

**ANÁLISIS DEL CONSUMO SEMANAL DE
CIGARRILLOS EN ESPAÑA, UTILIZANDO LA
VEROSIMILITUD EMPÍRICA.***

M^ª del Pilar Fernández Sánchez**

e-mail: pilarfs@ugr.es

Agustín Hernández Bastida

e-mail: bastida@ugr.es

Carlos Sánchez González

e-mail: csanchez@ugr.es

Departamento de Métodos Cuantitativos para la Economía y la Empresa.

Universidad de Granada.

Resumen

En este trabajo se estudia el consumo semanal de cigarrillos, con datos procedentes de la Encuesta Continua de Presupuestos Familiares. Empleamos el método de la Verosimilitud empírica. El consumo se estudia en general y analizando su relación con la renta.

Palabras clave: Verosimilitud empírica, unidades físicas, tabaco.

** El orden de los autores es alfabético.

1. Introducción.

Son varios los objetivos que nos proponemos en este trabajo. Por un lado, estimar y comparar para los hogares españoles el consumo de tabaco negro y rubio. Por otro, analizar si ese consumo presenta diferencias al tener en cuenta el nivel de renta de los hogares.

Hemos utilizado datos procedentes de la Encuesta Continua de Presupuestos Familiares (ECPF), relativos al consumo semanal, medido en cajetillas de cigarrillos, de los dos tipos de tabaco considerados.

El hecho de utilizar como variable unidades físicas, y no el gasto expresado en unidades monetarias, supone una primera diferencia con la mayoría de los trabajos sobre consumo. La segunda diferencia radica en estudiar el consumo a partir de la función de verosimilitud empírica, y realizar la inferencia a partir de ella.

La técnica que utilizamos nos permite representar formalmente el consumo de los hogares españoles considerando únicamente esa variable. Este método se caracteriza por no usar variables explicativas y por no estimar función de regresión alguna. Eso significa que directamente no se pueden obtener conclusiones en base a variables explicativas, ni es posible conocer elasticidades de demanda.

La verosimilitud empírica consiste fundamentalmente en calcular el perfil de verosimilitud de una distribución general multinomial cuya masa de probabilidad se encuentra localizada en datos puntuales, y fue introducida para construir intervalos de confianza, aplicándose en numerosos contextos como modelos aleatorios, modelos autorregresivos, modelos de regresión,...etc. Sugiere un modelo para un conjunto de parámetros, en el que no son necesarias hipótesis de partida sobre la distribución de la población de la que procede la muestra, ni de la forma de la función de verosimilitud, ni de la correlación entre variables. La verosimilitud y sus regiones son una función sólo de los datos muestrales.

Hay numerosos trabajos tanto en un escenario no paramétrico, no bayesiano y en estudios asintóticos, (ver Hall (1986, 1987); DiCiccio et al (1989); Hall y La Scala (1990); Owen (1991); Chen (1993); Chen y Hall (1993); Zhang (1996); Monti (1997)),

como en un escenario de análisis bayesiano (ver McCray (1984, 1986)), que recogen la utilización de esta técnica.

Más concretamente asociada a problemas de naturaleza económica, destacamos trabajos en auditoría contable (ver Hernández et al (1997,1998)), valoración contingente de parques naturales, (ver Vázquez y León (1998)), consumo de diversos productos en unidades físicas, (ver Fernández et al (2002)), y en renta y ocultación (ver Fernández et al (2002, 2003)).

El buen comportamiento que ha demostrado en estos campos y las diferentes posibilidades de análisis que facilita, justifican su utilización en esta rama de la Economía.

El resto del trabajo se estructura de la siguiente forma: el apartado 2 expone los datos y la técnica utilizados; el 3 los resultados y las conclusiones obtenidas y en el 4 se recogen las referencias bibliográficas.

2.Datos y Técnica.

2.1.- Datos.

En este apartado comentamos los datos utilizados en este trabajo, señalando especialmente su origen y lo que significa. Los datos proceden de la Encuesta Continua de Presupuestos Familiares (ECPF).

El objetivo¹ fundamental de esta encuesta es proporcionar estimaciones acerca de los gastos de consumo y de los ingresos, trimestrales y anuales, para el conjunto nacional, según diversas variables de clasificación. Sin embargo, ofrece también valores del consumo, en cantidades físicas, de algunos artículos como alimentos, bebidas o tabaco. Es de esa información de donde se han extraído los datos utilizados en este trabajo.

