all the signatory countries). Because the reduction efforts increase production costs¹, the Protocol defines three flexible mechanisms to cut down these costs: Joint Implementation, Clean Development Mechanism and Emission Trading.

Countries like USA and Canada have implemented their own emission trading systems. However, the European Emission Trading Scheme (EETS), created by Directive 2003/87/EC for the UE members and started in 2005, is the prime candidate to become the emission trading regime internationally accepted, since it encompasses emission rights buyers' and sellers'.

The basic idea of this new mechanism was given by Coase (1960), who set the idea of assigning property rights to the environment. Nevertheless, it was applied to environmental problems in Dales (1968) and Montgomery (1972).

According to the EETS, the transfer of unused allowances from the first period, 2005-2007, to the Kyoto Protocol commitment, 2008-2012, is banned. Nevertheless, banking is allowed within each compliance period. This Scheme has forbidden banking by means of limiting the life of allowances to the period for which they are issued.

¹ Memo/04/44.

Our paper studies a dynamic economic growth model with emission trading taking into account the fact that banking between compliance periods may or may not be banned. Among the authors who analyze the relationship between the emission rights trading and the economic growth, Stockey (1998) and Grimaud (1999) can be cited. These authors analyze how emission trading, as well as other control instruments, affects growth. In this kind of literature, which studied economic growth and emission trading, we find authors like Ono (2002), who in an overlapping generations model of growth, considers a government that controls emissions by assigning emission quotas and permits that firms can freely trade. This author finds that although the decrement in quotas reduces current emissions, it could result in environmental deterioration in the long run.

There is a huge theoretical literature about the economic efficiency of allowing banking. Many of the works show that the use of banking reduces overall compliance costs by allowing for inter-temporal flexibility, (see, for example, Boemare & Quirion (2002) or Ellerman et al. (2003)). Haites (2006) sets out that without banking the firms make their abatement investments “just in time” to reduce emissions during the compliance period, the rate of non-compliance period is likely to be higher and the price rises sharply or falls to zero at the end of each compliance period.

An empirical analysis is carried out by Schleich et al. (2006). These authors analyze the implications of a ban on banking in the EETS when initial emissions targets are lenient by means of a simulation carried out in Germany. They show that a ban on banking would lead to efficiency losses in addition to those losses which arise from the lack of inter-temporal flexibility.

Dynamic approaches to model the emission rights trading and banking can be found in Rubin (1996) and Cronshaw & Kruse (1996). Both papers analyze that under certain conditions, like absence of uncertainty or competitive behaviour, an emission trading mechanism with banking and borrowing, achieves the emissions reduction at a lower cost. Stevens & Rose (2002) constitutes an extension to these studies, incorporating an explicit restriction on emission trading. They observe that the decentralized behaviour of the firm leads to the least cost solution.

In this paper we analyze, by means of an endogenous economic growth model with emission trading, the optimal decisions of a firm that belongs to an annex B country of the Kyoto Protocol. This company has to determine the optimal trade-off between investment in physical capital, which fuels production and hence emissions, and investment in technological capital, which enhances final output but reduces the emissions per unit of output. Furthermore, this company has to determine the timing for buying or selling additional emission rights in the emission permits market, in order to reduce the cost of compliance.

As we previously stated, the transfer of unused allowances from one compliance period to another is banned. However, we add this possibility in our analysis, solving the model under two alternative scenarios, with and without banking between periods. The optimal time paths for investment in physical and technological capital, as well as for the amount of emission rights traded in the market are obtained. The comparison of the results under both scenarios illustrates the effect of allowing banking from one compliance period to another on the optimal capital investment and emissions. A sensitivity analysis with respect to different model's parameters is carried out for two scenarios that depend on whether transfer of unused allowances from one compliance period to another is or is not banned.

The paper is organized as follows. Section 2 presents the model of economic growth with emission trading. In section 3 we develop a numerical algorithm to solve the model under two scenarios: with and without banking between compliance periods. In section 4 a sensitivity analysis of the model outcomes has been carried out for each scenario. Finally the conclusions are showed in section 5.

2. The model

We consider a firm located in a signatory country of the annex B part of the Kyoto Protocol, and as such it is required to reduce its CO_2 emissions. Two types of inputs are involved in the productive process of this company: physical capital, K , and technological capital, h . The latter is also used as abatement capital to improve the efficiency of physical capital in terms of generating lower emissions. In other words, an increment in technological capital reduces the emissions per unit of output (see, for example, Van der Ploeg & Zeeuw (1992) or Jorgensen & Zaccour (2001)). We assume a

Cobb-Douglas production function in these two types of capital $Y(K(t), h(t)) = AK(t)^\alpha h(t)^{1-\alpha}$, $A > 0$, where $\alpha \in (0,1)$.

The accumulation through time of both capital stocks depends on the investments carried out by the firm and the depreciation rates. Thus, the temporal evolutions of these capital stocks are described by:

$$\dot{K}(t) = I^k(t) - \delta_k K(t), \quad K(0) = K_0 > 0, \quad (1)$$

$$\dot{h}(t) = I^h(t) - \delta_h h(t), \quad h(0) = h_0 > 0, \quad (2)$$

where, $I^k(t)$, $I^h(t)$ denote investment in physical and technological capital, respectively, while constants δ_k and δ_h are the positive depreciation rates.

Quadratic costs are associated with the investment in both types of capital. These investment cost functions are defined as $C_j(I^j) = c_j(I^j)^2/2$, $c_j > 0$, $\forall j = K, h$.

The productive process of the final good is associated with an undesirable by-product: the emission of pollutants. Emissions are considered proportional to the final output. However, the proportionality factor is not an exogenous constant, but a downward sloping and convex curve of the technological (and abatement) capital. Thus, emissions are defined as $e(t) = f(K(t), h(t)) = h(t)^{-\beta} y(t) = AK(t)^\alpha h(t)^{1-\alpha-\beta}$, where $\beta > 1-\alpha$.

In order to reduce its emissions and achieve the compromise assumed with the signature of the Kyoto Protocol, the firm has two possibilities. On one hand, it can invest in abatement capital to reduce emissions per unit of output. On the other hand, the firm can participate in the international emission trading so as to reduce the cost of Kyoto.

Taking into account a *Cap and trade system*, the signatory country fixes an emission quota for the company at a level S . This level is constant through the planning period, as UE members have fixed for the pre-Kyoto period, 2005-2007. Moreover, it is assumed a *Grandfathering* allocation system so the firms receive emission permits free of charge according with their past emissions.

Taking into account this mechanism, a company that emits below its quota generates a surplus of emission rights². This surplus can be either accumulated for the future, or

² According to the law 1/2005, which regulates the Spanish emission trading system, an emission right is defined as “the subjective right to emit one equivalent tone of CO₂ during a certain time”.

traded to other companies that demand additional emission rights. The purchase (sale) of emission rights in this market is represented by a positive (negative) value of the variable $z(t)$. We assume that the amount of emission rights that can be purchased (sold) in this market is bounded by $z_{\max} > 0$ (resp. $z_{\min} < 0$). We consider that the price of the emission rights, $p(t)$, is exogenous to the model. Its time path is described by an affine and upward sloping function, $p(t) = p_0 + p_1 t$, $p_0, p_1 > 0$. This function represents the company's expectatives about the evolution of the price of these rights. Although for the European Emission Trading Scheme the price of the emission rights has experienced a decreasing path as far as we are approaching the end of the experimental period, 2005-2007, the hypothesis of a growing price is supported by several studies like Bernard & Vielle (2002) or Soberg (2000), between others. Statistical analyses also support this hypothesis, (see, for example, the actual evolution of emission rights prices to be traded in the future international market by CIMD Agencia de Valores).

Following Rubin (1996), emission rights can be accumulated leading to the dynamic equation:

$$\dot{B}(t) = S - e(t) + z(t), \quad B(0) = 0, B(t) \geq 0, B(T) \geq 0, \quad (3)$$

where $B(t) \geq 0$ implies that firms can accumulate, but they are not allowed to borrow emission permits when an ETM exists.

According to this equation, the accumulated amount of emission rights at time t , $B(t)$, depends on the emission quota imposed by the government of the signatory country, its actual emissions, $e(t)$, and the emission rights traded in this market. In the current regulations of ET in the UE as well as in the Marrakech accords³ unused allowances at the end of the current compromise period of five years cannot be transferred into the following period, in other words, the possibility of transfers between compromise periods is not allowed. Since the stock of the accumulated emission rights cannot be negative, $B(t) \geq 0$, the optimal management of this stock is linked to a null stock at the end of the planning period, $B(T) = 0$. The main question of this paper is how this

³ The Marrakech accords are a set of agreements which gives all the details about the Kyoto Protocol application. In particular, we point out the decision 18/CP.7.

constraint of no emission rights transfers between periods may affect the optimal decisions of the company and thus, the accumulation of both capital stocks and the emissions of pollutants through time. Thus, we analyze and compare two scenarios depending on whether this restriction is fulfilled.

The company seeks to maximize the net benefit through the compromise period. At any time, t , this net benefit can be written as:

$$\Pi(t) = Y(K(t), h(t)) - C_K(I^K(t)) - C_h(I^h(t)) - z(t)p(t).$$

This net benefit is defined as the production income, considering a unitary price for the production good, minus the cost of investment in both types of capital. The last term represents the cost of the emission rights traded by the firm, if the company is a net buyer of emission rights, or a revenue, if a net seller.

Summarizing, the economic growth and emission rights trading problem that the firm has to solve, can be written as:

$$\underset{z, I^K, I^h}{\text{Max}} \int_0^T \Pi(t) dt + g_K K(T)^\alpha + g_h h(T)^{1-\alpha} + g_B \ln(1+B(T)), \quad (4)$$

$$\text{s.t: } (1), (2), (3), z_{\min} \leq z(t) \leq z_{\max}.$$

Along the planning period, $[0, T]$, the firm determines its investments in physical and technological capitals, $I^K(t)$ and $I^h(t)$, as well as the amount of emission rights to buy or sell, $z(t)$, at any time. The last three terms in the objective functional are the salvage functions, which represent the value that the entrepreneur gives to the physical and technological capital stocks and to the possession of emission rights at the end of the planning period, respectively. Constants g_K , g_h and g_B are the subjective value that the company gives to the possession of each type of capital and emission rights at the end of the compromise period. The value of last parameter, g_B , depends on whether banking between periods is or is not allowed. When banking between periods is banned, then $g_B = 0$. By contrast, when banking is permitted, $g_B \geq 0$ and its value represents the expectations about how restrictive the next compromise period will be: the more restrictive the firm's expectation about the tightness of next compromise period, the higher the value of g_B .

The Hamiltonian function for the company reads⁴:

$$H(z, I^K, I^h, K, h, B, \lambda_K, \lambda_h, \lambda_B) = AK^\alpha h^{1-\alpha} - \frac{c_K}{2}(I^K)^2 - \frac{c_h}{2}(I^h)^2 - z(p_0 + p_1 t) + \lambda_K(I^K - \delta_K K) + \lambda_h(I^h - \delta_h h) + \lambda_B(S - AK^\alpha h^{1-\alpha-\beta} + z),$$

where λ_K, λ_h and λ_B are the costate variables associated with physical capital, technological capital and accumulated emission rights. These shadow prices represent the marginal value of additional units of the corresponding stock.

In this problem we have non-negativity constraints in the accumulated emission rights and in the number of rights traded in the ETM. For constraint $B(t) \geq 0$, let us note that at any time when $B(t) > 0$ the constraint is not binding and it can be ignored. Nevertheless, when $B(t) = 0$ we have to be sure that the amount of accumulated emission rights does not decrease further. Thus, we use $\dot{B}(t) \geq 0$ as a constraint so that B does not become negative. When $B = 0$ we associate the multiplier η to restriction $\dot{B}(t) \geq 0$. To make things simple, we impose an “either or” condition, $\eta B = 0$, so the multiplier take a null value whenever $B > 0$. Therefore, the Lagrangian function can be written as:

$$L(z, I^K, I^h, K, h, B, \lambda_K, \lambda_h, \lambda_B) = H + \eta \dot{B} + \mu_1(z - z_{\min}) + \mu_2(z_{\max} - z).$$

Assuming interior solutions, to obtain the first order conditions for both models, we equal to zero the partial derivatives of the Hamiltonian with respect the control variables. The number of emission rights traded in the ETM is a bang-bang function:

$$z = \begin{cases} z_{\max} & -(p_0 + p_1 t) + \lambda_B > 0, \\ \hat{z} \in [z_{\min}, z_{\max}] & -(p_0 + p_1 t) + \lambda_B = 0, \\ z_{\min} & -(p_0 + p_1 t) + \lambda_B < 0, \end{cases} \quad (5)$$

and the investment in each capital stock reads:

$$I^K = \frac{\lambda_K}{c_K} \quad ; \quad I^h = \frac{\lambda_h}{c_h}. \quad (6)$$

Taking partial derivatives of the Lagrangian with respect to the costate variables. the first order conditions and the transversality conditions are:

$$\dot{\lambda}_K = \delta_K \lambda_K - A\alpha K^{\alpha-1} h^{1-\alpha} + (\lambda_B + \eta) A\alpha K^{\alpha-1} h^{1-\alpha-\beta}, \quad \lambda_K(T) = g_K \alpha K(T)^{\alpha-1}, \quad (7)$$

⁴ Henceforth we omit time arguments when no confusion is caused by doing so.

$$\dot{\lambda}_h = \delta_h \lambda_h - A\rho K^\alpha h^{-\alpha} + (\lambda_b + \eta)A(\rho - \beta)K^\alpha h^{-\beta-\alpha}, \quad \lambda_h(T) = g_h(1 - \alpha)h(T)^{-\alpha}, \quad (8)$$

$$\dot{\lambda}_b = 0, \quad \lambda_b(T) \geq \frac{g_b}{1 + B(T)}, \quad \left(\lambda_b(T) = \frac{g_b}{1 + B(T)} \text{ if } B(T) > 0 \right), \quad (9)$$

$$\eta = p - \lambda_b - \mu_1 + \mu_2, \quad (10)$$

together with the complementary slackness conditions:

$$\mu_1 \geq 0, \quad \mu_1(z - z_{\min}) = 0,$$

$$\mu_2 \geq 0, \quad \mu_2(z_{\max} - z) = 0,$$

$$\eta \geq 0, \quad \eta B = 0, \quad \eta \dot{B} = 0. \quad (11)$$

The optimal amount of emission rights traded in the ETM, z , is given by the bang-bang solution in (5). Its value is determined by the price of the emission rights and the marginal value of the accumulated emission rights. The optimal behaviour for this company is to choose the maximum (minimum) value of z when $p < \lambda_b$ ($p > \lambda_b$). The firm purchases (sells) emission rights at the maximum rate when the price of the emission rights is below (above) the marginal value of B . From equation (9) follows that λ_b is constant. Thus, the higher this value, the longer the firm is buying emission rights and hence, the larger this firm is accumulating emission permits. Within this context we distinguish two scenarios depending on whether banking between periods is or is not allowed. The optimal solution of z is different under both scenarios.

From condition (11) it follows that the optimal number of emission rights traded in the ETM is as given in expression (5) as long as the amount of accumulated emission rights is positive. However, when there are no emission rights at stock, $B(t) = 0$, this stock cannot decrease further. Therefore, the minimum value of z is no longer, $z_{\min} < 0$, but a new value $\hat{z} = e - S$, which makes $\dot{B} = 0$, i.e. B remains motionless, no emission rights are accumulated.

First order conditions (6) show that the optimal investments in physical and technological capital are linear functions of the shadow prices of their respective stocks.

From equation (9) the costate variable λ_b is a constant, which value has to fulfil the transversality condition. That condition states that $\lambda_b \geq \frac{g_b}{1 + B(T)}$ (= if $B(T) > 0$).

Therefore, when banking between periods is not allowed it is not optimal for the firm to maintain a positive stock of permits at the end of the compromise period. Thus, $B(T)=0$ and the transversality condition reads $\lambda_B \geq g_B$. Conversely, when banking is permitted the firm could be interested in a positive stock of permits at the end of the planning horizon. In that case, with $B(T)>0$, the transversality condition implies $\lambda_B = \frac{g_B}{1+B(T)} < g_B$.

We have to solve a two-point boundary value problem with six non-linear differential equations, (1), (2), (3), (7), (8) and (9). No analytical solution has been found for this complex problem. Thus, numerical integration has been carried out as it is explained in the following section.

3. Numerical characterization of optimal time paths

In this subsection we develop a numerical algorithm in order to solve the dynamic model of economic growth and emission trading. The optimal time paths of the control, state, and costate variables are obtained under both scenarios, with and without banking between periods. This is a two-point boundary value problem with initial values for the state variables and final values for the accumulated amount of emission rights as well as for the costate variables (given by the transversality conditions). The `bvp4c` routine in MATLAB is a well known tool to solve this type of problems.

We have considered the following parameter's values:

Table 1: *Initial parameter's values*

$T = 2$	$A = 0.3$	$S = e(0) 0.66$	$\alpha = 0.6$	$\beta = 0.6$	$K_0 = 2.2$	$h_0 = 0.9$	$\delta_K = 0.1$	$\delta_h = 0.1$
$c_K = 1$	$c_h = 1$	$p_0 = 0.001$	$p_1 = 0.03$	$g_K = 0.5$	$g_h = 0.5$	$z_{\max} = 1$	$z_{\min} = -1$	

These values fit a firm with an initial physical capital stock, $K_0 = 2.2$, above the initial technological capital stock, $h_0 = 0.9$. The response of emissions to increments in technological capital, $\beta = 0.6$, is higher than the output elasticity of this capital, $1 - \alpha = 0.4$. Thus, the direct reduction in emission due to a higher technological capital is stronger than the increment in emissions associated with the increment in output

produced by this higher h . Initial emissions, $e(0) = AK_0^\alpha h_0^{1-\alpha-\beta} = 0.4917$, are above its quota, $S = e(0)0.66 = 0.3243$. The firm has to make an effort in pollution cutting investment and or it must buy emission rights in the ETM. The depreciation rates, $\delta_k = \delta_h$, the investment costs, $c_k/2 = c_h/2$, and the value given by the entrepreneur at the end of the time period, $g_k = g_h$, are identical for both types of capital. We finally assume an entrepreneur that expects an increment in the prices of the emission rights through time. For simplicity a constant growth rate, p_1 , is considered.

Under the first scenario, where banking between periods is banned, the company does not value the possession of emission rights at the end of the compromised period, so $g_B = 0$. According to equation (9) $\dot{\lambda}_B = 0$, and thus, the optimal path of λ_B is a constant through the whole time horizon. To determine this constant, we integrate the system taking the shadow price $\lambda_B(t)$ as a constant. For each constant λ_B we defined the value function as $V(\lambda_B) = \sum_0^T \Pi(t) + g_k K(T)^\alpha + g_h h(T)^{1-\alpha}$. With MATLAB we obtain $V(\lambda_B)$ for each value of λ_B within interval $[0, \bar{\lambda}_B]$, where $\bar{\lambda}_B$ is the maximum value of λ_B for which the accumulated amount of emission rights is null at the end of the compromise period, $B(T) = 0$. This value is defined as the marginal value of emission rights when banking between periods is banned, and it is constant through the planning period. The numerical integration shows that $V(\lambda_B)$ is a monotonous growing function. Therefore, the maximum value is obtained for the maximum $\lambda_B = \bar{\lambda}_B$. Since banking between period is not allowed, the optimal value of λ_B is that for which the optimal path of B grows at a maximum rate within a first subperiod and decreases at the fastest rate within a second subinterval, reaching zero only at the end of the compromise period $\bar{\lambda}_B = 0.03575$, which complies with the transversality condition. Once the optimal constant value of λ_B has been computed, the two-point boundary value problem can be integrated using the `bvp4c` routine in MATLAB.

In a second scenario, we consider that banking between compromise periods is allowed. Under this assumption there are two possibilities depending on how the firm values the possession of emission rights at the end of the planning period. If the company expects a soft compromise after Kyoto, i.e. if the company values the possession of unused

allowances to transfer to the next period below or above a given bound, given by $\bar{\lambda}_B$. If the value that the company gives to emission rights at time T is not high enough, $g_B < \bar{\lambda}_B$, the firm behaviour is equal to the case in which banking is banned. The firm accumulates rights within a first subperiod and reduces the stock of emission permits henceforth up until the stock B vanishes at the final time, T . Conversely, if the company values the possession of unused allowances above $\bar{\lambda}_B$ the company does not deplete the stock of rights completely, but it saves some emission rights for the next compromise period, $B(T) > 0$. Transversality condition in (9) now reads $\lambda_B(T) = \frac{g_B}{1+B(T)}$, where $B(T)$ is the final value of the optimal trajectory of B . In this situation, the constant optimal value of λ_B is directly obtained from the numerical integration through the `bvp4c` routine in MATLAB. For $g_B = 0.08$, the optimal value of λ_B is 0.046534.

In this second scenario, regardless of whether g_B is greater or lower than $\bar{\lambda}_B$ for these parameter's values, the initial investment in physical capital is larger than the investment in technological capital, decreasing the investment in both capitals through time. When the firm values the possession of emission rights above $\bar{\lambda}_B$, the higher this salvage value, g_B , the lower the growth in the physical capital stock (see Figure 1)⁵. Conversely, technological capital displays the opposite behaviour (see Figure 2). When banking is not banned and the company sufficiently values the possession of a surplus of emission rights at the end of the planning period, the firm increases its effort in abatement capital and invests less in physical capital to reduce emissions and attain a surplus in emission rights. Thus, when the company values the possession of emission rights below the given bound, $g_B \leq \bar{\lambda}_B$, production is higher (see Figure 3), and so is the value function. In our numerical integration for $g_B = 0.032 < \bar{\lambda}_B$ the value function is $V = 44.25905$, while for $g_B = 0.08 > \bar{\lambda}_B$ the value function is $V = 42.89655$.

⁵ All figures present a double scales for the two plots depicted on each of them.

Figure 1

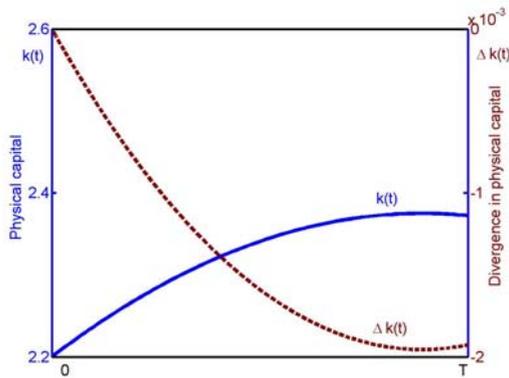


Figure 2

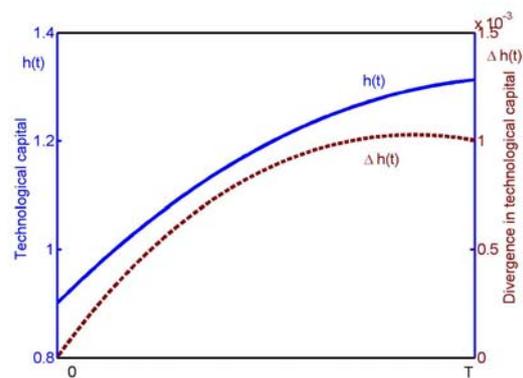
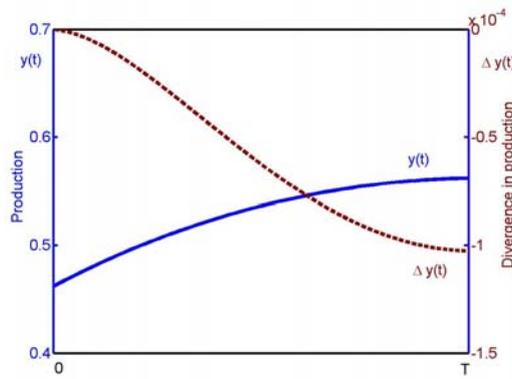


Figure 3



With respect to the emissions, under both scenarios they are decreasing through time, although they are above the emission quota, S . Furthermore, emissions are higher for $g_b \leq \bar{\lambda}_b$ than for $g_b > \bar{\lambda}_b$ (see Figure 4). Initially, with initial emissions, e_0 , above its quota, S , and with no initial emission rights, $B(0)=0$, the firm must buy permits to achieve its quota. Furthermore, with a growing expected price of the emission rights, the firm purchasing effort will be stronger at the beginning, buying at the maximum level, $z=1$. When the price of the emission rights reaches the marginal value of the emission rights, λ_b , the firm's position in the ETM turns into a net seller, being $z=-1$ (see Figure 5). When banking between periods is feasible and the firm highly values the possession of a surplus in emission rights at the end of the planning horizon, the period in which the firm acts as a net buyer last longer than when this surplus is not valued. Within this buying period, not used emission rights are accumulated to be used or sold in the future, $\dot{B}(t) > 0$. After this first period, the amount of these accumulated emission rights starts to fall until the end of the planning period. When the firm highly values the

possession of the surplus of emission rights at the end of the planning period, $g_B > \bar{\lambda}_B$, it keeps a surplus of emission rights at time T , $B(T) > 0$. Conversely, when $g_B \leq \bar{\lambda}_B$ then $B(T) = 0$ (see Figure 6). The change from a positive to a negative slope in B is related with the change from a rights-buyer position, $z > 0$, to a net seller of permits, $z < 0$. This switching time matches in both cases, as we can see comparing Figures 2 and 3.

Figure 4

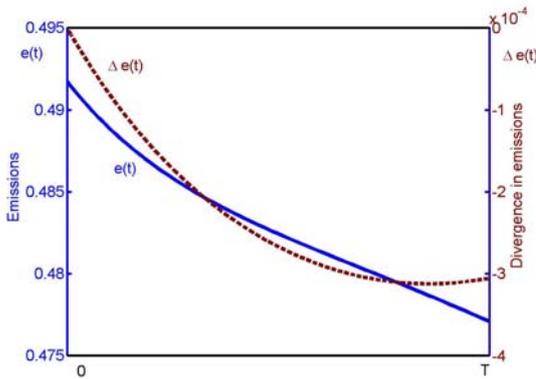


Figure 5

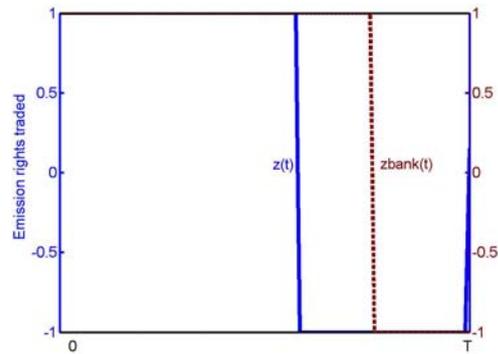
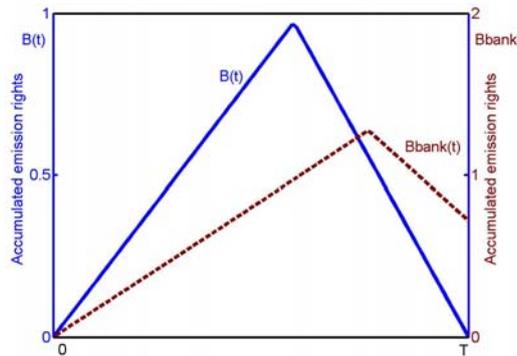


Figure 6



4. Sensitivity analysis

In this subsection we carry out a sensitivity analysis of the model outcomes, with and without banking between compliance periods. To that end we study the firm reaction to a 10%-change in the value of the more relevant parameters under two scenarios. The first scenario assumes that either banking between periods is banned or is allowed, the salvage value of B , g_B , is so small that no emission rights are accumulated at the end of the period. Under the second scenario, banking between periods is feasible, and the salvage value of B is high enough to lead to a positive $B(T)$.

In general, this 10%-change affects most variables in the same direction and with similar intensity under both scenarios. We can find some exceptions to this behaviour. On one hand, the accumulated amount of emission rights, B , is differently affected under each scenario. On the other hand, although a change in the salvage value of accumulated emission permits, g_B , does not generate changes in the model outcomes when banking between periods is forbidden, it does modify these results when banking is allowed and g_B is sufficiently large.

An increment in the depreciation rates, δ_k and δ_h , or in the investment costs of physical and technological capital, c_k and c_h , reduces investment in both types of capital, so the output of the firm decreases at a high speed. With respect to the emissions, they decrease faster when the investment cost or the depreciation rate of physical capital are higher. In contrast, an increment in the investment cost or in the depreciation rate of technological capital leads to higher emissions. With respect to the emission rights accumulated along the period we have to distinguish between the two scenarios. In the first scenario $g_B < \bar{\lambda}_B$, and no permit is maintained as a stock at the end of the planning period, $B(T)=0$. In this case, a higher δ_k or c_k reduces emissions and then, the firm accumulates emission rights faster at the beginning of the period, when it is a net buyer, while the stock of emission rights is faster reduced within the second subperiod which ends up with $B(T)=0$. By contrast, a higher δ_h or c_h increases emissions, and the effect on the accumulated emission rights is the reverse. In the second scenario, the amount of emission rights accumulated at the end of the period grows with δ_k or c_k and decreases with δ_h or c_h . Finally, the marginal value of the accumulated emission rights in the first scenario, $\bar{\lambda}_B$, behaves likewise as emissions. Therefore, a higher δ_k or c_k reduces $\bar{\lambda}_B$ and thus, increases the probability of $g_B > \bar{\lambda}_B$, i.e. the probability of a positive stock of emission rights at the end of the compromise period when banking between periods is allowed.

An increment in initial physical and technological capital, K_0 and h_0 , and in the marginal value that the entrepreneur gives to the possession of both types of capital at the end of the compliance period, g_k and g_h , produces the opposite response of variables in the model, as the one described to increments in δ_k , δ_h , c_k and c_h .

A 10% increment in the elasticity of the physical capital in the production function, α , implies higher physical and technological capital stocks and also a higher production. As a consequence of greater production, emissions are higher. This implies that the accumulation of emission rights is lower under both scenarios and $\bar{\lambda}_b$ is higher too. Therefore, the likelihood of a positive stock of emission rights at the end of the period is smaller.

The stronger the response of emissions to increments in abatement capital, β , the greater the stock of physical and abatement capital and, as a result, the higher the final production. At the beginning of the compromise period the increment in emissions associated with a higher production is stronger than a higher abatement capital. However, from a certain moment on, the higher abatement capital offsets the greater production and thus, emissions are lower. As a consequence, the firm accumulates a higher amount of emission rights. On the other hand, the marginal value of emission rights when banking between periods is banned, $\bar{\lambda}_b$, is lower, and so, the possibility of a positive $B(T)$ increases.

If the emission objective S is relaxed from $e(0)0.66$ to $e(0)0.726$, less effort is carried out in technological capital, decreasing this stock. On the other hand, a less restrictive objective leads the firm to enlarge production by increasing physical capital. Higher physical capital and lower technological capital increase the flow of emissions. However, the accumulated amount of permits is higher along the whole time period, due to the increment in its quota. The marginal value of the accumulated emission rights in the first scenario, $\bar{\lambda}_b$, is lower and so, a positive final stock of emission rights is more likely.

The effect of a change in the expectations about the price of the emission right price can be decomposed in two parts: first, a variation in the initial price, p_0 , from 0.001 to 0.0011; second, a variation in the expected growth rate of this price, p_1 , from 0.03 to 0.033. An increment in any of these parameters leads the firm to invest more in technological capital and less in physical capital. The combination of these two effects leads to lower emissions. At the same time $\bar{\lambda}_b$ is higher, decreasing the probability of a positive amount of emission rights stored at the end of the compromise period to be

used in the next compromise period when this transfer between periods is allowed. If the firm expects that the price will grow faster, the firm initially buys a greater amount of permits. Under the second scenario, with $g_B > \bar{\lambda}_B$, the period of rights-buyer is shorter and thereafter permits are sold to a higher extent. In consequence, the accumulated amount of permits is greater at the beginning and lower at the end, finishing with a lower $B(T)$.

When banking between periods is forbidden, the salvage value of accumulated emission permits, g_B , is zero, and the sensitivity of the model to changes in this parameter has no sense. However, when banking is allowed, two situations can be studied. If the salvage value of B is lower than its marginal value when banking was not allowed, $g_B < \bar{\lambda}_B$, the value of g_B has no relevance as long as it remains below $\bar{\lambda}_B$. Nevertheless, when banking is allowed from one compliance period to another and $g_B > \bar{\lambda}_B$, an increment in g_B has the same effect as the increment in prices previously analyzed.

5. Conclusions

This paper analyzes the optimal decisions of a firm that belongs to an annex B country of the Kyoto Protocol. This company has to determine the optimal trade-off between investments in physical and in technological capital. Both capitals are essential inputs in the production of final output, but the latter also reduces emissions per unit of output, helping the firm to fulfil with its quota. In order to reduce the cost of compliance, the company has to determine the amount of emission permit to buy or sell additionally in the emission permits market.

The transfer of unused allowances is allowed within the compliance period, although it is banned from one compliance period to another. We analyze how the relaxation of this prohibition may affect our analysis. Thus, we solve the model under two alternative scenarios, obtaining the optimal time paths for investment in physical and in technological capital, as well as for the amount of emission rights traded in the market.

The economic growth model with emission trading has been solved numerically using an algorithm implemented in MATLAB. When banking is banned the firm does not value the possession of emission rights at the end of the planning period, so the accumulated amount of permits at that time is zero. In contrast, when banking between

compromise periods is allowed two possibilities arise, depending on the firm's salvage value of accumulated emission permits with respect to its marginal value when banking between periods is banned. When the former is lower, the firm's behaviour matches with and without banking. In contrast, when it is higher we obtain a different behaviour. The comparison of these two different outcomes illustrates the effect of allowing banking from one compliance period to another in the economic growth as well as in the compliance of the Kyoto Protocol objective.

Allowing banking between compliance periods, when the salvage value of accumulated emission permits is sufficiently higher, implies a larger technological capital and a lower physical capital, obtaining lower production. The possibility of banking between periods implies less emissions and higher accumulation of emissions permits. In fact, the amount of accumulated emissions rights at the end of the planning period is positive.

Finally, a sensitivity analysis has been developed for both scenarios. In general, changes in parameters' values have the same effect and similar intensity under both scenarios. In both cases, a 10%-increment in the investment costs and depreciation rates of both capitals implies lower capital stocks and production. By contrast, an increment in the initial values of physical and technological capital; in the salvage value of both capitals; in the technological capital elasticity; in the marginal productivity of physical capital, and in the response of emissions to increments in abatement capital have a positive effect on both capital stocks and in production. This sensitivity analysis has showed that the effect on emissions of a change in parameters' values is the opposite of the effect on the marginal value of emission rights when banking between periods is banned. Therefore, when emissions fall, the marginal value increases, making a positive final stock of emission rights less likely.

To summarize, allowing banking between periods leads the firm to reduce investment in physical capital and to increment investment in technological capital. In consequence, emissions are lower. Further, if the firm expects a restrictive future agreement, it may keep a positive stock of emission rights at the end of the planning period which will be used in subsequent compromise periods.

References

Bernard, A.L. y Vielle, M. (2002): "Does Non Ratification of the Kyoto Protocol by the US Increase the Likelihood of Monoplistic Behaviour by Russia in the Market of

Tradable Permits?”, *Fifth Annual Conference on Global Economic Analysis*, June 5-7, Taipei, Taiwan.

Boemare, C. y Quirion, P. (2002): “Implementing Greenhouse Gas Trading in Europe: Lessons From Economic Theory and International Experiences”, *Ecological Economics*, 43, pp. 213–230.

CIMD Agencia de Valores, www.grupocimd.com.

Coase, R.H. (1960): “The Problem of Social Cost”, *Journal of Law and Economics*, Vol. III, pp.1-44.

Cronshaw, M.B. y Kruse J.B. (1996): “Regulated Firms in Pollution Permit Markets with Banking”, *Journal of Regulatory Economics*, 9, pp. 179-189.

Dales, J. H. (1968): *Pollution, Property and Prices*, University Press, Toronto.

Decision 18/CP.7 (2002), “Modalities, rules and guidelines for emissions trading under Article 17 of the Kyoto Protocol”, FCCC/CP/2001/13/Add.2, pp.55 .

Directive 2003/87/EC, of the European Parliament and of the Council of 13 October 2003 establishing a scheme for greenhouse gas emission allowance trading within the community.

Ellerman, A.D., Harrison, D. y Joskow, P.L. (2003): *Emissions Trading: Experience, Lessons, and Considerations for Greenhouse Gases*, Pew Center for Global Climate Change, Arlington, VA.

Grimaud, A. (1999): “Pollution Permits and Sustainable Growth in a Schumpeterian Model”, *Journal of Environmental Economics and Management*, 38, pp.249-266.

Haites, E. (2006): “Allowance Banking in Emissions Trading Schemes: Theory and Practice”, MIMEO, <http://www.margaree.ca/papers.html>.

Jorgensen, S. y Zaccour, G. (2001): “Incentive Equilibrium Strategies and Welfare Allocation in a Dynamic Game of Pollution Control”, *Automatica*, 37, pp. 29-36.

Kyoto Protocol to the United Nations Framework Convention on Climate Change (1997), FCCC/CP/1997/L.7/Add.1. <http://unfccc.int/resource/docs/convkp/kpspan.pdf>.

Law 1/2005, March 9, regulating an emission trading scheme for greenhouse gas emission.

Memo/04/44, de 4 de marzo de 2004, “Preguntas y respuestas sobre el comercio de derechos de emisión y los planes nacionales de asignación”, (Versión actualizada a 6 de Enero de 2005).

Montgomery, W.D. (1972): “Markets in Licenses and Efficient Pollution Control Programs”, *Journal of Economic Theory*, 5, pp. 395-418.

Ono, T. (2002): “Effects of Emission Permits on Growth and the Environment”, *Environmental and Resource Economics*, 21, pp.75-87.

Protocolo de Kyoto de las Naciones Unidas sobre el Cambio Climático, (1997), FCCC/CP/1997/L.7/ Add.1. <http://unfccc.int/resource/docs/convkp/kpspan.pdf>

Rubin, J.D. (1996): “A Model of Intertemporal Emission Trading, Banking, and Borrowing”, *Journal of Environmental Economics and Management*, 31, pp. 269-286.

Schleich, J., et al (2006): “Banning Banking in EU Emissions Trading?”, *Energy Policy*, 34, pp. 112-120.

Soberg, M. (2000): “Price Expectations and International Quota Trading: an Experimental Evaluation”, *Environmental and Resource Economics*, 17, pp. 259-277.

Stevens, B., y Rose, A. (2002): “A Dynamic Analysis of the Marketable permits Approach to Global Warming Policy: a Comparison of Spatial and Temporal Flexibility”, *Journal of Environmental Economics and Management*, 44, pp. 45-69.

Stokey, N.L. (1998): “Are There Limits to Growth?”, *International Economic Review*, 39, pp.1-31.

United Nations Framework Convention on Climate Change (1992). <http://unfccc.int>.

Van der Ploeg, F., y Zeeuw, A.J. (1992): “International Aspects of Pollution Control”, *Environmental and Resource Economics*, 2, pp. 117-139.

APLICACIÓN DEL MÉTODO DE VALORACIÓN CONTINGENTE A BIENES CULTURALES: ESTIMACIONES DE VALOR DEL MUSEO PATIO HERRERIANO

Ana María BEDATE CENTENO

e-mail: ana@emp.uva.es
Departamento de Economía Financiera y Contabilidad
Universidad de Valladolid

Luis César HERRERO PRIETO

e-mail: herrero@emp.uva.es
Departamento de Economía Aplicada
Universidad de Valladolid

José Ángel SANZ LARA

e-mail: angel@emp.uva.es
Departamento de Economía Financiera y Contabilidad
Universidad de Valladolid

Área temática: Métodos cuantitativos.

Resumen

La utilización del método de valoración contingente es bastante frecuente en el campo de los bienes públicos, pero para el caso de los bienes culturales es relativamente reciente y, en particular, en España el número de aplicaciones hasta la fecha es muy limitado. Entre las ventajas de este método tenemos el hecho de que permite realizar estimaciones del valor de uso directo y de uso pasivo, en función de que la encuesta de valoración se dirija al público visitante o al público potencial. Sin embargo, hay una población de interés, difícil de delimitar en la práctica, que está formada por público experto o especialmente atraído por el objeto cultural. En el presente trabajo, se ha hecho un estudio de valoración económica del Museo Patio Herreriano de Arte Contemporáneo Español, en el que se refleja el valor de uso directo, el valor de uso pasivo, y el valor asignado por el público interesado, con estimaciones comparadas entre visitantes explícitos del museo, ciudadanos de Valladolid y asistentes a la Feria de Arte Contemporáneo de Madrid.

Palabras clave: Economía de la Cultura, Economía del Patrimonio, Valoración Contingente, Disposición a pagar.

Abstract

The use of the contingent valuation method is quite frequent in the field of the public goods, but for the case of the cultural goods it is relatively recent and, in particular, in Spain the number of applications so far is very limited. Among the advantages of this method we have the fact that it allows to carry out estimates direct use and passive use value, when the survey of valuation goes to the public visitor or the potential public. However, a population of interest -difficult to define in the practice- exists, formed by expert public or specially attracted by the cultural object. In this paper, a study of economic valuation of the Museum Patio Herreriano of Spanish Contemporary Art has been made, in which the direct use value, the passive use value, and the value assigned by the interested public is reflected, with compared estimates among explicit visitors of the museum, citizens of Valladolid and assistants to the Fair of Contemporary Art of Madrid.

Key words: Cultural Economics, Heritage Economics, Contingent Valuation, Willingness to Pay.

1. Introducción

La *Economía de la Cultura* se está consolidando como una rama disciplinar destacada dentro del análisis económico, tanto porque constituye un terreno muy fértil para el razonamiento teórico, como por sus posibilidades de verificación empírica acerca del comportamiento de los hombres y las instituciones respecto de la cultura presente y acumulada. Dentro del amplio espectro analítico de esta nueva disciplina¹ uno de los aspectos que más aportes teóricos y aplicados está suscitando es el de la valoración de bienes culturales en general y, particularmente, los que pertenecen al ámbito del *Patrimonio Histórico*. Este tipo de elementos constituye un grupo de bienes de naturaleza muy variada, abarcando una gran cantidad de categorías genéricas, como los conjuntos históricos, edificios singulares, museos, sitios arqueológicos, obras de arte; así como también podríamos entender la música heredada, las tradiciones, el folclore, etc. Los principales problemas que surgen a la hora de realizar su valoración económica es que se trata de *bienes de no mercado*, primero, por la condición de propiedad pública de muchos de estos elementos, y, segundo, porque algunos beneficios asociados al consumo de patrimonio histórico como bien intangible no son comercializables, aún cuando exista una voluntad de pago latente. De esta forma, se suele distinguir entre *valor de uso directo* y *valor de uso pasivo*, donde este último se corresponde con el valor del patrimonio histórico como atributo de identificación o elemento de prestigio de una sociedad²; y la mayor parte de los métodos de valoración persiguen la estimación de la *Disposición a Pagar* (en adelante DAP) del público por el patrimonio histórico acudiendo a mercados indirectos o mediante la creación de mercados hipotéticos donde se simulen las transacciones que se producirían en la realidad.

A efectos de esta investigación aplicada, que se centra en la valoración económica de un museo, vamos a considerar que un establecimiento museístico es un elemento cultural acotado en el que normalmente se realiza un recuento de las visitas, a diferencia, por ejemplo, de los centros históricos y algunos edificios de atractivo

¹ Como referentes de la delimitación analítica de la *economía de la cultura* pueden verse Throsby (1994, 2001), Benhamou (1996) y Towse (1997); y más recientemente un análisis sobre el estado de la cuestión en Blaug (2001).

² De aquí se deducen, a su vez, los denominados *valor de opción, de legado y de prestigio* característicos del patrimonio histórico. Ver Frey (2000) y Herrero (2001).

artístico en los que resulta difícil controlar el número de visitas y, sobre todo, calibrar el interés del público en el uso directo de los mismos. Otra característica destacada de los museos es que, no sólo son poseedores de un patrimonio heredado, sino que además constituyen un producto cultural en sí mismos, ya que la forma en que se disponen las obras también puede considerarse, en ocasiones, como una creación cultural específica que es capaz de atraer a los visitantes. Una última reflexión que hemos tenido en cuenta en este trabajo sobre valoración económica de museos, es que su oferta es flexible³, pues nunca exponen la totalidad de sus obras, sino que presentan al público una parte de las mismas, lo cual les permite modificar las colecciones expuestas, realizar exposiciones temporales y/o itinerantes, prestar algunos de sus fondos a otros museos, etc., con el fin de dar a conocer sus obras y captar más visitantes. Por esta razón, no se trata de la gestión de un recurso único y fijo sino, en cierta manera, una oferta cultural flexible y variable.

En este trabajo se presenta una propuesta analítica en un tema bastante específico dentro del amplio campo que abarca la economía de los museos, como es el de la *valoración de las funciones expositivas de un museo*. No debemos olvidar en ningún momento que los museos representan un prototipo especial de patrimonio cultural que combina los aspectos inertes con los de creación viva, lo que comporta una experiencia cultural de carácter cualitativo y lleva asociados múltiples valores (de opción, existencia, legado, prestigio, etc.), de ahí la dificultad de su valoración. En esta investigación, concretamente se pretende lograr la estimación de funciones de demanda del Museo Patio Herreriano de Arte Contemporáneo Español en dos momentos distintos de tiempo. Las valoraciones económicas que hemos llevado a cabo presentan una doble vertiente: por un lado, la estimación de la voluntad de pago por un uso puntual del museo considerado como elemento de ocio, es decir, la valoración del consumo cultural a través del precio de entrada; y, por otro, la estimación de una cuota de donación en forma de contribución voluntaria dedicada a los gastos de mantenimiento de la colección, exposiciones y demás funciones desempeñadas por el museo. Para realizar esta última estimación se ha empleado el método de valoración contingente con un

³ Cf. Frey (2000, pp. 35 y 36).

formato de elección dicotómica de doble acotación seguido de una pregunta abierta⁴ y se han aplicado técnicas no paramétricas, en concreto el algoritmo de An y Ayala.

Las razones particulares de la elección del Museo Patio Herreriano de Arte Contemporáneo Español han sido fundamentalmente dos: una tiene que ver con la propia relevancia de la colección artística de este museo y las expectativas que se han formado alrededor de la nueva institución; y la otra, está relacionada con la propia creación de una dotación cultural *ex novo* para la ciudad de Valladolid y, por lo tanto, un elemento más dentro de su oferta cultural y conjunto de atractivos turísticos. Este segundo aspecto de variación de la oferta cultural y del equipamiento público resulta imprescindible para plantear, como es nuestro caso, estimaciones creíbles de la función de demanda de objetos no mercantilizables, pues los individuos juzgan su cambio de bienestar no por modificaciones en los precios, sino por cambios (reales o potenciales) en la cantidad, es decir, de la oferta.

Por lo que se refiere al primer aspecto, la entidad y trascendencia de la colección museística está de sobra acreditada, pues constituye en realidad una de las operaciones de coleccionismo privado más interesantes y eficaces de la historia reciente en España. En efecto, la Colección de Arte Contemporáneo, está formada por 851 obras hasta el momento, que se han ido comprando desde mediados de los años ochenta por un grupo de 23 empresas españolas, bajo la supervisión artística independiente de tres acreditados expertos en arte contemporáneo español. El fondo artístico ofrece una visión coherente y valiosa del desarrollo histórico del arte español desde 1918 hasta hoy, y constituye una colección viva con identidad propia⁵. La particularidad institucional de este nuevo museo resulta también bastante novedosa, pues consiste en una combinación de utilidad privada, la propiedad de la colección, que pertenece al grupo de empresas de la Asociación Colección Arte Contemporáneo; y utilidad pública, el Ayuntamiento de Valladolid, que cede uno de los claustros del antiguo Monasterio de San Benito para formar la sede del museo y propicia la creación de una Fundación Patio Herreriano que

⁴ En otros trabajos similares en el ámbito de la Economía de la Cultura se han empleado diferentes formatos de encuesta: elección múltiple (Martin,1994), pregunta abierta (Bille Hansen, 1997), tarjeta de pago (Chambers *et al.*, 1998), o elección dicotómica de acotación simple (Santagata y Signorello, 2000; Bravi *et al.*, 2002 y Thompson *et al.*, 2002).

tiene como objetivo la promoción y difusión del arte contemporáneo español a través del museo⁶. Por último, la rehabilitación del Patio Herreriano ha obtenido un claro y amplio reconocimiento, pues la intervención arquitectónica ha logrado combinar el respeto a la arquitectura original, destacando los elementos de interés que han llegado hasta nosotros, con las exigencias del nuevo uso del edificio en su condición de museo, y la recuperación de un área urbana del centro de Valladolid hasta ese momento bastante degradada y aislada del público. Todo este cúmulo de razones acreditan, en nuestra opinión, la elección del Museo Patio Herreriano de Arte Contemporáneo Español como prototipo de aplicación de la metodología de valoración económica propuesta.

2.- Metodología de trabajo

Esta investigación, está planteada en dos etapas. Una fase inicial en la que se pretenden obtener las valoraciones correspondientes a los visitantes y ciudadanos de Valladolid en los primeros meses de funcionamiento del museo, y una segunda fase, en la que se repiten las valoraciones con datos del tercer año de andadura del museo y además, se intenta recabar la opinión de público interesado en arte de vanguardia.

En la primera etapa del trabajo, con el fin de lograr el doble objetivo de presentar resultados acerca de la estimación del precio de entrada que el público consideraría oportuno abonar por la visita y de la obtención de la *disposición a pagar* (DAP) como cuota de financiación por parte de los ciudadanos de Valladolid y de los visitantes al museo, se planteó la realización de dos encuestas. La primera, denominada *Valladolid 1*, fue dirigida a los habitantes de la ciudad antes de la inauguración del museo el día 4 de junio de 2002, coincidiendo con la campaña de información y difusión llevada a cabo por el Ayuntamiento de Valladolid y el propio Museo Patio Herreriano. Concretamente, la recogida de información tuvo lugar entre los días 12 de mayo y 1 de junio de 2002 y se logró entrevistar a 911 personas que, con el apoyo visual de una serie de fotos del edificio y de algunas obras de arte, fueron capaces de colaborar facilitando los datos que se les requería. La explotación de los datos de esta primera encuesta permitió realizar

⁵ Puede verse a este respecto el Catálogo actual del Museo Patio Herreriano de Arte Contemporáneo Español en Torrente *et al.* (2002) y el Catálogo previo de la Colección de Arte Contemporáneo en Abad *et al.* (1995).

⁶ *Ibidem* para los aspectos institucionales y de gestión del Museo Patio Herreriano

diversos análisis, entre los cuales se encuentra el estudio de la predicción de los precios de entrada y la estimación del *valor de uso pasivo* del museo pues por aquel entonces aún no era posible visitarlo. La segunda encuesta, designada *Visitantes MPH 1*, estaba destinada únicamente a las personas que visitaban el museo con el fin de centrarnos en el *valor de uso directo* del mismo. Este segundo acopio de información tuvo lugar entre los meses de agosto y octubre del mismo año, recabando un total de 579 encuestas.

Un aspecto importante que debemos resaltar a este respecto es el efecto *impacto de la inauguración*, que presenta sus ventajas e inconvenientes para este estudio. Por un lado resulta beneficioso en el sentido de que se producen muchas visitas, tanto individuales como colectivas, y esto da pie a conocer la opinión de un público muy amplio y variado compuesto por gente de la calle, turistas, estudiantes, grupos de empresas, expertos en arte contemporáneo, personas relacionadas con la cultura, etc.; sin embargo, hay que tener en cuenta que éste no tiene por qué coincidir con el espectro habitual de visitantes del museo una vez se estabilice la situación, por lo que los resultados se tomaron con una razonable cautela⁷.

En la segunda fase de la investigación se repitió el esquema de trabajo de la primera etapa, recogiendo información entre los meses de julio de 2004 hasta junio de 2005. La encuesta dirigida a los ciudadanos de Valladolid para detectar el *valor de uso pasivo* se denominó *Valladolid 2* y con ella se logró entrevistar a 813 personas, mientras que la realizada a los visitantes al museo, denominada *Visitantes MPH 2*, consiguió información de 592 individuos. Existe sin embargo, una población de interés difícil de delimitar que está formada por el público experto o interesado, y para recabar su opinión, lanzamos una encuesta de valoración a la Feria de Arte Contemporáneo ARCO de Madrid, en la edición del año 2005.

En el estudio sobre el precio de entrada, se proporcionan estimaciones estadísticas puntuales y por intervalo, y, además, hemos pretendido detectar también las diferencias existentes entre distintos grupos de población atendiendo a la edad, ingresos, etc. Por lo que respecta a la valoración de uso pasivo por parte de los ciudadanos de Valladolid y de los visitantes a ARCO, y de uso directo por parte de los visitantes del museo, se

⁷ Al repetir las valoraciones una vez pasado el efecto *impacto de la inauguración* del museo, los resultados fueron algo diferentes.

aplica un ejercicio de valoración contingente, estimando la disposición a pagar por métodos no paramétricos, concretamente utilizando el algoritmo de An y Ayala, que supone una mejora del algoritmo de Turnbull habitualmente empleado en estas situaciones.

2.1.- Método de valoración contingente

El método de valoración contingente consiste en crear un mercado hipotético —contingente— y obtener mediante una encuesta la máxima DAP en valor monetario que el entrevistado otorga al bien que se está valorando en cuestión, o a una modificación en la calidad o cantidad del mismo, donde la oferta está representada por la persona que realiza la entrevista y la demanda por la persona entrevistada. En la encuesta, después de proporcionar la información específica sobre la finalidad del estudio y la situación que se va a valorar, se pregunta al encuestado por su DAP respecto del bien objeto de análisis. Esta pregunta se puede realizar de varias formas, entre las que destacamos las siguientes:

- (i) Una pregunta abierta, por ejemplo, *¿cuál es su máxima disposición a pagar por ...?*
- (ii) Un referéndum, en el que el encuestado contesta con un “sí” o con un “no” si su disposición a pagar es la que se le presenta.

Este último formato se puede llevar a cabo realizando una única pregunta o una serie de preguntas encadenadas con diferentes ofertas que dependen de las respuestas anteriores, por ejemplo: *¿estaría usted dispuesto a pagar X pesetas por ...?*; en caso afirmativo, *¿estaría usted dispuesto a pagar 2X pesetas por ...?*; en caso negativo, *¿estaría usted dispuesto a pagar X/2 pesetas por ...?* Esta configuración de dos preguntas enlazadas se conoce con el nombre de *formato de elección dicotómica de doble acotación* y, en el fondo, hacen que la persona encuestada se encuentre en una situación parecida a la de ir a un mercado real y tomar la decisión de comprar o no, en función del precio que le ofrecen, con una pequeña salvedad, y es que en este caso el individuo no tiene que abonar la cantidad que propone⁸.

⁸ Un estudio más detallado de la metodología a emplear en un estudio de valoración contingente puede consultarse Mitchell y Carson (1989), Riera (1994) y Sanz *et al.* (2003).

En los estudios de valoración contingente que emplean el formato de elección dicotómica de acotación doble, al formular consecutivamente las dos preguntas de valoración, tenemos cuatro posibles respuestas (No - No, No - Sí, Sí - No, Sí - Sí) que dividen al intervalo monetario $[0, \infty)$ en cuatro intervalos más pequeños ($[0, O_b)$, $[O_b, O_i)$, $[O_i, O_a)$, $[O_a, \infty)$), con lo que la única información de la que disponemos finalmente es que la verdadera DAP cae dentro de uno de ellos, es decir, los datos están agrupados en intervalos.

De esta forma, sea $\{X_1, X_2, \dots, X_N\}$ una muestra aleatoria de una función de distribución absolutamente continua $F(x)$ definida en el intervalo $[0, \infty)$, como puede ser la de la variable DAP por un bien público. En los estudios de valoración contingente en los que se emplea el formato dicotómico de doble acotación, la información disponible se traduce en que para cada individuo n , existe un intervalo $A_n = [a_n, b_n)$, tal que $X_n \in A_n$, luego la función del logaritmo de la verosimilitud en términos de la función de supervivencia es,

$$l(S) = \sum_{n=1}^N \log P(A_n) = \sum_{n=1}^N \log [S(a_n) - S(b_n)].$$

Puede observarse que la función del logaritmo de la verosimilitud sólo depende de S a través de sus valores en los puntos de corte que definen todos los intervalos A_n . Para explorar más allá de este punto, sea $0 = t_0 < t_1 < \dots < t_W < t_{W+1} = \infty$ la permutación ordenada de todos los puntos de corte, que representa el conjunto de W ofertas distintas que se plantean en cada estudio. Estos t_k constituyen una partición del soporte de la distribución $[0, \infty)$ en $W + 1$ intervalos básicos más pequeños de la forma $[t_{k-1}, t_k)$. Así, para cada encuestado n de la muestra, existen dos enteros i_n y j_n , verificando $0 \leq i_n < j_n \leq W+1$ tales que la DAP del individuo n cae en un intervalo $[t_{i_n}, t_{j_n})$. Además, en cada entrevistado, observamos una serie de características tales como edad, ingresos, nivel de educación, sexo, etc., y la recogemos en un vector al que denotamos por x_n . Entonces podemos resumir la información obtenida como:

$$\{(i_n, j_n, x_n) \quad \text{con } n = 1, 2, \dots, N\}$$

En el presente trabajo, el número de ofertas empleado para la obtención de la DAP es de $W = 9$, recogidas en siete adaptaciones del modelo de encuesta y definidas a partir de

las distintas ofertas iniciales⁹. Una vez expuesta la situación de partida y qué tipo de información está disponible, pasamos a describir de forma breve el *algoritmo no paramétrico de An y Ayala*.

2.2.- Algoritmo no paramétrico de An y Ayala

Los primeros algoritmos no paramétricos propuestos para estimar la función de supervivencia para muestras con información incompleta, como ocurre por ejemplo cuando los datos vienen agrupados en intervalos, son los sugeridos por Ayer *et al.* (1955) para datos binarios simples, y el de Kaplan y Meier (1958) para datos censurados por la derecha o censurados en intervalos regularmente. El algoritmo autoconsistente de Turnbull (1974) y especialmente su propia generalización (Turnbull, 1976), han supuesto un avance definitivo para este tipo de técnicas cuando los datos están censurados por la izquierda, censurados por la derecha, o censurados en intervalos regulares. Finalmente, el algoritmo propuesto por An y Ayala (1996) generaliza el algoritmo autoconsistente de Turnbull para poder tratar datos agrupados arbitrariamente, lo que constituye un esquema muy común dentro del análisis económico de valoración de bienes públicos.

El *algoritmo de An-Ayala* se basa en el sistema de ecuaciones de autoconsistencia que se deriva de las condiciones necesarias de primer orden que caracterizan la maximización del logaritmo de la verosimilitud no paramétrica¹⁰. En cada iteración, las estimaciones actuales se usan para convertir datos arbitrariamente agrupados, en datos censurados en intervalos regularmente (el *E-paso*), y entonces aplicar el estimador de Kaplan-Meier para obtener una nueva ronda de estimaciones (el *M-paso*).

De esta manera y empleando la notación citada anteriormente, para cada individuo n sean i_n y j_n valores tales que $a_n = t_{i_n}$ y $b_n = t_{j_n}$. Entonces, para $i_n, j_n = 0, 1, 2, \dots, W + 1$ e $i_n < j_n$, denotemos al número de observaciones agrupadas en el intervalo $[t_i, t_j)$ como:

⁹ El modelo de encuesta es único pero aparecen de forma aleatoria siete posibles pujas de partida, con el fin de evitar el *efecto anclaje* en una única cifra.

¹⁰ El algoritmo puede consultarse de forma íntegra en An y Ayala (1996).

$$\gamma_{ij} = \sum_{n=1}^N 1_{(i_n=i, j_n=j)}$$

En estas condiciones, el algoritmo autoconsistente de An-Ayala itera los dos pasos siguientes hasta alcanzar la convergencia:

1°) *E-paso*: Sea $1 = \hat{S}_0^0 \geq \hat{S}_1^0 \geq \dots \geq \hat{S}_W^0 \geq \hat{S}_{W+1}^0 = 0$ el conjunto actual de estimaciones. Para todos los $k = 1, \dots, W+1$, calculamos el número de “muertes” entre t_{k-1} y t_k :

$$\delta_k^0 = \sum_{r=0}^{k-1} \sum_{s=k}^{W+1} \gamma_{rs} \frac{\hat{S}_{k-1}^0 - \hat{S}_k^0}{\hat{S}_r^0 - \hat{S}_s^0}$$

y el número de “en riesgo” en t_{k-1} :

$$n_k^0 = \sum_{s=k}^{W+1} \delta_s^0,$$

2°) *M-paso*: Aplicar el estimador de Kaplan-Meier a los datos $(\delta_1^0, \delta_2^0, \dots, \delta_W^0)$ para obtener una nueva ronda de estimaciones, esto es,

$$\begin{cases} \hat{S}_0^1 = 1 \\ \hat{S}_k^1 = \left[1 - \frac{\delta_k^0}{n_k^0} \right] \hat{S}_{k-1}^1, & k = 1, 2, \dots, W \\ \hat{S}_{W+1}^1 = 0. \end{cases}$$

Una vez que se alcanza la convergencia, la función de supervivencia obtenida expresa, en nuestro caso, la disposición a pagar de los individuos por el objeto del patrimonio cultural que estamos analizando, la cual es asimilable a la función de demanda individual y sirve como base para obtener el excedente del consumidor y, por lo tanto, el valor económico asignado al patrimonio cultural¹¹.

2.3.- Estimación de la disposición media a pagar

Para realizar la estimación de la disposición media a pagar (μ) en este contexto se pueden efectuar dos posibles aproximaciones:

¹¹ Una justificación microeconómica más detenida puede verse en Mitchell y Carson (1989), Chambers *et al.* (1998) y Sanz (2004).

- a) Una alternativa empleada por los investigadores de valoración contingente es la denominada “conservadora” que asigna toda la probabilidad $S_k - S_{k-1}$ al extremo inferior t_{k-1} del intervalo correspondiente. Luego, en este caso, el valor de μ es la suma del área de W rectángulos:

$$\mu = \sum_{k=1}^W (t_k - t_{k-1}) S_k$$

- b) La segunda opción es considerar la interpolación lineal de la función de supervivencia, con lo cual μ es la suma de $W + 1$ trapezoides:

$$\mu = \frac{1}{2} \sum_{k=1}^{W+1} (t_k - t_{k-1}) (S_k + S_{k-1})$$

donde t_{W+1} es un valor razonable que se debe elegir con la convicción de que a partir de dicho valor la probabilidad de la DAP es cero.

Los resultados de esta investigación se presentarán solo con la alternativa conservadora, dada la tendencia bastante habitual a la sobrestimación de la DAP en los estudios de valoración contingente. En los siguientes apartados, vamos a detallar concretamente las aplicaciones realizadas¹².

3. Estimación de los precios de entrada

En la encuesta realizada a los ciudadanos de Valladolid (*Valladolid I*), antes de la inauguración del museo, se preguntaba a los encuestados por el precio de entrada que estarían dispuestos a pagar por una visita, es decir, se pedía la valoración estrictamente del consumo de ocio que justifica la realización de una visita al museo. Aunque esta primera encuesta estaba enfocada a realizar diversos estudios de valoración y tipología del turista cultural, los datos recogidos a este respecto fueron sumamente interesantes¹³.

En este apartado, lo que se valora exclusivamente es el uso de una de las funciones del museo, la exhibición de la colección museística por el visitante potencial. De esta

¹² Un mayor detalle de los pormenores teóricos y aplicados de esta investigación puede verse en Herrero *et al.* (2004).

¹³ Para la realización del trabajo de campo contamos con la colaboración de dos agentes encuestadores, que recogieron información en lugares muy frecuentados y céntricos de Valladolid. Las entrevistas se efectuaron en cinco zonas, a saber: i) Plaza de Poniente e Iglesia de San Benito; ii) Plaza de San Pablo y

manera, como podemos ver en la Tabla 1, el precio medio estimado con 893 respuestas válidas, es de 4 euros, con una dispersión bastante elevada. La cuarta parte de los entrevistados declararon un precio de entrada inferior a 2 euros, la mitad de la muestra se inclinó por menos de 3 euros y las tres cuartas partes de la misma por un precio a lo sumo igual a seis euros¹⁴. El intervalo de confianza al 95% para el precio medio es (3,72, 4,28). Este es un dato significativo de la valoración y las expectativas creadas por el Museo Patio Herreriano, pues esta estimación de precios superaba aproximadamente en un 50% a la media de precios de los museos vallisoletanos en funcionamiento hasta ese momento, que podían constituir un referente cierto para el público encuestado¹⁵; y superaba también los precios de algunos otros museos del mismo tipo a nivel nacional, como el Museo Nacional Centro de Arte Reina Sofía (3 €), el Instituto Valenciano de Arte Moderno (2 €), y el Museo de Arte Abstracto de Cuenca (3 €)¹⁶. En todo caso, la estimación de precios realizada por el público vallisoletano se acercaba más al precio del *día del visitante* (media entrada) que al precio habitual de un día ordinario que era de 6 €.

Tabla 1. Descripción de la variable Precio de entrada

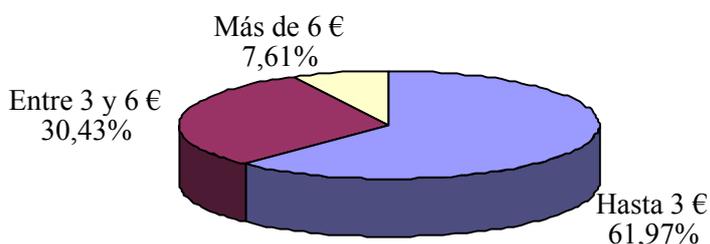
Tamaño muestral	893
Media	4,00
Desviación típica	4,22
Coefficiente de variación	105,62
Primer cuartil	2,00
Mediana	3,00
Tercer cuartil	6,00

Plaza de San Miguel; iii) Plaza Mayor y Plaza de la Fuente Dorada; iv) La Antigua, Plaza de la Universidad y Plaza de Santa Cruz; v) Plaza de España.

¹⁴ Una representación de la distribución de precios puede verse en el Gráfico 1

¹⁵ Los ejemplos más representativos, en este sentido, pueden ser el Museo Nacional de Escultura (2,40 €), el Museo Catedralicio (2,10 €), el Museo Oriental (2,70 €) y el Museo de Valladolid (1,20 €).

Gráfico 1. Distribución del precio de entrada



Las variables de carácter sociodemográfico como sexo, edad, estudios y nivel de ingresos, permiten hacer estudios particulares en lo referente a las estimaciones realizadas, y para ello las hemos codificado en distintas categorías, de la forma que se indica en la Tabla 2. La comparación de medias para los distintos grupos, habitualmente se efectúa utilizando un análisis de la varianza (ANOVA), siempre que se cumplan los requisitos de este análisis¹⁷ o, en su defecto, se opta por hacer un estudio comparativo de las medianas basado en el contraste no paramétrico de Kruskal Wallis.

Tabla 2. Codificación de las variables sociodemográficas

Sexo	Edad	Estudios	Ingresos
Hombre	Hasta 26	No universitarios	Sin ingresos
Mujer	27 – 50	Universitarios	Menos de 1803 €
	Más de 51		Más de 1803 €

Los resultados obtenidos a partir de la encuesta *Valladolid 1*, únicamente muestran diferencias significativas al comparar el precio de entrada por grupos de edad (Tabla 3). La Tabla 4 referente al contraste de diferencia de medianas, muestra diferencias significativas en los rangos promedio y el Gráfico 2, indica que las personas de mayor edad dan valores más bajos para el precio de entrada, que cualquiera de los otros dos grupos, puesto que la caja correspondiente al grupo de mayor edad está desplazada hacia la izquierda.

¹⁶ No ocurre así con el Museo Guggenheim y el Museo de Arte Contemporáneo de Barcelona cuyos precios eran 7 € y 6 € respectivamente.

¹⁷ Los supuestos básicos del ANOVA son: varianzas iguales dentro de los grupos considerados y residuos independientes idénticamente distribuidos según una normal. En nuestro caso ambas condiciones no se cumplen.

Tabla 3. Medias de Entrada por grupo de Edad (Valladolid 1)

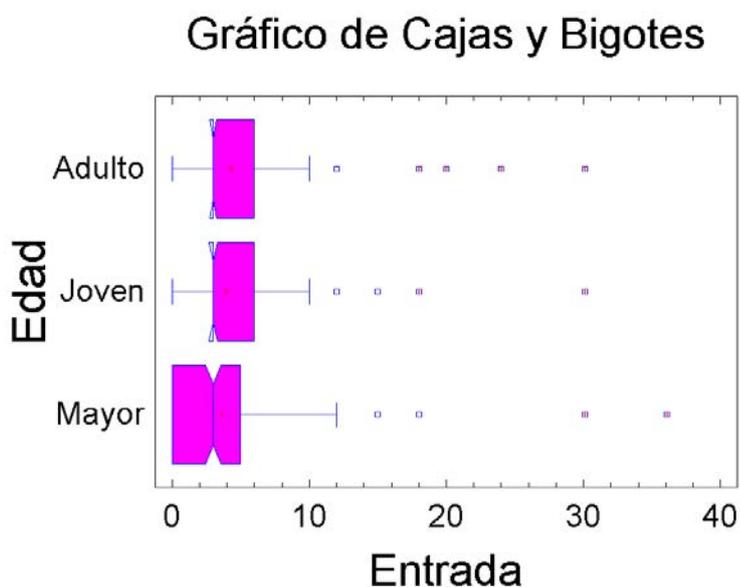
Edad	Tamaño muestral	Medias
Joven	262	3,91
Adulto	425	4,29
Mayor	203	3,55
Total	890	4,00

Tabla 4. Contraste de Kruskal-Wallis para la diferencia de medianas de Entrada según grupo de Edad (Valladolid 1)

Edad	Tamaño muestral	Rango Promedio
Joven	262	450,939
Adulto	425	473,786
Mayor	203	379,261

Estadístico = 19,7725
P-valor = 0,0000508702

Gráfico 2. Gráfico de caja y bigotes para Entrada según grupo de Edad (Valladolid 1)



4.- Estimaciones de la DAP: valor de uso pasivo y valor de uso directo

Nuestro propósito, en esta parte del estudio, es obtener distintas valoraciones referidas al Museo Patio Herreriano, pero no desde el punto de vista del consumo puntual, como hacíamos al estimar el precio de entrada; si no que, en este caso, intentaremos obtener la valoración global de la dotación cultural (creación, mantenimiento y ejecución de sus actividades), representada en forma de asignación alícuota de un individuo, al estilo de lo que serían los precios revelados en la financiación de un bien público.

La información utilizada para estimar el valor de *uso pasivo*, procede de las encuestas realizadas a los ciudadanos de Valladolid (*Valladolid 1* y *Valladolid 2*), mientras que para estimar *el uso directo* se emplean las encuestas realizadas a los visitantes (*Visitantes MPH 1* y *Visitantes MPH 2*) a la salida del museo. Todas las encuestas tienen una estructura similar dividida en cuatro partes, la primera de las cuales es introductoria y en ella se realizan preguntas relativas a los hábitos de consumo cultural del encuestado para detectar su afinidad con el objeto de consumo, pero sobretodo con el fin de que tome contacto con el tema a tratar, que es su voluntad de pago. El segundo bloque de la encuesta constituye la parte central de la misma, donde se abordan las preguntas sobre la disposición a pagar en términos monetarios. En esta investigación se utiliza la ventaja estratégica de que la situación planteada va ligada a un incremento de la dotación cultural de la ciudad por tratarse de un museo nuevo con un contenido sin precedentes en Valladolid. Tras una breve introducción, se presenta de forma clara y concreta el escenario hipotético de valoración contingente, especificando bien qué se valora y cómo se efectuaría el supuesto pago, que consistía en una cuota de donación anual¹⁸. Evidentemente, en la encuesta realizada antes de la inauguración del museo, la pregunta de valoración de algo que aún no se ha podido visitar se realiza con el apoyo de una selección de fotografías tanto del edificio como de alguna de las obras más representativas de la exposición.

¹⁸ En otras investigaciones similares se hacen propuestas como la aportación económica a una fundación que gestione las actividades propias del bien cultural a valorar, el incremento de los impuestos para contribuir a fines de carácter cultural, o el establecimiento de precios adecuados para estos gastos. Nosotros nos hemos inclinado por una proposición parecida al primer tipo, que se corresponde más a la particularidad jurídica de la institución gestora del Museo Patio Herreriano.

En esta investigación, la formulación de la pregunta de valoración se ha realizado mediante el formato dicotómico de doble acotación seguida de una pregunta abierta. Las ofertas iniciales presentadas a los entrevistados a través de una pregunta cerrada, son diferentes en cada caso y se asignan de forma aleatoria, con el fin de evitar, en lo posible, el sesgo de anclaje en las estimaciones. Las cantidades propuestas en esta primera pregunta, se calcularon en función de las contribuciones anuales realizadas por las asociaciones de *Amigos de los Museos* más representativas de nuestro país¹⁹ y fueron concretamente: 6, 15, 30, 45, 60, 90 y 150 euros como escala de ofertas iniciales. En función de la primera respuesta dada por el entrevistado, se formulaba la segunda pregunta cerrada proponiendo cantidades inmediatamente inferiores: 3, 6, 15, 30, 45, 60 y 90 euros respectivamente, si la primera pregunta obtuvo respuesta negativa; e inmediatamente superiores: 15, 30, 45, 60, 90, 150 y 240 euros, en caso contrario. Finalmente a todos los encuestados se les hacía una pregunta abierta en la que debían expresar su máxima DAP final.

En la aplicación del método de valoración contingente, un aspecto muy importante es distinguir entre aquellos ciudadanos que registran una DAP nula pero mantienen una voluntad latente favorable a la contribución a través de otros canales de financiación (por ejemplo, impuestos); de aquellos otros que responden a la pregunta abierta de valoración con cero porque rechazan el escenario hipotético que se les ha planteado en la investigación y que, por tanto es recomendable excluirlos para hacer una estimación certera. Una forma habitual de resolver este problema es mediante la inclusión de una pregunta complementaria para todos aquellos que declaran cero como máxima DAP, acerca de los posibles motivos de su respuesta para poder seleccionar posteriormente entre los que rechazan el escenario de valoración y aquellos que lo aceptan pero registran en ese momento una DAP nula. De esta forma, en las encuestas que hemos realizado en esta investigación, hemos tomado como ceros reales aquellos que consideran que ya pagan impuestos o que les gustaría contribuir pero no pueden en ese momento, porque todos ellos manifiestan de alguna manera una voluntad latente de pago.

¹⁹ Se consultó en este sentido, la cuota de amigos del Museo Nacional Centro de Arte Reina Sofía (65 €), Museo Nacional de Escultura (30 €), Instituto Valenciano de Arte Moderno (36 €), Museo del Prado (70 €), etc.

El tercer bloque de preguntas del cuestionario tiene que ver con el origen de los visitantes, el medio de transporte utilizado, el número de acompañantes, los motivos de la visita, el grado de satisfacción alcanzado en la misma y la importancia de este museo en la vida cultural de la ciudad. Lógicamente, en las encuestas realizadas a ciudadanos de Valladolid este bloque de preguntas se omite. El grado de dificultad de las preguntas se va relajando, y se termina con un cuarto bloque formado por una serie de preguntas de carácter sociodemográfico que proporcionan información adicional para realizar estudios comparativos particulares²⁰.

4.1.- Valor de uso pasivo (encuestas *Valladolid 1* y *Valladolid 2*)

El valor de uso pasivo, es el que se asigna por la propia existencia del bien, por la posibilidad de utilizarlo en otro momento, por poder legarlo a generaciones futuras, etc. Por esta razón hemos considerado que podríamos obtenerlo encuestando a ciudadanos de Valladolid, que pueden tener cierto conocimiento de la existencia e importancia de este museo, pero que en el momento en el que se les entrevista no están haciendo uso del mismo. Tras eliminar los ceros protesta²¹, el número de encuestas utilizadas fue 766 en la primera fase y 588 en la segunda. Con ellas la aplicación del algoritmo de An y Ayala, permite calcular la función de supervivencia empírica, que viene a representar la curva de demanda del Museo Patio Herreriano revelada por los ciudadanos de Valladolid y así determinar el valor del excedente del consumidor como estimación de la máxima disposición al pago.

Si nos centramos en los aspectos matemáticos a la hora de calcular el excedente del consumidor, la estimación se ha realizado mediante la suma de las áreas de los rectángulos que podríamos formar por debajo de la función de supervivencia empírica que da lugar a una curva de demanda discreta. Esta estimación, como ya se indicó en el apartado 2.3 es conservadora, pero nos parece la más conveniente dado que frecuentemente se producen sesgos tendentes a sobrestimar la DAP debidos el carácter hipotético de la pregunta de valoración.

²⁰ Para mayor detalle ver Herrero *et al.* (2004).

²¹ En la primera encuesta el porcentaje de ceros protesta fue el 18,02% de los datos ya depurados y, en la segunda, el 26,04%.

Las estimaciones obtenidas para el valor general de la DAP son de 21,39 euros en la primera encuesta y 14,14 euros en la segunda. La Tabla 5, la Tabla 6, el Gráfico 3 y el Gráfico 4 muestran los principales resultados acerca de las funciones de supervivencia.

Tabla 5. Distribución de la DAP general. (Valladolid 1)

Oferta inicial	Función de densidad $f(x)$	Función de distribución $F(x)$	Función de supervivencia $S(x)$
0	0,3241	0,3241	0,6759
3	0,0226	0,3467	0,6533
6	0,1806	0,5272	0,4728
15	0,1412	0,6684	0,3316
30	0,1682	0,8366	0,1634
45	0,0379	0,8745	0,1255
60	0,0576	0,9321	0,0679
90	0,0385	0,9706	0,0294
150	0,0294	1,0000	0,0000
240	0,0000	1,0000	0,0000

Gráfico 3. Función de supervivencia de la DAP general (Valladolid 1)

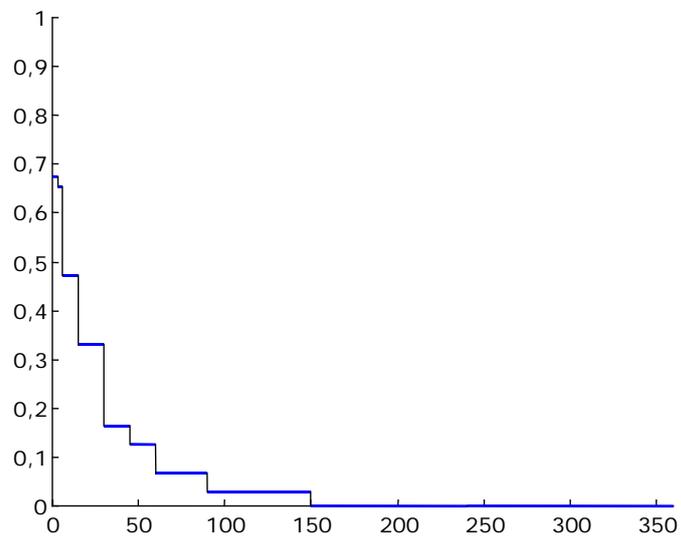
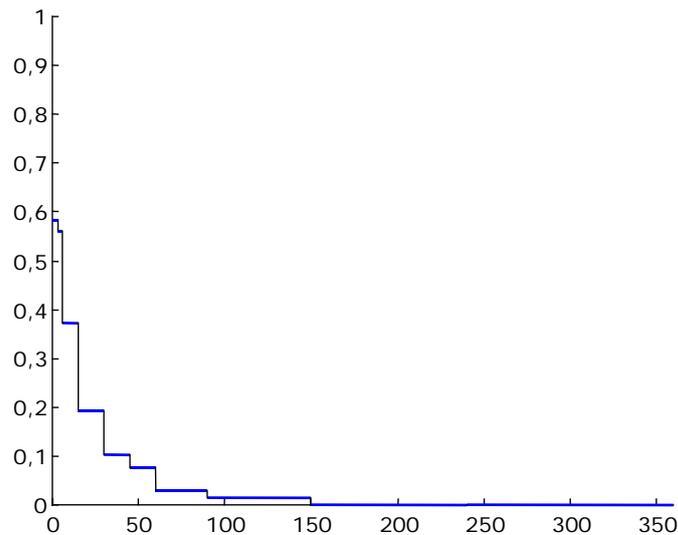


Tabla 6. Distribución de la DAP general. (Valladolid 2)

Oferta inicial	Función de densidad $f(x)$	Función de distribución $F(x)$	Función de supervivencia $S(x)$
0	0,4172	0,4172	0,5828
3	0,0237	0,4409	0,5591
6	0,1868	0,6277	0,3723
15	0,1789	0,8066	0,1934
30	0,0914	0,8980	0,1020
45	0,0243	0,9223	0,0777
60	0,0482	0,9705	0,0295
90	0,0147	0,9853	0,0147
150	0,0147	1,0000	0,0000
240	0,0000	1,0000	0,0000

Gráfico 4. Función de supervivencia de la DAP general (Valladolid 2)



4.2.- Valor de uso directo (*Visitantes MPH 1* y *Visitantes MPH 2*)

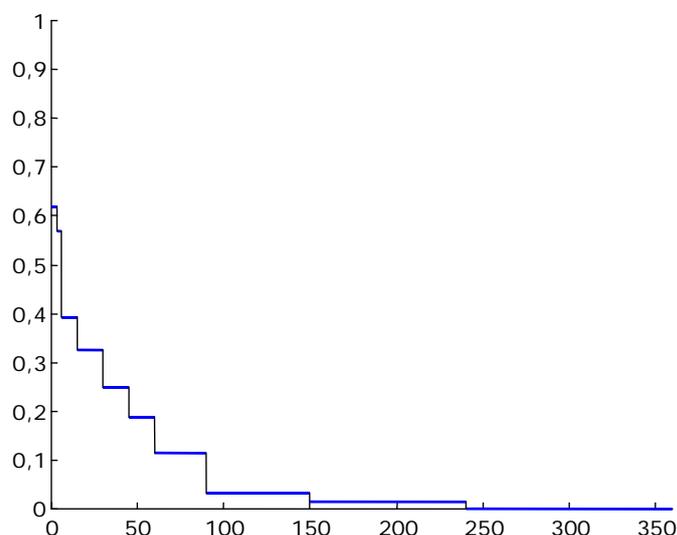
El valor de uso directo es el atribuido por los usuarios del museo en razón de los beneficios en su mayoría intangibles (educación, emoción estética, ocio, etc.), que obtienen de la visita. Del mismo modo que en el apartado anterior, tras eliminar los ceros protesta, el tamaño muestral se redujo a 485 y 437 en las encuestas *Visitantes MPH 1* y *Visitantes MPH 2* respectivamente. Las estimaciones medias obtenidas para la DAP, fueron de 25,32 euros en la primera fase del estudio y 23,88 euros en la

segunda²². La Tabla 7, la Tabla 8, el Gráfico 5, y el Gráfico 6 muestran los principales resultados de la estimación de la curva de demanda cuando se aplica el algoritmo no paramétrico de An y Ayala.

Tabla 7. Distribución de la DAP general (*Visitantes MPH 1*)

Oferta inicial	Función de densidad $f(x)$	Función de distribución $F(x)$	Función de supervivencia $S(x)$
0	0,3804	0,3804	0,6196
3	0,0505	0,4309	0,5691
6	0,1758	0,6067	0,3933
15	0,0687	0,6754	0,3246
30	0,0757	0,7510	0,2490
45	0,0610	0,8120	0,1880
60	0,0728	0,8848	0,1152
90	0,0819	0,9666	0,0334
150	0,0191	0,9857	0,0143
240	0,0143	1,0000	0,0000

Gráfico 5. Función de supervivencia de la DAP general (*Visitantes MPH 1*)

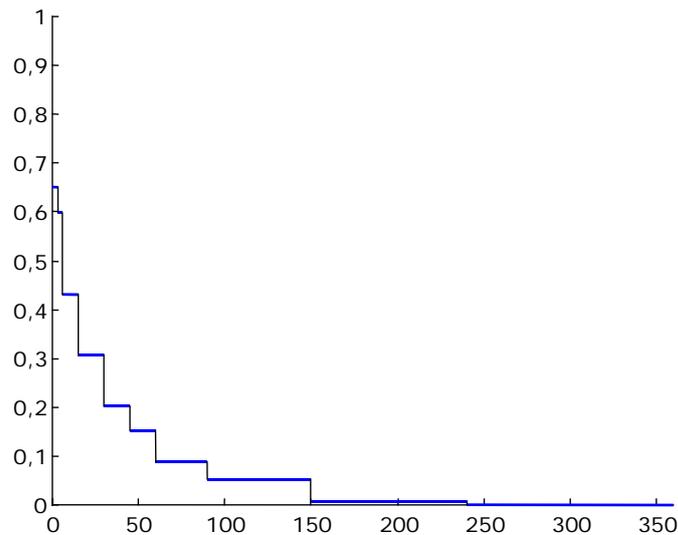


²² Curiosamente, en un estudio con el mismo planteamiento metodológico realizado sobre el Museo Nacional de Escultura de Valladolid (Sanz *et al.* 2003), la DAP de los usuarios directos resulta ser muy parecida, 25,13 € (4.181 ptas.); lo cual puede ser indicativo, tanto de la calidad del contenido y continente del Museo Patio Herreriano, como de la satisfacción revelada por los visitantes por contar con una dotación cultural de arte de vanguardia.

Tabla 8. Distribución de la DAP general (*Visitantes MPH 2*)

Oferta inicial	Función de densidad $f(x)$	Función de distribución $F(x)$	Función de supervivencia $S(x)$
0	0,3501	0,3501	0,6499
3	0,0523	0,4025	0,5975
6	0,1661	0,5685	0,4315
15	0,1248	0,6934	0,3066
30	0,1031	0,7964	0,2036
45	0,0497	0,8461	0,1539
60	0,0650	0,9111	0,0889
90	0,0364	0,9475	0,0525
150	0,0473	0,9947	0,0053
240	0,0053	1,0000	0,0000

Gráfico 6. Función de supervivencia de la DAP general (*Visitantes MPH 2*)



5.- Estimaciones de la DAP del público interesado: valor de uso pasivo

En muchas ocasiones, los encuestados rechazan el ejercicio de valoración contingente porque consideran que el objeto a valorar debe financiarlo quien verdaderamente lo usa o tiene un interés específico. Evidentemente, se intuye que este grupo de personas tiene una mayor disposición al pago por este tipo de bienes que el resto de los individuos, aunque son un sector reducido de la población que, por otra parte, no resulta fácil de localizar a la hora de hacer una encuesta.

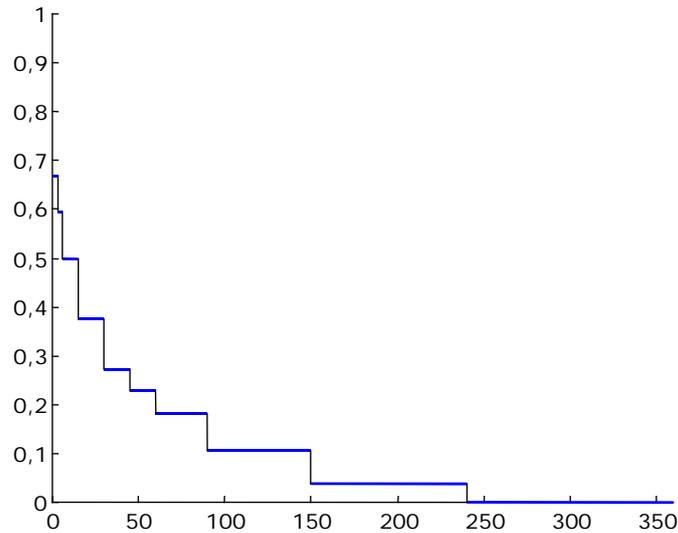
Los amigos del museo son claramente un grupo de interés pero no de carácter general, ya que manifiestan una preferencia clara por el museo muy unida a otras circunstancias como un gran conocimiento de la colección, el hecho de vivir en la ciudad o tener la posibilidad de hacer visitas frecuentes, el reconocimiento de señas de identidad, etc. La población de expertos o interesados específicamente en el arte contemporáneo, es más amplia y no está necesariamente vinculada a la ciudad donde se ubica el museo objeto de estudio, por esta razón hemos dirigido la encuesta a la Feria de Arte Contemporáneo de Madrid ARCO, donde acuden artistas, galeristas, críticos, compradores, amantes del arte de vanguardia, etc., es decir, un espectro amplio de individuos implicados o no en el arte contemporáneo, pero con la formación, la sensibilidad, y el gusto desarrollados en este tipo de manifestaciones artísticas.

La encuesta de valoración contingente realizada en febrero de 2005, nos permitió obtener una valoración de uso pasivo del Museo Patio Herreriano correspondiente a la población de expertos. En los cinco días que duró la feria, se recogieron 478 encuestas, que después de eliminar los ceros protesta (el 39,83% de los datos) redujeron el tamaño de la muestra a 287, obteniéndose 36,78 euros como estimación del valor medio de la DAP. La Tabla 9 y el Gráfico 7 ilustran este valor. Como puede apreciarse, es una DAP muy elevada, sobretodo si la comparamos con los demás resultados obtenidos en este trabajo.

Tabla 9. Distribución de la DAP general (*Visitantes ARCO*)

Oferta inicial	Función de densidad $f(x)$	Función de distribución $F(x)$	Función de supervivencia $S(x)$
0	0,3324	0,3324	0,6676
3	0,0719	0,4043	0,5957
6	0,0957	0,5000	0,5000
15	0,1250	0,6250	0,3750
30	0,1030	0,7281	0,2719
45	0,0441	0,7721	0,2279
60	0,0451	0,8173	0,1827
90	0,0761	0,8934	0,1066
150	0,0678	0,9612	0,0388
240	0,0388	1,0000	0,0000

Gráfico 7. Función de supervivencia de la DAP general (*Visitantes ARCO*)



6.- Conclusiones

Las razones para la aplicación de este tipo de metodología de valoración contingente a los elementos del patrimonio cultural se justifican por la inexistencia de mercados relevantes que canalicen de forma conveniente el grado de escasez, de coste o de deseabilidad de estos bienes, debido a las particularidades por las que se caracterizan, como son su condición de bienes públicos, la ausencia de valores de cambio representativos, la coexistencia de valores intangibles latentes que no se manifiestan en forma de precios y el carácter no comercializable de muchos de estos elementos. Para estos casos, existen procedimientos estadísticos con fuerte fundamentación microeconómica que calculan el valor económico asignado al patrimonio cultural, cuya utilidad sirve, al menos, de jerarquía ordinal comparativa entre diferentes ofertas culturales²³. Cabe decir que esta medición monetaria de los valores que emanan del patrimonio cultural es considerada por muchos como escandalosa, o al menos, bastante inconveniente, pues se piensa que degrada el sentido de la creación artística y del disfrute estético. Sin embargo, debemos señalar una ventaja esencial de estos procedimientos, como es la asimilación de todos estos valores en una misma regla

²³ A este respecto podemos citar algunos ensayos de metodología similar como Martin (1994), Chambers *et al.* (1998), Cuccia y Signorello (2000), Santagata y Signorello (2000) y Sanz (2004).

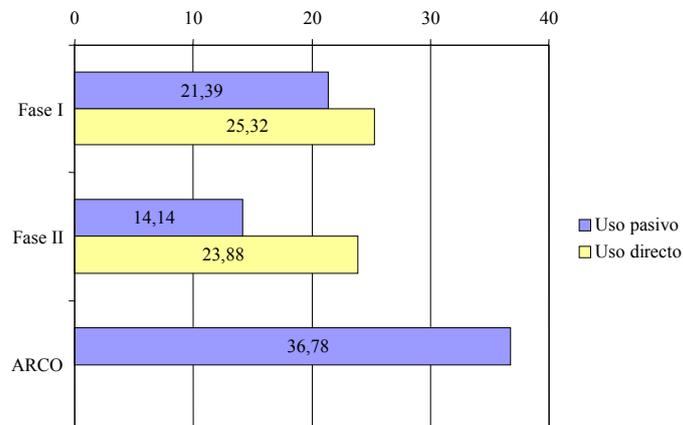
cuantitativa, *el dinero*, que constituye un patrón de referencia comprensible por todos, a la vez que una mercancía sustitutiva por otros bienes u otras opciones. De este modo, así se trate de los costes imputados al uso del patrimonio cultural, o de la disposición a pagar por los valores intangibles, o del conjunto de gastos derivados de las actividades ligadas; todas estas partidas pueden reducirse a una misma escala de medida como son las unidades monetarias.

Las valoraciones realizadas han tenido dos enfoques: uno como consumo cultural del momento de la visita (*precios de entrada*) y otro como donación voluntaria individual para la conservación y mantenimiento de las actividades del museo (*cuota de financiación*). De esta forma, en el caso del precio de entrada al museo, la valoración ha sido de 4 euros, lo que es indicativo de un alto aprecio de los ciudadanos por este tipo de bien o servicio cultural, vanguardista y escaso dentro de la oferta museística de Valladolid. En cuanto al enfoque de la contribución voluntaria, la Tabla 10 y el Gráfico 8 muestran de forma comparada los principales resultados, observándose que una vez pasado el efecto *impacto de la inauguración* del museo, las disposiciones a pagar se reducen ligeramente en el caso de los usuarios o visitantes al museo y algo más entre los ciudadanos de Valladolid. Por otra parte, se confirma la suposición de que el público experto o interesado tiene una disposición al pago muy superior a la del resto de los individuos (36,78 euros), que supera en un 160% la DAP de uso pasivo registrada por los ciudadanos de Valladolid en ese mismo año.

Tabla 10. Estimaciones generales de la DAP

	Fase I		Fase II		ARCO	
	n	DAP media	n	DAP media	n	DAP media
Uso pasivo	766	21,39	588	14,14	287	36,78
Uso directo	485	25,32	437	23,88	-	-

Gráfico 8. Estimaciones generales de la DAP



En definitiva, los resultados de este trabajo, ponen de manifiesto tres hechos importantes: el elevado precio de entrada que el visitante potencial estaría dispuesto a pagar, ya que es superior al del resto de los museos de Valladolid e incluso al de varios museos de arte moderno españoles; el hecho de que la cuota de donación estimada sea próxima a la aportación habitual de los amigos de los museos en nuestro país, teniendo en cuenta el corto periodo de funcionamiento del museo y las peculiaridades de la propiedad de la colección y de la gestión; y, por último, la elevada DAP que presenta el público interesado. Esto revela sin duda alguna, un interés por las manifestaciones artísticas de vanguardia, sensibilidad y emoción por la contemplación del arte en vivo; y, por otra parte, la utilidad de la aplicación de estas técnicas de valoración económica a los bienes culturales.

Bibliografía

- Abad, M.J., Fajardo, M. y Torrente, V. (Coord.) (1995): *Arte en España 1918-1994 en la Colección de Arte Contemporáneo*, Alianza Editorial, Madrid.
- An, M.Y. y Ayala, R.A. (1996): “A Simple Algorithm for Nonparametric Estimation of Distribution Functions with Arbitrarily Grouped Data”, Working Paper 9602, *Department of Economics, Duke University*, Durham.
- Benhamou, F. (1996): *L'économie de la culture*, Editions La Découverte, Paris.
- Blaug, M. (2001): “Where are we now on Cultural Economics?”, *Journal of Economic Surveys*, 15 (2), pp. 123-143.
- Chambers, C.M., Chambers, P.E. y Whitehead, J.C. (1998): “Contingent Valuation of Quasi-Public Goods: Validity, Reliability, and Application to Valuing a Historic Site”, *Public Finance Review*, 26 (2), pp. 137-154.

- Champ, P.A. y Bishop, R.C. (2001): "Donation Payment Mechanisms and Contingent Valuation: An Empirical Study of Hypothetical Bias", *Environmental and Resource Economics*, 19 (4), pp. 383-402.
- Cuadras, C.M., Echeverría, B., Mateo, J. y Sánchez, P. (1984): *Fundamentos de Estadística. Aplicación a las Ciencias Humanas*, P.P.U., Barcelona.
- Cuccia, T. y Signorello, G. (2000): "A Contingent Valuation Study of Willingness to Pay for Visiting a City of Art: The Case Study of Noto", *11th International Conference on Cultural Economics*, Minneapolis.
- Frey, B.S. (2000): *Arts and Economics: Analysis and Cultural Policy*, Springer-Verlag, Berlín.
- Herrero, L.C. (2001): "Economía del Patrimonio Histórico", *Información Comercial Española*, 792, pp. 151-168.
- Herrero, L.C., Bedate, A. y Sanz, J.A. (2004): "Valoración económica de bienes públicos relativos al patrimonio cultural de Castilla y León: metodología y resultados", *Revista de Investigación Económica y Social del Castilla y León*, 6.
- Krström, B. (1997): "Spike Models in Contingent Valuation", *American Journal of Agricultural Economics*, 79 (3), pp. 1.013-1.023.
- Martin, F. (1994): "Determining the Size of Museum Subsidies", *Journal of Cultural Economics*, 18, pp. 255-270.
- Mitchell, R.C. y Carson, R.T. (1989): *Using Survey to Value Public Goods: the Contingent Valuation Method*, John Hopkins University Press for Resources for the Future, Baltimore.
- Riera, P. (1994): *Manual de Valoración Contingente*, Instituto de Estudios Fiscales, Madrid.
- Santagata, W. y Signorello, G. (2000): "Contingent Valuation of a Cultural Public Good and Policy Design: The Case of «Napoli Musei Aperti»", *Journal of Cultural Economics*, 24 (3), pp. 181-204.
- Sanz, J.A. (2004): *Valoración económica del patrimonio cultural*, Trea, Gijón.
- Sanz, J.A. y Castro, R. (2002): "Estudio de la no disposición a pagar por bienes culturales", *Mercurio. Revista de Economía y Empresa*, 4, pp. 255-268.
- Sanz, J.A., Herrero, L.C. y Bedate, A. (2003): "Contingent Valuation and Semiparametric Methods. A case study of the National Museum of Sculpture in Valladolid, Spain", *Journal of Cultural Economics*, 27 (3-4), pp. 241-257.
- Throsby, D. (1994): "The Production and Consumption of the Arts: A View of Cultural Economics", *Journal of Economic Literature*, XXXII, pp. 1-29.
- Throsby, D. (2001): *Economics y Culture*, Cambridge University Press, Cambridge.
- Torrente, V., Fernández, O. y Del Río, V. (Coord.) (2002): *Catálogo del Museo Patio Herreriano de Arte Contemporáneo Español*, Fundación Museo Patio Herreriano de Arte Contemporáneo Español de Valladolid, Valladolid.
- Towse, R. (ed.) (1997): *Cultural economics: the arts, the heritage and the media industries*, Volúmenes I y II, Edward Elgar, Cheltenham.

JUICIOS INFLUYENTES EN EL PROCESO ANALÍTICO JERÁRQUICO (AHP)*

Altuzarra Casas, Alfredo
e-mail: altuzarr@unizar.es

Moreno Jiménez, José María
e-mail: moreno@unizar.es

Salvador Figueras, Manuel
e-mail: salvador@unizar.es

Departamento de Estructura e Historia Económica y Economía Pública
Grupo de Decisión Multicriterio (GDMZ), UNIVERSIDAD DE ZARAGOZA

Área temática: Métodos Cuantitativos.

Resumen

Un aspecto fundamental en la identificación de los patrones de comportamiento que guían la resolución científica de los problemas bajo el constructivismo cognitivo característico de la racionalidad procedimental multicriterio, es la determinación de los puntos críticos y los elementos de mayor influencia en los modelos decisionales utilizados. En este trabajo se presentan, siguiendo un enfoque bayesiano en el modelo estocástico tradicionalmente empleado, diferentes medidas para analizar la influencia de los juicios en la obtención de las prioridades alcanzadas mediante el proceso analítico jerárquico (AHP). Estas medidas abordan tanto la resolución de los problemas tipo α (P. α), selección de la mejor alternativa, como los problemas tipo γ (P. γ), ordenación de todas las alternativas.

Palabras clave: AHP, Influencia, Prioridades, Bayesiano.

* Este trabajo ha sido parcialmente sufragado por el proyecto del Ministerio de Educación y Ciencia “Herramientas Decisionales para la e-cognocracia” (ref. TSI2005-02511)

1. Introducción

El Proceso Analítico Jerárquico (AHP) es una técnica de decisión multicriterio discreta que permite obtener, en una escala de razón, las prioridades asociadas a un conjunto de alternativas que van a ser comparadas. Desarrollada por Saaty (1977, 1980), esta técnica aborda la resolución de problemas multicriterio, multientorno y multiactores, incorporando en el modelo los aspectos subjetivos y la incertidumbre inherente a todo Proceso de Toma de Decisiones.

AHP consta de cuatro etapas: modelización, valoración, priorización y síntesis. En la tercera etapa, se calculan, a partir de las matrices de comparaciones pareadas emitidas por el decisor, los vectores de prioridades locales que determinan la importancia relativa de los elementos de un nivel respecto a un nodo del nivel superior.

Para calcular el vector de prioridades se han propuesto en la literatura distintos métodos. Saaty (1980) propuso inicialmente el Eigenvector Method (EVM), o del Autovector Principal por la derecha, el más utilizado en un principio. En los últimos años ha tenido gran auge la utilización del Método de la Media Geométrica por Filas, Row Geometric Mean (RGM), (Saaty, 1980; Crawford y Williams, 1985; Aguarón y Moreno-Jiménez, 2000). En este trabajo se utiliza un procedimiento de priorización distinto, basado en técnicas bayesianas. Para ello se supone un modelo multiplicativo con errores lognormales en el que las estimaciones del vector de prioridades se obtienen a partir de las distribuciones a posteriori de los parámetros correspondientes. Este modelo, mediante una transformación logarítmica, se convierte en un modelo lineal de regresión.

Por otra parte, uno de los tópicos más estudiados en los modelos de regresión es la existencia y detección de atípicos (outliers) y observaciones influyentes que puedan distorsionar la estimación de los parámetros del modelo. Dado que el problema de estimación de los vectores de prioridades es un problema de regresión, resulta interesante el tratamiento de este tópico dentro del Proceso Analítico Jerárquico.

En este trabajo se analizan diferentes medidas de influencia utilizadas habitualmente en la literatura estadística y se muestra cómo se aplicarían al ámbito de AHP. Dicho estudio proporcionará un mejor conocimiento del proceso de toma de decisiones al permitir identificar, en particular, cuáles de los juicios emitidos son los juicios críticos responsables de las decisiones tomadas y cuáles no.

Las medidas de influencia utilizadas tradicionalmente no están orientadas a problemas que, dentro del ámbito de la Decisión Multicriterio, son muy relevantes, como es el caso de los problemas P. α (selección de la mejor alternativa) o los problemas P. γ (ordenación de la totalidad de las alternativas) Roy (1985). Por ello se proponen medidas de influencia relativas a estos dos problemas.

El trabajo se ha estructurado como sigue: en la siguiente sección se muestra el modelo bayesiano utilizado, en la Sección 3 se recogen las diferentes medidas de influencia utilizadas habitualmente en la literatura estadística y se muestra su aplicación al ámbito de AHP. La Sección 4 se dedica a presentar nuevas medidas de influencia, adaptadas, fundamentalmente, a problemas de tipo P. α ., mientras que en la Sección 5 se aplican estas medidas a un ejemplo concreto. Finalmente en la Sección 6 se comentan las conclusiones y se indican posibles extensiones.

2. Modelo bayesiano propuesto

Sea un decisor **D** que debe ordenar/seleccionar un conjunto compuesto por n alternativas $\mathbf{H} = \{A_1, A_2, \dots, A_n\}$ tomando como base un criterio de comparación C . Para ello, y siguiendo la metodología AHP emite un conjunto de juicios que dan como resultado una matriz A de comparaciones pareadas.

El objetivo del problema consiste en determinar, a partir de dicha información, un vector de prioridades $\mathbf{w} = (w_1, \dots, w_n)$ con $w_i \geq 0$ $i = 1, \dots, n$, donde w_i represente la prioridad relativa de la alternativa A_i respecto al criterio C .

La aproximación seguida en este trabajo se encuadra dentro del denominado AHP estocástico y toma como punto de partida el modelo multiplicativo dado por:

$$a_{ij} = \frac{w_i}{w_j} e_{ij} \quad (2.1)$$

donde e_{ij} es el error cometido al comparar las alternativas A_i y A_j .

Si en este modelo se toman logaritmos y se denota:

$$Y_{ij} = \log a_{ij} \quad \mu_i = \log w_i \quad \varepsilon_{ij} = \log e_{ij} \quad i, j \in \{1, \dots, n\} \quad i < j$$

la expresión (2.1) queda de la forma:

$$Y_{ij} = \mu_i - \mu_j + \varepsilon_{ij} \quad (2.2)$$

Además se supone que los errores ε_{ij} siguen distribuciones $N(0, \sigma^2)$ independientes.

Este modelo ha sido ampliamente utilizado en la literatura (De Jong, 1984; Crawford y Williams, 1985; Fichtner, 1986; Alho y Kangas, 1997). Desde un punto de vista bayesiano su tratamiento ha sido menor y generalmente centrado en el caso de la decisión en grupo (Alho y otros, 1996; Alho y Kangas, 1997; Leskinen y Kangas, 1998; Basak, 1998 o Altuzarra y otros, 2006).

Para evitar problemas de identificabilidad se toma $\mu_n = 0$ (equivalentemente $w_n = 1$), esto es, se fija la última alternativa (A_n) como la de referencia. Además, se supondrá que la varianza, σ^2 , es conocida.

2.1. Distribución a priori

La distribución a priori de las prioridades μ_i $i = 1, \dots, n-1$ viene dada por:

$$\mu_i \sim N(0, \lambda\sigma^2) \quad \forall i = 1, \dots, n-1 \text{ independientes con } \lambda > 0$$

Esto significa suponer que, a priori, la creencia acerca de las alternativas es que todas tienen la misma prioridad $w_1 = \dots = w_n = 1$. Además, cuanto mayor es λ , más difusa es la distribución a priori. En el límite, si $\lambda \rightarrow \infty$ la distribución a priori para cada μ_i se hace impropia y es la distribución de referencia no informativa para $\theta = (\mu_1, \dots, \mu_{n-1})'$.

