

# SEGMENTACIÓN DEL MERCADO DE TURISMO RURAL EN EL NOROESTE DE MURCIA<sup>1</sup>

**Lourdes Molera Peris**

Departamento de Métodos Cuantitativos para la Economía  
Universidad de Murcia  
e-mail: lmolera@um.es

**M<sup>a</sup> Teresa Díaz Delfa**

Departamento de Métodos Cuantitativos para la Economía  
Universidad de Murcia  
e-mail: mdd@um.es

## **Resumen**

El desarrollo espectacular experimentado por el turismo rural en la última década ha dado lugar a diversos estudios sobre el tema, tanto desde la perspectiva de la oferta como de la demanda. Siguiendo este último enfoque, en este trabajo se analizan las características y motivaciones de los turistas rurales que eligen como destino la comarca del Noroeste de la región de Murcia, donde se concentra la mayor parte de la oferta de alojamientos rurales de la región. Además, para disponer de un conocimiento más profundo sobre la demanda, se obtienen diferentes segmentos con motivaciones específicas para la práctica del turismo rural. Dicha segmentación se lleva a cabo utilizando conjuntamente redes neuronales autoorganizadas y el algoritmo de análisis cluster no jerárquico *K*-means, combinación que ha proporcionado buenos resultados en algunos estudios anteriores.

*Palabras clave:* segmentación de mercados, análisis cluster, redes neuronales autoorganizadas.

---

<sup>1</sup> Este trabajo forma parte del proyecto “Crecimiento Económico y Modelos de Preferencias en el Estudio del Turismo” n° 4628 financiado por la función Séneca en el bienio 2002-2004.

## **1. Introducción.**

El turismo rural es una noción muy amplia que abarca cualquier actividad turística que se desarrolle en el medio rural (Bardón Fernández, 1990). Como notan algunos autores (Martín Gil, 1994; Kastenholtz, Davis y Paul, 1999; Yagüe Perales, 2002), en sí no es un fenómeno totalmente nuevo, puesto que ya en el siglo XIX era habitual, por una parte, el desplazamiento de los aristócratas a sus casas solariegas de las zonas rurales para pasar en ellas los meses de verano y, por otra, el retorno en vacaciones a los pueblos de origen de los individuos que con la revolución industrial habían emigrado a las ciudades. Sin embargo, este turismo rural tradicional no creaba oportunidades de empleo y negocio para retener a la población, que, en general, acababa emigrando en busca de mejores posibilidades económicas (Yagüe Perales, 2002).

Ahora bien, a finales del siglo XX, con el declinar de la agricultura, surge una nueva modalidad de turismo rural, cuyo principal objetivo es potenciar el desarrollo económico y social de las zonas rurales más deprimidas (Blaine y Golan, 1993; Cavaco, 1995). Varios factores han sido determinantes en la rápida evolución experimentada por este tipo de turismo en los últimos años. Entre otros, cabe destacar el apoyo de las administraciones públicas –por ejemplo, a través de los programas Leader o de mejoras en infraestructuras–, el fuerte grado de cohesión y asociacionismo de los agentes implicados –con objeto de intercambiar información, coordinar la comercialización de la oferta y la calidad de la misma, etc.–, y, sobre todo, el deseo creciente de los individuos que viven en el medio urbano de pasar pequeños períodos de tiempo en zonas donde prime el contacto con la naturaleza, la tranquilidad y la calidad medioambiental que cada vez escasea más en las ciudades (Bote, 1987; Sharpley y Sharpley, 1997).

En este contexto, se han llevado a cabo estudios sobre los diferentes aspectos del turismo rural, analizando, por ejemplo, las implicaciones de esta actividad en el desarrollo socioeconómico de las áreas rurales deprimidas (Bardón, 1990; Blaine y

Golan, 1993), las actitudes de los residentes de las zonas rurales en las cuales se han ido instalando los alojamientos (Allen et al., 1988; Davis, Allen y Cosenza, 1988; Allen et al., 1993), o las características, necesidades y deseos del turista rural (Kastenholz, Davis y Paul, 1999; Fuentes García, 1995).

En esta última línea, este trabajo pretende estudiar las principales características (socioeconómicas, demográficas y de comportamiento de viaje) y motivaciones para la práctica del turismo rural de los individuos que eligen como destino la comarca del Noroeste de Murcia. La elección de esta zona se debe a que fue la primera en instalar alojamientos rurales en la comunidad y a que en ella se concentra la mayor parte de la oferta (en concreto, 962 plazas frente a las 1.490 de toda la región).