La encuesta se centra en los hogares privados que residen en viviendas familiares principales. Define el hogar como la persona o conjunto de personas que

¹ Para una exposición detallada de la metodología ver “Encuesta Continua de Presupuestos Familiares. Metodología”. Instituto Nacional de Estadística. Madrid. 1992.

ocupan en común una vivienda familiar principal o parte de ella, y consumen y comparten alimentos y otros bienes con cargo a un mismo presupuesto.

La muestra la constituyen hogares de todo el territorio nacional español, incluyendo Ceuta y Melilla. El ámbito temporal de estudio es cada trimestre del año.

Para alimentos, bebidas y tabaco, el período muestral, tiempo durante el cual colabora realmente el hogar seleccionado, es de una semana.

Conviene precisar que la ECPF utiliza el criterio del pago para medir el consumo de un bien o servicio. Así, recogerá, por tanto, los pagos efectuados durante la semana, independientemente de que se hayan consumido realmente los bienes y servicios adquiridos. En este sentido, la encuesta responde mejor a cantidad “comprada”, que a cantidad “consumida”.

El tamaño de la muestra es de 3.114 hogares y corresponde al primer trimestre de 1997.

Son muchos los autores (ver entre otros Blundell (1988); Blundell et al (1993); López (1995); Ramajo (1995, 1996)) que han señalado un problema existente en este tipo de encuestas: la presencia generalizada de ceros en las respuestas de los hogares, llegando en algunas ocasiones a representar más del 50% de las observaciones muestrales. Esto crea problemas de inconsistencia en la estimación econométrica, y ha impulsado el estudio de diversos modelos para analizar la demanda en función de las posibles causas de estos ceros (ver Deaton e Irish (1984), Atkinson et al (1989), Labeaga (1991), García y Labeaga (1991), Jones (1989), Pudney (1989), y Meghir y Robin (1992)).

El método que describimos más adelante no plantea conocer el origen de esos ceros, utilizando los datos tal y como han sido suministrados, sin incurrir en tratamientos diferenciados en función de sus posibles causas. Eso sí, se han depurado los datos eliminando algunos valores de la muestra que presentaban magnitudes irreales, y suprimiendo de la misma aquellas familias que declaraban no tener gasto en ningún artículo en ese trimestre.

2.2. Técnica.

Owen (1.988; 1.990) introdujo la noción de verosimilitud empírica realizando en sus trabajos una aproximación intuitiva que exponemos a continuación:

Sean x_1, \dots, x_n observaciones independientes de una función de distribución F_0 .

La función de distribución empírica (ver Kiefer y Wolfowitz (1956), Kaplan y Meier (1958), Bailey (1984) o Vardi (1985)), definida como,

$$F_n = \frac{1}{n} \sum_{i=1}^n \mathbf{d}_{x_i} \quad [1]$$

es un estimador de máxima verosimilitud no paramétrico de F_0 basado en x_1, \dots, x_n , ya que maximiza la función de verosimilitud

$$L(F) = \prod_{i=1}^n F\{x_i\}. \quad [2]$$

En esta expresión, $F\{x_i\}$ es la probabilidad del conjunto $\{x_i\}$ bajo F , x_i es el valor observado de X_i y F es cualquier medida de probabilidad.

Si en esa función de verosimilitud se tienen en cuenta restricciones obtenemos la verosimilitud empírica.

Consideremos una situación en la que pueden establecerse un número determinado de categorías T_1, T_2, \dots, T_k ; con probabilidades asociadas p_1, \dots, p_k desconocidas, y verificando,

$$p_i \in [0, 1], i = 1, \dots, k; \sum_{i=1}^k p_i = 1. \quad [3]$$

Notaremos $X_i, i = 1, \dots, k$, el representante de la categoría i -ésima y $\vec{p} = (p_1, \dots, p_k)$, el vector de probabilidades en el que cada uno de los p_i representa la probabilidad de que una realización concreta pertenezca a la i -ésima categoría.

² Aquí \mathbf{d}_x denota la distribución degenerada en x , es decir, en un punto de masa en x .