Sea $\mathbf{Y}_N = (Y_{i_1, j_1}, \dots, Y_{i_N, j_N})'$ el vector de juicios emitidos, donde se supondrá, sin pérdida de generalidad, que $1 \leq i_k < j_k \leq n \quad \forall k = 1, \dots, N$, siendo N el número de comparaciones realizadas por el decisor, cumpliendo $1 \leq N \leq \frac{n(n-1)}{2}$. Si $N = \frac{n(n-1)}{2}$ la información proporcionada por el decisor \mathbf{D} es completa. En caso contrario no se han realizado todas las comparaciones entre las alternativas y la información disponible es incompleta.

Sean $\boldsymbol{\theta} = (\mu_1, \dots, \mu_{n-1})'$ el vector del logaritmo de las prioridades y $\boldsymbol{\varepsilon}_N = (\varepsilon_{i_1, j_1}, \dots, \varepsilon_{i_N, j_N})'$ el vector de los errores.

El modelo puede expresarse en forma matricial como:

$$\mathbf{Y}_N = \mathbf{X}_N \boldsymbol{\theta} + \boldsymbol{\varepsilon}_N \quad \text{con } \boldsymbol{\varepsilon}_N \sim N_N(\mathbf{0}_N, \sigma^2 \mathbf{I}_N) \quad (2.3)$$

$$\boldsymbol{\theta} \sim N_{n-1}(\mathbf{0}_{n-1}, \lambda\sigma^2 \mathbf{I}_{n-1}) \quad (2.4)$$

donde $\mathbf{X}_N = (x_{kh})$ es una matriz de dimensión $(N \times n-1)$ de forma que:

$$x_{k,h} = \begin{cases} 1 & h = i_k \neq n \\ -1 & h = j_k \neq n \\ 0 & \text{resto} \end{cases} \quad (2.5)$$

Cada comparación realizada se representa mediante una fila de la matriz \mathbf{X}_N . Dicha fila contiene, una sola vez, los valores 1 y -1 en las columnas correspondientes a las alternativas comparadas, excepto que una de ellas sea la de referencia. En ese caso, si una de las alternativas comparadas es la de referencia, el juicio emitido será de la forma Y_{in} con $i < n$ y su fila tan sólo contendrá el valor 1 en la columna i -ésima.

2.2. Distribución a posteriori

Sabiendo que de (2.3) se sigue que $\mathbf{Y}_N | \boldsymbol{\theta} \sim N_N(\mathbf{X}_N \boldsymbol{\theta}, \sigma^2 \mathbf{I}_N)$ y que la distribución a priori (2.4) es conjugada se obtiene, aplicando el Teorema de Bayes, la distribución a posteriori del vector de parámetros $\boldsymbol{\theta}$:

$$\boldsymbol{\theta} | (\mathbf{X}_N, \mathbf{Y}_N) \sim N_{n-1} \left(\left(\mathbf{X}'_N \mathbf{X}_N + \frac{1}{\lambda} \mathbf{I}_{n-1} \right)^{-1} \mathbf{X}'_N \mathbf{Y}_N, \sigma^2 \left(\mathbf{X}'_N \mathbf{X}_N + \frac{1}{\lambda} \mathbf{I}_{n-1} \right)^{-1} \right) \quad (2.6)$$

En particular si $\lambda \rightarrow \infty$ se tiene que:

$$\boldsymbol{\theta} | (\mathbf{X}_N, \mathbf{Y}_N) \sim N_{n-1} \left((\mathbf{X}'_N \mathbf{X}_N)^{-1} \mathbf{X}'_N \mathbf{Y}_N, \sigma^2 (\mathbf{X}'_N \mathbf{X}_N)^{-1} \right) \quad (2.7)$$

A partir de estas distribuciones es posible realizar inferencias acerca del vector de parámetros $\boldsymbol{\theta}$ o de transformaciones suyas como, el vector de prioridades \mathbf{w} .

3. Estudio de la influencia

Según Cook y Weisberg (1982) “*el estudio de la influencia es el estudio de la dependencia de las conclusiones e inferencias según diversos aspectos de la formulación de un problema*”. Dicho estudio proporciona, en particular, información acerca de la fiabilidad que tienen las conclusiones y ayuda a conocer la dependencia de éstas respecto al modelo asumido.

Así, dentro del contexto de los modelos de regresión, Bradlow y Zaslavsky (1997) definen la influencia de un observación en la estimación puntual de un parámetro, como “*la diferencia entre la estimación de dicho parámetro basada en la totalidad de los datos y la nueva estimación producida cuando se elimina dicho juicio*”.

Mediante este tipo de análisis se estudia, en particular, la bondad del ajuste del modelo utilizado a los datos observados así como el grado de robustez del proceso de estimación seguido, todo lo cual redundará en una mayor calidad del proceso de inferencia seguido en el análisis.

Dentro del campo bayesiano, uno de los primeros trabajos en aplicar ideas de influencia fue el de Johnson y Geisser (1983), quienes presentan una aproximación basada en el estudio de las densidades predictivas a posteriori, utilizando las discrepancias entre las predicciones obtenidas con todo el conjunto de datos y las obtenidas con algún subconjunto de la muestra inicial. Por su parte Pettit y Smith (1985) definen y calculan diversas medidas de influencia para modelos lineales basadas en un análisis bayesiano del problema. Guttman y Peña (1988) proponen un marco unificado para detectar observaciones influyentes y atípicos (outliers) mediante el uso de la distancia de Kullback-Leibler, aplicándolo al caso de modelos lineales.

En los últimos años, Bayarri y Morales (2003) hacen una amplia revisión de trabajos anteriores dedicados a la detección de observaciones anómalas.

Ya en el ámbito de la Decisión Multicriterio, existen otras técnicas relacionadas con el estudio de observaciones influyentes como son los índices de estabilidad (Aguarón y Moreno, 2000) o el análisis de sensibilidad inverso (Aguarón y Moreno, 2000), que permiten dar una idea de la estabilidad de las soluciones obtenidas calculando, para cada observación o conjunto de observaciones, intervalos o regiones entre los cuales puedan tomar valores sin que cambien significativamente los resultados obtenidos.

A continuación se revisan algunas de las medidas de influencia usadas habitualmente en la literatura estadística para modelos de regresión.

3.1. Medidas de influencia

Sea un modelo lineal $\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$ donde:

\mathbf{Y} representa la variable respuesta (vector $n \times 1$)

$\boldsymbol{\beta}$ representa los p parámetros del modelo (vector $p \times 1$)

\mathbf{X} representa la matriz de datos (vector $n \times p$)

Para realizar un estudio acerca de las observaciones influyentes hay dos elementos que juegan un papel fundamental:

- Los residuos, que se calculan, obviamente, como la diferencia entre el valor estimado y el verdadero valor; así: $e_i = \hat{Y}_i - Y_i$

- La matriz \mathbf{H} o matriz de “apalancamiento” (leverage); también denominada como matriz “sombrero” (hat) que viene dada por:

$$\mathbf{H} = \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}' \quad (3.1)$$

Esta matriz desempeña un papel muy importante en el estudio de las observaciones influyentes por varias razones: en primer lugar, porque la matriz \mathbf{H} puede verse como la transformación lineal que proyecta ortogonalmente cualquier vector n-dimensional en el espacio de variables construido por las columnas de \mathbf{X} . Por otra parte, los elementos de la diagonal principal de \mathbf{H} (h_{ii}) describen lo lejos que está cada una de las observaciones individuales respecto al centroide calculado para todos los datos, tomando como espacio de variables independientes X , es decir, que h_{ii} mide la distancia entre x_i y \bar{x} .

Por lo tanto, si una observación tiene un valor de h_{ii} alto, querrá indicar, al menos, que dicha observación está alejada del resto de los datos. Por otra parte, un atípico será aquel punto que no se ajuste bien al modelo estudiado, es decir, que tenga un residuo inusualmente alto.

A partir de estos elementos se han propuesto diversas medidas de influencia, algunas de las cuales se muestran a continuación.

- DFBETAS (Influencia en los coeficientes de la regresión)

Este estadístico estudia cuánto van a cambiar las estimaciones de los coeficientes de la regresión cuando se elimina una observación.. Para el cálculo de la influencia de la observación i-ésima, se parte de:

$$DFBETA_i = \mathbf{b} - \mathbf{b}_{(i)} = \frac{(\mathbf{X}'\mathbf{X})\mathbf{x}_i'e_i}{1 - h_{ii}} \quad (3.2)$$

donde $\mathbf{b} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y}$ es la estimación del vector $\boldsymbol{\beta}$ con todas las observaciones y $\mathbf{b}_{(i)}$ es esa misma estimación, pero habiendo eliminado la observación i-ésima, es decir eliminando la fila i-ésima de la matriz \mathbf{X} .

En particular, si se quiere calcular la influencia de la observación i-ésima en el coeficiente β_j , resulta:

$$DFBETA_{ij} = \mathbf{b}_j - \mathbf{b}_{(i),j} = \frac{a_{ji}e_i}{1 - h_{ii}} \quad (3.3)$$

donde a_{ji} es el elemento j-ésimo del vector $(\mathbf{X}'\mathbf{X})^{-1}\mathbf{x}_i$.

Esta medida se estandariza para evitar problemas de cambios de escala y, usando el verdadero valor de la varianza σ^2 , o estimándola mediante $s_{(i)}^2$ (cuasi-varianza muestral, pero eliminando el dato i-ésimo), resulta:

$$DFBETAS_{ij} = \frac{a_{ji}}{\sigma(1-h_{ii})\sqrt{q_{jj}}} e_i = \frac{a_{ji}}{\sqrt{q_{jj}(1-h_{ii})}} e_i^* \quad (3.4)$$

donde $(\mathbf{X}'\mathbf{X})^{-1} = (q_{ij})$ y e_i^* es el residuo estudentizado:
$$e_i^* = \frac{e_i}{s_{(i)}\sqrt{(1-h_{ii})}}$$

Esta medida de influencia ($DFBETAS_{ij}$) proporciona el número de errores estándar que cambia el coeficiente j -ésimo si se elimina la observación i -ésima. Lógicamente, un valor alto indicará que la observación i -ésima ejerce una influencia notable en la estimación del coeficiente i -ésimo.

- DFFITS (Influencia en la predicción)

Este estadístico estudia cómo va a cambiar la predicción \hat{y}_i (para el valor x_i) si la observación i -ésima se eliminara.

$$DFFIT_i = \hat{y}_i - \hat{y}_{(i),i} = \mathbf{x}_i [\mathbf{b} - \mathbf{b}_{(i)}] = \frac{h_{ii} e_i}{1-h_{ii}} \quad (3.5)$$

donde $\hat{y}_{(i),i}$ denota la predicción del valor de y_i basado en las estimaciones del modelo utilizando el resto de observaciones.

Esta medida también se estandariza para evitar problemas de cambios de escala y resulta, usando que $\text{Var}(\hat{y}_i) = \sigma^2 h_{ii}$ (tomando el valor σ^2 , o su estimación $s_{(i)}^2$):

$$DFFITS_i = \frac{\sqrt{h_{ii}} e_i}{s_{(i)}(1-h_{ii})} = \sqrt{\frac{h_{ii}}{1-h_{ii}}} e_i^* \quad (3.6)$$

Esta medida representa el número de errores estándar en los que el valor estimado \hat{y}_i cambiaría si se eliminara la observación i -ésima.

Otra forma de interpretar $DFFITS_i$ es como el valor del residuo estudentizado, corregido (puede tanto aumentar como disminuir) por la medida del “apalancamiento” (de la matriz “sombrero” \mathbf{H}) $[h_{ii}/(1-h_{ii})]^{1/2}$.

Si la observación x_i es atípica, tendrá un gran residuo. Si además su valor de h_{ii} es alto, entonces será un punto con un valor de $DFFITS_i$ muy alto. Si por el contrario el valor del residuo es bajo, aunque el valor del h_{ii} sea alto, tendrá poco efecto en el valor final del estadístico $DFFITS_i$; y si el valor de h_{ii} es muy pequeño, entonces el resultado final se modera sustancialmente.

3.2. Aplicación al ámbito de AHP

En esta sección se aplican las medidas de influencia descritas en la sección anterior, al caso particular de AHP, considerando el modelo dado por las expresiones (2.3) a (2.5).

El propósito de estas medidas es cuantificar la influencia que tienen las N comparaciones pareadas efectuadas, denotadas como $\mathbf{Y}_N = (Y_{i_1, j_1}, \dots, Y_{i_N, j_N})'$, sobre el vector final de prioridades, denotado por $\boldsymbol{\theta} = (\mu_1, \dots, \mu_{n-1})'$.

Se van a mostrar los cálculos realizados para el caso de que el decisor haya emitido todos los juicios, es decir, cuando $N = \frac{n(n-1)}{2}$. A partir de la expresión (2.5) se llega a

que: $(\mathbf{X}'_N \mathbf{X}_N)^{-1} = \frac{1}{n}(\mathbf{I}_{n-1} + \mathbf{J}_{n-1})$, y, por lo tanto:

$$h_{ii} = \frac{2}{n} \quad \forall i = 1, \dots, N \quad (3.7)$$

Es decir, que en el caso de que se conozcan todas las observaciones se obtiene que todas ellas tienen el mismo “apalancamiento” o la misma distancia al centroide.

Además se obtienen las siguientes expresiones:

$$a_{ji} = \begin{cases} \frac{2}{n} & j = \alpha_i \text{ y } \beta_i = n \\ \frac{1}{n} & j = \alpha_i \text{ y } \beta_i \neq n \\ -\frac{1}{n} & j = \beta_i \\ 0 & \text{otro caso} \end{cases} \quad (3.8)$$

$$\mathbf{e}_i = Y_{\alpha_i, \beta_i} - \hat{Y}_{\alpha_i, \beta_i} = Y_{\alpha_i, \beta_i} - \left[\frac{1}{n} \sum_{t=1}^n (Y_{\alpha_i, t} - Y_{\beta_i, t}) \right] \quad (3.9)$$

$$\mathbf{e}_i^* = \frac{1}{\sigma} \sqrt{\frac{n}{n-2}} \mathbf{e}_i \quad (3.10)$$

Llamando $K = \frac{1}{\sqrt{2(n-2)}}$, el estadístico $DFBETAS_{ij}$, para el caso de un decisor que

haya emitido todos los juicios, queda:

$$DFBETAS_{ij} = \begin{cases} 2K e_i^* & j = \alpha_i \text{ y } \beta_i = n \\ K e_i^* & j = \alpha_i \text{ y } \beta_i \neq n \\ -K e_i^* & j = \beta_i \\ 0 & \text{otro caso} \end{cases} \quad (3.11)$$

De este resultado se pueden destacar las siguientes conclusiones:

- Dado un coeficiente cualquiera, por ejemplo μ_j , la influencia de las observaciones en esa prioridad va a depender únicamente del residuo estandarizado de los juicios del tipo $Y_{jt} \forall t$.

- Dado un juicio cualquiera, (sin pérdida de generalidad se supone Y_{13}), este juicio solo va a ejercer influencia en la estimación de los parámetros μ_1 y μ_3 . Además, los juicios del tipo Y_{1n} ejercerán doble influencia sobre μ_1 que los juicios del tipo $Y_{1t} \forall t$.

Para el estadístico DFFITS, aplicado al caso de matrices completas, se obtiene:

$$DFFITS_i = \sqrt{\frac{2/n}{1-2/n}} e_i^* = \sqrt{\frac{2}{n-2}} e_i^* \quad (3.12)$$

Si se tiene en cuenta la constante K, calculada para el estadístico DFBETAS:

$$DFFITS_i = 2K e_i^*$$

Y, por lo tanto, este estadístico, recordando la expresión (3.11) es idéntico al $DFBETAS_{ij}$ para el caso $j = \alpha_i$ y $\beta_i = n$.

4. Análisis bayesiano de la influencia

El gran inconveniente de las medidas de influencia analizadas en la sección anterior radica, en nuestra opinión, en la falta de una calibración de las mismas, es decir, del establecimiento de límites que indiquen si una observación es significativamente influyente o no. Además, y para el caso particular de AHP, aparece un problema adicional, ya que la estructura de la matriz de diseño (\mathbf{X}) hace que las medidas de apalancamiento, basadas en la matriz \mathbf{H} , pierden parte de su interés debido a que cuando se trabaja con matrices completas, \mathbf{H} es siempre la misma. Por otra parte, en AHP el interés no se centra en la predicción de futuras observaciones, como ocurre, por ejemplo, con el estadístico DFFITS.

Además, las medidas de influencia analizadas no dirigen directamente su atención a problemas que, en Teoría de la Decisión son muy relevantes como, por ejemplo, los

problemas P. α y P. γ . En particular, en AHP, tras la obtención del vector de prioridades a partir de una matriz de juicios, uno de los aspectos que se plantea es el de determinar si el decisor ha sido consistente a la hora de emitir sus juicios, es decir, hasta qué punto el vector \mathbf{w} refleja la opinión real del decisor (si se verifica la transitividad en el proceso de emisión de juicios).

Por lo tanto, el estudio de la influencia dentro del ámbito de AHP se puede aplicar a dos atributos, la priorización (problemas P. α y P. γ) y la consistencia. Este trabajo se va a centrar en analizar la influencia que ejerce la emisión de un juicio o de un conjunto de juicios en las prioridades obtenidas, restringiendo el estudio al caso de analizar un problema de tipo P. α . En futuros trabajos se analizarán los otros tipos de influencia.

Se presentan varios procedimientos para determinar tanto las observaciones influyentes, como los intervalos de estabilidad de los juicios; es decir, el conjunto de valores de los juicios emitidos que no harían cambiar las conclusiones del problema. Estos procedimientos se aplicarán a un ejemplo desarrollado por Saaty y Kearns (1985).

4.1 Análisis bayesiano de la influencia para problemas de tipo P. α

El estudio de la influencia aplicado a un problema P. α se puede orientar de dos formas:

- La primera consiste en caracterizar aquellas observaciones influyentes en la ordenación obtenida, fijándose, únicamente, en la alternativa que ocupa el primer lugar.

- En la segunda se busca caracterizar las observaciones influyentes en cuanto al cociente de dos prioridades cualesquiera $\frac{w_i}{w_j}$, o equivalentemente, a la diferencia de las

log-prioridades $\mu_i - \mu_j$. Evidentemente, al tratar un problema P. α , es lógico encaminar su estudio sólo al caso donde el subíndice i corresponda a la alternativa seleccionada en primer lugar, $i = (1)$.

A continuación se presentan los procedimientos que reflejan estas dos orientaciones.

i) Contraste χ^2 (P. α). Si sólo se pretende estudiar la ordenación obtenida, entonces la cantidad en la que se centra el interés será la distribución a posteriori de la alternativa más preferida $A_{(1)}$. Por lo tanto, habrá que comparar la distribución de $A_{(1)} \mid (\mathbf{X}_N, \mathbf{Y}_N)$ con la de $A_{(1)} \mid (\mathbf{X}_N, \mathbf{Y}_N) \cup (\mathbf{X}_{\text{nuevo}}, \mathbf{Y}_{\text{nuevo}})$. Dado que ambas son discretas, su comparación se puede establecer por medio de la simulación utilizando un contraste tipo χ^2 de homogeneidad para decidir si un juicio o grupo de juicios es influyente o no.

Este método tiene el inconveniente de no posibilitar el cálculo de intervalos ni de regiones de estabilidad.

Bastará con construir una tabla del tipo:

Tabla 4.1. Construcción del contraste $\chi^2(P.\alpha)$

	A_1	A_2	...	A_{n-1}	A_n
$(\mathbf{X}_N, \mathbf{Y}_N)$	$P[A_{(1)} = A_i (\mathbf{X}_N, \mathbf{Y}_N)]$				
$(\mathbf{X}_N, \mathbf{Y}_N) \cup (\mathbf{X}_{nuevo}, \mathbf{Y}_{nuevo})$	$P[A_{(1)} = A_i (\mathbf{X}_N, \mathbf{Y}_N) \cup (\mathbf{X}_{nuevo}, \mathbf{Y}_{nuevo})]$				

donde $(\mathbf{X}_{nuevo}, \mathbf{Y}_{nuevo})$ contiene el juicio o grupo de juicios sobre los que se quiere estudiar su influencia. Se procede a comparar las frecuencias obtenidas cuando se dispone de la totalidad de los juicios (N) con las obtenidas si se elimina $(\mathbf{X}_{nuevo}, \mathbf{Y}_{nuevo})$. Después, se calcula el estadístico χ^2 correspondiente y se decide si el o los juicios contenidos en $(\mathbf{X}_{nuevo}, \mathbf{Y}_{nuevo})$ son influyentes para un nivel de significación $(1 - \alpha)$ dado. También se puede calcular, a partir de estos programas, el conjunto mínimo de juicios para los que se mantiene la concordancia entre las dos distribuciones, la primera con $(\mathbf{X}_{nuevo}, \mathbf{Y}_{nuevo})$, y la segunda sin $(\mathbf{X}_{nuevo}, \mathbf{Y}_{nuevo})$, dado un nivel de credibilidad $(1 - \alpha)$.

ii) Intervalos de estabilidad (P. α). Si se quieren calcular intervalos de estabilidad habría que recurrir a analizar distribuciones de comparaciones pareadas entre prioridades de alternativas, tipo $\mu_i - \mu_j | (\mathbf{X}_N, \mathbf{Y}_N)$, y de esta forma se caracterizarían las observaciones influyentes en la estimación de la diferencia de log-prioridades.

La utilización de regiones HPD para el cálculo de los intervalos se puede hacer de forma genérica, analizando la siguiente expresión:

$$\begin{aligned}
 & P\{f[\mu_i - \mu_j^N | (\mathbf{X}_N, \mathbf{Y}_N) \cup (\mathbf{X}_{nuevo}, \mathbf{Y}_{nuevo})] \geq \\
 & \geq f[m_i^N - m_j^N | (\mathbf{X}_N, \mathbf{Y}_N) \cup (\mathbf{X}_{nuevo}, \mathbf{Y}_{nuevo})] | (\mathbf{X}_N, \mathbf{Y}_N) \cup (\mathbf{X}_{nuevo}, \mathbf{Y}_{nuevo})\} \geq \alpha
 \end{aligned}
 \tag{4.1}$$

En esta expresión se compara la distribución a posteriori de $\mu_i - \mu_j$ cuando se ha incorporado el juicio, o conjunto de juicios, $(\mathbf{X}_{nuevo}, \mathbf{Y}_{nuevo})$, con su estimación dada por la media de su distribución a posteriori $(m_i^N - m_j^N)$. Es decir, se busca saber si la estimación obtenida para $\mu_i - \mu_j$ antes de haber observado $(\mathbf{X}_{nuevo}, \mathbf{Y}_{nuevo})$ se puede considerar atípica, tras observar $(\mathbf{X}_{nuevo}, \mathbf{Y}_{nuevo})$.

En el caso de estudiar la influencia de una sola observación, $Y_{\text{nuevo}} = Y_{k\ell}$, se tiene que, (suponiendo la varianza desconocida), el vector de prioridades sigue una distribución t de Student. De esta forma, el cálculo de la región HPD queda:

$$P[|\mu_i - \mu_j| \geq m_i^N - m_j^N \mid (\mathbf{X}_N, \mathbf{Y}_N) \cup \{Y_{k\ell}\}] \geq \alpha \quad (4.2)$$

Utilizando que $\frac{\mu_i - \mu_j - M^{N+1}}{S^{N+1}} \sim t_{N+1-(n-1)}$, donde:

$$M^{N+1} = m_i^N - m_j^N \quad S^{N+1} = s_{1,N+1} \sqrt{b_{ii}^{N+1} + b_{jj}^{N+1} - 2b_{ij}^{N+1}} \quad \text{siendo } s_{1,N+1} \text{ la estimación de la varianza tras haber realizado } N \text{ comparaciones y siendo } b_{ij}^N = (\mathbf{X}'_N \mathbf{X}_N)^{-1}.$$

Desarrollando la expresión (4.2) es posible calcular un intervalo de estabilidad, compuesto por aquellos valores de Y_{nuevo} que, bajo la distribución de $\mu_i - \mu_j$, tienen una probabilidad de ocurrencia igual a $1-\alpha$. Se obtiene, así, el correspondiente intervalo de estabilidad, dado por:

$$m_k^N - m_\ell^N \pm t_{N-n+2, \frac{\alpha}{2}} s_{1,N+1} \frac{\sqrt{D_{k\ell}^N}}{\rho_{ij,k\ell}^N} \sqrt{1 - (\rho_{ij,k\ell}^N)^2} \quad (4.3)$$

donde $C_{ij,k\ell}^N = \text{Cov}[\mu_i - \mu_j, Y_{k\ell} \mid (\mathbf{X}_N, \mathbf{Y}_N)]$

$e_{k\ell}^N = Y_{k\ell} - (m_k^N - m_\ell^N)$ es el error de predicción tras N juicios

$$\rho_{ij,k\ell}^N = \text{Corr}[\mu_i - \mu_j, Y_{k\ell} \mid (\mathbf{X}_N, \mathbf{Y}_N)] \quad D_{k\ell}^N = 1 + b_{kk}^N + b_{\ell\ell}^N - 2b_{k\ell}^N$$

5. Ejemplo

La metodología expuesta en la sección anterior se ilustra con un ejemplo clásico de la literatura de AHP. El problema consiste en la elección de una casa de acuerdo a ocho criterios, y fue expuesto inicialmente en Saaty y Kearns (1985).

Tabla 5.1. Criterios de elección de una vivienda

	1	2	3	4	5	6	7	8
1	1	5	3	7	6	6	1/3	1/4
2	1/5	1	1/3	5	3	3	1/5	1/7
3	1/3	3	1	6	3	4	6	1/5
4	1/7	1/5	1/6	1	1/3	1/4	1/7	1/8
5	1/6	1/3	1/3	3	1	1/2	1/5	1/6
6	1/6	1/3	1/4	4	2	1	1/5	1/6
7	3	5	1/6	7	5	5	1	1/2
8	4	7	5	8	6	6	2	1

En primer lugar se ha calculado la estimación (media de la distribución a posteriori) del vector normalizado w :

$$w = (0.175, 0.063, 0.149, 0.019, 0.036, 0.042, 0.167, 0.35)$$

Las alternativas quedan ordenadas, según esta estimación, de la siguiente forma:

$$A_8 \geq A_1 \geq A_7 \geq A_3 \geq A_2 \geq A_6 \geq A_5 \geq A_4$$

5.1 Contraste χ^2 (P. α)

A continuación se muestra una tabla con los resultados obtenidos tras realizar este contraste. Se indica el p-valor obtenido para cada uno de los juicios. Se ha utilizado una muestra de tamaño 10000 obtenida a partir de un proceso de simulación como el descrito en la sección 4.1.

Tabla 5.2. p-valores de cada juicio χ^2 (P. α)

Juicios	Y ₁₂	Y ₁₃	Y ₁₄	Y ₁₅	Y ₁₆	Y ₁₇	Y ₁₈
p-valor	1	0.951	0.9996	1	1	0.7926	0.4381
Juicios	Y ₂₃	Y ₂₄	Y ₂₅	Y ₂₆	Y ₂₇	Y ₂₈	Y ₃₄
p-valor	1	1	1	1	1	0.9987	0.9999
Juicios	Y ₃₅	Y ₃₆	Y ₃₇	Y ₃₈	Y ₄₅	Y ₄₆	Y ₄₇
p-valor	0.9999	1	0.5444	0.6092	1	1	0.9998
Juicios	Y ₄₈	Y ₅₆	Y ₅₇	Y ₅₈	Y ₆₇	Y ₆₈	Y ₇₈
p-valor	0.9958	1	1	0.9999	1	1	0.9992

Todos los juicios tienen p-valores muy altos, por lo tanto no hay ningún juicio que pueda afectar la ordenación de la primera alternativa para los niveles de significación utilizados habitualmente.

Incluso el juicio Y₁₈ que compara las alternativas ordenadas en primer y segundo lugar, tampoco es influyente para cambiar dicha ordenación. Además, este es el juicio con un p-valor más pequeño (0.4381).

También se ha calculado el conjunto mínimo de juicios que hace que se mantenga la concordancia entre las distribuciones (con y sin el nuevo juicio). Estos juicios, así como sus correspondientes p-valores aparecen en la siguiente tabla. El nivel de significación que se ha adoptado es del 5%. En total se necesitan 13 juicios, como mínimo, para que no cambie la alternativa que está ordenada en primer lugar.

Tabla 5.3. Conjunto mínimo de juicios. $\chi^2(P,\alpha)$

Juicios	Y ₁₃	Y ₁₆	Y ₁₇	Y ₁₈	Y ₂₅	Y ₂₆	Y ₂₇
p-valor	6.83x10 ⁻⁶	8x10 ⁻¹²	0.0105	7.95x10 ⁻⁷	0	0	0
Juicios	Y ₃₄	Y ₃₇	Y ₃₈	Y ₅₆	Y ₅₈	Y ₆₈	
p-valor	1.23x10 ⁻⁶	8.34x10 ⁻⁷	0.033	1.11x10 ⁻⁶	7.8x10 ⁻⁴	0	

El proceso seguido para el cálculo de este conjunto mínimo es el siguiente: dados los N juicios iniciales, se calculan los p-valores correspondientes a cada uno de dichos N juicios. Se elimina el juicio con mayor p-valor (menor probabilidad de cambio de rango). Se continúa con este proceso hasta que todos los p-valores calculados sean mayores que el nivel de significación adoptado (1- α), siempre que la distribución a posteriori sea propia.

5.2 Intervalos de estabilidad (P. α)

En la Tabla 5.4 se presentan, para cada juicio, sus intervalos de estabilidad, así como el correspondiente intervalo de estabilidad mínimo. El encabezamiento de la tabla representa cada una de las posibles diferencias entre log-prioridades, es decir, entre la ordenada en primer lugar y las restantes. Cada una de las filas se corresponde con cada uno de los juicios, donde se muestran, los extremos inferior y superior de los intervalos de credibilidad. La última columna hace referencia al intervalo mínimo para cada juicio. Existen muchas celdas vacías, indicando que dicho juicio no es influyente para provocar el cambio de rango correspondiente al par de alternativas a las que hace referencia su columna. Por ejemplo, en la primera fila de los juicios, en la que están los intervalos del juicio Y₁₂, sólo aparecen valores numéricos en las prioridades correspondientes a la primera y cuarta columna; es decir, las que hacen referencia al posible cambio de rango entre las alternativas situadas en primera y segunda posición ($\mu_8 - \mu_1$) y entre las alternativas situadas en primera y quinta posición ($\mu_8 - \mu_2$). Los valores que aparecen en ambos intervalos son los mismos (0.1963, 26.8866), queriendo indicar que el juicio Y₁₂ puede variar entre dichos valores sin que se produzca cambio en la ordenación de las prioridades.

En el resto de celdas no aparece valor alguno, lo que es equivalente a indicar que el intervalo de variación sería de la forma (0, ∞).

Únicamente el juicio Y_{37} resulta influyente en la ordenación de las prioridades μ_8 y μ_3 y en las prioridades μ_8 y μ_7 , que corresponden al posible cambio de rango de la prioridad ordenada en primer lugar con las ordenadas en segundo y tercero, respectivamente. El valor de este juicio es $Y_{37} = 6$, que no entra dentro del intervalo de estabilidad calculado (0.0403, 5.5174).

Por otra parte, el tratar con información completa hace que los intervalos calculados sean los mismos para cada juicio, excepto cuando interviene la alternativa de referencia, A_8 . Esto es debido a la forma que tiene la matriz de varianzas-covarianzas de la distribución a posteriori para el caso de disponer de todas las observaciones. Por lo tanto, se puede decir, que en este caso, la influencia que se ejerce es global.

Tabla 5.4. Intervalos de estabilidad ($P.\alpha$)

	$\mu_{(1)} - \mu_{(2)}$ $\mu_8 - \mu_1$	$\mu_{(1)} - \mu_{(3)}$ $\mu_8 - \mu_7$	$\mu_{(1)} - \mu_{(4)}$ $\mu_8 - \mu_3$	$\mu_{(1)} - \mu_{(5)}$ $\mu_8 - \mu_2$	$\mu_{(1)} - \mu_{(6)}$ $\mu_8 - \mu_6$	$\mu_{(1)} - \mu_{(7)}$ $\mu_8 - \mu_5$	$\mu_{(1)} - \mu_{(8)}$ $\mu_8 - \mu_4$	Interv. mínimos
Y_{12}	0.1963			0.1963				0.1963
	26.8866			26.8866				26.8866
Y_{13}	0.0735		0.0735					0.0735
	10.0674		10.0674					10.0674
Y_{14}	0.8406						0.8406	0.8406
	115.1507						115.1507	115.1507
Y_{15}	0.3928					0.3928		0.3928
	53.8122					53.8122		53.8122
Y_{16}	0.3118				0.3118			0.3118
	42.7115				42.7115			42.7115
Y_{17}	0.1310	0.1310						0.1310
	17.9388	17.9388						17.9388
Y_{18}	0.1106	0.0538	0.0538	0.0538	0.0538	0.0538	0.0538	0.1106
	3.5874	7.3739	7.3739	7.3739	7.3739	7.3739	7.3739	3.5874
Y_{23}			0.0389	0.0389				0.0389
			5.3312	5.3312				5.3312
Y_{24}				0.2394			0.2394	0.2394
				32.7935			32.7935	32.7935
Y_{25}				0.1260		0.1260		0.1260
				17.2598		17.2598		17.2598
Y_{26}				0.1000	0.1000			0.1000
				13.6994	13.6994			13.6994
Y_{27}		0.0395		0.0395				0.0395
		5.4143		5.4143				5.4143
Y_{28}	0.0165	0.0165	0.0165	0.0339	0.0165	0.0165	0.0165	0.0339
	2.2619	2.2619	2.2619	1.1004	2.2619	2.2619	2.2619	1.1004
Y_{34}			0.7133				0.7133	0.7133
			97.7156				97.7156	97.7156
Y_{35}			0.3990			0.3990		0.3990
			54.6520			54.6520		54.6520

	$\mu_{(1)} - \mu_{(2)}$ $\mu_8 - \mu_1$	$\mu_{(1)} - \mu_{(3)}$ $\mu_8 - \mu_7$	$\mu_{(1)} - \mu_{(4)}$ $\mu_8 - \mu_3$	$\mu_{(1)} - \mu_{(5)}$ $\mu_8 - \mu_2$	$\mu_{(1)} - \mu_{(6)}$ $\mu_8 - \mu_6$	$\mu_{(1)} - \mu_{(7)}$ $\mu_8 - \mu_5$	$\mu_{(1)} - \mu_{(8)}$ $\mu_8 - \mu_4$	Interv. mínimos
Y ₃₆			0.2877			0.2877		
			39.4116			39.4116		
Y ₃₇			0.0403			0.0403		
			5.5174			5.5174		
Y ₃₈	0.0467	0.0467	0.0961	0.0467	0.0467	0.0467	0.0467	0.0961
	6.4029	6.4029	3.1151	6.4029	6.4029	6.4029	6.4029	3.1151
Y ₄₅						0.0547	0.0547	0.0547
						7.4936	7.4936	7.4936
Y ₄₆					0.0478			0.0478
					6.5462			6.5462
Y ₄₇	0.0092						0.0092	0.0092
	1.2641						1.2641	1.2641
Y ₄₈	0.0036	0.0036	0.0036	0.0036	0.0036	0.0036	0.0074	0.0074
	0.4936	0.4936	0.4936	0.4936	0.4936	0.4936	0.2402	0.2402
Y ₅₆					0.0854	0.0854		
					11.7043	11.7043		
Y ₅₇	0.0186						0.0186	0.0186
	2.5457						2.5457	2.5457
Y ₅₈	0.0074	0.0074	0.0074	0.0074	0.0074	0.0152	0.0074	0.0152
	1.0103	1.0103	1.0103	1.0103	1.0103	0.4915	1.0103	0.4915
Y ₆₇	0.0234					0.0234		
	3.2073					3.2073		
Y ₆₈	0.0093	0.0093	0.0093	0.0093	0.0191	0.0093	0.0093	0.0191
	1.2728	1.2728	1.2728	1.2728	0.6192	1.2728	1.2728	0.6192
Y ₇₈	0.0402	0.0826	0.0402	0.0402	0.0402	0.0402	0.0402	0.0826
	5.5074	2.6794	5.5074	5.5074	5.5074	5.5074	5.5074	2.6794

6. Conclusiones

En este trabajo se han presentado diferentes medidas de influencia utilizadas habitualmente en la literatura estadística y se ha mostrado su utilización dentro del ámbito de AHP, a partir de un modelo multiplicativo para los errores, en el que la obtención del vector de prioridades se realiza mediante técnicas bayesianas.

Como estas medidas no responden a problemas que son habituales dentro del campo de la Teoría de la Decisión se han propuesto nuevas medidas de influencia, como el contraste χ^2 ($P.\alpha$) o la construcción de intervalos de estabilidad para cada uno de los juicios emitidos por el decisor.

Las medidas propuestas se han centrado únicamente en la resolución de un problema de tipo $P.\alpha$, es decir, seleccionar la mejor alternativa. En posteriores trabajos se piensa ampliar este estudio al caso de la resolución de problemas de tipo $P.\gamma$ (ordenación de todas las alternativas), así como analizar la influencia que pueden ejercer los juicios

sobre la consistencia (transitividad) del problema analizado, puesto que éste es un tópico propio de AHP.

Bibliografía

- Aguarón, J.; Moreno-Jiménez, J. M. (2000): Local Stability Intervals In The Analytic Hierarchy Process. *European Journal of Operational Research*, 125, pp.113 – 132.
- Alho, J.; Kangas, J. (1997): Analyzing Uncertainties in Expert's Opinions of Forest Plan Performance. *Forest Science*, 43 (4), pp.521 - 528.
- Alho, J.; Kangas, J.; Kolehmainen, O. (1996): Uncertainty in Expert Predictions of the Ecological Consequences of Forest Plans. *Applied Statistics* 45 (1), pp.1 - 14.
- A. Altuzarra, J.M. Moreno-Jiménez and M. Salvador, A Bayesian Priorization Procedure for AHP-Group Decision Making, Forthcoming in the European Journal of Operational Research (2006). DOI:10.1016/j.ejor.2006.07.025.
- Basak, I. (1998): Probabilistic Judgments Specified Partially in the Analytic Hierarchy Process. *European Journal of Operational Research*, 108, pp.152 - 164.
- Bradlow, E. T.; Zaslavsky, A. M. (1997): Case Influence Analysis in Bayesian Inference. *Journal of Computational and Graphical Statistics*, 6 (3), pp.314 - 331.
- Bayarri, M. J.; Morales, J. (2003): Bayesian Measures of Surprise for Outlier Detection. *Journal of Statistical Planning and Inference*, 111, pp.3 - 22.
- Cook, R. D.; Weisberg, S. (1982): *Residuals and Influence in Regression*. New York, London: Chapman-Hall.
- Crawford, G.; Williams, C. (1985): A Note on the Analysis of Subjective Judgment Matrices. *Journal of Mathematical Psychology*, 29, pp.387 - 405.
- De Jong, P. (1984) : A Statistical Approach to Saaty's Scaling Method for Priorities. *Journal of Mathematical Psychology*, 28 pp.467 – 478.
- Fichtner, J. (1986): On Deriving Priority Vectors from Matrices of Pairwise Comparations. *Socio-Economic Planning Sciences*, 20, pp.399 - 405.
- Guttman, I.; Peña, D. (1988): Outliers and Influence: Evaluation by Posteriors of Parameters in the Linear Model. *Bayesian Statistics*, 3, pp.631 - 640.
- Johnson, W.; Geisser, S. (1983): A Predictive View of the Detection and Characterization of Influential Observations in Regression Analysis. *Journal of the American Statistical Association*, 78 (381), pp.137 - 144.
- Leskinen, P.; Kangas, J. (1998): Analysing Uncertainties of Interval Judgment Data in Multi-Criteria Evaluation of Forest Plans. *Silva Fennica*, 32 (4), pp.363 - 372.
- Pettit, L. I.; Smith, A. F. M. (1985): Outliers and Influential Observations in Linear Models. En *Bayesian Statistics 2* (J. M. Bernardo, M. H. De Groot, A. F. M. Smith D. V. Lindley eds.). 473 – 494. North Holland, Amsterdam.
- Roy, B. (1985): *Méthodologie Multicritère d'aide à la Décision*. Gestión Económica.
- Saaty, T. L. (1977): A Scaling Method For Priorities In Hierarchical Structures. *Journal Of Mathematical Psychology*, 15 (3), pp.234-281.
- Saaty, T. L. (1980): *Multicriteria Decision Making: The Analytic Hierarchy Process*. Mc Graw-Hill, New York. (2ª Impresión 1990, Rsw Pub. Pittsburgh).
- Saaty, T. L.; Kearns, K. P. (1985): *Analytical Planning*. Pergamon Press, New York.

HERRAMIENTAS DE VISUALIZACIÓN GRÁFICA PARA EL ANÁLISIS DE LA EVOLUCIÓN DINÁMICA EN LA E-COGNOCRACIA¹

ALBERTO TURÓN LANUZA
turon@unizar.es

JOSÉ MARÍA MORENO JIMÉNEZ
moreno@unizar.es

Grupo Decisión Multicriterio Zaragoza
FACULTAD DE ECONÓMICAS. UNIVERSIDAD DE ZARAGOZA.

Área temática: Métodos Cuantitativos.

Resumen

La e-cognocracia es un nuevo sistema democrático dirigido a la creación y difusión social del conocimiento relativo a la resolución científica de problemas altamente complejos asociados a la toma de decisiones públicas. Basado en el uso de las técnicas de decisión multicriterio como soporte metodológico, Internet como soporte de comunicación y la democracia como el elemento catalizador de la creación del conocimiento, este sistema de representación democrática estimula el aprendizaje que guía el proceso cognitivo característico del ser humano.

Para evaluar el “aprendizaje” individual y colectivo derivado de su aplicación, en lo que sigue se presentan dos nuevas herramientas de visualización gráfica que ayudarán al análisis visual de la evolución temporal de las preferencias de los usuarios. Estos desarrollos se han aplicado a los resultados obtenidos en un experimento de votación electrónica con dos rondas y discusión intermedia. La decisión considerada se refiere a un problema de localización y los participantes han sido los asistentes al taller sobre “voto electrónico y e-cognocracia” celebrado dentro de las II Jornadas sobre Comercio electrónico celebradas en Zaragoza del 28 de febrero al 2 de marzo de 2007.

Palabras clave: E-cognocracia, E-votación, E-participación, Multicriterio, Visualización Gráfica.

Abstract

E-cognocracy is a new democratic system oriented to the extraction and diffusion of knowledge derived from the scientific resolution of highly complex public decision making problems related to the governance of society. Based on multicriteria decision making techniques as methodological support, Internet as communicative support and democracy as catalytic element of the knowledge creation, this representative system stimulates the learning inherent to the cognitive process that characterises human beings.

To evaluate the individual and collective “learning” derived from its application, in what follows two new visualization tools are introduced in order to aid to the visual analysis of temporal evolution of users’ preferences. Those developments have been applied to the results obtained in an electronic voting experiment with two rounds and intermediate discussion. The decision relates to a location problem and the participants were the audience of a workshop on “electronic voting and e-cognocracy” in the framework of the II Conference on Electronic Commerce that took place in Zaragoza from february 28th to march 2nd 2007.

Key words: E-cognocracy, E-voting, E-participating, Multicriteria, Graphical Visualization.

¹ Este trabajo ha sido parcialmente sufragado por el proyecto del Ministerio de Educación y Ciencia “Herramientas Decisionales para la e-cognocracia” (ref. TSI2005-02511).

1. Introducción

La e-cognocracia (Moreno-Jiménez, 2003, 2004, 2006; Moreno-Jiménez y Polasek, 2003, 2005) es un nuevo sistema democrático dirigido a la creación y difusión social del conocimiento relativo a la resolución científica de problemas altamente complejos planteados en el ámbito de la toma de decisiones públicas relativas al gobierno de la sociedad. Basado en el uso de las técnicas de decisión multicriterio como soporte metodológico, Internet como soporte de comunicación y la democracia como el elemento catalizador de la creación del conocimiento, este sistema de representación democrática estimula el aprendizaje que guía el proceso cognitivo característico del ser humano (Capra, 2000).

Además de solventar algunas de las limitaciones de la democracia tradicional (falta de transparencia, control y participación de los ciudadanos), la e-cognocracia persigue la implicación de todos los ciudadanos en la creación de una sociedad mejor, más culta y formada (Moreno-Jiménez, 2006). En este sentido, se va a aprovechar uno de los modelos de representación más extendidos en el mundo occidental, la democracia, para crear y difundir socialmente el conocimiento derivado de la resolución científica y colectiva de los problemas. No hay que perder de vista una de las ideas esenciales de su propuesta, “No hay democracia sin libertad, ni libertad sin conocimiento” (Moreno-Jiménez, 2003).

Para favorecer la extracción de conocimiento, la discusión a través de herramientas colaborativas, los procesos de negociación entre los actores implicados en la resolución del problema y, en general, el aprendizaje tanto individual como colectivo, se va a recurrir a la utilización de herramientas de visualización gráfica en línea con lo planteado en Turón y Moreno-Jiménez (2006). En este caso, el trabajo se centra en ofrecer una visualización gráfica de la evolución dinámica de las prioridades asociadas a los elementos comparados en dos instantes de tiempo (Moreno-Jiménez y otros, 2007): antes de que se produjera el proceso de discusión entre los participantes a través de la herramienta colaborativa y después de hacerlo.

Para ilustrar las nuevas herramientas de visualización gráfica planteadas en este trabajo, se consideran los resultados alcanzados en el taller sobre “voto electrónico y e-

cognocracia” que tuvo lugar dentro de las II Jornadas de Comercio Electrónico celebradas en Zaragoza del 28 de febrero al 2 de marzo (Moreno-Jiménez y otros, 2007). El caso tratado (Escobar y Moreno-Jiménez, 2007), se refiere a la localización de una base de inteligencia de la OTAN en Zaragoza. En la modelización del problema se han contemplado, exclusivamente, cuatro criterios (C1: Beneficios, C2: Costes, C3: Oportunidades y C4: Riesgos) y tres alternativas (A1: Sí; A2: Sí, a más de 50 km de la ciudad y A3: No) que han sido evaluadas por 21 de los asistentes al taller. Para ello, han utilizado tarjetas digitales que han garantizado en todo momento las propiedades habitualmente contempladas en los procesos de e-voto y e-democracia.

El trabajo ha quedado estructurado como sigue. La Sección 2 introduce las herramientas de visualización gráficas (véase Turón y Moreno-Jiménez, 2006) y presenta los *diagramas ternarios* empleados para representar las prioridades individuales y colectivas, los *diagramas de áreas* usados para ver la evolución de la dispersión de las opiniones y los *diagramas dinámicos* utilizados para analizar la evolución temporal de las preferencias. La Sección 3 recoge las especificaciones tecnológicas de las herramientas desarrolladas. La Sección 4 incluye el caso analizado y, por último, la Sección 5 resalta las conclusiones más destacadas del trabajo.

2. Descripción de los gráficos

La evolución de las tecnologías de tratamiento de la información (*data mining*, *data warehousing*, inteligencia artificial...) trae consigo una mayor facilidad a la hora de abordar problemas de toma de decisiones en situaciones complejas o mal estructuradas, y permite la creación, representación, almacenamiento, difusión y gestión del conocimiento relativo a los problemas en estudio.

Con el fin de aprovechar estas nuevas ventajas se hace necesario proporcionar al decisor herramientas potentes y de fácil manejo que permitan explotar la flexibilidad, creatividad y capacidad de análisis del cerebro humano, de manera que los diferentes *tipos* (tácito y explícito, etc.) y *formas* (texto, cadenas de caracteres, sonidos, audio, reglas, modelos, etc.) de conocimiento puedan ser extraídos, codificados, catalogados y gestionados de manera efectiva.

Con ese fin se vienen desarrollando desde hace algunos años en el seno del Grupo Decisión Multicriterio Zaragoza (GDMZ) diversas herramientas decisionales que muestran el potencial de las técnicas de visualización gráfica (Turón y otros, 2003, 2005; Moreno–Jiménez y otros, 2005; y Turón y Moreno–Jiménez, 2004). Entre estas herramientas se encuentran: el *diagrama de densidad de la inconsistencia de los juicios*, el *gráfico de evolución de las estructuras de preferencia*, el *diagrama de localización de las estructuras de preferencia*, los *diagramas de densidades de consenso*, los *gráficos MDS*, los *diagramas ternarios* y los *biplots*. En este trabajo se muestra una aplicación de los diagramas ternarios a la visualización de la evolución dinámica de las preferencias de los decisores, así como dos nuevas herramientas gráficas, los *diagramas de áreas* y los *diagramas dinámicos o de evolución*, que han sido desarrolladas recientemente. Se presenta también una aplicación de estas herramientas a un experimento de votación electrónica con dos rondas y discusión intermedia.

2.1 Diagramas ternarios

Anteriormente hemos usado los *diagramas ternarios* (Turón y Moreno–Jiménez, 2006) para representar las estructuras de preferencia de un problema de decisión en AHP con tres alternativas y varios decisores, con el fin de facilitar la rápida identificación de las preferencias de cada uno de ellos y la detección de tendencias o patrones de comportamiento representativos de las diferentes posturas existentes en el grupo de decisores.

En el problema que nos ocupa, la inspección de los diagramas ternarios en la primera y la segunda vuelta de la votación permite descubrir cómo han variado las preferencias expresadas por los decisores.

Se han obtenido tres grupos diferentes de decisores, dependiendo de cuál ha sido la opción elegida en primer lugar. Cada grupo se ha señalado con un color diferente (rojo para quienes prefieren la alternativa 1, verde para quienes prefieren la 2 y azul para quienes se inclinan por la 3), representándose por triángulos los individuos y por un cuadrado el centro de masas del grupo. Un asterisco indica cuál es el centro de masas del grupo completo de decisores.

Cuando un decisor muestra la misma preferencia por dos alternativas se le ha representado por un color intermedio (amarillo cuando el decisor cambia la alternativa 1 por la 2 o viceversa, rosa cuando cambia la 1 por la 3 y viceversa y azul turquesa cuando cambia la 2 por la 3 o viceversa). Si un decisor se muestra igualmente interesado por las tres se le ha representado con color blanco.

2.2 Diagrama de áreas

Se ha obtenido un nuevo diagrama ternario, el *diagrama de áreas*, en el que se ha representado la clausura convexa de cada grupo de decisores, calculándose su área. La diferencia de esta área entre la primera y la segunda votación puede interpretarse como una medida de la variación de la preferencia que cada alternativa recibe por parte de los decisores. Se dispone así de una medida numérica y una representación visual.

2.3 Diagrama dinámico o de evolución

Para conocer con detalle la forma en que la discusión intermedia influye en cada decisor es interesante conocer cómo ha variado la posición de cada uno de ellos de un gráfico a otro. Para ello se han desarrollado dos *diagramas de evolución dinámica*, en los que esta variación se representa mediante flechas que tienen asignado un código de colores.

En el primero se representa cada flecha con un color de intensidad proporcional a su longitud, dentro de una escala que varía entre el amarillo para las variaciones más cortas y el rojo para las máximas variaciones.

En el segundo se ha asignado un código de colores a las flechas que permite destacar en qué casos ha habido cambio en la prioridad del decisor y en qué casos no lo ha habido. En los decisores cuya alternativa preferida en primer lugar es la misma en ambas votaciones, se les ha representado por una flecha del color correspondiente a su preferencia. A los decisores cuya alternativa preferida cambia durante la discusión intermedia se les ha asignado una flecha con un color de fondo representativo del cambio de preferencia (amarillo si pasa de la 1 a la 2 o viceversa, rosa si pasa de la 1 a la 3 o viceversa y azul turquesa si pasa de la 2 a la 3 o viceversa) y una trama del color correspondiente a la alternativa elegida definitivamente.

3. Implementación

Los gráficos propuestos se han incorporado al módulo AHP-GDM (Turón y Moreno–Jiménez, 2004, 2006; Turón y otros, 2003, 2005), desarrollado en lenguaje Visual Basic sobre la hoja de cálculo Excel por el Grupo de Decisión Multicriterio Zaragoza.

Los diagramas ternarios se obtienen con las herramientas de creación de gráficos (*Charts*) de Excel, a partir de las prioridades de los criterios para cada votante en ambas vueltas.

La clausura convexa de los grupos se ha obtenido adaptando el algoritmo QuickHull3D de J.Smith². Los polígonos se han trazado obteniendo, para cada vértice, las coordenadas equivalentes sobre el lienzo (coordenadas (x,y) en píxeles a partir de la esquina superior izquierda).

Los diagramas de evolución se han creado directamente a partir de las coordenadas de cada decisor en ambas vueltas, transformándolas previamente a coordenadas sobre el lienzo. Las áreas se han calculado utilizando los valores originales (proyecciones de las prioridades de las alternativas sobre el plano del simplex).

4. Aplicaciones

Los datos utilizados para ilustrar las nuevas herramientas de visualización gráfica propuestas en este caso, provienen del taller sobre “voto electrónico y e-cognocracia” celebrado la tarde del 1 de marzo (4 horas) dentro de las II Jornadas sobre Comercio Electrónico³. Una exposición detallada del problema puede verse en otro de los trabajos presentados a este congreso (Moreno y otros, 2007). Metodológicamente se utiliza una de las técnicas multicriterio más extendidas, el Proceso Analítico Jerárquico (AHP), para abordar la resolución del problema.

Fijada una jerarquía común para el problema, en la que se contemplan cuatro criterios (C1: Beneficios, C2: Costes, C3: Oportunidades y C4: Riesgos) y tres alternativas (A1: Sí; A2: Sí, a más de 50 km de la ciudad y A3: No), e incorporando las valoraciones dadas en un proceso de arriba hacia abajo por los 21 decisores participantes en el taller,

² <http://algotlist.manual.ru/math/geom/convhull/qhull3d.php>

³ <http://jcel.unizar.es/index.php>

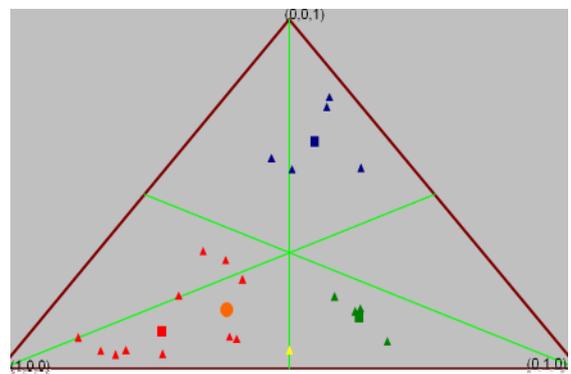
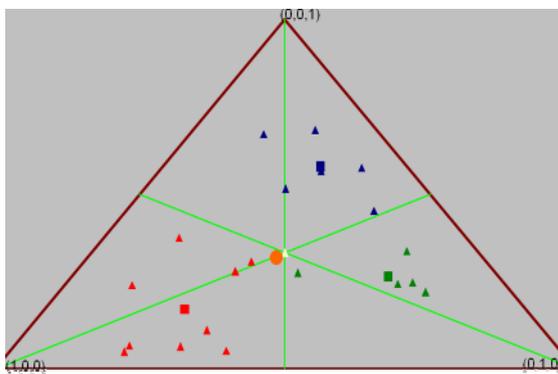
los resultados correspondientes a prioridades totales de las alternativas comparadas pueden verse en la Tabla 1.

Tabla 1. Prioridades totales de las alternativas comparadas.

Votante	Primera Vuelta			Segunda Vuelta			1ª vuelta	2ª vuelta
	Prioridades de las alternativas			Prioridades de las alternativas			ranking	ranking
1	0,3333	0,3333	0,3333	0,5850	0,2045	0,2105	1=2=3	123
2	0,4036	0,2884	0,3080	0,4536	0,2337	0,3127	132	132
3	0,1564	0,5950	0,2487	0,4790	0,1822	0,3388	231	132
4	0,5735	0,3726	0,0538	0,6966	0,2599	0,0435	123	123
5	0,0795	0,3440	0,5765	0,2097	0,2180	0,5722	321	321
6	0,6422	0,1176	0,2402	0,8182	0,0909	0,0909	132	12=3
7	0,4934	0,1300	0,3767	0,0416	0,1799	0,7785	132	321
8	0,2412	0,2431	0,5157	0,3182	0,4741	0,2076	31=2	213
9	0,1057	0,2096	0,6847	0,0606	0,1886	0,7508	321	321
10	0,1547	0,2786	0,5667	0,2898	0,5352	0,1750	321	213
11	0,6472	0,2878	0,0650	0,7987	0,1480	0,0532	123	123
12	0,1998	0,1273	0,6729	0,2282	0,1677	0,6041	321	312
13	0,5783	0,3097	0,1120	0,5467	0,3657	0,0875	123	123
14	0,7337	0,1985	0,0678	0,7536	0,1913	0,0551	123	123
15	0,3401	0,3843	0,2756	0,4523	0,2911	0,2566	213	123
16	0,1210	0,5407	0,3383	0,0893	0,3350	0,5758	231	321
17	0,1474	0,6324	0,2202	0,7785	0,1799	0,0416	231	123
18	0,7509	0,1991	0,0500	0,4728	0,4728	0,0544	123	1=23
19	0,1198	0,4274	0,4528	0,2911	0,6284	0,0805	321	213
20	0,4454	0,2751	0,2795	0,5563	0,3498	0,0940	132	123
21	0,1830	0,5717	0,2453	0,3038	0,5312	0,1651	231	213
22 (cdg)	0,3441	0,3570	0,2989	0,4434	0,3380	0,2187	213	123

Estos datos se han representado en los gráficos siguientes:

Diagramas ternarios

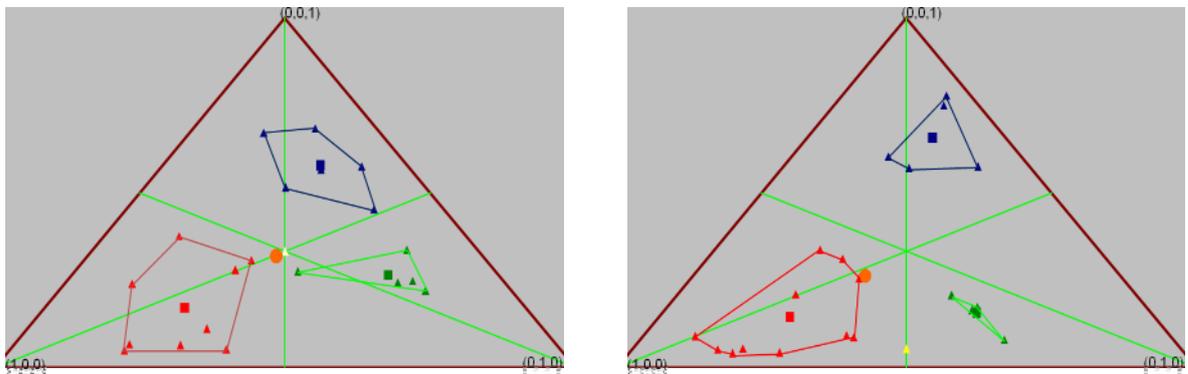


El código de colores permite apreciar la concentración de votos a favor de la primera alternativa (1,0,0), la segunda (0,1,0) o la tercera (0,0,1), así como su mayor o menor proximidad (intensidad con que la alternativa elegida en primer lugar es preferida a las demás).

Diagramas de áreas

Se han obtenido las áreas de cada uno de los polígonos representados, obteniéndose los valores siguientes:

	Primera vuelta	Segunda vuelta
(1,0,0) rojo	0,08615	0,07385
(0,1,0) verde	0,02308	0,00227
(0,0,1) azul	0,05608	0,03077

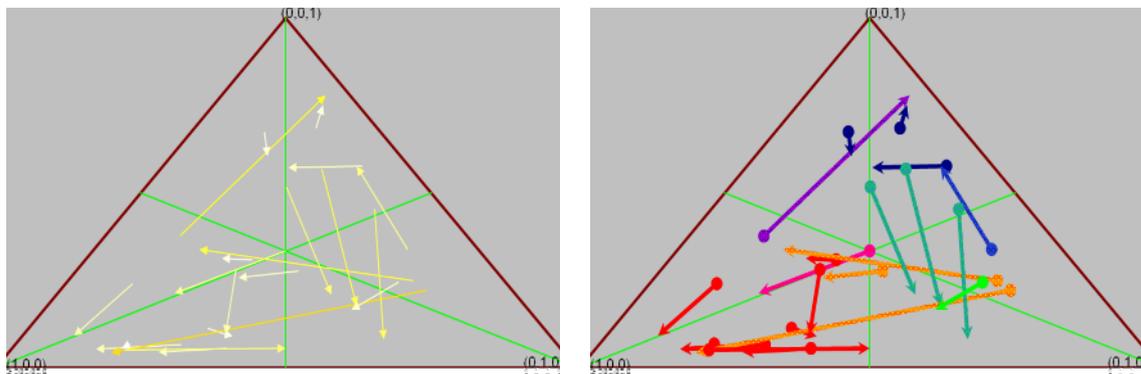


Se aprecia, al igual que en el gráfico anterior, cómo varía la concentración de votantes alrededor de cada una de las alternativas antes y después de la discusión. La representación de la clausura convexa de los votantes de cada una de las alternativas da una idea visual muy clara de cómo varía la preferencia global por cada una de las opciones tras la discusión.

Diagramas dinámicos o de evolución

El diagrama de la izquierda muestra cómo ha variado la posición de cada decisor después de la discusión previa a la segunda votación. La intensidad de esta variación

está representada por la magnitud de la flecha y también por el color, más intenso cuanto mayor sea dicha magnitud. El diagrama de la derecha permite no sólo ver fácilmente cómo han variado las preferencias de cada decisor, sino también detectar a simple vista en qué casos ha habido un cambio en la alternativa preferida, y cuál ha sido el sentido de este cambio. Las flechas rojas, verdes o azules simbolizan votantes que no han cambiado sus preferencias hacia la primera, segunda o tercera alternativas, respectivamente, mientras que las restantes flechas indican cambios en estas preferencias, cambios cuyo valor se deduce del código de colores (se aprecia abundancia de cambios de la segunda alternativa por la primera, y de la tercera por la segunda).



5. Conclusiones

En este trabajo se ha presentado una aplicación de una herramienta de visualización gráfica, los *diagramas ternarios*, a la evolución de las preferencias de los participantes en un problema de votación a dos rondas con discusión intermedia, en el que los participantes pueden variar sus preferencias de una a otra votación. También se han presentado dos nuevas herramientas, los *diagramas de áreas* y los *diagramas dinámicos o de evolución*, que permiten analizar visualmente estas variaciones de las preferencias de los individuos, tanto individual como globalmente. Se facilita así la tarea de localizar patrones de comportamiento presentes en la resolución del problema.

Los gráficos se han implementado en el módulo AHP-GDM, desarrollado en Visual Basic para la hoja de cálculo Microsoft Excel. Como ya se ha mostrado en otras herramientas gráficas incorporadas a este módulo, la interactividad de la hoja de cálculo

Excel hace de estos gráficos una potente herramienta para extraer y analizar información y, por tanto, para entender el comportamiento de las estructuras de preferencia y para explorar las interrelaciones entre ellas.

Bibliografía

Capra, F. (2000): *La trama de la vida. Una nueva perspectiva de los sistemas vivos*. Anagrama, 204.

Escobar, M. T.; Moreno-Jiménez, J. M. (2007): “Agregación de Estructuras de Preferencia en la e-cognocracia”. *XXII Anales de Economía Aplicada*, Valladolid.

Moreno-Jiménez, J. M. (2003): “Las Nuevas Tecnologías y la Representación Democrática del Inmigrante”, *IV Jornadas Jurídicas de Albarracín. Consejo General del Poder Judicial (TSJA)*.

Moreno-Jiménez, J. M. (2004): “E-cognocracia y Representación Democrática del Inmigrante”, *XIX Anales de Economía Aplicada* (CD). León.

Moreno-Jiménez, J. M. (2006): “E-cognocracia: Nueva Sociedad, Nueva Democracia”, *Estudios de Economía Aplicada*, 24(1-2), 559-581.

Moreno-Jiménez, J. M.; Piles, J.; Ruiz, J.; Salazar, J. L.; Sanz, A. (2007): “E-cognocracia y voto electrónico. Aplicación a un problema de localización”, *XXII Anales de Economía Aplicada*. Valladolid.

Moreno-Jiménez, J. M.; Polasek, W. (2003): “E-democracy and Knowledge. A Multicriteria Framework for the New Democratic Era”, *Journal Multicriteria Decision Analysis*, 12, 163-176.

Moreno-Jiménez, J. M.; Polasek, W. (2005): “E-cognocracy and the participation of immigrants in e-governance”. En Böhlen y otros (eds.): *TED Conference on e-government 2005. Electronic democracy: The challenge ahead*. University Rudolf Trauner-Verlag, *Schriftenreihe Informatik*, 13, 18-26.

Moreno-Jiménez, J. M.; Salvador, M. y Turón, A. (2005): “Group preference structures in AHP-group decision making”. *Proceedings of the 2nd Compositional Data Analysis Workshop*. University of Girona Foundation for Innovation and Training, Gerona.

Turón, A. y Moreno-Jiménez, J. M. (2004): “Visualización de información en el Proceso Analítico Jerárquico (AHP)”. *Actas de la XVIII Reunión ASEPELT-España*, León.

Turón, A. y Moreno-Jiménez, J. M. (2006): “Nuevas herramientas de visualización gráfico en Excel para el Proceso Analítico Jerárquico”. *Actas de la XX Reunión ASEPELT-España*, Tenerife.

Turón, A.; Aguarón, J.; Moreno-Jiménez, J. M. y Raluy, A. (2003): “Visualización gráfica de las estructuras de preferencias en AHP”. *Actas de la XVII Reunión ASEPELT-España*, Almería.

Turón, A.; Moreno-Jiménez, J. M. y Salvador, M. (2005): “Intensidad de las preferencias en votaciones. Visualización gráfica de su importancia”. *Actas de la XIX Reunión ASEPELT-España*, Badajoz.

MODELS OF RATIONAL BEHAVIOR WITH LINGUISTIC PREFERENCES

JOSÉ LUIS GARCÍA LAPRESTA
lapresta@eco.uva.es

LUIS CARLOS MENESES PONCIO
lmeneses@eco.uva.es

Departamento de Economía Aplicada
Grupo de Investigación PRESAD
UNIVERSIDAD DE VALLADOLID

Área temática: Métodos Cuantitativos.

Resumen

En la teoría de la decisión clásica existe una amplia gama de modelos de racionalidad que intentan capturar diferentes clases de comportamiento cuando los individuos comparan por pares los elementos de un conjunto de alternativas. Todos estos modelos asumen que los decisores tienen preferencias dicotómicas, de forma que ante cualquier par de alternativas un individuo solamente indica si una alternativa es preferida o indiferente a la otra. Sin embargo, en muchas decisiones reales los individuos sienten diferentes grados en sus preferencias. Por ello, en este trabajo hemos examinado alguno de estos modelos clásicos en un caso real en donde se permitía matizar las preferencias mediante un conjunto de términos lingüísticos. La principal conclusión del análisis empírico realizado es que el cumplimiento de las condiciones racionales decrece cuando los individuos tienen preferencias no extremas. Ante estas evidencias empíricas, proponemos algunos modelos de comportamiento racional, dentro del esquema de las preferencias lingüísticas, que permiten tener en cuenta las intensidades de preferencia que sienten los individuos.

Palabras clave: racionalidad, preferencias lingüísticas, transitividad, operadores monótonos.

Abstract

In classical decision theory there exist a large class of rationality models which try to capture different kinds of behavior when individuals compare by pairs a set of alternatives. All these models assume that decision makers have dichotomous preferences. However, in real decisions individuals feel different degrees of preference. In this paper we have checked the mentioned models in a real case where different kinds of linguistic preferences are allowed. After the empirical analysis, the main conclusion is that the fulfillment of rational conditions decreases when individuals have non-extreme preferences. Based on the obtained empirical evidences, we propose some classes of transitivity conditions in the framework of linguistic preferences.

Key words: rationality, linguistic preferences, transitivity, monotonic operators.

1. Introduction

Classical models of rational behavior only consider dichotomous assessments among pairs of alternatives (preference or not preference). Thus, intensities of preference are not allowed and all the modalities of preference are identified. In this sense, fuzzy preferences are a relevant tool for modeling preference intensities (see Nurmi (1981), Tanino (1984) and De Baets and Fodor (1997), among others). Fuzzy preferences represent intensities of preference by means of numerical values within the unit interval. Due to human beings have usually difficulties in showing their preferences with exact numerical values, linguistic preferences could be appropriate for capturing the lack of precision in human behavior (see Zadeh (1975)).

This paper deals with the fulfillment of some rationality conditions when the agents show their preferences in a linguistic manner. These conditions are based on linear orders, complete preorders, transitivity of preference and indifference, and triple-acyclicity. We have checked these models of rational behavior in a real case where individuals could show their preferences by means of linguistic terms.

Our empirical research is based on the data included in García-Lapresta and Meneses (2003, 2005), where groups of 50 students compared by pairs the alternatives of 3 different sets of 6 alternatives by means of 4 modalities of preference (totally, highly, rather and slightly); in absence of preference they declare indifference. We analyze the fulfillment of the above mentioned classical models of rational behavior with respect to the obtained data but taking into account the modalities of preference used by the individuals.

The paper is organized as follows. Section 2 is devoted to introduce terminology, the classical models of rational behavior we check in the empirical analysis and linguistic preferences. In Section 3 we include the empirical research. Section 4 contains a proposal of rationality models within the framework of linguistic preferences based on monotonic operators. Finally, Section 5 includes some conclusions.

2. Preliminaries

Let S, T be two binary relations on a set A . Δ is the *identity relation* of A [$a\Delta b \Leftrightarrow a=b$]; S^{-1} is the *inverse relation* of S [$aS^{-1}b \Leftrightarrow bSa$]; S^c is the *complement relation* of S [$aS^c b \Leftrightarrow \text{not } (aSb)$]; $S \cap T$ is the *intersection relation* of S and T [$a(S \cap T)b \Leftrightarrow (aSb \text{ and } aTb)$]; $S \cup T$ is the *union relation* of S and T [$a(S \cup T)b \Leftrightarrow (aSb \text{ or } aTb)$]; $S \circ T$ is the *composition relation* of S and T [$a(S \circ T)b \Leftrightarrow \exists c \in A (aS^c c \text{ and } cTb)$].

A way of introducing classical preference and indifference concepts is taking the (strong) *preference* as primitive notion, through an ordinary binary relation P on a set of alternatives A , where aPb means “ a is preferred to b ” or “ a is better than b ”. A basic assumption for P is *asymmetry*: $P \cap P^{-1} = \emptyset$. In this case the *indifference relation*, I , is defined by absence of preference, i.e., a is indifferent to b when neither a is preferred to b nor b is preferred to a : $I = (P \cup P^{-1})^c = P^c \cap (P^{-1})^c$. Then, the *weak preference* relation is defined by $P \cup I = (P^{-1})^c$. We note that from the asymmetry of P , we obtain that I is reflexive [$\forall a \in A aIa$] and symmetric [$\forall a, b \in A aIb \Rightarrow bIa$]; and the completeness of $P \cup I$ [$\forall a, b \in A a(P \cup I)b \text{ or } b(P \cup I)a$]. Consequently, for all $a, b \in A$, one and only one of the following facts holds: aPb or bPa or aIb .

2.1 Classical models of rationality

Among the wide variety of rationality models within the classical approach to preference modeling, we have considered triple-acyclicity, transitivity of preference and indifference relations, complete preorders and linear orders. For other models, see Roubens and Vincke (1985) and García-Lapresta and Rodríguez-Palmero (2004), among others.

Definition 1. Let P be a preference relation on A .

1. P is *triple-acyclic* if $P \circ P \subseteq P \cup I$, i.e.,

$$(aPb \text{ and } bPc) \Rightarrow (aPc \text{ or } aIc)$$

for all $a, b, c \in A$.

2. P is *I-transitive* if $I \circ I \subseteq I$, i.e.,

$$(aIb \text{ and } bIc) \Rightarrow aIc$$

for all $a, b, c \in A$.

3. P is *P-transitive* if $P \circ P \subseteq P$, i.e.,

$$(aPb \text{ and } bPc) \Rightarrow aPc$$

for all $a, b, c \in A$.

4. P is a *complete preorder* if $P \circ P \subseteq P$ and $I \circ I \subseteq I$.

5. P is a *linear order* if $P \circ P \subseteq P$ and $I = \Delta$.

In Figure 1 we show the implications among the considered classical models of rationality (see, for instance, García-Lapresta and Rodríguez-Palmero (2004)).

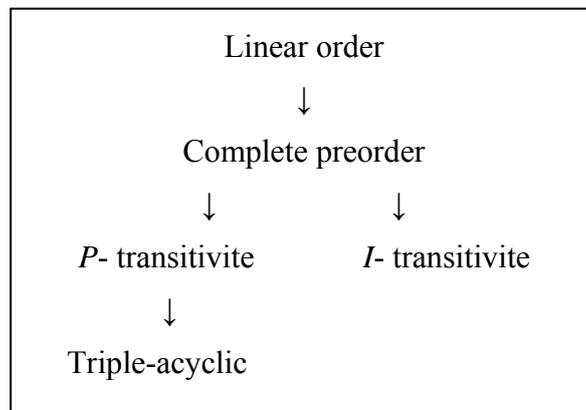


Figure 1. Relationships among rationality models.

2.2 Linguistic preferences

Let A be a finite set of alternatives and $\mathcal{L}_n = \{l_0, l_1, \dots, l_{2n}\}$ be a set of linguistic labels, where $n \geq 1$, ranked by a linear order: $l_0 < l_1 < \dots < l_{2n}$. The intermediate label l_n represents indifference, and the rest of labels are defined around it symmetrically. We denote $\mathcal{L}_n^+ = \{l_{n+1}, \dots, l_{2n}\}$.

Definition 2. A *linguistic preference* relation on A based on \mathcal{L}_n is a mapping $R: A \times A \rightarrow \mathcal{L}_n$ satisfying the *reciprocity*¹ condition:

$$R(a, b) = l_k \Rightarrow R(b, a) = l_{2n-k}$$

for all $a, b \in A$ and $k \in \{0, 1, \dots, 2n\}$.

We denote by $\mathcal{L}_n(A)$ the set of the linguistic preference relations on A based on \mathcal{L}_n .

The meaning of the labels can be summarized in this way: $R(a, b) = l_{2n}$, if a is totally preferred to b ; $l_n < R(a, b) < l_{2n}$, if a is somewhat preferred to b ; $R(a, b) = l_n$, if a is indifferent to b ; $l_0 < R(a, b) < l_n$, if b is somewhat preferred to a ; $R(a, b) = l_0$, if b is totally preferred to a . Since l_k and l_{2n-k} represent the same modality of preference, used in a symmetric way, in \mathcal{L}_n there are $n + 1$ possibilities for declaring opinions between pairs of alternatives: n kinds for preference and 1 for indifference.

We note that $\mathcal{L}_1 = \{l_0, l_1, l_2\}$ consists on the classical preference model. When other modalities of preference are allowed, we should use \mathcal{L}_n for $n > 1$. In fact, the sets of linguistic labels \mathcal{L}_2 , \mathcal{L}_3 and \mathcal{L}_4 have been widely used in the literature. In the empirical research we use \mathcal{L}_4 .

¹ On reciprocity in fuzzy binary relations, see Nurmi (1981), Tanino (1984) and García-Lapresta and Llamazares (2000), among others.

Given $R \in \mathcal{L}_n(A)$, let P_R the ordinary binary relation on A defined by aP_Rb if $R(a,b) > l_n$. Clearly P_R is asymmetric, hence an ordinary preference relation. We say that P_R is the *ordinary preference relation associated with R* . The *ordinary indifference relation associated with R* , I_R , is just the indifference relation associated with P_R : aI_Rb if $R(a,b) = l_n$.

3. The empirical analysis

We have considered the data included in García-Lapresta and Meneses (2003). In that paper two sets of alternatives were analyzed with respect to the fulfillment of several classes of fuzzy transitivity. A group of 50 students compared by pairs 6 different trips. We note that in Spain before students finish their degree, they usually go together for a trip. Previously they have to decide the trip among different possibilities. We asked them about a first set of alternatives with 6 destinations: China, Egypt, Prague-Budapest, Italy, Cuba and Russia. In a second stage, the same group of students compared the same trips but now taking into account the final price of these trips (in euros). This is the second set of alternatives. In García-Lapresta and Meneses (2005) we analyzed individual and collective rationality taking into account the preferences of a group of students with respect to 6 degrees just in the moment they registered in their first year in our Faculty. We have considered the data included in that paper for our present analysis of this third set of alternatives.

The sets of alternatives involved in our empirical research are:

$$X_1 = \{\text{China, Egypt, Prague-Budapest, Italy, Cuba, Russia}\}.$$

$$X_2 = \{(\text{China, 630}), (\text{Egypt, 450}), (\text{Prague-Budapest, 270}), (\text{Italy, 150}), (\text{Cuba, 540}), (\text{Russia, 330})\}.$$

$$X_3 = \{\text{Business Administration and Management (5 years), Business Administration (3 years), Law (5 years), Business Administration, Management and Law (6 years), Labor Relations (3 years), Economics (5 years)}\}.$$

Table 1. The semantics of \mathcal{L}_4

$R(a,b)$	$R(b,a)$	Meaning
l_8	l_0	a is totally preferred to b
l_7	l_1	a is highly preferred to b
l_6	l_2	a is rather preferred to b
l_5	l_3	a is slightly preferred to b
l_4	l_4	a is indifferent to b
l_3	l_5	b is slightly preferred to a
l_2	l_6	b is rather preferred to a
l_1	l_7	b is highly preferred to a
l_0	l_8	b is totally preferred to a

In all the cases, students compare by pairs the feasible alternatives by means of 4 modalities of linguistic preference (totally, highly, rather, slightly) and indifference (see Table 1).

We note that $aP_R b$ if $R(a,b) > l_4$; $aI_R b$ if $R(a,b) = l_4$; and $bP_R a$ if $R(a,b) < l_4$.

In Table 2 and Figure 2 we show the total percentages of individuals who fulfill the rationality models appearing in Figure 1. According to the implications showed in Figure 1, these percentages decrease when the conditions are more restrictive. We note that these percentages are higher in the set X_3 than in X_2 , excepting for I -transitivity, and in X_2 than in X_1 . This fact could be explained by a more accurate information and interest for the students.

Table 2. Percentages of individuals who fulfill the rationality models

	T-A	I-T	P-T	C.P.	L.O.
X_1	90%	64%	62%	46%	12%
X_2	90%	82%	76%	62%	28%
X_3	90%	76%	82%	62%	36%

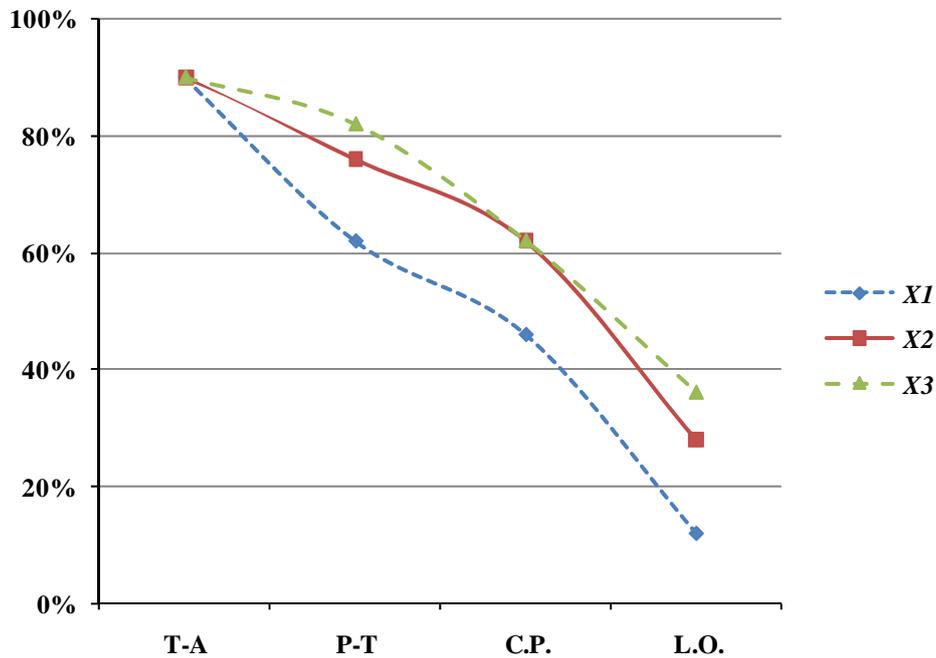


Figure 2. Percentages of individuals who fulfill the rationality models

It is important to note that no property has been fulfilled by all the individuals. Obviously, the smaller percentages of unfulfillment correspond to triple-acyclicity: in the three sets, 10% of individuals do not satisfy this property. Moreover, according to the empirical data, linear orders, even complete preorders, can not be consider as realistic models of rational behavior.

Now we deal with the relative fulfillment of some rationality models taking into account the different modalities of preference the students showed in the pairwise comparisons

of alternatives. This relative fulfillment means the percentages of cases where the corresponding properties are satisfied with respect to all the possible cases. For instance, P_R is P -transitive if $(aP_Rb \text{ and } bP_Rc) \Rightarrow aP_Rc$, for all $a, b, c \in A$. However, it is not the same to violate the above implication in one triplet than in almost all possible cases. The property is not satisfied in both situations, but in a different degree.

Table 3. Triple-acyclicity and P -transitivity in X_2

$R(a,b)$	$R(a,c)$	$R(a,c)$	$R(a,c)$
$R(b,c)$	l_0, l_1, l_2, l_3	l_4	l_5, l_6, l_7, l_8
l_5, l_5	11.63%	9.30%	79.07%
l_5, l_6	5.11%	2.19%	92.70%
l_5, l_7	2.02%	3.03%	94.95%
l_5, l_8	3.09%	0.00%	96.91%
l_6, l_6	1.08%	2.15%	96.77%
l_6, l_7	0.00%	2.44%	97.52%
l_6, l_8	1.30%	0.00%	98.70%
l_7, l_7	0.00%	0.00%	100%
l_7, l_8	0.00%	0.00%	100%
l_8, l_8	0.00%	0.00%	100%
Total	3.03%	2.28%	94.69%

In Table 3 and Figure 3 we include the percentages of relative fulfillment in X_2 by rows. For instance, the second row shows how is $R(a,c)$ whenever $R(a,b)=l_5$ and

$R(b,c)=l_6$, or $R(a,b)=l_6$ and $R(b,c)=l_5$. The first column contains the relative unfulfillment of triple-acyclicity ($R(a,c) < l_4$); the sum of the second and third columns correspond to the relative fulfillment of triple-acyclicity ($R(a,c) \geq l_4$); the sum of the first and second columns correspond to the relative unfulfillment of P -transitivity ($R(a,c) \leq l_4$); and the third column shows the relative fulfillment of P -transitivity ($R(a,c) > l_4$). All these percentages are relative to the number of cases where individuals can fulfill the specific condition. For instance, in column 1 row 1, 11.63% is the percentage of cases where individuals declare that a is slightly preferred to b ($R(a,b)=l_5$), b is slightly preferred to c ($R(b,c)=l_5$) and c is somewhat preferred to a ($R(a,c) < l_4$, i.e., $R(c,a) > l_4$). The last row includes the global information without paying attention to preference intensities $R(a,b)$ and $R(b,c)$. Notice that only the 3.03% of the triplets unfulfill triple-acyclicity and 5.31% of the triplets unfulfill P -transitivity. It is important to note that the fulfillment's percentages increase when individuals show higher intensities of preference in almost all the cases.

In the Appendix we include the data for X_1 and X_3 .

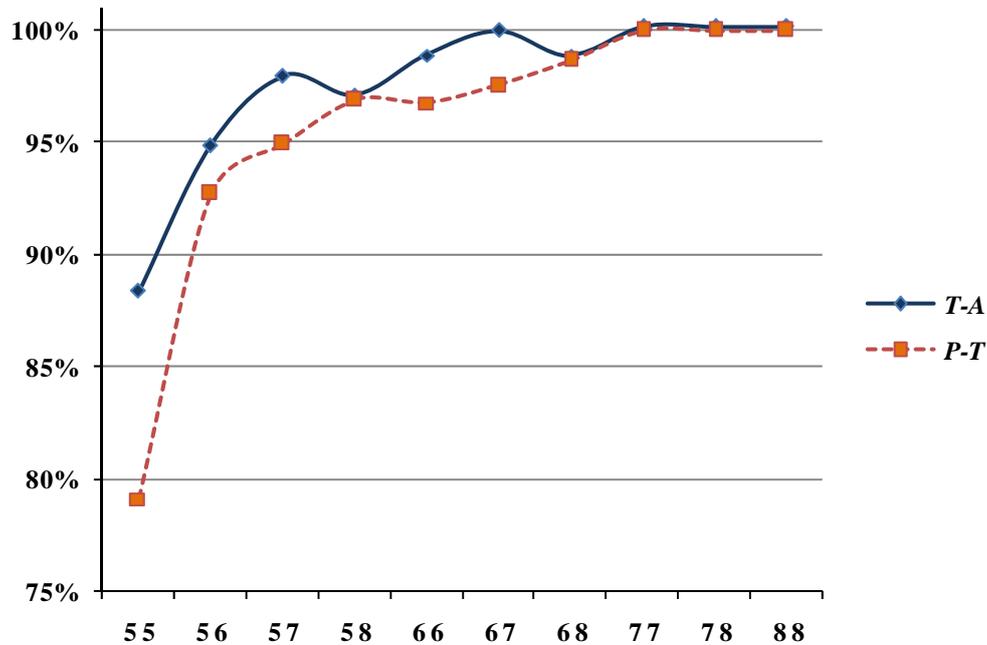


Figure 3. Fulfillment of triple-acyclicity and P -transitivity in X_2

In García-Lapresta and Meneses (2006) other classical rationality models has been analyzed: semiorders, interval orders and semitransitive relations. We note that the fulfillment's percentages of these models are very close to that of P -transitivity.

4. F -transitivity

In ordinary preferences, transitivity is the crucial point for modeling rationality. The same happens in fuzzy preferences, but there exist a wide class of transitivity notions for extending this property to the fuzzy framework (see Dasgupta and Deb (1996), Switalski (1999, 2001, 2003), and García-Lapresta and Meneses (2003, 2005), among others).

According to the empirical analysis, a clear conclusion is that percentages of unfulfillment of the considered rationality models decrease whenever the preference intensities increase. In fact, the case of extreme preference (l_8) hardly produces inconsistencies. In this sense, we can state that the classic models of rational behavior are appropriate for extreme preferences, but not for slight preferences. Thus, it is necessary to consider new models of rational behavior which take into account modalities of preference. In order to define the concept of transitivity in the framework of linguistic preferences, we could consider the idea of t-norm on a finite scale (see Mayor and Torrens (2005)). However, for this purpose, associativity and boundary conditions are not essential. Without these properties, we are going to work only with the commutative and monotonicity conditions by means of the class of monotonic operators, a wider class than the class of t-norms, which allows us to define transitivity in a very general way within the linguistic framework.

Definition 3. A *monotonic operator* on \mathcal{L}_n^+ is a mapping $F : \mathcal{L}_n^+ \times \mathcal{L}_n^+ \rightarrow \mathcal{L}_n$ satisfying the following conditions:

1. *Commutativity:* $F(l_i, l_j) = F(l_j, l_i)$, for all $l_i, l_j \in \mathcal{L}_n^+$.
2. *Monotonicity:* $F(l_i, l_j) = F(l_i, l_k)$, for all $l_i, l_j, l_k \in \mathcal{L}_n^+$ such that $l_j \leq l_k$.
3. $F(l_{2n}, l_{2n}) > l_n$.

Notice that the last condition is weaker than the usual boundary condition (adapted to our linguistic framework) $F(l_{2n}, l_{2n}) = l_{2n}$. On the other hand, it is important to note that associativity, $F(l_i, F(l_j, l_k)) = F(F(l_i, l_j), l_k)$, for all $l_i, l_j, l_k \in \mathcal{L}_n^+$, has not sense in our definition of monotonic operator. Since the codomain of these operators is \mathcal{L}_n , if $F(l_j, l_k) \leq l_n$ or $F(l_i, l_j) \leq l_n$, then the above equality is meaningless.

Examples of monotonic operators are the minimum and the maximum operators.

Definition 4. Let F be a monotonic operator on \mathcal{L}_n^+ and let $R \in \mathcal{L}_n(A)$. R is F -transitive if $R(a, b) \geq F(R(a, c), R(c, b))$ for all $a, b, c \in A$ such that $R(a, c), R(c, b) \in \mathcal{L}_n^+$.

In the next Proposition we show some interesting facts for two specific constant monotonic operators, those which satisfy $F(l_{n+1}, l_{n+1}) \in \mathcal{L}_n^+$, and one for the case of $n = 1$ (the classical preference model).

Proposition 5. Let F be a monotonic operator on \mathcal{L}_n^+ and let $R \in \mathcal{L}_n(A)$. The following statements hold:

1. If $F(l_i, l_j) = l_n$ for all $l_i, l_j \in \mathcal{L}_n^+$, then R is F -transitive if and only if P_R is triple-acyclic.
2. If $F(l_i, l_j) = l_{n+1}$ for all $l_i, l_j \in \mathcal{L}_n^+$, then R is F -transitive if and only if P_R is P -transitive.
3. If $F(l_{n+1}, l_{n+1}) \geq l_{n+1}$ and R is F -transitive, then P_R is P -transitive.
4. For $n = 1$, R is F -transitive if and only if P_R is P -transitive.

Proof.

- 1) Suppose R is F -transitive, $aP_R b$ and $bP_R c$. Since $R(a, b), R(b, c) \in \mathcal{L}_n^+$, we have $R(a, c) \geq F(R(a, b), R(b, c)) = l_n$ and, consequently, $aP_R c$ or $aI_R c$.

Conversely, suppose P_R is triple-acyclic and $R(a,b), R(b,c) \in \mathcal{L}_n^+$. Then, $aP_R b$ and $bP_R c$. Therefore, $aP_R c$ or $aI_R c$, i.e., $R(a,c) \geq l_n = F(R(a,b), R(b,c))$.

- 2) Suppose R is F -transitive, $aP_R b$ and $bP_R c$. Since $R(a,b), R(b,c) \in \mathcal{L}_n^+$, we have $R(a,c) \geq F(R(a,b), R(b,c)) = l_{n+1}$ and, consequently, $aP_R c$.

Conversely, suppose P_R is P -transitive and $R(a,b), R(b,c) \in \mathcal{L}_n^+$. Since $aP_R b$ and $bP_R c$, we have $aP_R c$, i.e., $R(a,c) \geq l_{n+1} = F(R(a,b), R(b,c))$.

- 3) Suppose R is F -transitive, $aP_R b$ and $bP_R c$. Since $R(a,b), R(b,c) \in \mathcal{L}_n^+$, then we have $R(a,c) \geq F(R(a,b), R(b,c)) \geq F(l_{n+1}, l_{n+1}) \geq l_{n+1}$ and, consequently, $aP_R c$.

- 4) Suppose Suppose R is F -transitive, $aP_R b$ and $bP_R c$. Since $R(a,b), R(b,c) \in \mathcal{L}_1^+$, we have $R(a,c) \geq F(R(a,b), R(b,c)) \geq F(l_2, l_2) > l_1$ and, consequently, $aP_R c$.

Conversely, suppose P_R is P -transitive and $R(a,b), R(b,c) \in \mathcal{L}_1^+ = \{l_2\}$. Then, $aP_R b$ and $bP_R c$. Therefore, $aP_R c$, i.e., $R(a,c) = l_2 = F(R(a,b), R(b,c))$.

4.1 Some examples

Now we are interested in searching some classes of monotonic operators which could be suitable for modeling the rational behavior through the corresponding linguistic transitivity. For this task we will consider the empirical research provided in the previous section.

Example 6. A well known rationality model in the framework of fuzzy preferences is that of *min-transitivity*. This model can be easily defined in \mathcal{L}_n^+ through

$$F_1(l_i, l_j) = \min(l_i, l_j) = l_{\min\{i, j\}} :$$

$$R(a, b) \geq \min\{R(a, c), R(c, b)\}$$

whenever $R(a, c), R(c, b) \in \mathcal{L}_n^+$.

Since $F_1(l_{n+1}, l_{n+1}) = l_{n+1}$, by Proposition 5, P_R is P -transitive for every $R \in \mathcal{L}_n(A)$ satisfying F_1 -transitivity.

Example 7. We now consider the monotonic operators $F_k : \mathcal{L}_n^+ \times \mathcal{L}_n^+ \rightarrow \mathcal{L}_n$, for $k = 2, 3, 4$, defined by

$$F_2(l_i, l_j) = \max\{l_{i+j-2n}, l_{n+1}\}$$

$$F_3(l_i, l_j) = \max\{l_{i+j-2n}, l_n\}$$

$$F_4(l_i, l_j) = l_{i+j-2n}$$

Notice that they have some similarities with the Łukasiewicz t-norm². On the other hand, we have

$$F_1(l_i, l_j) \geq F_2(l_i, l_j) \geq F_3(l_i, l_j) \geq F_4(l_i, l_j)$$

for all $l_i, l_j \in \mathcal{L}_n^+$.

In Table 4 we show these monotonic operators for $n = 4$.

Table 4. F_2, F_3, F_4 monotonic operators on \mathcal{L}_4^+

F_2	l_5	l_6	l_7	l_8	F_3	l_5	l_6	l_7	l_8	F_4	l_5	l_6	l_7	l_8
l_5	l_5	l_5	l_5	l_5	l_5	l_4	l_4	l_4	l_5	l_5	l_2	l_3	l_4	l_5
l_6	l_5	l_5	l_5	l_6	l_6	l_4	l_4	l_5	l_6	l_6	l_3	l_4	l_5	l_6
l_7	l_5	l_5	l_6	l_7	l_7	l_4	l_5	l_6	l_7	l_7	l_4	l_5	l_6	l_7
l_8	l_5	l_6	l_7	l_8	l_8	l_5	l_6	l_7	l_8	l_8	l_5	l_6	l_7	l_8

According to Proposition 5, P_R is P -transitive for every $R \in \mathcal{L}_n(A)$ satisfying F_2 -transitivity; and P_R is triple-acyclic for every $R \in \mathcal{L}_n(A)$ satisfying F_3 -transitivity.

² The Łukasiewicz t-norm is defined on $[0,1]^2$ through $T_L(x, y) = \max\{x + y - 1, 0\}$

Clearly, F_4 has a different behavior depending on if it aggregates high or low intensities of preference. For instance, if R is F_4 -transitive, $R(a,b)=l_6$ and $R(b,c)=l_7$, then $R(a,c)$ should be at least l_5 . If $R(a,b)=l_7$ and $R(b,c)=l_8$, then $R(a,c)$ should be at least l_7 . Even more, if $R(a,b)=l_5$ and $R(b,c)=l_6$, then $R(a,c)$ should only be at least l_3 . Thus, F_4 -transitivity does not guarantee triple-acyclicity of P_R .

In Table 5 and Figure 4 we show the percentages of individuals who fulfill F_k -transitivity for $k = 1, 2, 3, 4$ in the three sets of alternatives.

Table 5. Percentages of individuals who fulfill F_k -transitivity for $k = 1, 2, 3, 4$

	F_1	F_2	F_3	F_4
X_1	52%	60%	80%	84%
X_2	48%	66%	78%	82%
X_3	54%	82%	86%	88%

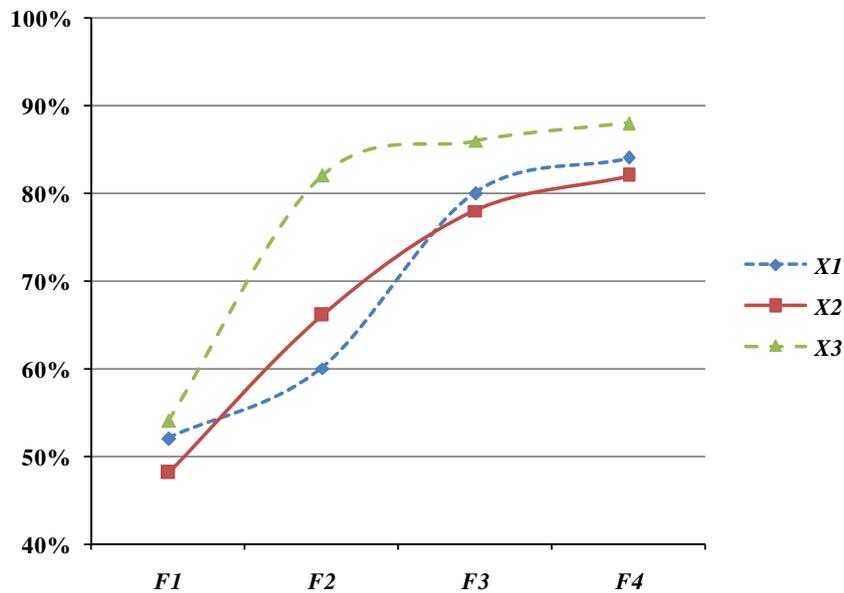


Figure 4. Percentages of individuals who fulfill F_k -transitivity for $k = 1, 2, 3, 4$

It is important to note the low fulfillment of F_1 -transitivity in all the cases. Thus, taking into account our empirical data, the well known rationality condition of min-transitivity (translated to the linguistic framework) cannot be considered as a realistic model of rational behavior.

F_1 and F_2 transitivities ensure P -transitivity to the associated ordinary preference relations. Notice that the only differences between F_1 and F_2 are in $\{l_6, l_7\}$, being $F_2(l_i, l_j) = l_{p-1}$ and $F_1(l_i, l_j) = l_p$ for all $i, j \in \{6, 7\}$. Obviously, F_2 -transitivity is fulfilled in a bigger degree than F_1 -transitivity; then F_2 -transitivity is preferable to F_1 -transitivity for modeling rational behavior.

If we compare F_2 and F_3 , the only differences appear in low preferences: $F_3(l_i, l_j) = l_{p-1}$ and $F_2(l_i, l_j) = l_p$ whenever $i + j < 3n + 1$. Consequently, F_3 -transitivity does not ensure P -transitivity to the associated ordinary preference relations, but only triple-acyclicity.

Since there exist clear differences between the fulfillment of F_3 -transitivity and F_2 -transitivity, it seems that F_3 -transitivity is a more reasonable coherence model than F_2 -transitivity.

If we now compare F_3 and F_4 , again the only differences appear in low preferences. So, for $n = 4$, F_3 and F_4 coincides excepting in the cases $F_4(l_5, l_5) = l_2 < l_4 = F_3(l_5, l_5)$ and $F_4(l_5, l_6) = F_4(l_6, l_5) = l_3 < l_4 = F_3(l_5, l_6) = F_3(l_6, l_5)$. Clearly, the weaker model of F_4 -transitivity is more realistic for modeling rationality than the previous ones. However, F_4 -transitivity does not ensure triple-acyclicity to the associated ordinary preference relations.

4.2 Our proposal

Taking into account the previous examples and Table 3, we can establish that F -transitivity should be defined through monotonic operators whose behavior be

different depending on the strongness of preferences they aggregate. Since for every $(l_i, l_j) \in \mathcal{L}_n^+ \times \mathcal{L}_n^+$ we have $i + j \in [2n + 2, 4n] \cap \mathbb{N}$, the middle point in this interval of positive integers is $3n + 1$. So, we differentiate in $\mathcal{L}_n^+ \times \mathcal{L}_n^+$ three disjoint subsets, L , M and H for low, medium and high preferences, respectively:

$$L = \{(l_i, l_j) \in \mathcal{L}_n^+ \times \mathcal{L}_n^+ \mid i + j < 3n + 1\}.$$

$$M = \{(l_i, l_j) \in \mathcal{L}_n^+ \times \mathcal{L}_n^+ \mid i + j = 3n + 1\}.$$

$$H = \{(l_i, l_j) \in \mathcal{L}_n^+ \times \mathcal{L}_n^+ \mid i + j > 3n + 1\}.$$

If $F : \mathcal{L}_n^+ \times \mathcal{L}_n^+ \rightarrow \mathcal{L}_n$ is a monotonic operator, we can decompose F in three mappings, $F^L : L \rightarrow \mathcal{L}_n$, $F^M : M \rightarrow \mathcal{L}_n$ and $F^H : H \rightarrow \mathcal{L}_n$, one for each kind of preferences in such a way that

$$F(l_i, l_j) = \begin{cases} F^L(l_i, l_j) & \text{if } (l_i, l_j) \in L, \\ F^M(l_i, l_j) & \text{if } (l_i, l_j) \in M, \\ F^H(l_i, l_j) & \text{if } (l_i, l_j) \in H. \end{cases}$$

In Figure 5 we show the decomposition of F in F^L , F^M and F^H .

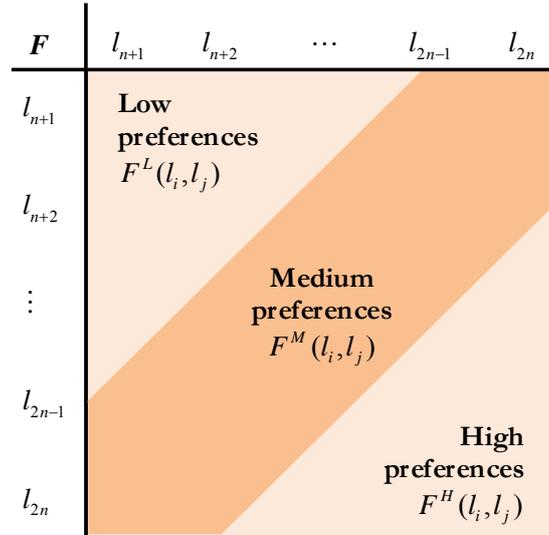


Figure 5. Decomposition of F in F^L , F^M and F^H .

For $n = 4$, we have:

$$L = \{(l_5, l_5), (l_5, l_6), (l_5, l_7), (l_6, l_5), (l_6, l_6), (l_7, l_5)\}.$$

$$M = \{(l_5, l_8), (l_6, l_7), (l_7, l_6), (l_8, l_5)\}.$$

$$H = \{(l_6, l_8), (l_7, l_7), (l_7, l_8), (l_8, l_6), (l_8, l_7), (l_8, l_8)\}.$$

In Table 6 we show the decomposition of F for $n = 4$

Table 6. Decomposition of F for $n = 4$

F	l_5	l_6	l_7	l_8
l_5	$F^L(l_5, l_5)$	$F^L(l_5, l_6)$	$F^L(l_5, l_7)$	$F^M(l_5, l_8)$
l_6	$F^L(l_6, l_5)$	$F^L(l_6, l_6)$	$F^M(l_6, l_7)$	$F^H(l_6, l_8)$
l_7	$F^L(l_7, l_5)$	$F^M(l_7, l_6)$	$F^H(l_7, l_7)$	$F^H(l_7, l_8)$
l_8	$F^M(l_8, l_5)$	$F^H(l_8, l_6)$	$F^H(l_8, l_7)$	$F^H(l_8, l_8)$

In Table 7 we show the decomposition of F_k for $k = 1, 2, 3, 4$.

Table 7. Decomposition of F_1, F_2, F_3 and F_4

k	$F_k^L(l_i, l_j)$	$F_k^M(l_i, l_j)$	$F_k^H(l_i, l_j)$
1	$\min\{l_i, l_j\}$	$\min\{l_i, l_j\}$	$\min\{l_i, l_j\}$
2	l_{n+1}	l_{n+1}	l_{i+j-2n}
3	l_n	l_{n+1}	l_{i+j-2n}
4	l_{i+j-2n}	l_{n+1}	l_{i+j-2n}

We now consider some monotonic operators which can be considered as appropriate for modeling rationality in the linguistic framework. The differences with respect to the monotonic operators included in 4.1 are double: weaker requirements in H and stronger requirements in L . These changes guarantee P -transitivity in F_k^H for $k = 5, 6, 7, 8$ and in F_k^M for $k = 5, 6, 7$. Triple-acyclicity is guaranteed in F_6^L , but not in F_k^L for $k = 5, 7, 8$.

Table 8 shows the decomposition of F_k for $k = 5, 6, 7, 8$. Table 9 and Figure 6 show percentages of individuals satisfying F_k -transitivity for $k = 5, 6, 7, 8$.

Table 8. Decomposition of F_5 , F_6 , F_7 and F_8

k	$F_k^L(l_i, l_j)$	$F_k^M(l_i, l_j)$	$F_k^H(l_i, l_j)$
5	l_{i+j-2n}	l_{n+1}	l_{n+2}
6	l_n	l_{n+1}	l_{n+2}
7	$\begin{cases} l_{n-1}, & \text{if } i = j = 3n \text{ and } i \neq 2n-1 \neq j \\ l_{i+j-2n} & \text{otherwise} \end{cases}$	l_{n+1}	l_{n+2}
8	$\begin{cases} l_{n-1}, & \text{if } i = j = 3n \text{ and } i \neq 2n-1 \neq j \\ l_{i+j-2n} & \text{otherwise} \end{cases}$	$\begin{cases} l_{n+1}, & \text{if } i = 2n \text{ or } j = 2n \\ l_n & \text{otherwise} \end{cases}$	l_{n+2}

Among the considered monotonic operators which ensure P -transitivity in M and H , F_7 -transitivity reaches the highest fulfillment percentages. F_8 -transitivity is the only F_k -transitivity property which does not guarantee P -transitivity in the associated ordinary preferences for M . However, F_8 -transitivity ensures P -transitivity in the associated ordinary preferences for H . This smaller requirement entails the highest fulfillment percentages.