En los dos primeros puntos del siguiente apartado se detallan, respectivamente, la recogida de datos y el análisis descriptivo de los mismos; en el tercero se resume brevemente el funcionamiento de las redes neuronales autoorganizadas y su aplicación en análisis cluster y segmentación de mercados; y, finalmente, en el último punto se describen los diferentes segmentos de turistas rurales obtenidos mediante la aplicación combinada de redes neuronales autoorganizadas (concretamente un mapa de rasgos autoorganizado o SOM) y el algoritmo de análisis cluster no jerárquico *K*-means. El último apartado del trabajo, como es habitual, recoge las principales conclusiones obtenidas.

## **2. Metodología y análisis empírico.**

### **2.1. Encuesta.**

Los datos proceden de una encuesta realizada en Semana Santa de 2002 a cargo del proyecto de investigación “Crecimiento Económico y Modelos de Preferencias en el Estudio del Turismo”. Esta encuesta se planteó como una entrevista personal con cuestionario estructurado, e iba dirigida a los individuos que pernoctaron una o más noches durante Semana Santa en alojamientos rurales de la comarca del Noroeste de Murcia. Mediante un muestreo por rutas aleatorias se obtuvieron 335 encuestas, todas ellas con datos válidos para el estudio.

El hecho de desarrollar la encuesta en Semana Santa que, en principio, plantea una limitación temporal, se debe a que ésta es una de las épocas del año con mayor ocupación en los alojamientos rurales, lo que permitía un mejor aprovechamiento de los fondos destinados a la encuesta. Además, en estas fechas suelen estar presentes tanto turistas rurales habituales como otros que sólo lo hacen en ocasiones puntuales, con lo que el abanico de posibles tipologías era, en principio, más amplio.

Como ya se ha comentado, el cuestionario contenía preguntas relativas a características demográficas y socioeconómicas de los entrevistados (provincia de residencia, edad, estudios, profesión,...) y al comportamiento en relación a la actividad turística (actividades a desarrollar durante la estancia en el alojamiento rural, frecuencia en la realización de turismo rural y de otro tipo de turismo, etc.), así como otras en relación a las motivaciones para la práctica de turismo rural. A partir de las respuestas a estas últimas se intentan detectar diferentes segmentos en la demanda de turismo rural, cuyos perfiles son completados gracias a las variables demográficas, socioeconómicas y de comportamiento anteriores.

## ***2.2. Análisis descriptivo.***

En primer lugar, llevamos a cabo un análisis descriptivo con objeto de descubrir las características generales de la demanda de turismo rural en el Noroeste de Murcia. Algunos de los resultados más destacados son los siguientes:

- Alicante y Murcia son las provincias de procedencia de la mayoría de los turistas (66,9% y 19,4%, respectivamente).
- Todos los individuos entrevistados son menores de 50 años y el 63,6% tienen menos de 30 años.
- Sólo el 35,8% están casados. En cuanto a la tipología familiar, el 37,9% de los individuos de la muestra viven en pareja, con o sin descendientes a su cargo (26,3% y 11,6%, respectivamente), el 7,5% son individuos que viven solos, con amigos o con sus hijos, y el 54,6% restante viven con sus progenitores pero han viajado sin ellos.

- El 38,5% tiene sólo estudios primarios y únicamente el 26,6% ha cursado estudios universitarios.
- Aproximadamente la mitad (53,13%) son trabajadores por cuenta ajena.
- El 83% declara unos ingresos familiares mensuales inferiores a 2.405 € la mayoría de los cuales (57,6%) se sitúan en el tramo entre 1.200 €y 2.405 €
- Más de la mitad (59,7%) alquila directamente el alojamiento a los propietarios y en torno a una cuarta parte (28,1%) utiliza la Central de Reservas de NORATUR, asociación a la que pertenecen la mayoría de los alojamientos de la zona.
- La gran mayoría (93,1%) de los turistas viaja con algún amigo (aunque posiblemente también con alguien de la familia). El 40% se desplaza con algún individuo de su familia y el 39,1% lo hace con algún niño en el grupo.
- El gasto por persona y día en la zona turística (sin contar el alojamiento) es inferior a 30 €para el 80,6% de los encuestados.
- El 68,4% había hecho turismo rural anteriormente.
- El 49% obtiene referencias del lugar a través de conocidos y el 24% lo hacen a través de Internet.
- En cuanto a las actividades que piensan realizar, las más demandadas son pasear por los alrededores (57,9%) y senderismo (53,7%).