Nuestro interés está en una magnitud, \mathbf{q} , que es una combinación lineal de los parámetros del problema, escribiéndola de la siguiente forma:

$$\mathbf{q} = \sum_{i=1}^k X_i p_i . \quad [4]$$

Notaremos por $L(\mathbf{q})$ a la verosimilitud empírica, obteniéndola como el máximo en \bar{p} del producto de los p_i sujeto a las restricciones que se indican:

$$L(\mathbf{q}) = \underset{\bar{p}}{\text{Max}} \prod_{i=1}^k p_i, \text{ sujeto a } \left[\begin{array}{l} p_i \in [0,1], i = 1, \dots, k; \\ \sum_{i=1}^k p_i = 1 \\ \sum_{i=1}^k X_i p_i = \mathbf{q} \end{array} \right] \quad [5]$$

Para cada valor de \mathbf{q} obtenemos como resultado de este proceso de maximización el correspondiente valor $L(\mathbf{q})$, y así se construye la función de verosimilitud empírica, (ver Fernández-Sánchez (2001), para más detalles).

Será interesante para nosotros la noción de intervalo de verosimilitud. Este concepto se relaciona directamente con el de verosimilitud empírica. Entendemos por intervalos de verosimilitud un conjunto de valores del espacio paramétrico en los podemos decir que la función de verosimilitud empírica toma un valor grande. El concepto de grande se establece en relación a los valores que toma el máximo de la función. Dichos intervalos nos permitirán establecer comparaciones en relación a este máximo y entre valores tanto de dentro como de fuera del intervalo.

Se define el intervalo de verosimilitud para una fuerza de evidencia k de la siguiente forma:

$$\begin{aligned}
(v_1, v_2) / \forall v \in (v_1, v_2) \quad \frac{L(a)}{L(v)} < k; \\
\forall v \notin (v_1, v_2) \quad \frac{L(a)}{L(v)} > k
\end{aligned}
\tag{6}$$

En esta expresión, “ a ” es la abcisa del valor máximo de la función de verosimilitud empírica, y como valor de k , en la literatura se utilizan generalmente los valores de 8 y 32, identificándolos con una fuerza de evidencia fuerte y muy fuerte respectivamente. Para más detalles en relación a los intervalos de verosimilitud puede consultarse Royall (1997).

Consideremos \mathbf{q} , como variable aleatoria, la magnitud sobre la que deseamos realizar inferencia y para ella es posible especificar una distribución a priori $\mathbf{x}(\mathbf{q})$. Combinando esta distribución a priori con la función de verosimilitud empírica se obtiene una distribución a posteriori:

$$\mathbf{x}(\mathbf{q} / \text{datos}) = \frac{L(\mathbf{q})\mathbf{x}(\mathbf{q})}{\int_{\Theta} L(\mathbf{q}) d\mathbf{x}(\mathbf{q})} .
\tag{7}$$

Si utilizamos como distribución a priori $\mathbf{x}(\mathbf{q})$ una distribución Uniforme, se consigue una normalización de esa verosimilitud. Esto permite disponer de una distribución de probabilidad para el parámetro \mathbf{q} , que es el objetivo que se persigue.

El método señalado tiene una serie de ventajas comparándolo con los métodos clásicos de estimación multiparamétrica, ventajas que resultan de sugerir un modelo en el que no se requieren hipótesis de partida sobre la distribución de la población de la que procede la muestra, ni de la forma de la función de verosimilitud, ni de la correlación entre variables.

3.Resultados y conclusiones.

Un aspecto a señalar del método que hemos descrito, es la facilidad para obtener muchos e importantes resultados de inferencia para el valor medio de la variable que se considera. En este caso, esa variable es el consumo medio semanal de tabaco de los hogares españoles, medido en cajetillas de cigarrillos.

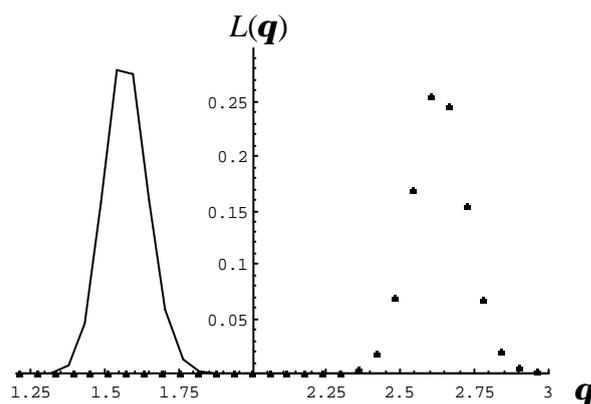
En este apartado recogemos los resultados y las conclusiones obtenidas al utilizar una muestra de 3.114 hogares correspondiente al primer trimestre de 1997.