Table 9. Percentages of individuals who fulfill F_k -transitivity for $k = 5, 6, 7, 8$

	F_5	F_6	F_7	F_8
X_1	84%	80%	84%	86%
X_2	92%	88%	94%	96%
X_3	90%	88%	92%	94%

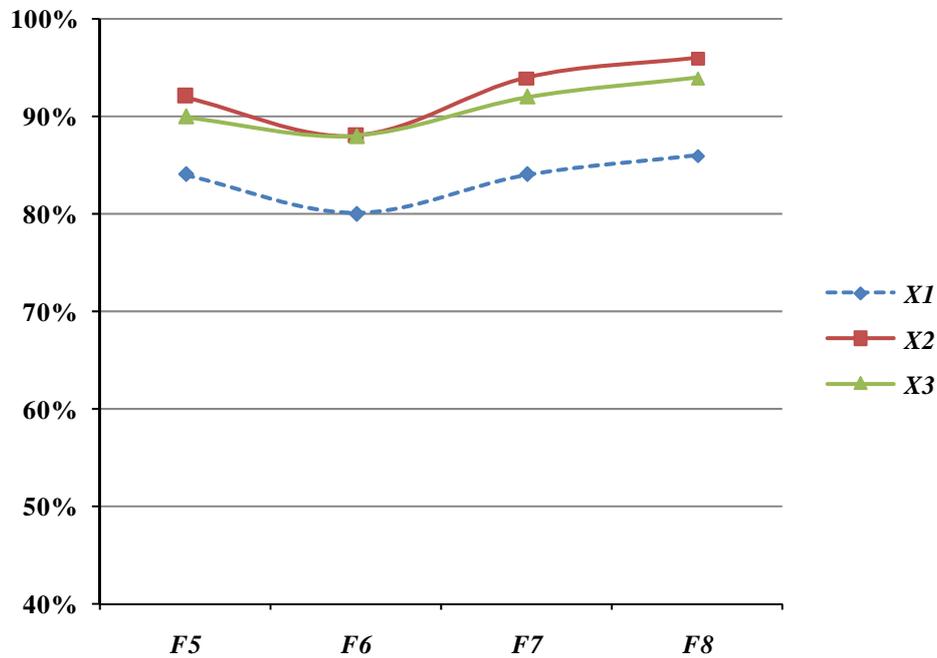


Figure 6. Percentages of individuals who fulfill F_k -transitivity for $k = 5, 6, 7, 8$

5. Concluding remarks

Since usually human being have difficulties for assigning exact numerical values to intensities of preference, linguistic preferences have the advantage of dealing with linguistic expressions such as occur in real life. In this paper we have considered that possibility and we have introduced a class of consistency models based on monotonic

operators, which generalizes the class of t-norms in discrete settings. So, given a monotonic operator F on a subset of linguistic preferences, our model of coherence requires that if a is preferred to b and b is preferred to c , then the intensity of preference between a and c should be at least the outcome given by F to the preference intensities between a and b and b and c . Then, the problem is to determine a class of suitable monotonic operators for modeling rationality in the linguistic framework.

It is important to emphasize that our proposal takes into account real data provided by individuals in real decision problems. First of all, we have considered some classical models of rationality, and we can conclude that, according to the empirical data, neither linear orders nor complete preorders capture the real behavior of individuals. Moreover, the fulfillment of these rationality models decreases whenever individuals show low preferences among alternatives. Subsequently, we have introduced other models of rational behavior which take into account intensities of preference. In this sense, F -transitivity (based on monotonic operators) allows us to define a wide class of rationality models which differentiate low, medium and high preferences by considering different requirements in each kind of preference modalities. According to our empirical data, those F -transitivity properties which guarantee P -transitivity in medium and high preferences have obtained very good results.

Acknowledgements

This research is partially supported by the Spanish Ministerio de Educación y Ciencia (Project SEJ2006-04267/ECON), the Junta de Castilla y León (Consejería de Educación y Cultura, Project VA040A05) and ERDF.

References

Dasgupta, M. and Deb, R. (1996): "Transitivity and fuzzy preferences", *Social Choice and Welfare*, 13, pp. 305-318.

De Baets, B. and Fodor, J. (1997): “Twenty years of fuzzy preference structures”, *Rivista di Matematica per le Scienze Economiche e Sociali*, 20, pp. 45-66.

García-Lapresta, J.L. and Llamazares, B. (2000): “Aggregation of fuzzy preferences: Some rules of the mean”, *Social Choice and Welfare*, 17, pp. 673-690.

García-Lapresta, J.L. and Meneses, L.C. (2003): “An empirical analysis of transitivity with four scaled preferential judgment modalities”, *Review of Economic Design*, 8, pp. 335-346.

García-Lapresta, J.L. and Meneses, L.C. (2005): “Individual valued preferences and their aggregation: Analysis of consistency in a real case”, *Fuzzy Sets and Systems* 151, pp. 269-284.

García-Lapresta, J.L. and Meneses, L.C. (2006): “Classical models of rational behavior and intensities of preference”, *Proceedings of the XIII Congress of SIGEF*, Hammamet (Tunisia), pp. 609-628.

García-Lapresta, J.L. and Rodríguez-Palmero, C. (2004): “Some algebraic characterizations of preference structures”, *Journal of Interdisciplinary Mathematics*, 7, pp. 233-254.

Mayor, G. and Torrens, J. (2005): “Triangular norms on discrete settings”, In: E.P. Klement and R. Mesiar (eds.), *Logical, Algebraic, Analytic and Probabilistic Aspects of Triangular Norms*. Elsevier, pp. 189-230.

Nurmi, H. (1981): “Approaches to collective decision making with fuzzy preference relations”, *Fuzzy Sets and Systems*, 6, pp. 249-259.

Roubens, M. and Vincke, P. (1985): *Preference Modelling*. Lecture Notes in Economics and Mathematical Systems, vol. 250. Springer-Verlag, Berlin.

Switalski, Z. (1999): “Rationality of fuzzy reciprocal preference relations”, *Fuzzy Sets and Systems*, 107, pp. 187-190.

Switalski, Z. (2001): "Transitivity of fuzzy preference relations - an empirical study", *Fuzzy Sets and Systems*, 118, pp. 503-508.

Switalski, Z. (2003): "General transitivity conditions for fuzzy reciprocal preference matrices", *Fuzzy Sets and Systems*, 137, pp. 85-100.

Tanino, T. (1984): "Fuzzy preference orderings in group decision making", *Fuzzy Sets and Systems*, 12, pp. 117-131.

Zadeh, L.A. (1975): "The concept of a linguistic variable and its applications to approximate reasoning", *Information Sciences*, Part I: 8, pp. 199-249. Part II: 8, pp. 301-357. Part III: 9, pp. 43-80.

Appendix. Tables for X_1 and X_3

Table 10. Triple-acyclicity and P-transitivity in X_1

$R(a,b)$	$R(a,c)$	$R(a,c)$	$R(a,c)$
$R(b,c)$	l_0, l_1, l_2, l_3	l_4	l_5, l_6, l_7, l_8
l_5, l_5	2.53%	6.33%	91.14%
l_5, l_6	5.85%	5.26%	88.89%
l_5, l_7	3.80%	2.53%	93.67%
l_5, l_8	5.36%	5.36%	89.29%
l_6, l_6	1.39%	5.56%	93.06%
l_6, l_7	4.63%	1.85%	93.52%
l_6, l_8	4.84%	3.23%	91.94%
l_7, l_7	0.00%	3.13%	100%
l_7, l_8	0.00%	0.00%	100%
l_8, l_8	0.00%	0.00%	100%
Total	3.86%	4.01%	92.13%

Table 11. Triple-acyclicity and P-transitivity in X_3

$R(a,b)$	$R(a,c)$	$R(a,c)$	$R(a,c)$
$R(b,c)$	l_0, l_1, l_2, l_3	l_4	l_5, l_6, l_7, l_8
l_5, l_5	9.68%	6.45%	83.87%
l_5, l_6	1.54%	3.08%	95.38%
l_5, l_7	0.00%	1.30%	98.70%
l_5, l_8	0.00%	0.00%	100%
l_6, l_6	3.25%	0.00%	96.75%
l_6, l_7	1.14%	0.57%	98.30%
l_6, l_8	2.27%	0.00%	97.73%
l_7, l_7	1.85%	0.00%	98.15%
l_7, l_8	0.00%	0.00%	100%
l_8, l_8	0.00%	0.00%	100%
Total	1.86%	0.99%	97.15%

UN PROCEDIMIENTO DE DECISIÓN COLECTIVA SESGADO HACIA EL CONSENSO

José Luis García Lapresta

lapresta@eco.uva.es

Departamento de Economía Aplicada, Grupo de Investigación PRESAD

David Pérez Román

david@emp.uva.es

Departamento de Organización de Empresas y Comercialización e Investigación de Mercados

UNIVERSIDAD DE VALLADOLID

Área temática: Métodos cuantitativos

Resumen

En este trabajo presentamos un procedimiento secuencial de decisión en el cual se obtiene una ordenación colectiva de alternativas a partir de las preferencias individuales, teniendo en cuenta las aportaciones al consenso de cada uno de los miembros del grupo. En un primer paso, se determinan las distancias entre las preferencias individuales y la preferencia colectiva correspondiente a la regla de Borda. A partir de estas distancias, se obtiene una medida del grado de consenso existente en cada subconjunto de votantes y asignamos un peso a cada uno de ellos, correspondiente a su contribución marginal al consenso, de forma que aquellos votantes que contribuyan negativamente al consenso son eliminados del grupo. A continuación se reitera el proceso hasta conseguir que todos los miembros del grupo contribuyan positivamente al consenso. Llegado este momento, se ponderan las opiniones de los votantes por los pesos correspondientes y se genera una ordenación colectiva de alternativas a través de una regla de Borda ponderada. Una vez introducido el procedimiento, se analizan algunas de sus propiedades.

Palabras clave: Toma de decisiones, Consenso, Regla de Borda.

Abstract

In this paper we introduce a sequential decision procedure that provides a collective ranking of alternatives from a set of individual preferences. This decision procedure takes into account the contributions to the consensus of each member of the group. In the first stage we consider the collective preference provided by the Borda rule and the distances between individual and collective preferences. Taking into account these distances, we measure the agreement in each subset of voters, and a weight is assigned to each voter: his/her marginal contribution to the agreement. Those individuals which negatively contribute to the consensus are expelled of the group. Then we reinitiate the decision procedure until all the individuals positively contribute to the consensus. Subsequently, we apply a weighted Borda rule where the opinions of each voter are multiplied by his/her weight which generates the final collective preference. We analyze some features of this decision procedure.

Key Words: Collective decision making, Consensus, Borda rule.

1. Introducción

Cuando un grupo de votantes muestra sus preferencias sobre un conjunto de alternativas, resulta interesante conocer el grado de acuerdo existente entre los miembros del grupo y en cada uno de los subgrupos que puedan formar. En este trabajo consideramos que los votantes ordenan las alternativas mediante preórdenes completos y que la preferencia colectiva viene dada mediante la regla de Borda [3], la cual genera un nuevo preorden completo que representa las preferencias de los agentes.

Existen poderosos argumentos a favor de la regla de Borda (algunos de ellos pueden encontrarse en Dummett [8, 9], Saari [16] y Young [19]), así como algunas caracterizaciones axiomáticas de la misma (véanse Gärdenfors [12], Young [18] y Fine y Fine [11]). Aquí nos interesa resaltar que la regla de Borda realiza un papel similar al que tiene la media aritmética cuando se trata de obtener un valor representativo de un conjunto de números reales. De hecho, la posición que obtiene una alternativa dentro del preorden completo generado por la regla de Borda corresponde a la media aritmética de las posiciones asignadas por los votantes a dicha alternativa en sus preferencias individuales.

Inicialmente, la regla de Borda [3] consideraba la hipótesis de que las preferencias individuales eran órdenes lineales. Sin embargo, existen extensiones de la regla de Borda a preórdenes completos (véanse Black [2], Gärdenfors [12] y Young [18], entre otros). Nosotros consideraremos una generalización de la regla de Borda utilizada por Smith [17], Black [2] y Cook y Seiford [7], en la cual las alternativas indiferentes obtienen la puntuación promedio (en nuestro enfoque, la posición promedio) de las que conseguirían las de su clase de equivalencia si se linealizara el preorden completo inicial.

Con objeto de tener un marco para establecer el grado de acuerdo existente en cada grupo de votantes, hemos introducido un concepto general de *medida de consenso*, el cual guarda relación con el dado por Bosch [4]. Dentro de este marco, hemos realizado una propuesta concreta de medida de consenso que tiene en cuenta la distancia entre las preferencias individuales y la preferencia colectiva generada por la regla de Borda.

La distancia utilizada está basada en la distancia euclídea entre aquellos vectores que representan las posiciones de las alternativas en los correspondientes preórdenes completos.

A partir de las distancias entre las preferencias individuales y la preferencia colectiva obtenida mediante la regla de Borda, hemos asignado un peso a cada votante en función de su contribución marginal al consenso y hemos introducido una regla de Borda ponderada que multiplica la posición original de las alternativas por el correspondiente peso asociado a cada votante. Este nuevo sistema de votación penaliza a aquellos votantes cuyas opiniones difieren considerablemente de la opinión media. Con ello, priorizamos la opinión de los votantes en función de su contribución al consenso (véase Cook, Kress y Seiford [5]).

Conviene señalar que el término *consenso* tiene diferentes significados en la literatura. Uno de ellos hace referencia a un procedimiento iterativo de decisión colectiva en el que los votantes han de modificar sus preferencias con el objetivo de llegar al máximo acuerdo posible. Habitualmente, existe la figura de un moderador que orienta a los votantes sobre cómo cambiar sus preferencias para intentar conseguir el mejor acuerdo posible (véanse, por ejemplo, Zadrozny [20] y Eklund, Rusinowska y de Swart [10]).

Nuestra propuesta consiste en proporcionar un procedimiento secuencial de decisión colectiva que prime en cada momento a aquellos votantes cuyas opiniones estén más cercanas a la opinión colectiva generada por la regla de Borda. En primer lugar, los votantes muestran sus preferencias (preórdenes completos) y se determina la preferencia colectiva a través de la regla de Borda. A continuación se calcula la contribución marginal al consenso de cada votante y se eliminan las opiniones de aquellos votantes cuya contribución al consenso sea negativa. Se reinicia el proceso de forma iterativa hasta disponer de un subconjunto de votantes en el que todos contribuyan positivamente al consenso. Una vez conseguido tal subconjunto, se utilizan los correspondientes índices de contribución marginal al consenso para ponderar las opiniones individuales. La regla de Borda ponderada proporciona la ordenación final de las alternativas. Hay que destacar que los votantes expresan su opinión una

única vez, por lo que no pueden modificarla una vez iniciado el proceso.

La organización del trabajo es la siguiente. En la Sección 2 introducimos la notación y las principales nociones que utilizaremos a lo largo del trabajo. En la Sección 3 presentamos un concepto general de medida del consenso, así como una medida particular basada en la regla de Borda y la distancia euclídea. También introducimos la contribución marginal al consenso de cada votante y una nueva regla de Borda ponderada, en la cual las opiniones individuales quedan afectadas por sus contribuciones marginales al consenso. La Sección 4 contiene el procedimiento recursivo de decisión colectiva, algunos ejemplos y las futuras líneas de investigación.

2. Preliminares

Sea $V = \{v_1, \dots, v_m\}$ ($m \geq 3$) un conjunto de votantes que muestran sus preferencias sobre un conjunto de alternativas $X = \{x_1, \dots, x_n\}$ ($n \geq 3$). Con $\mathcal{P}(V)$ se denotará el conjunto potencia de V (es decir, $I \in \mathcal{P}(V) \Leftrightarrow I \subseteq V$), con $L(X)$ el conjunto de órdenes lineales sobre X (relaciones binarias sobre X que son reflexivas, antisimétricas y transitivas), y con $W(X)$ el conjunto de preórdenes completos sobre X (relaciones binarias sobre X que son completas y transitivas). Un perfil de preferencias individuales es un vector $\mathbf{R} = (R_1, \dots, R_m)$. Dado $R \in W(X)$, el inverso de R es el preorden completo R^{-1} definido por $x_i R^{-1} x_j \Leftrightarrow x_j R x_i$, para cualesquiera $x_i, x_j \in X$. Con $|I|$ denotaremos el cardinal de I .

2.1. Codificación

A continuación introducimos un sistema de codificación de órdenes lineales y preórdenes completos mediante vectores que representan las posiciones relativas de las alternativas. Procedimientos similares a los aquí expuestos han sido utilizados en la generalización de reglas de puntuación (*scoring rules*) de su concepción original sobre órdenes lineales al caso de preórdenes completos (véanse Smith [17], Black [2] y Cook y Seiford [7], entre otros).

2.1.1. Órdenes lineales

Dado un perfil de órdenes lineales $(R_1, \dots, R_m) \in L(X)^m$, consideramos la aplicación $o_i : X \longrightarrow \{1, \dots, n\}$ que muestra la posición que ocupa cada alternativa en la opinión del votante v_i , es decir, $o_i(x_j)$ es la posición de la alternativa x_j en R_i . Por tanto, el vector $(o_i(x_1), \dots, o_i(x_n)) \in \{1, \dots, n\}^n$ determina el orden lineal subyacente. En consecuencia, existe una equivalencia entre $L(X)$ y las permutaciones en $\{1, \dots, n\}$. Así, para $n = 4$, el vector $(3, 1, 4, 2)$ denota que x_1 ocupa la tercera posición, x_2 la primera, x_3 la cuarta y x_4 la segunda.

2.1.2. Preórdenes completos

Para el caso de preórdenes completos no existe un único procedimiento de codificación. Nosotros proponemos uno que corresponde a linealizar el preorden completo y repartir las posiciones dentro de un mismo estrato (clase de equivalencia) mediante la media aritmética. Consideremos el siguiente ejemplo, donde se ordenan 7 alternativas mediante el siguiente preorden completo (a la derecha se consigna un orden lineal asociado al preorden completo):

$$\begin{array}{ccc}
 & & x_2 \\
 & & x_3 \\
 x_2 & x_3 & x_5 \\
 & x_1 & \\
 x_4 & & x_7 \\
 & x_6 & \\
 & & x_2 \\
 & & x_3 \\
 & & x_5 \\
 & & x_1 \\
 & & x_4 \\
 & & x_7 \\
 & & x_6
 \end{array} \longrightarrow$$

Tomando la media aritmética de las posiciones que ocupan las alternativas de la misma clase de equivalencia en el orden lineal asociado, obtenemos la siguiente codificación:

$$\begin{aligned}
 o_i(x_2) = o_i(x_3) = o_i(x_5) &= \frac{1 + 2 + 3}{3} = 2; & o_i(x_1) &= 4; \\
 o_i(x_4) = o_i(x_7) &= \frac{5 + 6}{2} = 5,5; & o_i(x_6) &= 7.
 \end{aligned}$$

De esta forma, el preorden completo queda codificado mediante el vector

$$(4, 2, 2, 5,5, 2, 7, 5,5).$$

Basándonos en esta idea, podemos introducir una aplicación que nos permita codificar preórdenes completos. Sea $(R_1, \dots, R_m) \in W(X)^m$ un perfil de preórdenes completos. La aplicación

$$o_i : X \longrightarrow \{1, 1,5, 2, 2,5, \dots, n - 0,5, n\}$$

asigna la posición relativa de cada alternativa para el votante v_i , es decir, $o_i(x_j)$ representa la posición relativa de la alternativa x_j en R_i . Teniendo en cuenta

$$\sum_{j=1}^n o_i(x_j) = 1 + 2 + \dots + n = \frac{n(n+1)}{2},$$

definimos

$$\mathcal{N}_n = \left\{ (t_1, \dots, t_n) \in \{1, 1,5, 2, 2,5, \dots, n - 0,5, n\}^n \mid \sum_{j=1}^n t_j = \frac{n(n+1)}{2} \right\}.$$

Análogamente, dado un vector $(t_1, \dots, t_n) \in \mathcal{N}_n$, éste define de manera unívoca un preorden completo R de la siguiente forma: $x_j R x_k \Leftrightarrow t_j \leq t_k$. Por ejemplo, para $n = 4$, el vector $(4, 1, 2,5, 2,5)$ indica que x_1 ocupa la última posición, x_2 la primera y x_3 y x_4 posiciones intermedias con empate entre ellas.

2.2. Preferencia colectiva

Dado un perfil de preórdenes completos $(R_1, \dots, R_m) \in W(X)^m$, consideramos la aplicación $o^* : X \longrightarrow \mathbb{Q}$ definida por

$$o^*(x_j) = \frac{o_1(x_j) + \dots + o_m(x_j)}{m},$$

que asigna la posición media que ocupa cada alternativa. Conviene destacar que el preorden completo R^* definido por

$$x_j R^* x_k \Leftrightarrow o^*(x_j) \leq o^*(x_k)$$

corresponde a la regla de Borda.

2.3. Distancias

Consideremos la distancia euclídea $d : \mathbb{R}^n \times \mathbb{R}^n \longrightarrow \mathbb{R}$, definida por

$$d((a_1, \dots, a_n), (b_1, \dots, b_n)) = \sqrt{\sum_{i=1}^n (a_i - b_i)^2}.$$

A partir de ella, introducimos una distancia en el conjunto de preórdenes completos.

Definición 2.1 La aplicación $\bar{d} : W(X) \times W(X) \longrightarrow \mathbb{R}$ definida por

$$\begin{aligned} \bar{d}(R_1, R_2) &= d((o_1(x_1), \dots, o_1(x_n)), (o_2(x_1), \dots, o_2(x_n))) \\ &= \sqrt{\sum_{j=1}^n (o_1(x_j) - o_2(x_j))^2}, \end{aligned}$$

para cualesquiera $R_1, R_2 \in W(X)$, se denomina *distancia euclídea entre preórdenes completos*.

Ejemplo 2.2 Para ilustrar lo anterior, consideremos los siguientes preórdenes completos sobre $X = \{x_1, \dots, x_7\}$ correspondientes a tres votantes:

R_1	R_2	R_3
x_2 x_3 x_5	x_1 x_4	x_2
x_1	x_7	x_1 x_3
x_4 x_7	x_2	x_5
x_6	x_3 x_5 x_6	x_4 x_6 x_7

Estos preórdenes completos quedan codificados de la siguiente manera:

$$R_1 \equiv (o_1(x_1), o_1(x_2), o_1(x_3), o_1(x_4), o_1(x_5), o_1(x_6), o_1(x_7)) = (4, 2, 2, 5, 5, 2, 7, 5, 5),$$

$$R_2 \equiv (o_2(x_1), o_2(x_2), o_2(x_3), o_2(x_4), o_2(x_5), o_2(x_6), o_2(x_7)) = (1, 5, 4, 6, 1, 5, 6, 6, 3),$$

$$R_3 \equiv (o_3(x_1), o_3(x_2), o_3(x_3), o_3(x_4), o_3(x_5), o_3(x_6), o_3(x_7)) = (2, 5, 1, 2, 5, 6, 4, 6, 6).$$

Entonces, la distancia euclídea entre ellos es:

$$\bar{d}(R_1, R_2) = d((4, 2, 2, 5, 5, 2, 7, 5, 5), (1, 5, 4, 6, 1, 5, 6, 6, 3)) = \sqrt{65,5} = 8.,093 ,$$

$$\bar{d}(R_1, R_3) = d((4, 2, 2, 5, 5, 2, 7, 5, 5), (2, 5, 1, 2, 5, 6, 4, 6, 6)) = \sqrt{9} = 3 ,$$

$$\bar{d}(R_2, R_3) = d((1,5, 4, 6, 1,5, 6, 6, 3), (2,5, 1, 2,5, 6, 4, 6, 6)) = \sqrt{55,5} = 7,449 .$$

Hemos de destacar que la distancia máxima entre preórdenes completos sobre X se corresponde con la distancia entre un orden lineal y su inverso. Por tanto,

$$d((1, 2, 3, 4, 5, 6, 7), (7, 6, 5, 4, 3, 2, 1)) = \sqrt{112} = 10,583.$$

Mediante la regla de Borda, obtenemos la preferencia colectiva R^* :

$$(o^*(x_1), o^*(x_2), o^*(x_3), o^*(x_4), o^*(x_5), o^*(x_6), o^*(x_7)) = (2,666, 2,333, 3,5, 4,333, 4, 6,333, 4,833).$$

Estos valores se corresponden con el orden lineal $x_2, x_1, x_3, x_5, x_4, x_7, x_6$, el cual queda codificado por $R^* \equiv (2, 1, 3, 5, 4, 7, 6)$.

Si procedemos a calcular las distancias entre los tres preórdenes completos iniciales y la preferencia colectiva obtenida:

$$\bar{d}(R_1, R^*) = d((4, 2, 2, 5,5, 2, 7, 5,5), (2, 1, 3, 5, 4, 7, 6)) = \sqrt{10,5} = 3,240 ,$$

$$\bar{d}(R_2, R^*) = d((1,5, 4, 6, 1,5, 6, 6, 3), (2, 1, 3, 5, 4, 7, 6)) = \sqrt{44,5} = 6,670 ,$$

$$\bar{d}(R_3, R^*) = d((2,5, 1, 2,5, 6, 4, 6, 6), (2, 1, 3, 5, 4, 7, 6)) = \sqrt{2,5} = 1,581 .$$

Por tanto, $\bar{d}(R_3, R^*) < \bar{d}(R_1, R^*) < \bar{d}(R_2, R^*)$, es decir, la opinión del tercer votante es la más próxima a la preferencia colectiva.

Observación 2.3 Hemos de señalar que la utilización de diferentes métricas con el fin de agregar preferencias individuales ha sido analizada por diferentes autores (véanse, por ejemplo, Cook, Kress y Seiford [5] y Meskanen y Nurmi [15]). En 1959, Kemeny [13] introdujo una distancia entre preórdenes completos y proporcionó una caracterización axiomática para dicha distancia. Además, Kemeny definió la preferencia colectiva como aquella que minimiza la suma de los cuadrados de las distancias con las preferencias individuales. La regla de Kemeny determina cuál es el orden que requiere el menor número de cambio de pares en las preferencias individuales para obtener dicho orden. En 1978, Cook y Seiford [6] comprobaron que la métrica

l^1 aplicada sobre el conjunto de órdenes lineales es la única que satisface ciertos axiomas, y demostraron que el problema de encontrar un orden lineal que minimice la distancia con las preferencias individuales (órdenes lineales también) podía ser considerado como un problema de programación lineal. En 1982, Armstrong, Cook y Seiford [1] extendieron los resultados anteriores al caso de preórdenes completos. Cook y Seiford [7] propusieron el método de la *varianza mínima* en el caso de que las preferencias individuales vengan dadas por preórdenes completos (para órdenes lineales su método coincide con la regla de Borda). Conviene señalar que los procedimientos mencionados, basados en distancias, conllevan un elevado esfuerzo computacional.

3. Regla de Borda ponderada

En esta sección introducimos una regla de Borda modificada que resulta de ponderar las opiniones de los votantes mediante índices que miden la contribución al consenso de cada uno de ellos. La opinión de aquellos votantes cuyas preferencias se encuentren más próximas a la media (obtenidas mediante la regla de Borda) tendrán un mayor peso que la de aquéllos que estén más alejados de la preferencia colectiva. Cabe sospechar que, desde un punto de vista conductista, esto implicará que los votantes moderarán sus preferencias para no resultar penalizados.

3.1. Medidas de consenso

Cuando un grupo de votantes muestra sus preferencias sobre un conjunto de alternativas es claro que el máximo consenso se consigue cuando todas las opiniones son coincidentes. Puesto que no es frecuente que exista unanimidad entre los votantes, resultaría interesante disponer de alguna medida que indicara el grado de acuerdo existente en cada grupo de votantes respecto de cada perfil de preferencias individuales. Dado que existen muy diversas formas de llevar a cabo esta tarea, presentamos una noción general de medida de consenso, basada en el cumplimiento de varios axiomas, en el supuesto de que los votantes muestren sus preferencias median-

te preórdenes completos y que las posiciones que ocupe cada una de las alternativas en las preferencias individuales sea relevante.

Con el fin de introducir la noción de medida de consenso, será necesaria la siguiente notación. Dados un perfil $\mathbf{R} = (R_1, \dots, R_m) \in W(X)^m$, una permutación $\pi : \{1, \dots, m\} \longrightarrow \{1, \dots, m\}$ y $\emptyset \neq I \subseteq V$, denotaremos $\mathbf{R}_\pi = (R_{\pi(1)}, \dots, R_{\pi(m)})$ e $I_\pi = \{v_{\pi(i)} \mid v_i \in I\}$. Dada una permutación $\sigma : \{1, \dots, n\} \longrightarrow \{1, \dots, n\}$, denotaremos con $\mathbf{R}^\sigma = (R_1^\sigma, \dots, R_m^\sigma)$ el perfil que resulta de renombrar en \mathbf{R} las alternativas de acuerdo con σ , es decir, $x_i R_k x_j$ si y sólo si $x_{\sigma(i)} R_k^\sigma x_{\sigma(j)}$ para cualesquiera $i, j \in \{1, \dots, n\}$ y $k \in \{1, \dots, m\}$.

A continuación introducimos una noción de medida de consenso que guarda relación con la dada por Bosch [4].

Definición 3.1 Una *medida de consenso* es una aplicación

$$\mathcal{M} : W(X)^m \times \mathcal{P}(V) \longrightarrow [0, 1]$$

que satisface las siguientes condiciones:

1. *Unanimidad.* Para cualesquiera $\mathbf{R} \in W(X)^m$ y $\emptyset \neq I \subseteq V$ se verifica

$$\mathcal{M}(\mathbf{R}, I) = 1 \quad \Leftrightarrow \quad R_i = R_j \quad \text{para cualesquiera } v_i, v_j \in I.$$

2. *Anonimato.* Para cualesquiera permutación $\pi : \{1, \dots, m\} \longrightarrow \{1, \dots, m\}$, $\mathbf{R} \in W(X)^m$ y $\emptyset \neq I \subseteq V$ se verifica

$$\mathcal{M}(\mathbf{R}, I) = \mathcal{M}(\mathbf{R}_\pi, I_\pi).$$

3. *Neutralidad.* Para cualesquiera permutación $\sigma : \{1, \dots, n\} \longrightarrow \{1, \dots, n\}$, $\mathbf{R} \in W(X)^m$ y $\emptyset \neq I \subseteq V$ se verifica

$$\mathcal{M}(\mathbf{R}, I) = \mathcal{M}(\mathbf{R}^\sigma, I).$$

4. *Reciprocidad.* Para cualesquiera $\mathbf{R} \in W(X)^m$ y $\emptyset \neq I \subseteq V$ se verifica

$$\mathcal{M}(\mathbf{R}, I) = \mathcal{M}(\mathbf{R}^{-1}, I).$$

Los axiomas recién introducidos tienen un claro significado. Así, el de unanimidad quiere decir que el máximo consenso, 1, sólo se consigue cuando todas las preferencias individuales coinciden. El anonimato refleja el hecho de que el consenso no depende de qué votantes muestran sus preferencias sino de estas mismas. La neutralidad refleja un tratamiento equitativo para las alternativas, de forma que el consenso no dependa del nombre de éstas. Finalmente, el axioma de reciprocidad muestra que el consenso existente entre un grupo de votantes es el mismo que el correspondiente al que hay cuando este grupo invierte sus preferencias entre las alternativas.

Conviene destacar que Bosch [4] considera únicamente perfiles de órdenes lineales y no exige la condición de reciprocidad en su definición de medida de consenso. De hecho, la reciprocidad sólo tiene sentido cuando son relevantes todas las posiciones de las alternativas en cada una de las preferencias individuales. Por ejemplo, si las preferencias individuales se agregan mediante la regla de pluralidad (vencen aquellas alternativas que obtienen mayor número de primeras posiciones), entonces la única información que interesa conocer para determinar el consenso es la alternativa que ocupa la primera posición en cada votante.

De acuerdo con Bosch [4], máximo consenso y mínimo consenso (máximo disenso) no son conceptos simétricos. Por este motivo no se determina cuándo la medida de consenso es nula. Por otra parte, hemos de señalar que del axioma de unanimidad se deriva $\mathcal{M}(\mathbf{R}, \{v_k\}) = 1$, para cualesquiera $\mathbf{R} \in W(X)^m$ y $v_k \in V$.

A continuación proponemos una medida de consenso para perfiles de preórdenes completos, bajo el supuesto de que todas las posiciones que ocupan las alternativas en cada preferencia individual son relevantes. Está basada en la regla de Borda, la cual genera la preferencia colectiva (que también es un preorden completo) y la distancia euclídea entre preórdenes completos (Definición 2.1), gracias a la cual se compararán las preferencias individuales con la colectiva.

Definición 3.2 La *medida de consenso de Borda* es la aplicación

$$\mathcal{M}_B : W(X)^m \times \mathcal{P}(V) \longrightarrow [0, 1]$$

definida por

$$\mathcal{M}_B(\mathbf{R}, I) = \begin{cases} 1 - \frac{\sum_{v_i \in I} \bar{d}(R_i, R^*)}{|I| \cdot \Delta_n}, & \text{if } I \neq \emptyset, \\ 0, & \text{if } I = \emptyset, \end{cases}$$

siendo R^* la preferencia colectiva obtenida mediante la regla de Borda y

$$\Delta_n = \text{máx}\{\bar{d}(R_i, R_j) \mid R_i, R_j \in W(X)\}.$$

Conviene señalar que $\Delta_n = \bar{d}(R, R^{-1})$, para cada $R \in L(X)$. El Cuadro 1 muestra algunos de los valores que toma Δ_n .

Cuadro 1: Algunos valores de Δ_n

n	Δ_n
2	$\sqrt{2} = 1,414$
3	$\sqrt{8} = 2,828$
4	$\sqrt{20} = 4,472$
5	$\sqrt{40} = 6,324$
6	$\sqrt{70} = 8,366$
7	$\sqrt{112} = 10,583$
8	$\sqrt{168} = 12,961$
9	$\sqrt{240} = 15,491$
10	$\sqrt{330} = 18,165$

Observación 3.3 La utilización de la métrica euclídea en la Definición 3.2 está directamente relacionada con el hecho de haber elegido la regla de Borda como método inicial de agregación de preferencias individuales. Tal como ha sido advertido anteriormente, la regla de Borda proporciona un preorden completo que ordena las alternativas a través de la media aritmética de las posiciones que ocupan las alternativas en las preferencias individuales. Dado que la media aritmética minimiza la suma de las distancias a los valores individuales respecto de la métrica euclídea, parece lógico utilizar dicha métrica para calcular las distancias entre las preferencias individuales y la colectiva.

Proposición 3.4 \mathcal{M}_B es una medida de consenso.

DEMOSTRACIÓN: Una comprobación rutinaria permite asegurar que \mathcal{M}_B toma sus valores en el intervalo $[0, 1]$ y verifica los axiomas de unanimidad, anonimato, neutralidad y reciprocidad. ■

3.2. Contribución marginal al consenso

A continuación introducimos un índice que mide la contribución marginal al consenso de cada votante respecto de un perfil de preórdenes completos dado. Este índice se obtiene al sumar las diferencias entre el consenso que hay entre cada grupo de votantes al que pertenezca el votante en cuestión y el consenso que hay si tal votante deja de pertenecer al grupo. Nótese que dicho índice es válido para cualquier medida de consenso.

Definición 3.5 Dada una medida de consenso \mathcal{M} , la *contribución marginal al consenso* del votante v_i respecto de un perfil $\mathbf{R} \in W(X)^m$ se define como:

$$w_i = \sum_{I \subseteq V} \left(\mathcal{M}(\mathbf{R}, I) - \mathcal{M}(\mathbf{R}, I \setminus \{v_i\}) \right).$$

Obviamente, si $v_i \notin I$, entonces $\mathcal{M}(\mathbf{R}, I) - \mathcal{M}(\mathbf{R}, I \setminus \{v_i\}) = 0$. Si $w_i > 0$, diremos que el votante v_i contribuye positivamente al consenso; por el contrario, si $w_i < 0$, diremos que el votante v_i contribuye negativamente al consenso.

A partir de ahora consideraremos siempre la media de consenso de Borda \mathcal{M}_B .

3.3. Reglas de Borda ponderadas

A continuación introducimos una nueva clase de reglas de Borda ponderadas, a través de la cual se determinará una preferencia colectiva que tiene en cuenta la contribución marginal al consenso de cada uno de los votantes.

Definición 3.6 La regla de Borda ponderada asociada a $\mathbf{w} = (w_1, \dots, w_m) \in \mathbb{R}^m$, $R^{\mathbf{w}}$, se define como

$$x_j R^{\mathbf{w}} x_k \Leftrightarrow o^{\mathbf{w}}(x_j) \leq o^{\mathbf{w}}(x_k),$$

donde

$$o^{\mathbf{w}}(x_j) = \frac{w_1 \cdot o_1(x_j) + \dots + w_m \cdot o_m(x_j)}{m}.$$

Resulta inmediato comprobar que $R^{\mathbf{w}}$ es siempre un preorden completo.

Ejemplo 3.7 Consideremos el perfil sobre $X = \{x_1, \dots, x_7\}$ incluido en el Ejemplo 2.2:

R_1			R_2			R_3		
x_2	x_3	x_5	x_1	x_4	x_2			
	x_1			x_7	x_1	x_3		
x_4		x_7		x_2		x_5		
	x_6		x_3	x_5	x_6	x_4	x_6	x_7

Recordemos que aplicando la regla de Borda, obtenemos el orden lineal R^* :

$$x_2, x_1, x_3, x_5, x_4, x_7, x_6.$$

Las contribuciones marginales al consenso de cada uno de los votantes, vienen representadas por el vector $\mathbf{w} = (w_1, w_2, w_3) = (0,805, -0,167, 1,275)$. Hemos de destacar que los votantes primero y tercero tienen una contribución marginal al consenso positiva, mientras que la contribución del segundo votante es negativa, lo cual proporcionará un efecto negativo sobre su elección al multiplicar las posiciones de las alternativas por $-0,167$. Además, la opinión del tercer votante cuenta $1,275/0,805 = 1,583$ veces más que la del primero.

A partir de la Definición 3.6, la regla de Borda ponderada asociada a \mathbf{w} nos proporciona el orden lineal $R^{\mathbf{w}}$: $x_2, x_3, x_5, x_1, x_7, x_4, x_6$. Si bien x_2 se mantiene como vencedora, se producen ciertos cambios, como por ejemplo que x_1 pasa de la segunda posición a ocupar la cuarta y x_3 de la tercera a la segunda.

Dado que los valores negativos de w_i pueden alterar de forma artificial el ordenamiento de $R^{\mathbf{w}}$, tomaremos como vector de ponderación $\mathbf{w}' = (w'_1, \dots, w'_m)$, donde $w'_i = \max\{w_i, 0\}$.

Ejemplo 3.8 Consideremos los ejemplos 2.2 y 3.7. En este caso, la regla de Borda ponderada asociada a $\mathbf{w}' = (w'_1, w'_2, w'_3) = (0,805, 0, 1,275)$, $R^{\mathbf{w}'}$, nos proporciona el preorden $x_2, x_3, x_1, x_5, x_4 - x_7, x_6$, donde $x_4 - x_7$ significa que las alternativas x_4 y x_7 empatan. En este caso no se ha considerado la opinión del segundo votante y se han mantenido las preferencias expresadas al inicio por el primer y el tercer votante. El preorden obtenido en este caso es diferente al obtenido para $R^{\mathbf{w}}$.

3.4. Procedimiento secuencial

A partir de las ideas anteriores, proponemos un procedimiento secuencial en el cual se determina un preorden completo final sobre el conjunto de alternativas a partir de las opiniones facilitadas por los votantes. Este procedimiento consta de los siguientes pasos:

1. Dado un conjunto de votantes $V = \{v_1, \dots, v_m\}$ que muestran sus preferencias sobre un conjunto de alternativas $X = \{x_1, \dots, x_n\}$ mediante preórdenes completos, se codifican sus opiniones a través de vectores que indican las posiciones relativas de cada alternativa (2.1).
2. Se agregan las posiciones que ocupan las alternativas en los preórdenes completos individuales a través de la media aritmética para obtener el preorden completo colectivo (2.2). Esto equivale a utilizar la regla de Borda extendida utilizada, entre otros, por Smith [17].
3. Se calculan las distancias entre los vectores que representan las preferencias individuales y la colectiva. A partir de estas distancias se calcula la contribución marginal al consenso de cada uno de los votantes (Definición 3.5): w_1, \dots, w_m .
 - a) Si $w_i \geq 0$ para cualquier $i \in \{1, \dots, m\}$, entonces se ponderan las posiciones que ocupan las alternativas por los índices w_i correspondientes a los votantes y se determina el preorden completo que representa la preferencia colectiva final mediante la regla de Borda ponderada (Definición 3.6).

- b) En caso contrario, se eliminan aquellos votantes cuya contribución marginal al consenso sea negativa, reiniciando el procedimiento de decisión para el resto de votantes $V' = \{v_i \in V \mid w_i \geq 0\}$.

Conviene destacar que la convergencia de este procedimiento está garantizada, puesto que en cada fase del proceso secuencial disminuye o se estabiliza el conjunto de votantes cuyas opiniones son tenidas en cuenta.

3.5. Un ejemplo ilustrativo

Consideremos que el conjunto de votantes $V = \{v_1, v_2, v_3, v_4, v_5\}$ ha ordenado el conjunto de alternativas $X = \{x_1, \dots, x_7\}$ mediante los órdenes lineales mostrados en el Cuadro 2.

Cuadro 2: Órdenes individuales

R_1	R_2	R_3	R_4	R_5
x_1	x_2	x_1	x_1	x_7
x_4	x_1	x_6	x_6	x_6
x_2	x_7	x_7	x_5	x_5
x_3	x_4	x_4	x_4	x_3
x_7	x_3	x_2	x_3	x_4
x_6	x_6	x_5	x_7	x_2
x_5	x_5	x_3	x_2	x_1

La relación de preferencia colectiva generada por la regla de Borda puede obtenerse calculando la media aritmética de las posiciones que ocupan las alternativas en las preferencias individuales, que viene representada por el vector:

$$(o^*(x_1), o^*(x_2), o^*(x_3), o^*(x_4), o^*(x_5), o^*(x_6), o^*(x_7)) = (2,4, 4,4, 5, 3,8, 5,2, 3,6, 3,6).$$

Estos valores se corresponden con el preorden completo:

$$x_1, x_6 - x_7, x_4, x_2, x_3, x_5. \tag{1}$$

A partir de de las opiniones individuales mostradas en el Cuadro 2 y el orden que representa la preferencia colectiva (1), las contribuciones marginales al consenso son:

$$\mathbf{w}^1 = (w_1^1, w_2^1, w_3^1, w_4^1, w_5^1) = (0,363, 0,363, 3,157, -1,234). \quad (2)$$

Puesto que la contribución marginal al consenso del quinto votante es negativa, en el siguiente paso su opinión no será considerada. Tomando las columnas del Cuadro 2 correspondientes al nuevo conjunto de votantes $V' = \{v_1, v_2, v_3, v_4\}$, obtenemos el vector:

$$\begin{aligned} & (o^*(x_1), o^*(x_2), o^*(x_3), o^*(x_4), o^*(x_5), o^*(x_6), o^*(x_7)) = \\ & (1,25, 4, 5,25, 3,5, 5,75, 4, 4,25). \end{aligned}$$

Estos valores se corresponden con el preorden completo:

$$x_1, x_4, x_2 - x_6, x_7, x_3, x_5. \quad (3)$$

A partir de las preferencias de V' y la preferencia colectiva (3), la contribución marginal al consenso de los votantes considerados es ahora:

$$\mathbf{w}^2 = (w_1^2, w_2^2, w_3^2, w_4^2) = (1,115, 0,453, 0,865, -0,119). \quad (4)$$

De nuevo aparece un votante cuya contribución marginal al consenso es negativa, por lo que se procede a su eliminación. Reiniciamos el proceso de decisión considerando el nuevo conjunto de votantes $V'' = \{v_1, v_2, v_3\}$, para los cuales se obtiene el vector:

$$\begin{aligned} & (o^*(x_1), o^*(x_2), o^*(x_3), o^*(x_4), o^*(x_5), o^*(x_6), o^*(x_7)) = \\ & (1,333, 3, 5,333, 3,333, 6,666, 4,666, 3,666). \end{aligned}$$

La regla de Borda nos proporciona el orden lineal:

$$x_1, x_2, x_4, x_7, x_6, x_3, x_5. \quad (5)$$

Ahora consideramos las preferencias individuales de V'' y la preferencia colectiva (5). En este caso, las contribuciones marginales al consenso son:

$$\mathbf{w}^3 = (w_1^3, w_2^3, w_3^3) = (0,826, 0,933, 0,298). \quad (6)$$

Puesto que todas ellas son positivas, utilizamos la Definición 3.6 para obtener la relación de preferencia colectiva generada por la regla de Borda ponderada asociada a \boldsymbol{w}^3 , la cual finaliza el proceso de decisión mediante el siguiente orden lineal:

$$x_1, x_2, x_4, x_7, x_3, x_6, x_5. \quad (7)$$

Conviene destacar la significativa diferencia existente entre las preferencias colectivas obtenidas en cada uno de los pasos del proceso de decisión: (1), (3), (5) y (7).

Inicialmente, (1) señala el orden suministrado por la regla de Borda clásica (todas las opiniones de los votantes tienen el mismo peso). Tomando este preorden completo como referente de la opinión colectiva de todos los votantes, en (2) observamos que el quinto votante contribuye negativamente al consenso. Después de su eliminación, obtenemos el nuevo orden (3) así como los nuevos pesos en (4), los cuales son bien distintos a los obtenidos en (2). La razón reside en que la opinión de v_5 distorsionaba sustancialmente los cálculos.

En la siguiente iteración, hemos eliminado al votante v_4 debido a su contribución negativa al consenso en el nuevo grupo de votantes $V' = \{v_1, v_2, v_3, v_4\}$, se ha obtenido el nuevo orden colectivo (5) y los nuevos pesos (6). Puesto que todos ellos son positivos, es el momento de aplicar nuestra regla de Borda ponderada (Definición 3.6), y así obtenemos (7) como resultado final.

Es importante destacar que la contribución marginal al consenso varía en cada iteración, en función del nuevo orden colectivo obtenido por la regla de Borda. Por ejemplo, v_3 posee la mayor contribución marginal en (2), $w_3^1 = 3,157$; sin embargo, este votante tiene su menor contribución marginal al consenso en (6), $w_3^3 = 0,298$.

4. Conclusiones y futuras líneas de investigación

En este trabajo hemos introducido una medida de consenso que permite conocer el grado de acuerdo existente en cada grupo de votantes respecto de la preferencia colectiva obtenida a través de la regla de Borda. A partir de esta medida hemos considerado la contribución marginal al consenso de cada votante y hemos introducido

una regla de Borda ponderada en la que las opiniones individuales son ponderadas por las correspondientes contribuciones marginales al consenso. Esta regla se aplica una vez finalizado un proceso iterativo en el que se van eliminando paulatinamente las opiniones de aquellos votantes que no contribuyen al consenso.

Hemos de remarcar que la secuencialidad del procedimiento de decisión colectiva introducido no requiere de los votantes sucesivas manifestaciones de sus preferencias ni se solicita de ellos que las modifiquen. De hecho los votantes sólo han de mostrar sus preferencias sobre las alternativas una sola vez, proporcionando un preorden completo, y es el proceso recursivo el que se encarga de llevar a cabo la selección de votantes que aportan consenso al grupo y de ponderar sus opiniones según su contribución marginal al consenso.

Está claro que nuestro procedimiento de votación tiene un efecto de moderación en los votantes, ya que si éstos conocen su funcionamiento intentarán mostrar opiniones no excesivamente discrepantes, pues de lo contrario corren el riesgo de ser eliminados del proceso o de ver penalizada su influencia en la preferencia colectiva que se trata de determinar. Esto no ocurre con la regla de Borda ni con otros sistemas de votación, en los cuales los agentes pueden sentir la tentación de mostrar preferencias muy discrepantes respecto de la opinión mayoritaria con objeto de modificar artificialmente la preferencia colectiva.

Una pregunta natural que nos podemos plantear es si este nuevo sistema de votación mejora la regla de Borda original aportando nuevas características dentro del ámbito de la Elección Social.

A continuación llevamos a cabo un análisis preliminar sobre dos importantes condiciones a tener en cuenta, como son la eficiencia de Condorcet y la propiedad denominada *no veto power*. Esta última significa que si x_i es la alternativa mejor valorada para al menos $m - 1$ votantes ($m \geq 3$), entonces x_i debería ser la alternativa ganadora. La regla de Borda, como es bien sabido, no cumple esta condición. Por ejemplo (véase Maskin y Sjöström [14]), si se considera que un votante ordena las alternativas de forma que x_2 es la primera y x_1 la última, y el resto de los votantes ordenan x_1 en primer lugar y x_2 en la segunda posición: $R_1 \equiv (n, 1, \dots)$,

$R_2 = \dots = R_m \equiv (1, 2, \dots)$. Si $n > m$, entonces x_2 vence a x_1 por la regla de Borda:

$$o^*(x_2) = \frac{1 + (m-1) \cdot 2}{m} < \frac{n + (m-1) \cdot 1}{m} = o^*(x_1).$$

En cuanto a la eficiencia de Condorcet, una alternativa se dice que es vencedora de Condorcet si ésta vence al resto de las alternativas por mayoría simple (cuando dicha alternativa se confronta por pares a cada una de las demás, ésta tiene en todos los casos más votos que las restantes). La regla de Borda, así como otros muchos sistemas de votación, no dan como vencedora necesariamente a la alternativa vencedora de Condorcet (cuando esta existe).

Ejemplo 4.1 Sea $(R_1, R_2, R_3) \in L(\{x_1, x_2, x_3, x_4\})^3$ el perfil de preferencias detallado más adelante. Dicho perfil, es un ejemplo simple de violación de la condición *no veto power* por la regla de Borda. Además, x_1 es la ganadora de Condorcet, mientras que x_2 vence a x_1 aplicando la regla de Borda, R^* . En cambio, si consideramos la regla de Borda ponderada (Definición 3.6) asociada al vector de contribuciones marginales al consenso, en este caso $\mathbf{w} = (0,838, 0,838, 0,143)$, entonces esta regla, R^w , selecciona como ganadora a x_1 , que es precisamente la alternativa vencedora de Condorcet.

R_1	R_2	R_3	R^*	R^w
x_1	x_1	x_2	x_2	x_1
x_2	x_2	x_3	x_1	x_2
x_3	x_3	x_4	x_3	x_3
x_4	x_4	x_1	x_4	x_4

En este ejemplo, la alternativa vencedora de Condorcet gana a las restantes alternativas y, además, no se viola la condición *no veto power*. Conviene analizar si estas características pueden generalizarse o si el procedimiento secuencial introducido mejora la regla de Borda original respecto a estas dos condiciones.

Nuestra propuesta considera la regla de Borda como procedimiento inicial de agregación de las preferencias individuales. Puesto que la regla de Borda ordena las alternativas considerando la media aritmética de sus posiciones y la media aritmética

está asociada a la distancia euclídea, parece razonable extender nuestro procedimiento secuencial de decisión colectiva a otros operadores de agregación y sus posibles distancias asociadas (si las tienen) u otras pertinentes.

Por otro lado, la regla de Borda es un caso particular de regla de puntuación (*scoring rule*), en donde cada votante asigna una determinada puntuación a cada alternativa de acuerdo con la posición que ésta ocupe en sus preferencias. Una propuesta interesante a estudiar consiste en generalizar nuestro procedimiento secuencial de decisión colectiva al caso general de reglas de puntuación. No obstante, conviene tener presente que la medida de consenso definida en cada caso ha de ser coherente con la regla de puntuación considerada.

Agradecimientos

Este trabajo tiene su inicio en una estancia del primer autor en la Universidad de Tilburg, el cual agradece la hospitalidad de Harrie de Swart y a Rob Bosch por facilitarle su tesis doctoral. El primer autor agradece la ayuda financiera de la Junta de Castilla y León (Consejería de Educación y Cultura, Proyecto VA040A05), del Ministerio de Educación y Ciencia (Plan Nacional de Investigación Científica, Desarrollo e Innovación Tecnológica, Proyecto SEJ2006-024267/ECON) y de FEDER. Los autores desean agradecer a Miguel Ángel Ballester, Carmen Beviá, Luis Corchón y Miguel Martínez Panero sus comentarios y sugerencias.

Referencias

- [1] Armstrong, R.D., Cook, W.D., Seiford, L.M. (1982): Priority ranking and consensus formation: The case of ties. *Management Science* 8, pp. 638-645.
- [2] Black, D. (1976): Partial justification of the Borda count. *Public Choice* 28, pp. 1-15.
- [3] Borda, J.C. de (1784): *Mémoire sur les élections au scrutin*. Histoire de l'Académie Royale des Sciences, Paris.

- [4] Bosch, R. (2005): *Characterizations of Voting Rules and Consensus Measures*. Ph. D. Dissertation, Tilburg University.
- [5] Cook, W.D., Kress, M., Seiford, L.M. (1996): A general framework for distance-based consensus in ordinal ranking models. *European Journal of Operational Research* 96, pp. 392-397.
- [6] Cook, W.D., Seiford, L.M. (1978): Priority ranking and consensus formation. *Management Science* 24, pp. 1721-1732.
- [7] Cook, W.D., Seiford, L.M. (1982): On the Borda-Kendall consensus method for priority ranking problems. *Management Science* 28, pp. 621-637.
- [8] Dummett, M. (1984): *Voting Procedures*. Clarendon Press, Oxford.
- [9] Dummett, M. (1998): The Borda Count and Agenda Manipulation. *Social Choice and Welfare* 15, pp. 289-296.
- [10] Eklund, P., Rusinowska, A., de Swart, H. (2007): Consensus reaching in committees. *European Journal of Operational Research* 178, pp. 185-193.
- [11] Fine, B., Fine, K. (1974): Social choice and individual ranking I. *Review of Economic Studies* 41, pp. 459-475.
- [12] Gärdenfors, P. (1973): Positionalist voting functions. *Theory and Decision* 4, pp. 1-24.
- [13] Kemeny, J. (1959): Mathematics without numbers. *Daedalus* 88, pp. 571-591.
- [14] Maskin, E.S., Sjöström, T. (2002): Implementation theory. En: *Handbook of Social Choice and Welfare* Vol. 1 (eds.: K.J. Arrow, A.K. Sen, K. Suzumura). Elsevier, pp. 237-288.
- [15] Meskanen, T., Nurmi, H. (2006): Distance from consensus: A theme and variations. En: *Mathematics and Democracy, Recent Advances in Voting Systems and Collective Choice*. Studies in Choice and Welfare. (eds.: B. Simeone, F. Pukelsheim). Springer Verlag, pp. 117-132.
- [16] Saari, D.G. (1995): *Basic Geometry of Voting*. Springer Verlag, Berlin.
- [17] Smith, J. (1973): Aggregation of preferences with variable electorate. *Econometrica* 41, pp. 1027-1041.

- [18] Young, H.P. (1974): An axiomatization of Borda's rule. *Journal of Economic Theory* 9, pp. 43-52.
- [19] Young, H.P. (1997): Group choice and individual judgements. En: *Perspectives on Public Choice: A Handbook* (ed.: D. Mueller). Cambridge University Press, Cambridge, pp. 181-200.
- [20] Zadrozny, S. (1997): An approach to the consensus reaching support in fuzzy environment. En: *Consensus under Fuzziness* (eds. J.Kacprzyk, H. Nurmi, M. Fedrizzi). Kluwer Academic Publishers, pp. 83-109.

CLASIFICACION DE EMPRESAS SEGÚN SU SOLVENCIA A PARTIR DE INFORMACION CONTABLE

ERNESTO JESUS VERES FERRER)
e-mail: Ernesto.Veres@uv.es

Departamento de ECONOMIA APLICADA
UNIVERSIDAD DE VALENCIA

Área temática: Métodos cuantitativos.

Resumen

La correcta clasificación de la solvencia de una empresa que solicita un crédito es relevante para la entidad crediticia que lo concede. A partir de la información contable de la situación empresarial puede obtenerse una gran batería de ratios. Debidamente utilizados, el trabajo presenta un proceso de decisión para clasificar las empresas según su situación de solvencia, permitiendo asignar su calificación como sana o posiblemente fracasada. El criterio básico atiende al número de ratios que clasifican a la empresa en una u otra categoría, entendiendo que un elevado número de ellos que apunten hacia una situación o no de solvencia implican, probablemente, la confirmación de esa situación.

El trabajo finaliza aplicando el proceso de clasificación a una muestra de pequeñas y medianas empresas con actividad o que cesaron en la misma durante el año 2004 en la Comunidad Valenciana.

Palabras clave: clasificación, PYME, ratio contable, solvencia .

Abstract

The correct classification of the company's solvency that asks for a credit is outstanding for the credit organization that grants it. From the countable information of the enterprise situation a great battery of ratios can be obtained. Properly used, the work presents a decision process to classify the companies according to its solvency situation, being allowed to possibly assign its failed qualification as it heals or. The basic criterion takes care of the number of ratios that they classify to the company in one or another category, understanding that a high number of them who aim towards a situation or not of solvency they imply, probably, the confirmation of that situation.

The work finalizes applying the process of classification to a sample of small and medium companies with activity or that stopped in the same one during year 2004 in the Valencian Community.

Key words: accounting ratio, classification, PYME, solvency.

1. Introducción

La clasificación de solvencia o insolvencia de una empresa a partir de la información contable disponible para la misma sigue siendo un problema abierto, ante la ausencia de una teoría económica que la sustente. Esta circunstancia sitúa a los investigadores ante el dilema de partir de un elevado número de variables con las que captar la estructura multidimensional de la empresa –lo que puede provocar problemas operativos y reducir la utilidad de los modelos propuestos-, o bien considerar un número reducido de variables aunque sea a costa de descuidar numerosos aspectos de la empresa. La evidencia demuestra que en la mayoría de los estudios se opta por considerar un amplio número de variables para luego proceder a su reducción. En este punto debería procederse con mucho cuidado procurando seguir una metodología adecuada ya que la exclusión de una variable relevante puede provocar graves sesgos en el modelo desvirtuando en este caso los resultados obtenidos.

Los comentarios anteriores se enfatizan todavía más en el caso de las pequeñas y medianas empresas, no sujetas a auditoría externa, debido a la calidad deficiente – información no completa o con inconsistencias internas, referencia temporal no exacta para muchas variables contables, etc.- de la información contable de las mismas.

Sin embargo, aún siendo deficiente la información contable, a partir de ella pueden elaborarse una gran batería de ratios contables que miden diversos aspectos de la realidad empresarial. Es evidente que muchas de ellas tienen relación muy estrecha, al estar definidas a partir de las mismas variables, pero su comportamiento no es siempre uniforme o con el mismo sentido. El hecho de que un número elevado de estas ratios apunten hacia una situación o no de insolvencia da pie a confirmar, probablemente, esa situación.

Este trabajo presenta un proceso de decisión para clasificar las empresas según su situación de solvencia, permitiendo asignar su calificación como sana o posiblemente fracasada. El criterio básico atiende al número de ratios que clasifican a la empresa en una u otra categoría. La clasificación se efectúa en etapas. Conforme aumentan éstas la incertidumbre de la clasificación también aumenta. Por ello, es posible detener el proceso, dejando un conjunto de empresas en un “limbo” de no clasificables y que

puedan exigir comprobaciones adicionales no directamente relacionadas con los ratios contables.

El proceso es general, si bien cada aplicación concreta dará lugar a un criterio de clasificación específico. El trabajo finaliza aplicando el proceso de clasificación a una muestra de pequeñas y medianas empresas con actividad o que cesaron en la misma durante el año 2004 en la Comunidad Valenciana.

2. Definición de ratios

En la aplicación realizada en el apartado sexto, se han utilizado un total de 73 ratios, agrupados en 7 categorías según la finalidad contable de los mismos: rentabilidad, estructura financiera, actividad, rotación, liquidez, solvencia a largo plazo y estructura económica. Se pretende que en su conjunto describan suficientemente la situación financiera y contable de las empresas. Se recogen en la tabla siguiente:

RATIO	DEFINICION
Rentabilidad	
REN1	Resultado antes de intereses e impuestos / Activo total
REN2	Resultado del ejercicio / Activo total
REN3	Resultado del ejercicio / Fondos propios
REN4	Resultado del ejercicio / Pasivo total
REN5	Resultado de actividades ordinarias / Activo total
REN6	Resultado antes de intereses e impuestos / Ventas
REN7	Resultado de actividades ordinarias / Pasivo total
REN8	Resultado de actividades ordinarias / Fondos propios
REN9	Resultado del ejercicio / Ventas
REN10	Cash Flow recursos generados / Fondos propios
REN11	Cash Flow recursos generados / Activo total
REN12	Cash Flow recursos generados / Pasivo total
REN13	Cash Flow recursos generados / Pasivo circulante
REN14	Resultado antes de impuestos / Activo total
REN15	Resultado neto – Realizable – Existencias / Activo total
REN16	Cash Flow recursos generados / Ventas
REN17	Resultado antes de intereses e impuestos / Gastos financieros
REN18	Resultado antes de impuestos / Fondos propios
REN19	Resultado antes de impuestos / Ventas
REN20	Resultado antes de impuestos / Pasivo total
REN21	Resultado antes de int. e imp. / Gastos financieros + Pasivo circulante
REN22	Pasivo fijo / Cash Flow recursos generados
REN23	Resultado de actividades ordinarias / Ventas
Estructura financiera	
EF1	Pasivo total / Fondos propios
EF2	Pasivo fijo / Fondos propios
EF3	Pasivo circulante / Fondos propios
EF4	Gastos financieros / Pasivo total

EF5	Pasivo circulante / Pasivo total
EF6	Pasivo fijo / Activo total
EF7	Pasivo circulante / Activo total
EF8	Fondos propios / Pasivo total
EF9	Gastos financieros / Ventas
EF10	Pasivo total / Activo total
EF11	Pasivo fijo / Capital social
Actividad	
ACT1	Valor añadido / Ventas
ACT2	Gastos de personal / Valor añadido
ACT3	Gastos financieros / Valor añadido
ACT4	Ingresos de explotación / consumos de explotación
ACT5	Resultado del ejercicio / Valor añadido
ACT6	Pasivo fijo / Ventas
ACT7	Gastos de personal / Ventas
ACT8	Gastos de personal / Activo fijo
Rotación	
ROT1	Ventas / Activo total
ROT2	Ventas / Activo circulante
ROT3	Ventas / Activo fijo
ROT4	Ventas / Fondos propios + Pasivo fijo
ROT5	Ventas / Pasivo circulante
ROT6	Activo total / Ingresos de explotación
ROT7	Ventas / Capital circulante
ROT8	Ventas / Existencias
ROT9	Ventas / Realizable
Solvencia (liquidez)	
SOLV1	Activo circulante / Pasivo circulante
SOLV2	Activo circulante – Existencias / Pasivo circulante
SOLV3	Disponible / Pasivo circulante
SOLV4	Capital circulante / Pasivo circulante
SOLV5	Capital circulante / Activo total
SOLV6	Capital circulante / Ventas
SOLV7	Realizable / Capital circulante
SOLV8	Activo circulante – Existencias / Ventas
SOLV9	Recursos generados antes de impuestos / Pasivo circulante
SOLV10	(Activo circulante – Existencias) – Pasivo circulante / Gastos de explotación – Amortizaciones – Provisiones (Intervalo sin crédito)
SOLV11	Capital circulante / Ingresos de explotación
SOLV12	Realizable / Ingresos de explotación
Solvencia a largo plazo	
SOLVLP1	Activo fijo / Fondos propios
SOLVLP2	Recursos generados antes de impuestos / Pasivo total
SOLVLP3	Activo total neto / Pasivo total
Estructura económica	
CF1	Existencias / Activo circulante
CF2	Activo circulante / Activo total
CF3	Activo circulante / Activo fijo
CF4	Existencias / Capital circulante
CF5	Existencias / Activo total
CF6	Realizable / Activo total
CF7	Disponible / Activo total

Tabla 1. Ratios contables

Dada la naturaleza del algoritmo que va a proponerse, cuantos más ratios se utilicen más aumenta la capacidad clasificatoria del mismo. Aunque algunos de estos ratios estén correlados entre sí, la no existencia de una correlación teórica perfecta entre ellos determina que todos sean susceptibles de utilizarse en la metodología posterior.

3. Intervalos de clasificación

Para cada uno de los ratios contables se construye un intervalo para la clasificación de una empresa en *sana* (0), *quebrada* (1) o *dudosa* (2). Así, siendo para cada uno de los ratios

- m_0 mínimo del ratio para las empresas sanas
- M_0 máximo del ratio para las empresas sanas
- m_1 mínimo del ratio para las empresas quebradas
- M_1 máximo del ratio para las empresas quebradas

se verifica:

Condición	Intervalo empresa sana: la empresa se clasifica sana si el correspondiente <i>ratio</i> pertenece a	Intervalo empresa quebrada: la empresa se clasifica quebrada si el correspondiente <i>ratio</i> pertenece a	Intervalo empresa no clasificable: la empresa no queda clasificada si el correspondiente <i>ratio</i> pertenece a
Si $M_0 \leq m_1$	$]-\infty, \frac{M_0 + m_1}{2}[$	$]\frac{M_0 + m_1}{2}, +\infty[$	es igual a $\frac{M_0 + m_1}{2}$
Si $M_1 \leq m_0$	$]\frac{M_1 + m_0}{2}, +\infty[$	$]-\infty, \frac{M_1 + m_0}{2}[$	es igual a $\frac{M_1 + m_0}{2}$
Si $m_0 \leq m_1 \wedge M_0 \leq M_1$	$]-\infty, m_1[$	$]M_0, +\infty[$	Resto
Si $m_1 \leq m_0 \wedge M_1 \leq M_0$	$]M_1, +\infty[$	$]-\infty, m_0[$	Resto
Si $m_0 < m_1 \wedge M_1 < M_0$	$]-\infty, m_1[\cup]M_1, +\infty[$		Resto
Si $m_1 < m_0 \wedge M_0 < M_1$		$]-\infty, m_0[\cup]M_0, +\infty[$	Resto

Tabla 2. Definición intervalos de clasificación

El anterior criterio de clasificación se establece, pues, en términos de la coincidencia o no de intervalos donde coexistan ratios para ambos tipos de empresas, sanas o quebradas. En aquellos subintervalos en los que todos los ratios deducidos de la muestra de empresas pertenecen a una de las categorías, la empresa queda clasificada en ella. Y para aquellos subintervalos en donde coexisten ratios de ambos tipos de empresas, la clasificación queda dudosa y, por tanto, da lugar a la no clasificación de la empresa correspondiente.

4. El vector de clasificación

Una vez clasificado cada ratio de cada empresa se calcula, para éstas, su correspondiente vector de clasificación, compuesto por tres componentes que recogen, respectivamente, el número de ratios que clasifican a la empresa como sana n_0 , el número de ratios que clasifican a la empresa como quebrada n_1 , y finalmente, el número de ratios que no clasifican a la empresa en alguna de las dos categorías anteriores, n_D . Para cada empresa i se dispone, pues, del siguiente vector clasificatorio:

$$V_i = (n_{i0}, n_{i1}, n_{iD})$$

donde $n_{i0} + n_{i1} + n_{iD} = N$ con $N =$ número total de ratios.

Este vector clasificatorio es el finalmente utilizado en el proceso de clasificación final, en cuanto que la dominancia de una u otra de sus componentes es la que determinará la clasificación de la empresa en sana o quebrada.

5. Algoritmo de clasificación

El algoritmo de decisión para la clasificación final de una empresa como sana o quebrada es un proceso iterativo en el que, en cada etapa, se pretende minimizar el número de errores cometidos. Se parte siempre de la hipótesis de que de los dos posibles errores de clasificación –clasificar una empresa como quebrada, cuando es sana (error tipo I); o clasificar una empresa como sana, cuando es quebrada (error tipo II)- es este último el de mayor importancia y, por ello, el que hay que intentar minimizar con mayor intensidad y cuidado.

El algoritmo se realiza en los siguientes pasos:

PASO 1°. Se ordenan todas las empresas, de más a menos, según el valor $|n_0 - n_1|$. Siendo x un valor absoluto cualquiera para esa diferencia, se determina un punto de corte c que cumple la siguiente doble condición:

$$\forall x \geq c \wedge n_0 > n_1 \text{ las empresas son sanas}$$

$$\forall x \leq -c \wedge n_1 > n_0 \text{ las empresas son quebradas}$$

El criterio de clasificación para una empresa concreta es:

$\forall x \geq c$ la empresa se clasifica en sana o en quebrada atendiendo

al mayor número de ratios n_0 ó n_1 , respectivamente

Actuando de esta forma, en este primer paso no se cometen errores de clasificación para el conjunto de empresas de la muestra utilizada.

PASO 2°. Para el resto de empresas, se considera el primer valor de $k = |n_0 - n_1|$ no considerado en el paso anterior, esto es, $k=c-1$. Para ese valor se determinan el número de empresas incorrectamente clasificadas si se opta por considerarlas todas sanas o por considerarlas todas fracasadas. La decisión a tomar es aceptar el criterio que minimiza los errores cometidos. Por tanto, todas las empresas que cumplen la condición de tener el mismo valor de $|n_0 - n_1| = c-1$ se clasificarán como sanas si el número de empresas quebradas incorrectamente clasificadas como sanas es menor que el número de empresas sanas incorrectamente clasificadas como quebradas, y al revés en caso contrario. En caso de empate entre los dos posibles errores se opta por clasificarlas como quebradas, atendiendo a la gradación de importancia de los errores comentada con anterioridad. En el caso de un valor de $|n_0 - n_1|$ para el que sólo existan empresas de una naturaleza –sanas o fracasadas-, la decisión a adoptar es la de la clasificación en ese mismo tipo. Actuando de esta forma, se minimiza la probabilidad de una mala clasificación.

Formalizando este paso, si denotamos por n_{0k} el número de empresas mal clasificadas como sanas, y por n_{1k} el número de empresas mal clasificadas como quebradas, siendo $k = |n_0 - n_1|$ y $k < c$, el algoritmo propuesto es el siguiente:

- Si $n_{0k} > n_{1k} > 0$, todas las empresas para las que $k = |n_0 - n_1|$ se clasifican como quebradas. En términos de probabilidad, si n_k es el número total de empresas resultantes existentes cuando $k = |n_0 - n_1| < c$, $n_k = n_{0k} + n_{1k}$, resulta:

$$p_{0k} = P(\text{mala clasificación utilizando criterio de empresa sana} / |n_0 - n_1| = k) = \frac{n_{0k}}{n_k}$$

$$p_{1k} = P(\text{mala clasificación utilizando criterio de empresa quebrada} / |n_0 - n_1| = k) = \frac{n_{1k}}{n_k}$$

Y, consecuentemente, $p_{ok} > p_{1k}$.

- Si $0 < n_{0k} < n_{1k}$ todas las empresas para las que $k = |n_0 - n_1|$ se clasifican como sanas. Consecuentemente, en términos de probabilidad y con la misma notación anterior, resulta $p_{ok} < p_{1k}$.
- Si $n_{0k} = n_{1k} \neq 0$ todas las empresas para las que $k = |n_0 - n_1|$ se clasifican como quebradas, atendiendo a la mayor importancia dada al error de tipo II. En este caso, $p_{ok} = p_{1k}$.
- Finalmente, cuando $n_{0k} = 0$ y $n_{1k} \neq 0$, resulta ser n_{1k} igual al número total de empresas resultantes, que son sanas, por lo que se clasifican como tales. Consecuentemente, en términos de probabilidad y con la misma notación anterior, resulta:

$$p_{0k} = P(\text{mala clasificación utilizando criterio de empresa sana} / |n_0 - n_1| = k) = 0$$

$$p_{1k} = P(\text{mala clasificación utilizando criterio de empresa quebrada} / |n_0 - n_1| = k) = 1$$

Consecuentemente, $p_{ok} < p_{1k}$.

Y si $n_{0k} \neq 0$ y $n_{1k} = 0$, resulta ser n_{0k} igual al número total de empresas resultantes, que son quebradas, por lo que se clasifican como tales. Consecuentemente, en términos de probabilidad y con la misma notación anterior, resulta:

$$p_{0k} = P(\text{mala clasificación utilizando criterio de empresa sana} / |n_0 - n_1| = k) = 1$$

$$p_{1k} = P(\text{mala clasificación utilizando criterio de empresa quebrada} / |n_0 - n_1| = k) = 0$$

Y, consecuentemente, $p_{ok} > p_{1k}$.

En ambos casos finaliza el proceso de clasificación.

PASOS SUCESIVOS. El proceso es repetitivo para las empresas resultantes, finalizando el criterio una vez clasificadas las empresas para las que la diferencia

$|n_0 - n_1|$ es nula. En caso de que en dos ó más de estos pasos los criterios de clasificación coincidan, el algoritmo resultante los subsume en una única decisión.

El proceso se fundamenta en el hecho –de difícil cuantificación concreta en términos probabilísticos, dado que la distribución de probabilidad asociada a cada una de las ratios contable R_i es Multinomial de tres dimensiones, pero con distinto vector de probabilidades, perdiéndose la aditividad- de que, supuestamente, un alto número de ratios sanos en una empresa da pie a pensar en su situación sana, mientras que un alto número de ratios malos da pie a pensar en la situación de quebrada para la empresa correspondiente. Al ser un proceso iterativo, la previa clasificación de empresas da lugar a que vayan primeramente clasificándose las empresas con situación más clara, dejando las más dudosas para los últimos pasos. La calidad de la clasificación es, pues, mejor en las primeras etapas que en las últimas.

Una modificación ligeramente más compleja del criterio de clasificación propuesto consiste en volver a calcular los intervalos de clasificación del apartado 3 y los vectores de clasificación del apartado 4, *para aquellas empresas que no quedan clasificadas en el Paso 1º*. Con ellas y los nuevos intervalos se aplica otra vez el Paso 1º del criterio de decisión, para volver a calcular los intervalos y vectores de clasificación con las empresas resultantes. Solamente cuando tras este proceso no hay posibilidad de clasificar empresas en el Paso 1º, se pasa a clasificarlas en los pasos 2º y sucesivos.

Dado que las condiciones económicas y fiscales fluctúan con el tiempo, el proceso descrito debe calcularse sobre la información contable existente en cada ejercicio, por lo que el algoritmo obtenido también fluctúa con el tiempo y tiene validez de aplicación en el año más reciente al que está referida la información contable utilizada. Conforme nos alejamos en el tiempo de ella la capacidad discriminatoria del proceso disminuye.

6. Aplicación

La aplicación de la metodología propuesta consta de dos fases. En primer lugar, definir el criterio de clasificación sobre una muestra de empresas cuya situación de sana o fracasada es conocida. En segundo lugar, y para constatar la calidad del algoritmo de decisión, se aplica éste a una muestra de validación, para determinar el nivel de coincidencias logrado.

a) Muestra para definir el criterio de decisión

Para la selección y obtención de la muestra de empresas se ha recurrido inicialmente a la base de datos financieros SABI (Sistema de Análisis de Balances Ibéricos), cuyo uso ha sido cedido a la Universidad de Valencia por la empresa INFORMA, S.A. Se ha utilizado concretamente el disco 72 correspondiente a la actualización de agosto de 2005, y que contiene datos cualitativos y cuantitativos de un total de 832083 empresas españolas desde 1995.

Para seleccionar la muestra de empresas de este trabajo se exigió primeramente dos requisitos: uno geográfico, el que las empresas desarrollaran su actividad en la Comunidad Valenciana; el segundo temporal, restringiendo las empresas, sanas o fracasadas, a aquellas que tuvieron actividad o que cesaron en la misma durante el año 2004, habiendo estado activas al menos en los últimos cinco ejercicios económicos. Esta última exigencia determina la disminución del número de posibles empresas, sobre todo fracasadas, susceptibles de aparecer en la muestra.

Respecto la definición de PYME utilizada se han utilizado los criterios básicos referidos a su plantilla, el volumen de facturación y el volumen de activo fijado por el balance general anual. A estos tres parámetros se añadió un cuarto, el de independencia, excluyendo del estudio a aquellas empresas a las que su pertenencia a un grupo mayor les permitía disponer de medios superiores a los de su propia dimensión.

Los límites cualitativos empleados en el trabajo para definir a una PYME –según la Recomendación de la Comisión Europea de 6 de mayo de 2003- han sido los siguientes:

CRITERIOS	EMPRESA MEDIANA	EMPRESA PEQUEÑA	MICROEMPRESA
PLANTILLA	< 250	< 50	< 10
VOLUMEN DE FACTURACION	< 50.000.000 €	< 10.000.000 €	< 2.000.000 €
BALANCE GENERAL ANUAL	< 43.000.000 €	< 10.000.000 €	< 2.000.000 €

Tabla 3. Criterios de definición para las PYMES

El resultado de la selección supuso la consideración de 158 empresas, de las que 81 eran sanas y 77 fracasadas. Sus principales descriptivos se recogen en el Anexo I.

b) Muestra para la validación del algoritmo

Las empresas pertenecen a la misma base de datos de la muestra anterior. La muestra de validación consta de un total de 203 empresas, 172 sanas y 31 fracasadas, obtenidas utilizando idénticos criterios restrictivos que los empleados en la selección de la muestra para la elaboración del algoritmo de clasificación. Sus principales descriptivos se recogen en el Anexo II.

c) *Límites de los intervalos de clasificación*

Teniendo en cuenta los estadísticos *max* y *min* de la tabla del Anexo I, se construyen los intervalos de clasificación definidos en la Tabla 2:

RATIO	Empresas sanas	Empresas quebradas
Rentabilidad		
REN1	$]0'24732, +\infty[$	$] -\infty, -0'02268[$
REN2		$] -\infty, -0'01366[\cup]0'19678, +\infty[$
REN3		$] -\infty, -0'09446[\cup]1'22592, +\infty[$
REN4	$]0'52330, +\infty[$	$] -\infty, -0'01597[$
REN5	$]0'07004, +\infty[$	$] -\infty, -0'01951[$
REN6	$]0'13107, +\infty[$	$] -\infty, -0'018245[$
REN7	$]0'07120, +\infty[$	$] -\infty, -0'02844[$
REN8		$] -\infty, -0'19333[\cup]1'84057, +\infty[$
REN9		$] -\infty, -0'01033[\cup]0'13369, +\infty[$
REN10		$] -\infty, 0'04058[\cup]3'89845, +\infty[$
REN11		$] -\infty, 0'02695[\cup]0'24576, +\infty[$
REN12	$]0'52330, +\infty[$	$] -\infty, 0'02892[$
REN13	$]0'71927, +\infty[$	$] -\infty, 0'04144[$
REN14		$] -\infty, -0'01951[\cup]0'30217, +\infty[$
REN15	$] -0'19161, +\infty[$	$] -\infty, -0'83269[$
REN16		$] -\infty, 0'01064[\cup]0'31835, +\infty[$
REN17	$]22'78351, +\infty[$	$] -\infty, -3'27768[$
REN18		$] -\infty, -0'13495[\cup]1'88603, +\infty[$
REN19		$] -\infty, -0'01476[\cup]0'20568, +\infty[$
REN20	$]0'52330, +\infty[$	$] -\infty, -0'02281[$
REN21	$]0'20206, +\infty[$	$] -\infty, -0'04467[$
REN22		$] -\infty, 0[\cup]1'55361, +\infty[$
REN23	$]0'04692, +\infty[$	$] -\infty, -0'01785[$
Estructura financiera		
EF1		$] -\infty, 0'19698[\cup]22'52353, +\infty[$
EF2		$] -\infty, 0[\cup]7'05382, +\infty[$

EF3		$]-\infty, 0'19180[\cup]20'04448, +\infty[$
EF4	$]0'21753, +\infty[$	$]-\infty, 0[$
EF5	$]-\infty, 0'27650[$	$]1, +\infty[$
EF6	$]-\infty, 0[$	$]0'74435, +\infty[$
EF7	$]-\infty, 0'24033[$	$]0'92403, +\infty[$
EF8	$]0'94436, +\infty[$	$]-\infty, 0'044398[$
EF9	$]-\infty, 0[$	$]0'07959, +\infty[$
EF10	$]-\infty, 0'51431[$	$]0'95749, +\infty[$
EF11	$]108'07266, +\infty[$	$]-\infty, 0[$
Actividad		
ACT1		$]-\infty, 0'05074[\cup]0'84268, +\infty[$
ACT2		$]-\infty, 0'26958[\cup]0'96216, +\infty[$
ACT3		$]-\infty, 0[\cup]0'24521, +\infty[$
ACT4	$]1'15084, +\infty[$	$]-\infty, 0'98208[$
ACT5		$]-\infty, -0'0564331[\cup]0'38431, +\infty[$
ACT6	$]-\infty, 0'13451[$	$]1'63349, +\infty[$
ACT7	$]-\infty, 0'02038[$	$]0'81080, +\infty[$
ACT8	$]20'34262, +\infty[$	$]-\infty, 0'098789[$
Rotación		
ROT1	$]7'19032, +\infty[$	$]-\infty, 0'50330[$
ROT2	$]1'102319, +\infty[$	$]-\infty, 0'73937[$
ROT3	$]112'61601, +\infty[$	$]-\infty, 0'72105[$
ROT4		$]-\infty, 0'77341[\cup]50'75000, +\infty[$
ROT5	$]7'43444, +\infty[$	$]-\infty, 0'90286[$
ROT6	$]-\infty, 0'13908[$	$]1'76076, +\infty[$
ROT7	$]219'81479, +\infty[$	$]-\infty, -226'48413[$
ROT8	$]312'42495, +\infty[$	$]-\infty, 1'16401[$
ROT9	$]218'82739, +\infty[$	$]-\infty, 1'18123[$
Solvencia (liquidez)		
SOLV1	$]1'93908, +\infty[$	$]-\infty, 0'47624[$
SOLV2	$]-\infty, 0'05867[\cup]1'63422, +\infty[$	
SOLV3	$]1'03869, +\infty[$	$]-\infty, 0'00193[$
SOLV4	$]0'93908, +\infty[$	$]-\infty, -0'52376[$
SOLV5	$]0'40037, +\infty[$	$]-\infty, -0'43209[$
SOLV6	$]0'56032, +\infty[$	$]-\infty, -0'33209[$
SOLV7	$]-\infty, 0'00409[\cup]1'62409, +\infty[$	
SOLV8	$]-\infty, 0'01795[$	$]1'24096, +\infty[$
SOLV9	$]0'71927, +\infty[$	$]-\infty, 0'05180[$

SOLV10	$]0'35202, +\infty[$	$] -\infty, -0'94956[$
SOLV11	$]0'54314, +\infty[$	$] -\infty, -0'32357[$
SOLV12	$] -\infty, 0'00456[$	$]0'84658, +\infty[$
Solvencia a largo plazo		
SOLVLP1		$] -\infty, 0'05213[\cup]13'00971, +\infty[$
SOLVLP2	$]0'52330, +\infty[$	$] -\infty, 0'03103[$
SOLVLP3	$]1'94436, +\infty[$	$] -\infty, 1'04440[$
Estructura económica		
CF1	$]0'93762, +\infty[$	$] -\infty, 0[$
CF2	$] -\infty, 0'14151[\cup]0'97974, +\infty[$	
CF3	$] -\infty, 0'16483[\cup]48'36209, +\infty[$	
CF4	$] -\infty, -78'87214[$	$]37'33042, +\infty[$
CF5	$] -\infty, 0[$	$]0'76405, +\infty[$
CF6	$] -\infty, 0'00335[$	$]0'88007, +\infty[$
CF7	$]0'51331, +\infty[$	$] -\infty, 0'00061[$

Tabla 4. Intervalos de clasificación

d) Algoritmo o criterio de clasificación

Considerando los ratios contables estudiados, la aplicación de la metodología da lugar al siguiente criterio para clasificar una empresa en sana o fracasada, basado en la diferencia entre las componentes n_0 y n_1 del vector de clasificación:

Paso	Criterio
1º	Para diferencias iguales o mayores que 5 la empresa se clasifica según que el mayor número de ratios lo sea en 0 ó 1, respectivamente
2º	Si la diferencia es igual a 4 la empresa se clasifica como 0
3º	El resto de empresas, con valores de la diferencia iguales o menores que 3, también se clasifican como 0
RESULTADO	2 errores, ambas de empresas quebradas clasificadas como sanas

Tabla 5. Algoritmo de decisión

Para adoptar la decisión del segundo paso se ha tenido en cuenta que la clasificación como sanas de las seis empresas con diferencia $|n_0 - n_1|$ igual a cuatro da lugar a dos errores, frente a los cuatro errores que supondría clasificarlas como quebradas:

$$n_{0k=4} = 2 < n_{1k=4} = 4 \rightarrow \text{las empresas con } |n_0 - n_1| = 4 \text{ se clasifican como sanas}$$

En términos de probabilidad

$$P(\text{mala clasificación utilizando criterio de empresa sana} / |n_0 - n_1| = 4) = 0'3$$

$$P(\text{mala clasificación utilizando criterio de empresa quebrada} / |n_0 - n_1| = 4) = 0\hat{6}$$

En cuanto al paso tercero, la decisión es inmediata al ocurrir que las cuarenta y cinco empresas resultantes de los dos pasos anteriores son todas sanas:

$$n_{01k} = 0 \text{ y } n_{1k} = 45 \rightarrow \text{las empresas con } |n_0 - n_1| < 4 \text{ se clasifican como sanas}$$

En términos de probabilidad

$$P(\text{mala clasificación utilizando criterio de empresa sana}) = 0$$

$$P(\text{mala clasificación utilizando criterio de empresa quebrada}) = 1$$

e) *Resultado*

El algoritmo propuesto, aplicado sobre la muestra de empresas, arroja 2 errores, que corresponden a dos empresas quebradas que se clasifican como sanas. Aunque este es el error de mayor incidencia en la decisión de una entidad de crédito, su mínima entidad avala la bondad de una posible clasificación utilizando el algoritmo. En resumen:

		Aplicación del algoritmo		
		Sana	Fracasada	Total
Muestra	Sana	81 (100 %)	0 (0 %)	81 (100 %)
	Fracasada	2 (2'6 %)	75 (97'4 %)	77 (100 %)
	Total	83	75	158

Tabla 6. Resultado de la aplicación del algoritmo sobre la muestra de empresas

La aplicación sobre la muestra de validación arroja 4 errores, pero en este caso corresponden a empresas sanas que se clasifican como quebradas (error de tipo I). Todas las empresas quebradas se clasifican como tales. Este resultado confirma la bondad del algoritmo en esta muestra de validación, al clasificar correctamente como fracasadas todas las empresas que efectivamente lo son. El resultado se resume en la tabla siguiente:

		Aplicación del algoritmo		
		Sana	Fracasada	Total
Muestra de validación	Sana	168 (97'7 %)	4 (2'3 %)	172 (100 %)
	Fracasada	0 (0 %)	31 (100 %)	31 (100 %)
	Total	168	35	203

Tabla 7. Resultado de la aplicación del algoritmo sobre la muestra de validación

Dada la bondad de los resultados obtenidos no hemos procedido a aplicar la modificación del criterio de clasificación propuesto al final del apartado 5.

7. Conclusiones

El algoritmo ha proporcionado buenos resultados de clasificación en las dos muestras de empresas utilizadas. Es de fácil aplicación y no utiliza un lenguaje estadístico/matemático complicado, de ahí la posible eficacia de su uso. Si bien los errores de clasificación que ha habido que aceptar en su desarrollo afectaban a las empresas quebradas que se han clasificado como sanas (error de tipo II), que son los de mayor gravedad, su escasa incidencia en el conjunto de toda la muestra hace que no pierda interés. Esto se aprecia claramente en la muestra de validación, en la que los errores cometidos, también escasos, son todos del tipo I - error menos relevante para una entidad de crédito que el cometido al aceptar como sana una empresa que realmente es fracasada-, por lo que con las muestras de empresas utilizadas en este trabajo la capacidad clasificatoria del procedimiento resulta claramente manifiesta.

Sin embargo dejamos constancia de las limitaciones del procedimiento. En primer lugar, el algoritmo de decisión es único para cada situación planteada, esto es, su validez se refiere exclusivamente a la población de la que se ha extraído la correspondiente muestra. Eso supone que un criterio de clasificación válido para un conjunto de datos no tiene por qué serlo para otro. Lo que resulta de aplicabilidad general es la filosofía del procedimiento, no los resultados concretos de su aplicación a cada caso concreto de datos, diferentes espacial y/o temporalmente. En segundo lugar, no puede cuantificarse el efecto que sobre la aplicación del algoritmo, sobre todo en sus últimas etapas, tiene la presencia de posibles ratios redundantes, por estar entre sí muy correlados o por responder en su definición a conceptos muy parecidos. En cualquier caso el procedimiento se basa en la cantidad de ratios utilizados, por lo que a mayor número de ellos mayor es la capacidad discriminante del mismo. Finalmente, hasta qué punto la consideración de más ratios de una u otra categoría pueda afectar a las decisiones de cada etapa del algoritmo queda también como problema a estudiar en el futuro.

8. Bibliografía

Altman, E.I. (1968): “Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy”, *The Journal of Finance*, Vol. XIII, 4, September, pp. 589-609.

Dieguez, J. *et al* (2006): “Modelos de predicción de la insolvencia empresarial: La incorporación de ratios a partir de un marco teórico”. Comunicación presentada en la VI

Jornadas sobre Predicción de Insolvencia Empresarial. Carmona (Sevilla), 9 y 10 de noviembre.

Karels, G.V. y Prakash, A.J. (1987): “Multivariate Normality and Forecasting of Business Bankruptcy”, *Journal of Business Finance and Accounting*, Vol. 14, pp. 573-592.

Labatut, G. *et al* (2006). “La Ley Concursal en los procesos de insolvencia empresarial de las Pymes: Evolución histórica y valoración”, *Técnica Contable*, 686, pp. 9-18.

López, J. *et al* (1998). “La suspensión de pagos en las Pymes: Una aproximación empírica”. *Revista Española de Financiación y Contabilidad*, Vol. XXVII, 94, pp. 71-97.

Molina, R. (194). *La selección de ratios contables en el análisis de solvencia*. Trabajo de investigación. Departamento de Contabilidad, Universidad de Valencia.

Molina, R. (2003). *La evolución de la presión fiscal en las PYME's: una aproximación a su incidencia sobre la situación financiera de las empresas de la Comunidad Valenciana*. Tesis Doctoral, Universidad de Valencia.

Pozuelo, J. (2007). *La predicción del fracaso empresarial: un estudio empírico de la pequeña empresa de la Comunidad Valenciana*. Tesis Doctoral, Universidad de Valencia.

ANEXO I. Muestra de aplicación

Estadísticos descriptivos empresas sanas (81 empresas)

RATIO	Mínimo	Máximo	Media	Desv. típ.
Rentabilidad				
REN1	-,022682	,352048	,09556658	,060928344
REN2	-,013658	,196784	,05178960	,038370767
REN3	-,094463	1,225925	,18257539	,186825684
REN4	-,015967	,620566	,11612748	,129348970
REN5	-,019513	,334071	,06888092	,059153742
REN6	-,018245	,205268	,05214865	,039817490
REN7	-,028435	1,053507	,15380602	,191106414
REN8	-,193335	1,840573	,24048956	,269658132
REN9	-,010330	,133691	,02968872	,027248519
REN10	,040577	3,898446	,44248959	,578385255
REN11	,026952	,245759	,09810759	,053535283
REN12	,028924	,831326	,20288433	,173371250
REN13	,041436	2,129062	,28839837	,365900803
REN14	-,019513	,302171	,07172456	,055140948
REN15	-,832693	,112327	-,49470147	,229194917
REN16	,010644	,318349	,06129266	,053725376
REN17	-3,277684	684,700000	16,23414202	76,813785069
REN18	-,134952	1,886038	,25500553	,271211748
REN19	-,014758	,205679	,04042291	,038118026
REN20	-,022811	,952907	,15963469	,183779782
REN21	-,044670	1,057883	,23146322	,241509727
REN22	,000000	11,553612	1,20080030	2,205184381

REN23	-,017854	,205640	,03781803	,039637728
Estructura financiera				
EF1	,196977	22,523527	3,66603512	4,590207409
EF2	,000000	7,053825	,65148578	1,374968222
EF3	,191796	20,044484	3,01454934	3,786590255
EF4	,000000	,379318	,05032888	,052144605
EF5	,113323	1,000000	,85882356	,187147695
EF6	,000000	,744352	,09425448	,140899658
EF7	,095133	,924034	,52641271	,224492519
EF8	,044398	5,076742	1,01359413	1,207253036
EF9	,000000	,079594	,01687560	,015368141
EF10	,164562	,957489	,62066719	,232711602
EF11	,000000	208,203946	9,95884997	26,758174657
Actividad				
ACT1	,050737	,842684	,30804153	,173024798
ACT2	,269585	,962163	,68748712	,156022384
ACT3	,000000	,245206	,06557550	,061761865
ACT4	,982082	1,258285	1,05643605	,047848877
ACT5	-,056433	,384311	,10845534	,085128091
ACT6	,025254	1,633493	,38713257	,288309896
ACT7	,019220	,810799	,21923220	,144027712
ACT8	,097889	22,058434	2,75084499	3,946997166
Rotación				
ROT1	,503304	12,413087	2,24218593	1,787218643
ROT2	,739375	28,240608	3,39982095	3,329788868
ROT3	,721046	179,307080	16,79959843	26,798525135
ROT4	,773414	50,750000	6,91543769	7,976272791
ROT5	,902861	76,484497	5,45765427	8,467243076
ROT6	,080560	1,760760	,62316964	,345335133
ROT7	-226,484127	268,720009	11,78568067	62,937208684
ROT8	1,164011	2120,445545	51,42650031	238,403263580
ROT9	1,181229	332,570982	17,51406308	43,388419729
Solvencia (liquidez)				
SOLV1	,476243	5,334675	1,68404920	1,108602546
SOLV2	,021571	4,993980	1,25280523	1,035455769
SOLV3	,001927	3,818348	,48821356	,701019544
SOLV4	-,523759	4,334675	,68404870	1,108603325
SOLV5	-,432092	,755318	,18045189	,261298637
SOLV6	-,332091	1,081265	,10611457	,201399550
SOLV7	,000000	3,350216	,75054968	,653026748
SOLV8	,007295	1,240946	,29047104	,204296371
SOLV9	,051796	2,318436	,34363347	,422420374
SOLV10	-,949565	1,075428	-,01777833	,267158816
SOLV11	-,323570	1,080308	,10463009	,198493282
SOLV12	,000000	,846576	,18876859	,147692982
Solvencia a largo plazo				
SOLVLP1	,052129	13,009710	1,52141963	2,272698982
SOLVLP2	,031032	1,086594	,24639154	,222561166
SOLVLP3	1,044398	6,076742	2,01359466	1,207252551
Estructura económica				
CF1	,000000	,977141	,29538801	,262355143

CF2	,135578	,980850	,70686499	,206972693
CF3	,156843	51,221801	5,71455733	7,610278604
CF4	-182,113250	37,330417	-1,15295950	21,989621238
CF5	,000000	,764051	,22136124	,209827214
CF6	,000000	,880067	,32512983	,214101535
CF7	,000611	,723450	,15577642	,148217117

Estadísticos descriptivos empresas fracasadas (77 empresas)

RATIO	Mínimo	Máximo	Media	Desv. típ.
Rentabilidad				
REN1	-1,341555	,247321	-,11769697	,251180482
REN2	-1,427995	,495722	-,13929607	,266049718
REN3	-8,913802	95,366617	1,46295668	11,198871003
REN4	-,594367	,523298	-,10894216	,174041387
REN5	-1,425427	,070035	-,16347842	,257215247
REN6	-3,115953	,131070	-,23311711	,571531420
REN7	-,775183	,071200	-,13733997	,171246429
REN8	-10,699398	94,407519	1,33017952	11,380246857
REN9	-5,991550	,912479	-,28469779	,934304272
REN10	-8,913802	85,912180	1,25289242	10,110297156
REN11	-1,421840	,495722	-,10431002	,262764790
REN12	-,591805	,523298	-,07339363	,166799515
REN13	-,753525	,719270	-,09847546	,242486779
REN14	-1,427995	,495722	-,14894140	,273694007
REN15	-2,136241	-,191616	-,79417414	,309473837
REN16	-5,327987	,912479	-,21545074	,843381106
REN17	-1373,275862	22,783515	-29,87779823	161,402457586
REN18	-8,913802	95,366617	1,40688259	11,231391871
REN19	-8,559346	1,303544	-,32589597	1,175218297
REN20	-,775183	,523298	-,12027384	,195955284
REN21	-1,071435	,202060	-,12585283	,229411786
REN22	-475,010000	1313,370504	12,36075112	160,734350662
REN23	-3,118222	,046921	-,28118402	,609281241
Estructura financiera				
EF1	-648,346165	121,426044	-6,71773972	96,858339846
EF2	-19,914925	18,155810	,36887396	4,462080311
EF3	-641,747368	107,752893	-7,08661368	94,375646424
EF4	,000000	,217533	,04598600	,043553879
EF5	,276500	1,000000	,86828915	,169184826
EF6	,000000	,990890	,16021099	,231147527
EF7	,240328	2,209040	,90802163	,313287441
EF8	-,615445	,944356	,02996163	,294696053
EF9	,000000	,374918	,04146628	,048444754
EF10	,514309	2,600410	1,06823262	,390308787
EF11	,000000	108,072657	5,33856297	14,501890301
Actividad				
ACT1	-6,691260	2,168371	,12386478	1,002355637
ACT2	-8,512352	92,501559	1,77136410	10,653810670
ACT3	-,950026	11,762971	,26608073	1,355085274
ACT4	,000000	1,150841	,88333982	,212256342

ACT5	-106,886306	10,190914	-1,12348710	12,445008949
ACT6	,134509	32,773453	2,23578028	4,773109214
ACT7	,020379	1,202253	,34057731	,232456688
ACT8	,042587	20,342615	2,55995442	3,492824249
Rotación				
ROT1	,000000	7,190321	1,45444978	1,235802453
ROT2	,000000	11,023190	2,24403128	2,110996930
ROT3	,000000	112,616012	11,84206172	18,143105515
ROT4	-355,230648	256,260174	17,43862598	68,045052573
ROT5	,000000	7,434443	1,72154217	1,453445735
ROT6	,139076	18,968821	1,86124226	3,314546550
ROT7	-374,152089	219,814788	-1,94389514	56,152363484
ROT8	,000000	312,424953	19,21015297	52,282518016
ROT9	,000000	218,827394	10,90357164	28,382078760
Solvencia (liquidez)				
SOLV1	,230268	1,939079	,85523201	,345977292
SOLV2	,058674	1,634219	,48440594	,330467292
SOLV3	,000000	1,038689	,06837846	,144953696
SOLV4	-,769732	,939079	-,14476805	,345977429
SOLV5	-1,417139	,400368	-,19013078	,348674049
SOLV6	-10,953290	,560315	-,53154041	1,698218777
SOLV7	,004095	1,624095	,40894774	,317434632
SOLV8	,017952	18,499218	,83434660	2,238867211
SOLV9	-,753525	,719270	-,11046803	,255408513
SOLV10	-16,254967	,352023	-,93433102	2,245214766
SOLV11	-10,953290	,543140	-,50477176	1,645428453
SOLV12	,004564	18,041820	,65805126	2,112970200
Solvencia a largo plazo				
SOLVLP1	-91,727519	34,941737	,42261811	13,397384614
SOLVLP2	-,591805	,523298	-,08472531	,183618781
SOLVLP3	,384555	1,944356	1,02996163	,294695886
Estructura económica				
CF1	,000000	,937623	,41737595	,286109623
CF2	,141507	,979742	,71789097	,225664024
CF3	,164832	48,362091	7,46139449	9,602245857
CF4	-78,872145	97,651126	,03734262	16,007095125
CF5	,000000	,900331	,32314705	,254263149
CF6	,003352	,951130	,33173102	,219954797
CF7	,000000	,513309	,05750091	,102025378

ANEXO II. Muestra de validación

Estadísticos descriptivos empresas sanas (172 empresas)

RATIO	Mínimo	Máximo	Media	Desv. típ.
Rentabilidad				
REN1	-,008647	,351615	,08904217	,064936486
REN2	,001519	,234011	,05233264	,045124228
REN3	,003342	,784532	,15430848	,109550822
REN4	,002264	1,479369	,14111302	,227325279
REN5	-,008415	,351802	,07485089	,067061977

REN6	-,004592	,357186	,05990730	,057207730
REN7	-,023815	2,209542	,20254888	,333488447
REN8	-,013014	1,120711	,22015980	,153131582
REN9	,000197	,368631	,03688672	,047417755
REN10	,034453	1,509130	,32253445	,252919334
REN11	,012776	,331771	,09019642	,054475968
REN12	,014530	1,856180	,21854075	,288879072
REN13	,019224	1,856180	,27880983	,328284685
REN14	,002498	,355248	,07538874	,066890961
REN15	-,946866	,052329	-,47641073	,244877110
REN16	,001809	,479234	,06293142	,061876439
REN17	-1,686310	58642,00000	370,65096696	4471,75516528
REN18	,005497	1,120757	,22050803	,154890578
REN19	,000701	,528266	,05332419	,070118658
REN20	,003532	2,248306	,20448555	,336027510
REN21	-,034321	2,163507	,25995898	,340850213
REN22	,000000	50,822006	2,45362621	5,475039092
REN23	-,004469	,532095	,05298269	,069640118
Estructura financiera				
EF1	,061208	41,862927	3,63639860	5,057971327
EF2	,000000	14,759585	,89931923	2,235154528
EF3	,061208	31,370209	2,73025185	3,542919499
EF4	,000000	,158876	,02970776	,027664205
EF5	,013165	1,000000	,82100768	,218043474
EF6	,000000	,890757	,12513478	,166979813
EF7	,011883	,916452	,50582473	,221151053
EF8	,023887	16,337657	1,08838780	1,939631843
EF9	,000000	,062418	,01160206	,010999885
EF10	,057678	,976670	,63206466	,222555210
EF11	,000000	153,475622	11,86934885	27,265595499
Actividad				
ACT1	,010507	1,049829	,27321112	,146505882
ACT2	,151668	,935978	,65673424	,178854399
ACT3	,000000	,236961	,05088244	,048501067
ACT4	,995503	1,532169	1,06698619	,073970231
ACT5	,002365	,508484	,13752516	,115348917
ACT6	,024046	6,437589	,45340075	,541501662
ACT7	,007004	,449407	,18201460	,109172249
ACT8	,087558	37,680619	2,61845319	5,034254719
Rotación				
ROT1	,072034	11,555435	1,95422380	1,343865218
ROT2	,107686	16,694712	3,13462193	2,342613576
ROT3	,217578	340,412712	18,78925293	37,739172775
ROT4	,134324	126,067913	6,21268958	10,858449554
ROT5	,155338	54,491116	5,13794755	6,991722678
ROT6	,085710	13,882252	,76654540	1,102406137
ROT7	-345,391260	1899,557647	12,00973366	159,458880807
ROT8	,545055	2572,555556	44,90309792	201,328944600
ROT9	1,176003	1428,708688	32,31970310	151,333743484
Solvencia (liquidez)				
SOLV1	,206228	9,842460	1,71334197	1,390466034

SOLV2	,000000	9,722488	1,28809675	1,377492050
SOLV3	,000000	9,568019	,50467790	,970927883
SOLV4	-,793772	8,842460	,70895725	1,390626236
SOLV5	-,618011	,802145	,18038830	,226986510
SOLV6	-,896321	2,848619	,13795214	,304451677
SOLV7	,000000	9,540335	,78089903	,874799425
SOLV8	,000000	9,286223	,36098207	,716035661
SOLV9	,024901	2,625118	,35644270	,444605733
SOLV10	-1,205407	3,801208	,01886110	,396411174
SOLV11	-,864671	2,848619	,13560033	,297411198
SOLV12	,000000	,655960	,19538027	,138544549
Solvencia a largo plazo				
SOLVLP1	,008803	15,309751	1,50976860	2,266510889
SOLVLP2	,018660	2,625118	,28191405	,393751385
SOLVLP3	1,023887	17,337657	2,08838779	1,939631772
Estructura económica				
CF1	,000000	1,000000	,29231412	,248496361
CF2	,098259	,995444	,68804154	,226555876
CF3	,108966	218,492147	8,90769599	22,204433410
CF4	-84,980319	232,622941	1,16034855	21,257201890
CF5	,000000	,735965	,19945960	,181906616
CF6	,000000	,867911	,32682958	,212339254
CF7	,000000	,693390	,16272219	,169487266

Estadísticos descriptivos empresas fracasadas (31 empresas)

RATIO	Mínimo	Máximo	Media	Desv. típ.
Rentabilidad				
REN1	-1,745654	,136786	-,21180075	,413233130
REN2	-1,817110	,020549	-,22599541	,381058082
REN3	-18,541551	32,548976	,45585198	7,141097055
REN4	-1,066479	,021396	-,17628599	,244875755
REN5	-1,745654	,021958	-,25565232	,408302075
REN6	-,923823	,065697	-,16478137	,285326377
REN7	-1,332345	,022863	-,20238754	,285430114
REN8	-87,425669	44,398515	-1,50983559	17,961175149
REN9	-,800306	,016561	-,17459255	,246181516
REN10	-11,102493	25,725467	,81934063	5,214655768
REN11	-1,800879	,061675	-,19431849	,378113511
REN12	-1,031405	,064219	-,14382859	,240942460
REN13	-1,070051	,077069	-,17801785	,264651652
REN14	-1,817110	,026340	-,24358268	,405668181
REN15	-1,950600	-,386520	-,89298257	,333732523
REN16	-,706494	,096684	-,14519645	,233019346
REN17	-27,607887	1,676497	-3,32994111	6,413919659
REN18	-18,930748	48,809375	,94710489	9,746717352
REN19	-1,009276	,023659	-,18938181	,270891556
REN20	-1,066479	,027427	-,18748517	,249084620
REN21	-1,102308	,147435	-,15599079	,272465488
REN22	-462,761905	5,143113	-17,94578141	82,860549047
REN23	-,973560	,023659	-,20080152	,285064158

Estructura financiera				
EF1	-253,153055	382,098338	18,77080565	95,677161563
EF2	-87,335003	98,121422	3,20730405	25,647969036
EF3	-165,818052	283,976916	15,56350160	72,086102564
EF4	,000000	,139460	,05306873	,034996544
EF5	,375589	1,000000	,80336480	,169841576
EF6	,000000	,743067	,21383777	,207642230
EF7	,369779	2,245010	,85078174	,376813142
EF8	-,599689	,665703	,03090016	,273327165
EF9	,000000	,160152	,04385880	,032877953
EF10	,600347	2,498056	1,06461952	,411794891
EF11	,000000	170,663561	13,24539360	31,899834711
Actividad				
ACT1	-,660423	,792475	,23002105	,296196369
ACT2	-4,548441	6,565917	,63666803	1,945344057
ACT3	-,967607	2,432045	,22353525	,551692678
ACT4	,000000	1,069997	,86534669	,228802290
ACT5	-8,441011	7,510169	,08972556	2,660412399
ACT6	,308546	2,401011	,92889984	,577285023
ACT7	,008552	1,263533	,34836417	,242886578
ACT8	,000000	11,184894	2,16884749	2,149013677
Rotación				
ROT1	,000000	3,091965	1,42366042	,768752489
ROT2	,000000	19,691765	3,18719986	4,433752971
ROT3	,000000	258,765943	14,50095614	45,527346796
ROT4	-19,028384	71,431472	7,13079331	14,686654694
ROT5	,000000	4,445232	1,81088822	1,078268504
ROT6	,323419	2,858258	,88089606	,533850085
ROT7	-983,706731	324,918869	-3,88543892	200,833559338
ROT8	,000000	46,714002	8,29121408	11,282368427
ROT9	,000000	131,473392	17,50824488	33,116564437
Solvencia (liquidez)				
SOLV1	,032956	2,187191	,99421020	,495193823
SOLV2	,017467	1,425827	,44601788	,287107233
SOLV3	,000000	,202274	,03911531	,045701046
SOLV4	-,967048	1,187190	-,00579058	,495194271
SOLV5	-2,091720	,492607	-,14031610	,531851995
SOLV6	-1,490134	,669641	-,07309413	,425758283
SOLV7	,004095	1,352701	,38880705	,289478894
SOLV8	,012125	,915608	,30632300	,228884560
SOLV9	-1,070051	,084305	-,19385289	,276913210
SOLV10	-1,914664	15,793388	,16315671	2,921125293
SOLV11	-1,490134	,654631	-,07340807	,422700002
SOLV12	,007606	,781960	,25246644	,210313029
Solvencia a largo plazo				
SOLVLP1	-116,479113	93,580794	2,02484409	28,667735268
SOLVLP2	-1,031405	,070249	-,15502777	,246033405
SOLVLP3	,400311	1,665703	1,03090032	,273327006
Estructura económica				
CF1	,000000	,876334	,50956094	,263875236
CF2	,049221	,991860	,71046643	,233489109

CF3	,051769	121,849107	10,86923918	25,225257538
CF4	-234,586538	96,429502	,59928623	48,998430381
CF5	,000000	,851626	,37908143	,236821631
CF6	,003352	,782761	,28790573	,194953285
CF7	,000000	,165571	,03054443	,037898663

Análisis estadístico Bayesiano de las rentabilidades empresariales en países de la Unión Europea

José Luis GALLIZO

e-mail: gallizo@aegern.udl.es

Departamento de Administración de Empresas y Gestión Económica de los Recursos Naturales
UNIVERSIDAD LLEIDA

Pilar GARGALLO

e-mail: pigarga@unizar.es

Departamento de Estructura e Historia Económica y Economía Pública
UNIVERSIDAD ZARAGOZA

Ramón SALADRIGUES

e-mail: ramon.saladrigues@aegern.udl.cat

Departamento de Administración de Empresas y Gestión Económica de los Recursos Naturales
UNIVERSIDAD LLEIDA

Manuel SALVADOR

e-mail: salvador@unizar.es

Departamento de Estructura e Historia Económica y Economía Pública
UNIVERSIDAD ZARAGOZA

Área temática: Métodos Cuantitativos.