Estos datos coinciden, en líneas generales, con otros estudios que revelan algunas características de los turistas rurales españoles (Bote, 1987; Fuentes García, 1995). En ellos se establece que son individuos que viven en zonas urbanas, jóvenes (entre 25 y 45 años), con nivel sociocultural alto y poder adquisitivo medio-alto, y que presentan una actitud activa hacia el disfrute de actividades típicamente rurales, incluyendo deportes, trabajos relacionados con la agricultura y contemplación del paisaje. También destacan que más de la mitad de las personas que practican turismo rural lo hace acompañado de familiares y que la mayoría conoce su destino por medio de amigos y familiares. Quizás lo único que se desvíe un poco de este perfil en nuestro caso sea el elevado porcentaje de individuos con nivel de estudios primarios (38,5%).

Una vez analizadas las características generales de la demanda, para segmentar ésta utilizamos las valoraciones de los encuestados sobre dieciocho motivos posibles para la práctica de turismo rural (tabla 1), medidos en una escala Likert de 1 (muy poco valorado) a 5 (muy valorado).

<b>Motivaciones para la práctica del turismo rural</b>
1. Atmósfera relajada
2. Aire fresco y contacto con la naturaleza
3. Belleza del entorno
4. Lugar poco aglomerado
5. Posibilidad de realizar actividades al aire libre
6. Posibilidad de contemplar monumentos o arquitectura típica
7. Actividades culturales
8. Agroactividades (hacer queso, amasar pan,...)
9. Buena comida típica o tradicional
10. Posibilidad de convivir con los amigos
11. Posibilidad de convivir con la familia
12. Posibilidad de relacionarse con la gente del lugar
13. Posibilidad de conocer gente nueva
14. Posibilidad de viajar con niños
15. Posibilidad de descansar
16. Independencia y flexibilidad
17. Distancia de viaje razonable
18. Precio

Tabla 1

Las motivaciones más valoradas en término medio son la posibilidad de convivir con los amigos y la posibilidad de descansar, siendo las agroactividades y las actividades culturales las menos valoradas.

Por otra parte, al considerar bastantes motivaciones, algunas de las cuales con correlaciones altas entre sí, y teniendo en cuenta el valor de la medida de adecuación muestral de Kaiser-Meyer-Olkin y el resultado del contraste de esfericidad de Barlett, planteamos antes de la segmentación un análisis de componentes principales.

Aunque es una práctica habitual, este enfoque de aplicación del análisis cluster sobre las puntuaciones estandarizadas de un análisis factorial previo ha sido criticado por algunos investigadores en marketing que piensan que puede descartar información relevante y distorsionar la verdadera estructura de los datos (Arabie & Hubert, 1994). Otros autores (Green y Krieger, 1995; Schaffer y Green, 1998) recomiendan emplear variables externas no utilizadas en el proceso de agrupación para establecer los perfiles de los segmentos, de modo que si hay diferentes soluciones disponibles se elija la que presente asociaciones más significativas con estas variables. Siguiendo este razonamiento, en nuestro caso se realiza previamente un análisis factorial porque el resultado que se obtiene tras la aplicación combinada de SOM y *K*-means a las puntuaciones factoriales se puede interpretar bastante bien en relación a las variables exógenas existentes (características socioeconómicas, demográficas y pautas de comportamiento en el viaje).

En cuanto a la solución factorial obtenida, los siete factores con valor propio mayor que 1 explicaban conjuntamente el 68,61% de la varianza total de los datos. No obstante, el factor que englobaba las agroactividades y la comida típica o tradicional presentaba una baja consistencia interna, con un valor del coeficiente de fiabilidad  $\alpha$  de Cronbach inferior al valor 0,6 recomendado. Esto nos llevó al planteamiento de un segundo análisis de componentes principales sólo con dieciséis variables, una vez eliminadas las dos anteriores, solución que se recoge en la tabla 2, donde aparecen destacadas las correlaciones más altas. Se utilizó la rotación Varimax con objeto de obtener una interpretación más clara de los factores, de modo que cada uno de ellos tuviera, en la medida de lo posible, sólo cargas factoriales muy altas o muy bajas. En este caso se consideró una solución con seis factores que explica el 68,46% de la varianza total y en la que todos ellos presentan un valor apropiado del  $\alpha$  de Cronbach.