3.1.Consumo medio de tabaco.

El primer resultado que recogemos es una gráfica que contiene las funciones de verosimilitud empírica para el consumo medio de los dos tipos de cigarrillos considerados.

Exponemos conjuntamente las dos curvas para facilitar su comparación. La gráfica de la izquierda, en trazo continuo, es la obtenida para el consumo medio de tabaco negro, mientras que la de puntos discontinuos corresponde al tabaco rubio.

Gráfico 1: Funciones de verosimilitud empírica para el consumo semanal de cajetillas de tabaco negro y rubio.



Fuente: Elaboración propia.

Aunque a primera vista parezcan funciones de densidad, no lo son. En este caso, el eje de abscisas recoge los valores del parámetro, y el de ordenadas, los valores de la función de verosimilitud empírica. Para cada q , resolviendo el problema de

maximización planteado anteriormente, se obtiene su $L(\mathbf{q})$ correspondiente. Así, punto a punto construimos las funciones. Aquí se ha hecho para 1.000 puntos.

Las gráficas muestran valores del consumo medio para los que la función de verosimilitud no es cero, y ahí se definen los intervalos de verosimilitud.

Ya hemos comentado que la función de verosimilitud no es una función de densidad, pero si se utiliza una distribución a priori Uniforme, junto con la verosimilitud empírica calculada, obtendremos una distribución a posteriori. Esta distribución a posteriori supone una normalización de la verosimilitud y refleja el comportamiento en probabilidad de la variable considerada.

Toda la inferencia se realiza a partir de esa distribución a posteriori, obteniendo estimaciones puntuales y por intervalo para ese consumo medio.

Para la estimación puntual se utiliza la moda de la distribución a posteriori, que es un estimador de máxima verosimilitud. Nos fijamos para ello en aquel valor que presenta el más grande para $L(\mathbf{q})$. En este caso, el consumo medio estimado es de 1,54 cajetillas por semana para el tabaco negro, y 2,6 para el tabaco rubio. Estos son los consumos semanales más probables de los hogares españoles.

La estimación por intervalo la realizamos calculando previamente los percentiles de la distribución a posteriori. Con ellos determinamos algunas regiones de probabilidad. Así, se establecen los valores entre los que existe una probabilidad, de la cuantía deseada, de que se encuentre el verdadero valor del consumo medio de tabaco.

El siguiente cuadro resume los principales resultados que hemos obtenido:

Cuadro 1: Resultados de inferencia para el consumo medio semanal de tabaco, en cajetillas.					
Tabaco	Estimador puntual	R. confianza (90%)	R. confianza (95%)	Int verosimilitud (8)	Int verosimilitud (32)
Negro	1,54	(1,43 ; 1,62)	(1,4 ; 1,7)	(1,43 ; 1,70)	(1,43 ; 1,7)
Rubio	2,6	(2,45 ; 2,78)	(2,42 ; 2,78)	(2,48 ; 2,7)	(2,42 ; 2,8)

Fuente: Elaboración propia.

Observando la tabla podemos afirmar que el consumo semanal estimado de tabaco rubio está claramente por encima del de tabaco negro. Comparando las estimaciones puntuales, hay un 68% más de consumo de tabaco rubio. Esto confirma la idea inicial que nos habíamos creado al observar las funciones de verosimilitud

empírica, ya que la gráfica de tabaco rubio está bastante a la derecha respecto a la de negro.

Las regiones de confianza determinan los valores entre los que se encuentra el valor del consumo medio con una probabilidad del 90 o del 95%. Podemos afirmar que hay una probabilidad del 90% de que el consumo medio de tabaco negro esté entre 1,43 y 1,62 cajetillas por semana, o que el número medio de cajetillas de tabaco rubio está entre 2,42 y 2,78, con una probabilidad del 95%. Al igual que ocurría con la gráfica, los intervalos para el consumo de rubio están sensiblemente a la derecha de los correspondientes para negro.

Los intervalos de verosimilitud son un rango de valores verosímiles, y además pueden reinterpretarse como conjunto de valores con gran probabilidad. ¿Cuál sería el nivel de confianza que correspondería a esos intervalos de verosimilitud?