Resumen

En este trabajo se realiza un análisis estadístico de la persistencia de las rentabilidades de una muestra de empresas de diversos países de la Unión Europea. Para ello se utilizan modelos dinámicos de ecuaciones estructurales que permiten descomponer la evolución anual de dichas rentabilidades en una componente permanente de carácter estructural y otra componente transitoria de carácter coyuntural. Esta descomposición posibilita la evaluación de la importancia relativa de dichas componentes fundamentales. El análisis se realiza desde una óptica bayesiana lo cual dota al estudio realizado de mayor flexibilidad y realismo al no depender las estimaciones obtenidas de resultados asintóticos y permitir un tratamiento natural de datos ausentes muy habituales en este tipo de estudios

Palabras clave: Ecuaciones Estructurales, Inferencia Bayesiana, MCMC, Beneficios Anormales, Persistencia de beneficios, Rentabilidad de Activos.

Abstract

In this paper a statistical analysis of the persistence of profits in a sample of firms of the UE is made. To that end we use a dynamic structural equation model that let us to decompose the annual evolution of the firm's returns in a permanent, a transitory and a specific component. This decomposition makes possible, in particular, to evaluate their relative importance. We adopt a Bayesian approach which gives more flexibility and realism to the analysis, given that the estimations do not depend on asymptotic results and it makes possible a natural treatment of missing data.

Keywords: Structural Equations Models, Bayesian Inference, MCMC, Abnormal Returns, Persistence of Profits, Return of Assets

1. Introducción

Es un hecho empíricamente comprobado que los niveles de rentabilidad de las empresas difieren significativamente de los valores medios de equilibrio (Jacobson, 1988). Este hecho ha motivado la aparición en la literatura de diversos trabajos que buscan describir la evolución temporal de dichas rentabilidades con el fin de analizar si dichas diferencias tienen un carácter coyuntural o transitorio, tendiendo a desaparecer a lo largo del tiempo, o bien tienen un carácter permanente (ver, por ejemplo, Waring, 1996; McGahan y Porter, 2003; Bou y Satorra, 2006).

Recientemente Bou y Satorra (2006) han propuesto un modelo de ecuaciones estructurales dinámico que permite una descomposición de la evolución de las rentabilidades anormales de las empresas en una componente permanente, otra transitoria y otra específica. Su metodología permite, en particular, evaluar la importancia de cada una de ellas así como distinguir entre efectos que afectan al sector en el que se enmarca la actividad de la empresa y efectos que son específicos de la misma. Dichos autores analizan la evolución de 5000 empresas españolas en el periodo 1995-2000 y concluyen que existen diferencias permanentes significativas, tanto a nivel sectorial como particular de cada empresa, siendo éstas últimas de mayor magnitud; dichos autores no aprecian, sin embargo, diferencias significativas entre los procesos de ajuste de las componentes transitorias que operan tanto a nivel de sector como a nivel de empresa.

En esta comunicación extendemos el estudio realizado por dichos autores a diversos países europeos utilizando una muestra más grande de empresas. A diferencia ellos, sin embargo, permitimos la existencia de datos ausentes en las series analizadas. Para tratar este tipo de información adoptamos un planteamiento bayesiano del proceso de estimación. Este planteamiento permite un tratamiento natural de los datos ausentes mediante la realización de imputaciones aleatorias, dotando al estudio realizado de mayor flexibilidad y realismo al no depender las estimaciones obtenidas de aproximaciones asintóticas.

El plan del trabajo es como sigue: en la sección 2 se describen los datos utilizados en el mismo; en la sección 3 se describe el modelo así como el proceso de

estimación de sus parámetros; en la sección 4 se exponen los resultados obtenidos y, finalmente, la sección 5 concluye.

1.1 Notación

En lo que sigue $[X]$ denota la densidad de la variable aleatoria X y $[X|Y]$ la densidad de la distribución condicional X dado Y . $\text{Diag}(\mathbf{v})$ denota la matriz cuadrada que contiene las componentes del vector \mathbf{v} en la diagonal y $\text{blockdiag}(\mathbf{A}, \mathbf{B})$ es la matriz diagonal por bloques formada a partir de las matrices \mathbf{A} y \mathbf{B} . $N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ denota la distribución normal multivariante p -dimensional con vector de medias $\boldsymbol{\mu}$ y matriz de varianzas $\boldsymbol{\Sigma}$. $\text{Gamma}(p, a)$ denota la distribución gamma con media $\frac{p}{a}$ y varianza $\frac{p}{a^2}$. $U(a, b)$ denota la distribución uniforme en (a, b) .

2. Los datos

Los datos analizados corresponden a una muestra de 85832 grandes empresas de 9 países de la Unión Europea extraídas de la base de datos AMADEUS. El periodo analizado abarca desde 1998 al 2004 siendo la frecuencia de observación anual.

El ratio de rentabilidad analizado es la rentabilidad por activo (*Return on Assets*) (ROA). En la Tabla 1 se muestra la composición de la muestra por países y sectores. Así mismo, en la Tabla 2 se muestra el porcentaje de datos ausentes por sector y país. Se observa, en particular, que un 8.54% de las empresas analizadas tenían algún dato ausente oscilando dicho porcentaje entre un 7.51% y un 12.70% para países y un 7.60% y 12.52% para sectores. A diferencia de Bou y Satorra (2006), sin embargo, las empresas con datos ausentes no fueron eliminadas del análisis siendo sus valores imputados de forma aleatoria de acuerdo al proceso de estimación descrito en el Apéndice.

En la Tabla 3 se muestran los valores medianos del ROA temporal medio de cada empresa analizada según el país y el sector.

El ROA medio en el periodo analizado fue de un 0.08% siendo los sectores la Construcción, Hoteles y Restaurantes, Comercio, Electricidad y Gas y Manufactureros los más rentables mientras que los sectores de Transporte y Comunicaciones, Vivienda y, sobre todo, Otros Servicios (Educación, Sanidad, Administración Pública, Defensa, Actividades Sociales) el menos rentable. Por países se observa que las empresas de

Finlandia, Grecia y España han tendido a ser las más rentables en el periodo analizado mientras que las empresas de Italia, Bélgica y Alemania han sido las menos rentables. Particularizando a sectores de cada país se observa que en España y Finlandia la alta rentabilidad tiene un carácter general mientras que en Grecia son los sectores de Agricultura-Minería, Manufacturas, Electricidad y Gas, Comercio y, sobre todo, el de la Construcción los responsables de la alta rentabilidad de las empresas del país. En los países con baja rentabilidad (Bélgica, Alemania e Italia) dicho comportamiento se da en la mayor parte de los sectores con la única excepción del sector Electricidad y Gas en Bélgica que mostró una alta rentabilidad en el periodo. Llama en particular, la atención, el bajo nivel de rentabilidad de los Hoteles y Restaurantes en Alemania. Sin embargo, esta cifra tiene muy baja fiabilidad dado el escaso número de empresas de este sector presentes en la muestra (ver Tabla 1).

Finalmente se observa que existe más homogeneidad en las rentabilidades medianas a nivel de sectores que de países hecho que parece poner de manifiesto la mayor importancia del efecto país que del efecto sector en la rentabilidad de las empresas. Por dicha razón en el modelo que se presenta a continuación se decidió realizar el análisis por países en lugar de por sectores.

Tabla 1: Número de empresas analizadas por sector y país

Sector/País	Bélgica	Finlandia	Francia	Alemania	Grecia	Italia	España	Suecia	Reino Unido	Total
Agricultura-Minería	59	28	179	11	102	154	176	59	497	1265
Manufacturas	1757	782	4978	548	805	5081	2923	1546	5494	23914
Electricidad y Gas	64	83	120	258	12	83	157	182	163	1122
Construcción	391	118	847	42	124	697	849	251	1548	4867
Comercio	2564	801	6575	308	774	3665	3315	1958	4666	24626
Hoteles y Restaurantes	54	43	190	4	80	93	256	39	567	1326
Transporte y Comunicaciones	498	211	983	71	85	619	574	399	1510	4950
Inmobiliarias, renting y servicios empresariales	1457	410	4776	718	154	1095	2333	2292	6765	20000
Otros Servicios	182	61	578	181	64	366	377	175	1778	3762
Total	7026	2537	19226	2141	2200	11853	10960	6901	22988	85832

Tabla 2: Porcentaje de datos ausentes por sector y país

Sector/País	Bélgica	Finlandia	Francia	Alemania	Grecia	Italia	España	Suecia	Reino Unido	Total
Agricultura-Minería	10,17%	10,71%	11,73%	18,18%	10,78%	9,74%	10,80%	10,17%	13,08%	11,70%
Manufacturas	8,42%	10,23%	7,55%	12,23%	9,94%	7,54%	9,13%	8,21%	7,21%	8,05%
Electricidad y Gas	9,38%	10,84%	10,00%	15,50%	16,67%	10,84%	10,83%	12,09%	11,04%	12,03%
Construcción	15,86%	10,17%	9,56%	11,90%	10,48%	10,47%	9,54%	14,74%	8,20%	10,09%
Comercio	8,27%	9,99%	6,27%	17,53%	10,34%	9,03%	9,68%	8,58%	7,95%	8,24%
Hoteles y Restaurantes	11,11%	11,63%	12,63%	50,00%	11,25%	9,68%	14,84%	10,26%	12,17%	12,52%
Transporte y Comunicaciones	13,05%	13,74%	8,44%	9,86%	10,59%	11,31%	12,20%	15,54%	7,68%	10,32%
Inmobiliarias, renting y servicios empresariales	7,34%	15,37%	7,79%	10,31%	9,74%	7,67%	8,32%	8,42%	6,16%	7,60%
Otros Servicios	12,09%	9,84%	12,11%	11,60%	9,38%	16,12%	15,92%	10,86%	8,32%	10,93%
Total	9,02%	11,31%	7,55%	12,70%	10,23%	8,72%	9,74%	9,25%	7,51%	8,54%

Tabla 3: Valores medianos del ROA medio

Sector/País	Bélgica	Finlandia	Francia	Alemania	Grecia	Italia	España	Suecia	Reino Unido	Total
Agricultura-Minería	-0,33%	1,60%	0,12%	-0,12%	1,48%	-0,71%	0,95%	0,31%	-0,35%	0,04%
Manufacturas	-0,49%	1,54%	0,18%	-0,07%	1,92%	-1,35%	1,72%	1,43%	0,95%	0,19%
Electricidad y Gas	4,88%	0,11%	0,05%	0,06%	2,16%	-0,56%	0,66%	0,91%	0,19%	0,28%
Construcción	-0,29%	2,90%	0,16%	-0,77%	3,46%	-0,81%	1,11%	1,27%	1,42%	0,56%
Comercio	-0,48%	2,33%	0,44%	-0,07%	0,88%	-0,98%	1,03%	1,60%	0,94%	0,31%
Hoteles y Restaurantes	-0,56%	0,46%	-0,53%	-9,65%	-0,12%	-2,10%	1,74%	0,66%	0,67%	0,40%
Transporte y Comunicaciones	-0,75%	0,66%	-0,48%	-0,70%	0,15%	-2,02%	1,00%	0,57%	0,18%	-0,21%
Inmobiliarias, renting y servicios empresariales	-0,70%	0,47%	-0,38%	-0,79%	0,68%	-1,69%	0,93%	0,09%	-0,28%	-0,22%
Otros Servicios	-0,61%	1,44%	-0,32%	-0,50%	-0,13%	-2,40%	1,20%	0,12%	-1,51%	-0,72%
Total	-0,51%	1,49%	0,07%	-0,31%	1,31%	-1,21%	1,19%	0,86%	0,31%	0,08%

3. El modelo

Para cada país, los datos corresponden a una muestra de N empresas sobre las que se mide el valor de un ratio financiero R a lo largo del periodo $\{1, \dots, T\}$.

Sean $\{R_{it}; t \in T_i \subseteq \{1, \dots, T\}; i=1, \dots, N\}$ donde $T_i = \{t_{i1} < t_{i2} < \dots < t_{im_i} \leq T\}$ denotan los periodos de tiempo para los que se dispone de información del ratio R en la empresa i -ésima.

Sea $y_{it} = \log R_{it} - \log R_{\text{mean}, t}$ donde $R_{\text{mean}, t} = \sqrt[N_t]{\prod_{i \in N_t} R_{it}}$ es la media geométrica del

ratio R en el periodo t obtenida utilizando los datos de las N_t empresas para las que se dispone de información en dicho periodo;

El modelo viene dado por:

$$y_{it} = P_i + A_{it} + U_{it}; t \in I_i = \{t_{i1}, t_{i1}+1, \dots, t_{i,m_i}\}; i=1, \dots, N \quad (1)$$

donde:

$$P_i \sim N(0, \sigma_p^2) \quad i=1, \dots, N \quad (2)$$

es la componente permanente (de carácter estructural) de la evolución del ratio R en la empresa i -ésima a lo largo del tiempo. Dicha componente cuantifica las diferencias que se mantienen entre las empresas a largo plazo en el periodo de observación.

$$A_{it} = \beta A_{i,t-1} + D_{it}; t = t_{i1}+1, \dots, t_{i,m_i} \quad (3)$$

con $D_{it} \sim N(0, \sigma_A^2)$ y $A_{i,t_{i1}} \sim N\left(0, \frac{\sigma_A^2}{1-\beta^2}\right)$ $i=1, \dots, N$, donde A_{it} es la componente

transitoria (de carácter coyuntural) de la evolución del ratio financiero R en la empresa i -ésima a lo largo del tiempo. Esta componente cuantifica los efectos a corto plazo que explican las diferencias existentes en la evolución de las empresas. Dichos efectos tienden a desaparecer en el tiempo a la velocidad marcada por el coeficiente de ajuste β .

$$U_{it} \sim N(0, \sigma_{Ut}^2); t \in I_i = \{t_{i1}, t_{i1}+1, \dots, t_{i,m_i}\} = \{t_{i1}, \dots, t_{i,n_i}\}; i=1, \dots, N \quad (4)$$

donde U_{it} son las componentes específicas de cada empresa que recogen la influencia de shocks que afectan a la evolución de cada empresa individual en un periodo de tiempo concreto pero que, a diferencia de A_{it} , no tienen ninguna persistencia en el tiempo.

3.1 Distribución a priori

Viene dada por las siguientes distribuciones que se suponen independientes.

$$\tau_P = \frac{1}{\sigma_P^2} \sim \text{Gamma}(n_P, n_P s_P^2) \quad (5)$$

$$\tau_A = \frac{1}{\sigma_A^2} \sim \text{Gamma}(n_A, n_A s_A^2) \quad (6)$$

$$\beta \sim U(-1,1) \quad (7)$$

$$\tau_{U,t} = \frac{1}{\sigma_{U,t}^2} \sim \text{Gamma}(n_U, n_U s_U^2); t=1, \dots, T \quad (8)$$

con n_P, s_P, n_A, s_A, n_U y s_U constantes positivas conocidas.

3.2. Distribución a posteriori

Sea $\theta = (\tau_P, \tau_A, \tau_{U1}, \dots, \tau_{UT}, \beta)$ y $\mathbf{D} = (y_{it}; t \in T_i; i=1, \dots, N)$

El cálculo de la distribución a posteriori se realiza mediante el mecanismo de aumento de datos (Tanner y Wong, 1987). Para ello introducimos como parámetros adicionales

$$\mathbf{P} = (P_1, \dots, P_N)', \mathbf{A} = (A_{it}; t \in I_i; i=1, \dots, N), \mathbf{y}_{\text{miss}} = (y_{it}; t \in I_i - T_i; i=1, \dots, N)$$

Se tiene que:

$$\begin{aligned} & [\theta, \mathbf{P}, \mathbf{A}, \mathbf{y}_{\text{miss}} | \mathbf{D}] \propto \\ & \prod_{i=1}^N \left\{ \prod_{j=1}^{n_i} [y_{it_{ij}} | P_i, A_{it_{ij}}, \tau_{U_{t_{ij}}}] [P_i | \tau_P] \prod_{j=2}^{n_i} [A_{it_{ij}} | A_{it_{ij-1}}, \tau_A, \beta] [A_{it_{i1}} | \tau_A, \beta] \right\} \prod_{t=1}^T [\tau_{U_t}] [\tau_P] [\tau_A] [\beta] \propto \\ & \prod_{i=1}^N \left\{ \prod_{j=1}^{n_i} \tau_{U_{t_{ij}}}^{1/2} \exp \left[-\frac{\tau_{U_{t_{ij}}}}{2} (y_{it_{ij}} - P_i - A_{it_{ij}})^2 \right] \tau_P^{1/2} \exp \left[-\frac{\tau_P}{2} P_i^2 \right] \right\} \tau_P^{n_P-1} \exp[-\tau_P n_P s_P^2] I_{(0, \infty)}(\tau_P) \\ & \prod_{i=1}^N \left\{ \prod_{j=2}^{n_i} \tau_A^{1/2} \exp \left[-\frac{\tau_A}{2} (A_{it_{ij}} - \beta A_{it_{ij-1}})^2 \right] (\tau_A (1 - \beta^2))^{1/2} \exp \left[-\frac{\tau_A (1 - \beta^2)}{2} A_{it_{i1}}^2 \right] \right\} \end{aligned}$$

$$\tau_A^{n_A-1} \exp[-\tau_A n_A s_A^2] \mathbb{I}_{(0,\infty)}(\tau_A) \prod_{t=1}^T \tau_{U_t}^{n_{U_t}-1} \exp[-\tau_{U_t} n_{U_t} s_{U_t}^2] \mathbb{I}_{(0,\infty)}(\tau_{U_t}) \mathbb{I}_{(-1,1)}(\beta) \quad (9)$$

donde \mathbb{I}_A denota la función indicador de A .

3.3. Estimación de los parámetros

Dado que la distribución (9) no es tratable analíticamente recurrimos a métodos MCMC (ver, por ejemplo, Robert y Casella, 2004) y, más concretamente, al *Gibbs sampling* para su cálculo (ver Apéndice para más detalles). Como consecuencia de dicho algoritmo se obtiene una muestra

$$\{\boldsymbol{\theta}^{(s)} = (\tau_P^{(s)}, \tau_A^{(s)}, \tau_{U_1}^{(s)}, \dots, \tau_{U_T}^{(s)}, \beta^{(s)}), \mathbf{P}^{(s)}, \mathbf{A}^{(s)}, \mathbf{y}_{\text{miss}}^{(s)}; s = 1, \dots, S\} \quad (10)$$

de la distribución (9). A partir de dicha muestra se podrán realizar inferencias acerca de los parámetros del modelo (1) utilizando los momentos y cuantiles adecuados. En particular se podrán obtener estimaciones puntuales de dichos parámetros mediante las medias o medianas muestrales de las componentes correspondientes de (10) o construir intervalos de credibilidad del 95% utilizando los cuantiles 2.5 y 97.5 de dichas componentes.

4. Resultados

Con el fin de que la distribución a priori fuera no informativa se tomó $n_p = n_A = n_U = s_p = s_A = s_U = 0$. El algoritmo descrito en el Apéndice se ejecutó durante un número de iteraciones $S_{\text{max}} = 5000$ estimándose la convergencia de la cadena en $s_0 = 1000$ iteraciones. Se tomó finalmente una muestra de cada 5 ($L=5$) con el fin de reducir significativamente la autocorrelación de los componentes de la misma. En la Tabla 4 se muestran las estimaciones obtenidas para cada uno de los parámetros del modelo (1). Más concretamente se ha calculado, para cada parámetro, la mediana a posteriori así como los límites del intervalo de credibilidad bayesiano del 95% obtenido a partir de los cuantiles 2.5 y 97.5 de la muestra (10). Así mismo, en la Tabla 5 se muestra, para cada país, el porcentaje de la evolución total explicado por cada una de las componentes (permanente, transitoria y específica). Este porcentaje ha sido calculado a partir de las medianas a posteriori mostradas en la Tabla 4. En el caso particular de las componentes específicas dicho porcentaje se ha obtenido a partir de la variabilidad media en el periodo analizado.

Se observa, en primer lugar, que todas las componentes ejercen una influencia significativa en la evolución del ROA para las empresas analizadas. La componente permanente tiene especial importancia en la evolución de las empresas de Reino Unido, Suecia y, en menor medida, en España (ver Tabla 5). En el resto de los países la importancia de las 3 componentes es similar a excepción de Francia en el que la componente transitoria tiene mayor importancia, y Grecia en el que predomina la componente específica. Este último hecho hace pensar que algunos países como Grecia se adaptan con mayor dificultad a los ciclos económicos, con lo que su actividades económicas presentan un comportamiento más coyuntural.

La componente transitoria tiene, en general, un carácter poco persistente, aunque el coeficiente β se desvía significativamente de 0 en todos los países. Su valor oscila en torno a un 15%-20% (ver Tabla 4) para la mayor parte de los países. Este hecho refleja el alto grado de competitividad existente en el periodo analizado, que elimina rápidamente, las ventajas competitivas a corto plazo existentes en el mercado. Este proceso es especialmente intenso en Grecia y España en los que el coeficiente β es significativamente menor que en el resto de los países y es más suave en Francia, Reino Unido y Finlandia en los que es mayor.

Respecto a la componente específica se observa una clara tendencia decreciente en su importancia a lo largo del tiempo (ver Figura 1). Tan sólo en el año 2004 se aprecia un repunte en Finlandia y Alemania que podría deberse (sobre todo en el caso de Alemania) a la menor disponibilidad de datos en dicho año para ambos países.

Tabla 4: Estimaciones de los parámetros

Parámetro	País	Bélgica	Finlandia	Francia	Alemania	Grecia	Italia	España	Suecia	Reino Unido
β	Cuantil2.5	0.1188	0.1853	0.2105	0.1260	0.0478	0.1465	0.1007	0.1577	0.2061
	Mediana	0.1370	0.2165	0.2145	0.1508	0.0842	0.1570	0.1123	0.1726	0.2096
	Cuantil97.5	0.1519	0.2383	0.2241	0.1749	0.1140	0.1705	0.1223	0.1926	0.2144
σ_p	Cuantil2.5	11.22	7.65	12.03	10.44	8.33	6.25	8.58	13.09	29.47
	Mediana	11.48	7.88	12.18	10.88	8.63	6.34	8.72	13.38	29.78
	Cuantil97.5	11.74	8.15	12.34	11.30	8.97	6.44	8.88	13.64	30.06
$\frac{\sigma_A}{\sqrt{1-\beta^2}}$	Cuantil2.5	11.76	5.80	13.97	10.72	5.96	6.16	8.16	8.77	19.35
	Mediana	12.02	6.06	14.10	11.19	6.59	6.30	8.31	9.00	19.48
	Cuantil97.5	12.28	6.29	14.25	11.58	7.31	6.40	8.50	9.30	19.60
$\sigma_{U,1998}$	Cuantil2.5	11.38	4.20	12.75	16.98	24.48	8.25	11.68	12.83	19.70
	Mediana	11.93	4.65	13.13	17.95	25.34	8.51	12.02	13.31	20.17
	Cuantil97.5	12.51	5.15	13.51	18.98	26.23	8.74	12.33	13.77	20.61
$\sigma_{U,1999}$	Cuantil2.5	22.10	1.89	10.74	3.25	24.11	7.88	9.37	14.93	19.81
	Mediana	22.62	2.26	11.18	4.09	24.93	8.07	9.68	15.37	20.23
	Cuantil97.5	23.09	2.63	11.52	5.18	25.81	8.30	9.94	15.78	20.61
$\sigma_{U,2000}$	Cuantil2.5	6.29	3.74	5.23	4.54	10.76	2.14	3.17	12.46	8.90
	Mediana	7.02	4.23	5.77	5.38	11.54	2.39	3.57	12.93	9.43
	Cuantil97.5	7.73	4.68	6.45	6.45	12.19	2.81	3.91	13.35	9.90
$\sigma_{U,2001}$	Cuantil2.5	2.55	4.55	7.50	3.81	5.68	1.73	4.65	17.09	6.62
	Mediana	3.23	5.03	7.84	4.71	6.55	1.88	5.02	17.52	7.15
	Cuantil97.5	3.90	5.43	8.18	6.08	7.26	2.09	5.39	17.93	7.70
$\sigma_{U,2002}$	Cuantil2.5	11.89	4.33	2.64	8.18	3.48	3.15	2.71	12.12	3.47
	Mediana	12.39	4.80	3.00	9.06	4.63	3.44	3.24	12.53	4.02
	Cuantil97.5	12.91	5.25	3.54	9.99	5.42	3.74	3.67	12.90	4.36
$\sigma_{U,2003}$	Cuantil2.5	13.04	5.76	2.29	7.85	5.54	4.82	5.59	2.01	1.96
	Mediana	13.60	6.25	2.69	8.96	6.59	5.04	6.01	2.25	2.49
	Cuantil97.5	14.05	6.68	3.04	9.87	7.28	5.30	6.37	2.67	2.71
$\sigma_{U,2004}$	Cuantil2.5	2.77	17.22	1.99	19.06	6.25	6.59	2.91	3.41	2.75
	Mediana	3.46	17.81	2.24	20.26	7.18	6.80	3.33	4.07	3.03
	Cuantil97.5	4.38	18.43	2.51	21.61	7.96	7.02	4.01	4.51	3.61

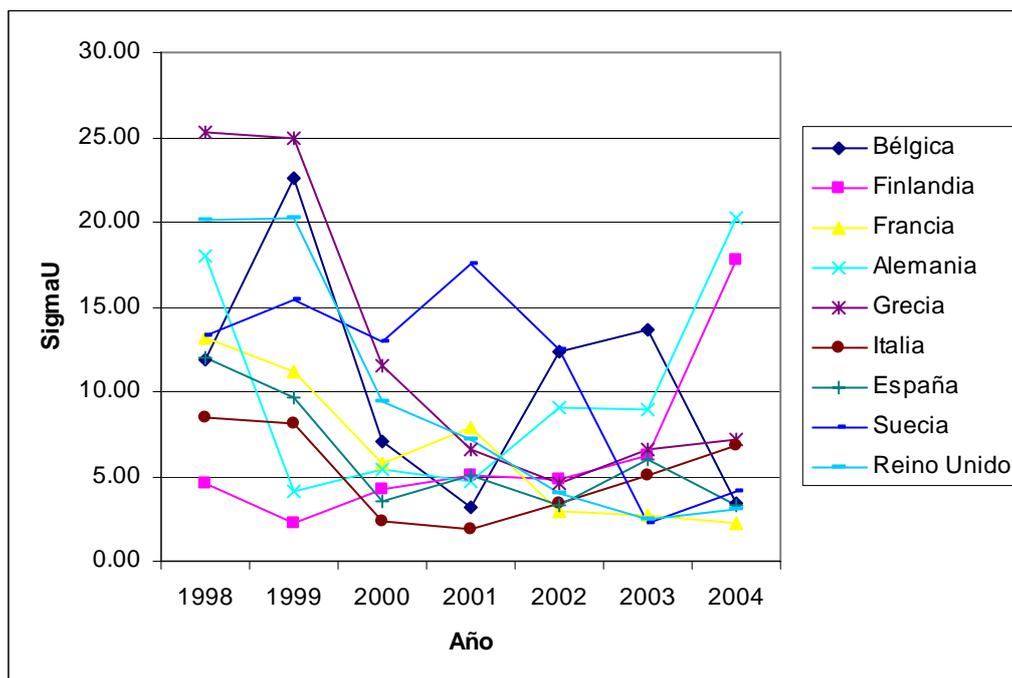


Figura 1: Estimaciones de $\sigma_{U,t}$; $t=1998, \dots, 2004$ para cada país

En la Tabla 6 se muestran los resultados obtenidos al realizar un estudio comparativo por sector y país de las componentes permanentes. Más concretamente se muestran los valores medianos de dichas componentes calculados a partir de las medianas a posteriori de los parámetros individuales $\{P_i; i=1, \dots, N\}$. Se observa la existencia de diferencias significativas tendiendo a tomar valores más altos en los sectores más rentables de cada país (comparar Tablas 3 y 6).

Realizando un estudio comparativo mediante contrastes de Kruskal-Wallis en cada país, se observa que en una gran cantidad de ellos los sectores en los que existe un alto grado de competitividad y propiedad privada (Manufacturas, Comercio al por menor, Construcción y, en el caso de España el sector de Hoteles y Restaurantes debido a la importancia del turismo), han tendido a ser los más rentables de forma permanente. Por el contrario los sectores con una mayor presencia del sector público -sea esta por regulación o por financiación- como Agricultura-Minería, Transporte y Comunicaciones y Otros Servicios, tienden a ser los menos rentables. Así, en Agricultura hay fuertes limitaciones a la competencia (barreras de entrada) y es un sector altamente

subvencionado. En Minería, en cambio, el problema reside en la difícil competencia con el exterior, que produce a costes menores. Lo mismo sucede con el sector del Transporte, un sector que trabaja con márgenes muy bajos (y que han ido decreciendo en los últimos años) debido a una gran competencia. Por último en Otros Servicios son empresas con un importante peso de los costes de personal y que cuentan con más financiación pública que el resto de sectores (educación, sanidad, actividades sociales).

Finalmente, en la Tabla 7 se analiza la bondad de ajuste de los modelos estimados. En particular se muestra el coeficiente de correlación múltiple R^2 así como los niveles de cubrimiento de los intervalos predictivos bayesianos con un nivel del 95% calculados, estos últimos, a partir de la muestra (10) mediante el muestreo de composición. Se observa que la bondad de ajuste es de tipo medio con un R^2 global del 68.38%. Dicha bondad de ajuste no es homogénea. Así en Francia y Reino Unido el grado de ajuste del modelo a los datos es medio-alto con R^2 de alrededor de alrededor del 75%. En Grecia, Finlandia y Suecia, por el contrario, el grado de bondad de ajuste es bajo con R^2 que oscilan entre un 20%-40%. Finalmente los intervalos predictivos bayesianos muestran una clara tendencia al sobrecubrimiento debido, fundamentalmente, al alto nivel de incertidumbre asociado a la estimación de los parámetros P_i y A_{it} de las empresas individuales al disponer a lo más de 7 datos por serie.

Tabla 5: Porcentaje de la evolución del ROA explicado por cada componente para cada país

Componente	País	Bélgica	Finlandia	Francia	Alemania	Grecia	Italia	España	Suecia	Reino Unido
Permanente	Cuantil2.5	29.89%	36.56%	35.90%	29.18%	20.58%	34.98%	38.61%	42.31%	62.51%
	Media	30.75%	38.14%	36.51%	31.02%	21.90%	35.70%	39.41%	43.37%	63.01%
	Cuantil97.5	31.77%	39.79%	37.13%	32.76%	23.32%	36.48%	40.22%	44.39%	63.49%
Transitoria	Cuantil2.5	32.32%	20.54%	48.02%	30.08%	10.42%	33.69%	34.54%	18.57%	26.52%
	Media	33.72%	22.48%	48.94%	32.78%	12.90%	35.08%	35.91%	19.72%	26.96%
	Cuantil97.5	35.03%	24.36%	49.75%	34.99%	15.73%	36.27%	37.39%	20.96%	27.43%
Específica	Cuantil2.5	33.92%	37.10%	13.73%	33.26%	61.88%	28.08%	23.17%	35.38%	9.70%
	Media	35.53%	39.38%	14.55%	36.20%	65.20%	29.22%	24.68%	36.91%	10.02%
	Cuantil97.5	37.12%	41.90%	15.32%	39.28%	68.24%	30.76%	26.04%	38.35%	10.37%

Tabla 6: Valores medianos de la componente permanente por sector y país

(en negrita los sectores significativamente más rentables)

Sector/País	Bélgica	Finlandia	Francia	Alemania	Grecia	Italia	España	Suecia	Reino Unido
Agricultura-Minería	2.17%	-0.25%	1.65%	1.43%	0.55%	1.22%	0.22%	-0.19%	2.90%
Manufacturas	1.44%	-0.26%	1.60%	1.31%	0.95%	0.55%	0.71%	1.17%	4.02%
Electricidad y Gas	6.21%	-1.61%	1.47%	1.39%	0.06%	1.29%	-0.27%	0.66%	3.29%
Construcción	1.64%	1.51%	1.59%	0.60%	1.25%	1.12%	0.24%	0.65%	4.47%
Comercio	1.47%	0.53%	1.90%	1.37%	0.23%	0.93%	0.09%	1.21%	4.02%
Hoteles y Restaurantes	1.46%	-1.83%	1.07%	-8.61%	-0.95%	-0.33%	0.48%	-0.30%	3.72%
Transporte y Comunicaciones	1.18%	-1.01%	0.87%	0.75%	-0.69%	-0.02%	0.02%	0.17%	3.29%
Vivienda	1.19%	-1.35%	0.97%	0.49%	-0.41%	0.20%	0.02%	-0.28%	2.74%
Otros Servicios	1.38%	-0.39%	1.17%	0.68%	-0.86%	-0.46%	0.22%	0.30%	1.53%

Tabla 7: Bondad de ajuste de los modelos estimados

País	Bélgica	Finlandia	Francia	Alemania	Grecia	Italia	España	Suecia	Reino Unido	Total
CUB95	99.49	99.16	99.71	99.36	98.95	99.54	99.59	99.16	99.66	99.55
R ²	0.5501	0.3700	0.7855	0.5485	0.2007	0.5939	0.6228	0.3839	0.7544	0.6838

5. Conclusiones

En este trabajo se ha realizado un análisis estadístico bayesiano de la evolución de las rentabilidades anormales de una muestra de empresas de la Unión Europea. Para ello se ha utilizado un modelo de ecuaciones estructurales dinámico propuesto por Bou y Satorra (2007) que permite descomponer dicha evolución en sus componentes permanente (de carácter estructural), transitoria y específica (ambas de carácter más coyuntural) así como evaluar su importancia. El análisis se ha realizado desde una óptica bayesiana que, al no depender de resultados asintóticos en su proceso de estimación, y permitir un tratamiento natural de datos ausentes mediante imputaciones aleatorias, dota de mayor robustez, flexibilidad y realismo al estudio realizado.

Los resultados obtenidos ponen de manifiesto la significación estadística de las 3 componentes. El comportamiento no es, sin embargo, homogéneo y depende del sector y país analizado. Se observa, sin embargo, con carácter general, un bajo nivel de persistencia en la componente transitoria, fruto del creciente nivel de competitividad presente en el panorama empresarial de los países analizados. Por sectores, son aquéllos más relacionados con el sector privado (Comercio, Manufacturas, Construcción y, en el caso de España, el de Hoteles y Restaurantes) lo que muestran una tendencia de carácter más permanente a obtener rentabilidad más altas. Por el contrario los sectores más subvencionados y/o relacionados con el Sector Público (Agricultura-Minería, Transporte y Comunicaciones, Educación, Sanidad, etc) muestran una tendencia a obtener rentabilidades más reducidas.

En el análisis realizado se ha supuesto que el proceso de ajuste es el mismo para todos los sectores. La evidencia obtenida sugiere que esta hipótesis puede ser demasiado restrictiva. Extender el modelo permitiendo la existencia de efectos específicos del sector en el que la empresa desarrolla su actividad es el siguiente objetivo de nuestra agenda de investigación. Los resultados obtenidos serán expuestos en otra parte.

6. Apéndice

En este apéndice se describe el algoritmo utilizado para la obtención de la muestra (10) utilizada para realizar inferencias acerca de los parámetros del modelo (1).

El algoritmo se basa en el *Gibbs sampling* y consta de los siguientes pasos:

Paso 0. (Inicialización)

Se extrae una muestra $\{\boldsymbol{\theta}^{(0)} = (\tau_p^{(0)}, \tau_A^{(0)}, \tau_{U_1}^{(0)}, \dots, \tau_{U_T}^{(0)}, \beta^{(0)}), \mathbf{P}^{(0)}, \mathbf{A}^{(0)}\}$ de las distribuciones a priori (2)-(3) y (5)-(8). Se fija el número máximo de iteraciones S_{\max} y se coloca el contador $s=1$.

Paso 1 (Imputación de datos ausentes)

Se extraen $y_{it}^{(s)}$ de $N\left(P_i^{(s-1)} + A_{it}^{(s-1)}, \frac{1}{\tau_{Ut}^{(s-1)}}\right)$ para $t \in I_i - T_i$ e $i=1, \dots, N$

Paso 2 (Distribución completamente condicionada de $(P_i, A_{it}; t \in I_i) \mid \mathbf{D}, \text{resto}$)

Extraer $(P_i^{(s)}, A_{it}^{(s)}, t \in I_i)$ de $N_{n_i}(\mathbf{m}_i^{(s)}, \mathbf{S}_i^{(s)})$ para $i=1, \dots, N$ donde

$$\mathbf{m}_i^{(s)} = \mathbf{S}_i^{(s)} \left(\mathbf{X}_i' (\mathbf{D}_i^{(s-1)})^{-1} \mathbf{y}_i^{(s)} \right), \quad \mathbf{S}_i^{(s)} = \left(\mathbf{X}_i' (\mathbf{D}_i^{(s-1)})^{-1} \mathbf{X}_i + (\boldsymbol{\Sigma}_i^{(s-1)})^{-1} \right)$$

donde $\mathbf{X}_i = (\mathbf{1}_{n_i \times 1}, \mathbf{I}_{n_i})$, $(\mathbf{D}_i^{(s)})^{-1} = \text{diag}(\tau_{U_t}^{(s)}; t \in I_i)$ y $(\boldsymbol{\Sigma}_i^{(s)})^{-1} = \text{blockdiag}(\tau_p^{(s)}, \mathbf{M}(n_i, \tau_A^{(s)}, \beta^{(s)}))$ con $\mathbf{M}(n, \tau, \beta)$ denotando la inversa de la matriz $n \times n$ $\mathbf{B} = (b_{ij})$ con $b_{ij} = \sigma_{ij} = \frac{1}{\tau(1-\beta^2)} \beta^{|i-j|}$

Paso 3: (Distribución completamente condicionada de $\tau_{Ut} \mid \mathbf{D}, \text{resto}$)

Extraer $\tau_{U_t}^{(s)}$ de $\text{Gamma}\left(\frac{|M_t|}{2} + n_U, n_U s_U^2 + \frac{1}{2} \sum_{i \in M_t} (y_{it}^{(s)} - P_i^{(s)} - A_{it}^{(s)})^2\right)$ donde $M_t =$

$\{i \in \{1, \dots, N\} : t \in N_i\}$ para $t=1, \dots, T$

Paso 4: (Distribución completamente condicionada de $\tau_p \mid \mathbf{D}, \text{resto}$)

Extraer $\tau_p^{(s)}$ de $\text{Gamma}\left(\frac{N}{2} + n_p, n_p s_p^2 + \frac{1}{2} \sum_{i=1}^N P_i^{(s)2}\right)$

Paso 5: (Distribución completamente condicionada de $\tau_A \mid \mathbf{D}, \text{resto}$)

Extraer $\tau_A^{(s)}$ de

$$\text{Gamma} \left(\frac{\sum_{i=1}^N n_i}{2} + n_A, n_A s_A^2 + \frac{1}{2} \left((1 - \beta^{(s-1)2}) \left(\sum_{i=1}^N A_{it_i}^{(s)2} \right) + \sum_{i=1}^N \sum_{j=2}^{n_i} (A_{it_j}^{(s)} - \beta A_{it_{j-1}}^{(s)})^2 \right) \right)$$

Paso 6: (Distribución completamente condicionada de β | \mathbf{D} , resto)

Extraer u de $U\left(0, (1 - \beta^{(s-1)2}) \frac{N}{2}\right)$ y luego $\beta^{(s)}$ de $N(m_\beta^{(s)}, s_\beta^{(s)})$ truncada en

$(\beta_{\text{inf}}, \beta_{\text{sup}})$ donde:

$$m_\beta^{(s)} = \frac{\sum_{i=1}^N \sum_{j=2}^{n_i} A_{it_j}^{(s)} A_{it_{j-1}}^{(s)}}{\sum_{i=1}^N \sum_{j=2}^{n_i-1} A_{it_j}^{(s)2}}, \quad s_\beta = \frac{1}{\tau_A \sum_{i=1}^N \sum_{j=2}^{n_i-1} A_{it_j}^{(s)2}}, \quad \beta_{\text{inf}} = -\sqrt{1 - u \frac{2}{N}} \quad \text{y} \quad \beta_{\text{sup}} = \sqrt{1 - u \frac{2}{N}}.$$

Paso 7: (Incremento de iteración)

Poner $s = s+1$. Si $s \leq S_{\text{max}}$ ir al paso 1. En otro caso acabar.

Como consecuencia del algoritmo se obtiene una muestra

$$\left\{ \boldsymbol{\theta}^{(s)} = (\tau_P^{(s)}, \tau_A^{(s)}, \tau_{U_1}^{(s)}, \dots, \tau_{U_T}^{(s)}, \beta^{(s)}), \mathbf{P}^{(s)}, \mathbf{A}^{(s)}, \mathbf{y}_{\text{miss}}^{(s)}; s = s_0, s_0 + L, \dots, s_0 + (S-1)L = S_{\text{max}} \right\}$$

donde s_0 es el número de iteraciones estimado para que la cadena de Markov asociada al Gibbs sampling converja a su distribución estacionaria (9) y L se elige de forma que la autocorrelación de las componentes de $\boldsymbol{\theta}$ sea despreciable para retardos superiores a L . En nuestro caso utilizamos el procedimiento de Geweke (1992) para establecer s_0 y L se determinó mediante inspección visual de los correlogramas de las componentes de $\boldsymbol{\theta}$.

Bibliografía

Bou, J.C. y Satorra, A. (2007): "The persistence of abnormal returns at industry and firm levels: evidence from Spain" (forthcoming in *Strategic Management Journal*).

Geweke J. (1992): "Evaluating the Accuracy of Sampling-Based Approaches to the Calculation of Posterior Moments", *Bayesian Statistics 4*, (eds. Bernardo, J.M; Berger, J.O.; Dawid, A.P. and Smith, A.F.M.), 169-194.

Jacobson, R. (1988). "The persistence of abnormal returns". *Strategic Management Journal*, **9**, 415-430.

McGanan A.M. y Porter, M.E. (2003). “The emergente and sustainability of abnormal profits”. *Strategic Organization*, **1**, 79-108.

Robert, C.P. y Casella, G. (2004): *Monte Carlo Statistical Methods. Second Edition*. New-York, Springer-Verlag.

Tanner, M.A. y Wong, W.H. (1987): “The calculation of posterior distributions by data augmentation” (with discussion). *Journal of the American Statistical Association*, **82**: 528-550.

Waring, G. (1996). “Industry differences in the persistence of firm-specific returns”. *American Economic Review*, **86**, 1253-1265.

CONSIDERACIONES EN TORNO AL ANÁLISIS MULTIRRESOLUCIÓN DE SHANNON EN SERIES TEMPORALES.

FEDERICO PALACIOS GONZÁLEZ
e-mail: fpalacio@ugr.es

Departamento de METODOS CUANTITATIVOS PARA LA ECONOMÍA Y LA EMPRESA
UNIVERSIDAD DE GRANADA

Área temática: Métodos cuantitativos.

Resumen

El desarrollo tecnológico de las últimas décadas, en lo que respecta a sistemas de compresión de ficheros de video y audio, ha impulsado los métodos de construcción de sistemas ortogonales en el espacio de Hilbert de las funciones de cuadrado integrable, que se aplican en el ámbito del proceso digital de señales bajo las etiquetas de Análisis Multirresolución y Análisis Wavelet.

Los resultados obtenidos al respecto en la pasada década, por una parte, comienzan a encontrar en el presente, aplicación para el tratamiento de series económicas; por otra parte, se están relacionando con los métodos tradicionales del análisis de series temporales en el dominio de la frecuencia.

En este trabajo se hace una adaptación del Análisis Multirresolución de Shannon para el tratamiento de series que previamente son convertidas a señales en tiempo continuo mediante interpolación. Cada aproximación de dicha señal es el resultado de aplicarles un filtro variable en el tiempo definido como un operador lineal de núcleo.

La metodología se aplica a datos simulados y a la serie económica de la variación porcentual trimestral del PIB en el periodo 1980.4 - 2004.4

Palabras clave: Análisis Multirresolución, Análisis Wavelets, Filtro Lineal Variable en el Tiempo, Operador Lineal de Núcleo, Series temporales.

Abstract

Technological progress over the past two decades, with regard to video and audio file compressing systems, has enhanced the orthogonal system construction methods in the $L_2(\mathbb{R})$ Hilbert space, being applied in the digital signal processing (DSP) under the labels of Multi-resolution Analysis and Wavelet Analysis.

The results obtained in recent years have, on the one hand, begun to be used in economic series treatment, and, on the other, are being related to traditional methods of time series analysis in the frequency domain.

This paper shows how the Shannon Multi-resolution Analysis is applicable to the treatment of series that have previously been transformed into continuous time signals by means of interpolation. Each approximation of these signals is the output of a Time Varying Lineal Operator. The methodology is applied to simulated data and to a quarterly percentage variation economic series of the Gross Domestic Product (GDP) from 1980.4 to 2004.4.

Key words: Multi-Resolution Analysis, Wavelet Analysis, Time Varying Linear Filter, Kernel Linear Operator, Time Series.

1. Introducción

El desarrollo de la teoría de “Multi-Resolution Analysis” (MRA) y “Wavelet Analysis” (WA), Hernández y Weiss (1996), para la aproximación y descomposición en sumas ortogonales de funciones en el espacio de Hilbert de las funciones de cuadrado integrable, ha proporcionado una infraestructura teórica utilizada como marco universal ideal para el análisis de todo tipo de señales y en particular, de las que provienen de los distintos sistemas económicos. Los prolegómenos del MRA se remontan a 1981/83, en el ámbito del análisis microlocal, dentro de la teoría de ecuaciones diferenciales. Sus progenitores son Peter J. Burt, Edward H. Adelson and James Crowley. El MRA adquiere gran difusión cuando se introduce como herramienta para la construcción de Wavelets en 1987/89, inicialmente por Stephane Mallat(1987-1989) e Yves Meller (1988 y 1989c) que son los que introducen el MRA en el campo de las señales digitales, adaptando la teoría de tiempo continuo a tiempo discreto dentro de lo que se conoce como Proceso Digital de Señales (DSP). La gran popularidad de estas herramientas, principalmente la versión en tiempo discreto, es debida a que el análisis Wavelet (WA) es la base que subyace en los sistemas de compresión de ficheros utilizados en la tecnología digital. Ambos, MRA y WA, han demostrado ser herramientas muy eficaces para el análisis de señales tanto en tiempo discreto como continuo y tienen gran aplicación en el ámbito de las series económicas; véase, por ejemplo, Percival y Walden (2000).

En este trabajo se esboza una propuesta para el uso de tal tecnología para el análisis, en tiempo continuo, de señales que se obtiene mediante interpolación temporal de la información en tiempo discreto procedente de una serie temporal. El objetivo es obtener un mayor rendimiento en el análisis de series de corta longitud como son, generalmente, las que provienen del ámbito económico.

Por otra parte, con objeto de explicitar los mecanismos que actúan en los filtros subyacentes en estos tipos de análisis, se establece una relación con los operadores lineales, Meyer (1989a y 1989b), definidos mediante una función núcleo sobre un espacio de Hilbert, harto conocidos en el seno del análisis funcional en la primera mitad del siglo XX (véase por ejemplo Bachman y Narici (1981)). Concretamente se

explicitan estos filtros, variables en el tiempo, como tales operadores para el caso del MRA y WA de Shannon originados por la función $\text{sinc}(t) = \text{sen } \pi t / \pi t$.

Cuando se trabaja en tiempo discreto mediante convolución de series finitas con un filtro cuya función impulso-respuesta contiene un número finito términos (FIRF), es conveniente tener presente, Iacobucci y Noullez (2004), que es imposible diseñar un filtro que preserve o elimine todas las frecuencias en un rango dado. El llamado filtro ideal es inalcanzable en este caso. De hecho, especialmente cuando se trata de series de corta longitud, variaciones abruptas en la frecuencia de respuesta produce lo que se conoce como fenómeno de Gibbs; la aparición de fluctuaciones espúreas y artificiales en la señal filtrada. No obstante, el tratamiento en tiempo continuo mediante interpolación mitiga considerablemente tal fenómeno de Gibbs que casi se reduce al “efecto borde”¹. Dicho efecto se hace más patente cuando los valores inicial y final de la serie difieren excesivamente creando una discontinuidades manifiestas en los puntos kT , $k=0,1,2,\dots$ al extender periódicamente la señal, desde el intervalo de observación $[0, T]$ a toda la recta real, o bien, cuando la señal se prolonga con un valor constante en el exterior del intervalo $[0, T]$. Un análisis detallado de este hecho puede verse, por ejemplo, en Hamming (1998), Oppenheim and Schaffer (1999) y Priestly (1981).

En el trabajo de Iacobucci y Noullez (2004) se hace una comparación de los filtros existentes, todos ellos FIRF en tiempo discreto. En este trabajo se comparan los mejores resultados allí obtenidos con los que producen, sobre los mismos datos, las técnicas aquí descritas.

1.1 Organización del trabajo

En la sección 2 se describen las estructuras MRA y la forma de construir sucesiones de niveles de resolución alternativas a las de tipo binario (potencias de 2) para conseguir estructuras de MRA compatibles con el intervalo $[0, T]$ de observación de la serie temporal.

En la sección 3 se muestra cómo hacer una interpolación $y(t)$ de la serie en el seno del MRA, de forma que ésta pertenezca al espacio V_0 . También se explicita la

¹ Lo distorsión se hace notable en cuando t se acerca a los extremos o “bordes” del intervalo $[0, T]$

forma de las proyecciones de $y(t)$ sobre los distintos subespacios del MRA de Shannon que tiene como función de escalado la función

$$\text{sinc}(t) = \frac{\text{sen}(\pi t)}{\pi t}.$$

En la sección 4 se muestra la transformada de Fourier de dichas proyecciones, quedando de manifiesto que éstas son aproximaciones de $y(t)$ y que, en ellas, el espectro de frecuencias queda recortado y reducido a frecuencias dentro de la banda $[-\pi s_j, \pi s_j]$.

En la sección 5 se muestra la forma de tales proyecciones como operadores de núcleo que se definen a partir de la propia función sinc.

En la sección 6 se muestra cómo la señal interpolada admite una descomposición finita en suma de funciones ortogonales: Una aproximación en la banda de bajas frecuencias $[0, \omega_c]$ y detalles en las sucesivas bandas de frecuencias $(\omega_{j+1}, \omega_j]$ con $j = 0, 1, \dots, c-1$ siendo c el número de divisores primos de T , contando las multiplicidades de cada uno de ellos. Tal descomposición pertenece al Wavelet de Shannon asociado al MRA.

En la sección 7 se aplica, la metodología descrita, a diversos ejemplos con datos simulados, y a la serie trimestral² de las variaciones porcentuales del PIB durante el periodo 1980.4 a 2004.4.

En la sección 8 se hace una comparación de filtros utilizando los resultados obtenido por Iacobucci y Noullez (2004).

La sección 9 contiene las conclusiones.

² Datos obtenidos de la base de datos del INE: Contabilidad Nacional Trimestral de España

2. Análisis Multirresolución.

En general, Mallat (1999), Hernandez y Weiss (1999), un MRA es una sucesión de subespacios cerrados de $L^2(\mathbb{R})$ $\{V_j\}_{j \in \mathbb{Z}}$ para la cual se verifican las siguientes 6 propiedades

$$(i) \quad \forall (j, k) \in \mathbb{Z}^2, \quad f(t) \in V_j \Leftrightarrow f(t - 2^j k) \in V_j$$

$$(ii) \quad \forall j \in \mathbb{Z}, \quad V_j \supset V_{j+1}$$

$$(iii) \quad \forall j \in \mathbb{Z}, \quad f(t) \in V_j \Leftrightarrow f(t/2) \in V_{j+1}$$

$$(iv) \quad \lim_{j \rightarrow -\infty} V_j = \bigcap_{j=-\infty}^{+\infty} V_j = \{0\}$$

$$(v) \quad \lim_{j \rightarrow -\infty} V_j = \text{Cierre} \left(\bigcup_{j=-\infty}^{+\infty} V_j \right) = L^2(\mathbb{R})$$

$$(vi) \quad \text{Existe una función } \theta \text{ tal que } \{\theta(x-k)\}_{k \in \mathbb{Z}} \text{ es una base de Riesz en } V_0$$

La función θ recibe el nombre de función de escalado (scaling function, father wavelet).

Como consecuencia de las anteriores condiciones, cada subespacio V_j está engendrado por la base de Riesz

$$\{\theta_{j,k}(t) = 2^{-j/2} \theta(2^{-j}t - k)\}_{k \in \mathbb{Z}}$$

y cualquier función $f \in L_2(\mathbb{R})$ puede ser aproximada a distintos niveles de resolución mediante proyecciones ortogonales de f sobre los subespacios V_j . Dichas proyecciones pueden escribirse como

$$f_j(t) = \sum_{k \in \mathbb{Z}} a_{j,k} \theta_{j,k}(t).$$

Si las bases de Riesz son ortonormales, entonces

$$a_{j,k} = \int_{-\infty}^{+\infty} \theta_{j,k}(t) f(t) dt$$

En particular, cuando $\theta(t) = \text{seno}(\pi t)/\pi t$, se verifica que $\{\theta(x-k)\}_{k \in \mathbb{Z}}$ es una base de ortonormal en V_0 , Mallat (1999). Y las bases definidas en (1) son ortonormales en sus respectivos espacios V_j . Se obtiene, de este modo, el MRA de Shannon.

La sucesión de valores $s_j = 2^{-j}$ $j \in \mathbb{Z}$ recibe el nombre de niveles de resolución y sus inversos el de valores de escala.

La elección de sucesivas potencias de 2 como valores de escala es algo que se ha impuesto en las aplicaciones y en la mayoría de las aproximaciones teóricas al tema. Pero esto no ha de ser necesariamente así, Pollock y Lo Casio (2005). Basta que se verifique, para todo j , que s_j es múltiplo de s_{j+1} ; simbólicamente $s_j = \dot{s}_{j+1}$. En este caso, la estructura para el análisis multiresolución (MRA) quedaría reformulada de la siguiente manera

$$(i) \quad \forall (j,k) \in \mathbb{Z}^2, \quad f(t) \in V_j \Leftrightarrow f\left(t - \frac{k}{s_j}\right) \in V_j$$

$$(ii) \quad \forall j \in \mathbb{Z}, \quad V_j \supset V_{j+1}$$

$$(iii) \quad \forall j \in \mathbb{Z}, \quad f(t) \in V_j \Leftrightarrow f\left(\frac{s_{j+1}}{s_j} t\right) \in V_{j+1}$$

$$(iv) \quad \lim_{j \rightarrow \infty} V_j = \bigcap_{j=-\infty}^{+\infty} V_j = \{0\}$$

$$(v) \quad \lim_{j \rightarrow -\infty} V_j = \text{Cierre}\left(\bigcup_{j=-\infty}^{+\infty} V_j\right) = L^2(\mathbb{R})$$

$$(vi) \quad \text{Existe una función } \theta \text{ tal que } \{\theta(x-k)\}_{k \in \mathbb{Z}} \text{ es una base de Riesz en } V_0$$

Cada subespacio V_j está engendrado por la base de Riesz

$$\left\{ \theta_{j,k}(t) = \sqrt{s_j} \theta(s_j t - k) \right\}_{k \in \mathbb{Z}} \quad (1)$$

En este trabajo se utilizarán secuencias con la forma ³

$$s_j = \frac{m_j}{T} \quad \forall j > 0 \quad \text{y} \quad s_{-j} = \frac{1}{s_j} \quad (2)$$

donde $m_0 = T$ y $m_j = m_{j+1}$ y donde $[0, T]$ es el rango de tiempo de observación de la serie temporal que se ha de analizar.

Para la construcción de los distintos m_j se utilizarán los divisores primos de T de la siguiente forma:

Supóngase que hay $d_1 < d_2 < \dots < d_p$ divisores primos de T con sendas multiplicidades r_1, r_2, \dots, r_p . Sean

$$c = \sum_{i=1}^p r_i \quad \text{y} \quad \mathbf{b} = (b_j)_{j=1,2,\dots,c} \quad (3)$$

donde

$$b_j = \begin{cases} d_1 & \text{si } 1 \leq j \leq r_1 \\ d_2 & \text{si } r_1 + 1 \leq j \leq r_1 + r_2 \\ \vdots & \vdots \\ d_p & \text{si } \sum_{j=1}^{p-1} r_j + 1 \leq j \leq c \end{cases}$$

Sea $\eta(\mathbf{b}) = (b_{\eta(j)})_{j=1,2,\dots,c}$ una permutación cualquiera η de las componentes de \mathbf{b} . Dado un valor $j > 0$ arbitrario, se calcula

³ Los valores s_j deben estar definidos para que el MRA quede completo, pero en realidad aquí o serán usados. Lo mismo sucede con los s_j para $j > c$

$$m_{j+1} = \frac{m_j}{b_{\eta(j)}} \quad (4)$$

teniendo en cuenta que $m_0 = T$. Nótese que siempre se verifica que $m_c = 1$

Este modo de construir los distintos niveles de resolución permite afirmar que a cualquier nivel de resolución s_j , siempre existe un valor k , tal que $k/s_j = T$ y, tal que la sucesión de puntos k/s_j $k = 0, 1, 2, \dots, m_j$ define una partición regular en m_j segmentos del intervalo $[0, T]$. Se dirá, en este caso, que la sucesión de niveles de resolución s_j es compatible con el intervalo de observación $[0, T]$.

Dada una permutación particular $\eta(\mathbf{b})$ del vector \mathbf{b} , mediante (2) y (4), se obtiene la correspondiente sucesión de valores m_j y de niveles de resolución s_j tal y como se muestra en la Tabla 1. Las distintas permutaciones permiten definir distintas sucesiones de valores m_j , lo cual enriquece las posibilidades de análisis (véanse secciones 6 y 7).

Tabla 1

j	$d_{\eta(j)}$	m_j	$s_j = m_j/T$
0	-----	T	1
1	$d_{\eta(1)}$	$T / d_{\eta(1)} = \prod_{i=2}^c d_{\eta(i)}$	$1/d_{\eta(1)}$
..	
j	$d_{\eta(j)}$	$T / \prod_{i=1}^j d_{\eta(i)} = \prod_{i=j+1}^c d_{\eta(i)}$	$1 / \prod_{i=1}^j d_{\eta(i)}$
..		
$c-1$	$d_{\eta(c-1)}$	$T / \prod_{i=1}^{c-1} d_{\eta(i)} = d_{\eta(c)}$	$1 / \prod_{i=1}^{c-1} d_{\eta(i)}$
c	$d_{\eta(c)}$	1	$1/T$

3. Función de interpolación de una serie temporal y sus proyecciones.

Dada una serie observada $y_0, y_1, y_2, \dots, y_T$, ésta se convertirá en una señal en tiempo continuo, definida sobre el intervalo $[0, T]$, mediante la siguiente fórmula de interpolación

$$y(t) = \sum_{k=0}^T c_k \theta(t-k) \quad (5)$$

donde los coeficientes c_k deben ser tales que $y_t = y(t) \forall t = 0, 1, 2, \dots, T$.

Llamando

$$\mathbf{X} = \begin{pmatrix} \theta(0) & \theta(-1) & \dots & \theta(-T) \\ \theta(1) & \theta(0) & \dots & \theta(1-T) \\ \dots & \dots & \dots & \dots \\ \theta(T) & \theta(T-1) & \dots & \theta(0) \end{pmatrix} \quad \mathbf{Y} = \begin{pmatrix} y_0 \\ y_1 \\ \dots \\ y_T \end{pmatrix} \quad \mathbf{c} = \begin{pmatrix} c_0 \\ c_1 \\ \dots \\ c_T \end{pmatrix}$$

no es difícil comprobar que la expresión (5) queda determinada mediante el vector de coeficientes.

$$\mathbf{c} = \mathbf{X}^{-1} \mathbf{Y}$$

Dada la forma (5) para la función interpolante $y(t)$, es claro que ésta pertenece a un subespacio V'_0 de dimensión finita $m_0 + 1 = T + 1$ del subespacio V_0 del MRA de Shannon. Si representamos mediante $y_j(t)$ a la proyección ortogonal de $y(t)$ sobre el espacio V_j , entonces, teniendo en cuenta (ii), se verifica que $V_0 \subset V_j \forall j < 0$ y, consecuentemente, $y_j(t) \equiv y(t) \forall j < 0$. Solo las aproximaciones sobre V_j con $j > 0$ son de interés para las funciones de V_0 . Estas aproximaciones son de la forma

$$y_j(t) = \sum_{k \in \mathbb{Z}} a_{j,k} \theta_{j,k}(t) \quad (6)$$

donde

$$a_{j,i} = \int_{-\infty}^{+\infty} \theta_{j,i}(t) y(t) dt = \sum_{k=0}^T c_k \int_{-\infty}^{+\infty} \theta_{j,i}(t) \theta_{1,k}(t) dt \quad \forall i \in \mathbb{Z}$$

Por otra parte, dado $m_c = 1$ con c definido en (3), se tiene que V'_c está engendrado por dos funciones $\theta_{c,0}$ y $\theta_{c,1}$; la segunda, es trasladada de θ a $t = T$, y ambas dilatadas de θ mediante el factor de escala T . Para toda $j > c$, se tiene un valor de escala superior a la amplitud del intervalo $[0, T]$ y, por tanto, un nivel de resolución muy bajo. En este caso las aproximaciones de $y(t)$ son prácticamente coincidentes sobre el intervalo de observación $[0, T]$ y, por tanto, carecen de interés. Esto reduce el marco de análisis a la sucesión finita de subespacios

$$V_0 \supset V_1 \supset \dots \supset V_c$$

siendo c , definido en (3), la suma de las multiplicidades de los divisores primos de T

4. Transformadas de Fourier de las proyecciones.

Se representará mediante $F_g(\omega)$ la transformada de Fourier de la función g . Es bien conocido que si $\theta(t) = \text{sen } \pi t / \pi t$, entonces, Pollock (1999)

$$F_\theta(\omega) = I_{[-\pi, \pi]}(\omega) = \begin{cases} 1 & \text{si } \omega \in [-\pi, \pi] \\ 0 & \text{si } \omega \notin [-\pi, \pi] \end{cases}$$

También es perfectamente conocido el efecto que tienen las traslaciones y dilataciones de una función sobre su transformada de Fourier. Más concretamente, mediante un simple cambio de variable, teniendo en cuenta (1) se obtiene

$$F_{\theta_{j,k}}(\omega) = \int_{-\infty}^{+\infty} \theta_{j,k}(t) e^{-i\omega t} dt = \frac{1}{\sqrt{s_j}} e^{\frac{-ik\omega}{s_j}} F_\theta\left(\frac{\omega}{s_j}\right) = \frac{1}{\sqrt{s_j}} e^{\frac{-ik\omega}{s_j}} I_{[-\pi s_j, \pi s_j]}(\omega) \quad (7)$$

Este resultado permite calcular

$$F_{y_j}(\omega) = \left\{ \frac{1}{\sqrt{s_j}} \sum_{k \in \mathbb{Z}} a_{j,k} e^{\frac{-ik\omega}{s_j}} \right\} I_{[-\pi s_j, \pi s_j]}(\omega) \quad (8)$$

La expresión (8) muestra que la transformada de Fourier de la proyección ortogonal $y_j(t)$ (de la señal $y(t)$ sobre V_j) es el producto de una función periódica de periodo $2\pi/s_j$ y la función indicador del intervalo $[-\pi s_j, \pi s_j]$. Esto es el resultado de

aplicar un filtro de paso de bajas frecuencias (Low-Pass Filter) a la señal original, siendo la frecuencia de corte del filtro

$$\omega_j = \pi s_j = \frac{\pi m_j}{T} \quad (9)$$

que se corresponde con un periodo

$$T_j = \frac{2\pi}{\omega_j} = \frac{2}{s_j} = \frac{2T}{m_j} \quad (10)$$

Completando la Tabla 1 con (9) y (10) se obtiene la Tabla 2.

En ocasiones, solo interesan parte de las bandas de frecuencias que se muestran en la Tabla 2. En estos casos pueden omitirse filas de la tabla manteniendo solo las que definen las bandas de interés. En otras ocasiones, las bandas de frecuencia resultantes con las distintas permutaciones del vector \mathbf{b} no proporcionan una imagen óptima de la banda o bandas de frecuencia de mayor interés. En estos casos puede recurrirse a sucesiones de niveles de resolución no compatibles con la amplitud del intervalo $[0, T]$ siendo conscientes de que ya la señal interpolada deja de pertenecer al espacio V_0 .

Tabla 2

j	$d_{\eta(j)}$	m_j	$s_j = m_j/T$	$\omega_j = \pi s_j$	$T_j = \frac{2\pi}{\omega_j} = \frac{2}{s_j}$
0	-----	T	1	π	2
1	$d_{\eta(1)}$	$T/d_{\eta(1)} = \prod_{i=2}^c d_{\eta(i)}$	$1/d_{\eta(1)}$	$\frac{\pi}{d_{\eta(1)}}$	$2d_{\eta(1)}$
..
j	$d_{\eta(j)}$	$T/\prod_{i=1}^j d_{\eta(i)} = \prod_{i=j+1}^c d_{\eta(i)}$	$1/\prod_{i=1}^j d_{\eta(i)}$	$\pi/\prod_{i=1}^j d_{\eta(i)}$	$2\prod_{i=1}^j d_{\eta(i)}$
..
$c-1$	$d_{\eta(c-1)}$	$T/\prod_{i=1}^{c-1} d_{\eta(i)} = d_{\eta(c)}$	$1/\prod_{i=1}^{c-1} d_{\eta(i)}$	$\pi/\prod_{i=1}^{c-1} d_{\eta(i)}$	$2\prod_{i=1}^{c-1} d_{\eta(i)}$
c	$d_{\eta(c)}$	1	$1/T$	π/T	$2T$

5. Filtros lineales variables en el tiempo.

Un filtro variable en el tiempo, Hlawatsch et al (2000), Bachman y Narici (1981), Kolmogorov y Fomin (1975), está definido mediante una función impulso-respuesta o núcleo $h(t, t')$. El resultado de filtrar una señal $f(t)$ perteneciente al espacio de Hilbert de las funciones de cuadrado integrable $L_2(R)$ es otra función del mismo espacio que se define como

$$g(t) = \int_{-\infty}^{+\infty} h(t, t') f(t') dt' \quad (11)$$

Esto es, un operador lineal P continuo definido sobre un espacio de Hilbert. El operador adjunto P^* está definido mediante la función núcleo

$$h^*(t, t') = h(t', t).$$

Dado $\theta(t) = \text{sen } \pi t / \pi t$, si se define la función núcleo

$$h_j(t, t') = \sum_{k \in Z} \theta_{j,k}(t) \theta_{j,k}(t') \quad (12)$$

se obtienen operadores P_j que evidentemente son autoadjuntos. Es decir

$$\forall t, t' \in R \quad h_j(t, t') = h_j^*(t, t')$$

y, por tanto

$$\forall g \in L_2(R) \quad P_j(g) = P_j^*(g)$$

5.1 Proposición

Se verifica que $P_j(L_2(R)) \subset V_j$

En efecto, dada una función arbitraria g perteneciente a $L_2(R)$, sea

$$\begin{aligned} g_j(t) &= P_j(g)(t) = \int_{-\infty}^{+\infty} h(t, t') g(t') dt' = \int_{-\infty}^{+\infty} \sum_{k \in Z} \theta_{j,k}(t) \theta_{j,k}(t') g(t') dt' = \\ &= \sum_{k \in Z} \theta_{j,k}(t) \int_{-\infty}^{+\infty} \theta_{j,k}(t') g(t') dt' = \sum_{k \in Z} a_{j,k} \theta_{j,k}(t) \end{aligned} \quad (13)$$

donde $a_{j,k} = \int_{-\infty}^{+\infty} \theta_{j,k}(t') g(t') dt'$.

Según el resultado obtenido en (13) es evidente que $g_j = P_j(g) \in V_j$ con lo cual la proposición es cierta.

5.2 Proposición

Se verifica que P_j es un operador idempotente. En efecto, para todo $i \in Z$

$$P_j(P_{j,i})(t) = \int_{-\infty}^{+\infty} \sum_{k \in Z} \theta_{j,k}(t) \theta_{j,k}(t') \theta_{j,i}(t') dt = \sum_{k \in Z} \theta_{j,k}(t) \int_{-\infty}^{+\infty} \theta_{j,k}(t') \theta_{j,i}(t') dt = \theta_{j,i}(t)$$

ya que al ser $\{\theta_{j,k}\}_{k \in Z}$ base ortonormal de V_j se verifica que

$$\int_{-\infty}^{+\infty} \theta_{j,k}(t') \theta_{j,i}(t') dt = \delta_{i,k} = \begin{cases} 1 & \text{si } i = k \\ 0 & \text{si } i \neq k \end{cases}$$

Por tanto, $\forall g \in L_2(R)$

$$P_j^2(g) = P_j(P_j(g)) = P_j\left(\sum_{k \in Z} a_{j,k} \theta_{j,k}\right) = \sum_{k \in Z} a_{j,k} P_j(\theta_{j,k}) = \sum_{k \in Z} a_{j,k} \theta_{j,k} = P_j(g)$$

Luego $P_j^2 = P_j$

La restricción de este operador al espacio V_0' es un operador autoadjunto definido sobre un espacio de dimensión finita e idempotente. Por tanto, es la proyección ortogonal del espacio V_0' sobre el espacio V_j (Bachman y Narici (1981)).

Es evidente que si $y(t) \in V_0'$ es una señal en tiempo continuo que se obtiene mediante interpolación de una serie observada en el intervalo $[0, T]$, tal y como se ha definido en (10), entonces

$$y_j(t) = P_j(y)(t)$$

es la misma aproximación de $y(t)$ definida en (6)

Por tratarse de proyecciones ortogonales, los operadores P_j constituyen filtros de tipo “Low-Pass” con frecuencias de corte ω_j especificadas en (9). Dado que $V_j \supset V_{j+1}$,

los operadores $P_j - P_{j+1}$ son proyecciones sobre el complemento ortogonal W_{j+1} de V_{j+1} en V_j y definen filtros de tipo “Band-Pass” cuya banda de frecuencia es el intervalo⁴

$$\left(\omega_{j+1}, \omega_j\right] = \left(\pi s_{j+1}, \pi s_j\right] = \left(\frac{\pi m_{j+1}}{T}, \frac{\pi m_j}{T}\right] \quad (14)$$

5. Descomposición Wavelets de la señal $y(t)$.

En general el análisis MRA es un buen instrumento para la construcción de Wavelets, Wojtaszyk (1997).

El análisis Wavelet es bien sencillo una vez que se obtienen las distintas aproximaciones en el seno del MRA:

Se llama detalle $d_j(t)$ a la diferencia entre las dos aproximaciones sucesivas siguiente

$$d_j = y_j - y_{j+1} = (P_j - P_{j+1})(y). \quad (15)$$

Si se tiene en cuenta que $y(t) \equiv y_0(t)$, es obvio que

$$y(t) = y_c(t) + \sum_{j=c-1}^0 d_j(t) \quad (16)$$

La expresión (14) es una descomposición en suma de funciones que son ortogonales entre sí, Wojtaszyk (1997). Mas concretamente, la señal original se descompone en suma de una aproximación conteniendo las bajas frecuencias y c detalles que recogen la aportación a la señal, a lo largo del tiempo, en las distintas bandas de frecuencia definidas mediante (14) para $j = c-1, c-2, \dots, 1$. Es decir, aquellas bandas de frecuencias cuyos extremos son los valores recogidos en la cuarta columna de la Tabal 2.

6. Ejemplos y aplicación.

Ejemplo 1: Se considera una serie simulada de longitud $T = 120$, mediante el modelo

$$y_t = 6,457310029 \cos(0,314t - \pi/2) + u_t$$

⁴ Se ha omitido la parte de frecuencias negativas

con $u_t \rightarrow N(0, 2^2)$. Alternativamente, este modelo, puede escribirse como

$$y_t = f(t) + u_t$$

donde⁵

$$f(t) = 1,90221144 f(t-1) - f(t-2)$$

y donde

$$f(0) = 0 \quad y \quad f(1) = 2$$

Se trata de una señal periódica de frecuencia 0,314 y periodo $2\pi/0,314 \approx 20$, con una cierta dosis de ruido blanco proporcionada por la perturbación u_t .

Dado el vector $\mathbf{b} = (2, 2, 2, 3, 5)$ de divisores primos de $T = 120$, utilizando la permutación $\eta(\mathbf{b}) = (2, 2, 5, 3, 2)$, se obtiene la sucesión de valores de m_j según se especifica en (4), la de correspondientes valores s_j según (2), los ω_j definidos en (9) que son los límites de las bandas de frecuencias especificadas en (14) y, por último, pueden calcularse los periodos de corte correspondientes a tales frecuencias usando (10). Todos ellos quedan recogidos en Tabla 3.

Tabla 3.

j	$d_{\eta(j)}$	m_j	s_j	ω_j	T_j
0	-----	120	1	3,1416	2
1	2	60	1 / 2	1,5708	4
2	2	30	1 / 4	0,7854	8
3	5	6	1 / 20	0,1571	40
4	3	2	1 / 60	0,0524	120
5	2	1	1 / 120	0,0262	240

En Fig. 1 se muestra la gráfica de la serie simulada junto con la aproximación obtenida al nivel de resolución más bajo $s_5 = 1/120$. Resulta imperceptible en el gráfico por ser casi coincidente con el eje de abscisas. En Fig. 2 se muestran los detalles que se obtienen según (15), y que son el resultado de filtrar la serie original mediante filtros de

⁵ $1,90221144 = 2 \cos(0,314)$

tipo paso-banda cuyas respectivas bandas de frecuencia (sus extremos) se recogen en la cuarta columna de la Tabla 3.

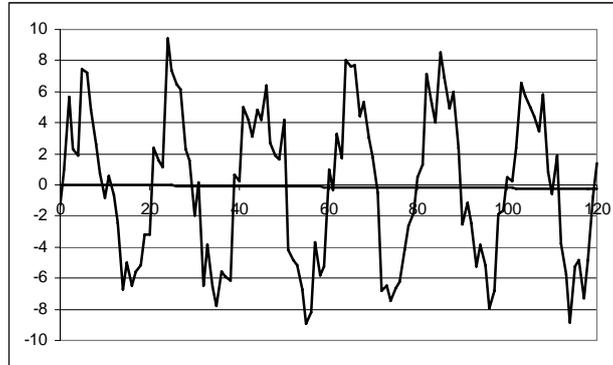


Fig. 1

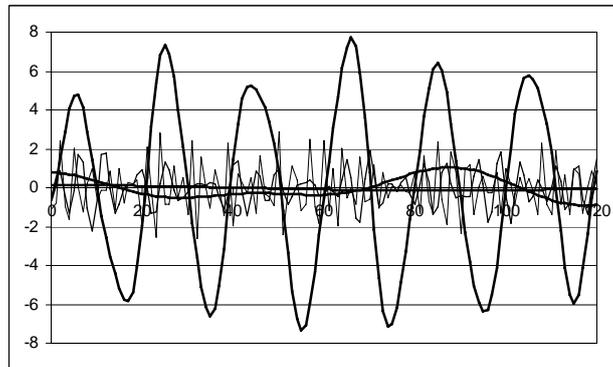


Fig. 2

La componente de mayor amplitud se corresponde con la frecuencia 0,314 y periodo

$$\frac{2\pi}{0,314} \approx 20 \in [8, 40)$$

En Fig.3 puede verse el análisis realizado para la misma señal despojada de ruido; se incluye en el gráfico en trazo de color gris. Puede apreciarse un efecto de distorsión en los bordes, debido a que se está filtrando una señal de la cual sólo se dispone información en un intervalo de tiempo finito $[0, T]$. Dicho “efecto borde” es también apreciable en Fig. 2. Las diferencias que se aprecian entre Fig. 2 y 3, en las distintas bandas de frecuencia, pueden ser atribuidas al ruido que se ha incorporado en la señal que se analiza en Fig. 2

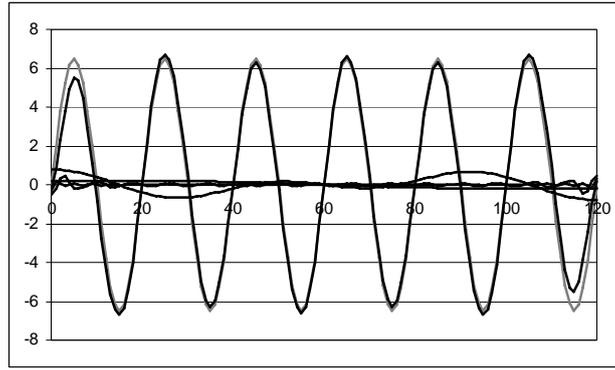


Fig. 3

Ejemplo 2: Se considera la serie simulada mediante el modelo AR(2) siguiente

$$y_t = 1,90221144y_{t-1} - y_{t-2} + u_t$$

donde $u_t \rightarrow N(0, 2^2)$ e incorreladas, y donde

$$y_0 = 0 + u_0 \quad y_1 = 2 + u_1$$

La pauta cíclica es la misma que en el caso anterior pero aquí las perturbaciones se involucran dentro del proceso autorregresivo. Los valores simulados para las perturbaciones que se incorporan al modelo son exactamente los mismos que se usaron en 6.1. En Fig. 4 se muestra la serie simulada y la aproximación a frecuencias más bajas. En Fig 5 se encuentran los detalles que se corresponden con los mostrados en la Fig. 2

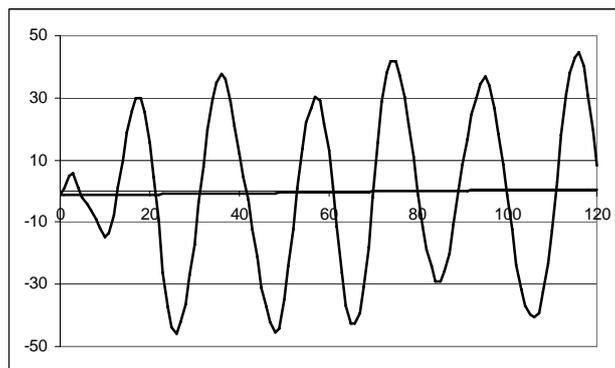


Fig. 4

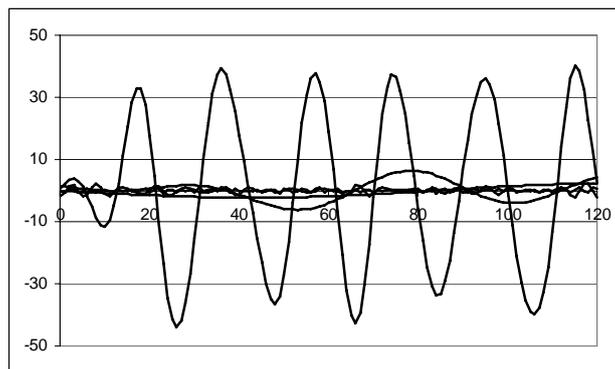


Fig. 5

Aplicación: Consideremos la serie de los porcentajes de variación trimestral del PIB en España durante el periodo 1980.4 a 2004.4 (Datos INE Contabilidad Trimestral Nacional Base 1995). Se tiene una serie de longitud de 97 datos en el intervalo $[0, 96]$, donde el origen de tiempos ($t = 0$) se ha desplazado al momento de observación 1980.4

El vector de divisores primos de T es $\mathbf{b} = (2, 2, 2, 2, 2, 3)$ utilizando la permutación $\tau(\mathbf{b}) = (3, 2, 2, 2, 2, 2)$ se obtiene la siguiente Tabla 4.

Tabla 4.

j	$d_{\eta(j)}$	m_j	s_j	ω_j	T_j
0	-----	96	1	3,1416	2
1	3	32	1/3	1,0472	6
2	2	16	1/6	0,5236	12
3	2	8	1/12	0,2618	24
4	2	4	1/24	0,1309	48
5	2	2	1/48	0,0654	96
6	2	1	1/96	0,0327	192

En Fig. 6 se observa la señal original y la aproximación a mínima resolución. En Fig. 7 pueden verse los distintos detalles según las bandas de frecuencia especificadas en Tabla 2. Se observa que la componente de mayor importancia (amplitud) corresponde a la banda de frecuencia $[1,0472; 3,1416]$ y, en términos de periodo, a la banda $[2; 6)$, que contiene el ciclo anual (4 trimestres) y en el que se percibe un primer armónico, de frecuencia doble y periodo la mitad (2 trimestres)

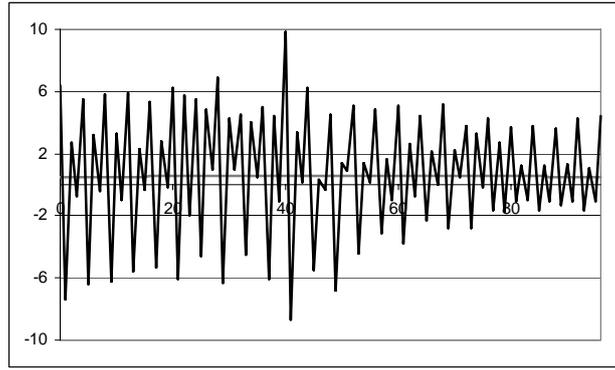


Fig. 6

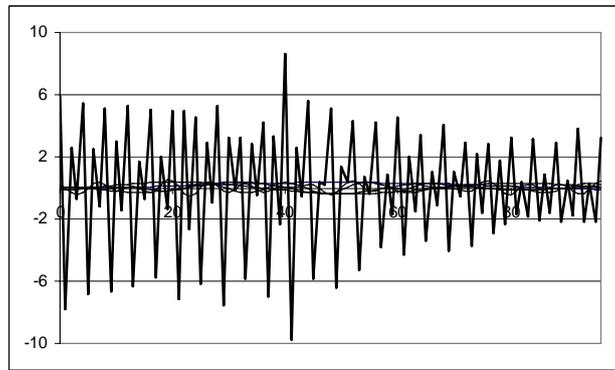


Fig. 7

El gran parecido entre los dos gráficos mostrados en Fig. 6 y Fig. 7, corrobora que las frecuencias realmente significativas en esta señal corresponden al ciclo anual y semianual, ambos recogidos en la banda [2, 6). Al objeto de separar estas señales se procede con la secuencia de valores m_j reflejados en la Tabla 5

Tabla 5

j	m_j	s_j	ω_j	T_j
0	160	4/3	4,1888	1,5
1	80	2/3	2,0944	3
2	40	1/3	1,0472	6

Este modo de seleccionar los m_j presenta el inconveniente de que la señal $y(t)$ no pertenece al espacio V_0 sino que está en un espacio de resolución ligeramente inferior a la de V_0 . Considerando que la proyección $P_0(y) = y_0$ es una buena aproximación de y , puede mantenerse válida (al menos en forma aproximada) la descomposición formulada en (16) y, sobre todo, que las dos bandas de frecuencias, cuyos extremos son los valores de la cuarta columna de la Tabla 5, separan las frecuencias $2\pi/4 = 1,5708$ y

$2\pi/2 = 3,1416$. Estas se corresponden con un periodo anual (4 trimestres) y otro semianual (2 trimestres), respectivamente, y aparecen agregadas conformando una sola onda en Fig. 7. Pueden apreciarse las dos componentes por separado en Fig. 8 y Fig. 9.

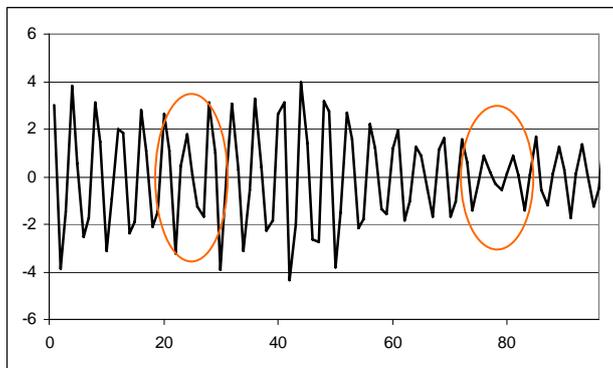


Fig. 8

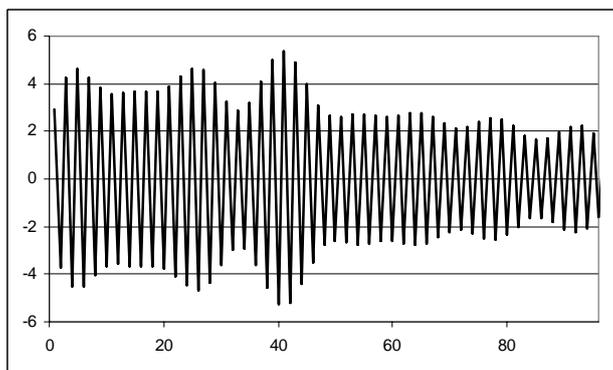


Fig. 9

En Fig. 9 puede verse la naturaleza pulsante del ciclo semianual, y una disminución de la varianza a partir de $t = 49$ correspondiente a 1993.1. Algo análogo se observa en el ciclo anual, aunque la disminución de la varianza se retrasa a $t = 58$ correspondiente a 1995.2. Las dos pequeñas deformaciones que se observan en la onda en $t=27$ y $t=79$ se corresponden, respectivamente, con 1987.3 y 2000.3. Posiblemente puedan considerarse como un efecto de los “shocks” producidos sobre la producción por la crisis de 1987 y el comienzo de la implantación del euro, respectivamente.

8 Comparación con otros filtros.

En Iacobucci y Noullez (2004) puede encontrarse una revisión crítica de los filtros de Hodrick-Prescott, Baxter-King, y Christiano-Fitzgerald y los filtros de tipo “windowed”, en comparación con un filtro propio propuesto por los autores en el

trabajo. Para la comparación se utiliza como banco de pruebas la serie generada por el modelo

$$y_t = \text{Sen}\left(\frac{2\pi t}{24}\right) - 0,15 \text{Sen}\left(\frac{2\pi t}{6}\right)$$

el mejor resultado allí obtenido intentando separar la componente de periodo 24 de la serie y_t , se muestra en Fig. 10. En Fig. 11 puede verse el resultado de separar las dos componentes de y_t , mediante la metodología aquí descrita, utilizando la secuencia de niveles de resolución y bandas de frecuencia mostrados en la Tabla 6.

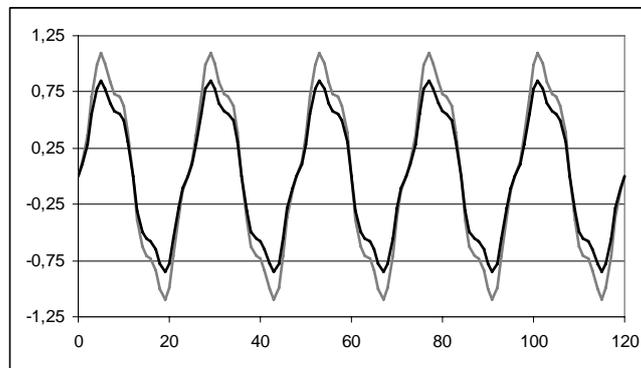


Fig. 10

Tabla 6.

j	m_j	s_j	ω_j	T_j
0	120	1	3,1416	2
1	30	1/4	0,7854	8
2	1	1/120	0,0262	240

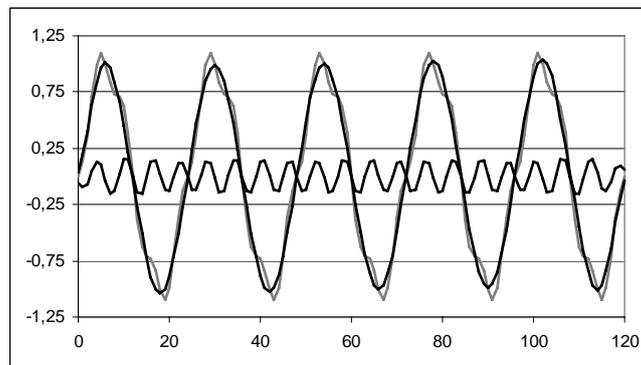


Fig. 11

9. Conclusiones.

Una serie temporal puede ser convertida en una señal en tiempo continuo, utilizando una función de interpolación perteneciente al espacio V_0 de un Análisis Multirresolución (MRA). Como tal, puede ser analizada dentro de dicha estructura MRA y de su correspondiente Análisis Wavelet (WA), obteniéndose mejores resultados que si es tratada como una señal digital (discreta), especialmente en el caso de series de datos económicos, que suelen ser de corta longitud. En este trabajo se ha utilizado el MRA y WA de Shannon para tal fin.

La pertenencia a V_0 de la señal interpolada se consigue mediante la adecuada selección de valores de escala, compatible con el intervalo de observación $[0, T]$. Como consecuencia de tal selección, toda aproximación a nivel de resolución inferior a s_0 coincide con la propia señal. Por otra parte, niveles de resolución muy bajos, tales que $T < 1/s_j$, generan aproximaciones prácticamente iguales sobre el intervalo $[0, T]$ y detalles nulos, a efectos prácticos, sobre el citado intervalo y, por tanto, carentes de interés. Esto permite una descomposición exacta de la señal que se circunscribe a una sucesión finita de subespacios V_j (epígrafes 2 y 3).

Utilizando el MRA de Shannon originado por la función sinc que es una función de escalado (“scaling function”) ortogonal, queda patente que las aproximaciones se obtienen mediante filtros variables en el tiempo cuya función de impulso es una función núcleo que se construye a partir de la propia función sinc.

Esta metodología de análisis es aplicable a series estacionarias y no estacionarias, aunque hay que considerar el efecto borde que se produce como consecuencia de ser finito el intervalo de observación. Dicho efecto es más patente en las aproximaciones, y detalles, de más baja resolución, sobre todo, en series que muestran una tendencia no constante y, por tanto, unos valores en los extremos ($t = 0$ y $t = T$) muy diferentes entre sí, y muy diferentes del valor medio de la serie en el intervalo de observación.

Bibliografía.

- Bachman G. y Narici L. (1981): *Análisis Funcional*, Editorial Tecnos S.A., Madrid.
Hamming, R. W. (1998): *Digital Filters*, Dover Publications, Inc., New York.

- Hernández y Weiss(1996): *A First course on Wavelets*, CRC Press London.
- Iacobucci A. and Noullez A. (2004): “A Frequency Selective Filter for Short-Length Time Series”, *Document de Travail n° 5 Mai 2004*, <http://www.ofce.sciences-po.fr/publications/lettre2004.htm>.
- Kolmogorov A.N. y Fomín S.V. (1975): *Elementos de la teoría de funciones y del análisis funcional*. Editorial Mir Moscu.
- Mallat S. (1987): “An efficient image representation for multiescale analysis”, in *Proc. Of Machine Vision Conference*, Lake Tahoe.
- Mallat S. (1989): “Multiresolution approximation and wavelet orthonormal bases of $L^2(R)$ ”, *Trans. Math. Soc.*, 315 pp. 69-87.
- Mallat S.(1989): “A Theory for multiresolution signal decomposition: the wavelet representation”, *IEEE Trans. Patt. Anal. And Mach. Intell.*, 11(7) pp. 674-693
- Mallat S. (1999): *A wavelet tour of signal processing*. Academic Press, New York
- Meyer Y.(1988) : “Constructions de bases orthonormées d'ondelettes ”. *Rev. Mat. Iberoamericana*, 4(1):31-39.
- Meyer Y. (1989a): “Orthonormal wavelets”, In *IXth International Congress on Mathematical Physics (Swansea, 1988)*, pp. 38-47, Hilger, Bristol.
- Meyer Y. (1989b): “Wavelets and operators”. In *Analysis at Urbana, Vol. I (Urbana, IL, 1986--1987)*, volume 137 of *London Math. Soc. Lecture Note Ser.*, pp. 256-365. Cambridge Univ. Press, Cambridge.
- Meyer Y. (1989c) : “Ondelettes, filtres miroirs en quadrature et traitement numérique de l'image ”, *Gaz. Math.*, (40):31-42.
- Meyer Y.(1993), “Wavelets and operators”. In *Different perspectives on wavelets (San Antonio, TX, 1993)*, volume 47 of *Proc. Sympos. Appl. Math.*, pages 35-58. Amer. Math. Soc., Providence, RI.
- Oppenheim, A. V. and Schafer, R. V. (1999). *Discrete-Time Signal Processing*. Prentice-Hall, New Jersey.
- Percival D. B. y Walden A. T (2000). *Wavelets Methods for Time Series Análisis*. Cambridge University Press, Cambridge U.K.
- Priestley, M. B. (1981). *Spectral Analysis and Time Series*. Academic Press, Orlando, Florida.
- Pollock D.S.G. (1999): *A Handbook of time-series analysis, signal processing and dynamics*, Academic Press, London
- Pollock D.S.G. y Lo Casio I. (2005): Orthogonality Conditions for Non-Dyadic Wavelet Analysis, Working Paper, <http://www.gmw.ac.uk/~ugte133/PAPERS/NONDWAVE.PDF>.
- Wojtaszczyk P. (1997): *A Mathematical introduction to wavelets*, Cambridge University Press. Cambridge.

ANÁLISIS DE LA IMAGEN DE UNA UNIVERSIDAD PÚBLICA Y DE SU EFECTO SOBRE LA INTENCIÓN DE COMPRA DE ARTÍCULOS CON SU LOGOTIPO

ANA MARIA MARTIN ARROYUELOS

e-mail: ana.martin@ehu.es

Departamento de Economía Aplicada III

JUAN JOSÉ MIJANGOS DEL CAMPO

e-mail: juanjo.mijangos@ehu.es

Departamento de Economía Industrial

UNIVERSIDAD DEL PAÍS VASCO (UPV/EHU)¹

Área temática: Métodos Cuantitativos.

Resumen

Este trabajo surge en un contexto en el que la Universidad del País Vasco (UPV/EHU) se plantea la apertura de una tienda para comercializar productos con su logotipo con el doble objetivo de conseguir por un lado una rentabilidad económica y por otro una mayor proximidad con la sociedad en general. Con el objeto de analizar la viabilidad de dicha tienda se realizaron dos encuestas: una iba dirigida al personal vinculado a la propia universidad y otra al público en general. Este trabajo se centra en el análisis de la segunda encuesta. En primer lugar se realiza un diagnóstico acerca de la imagen que el público en general tiene sobre la UPV/EHU. Dentro de este análisis se tratará de establecer un indicador de imagen corporativa de la universidad mediante la utilización de un análisis factorial. Asimismo, se determina cuáles son los factores que contribuyen a que un determinado individuo incremente su intención de compra de artículos con el logotipo de la UPV/EHU mediante la utilización de un modelo de regresión logit. Se incidirá especialmente en el efecto que la imagen corporativa de la institución tiene sobre dicha intención de compra.

Palabras clave: Análisis Factorial, Modelo de Regresión Logit, Imagen Corporativa.

Abstract

This work is placed in a context in which the University of the Basque Country is planning to open a corporate shop to sell products featuring the University logo. One of the aims of the University is to obtain an economic profitability but the most important objective is to get closer to the society in order to improve its corporate image. To analyze the viability of this gift shop, two surveys have been developed. The first one has been addressed to university-related people and the second one to the general public. In this work we will analyze the results of this last survey. First of all we have made a diagnosis of the corporate image that the general public have about the University of the Basque Country. In this part we have tried to propose an index of the corporate image with the application of a factor analysis. Furthermore we have used a logit regression model in order to determine the factors that increase individual probability of buying corporate products with the University logo. We will pay special attention to the effect of the institutional image on this probability.

Key words: Factor Analysis, Logit Regression Model, Corporate Image.

¹ Este trabajo ha sido realizado con apoyo financiero del Grupo de Investigación Consolidado DEC de la Universidad del País Vasco, código UPV/EHU GIU06/53.

1. Introducción

El origen de este trabajo se sitúa en un contexto en el que la UPV/EHU, dentro de un proyecto más amplio, que tiene como objeto mejorar su imagen como institución pública, se plantea la creación de una tienda dedicada a la comercialización de artículos con su logotipo.

Dado que actualmente el comportamiento de los consumidores cuando se enfrentan a la decisión de comprar o no comprar un determinado producto no depende tanto del producto en sí mismo sino de las percepciones que se tengan sobre el mismo, se considera que la imagen corporativa de la UPV/EHU entre el público en general tendrá un efecto decisivo a tener en cuenta al analizar la viabilidad de una tienda de esta naturaleza, ya que de ello dependerá que los potenciales compradores estén dispuestos a adquirir artículos con el logotipo que la identifica.

Por lo tanto, si se quiere realizar un diagnóstico sobre la capacidad que tiene la UPV/EHU de comercializar artículos con su logotipo entre el público en general, un primer paso será realizar un diagnóstico acerca de la imagen que dicha institución tiene en la sociedad que la rodea.

Con este motivo se han realizado dos encuestas. La primera encuesta se dirigió a personas que actualmente están vinculadas con la UPV/EHU: estudiantes, profesores y personal de administración y servicios. La segunda se realizó a pie de calle, y estaba dirigida a personas no vinculadas actualmente con la institución.

En este trabajo se realizará el estudio de la segunda de las dos encuestas mencionadas. Con los datos obtenidos en la misma se han realizado los análisis que pasamos a detallar a continuación.

En primer lugar se ha realizado un estudio acerca de cuál es la imagen que la Universidad del País Vasco tiene entre el público en general. Se han determinado cuáles son las dimensiones que se pueden encuadrar en dicha imagen, mediante la utilización de un análisis factorial, con el fin de elaborar indicadores de dicha imagen; además, se han encontrado algunas pautas de comportamiento respecto de dichos indicadores.

En segundo lugar, se ha llegado a determinar qué influencia tienen las diferentes características de los individuos sobre la propensión de compra de artículos con el logotipo de la UPV/EHU. Dado que la variable que se pretende explicar es categórica (compra o no compra), se ha realizado una regresión con variable dependiente discreta, en concreto un modelo logit.

El desarrollo del trabajo es el que se presenta a continuación. En primer lugar se exponen los detalles del diseño de la encuesta, tanto en lo que se refiere al método de muestreo utilizado como al diseño del cuestionario. Posteriormente se mostrarán los resultados obtenidos en el análisis realizado en referencia al estudio de imagen corporativa de la UPV/EHU. En la siguiente sección se exponen los resultados obtenidos acerca de la estimación de la probabilidad de compra de artículos con el logotipo mediante la utilización de una regresión con variable dependiente discreta. Finalmente se expondrán las conclusiones generales del trabajo.

2. Diseño de la encuesta

La encuesta se realizó mediante muestreo por cuotas cruzadas para género y edad (de 14 a 29, de 30 a 44 y más de 44 años). La información se recogió mediante encuesta realizada a pie de calle en las tres capitales vascas. En concreto, en Bilbao se decidió tomar una muestra de 200 individuos, seleccionándose 14 puntos para la realización de las encuestas. En San Sebastián se realizaron 160 encuestas con 12 puntos de recogida y en Vitoria se realizaron 120 encuestas en 9 puntos de recogida. La recogida de esta información tuvo lugar entre los meses de noviembre y diciembre de 2004.

Dado que el muestreo realizado es de tipo no probabilística, somos conscientes de que no queda garantizada la representatividad de la población en estudio. Ahora bien, dado el tema objeto del estudio, que no es delicado o comprometido, consideramos que las respuestas dadas por los individuos que han cumplimentado las encuestas no debieran diferir en exceso de las del resto de población que querríamos representar, con lo que el sesgo de los resultados no será considerable. Debido a ello trataremos la información de manera similar a la correspondiente al muestreo probabilística, aun siendo conscientes de sus limitaciones. De hecho, se observa bastante coherencia entre nuestras

conclusiones y las obtenidas a partir de la otra encuesta a la que hemos hecho referencia anteriormente, que estaba dirigida al personal vinculado con la Universidad, y que sí se había realizado mediante la utilización de técnicas probabilísticas (Fernández y otros, 2005).

A continuación mostramos la estructura de la encuesta. Las preguntas que la componen están agrupadas en torno a tres partes².

En la primera se incluyeron preguntas sobre la percepción que los encuestados tenían sobre la Universidad del País Vasco. En concreto se recogían preguntas acerca del grado de conocimiento que el encuestado tiene sobre la UPV/EHU: si recuerda haber visto su logotipo y sabe a qué institución pertenece, o sobre el grado de conocimiento entre la sociedad tanto de la Universidad del País Vasco, como de otras Universidades del entorno. Posteriormente, y dentro de este primer grupo de preguntas se recogía información sobre la opinión del encuestado en referencia a determinados aspectos de la actividad que la UPV/EHU desarrolla. Dado que en el siguiente apartado nos centraremos en estas últimas preguntas, expondremos allí cuáles son los conceptos a los que nos referimos.

A continuación, la segunda parte de la encuesta incluye preguntas que se refieren a artículos con el logotipo de la UPV/EHU. La primera pregunta era de tipo dicotómico (Sí/No) y se refería a si el individuo estaría interesado en comprar un producto con el logotipo de la UPV/EHU. Esta pregunta se realiza antes de mostrar ningún producto corporativo, por lo que hablaremos de la intención a priori del individuo de comprar artículos con el logotipo de la UPV/EHU. Los grupos de productos se muestran posteriormente, y se pide valorar con una escala de “muy poco probable”, “poco probable”, “probable” o “muy probable” la posibilidad de adquirir algún artículo del grupo de que se trate. Después de esto, se pide que se valore, en una escala de 1 a 7, la cuantía en que cada una de las características que se citan a continuación debieran aparecer en un artículo con el logotipo de la UPV/EHU: Original, Audaz, Práctico, Tradicional, Artístico, Elegante, Serio y Moderno.

² La encuesta completa aparece en (Fernández y otros, 2005).

Donde F_1, F_2, \dots, F_m son los factores comunes, e_1, e_2, \dots, e_n son los factores únicos y l_{jh} es el peso del factor común h en la variable j . Debe tenerse en cuenta que ni los factores comunes ni los únicos son observables.

El objetivo del análisis factorial es la extracción de los factores comunes o latentes. Se pueden plantear diferentes métodos de extracción de tales factores. Nosotros hemos considerado el método de componentes principales, cuyo objetivo es explicar la mayor cantidad de varianza de las variables originales a través del menor número de factores. Al realizar dicha extracción, los valores propios asociados a cada factor nos indicará el porcentaje de varianza explicada por el mismo. Dichos valores propios tienen un orden decreciente, de tal modo que el primer factor será el que mayor varianza recoja, y por lo tanto al que mayor valor propio corresponda, y así sucesivamente para los demás factores.

Para determinar cuántos factores comunes vamos a considerar en el modelo existen diferentes criterios. De entre ellos, el que nosotros vamos a utilizar es el criterio del valor propio o regla de Kaiser, que consiste en retener aquellos factores cuyo valor propio asociado sea mayor o igual que 1, aunque como vamos a ver relajaremos ligeramente dicho criterio, y admitiremos un factor cuyo valor propio está muy próximo a 1, pero sin llegar a dicho valor, para conseguir una mejor interpretación del modelo.