El primer factor recoge las variables que hacen referencia a la búsqueda de tranquilidad y contacto con la naturaleza, mientras que el segundo está relacionado con el desarrollo de actividades, principalmente de carácter cultural. La belleza del entorno y la posibilidad de desarrollar actividades al aire libre presentan relación con

ambos factores, como es lógico teniendo en cuenta los conceptos que éstos representan. Por otra parte, las variables relacionadas con la vida social y la búsqueda de nuevas amistades son las que más importancia tienen en el tercer factor, frente a las relaciones familiares que determinan el cuarto, en el que también influye la distancia de viaje razonable al lugar turístico elegido, algo que es lógico que suceda si, por ejemplo, se viaja con niños. En el quinto factor están incluidas las características del viaje y de la estancia y, por último, el sexto tiene en cuenta, sobre todo, la convivencia con los amigos y la posibilidad de realizar actividades al aire libre.

	FACTOR 1	FACTOR 2	FACTOR 3	FACTOR 4	FACTOR 5	FACTOR 6
Atmósfera relajada	<b>0,807</b>	0,020	-0,009	0,066	0,091	-0,030
Aire fresco y naturaleza	<b>0,762</b>	0,178	0,058	0,198	-0,093	0,193
Belleza del entorno	<b>0,642</b>	<b>0,325</b>	0,028	0,046	-0,089	0,228
Posibilidad de descansar	<b>0,616</b>	-0,086	0,195	-0,095	<b>0,298</b>	-0,271
Lugar poco aglomerado	<b>0,601</b>	-0,013	-0,067	0,119	0,238	0,054
Actividades culturales	0,015	<b>0,853</b>	0,084	0,066	0,105	-0,021
Monumentos o arquitect. típica	0,059	<b>0,839</b>	0,129	0,008	0,062	-0,077
Actividades al aire libre	<b>0,310</b>	<b>0,475</b>	0,102	0,110	-0,052	<b>0,398</b>
Relacionarse con gente del lugar	0,067	0,112	<b>0,934</b>	0,080	0,016	0,075
Conocer gente nueva	0,018	0,147	<b>0,934</b>	0,045	-0,026	0,084
Convivir con la familia	0,107	0,038	0,092	<b>0,885</b>	0,045	-0,025
Viajar con niños	0,112	0,069	0,030	<b>0,869</b>	0,019	-0,076
Precio asequible	0,073	0,177	-0,010	-0,108	<b>0,762</b>	-0,043
Distancia de viaje razonable	0,085	0,028	-0,072	<b>0,341</b>	<b>0,739</b>	0,206
Independencia y flexibilidad	<b>0,455</b>	-0,110	0,121	-0,045	<b>0,479</b>	0,052
Convivir con los amigos	0,077	-0,081	0,134	-0,137	0,141	<b>0,840</b>
VARIANZA EXPLICADA	17,11%	11,83%	11,71%	11,13%	9,79%	6,89%
VARIANZA ACUMULADA	17,11%	28,94%	40,65%	51,78%	61,57%	68,46%

Tabla 2

Las puntuaciones factoriales tipificadas de los 335 individuos encuestados en los seis factores retenidos se utilizan como variables de partida para la segmentación de mercados, esto es, para determinar grupos de turistas homogéneos internamente y heterogéneos entre sí. Dicha segmentación se llevará a cabo, como ya hemos comentado, mediante un análisis cluster en dos pasos que combina una técnica ámbito de la inteligencia artificial, como son las redes neuronales, y otra técnica de análisis cluster tradicional (en este caso el algoritmo no jerárquico *K-means*).

### ***2.3. Redes neuronales autoorganizadas para análisis cluster y segmentación de mercados.***

Una red neuronal es un modelo matemático que intenta reproducir el funcionamiento del cerebro, imitando su proceso de aprendizaje a partir de ejemplos. En una red neuronal se distinguen dos elementos: las neuronas o unidades que procesan la información, y las conexiones entre ellas, que permiten la transmisión de ésta. Las neuronas de una red se organizan generalmente en capas: una de entrada, donde se sitúan las neuronas que reciben los datos o patrones de aprendizaje; a continuación, posibles capas ocultas; y, finalmente, una capa de salida donde se obtiene el resultado del proceso. Las conexiones de la red llevan asociadas unos pesos o ponderaciones que se van modificando a medida que la red va aprendiendo de la información que le llega del exterior.