Para el tabaco negro, el intervalo para una fuerza de evidencia 8, (1,43 ; 1,70) corresponde a un nivel de confianza del 92,5%, aproximadamente. El intervalo de verosimilitud obtenido para una fuerza de evidencia 32, (1,43 ; 1,72) se identifica con un intervalo del 94%.

Para el tabaco rubio, los intervalos de verosimilitud correspondientes (2,48 ; 2,78) y (2,42 ; 2,84) son equiparables a intervalos de confianza del 87,5 y 96,5% respectivamente.

3.2 Consumo medio de tabaco por niveles de ingresos.

El segundo de los objetivos que nos planteamos en este trabajo, es analizar si el comportamiento en el consumo de tabaco se modifica o no al considerar la variable renta. Nos preguntamos si al considerar la población por niveles de renta, variará o no el esquema de consumo.

Para ello, usando la variable ingreso trimestral del hogar, proporcionada por la encuesta, establecemos 10 grupos de renta a partir de los deciles de esa variable. Para los hogares incluidos en esos grupos identificamos los consumos semanales de tabaco, medidos en unidades físicas (cajetillas).

En cada grupo de rentas consideramos las familias cuyos ingresos trimestrales se encuentran entre las siguientes cantidades, obtenidas con los deciles de la variable ingreso.

Cuadro 2: Grupos de rentas según deciles de ingresos			
Grupos de rentas	Euros /trimestre	Grupos de rentas	Euros /trimestre
g1	240,40 – 1.816,31	g6	4.042,83 – 4.681,88
g2	1.816,31 – 2.468	g7	4.681,88 – 5.392,28
g3	2.468 – 3.003,63	g8	5.392,28 – 6.340,68
g4	3.003,63 – 3.515,92	g9	6.340,68 – 8.093,62
g5	3.515,92 – 4.042,83	g10	8.088,40 – 61.315,25

Fuente: Elaboración propia.

Reproducimos a continuación los resultados obtenidos por el método al aplicarlo a distintos grupos de consumo, para el caso del tabaco negro. Las comparaciones se llevan a cabo usando la moda de la distribución a posteriori, como estimador puntual.

Cuadro 3: Principales resultados de inferencia para consumo medio de tabaco negro por niveles de renta.					
Grupos de renta	Estimador puntual	R. confianza (90%)	R. confianza (95%)	Int verosimilitud (8)	Int verosimilitud (32)
g1	1,15	(0,9 ; 1,47)	(0,85 ; 1,54)	(0,85 ; 1,53)	(0,76 ; 1,66)
g2	1,66	(1,3 ; 2,1)	(1,24 ; 2,2)	(1,23 ; 2,19)	(1,14 ; 2,37)
g3	1,4	(1,08 ; 1,89)	(1,02 ; 2)	(1,01 ; 1,94)	(0,92 ; 2,2)
g4	1,39	(1,09 ; 1,75)	(1,64 ; 1,82)	(1,03 ; 1,83)	(0,94 ; 1,96)
g5	2,06	(1,67 ; 2,55)	(1,6 ; 2,65)	(1,59 ; 1,69)	(1,46 ; 2,83)
g6	1,69	(1,35 ; 2,11)	(1,29 ; 2,2)	(1,29 ; 2,19)	(1,19 ; 2,38)
g7	1,81	(1,43 ; 2,36)	(1,36 ; 2,48)	(1,34 ; 2,47)	(1,23 ; 2,69)
g8	1,63	(1,27 ; 2,15)	(1,21 ; 2,27)	(1,24 ; 2,23)	(1,14 ; 2,47)
g9	1,32	(1,02 ; 1,91)	(0,96 ; 2,05)	(0,99 ; 1,98)	(0,88 ; 2,25)
g10	1,62	(1,28 ; 2,09)	(1,21 ; 2,18)	(1,21 ; 2,18)	(1,09 ; 2,25)

Fuente: Elaboración propia.

Así podemos establecer las siguientes conclusiones:

.- El mayor consumo semanal corresponde al grupo de renta 5, con un poco más de 2 cajetillas por semana, y supone un 55% más que el mínimo consumo. Éste último tiene lugar en el grupo de rentas más bajas, siendo la estimación de 1,15 cajetillas por semana.