3.1 Principales dimensiones de la imagen de la UPV/EHU

En este apartado pretendemos llegar a establecer un indicador de la imagen que el público en general tiene de la Universidad del País Vasco. Para ello se utilizarán las cuestiones de la encuesta que se refieren a la opinión de los individuos acerca de la actividad de la UPV/EHU.

Con tales cuestiones se pretende cubrir algunos de los aspectos que se considera pueden incluirse dentro del concepto más amplio de imagen corporativa de la institución. En concreto se pretendía cubrir por un lado aspectos referentes a la proyección exterior que la universidad tiene sobre la sociedad en general, y en segundo lugar aspectos relativos a la organización interna de la misma, en lo que se refiere tanto a sus obligaciones docentes y de investigación como a su dotación de recursos. Se trata de las siguientes cuestiones, que hay que valorar de 1 a 5:

Institución encerrada en sí misma	1 2 3 4 5	Institución abierta a la sociedad
No es de utilidad en la sociedad	1 2 3 4 5	Ofrece un servicio útil a la sociedad
No contribuye a la difusión de la cultura	1 2 3 4 5	Contribuye a la difusión de la cultura
No contribuye al progreso e innovación	1 2 3 4 5	Contribuye al progreso e innovación
Docencia: no es función fundamental	1 2 3 4 5	Docencia: es función fundamental
Investigación: no es una función fundamental	1 2 3 4 5	Investigación: es una función fundamental
Los recursos asignados a la UPV/EHU son escasos	1 2 3 4 5	Los recursos asignados a la PV/EHU son excesivos

Al objeto de determinar si tal como se han planteado las cuestiones existen dos dimensiones diferentes: por un lado la externa (cuatro primeros ítems), y por otro lado la interna (tres últimos ítems), se ha realizado un análisis factorial de donde se han obtenido los siguientes resultados. En la Tabla 1 aparecen los resultados referentes a los valores propios correspondientes a todos los factores que pueden deducirse del análisis y a las sumas de los porcentajes de varianzas recogido con los tres primeros factores considerados. Hemos decidido incluir el tercer factor, dado que su valor propio está muy próximo a uno, y si excluimos dicho factor, la última variable (recursos escasos frente a suficientes) no quedaría bien representada en ninguno de los factores, como veremos después.

Tabla 1

Componente	Autovalores iniciales			Sumas de las saturaciones al cuadrado de la extracción		
	Total	% de la varianza	% acumulado	Total	% de la varianza	% acumulado
1	2,679	38,272	38,272	2,679	38,272	38,272
2	1,285	18,358	56,630	1,285	18,358	56,630
3	,970	13,863	70,493	,970	13,863	70,493
4	,703	10,036	80,529			
5	,515	7,362	87,891			
6	,461	6,580	94,471			
7	,387	5,529	100,000			

Se ha realizado posteriormente una rotación Varimax, ya que con ello se facilita la interpretación de los factores y los resultados obtenidos sobre las correlaciones de las

variables con cada uno de los factores considerados aparecen recogidos en la Tabla 2. Se observa claramente cómo el primer factor guarda correlaciones altas con los cuatro primeros ítems considerados, esto es, con todos aquellos que podríamos considerar que recogen la proyección externa de la universidad hacia la sociedad.

Los dos siguientes ítems guardan una correlación muy alta con la segunda componente y podrían entenderse como la proyección interna de la universidad: en qué medida debe la universidad dedicarse a la investigación y a la docencia. Finalmente, la cuestión sobre la suficiencia de los recursos no quedaría bien reflejada en ninguno de esos dos factores, y tendríamos que ir al tercero para ver que sólo aparece bien reflejada esta variable. Por lo tanto, la primera dimensión, proyección externa de la universidad, sí queda diferenciada en un único factor, pero la segunda dimensión se separa a su vez en dos subfactores, es decir, no habría un único factor que representara el funcionamiento interno de la universidad.

Tabla 2

	Componente		
	1	2	3
Abierta	,752	-,004	,013
Util	,797	,121	-,089
DifCultu	,823	,132	,049
ProInnov	,609	,407	,189
Docencia	,221	,792	,001
Investig	,012	,885	-,091
Recursos	,019	-,062	,984

Cabría ahora plantearse si existen diferencias significativas en las puntuaciones obtenidas para los tres factores obtenidos de acuerdo con las características personales de los individuos. Para comprobarlo realizaremos un Análisis de Varianza (ANOVA) (Hair y otros, 1999). Este análisis nos permite estudiar el efecto que una variable independiente categórica (en nuestro caso puede ser la capital, el sexo del encuestado, ...), al que llamamos factor, tiene sobre una variable cuantitativa. En los casos en que de la realización de dicho análisis se deduzca que existen diferencias significativas para los diferentes valores del factor, nos debemos plantear cuáles son los grupos (valores del factor) que tiene más responsabilidad en dicha conclusión, es decir, entre qué grupos

aparecen las diferencias más relevantes. Para llegar a determinarlo se realizan pruebas de comparaciones múltiples. De entre las diferentes pruebas para realizar las comparaciones múltiples, hemos utilizado las que considerábamos más apropiadas para nuestro análisis dado que en los diferentes grupos que consideramos los tamaños muestrales son diferentes y en ocasiones no se admite la homogeneidad de varianzas (Hochberg y Tamhane, 1987), (Klockars y Sax, 1986). Las conclusiones obtenidas son las que aparecen a continuación.

Si nos fijamos en el factor que refleja la proyección externa de la universidad, se observan diferencias significativas para la capital en que se ha realizado la encuesta. En concreto, son los encuestados en San Sebastián los que presentan diferencias positivas significativas respecto de los encuestados en Bilbao, mientras que para los encuestados en Vitoria la media toma un valor intermedio, sin presentar diferencias significativas respecto a ninguno de los otros dos grupos.

También aparecen diferencias significativas en función de la opinión que el encuestado tiene acerca del grado de conocimiento de la UPV/EHU en la sociedad; en este caso, las diferencias son positivas en el caso en que el individuo cree que la UPV/EHU es muy conocida frente a todas las demás opiniones (muy poco, poco o conocida sin más).

En lo que se refiere a la intención de compra, vemos que también aparecen diferencias significativas; en concreto dicha diferencia significativa se da a favor de quienes sí están dispuestos a comprar artículos con el logotipo de la UPV/EHU. Parece lógico que esto sea así, dado que un individuo estará más dispuesto a adquirir artículos con el logotipo de una entidad cuanto más cercana sienta a dicha entidad.

La edad también aparece como una variable significativa. En concreto se ha procedido a recodificarla en dos grupos: los dos tramos más jóvenes irían en el primero y los dos de mayor edad en el segundo. La razón es que no se encuentran diferencias significativas entre los tramos superiores, por lo que se podrían agrupar, y entre los más jóvenes, el problema es que el grupo de menor edad recoge muy pocos individuos. Tras realizar esta recodificación se llega a que las diferencias son positivas para los individuos de mayor edad respecto de los más jóvenes.

Si consideramos ahora el segundo factor, se obtienen los siguientes resultados. Si nos fijamos en la capital aparecen diferencias significativas entre las tres capitales obteniéndose las mayores puntuaciones en San Sebastián, posteriormente en Bilbao y en último lugar en Vitoria.

En cuanto a la opinión que el encuestado tiene acerca de cómo de conocida es la UPV/EHU, también aparecen diferencias significativas para este segundo factor, en el mismo sentido que para el factor de proyección externa, esto es, las diferencias significativas son positivas en aquellos casos en que el individuo cree que la UPV/EHU es muy conocida frente a todas las demás opiniones.

También para este segundo factor la edad es significativa en el mismo sentido en que lo era para el primer factor, es decir, las diferencias son positivas para los individuos de mayor edad.

Cabe mencionar que la intención de compra de artículos con el logotipo de la UPV/EHU a priori no aparece como una variable significativa cuando nos referimos a este segundo factor.

Si nos fijamos en el último factor (el que refleja la opinión sobre el nivel de recursos de la UPV/EHU), se observa que la capital sí es una variable relevante, en tanto que aparecen diferencias significativas negativas de Bilbao, tanto respecto a Vitoria como a San Sebastián, no habiendo entre estas últimas diferencias significativas.

La edad también es una variable significativa, pero el comportamiento en este caso es el contrario al de los otros dos factores. Así, las diferencias son negativas para los individuos de mayor edad respecto de los más jóvenes.

En el caso de este último factor, ni el grado de conocimiento que el individuo cree que la sociedad tiene de la UPV/EHU, ni la intención a priori de compra de artículos con el logotipo de la UPV/EHU son variables relevantes.

Finalmente, cabe comentar que el sexo y el nivel máximo de estudios realizados no tienen efectos significativos sobre ninguno de los tres factores. Tampoco es relevante en el análisis el hecho de que el individuo reconozca haber visto con anterioridad el logotipo de la UPV/EHU.

3.2 Indicador de la imagen de la proyección externa de la UPV/EHU

A continuación, y dado que parece que nos encontramos ante tres dimensiones diferentes que podrían configurar la imagen de la universidad por parte del público, realizamos por separado el análisis factorial para los cuatro primeros ítems, con lo cual hablaríamos de un indicador de la imagen de proyección externa de la universidad. También podríamos realizar el mismo análisis con las otras variables referentes a la proyección interna, pero nos interesa básicamente la proyección externa, porque cuando vayamos a realizar un análisis de la aceptación que podría tener entre el público en general una tienda de artículos con el logotipo de la UPV/EHU, que es el objeto de este estudio, la variable que va a ser relevante es la proyección externa, y no la interna.

De hecho, cuando hemos realizado los análisis de varianza sobre los tres factores obtenidos, se ha llegado a la conclusión de que la variable que recoge la intención a priori de compra de artículos con el logotipo de la UPV/EHU, que sería la que mayor proximidad tiene con el estudio que estamos realizando, únicamente afecta a la puntuación del primer factor obtenido, esto es, el de proyección externa de la universidad, y no tiene efecto sobre ninguno de los otros dos factores considerados.

Este resultado podría estar indicando que lo más relevante para el público en general es en qué medida la sociedad en general se beneficia del trabajo de ésta, en la medida en que contribuye al progreso, a la innovación, a la difusión de la cultura, etc....

Si realizamos el análisis factorial sobre esas cuatro variables, obtenemos los resultados que aparecen en la Tabla 3 referentes a los valores propios y a los porcentajes de varianza total explicada.

Tabla 3

Componente	Autovalores iniciales			Sumas de las saturaciones al cuadrado de la extracción		
	Total	% de la varianza	% acumulado	Total	% de la varianza	% acumulado
1	2,361	59,023	59,023	2,361	59,023	59,023
2	,709	17,726	76,749			
3	,527	13,180	89,930			
4	,403	10,070	100,000			

Se observa cómo con un solo factor queda recogido el 59% de la varianza total. En la primera columna de la Tabla 4 aparecen las correlaciones entre el factor y cada una de las variables. Se observa que en todos los casos las correlaciones son altas, lo que indicaría que todas las variables se recogen de forma adecuada con dicho factor. En la segunda columna de la misma tabla aparecen las puntuaciones de cada variable en las componentes.

Tabla 4

	Correlación	Puntuaciones
	1	1
Abierta	,707	,300
Util	,791	,335
DifCultu	,840	,356
ProInnov	,728	,308

Con los valores que aparecen en la columna de puntuaciones, podríamos determinar para cada individuo cuál es su puntuación en el factor, mediante la siguiente expresión.

$$Z_i = 0.3 * abierta + 0.335 * util + 0.356 * difcultu + 0.308 * proinnov$$

donde cada una de las variables que recogemos son las tipificaciones de las variables originales.

Si tenemos en cuenta que los valores para la media y la desviación típica obtenida para cada una de las variables originales son las que aparecen a continuación,

	Media	Desv. típ.
Abierta	3,66	1,076
Util	3,97	,959
DifCultu	3,92	1,001
ProInnov	3,68	1,007

La ecuación, expresada en función de dichas variables originales, quedaría del siguiente modo:

$$\begin{aligned}
 Z_i &= 0.3 * \left(\frac{abierta - 3.66}{1.076} \right) + 0.335 * \left(\frac{util - 3.97}{0.959} \right) + 0.356 * \left(\frac{difcultu - 3.92}{1.001} \right) + \\
 &+ 0.308 * \left(\frac{proinnov - 3.68}{1.007} \right) = \quad (1) \\
 &= 0.2788 * abierta + 0.3493 * util + 0.3556 * difcultu + 0.3059 * proinnov - 4.92695
 \end{aligned}$$

A partir de esta expresión podremos determinar en qué forma se va a calcular el indicador de imagen de proyección externa de la universidad. Para ello deberemos establecer cuál será la contribución de cada uno de los cuatro ítems considerados en la construcción del indicador. Si quisiéramos que la escala del indicador fuese la misma que la de los ítems, esto es, desde 1 (la peor imagen) hasta 5 (la mejor imagen), bastaría con calcularlo como:

$$Y_i = \left(\frac{Z_i - \min\{Z_j\}}{\max\{Z_j\} - \min\{Z_j\}} \right) \times (5 - 1) + 1 \quad (2)$$

Se podría también plantear un indicador que tomase valores de 1 a 10 o de 1 a 100, sin más que cambiar en la expresión (2) el valor (5-1) por (10-1) o (100-1) respectivamente, pero dado que la escala original para cada ítem estaba entre 1 y 5, ésta nos ha parecido adecuada.

Si tenemos en cuenta que $\min\{Z_j\} = -3.63337$ y $\max\{Z_j\} = 1.52335$, se llega a:

$$Y_i = 0.2162 * \text{abierta} + 0.2709 * \text{util} + 0.2757 * \text{dificultu} + 0.2372 * \text{proinnov}$$

En realidad los coeficientes que ahora multiplican a cada uno de los ítems considerados, y que reflejan la contribución de cada uno de dichos ítems a dicho indicador vendrán determinado por los coeficientes que en la expresión (1) multiplicaban a cada ítem dividido entre la suma de todos ellos. Nótese que ahora la suma de los coeficientes debe ser la unidad, ya que en el caso en que un individuo hubiese valorado con 1 a todos los ítems considerados, el indicador resultante también valdría 1 y, si en el caso opuesto hubiese valorado con 5 a todos los ítems, también el indicador valdría 5.

Aunque las contribuciones de los cuatro ítems son bastante similares, se observa que éstos pueden ser ordenados de mayor a menor importancia de la siguiente forma: en primer lugar la variable Difusión de la Cultura (27.58%), en segundo lugar, la variable Útil (27.09%), en tercer lugar la variable Progreso e Innovación (23.72%). Y finalmente la variable Abierta (21.62%).

4. Intención de compra de artículos con logotipo de la UPV/EHU: modelo de regresión logit

Los modelos de elección discreta se comenzaron a utilizar de forma habitual en el campo de la salud en los años sesenta. Posteriormente se han ido introduciendo en muchos otros campos, como la sociología o las ciencias empresariales, donde se ha utilizado en muchas aplicaciones. Las razones de que este tipo de modelos se utilicen mucho en marketing son las siguientes (Grande y Abascal, 2003): por un lado permiten encontrar tipologías de consumidores, lo que permitirá determinar diferentes perfiles; en segundo lugar permiten calcular las probabilidades de un determinado comportamiento por parte de los consumidores en función de sus características, y finalmente, si se conoce la magnitud de la población, se podrá obtener el tamaño de un segmento multiplicando la probabilidad de comportamiento por el tamaño de la población.

El objetivo de este apartado es establecer un modelo que nos permita estimar la probabilidad de compra de artículos con el logotipo de la UPV/EHU por parte de la población en general, en función de sus características personales. El interés radica en que si somos capaces de conocer tales características sabremos cuál es mayoritariamente el tipo de individuo más proclive a comprar nuestros productos y, evidentemente, a no hacerlo. Al tener información acerca de las características de los consumidores potenciales menos motivados, se podrían adoptar decisiones con el objetivo de aumentar la probabilidad de compra de tales consumidores.

Dado que la variable cuyo comportamiento queremos explicar es la elección del individuo entre dos opciones, en este caso comprar artículos con el logotipo de la UPV/EHU o no hacerlo, y de que en consecuencia se trata de una variable cualitativa, utilizaremos un modelo de elección discreta; en concreto un modelo de regresión logit (Maddala, 1983), (Hosmer-Lemeshow, 2000).

La variable a explicar tomará el valor 0 si el individuo decide no comprar y el valor 1 si el individuo decide sí hacerlo. Al realizar la regresión, planteamos la siguiente relación entre la variable a explicar y las explicativas

$$y_i^* = \sum x_{ij} \beta_j + u_i$$

donde y_i^* es una variable latente, no observable.

Por otro lado se dispone de una variable indicador, y_i , que guarda la siguiente relación con la variable subyacente:

$$y_i = \begin{cases} 1 & \text{si } y_i^* > 0 \\ 0 & \text{en caso contrario} \end{cases}$$

A partir de esta relación se llega a obtener la probabilidad de que el individuo tome la decisión de comprar ($y_i=1$), o de no hacerlo, ($y_i=0$), del siguiente modo:

$$\begin{aligned} P(y_i = 1) &= \pi_i = P(y_i^* > 0) = P(\sum x_{ij}\beta_j + u_i > 0) = \\ &= P(u_i > -\sum x_{ij}\beta_j) = 1 - F_{u_i}(-\sum x_{ij}\beta_j) = F_u(\sum x_{ij}\beta_j) \end{aligned}$$

$$P(y_i = 0) = 1 - \pi_i = 1 - F_u(\sum x_{ij}\beta_j)$$

Tras estimar los coeficientes de regresión seremos capaces de calcular la probabilidad de que un determinado individuo compre o no lo haga en función de sus características personales. Así, dado que estamos utilizando un modelo logit, y en consecuencia la distribución utilizada es la logística, se tiene que:

$$\begin{aligned} \pi_i = P(y_i = 1) &= \frac{e^{\sum x_{ij}\hat{\beta}_j}}{1 + e^{\sum x_{ij}\hat{\beta}_j}} \\ 1 - \pi_i = P(y_i = 0) &= \frac{1}{1 + e^{\sum x_{ij}\hat{\beta}_j}} \end{aligned}$$

4.1 Intención de compra de productos con el logotipo de la UPV/EHU

Dado que vamos a trabajar con modelos de elección binaria, la variable que queremos explicar, la decisión de comprar artículos con logotipo de la UPV/EHU o de no hacerlo, deberá tomar dos valores únicamente. En la encuesta la información que aparece es una valoración (en una escala del 1 al 4) del grado de probabilidad que el individuo cree que tiene de comprar alguno de los artículos que se le presentan (desde muy poco probable a muy probable) para determinados grupos de artículos. Para definir nuestra variable consideraremos que el individuo tiene intención de comprar algún

artículo del grupo en cuestión si cree que es probable o muy probable que lo haga y consideraremos que el individuo tiene intención de comprar artículos con el logotipo de la UPV/EHU, en general, si tiene intención de comprar al menos de uno de los grupos de artículos que se le presentan.

Si nos referimos ahora a las variables explicativas utilizadas en la estimación, han sido las siguientes:

- Sexo, dividido en hombre y mujer.
- Ciudad de realización de la encuesta: Bilbao, Vitoria, San Sebastián
- Edad: aunque inicialmente había cuatro tramos, se ha recodificado y los dos grupos finales son los siguientes: de 14 a 29 y de 30 o más de 30 años.
- Indicador de imagen externa de la UPV/EHU.
- Intención de compra a priori.
- Opinión acerca de la medida en que las siguientes características deben aparecer en artículos con el logotipo de la UPV/EHU: Original, Audaz, Práctico, Tradicional, Artístico, Elegante, Serio, Moderno.

En los casos de Sexo, Ciudad, Edad e Intención de compra a priori, al plantear el modelo dejamos una de las categorías de cada variable fuera, dado que de no hacerlo así, tendríamos un problema de multicolinealidad perfecta entre las variables explicativas. La que hemos dejado fuera en todos los casos ha sido la última. De este modo, a la hora de extraer las conclusiones, debemos ser conscientes de que siempre deberán referirse a aquella categoría que hemos dejado sin recoger en el modelo (la de referencia).

Si nos referimos a la opinión acerca de en qué medida las características original, audaz, práctico, tradicional, artístico, elegante, serio y moderno debían aparecer en un artículo con logotipo de la UPV/EHU, se planteó incluir cada una de ellas como una variable explicativa. Se observó que existía entre algunas de ellas una correlación elevada que podría provocar problemas de multicolinealidad. Por ello se procedió a realizar un análisis factorial previo, y considerar como variables explicativas de la regresión los dos primeros factores obtenidos de dicho análisis, tras realizar una rotación Varimax, que recogían un 52% de la inercia total.

En la tabla 2.1 aparecen las correlaciones de las variables originales con cada uno de los factores. Se observa que las características que tienen proyección alta sobre el primer factor son: original, audaz, artístico, moderno y práctico. Las que tienen proyección elevada sobre el segundo factor son tradicional, elegante y serio. Podríamos por lo tanto considerar que el primer eje recoge características más innovadoras y/o prácticas y el segundo más clásicas.

Tabla 2.1

CARACTERISTICAS CONSIDERADAS	FACTORES	
	1	2
ORIGINAL	,762	-,111
AUDAZ	,664	-,050
PRACTICO	,511	,242
TRADICIONAL	-,024	,775
ARTISTICO	,639	,123
ELEGANTE	,316	,691
SERIO	-,127	,833
MODERNO	,706	-,001

Finalmente cabe mencionar que también se incluyó inicialmente en la regresión el factor que recogía la imagen de actividad interna de la Universidad, así como la variable referente a si los recursos eran escasos o excesivos. Ahora bien, se han excluido del modelo dado que no son significativas. Lo mismo ha sucedido con la variable que recoge el nivel máximo de estudios alcanzado.

Tras realizar estas agrupaciones se llega a la siguiente ecuación de regresión, donde ahora todas las variables son significativas al nivel de significación del 5%.

$$X'\beta = -3.799 - 0.502 * \text{hombre} + 2.965 * \text{Bilbao} + 2.113 * \text{Vitoria} + 0.861 * \text{joven} + \\ + 2.332 * \text{intención inicial de compra} + 0.321 * \text{factor innopract} + \\ + 0.523 * \text{factor clásico} + 0.395 * \text{imagen externa}$$

(R2 de Nagelkerke³: 0.443)

recordando que si queremos estimar la probabilidad de que el individuo esté dispuesto a comprar artículos con el logotipo se calculará utilizando la expresión:

³ (Nagelkerke, N. J. D., 1991)

$$\hat{\pi}_i = \hat{P}(y_i = 1) = \frac{e^{x_i'\hat{\beta}}}{1 + e^{x_i'\hat{\beta}}}$$

A la vista de los resultados de la regresión se pueden extraer las siguientes conclusiones.

Si nos fijamos en el sexo se observa que las mujeres, tienen una probabilidad estimada de comprar artículos con el logotipo de la UPV/EHU mayor que los hombres.

La capital en que se realiza la encuesta también tiene relevancia; de hecho la magnitud de los coeficientes nos lleva a concluir que en Bilbao la probabilidad estimada de compra es significativamente mayor que en Vitoria, y en ambos casos mucho mayor que en San Sebastián. Este resultado también se obtenía al analizar la encuesta realizada al personal que mantiene vinculación con la UPV/EHU si comparamos los campus de Vizcaya, Alava y Guipúzcoa.

La siguiente variable que aparece como significativa es la edad del individuo; en este caso, dado que el grupo de referencia son los individuos de 30 o más años, lo que indicaría el resultado obtenido es que los de menor edad tienen una mayor probabilidad de adquirir productos con el logotipo de la UPV/EHU.

A continuación nos fijamos en la variable que recoge la intención inicial que el encuestado tenía de comprar artículos con el logotipo de la UPV/EHU. El signo positivo indica que tienen mayor probabilidad estimada de comprar artículos con el logotipo aquellos individuos que a priori ya tienen intención de hacerlo. Se observa que se trata de una variable muy relevante en el estudio y es lógico que sea así, dado que el individuo ya está mostrando su predisposición sin haber visto ninguno de los artículos que posteriormente se pondrán a la venta.

Las siguientes variables que aparecen en la ecuación de regresión se refieren a la medida en que determinadas características debieran aparecer en los artículos con el logotipo de la UPV/EHU. Tras la realización del análisis factorial, se comprobaba que había dos factores que agrupaban por un lado las características de tipo innovador o práctico, y por otro las de tendencia clásica. En la regresión se observa que ambos factores son relevantes, lo que indica que cuanto más considere el individuo que deban

darse las características correspondientes, mayor probabilidad de compra tiene; si nos fijamos en la magnitud de los coeficientes, lo que observamos es que la línea clásica tiene un mayor efecto sobre la probabilidad estimada de compra.

Finalmente llegamos a la variable que recoge el indicador de la imagen de la proyección externa de la universidad. Se ha considerado el indicador con una puntuación de entre 1 (peor imagen posible) y 5 (mejor imagen posible). Se observa que su efecto es significativamente positivo sobre la probabilidad estimada de compra de artículos con el logotipo de la UPV/EHU. Nos parece un resultado acorde con lo esperado, dado que los individuos están más dispuestos a adquirir artículos con logotipo de aquellas instituciones que les parecen cercanas, prestigiosas, ..., en definitiva, de las que tienen una buena imagen.

4.2 Efecto de la imagen de la universidad sobre la probabilidad estimada de compra de artículos con el logotipo de la UPV/EHU

Nos podemos plantear evaluar cuál es el efecto de un incremento unitario de la imagen de la proyección externa de la universidad sobre la probabilidad estimada de compra de artículos con el logotipo de la UPV/EHU. Dado que el modelo de regresión logit no es lineal, la derivada de dicha probabilidad estimada respecto de la variable en cuestión depende, además del coeficiente que corresponde en la regresión a dicha variable, de los valores de todas las variables explicativas. Debido a ello no se puede hablar de un incremento en la probabilidad estimada común para todos los individuos. Por ello, lo que haremos es hablar de un incremento medio de probabilidad estimada; es decir, calcularemos el incremento para cada individuo y posteriormente lo promediamos. En concreto dicho incremento sería de un 6.4%, ya que la probabilidad media pasaría de ser un 49.9% a un 56.3%. Ahora bien, se trataría de una estimación sesgada a la baja, porque en realidad un incremento en esta variable actúa doblemente sobre la probabilidad estimada de compra.

Así, el efecto que se acaba de calcular es el directo, pero además debemos tener en cuenta que cuando hemos analizado qué variables tenían alguna relación con este indicador, una de ellas era la intención inicial de compra. La razón es que cabe esperar

que cuanto mejor sea la imagen de la universidad percibida por un individuo, mayor será su intención de compra de artículos con su logotipo.

De hecho, si separamos a los encuestados de acuerdo con el valor del indicador en tres grupos del mismo tamaño, se observa que para los individuos que tienen una buena imagen de la proyección externa de la universidad el porcentaje de los que están dispuestos a adquirir artículos con el logotipo de la universidad sin ver los artículos es del 27.2%, para los que tienen una imagen intermedia dicho porcentaje pasa a ser del 22.9%, y para los que tienen una mala imagen el porcentaje desciende hasta el 14.5%, siendo estas diferencias significativas. Es evidente por lo tanto que la intención de compra de artículos con el logotipo de la UPV/EHU que tiene un individuo depende de la imagen que percibe de la proyección externa de la universidad. Si tenemos en cuenta que la intención de compra a priori es muy significativa en la decisión final de compra (una vez vistos los grupos de artículos), es evidente la existencia de un efecto indirecto de la imagen de la proyección externa de la universidad a través de la intención de compra a priori. Además, si la universidad tiene éxito en su objetivo y la creación de la tienda universitaria contribuye a una mejora de su imagen entre la sociedad, esto a su vez supondría un incremento en la probabilidad estimada de compra en un futuro.

5. Conclusiones generales

En este trabajo hemos conseguido, mediante la utilización del análisis factorial, determinar cuáles son las principales dimensiones de la imagen de una universidad pública, utilizando como referencia la Universidad del País Vasco. Así, hemos obtenido una dimensión que reflejaría la proyección externa de la universidad, otra que reflejaría la organización interna de la universidad y la última que recogería la percepción que los individuos tienen de la escasez o exceso de los recursos asignados a la universidad. Además, dado que cuando analicemos el comportamiento de los individuos respecto de la posible adquisición de artículos con el logotipo de dicha institución el factor relevante es el referente a la proyección externa de la universidad, hemos elaborado un indicador de dicha proyección externa, señalando cuál es la aportación a dicho indicador de cada uno de los cuatro ítems que lo componen.

Por otro lado, hemos comprobado cómo la utilización del modelo de regresión logit nos ha permitido determinar cuáles son las características personales que hacen variar la probabilidad de compra de artículos con el logotipo de la UPV/EHU por parte de un individuo. Así, las variables que se han mostrado como relevantes son el sexo, la edad, la capital en que se ha realizado la encuesta, la intención que los individuos tenían a priori de comprar artículos con el logotipo, la opinión sobre la medida en que determinadas características deben reflejarse en artículos con el logotipo y finalmente la imagen de la proyección externa de la universidad. De hecho, hemos llegado a cuantificar cuál sería el incremento en la probabilidad estimada de artículos con el logotipo de una mejora en la imagen de la proyección externa de la universidad que conlleve un aumento de un punto en el indicador que la mide.

Bibliografía

Fernández, K., Martín, A.M., Modroño, J.I., Zorrilla, P. (2005): *Análisis de la viabilidad de la tienda universitaria de la UPV/EHU “ehuDENDA” de acuerdo con la imagen percibida de la institución*. Informe Técnico, Universidad del País Vasco/Euskal Herriko Unibertsitatea, Bilbao

Grande, I., Abascal E. (2003): *Fundamentos y técnicas de investigación comercial*. 7ª edición. ESIC Editorial

Hair J. F., Anderson R. E. , Tatham R. L. , Black W. C. (1999): *Análisis Multivariante*, quinta edición, Prentice Hall.

Hochberg, Y. , Tamhane, A.C. (1987): *Multiple Comparison Procedures*, New York, John Wiley and Sons.

Hosmer, D.W.; Lemeshow, S. (2000): *Applied Logistic Regression, 2nd Edition*, Wiley & Sons Inc, USA.

Klockars, Alan J. and Sax, G. (1986): *Multiple Comparison*, SAGE Quantitative Applications Series.

Luque, T. (2000): *Técnicas de análisis de datos en investigación de mercado*, Ed. Pirámide.

Maddala, G. S. (1983): *Limited-dependent and qualitative variables in econometrics*. Econometric Society Monographs.

Nagelkerke, N. J. D. (1991): "A note on a general definition of the coefficient of determination", *Biometrika*, 78, pp. 691-692

Peña, D. (2002): *Análisis de Datos Multivariantes*. Ed. Mc Graw-Hill.

LA PREDICCIÓN DEL ÍNDICE GENERAL DE LA BOLSA DE MADRID (IGBM): Incidencia de la distribución residual y efectos estacionales.

Israel Senra Díaz

e-mail: israelsenra@ine.es

Instituto Nacional de Estadística (INE)

J. Bernardo Pena Trapero

e-mail: bernardo.pena@uah.es

Departamento de Estadística, Estructura Económica y O.E.I.

Universidad de Alcalá (UAH)

Área temática: Métodos Cuantitativos

RESUMEN

Las distribuciones de las series financieras suelen tener colas pesadas y distribuciones residuales distintas de la normal. En el caso de los datos diarios del Índice General de la Bolsa de Madrid (IGBM), Pena y Senra (2006) estudiaron la posibilidad de recoger este comportamiento mediante la modelización de datos atípicos. Alternativamente en este trabajo se analizan las distribuciones diferentes a la normal consideradas en este tipo de literatura y se realiza un estudio de la capacidad predictiva del modelo GARCH y los asimétricos atendiendo a esta situación, mediante su aplicación al IGBM.

Otro objetivo de este trabajo es analizar la existencia de efectos estacionales mediante el análisis del posible comportamiento diferenciado tanto en los meses de verano como de los días de la semana en el IGBM, tal y como apunta Tsay (2002) que puede ocurrir en otras series de estas características.

ABSTRACT

Financial time series usually show heavy tails and residual distributions different than normal. Pena and Senra (2006) studied the possibility to capture this effect through outliers with daily data for the the Índice General de la Bolsa de Madrid (IGBM). Alternatively, this paper considers residual distributions different than normal and evaluate empirically their performance with GARCH and assymetric models applied to IGBM.

Another objective is to analyse the existence in IGBM of seasonal effects through a differenced behaviour in summer months or depending on the day of the week as it is suggested by Tsay (2002).

Palabras Clave: Efectos estacionales, asimetría, evaluación de predicciones, GARCH.

Clasificación JEL: C52; C53

1. INTRODUCCIÓN.

La modelización y predicción de la volatilidad del mercado ha sido un punto central de la investigación empírica y teórica durante la última década por parte de investigadores y profesionales del sector. La volatilidad es uno de los conceptos más importantes en el análisis de datos financieros. La volatilidad medida como la desviación típica o la varianza de los rendimientos, se utiliza frecuentemente como una estimación del riesgo total del activo financiero. Muchos modelos de valoración de activos requieren la estimación o predicción de la volatilidad (inobservada). La volatilidad de los precios de mercado también es parte fundamental de la fórmula de Black-Scholes para valorar el precio de las opciones negociadas. En este sentido el artículo de González-Rivera, Lee y Mishra (2004) analiza la caracterización de la volatilidad mediante funciones de pérdida económica basadas en el valor del riesgo. Otra forma de caracterizar la volatilidad es mediante la denominada “Realized volatility” que considera los rendimientos intradiarios al cuadrado. El artículo de Hansen y Lunde (2004) estudia las características de este tipo de aproximación y muestra que cuando el intervalo de tiempo tiende a cero las estimaciones de la volatilidad son consistentes. No obstante, esta alternativa queda descartada, puesto que en este trabajo se consideran datos diarios. Por último, un enfoque ampliamente utilizado se basa en considerar una función de los rendimientos al cuadrado, véase Andersen y Bollerslev (1998), que resulta en un estimador insesgado aunque poco preciso.

Existe un gran número de diferentes tipos de modelos no lineales, pero sólo un pequeño número de ellos han resultado útiles para modelizar datos financieros. De estos modelos los más populares son los ARCH de Engle (1982) y los GARCH de Bollerslev (1986), que utilizan una medida de volatilidad dependiente de los rendimientos al cuadrado. Los modelos de la familia ARCH-GARCH son capaces de recoger características habituales en las series de datos de alta frecuencia como la leptocurtosis o el agrupamiento de la volatilidad. Estos modelos adolecen, sin embargo, del defecto de que no captan comportamientos asimétricos que presenta la volatilidad en numerosos casos debido al hecho observado en las finanzas empíricas de la correlación negativa entre los rendimientos de los valores bursátiles y la

volatilidad. Existen dos explicaciones básicas a este hecho. Por una parte el llamado efecto apalancamiento, según el cual un rendimiento negativo hace que el activo bursátil tenga un mayor riesgo y por ello incrementa la volatilidad aumentando el apalancamiento financiero. Por otra parte un aumento anticipado de la volatilidad eleva el rendimiento referido sobre el activo. De esta forma, como señalan Awartani y Corradi (2005), el efecto apalancamiento y la explicación proporcionada por la retroalimentación de la volatilidad conducen a nexus causales diferentes; de hecho el primero prescribe un nexus causal de los rendimientos a la volatilidad condicional, mientras este último lo hace de la volatilidad condicional a los rendimientos.

En junio de 2005 Pena y Senra, en su trabajo “Predicción de la volatilidad de los rendimientos del Índice de la General de la Bolsa de Madrid (IGBM)” que presentaron en la Asamblea General de Asepelt, se examinaba la capacidad predictiva de distintos modelos de la familia ARCH para caracterizar la volatilidad de los rendimientos del Índice General de la Bolsa de Madrid (IGBM). En dicho trabajo se constató que el modelo GARCH(1,1) es el que mejor se comporta en el análisis de los errores de predicción un período hacia delante, mientras que en los modelos asimétricos se consiguen ganancias en los errores de predicción a dos y cinco períodos. Una extensa revisión de los modelos asimétricos se puede encontrar en el artículo de Poon y Granger (2003). Mientras que artículos como los de Loudon, Watt y Yadav (200), Taylor (2001) o Awartani y Corradi (2005), entre otros, encuentran evidencia a favor del uso de modelos asimétricos, otros artículos como los de Bluhm y Yhu (2000) o Brooks (1998) la encuentran en contra.

Posteriormente, en Junio de 2006 Pena y Senra, también con motivo de la Asamblea General de Asepelt presentaron otro trabajo “Análisis del efecto de los datos atípicos y la asimetría en la predicción de la volatilidad del Índice de la General de la Bolsa de Madrid (IGBM)” que continuaba con el anteriormente indicado, estudiando la presencia de atípicos que incluía el IGBM y los efectos de aquéllos sobre la capacidad predictiva de los modelos de la familia ARCH. En este trabajo se llega a la conclusión de que los resultados fuera de la muestra indican que la inclusión de atípicos no consiguen mejorar la capacidad predictiva del modelo GARCH(1,1),

mientras que la combinación de las alternativas asimétricas, como la formulación de los modelos EGARCH, junto con la inclusión de atípicos mejoran las predicciones del GARCH(1,1), siendo aquél el único modelo, entre los asimétricos estudiados, que da mejores resultados para todos los horizontes de predicción (a 1,2,3,4,5 y 10 períodos hacia delante). La presencia de atípicos en series temporales puede producir diversos efectos en la identificación y estimación de modelos, tanto para la media como para la varianza condicional. Por otra parte, tanto la heterocedasticidad condicional como la presencia de atípicos pueden generar el exceso de curtosis, véase por ejemplo, Carnero, Peña y Ruiz (2001).

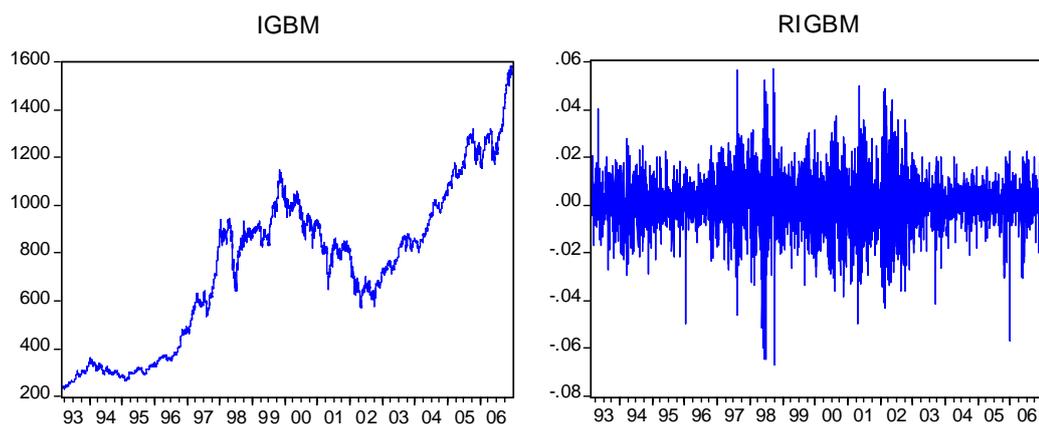
En el presente trabajo amplían los estudios anteriores mediante las siguientes aportaciones. En la sección 2 se realiza una revisión del Índice General de la Bolsa de Madrid con información actualizada del mismo hasta Diciembre de 2006, que será utilizado como referencia para evaluar los resultados en predicción de los distintos modelos. En la sección 3 se presenta la actualización de los datos, así como un contraste de existencia de asimetrías tanto en signo como del tamaño del sesgo de dichas asimetrías en los modelos GARCH, además se presentan las curvas de impacto motivadas por nueva información. En la sección 4, dado que en los modelos anteriores se había supuesto que la distribución de los rendimientos eran normales, y teniendo en cuenta que según Rue Tsay (2002), si el modelo GARCH(1,1), presenta el coeficiente autorregresivo positivo, entonces era un indicio de colas pesadas en la distribución residual, se utilizarán distribuciones distintas a la normal (t-Student y la Generalized Error Distribution (GED)). Partiendo de esta nueva hipótesis en la distribución residual se analizará la posible presencia de estacionalidad en los rendimientos tanto dentro de la semana como dentro del año para datos mensuales, en particular para poder contrastar la hipótesis del efecto de los meses de verano en relación con el resto del año. En la sección 5, se realiza un ejercicio de evaluación de las predicciones y por último en la sección 6 se ofrecen las principales conclusiones.

2. EVOLUCIÓN DEL ÍNDICE IGBM.

El IGBM es el índice al que se refiere el presente trabajo. Con este índice se pretende obtener una visión lo más comprensiva posible del mercado de acciones español y cada año actualiza su composición teniendo en cuenta la capitalización bursátil de los valores admitidos o negociados en la Bolsa de Madrid a 31 de diciembre de cada ejercicio, y su composición y ponderación permanecen inalterados durante todo el año.

Se dispone de 13 años de observaciones de los datos diarios de cierre del Índice General de la Bolsa de Madrid, desde el 17 de febrero de 1993 hasta el 29 de diciembre de 2006. El objetivo es caracterizar y predecir la volatilidad de los rendimientos del IGBM (RIGBM) calculados a partir de la primera diferencia del logaritmo del índice. La figura 1 muestra la evolución temporal tanto del índice IGBM como de sus rendimientos para el período muestral considerado.

Figura 1: Evolución del índice IGBM y de sus rendimientos para el período 1993-2006.



Los gráficos son suficientemente elocuentes para comprender el comportamiento de este índice desde 1993, no obstante es conveniente señalar algunos rasgos históricos para la mejor comprensión de lo acaecido en este lapso de tiempo. El periodo comprendido entre los años 1993 y 1996 presenta una etapa de relativa atonía en la actividad económica española. Este último año representa el inicio de un periodo de

dicha actividad en el marco de una economía europea en expansión, además se intensifican las inversiones de empresas españolas cotizadas en América Latina, registrándose tasas de crecimiento económico importantes también en dicha área.

La estabilidad alcista controlada por el indicador hasta el verano de 1998 parece interrumpirse como consecuencia de la crisis de las economías emergentes. Desde septiembre de 2000 –crisis de los mercados tecnológicos-, los mercados bursátiles de los países desarrollados inician una tendencia bajista que afecta a la generalidad de las empresas cotizadas en un contexto económico menos dinámico. En 2001 los países latinoamericanos vieron truncada la recuperación registrada por el adverso contexto externo aunque sus efectos fueron absorbidos sin pasar por crisis internas, a excepción de Argentina. Su impacto fue indudable en la bolsa española debido a los vínculos existentes entre las empresas españolas cotizadas y el tejido productivo y financiero argentino. La fase expansiva de la economía española se frenó en 2001. La desaceleración fue menor que la registrada como promedio en la zona euro por lo que el diferencial de crecimiento se mantuvo positivo en relación con dichas economías. No obstante, la pauta inversora de las familias españolas en activos financieros se iba haciendo más conservadora en consonancia con el empeoramiento de las perspectivas económicas.

El ejercicio de 2001 se caracteriza por la incertidumbre –que crece a raíz de los acontecimientos del 11 de septiembre, que llevan a cuestionar las perspectivas de rentabilidad empresarial, especialmente en actividades turísticas y aseguradoras- y su reflejo en los mercados financieros es notable. El comportamiento del IGBM acusa la incertidumbre. Asimismo, crece la desconfianza de los inversores en el funcionamiento de los mercados, en particular los norteamericanos como consecuencia de la suspensión de pagos de la empresa de energía Enron. Sin embargo, en el último trimestre del ejercicio se produce una recuperación gradual de las cotizaciones bursátiles, que probaría que la percepción del impacto sobre los resultados empresariales no iba a ser tan negativa como inicialmente se había previsto. Además, las autoridades monetarias y económicas reaccionaron con la

adopción de medidas, tanto en USA como en la zona euro, contribuyendo a mejorar la confianza en los mercados¹.

No obstante, 2002 sigue caracterizándose por la inestabilidad económica internacional, y el deterioro de los beneficios empresariales, lo que provoca cierta pérdida de confianza de los inversores, a pesar de constatarse un notable posicionamiento de inversores no residentes en el mercado de acciones español. Habrá que esperar a 2003 para apreciar la recuperación bursátil, en consonancia con una recuperación económica.

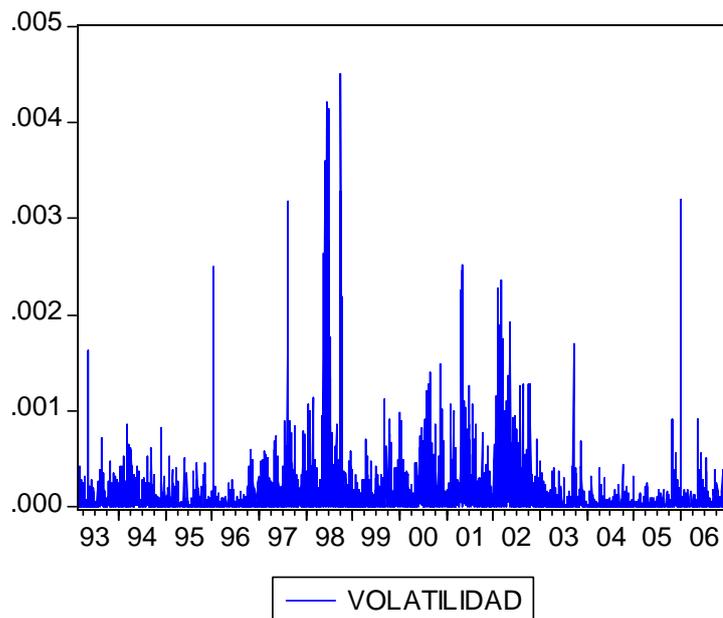
En 2004 ya se puede hablar de un ejercicio favorable para los mercados de valores españoles y dicha circunstancia es claramente apreciable en el comportamiento del IGBM, a pesar de la volatilidad de los precios del petróleo. En cuanto al impacto del 11-M en la economía española, éste no fue tan significativo como el que podía esperarse, aunque resulta perceptible su influencia, si bien no es comparable a los efectos del 11-S. Por último, en líneas generales el comportamiento del índice en 2005 ha sido favorable y ha presentado un espectacular repunte a finales del 2006.

Dado que la volatilidad no es una variable observable, se necesita una buena aproximación de la misma. Si la media condicional es cero, los rendimientos al cuadrado proporcionan un estimador insesgado del verdadero proceso subyacente que sigue la volatilidad. Si la media condicional no es cero, entonces se deben utilizar los residuos al cuadrado de la regresión de los rendimientos sobre una constante, su pasado y otros regresores si fueran necesarios. Si la media condicional estuviera mal especificada, los residuos al cuadrado no serían un estimador insesgado de la varianza condicional. En el caso de los rendimientos del IGBM, la estructura de la media condicional depende de la inclusión ó no de atípicos en la misma, es por

¹ Un tratamiento más exhaustivo sobre la evolución económica en este período puede verse, por ejemplo en www.cmnv.es

ello, que en cualquier caso, se consideran los rendimientos al cuadrado como estimador de la volatilidad². La figura 2 muestra la volatilidad estimada.

Figura 2: *Estimación de la volatilidad mediante la serie de rendimientos al cuadrado.*



Independientemente de la insesgadez de la estimación de la volatilidad mediante los rendimientos al cuadrado, la capacidad predictiva de la misma es muy baja (el coeficiente de determinación en la regresión de los rendimientos al cuadrado frente a la volatilidad, no puede exceder $1/3$), tal y como señalan Andersen y Bollerslev (1998). No obstante, Awartani y Corradi (2005) señalan la validez de esta aproximación cuando el objetivo es la comparación de la capacidad predictiva de distintos modelos, basándose en que si la función de pérdida es cuadrática, el uso de los rendimientos al cuadrado asegura que se obtiene una ordenación correcta de los modelos respecto a su capacidad predictiva.

² Obsérvese que al considerar distintas especificaciones, los distintos modelos dan lugar a distintas series de residuos y distintas estimaciones de la volatilidad. Es por ello, que, a efectos de comparación de las predicciones de los distintos modelos se utilizarán los rendimientos al cuadrado.

3. CONTRASTE DE EXISTENCIA DE ASIMETRÍAS.

Una de las principales características de los rendimientos de activos, es una respuesta asimétrica de la volatilidad a nivel de los rendimientos. Este hecho se conoce como efecto “apalancamiento” Ruiz (1994), y consiste en que la volatilidad se incrementa en mayor medida cuando el rendimiento es negativo que cuando el rendimiento es positivo Black (1976) y Christie(1982). Los modelos simétricos, sin embargo, tienen el defecto de que no captan los comportamientos asimétricos que presenta la volatilidad, debido entre otros aspectos al hecho observado en las finanzas empíricas de la correlación negativa de los rendimientos bursátiles y con la volatilidad.

Para contrastar que IGBM posee comportamientos asimétricos se aplica los contrastes de Engle y Ng (1993), conocidos más comúnmente como los “tests de signos y del tamaño del sesgo”.

Estos contrastes sirven para determinar si existe la necesidad de estimar un modelo asimétrico para modelizar la serie o bien si se requiere utilizar el modelo simétrico GARCH.

En la práctica, estos contrastes³ se aplican a los residuos del ajuste de un GARCH (U_t) a los rendimientos.

Previamente a la realización del mismo se definen dos variables ficticias S_{t-1}^- y S_{t-1}^+ , de la siguiente manera

$$S_{t-1}^- = \begin{cases} 1 & \text{si } \hat{U}_{t-1} < 0 \\ 0 & \text{en el resto} \end{cases} \quad \text{y} \quad S_{t-1}^+ = 1 - S_{t-1}^-$$

y a continuación se estima el siguiente modelo de regresión :

$$\hat{U}_t^2 = \phi_0 + \phi_1 S_{t-1}^- + \phi_2 S_{t-1}^- U_{t-1} + \phi_3 S_{t-1}^+ U_{t-1} + V_t$$

donde V_t es el término de error idénticamente distribuido.

Se estima por MCO y de ahí se obtienen las siguientes conclusiones:

- Si el signo de los impactos sobre \hat{U}_{t-1} , tanto positivos como negativos, influye sobre la varianza condicionada, se tendrá que el coeficiente ϕ_0 será estadísticamente significativo.
- Si ϕ_1 es significativo, la magnitud del sesgo provocado por los signos influye en magnitud sobre la varianza condicionada.
- Si ϕ_2 y/o ϕ_3 son estadísticamente significativos, nos estaría indicando que los shocks de signo negativo y/o positivos influyen en magnitud sobre la varianza condicional.

Realizada la estimación se obtiene que:

$$\hat{U}_t^2 = 0,000107 - 3,08E-0,5 S_{t-1}^- - 0,011086 S_{t-1}^- U_{t-1} + 0,003088 S_{t-1}^+ U_{t-1}$$

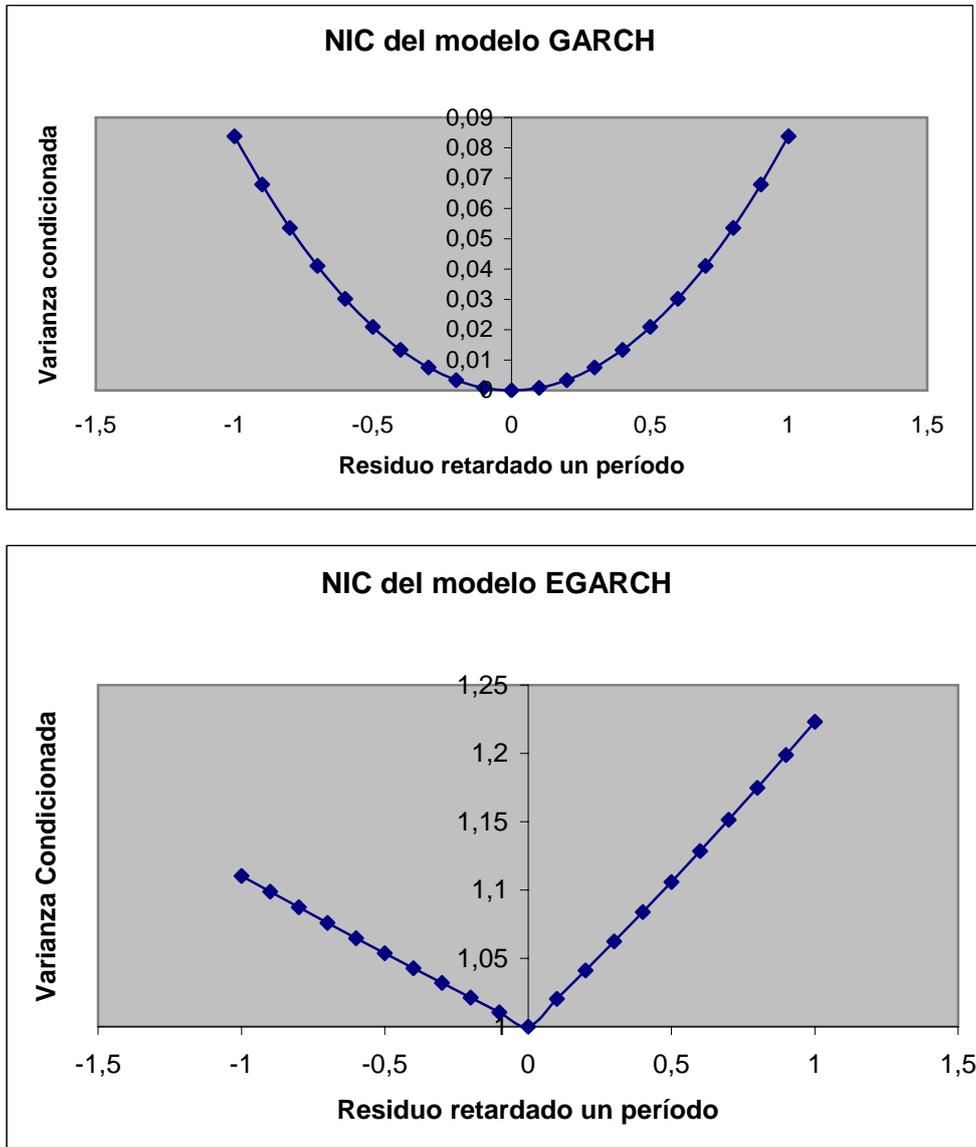
(13,52) (-2,32) (-11,57) (3,90)

Como se puede comprobar, todos los coeficientes son significativos por lo tanto la serie presenta efectos asimétricos, tanto en signo como en magnitud de dichos impactos sobre la volatilidad condicionada.

Una manera de poder observar la magnitud de estos impactos se puede realizar de manera gráfica a través de las llamadas curvas de impacto que se deben a Pagan y Schwert (1990), mas conocidas como NIC (News Impact Curves). A continuación se adjuntan las mismas de los modelos que se analizarán posteriormente, el modelo simétrico GARCH y el asimétrico que por las características e importancia resaltadas en el trabajo Pena y Senra (2006) se utilizará en este trabajo, el EGARCH.

³ Se puede encontrar un desarrollo mas detallado de estos contrastes en Brooks (2002)

Figura 3: Representación gráfica de las curvas de impacto de los modelos GARCH y EGARCH.



Fuente: Elaboración propia.

Las NIC se han representado utilizando la varianza condicional del modelo estimado correspondiente, tomando como varianza incondicional la varianza condicional retardada. Para representar las NIC se ha considerado un rango para los valores de U_{t-1} entre -1 y $+1$, para que sea mas comprensible la interpretación y la comparación entre ambas curvas de impacto.

Como se puede observar la NIC del modelo GARCH es simétrica con respecto al cero, por lo tanto un impacto de una magnitud dada tendrá el mismo efecto sobre la varianza condicional sea cual sea el signo de ésta. Sin embargo, en el modelo EGARCH se puede ver que ante impactos de la misma magnitud, aquellos que tienen signo positivo tienen más repercusión sobre la volatilidad condicionada que los negativos.

4. ANÁLISIS DE LA ESTACIONALIDAD EN EL IGBM

En el contexto de series financieras se pueden observar efectos estacionales relacionados con el día de la semana o la estación del año. Evidencia en este sentido se encuentra en French (1980), Gibbons y Hess (1981), Keim y Stambaugh (1984) o Jaffe y Westerfield (1985) entre otros. Estos resultados parecen contradecir la hipótesis de eficiencia de los mercados, pero se argumenta que no tienen por qué estar directamente asociados a ganancias y pueden estar relacionados con primas de riesgo variables en el tiempo. En cualquier caso su omisión puede conducir a un modelo mal especificado.

Teniendo en cuenta los posibles efectos estacionales (tanto desde el punto de vista diario como de los meses), así como la estructura en niveles de la serie, se ha realizado la modelización considerando el efecto que tiene sobre el nivel y la volatilidad condicionada de la misma.

Se ha modelizado la media mediante un modelo ARIMA(1,0,0) sin constante y GARCH(1,1) y otro con igual modelización para la media y EGARCH(1,1). Se han introducido todas las variables ficticias correspondientes en cada caso, e igualmente en la ecuación de la varianza sin constante. Para la estimación de los modelos se ha considerado como distribución residual la *t* de Student y la GED (Generalized Error Distribution)

B.1.2) Distribución residual: GED

$$(1 - 0,053146) \text{rigbm}_t = 0,000836 L + 0,001128 M + 0,000952 X + 0,000699 J + \varepsilon_t \sigma_t$$

(2,9011) (2,3404) (3,2744) (2,8264) (2,0219)

$$\log(\sigma_t^2) = -0,282536 V + 0,154529 \left| \frac{\varepsilon_{t-1}}{\sigma_{t-1}} \right| - 0,045078 \frac{\varepsilon_{t-1}}{\sigma_{t-1}} + 0,983323 \log(\sigma_{t-1}^2)$$

(-2,8398) (9,3501) (-4,7413) (266,5186)

AIC=-6,2991

BIC=-6,2909

En cuanto a la consideración del modelo GARCH asimétrico, ambas formulaciones presentan especificaciones similares. En ambas el efecto día de la semana resulta significativo y el mejor ajuste se vuelve a alcanzar con el modelo con distribución t de Student.

C) MODELOS QUE RECOGEN EL EFECTO DE LOS MESES

C) Modelos GARCH

C.1.1) Distribución residual: t-student

$$(1 - 0,0433) \text{rigbm}_t = 0,0016 \text{Enero} + 0,0017 \text{Febrero} + 0,0015 \text{Noviembre} +$$

(2,35) (2,91) (2,69) (2,87)

$$+ 0,047073 \text{Diciembre} + \varepsilon_t \sigma_t$$

(3,03)

$$\sigma^2 = 1,67e - 06 \text{Diciembre} + 0,0844 \varepsilon_{t-1}^2 + 0,902912 \sigma_{t-1}^2$$

(8,15) (85,89)

AIC=-6,31

BIC=-6,26

C.1.2) Distribución residual: GED

$$(1 - 0,0415) \text{rigbm}_t = 0,0016 \text{Enero} + 0,0017 \text{Febrero} + \varepsilon_t \sigma_t$$

(2,25) (2,91) (2,69)

$$\sigma^2 = -2,49e - 06 \text{Noviembre} + 2,56e - 06 \text{Diciembre} + 0,0809 \varepsilon_{t-1}^2 + 0,902912 \sigma_{t-1}^2$$

(-2,08) (2,85) (8,46) (89,12)

AIC=-6,31

BIC=-6,25

La especificación alternativa del modelo GARCH con distribución residual t de Student presenta un menor criterio de ajuste. Centrándonos en la ecuación con mejor ajuste, se ve que en la especificación de la media presentan rendimientos más elevados los meses de enero, febrero, noviembre y diciembre, mientras que en la realización de la varianza sólo diciembre resulta significativo.

D) MODELOS QUE RECOGEN EL EFECTO DE LOS MESES

D.1) Modelos EGARCH

D.1.1) Distribución residual: t-student

$$\begin{aligned}
 (1 - 0,051820) \text{rigbm}_t &= 0,0014 \text{ Enero} + 0,0015 \text{ Febrero} + 0,0012 \text{ Noviembre} + \\
 &\quad (2,8757) \quad (2,71) \quad (2,69) \quad (2,23) \\
 &+ 0,0014 \text{ Diciembre} + \varepsilon_t \sigma_t \\
 &\quad (2,43) \\
 \log(\sigma^2) &= -0,0439 \text{ Enero} - 0,0032 \text{ Marzo} - 0,0338 \text{ Abril} - 0,0043 \text{ Junio} - \\
 &\quad (-2,20) \quad (-1,95) \quad (-2,01) \quad (-2,58) \\
 &- 0,0047 \text{ Noviembre} - 0,2348 + 0,1577 \left| \frac{\varepsilon_{t-1}}{\sigma_{t-1}} \right| - 0,0549 \frac{\varepsilon_{t-1}}{\sigma_{t-1}} + 0,983323 \log(\sigma_{t-1}^2) \\
 &\quad (-2,42) \quad (-5,86) \quad (8,85) \quad (-5,47) \quad (259,6057)
 \end{aligned}$$

D.1.2) Distribución residual: GED

$$\begin{aligned}
 (1 - 0,047) \text{rigbm}_t &= 0,0014 \text{ Enero} + 0,0017 \text{ Febrero} + 0,0011 \text{ Noviembre} + \\
 &\quad (2,57) \quad (2,60) \quad (2,83) \quad (2,12) \\
 &+ 0,0013 \text{ Diciembre} + \varepsilon_t \sigma_t \\
 &\quad (2,20) \\
 \log(\sigma^2) &= -0,0582 \text{ Enero} - 0,0041 \text{ Marzo} - 0,0323 \text{ Abril} - 0,0536 \text{ Mayo} - 0,0536 \text{ Junio} - \\
 &\quad (-3,15) \quad (-2,75) \quad (-2,16) \quad (-3,52) \quad (-3,52) \\
 &- 0,0040 \text{ Agosto} - 0,0321 \text{ Septiembre} - 0,0313 \text{ Octubre} - 0,0589 \text{ Noviembre} - 0,2355 + \\
 &\quad (-2,52) \quad (-1,96) \quad (-1,98) \quad (-3,20) \quad (-6,19) \\
 &+ 0,1493 \left| \frac{\varepsilon_{t-1}}{\sigma_{t-1}} \right| - 0,054 \frac{\varepsilon_{t-1}}{\sigma_{t-1}} + 0,9832 \log(\sigma_{t-1}^2) \\
 &\quad (9,03) \quad (-5,73) \quad (275,13)
 \end{aligned}$$

Centrándonos en la especificación con mejor ajuste se ve que nuevamente corresponde a la t de Student y que el efecto mes es significativo tanto en la ecuación de la media como en la de la varianza. En la ecuación de la media todos los meses presentan una media positiva, pero significativamente distinta de cero tan sólo en enero, febrero, noviembre y diciembre. En la ecuación de la varianza sorprende el

resultado, pues todos los meses impactan negativamente y lo hacen de forma significativa los meses de enero, marzo, abril, junio y diciembre.

5. EVALUACIÓN DE LA CAPACIDAD PREDICTIVA DE LA VOLATILIDAD DEL IGBM.

Los modelos utilizados para la evaluación de las predicciones son los modelos ARIMA(1,0,0) y GARCH con distribución residual t-Student y el modelo ARIMA(1,0,0) y EGARCH con distribución residual t-Student

El análisis de capacidad predictiva con cada uno de los modelos se realiza de acuerdo con el siguiente esquema⁴:

- 1) Se estima el modelo con información hasta el 31 de diciembre de 2004.
- 2) Se generan predicciones con horizonte a 1, 2 ,3, 4, 5 y 10 períodos por delante.
- 3) Se añade una observación más al modelo y se vuelve a estimar y a generar nuevas predicciones.
- 4) Se calculan las series de errores de predicción a 1, 2 ,3 ,4, 5 y 10 períodos por delante obtenidas con las distintas predicciones realizadas.
- 5) Se calculan los estadísticos descriptivos de los errores de predicción. En particular, se calcula la *Raíz Cuadrada del Error Cuadrático Medio* (RECM) para comparar la precisión de las predicciones con los distintos modelos.

En total se han generado un total de 520 sendas de predicción que han supuesto 520 errores de predicción a un período, 519 a dos períodos y así sucesivamente hasta 5 errores de predicción a 10 períodos.

En los cuadros 1 y 2 se recoge información estadística de los errores de predicción a 1, 2, 3, 4, 5 y 10 períodos hacia delante de los modelos GARCH y EGARCH respectivamente, cabe destacar de ambos la raíz del error cuadrático medio (RECM),

⁴ Dada el gran número de sendas de predicción que se deben generar al trabajar con datos diarios, se ha elaborado un programa de EVIEWS que está a disposición de aquellas personas que lo soliciten.

que posteriormente en el cuadro 3 sirve para establecer una comparación en predicción entre ambos modelos.

Cuadro 1: Análisis estadístico de los errores de predicción a 1,2,3,4,5 y10 períodos hacia delante del modelo GARCH(1,1)

GARCH	Errores 1p	Errores 2p	Errores 3p	Errores 4p	Errores 5p	Errores 10p
Promedio	2.82E-05	1.82E-05	1.26E-05	-1.92E-06	-8.77E-06	-1.12E-05
ECM	3.22E-08	3.72E-08	3.72E-08	3.71E-08	3.56E-08	3.65E-08
RECM	0.000179457	0.00019299	0.00019287	0.00019259	0.00018856	0.00019114
RECM*1000	0.17945665	0.19299469	0.19287206	0.19258527	0.1885604	0.19113628

Fuente: Elaboración propia. ECM (error cuadrático medio), RECM (raíz cuadrada del ECM)

Cuadro 2: Análisis estadístico de los errores de predicción a 1,2,3,4,5 y10 períodos hacia delante del modelo EGARCH(1,1)

EGARCH	Errores 1p	Errores 2p	Errores 3p	Errores 4p	Errores 5p	Errores 10p
Promedio	3.58E-05	2.68E-05	2.15E-05	6.99E-06	8.32E-07	-5.33E-07
ECM	3.32E-08	3.69E-08	3.70E-08	3.68E-08	3.54E-08	3.62E-08
RECM	0.000182133	0.0001922	0.00019233	0.00019196	0.00018803	0.00019035
RECM*1000	0.182132701	0.19220162	0.19233001	0.19195757	0.18803238	0.19034795

Fuente: Elaboración propia. ECM (error cuadrático medio), RECM (raíz cuadrada del ECM)

El cuadro 3 recoge la comparación de los resultados en predicción en términos del ratio del RECM de los diferentes modelos asimétricos sobre el RECM del modelo GARCH (utilizado como referencia) para los distintos horizontes de predicción. Un ratio menor que la unidad significa que el modelo en cuestión mejora las predicciones

Cuadro 3. Ratio del RECM del modelo asimétrico sobre el RECM del modelo GARCH (1,1)

Horizonte de Predicción	1 período	2 períodos	3 períodos	4 períodos	5 períodos	10 períodos
EGARCH	1.01491165	0.995906524	0.997200187	0.996728802	0.997189224	0.995866904

Del análisis del cuadro 3 se observa que en relación con el modelo simétrico, la inclusión de los efecto estacional ocasionado por los días no consigue mejorar en

general la predicción de la volatilidad del modelo GARCH(1,1) a 1 período, mientras que a 2, 3,4,5 y 10 períodos el modelo EGARCH consigue mejorar ligeramente la predicción del modelo simétrico.

Por ultimo, para evaluar si las diferencias son estadísticamente significativas con respecto al modelo de referencia se aplica el contraste debido a Diebold y Mariano (1995) (DM de ahora en adelante). La hipótesis nula establece la igualdad de los dos procedimientos de predicción en términos de una función de pérdida cuadrática, $g(e)$, sobre sus errores, es decir, $E(d_t) = 0$, donde $d_t = g(e_{GARCH_t}) - g(e_{ASIM_t})$. El estadístico del contraste de Diebold y Mariano (1995) viene dado por:

$$DM = \frac{\bar{d}}{\sqrt{\text{var}(\bar{d})}}$$

donde \bar{d} es la media muestral de la serie d_t y $\text{var}(\bar{d})$ se calcula teniendo en cuenta la correlación serial en d_t . El estadístico DM sigue una distribución normal estándar bajo la hipótesis nula y su corrección por muestras finitas sigue una t de Student con $l-1$ grados de libertad, donde l es el número de predicciones. El estadístico DM, así como el t-valor asociado para los distintos modelos considerados se presenta en el cuadro 4.

Cuadro 4. Estadístico DM y t-valores asociados.

Horizonte de Predicción	C	t-Statistic
1 período	-1.01E-09	-0.9218
2 períodos	2.96E-10	0.6609
3 períodos	2.08E-10	0.588
4 períodos	2.35E-10	0.5527
5 períodos	1.93E-10	0.5081
10 períodos	3.01E-10	0.9748

El cuadro 4 muestra como los resultados de predicción obtenidos por los distintos modelos no presentan diferencias significativas.

6. CONCLUSIONES.

En este artículo se ha presentado un análisis del papel de las asimetrías en el IGBM teniendo en cuenta como primera aproximación el contraste de su existencia mediante la aplicación del contraste de Engle y Ng (1993). Se ha comprobado que existen tanto efectos de signo como de magnitud de los impactos de distinto signo en la serie del IGBM y por tanto, existen efectos asimétricos.

Además se han considerado dos aspectos adicionales. Primero la utilización de distintas distribuciones residuales en la estimación de los distintos modelos GARCH y EGARCH, eligiendo aquél que mejor AIC y BIC, en este caso se trata de la distribución t de Student y GED respectivamente, y el segundo la captación de efectos diarios y mensuales en el análisis de la volatilidad de los rendimientos del índice IGBM. Los modelos estimados muestran la pertinencia de la consideración de estas dos características, pero a nivel de evaluación de la capacidad predictiva, no se han observado diferencias significativas entre las alternativas consideradas.

Estos resultados confirman los obtenidos en anteriores trabajos, en el sentido en que el modelo GARCH(1,1) tiene mejor capacidad predictiva a un período hacia delante, mientras que el modelo asimétrico tiene mejor capacidad predictiva a dos, tres, cuatro, cinco y diez períodos hacia delante.

Sin embargo estas diferencias entre ambos tipos de modelos son muy pequeñas y no significativas. Por otra parte teniendo en cuenta que el contraste de simetría se decanta por el modelo asimétrico, creemos que este es el adecuado.

Distintos aspectos de este trabajo necesitarán ser tratados más a fondo en futuras versiones del mismo como la posibilidad de la existencia de un cambio estructural o el análisis de su relación con otros índices, lo que suscitaría la necesidad de una modelización multivariante.

7. BIBLIOGRAFÍA.

1. Andersen, T.G, and Bollerslev, T. (1998), “DM-Dollar Volatility: Intraday Activity Patterns, Macroeconomic Announcements and Longer-Run Dependencies. *Journal of Finance*, **53**, 2190-265.
2. Awartani, B.M.A. and Corradi, V. (2005), “Predicting the volatility of the S&P-500 stock index via GARCH models: the role of asymmetries”. *International Journal of Forecasting*, **21**, 167-183.
3. Black, F. (1976), “Studies of stock prices volatility changes”. *Proceedings of the 976 Meeting of the American Statistical Association, Business and Economic Statistics Section*, 177-181.
4. Bollerslev, T. (1986), “Generalized Autoregressive Conditional Heteroskedasticity”, *Journal of Econometrics*, **31**, 307-327.
5. Brooks, C. (1998), “Predicting stock market volatility: can market volume help?”, *Journal of Forecasting*, **17**, 59-80.
6. Brooks, C. (2002): *Introductory econometrics for finance*, Cambridge University Press.
7. Bluhm, H.H.W. and Yu, J. (2000), “Forecasting volatility: evidence from the German stock market”, *Working Paper, University of Auckland*.
8. Carnero, A., Peña, D. and E. Ruiz (2001), “Outliers and conditional autoregressive heteroscedasticity”, *Estadística*, **53**, 143-213.
9. Chen, C. and J. Liu (1993), “Joint estimation of model parameters and outlier effects in time series”, *Journal of the American Statistical Association*, **88**, 284-296.

10. Christie, A. A. (1982) "The Stochastic Behaviour of Common Stock Variances: Value, Leverage and Interest Effects", *Journal of Financial Economics*, **10**, 407-432.
11. Diebold, F. X. and Mariano, R. S. (1995), "Comparing Predictive Accuracy," *Journal of Business and Economic Statistics*, **13**, 253-263.
12. Glosten, L.R., Jagannathan, R. & Runkle, D.E. (1993). On the relation between the expected value and the volatility of the nominal excess return on stocks. *Journal of Finance*, **48**, 1779-801.
13. González-Rivera, G., Lee, T.H. & Mishra, S. (2004). Forecasting volatility: A reality check based on option pricing, utility function, value-at-risk, and predictive likelihood, *International Journal of Forecasting*, **20**, 629-645.
14. Engle, R.F. (1982). Autoregressive Conditional Heteroskedasticity with Estimates of Variance of U.K. Inflation. *Econometrica*, **50**, 987-1007.
15. Engle, R.F. and Ng, V.K. (1993). Measuring and Testing the Impact of News on Volatility, *Journal of Finance* **48**, 1749-78
16. Franses, P.H. and Ghijssels, H. (1999), "Additive outliers, GARCH and forecasting volatility", *International Journal of Forecasting*, **15**, 1-9.
17. French, K.R. (1980), Stock returns and the Weekend effect, *Journal of Financial Economics*, **8**(1), 55-69.
18. Gibbons, M.R. and Hess, P. (1981), Day of the Week effects an asset returns, *Journal of Business*, **54**, 579-96.
19. Hansen, P.R. and Lunde, A. (2003), A forecast comparison of volatility models: Does anything beat a GARCH(1,1), *Journal of Applied Econometrics*.
20. Jaffe, J., and Westerfield, R. (1985), Patterns in Japanese Common Stock Returns: Day of the week and turn of the year effects, *Journal of Financial and Quantitative analysis* **20**, 261-72.

21. Keim, D.B. and Stambaugh, R.F. (1984), A further investigation of the weekend effect in stock returns, *Journal of Finance*, 39, 819-835.
22. Loudon, G.F., Watt, W.H. and Yadav, P.K. (2000), “An empirical analysis of alternative parametric ARCH models”. *Journal of Applied Econometrics*, 2, 117-136.
23. Pagan, A.R. and Schwert, G.W. (1990) “Alternative Models for Conditional Stock Volatilities”, *Journal of Econometrics* 45, 267-90.
24. Pena, B. y Senra, I (2006). “Análisis de los datos atípicos y la asimetría en la predicción del Índice General de la Bolsa de Madrid (IGBM)”, Comunicación presentada en la Reunión Anual de Asepelt, La Laguna, 2006.
25. Poon, S.H y Granger C.W.J (2003) “Forecasting Volatility in Financial Markets, *Journal of Economic Literature*, Vol XLI (June 2003)
26. Ruiz, E (1994). “Modelos para series temporales heterocedásticas”.*Cuadernos Económicos del ICE* N° 56/1994/1.
27. Senra, I. y Pena, B. (2005), “Predicción de la volatilidad de los rendimientos del Índice General de la Bolsa de Madrid (IGBM): el papel de las asimetrías”, Comunicación presentada en la Reunión Anual de Asepelt, Badajoz, 2005.
28. Ruey S. Tsay (2002): *Analysis of Financial Time Series*, John Wiley & Sons, INC.

UN ESTUDIO COMPARATIVO DE METODOLOGÍAS MULTICRITERIO EN LA SELECCIÓN DE UN SISTEMA ERP PARA UNA EMPRESA DE PAQUETERÍA *

JOSÉ MARÍA MORENO JIMÉNEZ

e-mail: moreno@unizar.es

Grupo Decisión Multicriterio Zaragoza
UNIVERSIDAD DE ZARAGOZA

ADRIÁN TONCOVICH

e-mail: toncovic@unizar.es

Grupo Decisión Multicriterio Zaragoza
UNIVERSIDAD DE ZARAGOZA

NANCY LÓPEZ

e-mail: nblopez@uns.edu.ar

Departamento de Ingeniería
UNIVERSIDAD NACIONAL DEL SUR (ARGENTINA)

Área temática: Métodos Cuantitativos.

Resumen

Las metodologías multicriterio se han empleado muchas veces para resolver problemas de selección en distintas áreas de aplicación, tales como la administración pública, finanzas, planificación estratégica, etc. En este caso se presentan los resultados derivados de la aplicación de las metodologías AHP, ANP y MAUT al problema de selección de una herramienta ERP para una empresa dedicada a la distribución de paquetería urgente. Los programas informáticos utilizados al aplicar cada una de las tres técnicas multicriterio han sido respectivamente, Expert Choice, Super Decisions y Análisis Multiatributo Genérico (AMG). Finalmente se realiza un estudio comparativo de los resultados obtenidos con el fin de obtener conclusiones relativas a las ventajas e inconvenientes de cada uno de los enfoques utilizados.

Palabras clave: Decisión Multicriterio, Selección, AHP, ANP, MAUT, ERP.

Abstract

Multicriteria methodologies have been widely used in solving selection problems raised in different application areas, such as public administration, finance, strategic planning, etc. In this case the results derived from the application of AHP, ANP and MAUT methodologies to an ERP system selection problem faced by an express parcel distribution company are presented. The software used to implement each one of the three multicriteria techniques have been, respectively, Expert Choice, Super Decisions and Generic Multiattribute Analysis (AMG). Finally, a comparative study is performed with the aim of drawing conclusions regarding the advantages and shortcomings of each one of the approaches used in this case.

Key words: Multicriteria Decision Making, Selection Problem, AHP, ANP, MAUT, ERP.

* Este trabajo ha sido parcialmente sufragado por el proyecto del Ministerio de Educación y Ciencia "Herramientas Decisionales para la e-cognocracia" (ref. TSI2005-02511) y una beca del segundo autor financiada por la Universidad de Zaragoza y el Banco Santander Central Hispano.

1. Introducción

Las técnicas de decisión multicriterio permiten resolver de forma realista y efectiva de los problemas complejos que se plantean tanto en el dominio de las decisiones públicas como en el de las privadas. Estas técnicas se han venido aplicando en los diferentes ámbitos y actividades de la gestión.

En el caso concreto de las organizaciones empresariales, estas técnicas se han venido aplicando a las diferentes áreas funcionales. En lo que sigue, se analiza como se comportan en la selección de un sistema ERP para una empresa de paquetería urgente, dos de las aproximaciones multicriterio más extendidas: la Teoría de Utilidad Multiatributo (MAUT) y el Proceso Analítico Jerárquico (AHP). Para la primera utilizaremos el software *Análisis Multiatributo Genérico* (AMG), mientras que para la segunda emplearemos dos herramientas informáticas diferentes: por un lado, *Expert Choice* (EC) para abordar la resolución del problema con AHP (*analytic hierarchy process*), y por otra, *Super Decisions* (SD) para hacerlo conforme a la extensión de AHP conocida como Proceso Analítico Sistémico, ANP (*analytic network process*), que permite contemplar interdependencias entre los elementos considerados.

La selección de un ERP para cualquier empresa es una de las decisiones estratégicas que en gran parte van a condicionar su desarrollo futuro. Es por ello necesario que la resolución de esta cuestión se efectúe con las máximas garantías. En nuestro caso, tras un análisis del mercado, se han considerado seis alternativas, que van a ser evaluadas conforme a 17 atributos, agrupados en cinco criterios, lo que da lugar a una estructura jerárquica de tres niveles.

Los resultados reflejan que aún cuando las tres metodologías permiten modelar de forma adecuada el problema bajo análisis, las particularidades de cada una de ellas hacen que resulte más sencillo tener en cuenta algunas características del mismo (tratamiento de la incertidumbre en la emisión de juicios, la realización de análisis de sensibilidad, la modelización de relaciones de interdependencia y retroalimentación,...).

El trabajo se ha estructurado como sigue. La Sección 2 introduce los conceptos fundamentales relativos a las técnicas multicriterio que serán utilizadas en esta oportunidad. La Sección 3 presenta el problema que debe resolverse, partiendo de una descripción de las actividades de la empresa. La Sección 4 detalla la aplicación de las

tres técnicas multicriterio al problema de interés a través de las aplicaciones informáticas correspondientes. Por último, la Sección 5 enumera las conclusiones más relevantes del trabajo.

2. Técnicas de Decisión Multicriterio

Las técnicas de decisión multicriterio (TDM) son un conjunto de herramientas que permiten una resolución realista y efectiva de los problemas al considerar simultáneamente múltiples escenarios, actores y criterios, que pueden ser tanto tangibles como intangibles. Este tipo de técnicas pueden clasificarse (Cohon, 1978; Hwang y Masud, 1979; Zionts, 1978; Rietveld, 1980; Evans, 1984) conforme a numerosos criterios. Fijándonos en el flujo de información entre el analista y el decisor las técnicas de decisión multicriterio se clasifican en: (1) sin información a priori o generadoras, (2) con información a priori y (3) interactivas.

Dentro de las técnicas con información a priori y según cuál sea el conjunto de soluciones considerado estas técnicas pueden separarse, a su vez, en TDM continuas (programación multiobjetivo) y TDM discretas (programación multiatributo). En este último caso, las aproximaciones más conocidas son:

- i. Técnicas basadas en funciones de agregación: (a) directas (MAUT) y (b) jerárquicas (AHP).
- ii. Técnicas basadas en relaciones binarias de orden (ELECTRE, PROMETHEE).

En lo que sigue nos vamos a centrar exclusivamente en las técnicas del primer grupo.

2.1 Teoría de la Utilidad Multiatributo (MAUT)

El origen de MAUT está vinculado con el desarrollo de la Teoría de la Utilidad de Von Neumann y Morgenstern (1944), que fuera posteriormente extendida a problemas multicriterio por Keeney y Raiffa (1976) y Fishburn (1988). Esta metodología se basa en la agregación de los distintos criterios en una única función de utilidad, que debe optimizarse. Los métodos MAUT emplean funciones de utilidad para transformar los valores brutos (sin procesar) de rendimiento de las alternativas respecto de distintos criterios, tanto cuantitativos (objetivos) como cualitativos (subjetivos), a una escala común adimensional. Las funciones de utilidad desempeñan otro papel esencial, al transformar los valores brutos de rendimiento, de modo que un rendimiento más deseable tiene un mayor valor de la función de utilidad.

La idea fundamental del método radica en expresar las preferencias del decisor en términos de la utilidad que le reporta (principio de racionalidad). Para su desarrollo se introducen los conceptos de utilidad, independencia de atributos y loterías.

Desde el punto de vista metodológico, debe dedicarse un esfuerzo considerable a la obtención de una función de utilidad multiatributo. En primer lugar se debe introducir al decisor en la terminología, de forma que pueda transferir sus percepciones acerca del problema a la técnica de resolución. Luego se deben comprobar las condiciones de independencia e identificar la forma apropiada de descomposición de la función de utilidad. A continuación, se especifican las funciones de utilidad para cada atributo y se obtienen los pesos de cada criterio. Finalmente, se debe comprobar la consistencia de la función de utilidad obtenida.

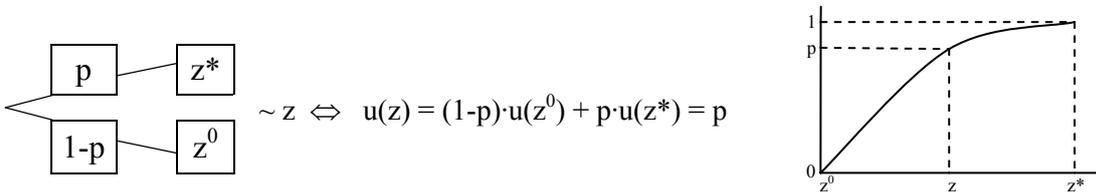
Con respecto a las condiciones de independencia, puede distinguirse una *independencia preferencial*, que está asociada a preferencias de tipo ordinal: el método hace una distinción cuando las preferencias que afectan a niveles de un conjunto de atributos no dependen de los niveles fijados para los demás atributos. También se establece una *independencia en utilidad*, que tiene vinculación con las preferencias de tipo cardinal: MAUT requiere que las preferencias en loterías que afectan a niveles de un conjunto de atributos no dependan de los niveles fijados para el resto. Así se definen los atributos mutuamente independientes en utilidad (MUI) de modo que cada subconjunto de atributos es independiente en utilidad de su complementario. Por último, se identifica una *independencia aditiva*: dos atributos se dicen aditivamente independientes si la comparación en preferencias pareadas de dos loterías cualesquiera, definida por la probabilidad conjunta, sólo depende de las distribuciones marginales.

Con relación a la forma funcional de la función de utilidad se han propuesto distintos esquemas de descomposición: aditivo, aditivo ponderado, multiplicativo y multilineal.

Obedeciendo a las condiciones de independencia se asignan las diferentes formas funcionales. En el caso de que las utilidades individuales de cada atributo fueran independientes con respecto al resto se utiliza una descomposición multilineal, cuando los atributos son MUI se emplea la descomposición multiplicativa y si existe independencia aditiva se utiliza la descomposición aditiva.

Para la obtención de la función de utilidad para cada atributo z , $u(z)$ se procede de la

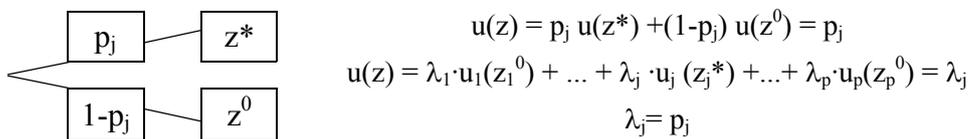
siguiente forma. Para el mejor valor posible z^* se fija un valor de utilidad $u(z^*) = 1$ y para el peor valor posible z^0 se asigna una utilidad $u(z^0) = 0$. Para el resto de valores $z \in (z^0, z^*)$ se plantea la equivalencia entre una lotería y la consecuencia segura z :



En el caso de la equivalencia en probabilidad se busca un valor de p y en el caso de la equivalencia en certidumbre se busca el nivel z del atributo.

En MAUT, los pesos reflejan la importancia relativa de los atributos en una situación dada. La determinación de los pesos implica la asignación de valores a las constantes de escala λ_j de la función de utilidad conjunta. Suponiendo una función de utilidad aditiva ponderada: $u(z) = \lambda_1 \cdot u_1(z_1) + \dots + \lambda_j \cdot u_j(z_j) + \dots + \lambda_p \cdot u_p(z_p)$, se consideran los puntos ideal ($z^* = (z_1^*, \dots, z_p^*)$; $u(z^*) = 1$) y anti-ideal ($z^0 = (z_1^0, \dots, z_p^0)$; $u(z^0) = 0$) como puntos de referencia, donde $z_j^* = \text{Max } z_j(x^i)$ y $z_j^0 = \text{Min } z_j(x^i)$, si los atributos son de máximo.

Para determinar la constante λ_j se busca la probabilidad p_j que haga al decisor indiferente entre $z = (z_1^0, \dots, z_j^*, \dots, z_p^0)$ y la lotería:



En lo que respecta al estudio de la consistencia en MAUT, una vez determinada la función de utilidad, aquélla se evalúa con las preferencias del decisor. Si existieran discrepancias, las mismas puede deberse a que el analista cometió algún error, el decisor cometió algún error, las actitudes del decisor han cambiado, o la situación ha cambiado.

2.2 El Proceso Analítico Jerárquico (AHP)

Los procesos analíticos jerárquico y sistémico fueron propuestos por Thomas Saaty (1980, 1996). En primer lugar desarrolló el proceso analítico jerárquico (AHP) y posteriormente propuso, como una extensión del mismo, el proceso analítico sistémico (ANP), para tener en cuenta las relaciones de interdependencia y retroalimentación.

La metodología AHP constituye uno de los métodos más ampliamente aplicados en la resolución de problemas de toma de decisiones multicriterio. AHP realiza una

modelización del problema que da lugar a la formación de una jerarquía representativa del esquema decisional asociado. Dicha jerarquía presenta en el nivel superior la meta que se persigue en la resolución del problema y en el nivel inferior se incluyen las distintas alternativas a partir de las que debe realizarse una selección. Los niveles intermedios se emplean para detallar el conjunto de criterios y atributos considerados en la toma de decisiones. Una de las ventajas de AHP reside en su habilidad para incorporar las preferencias de los decisores a través de la incorporación de criterios y atributos subjetivos. La evaluación de la jerarquía implica la realización de comparaciones pareadas de elementos de la jerarquía respecto del elemento del que dependen. El proceso de síntesis que utiliza esta información da lugar a las prioridades para cada una de las alternativas. El resultado final del método permite entonces seleccionar como la mejor alternativa, a aquella que tiene el valor de prioridad más elevado, o realizar un ordenamiento de alternativas que facilite, por ejemplo, la asignación de recursos. Además de estos resultados el método permite efectuar un análisis de sensibilidad, que facilita un mayor entendimiento del problema. Por otra parte, se puede efectuar un análisis del problema desde el punto de vista de la toma de decisiones grupal, a los fines de proporcionar un resultado global que aglutine las preferencias individuales de los actores interesados en la resolución del problema (Moreno-Jiménez, 2002).

La metodología AHP está basada en la realización de comparaciones pareadas del tipo ¿qué tan importante es el elemento C_i con relación al elemento C_j ?, que se realizan con el fin de establecer las ponderaciones para los criterios y evaluar el rendimiento de las alternativas respecto de los atributos incorporados en el planteamiento del problema.

Para obtener las ponderaciones asociadas a los criterios y atributos, se supone en primer lugar que los m criterios están organizados siguiendo una estructura jerárquica. Para cada par de criterios, se compara la importancia relativa entre ellos. Para ello, se utilizan juicios de la escala fundamental de Saaty (Saaty, 1980), donde los valores utilizados al comparar un elemento con otro significan: 1 (igual), 3 (moderada), 5 (fuerte), 7 (muy fuerte) y 9 (extrema). Los valores pares se utilizan para expresar valoraciones intermedias, los decimales para ganar precisión y los recíprocos para reflejar la mayor importancia del segundo elemento comparado.

Sea a_{ij} el valor obtenido de la comparación del criterio C_i con respecto al criterio C_j . Suponiendo que el decisor es consistente en la emisión de juicios sobre cualquier par de criterios y sabiendo, además, que todos los criterios siempre tendrán una misma evaluación al compararse a sí mismos, se tiene que: $a_{ij} \cdot a_{ji} = 1$ y $a_{ii} = 1$. Esto significa que solamente se deben efectuar $[\frac{1}{2} \cdot m \cdot (m - 1)]$ comparaciones a los efectos de determinar el conjunto completo de juicios para los m criterios. Los datos correspondientes proporcionan una matriz recíproca de comparaciones pareadas de dimensiones $m \times m$.

A continuación se debe estimar el conjunto de prioridades que refleje mejor los juicios volcados en la matriz de comparaciones pareadas \mathbf{C} . Debe notarse que la consistencia de la matriz \mathbf{C} , es decir $a_{ij} \cdot a_{jk} = a_{ik}$, $\forall i, j, k$, no está garantizada. En consecuencia, se debe buscar un vector de m componentes tal que la matriz $\mathbf{W}_{(m \times m)} = (w_{ij} = w_i/w_j)$ proporcione el mejor ajuste para los juicios registrados en la matriz \mathbf{C} . Se han propuesto diversas técnicas para conseguir este objetivo.

En el método original de Saaty, basado en el álgebra matricial, las ponderaciones vienen dadas por los elementos del autovector principal (EGV) asociado con el máximo autovalor de la matriz de juicios. El método EGV ha recibido diversas críticas, tanto desde la perspectiva de la determinación de las prioridades como desde el punto de vista de la consistencia, lo que ha provocado el desarrollo de numerosas variantes. Una de estas técnicas, el método de los mínimos cuadrados logarítmicos, da lugar a una forma directa de calcular el vector \mathbf{w} de prioridades locales. Para ello, se calcula la media geométrica de cada fila en la matriz \mathbf{C} , luego se calcula la suma de las medias geométricas, y se normaliza cada una de las medias geométricas dividiendo por esta suma (Saaty y Vargas, 1984).

En la mayor parte de las aplicaciones, los criterios suelen disponerse siguiendo una estructura jerárquica. En estos casos, AHP efectúa las comparaciones pareadas, bien de arriba hacia abajo (*top-down*) o de abajo hacia arriba (*botton-up*). AHP utiliza la misma técnica de comparaciones pareadas para establecer la calificación o importancia relativa de rendimiento para cada par de alternativas sobre cada uno de los atributos. Al comparar se emplea el mismo conjunto de juicios de la evaluación anterior (escala de Saaty), y las mismas técnicas pueden emplearse para calcular las prioridades de los elementos comparados.

Una vez que han quedado determinadas las ponderaciones y las calificaciones de rendimiento, y luego de realizar una posible normalización, las alternativas son evaluadas mediante técnicas de agregación similares a las que emplean los métodos MAUT. La variante aditiva utiliza la media aritmética ponderada que emplea MAUT, y la multiplicativa se funda, en esencia, en el cálculo de medias geométricas ponderadas.

2.2 El Proceso Analítico Sistemático (ANP)

ANP fue propuesto originalmente por Thomas Saaty con la finalidad de extender las capacidades de AHP al caso de situaciones en las que existe dependencia y retroalimentación entre los elementos del modelo que representa a un cierto sistema (Saaty, 1980, 1996).

El método ANP está compuesto por jerarquías de control, conglomerados (*clusters*, en inglés), nodos, interrelaciones entre nodos e interrelaciones entre conglomerados. Para facilitar la comprensión de la metodología, el proceso de modelización puede descomponerse en cuatro etapas que se describirán seguidamente.

En primer lugar, se realizan las comparaciones pareadas y se valoran las ponderaciones relativas. La jerarquía de control de ANP proporciona una base de comparación para cada tipo de interacción dentro de la red. Saaty (1996) planteó cuatro jerarquías de control básicas: Beneficios, Oportunidades, Costes y Riesgos (BOCR). No obstante, no es preciso incorporar todas las jerarquías de control en el modelo, todo depende de la relevancia de los criterios. La determinación de las ponderaciones relativas en ANP está basada en la realización de comparaciones pareadas al igual que en AHP. Las comparaciones pareadas, basadas en el principio de AHP, de los nodos en cada nivel se efectúan con respecto a su importancia relativa hacia su criterio de control, empleando, en general, la escala de Saaty. La calificación a_{ij} en la matriz de comparaciones pareadas representa la importancia relativa del elemento situado en la fila i -ésima respecto del componente situado en la columna j -ésima, es decir, $a_{ij} = w_i / w_j$.

Si existen n elementos que deben compararse, entonces la matriz \mathbf{A} se define como:

$$\mathbf{A} = \begin{bmatrix} 1 & a_{12} & \cdots & a_{1n} \\ 1/a_{12} & 1 & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 1/a_{1n} & 1/a_{2n} & \cdots & 1 \end{bmatrix}, \text{ que es una aproximación de: } \mathbf{W} = \begin{bmatrix} w_1/w_1 & w_1/w_2 & \cdots & w_1/w_n \\ w_2/w_1 & w_2/w_2 & \cdots & w_2/w_n \\ \vdots & \vdots & \ddots & \vdots \\ w_n/w_1 & w_n/w_2 & \cdots & w_n/w_n \end{bmatrix}$$

Siguiendo la idea fundamental de AHP, en lugar de realizar una asignación de las

ponderaciones w_i y w_j a los elementos i y j , se evalúa el peso relativo w_i / w_j para cada par de elementos. Una vez realizadas todas las comparaciones, el vector de prioridades \mathbf{w} se calcula a través de la solución única del sistema: $\mathbf{A}\mathbf{w} = \lambda_{\max}\mathbf{w}$, donde λ_{\max} es el autovalor principal de \mathbf{A} y $\sum_{i=1}^n w_i = 1$, la normalización modo distributiva que garantiza la unicidad (Saaty, 1980). AHP y ANP permiten valorar la inconsistencia en la emisión de juicios mediante el índice de consistencia (IC) que viene dado como: $IC = \frac{\lambda_{\max} - n}{n - 1}$. Si la razón de consistencia, $RC = IC/IC(A)$, donde $IC(A)$ es el índice de consistencia aleatorio (promedio) obtenido al simular juicios de \mathbf{A} en la escala de Saaty, es menor que 0,1, la consistencia suele considerarse aceptable (Saaty, 1980).

A continuación debe obtenerse la supermatriz inicial. La determinación de las ponderaciones relativas antes citadas está apoyada en las comparaciones pareadas al igual que en el caso del método AHP clásico. Luego, las ponderaciones se colocan dentro de la supermatriz que traduce las interrelaciones entre los elementos del sistema. El aspecto general de la supermatriz se ilustra en la Tabla 1 donde C_N indica el N-ésimo conglomerado, e_{Nn} indica el n-ésimo elemento en el N-ésimo conglomerado, y \mathbf{W}_{ij} es la matriz bloque conformada por los vectores de prioridades \mathbf{w} asociados a la importancia de los elementos en el i-ésimo conglomerado respecto del j-ésimo conglomerado.

Tabla 1: Configuración general de la supermatriz.

		C_1				C_2				...	C_N			
		e_{11}	e_{12}	...	e_{1n_1}	e_{21}	e_{22}	...	e_{2n_2}		...	e_{N1}	e_{N2}	...
C_1	e_{11}	\mathbf{W}_{11}				\mathbf{W}_{12}				...	\mathbf{W}_{1N}			
	e_{12}													
	\vdots													
	e_{1n_1}													
C_2	e_{21}	\mathbf{W}_{21}				\mathbf{W}_{22}				...	\mathbf{W}_{2N}			
	e_{22}													
	\vdots													
	e_{2n_2}													
\vdots	\vdots	\vdots				\ddots				\vdots				
C_N	e_{N1}	\mathbf{W}_{N1}				\mathbf{W}_{N2}				...	\mathbf{W}_{NN}			
	e_{N2}													
	\vdots													
	e_{Nn_N}													

Luego, debe generarse la supermatriz ponderada. La supermatriz inicial está constituida por distintos autovectores cada uno de los cuales suma uno. La supermatriz inicial debe transformarse en una matriz donde cada una de sus columnas sume uno. A los efectos

de reducir la suma de las columnas a la unidad cada uno de los elementos incluidos en un bloque de la supermatriz se multiplica por la ponderación asociada al conglomerado.

El autovector obtenido a partir de la comparación del nivel conglomerado con respecto al criterio de control se emplea como ponderaciones de los conglomerados. Esto da lugar a una matriz estocástica por columnas. Si algún bloque de la supermatriz tiene una columna con todos sus elementos nulos, debe normalizarse usando las ponderaciones de los conglomerados para asegurar que la suma de la columna sea igual a la unidad.

Finalmente, se tienen que calcular de los vectores y ponderaciones de prioridad globales. La supermatriz ponderada se eleva (Saaty, 1996) a potencias sucesivas con el fin de obtener los vectores de prioridad globales: $\lim_{k \rightarrow \infty} \mathbf{W}^k$.

3. Selección de un sistema ERP

En este trabajo se ha considerado un problema asociado a la actividad de una empresa dedicada a la distribución y transporte de paquetería. Dicho problema surgió en el contexto del proceso de reestructuración de las operaciones de negocio desarrollado con el propósito de lograr una adecuada alineación a sus necesidades estratégicas de crecimiento. Durante este proceso se identificó la necesidad de actualizar el sistema informático empleado en la gestión de las diversas funciones y áreas de negocio. Dicha renovación implica la sustitución del actual sistema informático, que ha quedado obsoleto, y la implantación de una nueva aplicación informática, sistema ERP (*Enterprise Resource Planning*), que facilite una gestión integrada de las actividades. A tal fin, la dirección decidió llevar a cabo un proceso de selección que permitiera identificar entre las alternativas disponibles cuál es aquella que se adapta mejor a las características de su actividad y que representa una solución adecuada desde el punto de vista de los distintos actores participantes en el proceso de toma de decisiones.

Para mayores detalles, la empresa analizada se dedica al transporte de paquetería en el contexto regional de Aragón y regiones vecinas (La Rioja y Navarra). La principal actividad está centrada en el transporte de documentación de diversas entidades bancarias (Cajas rurales y bancos propiamente dichos), que tiene un patrón muy claro de funcionamiento. Así mismo, siguiendo una estrategia orientada a conseguir una fidelización de los clientes, la empresa realiza todo tipo de servicio logístico que pueda efectuarse aprovechando la red de que dispone.

La compañía ha tenido una evolución irregular hasta hace algunos años, pero a partir de la adquisición del paquete accionario principal por parte de su actual dueño, se ha iniciado un proceso de saneamiento, consolidación, y actualmente expansión, que requiere una redefinición de estrategias y tácticas de negocio, así como también una revisión de las bases operativas sobre las que se apoya su proceso de negocio. En este contexto y luego de cierto análisis, ha quedado patente la necesidad de modernizar los sistemas de información de la empresa, no solamente para proporcionar un mejor servicio al cliente, sino también para favorecer el intercambio de información y mejorar la gestión global dentro de la empresa.

Aunque un análisis preliminar del problema no necesariamente llevaría a considerar la adquisición de la solución integral asociada a un sistema ERP, las mejoras sustantivas que pueden generarse a partir de su implantación, unido a la complejidad creciente del negocio y las perspectivas futuras de crecimiento, hicieron recomendable realizar una inversión de este tipo. Sin embargo, la elección de una aplicación ERP para una empresa no constituye, en la práctica, una decisión directa y requiere un análisis cuidadoso que evite errores costosos en el largo plazo. Por ello, la empresa decidió emprender este proceso utilizando una metodología racional que sirviera de apoyo en la adopción de la mejor decisión. Dicha metodología incluyó la aplicación de las tres técnicas multicriterio discretas antes citadas. En lo que sigue se resumirán los principales aspectos del procedimiento realizado.

En primer lugar se constituyó un grupo de trabajo que tuvo a su cargo la ejecución del proceso decisional asociado. Dicho grupo estuvo formado por cinco personas de los niveles jerárquicos superiores, provenientes de distintos sectores de la empresa: el responsable económico (nivel gerencial), el director de operaciones, el responsable de administración, el responsable de nuevas tecnologías y el adjunto de logística.

Como siguiente paso, a partir del trabajo grupal, se identificaron de forma consensuada seis criterios principales y 17 atributos relevantes para la decisión, que se detallan en la Tabla 2, representativos del funcionamiento global del sistema, que traducen adecuadamente las necesidades de cada una de las áreas clave de gestión.

En la siguiente fase el equipo decisor realizó una exploración del mercado con la intención de recopilar suficiente información acerca de la oferta existente, e identificar

en las herramientas disponibles los elementos que resultan de interés de cara a la meta establecida. La conformación de un grupo de trabajo multidisciplinar resulta clave en la definición de las características que debe poseer la herramienta. El trabajo en equipo debe permitir arribar a una clara definición de los elementos que darán forma a la solución ideal a la que deberá acercarse la alternativa que será finalmente seleccionada.

Tabla 2: Criterios y atributos del problema.

CRITERIO	ATRIBUTO	DESCRIPCIÓN
Económico	Adquisición	Asociado a los costes de licencias y producto.
	Mantenimiento	Coste anual, importe y descuentos.
Administrativo	Facturación	Sistema, procesos, gestión de clientes, centros de coste, etc.
	Tarifas	Estructura afín a las necesidades.
	Gestión de costes	Características, enlaces contables, informes, etc.
	Gestión documental	Vinculada a procesos administrativos (digitalización de documentos, documentación de chóferes, vehículos, etc.).
Logístico	Almacén	Gestión global (costes, tiempos, etc.) desde la llegada de mercancía a las instalaciones hasta la salida de bultos para su distribución.
	Tráfico	Gestión del transporte de mercancías y rutas.
	Trazabilidad	Rastreabilidad de cada expedición desde el instante de recogida hasta el de entrega.
	Gestión de incidencias	Manejo de los fallos derivados de las operaciones con clientes internos, externos y proveedores.
Tecnológico	Plataformas	Gestión de bases de datos y arquitectura de sistemas.
	Accesibilidad	Capacidad de importar y exportar datos, factibilidad de desarrollar aplicaciones propias e integrarlas al sistema.
	Potencial de desarrollo	Capacidad de incorporar nuevas prestaciones (gestión de almacenes múltiples, gestión avanzada de delegaciones) generadas por la inclusión de nuevos servicios y clientes.
	Comunicación y trazabilidad	Fiabilidad, prestaciones y formas de actualizar los sistemas tecnológicos implicados.
Comercial	Interfaz	Interacción y comunicación con clientes y colaboradores.
	Portal corporativo	Desarrollo y gestión de sitio web de representación y contacto.
	Gestión comercial	Sistema de gestión de las relaciones con los clientes. Administración eficiente de la documentación comercial (propuestas, presupuestos, aceptaciones, etc.).

Como resultado del análisis anterior se determinó un conjunto de alternativas que verificaban los requisitos esenciales exigidos a una alternativa de solución para el problema. Una vez determinado este grupo inicial, se procedió a recabar más información, para adquirir un mayor conocimiento en el que apoyar el proceso de selección en las etapas siguientes. Esta profundización se consiguió esencialmente por medio del contacto directo con los proveedores y la realización de presentaciones y demostraciones que proporcionaron una idea bastante clara acerca del funcionamiento real de los sistemas y las posibles ventajas y desventajas de su implantación en la empresa. Las alternativas seleccionadas en esta etapa se muestran en la Tabla 3.

Tabla 3: Sistemas ERP considerados en el estudio.

Nombre de la aplicación	Proveedor
Aril	Logistec
WinTrans	Itranser
QTrans	Advantur Technologies, S.L.
Alertran - Tamesis	Alerce, Informática aplicada, S.A.
Ekon	CCS Agresso
Iris	Redyser

Seguidamente se realizó una valoración de la importancia relativa de los criterios y atributos y se calificó cada una de las alternativas respecto de cada uno de ellos. En esta etapa se estimaron las ponderaciones correspondientes a los criterios y atributos considerando su aportación a la meta. La ejecución de esta actividad estuvo a cargo de las personas más implicadas con el proyecto y que al mismo tiempo poseen una perspectiva amplia del negocio de modo que resultan capaces de percibir cuáles son los factores que contribuyen a la mejora global de la empresa en el medio y largo plazo.

La información recopilada en las etapas anteriores fue empleada como entrada de las tres metodologías multicriterio citadas, implementadas por medio de aplicaciones específicas, que dieron lugar a tres ordenaciones de las alternativas.

4. Aplicación de las Técnicas de Decisión Multicriterio

Seguidamente se expondrán los detalles relativos a la resolución del problema introducido en la sección anterior mediante las tres aplicaciones informáticas que tienen implementadas las técnicas multicriterio consideradas en este caso.

4.1 Utilización AMG para la selección de la herramienta ERP

La aplicación del método MAUT se llevó a cabo mediante el software AMG (Análisis Multiatributo Genérico). Dicha herramienta fue desarrollada en el Departamento de Inteligencia Artificial de la Facultad de Informática de la Universidad Politécnica de Madrid, bajo la responsabilidad de los profesores Sixto Ríos Insúa, Alfonso Mateos Caballero y Antonio Jiménez Martín (www.dia.fi.upm.es/~ajimenez/GMAA).

La creación de un espacio de trabajo en AMG implica la ejecución de cinco pasos: (1) estructuración del problema (construcción de la jerarquía decisional), (2) identificación de alternativas factibles, consecuencias e incertidumbre, (3) cuantificación de preferencias, (4) evaluación de alternativas, y (5) análisis de sensibilidad.

La Figura 1 muestra el aspecto de la jerarquía asociada al problema. Para el procesamiento de alternativas, primero se introducen los rangos y unidades de los

critérios vinculados con los nodos de menor nivel dentro de la jerarquía, luego, se cuantifican las utilidades individuales en los criterios y por último, deben incorporarse las alternativas y sus consecuencias para los distintos criterios. La Figura 2 (izquierda) muestra las alternativas y sus consecuencias para el problema que se analiza.

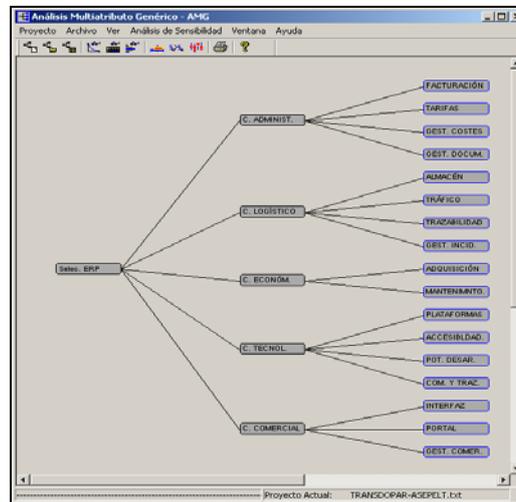


Figura 1: Vista de Jerarquía decisional en AMG.

La incertidumbre sobre las consecuencias o impactos de las alternativas puede considerarse con ayuda de intervalos de valores uniformemente distribuidos.

La cuantificación de las preferencias implica calcular las utilidades individuales en los atributos y la importancia relativa de los objetivos de la jerarquía, valiéndose de ponderaciones. Esta información se utiliza para evaluar las alternativas a través de una función de utilidad aditiva. AMG permite que el decisor introduzca información incierta por medio de intervalos de valores en respuesta a ciertas cuestiones probabilísticas.

En AMG existen tres formas de introducir y modificar las utilidades individuales de cualquier atributo: construir una función de utilidad lineal a trozos imprecisa, asignar utilidades imprecisas para valores discretos del atributo o utilizar una escala subjetiva, en la que los valores se traducen a utilidades directamente. En el caso analizado se consideraron únicamente atributos del segundo tipo. La Figura 2 (derecha) detalla la función de utilidad resultante para el atributo Gestión de Costes.

Con el fin de construir la función de utilidad aditiva se requiere conocer las contribuciones de cada uno de los criterios y atributos a la utilidad total de las alternativas, que vienen dadas por las ponderaciones respectivas (Figura 3, izquierda).

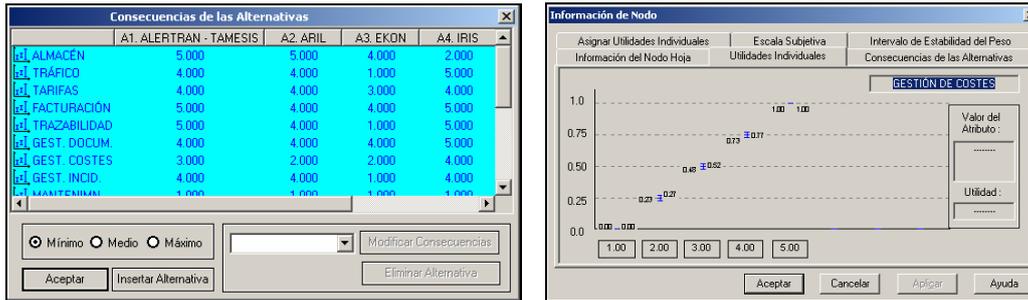


Figura 2: Alternativas y consecuencias para los diferentes atributos y asignación de consecuencias.

AMG permite incorporar imprecisión en las ponderaciones, por lo tanto, para cada uno de los criterios y atributos se obtendrán, en general, ponderaciones medias normalizadas e intervalos de ponderaciones normalizadas. En la Figura 3 (derecha) se muestra el caso de los nodos incluidos en el criterio logístico. La asignación de ponderaciones se inicia con los atributos del problema y prosigue hacia arriba cubriendo toda la jerarquía.

Las ponderaciones de los atributos se obtienen multiplicando las ponderaciones medias normalizadas y las extremas de los intervalos de ponderaciones normalizadas locales que se encuentran en el camino desde el objetivo global hasta el atributo considerado.

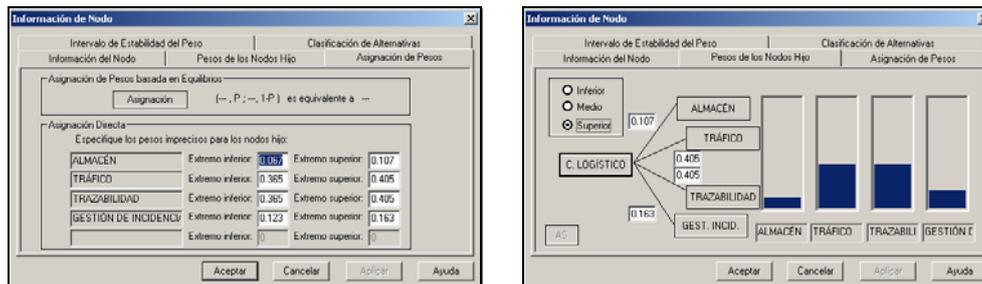


Figura 3: Asignación y visualización de las ponderaciones.

Una vez que se han calificado todas las alternativas, AMG permite evaluarlas por medio de una función de utilidad aditiva ponderada. De esta forma se obtienen por un lado, las utilidades globales medias, que determinan la ordenación de alternativas, y, por otro, las mínimas y máximas, que permitirán evaluar la estabilidad de esta ordenación.

Los resultados de la evaluación para el problema analizado se ven en la Figura 4. Allí se muestran los intervalos de utilidades globales y la ordenación correspondiente de las alternativas. Los segmentos amarillos indican las utilidades medias.

El solapamiento de los intervalos puede dificultar la adopción de una decisión. Por tanto, debe recurrirse a análisis de sensibilidad ulteriores que permitan extraer más información y faciliten, en lo posible, la obtención de una alternativa dominante.



Figura 4: Ordenación de las alternativas.

La aplicación facilita la presentación de los resultados mediante diversos diagramas. Por ejemplo, la Figura 5 (izquierda) desglosa la contribución de cada uno de los atributos a la función de utilidad para cada alternativa. En la misma figura (derecha) aparece retratado el gráfico de comparación de alternativas.

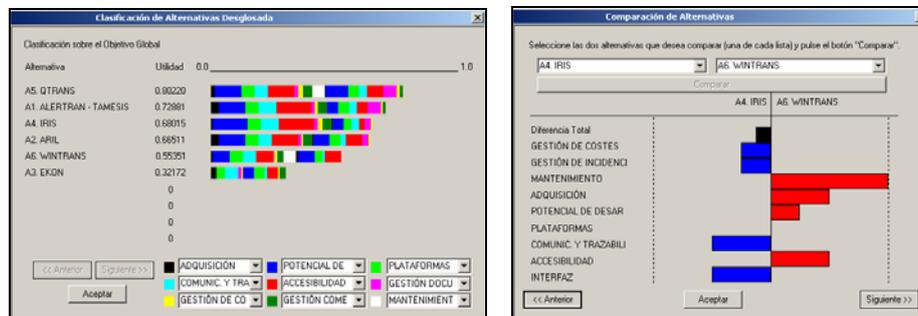


Figura 5: Desglose de la contribución de los atributos a la utilidad global y comparación de alternativas.

Puede efectuarse un análisis de sensibilidad simple alterando los parámetros y registrando el impacto sobre la ordenación de las alternativas. No obstante, este análisis no es lo suficientemente sistemático como para asistir en el proceso de toma de decisiones. Mediante un procedimiento apropiado (Ríos Insua y French, 1991) se pueden rechazar de plano malas estrategias descartando alternativas dominadas y/o no óptimas y definir una estrategia dominante (Figura 6).



Figura 6: Pantalla de alternativas no dominadas y potencialmente óptimas y detalles.

Por medio de una estrategia de simulación se puede efectuar un análisis de sensibilidad que actúa simultáneamente sobre las ponderaciones, dando lugar a resultados que pueden analizarse estadísticamente para generar conocimiento y proponer

recomendaciones sobre el modelo multicriterio. Como caso extremo, las ponderaciones para los criterios pueden generarse de forma totalmente aleatoria. La Figura 7 (izquierda) muestra los resultados de un experimento de simulación de 10000 réplicas y las medidas estadísticas asociadas (derecha).

Además, AMG permite generar aleatoriamente las ponderaciones manteniendo el orden asociado a la importancia del criterio o a partir de los intervalos definidos teniendo en cuenta el error de la estimación. Esto impone restricciones importantes sobre el dominio de ponderaciones, que resultan consistentes con los juicios del decisor.

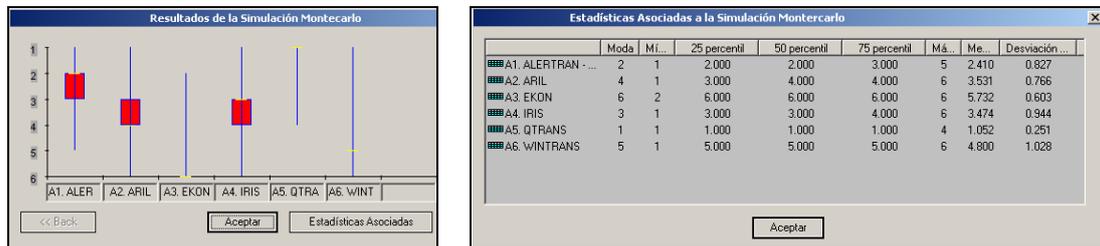


Figura 7: Resultado y análisis estadístico de la simulación con ponderaciones aleatorias.

4.2 Resolución mediante Expert Choice

La herramienta Expert Choice fue utilizada para aplicar el método AHP al problema en estudio. La aplicación de AHP a un problema particular implica, en primer lugar, la construcción de la jerarquía de elementos que lo represente adecuadamente. En este caso, en el nivel superior se coloca el nodo meta que viene dado por la determinación de la herramienta ERP que mejor se adapte a las necesidades de la empresa. Luego deben determinarse los criterios y atributos que dependen del nodo meta. La Figura 8 muestra la jerarquía decisional asociada al problema, descrita mediante una pantalla capturada del programa Expert Choice, así mismo aparecen representadas en la misma figura las alternativas de decisión identificadas. Una vez diseñada la jerarquía del problema, debe iniciarse el proceso de valoración.

En una primera instancia debe realizarse una evaluación de la importancia de los criterios con relación a su contribución a la consecución de la meta y luego respecto de cada criterio debe determinarse cuál es la importancia relativa de los atributos que dependen de él. El proceso de valoración debe proseguir con la apreciación de la importancia de las alternativas respecto de cada uno de los atributos existentes. En el caso de los criterios y atributos, las valoraciones se realizan a través de comparaciones de pares de elementos respecto del elemento del que dependen en la jerarquía.

Dichas comparaciones se efectúan teniendo en consideración los valores dados por la escala de Saaty. La Figura 9 muestra las comparaciones pareadas asociadas a los criterios respecto de su contribución a la meta. En este caso se hace uso de la escala numérica 1-9, pero alternativamente podría haberse utilizado un formulario gráfico o verbal equivalente o, incluso, se podrían haber ingresado directamente los valores de prioridad.

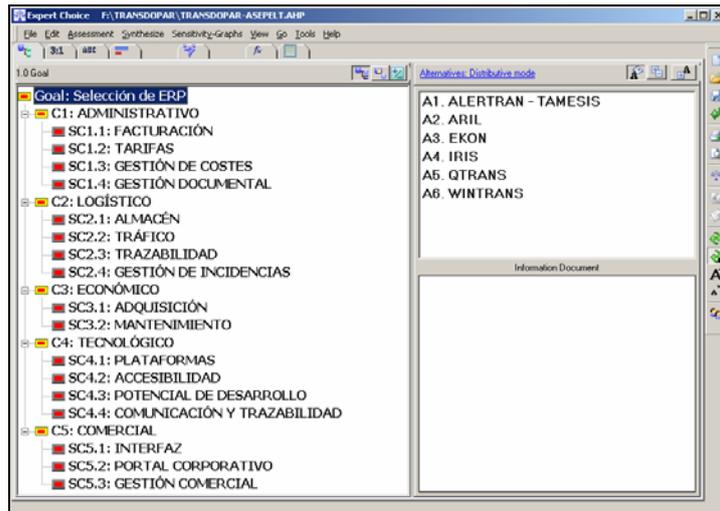


Figura 8: Jerarquía decisional.

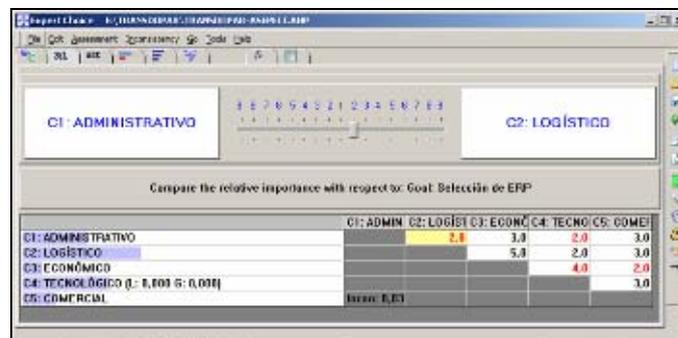


Figura 9: Comparaciones pareadas de los criterios respecto de su contribución a la meta.

Una vez concretado el proceso de valoración para los elementos superiores de la jerarquía, se deben evaluar las alternativas respecto de cada uno de los atributos identificados. Dicha valoración puede efectuarse mediante el procedimiento de comparaciones pareadas, o, en su lugar, puede utilizarse el método de las medidas relativas, que, en esencia, consiste en completar una tabla donde se vuelcan las valoraciones asociadas a las alternativas para cada uno de los atributos, utilizando una escala de medida predeterminada. En este caso se ha utilizado esta segunda opción, que aparece recogida en la Figura 10, donde puede apreciarse que cada alternativa ha sido evaluada en una escala del 0 al 1 que proporciona directamente un valor de prioridad.

Culminada la etapa de valoración, corresponde realizar la priorización y síntesis de resultados con la finalidad de determinar una ordenación para las alternativas, a partir de las prioridades globales resultantes de cada una de ellas (Figura 11).

Alternative	DIRECT C1: ADMINISTRATIVO SCI.1: FACTURACIÓN (L.:.385 G.:.075)	DIRECT C1: ADMINISTRATIVO SCI.2: TARIFAS (L.:.385 G.:.075)	DIRECT C1: ADMINISTRATIVO SCI.3: GESTIÓN DE COSTES (L.:.143 G.:.029)	DIRECT C1: ADMINISTRATIVO SCI.4: GESTIÓN DOCUMENTAL (L.:.087 G.:.017)	DIRECT C2: LOGISTICO SCI.1: ALMACEN (L.:.087 G.:.033)
A1 ✓A1 ALERTRAN	1	.75	.5	.75	1
A2 ✓A2 APRIL	.75	.75	.25	.75	1
A3 ✓A3 EKON	.75	.5	.25	.75	.75
A4 ✓A4 IRIS	1	.75	.75	1	.25
A5 ✓A5 QTRANS	.75	.75	.75	1	.5
A6 ✓A6 WINTRANS	.75	.75	.5	.25	.25

Figura 10: Valoraciones de las alternativas respecto de los atributos de decisión.

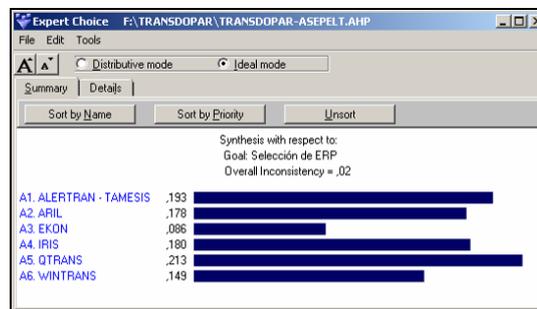


Figura 11: Prioridades globales de las alternativas.

A continuación, puede realizarse un análisis de sensibilidad para determinar si los resultados generados son lo suficientemente estables frente a variaciones de los parámetros como para recomendar la selección de una alternativa en particular. La Figura 12 y la Figura 13 muestran las pantallas correspondientes a las opciones que permiten realizar el análisis de sensibilidad mediante EC. Al aplicar estas herramientas pudo comprobarse que la alternativa mejor posicionada permanece en ese estado si el conjunto de parámetros del modelo se mantiene dentro de un entorno cercano al actual. Para que se altere la situación deben producirse cambios abruptos que no son factibles en el contexto de un proceso de toma de decisiones racionalmente dirigido.

4.3 Utilización de Super Decisions para la selección de la herramienta ERP

La aplicación de la metodología ANP al problema de selección de la herramienta ERP se realizó a través del programa Super Decisions, cuyo diseño se debe a William Adams y a la Fundación Creative Decisions, fundada por Thomas y Rozann Saaty. La aplicación de esta herramienta requiere primeramente la definición de una red asociada al problema que muestre todas las interrelaciones y dependencias que existen entre los componentes identificados, relevantes desde el punto de vista de la toma de decisiones.

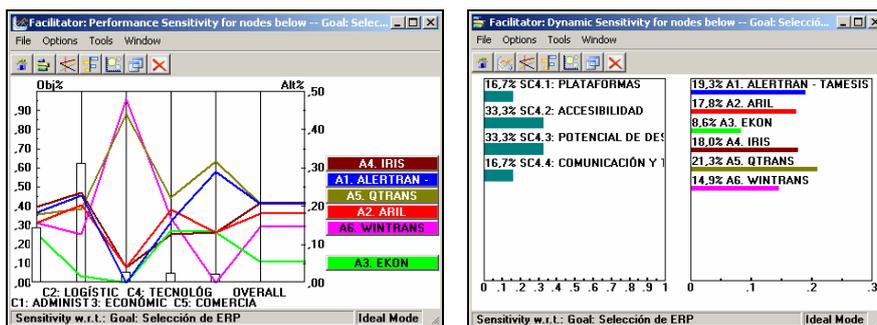


Figura 12: Análisis de sensibilidad del rendimiento de las alternativas y análisis de sensibilidad dinámico

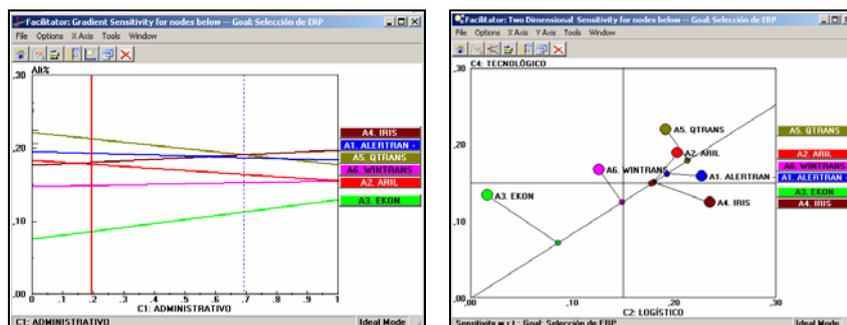


Figura 13: Análisis de sensibilidad del gradiente y análisis de sensibilidad bidimensional.

La Figura 14 muestra la red asociada al problema bajo estudio, en la que se aprecian las relaciones de interdependencia entre los elementos del sistema. Recuadrados en rojo se muestran los nodos que dependen del nodo trazabilidad. En particular, se ha considerado la importancia que tiene cada uno de los atributos respecto de los restantes atributos del modelo. Además, se ha tenido en cuenta la trascendencia que tienen los distintos conglomerados incluidos en la red desde la perspectiva del logro de la meta.

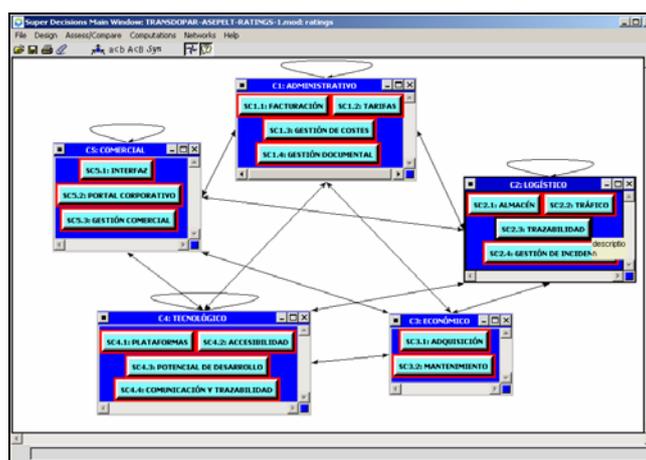


Figura 14: Modelo de red con visualización de conexiones.

Una vez construida la red del problema se debe iniciar el proceso de evaluación que permite obtener, mediante una síntesis de la información introducida, una definición de prioridades para las alternativas. Dicho proceso implica la realización de comparaciones entre pares de elementos del modelo, respecto de otro elemento del mismo.

Como primer paso se deben realizar las comparaciones asociadas a los conglomerados. A modo de ejemplo, la Figura 15 exhibe las comparaciones pareadas de los conglomerados del modelo con respecto al conglomerado Criterio Logístico. Así mismo, la aplicación facilita el cálculo de las prioridades normalizadas e ideales correspondientes, y de los índices de inconsistencia.

Figura 15: Ejemplo de comparaciones entre conglomerados.

Posteriormente, deben realizarse las comparaciones pareadas de los nodos incluidos dentro de cada uno de los conglomerados respecto de cada uno de los nodos con los que tienen vinculación. Por ejemplo, la Figura 16 muestra las comparaciones realizadas entre los nodos correspondientes al conglomerado del Criterio Administrativo, respecto del nodo Trazabilidad, que pertenece al conglomerado Criterio Logístico. Las comparaciones permiten identificar, en este caso, cuáles de los atributos del criterio Administrativo resultan más relevantes en la evaluación del atributo Trazabilidad.

Finalmente, se deben introducir las valoraciones de cada una de las alternativas respecto de los atributos considerados. Para ello se ha utilizado nuevamente el método de las medidas relativas. La Figura 17 muestra parcialmente la tabla asociada a dicho método.

Figura 16: Ejemplo de comparaciones entre atributos, en este caso respecto del atributo trazabilidad.

Terminada la fase de introducción de juicios, SD tiene información suficiente para realizar los cálculos que darán lugar a una asignación de prioridades para las alternativas, en las que se apoyará el decisor para tomar una decisión apropiada.

	SC1.1: ADQUISICIÓN 0.010089	SC1.2: MANTENIMIENTO 0.050444	SC1.1: FACTURACIÓN 0.058122	SC1.2: TARIFAS 0.058122	SC1.3: GESTIÓN DE COSTES 0.026360	SC1.4: GESTIÓN DOCUMENTAL 0.015202
A1. ALERTRAN	EXCESIVO	EXCESIVO	MUY BUENO	BUENO	REGULAR	BUENO
A2. ARIL	ACEPTABLE	EXCESIVO	BUENO	BUENO	MALO	BUENO
A3. EKON	EXCESIVO	EXCESIVO	BUENO	REGULAR	MALO	BUENO
A4. IRIS	ACEPTABLE	EXCESIVO	MUY BUENO	BUENO	BUENO	MUY BUENO
A5. QTRANS	ACEPTABLE	BAJO	BUENO	BUENO	BUENO	MUY BUENO
A6. WINTRANS	BAJO	BAJO	BUENO	BUENO	REGULAR	MALO

Figura 17: Valoraciones de las alternativas respecto de cada uno de los atributos del problema.

SD permite visualizar en primera instancia la supermatriz no ponderada de los juicios introducidos anteriormente. A partir de esta supermatriz no ponderada se calculará la supermatriz de juicios ponderada recurriendo para esto a la matriz conglomerado. Esta última se obtiene a partir de las comparaciones realizadas entre los conglomerados (Figura 18). En efecto, cada uno de los elementos de la supermatriz ponderada de juicios será resultado del producto entre el elemento de la matriz no ponderada de juicios y la ponderación perteneciente a la matriz conglomerado que le corresponde. De esta manera, la supermatriz resultante será estocástica por columnas. Elevando dicha supermatriz a potencias sucesivas se obtendrá la supermatriz límite, a partir de la que se extraen las prioridades correspondientes a cada uno de los nodos de la red, que intervendrán en el cálculo de las prioridades de las alternativas (Figura 19).

Cluster Name	C1-ADMINISTRATIVO	C2-LOGÍSTICO	C3-ECONÓMICO	C4-TECNOLOGICO	C5-COMERCIAL
C1-ADMINISTRATIVO	0.24225	0.12899	0.16296	0.17999	0.12799
C2-LOGÍSTICO	0.40999	0.28994	0.04292	0.30499	0.20499
C3-ECONÓMICO	0.06294	0.05429	0.00099	0.05429	0.05429
C4-TECNOLOGICO	0.20999	0.23999	0.20999	0.27999	0.27999
C5-COMERCIAL	0.09299	0.12499	0.00099	0.09499	0.12499

Figura 18: Pantalla correspondiente a la matriz conglomerado.

Las prioridades obtenidas para las alternativas corresponden a valores brutos, que pueden ser normalizados utilizando la suma de valores brutos de prioridades de las alternativas, o bien, a partir de ellos se pueden calcular valores de prioridad ideales, dividiendo el valor bruto de prioridad de cada alternativa por el mayor valor bruto de prioridad de todas las alternativas (Figura 20).

Se observa que la ordenación obtenida es análoga a las determinadas por AMG y EC. De esta forma, sobre la base de los resultados generados por SD, debería seleccionarse la alternativa QTrans, que ha conseguido un mayor valor de prioridad global.

Figura 19: Vista parcial de la supermatriz límite.

Con el fin de evaluar la estabilidad de los resultados ante variaciones de los parámetros del modelo, la aplicación permite realizar diversos análisis de sensibilidad. En este caso, al haber introducido las valoraciones con el método de medidas relativas no pueden obtenerse resultados de utilidad práctica.

Name	Graphic	Ideals	Normals	Raw
A1. ALERTRAN - TAMESIS		0.907469	0.200827	0.200827
A2. ARIL		0.798969	0.176815	0.176815
A3. EKON		0.393650	0.087116	0.087116
A4. IRIS		0.836645	0.185153	0.185153
A5. QTRANS		1.000000	0.221304	0.221304
A6. WINTRANS		0.581932	0.128784	0.128784

Figura 20: Prioridades finales globales de las alternativas.

4.4 Análisis de los resultados

Según puede comprobarse a partir de las ordenaciones de las alternativas obtenidas por medio de las tres técnicas aplicadas (Tabla 4), se han reproducido los mismos resultados, con variaciones insignificantes entre los valores de prioridad. Por otra parte, los análisis de sensibilidad realizados mediante AMG y EC permiten asegurar que existe suficiente estabilidad en la ordenación sugerida, y que la misma no se verá afectada por cambios acotados de las variables del modelo. Podría argüirse que la identidad de los resultados viene dada por la perspectiva utilizada por el analista, que tiende a realizar una adaptación de la metodología aplicada antes a la nueva y consecuentemente, no realiza una abstracción de las circunstancias pasadas, dando lugar a una repetición de resultados. No obstante, en este caso el problema presenta una estructura de preferencias claramente definida y, en consecuencia, resulta razonable que, a pesar de utilizar distintas técnicas, se arribe a una misma solución.

Distinto sería el caso si se tuviera un problema más indefinido que diera mayor juego en la formulación del modelo. En esta última situación es evidente que las particularidades que tiene cada una de las metodologías darían lugar a formulaciones más variadas y, en consecuencia, los resultados asociados deberían presentar, de igual modo, un mayor grado de diversidad.

Tabla 4: Prioridades normalizadas y ordenaciones finales de las alternativas.

Alternativas	MAUT – AMG		AHP - EC		ANP - SD	
	Prioridad media	Orden	Prioridad	Orden	Prioridad	Orden
A1	0,1943	2	0,1932	2	0,2007	2
A2	0,1773	4	0,1788	4	0,1767	4
A3	0,0858	6	0,0863	6	0,0873	6
A4	0,1813	3	0,1802	3	0,1850	3
A5	0,2138	1	0,2133	1	0,2214	1
A6	0,1475	5	0,1491	5	0,1289	5

5. Conclusiones

A partir del análisis realizado en la resolución del problema, pudo identificarse el grado de aplicabilidad de los resultados obtenidos a partir de una herramienta multicriterio específica. Considerando en primer lugar la aplicación de MAUT a un problema concreto, pudo realizarse el análisis de sensibilidad para evaluar la estabilidad de la solución ante cambios en los parámetros del modelo. Así se puede apreciar el grado de equilibrio interno que posee la solución proporcionada.

Posteriormente, se realizó un planteamiento del problema mediante AHP. La obtención de resultados coincidentes mediante ambos métodos, permitió confirmar la validez del enfoque multicriterio, ya que utilizando diferentes aproximaciones se pudieron reproducir los resultados de forma independiente.

Por último, se incrementó la robustez del modelo incorporando las interrelaciones y la retroalimentación entre los distintos elementos del sistema, que no pueden considerarse al utilizar una formulación jerárquica apoyada en las técnicas AHP o MAUT, recurriendo a un planteamiento basado en ANP, mediante el que se reprodujeron nuevamente los resultados anteriores.

La reproducción de resultados obtenida con las tres técnicas, sugiere que para conjuntos de problemas que presentan una estructura de preferencias definida, deberían obtenerse resultados similares independientemente de la aproximación multicriterio utilizada.

Bibliografía

- Cohon, J.L. (1978): *Multiobjective Programming and Planning*, Academic Press.
- Evans, G.W. (1984): "An overview of techniques for solving multiobjective mathematical programs", *Management Science*, 30(11), pp. 1268-1282.
- Fishburn, P.C. (1988): *Nonlinear Preference and Utility Theory*, John Hopkins University Press, Baltimore.
- Hwang, C. y Masud, A. (1979): "Multiple objective decision-making methods and applications", en *Lecture Notes in Economics and Mathematical Systems*, 164, Springer-Verlag, New York.
- Keeney, R.L. y Raiffa, H. (1976): *Decisions with Multiple Objectives: Performances and value trade-offs*, Wiley, New York.
- Moreno Jiménez, J.M. (2002): "El Proceso Analítico Jerárquico. Fundamentos, metodología y aplicaciones", en *Toma de decisiones con criterios múltiples. RECT@*, Caballero, R. y Fernández, G.M. (Eds.), serie monográfica nº 1, pp. 21-53.
- Rietveld, P. (1980): *Multiple Objective Decision Methods and Regional Planning*, North-Holland, New York.
- Ríos Insua, D. y French, D. (1991): "A framework for sensitivity analysis in discrete multi-objective decision-making", *European Journal of Operational Research*, 54, pp. 176-190.
- Saaty, T.L. (1980): *The Analytic Hierarchy Process*, McGraw-Hill, New York.
- Saaty, T.L. (1996): *Decision Making with Dependence and Feedback: The Analytic Network Process*, RWS Publications, Pittsburgh.
- Saaty, T. y Vargas, L. (1984): "Comparison of eigenvalue, logarithmic least squares and least squares methods in estimating ratios", *Mathematical Modelling*, 5, pp. 309-324.
- Vincke, P. (1992): *Multi-criteria Decision-aid*, John Wiley, Chichester.
- Von Neumann, J. y Morgenstern, O. (1944): *Theory of games and economic behaviour*, Princeton University Press, Princeton.
- Zionts, S. (Ed.) (1978): *Multiple Criteria Problem Solving*, Springer.

ESTIMACIÓN QMV DE MODELOS GARCH vs ESTIMACIÓN NO PARAMÉTRICA DE LA VOLATILIDAD LOCAL EN SERIES TEMPORALES FINANCIERAS CON OUTLIERS

JULIO ANGEL AFONSO RODRÍGUEZ

e-mail: jafonsor@ull.es

Departamento de Economía de las Instituciones, Estadística Económica y Econometría
Instituto Universitario de Desarrollo Regional (IUDR)
UNIVERSIDAD DE LA LAGUNA

Área temática: MÉTODOS CUANTITATIVOS.

Resumen

En este trabajo estudiamos el comportamiento de una clase de estimadores no paramétricos de la volatilidad local propuesta por Lally, Randal y Thomson (2001, 2004) en el contexto de distribuciones de colas anchas (leptocúrticas). Se trata de estimadores de escala localmente robustos basados en medias móviles finitas de las desviaciones cuadráticas de una serie temporal en torno a su nivel local posiblemente variable en el tiempo. Nuestro objetivo principal consiste, en primer lugar, en aportar nuevos resultados relativos a la no robustez de la estimación cuasi-máximo verosímil de modelos GARCH y, en segundo lugar, en estudiar la robustez de la estimación no paramétrica empleando como medida la magnitud del sesgo por contaminación del proceso subyacente de volatilidad por outliers (VO: volatility outliers). Los resultados analíticos obtenidos se completan con un estudio empírico comparativo de la estimación no paramétrica del proceso de volatilidad de varias series temporales financieras reales frente a la estimación por cuasi-máxima verosimilitud de un modelo GARCH.

Palabras clave: modelos GARCH, volatilidad local, estimación no paramétrica, outliers en volatilidad.

Abstract

In this paper, we study the robustness behavior of a class of non-parametric estimators of local volatility due to Lally, Randal and Thomson (2001, 2004) in the context of leptokurtic distributions. The non-parametric estimator is a local robust scale estimator based on finite moving-averages of the squared deviations of a time series from its local level, possibly time-variant. Our main object is to get new evidence about the non-robustness of the quasi-maximum likelihood estimator of a GARCH model, and to study the robustness of this non-parametric estimator through the boundness of the contamination bias due to the contamination of the underlying volatility process by volatility outliers. Theoretical results are complemented with empirical results of the non-parametric estimation of the volatility process for several financial time series.

Key words: GARCH models, local volatility, non-parametric estimation, volatility outliers

1. Introducción

El concepto de volatilidad juega un papel central en múltiples estudios, tanto empíricos como metodológicos, especialmente en el ámbito de la econometría financiera ya sea haciendo referencia a la varianza (condicional o incondicional) del proceso, al factor o parámetro de escala constante en una distribución o a la volatilidad local cambiante en el tiempo.

En el contexto de la caracterización y medida de la volatilidad cambiante en el tiempo la mayor cantidad de estudios emplean algún tipo de formulación perteneciente a la familia de modelos tipo GARCH o tipo Volatilidad Estocástica (SV), debidos inicialmente a Engle (1982) y Taylor (1986, 1994), respectivamente. La popularidad de estas especificaciones para el proceso de varianza condicional no constante reside básicamente en su clara especificación paramétrica y en la relativamente sencilla estimación dado el software disponible en la actualidad. Sin embargo, las propiedades asociadas a la mayoría de estas especificaciones habitualmente empleadas son generalmente altamente sensibles a la presencia de observaciones anómalas u outliers de la misma forma que lo son los procedimientos habituales de inferencia (identificación, estimación y predicción).

Así, especialmente en el contexto de los modelos tipo GARCH, se han realizado en los últimos años diversos estudios que han tratado de determinar las principales consecuencias de los procedimientos habituales de inferencia cuando la serie observada está contaminada por outliers que afectan tanto al proceso de media como de varianza condicional. Entre los mismos cabe citar las contribuciones de Franses, van Dijk y Lucas (1998), Carnero, Peña y Ruiz (2000) y Afonso, Cano y Guirao (2001) relativas al proceso de identificación del esquema de heterocedasticidad condicional, y de Li y Kao (2002) y Muller y Yohai (2002, 2005), en relación con algunas de las propuestas más recientes de estimación robusta de estos modelos.

Aún cuando alguna de estas nuevas propuestas de estimación paramétrica robusta parece presentar propiedades bastante satisfactorias, no están exentas del problema fundamental que subyace a toda modelización paramétrica, el riesgo de error de especificación. En cambio, la regresión no paramétrica no asume ninguna forma funcional específica para la función de regresión que describe el comportamiento tanto

de la media como de la varianza condicional. Además, un estimador no paramétrico de la volatilidad puede considerarse particularmente apropiado para la extracción y comprensión de la evolución de la volatilidad histórica como paso previo al ajuste de cualquier otro modelo paramétrico más sofisticado, además de proporcionar una referencia robusta para evaluar el comportamiento en la muestra y de predicción de procedimientos paramétricos alternativos.

Así, Heid (1996), Bossaerts, Hafner y Härdle (1996), Hafner (1998) y Pagan y Ullah (1999), entre otros, plantean la utilización de los denominados estimadores kernel introducidos por Nadaraya (1964) y Watson (1964)¹ para la estimación no paramétrica de momentos de distintos órdenes de la distribución condicional del proceso que permite así estimar indirectamente la volatilidad aproximada por la varianza de la distribución condicional.

Alternativamente, una aproximación distinta para la estimación no paramétrica de la volatilidad cambiante en el tiempo es la debida a Lally, Randal y Thomson (2001, 2004), basada en un esquema de media móvil ponderada de longitud finita de las desviaciones cuadráticas de la serie temporal respecto de su nivel local, posiblemente también variable en el tiempo. El sistema de ponderaciones considerado es doble: por un lado, un conjunto de ponderaciones de suavización del ajuste en cada instante del tiempo mediante la media móvil en torno a ese punto y, por otro lado, un conjunto de ponderaciones de robustez frente a distintos supuestos distribucionales y resistencia a outliers. Así mismo, estas ponderaciones de robustificación dependen de una estimación inicial o a priori de la volatilidad local, de forma que para obtener el estimador final será necesario realizar un proceso iterativo de ajuste hasta alcanzar una convergencia aproximada.

Así, de forma general, el estimador de la volatilidad local propuesto constituye una clase de estimadores dependiendo de la elección y combinación de ambos conjuntos de ponderaciones, de la determinación del *span* o amplitud de la media móvil, la constante de truncamiento, de la elección de la estimación inicial de la escala local y del proceso iterativo de ajuste considerado.

¹ Nadaraya, E. (1964): "On estimating regression". *Theory of Probability Applied*, 10, pp. 186-190.
Watson, G. (1964): "Smooth regression analysis". *Sankhya*, Series A, 26, pp. 359-372.

Siendo este trabajo una primera aproximación al estudio de algunas de las características y comportamiento del estimador no paramétrico descrito bajo ciertas condiciones, el objetivo del mismo es doble. En primer lugar realizar un estudio preliminar de la robustez de esta clase de estimadores no paramétricos en presencia de outliers en el proceso subyacente de volatilidad de la serie temporal y, en segundo lugar, examinar mediante simulación el ajuste que proporciona este estimador cuando se emplea como mecanismo generador de la volatilidad cambiante en el tiempo un esquema particular, en concreto un modelo GARCH(1,1) bajo distintos supuestos a cerca de la distribución del término de error y distintos tipos de outliers que afectan al mecanismo que genera la volatilidad siguiendo este esquema paramétrico.

La organización del trabajo es la siguiente. El epígrafe 2 introduce formalmente los conceptos de outlier en nivel (*LO: level outlier*) y outlier en volatilidad (*VO: volatility outlier*), lo que permite distinguir entre ambos tipos de perturbaciones y anticipar en general qué aspectos pueden verse afectados en un determinado tipo de análisis según esté presente uno u otro tipo de outlier. En particular, en el caso de un modelo GARCH(1,1), obtenemos la estructura dinámica del efecto que se genera a partir de la inclusión de un *VO*, y, de forma general, obtenemos algunas notas sobre el efecto de un *VO* en las estimaciones de un modelo tipo GARCH mediante cuasi-máxima verosimilitud. El epígrafe 3 describe el estimador no paramétrico de la volatilidad local, mientras que en la sección 4 se analiza su comportamiento en presencia de outliers en volatilidad. La sección 5 presenta los resultados de un análisis empírico basado en series de rendimientos diarios de 6 índices bursátiles y 2 series de tipos de cambio, empleando distintas configuraciones del estimador no paramétrico. Finalmente, la sección 6 concluye con las principales conclusiones del estudio y apunta una serie de aspectos a tratar en el futuro.

2. Outliers en volatilidad (VO)

2.1. Outliers en nivel (LO) y en volatilidad (VO)

A partir del trabajo de Hotta y Tsay (1998), Doornik y Ooms (2005) plantean una definición formal de outlier distinguiendo entre aquéllos que únicamente afectan al comportamiento y a las características del nivel de la serie y aquéllos otros que afectan únicamente al comportamiento de la varianza condicional. Introducen así una tipología

de outlier definidos como outliers en nivel (*LO*), en el primer caso, y outliers en volatilidad (*VO*) en el segundo. Formalmente, particularizando al caso de una regresión lineal, sea y_t la serie temporal observada, $\mathbf{x}_t = (x_{1,t}, \dots, x_{q,t})$ un vector de q regresores y $\boldsymbol{\zeta}$ un vector de q parámetros constantes, se tiene que en el caso de un único outlier en el proceso podemos escribir

$$y_t - \mathbf{x}_t \boldsymbol{\zeta} = \varepsilon_t(\theta_0) \quad (1)$$

donde $\varepsilon_t(\theta_0)$ denota el término de error contaminado por la presencia de una observación anómala u outlier, siendo $\theta_0 = (T_0, \omega_0)$, con T_0 indicando la localización del outlier y ω_0 la magnitud de la perturbación. De forma general, tenemos la siguiente representación para $\varepsilon_t(\theta_0)$

$$\varepsilon_t(\theta_0) = \varepsilon_t + \varphi_t(\theta_0) \quad (2)$$

donde ε_t se corresponde con el término de error libre de contaminación, con media incondicional nula, $E(\varepsilon_t) = 0$, y normalizado de forma que $Var(\varepsilon_t) = 1$, y $\varphi_t(\theta_0)$ es una función determinista que identifica el tipo de outlier. En el marco de esta formulación general, consideraremos los cuatro tipos habitualmente definidos en la literatura de series temporales: aditivo, innovacional, cambio temporal y cambio de nivel.

Considerando que el nivel de volatilidad local (es decir, en cada instante del tiempo t) puede aproximarse por la varianza de la distribución condicional de los errores dado un cierto conjunto de información condicionante (I_t), se tiene entonces la siguiente definición del tipo de outlier introducido por (1):

$$LO \text{ (level outlier):} \quad \sigma_t^2 = Var_t(\varepsilon_t) = Var(\varepsilon_t | I_t) \quad (3)$$

$$VO \text{ (volatility outlier):} \quad \sigma_t^2 = Var_t(\varepsilon_t(\theta_0)) = Var(\varepsilon_t(\theta_0) | I_t) = \sigma_t^2(\theta_0) \quad (4)$$

Así, un outlier en nivel surge al incorporar el efecto determinista de la observación anómala que describe la media condicional o nivel de la serie,

$$y_t - \mathbf{x}_t \boldsymbol{\zeta} - \varphi_t(\theta_0) = \varepsilon_t$$

de forma que el término de error se utiliza luego para describir el comportamiento de la varianza condicional. Por tanto, un outlier en volatilidad surge cuando el efecto de una observación anómala no se incorpora a la modelización del comportamiento del nivel de la serie y , por tanto, se traslada al comportamiento de la varianza condicional, perturbando dicho proceso.

En base a esta distinción y únicamente en el caso de outliers de tipo aditivo, Doornik y Ooms (2005) proponen un procedimiento iterativo para detectar outliers aditivos en nivel (*ALO*) y en volatilidad (*AVO*) en un modelo GARCH(1,1).

Asumiendo a priori un determinado esquema paramétrico para representar el mecanismo generador del proceso de varianza condicional es posible caracterizar el comportamiento de la volatilidad cuando se considera la ocurrencia de un *VO*.

2.2. Efectos de los outliers en volatilidad sobre la especificación GARCH

Consideramos en primer lugar el modelo GARCH(1,1) (Bollerslev (1987)), definido como

$$\varepsilon_t = \xi_t \cdot \sqrt{h_t}, \quad E(\xi_t) = 0, \quad Var(\xi_t) = 1 \quad (5.1)$$

$$h_t = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 h_{t-1} \quad (5.2)$$

donde $0 < \alpha_i < 1$ ($i = 0, 1$), $0 \leq \beta_1 < 1$ y $\alpha_1 + \beta_1 < 1$, y la secuencia ξ_t es de variables aleatorias independientes e independiente del proceso h_t . Mediante sustituciones recursivas podemos expresar (5.2) alternativamente de la forma

$$h_t = \left(\frac{\alpha_0(1-\beta_1^t)}{1-\beta_1} + \beta_1^t h_0 \right) + \alpha_1 \sum_{j=1}^t \beta_1^{j-1} \varepsilon_{t-j}^2 = \alpha_{0,t} + \alpha_1 \sum_{j=1}^t \beta_1^{j-1} \varepsilon_{t-j}^2 \quad (6)$$

En el caso de un único outlier en volatilidad, el modelo queda especificado en términos del proceso de error contaminado, ecuación (2), de la forma

$$h_t(\theta_0) = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2(\theta_0) + \beta_1 h_{t-1}(\theta_0) \quad (7)$$

es decir,

$$h_t(\theta_0) = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 h_{t-1}(\theta_0) + \alpha_1 \eta_{t-1}(\theta_0) \quad (8.1)$$

donde

$$\eta_t(\theta_0) = \varepsilon_t^2(\theta_0) - \varepsilon_t^2 = \varphi_t^2(\theta_0) + 2\varphi_t(\theta_0)\varepsilon_t, \quad (8.2)$$

De nuevo, empleando sustituciones recursivas en (8.1) es posible obtener la relación entre el esquema de volatilidad contaminado por el outlier y el del esquema GARCH(1,1) de volatilidad en una serie no contaminada, resultando así el efecto dinámico del outlier, de la forma

$$h_t(\theta_0) = h_t + F_t(\theta_0) \quad (9.1)$$

con h_t dado por la ecuación (6), y

$$F_t(\theta_0) = \alpha_1 \sum_{k=0}^{t-1} \beta_1^k \eta_{t-k}(\theta_0) \quad (9.2)$$

con $\eta_{t-k}(\theta_0)$ dado por (8.2), y donde este término sumatorio permite describir de forma general la dinámica del efecto del outlier en volatilidad. Así, en el caso de un outlier en volatilidad en el modelo GARCH(1,1), en general, se espera que, con independencia del tipo de outlier, el efecto sea transitorio aún cuando el efecto puede extenderse de forma considerable a lo largo del tiempo.

Introducimos a continuación de descripción de los cuatro tipos básicos de outlier habitualmente considerados en el contexto del análisis de intervención en la metodología ARIMA con el objeto de caracterizar más particularmente en cada caso la relación entre el proceso de volatilidad contaminado y libre de outlier. Estos se definen como

$$\begin{aligned} \varphi_t(\theta_0) &= \omega_0 \cdot I_t^{(T_0)} && AVO(\text{additive volatility outlier}) : \text{aditivo} \\ &= \omega_0 \cdot \pi_{t-T_0} \cdot S_t^{(T_0)} && IVO(\text{innovational volatility outlier}) : \text{inovacional} \\ &= \omega_0 \cdot S_t^{(T_0)} && LSVO(\text{level shift VO}) : \text{cambio de nivel} \\ &= \omega_0 \cdot \delta^{t-T_0} \cdot S_t^{(T_0)} && TCVO(\text{temporary change VO}) : \text{cambio temporal} \quad (0 < \delta < 1) \end{aligned} \quad (10)$$

donde los coeficientes $\pi_j = \pi_j(\alpha_1, \beta_1)$ son función de los parámetros de la especificación GARCH a través de la función de respuesta al impulso de la representación ARMA para la serie de cuadrados², $I_t^{(T_0)}$ es una variable impulso que toma valor 1 cuando $t = T_0$ y cero en caso contrario y $S_t^{(T_0)}$ es una variable escalón que toma valor 1 para todo $t \geq T_0$ y cero en caso contrario.

En primer lugar tenemos que en el caso de un **outlier aditivo en volatilidad (AVO)**, la diferencia entre la especificación GARCH(1,1) contaminada y no contaminada viene dada por

$$F_t^{AVO}(\theta_0) = \omega_0 \cdot \alpha_1 \beta_1^{t-T_0-1} (2\varepsilon_{T_0} + \omega_0) \cdot S_t^{(T_0+1)} \quad (11)$$

de forma que el outlier tiene un impacto sobre la volatilidad a partir del período siguiente a su ocurrencia con un efecto que se va reduciendo a lo largo del tiempo, asumiendo que $\beta_1 < 1$. Además, si $\varepsilon_{T_0} = 0$, tanto un outlier positivo como negativo producirán un incremento de la volatilidad (ver también Doornik y Ooms, 2005). La

² En el caso particular GARCH(1,1), tenemos que $\pi_{t-T_0} = \alpha_1 \cdot (\alpha_1 + \beta_1)^{t-T_0-1} = \alpha_1 \cdot (\delta_1^{t-T_0-1})$.

magnitud del impacto intermedio, $m_t(\theta_0)$, y la mediana, $m(T_0)$, del tiempo de supervivencia del efecto del *AVO* vienen dados por

$$m_t(\theta_0) = \omega_0(2\varepsilon_{T_0} + \omega_0) \left(\frac{1 - \beta_1^i}{1 - \beta_1} \right) \quad (i = 1, \dots, T - T_0), \quad m(T_0) = \frac{\ln((1 + \beta_1^{T-T_0})/2)}{\ln(\beta_1)}$$

En segundo lugar, en el caso de un **outlier en volatilidad innovacional (IVO)**, el efecto dinámico del outlier viene dado por

$$F_t^{IVO}(\theta_0) = \omega_0 \alpha_1 \beta_1^{t-T_0-1} (2\varepsilon_{T_0} + \omega_0) S_t^{(T_0+1)} + \omega_0 \alpha_1 \left[\left(2 \sum_{k=1}^{t-T_0-1} \beta_1^{k-1} \delta_1^{(t-T_0-k)-1} \varepsilon_{t-k} \right) + \omega_0 \alpha_1 \left(\sum_{k=1}^{t-T_0-1} \beta_1^{k-1} \delta_1^{2[(t-T_0-k)-1]} \right) \right] S_t^{(T_0+2)} \quad (12)$$

donde $\delta_1 = (\alpha_1 + \beta_1)$ y $\sum_{k=1}^{t-T_0-1} \beta_1^{k-1} \delta_1^{2[(t-T_0-k)-1]} = (\delta_1^{2(t-T_0-1)} - \beta_1^{t-T_0-1}) / (\delta_1^2 - \beta_1)$. En este caso, al efecto del *AVO* se añade un efecto diferencial que se propaga a lo largo de la serie a partir del segundo período siguiente al de ocurrencia del outlier.

En el caso de un **outlier en volatilidad de cambio de nivel (LSVO)**, tenemos que

$$F_t^{LSVO}(\theta_0) = \omega_0 \alpha_1 \left(\sum_{k=1}^{t-T_0} \beta_1^{k-1} (2\varepsilon_{t-k} + \omega_0) \right) S_t^{(T_0+1)} = \omega_0 \alpha_1 \left(2 \sum_{k=1}^{t-T_0} \beta_1^{k-1} \varepsilon_{t-k} + \frac{\omega_0(1 - \beta_1^{t-T_0})}{1 - \beta_1} \right) S_t^{(T_0+1)} \quad (13)$$

lo que resulta ser equivalente a un cambio estructural permanente en la volatilidad, mientras que finalmente en el caso de un **outlier en volatilidad de cambio temporal (TCVO)**, de efecto instantáneo y reducción gradual del impacto tenemos que

$$F_t^{TCVO}(\theta_0) = \omega_0 \alpha_1 \left(2 \sum_{k=1}^{t-T_0} \beta_1^{k-1} \delta^{t-T_0-k} \varepsilon_{t-k} + \frac{\omega_0(\beta_1^{t-T_0} - \delta^{2(t-T_0)})}{\beta_1 - \delta^2} \right) S_t^{(T_0+1)} \quad (14)$$

Estas expresiones pueden servir tanto para evaluar de forma aproximada el efecto de los outliers en volatilidad sobre la estimación de los parámetros de un modelo GARCH(1,1) como para la simulación de un modelo GARCH(1,1) con outlier en volatilidad, simplemente incorporando estos efectos aditivos a un esquema GARCH(1,1) generado inicialmente sin contaminación. Estas expresiones pueden generalizarse a otras especificaciones GARCH de forma similar. Así, por ejemplo, un modelo ARCH(p) contaminado por un único outlier en volatilidad, viene dado por

$$h_t(\theta_0) = \alpha_0 + \sum_{i=1}^p \alpha_i \varepsilon_{t-i}^2(\theta_0) \quad (15)$$

de forma que empleando la ecuación (2), podemos expresar el proceso $h_t(\theta_0)$ en

función del proceso libre de contaminación, h_t , de la forma

$$h_t(\theta_0) = \alpha_0 + \sum_{i=1}^p \alpha_i [\varepsilon_{t-i}^2 + \eta_{t-i}(\theta_0)] = h_t + \sum_{i=1}^p \alpha_i \eta_{t-i}(\theta_0) = h_t + F_t(\theta_0) \quad (16)$$

donde, la función $F_t(\theta_0)$ describe la dinámica del efecto de la contaminación del proceso por un outlier en volatilidad. En el caso de un outlier aditivo, tenemos que

$$F_t^{AVO}(\theta_0) = \alpha_{t-T_0} \omega_0 (2\varepsilon_{T_0} + \omega_0) \cdot I_t^{(T_0+i)} \quad i = 1, \dots, p \quad (17)$$

de forma que la influencia del AVO sólo se manifiesta en el período en el que ocurre. Si, en cambio, consideramos la ocurrencia de outlier innovacional, el efecto y la dinámica del mismo viene dado por

$$F_t^{IVO}(\theta_0) = \omega_0 \sum_{k=1}^{t-T_0} \alpha_k \pi_{t-T_0-k} (2\varepsilon_{t-k} + \omega_0 \pi_{t-T_0-k}) \cdot I_t^{(T_0+i)} \quad i = 1, \dots, p \quad (18.1)$$

$$= \omega_0 \sum_{k=1}^p \alpha_k \pi_{t-T_0+p-k} (2\varepsilon_{t-k} + \omega_0 \pi_{t-T_0+p-k}) \cdot S_t^{(T_0+p+1)} \quad (18.2)$$

donde (18.1) es aplicable si $t = T_0 + i$ ($i = 1, \dots, p$), y (18.2) si $t > T_0 + p$, y los coeficientes π se determinan empleando la función de respuesta al impulso de la representación media móvil infinita para la serie de cuadrados³.

Finalmente, indicamos la posibilidad de extender todo este análisis al caso de múltiples outliers en volatilidad, por ejemplo, combinando diferentes tipos de outliers, magnitudes y localizaciones no consecutivas de la forma general

$$\varepsilon_t(\Theta_m) = \varepsilon_t + \varphi_t(\Theta_m), \quad \varphi_t(\Theta_m) = \sum_{k=0}^m \varphi_{k,t}(\theta_k).$$

Un caso de interés podría ser el análisis de los denominados outliers de relocalización (ver Wu *et.al* (1993)), que generan un comportamiento en volatilidad similar al descrito por el fenómeno del agrupamiento de la volatilidad, definido como secuencia de observaciones con variabilidad inusualmente baja seguida por otras de variabilidad inusualmente elevada, de forma tal que la suma de los valores dentro del bloque es la misma que la que se podría esperar para una serie no contaminada. Este modelo puede describirse de la siguiente forma

$$\varphi_t(\Theta_m) = \sum_{k=0}^m \varphi_{k,t}(\theta_k) = \sum_{k=0}^m \omega_k \cdot I_t^{(T_0+k)}, \quad \sum_{k=0}^m \omega_k = 0, \quad \omega_0 = -\sum_{k=1}^m \omega_k$$

³ En este caso, tales coeficientes son: $\pi_0 = 1$, $\pi_k = \sum_{i=1}^k \pi_{k-i} \alpha_i$ ($k=1, \dots, p-1$), $\pi_k = \sum_{i=1}^p \pi_{k-i} \alpha_i$ ($k \geq p$).

es decir, como una secuencia de $(m+1)$ outliers aditivos consecutivos, con magnitudes que se compensan entre sí.

2.3. Efectos de los outliers en volatilidad sobre la estimación GARCH

El procedimiento estándar de estimación de modelos GARCH es el máximo verosímil basado en la distribución condicional del proceso ε_t dada la secuencia $E_{t-1} = (\varepsilon_0, \varepsilon_1, \dots, \varepsilon_{t-1})$, de forma que la densidad conjunta de $\varepsilon_1, \dots, \varepsilon_T$, viene dada por

$$f(\varepsilon_1, \dots, \varepsilon_T) = f(E_0) \prod_{t=1}^T f(\varepsilon_t | E_{t-1}), \quad f(\varepsilon_t | E_{t-1}) = \frac{f(\xi_t)}{h_t}$$

y así, la log-verosimilitud condicional a una realización inicial de h_0 viene dada por $\ell(\boldsymbol{\gamma}) = \sum_{t=1, T} \ell_t(\boldsymbol{\gamma})$, donde $\boldsymbol{\gamma}$ es el vector de parámetros de la ecuación GARCH. Si se asume que la distribución de las variables ξ_t en (5) es normal⁴, se tiene que

$$\ell_t(\boldsymbol{\gamma}) = -\ln(2\pi)/2 - (1/2) \cdot \left(\ln h_t(\boldsymbol{\gamma}) + \frac{\varepsilon_t^2}{h_t(\boldsymbol{\gamma})} \right) \quad t = 1, \dots, T.$$

Puesto que en presencia de un VO, se tiene de forma general que

$$\varepsilon_t^2(\theta_0) = \varepsilon_t^2 + \varphi_t(\theta_0)$$

$$\text{y } h_t(\boldsymbol{\gamma}, \theta_0) = h_t(\boldsymbol{\gamma}) + F_t(\boldsymbol{\gamma}, \theta_0) = h_t(\boldsymbol{\gamma}) \left(1 + \frac{F_t(\boldsymbol{\gamma}, \theta_0)}{h_t(\boldsymbol{\gamma})} \right)$$

entonces la log-verosimilitud gaussiana puede expresarse ahora como

$$\ell_t(\boldsymbol{\gamma}, \theta_0) = \ell_t(\boldsymbol{\gamma}) - \frac{1}{2} \left(\ln \left[1 + \frac{F_t(\boldsymbol{\gamma}, \theta_0)}{h_t(\boldsymbol{\gamma})} \right] + \frac{\eta_t(\theta_0) - \xi_t^2 F_t(\boldsymbol{\gamma}, \theta_0)}{h_t(\boldsymbol{\gamma}, \theta_0)} \right) \quad t = 1, \dots, T,$$

donde $\ell_t(\boldsymbol{\gamma}, \theta_0)$ es la log-verosimilitud bajo contaminación. Puesto que la maximización de esta verosimilitud se realiza habitualmente empleando el algoritmo iterativo BHHH, que se basa exclusivamente en el gradiente de $\ell_t(\boldsymbol{\gamma})$, el efecto del outlier se traslada al estimador de $\boldsymbol{\gamma}$ a través de los términos

⁴ El supuesto de normalidad es raramente soportado por los datos, de forma que ignorar la no normalidad resultará en una incorrecta especificación de la función de log-verosimilitud. Sin embargo, maximizar en este caso la log-verosimilitud gaussiana se justifica por la teoría de la cuasi-máxima verosimilitud (QMV), resultando estimadores consistentes y asintóticamente normales. Si se especifica una distribución paramétrica alternativa, la estimación MV resultará en estimadores más eficientes. Sin embargo, la estimación MV bajo incorrecta especificación de una distribución condicional no gaussiana puede resultar en estimadores inconsistentes (Newey, W., D. Steigerwald (1997): "Asymptotic bias for quasi maximum likelihood estimators in condicional heteroskedasticity models", *Econometrica*, 65, pp.587-599).

$$\frac{\partial}{\partial \gamma} \ln \left(1 + \frac{F_t(\gamma, \theta_0)}{h_t(\gamma)} \right) = \frac{F_t(\gamma, \theta_0)}{h_t(\gamma, \theta_0)} \left(\frac{F_t'(\gamma, \theta_0)}{F_t(\gamma, \theta_0)} - \frac{h_t'(\gamma)}{h_t(\gamma)} \right), F_t'(\gamma, \theta_0) = \frac{\partial F_t(\gamma, \theta_0)}{\partial \gamma}, h_t'(\gamma) = \frac{\partial h_t(\gamma)}{\partial \gamma}$$

y

$$\frac{\partial}{\partial \gamma} \frac{\eta_t(\theta_0) - \xi_t^2 F_t(\gamma, \theta_0)}{h_t(\gamma, \theta_0)} = \xi_t^2 \frac{F_t(\gamma, \theta_0)}{h_t(\gamma, \theta_0)} \left(\frac{h_t'(\gamma, \theta_0)}{h_t(\gamma, \theta_0)} - \frac{F_t'(\gamma, \theta_0)}{F_t(\gamma, \theta_0)} \right) - \frac{\eta_t(\theta_0)}{h_t(\gamma, \theta_0)} \frac{h_t'(\gamma, \theta_0)}{h_t(\gamma, \theta_0)}$$

Así, aún en el caso de un único *VO*, es de esperar, que su influencia global sobre la estimación de γ sea significativa, especialmente por el efecto de los términos $F_t'(\gamma, \theta_0)/F_t(\gamma, \theta_0)$ y $\eta_t(\theta_0)/h_t(\gamma, \theta_0)$. Así, en el caso GARCH(1,1) y en el caso de un único *AVO*, se tiene que

$$\frac{F_t'^{AVO}(\gamma, \theta_0)}{F_t^{AVO}(\gamma, \theta_0)} = \begin{pmatrix} 0 \\ 1/\alpha_1 \\ (t-T_0-1)/\beta_1 \end{pmatrix} \cdot S_t^{(T_0+1)}, \text{ y } \frac{\eta_t(\theta_0)}{h_t(\gamma, \theta_0)} = \omega_0 \left(\frac{\omega_0}{h_t(\gamma, \theta_0)} + 2\xi_t \frac{\omega_0}{\sqrt{h_t(\gamma, \theta_0)}} \right) \cdot I_t^{(T_0)}$$

3. Estimación no paramétrica de la volatilidad local

Consideremos siguiente la representación general de una serie temporal de rendimientos de algún activo financiero,

$$\tilde{y}_t = y_t - \mu_t = v_t \cdot \sigma_t \quad (19)$$

donde μ_t representa el nivel de la serie, posiblemente cambiante en el tiempo, y donde se asume que el término de error v_t , independiente de y_t , sigue una distribución de colas anchas. Lally *et al.* (2001, 2004) proponen como estimador de la volatilidad cambiante en el tiempo, σ_t^2 , una media móvil finita de amplitud $n = 2 \cdot r + 1$ dada por la siguiente expresión

$$\hat{\sigma}_t^2(n) = \sum_{j=-r}^r \varpi_j(n) \cdot q_{t+j} \cdot \tilde{y}_{t+j}^2 \quad (20)$$

donde las ponderaciones de suavización, $\varpi_j(n)$, satisfacen $\sum_{j=-r}^r \varpi_j(n) = 1$ y q_t son las ponderaciones de robustez que dependen de alguna estimación inicial o a priori s_t^2 de la volatilidad local, σ_t^2 , $q_t = Q((\tilde{y}_t/2s_t)^2)$. En la práctica, dependiendo de la elección de dicha estimación inicial, se tratará de determinar de forma aproximadamente óptima la magnitud de estas ponderaciones de robustez, lo que requerirá algún proceso iterativo de forma que la estimación inicial de σ_t^2 será sucesivamente refinada.

Lally *et al.* consideran dos posibles especificaciones alternativas para estos factores de ponderación, asociados una elección de la estimación inicial de la escala local, σ_t^2 : (A) ponderaciones de robustez basadas en un estimador robusto del parámetro de escala en distribuciones simétrica de colas anchas tipo-A con función de ponderación doble, como el propuesto por Lax (1985)⁵ y, (B) ponderaciones de robustez basadas en un estimador de máxima-verosimilitud del parámetro de escala relacionado con la distribución $T(\nu)$ -Student empleando el algoritmo *EM*. Describimos a continuación los detalles de ambos estimadores.

3.1. Estimador-A del parámetro de escala local de Lax (1985)

El estimador-A de la escala local con función de doble ponderación viene dado por

$$s_{A,t}^2(c, s_0) = \frac{k_{A,t}^2(c, s_0)}{t-1} \sum_{i=1}^t q_i(c, s_0) (y_i - \hat{M}_0)^2 \quad (21)$$

donde

$$q_i(c, s_0) = I_i(c, s_0) (1 - u_i^2(c, s_0))^4, \quad I_i(c, s_0) = \begin{cases} 1 & \text{si } |y_i - \hat{M}_0| \leq c \cdot s_0 \\ 0 & \text{si } |y_i - \hat{M}_0| > c \cdot s_0 \end{cases} \quad (22)$$

$$u_i(c, s_0) = I_i(c, s_0) \frac{(y_i - \hat{M}_0)}{c \cdot s_0} \quad (23)$$

$$k_{A,t}^{-1}(c, s_0) = \frac{1}{t} \sum_{i=1}^t (1 - u_i^2(c, s_0)) (1 - 5 \cdot u_i^2(c, s_0)), \quad (24)$$

con \hat{M}_0 un estimador-*M* de la localización, generalmente la mediana muestral, y s_0 una estimación auxiliar del parámetro de escala, generalmente la desviación absoluta mediana de la serie respecto de la mediana muestral, $s_0 = MAD(y_t, \hat{M}_0)$. La elección del valor de la constante positiva c , el parámetro de truncamiento en (18), dependerá de la elección de s_0 .

3.2. Estimador-T del parámetro de escala local

Si consideramos la siguiente representación alternativa a (19) para la serie temporal de rendimientos observada,

$$y_t = \mu_t + \sigma \cdot z_t / \sqrt{V_t} \quad (25)$$

⁵ Lax, D.A. (1985): "Robust estimators of scale: finite-sample performance in long-tailed symmetric

donde $z_i \sim i.i.d.N(0, 1)$ y V_i son variables aleatorias independientes en $[0, \infty)$ y con función de distribución $F_{\nu}(\nu)$ y las variables V_i se asumen con distribución proporcional a una variable aleatoria χ_{ν}^2 , entonces las variables $\sqrt{\nu} (y_i - \mu_i)$ tienen distribución $T(\nu)$ de Student. La esperanza de la log-verosimilitud gaussiana de la muestra aleatoria de (y_1, \dots, y_T) , (V_1, \dots, V_T) condicional a los valores observados de \mathbf{y}_T y evaluada en una estimación a priori de σ^2 , $\hat{\sigma}_0^2$, es aproximadamente igual a

$$E_0(\ln L_T | \mathbf{y}_T) = -\ln(\sigma^2) - \frac{1}{T \cdot \sigma^2} \sum_{t=1}^T E_0(V_t | \mathbf{y}_T) \tilde{y}_t^2 \quad (26)$$

de forma que la maximización respecto a σ^2 proporciona el siguiente estimador

$$\hat{\sigma}^2 = \frac{1}{T} \sum_{t=1}^T E_0(V_t | \mathbf{y}_T) \tilde{y}_t^2 \quad (27)$$

Si se considera ahora el caso donde $V_i \sim \chi_{\nu}^2/(\nu-2)$, de forma que $\sqrt{\nu} (y_i - \mu_i)/\sqrt{\nu-2}$ se distribuye $T(\nu)$ con $Var(y_i) = \sigma^2$. Entonces (27) resulta ser⁶

$$\hat{\sigma}_T^2(\nu) = \frac{1}{T} \sum_{t=1}^T q_t(\nu, \sigma_0^2) \tilde{y}_t^2 = \frac{1}{T} \sum_{t=1}^T q_{0,t}(\nu) \tilde{y}_t^2 \quad (28)$$

con

$$q_{0,t}(\nu) = q_t(\nu, \sigma_0^2) = \frac{\nu+1}{\nu-2} \left(1 + \frac{\tilde{y}_t^2}{(\nu-2)\sigma_0^2} \right)^{-1} \quad (29)$$

En la práctica, para un valor dado de ν , es preciso iterar este estimador para obtener el estimador máximo-verosímil final de σ^2 , es decir,

$$\hat{\sigma}_T^2(\nu, i) = \frac{1}{T} \sum_{t=1}^T \hat{q}_{i-1,t}(\nu) \tilde{y}_t^2, \quad \hat{q}_{i-1,t}(\nu) = \frac{\nu+1}{\nu-2} \left(1 + \frac{\tilde{y}_t^2}{(\nu-2)\hat{\sigma}_T^2(\nu, i-1)} \right)^{-1} \quad (30)$$

3.3. Estimadores no paramétricos alternativos de la volatilidad local

Así, si el estimador descrito en (20) se basa en las ecuaciones (21)-(23), entonces s_t será el estimador *MAD* móvil, requiriendo únicamente dos iteraciones, una para determinar s_t y otra para determinar la estimación final de σ_t^2 . Es decir,

distributions”, *Journal of the American Statistical Association*, 80, pp. 736-741.

⁶ La estructura de ponderaciones de robustez en (29) resulta de considerar que la esperanza condicional en (27) es una función cuadrática $Q(\cdot)$ de (y_t, σ_0) que admite una representación general mediante la transformada de Laplace, $M(\cdot)$. Es decir, $Q(t) = -\partial \ln(M(t))/\partial t$, donde $M(t) = \int_{0,\infty} e^{-ts} \sqrt{s} dF(s)$.

$$\hat{\sigma}_t^2(n|c, s_t) = \sum_{j=-r}^r \varpi_j(n) \cdot q_{t+j}(c, s_{t+j}) \cdot \tilde{y}_{t+j}^2 \quad (31)$$

donde las ponderaciones q_t vienen dadas por las ecuaciones (22) y (23). A este estimador de la volatilidad local lo denominaremos estimador-A.

Por otro lado, si el estimador de media móvil se basa en (28) y (29), hemos considerado dos posibles especificaciones alternativas dependiendo de la elección de la estimación a priori de la escala local. En primer lugar, considerar como estimador inicial de σ_0 para cada $t = 1, \dots, T$ el estimador dado en (21) basado en s_t , la desviación absoluta mínima móvil (*MAD*) respecto a la estimación local de la localización, y utilizar éste valor como inicial en el proceso iterativo de ajuste de (30) hasta alcanzar la convergencia aproximada, empleando como criterio de convergencia que el gradiente de la función objetivo sea aproximadamente cero, es decir,

$$\nabla_{\hat{E}_{(i)}}(\ln L_T | \mathbf{y}_T) = \left. \frac{\partial E_0(\ln L_T | \mathbf{y}_T)}{\partial \sigma^2} \right|_{\sigma^2 = \hat{\sigma}_t^2(v, i)} \approx 0 \quad (32)$$

En este caso, el estimador dependerá de la elección del parámetro de truncamiento c en el cálculo de (21).

Una segunda posibilidad consiste en emplear como estimador a priori de la escala la varianza muestral móvil para inicializar el proceso iterativo en (30), empleando el mismo criterio de convergencia.

Las estimaciones resultantes se toman entonces para construir las ponderaciones en (29) y el cálculo final del estimador no paramétrico en (20). Hemos denominado a estos estimadores como estimadores-*T(MAD)* y *T(VAR)*, respectivamente:

$$\hat{\sigma}_t^2(n|c, v) = \sum_{j=-r}^r \varpi_j(n) \cdot q_{t+j}^{(MAD)}(v) \cdot \tilde{y}_{t+j}^2 \quad (33)$$

y

$$\hat{\sigma}_t^2(n|v) = \sum_{j=-r}^r \varpi_j(n) \cdot q_{t+j}^{(VAR)}(v) \cdot \tilde{y}_{t+j}^2 \quad (34)$$

En el siguiente epígrafe examinaremos el comportamiento de esta clase de estimadores cuando el proceso inobservable de volatilidad se contamina por la ocurrencia de un outlier.

4. Función de influencia empírica en el caso de un outlier aditivo en volatilidad

La función de influencia empírica proporciona una idea de cómo se comporta un estimador cuando se cambia un punto en la muestra, de forma que depende exclusivamente de los datos, independientemente de posibles supuestos distribucionales. En el caso del estimador propuesto en (20), consideraremos la construcción de la función de influencia para cada punto de la muestra t . Asumiendo, sin pérdida de generalidad que el proceso tiene media nula, se tiene que los valores observados empleados para la estimación de la escala local vendrán dados por

$$y_{t+j}(\theta_0) = y_{t+j} + \varphi_{t+j}(\theta_0), \quad j = -r, \dots, r, t = 1, \dots, T \quad (35)$$

El estimador no paramétrico de la volatilidad local se define ahora como

$$\hat{\sigma}_t^2(n, \theta_0) = \sum_{j=-r}^r \varpi_j(n) \cdot q_{t+j}(\theta_0) \cdot y_{t+j}^2(\theta_0) \quad (36)$$

donde las ponderaciones de robustez, que son funciones dependientes de los datos, también resultarán afectadas por la ocurrencia del outlier en volatilidad.

Se tiene así que para cada $t = 1, \dots, T$ y $j = -r, \dots, r$, podemos escribir (36) como

$$\hat{\sigma}_t^2(n, \theta_0) = \hat{\sigma}_t^2(n) + \sum_{j=-r}^r \varpi_j(n) \cdot q_{t+j}(\theta_0) \left(A_{t+j}(\theta_0) + B_{t+j}(\theta_0) \right) \quad (37)$$

donde

$$A_{t+j}(\theta_0) = [\Delta_{t+j}(\theta_0) - 1] y_{t+j}^2, \quad (38)$$

$$B_{t+j}(\theta_0) = \Delta_{t+j}(\theta_0) \cdot \eta_{t+j}(\theta_0) = \Delta_{t+j}(\theta_0) \cdot [\varphi_{t+j}^2(\theta_0) + 2\varphi_{t+j}(\theta_0) y_{t+j}] \quad (39)$$

$$\text{y } \Delta_{t+j}(\theta_0) = q_{t+j}(\theta_0) / q_{t+j}. \quad (40)$$

En el caso de un outlier aditivo en volatilidad, tenemos que si para algún j , $t+j = T_0$,

$$\hat{\sigma}_t^2(n, \theta_0) = \hat{\sigma}_t^2(n) + \varpi_{T_0-t}(n) \cdot q_{T_0}(\theta_0) \left(A_{T_0}(\theta_0) + B_{T_0}(\theta_0) \right) I_{[T_0-r, T_0+r]}(t) \quad (37')$$

donde $I_{[T_0-r, T_0+r]}(t) = 1$ si $T_0-r \leq t \leq T_0+r$, y 0 en caso contrario. Se tiene así que la influencia de un outlier en volatilidad sobre el comportamiento del estimador no paramétrico es doble, a través de la propia observación contaminada y a través de las ponderaciones de robustez. En el caso de un VO de tipo innovacional, se tiene que si para algún t , $t+j \geq T_0$,

$$\begin{aligned}\sum_{j=-r}^r \varpi_j(n) \cdot q_{t+j} \cdot B_{t+j}(\theta_0) &= \omega_0 \sum_{\substack{j=T_0-t \\ t+r-T_0}}^r \varpi_j(n) \cdot q_{t+j} \cdot \Delta_{t+j}(\theta_0) (\omega_0 \pi_{t+j-T_0}^2 + 2\pi_{t+j-T_0} y_{t+j}) \\ &= \omega_0 \sum_{k=0}^{t+r-T_0} \varpi_{T_0-t+k}(n) \cdot q_{T_0+k} \cdot \Delta_{T_0+k}(\theta_0) (\omega_0 \pi_k^2 + 2\pi_k y_{T_0+k})\end{aligned}$$

de forma que la ecuación (37) viene dada ahora por

$$\hat{\sigma}_t^2(n, \theta_0) = \hat{\sigma}_t^2(n) + \sum_{k=0}^{t+r-T_0} \varpi_{T_0-t+k}(n) q_{T_0+k} [A_{T_0+k}(\theta_0) + \omega_0 \pi_k \Delta_{T_0+k}(\theta_0) (\omega_0 \pi_k + 2y_{T_0+k})] I_{[T_0-r, T_0+r]}(t)$$

Puesto que en este caso el efecto total del outlier viene dado por la suma de efectos similares al causado por un *AVO*, es de esperar que su influencia sea también limitada. Analizaremos a continuación el caso de un outlier en volatilidad aditivo para las dos opciones de ponderación analizadas en este trabajo.

4.1. Influencia de un *AVO* con ponderaciones q basadas en el estimador-A

En este caso tenemos, en primer lugar, que $B_{T_0}(\theta_0)$ queda de la forma

$$B_{T_0}(\theta_0) = \omega_0 \cdot \Delta_{T_0}(\theta_0) [\omega_0 + 2y_{T_0}]$$

mientras que, por otro lado, el término de ponderación viene dado por

$$q_{T_0}(\theta_0) = I_{T_0}^*(c, s_0) \left(1 - \frac{y_{T_0}^2(\theta_0) \cdot I_{T_0}^*(c, s_0)}{c^2 s_0^2} \right)^4,$$

donde la función indicador se define como

$$I_{T_0}^*(c, s_0) = \begin{cases} 1 & \text{si } |y_{T_0} + \omega_0| \leq c \cdot s_0 \\ 0 & \text{si } |y_{T_0} + \omega_0| > c \cdot s_0 \end{cases}$$

Por la desigualdad del triángulo, se tiene que $|y_{T_0} + \omega_0| \leq |y_{T_0}| + |\omega_0|$, de forma que si la función $I_{T_0}^*(c, s_0)$ toma valor uno, también lo hará $I_{T_0}(c, s_0)$, puesto que $|y_{T_0}| \leq c \cdot s_0 - |\omega_0| < c \cdot s_0$. Así, se tiene que (40) puede escribirse como

$$\Delta_{T_0}(\theta_0) = \frac{q_{T_0}(\theta_0)}{q_{T_0}} = \begin{cases} \left(1 - \frac{\omega_0(\omega_0 + 2y_{T_0})}{c^2 s_0^2 - y_{T_0}^2} \right)^4 & \text{si } |y_{T_0} + \omega_0| \leq c \cdot s_0 \\ 1 & \text{si } |y_{T_0} + \omega_0| > c \cdot s_0 \end{cases}$$

Por construcción, se tiene entonces que cuando $|y_{T_0} + \omega_0| \leq c \cdot s_0$,

$$0 \leq \frac{|\omega_0| (|\omega_0| + 2c \cdot s_0)}{c^2 s_0^2 - y_{T_0}^2} \leq 1$$

de forma que $\Delta_{T_0}(\theta_0) \leq 1$. Por tanto, en presencia de un outlier aditivo en volatilidad, siempre que la elección de las constantes (c, s_0) determinen que $|y_{T_0} + \omega_0| \leq c \cdot s_0$,

$$\hat{\sigma}_t^2(n, \theta_0) = \hat{\sigma}_t^2(n) + \varpi_{T_0-t}(n) \cdot \left(1 - \frac{y_{T_0}^2}{c^2 s_0^2}\right) \left(A_{T_0}(\theta_0) + B_{T_0}(\theta_0)\right) I_t^{(T_0-j)}$$

donde $0 \leq 1 - y_{T_0}^2 / c^2 s_0^2 \leq 1$,

$$A_{T_0}(\theta_0) = y_{T_0}^2 \left[\left(1 - \frac{\omega_0(\omega_0 + 2y_{T_0})}{c^2 s_0^2 - y_{T_0}^2}\right)^4 - 1 \right] \leq -a \cdot y_{T_0}^2 \quad 0 \leq a \leq 1$$

$$\text{y } B_{T_0}(\theta_0) = \omega_0(\omega_0 + 2y_{T_0}) \left(1 - \frac{\omega_0(\omega_0 + 2y_{T_0})}{c^2 s_0^2 - y_{T_0}^2}\right)^4 \leq \omega_0(\omega_0 + 2y_{T_0}).$$

Así, aún en el caso de ocurrencia de un *AVO*, si su magnitud no excesiva como para que no opere el truncamiento, es decir, que no se omita dicha observación en el cómputo de la correspondiente ponderación, el efecto sobre el estimador-A de la volatilidad local será en general de influencia limitada. Por otro lado, cuando la elección de las constantes (c, s_0) determine efectivamente el truncamiento de la observación influyente, es decir, cuando $|y_{T_0} + \omega_0| > c \cdot s_0$, entonces se tiene que ésta no ejercerá distorsión alguna, puesto que directamente se excluye del cómputo de la media móvil en torno al correspondiente punto muestral t .

4.2. Influencia de un *AVO* con ponderaciones q basadas en el estimador- T

En este caso se tiene por construcción que el factor dado en (40) está acotado superiormente

$$\Delta_{T_0}(\theta_0) = \frac{1}{1 + \frac{\omega_0(\omega_0 + 2y_{T_0})}{(\nu - 2)\sigma_0^2 + y_{T_0}^2}} < \frac{1}{q_{T_0}} \quad (q_{T_0} < 1), \quad \Delta_{T_0}(\theta_0) - 1 = -\frac{1}{1 + \frac{\omega_0(\omega_0 + 2y_{T_0})}{(\nu - 2)\sigma_0^2 + y_{T_0}^2}} < \frac{1 - q_{T_0}}{q_{T_0}}$$

Por tanto,

$$A_{T_0}(\theta_0) < y_{T_0}^2 \left(\frac{1 - q_{T_0}}{q_{T_0}} \right), \quad \text{y} \quad B_{T_0}(\theta_0) < \frac{\omega_0(\omega_0 + 2y_{T_0})}{q_{T_0}}$$

Así, el efecto diferencial de un *AVO* sobre el estimador- T no paramétrico de la volatilidad respecto del caso de no contaminación está acotado superiormente por

$$\hat{\sigma}_t^2(n, \theta_0) - \hat{\sigma}_t^2(n) < \varpi_{T_0-t}(n) \cdot [y_{T_0}^2 (1 - q_{T_0}) + \omega_0 (\omega_0 + 2y_{T_0})] I_{[T_0-r, T_0+r]}(t)$$

lo que representa también una influencia limitada de la magnitud de un outlier aditivo en volatilidad, pero que dependiendo de la misma podría llegar a ser elevada, al igual que en el caso anterior. Para ilustrar estos resultados se ha efectuado un extenso ejercicio de simulación que no se incluye en el presente trabajo⁷, pero que describimos a continuación así como las principales conclusiones que se extraen del mismo.

4.3. Evaluación numérica de la robustez

Con el objeto de evaluar el comportamiento de distintas configuraciones del estimador analizado y su robustez ante la inclusión de outliers en volatilidad se ha simulado un modelo básico de referencia GARCH(1,1) bajo distintos supuestos distribucionales (normalidad y distribución *T*-Student) y variantes de la inclusión de un outlier aditivo en volatilidad (localización y magnitud) y en cada caso hemos considerado el ajuste del proceso de volatilidad empleando el estimador no paramétrico.

La consideración de un esquema particular como el GARCH(1,1) para generar el proceso de volatilidad a partir de la siguiente representación $y_t - \mu_t = z_t \cdot \sqrt{h_t}$, nos permite obtener la relación exacta entre $\hat{\sigma}_t^2(n)$ y h_t . En primer lugar tenemos que

$$\hat{\sigma}_t^2(n) = \alpha_0 \sum_{j=-r}^r \varpi_j(n) \cdot q_{t+j} + (\alpha_1 + \beta_1) \sum_{j=-r}^r \varpi_j(n) q_{t+j} \cdot \tilde{y}_{t+j-1}^2 + \sum_{j=-r}^r \varpi_j(n) q_{t+j} (1 - \beta_1 L) v_{t+j}$$

donde $v_t = (z_t^2 - 1)h_t$ es un proceso con valor esperado nulo, incorrelacionado y condicionalmente heterocedástico. Considerando en segundo lugar que el modelo GARCH(1,1) puede escribirse también como

$$h_t = \alpha_0 + (\alpha_1 + \beta_1) \tilde{y}_{t-1}^2 - \beta_1 v_{t-1}$$

entonces, se tiene finalmente que

$$E[h_t - \hat{\sigma}_t^2(n)] = \alpha_0 \left(1 - \sum_{j=-r}^r \varpi_j(n) \cdot q_{t+j} \right) + (\alpha_1 + \beta_1) \left(\tilde{y}_{t-1}^2 - \sum_{j=-r}^r \varpi_j(n) q_{t+j} \cdot \tilde{y}_{t+j-1}^2 \right) \quad (41)$$

de forma que la magnitud del error cometido al aproximar la volatilidad generada por el esquema GARCH(1,1) empleando el estimador $\hat{\sigma}_t^2(n)$ dependerá básicamente de la magnitud del último término en (41) que mide el error del ajuste del cuadrado de cada

observación por una media móvil en un entorno del mismo. En la práctica, comparamos los resultados de la estimación no paramétrica con los de la estimación cuasi máximo-verosímil empleando los errores absoluto medio y absoluto relativo medio. Las principales conclusiones que pueden extraerse son las siguientes. En primer lugar, en relación con el estimador- A , bajo distribución condicional normal o T -Student, los resultados no son sensibles a la elección de la constante de truncamiento, c , y el error de ajuste es mayor cuanto menor es la amplitud de la media móvil, n . En el caso de un proceso contaminado por un AVO , el error de ajuste es sensiblemente inferior al caso de la estimación QMV y además se reduce cuanto mayor es el valor de c y en cualquier caso es bastante inferior cuando se reduce el número de términos en el cálculo de la media móvil. Por otro lado, las magnitudes de los errores no son sensibles a la localización del outlier.

En relación con el estimador- $T(MAD)$, en el caso sin outliers, el comportamiento bajo normalidad es el esperado, puesto que los errores se reducen al incrementar los grados de libertad (ν), mientras que con la distribución T -Student, $T(d)$, los errores del ajuste se reducen siempre que $\nu \geq d$. Los resultados no son sensibles al valor de c y, en general, son mayores cuando se reduce la amplitud de la media móvil. En presencia de un AVO , los niveles de error son superiores en comparación con los correspondientes al estimador- A , aproximándose sus magnitudes cuanto mayor es el valor de c , y reduciéndose significativamente cuanto mayor es el valor de ν . En general, los resultados del ajuste son mejores cuanto menor es la elección del valor de n .

Finalmente, en relación con el estimador- $T(VAR)$, los resultados en el caso de no contaminación son similares a los del estimador- A , mientras que con un AVO los errores son ligeramente superiores, especialmente cuando $c = 9$ ⁸. El comportamiento del estimador es más satisfactorio cuando se reduce el número de términos en la media móvil y, en general, el ajuste es satisfactorio para valores moderados de ν , pero empeora para valores grandes de ν . Los resultados no son en ningún caso sensibles a la localización del outlier.

⁷ Los resultados de este estudio de simulación pueden obtenerse mediante petición al autor.

⁸ El valor $c = 9$, es el recomendado por Lax a partir de su estudio de simulación cuando se utiliza como estimador del parámetro de escala la desviación absoluta mínima, mientras que Lally *et al.* recomiendan emplear valores ligeramente superiores, 10 u 11.

En todos los casos, los errores absoluto y relativo aumentan con la magnitud del efecto instantáneo del outlier, ω_0 , y son siempre inferiores a los valores que se alcanzan empleando el estimador QMV , indicando una mayor resistencia al efecto pernicioso de los outliers.

5. Análisis empírico

En este epígrafe examinamos el comportamiento de estos estimadores sobre datos reales de rendimientos de activos financieros. Las series que hemos considerado son rendimientos diarios de seis índices y de dos tipos de cambio que se describen en el siguiente cuadro junto con las estimaciones robustas del parámetro de escala basado en las propuestas de Lax (1985) y Lally *et.al.* (2001, 2004).

Cuadro 1. Series analizadas y estimación robusta del parámetro de escala

	$s_{A,T}^2(c, s_0)$		$\hat{\sigma}_T^2(\nu)$		$\hat{\sigma}_T^2(\nu, i)$	
	$c = 9$	$\nu = 3$	$\nu = 9$	$\nu = 3$	$\nu = 9$	
IBEX35 (Periodo: 27.12.95-28.12.01) T = 1497	0.18934	0.25704	0.20614	0.34632 (13)	0.21099 (5)	
CAC40 (Periodo: 04.01.00-20.02.07) T = 1822	0.15282	0.20716	0.17341	0.27529 (14)	0.17911 (6)	
DJIA (Dow Jones Industrial Average) (Periodo: 04.01.00-20.02.07) T = 1791	0.09025	0.12149	0.10094	0.16223 (15)	0.10407 (6)	
NASDAQ (Periodo: 05.01.99-20.02.07) T = 2043	0.27425	0.36612	0.30733	0.48923 (15)	0.31729 (6)	
SP500 (Periodo: 04.01.00-16.04.07) T = 1829	0.09742	0.13035	0.10877	0.17268 (14)	0.11187 (6)	
FTSE100 (Periodo: 03.01.98-16.04.07) T = 2343	0.10436	0.13978	0.11644	0.18500 (14)	0.11973 (6)	
TC \$/Libra Esterlina (Periodo: 04.01.00-17.04.07) T = 1832	0.02592	0.03425	0.02699	0.04579 (12)	0.02773 (5)	
TC Euro/\$ (Periodo: 04.01.00-17.04.07) T = 1832	0.03566	0.04712	0.03722	0.06311 (13)	0.03772 (5)	

Nota: Las estimaciones del parámetro de escala están multiplicadas por 1000. En las dos últimas columnas, entre paréntesis, el número de iteraciones hasta la convergencia aproximada.

Estas estimaciones del parámetro de escala de la distribución incondicional difieren de las estimaciones de la dispersión habituales en un análisis exploratorio, válidas en el caso de que la distribución fuese aproximadamente normal, lo que proporciona una justificación inicial para tratar la estimación robusta de la volatilidad local ante posibles observaciones anómalas. Los gráficos de estas series presentan la apariencia habitual, con bloques de alta y baja volatilidad y observaciones anómalas que podrían interpretarse con outliers en volatilidad. El siguiente cuadro presenta los resultados de la estimación QMV de un modelo GARCH(1,1) para cada serie y de la estimación no paramétrica empleando los estimadores denominados estimador-A y $T(VAR)$ para una selección de parámetros (n, c_0) y (n, ν) , respectivamente.

Los resultados que se presentan son los estadísticos error absoluto medio (*EAM*), error relativo absoluto medio (*ERAM*) y error cuadrático medio (*ECM*), definidos como

$$EAM_T = \frac{1}{T} \sum_{t=1,T} |\hat{\sigma}_t^2 - y_t^2|, ERAM_T = \frac{1}{T} \sum_{t=1,T} |1 - y_t^2 / \hat{\sigma}_t^2|, ECM_T = \frac{1}{T} \sum_{t=1,T} (\hat{\sigma}_t^2 - y_t^2)^2$$

donde $\hat{\sigma}_t^2$ es, en cada caso, el estimador considerado de la volatilidad local y se utiliza y_t^2 como aproximación de la serie inobservable de volatilidad.

Cuadro 2. Resultados comparativos de la estimación de la volatilidad

		QMV	Estimador-A $\hat{\sigma}_t^2(n/c, s_t)$		Estimador-T(VAR) $\hat{\sigma}_t^2(n/\nu)$	
		GARCH(1,1)	$(n, c_0) = (20, 9)$	$(n, c_0) = (20, 11)$	$(n, \nu) = (20, 5)$	$(n, \nu) = (20, 9)$
IBEX35	<i>EAM</i>	0.22479	0.18981	0.19164	0.21855	0.20607
	<i>ERAM</i>	0.98369	1.45902	1.30147	0.86724	0.95015
	<i>ECM</i>	1.87541	1.79904	1.74448	1.64580	1.62298
CAC40	<i>EAM</i>	0.19881	0.16971	0.17126	0.19721	0.18449
	<i>ERAM</i>	0.99006	1.40107	1.24897	0.85556	0.93354
	<i>ECM</i>	1.64652	1.59384	1.54457	1.50544	1.46122
DJIA	<i>EAM</i>	0.12145	0.10238	0.10325	0.11904	0.11187
	<i>ERAM</i>	1.01838	1.48978	1.31816	0.90245	0.98562
	<i>ECM</i>	0.75014	0.72915	0.70574	0.67493	0.66431
NASDAQ	<i>EAM</i>	0.35119	0.29943	0.30139	0.34781	0.32514
	<i>ERAM</i>	0.98499	1.51906	1.33819	0.89116	0.97243
	<i>ECM</i>	6.01311	5.81321	5.62610	5.34070	5.26324
SP500	<i>EAM</i>	0.12535	0.10685	0.10773	0.12395	0.11613
	<i>ERAM</i>	1.03175	1.58017	1.37989	0.91655	0.99637
	<i>ECM</i>	0.63094	0.61005	0.58935	0.56918	0.55962
FTSE100	<i>EAM</i>	0.13197	0.11221	0.11298	0.12893	0.12084
	<i>ERAM</i>	0.99881	1.42196	1.26670	0.86730	0.94600
	<i>ECM</i>	0.62084	0.58991	0.57231	0.54764	0.53259
TC \$/£	<i>EAM</i>	0.02762	0.02369	0.02394	0.02779	0.02612
	<i>ERAM</i>	1.02812	1.72407	1.47908	0.94843	1.03188
	<i>ECM</i>	0.01789	0.01821	0.01768	0.01700	0.01669
TC Euro/\$	<i>EAM</i>	0.03816	0.03269	0.03306	0.03872	0.03621
	<i>ERAM</i>	1.03371	1.71226	1.49785	0.94815	1.03240
	<i>ECM</i>	0.03703	0.03818	0.03726	0.03567	0.03507

Nota: En la primera fila de cada celda, la estimación del error absoluto medio ($\times 1000$), en la segunda fila la estimación del error relativo absoluto medio y en la última, la estimación del error cuadrático medio ($\times 10^7$).

En general, se observa que los resultados de la estimación no paramétrica de la volatilidad local proporciona un mejor ajuste que la estimación paramétrica de un modelo GARCH(1,1). Aunque en algunos casos la ganancia en el ajuste global de la serie de volatilidad no parece ser significativa, la naturaleza robusta del estimador no paramétrico previene frente a observaciones anómalas en el proceso de volatilidad en el ajuste observación por observación.

Finalmente, consideramos la re-estimación no paramétrica de la volatilidad local empleando el estimador $T(VAR)$, donde el parámetro ν de grados de libertad se estima

por máxima verosimilitud a partir de la serie de rendimientos estandarizados empleando el estimador-A. Los resultados se muestran en el siguiente cuadro.

Cuadro 3. Estimación no paramétrica de la volatilidad.

	$\hat{\sigma}_T = \hat{\sigma}_T(n, c_0)$		Estimador-T(VAR) $\hat{\sigma}_T^2(n / \hat{\sigma}_T)$	
	$(n, c_0) = (20,9)$	$(n, c_0) = (10,9)$	$(n, \nu) = (20,5)$	
IBEX35	6.03 (3) [4.09]	4.36 (6) [1.67]	<i>EAM</i>	0.20723
			<i>ERAM</i>	0.91775
			<i>ECM</i>	1.66454
CAC40	6.33 (5) [4.25]	4.46 (3) [1.59]	<i>EAM</i>	0.18585
			<i>ERAM</i>	0.91375
			<i>ECM</i>	1.44122
DJIA	5.52 (6) [2.95]	4.49 (7) [1.66]	<i>EAM</i>	0.11038
			<i>ERAM</i>	0.92845
			<i>ECM</i>	0.61831
NASDAQ	5.29 (6) [2.46]	4.03 (5) [1.14]	<i>EAM</i>	0.33041
			<i>ERAM</i>	0.88068
			<i>ECM</i>	5.20318
SP500	5.02 (5) [2.22]	4.20 (3) [1.35]	<i>EAM</i>	0.11773
			<i>ERAM</i>	0.89046
			<i>ECM</i>	0.52455
FTSE100	6.16 (7) [3.54]	4.29 (7) [1.28]	<i>EAM</i>	0.12051
			<i>ERAM</i>	0.91614
			<i>ECM</i>	0.51208
TC \$/£	4.24 (7) [1.39]	3.31 (8) [0.71]	<i>EAM</i>	0.02859
			<i>ERAM</i>	0.89760
			<i>ECM</i>	0.01739
TC Euro/\$	4.30 (6) [1.46]	3.45 (6) [0.79]	<i>EAM</i>	0.04003
			<i>ERAM</i>	0.88767
			<i>ECM</i>	0.03552

Nota: Entre paréntesis el número de iteraciones hasta la convergencia y entre corchetes los estadísticos *T*-ratio asintóticos para el contraste de la significación nula. Los resultados de los estadísticos *EAM* y *ECM* aparecen multiplicados por factores 10^3 y 10^7 , respectivamente.

6. Conclusiones y extensiones

En este trabajo hemos examinado el comportamiento del estimador no paramétrico de la volatilidad local propuesto por Lally, Randal y Thomson (2001, 2004) en presencia de outliers en el proceso de la volatilidad. Se trata en todo caso de un estudio preliminar en relación con la robustez de este estimador, como alternativa frente a otros procedimientos paramétricos y no paramétricos propuestos recientemente para estimar de forma robusta la volatilidad subyacente a una serie temporal financiera.

De forma general, podemos indicar que el estimador no paramétrico basado en una media móvil doblemente ponderada del cuadrado de las observaciones de la serie en cada punto de la muestra se caracteriza por una influencia limitada ante la presencia de outliers en volatilidad que, en algunos casos, no depende exclusivamente de la magnitud del outlier, de forma que el estimador resulta ser ciertamente resistente a este tipo de

contaminación. Este comportamiento robusto es confirmado a través del estudio de simulación realizado, al menos bajo el diseño considerado. Por otro lado, hemos realizado una aplicación particular a series reales de rentabilidades diarias de un tipo de cambio y de un índice bursátil, proporcionando resultados relativamente satisfactorios.

Consideramos finalmente las siguientes posibles extensiones del presente trabajo:

- θ Estudiar la posibilidad de determinar, o bien de forma automática o bien mediante algún procedimiento dependiente de los datos, el tamaño o extensión de la media móvil (n) puesto que, en general, para los distintas versiones estudiadas del estimador los errores del ajuste no paramétrico son tanto menores cuánto menor es el mismo.
- θ Determinación de la constante de truncamiento para las versiones A y $T(MAD)$ del estimador no paramétrico.
- θ Estudiar los resultados del ajuste dependiendo de la elección de los pesos $\varpi_j(n)$ para la suavidad del ajuste. La estructura de pesos uniformes se ha elegido en este estudio por simplicidad y porque supone una especificación que prima más la precisión que la suavidad del ajuste. Sería deseable entonces estudiar formas alternativas, pero atendiendo al compromiso entre precisión de la estimación y suavidad del ajuste puesto que la relación es inversa: al aumentar la suavidad del ajuste, disminuye la precisión.
- θ Consideración de múltiples outliers en volatilidad.
- θ Estudiar el grado de influencia sobre las estimaciones resultantes en el caso de outliers en volatilidad de tipo distinto al aditivo.
- θ Desarrollar métodos de inferencia apropiados para evaluar la calidad del ajuste, empleando procedimientos de remuestreo.

Referencias

Afonso Rodríguez, J., Cano Fernández, V., Guirao Pérez, G. (2001): Observaciones atípicas y contrastes de no linealidad. *Economía y Finanzas 2001. Libro Homenaje al Profesor D. Fco. P. Calatayud*, 43-61. ARTE Comunicación Visual, S.L.

Bollerslev, T. (1987): "A conditional heteroskedastic time series model for speculative prices and rates of return", *Review of Economics and Statistics*, 69, pp. 542-547.

Bossaerts, P., Hafner, C., Härdle, W. (1996): "Foreign exchange rates have surprising volatility". Discussion paper 68. Humboldt-Universität, Berlin.

Carnero, A., Peña, D., Ruiz, E. (2001): "Outliers and conditional autoregressive heteroscedasticity in time series". *Revista de Estadística*, 53, pp. 143-213.

Doornik, J.A., Ooms, M. (2005): "Outlier detection in GARCH models". Working Paper 24, Nuffield College, University of Oxford.

Engle, R.F. (1982): "Autoregressive conditional heteroskedasticity with estimates of the variance of the United Kingdom inflation", *Econometrica*, 50, pp. 987-1007.

Franses, P.H., van Dijk, D., Lucas, A. (1998): "Short patches of outliers, ARCH and volatility modelling". Econometric Institute report. Erasmus University Rotterdam.

Hafner, C.M. (1998): *Nonlinear Time Series Analysis with Applications to Foreign Exchange Rate Volatility*. Heidelberg: Physica-Verlag. Berlin.

Heid, F. (1996): "Non-parametric volatility estimation of exchange rates and stock prices". Discussion paper no.A-533. University of Bonn.

Hotta, L.K., Tsay, R.S. (1998): "Outliers in GARCH processes". Working paper, Graduate School of Business. University of Chicago.

Lally, M., Randal, J., Thomson, P. (2001): "Non-parametric volatility estimation". Proceedings of the 2nd International Symposium on Business and Industrial Statistics. Yokohama, Japan.

Lally, M., Randal, J., Thomson, P. (2004): "Non-parametric volatility estimation". *Quantitative Finance*, 4(4), 427-440.

Li, J., Kao, C. (2002): "Bounded influence estimation and outlier detection for ARCH/GARCH models with an application to foreign exchange rates". Working paper. Maxwell School of Syracuse University.

Muler, N., Yohai, V.J. (2005): "Robust estimates for GARCH models". Documento de Trabajo noviembre-2005. Instituto de Cálculo. Universidad de Buenos Aires.

Muller, N., Yohai, V.J. (2002): "Robust estimates for ARCH processes", *Journal of Time Series Analysis*, 23(3), pp. 341-375.

Pagan, A., Ullah, A. (1999). *Nonparametric Econometrics*. Cambridge University Press.

Taylor, S.J. (1986). *Modelling Financial Time Series*. John Wiley, Chichester.

Taylor, S.J. (1994). *Modelling stochastic volatility: a review and comparative study*. *Mathematical Finance*, 4, 183-204.

Wu, L.S.Y., J.R.M. Hosking y N. Ravishanker (1993): "Reallocation outliers in time series", *Applied Statistics*, 42(2), pp.301-313.

EL EFECTO DE LAS NOTICIAS FINANCIERAS EN EL MERCADO DE LA BOLSA

MARÍA DEL PILAR PÉREZ GONZÁLEZ

e-mail: mpperez@eco.uva.es

Departamento de Economía Aplicada
UNIVERSIDAD DE VALLADOLID

Área temática: Métodos cuantitativos.

Resumen

El objetivo de este trabajo es la construcción de un modelo para predecir los cambios de tendencia en los precios del mercado de la Bolsa de Madrid a partir de las noticias relativas a las empresas estudiadas. Para ello, extraemos patrones de los comentarios de prensa usando técnicas de procesamiento de lenguaje natural. El modelo se obtiene como una red neuronal entrenada usando estos patrones como inputs y la reacción del mercado como output. Los resultados obtenidos mejoran el paseo aleatorio.

Palabras clave: neural networks, mercado de la bolsa, NLP, predicción.

Abstract

The main of this work is to build a model to predict the evolution of the prices of the stock market by looking at appeared news articles. To this end, some features of news articles are chosen using the latest developments in Natural Language Processing. A predictive model is then constructed as a feedforward neural network trained using this templates as inputs and the reaction of the stock market as output. The model has been tested with a data set of news articles for which results are better than random walk.

Key words: neural networks, stock market, NLP, forecasting.

1. Introducción

Tradicionalmente el análisis del mercado de la bolsa se ha realizado mediante el análisis de la evolución de los precios pasados. Este trabajo parte de la idea de considerar la aparición de comentarios de prensa relativos al conjunto de empresas objeto de estudio como un procedimiento para mejorar las predicciones.

Existen publicaciones que justifican la importancia de considerar la aparición de noticias financieras como un hecho relevante en la evolución del mercado de la bolsa. Lo que pretendemos es construir un modelo que establezca una relación entre ciertas características encontradas en noticias pasadas y la evolución de los precios del mercado, para que al introducir las características de las nuevas noticias nos permita predecir los cambios de tendencia en la evolución de estos precios. Para determinar estas características empleamos patrones lingüísticos y para construir el modelo, redes neuronales.

El problema se plantea como un problema de clasificación, de modo que lo que el modelo predice es la reacción del mercado ante la publicación de una noticia, distinguiendo entre una reacción positiva, negativa o no relevante.

2. Trabajos relacionados

Uno de los primeros trabajos que clasifica las noticias para obtener información acerca de las empresas es el trabajo de Seo y otros (2002). Ellos las clasifican en 5 categorías: buenas, probablemente buenas, neutrales, probablemente malas y malas, en función de la información que proporcionan de la empresa a la que se refieren.

Ellos realizan la clasificación mediante la extracción de lo que denominan FCPs en las noticias. No vamos a entrar en el modo de obtenerlos ni en cómo los emplean, pero sí vamos a señalar una conclusión a la que llegan en su trabajo. Ellos observan que la mayoría de las noticias financieras contienen el nombre o las siglas de numerosas empresas y comprueban y que precisamente las frases en las que éstas aparecen son las más interesantes.

Lavrenko y otros (2000) intentan predecir en el mercado de la bolsa usando modelos de lenguaje en lo que ellos denominan \mathcal{A} nalyst. En primer lugar dividen las series de tiempo en lo que llaman *trends* (tramos con la misma tendencia) y plantean como objetivo de su análisis seleccionar de un conjunto de noticias, aquellas que coinciden con el comienzo de

un trend, es decir, las que determinan un cambio de tendencia.