.- Al pasar de un grupo de rentas al siguiente encontramos que:

- Uno de los mayores incrementos en el consumo tiene lugar al pasar del grupo de rentas más bajas al siguiente nivel de rentas, produciéndose un aumento del 44%.
- Entre el grupo 2 y el 3 el consumo decrece, con una reducción de más del 15%.
- Al pasar del tramo de renta 3 al 4, hay una disminución mínima (en torno al 0,7%) , con lo que podemos decir que el consumo se mantiene prácticamente constante.
- En el cambio del grupo 4 al 5 de rentas es cuando ocurre el mayor aumento en el consumo, alrededor del 48%.
- A continuación el consumo se reduce en un 17,96%, entre el grupo 5 y 6, aumentando un 7% entre el 6 y el siguiente.
- Decrece el consumo en los grupos 7 al 9, primero casi un 10% y luego un 19%.
- Hay una recuperación entre el penúltimo grupo de rentas y el siguiente, experimentando una subida el consumo cercana al 23%

.- Aunque el consumo correspondiente a las rentas más altas es un 40,86% más que el de las rentas más bajas, no es posible vislumbrar un patrón de comportamiento en el consumo a medida que aumenta la renta. Es más, hay cuatro ocasiones en que aumenta el consumo al aumentar la renta y cinco en que disminuye. Acudiendo a los conceptos de teoría económica hay ocasiones en que comportaría como un bien normal, pero en otras sería un bien inferior.

Además, comparando el incremento que se produce en la renta al pasar de un tramo al siguiente (medido por los puntos medios de los intervalos de renta), con el incremento producido en el consumo al variar de tramo, hay un único caso en el que el incremento observado en el consumo (48%) supera al producido en la renta (15,93%). Esta situación se produce entre el grupo 4 y el 5.

Podemos establecer que en este caso, la renta no aparece como un factor determinante en el consumo. Esto sugiere considerar otras variables como podrían ser la edad, el sexo, el tipo de trabajo del consumidor....

Veamos si esta situación es similar al considerar el consumo de tabaco rubio. Los resultados de inferencia se recogen en la tabla siguiente:

Cuadro 4: Principales resultados de inferencia para consumo medio de tabaco rubio por niveles de renta.					
Grupos de renta	Estimador puntual	R. confianza (90%)	R. confianza (95%)	Int verosimilitud (8)	Int verosimilitud (32)
g1	1,34	(1,05 ; 1,77)	(0,99 ; 1,87)	(0,98 ; 1,85)	(0,89 ; 2,01)
g2	1,51	(1,19 ; 1,93)	(1,14 ; 2,02)	(1,14 ; 1,99)	(1,02 ; 2,16)
g3	2,28	(1,89 ; 2,75)	(1,82 ; 2,84)	(1,81 ; 2,83)	(1,69 ; 3,02)
g4	2,74	(2,3 ; 3,23)	(2,22 ; 3,33)	(2,21 ; 3,33)	(2,08 ; 3,53)
g5	2,24	(1,87 ; 2,73)	(1,8 ; 2,82)	(1,8 ; 2,8)	(1,68 ; 3)
g6	2,3	(1,93 ; 2,76)	(1,86 ; 2,86)	(1,84 ; 2,85)	(1,72 ; 3,02)
g7	3,3	(2,85 ; 3,84)	(2,76 ; 3,95)	(2,77 ; 3,96)	(2,61 ; 4,16)
g8	3,3	(2,8 ; 3,95)	(2,71 ; 4,08)	(2,7 ; 4,08)	(2,53 ; 4,33)
g9	3,52	(3,01 ; 4,1)	(2,92 ; 4,22)	(2,9 ; 4,22)	(2,77 ; 4,44)
g10	3,87	(3,29 ; 4,61)	(3,19 ; 4,77)	(3,21 ; 4,72)	(3,02 ; 5,08)

Fuente: Elaboración propia.

.- Claramente, el consumo aumenta con el nivel de rentas. Al pasar de un nivel al siguiente aumenta el consumo, con excepción del grupo 5 que disminuye en un 18,24% en relación al consumo anterior.

.- El máximo consumo se alcanza en el grupo de rentas más altas, y su valor casi triplica al del grupo de rentas menores, que presenta el consumo más bajo de todos.

.- El mayor incremento en el consumo tiene lugar en el cambio del segundo grupo de rentas al siguiente, con un consumo 1,5 veces mayor para este último nivel.