Los inputs de su trabajo son los artículos relativos a 127 empresas y sus cotizaciones en el mismo período, tomados de biz.yahoo.com. Para determinar cuáles son los artículos relevantes para una empresa específica usan la información proporcionada por Yahoo. Para construir el modelo cuentan las veces que aparecen las palabras y emplean un algoritmo bayesiano para determinar que noticias coinciden con cambios de tendencia. En la evaluación del modelo plantea un sistema de compra y venta con el que concluye que se obtienen resultados positivos. Como sugerencias para futuros trabajos plantean la mejora en la obtención de las características asociadas a las noticias.

Otros autores que han analizado la influencia de las noticias en el mercado de la bolsa son Fawcett y Provost (1999) quienes usando como características de los artículos palabras y pares de palabras obtienen una predicción mejor que el paseo aleatorio.

En este trabajo, la predicción de la influencia de los nuevos artículos en el precio de mercado es vista como una “actividad de control”. Analizan un flujo de datos y asignan alarmas que indiquen actividad positiva en los datos. Consideran que un cambio superior al 10% en la cotización de la empresa es una actividad y una nueva noticia asociada a ese cambio, la alarma.

Ahmad y otros (2002) comprueban la relación entre las veces que aparecen en las noticias algunas palabras que representan buenas o malas noticias y la evolución del FTSE 100.

3. Características de las noticias

3.1 Las palabras como características de las noticias

Como hemos dicho en la introducción, lo que pretendemos es construir un modelo que establezca una relación entre ciertas características encontradas en noticias pasadas y la evolución de los precios del mercado, para que al introducir las características de las nuevas noticias nos permita predecir la evolución de estos precios. Veamos a continuación cómo vamos a asociar características a un artículo.

El método más común para asociar características a un documento es considerar cada palabra como una característica. Algunos autores, para disminuir el número de características proponen usar una lista de parada, pero este método puede eliminar palabras que para

nosotros pueden resultar importantes. Otra técnica que se utiliza es exigirles un mínimo de apariciones.

En lengua inglesa, se han desarrollado algoritmos que agrupan palabras en clases. Los algoritmos de contención, WordNet y The Integral Dictionary (TID), son algunos ejemplos. Los algoritmos de contención identifican palabras con raíz común. Tienen el inconveniente de que pueden agrupar palabras cuyo significado interesa diferenciar.

WordNet¹ es un sistema de referencia léxica en las que las palabras se organizan en nombres, verbos, adjetivos y adverbios, y dentro de cada clase en conjuntos de sinónimos. Una forma de reducir las características de un documento usando WordNet es identificar cada palabra con el conjunto de sinónimos que la incluye (lemma). Esta herramienta se ha desarrollado para otras lenguas incluida la española con el nombre de EuroWordNet², sin embargo este es un recurso limitado porque no contiene ni todas las palabras de cada lengua, ni todos los sentidos posibles.

The Integral Dictionary³ está organizado en conceptos anotados por unas pocas palabras que describen su contenido. Los conceptos son clasificados en categorías. La diferencia fundamental entre TID y WordNet es que en TID las clases no dependen de las palabras de una lengua concreta e incluso es posible crear un concepto sin palabras. Aunque ésta es una herramienta que surgió en primer lugar para la lengua inglesa, como WordNet, también se ha desarrollado para otras lenguas, entre ellas el español⁴

3.2 Relaciones entre las palabras. Patrones lingüísticos

En la sección anterior hemos visto diferentes métodos para asociar características a un documento, basados todos ellos en las palabras y la forma en que podemos agruparlas pero sin tener en cuenta el contexto en que éstas aparecen. Evidentemente, no puede tener el

¹MILLER, G. A., BECKWITH, R., FELLBAUM, C., GROSS, D. Y MILLER, K. J., *Introduction to wordnet: An on-line lexical database*, Journal of Lexicography, 1990.

²VOSSEN, P., *A multilingual database with lexical networks*, Kluwer Academic Publishers, 1998.

³DUTOIT, D. Y NUGUES, P., *A lexical network and an algorithm to find words form definitions*, Proceedings of the 15th European Conference on Artificial Intelligence, pp. 450-454, 2002.

⁴DUTOIT, D. Y POIBEAU, T., *Inferring knowledge from a large semantic network*, artículo online.

mismo efecto una noticia en la que se anuncia que *la Empresa 1 compra la Empresa 2* que otra noticia en la que aparece *la Empresa 2 compra la Empresa 1*, sin embargo las palabras son las mismas en las dos frases.

A continuación nos planteamos como podemos mejorar las características asociadas a una noticia considerando las relaciones entre las palabras mediante la búsqueda de patrones lingüísticos.

4. El modelo. Las redes neuronales

Establecido un método para determinar las características asociadas a una noticia, lo que nos planteamos a continuación es la construcción de un modelo, que tomando como inputs las características de las nuevas noticias nos proporcione información sobre las consecuencias que éstas tendrán en el mercado.

Este problema lo resolveremos entrenando una red neuronal que sirva para clasificar cada noticia en buena, mala o no relevante. En la sección siguiente explicaremos con más detalle el modo en que hemos llevado a cabo esto. A continuación vamos a realizar una breve introducción a las redes neuronales, la herramienta que emplearemos en la construcción del modelo.

Las redes neuronales son capaces de aprender a partir de ejemplos, imitando esquemáticamente el funcionamiento del cerebro, para tratar de reproducir sus capacidades. Para comprender mejor el paralelismo existente entre ambos repasemos brevemente el funcionamiento de los sistemas neuronales biológicos.

4.1 Breve introducción biológica

Santiago Ramón y Cajal en sus investigaciones descubrió que el sistema nervioso está compuesto por una red de células individuales, las neuronas, ampliamente interconectadas entre sí. Posteriormente, los avances de la técnica han permitido profundizar mucho más en el estudio de la neurona. Aunque estas células pueden presentar múltiples formas, muchas de ellas tienen el aspecto de la dibujada en la figura 1, es decir, poseen un cuerpo o soma del que surge un árbol de ramificaciones compuesto por las dendritas, y del que parte una fibra tubular denominada axón, que también se ramifica en su extremo final para conectar con otras neuronas.

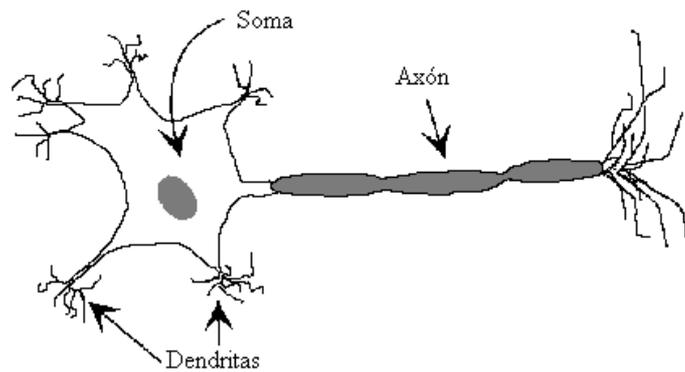


Figura 1: Estructura de una neurona biológica típica

La unión entre dos neuronas se denomina sinapsis y es direccional, es decir, la información fluye siempre en un único sentido. La forma de comunicación más habitual entre las neuronas tiene lugar como resultado de la liberación de unas sustancias por la célula llamada presináptica, denominadas neurotransmisores. Si la neurona postsináptica (la que recibe la señal) posee en las dendritas o en el soma canales sensibles a las sustancias liberadas, las fijará, provocando pequeños potenciales postsinápticos, que pueden ser excitadores o inhibidores. Cuando se ha producido un número suficiente de excitaciones, la suma de los potenciales positivos generados puede elevar el potencial de la neurona hasta sobrepasar lo que se denomina umbral de disparo. En ese momento, se produce un flujo de iones de sodio que provoca la despolarización brusca de la neurona. A continuación la neurona vuelve a la situación original de reposo. Este proceso constituye la generación de un potencial de acción que al propagarse a lo largo del axón da lugar a la transmisión eléctrica de la señal nerviosa.

Durante el desarrollo de un ser vivo, el cerebro se modela, es decir, el individuo adquiere cualidades por la información que le proporciona el medio externo. Esta transformación se realiza mediante el modelado de las intensidades sinápticas, el establecimiento de nuevas conexiones entre las neuronas, la ruptura de otras o la muerte neuronal. Este proceso constituye el aprendizaje.

4.2 Un sistema neuronal artificial

Veamos como se traduce este simplificado esquema del funcionamiento neuronal biológico en el diseño de una red neuronal artificial. Los tres conceptos claves de los sistemas

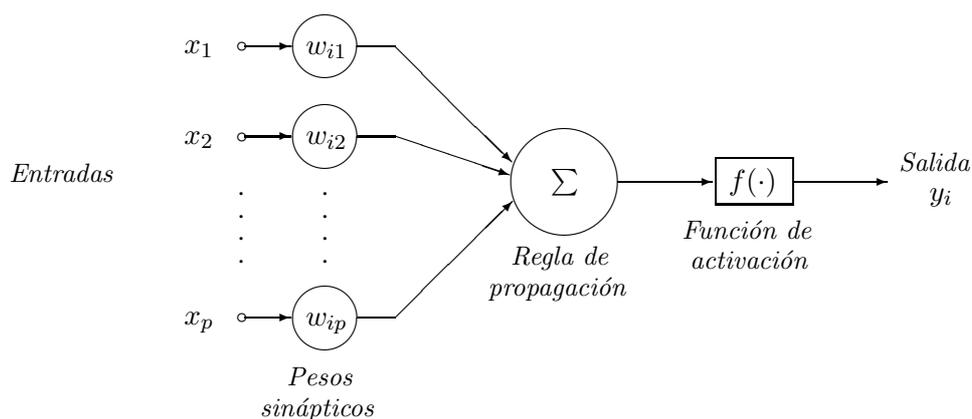


Figura 2: Modelo estándar de una neurona.

nerviosos que las redes neuronales pretenden imitar son: el paralelismo de cálculo, la memoria distribuida y la adaptabilidad al entorno. Para ello, se toma como unidad básica la neurona artificial que se organiza en capas, junto con los nodos de entrada y salida. Formalmente, una red neuronal está compuesta por los siguientes elementos: un conjunto de procesadores elementales o neuronas artificiales, un patrón de conectividad o arquitectura, una dinámica de activaciones, una regla o dinámica de aprendizaje y el entorno donde opera.

Una neurona general se puede representar tal como aparece en la figura 2. En términos matemáticos, podemos describir una neurona por la siguiente ecuación:

$$y_i = f \left(\sum_{j=1}^p w_{ij} x_j \right)$$

con y_i el output de la neurona, x_1, x_2, \dots, x_p los inputs, $w_{i1}, w_{i2}, \dots, w_{ip}$ los pesos sinápticos de la neurona i y $f(\cdot)$ la función de activación. Las entradas y salidas de las neuronas pueden tomar valores continuos o discretos. Dependiendo del tipo de salidas, las neuronas se pueden denominar de tipo McCulloch-Pitts, tipo Ising, de tipo Potts, etc.

La arquitectura es la estructura de conexionado de la red. En general, las neuronas se suelen agrupar en capas: una capa de entrada formada por las neuronas que reciben los datos, una capa de salida cuyas neuronas proporcionan la respuesta de la red, y una o varias capas ocultas en las que se procesa la información (figura 3). Las conexiones sinápticas son direccionales, es decir, la información se propaga en un único sentido, y se

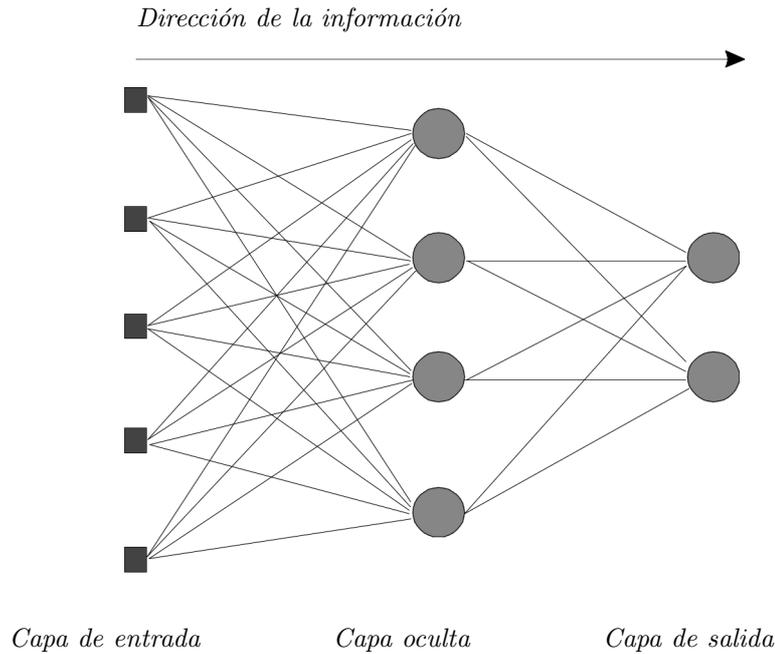


Figura 3: Estructura de una red neuronal

pueden producir entre neuronas de la misma capa o de capas diferentes. En algunas redes, se tienen conexiones en sentido contrario al flujo de información, o incluso de una neurona consigo misma, son las redes denominadas recurrentes. Para una red con una capa oculta, el output se puede representar como

$$\sum_{i=1}^q \xi_i f \left(\sum_{j=1}^p w_{ij} x_j \right).$$

En las redes neuronales puede definirse el aprendizaje como el proceso por el que se produce el ajuste de los parámetros libres de la red a partir de un proceso de estimulación por el entorno que rodea la red. El tipo de aprendizaje viene determinado por la forma en la que dichos parámetros son adaptados. El modo de aprendizaje más convencional consiste en la modificación de los pesos sinápticos siguiendo una regla de aprendizaje. En algunos casos, el aprendizaje conlleva la creación o destrucción de neuronas, con la consiguiente modificación de la estructura de la red.

4.3 El perceptrón multicapa y el algoritmo de retropropagación de errores

La red más comúnmente empleada es el perceptrón multicapa con el algoritmo de retropropagación de errores (*backpropagation*). Se trata de una red de aprendizaje supervisado,

con una estructura similar a la representada en la figura 3. Este tipo de red resulta especialmente interesante porque varios autores, entre los que se encuentran Funahashi (1989), Hornik, Stichcombe y White (1989), han demostrado que el perceptrón multicapa de una única capa oculta puede aproximar hasta el nivel deseado, dentro de un conjunto compacto, cualquier función continua.

El algoritmo de retropropagación de errores consiste, esencialmente, en, fijados unos pesos arbitrarios y dados unos valores iniciales y unos esperados, emplear la comparación entre el valor observado en las neuronas de salida de la red y los valores esperados para modificar los pesos que definen la red, de forma que disminuya el error.

La determinación de una estructura de la red conlleva la elección de una función de activación de las neuronas, además del número de nodos en la capa de entrada, capas ocultas y número de neuronas en cada capa. La aplicación del algoritmo de retropropagación de errores supone la elección de una función de coste que se desea minimizar, de un factor de aprendizaje y del momento. El factor de aprendizaje fija el nivel de tolerancia para determinar el momento en que termina el entrenamiento de la red y el momento es un parámetro entre 0 y 1 que relaciona la actualización de los pesos con la pasada, dando “suavidad” a los cambios de pesos.

5. Una aplicación práctica

5.1 Los datos

Los datos que usamos en nuestro trabajo corresponden a las noticias relacionadas con las empresas que componen el IBEX-35 durante cuatro meses y los precios de estas empresas durante el mismo período, extraídos de yahoo.com cada diez minutos.

Para cada empresa, sea componente del IBEX-35 o no, Yahoo proporciona un conjunto diario de noticias relacionadas. Seleccionar las noticias de esta forma reduce el coste de buscar entre todas las publicadas cada día aquellas que contienen el nombre de las empresas que nos interesan, pero hemos de confiar en que la selección hecha por Yahoo sea suficientemente completa.

Como inputs se toman los precios y las noticias pasadas. Para ello, transformamos los patrones lingüísticos en vectores numéricos, reemplazando cada letra por su posición (con

dos números) en el alfabeto. Un espacio entre palabras se reemplaza por 00.

El output es un vector con dos ceros y un 1 que, en función de la posición del 1, nos indica si el efecto de la noticia es positivo, negativo o no relevante.

5.2 La red neuronal

Para construir el modelo, en primer lugar debemos asociar a cada característica (un input del modelo) el output correspondiente, valor que calcularemos tras analizar los precios del mercado. En este análisis vamos a trabajar suponiendo que los nuevos artículos tienen un efecto inmediato en el mercado y que su influencia sólo afecta a corto plazo. Estas hipótesis de trabajo están basadas en los trabajos de autores como Breatley y Meyers (2000), Chan y otros (2001) o Patell y Wolfson (1984).

Establecidas estas hipótesis de trabajo, la primera cuestión que nos planteamos es cómo determinar cuál es la primera influencia. No sabemos cuando los agentes que están operando en el mercado han leído la noticia, por lo que no nos fijaremos en los precios de mercado sino en la tendencia de la serie. En segundo lugar, es importante aislar el efecto que una noticia ha producido en el mercado. De los métodos que proponen Brown y Warner (1980) hemos escogido comparar los precios de la empresa con los de otras empresas en el mismo momento.

La arquitectura que hemos elegido para obtener la red es el perceptrón multicapa. Consideramos redes con una capa de inputs, una capa oculta y una capa de outputs. Determinamos el número de nodos en la primera capa por el tamaño de los vectores que representan los patrones de las noticias. El número de outputs es 3, uno para cada clase de noticia. El número de nodos en la capa oculta varía entre 1 y el número de inputs. Las redes son entrenadas con el algoritmo de retropropagación de errores aplicado a inputs extraídos aleatoriamente del conjunto de entrenamiento con pesos elegidos aleatoriamente en el intervalo $[-0.5, 0.5]$ y tomando como función de activación la función sigmoidea. La red es seleccionada atendiendo al error cuadrático medio.

5.3 Los resultados

Tratamos de medir la influencia de los artículos en el mercado. Para ello lo que hacemos es clasificar sus características empleando la red neuronal entrenada.

Una primera cuestión a señalar es que el porcentaje de artículos considerados no relevantes es bastante elevado. Esto es debido en algunos casos, a que sus características no cumplen la condición del mínimo de ocurrencia, por lo que no son clasificadas. A pesar de ello, de los resultados obtenidos podemos concluir que nuestro modelo predice mejor que el paseo aleatorio.

Hemos seleccionado un conjunto de noticias que a priori se puede esperar tengan mucha influencia en la evolución de los precios, sin embargo los resultados de la aplicación del modelo a estos artículos no da unos resultados muy diferentes de los otros.

6. Conclusiones

El objetivo de este trabajo es predecir la influencia de las noticias financieras en la evolución de los precios del mercado de la bolsa. Para ello extraemos determinadas características de esas noticias usando técnicas de Procesamiento del Lenguaje Natural. Estas características son usadas como inputs en una red neuronal entrenada para que nos sirva para clasificar las noticias. Dividimos las noticias en buenas, malas y no relevantes, en función de los cambios en la tendencia del mercado como consecuencia de esa noticia.

La aplicación práctica está realizada para las empresas que componen el IBEX-35, pero los métodos desarrollados son aplicables a una única empresa u otros grupos de empresas. Los resultados obtenidos mejoran el paseo aleatorio.

El paso de los patrones lingüísticos a vectores supone un tiempo adicional de computación que podríamos eliminar diseñando una red que permita clasificar inputs dados por palabras o sentencias. En cualquier caso un paso previo que nos proponemos llevar a cabo es la sustitución de los patrones lingüísticos por vectores más pequeños, cambiando cada palabra por un número en lugar de cada letra.

Bibliografía

Ahmad, K., Oliveira, P. C. F., Casey, M. y Taskaya, T. (2002): "Description of events: an analysis of keywords and indexical names", *Third International Conference on Language Resources and Evaluation*.

Breatley, R. y Meyers, S. (2000): *Principles of corporate finance*, McGraw-Hill.

Brown, S. J. y Warner, J. B. (1980): "Measuring security price performance", *Journal of*

Financial Economics, pp. 205-258.

Chan, Y., Chui, A. C. W. y Kwok, C. C. Y. (2001): “The impact of salient political and economic news on the trading activity”, *Pacific-Basin Finance Journal*, Vol.9, pp. 195-217.

Dunham, M. H. (2002): *Datamining - Introductory and advanced topics*, Pearson Education.

Dutoit, D. y Nugues, P. (2002): “A lexical network and an algorithm to find words form definitions”, *Proceedings of the 15th European Conference on Artificial Intelligence*, pp. 450-454.

Dutoit, D. y Poibeau, T. (artículo online): “Inferring knowledge from a large semantic network”.

Fawcett, T. y Provost, F. (1999): “Activity monitoring: Noticing interesting changes in behavior”, *Proceedings on the Fifth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp.53-62.

Funahashi, K. I. (1989): “On the Approximate Realization of Continuous Mappings by Neural Networks”, *Neural Networks*, 2, pp. 183-192.

Hornik, K., Stichcombe, M. y White, H. (1989): “Multilayer Feedforward Networks are Universal Approximators”, *Neural Networks*, 2, pp. 359-366.

Lavrenko, V., Schmill, M. D., Lawrie, D., Ogilvie, P., Jensen, D. y Allan, J. (2000): “Language models for financial news recommendation”, *CIKM*, pp.389-396.

Martín del Brío, B. y Sanz Molina, A. (1997): *Redes Neuronales y Sistemas Borrosos*, Ra-Ma, Madrid.

Miller, G. A., Beckwith, R., Fellbaum, C., Gross, D. y Miller, K. J. (1990): “Introduction to wordnet: An on-line lexical database”, *Journal of Lexicography*.

Patell, J. M. y Wolfson, M. A. (1984): “The intraday speed of adjustment of stock prices to earnings and dividend announcements”, *Journal of Financial Economics*, 13, pp.223-252.

Seo, Y. W., Giampapa, J. A. y Sycara, K. P. (2002): “Text classification for intelligent

agent portfolio management”, *International Joint Conference on Autonomous Agents and Multi-Agent Systems*.

Vossen, P. (1998): *A multilingual database with lexical networks*, Kluwer Academic Publishers.

ANÁLISIS MATEMÁTICO-FINANCIERO DE LAS HIPOTECAS INVERSAS: EL CASO ESPAÑOL

Pablo Castillo Villegas
E-mail: pcastillo@ual.es

Salvador Cruz Rambaud
E-mail: scruz@ual.es

Departamento de Dirección y Gestión de Empresas
UNIVERSIDAD DE ALMERÍA

Área temática: Métodos cuantitativos.

Resumen

El proyecto de ley por el que se modifica la Ley 2/1981, de 25 de marzo, de regulación del mercado hipotecario y otras normas del sistema hipotecario y financiero, regula por primera vez en el mercado financiero español la denominada hipoteca inversa.

El legislador en el preámbulo al proyecto de ley define la hipoteca inversa “*como un préstamo o crédito hipotecario del que el propietario de la vivienda realiza disposiciones, normalmente periódicas, aunque la disposición pueda ser de una sola vez, hasta un importe máximo determinado por un porcentaje del valor de tasación en el momento de constitución. Cuando se alcanza dicho porcentaje, el mayor dependiente deja de disponer de la renta y la deuda sigue generando intereses. La recuperación por parte de la entidad del crédito dispuesto más los intereses se produce normalmente de una vez cuando fallece el propietario, mediante la cancelación de la deuda por los herederos o la ejecución de la garantía hipotecaria por parte de la entidad de crédito*”.

La restricción de este producto al mundo anglosajón, Reino Unido y EE.UU. principalmente, junto al reducido número de contrataciones, explican la escasa atención que ha suscitado el estudio del modelo al que responde la hipoteca inversa.

La experiencia internacional junto a la parquedad de la norma española desarrollada en sólo diez breves puntos incorporados a una Disposición Adicional, justifican el interés del estudio de un producto financiero de enorme impacto social por estar destinado a personas mayores de 65 años o afectadas de dependencia severa o gran dependencia.

Palabras clave: Hipoteca inversa, préstamo, amortización, renta actuarial, riesgo, tasación.

1. Introducción.

La hipoteca inversa es un préstamo o crédito hipotecario destinado a propietarios de inmuebles que desean acceder a un mayor nivel de renta sin enajenar su patrimonio inmobiliario. Producto financiero destinado, en su concepción original, a la población de mayor edad, en la actualidad se ha extendido a otros segmentos de la sociedad.

Aunque se asocia a la disposición de una renta periódica, la hipoteca inversa también presenta en otras modalidades. En todos los casos, el producto financiero se caracteriza por respetar la voluntad del propietario de no enajenar su vivienda habitual, ofreciendo a los herederos la posibilidad de recuperarla, bajo determinadas condiciones, al fallecimiento del beneficiario.

En 2010 la población mayor de 70 años supondrá el 12,16% del total nacional¹, constituyendo un grupo social que demanda soluciones específicas a problemas concretos. Dado que los valores inmobiliarios están fuertemente ligados a ciclos alcistas prolongados, algunos propietarios se enfrentan a la paradoja de ver mermados sus ingresos como consecuencia del acceso a la jubilación, mientras que la vivienda que habitan se revaloriza de modo continuado. La hipoteca inversa facilita al propietario disponer de las potenciales plusvalías generadas por la venta de su vivienda, sin desprenderse de ésta.

En la actualidad son tres las entidades financieras que comercializan hipotecas inversas (Caixa Terrassa, Caixa Sabadell e Ibercaja), número que debería incrementarse tras el proceso parlamentario de aprobación del Proyecto de Ley de reforma del mercado hipotecario. El Consejo de Ministros remitió el Proyecto a las Cortes el día 8 de marzo de 2007; en la Exposición de motivos expresa el ambicioso objetivo de alcanzar la *"satisfacción del incremento de las necesidades de renta durante los últimos años de vida"*². El Proyecto mejora respecto al Anteproyecto al incorporar en la misma Exposición de motivos una detallada definición de hipoteca inversa que reproducimos en su totalidad por su interés:

¹ Proyección de población sobre Base Censo 2001. Instituto Nacional de Estadística.

² Proyecto de Ley por la que se modifica la Ley 2/1981, de 25 de marzo, de regulación del mercado hipotecario y otras normas del sistema hipotecario y financiero, de regulación de las hipotecas inversas y el seguro de dependencia y por la que se establece determinada norma tributaria, de 2 de marzo de 2007.

“La hipoteca inversa regulada en esta Ley se define como un préstamo o crédito hipotecario del que el propietario de la vivienda realiza disposiciones, normalmente periódicas, aunque la disposición pueda ser de una sola vez, hasta un importe máximo determinado por un porcentaje del valor de tasación en el momento de la constitución. Cuando se alcanza dicho porcentaje, el mayor o dependiente deja de disponer de la renta y la deuda sigue generando intereses. La recuperación por parte de la entidad del crédito dispuesto más los intereses se produce normalmente de una vez cuando fallece el propietario, mediante la cancelación de la deuda por los herederos o la ejecución de la garantía hipotecaria por parte de la entidad de crédito”.

Pero, ¿es la hipoteca inversa un producto interesante para las entidades financieras? En la Sección 2, expondremos los tipos de riesgo asociados al producto y las posibles medidas de control que se pueden adoptar.

¿Es previsible el éxito de la fórmula en el caso español? Para responder a esta pregunta examinamos, en la Sección 3, los antecedentes internacionales, centrándonos en los casos del Reino Unido y los Estados Unidos.

Una vez tratadas sus características financieras y los modelos existentes, desarrollamos matemáticamente, en la Sección 4, su operativa, planteando el modelo general y su aplicación a un caso concreto.

Por último, en el apartado de conclusiones, sintetizamos las principales ideas articuladas a través de texto, con especial atención al caso español.

2. La hipoteca inversa: tipos de riesgo y control.

La hipoteca inversa es un préstamo con garantía de la vivienda habitual, amortizable sólo tras el fallecimiento del último de los propietarios o inquilinos. Los propietarios mantienen la propiedad y los herederos disponen de la opción de retener la propiedad satisfaciendo el total dispuesto más los intereses generados.

La disposición del préstamo se establece a partir de tres modalidades básicas:

1. De ocupación o tenencia: un pago periódico asociado a la ocupación de la vivienda habitual por el prestatario.

2. Por período limitado: mediante pagos periódicos por un espacio limitado de tiempo, o incluso toda la vida, posponiendo el cobro de cualquier importe adicional al fallecimiento del beneficiario.
3. De línea de crédito.

Los distintos tipos incluyen variantes tales como pagos mensuales, anuales o un solo pago. El riesgo es, en un entorno de información asimétrica, la posibilidad de daño o pérdida derivada de una operación o actividad. El objetivo de las medidas de control del riesgo es minimizar la probabilidad de que la pérdida se haga efectiva.

Las entidades financieras, como prestamistas, se enfrenta a una extensa y diversa tipologías de riesgos. Montero Pérez (1993)³ los agrupa en:

- Riesgos financieros, que incluyen los riesgos de crédito, interés, mercado, liquidez, cambio y solvencia. Son los riesgos propios del negocio bancario y aquellos en los que hemos de centrarnos.
- Riesgo de entorno, afectan al desarrollo de la actividad económica en un mayor o menor grado. Son factores externos de tipo económico, político, social o de regulación, entre otros.
- Riesgos de servicio, derivados de las relaciones que se establecen entre entidad y cliente (tecnológico, operacional, etcétera).

La hipoteca inversa, como producto financiero, combina tres tipos de riesgo financiero con las peculiaridades a las que haremos mención.

En primer lugar, el **riesgo de interés** asociado a la incertidumbre sobre la evolución futura de los tipos de interés. Siendo un riesgo tradicionalmente asociado a la actividad bancaria, la doctrina sobre su “*sensibilidad, medida, control y gestión*”⁴ es exhaustiva y las técnicas de medida, control y gestión, variadas.

La sensibilidad es el grado de variación del balance de la entidad ante alteraciones en las condiciones externas. La hipoteca inversa es una cartera que contiene un activo muy sensible al tipo de interés y un pasivo menos sensible, entendiendo el producto como la

³ Clasificaciones más recientes, como la de Pérez Ramírez y Calvo González-Vallinas (2006), coinciden en la enumeración de los riesgos financieros (riesgo de crédito, tipos de interés, mercado, cambio y liquidez) como los más destacables y característicos de la actividad bancaria.

⁴ Cardona (1993).

combinación de activo y un pasivo. El activo es el pago final y el pasivo las entregas al beneficiario.

El grado de sofisticación de las técnicas e instrumentos de medición del riesgo de interés es muy alto, como consecuencia de su importancia en el negocio bancario. Entre las principales técnicas, destacamos las denominadas de *gap* (brecha o desfase), duración o la utilización de modelos de simulación.

El control del riesgo de interés se formaliza a través de instrumentos de cobertura, arbitraje, gestión de conjuntos de instrumentos, etcétera.

Las entidades financieras deben simultáneamente establecer sistemas que adapten la composición de la estructura del balance a la evolución de los tipos de interés. Las dificultades a las que han de enfrentarse los prestamistas son comunes a otros productos financieros, siendo las modalidades de tipo de interés las habituales: tipo de interés fijo o variable.

La gestión interna del riesgo por parte de las entidades financieras implica establecer los necesarios controles que eviten situaciones de riesgo relacionadas con la actividad.

Boehm, Ehrhardt y Michael (1994) concluyen que, en el caso de tipos de interés fijos y en comparación a otros productos financieros, la hipoteca inversa presenta un riesgo de interés mucho mayor.

Las hipotecas inversas generan importantes pasivos fuera de balance (*off-balance-sheet liabilities*) en aquellos casos en los cuales los tipos interés crecen por encima de la tasa fija de ésta. Asimismo, las hipotecas inversas exponen a sus prestatarios a la volatilidad de los precios debido al riesgo de interés.

La edad del prestatario como variable asociada a la **longevidad del prestatario** y englobada en los riesgos de crédito. Dado que el préstamo no es reembolsable hasta que el propietario fallezca, es necesaria la aplicación de técnicas actuariales que estimen la longevidad del beneficiario, siendo evidente evaluar el riesgo inherente a un cálculo de este tipo. El prestatario puede vivir mucho más de lo estimado por el análisis actuarial, ¿qué consecuencias pueden derivar de esta situación? La suma del capital prestado, en cualquiera de las modalidades habituales, más los intereses puede llegar a superar el valor de la propiedad, dando lugar a un patrimonio negativo (*negative equity*). El riesgo

de crédito se define como la posibilidad de que los fondos prestados en la actividad financiera no sean reembolsados por el prestatario en la fecha y forma acordados según las condiciones contratadas, es por tanto un riesgo inherente a cualquier operación o producto financiero. La posibilidad de patrimonio negativo es un riesgo contemplado por las entidades financieras que, en el caso de la hipoteca inversa, tiene características particulares ya que se encuentra, en muchos casos, asociado a una menor o nula capacidad de generación de recursos por el beneficiario como clase pasiva (jubilado o dependiente) y que puede derivar en situaciones asimilables a la suspensión de pagos.

Las entidades financieras han optado en la práctica por la cobertura del riesgo asociado a la longevidad mediante la gestión por aseguradoras.

El **riesgo asociado al valor del inmueble** (riesgo de crédito) debe minimizarse previo análisis de los factores actuantes. Como consecuencia de una evolución negativa de los valores inmobiliarios junto a la supervivencia del prestatario por encima de los cálculos actuariales (riesgo de longevidad), es posible la existencia de patrimonios negativos en el momento de realización de la deuda.

En el caso de la hipoteca inversa el capital prestado actual equivale al valor futuro del bien que sirve como garantía de la deuda. A diferencia del préstamo hipotecario habitual en el que hay una disposición inicial del capital según un valor tasado y posteriores reintegros (sujetos a la evolución de los tipos de interés), en el caso de la hipoteca inversa las disposiciones son periódicas y el reembolso del importe de la deuda más los intereses generados se produce al final del período (bajo el supuesto de no supervivencia por encima de la edad fijada en los cálculos actuariales).

El saldo acumulado de un préstamo se incrementa gradualmente a través del tiempo siguiendo un patrón contrario a las hipotecas habituales.

Si el prestamista establece contractualmente un límite al valor total del préstamo más intereses, basado en una tasación inicial de mercado, los pagos cesarían al alcanzar el importe determinado (en general, debido a la supervivencia del prestatario más allá de la fecha considerada en los cálculos actuariales). Esta medida no resuelve la imposibilidad de ejecutar el préstamo hasta el fallecimiento del prestatario, lo que, de prolongarse y entrar en combinación con un ciclo de descenso de los valores inmobiliarios, puede derivar en una situación de patrimonio negativo.

El estudio del riesgo en la hipoteca inversa ha de comprender la combinación y posibles escenarios resultantes del riesgo de interés y longevidad del prestatario, e incorporar el vinculado al valor del inmueble.

El valor de una hipoteca inversa estará, en consecuencia, en función de los tipos de interés (incluyendo las provisiones por pagos anticipados), la probabilidad de amortización anticipada y el valor (y porcentaje de revalorización) de la vivienda.

3. Antecedentes internacionales: Reino Unido y EE.UU.

Las hipotecas inversas o *reverse mortgages* han sido comercializadas en el Reino Unido desde mediados de los años sesenta, habiéndose extendido por los países anglosajones. En este apartado, comparamos las variantes del Reino Unido y los EE.UU.; aunque en un principio se podría afirmar la relación inequívoca entre los hábitos hipotecarios anglosajones y el éxito de productos financieros innovadores, comprobaremos que existen diferencias considerables entre ambos modelos, así como la relativa popularidad de la pensión hipotecaria.

Aún así, el marco regulatorio y social es determinante para la difusión de la implantación de un producto financiero que provoca rechazo, en parte, por la asociación con la renta vitalicia de la venta y la renuncia a las potenciales plusvalías futuras. Es inevitable preguntarse, ¿por qué no se ha extendido la hipoteca inversa a la Europa continental? Un mayor índice de movilidad y una menor preferencia por la propiedad de la vivienda habitual, impulsaría el desarrollo de la hipoteca vitalicia como complemento de los ingresos de unas personas que se encuentran en situación de especial vulnerabilidad.

3.1. Comparativa de control de tipos de riesgo.

En el epígrafe anterior enumerábamos los tres tipos de riesgo asociados a los tipos de interés, longevidad del prestatario y valor de la propiedad. El siguiente cuadro sintetiza las diferencias básicas entre ambos modelos.

	Reino Unido	EE.UU.
Tipo de interés	Tipo fijo	Tipo variable
Longevidad	Gestionado por aseguradoras	Gestionado por aseguradoras
Valor de la propiedad	Prevalece el libre mercado	El gobierno central a través de la Federal Housing Administration (FHA) ha implantado un seguro de garantía adicional

Tabla 1. Fuente: Elaboración propia a partir de Taffin (2004).

La hipoteca inversa en EE.UU. se comercializa con un mayor grado de intervencionismo por parte de la administración, mientras que en el Reino Unido se opta por la autorregulación. Un ejemplo significativo son las medidas adoptadas en ambos casos para mitigar la imagen negativa asociada al producto, en parte debido a las negativas experiencias en los inicios con algunos tipos de hipotecas *flotantes* que dieron lugar a patrimonios negativos. En el Reino Unido las entidades financieras siguen un código de buena práctica, mientras que en los EE.UU. es un experto externo el responsable de asesorar y guiar al cliente en su negociación con la entidad.

El escaso número de contrataciones⁵ (ver tabla 2) desincentiva a las entidades, a la vez que aumenta el coste del producto y dificulta la implantación de controles y medidas específicas de gestión del riesgo.

País	Número de préstamos contratados (2004)
Reino Unido	26.000
EE.UU	35.000

Tabla 2: Elaboración propia a partir de Taffin (2004).

4. Modelo matemático-financiero: un caso práctico.

En un primer punto desarrollamos el planteamiento general de un modelo tipo y la tabla de símbolos de conmutación para datos de mortalidad españolas de ambos sexos.

A continuación, en un segundo apartado, exponemos un caso práctico basado en las anteriores tablas.

⁵ Los datos son muy significativos si consideramos la diferencia de población total existente entre ambos países.

4.1. Planteamiento general.

Consideremos el caso de una persona de edad actuarial x que tiene una vivienda cuyo valor de tasación es V . El banco X está dispuesto a concederle una hipoteca inversa por importe de un tanto por uno t de su valor de tasación ($P = t \cdot V$). El prestatario puede optar por cualquiera de las siguientes tres modalidades:

1. Disfrutar de una renta vitalicia y anual, que, a su vez, puede ser de los siguientes tipos:

- a. Pospagable. En este caso, planteamos $P = a \cdot a_x = a \frac{N_{x+1}}{D_x}$, de donde

podemos obtener el valor de a .

- b. Prepagable. En este caso, planteamos $P = a \cdot \ddot{a}_x = a \frac{N_x}{D_x}$, de donde

podemos obtener el valor de a .

2. Disfrutar de una renta temporal y anual (n años), que, a su vez, puede ser de los siguientes tipos:

- a. Pospagable. En este caso, planteamos $P = a \cdot a_{x:n} = a \frac{N_{x+1} - N_{x+n+1}}{D_x}$, de

donde podemos obtener el valor de a .

- b. Prepagable. En este caso, planteamos $P = a \cdot \ddot{a}_{x:n} = a \frac{N_x - N_{x+n}}{D_x}$, de

donde podemos obtener el valor de a .

3. Disfrutar de una renta vitalicia y fraccionada de frecuencia m , que, a su vez, puede ser de los siguientes tipos:

- a. Pospagable. En este caso, planteamos $P = a \cdot a_x^{(m)} \approx a \left(a_x + \frac{m-1}{2m} \right)$, de

donde podemos obtener el valor de a .

b. Prepagable. En este caso, planteamos $P = a \cdot \ddot{a}_x^{(m)} = a \left(\ddot{a}_x - \frac{m-1}{2m} \right)$, de

donde podemos obtener el valor de a .

Los símbolos D_x y N_x se llaman símbolos de conmutación y pueden calcularse a partir de las tablas de mortalidad de la población a la que va dirigida los cálculos actuariales. En este caso, hemos utilizado las tablas de mortalidad de la población española 1998-1999 para varones, mujeres y ambos sexos. En efecto, si l_x es el número de supervivientes de edad x de una población inicial de 100.000 personas, podemos calcular los símbolos de conmutación de la siguiente forma:

- $D_x = (1+i)^{-x} l_x$, siendo i el tipo de interés técnico.
- $N_x = D_x + D_{x+1} + \dots + D_w$, siendo w la edad máxima que ninguna persona podrá alcanzar.

A continuación, presentamos los símbolos de conmutación para las tablas de mortalidad españolas de ambos sexos (tipo de interés técnico: 5%):

x	l_x	D_x	N_x
0	100.000	100.000,00	2.028.688,38
1	99.526	94.786,67	1.928.688,38
2	99.483	90.234,01	1.833.901,71
3	99.457	85.914,70	1.743.667,70
4	99.438	81.807,89	1.657.753,00
5	99.419	77.897,39	1.575.945,11
6	99.401	74.174,56	1.498.047,73
7	99.386	70.631,77	1.423.873,17
8	99.372	67.258,88	1.353.241,39
9	99.357	64.046,41	1.285.982,51
10	99.342	60.987,37	1.221.936,11
11	99.325	58.073,27	1.160.948,74
12	99.309	55.298,97	1.102.875,46
13	99.294	52.657,73	1.047.576,50
14	99.278	50.142,14	994.918,77
15	99.255	47.743,35	944.776,63
16	99.226	45.456,57	897.033,28
17	99.188	43.275,40	851.576,71
18	99.140	41.194,72	808.301,31
19	99.087	39.212,09	767.106,59
20	99.033	37.324,50	727.894,50

21	98.978	35.527,40	690.570,01
22	98.919	33.815,45	655.042,61
23	98.859	32.185,65	621.227,16
24	98.798	30.634,09	589.041,51
25	98.737	29.157,31	558.407,42
26	98.676	27.751,71	529.250,11
27	98.612	26.413,06	501.498,40
28	98.544	25.137,95	475.085,34
29	98.469	23.922,68	449.947,39
30	98.390	22.765,23	426.024,71
31	98.304	21.662,22	403.259,48
32	98.214	20.611,80	381.597,27
33	98.115	19.610,49	360.985,47
34	98.010	18.656,67	341.374,98
35	97.902	17.748,68	322.718,30
36	97.791	16.884,34	304.969,62
37	97.674	16.061,09	288.085,28
38	97.549	15.276,70	272.024,19
39	97.414	14.529,10	256.747,50
40	97.272	13.817,07	242.218,40
41	97.121	13.138,68	228.401,33
42	96.964	12.492,81	215.262,64
43	96.795	11.877,17	202.769,84
44	96.611	11.290,09	190.892,67
45	96.409	10.729,99	179.602,58
46	96.195	10.196,35	168.872,59
47	95.963	9.687,39	158.676,24
48	95.715	9.202,24	148.988,85
49	95.449	8.739,68	139.786,61
50	95.153	8.297,70	131.046,93
51	94.833	7.875,99	122.749,23
52	94.497	7.474,37	114.873,24
53	94.138	7.091,40	107.398,87
54	93.739	6.725,09	100.307,47
55	93.297	6.374,65	93.582,38
56	92.833	6.040,90	87.207,73
57	92.361	5.723,99	81.166,83
58	91.826	5.419,84	75.442,84
59	91.225	5.127,97	70.023,01
60	90.615	4.851,12	64.895,04
61	89.973	4.587,38	60.043,92
62	89.208	4.331,79	55.456,53
63	88.377	4.087,08	51.124,75
64	87.502	3.853,92	47.037,66
65	86.559	3.630,85	43.183,74
66	85.519	3.416,40	39.552,90
67	84.395	3.210,95	36.136,49
68	83.183	3.014,13	32.925,54
69	81.855	2.824,77	29.911,41
70	80.405	2.642,60	27.086,64

71	78.856	2.468,28	24.444,03
72	77.178	2.300,72	21.975,75
73	75.343	2.139,07	19.675,03
74	73.334	1.982,88	17.535,97
75	71.174	1.832,84	15.553,08
76	68.807	1.687,51	13.720,25
77	66.233	1.547,03	12.032,74
78	63.445	1.411,34	10.485,71
79	60.526	1.282,29	9.074,37
80	57.455	1.159,27	7.792,07
81	54.205	1.041,61	6.632,81
82	50.723	928,29	5.591,19
83	47.064	820,31	4.662,91
84	43.192	716,97	3.842,60
85	39.202	619,75	3.125,63
86	35.126	528,87	2.505,87
87	31.034	445,01	1.977,00
88	26.940	367,91	1.532,00
89	22.958	298,6	1.164,09
90	19.158	237,31	865,49
91	15.652	184,65	628,18
92	12.594	141,5	443,53
93	9.856	105,46	302,04
94	7.412	75,53	196,57
95	5.291	51,35	121,04
96	3.542	32,74	69,69
97	2.195	19,32	36,95
98	1.242	10,41	17,63
99	632	5,05	7,21
100	285	2,17	2,17

Fuente: Instituto Nacional de Estadística (INE).

4.2. Caso práctico.

Supongamos que:

- $x = 65$ años.
- $i = 5\%$.
- Renta vitalicia prepagable.
- $P = 180.000$ euros.

En este caso, planteamos:

$$180.000 = a \frac{N_{65}}{D_{65}} = a \frac{43.183,74}{3.630,85},$$

de donde $a = 15.138,77$ euros anuales.

5. Conclusiones.

La reciente regulación en España de las hipotecas inversas va a suponer la implantación de un producto financiero que va a permitir complementar la pensión de jubilación de muchas familias que ven cómo, paulatinamente, su pensión de jubilación va perdiendo cada vez más poder adquisitivo, mientras que su vivienda habitual se va revalorizando como consecuencia del conocido *boom* inmobiliario. La hipoteca inversa viene a rellenar este desfase entre una pensión de jubilación a la baja (en términos reales) y el constante aumento del valor de la vivienda. Con este producto financiero, el propietario del inmueble puede obtener un préstamo o un crédito con garantía hipotecaria y continuar como titular del inmueble hasta su fallecimiento, a menos que no se produzca la devolución del principal más los intereses por parte de los herederos del inmueble.

En esta ponencia hemos presentado la descripción de este producto novedoso y su planteamiento matemático-financiero, quedando por resolver parte de la casuística que puede presentarse y los riesgos inherentes al producto.

Anexo: Tablas de mortalidad de la población de España 1998-1999.

	Varones			Mujeres		
	I_x	D_x	N_x	I_x	D_x	N_x
0	100.000	100.000,00	2.015.626,83	100.000	100.000,00	2.042.004,22
1	99.487	94.749,52	1.915.626,83	99.567	94.825,71	1.942.004,22
2	99.446	90.200,45	1.820.877,31	99.523	90.270,29	1.847.178,50
3	99.418	85.881,01	1.730.676,86	99.499	85.950,98	1.756.908,21
4	99.399	81.775,80	1.644.795,85	99.479	81.841,62	1.670.957,23
5	99.380	77.866,83	1.563.020,05	99.460	77.929,51	1.589.115,61
6	99.358	74.142,47	1.485.153,22	99.447	74.208,88	1.511.186,10
7	99.338	70.597,66	1.411.010,75	99.437	70.668,02	1.436.977,22
8	99.321	67.224,36	1.340.413,09	99.426	67.295,43	1.366.309,20
9	99.304	64.012,24	1.273.188,72	99.414	64.083,15	1.299.013,77
10	99.286	60.952,99	1.209.176,48	99.401	61.023,59	1.234.930,62
11	99.268	58.039,94	1.148.223,49	99.386	58.108,94	1.173.907,02
12	99.250	55.266,11	1.090.183,54	99.372	55.334,05	1.115.798,09
13	99.233	52.625,38	1.034.917,43	99.359	52.692,20	1.060.464,04
14	99.212	50.108,80	982.292,05	99.347	50.176,99	1.007.771,84
15	99.184	47.709,20	932.183,25	99.331	47.779,91	957.594,86
16	99.147	45.420,38	884.474,05	99.310	45.495,06	909.814,95
17	99.097	43.235,69	839.053,67	99.283	43.316,84	864.319,89
18	99.028	41.148,18	795.817,98	99.257	41.243,33	821.003,05
19	98.953	39.159,06	754.669,80	99.230	39.268,68	779.759,71
20	98.872	37.263,82	715.510,73	99.204	37.388,94	740.491,03
21	98.788	35.459,20	678.246,92	99.178	35.599,19	703.102,09
22	98.697	33.739,56	642.787,72	99.153	33.895,44	667.502,90
23	98.605	32.102,96	609.048,16	99.126	32.272,58	633.607,46
24	98.514	30.546,03	576.945,20	99.097	30.726,80	601.334,88
25	98.423	29.064,58	546.399,17	99.068	29.255,05	570.608,08
26	98.332	27.654,96	517.334,59	99.038	27.853,52	541.353,03
27	98.236	26.312,35	489.679,62	99.008	26.519,13	513.499,51
28	98.133	25.033,10	463.367,28	98.975	25.247,89	486.980,38
29	98.022	23.814,08	438.334,17	98.939	24.036,87	461.732,49
30	97.904	22.652,78	414.520,09	98.900	22.883,23	437.695,62
31	97.778	21.546,31	391.867,31	98.857	21.784,08	414.812,39
32	97.645	20.492,38	370.321,00	98.812	20.737,30	393.028,31
33	97.499	19.487,37	349.828,62	98.761	19.739,61	372.291,02
34	97.345	18.530,09	330.341,25	98.707	18.789,35	352.551,41
35	97.188	17.619,24	311.811,16	98.648	17.883,92	333.762,06
36	97.028	16.752,60	294.191,92	98.589	17.022,12	315.878,13
37	96.857	15.926,74	277.439,32	98.527	16.201,35	298.856,01
38	96.675	15.139,82	261.512,57	98.459	15.419,21	282.654,66
39	96.482	14.390,09	246.372,75	98.385	14.673,92	267.235,45
40	96.283	13.676,58	231.982,66	98.301	13.963,23	252.561,53
41	96.074	12.997,04	218.306,07	98.209	13.285,87	238.598,30
42	95.858	12.350,31	205.309,03	98.113	12.640,84	225.312,43
43	95.622	11.733,24	192.958,72	98.010	12.026,26	212.671,59
44	95.369	11.144,95	181.225,48	97.899	11.440,61	200.645,33
45	95.090	10.583,19	170.080,53	97.774	10.881,90	189.204,72
46	94.794	10.047,85	159.497,34	97.644	10.349,94	178.322,81
47	94.478	9.537,48	149.449,50	97.497	9.842,25	167.972,87

48	94.139	9.050,72	139.912,02	97.343	9.358,76	158.130,63
49	93.767	8.585,67	130.861,29	97.186	8.898,73	148.771,87
50	93.348	8.140,29	122.275,62	97.014	8.459,98	139.873,14
51	92.898	7.715,29	114.135,33	96.827	8.041,60	131.413,16
52	92.427	7.310,64	106.420,04	96.628	7.642,92	123.371,56
53	91.918	6.924,17	99.109,40	96.421	7.263,38	115.728,64
54	91.352	6.553,84	92.185,23	96.191	6.901,00	108.465,26
55	90.721	6.198,64	85.631,39	95.942	6.555,37	101.564,25
56	90.055	5.860,13	79.432,75	95.685	6.226,49	95.008,88
57	89.378	5.539,12	73.572,62	95.422	5.913,69	88.782,39
58	88.616	5.230,37	68.033,50	95.117	5.614,08	82.868,71
59	87.766	4.933,53	62.803,13	94.768	5.327,13	77.254,62
60	86.896	4.652,02	57.869,60	94.423	5.054,98	71.927,50
61	85.979	4.383,74	53.217,58	94.053	4.795,41	66.872,51
62	84.889	4.122,07	48.833,84	93.612	4.545,64	62.077,10
63	83.709	3.871,21	44.711,77	93.127	4.306,75	57.531,46
64	82.475	3.632,51	40.840,56	92.612	4.078,99	53.224,71
65	81.162	3.404,46	37.208,05	92.039	3.860,71	49.145,73
66	79.718	3.184,66	33.803,59	91.406	3.651,58	45.285,02
67	78.157	2.973,62	30.618,93	90.721	3.451,63	41.633,43
68	76.486	2.771,47	27.645,32	89.974	3.260,20	38.181,80
69	74.688	2.577,44	24.873,85	89.118	3.075,42	34.921,60
70	72.768	2.391,61	22.296,41	88.143	2.896,92	31.846,18
71	70.756	2.214,74	19.904,80	87.057	2.724,98	28.949,26
72	68.585	2.044,56	17.690,06	85.870	2.559,83	26.224,28
73	66.244	1.880,74	15.645,50	84.538	2.400,12	23.664,44
74	63.722	1.722,98	13.764,77	83.040	2.245,32	21.264,32
75	61.053	1.572,21	12.041,78	81.380	2.095,66	19.019,00
76	58.184	1.426,98	10.469,58	79.506	1.949,90	16.923,34
77	55.102	1.287,04	9.042,60	77.421	1.808,35	14.973,44
78	51.869	1.153,83	7.755,56	75.053	1.669,56	13.165,09
79	48.620	1.030,05	6.601,73	72.430	1.534,49	11.495,52
80	45.357	915,17	5.571,67	69.514	1.402,58	9.961,03
81	41.951	806,14	4.656,51	66.354	1.275,07	8.558,45
82	38.418	703,09	3.850,37	62.854	1.150,30	7.283,38
83	34.833	607,13	3.147,28	59.056	1.029,32	6.133,08
84	31.177	517,53	2.540,15	54.915	911,57	5.103,76
85	27.563	435,75	2.022,62	50.513	798,57	4.192,19
86	24.036	361,89	1.586,88	45.868	690,61	3.393,62
87	20.672	296,42	1.224,98	41.046	588,58	2.703,01
88	17.463	238,48	928,56	36.083	492,77	2.114,44
89	14.449	187,93	690,07	31.162	405,3	1.621,67
90	11.655	144,37	502,14	26.391	326,9	1.216,37
91	9.214	108,7	357,77	21.862	257,91	889,47
92	7.229	81,22	249,08	17.778	199,74	631,56
93	5.551	59,4	167,86	14.022	150,04	431,82
94	4.118	41,97	108,46	10.601	108,03	281,78
95	2.914	28,28	66,49	7.593	73,69	173,75
96	1.941	17,94	38,21	5.090	47,05	100,05
97	1.202	10,58	20,27	3.154	27,77	53
98	681	5,71	9,69	1.782	14,94	25,24
99	348	2,78	3,98	904	7,22	10,3

100	158	1,2	1,2	405	3,08	3,08
-----	-----	-----	-----	-----	------	------

Fuente: Instituto Nacional de Estadística (INE).

Bibliografía

Asociación Hipotecaria Española (2006): “Anteproyecto de Ley de Modernización del Mercado Hipotecario, de Regulación del Régimen de las Hipotecas Inversas y del Seguro de Dependencia. Posición de la AHE”, www.ahe.es.

Asociación Hipotecaria española (2006): “Resumen de las propuestas de modificación formuladas al Anteproyecto de Ley de Modernización del Mercado Hipotecario, de Regulación del Régimen de las Hipotecas Inversas y del Seguro de Dependencia”, www.ahe.es.

Asociación Hipotecaria Española (2007): “Nota informativa de la AHE en relación con el Proyecto de Ley de Reforma del Mercado Hipotecario”, www.ahe.es.

Boehm, T.P. y Ehrhardt, M.C. (1994): “Reverse mortgages and interest rate risk”, *Journal of the American Real Estate and Urban Economics Association*, 22, 25, pp. 387-408.

Cardona, M. (1993): “Sensibilidad de balance a las variaciones de tipo de interés. Control y gestión del riesgo del tipo de interés”, *Perspectivas del Sistema Financiero* 43; pp. 35-41.

Cruz Rambaud, S. y Valls Martínez, M^a. C. (2003): *Introducción a las matemáticas financieras*, Editorial Pirámide, Madrid.

Klein, L. y Sirmans, C.F. (1994): “Reverse mortgages and prepayment risk”, *Journal of the American Real Estate and Urban Economics Association*, 22, 25, pp. 409-431.

Miceli, T. y Sirmans, C.F. (1994): “Reverse mortgages and borrower maintenance risk”, *Journal of the American Real Estate and Urban Economics Association*, 22, 25, pp. 433-450.

Ministerio de Economía y Hacienda (2007): “Proyecto de Ley por el que se modifica la Ley 2/1981, de 25 de marzo, de regulación del mercado hipotecario y otras normas del sistema hipotecario y financiero, de regulación de las hipotecas inversas y el seguro de dependencia y por la que se establece determinada norma tributaria”, www.meh.es.

Ministerio de Economía y Hacienda (2006): “Anteproyecto de Ley de Modernización del mercado hipotecario, de regulación del régimen de las hipotecas inversas y del seguro de dependencia”, www.meh.es.

Montero Pérez, A. (1993): “Política de riesgos y control de riesgos”, *Perspectivas del Sistema Financiero*, 43; pp. 69-73.

Ortega, R. (1993): “El control de riesgo en las entidades financieras: ¿una tarea pendiente?”, *Perspectivas del Sistema Financiero* 43; pp. 60-65.

Pérez Ramírez, J. y Calvo González-Vallinas, J. (2006): *Instrumentos financieros*, Editorial Pirámide, Madrid.

Taffin, C. (2004): “La hipoteca inversa o vitalicia”, www.ahe.es.

ESTIMACION DE LOS TANTOS DE MORTALIDAD

Juan Escuder Bueno

e-mail: Juan.Escuder@uv.es

Roberto Escuder Vallés

e-mail: Roberto.Escuder@uv.es

Jose Manuel Pavia

e-mail: Jose.M.Paviar@uv.es

Departamento de Economía Aplicada

UNIVERSIDAD DE VALENCIA

Montserrat Guillén

e-mail: mguillen@ub.edu

Departamento de Econometría, Estadística y Economía Española

UNIVERSIDAD DE BARCELONA

Área temática: Métodos Cuantitativos

Resumen

En este trabajo pretendemos aclarar un problema real que se presenta en la construcción de cualquier tabla de mortalidad, sea para poblaciones genéricas o específicas (p.ej., colectivos de asegurados): el cálculo u obtención de los tantos brutos de mortalidad. En efecto, si los datos brutos o no ajustados no son adecuados, por muy sofisticados que sean los métodos de graduación, difícilmente las tablas construidas lo serán. Para ello, se realiza un pequeño bosquejo histórico de los diferentes métodos utilizados para la obtención de los tantos brutos de mortalidad. Además, se profundiza en los aspectos empíricos más relevantes del método actual del censo, con su aplicación al caso español. Se estudian sus propiedades, y se proponen pautas de actuación.

Palabras clave: Tanto bruto de mortalidad, método del censo, tantos ajustados de mortalidad, población de riesgo.

Abstract

In this paper, we try to clarify a question that emerges in the elaboration of almost any life table, both for specific and general populations: the estimation of the crude mortality rates. In fact, if the rough data, or not graduated values, are inadequate the graduation methods used is irrelevant, and its sophistication too. It is really difficult that the elaborated life tables be proper. So, a brief historical analysis about the several procedures used to obtain the crude mortality rates. Besides, we go deep on the more relevant practical issues of the current method, focused on its application on the Spanish case. Its statistical-mathematical properties are studied, and some practicalities are suggested.

Key words: (max. 6). Crude mortality rate, census method, graduated mortality rate, risk population.

1. Introducción

La construcción de cualquier tabla de mortalidad, general o específica (en concreto, de colectivos de asegurados) parte de la experiencia. En este trabajo se abordan los métodos para la obtención de datos muestrales necesarios para la su construcción. Asimismo, se analiza como los diversos métodos de recogida de datos han evolucionado a lo largo del tiempo. De hecho, este problema relegado (probablemente de una manera inconsciente) a segundo plano en otras épocas de pensamiento, debido a su tratamiento eminentemente teórico, está, sin embargo recuperando en la actualidad el puesto que se merece. Los datos muestrales tienen una importancia capital en cuanto a la consiguiente validez de las tablas a construir.

A grandes rasgos y siguiendo a Insolera (1950), los diferentes métodos se pueden clasificar en dos grandes grupos: **por generaciones y por contemporáneos**. Los primeros consisten en la observación longitudinal o de una generación desde su nacimiento hasta su extinción; mientras, los segundos se basan en observaciones transversales, es decir, consisten en la observación de todos los contemporáneos (supervivientes a cada edad) en un momento del tiempo físico (o momento censal). Ninguno de ellos, no obstante, está exento de críticas o debilidades.

Los métodos basados en observar un colectivo desde su nacimiento o constitución hasta su extinción, o sustentados en una **generación determinada**, precisan seguir a cada una de las cabezas, personas o individuos del colectivo desde su nacimiento hasta su muerte. Por lo que, desde el punto de vista de su posibilidad real, se puede afirmar que son muy poco viables. Si todas las dificultades inherentes por la movilidad de las personas (p.ej., pérdidas de su paradero) no fueran suficientes, baste comentar que habríamos de esperar alrededor de 100 o 110 años para construir la correspondiente tabla: hasta el fallecimiento del último superviviente. Por lo que, la tabla nacería ya **obsoleta**.

Ante la práctica imposibilidad real de seguir un grupo a lo largo de toda su existencia, u observación temporal completa, surgieron los criterios basados en los grupos de **contemporáneos**, cuya idea fundamental es partir de los datos proporcionados por una **observación transversal de todos de supervivientes existentes en un momento determinado del tiempo físico**. Sin embargo, los distintos grupos contemporáneos (o cabezas de la misma edad en el momento del censo u observación transversal), pertenecen a diferentes generaciones distintas, por lo que existe un posible error implícito consistente en considerarlos como pertenecientes a una “misma generación”, algo así como una “**generación ficticia**”, formada por aproximadamente un centenar más o menos de **generaciones reales**.

Las dificultades indicadas pueden soslayarse introduciendo como supuestos simplificadores, los conceptos de población **cerrada y estacionaria**. Una población es **cerrada**, cuando una vez constituido el grupo de recién nacidos I_0 , generación o cohorte, no hay nuevos ingresos ni bajas por causas distintas del fallecimiento; concepto que incorpora implícitamente la idea empírica de que a cada individuo se le puede seguir durante todo su ciclo vital. Por otra parte, el concepto de población **estacionaria**, es más teórico, pues su significado es el de población que alcanza la situación en la que el número de nacidos, así como el de supervivientes y fallecidos, a cada edad, se

mantiene constante en cada año a lo largo del tiempo físico. La estacionariedad de las poblaciones es una situación ideal o conceptual a la que se llegaría cuando pudieran constituirse grupos ideales donde todas las probabilidades posibles (de vida y/o muerte) de cada edad fuesen invariables temporalmente. Esta situación evidentemente es conceptual e imposible de alcanzar en la realidad, pero operativamente hablando simplificadora.

Admitiendo la posibilidad de seleccionar un método adecuado para el cálculo de los tantos brutos de mortalidad o estimaciones de los tantos anuales (netos) de mortalidad. Las probabilidades de supervivencia o muerte se obtendrán a partir de los primeros mediante criterios de ajuste o graduación, que deben someterse a diferentes criterios de ajuste estadístico y de contrastes, para su validez. Una vez aceptados se procede a la elaboración o construcción de las correspondientes tablas de mortalidad, en las que en general se parte de un número ficticio de recién nacidos l_0 (en general un múltiplo de 10, habitualmente $l_0 = 100.000$ ó $1.000.000$) y aparecen los valores de las funciones biométricas: l_x , d_x , p_x , q_x , y e_x , para todas las edades $x = 0, 1, 2, \dots, w$.

2.- Precursores.-

Siguiendo al profesor *Intolera (1950)* podemos citar como métodos precursores del actual método del censo los propuestos por Halley (1656-1742), Hermann (1678-1733) y Quetelet (1796-1874), que no eran propiamente actuarios ni incluso demógrafos, sino matemáticos, físicos o astrónomos preocupados por cuestiones demográficas. Todos ellos pretendían eludir las dificultades del método directo por excelencia pero no lo lograron realmente.

El método de Halley, (astrónomo que publicó su primera tabla de mortalidad en 1693 a partir de observaciones sobre defunciones en la ciudad de Breslau) o método de los fallecimientos, es muy simple. En esencia, consiste en observar los fallecidos a cada edad: $d_0, d_1, \dots, d_x, \dots, d_w$, durante un solo año o durante varios años (podríamos decir 4 o 5 años) promediándolos, para posteriormente hallar los supervivientes a cada edad mediante las sumas de fallecidos correspondientes, y por cociente las tasas o tantos brutos. El error implícito en este método elemental es el obtener los supervivientes sumando fallecidos de distintas generaciones, lo cual solo puede admitirse bajo el supuesto de estacionariedad.

Hermann introdujo en el cálculo del denominador de los tantos brutos de mortalidad, listados de nacidos (en vez de fallecidos). Sin embargo, en este caso aparece el problema de rastrearlos a lo largo del tiempo físico, pues para el cálculo de cada uno de los tantos de mortalidad brutos se requieren los listados de nacimientos de todos los años precedentes necesarios. Además en el cómputo de los denominadores dada la expresión encadenada que resulta, pueden introducirse errores acumulados de los diferentes años. No obstante, este criterio ha sido usado para obtener los tantos brutos de mortalidad, correspondientes a las edades infantiles (pues no se requieren muchos años precedentes).

Finalmente el método de Quetelet elimina la dificultad de tener que recurrir a una lista de nacimientos de los años precedentes necesarios (hoy se trataría de 100 o más),

sustituyendo a los nacidos en sucesivos períodos, por supervivientes clasificados por edades observados en un momento del tiempo físico; esto es, incorpora ya la idea de un *método censal o por contemporáneos*. En definitiva compara fallecidos con supervivientes.

El método de Quetelet es el antecedente del método actual del censo aunque contiene errores implícitos entre ellos: (a) Considerar a los supervivientes a cada edad x y momento t , $P^{(t)}(x)$ o supervivientes contemporáneos con edad cumplida x y sin cumplir $(x+1)$, como denominador de los tantos brutos. Ello es como si se tratara del valor empírico de la función biométrica I_x , o supervivientes que alcanzan la edad x procedentes del grupo inicial l_0 , que son dos conceptos distintos. Y, (b) Obtener los fallecidos por diferencia entre los propios stocks $P^{(t)}(x)$ para cada dos edades contiguas $P^{(t)}(x) - P^{(t)}(x+1)$ lo cual sólo tiene sentido en caso de población estacionaria.

Conviene, sin embargo, resaltar que salvo las deficiencias indicadas, podemos considerar al método de Quetelet como el referente inmediato inmediato del actual método del censo, y que propició la necesidad de elaborar censos y estadísticas de movimientos naturales de población correctas.

3.- Método del censo: diferentes criterios. problemáticas específicas que plantean

El método del censo se basa en la Teoría Formal de la Población y al menos inicialmente su interpretación se facilita muchísimo mediante el esquema de Lexis. En este trabajo vamos a considerar tres variantes o enfoques del mismo.

3.1- Método del censo criterio clásico (o criterio I)

La primera variante del método general del censo que vamos a considerar es el método seguido por el Instituto Nacional de Estadística de España (INE), que se basa en los datos por edades *de un solo censo y de las estadísticas de los Movimientos Naturales de la Población del año siguiente y del anterior al censo*. En los cálculos se presupone: (i) la no existencia de movimientos migratorios, o bien que se compensen para cada edad las inmigraciones con las emigraciones, y (ii) la uniformidad de los fallecimientos para edades iguales y superiores a uno, atribuyéndose durante el primer año de vida más intensidad de fallecimiento al principio. A este criterio podríamos calificarlo como algebraico o geométrico.

Inicialmente este criterio estuvo pensado para que los censos coincidieran con las cero horas del primer día de cualquier año, y los tantos brutos de mortalidad se obtenían por cociente para cada edad x , entre los fallecidos de cada generación, y los que superviven a cada edad. Esta situación es la que correspondió a los censos españoles fechados a 31 de diciembre (o bien 1 de enero) como ocurrió en España hasta el censo de 1970. Posteriormente hubo que introducir modificaciones ya que los censos dejaron de estar referidos a dichas fechas.

A) Cuando los censos corresponden a 31 de diciembre o 1 de enero (en España lo fueron los Censos de los años acabados en cero hasta 1970 inclusive), se puede obtener una estimación de los tantos brutos de mortalidad de una forma casi mecánica. En

concreto, denotemos por $P_x^{(31/12/t)}$, o, $P_x^{(1/1/t+1)}$ a los supervivientes a cada edad x en la fecha referenciada dada por el Censo; por $d_x^{(t)}$ y $d_x^{(t+1)}$, a los fallecidos a cada edad x , del año anterior y siguiente a la fecha del Censo, obtenidas de los correspondientes Movimientos Naturales de Población; y por $q_x^{(b)}$ a los correspondientes tantos brutos de mortalidad. Dado que para esta fecha censal los datos por edades coinciden con los datos por generaciones, podemos escribir que:

$$C_x^{31/12/t} \equiv C_x^{1/1/t+1} \equiv P_x^{01/01/t+1} \equiv P_x^{31/12/t} \equiv P_{g(t-x)}^{31/12/t} \equiv P_{g(t-x)}^{01/01/t+1}$$

Así bajo hipótesis de uniformidad de los fallecimientos a lo largo de cada año, los siguientes algoritmos proveerían las tasas brutas.

A.1) - Si los datos estén referidos a edades cumplidas, el algoritmo sería:

$$q_x^{(b)} = \frac{0,5 * [d_x^{(t)} + d_x^{(t+1)}]}{P_x^{(31/12/t)} + 0,5d_x^{(t)}} \quad \text{para } x = 1, 2, \dots, w; \text{ y}$$

$$q_x^{(b)} = \frac{0,7d_x^{(t)} + 0,3d_x^{(t+1)}}{P_x^{(31/12/t)} + 0,7d_x^{(t)}} \quad \text{para } x = 0,$$

La corrección para la edad cero es debido a que se considera que durante el primer año de vida hay mayor propensión al fallecimiento durante los primeros meses de vida. También en caso de disponer además de datos por generaciones, para las edades cero y uno, puede hacerse el correspondiente cálculo mediante generaciones (véase el apartado **A.2.**). No obstante, en algunos trabajos se ha utilizado las defunciones de dos años anteriores y dos siguientes, considerando como defunciones definitivas las obtenidas para cada edad por el promedio correspondiente.

A.2) Si los datos estuviesen dados por generaciones, teniendo en cuenta lo ya dicho en cuanto a la igualdad de los *stocks* de supervivientes por edades y por generaciones, por coincidir la fecha censal con el principio o final del año (esto es, $P_{g(t-x)}^{01/01/t+1} = P_x^{01/01/t+1}$) tendremos¹:

$$q_x^{(b)} = \frac{g(t-x)D_x}{1_{x,t}} = \frac{g(t-x)D_{x,t} + g(t+1-x)D_{x,t+1}}{1_{x,t}} = \frac{g(t-x)D_{x,t} + g(t-x)D_{x,t+1}}{P_{g(t-x)}^{01/01/t+1} + {}_{t-x}D_{x,t}}, \text{ para } x = 0, 1, 2, \dots, w.$$

B) Cuando los censos, como ocurrió en España a partir del de 1980, dejaron de coincidir con el 31 de diciembre de los años acabado en cero, la idea subyacente fue, mediante algún criterio, intentar desplazar los *stocks* censales al 1 de enero mas próximo o bien al 1 de enero del año en que se tengan Movimientos Naturales de Población más recientes. Por supuesto, la hipótesis de no existencia (o que se compensan) de movimientos migratorios continuó aceptándose, pues en otro caso necesitaríamos estadísticas diarias

¹ Nota.- La publicación del INE (Instituto Nacional de Estadística, Madrid 1999) de las Tablas de Mortalidad de la Población Española 1996-1997, sugiere para el caso **a.1.**-aplicar esta expresión por generaciones para las edades 0 y 1 (en sustitución de la expresión dada que consideraba coeficientes 0,7 y 0,3 para la edad cero.

o al menos mensuales de movimientos naturales de población para poder calcular los correspondientes desplazamientos. Asimismo, dado que no existen a nivel de población general, hay que recurrir de nuevo a más supuestos como el de uniformidad y proporcionalidad de los nacimientos y defunciones, ya que también es difícil conseguir estadísticas de los Movimientos Naturales de Población de los períodos un año anterior y posterior exacto respecto a la fecha del censo.

Actualmente los censos se refieren a las cero horas del primer día, de un mes concreto, que denotaremos por $(m+1)$, de un año concreto (en España los últimos tres censos se han realizado con referencia al 01/03/1981; 01/03/1991 y 01/11/2001). En estos casos se puede presuponer que en cada mes se producen las 12^a parte de las defunciones del año, y que los saldos migratorios más o menos se compensan. En ese caso, podemos proceder a estimar (mediante las correcciones o desplazamientos temporales necesarios) los valores censales al principio o al final del año censal, según los años disponibles de Estadísticas de Movimientos Naturales de Población disponibles. Por lo tanto, si la fecha censal fuese 01/ $m+1$ / $t+1$, y dispusiéramos de las estadísticas de los movimientos naturales de población de los años t y $(t+1)$ deberíamos calcular:

$P_x(31/12/t) \equiv P_x(1/1/t+1) = P_x(01/m+1/t+1) +$ Fallecidos durante los m primeros meses del año $(t+1)$ y podríamos proceder entre otros mediante el siguiente procedimiento geométrico-proporcional:

B.1) Si los datos están dados por edades:

$$P_x^{01/01/t+1} = \frac{12-m}{12} C_x^{01/m+1/t+1} + \frac{m}{12} C_{x+1}^{01/m+1/t+1} + \frac{[12+(12-m)]m}{144} D_{x,(t+1)} + \frac{m-m}{144} D_{x+1,(t+1)} =$$

$$\approx + \frac{12-m}{12} C_x^{01/m+1/t+1} + \frac{m}{12} C_{x+1}^{01/m+1/t+1} + \frac{(24-m)m}{288} D_{x,(t+1)} + \frac{m^2}{288} D_{x+1,(t+1)}$$

Estimados para todas las edades los valores de $P_x^{01/01/t+1}$ se podría aplicar el criterio clásico para obtener los tantos brutos de mortalidad:

$$q_x^{(b)} = \frac{0,5(d_x^{(t)} + d_x^{(t+1)})}{P_x^{01/01/t+1} + 0,5d_x^{(t)}} \text{ para } x = 1, 2, \dots, w;$$

Existe una salvedad en cuanto a las expresiones de los desplazamientos indicados para la edad $x = 0$, pues para estimar dichos *stocks* de población deberíamos proceder restando a los nacimientos del año inmediato anterior los fallecidos durante dicho año de dicha generación.

B.2) Si los datos estuvieran dados por generaciones, denotando por $D_{(t),g(t-x)}$ o $D_{(t)}^{g(t-x)}$, a los fallecidos durante el año t de la generación de $(t-x)$, evidentemente tendremos:

$$P_x^{01/01/t+1} = P_{g(t-x)}^{01/m+1/t+1} + \frac{12m}{144} D_{(t+1)}^{g(t-x)} = P_{g(t-x)}^{01/m+1/t+1} + \frac{12m}{144} ({}_{t-x}D_{x,t+1} + {}_{t-x}D_{x,t})$$

y por lo tanto:

$$q_x^{(b)} = \frac{{}_{t-x}D_x}{I_{x,g(t-x)}} = \frac{{}_{t-x}D_{x,t} + {}_{t-x}D_{x,t+1}}{P_{g(t-x)}^{01/01/t+1} + {}_{t-x}D_{x,t}} \approx \frac{{}_{t-x}D_{x,t} + {}_{t-x}D_{x,t+1}}{P_{g(t-x)}^{01/m+1/t+1} + \frac{12m}{144}({}_{t-x}D_{t+1} + {}_{t-x}D_{x,t})}$$

Como puede observarse el problema fundamental es el computo exacto del denominador de los tantos brutos de mortalidad, pues cuando obviamos el supuesto de estacionariedad, y más aun si introducimos movimientos migratorios, aparecen serios problemas de cómputo. Dichos denominadores son bastante más complejos que el mero cómputo de líneas de vida en el esquema de Lexis que verifiquen unas restricciones, y que a partir de ahora vamos a definirlos como *expuestos al riesgo de muerte, E_x , que generen a los correspondientes fallecidos o numeradores.*

3.2.- Método del censo (criterio II, o de cómputo exacto)

Este método hace especial hincapié en el concepto de exposición al riesgo. Lo denominaremos *criterio II o método de cómputo exacto*. Es el método que utilizaban y utilizan las compañías de seguros para elaborar sus tablas, pero que nosotros intentaremos aplicar a la construcción de tablas generales de mortalidad, considerando poblaciones abiertas, por lo tanto dando cabida a los movimientos migratorios (altas y bajas en general, y por lo tanto salidas por causas distintas de la muerte).

Comencemos prefijando el período de investigación (t_0, T) , que aunque no es necesario conviene comience un 1 de enero de un determinado año y concluya un 31 de diciembre, y que el período abarque al menos dos años (como es el período requerido por el método clásico) sin que sea superior a cuatro o cinco, dado el dinamismo y los cambios de pautas de la vida moderna. Además, si el período de observación es de dos años, podremos establecer comparaciones con el criterio anterior.

Recordemos que la idea subyacente es relacionar por cociente: a los fallecidos del período, que denotaremos por θ_x , y a los supervivientes, que estando expuestos al riesgo de muerte, hayan generado dichos fallecimientos, que denotaremos por E_x . Como seguidamente pasamos a dilucidar, los fallecidos del período θ_x , o numeradores de los tantos brutos de mortalidad, son fáciles de obtener; sin embargo la determinación de los denominadores E_x es bastante compleja.

Para abordar su cálculo, en primer lugar diferenciemos los diferentes subcolectivos con diferente contribución al riesgo. Evidentemente para cualquier edad x , podemos considerar:

- (i) El colectivo de personas que cumplen la edad x durante el período de investigación, que denotaremos por I_x^0 .
- (ii) Los existentes en el momento del comienzo del período de investigación, con (cada) edad x ya cumplida, a los que llamaremos iniciales (corresponden a los titulares de pólizas en vigor en caso de entidades aseguradoras) o también **beginners** b_x según la terminología anglosajona.
- (iii) Los que se incorporan durante el período de observación con edad x cumplida, o nuevas entradas de edad x durante el período de observación, n_x (inmigrantes o nuevas pólizas en el caso de entidades de seguros) o **news** en la terminología anglosajona.
- (iv) d.- Los que abandonan el colectivo por fallecimiento a la edad x , durante el período de observación,

o fallecidos θ_x (**deaths** en la terminología anglosajona). (v) Los que abandonan el colectivo a la edad x durante el período de observación por causas diferentes del fallecimientos, w_x (**withdrawers** en la terminología anglosajona, esto es emigrantes en el caso de población general; y, por rescates y/o finalizaciones de contratos en el caso de población de asegurados). Y, (vi) Los que en el momento de finalizar la investigación tienen la edad x cumplida, e_x (finalistas, o con pólizas en vigor en el momento final de la investigación; **enders**, en la terminología anglosajona).

Obsérvese como a partir de las Figuras 1 y 2 se puede visualizar la formación de dichos dichos subcolectivos.

Figura 1. Representación en el esquema de Lexis de los distintos tipos de individuos (edad 0)

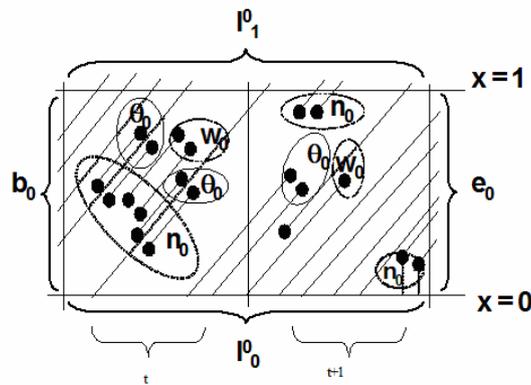
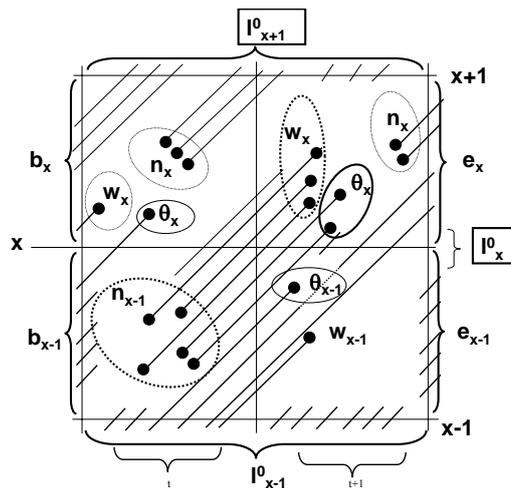


Figura 2. Representación en el esquema de Lexis de los distintos tipos de individuos (edad $x \neq 0$)



Evidentemente los b_x y e_x para todas las edades $x = 0, 1, \dots, w$, vienen proporcionados por un censo en el momento inicial y el momento final del período de investigación respectivamente. Luego este criterio requerirá entre otras cosas **dos censos u observaciones completas transversales**: uno en el **momento inicial** del período de

investigación u observación , y otro en el **final**. Por otra parte los datos correspondientes a los subcolectivos \mathbf{n}_x , θ_x , y \mathbf{w}_x vienen dados por las estadísticas diarias de las compañías de seguros, o bien deberían venir dados por las estadísticas de Movimientos Naturales de Población. Finalmente \mathbf{l}_x^0 , se pueden obtener por las siguientes ecuaciones de recurrencia:

$$\mathbf{l}_1^0 = \mathbf{l}_0^0 + \mathbf{b}_0 + \mathbf{n}_0 - \theta_0 - \mathbf{w}_0 - \mathbf{e}_0$$

$$\mathbf{l}_2^0 = \mathbf{l}_1^0 + \mathbf{b}_1 + \mathbf{n}_1 - \theta_1 - \mathbf{w}_1 - \mathbf{e}_1 = \mathbf{l}_0^0 + \sum_1^1 \mathbf{b}_x + \sum_1^1 \mathbf{n}_x - \sum_1^1 \mathbf{w}_x - \sum_1^1 \theta_x - \sum_1^1 \mathbf{e}_x$$

$$\mathbf{l}_3^0 = \mathbf{l}_2^0 + \mathbf{b}_2 + \mathbf{n}_2 - \theta_2 - \mathbf{w}_2 - \mathbf{e}_{21} = \mathbf{l}_0^0 + \sum_2^2 \mathbf{b}_x + \sum_2^2 \mathbf{n}_x - \sum_2^2 \mathbf{w}_x - \sum_2^2 \theta_x - \sum_2^2 \mathbf{e}_x$$

, y

$$\mathbf{l}_x^0 = \mathbf{l}_{x-1}^0 + \mathbf{b}_{x-1} + \mathbf{n}_{x-1} - \theta_{x-1} - \mathbf{w}_{x-1} - \mathbf{e}_{x-1} = \mathbf{l}_0^0 + \sum_0^{x-1} \mathbf{b}_t + \sum_0^{x-1} \mathbf{n}_t - \sum_0^{x-1} \mathbf{w}_t - \sum_0^{x-1} \theta_t - \sum_0^{x-1} \mathbf{e}_t$$

y así sucesivamente hasta \mathbf{l}_w^0 .

Obsérvese que para cada edad \mathbf{x} , \mathbf{l}_x^0 puede recomponerse o construirse a partir de \mathbf{l}_0^0 y de los subcolectivos \mathbf{b}_x , \mathbf{n}_x , θ_x , \mathbf{w}_x , y \mathbf{e}_x de las edades anteriores pero sin incluir la propia edad \mathbf{x} . Además, si se tratara de datos de una compañía de seguros en general no comenzaríamos en la edad cero, sino en otra posterior como matizaremos más adelante.

Una vez identificados los anteriores colectivos o subcolectivos, hay que proceder al cálculo de su contribución exacta a la exposición al riesgo, lo que nos proporcionará \mathbf{E}_x , o denominador de los tantos brutos de mortalidad.

En cuanto a la contribución al riesgo de los \mathbf{l}_x^0 , no es correcto que cada uno de ellos se compute como una unidad, es decir, considerar como si estuviera contribuyendo a la exposición al riesgo de muerte durante (toda) la edad \mathbf{x} , pues antes de alcanzar la edad $(\mathbf{x}+1)$ algunos de ellos fallecerán, otros abandonarán el colectivo por otras causas, y otros podrán llegar al momento final de la investigación sin haber alcanzado la edad $\mathbf{x}+1$. Por lo tanto, para incorporar correctamente \mathbf{l}_x^0 en el computo de \mathbf{E}_x , al valor \mathbf{l}_x^0 (que sería el resultante de computar cada individuo como una unidad) habrá que restar y también sumar ciertas cantidades como dilucidamos a continuación:

A) Los \mathbf{b}_x y \mathbf{n}_x (de edad \mathbf{x}) no estarán incluidos en los \mathbf{l}_x^0 (constate el lector mas que en la ecuación de \mathbf{l}_x^0 , sólo intervienen dichos subcolectivos pero con edades anteriores a \mathbf{x}), luego habrá que añadirlos por su contribución exacta a la exposición al riesgo, que tampoco será de una unidad por cabeza sino que dependerá de la edad exacta de entrada y/o de salida de cada cabeza en el correspondiente colectivo. Esto es que para cada \mathbf{r} , \mathbf{s} , y, tales que $0 < \mathbf{r}, \mathbf{s} < 1$, deberíamos tener en cuenta que:

A.1) Cada uno de los principiantes (o beginners), que tuvieran la edad exacta $(\mathbf{x}+\mathbf{r})$ en el momento del comienzo de la investigación o del período de investigación, que

denotaremos por b_{x+r} contribuirá a la exposición al riesgo a dicha edad x , en la cuantía $(1-r)$, en el caso que se mantuviera en dicho subcolectivo durante toda la edad x , de lo contrario habría que deducir además alguna cierta parte alícuota.

Por lo tanto $\sum_r (1-r)b_{x+r}$ será, la expresión del tiempo de exposición real al riesgo

por dichos **beginners** con x años cumplidos siempre que estuvieran o si se mantuvieran en el colectivo al menos durante toda la edad x , esto es mientras no salgan del colectivo por fallecimiento o por otras causas, o lleguen al final de la investigación con x años y sin cumplir $(x+1)$, y

A.2) Análogamente la contribución de los **news**, o nuevos entrantes, también dependerá de su edad exacta de incorporación. Si fuese $(x+r)$ su tiempo de exposición real a cada edad x vendrá dado por : $\sum_r (1-r)n_{x+r}$, siempre que siendo

n_{x+r} el numero de nuevos incorporados durante el período de observación con edad exacta $(x+r)$ no salgan del colectivo por fallecimiento o por otras causas o lleguen al final de la investigación con x años y sin cumplir $(x+1)$.

Por lo tanto las expresiones $\sum_r (1-r)b_{x+r}$ y $\sum_r (1-r)n_{x+r}$ se tendrán que añadir a

l_x^0 para ir aproximándonos al cómputo exacto de la exposición al riesgo a la edad x .

B) Por otra parte, hemos de tener en cuenta a los que abandonan el colectivo a la edad x (y sin cumplir la edad $x+1$) por fallecimiento θ_x , por causas diferentes al fallecimiento (esto es los w_x , emigrantes, por rescates etc.) y también a los que llegan al final del período de investigación con la edad x cumplida pero sin alcanzar la $(x+1)$, o e_x . Obsérvese que pueden proceder tanto de los que alcanzan la edad x en el período de investigación como de los beginners y nuevos entrantes de edad x . Por lo que en el calculo de E_x , al cómputo acumulado hasta este momento habrá que restar: $\sum_s (1-s)w_{x+s}$, por los retiros o abandonos a la edad $x+s$ (esto es antes de cumplir la edad $x+1$) por causas diferentes del fallecimiento, y $\sum_s (1-s)e_{x+s}$ por los finalistas o

personas que en el último instante de la investigación tienen la edad $x+s$.

Consecuentemente la expresión:

$$l_z^0 + \sum_r (1-r)b_{x+r} + \sum_r (1-r)n_{x+r} - \sum_r (1-s)w_{x+s} - \sum_r (1-s)e_{x+s}$$

representa al numero de cabezas o personas del colectivo de referencia que al inicio de la edad x generarían los θ_x fallecimientos a dicha edad. Dicha magnitud se denomina **exposición inicial al riesgo**, y denotaremos por E_x , esto es:

$$E_x = l_z^0 + \sum_r (1-r)b_{x+r} + \sum_r (1-r)n_{x+r} - \sum_r (1-s)w_{x+s} - \sum_r (1-s)e_{x+s}$$

Por lo tanto los cocientes:

$$q_x^{(b)} = \frac{\theta_x}{E_x} = \frac{\theta_x}{l_z^0 + \sum_r (1-r)b_{x+r} + \sum_r (1-r)n_{x+r} - \sum_r (1-s)w_{x+s} - \sum_r (1-s)e_{x+s}}$$

se consideran como tantos brutos de mortalidad o estimadores de los tantos anuales de mortalidad² q_x

No debemos olvidar que a partir de los archivos de compañías de seguros podremos obtener directamente E_x , pero de los Censos generales y Movimientos Naturales de Población, seguramente no, aunque si podríamos obtener E_x^c , por lo tanto como:

$$E_x^c = E_x - \sum_r (1-r)\theta_{x+r}$$

para los casos de tablas generales calcularíamos E_x , a partir de E_x^c , y $\sum_r (1-r)\theta_{x+r}$.

Además en el caso de uniformidad de los fallecimientos la expresión $\sum_r (1-r)\theta_{x+r}$, viene a ser equivalente la mitad de los fallecidos del período, esto es que se puede aproximar mediante $\frac{1}{2}\theta_x$, y por ello para los casos de población general podríamos utilizar la siguiente aproximación:

$$E_x = E_x^c + \sum_r (1-r)\theta_{x+r} \cong E_x^c + \frac{1}{2}\theta_x$$

Evidentemente este criterio, que hemos denominado método del censo (criterio II), es el más exacto siempre que se pueda aplicar. Hemos dicho que el período de investigación puede ser cualquiera y es recomendable que sea superior a dos años, pero dada la dinamicidad de la vida real deberíamos evitar que fuese superior a cuatro o cinco años. Puede constatarse que en el caso particular de un período de dos años (en general, el anterior y posterior al del censo) es muy parecido al criterio I, aunque no son exactamente iguales. Además como también hemos indicado, puede utilizarse tanto para la construcción de **tablas generales de mortalidad** como para específicas de **colectivos de asegurados**, pero sin lugar a dudas donde mejor se puede utilizar es para este último caso, ya que de los ficheros de las entidades aseguradoras pueden extraerse los datos requeridos.

En el caso de colectivos de asegurados:

- b_x son asegurados existentes con edad x , al principio de la investigación;
- n_x , son los nuevos contratantes de seguros durante la investigación;

² El concepto anterior de exposición inicial al riesgo no debe confundirse con el de exposición central al riesgo E_x^c , definido mediante:

$$E_x^c = l_z^0 + \sum_r (1-r)b_{x+r} + \sum_r (1-r)n_{x+r} - \sum_s (1-s)\theta_{x+s} - \sum_s (1-s)w_{x+s} - \sum_r (1-s)e_{x+s}$$

correspondería al tiempo de exposición real al riesgo de muerte por las cabezas o personas a la edad x del colectivo en cuestión incluidos también los fallecimientos a la edad x , en cuyo caso, los cocientes:

$$m_x^{(b)} = \frac{\theta_x}{E_x^c} = \frac{\theta_x}{l_z^0 + \sum_r (1-r)b_{x+r} + \sum_r (1-r)n_{x+r} - \sum_s (1-s)\theta_{x+s} - \sum_r (1-s)w_{x+s} - \sum_r (1-s)e_{x+s}}$$

son los estimadores de los correspondientes **tantos centrales de mortalidad**, m_x

- w_x , son los que abandonan, rescatan o rescinden el contrato por causas distintas del fallecimiento; y
- θ_x , son los que fallecen a la edad x

y se pueden obtener los cálculos de sus contribuciones exactas (pues en general en la ficha de cada asegurado, aparece la fecha exacta de su nacimiento y de su alta o ingreso, así como su rescisión o salida y su causa, con lo que se puede comprobar exactamente sus contribuciones al riesgo de muerte). Sólo hay que introducir una salvedad respecto de todo lo indicado y es que en general no hay datos de edades iguales o inferiores a una dada sobretodo en algunos tipos de seguros (por ejemplo 18, para los seguros de automóviles), por ello hay que considerar la cota inferior de edad "a", por debajo de la cual no hay póliza alguna contratada, por lo tanto la primera edad a considerar en la cadena de expresiones para el cálculo 1_x^0 en los casos de colectivos de asegurados, será la edad $(a+1)$. Por lo tanto 1_a^0 para las tablas de asegurados, es el equivalente a 1_0^0 en las tablas generales.

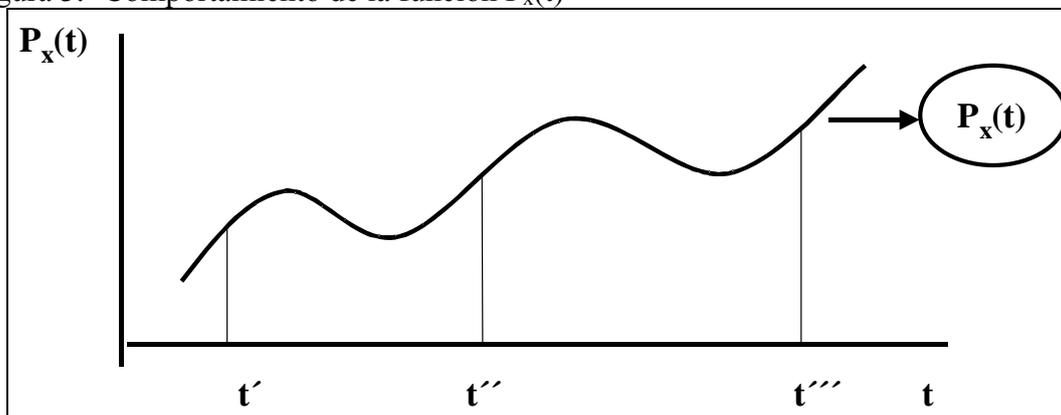
Aunque este método teóricamente también **puede ser aplicado a los casos de construcción de de tablas generales**, y sería deseable, es prácticamente imposible de llevarlo a la práctica, puesto que en los Censos y Movimientos Naturales de Población es bastante difícil por no decir imposible obtener directamente E_x , sino a través de E_x° lo que de hecho precisa como hemos visto supuestos simplificadores y/o aproximados.

3.3.- Metodo del censo (criterio III: Fundamentación del calculo de E_x)

Finalmente con este tercer y ultimo criterio abordamos el problema de la fundamentación estadístico-matemática del cálculo de la población o expuestos al riesgo E_x . Para ello partimos del supuesto que $P_x(t)$ constituye un proceso estocástico y que podría ser conocido para todo valor de x mediante observaciones muestrales en cualquier momento t .

Sea $P_x(t)$ a la población con edad x , de un colectivo (nación, comunidad, o región etc.) en el momento t . Si dicho momento o fecha t realizáramos un censo, la función $P_x(t)$ coincidiría con los *stocks* de población para cada edad x , proporcionada por el propio censo. Esto implicaría que al menos teóricamente que los censos se puedan elaborar en cualquier momento t del tiempo físico o de calendario, proporcionando información para toda edad x .

Figura 3.- Comportamiento de la función $P_x(t)$



Evidentemente al variar t , $P_x(t)$ (para cada edad x), será difícilmente constante, ya que continuamente recibirá nuevas incorporaciones (por nacimientos como por nuevas incorporaciones) y bajas (por defunciones u otras causas), esto es, que podíamos pensar que para cada edad x , la gráfica de $P_x(t)$ presentará un movimiento oscilante con una tendencia ascendente o descendente según las pautas de la población, como puede observarse en la figura 3.-

Si consideramos períodos arbitrariamente pequeños o infinitesimales, por ejemplo el período $[t, t+dt]$, cada uno de los individuos del colectivo $P_x(t)$ contribuirá al riesgo en la cantidad infinitesimal dt , y por lo tanto todos ellos en conjunto salvo infinitésimos de orden superior estarán sometidos al riesgo de muerte (o contribuyendo) desde el instante t al $t+dt$, esto es contribuyendo en la cantidad:

$$P_x(t) dt.$$

Durante un periodo cualquiera, pongamos de un año, (por ejemplo entre t' y $t'+1$) esto es entre dos observaciones (censales) anuales, una hecha a principio del año y otra hecha al final o principio del siguiente) su contribución será :

$$E_x^{t',t'+1} = \int_{t'}^{t'+1} P_x(t) dt$$

dicha expresión, si no se dispone de la forma analítica de $P_x(t)$ para todo t , como ocurre en la realidad, sólo se puede calcular numéricamente o con fórmulas aproximadas como por ejemplo la regla de Simpson o de los trapecios, que para dos informaciones censales de fechas t' y $t'+1$, sería:

$$\int_{t'}^{t'+1} P_x(t) dt \approx \frac{P_x(t') + P_x(t'+1)}{2} \cdot 1 = \frac{P_x(t') + P_x(t'+1)}{2}$$

Aplicando dicha regla reiterativamente, tendríamos:

A) En caso en caso de tres observaciones anuales:

$$\begin{aligned} \int_{t'}^{t'+2} P_x(t) dt &= \int_{t'}^{t'+1} P_x(t) dt + \int_{t'+1}^{t'+2} P_x(t) dt \approx \\ &\approx \frac{P_x(t') + P_x(t'+1)}{2} \cdot 1 + \frac{P_x(t'+1) + P_x(t'+2)}{2} \cdot 1 = \\ &= \frac{P_x(t') + 2P_x(t'+1) + P_x(t'+2)}{2} = \end{aligned}$$

$$= \frac{1}{2} P_x(t') + P_x(t'+1) + \frac{1}{2} P_x(t'+2)$$

y

B) Para el caso general de n observaciones anuales disponibles (por ejemplo desde t' la primera y $t'+n$, la última) análogamente resultaría la siguiente expresión aproximada:

$$\begin{aligned} \int_{t'}^{t'+n} P_x(t) dt &= \int_{t'}^{t'+1} P_x(t) dt + \int_{t'+1}^{t'+2} P_x(t) dt + \dots + \int_{t'+n-1}^{t'+n} P_x(t) dt \approx \\ &\approx \frac{P_x(t') + P_x(t'+1)}{2} \cdot 1 + \frac{P_x(t'+1) + P_x(t'+2)}{2} \cdot 1 + \dots + \frac{P_x(t'+n-1) + P_x(t'+n)}{2} = \\ &= \frac{1}{2} P_x(t') + [P_x(t'+1) + P_x(t'+2) + \dots + P_x(t'+n-1)] + \frac{1}{2} P_x(t'+n) \end{aligned}$$

Los mismos razonamientos se pueden extender a datos referidos a otros períodos de tiempo. Como lo más común son observaciones anuales, o posibilidad de al menos obtenerlas aproximadamente, podemos decir que según este criterio, la expresión general que utilizaremos para el cálculo aproximado de $E_x^{(c)}(t', t'+n)$ será:

$$E_x^c(t', t'+n) = \int_{t'}^{t'+n} P_x(t) dt = \frac{1}{2} P_x(t') + P_x(t'+1) + \dots + P_x(t'+n-1) + \frac{1}{2} P_x(t'+n)$$

debemos hacer constar que la expresión anterior $E_x^c(t', t'+n)$ no tiene por qué ser siempre así. Por ejemplo si los datos censales no estuvieran referidos al final o al principio de cada año sino a la mitad de cada año (estos es al 30 de junio, o al 1 de julio), tendríamos que la formula aproximada anterior utilizada para el cálculo de la integral podría ser:

$$\begin{aligned} E_x^c(t', t'+n) &= \int_{t'}^{t'+n} P_x(t) dt = P_x(t'+\frac{1}{2}) \cdot 1 + P_x(t'+1+\frac{1}{2}) \cdot 1 + \dots + P_x(t'+n-\frac{1}{2}) \cdot 1 = \\ &= P_x(t'+\frac{1}{2}) + P_x(t'+1+\frac{1}{2}) + \dots + P_x(t'+n-\frac{1}{2}) \end{aligned}$$

ya que el criterio de aproximación podría justificarse por las sumas de las áreas de los cuadrados de base la unidad y de altura la ordenada en el punto medio, en que puede descomponerse el recinto de integración.

Finalmente puesto que la diferencia entre $E_x^{(c)}$ y E_x , es $\sum_r (1-r)\theta_{x+r}$ y que en caso de distribución uniforme de los fallecimientos, si el período de investigación es (t_0, T) , podríamos aproximar la expresión anterior mediante: $0,5\theta_x(t', t'+n)$.

A partir de esta fundamentación, podríamos deducir otra expresión aproximada para el cálculo de los tantos brutos de mortalidad (o criterio III) como:

$$q_x^{(b)} = \frac{\theta_x}{E_x} = \frac{\theta_x}{E_x(t', t'+n)} = \frac{\theta_x}{E_x^c(t', t'+n) + \sum_r (1-r)\theta_{x+r}} = \frac{\theta_x}{E_x^c(t', t'+n) + 0,5\theta_x(t', t'+n)}$$

Obsérvese que en el caso particular de un período de observación o de investigación de dos años (que en este caso requeriría tres observaciones censales) necesitaremos los *fallecidos a cada edad durante dichos dos años, y tres experiencias censales anuales*, esto es que este nuevo criterio nos conduciría al siguiente criterio de cálculo:

$$q_x^b = \hat{q}_x = \frac{\theta_x}{E_x} = \frac{\theta_x}{\frac{1}{2}P_x(t) + P_x(t+1) + \frac{1}{2}P_x(t+2) + 0,5\theta_x} =$$

$$= \frac{\theta_{x,t} + \theta_{x,t+1}}{\frac{1}{2}P_x(t) + P_x(t+1) + \frac{1}{2}P_x(t+2) + 0,5(\theta_{x,t} + \theta_{x,t+1})}$$

siendo el numerador θ_x , el correspondiente a los fallecidos durante los dos años, t y $t+1$ de la investigación, y el denominador la expresión aproximada de los expuestos potenciales al riesgo según este criterio.

De nuevo aparece el problema real apuntado repetidamente: **que ocurrirá si solo se dispone de una única experiencia censal**. Evidentemente en el caso de disponer de un único censo, sería necesario realizar los “desplazamientos” de los datos censales necesarios, para calcular los *stocks* de población correspondientes a las fechas requeridas que son el 1 de enero de los años t , $t+1$, y $t+2$, a partir del Censo disponible y de los Movimientos Naturales de Población necesarios. lo cual introducirá errores al menos por las aproximaciones realizadas.

4.- Propiedades inferenciales

Analicemos por último cómo podemos obtener propiedades estadísticas inferenciales de los estimadores anteriores.. El hilo conductor de esta idea está basado en dos trabajos metodológicos en este sentido a nivel internacional que son: “Experiments in graduating the data for the English Life Tables N° 14)” del profesor J.J. McCutcheon³ y “On graduation by mathematical formula” de los profesores D.O. Forfar J.J. McCutcheon y D. Wilkie, publicado en el Journal of the Institute of Actuaries; así como en la obra clásica de Benjamin and Pollard (1989).

Para ello, fijado el período de observación, vamos a introducir la variable aleatoria número de fallecimientos de una cierta población o colectivo correspondientes al mismo período. Evidentemente el evento mortalidad-supervivencia se puede modelizar como un proceso aleatorio dicotómico. Denotando por E_x , al número de individuos expuestos al riesgo de muerte para cada edad x y por q_x a la probabilidad de que un individuo con edad exacta x años fallezca antes de alcanzar la edad $x+1$, el número de fallecimientos de individuos para cada edad x (esto es con edad: entre x –incluída- y $x+1$ –excluída-) durante el período de observación, puede considerarse como una variable aleatoria

³ McCutcheon, J.J. (1987). Experiments in graduating the data for the English Life Tables (N° 14). *Transactions of the Faculty of Actuaries. Vol. 40 (1)*

$\tilde{\theta}_x$ con una distribución de probabilidad asociada. La hipótesis más aceptable es presuponer que dicha distribución es binomial de parámetros \mathbf{E}_x y \mathbf{q}_x , esto es: $\tilde{\theta}_x \sim B(\mathbf{E}_x, \mathbf{q}_x)$, cuyos parámetros hay que estimar.

Obtengamos mediante el proceso binomial mortalidad-supervivencia indicado, la expresión de la probabilidad conjunta que de \mathbf{E}_x cabezas o personas realmente observadas fallezcan θ_x y sobrevivan $(\mathbf{E}_x - \theta_x)$. Entonces la función de verosimilitud $L(\mathbf{q}_x)$ es:

$$L(\mathbf{q}_x) = \binom{\mathbf{E}_x}{\theta_x} \mathbf{q}_x^{\theta_x} (1 - \mathbf{q}_x)^{\mathbf{E}_x - \theta_x}$$

y salvo el factor de proporcionalidad (que es el número combinatorio), podríamos formular así:

$$L(\mathbf{q}_x) = \mathbf{q}_x^{\theta_x} (1 - \mathbf{q}_x)^{\mathbf{E}_x - \theta_x}$$

Tomando logaritmos neperianos y derivando se sigue inmediatamente que:

$$\ln L(\mathbf{q}_x) = \theta_x \ln \mathbf{q}_x + (\mathbf{E}_x - \theta_x) \ln(1 - \mathbf{q}_x)$$

$$\frac{\partial L}{\partial \mathbf{q}_x} = \theta_x \frac{1}{\mathbf{q}_x} - (\mathbf{E}_x - \theta_x) \frac{1}{1 - \mathbf{q}_x}$$

igualando a cero y designando a la solución como $\mathbf{q}_x^{(b)}$, tendremos:

$$\theta_x \frac{1}{\mathbf{q}_x^{(b)}} = (\mathbf{E}_x - \theta_x) \frac{1}{1 - \mathbf{q}_x^{(b)}} \Rightarrow \frac{\mathbf{q}_x^{(b)}}{1 - \mathbf{q}_x^{(b)}} = \frac{\theta_x}{(\mathbf{E}_x - \theta_x)} \Rightarrow \mathbf{E}_x \mathbf{q}_x^{(b)} - \theta_x \mathbf{q}_x^{(b)} = \theta_x - \theta_x \mathbf{q}_x^{(b)} \Rightarrow$$

$$\Rightarrow \mathbf{E}_x \mathbf{q}_x^{(b)} = \theta_x \Rightarrow \mathbf{q}_x^{(b)} = \frac{\theta_x}{\mathbf{E}_x}$$

y por lo tanto:

$$\mathbf{q}_x^{(b)} = \frac{\theta_x}{\mathbf{E}_x}$$

serán estimadores máximo-verosimiles de los tantos anuales de mortalidad que suelen denotarse por:

$$q_x = \frac{d_x}{l_x}$$

Además dichos estimadores son insesgados, puesto que su esperanza matemática coincide con el parámetro poblacional a estimar. En efecto:

$$\mathbf{E}(\mathbf{q}_x^{(b)}) = \mathbf{E}\left(\frac{\theta_x}{\mathbf{E}_x}\right) = \frac{1}{\mathbf{E}_x} \mathbf{E}(\theta_x) = \frac{\mathbf{E}_x \mathbf{q}_x}{\mathbf{E}_x} = \mathbf{q}_x$$

y además su varianza es:

$$\mathbf{Var}(q_x^{(b)}) = \frac{q_x(1-q_x)}{E_x}$$

y que a su vez, en caso de que el número de fallecimientos sea lo suficientemente grande, puede estimarse mediante:

$$\mathbf{V\hat{a}r}(q_x^{(b)}) = \frac{q_x^{(b)}(1-q_x^{(b)})}{E_x}.$$

5.- Conclusiones.