- .- No hay prácticamente diferencias entre el tramo 7 y 8 de rentas.

- .- Podríamos distinguir tres niveles de consumo:
 - Consumo en torno a una cajetilla y media por semana, para los dos tramos de renta más bajas;

 - Más de dos cajetillas por semana para los niveles intermedios de rentas, grupos 3 al 6;

 - Consumo de más de tres cajetillas por semana, correspondientes a los niveles más altos de renta, grupos 7 al 10.

Si comparamos de nuevo el incremento medido en la renta con el incremento en el consumo, observamos que los grupos 2, 3 y 6 presentan un incremento en el consumo mayor que el incremento que experimentan en la renta.

En esta situación parece que la renta se muestra como un factor más determinante del consumo que en el caso anterior.

Por último, comparemos el consumo de tabaco para el mismo nivel de renta. En todos los tramos, salvo para el segundo grupo de rentas más bajas, el consumo de tabaco rubio está muy por encima del de tabaco negro. Estos porcentajes oscilan entre el casi 9% de más del grupo 5, a las dos veces y media de más del grupo 9.

A modo de conclusión final podríamos decir que el consumo medio en los hogares españoles de tabaco rubio, medido en cajetillas de cigarrillos, está por encima del consumo de tabaco negro. Esta situación se produce tanto al comparar niveles globales de consumo como al diferenciarlos por niveles de ingresos de esos hogares.

Además, si bien la renta es un factor importante en el consumo, a la hora de caracterizar el comportamiento de los hogares, sólo en el caso del tabaco rubio podemos considerarla un factor determinante. Esto nos lleva a pensar que otras variables, como edad, sexo, tipo de trabajo, o composición del hogar,..., tengan más relevancia en el consumo de dicho artículo.

4. Bibliografía.

1. Atkinson, A.B.; Gomulka, J. y Stern, N. (1989). "Spending on alcohol: Evidence form the Family Expenditure Survey 1970-1983". *Discussion paper of the London School of Economics*.
2. Bayley K.R., (1984). "Asymptotic equivalence between the cox estimator and the general ML estimators of regression and survival parameters in the cox model." *Ann. Stat.*, **12**, pp 730-736.
3. Blundell R., (1988). "Consumer behaviour: Theory and empirical evidence. A survey." *Economic Journal*, **98**, pp 16-65.
4. Blundell R., Pashardes P., Weber G., (1993). "What do we learn about consumer demand patterns from micro-data?." *American Economic Review* ,**83**, pp 570-597.
5. Chen S.X., (1993). "On the accuracy of empirical likelihood confidence regions for linear regression model." *Ann. Inst. Statist. Math.*, **45** (4), pp 621-637.
6. Chen S.X., Hall P., (1993). "Smoothed empirical likelihood confidence intervals for quantiles." *Ann. Stat.*, **21** (3), pp 1.166-1.181.
7. Deaton A., Irish M., (1984). "Statistical models for zero expenditures in household surveys." *Journal of Publics Economics*, **23**, pp 59-80.
8. Diccicio O T.J., Hall P., Romano J.P., (1989). "Comparison of parametric and empirical likelihood functions." *Biometrika*, **76**, pp 465-476.
9. Fernández J., (1998). "Un análisis de la oferta y la demanda en el mercado de girasol español". *Revista Española de Economía Agraria*, **183**, pp 153-172.
10. Fernández-Sánchez M. P., (2001). "*Utilización de la verosimilitud empírica en algunos problemas de interés económico*". Tesis doctoral. Universidad de Granada.
11. Fernández-Sánchez M. P., Hernández-Bastida A., Sánchez-González C., (2002). "Análisis de los ingresos y gastos trimestrales de los hogares españoles usando la Verosimilitud empírica". *Sometido a Estudios de Economía Aplicada*.

12. Fernández-Sánchez M. P., Hernández-Bastida A., Sánchez-González C., (2002).” Comparación del consumo en unidades físicas de tres tipos de carne por niveles de ingresos, utilizando la Verosimilitud empírica”. *Sometido a Revista Asturiana de Economía*.
13. Fernández-Sánchez M. P., Hernández-Bastida A., Sánchez-González C., (2003).” Una aproximación alternativa a la estructura de gastos de los hogares españoles: utilización de la Verosimilitud empírica”. *Sometido a Revista de Economía Aplicada*.
14. García J., Labeaga J.M., (1991). “A double hurdle dependent model for the demand for tobacco in Spain.” Mimeo.
15. Hall P., (1986). “On the bootstrap and confidence intervals.” *Ann. Stat.*, **14**, pp 1431- 1452.
16. Hall P., (1987). “On the bootstrap and likelihood based confidence regions.” *Biometrika*, **74**, pp 481-494.
17. Hall P., La Scala B., (1990). “Methodology and algorithms of the empirical likelihood.” *International Statistical Review*, **58** (2), pp 109-127.
18. Hernández-Bastida A., Vázquez-Polo J.M., (1997). “A note on the quasi-bayesian audit risk model for dollar-unit sampling.” *The European Accounting Review*, **6** (3), pp 501-507.
19. Hernández-Bastida A., Martel-Escobar M.C., Vázquez-Polo J.M. (1998). “On maximum entropy priors and a most likely likelihood in auditing.” *Questro*, **22** (2), pp 231-242.
20. Instituto Nacional de Estadística. (1992). ‘*Encuesta Continua de Presupuestos Familiares. Metodología*’, Madrid.
21. Jones A., (1989). “A double hurdle model of cigarette consumption.” *Journal of Applied Econometrics*, **4**, pp 23-39.
22. Kaplan E.L., y Meier R P., (1958). “Non parametric estimation from incomplete observations.” *J.A.S.A.*, **53**, pp 457-481.

23. Kiefer J., Wolfowitz J., (1956). "Consistency of the maximum likelihood estimator in the presence of infinitely many incidental parameters." *Ann. Math. Stat.*, **27**, pp 887-906.
24. Labeaga J.M., (1991). "A dynamic panel data model with limited dependent variables: an application to the demand for tobacco." *Documento de trabajo* nº 9.201, UNED.
25. López-Nicolás A., (1995). "Algunas consideraciones sobre el uso de datos microeconómicos en el análisis de la demanda. El caso de la Encuesta Continua de Presupuestos Familiares." *Hacienda Pública Española* **132**, pp 133-141.
26. Mc.Cray J.H., (1984). "A quasi-bayesian audit risk model for dollar-unit sampling." *The Accounting Review*, **59**, pp 35-51.
27. McCray J.H., (1986). "A general bayesian risk model for dollar-unit sampling and multiple populations." *Manuscrito no publicado*. The College of William and Mary. School of Business Administration, Virginia. U.S.A.
28. Meguir C., Robin J. M., (1992). "Frequency of purchase and the estimation of demand systems." *Journal of Econometrics*, **53**, pp 53-85.
29. Monti (1997). "Empirical likelihood confidence regions in time series models". *Biometrika*, **84** (2), pp 395-405.
30. Owen A.B., (1988). "Empirical likelihood ratio confidence intervals for a single functional." *Biometrika*, **75**, pp 237-249.
31. Owen A.B., (1990). "Empirical likelihood ratio confidence regions." *Ann. Stat.*, **18**, (1), pp 90-120.
32. Owen A.B., (1991). "Empirical likelihood for linear models." *Ann. Stat.*, **1.894**, pp 1.725-1.747.
33. Pudney S., (1989). · *Modelling individual choice. The econometrics of corners, kinks and holes*". Basil Blackwell, Oxford.

- 34 Ramajo J., (1995). "Análisis de los determinantes de la demanda en Extremadura". *Actas IX Reunión Asepelt-España*, Santiago de Compostela. Tomo 1, 397-407.
- 35 Ramajo J., (1996). "La demanda de productos alimenticios, bebidas y tabaco en Extremadura: un enfoque de regresiones censuradas". *Investigación Agraria: Economía*, **11** (3), pp 469-498.
36. Royall R.M., (1997). "*Statistical Evidence: A likelihood paradigm*". Ed. Chapman y Hall.
37. Vázquez F.J., León C., (1998). "Modelización del aprendizaje en valoración contingente." *Investigaciones Económicas*, **24**, pp 117-138.
38. Vardy Y. (1985). "Empirical distributions in selection bias models". *Ann. Stat.*, **13**, pp 178-203.
39. Zhang (1996). "Confidence intervals for a distribution function in the presence of auxiliary information." *Computational Statistics and Data Analysis*, **21**, pp 327-342.