En definitiva, después de analizar los diferentes criterios y sus propiedades consideramos que lo óptimo consistiría en aplicar el criterio II o de cómputo exacto, tanto para poblaciones generales como específicas.

Sin embargo, dados los requisitos exigidos por el criterio II, su aplicabilidad en poblaciones generales es prácticamente imposible, pues requiere de una contabilización diaria de los movimientos naturales y migratorios. Lo cual obviamente es imposible para un estado del tamaño español.

Como consecuencia, para el caso de poblaciones generales, lo más razonable sería aplicar el criterio III con todas las aproximaciones allí indicadas.

Bibliografía

Artis, M., Felipe, A. y Guillen, M. (1997) "Time-dependent mortality patterns in the Spanish population: 1975-1993", Seminar of Mortality. Edimburgh, Reino Unido, 4 de julio de 1997

Ayuso, M., Corrales, H., Guillen, M., Perez-Marin, A.M. y Rojo, J.L. (2001): *Estadística Actuarial Vida*. Edicions Universitat de Barcelona. Barcelona.

Benjamin, P. y Pollard, J. (1986): *The Analysis of Mortality and Other Actuarial Statistics*. Ed. Heinemann. Londres.

Bowers, J.R., Gerber, H., Hickman, J.C., Jones, D.A. y Nesbitt, C.J. (1986) : *Actuarial Mathematics*. The Society of Actuaries. Illinois.

Escuder, R. y Navarro, E. (2002): *Construcción de las tablas de mortalidad de colectivos de asegurados de la población española de 1997-98*. Proyecto de investigación financiado por Unión Española de Entidades Aseguradoras y Reaseguradoras (UNESPA) e Instituto de Actuarios Españoles (IAE).

Escuder, R., Mendez, S., Navarro, E. y otros (2003): *Tablas de mortalidad de la Comunidad Valenciana años 2000-2001. Evolución reciente de las tendencias de mortalidad y proyecciones*. Proyecto de investigación financiado por Generalitat valenciana. CITIB/2002/286.

Felipe, A. y Guillen, M. (1998) Evolucion y prediccion de las tablas de mortalidad dinamicas para la poblacion española. Madrid: Fundacionn MAPFRE Estudios. ISBN: 84-8942-9294. D. L.: M-6525-1999.

Felipe, A., Guillen, M. y Artis, M. (1998) "Recent mortality trends in the Spanish population", Transactions of the 26th International Congress of Actuaries, 9, 250-269. Brimingham: Reino Unido.

Forfar, D.O, McCutcheon, J.J. y Wilkie, D. (1988): On Graduation by Mathematical Formula. *Journal of the Institute of Actuaries*, 115, 1-149.

Heligman, L. y Pollard, J.H. (1980): The Age Pattern of Mortality. *Journal of the Institute of Actuaries*, 107, 49-80.

Hossack, I.B., Pollard, J.H. y Zehnwirt, B. (1999): *Introductory Statistics with Application in General Insurance*. Ed. Cambridge University Press. Cambridge.

I.N.E: Censos y Movimientos Naturales de Población.
[.http://www.ine.es/inebase/index.html](http://www.ine.es/inebase/index.html)

Insolera, F.(1950): *Curso de Matemática financiera y Actuarial*.
Ed. Aguilar. Madrid (traducción de la segunda edición italiana. Torino 1937)

London, D. (1985): *Graduation: The revision of Estimates*. Actex. Abington.

McCuttcheon, J.J. (1987): Experiments in graduating the data for the English Life Tables (N^o 14). *Transactions of the Faculty of Actuaries*. 40 (1), 135-147.

Navarro, E. (1991): *Tablas de mortalidad de la población española 1982. Metodología y Fuentes*. Editorial MAPFRE. Madrid.

Pavia J.M y Escuder, R. (2003): El Proceso Estocástico de Muerte. Diferentes estrategias para la elaboración de tablas recargadas. Análisis de sensibilidad. *Estadística Española*, 45; n^o 153, 253-274 .

Prieto, E. y Fernandez, M.J. (1994): *Tablas de mortalidad de la población española de 1950 a 1990. Tabla proyectada del año 2000*. UNESPA. Madrid

ESTUDIO DE LA SINIESTRALIDAD EN EL SEGURO DE AUTOMÓVILES: IDONEIDAD DE LOS MODELOS INFLADOS DE CEROS Y OTRAS EXTENSIONES*

MARÍA DEL CARMEN MELGAR HIRALDO
e-mail: mcmelhir@upo.es

JOSÉ ANTONIO ORDAZ SANZ
e-mail: jaordsan@upo.es

Departamento de Economía, Métodos Cuantitativos e Historia Económica
UNIVERSIDAD PABLO DE OLAVIDE

M. KAZIM KHAN
e-mail: kazim@math.kent.edu

Department of Mathematical Sciences
KENT STATE UNIVERSITY

Área temática: Métodos Cuantitativos.

Resumen

El conocimiento de los factores que intervienen en la siniestralidad del seguro de automóviles es un tema de gran interés para las compañías aseguradoras. La consideración de dichos factores podría eventualmente permitir a las entidades fijar con mayor precisión la prima a pagar por parte de sus clientes y alcanzar así un mayor grado de eficiencia.

Tradicionalmente, los estudios orientados al análisis del número de siniestros que los asegurados declaran a su compañía se han basado en el uso de modelos de recuento (*count data models*), fundamentalmente el modelo de Poisson y el modelo binomial negativo. Sin embargo, la posibilidad que el asegurado tiene de decidir si declara o no los siniestros que sufre, convierte la utilización de los modelos inflados de ceros (*zero-inflated models*) en una opción muy adecuada desde el punto de vista teórico.

El trabajo que presentamos se centra en la aplicación de dichos modelos en la determinación de las principales características relativas a los vehículos, a los clientes asegurados y a las pólizas contratadas que resultan relevantes en la siniestralidad del seguro de automóviles, usando una base de datos procedente de una entidad aseguradora privada española.

El estudio se completa con la estimación del número de clientes que, pese a sufrir siniestros, deciden no declararlos, así como de la cifra que éstos suponen; y con una discusión final, de carácter teórico, sobre las posibles ventajas que podrían derivarse de la consideración de otros modelos alternativos, como pueden ser los denominados modelos afinados (*thinned models*).

Palabras clave: Seguro de automóviles, siniestralidad, modelos inflados de ceros, modelos afinados.

Abstract

Knowing the factors that have influence on the accidents of the automobile insurance industry is an issue of a great relevance for insurance companies. The study of these factors could eventually allow insurers to fix more precisely the premiums their clients should pay for the policies and to reach higher levels of efficiency.

Traditionally, research oriented to analyze the number of declared accidents by the policy holders has been based on the use of count data models, basically Poisson and negative binomial models. However, policy holders have the option of not declaring the accidents they suffer which leads to consider the zero-inflated models as a very appropriate option, from a theoretical point of view.

The present work is focused on the application of this type of models to determine the main features of vehicles, insured clients and contracted policies that are relevant to explain why accidents happen. For this purpose we use a database provided by a Spanish private insurance company.

The study is completed with the estimations of the number of clients that have accidents but decide not to declare them, and the number of these accidents. We conclude with a final theoretical discussion on the possible advantages given by other alternative models, as the so-called thinned models.

Key words: Automobile insurance industry, accidents, zero-inflated models, thinned models.

* Este trabajo ha recibido ayudas del Ministerio de Educación y Ciencia y FEDER (Proyecto SEJ2005-00741/ECON), así como del Ohio Board of Regents (OBR-2006), USA.

1. Introducción

Los siniestros del seguro de automóviles que los asegurados sufren y declaran a sus compañías inciden en éstas de forma negativa. En consecuencia, resulta de gran importancia que las entidades aseguradoras realicen un ajuste preciso de las primas, de forma que permitan hacer frente a las coberturas elegidas por sus clientes. Para llevar a cabo esta tarea de forma eficiente, es fundamental estudiar aquellos factores que pueden influir en una mayor o menor siniestralidad.

Los modelos econométricos que se han venido utilizando tradicionalmente en los estudios orientados a analizar el número de siniestros son los de tipo recuento (*count data models*), que se aplican cuando la variable endógena toma valores enteros y no negativos. En particular, el modelo de regresión de Poisson y el modelo binomial negativo son los más empleados. Entre los trabajos que los usan, podemos citar, entre otros, los de Richaudeau (1999) y Cohen (2005), que tratan de detectar empíricamente la existencia de información asimétrica estudiando la posible correlación entre la siniestralidad observada de los asegurados y la cobertura de la que disfrutan¹.

A pesar del uso extendido de los dos modelos señalados, otros modelos más complejos, como son los llamados inflados de ceros (*zero-inflated models*), pueden convertirse en una opción adecuada desde el punto de vista teórico, ya que permiten tener en cuenta ciertos aspectos que los modelos tradicionales no contemplan. La gran cantidad de observaciones nulas presentes en la variable endógena, así como las distintas interpretaciones que pueden otorgarse a dichas observaciones, son algunos de los aspectos a los que nos referimos. En efecto, en el seguro de automóviles el porcentaje observado de asegurados que no presenta siniestros resulta extraordinariamente elevado. Pero, además, hay que tener en cuenta que el hecho de que un cliente no declare ningún siniestro no significa necesariamente que no lo haya sufrido; el asegurado dispone de la posibilidad de decidir si declara o no los siniestros que tiene.

Aunque su idoneidad parece evidente, los modelos inflados de ceros han sido poco utilizados en el seguro de automóviles hasta el momento. Shankar *et al.* (1997) y Lee *et al.* (2002) constituyen las referencias más destacables en este terreno a nivel

¹ Otros estudios destacables sobre información asimétrica en el seguro del automóvil son los de Boyer y Dionne (1989), Puelz y Snow (1994), Dionne *et al.* (1999) y Chiappori y Salanié (2000).

internacional y, en el caso de España, podemos resaltar los trabajos de Melgar y Guerrero (2005) y Melgar *et al.* (2004, 2005).

El objetivo principal que planteamos en este trabajo es mostrar la idoneidad de los modelos inflados de ceros para cuantificar el número de siniestros en el seguro de automóviles. Para ello, se llevará a cabo un análisis empírico en el que se utilizarán datos cedidos por una entidad aseguradora privada española que opera a nivel nacional. Señalaremos las características del vehículo asegurado, del tomador del seguro, así como de las pólizas, que influyen en la observación de un mayor número de siniestros declarados.² Asimismo, estimaremos el número de asegurados que han sufrido siniestros pero no los han declarado, así como el número total de siniestros no declarados que esto supone. Finalmente, trataremos, desde un punto de vista teórico, las posibles ventajas que podrían tener otros modelos alternativos, como son los denominados modelos afinados (*thinned models*).

Tras esta introducción, en la Sección 2 se exponen los aspectos más destacados de los modelos inflados de ceros. A continuación, se muestran los resultados empíricos del estudio en la Sección 3, tanto en lo que se refiere al análisis descriptivo inicialmente realizado de los datos utilizados, como los que se deducen de la aplicación del modelo econométrico finalmente elegido para la estimación. En la Sección 4, se comentan de forma teórica las características de los modelos afinados y las mejoras que podrían aportar a nuestro estudio. Concluimos resaltando los principales resultados en la Sección 5, a la que le siguen la Bibliografía y un Anexo final que incluye una breve descripción de las variables usadas en el análisis.

2. Formulación de los modelos inflados de ceros

Los modelos de regresión de Poisson y binomial negativo son los modelos de recuento que se usan habitualmente cuando la variable endógena objeto de estudio toma valores discretos no negativos. Éste es el caso del análisis que vamos a realizar en este trabajo, encaminado a la determinación de los factores que influyen en una mayor o menor siniestralidad en el seguro de automóviles.

² En Ordaz y Melgar (2006), se usa un modelo *probit* para señalar cuáles de estos factores resultan significativos en la ocurrencia o no de siniestros.

La variable dependiente que consideramos, esto es, el número de siniestros que el asegurado declara a su compañía aseguradora, presenta una gran sobredispersión, provocada por el elevado porcentaje de valores nulos observados, como se verá más adelante. Este hecho hace que, desde un punto de vista teórico, el modelo de regresión de Poisson no resulte el más adecuado. Se debe recurrir entonces a modelos alternativos, como el binomial negativo, entre otros, para abordar dicha sobredispersión. En nuestro caso, y debido a las distintas interpretaciones que puede tener el valor nulo de la variable endógena, la utilización de modelos inflados de ceros puede ser una buena elección. En efecto, el hecho de observar 0 siniestros declarados puede indicar que el asegurado no ha sufrido realmente ningún siniestro a lo largo del periodo considerado, pero también puede significar que, pese a haber tenido alguno, ha decidido no declararlo a su compañía. Estos aspectos se pueden reflejar de forma adecuada a través de un modelo inflado de ceros.³

Las distribuciones infladas de ceros se basan en la hipótesis de que hay dos tipos de individuos en la población: uno para el que la variable observada toma siempre el valor 0 y otro para el que los datos se generan según la distribución tradicional que se considere (Poisson o binomial negativa). De este modo, si denotamos por Z_i a la variable que en principio nos interesa y que toma valores discretos no negativos (en nuestro caso, el número de siniestros que tiene el asegurado i a lo largo del año) y por Y_i a su versión inflada de ceros que es la que observamos (para nuestro estudio, el número de siniestros declarados), se puede escribir $Y_i = Z_i I_i$, donde I_i es una variable aleatoria de tipo Bernoulli que tomará el valor 1 si el cliente i declara sus siniestros y el valor 0 en caso contrario. Si denotamos por q_i la probabilidad para el asegurado i de no declarar los siniestros que sufre (probabilidad de “no participación”), se tendrá que la probabilidad de que los declare es: $P(I_i = 1) = 1 - q_i$. Si las variables explicativas se denotan por X_{i1}, \dots, X_{in} , entonces se suele considerar que:

$$q_i = F(\tau(\beta_0 + \beta_1 X_{i1} + \dots + \beta_n X_{in})), \quad (1)$$

³ Pueden consultarse Cameron y Trivedi (1998) o Winkelmann (2003) para más detalles sobre estos modelos.

donde F se refiere, generalmente, a la función de distribución logística o la normal estándar [Greene (1999)], dando así lugar a modelos *logit* o *probit*, respectivamente, y $\tau, \beta_0, \beta_1, \dots, \beta_n$ son parámetros desconocidos a estimar.

Con este planteamiento, la distribución de probabilidad inflada de ceros vendrá dada por:

$$\begin{aligned} P(Y_i = 0) &= q_i + (1 - q_i)P(Z_i = 0) \\ P(Y_i = k) &= (1 - q_i)P(Z_i = k), \quad k = 1, 2, \dots \end{aligned} \quad (2)$$

o, equivalentemente:

$$P(Y_i = k) = q_i(1 - \min\{k, 1\}) + (1 - q_i)P(Z_i = k), \quad k = 0, 1, 2, \dots \quad (3)$$

Tanto de la definición de q_i como de la ecuación (3), se deduce que en el caso en que $q_i = 0$, es decir cuando todos los asegurados declaran los siniestros que sufren, el modelo inflado de ceros se reduce al modelo tradicional.

Por otra parte, si N es el número de clientes considerados de la compañía aseguradora, se pueden obtener el número de asegurados que no declaran siniestros (N_0), así como el número total de siniestros no declarados (A_0), a partir de las expresiones:

$$N_0 = \sum_{i=1}^N \chi_{\{z_i > 0, I_i = 0\}}, \quad A_0 = \sum_{i=1}^N Z_i \chi_{\{z_i > 0, I_i = 0\}}, \quad (4)$$

donde $\chi_{\{A\}}$ es la función característica del suceso A . Los valores esperados y las varianzas respectivas de dichas variables son:

$$E(N_0) = \sum_{i=1}^N q_i(1 - P(Z_i = 0)), \quad Var(N_0) = \sum_{i=1}^N q_i(1 - P(Z_i = 0))\{1 - q_i(1 - P(Z_i = 0))\} \quad (5)$$

$$E(A_0) = \sum_{i=1}^N q_i E(Z_i), \quad Var(A_0) = \sum_{i=1}^N q_i \{Var(Z_i) + E(Z_i)^2(1 - q_i)\}. \quad (6)$$

Una vez elegida la distribución de Z_i (Poisson o binomial negativa), se podrán estimar todos los parámetros del modelo, indicar las variables que resultan significativas en la determinación del número de siniestros que los asegurados declaran a su compañía, y, a partir de las expresiones dadas en las ecuaciones (5) y (6), obtener intervalos de

confianza para el número de clientes que no han declarado siniestros y el número que estos siniestros no declarados representan.

2.1 Modelo de Poisson inflado de ceros

A partir del planteamiento anterior, y suponiendo que la variable Z_i sigue una distribución de Poisson de parámetro λ_i , se obtiene la función de probabilidad del modelo de Poisson inflado de ceros (*Zero-Inflated Poisson – ZIP*), que viene dada por:

$$P(Y_i = k) = q_i(1 - \min\{k, 1\}) + (1 - q_i)e^{-\lambda_i} \frac{\lambda_i^k}{k!}, \quad k = 0, 1, 2, \dots, \quad (7)$$

donde $\lambda_i = \exp\{\beta_0 + \beta_1 X_{i1} + \dots + \beta_n X_{in}\}$ y los demás parámetros y variables tienen el significado ya indicado en la ecuación (1).

En cuanto a los valores esperados y las varianzas del número de clientes (N_0) que no declaran sus siniestros y del número de siniestros no declarados (A_0), sus expresiones son:

$$E(N_0) = \sum_{i=1}^N q_i(1 - e^{-\lambda_i}), \quad \text{Var}(N_0) = \sum_{i=1}^N q_i(1 - e^{-\lambda_i})\{1 - q_i(1 - e^{-\lambda_i})\} \quad (8)$$

$$E(A_0) = \sum_{i=1}^N q_i \lambda_i, \quad \text{Var}(A_0) = \sum_{i=1}^N q_i \lambda_i (1 + \lambda_i (1 - q_i)). \quad (9)$$

2.2 Modelo binomial negativo inflado de ceros

Si suponemos ahora que Z_i se distribuye según una binomial negativa de parámetros ν y $\frac{\nu}{\nu + \lambda_i}$, se obtiene la distribución de probabilidades del modelo binomial negativo inflado de ceros (*Zero-Inflated Negative Binomial – ZINB*) según la expresión:

$$P(Y_i = k) = q_i(1 - \min\{k, 1\}) + (1 - q_i) \frac{\Gamma(k + \nu)}{\Gamma(k + 1) \cdot \Gamma(\nu)} \left(\frac{\nu}{\nu + \lambda_i}\right)^\nu \left(\frac{\lambda_i}{\nu + \lambda_i}\right)^k \quad (10)$$

donde $\nu > 0$, $\lambda_i = \exp\{\beta_0 + \beta_1 X_{i1} + \dots + \beta_n X_{in}\}$ y los demás parámetros y variables se definen como ya se ha hecho en la expresión general del modelo.

Si se introduce el llamado “parámetro de precisión” α [Jones (2001)], dado por $\alpha = \frac{1}{\nu}$, se verifica que cuando α tiende a cero la distribución de probabilidades del modelo ZINB tiende a la del modelo ZIP. Por este motivo, α puede considerarse como una medida del nivel de sobredispersión que presentan los datos y el contraste de su significatividad nos ayudará a decidir entre uno u otro modelo inflado de ceros.

En lo que se refiere al número de clientes que no declaran los siniestros que sufren (N_0) y al número total de siniestros no declarados resultante (A_0), en el caso del modelo ZINB, sus valores esperados y varianzas adoptan las siguientes expresiones:

$$E(N_0) = \sum_{i=1}^N q_i \left(1 - \left(\frac{\nu}{\nu + \lambda_i} \right)^\nu \right), \quad Var(N_0) = \sum_{i=1}^N q_i \left(1 - \left(\frac{\nu}{\nu + \lambda_i} \right)^\nu \right) \left\{ 1 - q_i \left(1 - \left(\frac{\nu}{\nu + \lambda_i} \right)^\nu \right) \right\} \quad (11)$$

$$E(A_0) = \sum_{i=1}^N q_i \lambda_i, \quad Var(A_0) = \sum_{i=1}^N q_i \lambda_i \left(1 + \frac{\lambda_i}{\nu} + \lambda_i (1 - q_i) \right) \quad (12)$$

2.3 Elección entre modelos inflados de ceros y modelos tradicionales

Según se ha indicado en apartados anteriores, los modelos inflados de ceros coinciden con los tradicionales cuando la probabilidad de no participación es nula: $q_i = 0$. Sin embargo, la hipótesis que se hace en relación a que q_i sigue una distribución logística o normal tiene el inconveniente de que el modelo tradicional y el inflado de ceros no están anidados, con lo que decantarse por uno u otro modelo no se reduce a contrastar la nulidad de un parámetro. En tal caso, la elección puede hacerse usando el estadístico de Vuong, que viene dado por:

$$V = \frac{\sqrt{N} \left[\frac{1}{N} \sum_{i=1}^N m_i \right]}{\sqrt{\frac{1}{N} \sum_{i=1}^N (m_i - \bar{m})^2}}, \quad (13)$$

donde $m_i = \log \left(\frac{P_1(Y = y_i)}{P_2(Y = y_i)} \right)$, $P_1(Y = y_i)$ y $P_2(Y = y_i)$ son las distribuciones de probabilidad correspondientes al modelo inflado de ceros y al modelo tradicional, respectivamente, y \bar{m} es la media de m_i , $i = 1, \dots, N$.

Se trata de un estadístico bidireccional con una distribución asintótica normal tipificada [Vuong (1989)]. Considerando un nivel de significación del 5%, si V toma un valor mayor que 1,96, el modelo inflado de ceros será el elegido frente al tradicional; si V es menor que -1.96, la especificación tradicional del modelo será más adecuada; por contra, en el intervalo comprendido entre ambos valores la decisión no resulta clara.

3. Resultados empíricos

Se han aplicado los distintos modelos de recuento señalados para determinar los principales factores que resultan significativos al estimar el número de siniestros que los asegurados sufren y declaran a sus compañías. Con el modelo finalmente elegido, asimismo, se han estimado el número de asegurados que no declaran siniestros y el número de siniestros no declarados que corresponden a éstos. Para ello, se ha utilizado una base de datos proporcionada por una entidad aseguradora privada española presente en el mercado nacional.

3.1 Aspectos descriptivos de la base de datos

La información disponible hace referencia al periodo comprendido entre el 16 de junio de 2002 y el 15 de junio de 2003, e incluye datos relativos a características de los vehículos asegurados (tipo y uso), del titular de la póliza (edad, sexo, antigüedad del permiso de conducción y región de residencia) y de la propia póliza (prima anual y grado de cobertura aseguradora). Se conoce también el número de siniestros que cada asegurado ha declarado a la compañía durante el periodo considerado. La definición exacta de todas estas variables puede verse en el Anexo final del trabajo. Por su parte, la Tabla 1 y la Tabla 2 ofrecen un análisis descriptivo de las mismas.

En la Tabla 1 se indica, a nivel de cada variable, el porcentaje que cada categoría supone sobre el total de la muestra seleccionada, así como el porcentaje de cada categoría que sufre y declara siniestros.

Hay que señalar que, por motivos computacionales, se ha elegido finalmente una muestra aleatoria de 15.000 pólizas del total de 130.000 de que disponíamos inicialmente en la base de datos referida.

En cuanto a la Tabla 2, en ella se muestra la distribución del número de siniestros declarados por los 15.000 asegurados considerados, así como únicamente la de aquéllos

que han declarado algún siniestro. Destaca el hecho de que sólo 3.442 asegurados han declarado algún tipo de siniestro, lo que supone el 22,9% del total. De éstos, el 90,3% sólo declara 2 siniestros a lo sumo.

Tabla 1. Distribución de las categorías y de la siniestralidad por variable

<i>Variable</i> Categoría	Porcentaje sobre el total de la muestra	Porcentaje con siniestros, sobre el total de la categoría
<i>Tipo del vehículo asegurado</i>		
Turismo-furgoneta	80,5%	26,5%
Camión	1,3%	25,3%
Autocar	0,2%	52,2%
Ciclomotor-moto	7,7%	7,0%
Vehículo especial	10,3%	6,8%
<i>Uso del vehículo asegurado</i>		
Particular	79,8%	24,7%
Profesional	19,6%	16,3%
Otros	0,6%	12,0%
<i>Edad del asegurado</i>		
De 18 a 25 años	3,1%	23,4%
De 26 a 45 años	39,8%	24,2%
De 46 a 70 años	51,8%	22,7%
De 71 años y más	5,3%	15,9%
<i>Experiencia como conductor del asegurado</i>		
Menos de 2 años	0,8%	35,5%
2 años y más	99,2%	22,9%
<i>Sexo del asegurado</i>		
Hombre	85,3%	22,3%
Mujer	14,7%	26,5%
<i>Región de residencia del asegurado</i>		
Sur	46,3%	24,0%
Canarias	4,4%	21,3%
Centro	16,8%	19,0%
Ceuta y Melilla	0,4%	25,0%
Este	12,1%	24,4%
Madrid	1,4%	28,7%
Noreste	3,2%	24,4%
Noroeste	15,4%	22,7%
<i>Prima anual de la póliza</i>		
Hasta 300 €	32,2%	11,8%
De 301 a 400 €	26,8%	22,6%
De 401 a 600 €	23,2%	28,1%
Más de 600 €	17,8%	36,9%
<i>Cobertura de la póliza</i>		
Grado bajo	54,3%	16,1%
Grado medio	37,8%	29,3%
Grado alto	7,9%	39,4%
<i>Total</i>	100%	22,9%

Tabla 2. Distribución del número de siniestros declarados

Número de siniestros declarados	Frecuencia	Incluyendo 0		Sin incluir 0	
		Porcentaje	Porcentaje acumulado	Porcentaje	Porcentaje acumulado
0	11.558	77,1%	77,1%		
1	2.365	15,8%	92,9%	68,7%	68,7%
2	743	4,9%	97,8%	21,6%	90,3%
3	223	1,5%	99,3%	6,5%	96,8%
4	78	0,5%	99,8%	2,2%	99,0%
5	19	0,1%	99,9%	0,6%	99,6%
6	10	0,1%	100,0%	0,3%	99,9%
7	4	0,0%	100,0%	0,1%	100,0%
<i>Total</i>	15.000	100%		100%	

3.2 Estimación econométrica

La estimación del modelo se ha llevado a cabo con el software *Limdep 7.0*, usando la distribución logística para q_i , es decir $q_i = \Lambda(\tau(\beta_0 + \beta_1 X_{i1} + \dots + \beta_n X_{in}))$ [Greene (1995)]. Tras realizar varias simulaciones, nos hemos quedado finalmente con la especificación del modelo ZINB que puede observarse en la Tabla 3. En ella, figuran únicamente las categorías que han resultado significativas a un nivel de confianza del 95%.

Los motivos que nos han llevado a tal elección son varios.

En primer lugar, ya hemos indicado que, a nivel teórico, los modelos inflados de ceros parecen los más adecuados, debido a la sobredispersión existente en los datos y a la doble interpretación del valor 0 en los siniestros declarados. Además, como puede verse en la Tabla 3, el estadístico de Vuong tiene un valor de 9,6978; al ser superior a 1,96, este modelo sería preferido al tradicional binomial negativo, tal como se ha indicado en la Sección 2.3.

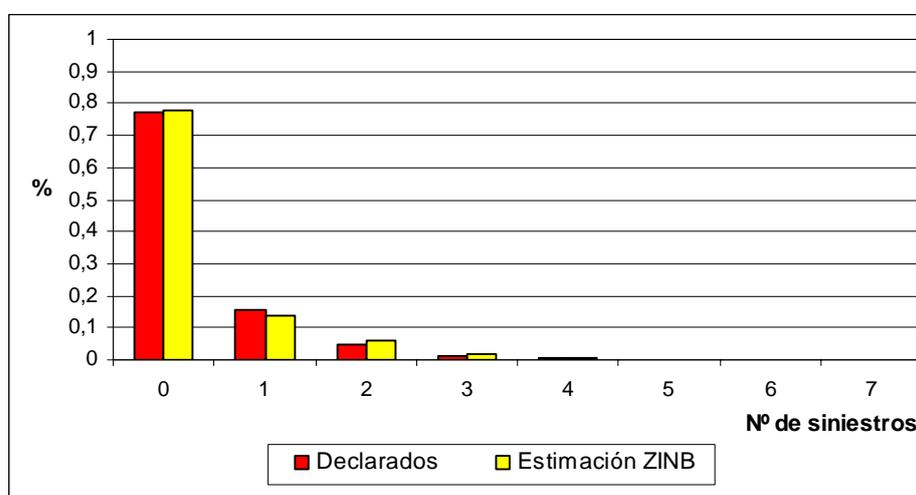
Si comparamos el modelo ZINB con el modelo ZIP, el primero parece ofrecer un mejor ajuste que el segundo, como así lo indica la significatividad del coeficiente de sobredispersión α (*Alpha*), cuyo p-valor es 0,0017. No obstante, hay que indicar que los resultados cualitativos que se obtienen, tanto en lo que hace referencia a las variables significativas como en lo relativo a su comportamiento frente a la siniestralidad, son bastante similares en ambos modelos.

Tabla 3. Resultados de la estimación del número de siniestros declarados, según el modelo ZINB

Modelo: Binomial Negativo Inflado de Ceros (Zero-Inflated Negative Binomial – ZINB)		
Método de estimación: Máxima verosimilitud		
Número de observaciones: 15.000		
Variable dependiente: NUMSIN		
Variables	Coficiente	P-valor
Constante	-0,36406	0,0000
<i>Tipo del vehículo asegurado</i>		
AUTOCAR	0,81861	0,0000
CICL_MOT	-0,93411	0,0000
VEH_ESP	-0,84024	0,0000
<i>Uso del vehículo asegurado</i>		
OTR_USOS	-0,62848	0,0026
<i>Experiencia como conductor del asegurado</i>		
ANTIG<2A	0,58427	0,0000
<i>Región de residencia del asegurado</i>		
CENTRO	-0,19388	0,0000
NOROESTE	-0,11234	0,0054
<i>Cobertura aseguradora de la póliza</i>		
GR_MEDIO	0,31675	0,0000
GR_ALTO	0,59024	0,0000
Log. de verosimilitud	-10.761,0	
Coficiente de sobredispersión (<i>Alpha</i>)	0,07268	0,0017
Parámetro de inflación de ceros (<i>Tau</i>)	-0,74384	0,0000
Estadístico de Vuong	9,6978	

En la Figura 1, puede apreciarse también, de manera gráfica, la alta capacidad predictiva de la estimación ZINB, lo que puede resultar un indicio más de la idoneidad de este modelo.

Figura 1. Distribución porcentual del nº de siniestros declarados y probabilidades estimadas por el modelo ZINB



Pasamos a continuación a analizar los resultados obtenidos.

En lo que se refiere al tipo de vehículo asegurado, tomando como base la categoría de los turismos y furgonetas, se observa que todas las demás han resultado ser significativas a excepción de los camiones. Los autocares muestran una mayor siniestralidad, mientras que los ciclomotores y motos y los vehículos especiales presentan una menor siniestralidad.

Si analizamos el uso principal al que se destina el vehículo, podemos comprobar que únicamente es significativa la categoría correspondiente a otros usos, siendo su siniestralidad menor que la que se observa en los usos particular y profesional.

La siguiente variable que resulta relevante es la relativa a la antigüedad en la posesión del permiso de conducir. Lo hace además en el sentido esperado, puesto que el estudio señala que los conductores con menos de 2 años de experiencia en la conducción presentan una mayor siniestralidad que los conductores con una experiencia más dilatada.

En cuanto a la región de residencia, que se supone de circulación habitual del asegurado, las Regiones Centro y Noroeste apuntan un comportamiento significativamente distinto al resto del territorio nacional, y lo hacen mostrando una menor siniestralidad.

Por último, en relación con la cobertura aseguradora de la póliza, hay que señalar que los grados medio y alto aparecen ambos como significativos frente al grado más bajo. Los coeficientes asociados a estos niveles de cobertura muestran una relación positiva y creciente con el número de siniestros declarados, lo que refleja el comportamiento esperado cuando existe información asimétrica en los datos, lo que es habitual en el seguro de automóviles. En efecto, en este ámbito el riesgo moral puede traducirse en el descuido del conductor al saber que disfruta de una amplia cobertura y, por otro lado, la selección adversa suele aparecer cuando un conductor, consciente de su propensión a sufrir siniestros, contrata una cobertura elevada para sentirse más protegido.

De las variables consideradas inicialmente en el análisis, algunas no aparecen finalmente entre las significativas a través de ninguna de sus categorías. En lo que respecta a la edad del asegurado y a la prima anual de la póliza, se han suprimido del estudio porque presentan una fuerte relación de endogeneidad con la antigüedad de

posesión del permiso de conducción y con los grados de cobertura, respectivamente. El efecto de la edad probablemente se vea reflejado en la variable relativa a la experiencia en la conducción. En cuanto a la prima, se ha preferido optar por analizar la influencia del grado de cobertura, ya que su comportamiento en relación a la siniestralidad tiene una interpretación más clara y directa, como se ha visto. Otra variable que tampoco ha resultado esencial en el estudio de la siniestralidad es el sexo del asegurado. No se han observado diferencias estadísticamente significativas entre hombres y mujeres.

En relación con las estimaciones del número de asegurados que no han declarado siniestros y del número de siniestros no declarados, éstas han podido calcularse a partir de las expresiones (11) y (12) dadas en la Sección 2.2. Los resultados pueden observarse en la Tabla 4.

Tabla 4. Resultados de las estimaciones del número de asegurados que no declaran siniestros y del número de siniestros no declarados

	Asegurados que no declaran siniestros	Siniestros no declarados
Número estimado	4.049	5.946
Intervalo de confianza al 95%	(3.941-4.157)	(5.761-6.130)

4. Posibles extensiones metodológicas

La formulación de los modelos inflados de ceros se basa, como hemos visto, en la siguiente relación entre el número de siniestros declarados (Y_i) y el número de siniestros sufridos por cada asegurado (Z_i): $Y_i = Z_i I_i$, de manera que el asegurado i declara los siniestros que sufre a lo largo del periodo considerado con probabilidad $P(I_i=1) = 1 - q_i$.

Con este planteamiento, se asume la hipótesis implícita de que el individuo o bien declara todos los accidentes que sufre, o bien no declara ninguno. Sin embargo, la hipótesis podría ser diferente y referirse a la posibilidad de que el asegurado decidiese, cada vez que sufre un siniestro, si declararlo o no.

En este caso, se podría proponer un modelo en el que el número de siniestros declarados viniese dado por $W_i = I_{i0} + I_{i1} + \dots + I_{iZ_i}$, donde Z_i es el número total de accidentes que el individuo i tiene a lo largo del periodo en estudio e I_{ij} indica si el j -ésimo siniestro

ha sido declarado o no (se tomará $I_{i0} = 0$ con probabilidad 1).⁴ Si se supone que $I_{i1}, I_{i2}, I_{i3}, \dots$ son variables Bernoulli de parámetro $1 - q_i$, independientes e idénticamente distribuidas, se pueden determinar las funciones de distribución de los llamados modelos afinados. Dependiendo de que Z_i siga una distribución de Poisson o una binomial negativa, se llegará, respectivamente, a la formulación del modelo de Poisson afinado (*Thinned Poisson model* – TP) o a la del modelo binomial negativo afinado (*Thinned Negative Binomial model* – TNB).

En cuanto al número de asegurados que no declaran siniestros (N_0) y al número de siniestros no declarados (A_0), se pueden expresar de la forma:

$$N_0 = \sum_{i=1}^N \mathcal{X}_{\{W_i < Z_i\}}, \quad A_0 = \sum_{i=1}^N (Z_i - W_i), \quad (14)$$

siendo sus valores esperados y varianzas respectivas:

$$E(N_0) = \sum_{i=1}^N P(W_i < Z_i), \quad \text{Var}(N_0) = \sum_{i=1}^N P(W_i < Z_i)P(W_i = Z_i) \quad (15)$$

$$E(A_0) = \sum_{i=1}^N q_i E(Z_i), \quad \text{Var}(A_0) = \sum_{i=1}^N \text{Var}(Z_i - W_i). \quad (16)$$

Deducimos, seguidamente, las funciones de probabilidad de ambos modelos afinados, así como las expresiones que adoptan las ecuaciones (15) y (16) en cada caso.

4.1 Modelo de Poisson afinado

Si Z_i se distribuye según una Poisson de parámetro λ_i , la distribución de probabilidades de W_i viene dada por:

$$P(W_i = k) = \sum_{j=k}^{\infty} P(Z_i = j) \binom{j}{k} (1 - q_i)^k q_i^{j-k} = e^{-\lambda_i(1-q_i)} \frac{(\lambda_i(1-q_i))^k}{k!}. \quad (17)$$

⁴ Los modelos inflados de ceros estándar se pueden considerar como un caso particular de este modelo más general, en los que se supone que $I_{i1} = I_{i2} = \dots = I_{iz_i} = I_i$ y, por tanto, la decisión para cada siniestro es la misma y es independiente del número total de siniestros sufridos, Z_i .

Es decir, W_i sigue una distribución de Poisson de parámetro $\lambda_i(1 - q_i)$.

Por otra parte, los valores esperados y las varianzas de N_0 y A_0 son:

$$E(N_0) = \sum_{i=1}^N (1 - e^{-\lambda_i q_i}), \quad \text{Var}(N_0) = \sum_{i=1}^N e^{-\lambda_i q_i} (1 - e^{-\lambda_i q_i}) \quad (18)$$

$$E(A_0) = \sum_{i=1}^N \lambda_i q_i, \quad \text{Var}(A_0) = \sum_{i=1}^N \lambda_i q_i \quad (19)$$

4.2 Modelo binomial negativo afinado

En el caso en que Z_i sigue una distribución binomial negativa de parámetros ν y

$\frac{\nu}{\nu + \lambda_i}$, se deduce la siguiente expresión para la función de probabilidad de W_i :

$$P(W_i = k) = \sum_{j=k}^{\infty} P(Z_i = j) \binom{j}{k} (1 - q_i)^k q_i^{j-k} = \frac{\Gamma(k + \nu)}{\Gamma(k + 1)\Gamma(\nu)} \left(\frac{\nu}{\nu + (1 - q_i)\lambda_i} \right)^\nu \left(\frac{(1 - q_i)\lambda_i}{\nu + (1 - q_i)\lambda_i} \right)^k \quad (20)$$

Resulta por tanto que W_i se distribuye según una binomial negativa de parámetros ν y

$$\frac{\nu}{\nu + (1 - q_i)\lambda_i}.$$

En cuanto a los valores esperados de N_0 y A_0 y sus varianzas respectivas, vienen dados por las siguientes expresiones:

$$E(N_0) = \sum_{i=1}^N \left\{ 1 - \left(\frac{\nu}{\nu + \lambda_i q_i} \right)^\nu \right\}, \quad \text{Var}(N_0) = \sum_{i=1}^N \left(\frac{\nu}{\nu + \lambda_i q_i} \right)^\nu \left\{ 1 - \left(\frac{\nu}{\nu + \lambda_i q_i} \right)^\nu \right\} \quad (21)$$

$$E(A_0) = \sum_{i=1}^N \lambda_i q_i, \quad \text{Var}(A_0) = \sum_{i=1}^N \left(\lambda_i q_i + \frac{(\lambda_i q_i)^2}{\nu} \right) \quad (22)$$

Estos modelos podrían generalizarse aún más, relajando las hipótesis que se hacen sobre las variables aleatorias I_{ij} y sobre la probabilidad q_i , pero darán lugar sin duda a expresiones menos manejables que las aquí obtenidas. Algunas opciones podrían ser considerar que $I_{i1}, I_{i2}, I_{i3}, \dots$ son variables intercambiables, o que la probabilidad de declarar cada siniestro es distinta. Todas estas posibilidades abren nuevas líneas de investigación susceptibles de ser abordadas en el futuro, con el objetivo de intentar

mejorar los análisis que llevamos a cabo sobre la siniestralidad en el seguro de automóviles.

5. Conclusiones

El principal objetivo de este trabajo consistía en mostrar la idoneidad de los modelos inflados de ceros a la hora de determinar las variables más significativas en la estimación del número de siniestros que los asegurados sufren y declaran a sus compañías.

Frente a los modelos de recuento tradicionales, como el de Poisson o el binomial negativo, los modelos inflados de ceros resultan ser una alternativa válida, ya que permiten tener en cuenta ciertas características que no son consideradas expresamente por los primeros, como son las diferencias cualitativas entre los asegurados que sufren siniestros y los que no, así como la doble interpretación de las observaciones nulas, que pueden indicar tanto que un asegurado no ha sufrido siniestros, como que los ha sufrido pero ha decidido no declararlos.

Esta idoneidad teórica se ha visto ratificada por el análisis empírico llevado a cabo, para el que se han utilizado datos procedentes de una compañía de seguros privada española. Las variables explicativas que se han considerado en las estimaciones hacen referencia a características tanto del vehículo asegurado, como del conductor y de la póliza contratada. El estudio pone de manifiesto la diferencia de comportamiento, en referencia a la siniestralidad declarada, de los distintos tipos y usos de vehículos, de los asegurados según su experiencia y región de residencia, así como de los grados de cobertura. Además, se ha podido apreciar que el sexo del conductor no parece influir en la siniestralidad.

El estudio se ha completado con la estimación del número de asegurados que, habiendo sufrido algún siniestro, no lo ha declarado. También se ha estimado el número total de siniestros que no han sido declarados a pesar de haber ocurrido.

Para finalizar, se apuntan las posibles ventajas teóricas que pueden tener para el estudio de la siniestralidad, la aplicación de otros modelos alternativos, los llamados modelos afinados. Su principal aportación consiste en permitir considerar la posibilidad de que el asegurado decida, cada vez que tiene un siniestro, si lo declara o no, frente al

planteamiento de declarar todos los siniestros o ninguno. Esta nueva línea que planteamos podría ampliar estos análisis en el futuro.

Bibliografía

Boyer, M. y Dionne, G. (1989): “An Empirical Analysis of Moral Hazard and Experience Rating”, *Review of Economics and Statistics*, 71, pp. 128-134.

Cameron, A.C. y Trivedi, P.K. (1998): *Regression Analysis of Count Data*, Cambridge University Press.

Chiappori, P.A. y Salanié, B. (2000): “Testing for Asymmetric Information in Insurance Markets”, *Journal of Political Economy*, 108, 1, pp. 56-78.

Cohen, A. (2005): “Asymmetric Information and Learning: Evidence from the Automobile Insurance Market”, *Review of Economics and Statistics*, 87, 2, pp. 197-207.

Dionne, G.; Gouriéroux, C. y Vanasse, C. (1999): “Evidence of Adverse Selection in Automobile Insurance Markets”, en *Automobile Insurance: Road Safety, New Drivers, Risks, Insurance Fraud and Regulation*, Dionne, G. y Laberge-Nadeau, C. (Ed.), Kluwer Academic Publishers, Boston, pp. 13-46.

Greene, W.H. (1995): *Limdep Version 7.0: User's Manual*, Econometric Software, Bellport, New York.

Greene, W.H. (1999): *Análisis Económico*, 3ª edición, Prentice Hall-Pearson Educación, Madrid.

Jones, A. M. (2001): *Applied Econometrics for Health Economists - A Practical Guide*, Office of Health Economics, London.

Lee, A.H.; Stevenson, M.R.; Wang, K. y Yau, K.K.W. (2002): “Modeling Young Driver Motor Vehicle Crashes: Data with Extra Zeros”, *Accident Analysis and Prevention*, 34, pp. 515-521.

Melgar, M.C. y Guerrero, F.M. (2005): “Los siniestros en el seguro del automóvil: un análisis econométrico aplicado”, *Estudios de Economía Aplicada*, 23, 1, pp. 355-375.

Melgar, M.C.; Ordaz, J.A. y Guerrero, F.M. (2004): “The Main Determinants of the Number of Accidents in the Automobile Insurance: an Empirical Analysis”, *Études et Dossiers*, 286, pp. 45-56.

Melgar, M.C.; Ordaz, J.A. y Guerrero, F.M. (2005): “Diverses Alternatives pour Déterminer les Facteurs Significatifs de la Fréquence d'Accidents dans l'Assurance Automobile”, *Assurances et Gestion des Risques – Insurance and Risk Management*, 73, 1, pp. 31-54.

Ordaz, J.A. y Melgar, M.C. (2006): “Analysis of the Main Influential Factors in the Accidents of the Automobile Insurance”, XVI Jornadas Luso-Espanholas de Gestão Científica *Inovação e Empreendedorismo para o Séc. XXI – Finanças e Contabilidade*, pp. 461-472.

Puelz, R. y Snow, A. (1994): "Evidence on Adverse Selection: Equilibrium Signalling and Cross-Subsidization in the Insurance Market", *Journal of Political Economy*, 102, 2, pp. 236-257.

Richaudeau, D. (1999): "Automobile Insurance Contracts and Risk of Accident: An Empirical Test Using French Individual Data", *Geneva Papers on Risk and Insurance Theory*, 24, pp. 97-114.

Shankar, V.; Milton, J. y Mannering, F. (1997): "Modeling Accident Frequencies as Zero-Altered Probability Processes: an Empirical Inquiry", *Accident Analysis and Prevention*, 29, 6, pp. 829-837.

Vuong, Q.H. (1989): "Likelihood Ratio Tests for Model Selection and Non-Nested Hypotheses", *Econometrica*, 57, pp. 307-333.

Winkelmann, R. (2003): *Econometric Analysis of Count Data*, Springer, Berlin.

ANEXO: Definición de las variables utilizadas en el análisis

Categoría del vehículo asegurado

TUR_FUR = 1 si el vehículo asegurado es un turismo o una furgoneta; 0 en caso contrario (categoría base)

CAMION = 1 si el vehículo asegurado es un camión; 0 en caso contrario

AUTOCAR = 1 si el vehículo asegurado es un autocar; 0 en caso contrario

CICL_MOT = 1 si el vehículo asegurado es un ciclomotor o una moto; 0 en caso contrario

VEH_ESP = 1 si el vehículo asegurado es un vehículo especial; 0 en caso contrario

Uso del vehículo asegurado

USO_PART = 1 si el uso del vehículo asegurado es el uso particular; 0 en caso contrario (categoría base)

USO_PROF = 1 si el uso del vehículo asegurado es profesional; 0 en caso contrario

OTR_USOS = 1 si el vehículo asegurado se destina a otros usos; 0 en caso contrario

Edad del asegurado

ED18_25 = 1 si el asegurado tiene entre 18 y 25 años; 0 en caso contrario (categoría base)

ED26_45 = 1 si el asegurado tiene entre 26 y 45 años; 0 en caso contrario

ED46_70 = 1 si el asegurado tiene entre 46 y 70 años; 0 en caso contrario

ED71_ = 1 si el asegurado tiene 71 años o más; 0 en caso contrario

Experiencia como conductor del asegurado

ANTIG<2A = 1 si el asegurado obtuvo el permiso de conducción hace menos de 2 años; 0 en caso contrario

Sexo del asegurado

MUJER = 1 si el asegurado es mujer; 0 en caso contrario

Región de residencia del asegurado

CANARIAS = 1 si el asegurado reside en la región de Canarias (Canarias); 0 en caso contrario

CENTRO = 1 si el asegurado reside en la región Centro (Castilla y León, Castilla-La Mancha, Extremadura); 0 en caso contrario

CEU_MEL = 1 si el asegurado reside en la región de Ceuta y Melilla (Ciudades Autónomas de Ceuta y Melilla); 0 en caso contrario

ESTE = 1 si el asegurado reside en la región Este (Cataluña, Comunidad Valenciana, Islas Baleares); 0 en caso contrario

MADRID = 1 si el asegurado reside en la región de Madrid (Comunidad Autónoma de Madrid); 0 en caso contrario

NORESTE = 1 si el asegurado reside en la región Noreste (Aragón, Comunidad Foral de Navarra, La Rioja, País Vasco); 0 en caso contrario

NOROESTE = 1 si el asegurado reside en la región Noroeste (Cantabria, Galicia, Principado de Asturias); 0 en caso contrario

SUR = 1 si el asegurado reside en la región Sur (Andalucía, Región de Murcia); 0 en caso contrario (categoría base)

Prima anual de la póliza

P0_300 = 1 si la prima anual pagada por el asegurado no supera los 300 € 0 en caso contrario (categoría base)

P301_400 = 1 si la prima anual pagada por el asegurado es mayor de 300 € y no supera los 400 € 0 en caso contrario

P401_600 = 1 si la prima anual pagada por el asegurado es mayor de 400 € y no supera los 600 € 0 en caso contrario

P601_ = 1 si la prima anual pagada por el asegurado es mayor de 600 € 0 en caso contrario

Cobertura aseguradora de la póliza

GR_BAJO = 1 si el asegurado disfruta del grado de cobertura bajo; 0 en caso contrario (categoría base)

GR_MEDIO = 1 si el asegurado disfruta del grado de cobertura medio; 0 en caso contrario

GR_ALTO = 1 si el asegurado disfruta del grado de cobertura alto; 0 en caso contrario

Siniestralidad

NUMSIN = número de siniestros sufridos y declarados por el asegurado a la compañía durante el periodo de tiempo considerado

TESTING FOR NONLINEARITY IN AN ARTIFICIAL FINANCIAL MARKET: A GENERAL CORRELATION INTEGRAL BASED APPROACH

MARIANO MATILLA GARCÍA
mmatilla@cee.uned.es

CARMEN GARCÍA LLAMAS
cgarcia@cee.uned.es

LAURA RODRÍGUEZ FERNÁNDEZ
lrodriguez@cee.uned.es

Departamento de ECONOMÍA APLICADA CUANTITATIVA I
UNIVERSIDAD NACIONAL DE EDUCACIÓN A DISTANCIA

Área temática: Métodos Cuantitativos.

Resumen

En este trabajo se aborda la cuestión de si el concepto original de integral de correlación, que es revisado en el trabajo, puede mejorar las conclusiones y resultados obtenidos por estadísticos de contraste fundados en dicho concepto al analizar realizaciones obtenidas a partir de un mercado financiero artificial. En este trabajo se aplica una versión generalizada del BDS a los datos financieros dados por Lux y Marchesi. Como estas series han sido motivo de estudio en otros trabajos relevantes y con otras técnicas, se ofrece una comparación entre nuestras herramientas y las de otros.

Palabras clave: BDS test, delay coordinates method, non-linear dynamics, artificial financial market, correlation integral.

Abstract

This article address the question of whether recent work on testing for non-linearity on realizations from an artificial market can improve by revisiting the original concept of correlation integral upon which several tests are based. A generalized version of BDS test is applied to Lux and Marchesi's financial data set. Results among alternative tests are compared.

Key words: BDS test, delay coordinates method, non-linear dynamics, artificial financial market, correlation integral.

Testing for non-linearity in an artificial financial market: a general correlation integral based approach

Abstract

This article address the question of whether recent work on testing for non-linearity on realizations from an artificial market can improve by revisiting the original concept of correlation integral upon which several tests are based. A generalized version of BDS test is applied to Lux and Marchesi's financial data set. Results among alternative tests are compared.

Keywords. BDS test, delay coordinates method, non-linear dynamics, artificial financial market, correlation integral.

JEL classification. C12, C14, D84.

1 Introduction

Recently there has been increasing interest in testing for non-linear structures in artificial financial markets. In Chen et al. (2001) the BDS test of Brock et al. (1996) and Kaplan's test (1994) are applied to the study of the artificial financial data (prices and returns) generated by Lux and Marchesi (1999, 2000), who showed that their artificial series of prices and returns shared some stylized facts of empirical data such as the presence of a unit root in the asset price dynamics and leptokurtic returns. Chen et al. provide a "lack of robustness" result between both tests since acceptances and rejections of the null of linearity are obtained with different versions of the model presented by Lux and Marchesi. Recently, Balaire (2004) has studied the same data set with other non-linear test procedures such as Hinich's (1982) bispectral test, Whites's (1989) neural network test and the recurrence quantification analysis introduced by Zbilut and Webber (1992) which in turn is based on Takens' (1981) reconstruction theorem. Balaire's results show that the BDS test and Kaplan's test both perform worse than the others in detecting non-linear behavior. This conclusion contrasts sharply with results given by Barnett et al.'s (1997) single blind controlled competition among those tests.

It is well known that different methods for detecting non-linear structure are available, some of which are based on the correlation integral introduced by Grassberger and Procaccia (G-P) (1983)¹. The G-P procedure is based upon the *method of delay time coordinates*, introduced by Takens in order to carry out the phase space reconstruction. The reconstruction will depend on the embedding dimension, usually referred as m , and on the delay time, denoted by τ . Among those tests based on the phase space reconstruction, the BDS test has proven to be highly suitable for economic time series analysis and therefore it is expected to perform accordingly for artificial economic time series. However, as stated earlier, when the BDS test is applied to the artificial series studied by Chen et al. the results are not as satisfactory as expected. A possible explanation of this fact might be that the selection of parameters that the BDS statistic requires is not fine enough. Even though standard values for m were considered, for the computations carried out by Chen

¹Liu et al. (1998) compare some of these methods.

et al. delay time was fixed at 1. This is certainly understandable because the distribution and properties of the BDS statistic under different delay times have not been studied until recently. In this regard, Barnett et al. highlight the arbitrariness when fixing delay time to one.

Recent research (Aparicio et al.,1999) shows that the choice of delay time is crucial when estimating the correlation dimension (a usual procedure based on the correlation integral). It is stated that an unfortunate choice of delay time yields misleading results concerning the dimension of well known attractors.

Matilla-Garcia et al. (2004) provide a statistic for independence known as BDS-G. BDS-G is also based on the correlation integral, but it incorporates delay time choice thereby allowing the statistic to capture structures that the BDS statistic (which fixes delay time to 1) might not reveal.

The aim of this paper is to use the BDS-G test to analyze the simulated artificial financial data provided in Chen et al., and also to compare the results thus obtained with the methodologies given in Barnett et al., that were also used for analysing the same data set more recently by Belaire and Chen et al. In so doing, the power of the BDS-G test against volatility clustering can be better explained.

The paper is structured as follows. Section 2 explains the basis of BDS-G and several deterministic and stochastic series are studied in order to show the benefits of this statistic. In section 3 the BDS-G test is used to test for hidden structures in different versions of Lux and Marchesi's model. Finally, section 4 concludes.

2 The Correlation integral and the BDS-G test

The basic idea behind state space reconstruction is that the past and the future of a time series both contain information about unobserved state variables that can be used to define a state at the present time. Reconstruction is done from a scalar time series and all relevant components (relative to the underlying dynamics) have to be extracted from it. Takens showed that this type of reconstruction yields a topologically equivalent attractor leaving the dynamic parameters invariant. The required reconstruction will embed the univariate

observations into a multivariate phase space. To that end information is encapsulated in the delay vector called the m -history. Takens' proof indicates that for a smooth dynamic system f on a d -dimensional manifold M

$$s_t = f^t(s_0)$$

which is related with the time series by

$$u_t = h(s_t), \quad \text{with } h : M \rightarrow \mathbb{R}$$

It is a generic property that the delay reconstruction map

$$\Phi(s) = (h(s), h(f^\tau(s)), \dots, h(f^{\tau(m-1)}(s)))$$

is an *embedding* (a one-to-one differentiable mapping with a differentiable inverse) if $m \geq 2n + 1$. This implies that there is a strong relationship (differentiable equivalence) between f^t on M and

$$G^t(u_t^m) = \Phi \circ f^t \circ \Phi^{-1}(u_t^m)$$

on image $\Phi(M) \subset \mathbb{R}^{2m+1}$. Because of this equivalence G may be used to compute statistical characteristics of the original system such as dimensions, prediction, etc. (Renyi, 1970).

As stated above the G-P correlation integral is based on Takens' theorem, and it consists of the two following steps: (1) for established values of m and τ , to convert the scalar time series $\{u_1, u_2, \dots, u_N\}$ into a set of m -histories: $u_i^{m,\tau} = (u_i, u_{i+\tau}, \dots, u_{i+(m-1)\tau})$; (2) to compute the 'correlation integral':

$$c_{m,n}^\tau(\varepsilon) = \frac{1}{\binom{n}{2}} \sum_{i=1}^n \sum_{j=i+1}^n H(\varepsilon - \|u_i^{m,\tau} - u_j^{m,\tau}\|) \quad (1)$$

where $n = N - (m - 1)\tau$ is the number of m -histories with τ -delay time that can be formed from N observations; and H is the Heavside function so that $H(u_i^{m,\tau}, u_j^{m,\tau})$ takes

the value 1 if both observations are within distance ε of each other, and 0 otherwise².

The correlation integral is crucial because it is a generalized U -statistic with symmetric kernel $H(\varepsilon - \|u - v\|)$, and hence the asymptotic distribution of the correlation integral can be developed.

Recent research (Aparicio et al.) shows that the choice of delay time is crucial when estimating the correlation dimension (based on the correlation integral), to the extent that an unfortunate delay time choice yields misleading results concerning the dimension of well known attractors. As Kantz and Schreiber (2004) indicate, in spite of the proper choice of delay time being quite important in applications, the relevant mathematical framework has not been convincingly studied.

The BDS statistic assumes $\tau = 1$, therefore m successive observations are stacked in producing the embedded phase space vectors. The time series obtained from Chen et al. are noisy and finite. These properties are shared by series found in the real-world. These restrictions make the selection of delay time crucial: For very small τ the dynamics in the space state take place in the reconstructed space so that $h(s_t)$ and $h(s_{t+\tau})$ are almost linearly dependent (which is not the case for the real observable of a nonlinear system). On the other hand, a large delay time will cause the coordinates to disjoin by stretching and folding³, so this will lead to vectors whose components are (seemingly) randomly distributed in the embedding space.

In order to overcome this problem Matilla-Garcia et al. introduce BDS-G which incorporates delay time. The idea behind BDS-G is to look for relevant deviations of the behavior of the correlation integral (1) from what is to be expected under IIDness of the data, but in a different way to what is done under BDS. The generalized version of the test allows the researcher to choose a proper delay time. Particularly, if the data under consideration are IID, then it is shown that $\lim_{n \rightarrow \infty} c_{m,n}^\tau(\varepsilon) = (\lim_{n \rightarrow \infty} c_{1,n}^\tau(\varepsilon))^m$ almost surely for all $\varepsilon > 0$ and $m = 2, 3, 4, \dots$. The statistic BDS-G is

²As in Brock *et al.* we take $\|x\| = \max_{1 \leq k \leq m} \{|x_k|\}$.

³When we use the descriptives ‘very small’ and ‘large’, we follow Buzug *et al.* (1990), so that these expressions are referred to in relative terms as compared to the characteristic recurrence time – the reciprocal value of the dominant frequency obtained from the power spectrum.

$$W_{m,n}^\tau(\varepsilon) = \sqrt{n} \frac{c_{m,n}^\tau(\varepsilon) - (c_1(\varepsilon))^m}{V_m^\tau} \quad (2)$$

whose⁴ limiting distribution is a standard Normal under the null of IID. The estimate of the standard deviation (V_m^τ), which now depends on the selected delay time, is given in Matilla-Garcia et al. The application of the test is therefore as straightforward as the standard BDS test.

As an example, Table 1 displays the BDS-G statistic (on prewhited data sets) for two well known deterministic and chaotic data from systems – the Logistic equation⁵ and the Hénon model⁶. The test has been conducted by setting several delay times ($\tau = 1, 2, 3$), four embedding dimensions and $\varepsilon =$ the standard deviation of data. The sample size was 1000 observations.

Table 1. BDS-G tests for Logistic and Hénon Maps

Logistic	$m = 2$	$m = 3$	$m = 4$	$m = 5$
$\tau = 1$	2.9	17.0	24.2	30.9
$\tau = 2$	17.0	15.8	13.3	10.3
$\tau = 3$	40.2	36.1	30.4	31.4
Hénon	$m = 2$	$m = 3$	$m = 4$	$m = 5$
$\tau = 1$	3.9	7.4	5.6	2.6
$\tau = 2$	6.8	2.8	1.8	2.9
$\tau = 3$	11.0	10	9.0	9.0

In the case of the Logistic equation it is clear that for all m and τ the test rejects the null of IIDness, although there is a stronger rejection for $\tau = 3$. The rejection is not so

⁴ $n \equiv N - (m - 1)\tau$.

⁵Given by

$$x_t = ax_{t-1}(1 - x_{t-1})$$

with $a = 3.85$ and $x_0 = 0.2$. It is well known that for the selected value of a the real dynamics of this equation (weakly chaotic) contain an attracting cycle of period 3, and therefore it seems reasonable to think that the reconstructed dynamic with $\tau = 3$ will capture dependence more clearly than with other delay times, especially with a small number of observations.

⁶Given by

$$\begin{aligned} x_{t+1} &= y_t + 1 - 1.4x_t^2 \\ y_{t+1} &= 0.3x_t \end{aligned}$$

with $x_0 = 1$ and $y_0 = 0$.

clear for the Hénon map. In fact, in this model τ selection is crucial because the test is ambiguous for $\tau = 2$. This is not true for $\tau = 3$, which produces a very clear rejection of the null.

The importance of choosing an adequate delay time can be observed even more clearly by analysing the Anosov map⁷. Taking the initial conditions $x_0 = 0.2$, $y_0 = 0.1$, from the evolution of x_t we generate prewhitened time series with 250 and 2500 data-points. Table 2 shows the results for $m = 2, 3, 4, 5$, $\varepsilon = 1.5\sigma$, and the most remarkable values of τ . As can be observed, the BDS test (corresponding to $\tau = 1$) would not reject the null of independence in any case and we would conclude that the series come from a IID process, thus rejecting the hypothesis that they were generated by a non-linear deterministic system. However, there are other choices of τ which do allow the existing nonlinearity to be identified.

Table 2. BDS-G tests for Anosov Map

Anosov		$m = 2$	$m = 3$	$m = 4$	$m = 5$
$N = 250$	$\tau = 1$	0.77	0.89	1.15	1.51
	$\tau = 3$	1.55	2.78	3.76	4.02
$N = 2500$	$\tau = 1$	0.17	-0.44	-0.44	-0.24
	$\tau = 2$	-3.83	-3.14	-2.36	-1.98
	$\tau = 14$	-2.64	-2.85	-2.90	-3.14

Frequently nonlinear stochastic models are used as models for logarithmic return time series in finance (see Brock et al. 1991). In this regard, it is acknowledged that the power of BDS ($\tau = 1$) against dependence in higher order moments, such as conditional heteroskedasticity, is small. Table 3 shows the results of several BDS-G tests on 500 data points generated by a GARCH model⁸. Additionally, results regarding the shuffled data

⁷Given by the equations:

$$\begin{cases} x_{t+1} = x_t + y_t \pmod{1} \\ y_{t+1} = x_t + 2y_t \pmod{1} \end{cases}$$

This is an ergodic, structurally stable automorphism of the two-dimensional torus obtained by identifying points whose coordinates differ by integers. Therefore all orbits remain inside the unit square and, unless periodic, fill up the entire square.

⁸The GARCH process is:

$$\begin{aligned} y_t &= h_t^{1/2} u_t \\ h_t &= 1 + 0.1y_{t-1}^2 + 0.8h_{t-1} \end{aligned}$$

set are also given. Interestingly, the BDS-G statistic at $\tau = 1$ does not reject the null of IIDness for the GARCH model, however, for $\tau = 2$ the statistic does reject the null. When the data are randomly shuffled, then the null is not rejected for both delay times because the structure has been melted. Once again, the importance of a proper delay time selection is clear. For that reason, the next section analyses artificial financial data taking into account the delay time.

Table 3. BDS-G statistics for GARCH data set

GARCH	$m = 2$	$m = 3$	$m = 4$	$m = 5$
$\tau = 1$	0.56	1.59	1.89	2.37
$\tau = 2$	2.26	3.04	3.45	3.89
SHUFFLED GARCH	$m = 2$	$m = 3$	$m = 4$	$m = 5$
$\tau = 1$	-1.11	-0.23	0.23	0.62
$\tau = 2$	-0.02	0.26	-0.20	-0.10

3 Testing for non-linear structure in the simulated data

In this section, BDS-G tests are applied to different versions of Lux and Marchesi's (1999) interacting agents model. Artificial data were generated without any change of the underlying mechanism.

In order to compute BDS-G tests one has to choose several parameters: (a) the delay time τ , (b) the embedding dimension m , and (c) the radius ε . In order to select the dimensional distance ε and the embedding dimension m , it is suggested that one follows the indications of Brock et al. (1991) and Kanzler (1999), who have found that the power of the BDS test is maximized when dimensional distance is selected between 1 and 2 times the standard deviation. Moreover Kanzler points out that as the embedding dimension m increases, the BDS distribution moves away from its asymptotic distribution. This result is also true for the BDS-G distribution. Several methods exist to select τ (see Aparicio et al.), and in this paper delay time selection has been based on the first minimum of the mutual information (see Fraser and Swinney, 1986).

with $h_0 = 1$ and $y_0 = 0$.

Practitioners of BDS and BDS-G tests usually consider different embedding dimensions (as we have done in Section 2). In the remainder of the paper this practice is followed and four embedding dimensions are studied. In this way the results can be easily compared with those provided by Chen et al. For the same reason we have set $\varepsilon = 1\sigma$.

Table 4 collects the outcomes of a sequence of BDS-G tests applied to the 20 data point windows used in Chen et al. To avoid rejections of the null hypothesis due to linear dependence, the BDS-G test has been applied to the estimated residuals of ARMA models. In order to compare results, the ARMA filters are those used by Chen et al.. The results are classified according to the following criteria: ‘acceptance’ of the null of IIDness, if (given the selected τ) none of the test statistics over $m = 2$ to $m = 5$ are significant at the 95% level; ‘rejection’ if all of them are significant; finally, mixed results are classified as ‘ambiguous’. The BDS column of Table 4 reproduces the results obtained by Chen et al.

Table 4. BDS-G test results

Subsample	ARMA	BDS-G	BDS	Subsample	ARMA	BDS-G	BDS
1	(0,0)	Reject	Reject	11	(0,2)	Reject	Ambiguous
2	(1,0)	Reject	Accept	12	(0,0)	Reject	Reject
3	(0,0)	Reject	Accept	13	(0,2)	Reject	Ambiguous
4	(1,0)	Ambiguous	Accept	14	(1,0)	Reject	Reject
5	(1,0)	Accept	Accept	15	(1,0)	Accept	Accept
6	(1,0)	Accept	Accept	16	(0,2)	Reject	Ambiguous
7	(1,0)	Accept	Accept	17	(2,1)	Reject	Reject
8	(1,0)	Ambiguous	Accept	18	(0,2)	Reject	Accept
9	(2,2)	Reject	Ambiguous	19	(2,2)	Reject	Ambiguous
10	(1,0)	Ambiguous	Accept	20	(1,0)	Accept	Accept

The results in Table 4 highlight the greater power of the BDS-G statistic over the BDS test. The BDS-G test rejects the null of IIDness in 12 out of 20 cases, while BDS rejects only 4. Interestingly, all of the subsamples that were classified as ‘ambiguous’ with the BDS test, are now transformed into rejections when a proper delay time is selected. Only 5 cases have been ‘accepted’, and 3 are now ambiguous.

As stated in the introduction, Chen et al. and Belaire conduct some other procedures complementary to BDS such as Kaplan's, Hinich's, White's and RQA's tests. Out of the 20 cases, the first three of these tests reject IIDness three, eight and nine times, respectively. Therefore the BDS-G test also out-performs these three complementary tests for detecting non-linear behavior in Lux and Marchesi's model.

RQA's test has been proved (see Belaire) to be a promising methodology for detection of non-linear structure. This procedure takes into account several variables, but it only focuses on three of them to classify subsamples. Strictly, these three variables measure the deterministic content of the time series. Therefore, one should expect that a clear rejection of the null of IIDness is obtained when *all* of the three estimated variables reject the null. Under this methodology, reviewing Belaire's results, 12 out of 20 subsamples strictly reject the null. In addition, it is worth noting that *four* parameters have to be chosen in order to compute RQA's test, three of which are those needed for the BDS-G test: the delay time, the embedding dimension m and the radius ε . Hence, the researcher needs to select an additional parameter in order to compute the RQA test. Interestingly, m is fixed when computing the RQA test in Belaire's paper, although BDS-G (and hence BDS as a particular case of it) tests along several ms . It is straightforward that if m were fixed beforehand then the BDS-G results would have been at least as powerful as those obtained here. In fact we have computed these calculations. We take $m = 15$ and the same delay time used in table Table 4, which allows for comparisons with RQA test. As was expected, under this embedding dimension the results improve: 17 out of 20 rejections. Now for subsamples 4, 6, 7, 8 and 10 the null is rejected. It is remarkable that all tests used for non-linearity (ie., BDS, BDSG, RQA, Kaplan, Hinich and White) are unable to reject the null for subsamples 5, 15 and 20, while these subsamples are also incompatible with GARCH structure.

It is well known that non-linearity in financial data seems to be contained in the second moments, while the power of the BDS test against dependence in higher order moments, such as conditional heteroskedasticity, is small. This seems to remain true for artificial financial data, since GARCH estimations carried out by Chen et al. show the presence of this kind of structure in 17 out of 20 subsamples, despite the fact that the

null of IIDness was only rejected in 4 subsamples. This poor performance of the BDS test in the presence of GARCH effects might be reconsidered by reviewing the nature of the correlation integral. Indeed, from the results shown in Table 4, firstly we can conclude that the BDS-G test rejects the null of independence even for tranquil subsamples like 2, 3 and 16 which were not rejected by considering $\tau = 1$. Secondly, more volatile subsamples like 9, 11 and 13 are now clearly ‘rejected’. In this way, the ability of BDS-G to deal with GARCH processes is also highlighted by reconsidering Table 3: for $\tau = 2$ the null of IIDness is rejected, while this conclusion is ambiguous for $\tau = 1$.

In addition, the presence of GARCH effects in so many subsamples contrasts with the results shown in Table 4. It seems reasonable to think that in most of the cases in which BDS-G rejects the null of IIDness it is because it detects GARCH effects. This is not the case of subsamples 5, 15 and 20, for which BDS-G is unable to detect any structure. These subsamples share a remarkable feature: the use of BDS and BDS-G tests fail even without taking the linear dependence out. That is, if we apply these tests to the original (non-filtered) subsamples 5, 15 and 20, we get the same final result: an acceptance of the null of IIDness. Therefore they reveal themselves to be unsuitable for detecting any kind of dependence (linear or nonlinear) in these cases, corresponding to turbulent periods. This could be due to the fact that the random switches between different kinds of agents modeled in the artificial financial market play a key role, especially over turbulent periods.

4 Conclusions

In this article, the results regarding Lux and Marchesi’s simulated artificial financial market have been revisited. To that end, the BDS-G test proposed by Matilla-García et al. has been used. One of the contributions of this article is to show that the use of this statistic, which is rooted on a more general interpretation of the correlation integral, clearly reduces the difficulties for classifying different versions of the artificial market. From a boarder perspective and by extension of Aparicio et al.’s empirical evidence, an adequate delay time selection leads to a more robust classification of results when tests based on correlation integral are involved.

The BDS-G test out-performs other non-linear tests such as Hinich's bispectral test, Whites's neural network test and Kaplan's test for detecting non-linear behavior in the data set. However, the procedure known as recurrence quantification analysis introduced by Zbilut and Webber seems to perform at least as well as the BDS-G method. Despite its power for detecting non-linear structure, RQA's test needs an additional parameter choice that might complicate the procedure.

References

- Albano, A.M., J. Muench, C. Schwartz, A.I. Mees, and P.E. Rapp (1988). "Singular-value decomposition and the Grassberger-Procaccia Algorithm." *Physical Review A* 38: 3017-3027.
- Aparicio, M.T., E. Pozo, and D. Saura (1999). "Investigating the chaotic behavior in Economic Series: the Delay-time in the Grassberger-Procaccia Algorithm." *International Journal of theoretical and Applied Finance*, vol 2.
- Barnett, W., A.R. Gallant, M.J. Hinich, J.A. Jungeilges, D.T. Kaplan, and M.J. Jensen (1997). "A single-bind controlled competition among tests for nonlinearity and chaos." *Journal of Econometrics* 82: 157-192.
- Belaire-Franch, J. (2004). "Testing for non-linearity in an artificial financial market: a recurrence quantification approach" *Journal of Economic Behavior and Organization* 54: 483-494
- Brock, W.A., W.D. Dechert, J.A. Scheinkman, and B. LeBaron (1996). "A test for independence based on the Correlation Dimension." *Econometric Reviews* 15: 197-235.
- Brock, W.A., D. Hsieh, and B. LeBaron (1991). *Nonlinear Dynamics, Chaos and Instability: Statistical Theory and Economic Evidence*. Cambridge, MA, MIT Press.
- Chen, S.-H., Lux, T. and M., Marchesi (2001). "Testing for non-linearity in an artificial financial market" *Journal of Economic Behavior and Organization* 46: 327-342.
- Buzug, Th., T. Reimeres, and G. Pfister (1990). "Optimal Reconstruction of strange attractors from purely geometrical arguments." *Europhysics Letters* 13: 605-610.

- de Lima, P.J.F. (1998) “Nonlinearities and nonstationarities in stock returns” *Journal of Business and Economic Statistics* 16: 227-236.
- Fraser, A.M. and H.L. Swinney (1986) Independent Coordinates for Strange Attractors from Mutual Information, *Physical Review Letters*, 33:1134-1140.
- Grassberger, P., and I. Procaccia (1982). “Characterization of Strange Attractors.” *Physical Review Letters* 50: 346-349.
- Hinich, M.J., (1982) “Testing for gaussianity and linearity of a stationary time series” *Journal of Time Series Analysis* 3:169-176.
- Kanz, H., and T. Schreiber (2004). *Nonlinear Time Analysis Series*. Cambridge, Cambridge University Press, 2th edition.
- Kanzler, L. (1999). *Very fast and correctly sized estimation of the BDS statistic*. Department of Economics of Oxford University. <http://users.ox.ac.uk/~econlrk>.
- Kaplan, D.T. 1994. “Exceptional events as evidence for determinism” *Physica D*. 73: 38-48
- Liu, T., C.W.J. Granger, and W.P. Heller (1998). “Using correlation exponent to decide whether an economic time series is chaotic.” *Journal of Applied Econometrics* 7: 25-39.
- Lux, T. Marchesi, M., (1999) “Scaling and criticality in a stochastic multi-agent model of a financial market” *Nature* 37: 498-500.
- Lux, T. Marchesi, M., (2000) “Volatility clustering in financial markets: a micro-simulation of interacting agents” *International Journal of Theoretical and Applied Finance* 3: 675-702.
- Matilla-García, M., Queralt, R., P. Sanz, and F.J. Vázquez (2004). “A Generalized BDS Statistic” *Computational Economics* 24: 277-300.
- Renyi, A. (1970). *Probability Theory*. Amsterdam, North-Holland.
- Takens, F. (1981). *Detecting Strange Attractors in Turbulence. Dynamical Systems and Turbulence*. Berlin, Springer-Verlag.

White, H., 1989. An additional hidden unit test for neglected nonlinearity in multilayer feedforward networks. In: Proceedings of the International Joint Conference on Neural Networks, Washintong D.C. IEEE Press, New York, pp. 451-455.

Zbilut, J.P. Webber, C.L., (1992). Embeddings and delays as derived from quantification of recurrence plots. Physics Letters A 171: 199-203.

LOS PROCESOS ESTABLES COMO GENERALIZACIÓN DEL MOVIMIENTO BROWNIANO: EL IBEX35

JESÚS MUÑOZ SAN MIGUEL
e-mail: jmiguel@us.es

JOSÉ JAVIER BUSTO GUERRERO
e-mail: jjbusto@us.es

Departamento de ECONOMÍA APLICADA I
UNIVERSIDAD DE SEVILLA

Área temática: Métodos cuantitativos.

Resumen

El movimiento browniano, caracterizado por la independencia y la normalidad de la distribución de sus incrementos, es uno de los modelos más utilizados para describir el precio de una acción. Sin embargo, su distribución empírica difiere de la distribución normal. En los años sesenta Benoît B. Mandelbrot (1924-) propuso como generalización del movimiento browniano los procesos estables como modelo de la evolución de los precios de un activo financiero manteniendo la independencia de los incrementos del proceso. En este trabajo se analiza algunas de las aportaciones de esta generalización tomando como modelo la serie de los cierres diarios del índice Ibex35 durante la década de los noventa.

Palabras clave: Proceso estocástico, Movimiento browniano, Proceso estable, Ibex35.

Abstract

Brownian motion, whose increments are Gaussian and independent, is one of the most used models to describe stock prices. However, its empirical distribution differs from the Gaussian one. During the sixties Benoit B. Mandelbrot (1924-) suggested stable processes, which generalize Brownian motion, as a model of price variation. In this paper is analyzed some features of stable processes taking the Ibex35 series as a model.

Key words: Stochastic process, Brownian motion, Stable process, Ibex35.

1.-Introducción

El movimiento browniano, caracterizado fundamentalmente por la independencia y normalidad de la distribución de sus incrementos, aparece por primera vez como modelo para describir la evolución de los precios de un activo financiero como un proceso estocástico en 1900 en la tesis de L. Bachelier [1], Sin embargo, no cobra importancia hasta que en 1965 el premio nobel en Economía P. A. Samuelson (1915-) propone el movimiento browniano geométrico o económico en el cual los logaritmos de los precios son los que siguen un movimiento browniano [9], para lo que se basa en la idea de trabajar con los logaritmos de los precios que aparece en 1959 en un artículo de M. F. M. Osborne[7]. En la misma década B. B. Mandelbrot (1924-) propuso como generalización del movimiento browniano los procesos estables, cuyos incrementos también son independientes pero donde las distribuciones que siguen estos incrementos son las distribuciones estables [3].

2.-Las distribuciones estables

Las distribuciones estables se caracterizan por el hecho de que la suma ponderada de dos variables aleatorias estables e independientes es, salvo centrado y escala, una variable aleatoria estable; propiedad que caracteriza a la distribución normal si añadimos el hecho de que tengan varianza finita. Excepto en casos muy concretos (distribuciones normal, de Cauchy y de Levy) no tienen una expresión para su función de densidad y se describen mediante su función característica [2, 8, 10]:

$$\varphi(t) = \exp\left(i\mu t - \gamma |t|^\alpha \left[1 + \beta \frac{t}{|t|} w(t, \alpha)\right]\right) \quad \text{con} \quad w(t, \alpha) = \begin{cases} \operatorname{tg}\left(\frac{\alpha\pi}{2}\right) & \alpha \neq 1 \\ \frac{2}{\pi} \ln |t| & \alpha = 1 \end{cases}$$

Los parámetros dan lugar a las distintas distribuciones estables, que notaremos por

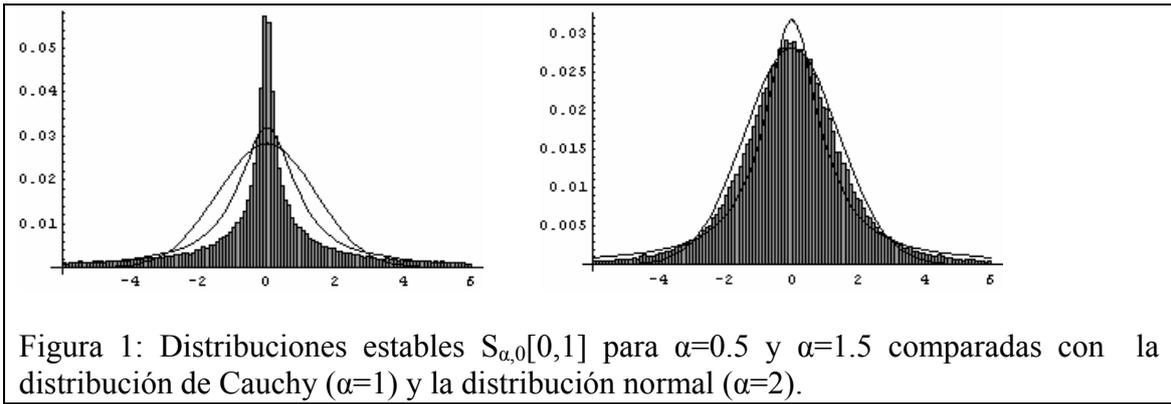
$S_{\alpha,\beta}[\mu,\gamma]$, y son:

- $\alpha \in (0,2]$ (exponente característico),
- $\beta \in [-1,1]$ (parámetro de asimetría),
- $\mu \in (-\infty, +\infty)$ (parámetro de localización)
- $\gamma \in [0, +\infty)$ (parámetro de escala).

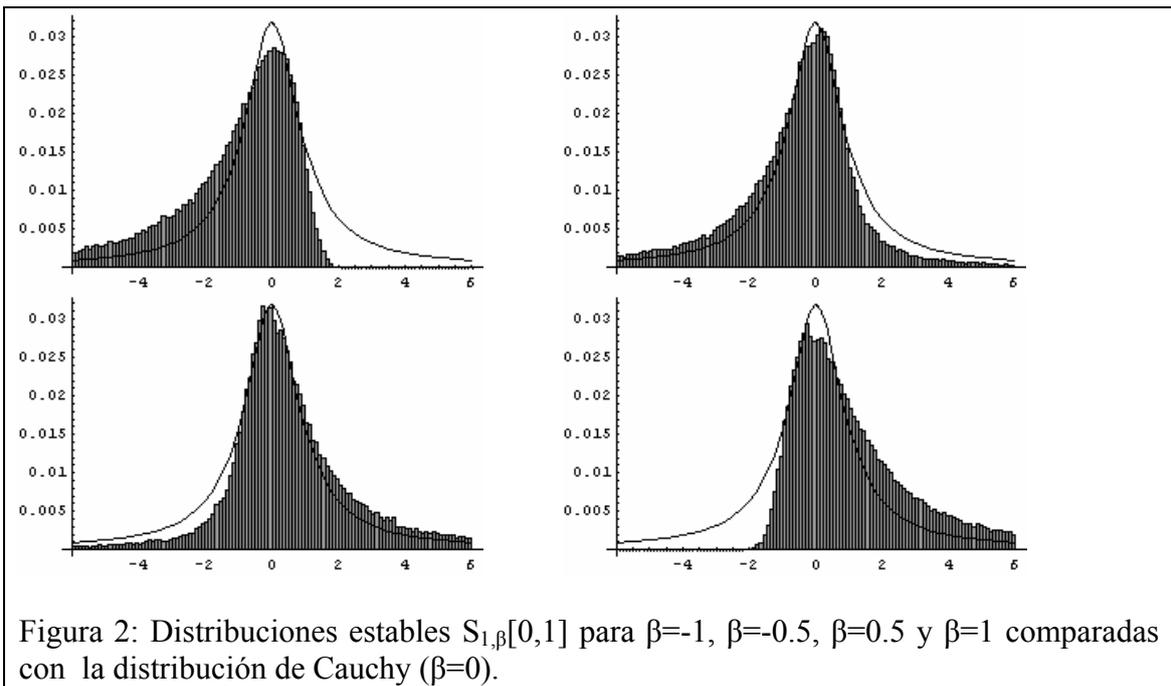
Para $\alpha=2$, la función característica resultante corresponde a una distribución normal con media μ y varianza 2γ (el valor de β es irrelevante, por lo que al ser simétrica se toma $\beta=0$). Para $\alpha=1$ y $\beta=0$ la función característica que se obtiene es la de una distribución de Cauchy con mediana μ y parámetro de escala γ . Ambas distribuciones son los casos límite para la finitud de la esperanza y de la varianza, ya que las distribuciones estables con esperanza finita están comprendidas entre la distribución normal, que tiene esperanza y varianza finita, y la distribución de Cauchy, cuya esperanza y varianza son infinitas.:

- si $\alpha \in (0,1]$ la distribución tiene media y varianza infinita,
- si $\alpha \in (1,2)$ la distribución tiene media finita pero varianza infinita,
- si $\alpha=2$ la distribución tiene tanto media como varianza finita.

El exponente característico determina el grado de apuntamiento de la distribución y el tamaño de las colas (figura 1):



El parámetro de asimetría, β , como su nombre indica, determina el grado de asimetría de la distribución. La distribución es simétrica para $\beta=0$ y para $\beta \neq 0$ es tanto más asimétrica cuanto mayor es el valor absoluto de β , estando sesgada a la derecha para $\beta > 0$ y a la izquierda para $\beta < 0$ (figura 2).



El parámetro de localización de la distribución, μ , tiene una interpretación clara cuando la media es finita, $\alpha > 1$, ya que en este caso es la media de la distribución y cuando la distribución es simétrica, $\beta=0$, ya que en este caso corresponde a la mediana. Sin embargo, cuando la media es infinita, $0 < \alpha \leq 1$, y no es simétrica, $\beta \neq 0$, μ no tiene una

interpretación tan clara y simplemente es un parámetro que describe la localización de la distribución (figura 3).

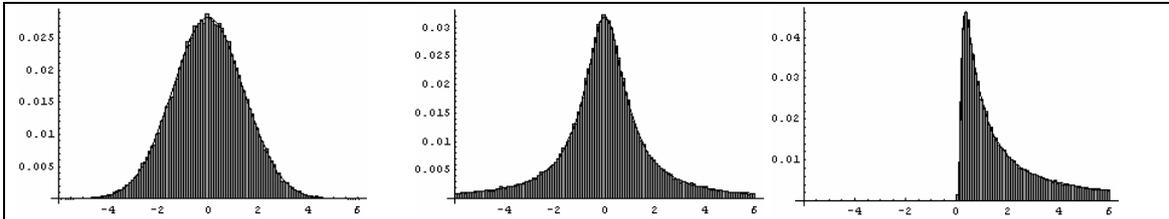


Figura 3: Las distribuciones estables que tienen función de distribución con parámetro de localización cero. El parámetro de localización de la distribución normal ($\alpha=2$) coincide con la media ya que tiene esperanza finita. El parámetro de localización de la distribución de Cauchy ($\alpha=1$ y $\beta=0$) coincide con la mediana ya que es simétrica. El parámetro de localización de la distribución de Levy ($\alpha=0.5$ y $\beta=1$), que tiene esperanza infinita y no es simétrica, no coincide ni con la media ni con la mediana.

El parámetro de escala de la distribución, γ , como su nombre indica, es un parámetro finito que define la escala de la distribución. En el único caso donde la varianza es finita, $\alpha=2$, corresponde a la mitad de la varianza.

3.-El movimiento browniano y los procesos estables

El movimiento browniano clásico se define como un proceso que comienza en el origen casi seguro, cuyos incrementos son independientes y están idénticamente distribuidos según una normal con media cero y varianza proporcional al incremento temporal, en el que el hecho de considerar la media como cero se debe a que en otro caso se descompone como la suma de una tendencia y un proceso con media cero [10]:

Definición 3.1: Un **movimiento browniano**, es un proceso estocástico, $\{B(t): t \geq 0\}$, que comienza en el origen casi seguro y verifica:

- El proceso tiene incrementos independientes.

- $\forall t \geq 0 \forall h \geq 0 \quad B(t+h)-B(t) \sim N(0, \sigma^2 h)$

Para $\sigma=1$ diremos que es un movimiento browniano estándar. ♣

Los incrementos del proceso son estacionarios, están idénticamente distribuidos y sólo dependen del incremento temporal. Además, al comenzar en el origen, el proceso coincide con su incremento respecto al origen y sigue una normal con media cero y varianza proporcional al incremento temporal:

$$B(t) \sim N(0, \sigma^2 t) \quad \forall t \geq 0.$$

Si generamos un ruido blanco gaussiano estándar siguiendo una distribución normal con media cero y varianza unitaria sus sumas acumuladas forman una aproximación al movimiento browniano (figura 4).

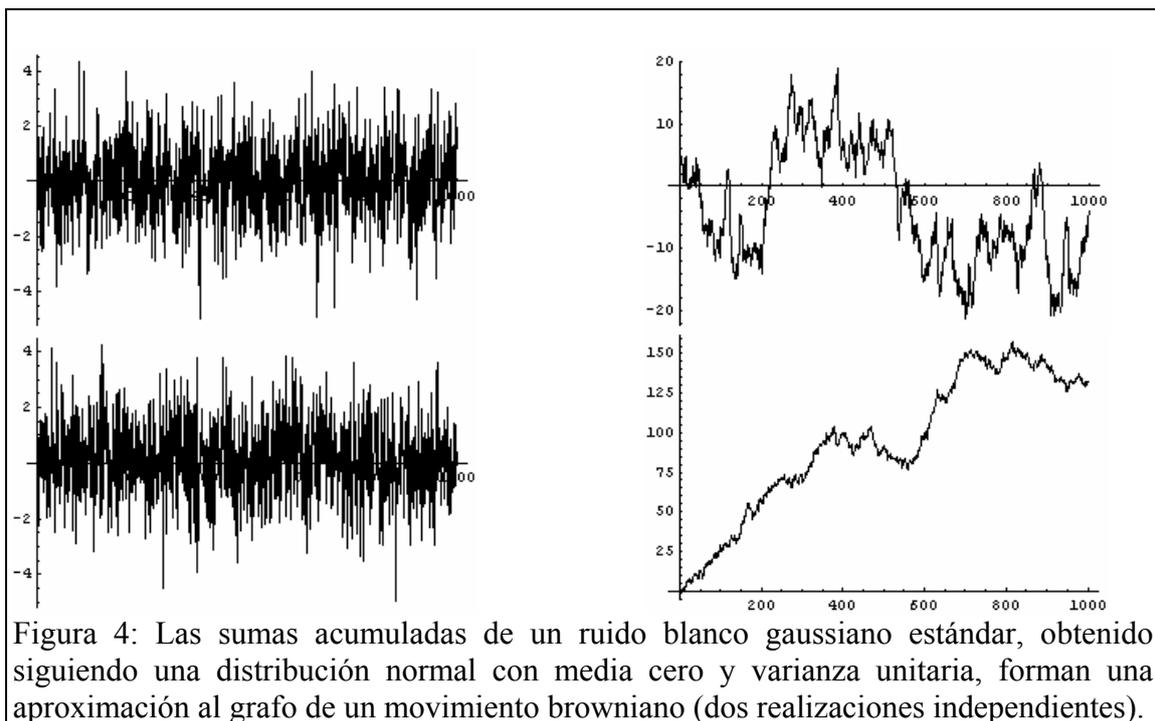


Figura 4: Las sumas acumuladas de un ruido blanco gaussiano estándar, obtenido siguiendo una distribución normal con media cero y varianza unitaria, forman una aproximación al grafo de un movimiento browniano (dos realizaciones independientes).

Los procesos estables se obtienen al considerar que sus incrementos siguen una distribución estable con exponente característico α , $0 < \alpha \leq 2$, en lugar de una distribución normal, que es estable con $\alpha=2$. Al igual que en el movimiento browniano se considera

que no tienen tendencia, por lo que el parámetro de localización, μ , que en la distribución normal es la media, se toma como cero [8, 10]:

Definición 3.2 Un proceso estable es un proceso estocástico $\{L_{\alpha,\beta}(t): t \geq 0\}$ que comienza en el origen (casi seguro) y verifica:

- El proceso tiene incrementos independientes.
- $\forall t \geq 0, \forall h \geq 0: L_{\alpha,\beta}(t+h) - L_{\alpha,\beta}(t) \sim S_{\alpha,\beta}(0, \gamma h)$. ♣

El parámetro de escala de la distribución, γ , va a establecer la proporcionalidad de los incrementos de un proceso estable con el incremento temporal, análogamente al movimiento browniano clásico único caso con varianza finita ($\alpha=2$) donde con la sustitución $\gamma=\sigma^2/2$ la distribución de los incrementos es normal con media cero y varianza $\sigma^2 h$. Como los incrementos del proceso están idénticamente distribuidos y sólo dependen del incremento temporal, los incrementos del proceso son estacionarios y, al comenzar en el origen, se tiene:

$$L_{\alpha,\beta}(t) \sim S_{\alpha,\beta}(0, \gamma t), \quad \forall t \geq 0.$$

Si en vez de un ruido blanco como en el movimiento browniano consideramos que los incrementos siguen distribuciones estables con parámetro de localización cero y parámetro de escala unitario obtenemos aproximaciones a distintos procesos estables que dependen fundamentalmente del exponente característico (figura 5).

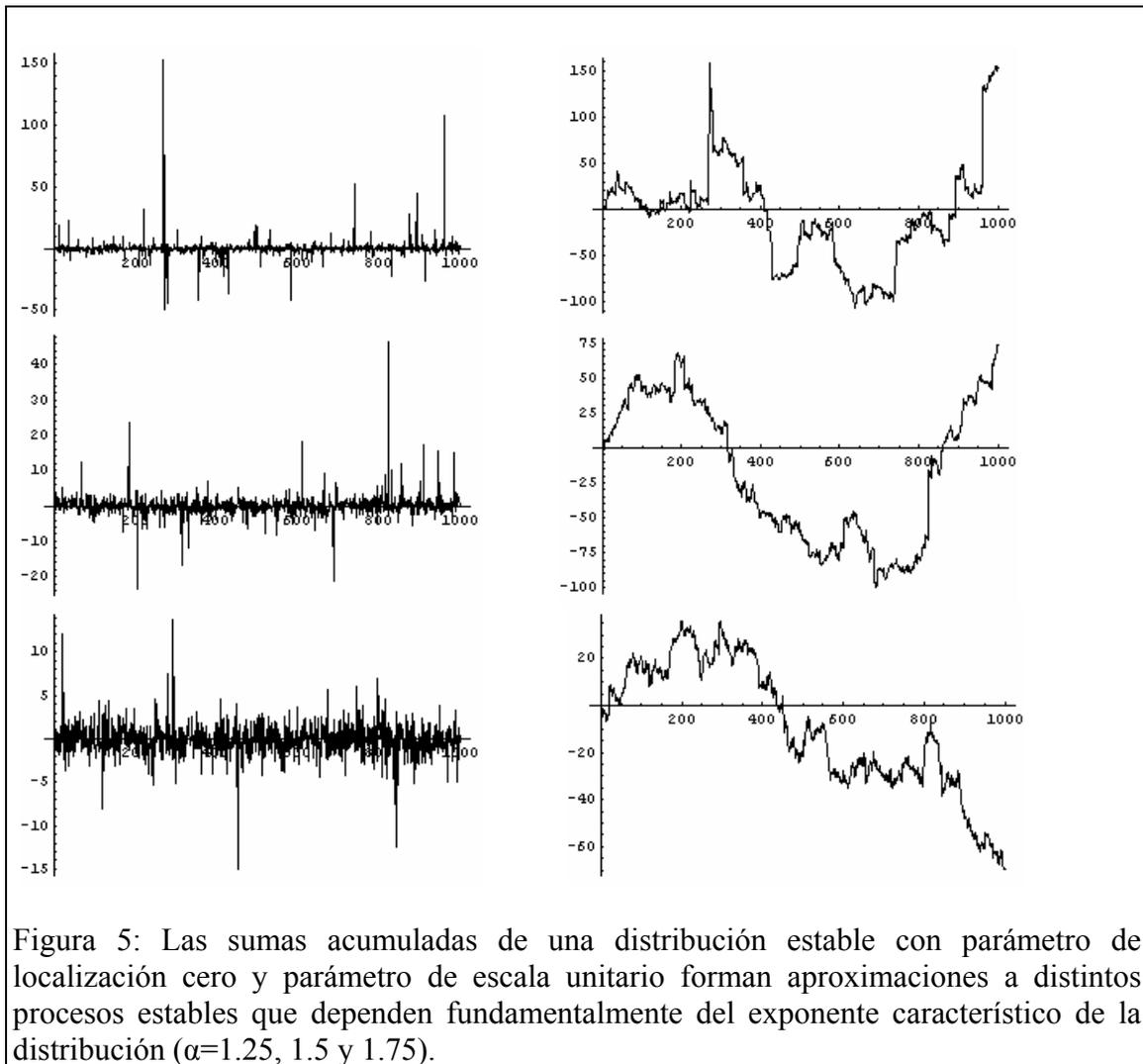


Figura 5: Las sumas acumuladas de una distribución estable con parámetro de localización cero y parámetro de escala unitario forman aproximaciones a distintos procesos estables que dependen fundamentalmente del exponente característico de la distribución ($\alpha=1.25, 1.5$ y 1.75).

4.-Conclusiones

Aunque el movimiento browniano es uno de los modelos más utilizados para describir el precio de una acción, en la mayoría de las acciones los datos empíricos no se ajustan bien al modelo, ya que la distribución empírica de los incrementos difiere de la normal, sobre todo en el tamaño de las colas. Por tanto, si en vez de utilizar el movimiento browniano como modelo de los precios de una acción utilizamos los procesos estables disponemos de dos nuevos parámetros para capturar la forma de la distribución de los incrementos del proceso: el exponente característico del proceso, que determina el grado

de apuntamiento de la distribución, y el parámetro de asimetría, que determina el grado de simetría de la distribución (figura 6).

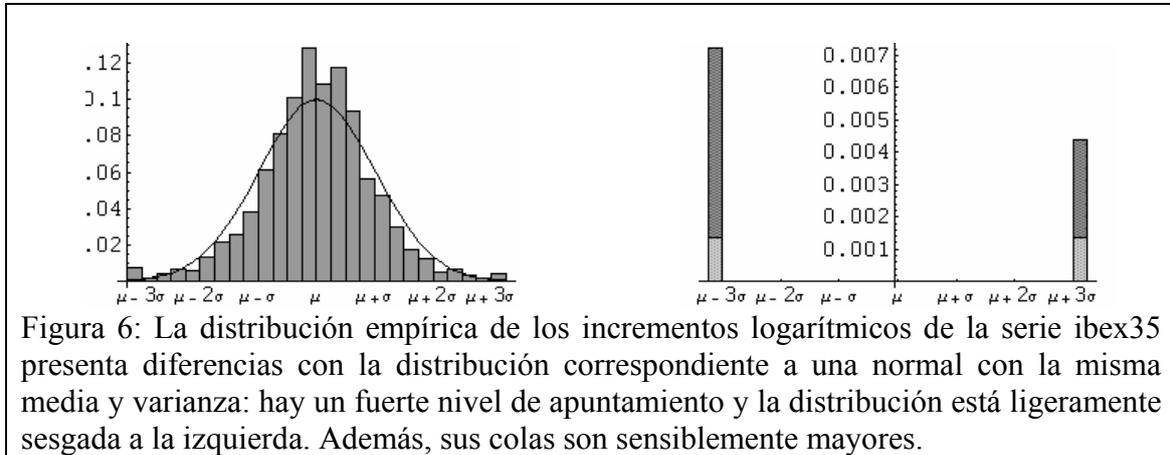


Figura 6: La distribución empírica de los incrementos logarítmicos de la serie ibex35 presenta diferencias con la distribución correspondiente a una normal con la misma media y varianza: hay un fuerte nivel de apuntamiento y la distribución está ligeramente sesgada a la izquierda. Además, sus colas son sensiblemente mayores.

El exponente característico del proceso se considera que varía entre $\alpha=1$ (distribución de Cauchy) y $\alpha=2$ (distribución normal), ya que en este caso la distribución tiene media finita y varianza infinita y, generalizando el caso del movimiento browniano, sería posible representar la serie como la suma de una tendencia y un proceso con media cero. Cuando se consideran los cierres correspondientes a distintos periodos, el comportamiento del proceso obtenido es compatible con el comportamiento lineal de la media muestral, al igual que en el movimiento browniano con tendencia, y además explica el comportamiento errático de la varianza muestral, que no es compatible con el movimiento browniano (figura 7).

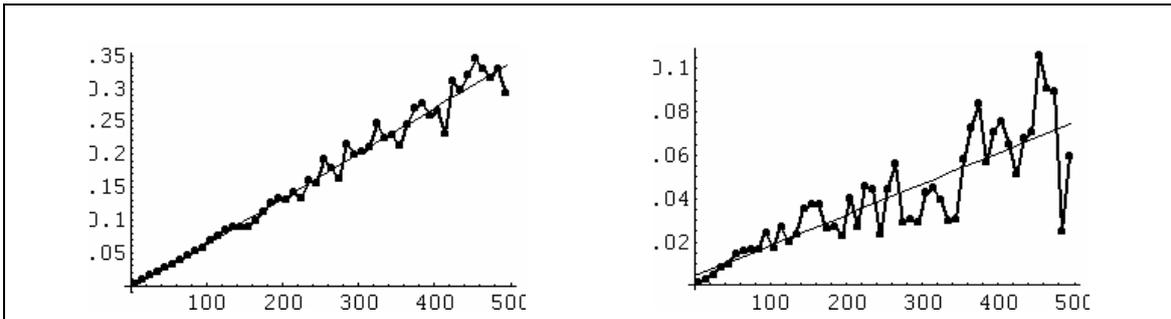


Figura 7: Si en la serie Ibx35 representamos la media muestral de las subseries que se obtienen al considerar distintos periodos de cierre se observa que estas medias muestran una tendencia lineal. Sin embargo, al representar la varianza muestral se observa que su comportamiento es errático, lo que no sucede en el movimiento browniano, ya que en este caso la varianza es proporcional al incremento temporal.

Esta generalización del movimiento browniano mantiene la independencia de los incrementos del proceso y sigue siendo adecuada para modelizar la apariencia errática e irregular que presentan las gráficas del precio de una acción y que se mantiene cuando se reduce la escala temporal en la que las representamos (esta característica común a las gráficas del precio de una acción, el movimiento browniano y los procesos estables recibe el nombre de autoafinidad y se analiza con detalle en [6]) (figura 8).

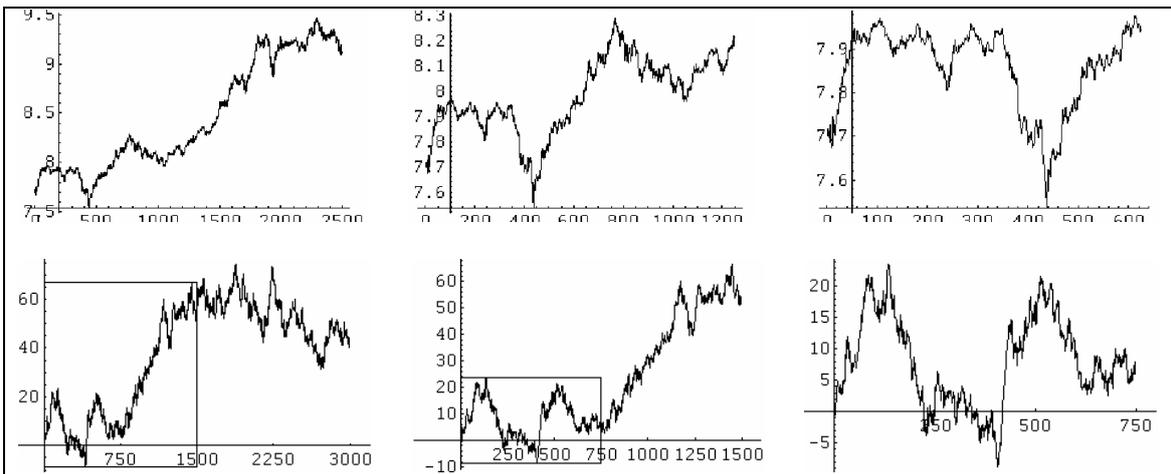


Figura 8: Cuando representamos la serie de los logaritmos de los cierres diarios del índice Ibx35 durante la década de los noventa en intervalos de tiempo con una duración cada vez menor los grafos siguen mostrando una apariencia errática e irregular (primera fila). Si simulamos un proceso estable y lo representamos en intervalos cada vez más pequeños los gráficos obtenidos mantienen el mismo aspecto y son iguales en sentido estadístico, ya que el proceso del que provienen es un proceso estable que sólo difiere del original en la escala vertical (segunda fila).

Bibliografía

1. Bachelier, L. Theory of speculation. En *The Random Character of Stock Market Prices*. Risk Books, 2000 (editor P. Cootner).
2. Feller, W. *Teoría de Probabilidades y sus Aplicaciones*, vol. II. Limusa, 1978.
3. Mandelbrot, B.B. The pareto-lévy law and the distribution of income. *Int. Economic Review*, 1:79--106, 1960.
4. Mandelbrot, B. B. *Fractals and Scaling in Finance*. Springer-Verlag, 1997.
5. Muñoz San Miguel, J. *La dimensión fractal en el mercado de capitales*. Tesis doctoral. Dpto. Economía Aplicada I, Universidad de Sevilla, 2002.
6. Muñoz San Miguel, J. Autoafinidad en series temporales. *Actas Asepelt XX*. Badajoz 2005.
7. Osborne, M. F. M. Brownian motion in the stock market. *Operations Research*, 7:145--173, 1959.
8. Samorodnitsky, G. y Taqqu, M.S. *Stable Non-Gaussian Random Processes*. Chapman and Hall, 1994
9. Samuelson, P.A.. Rational theory of warrant pricing. *Industrial Management Review*, 6:13--31, 1965.
10. Shiryaev, A. N. *Essentials of Stochastic Finance: Facts, Models, Theory*. World Scientific, 1999.