La utilización de redes neuronales ha permitido resolver problemas complejos, aplicándose en diferentes disciplinas con el fin de predecir, analizar similitudes, clasificar, etc. En particular, nosotros estamos interesados en las redes neuronales que permiten hacer análisis cluster, entre las que destacan las denominadas redes neuronales autoorganizadas (*self-organizing map*, SOM) desarrolladas por Kohonen (Kohonen, 1982; Kohonen, 2001). En ellas, las neuronas se disponen en sólo dos capas –una de entrada y otra de salida, siendo esta última un retículo generalmente unidimensional o bidimensional– y el aprendizaje es competitivo, esto es, a todas las neuronas de la capa de salida les llega la misma información, compiten entre sí y sólo una de ellas, llamada prototipo, es la ganadora. Durante el proceso de aprendizaje la red debe descubrir por sí misma rasgos comunes, regularidades o

agrupaciones en los datos, y reflejar dicha información en las ponderaciones asignadas a sus conexiones, de ahí su nombre de autoorganizada. Este proceso de aprendizaje es no supervisado, porque no hay un resultado final conocido de antemano con el que comparar para aprender, como sucede en el caso supervisado.

Formalmente, cada neurona  $j$  de la capa de salida está conectada con todas las neuronas de la capa de entrada mediante un vector de ponderaciones  $w_j$  y también presenta conexiones laterales con las neuronas de la propia capa de salida. Tras iniciar los vectores de ponderaciones con valores aleatorios, para cada ejemplo de entrenamiento o patrón de entrada  $x$  se calcula la distancia que lo separa de cada uno de los vectores de ponderaciones asignados a las distintas neuronas del retículo. La neurona ganadora o prototipo,  $g$ , será la más parecida al patrón de entrada según una medida de distancia, y su vector de ponderaciones y el de algunas neuronas vecinas se actualizará para parecerse todavía más a dicho patrón, con objeto de que ante el mismo ejemplo se active con mayor intensidad y ante otro ejemplo parecido al anterior se active o bien dicha neurona o bien otra cercana. De esta forma, patrones muy similares en las variables input serán asignados a la misma neurona o neuronas próximas en el espacio de salida. El proceso de aprendizaje consta de dos fases. En la primera fase, denominada de ordenación, se comienza con una tasa de aprendizaje y un tamaño de vecindad grandes, para luego ir reduciendo esos valores a medida que avanza el aprendizaje. Finalmente, en la segunda fase o fase de convergencia, se busca un ajuste más fino, haciendo que únicamente cambien los pesos de la neurona ganadora, permaneciendo inactivas todas las de su entorno.

Aunque no es la única forma de hacer análisis cluster con redes autoorganizadas, en este trabajo utilizamos como capa de salida un retículo unidimensional con tantas neuronas como el número de conglomerados  $K$  que se pretenden construir. En concreto, utilizamos un programa escrito en S-Plus por Waller, Kaiser, Illian y Manry<sup>2</sup> (1998) basado en el siguiente algoritmo (válido tanto para la fase de ordenamiento como para la de convergencia, con las correspondientes modificaciones sobre los valores iniciales de los parámetros), en el cual el proceso de

aprendizaje para cuando se ha realizado un número suficientemente elevado  $I$  de pasadas sobre todo el conjunto de datos de entrenamiento:

1. Inicialización: Para  $K$  clusters, inicializar los vectores de pesos  $\mathbf{w}_j$ ,  $j=1, \dots, K$ , que representan los centroides de los clusters, como vectores aleatorios de modo que cada componente tome valores según una distribución uniforme en el rango de la correspondiente variable input. Inicializar  $i=0$ . Inicializar el ratio de aprendizaje  $\alpha$  y el parámetro tamaño del entorno  $\eta$  en valores predeterminados.
2. Incremento: Hacer  $i = i+1$ . Abandonar si  $i > I$ .
3. Determinación de la neurona ganadora y clasificación: Para  $n=1, \dots, N$ , donde  $N$  es el número de patrones de aprendizaje, hacer  $t = n + (i-1)N$  y encontrar la neurona ganadora  $g_n$  en la capa de salida correspondiente a  $\mathbf{x}_n$  según la distancia euclídea,

$$\|\mathbf{x}_n - \mathbf{w}_{g_n}\| = \min_{j=1, \dots, K} \{\|\mathbf{x}_n - \mathbf{w}_j\|\}.$$

4. Actualización: Actualización de los pesos según la regla

$$\mathbf{w}_j = \mathbf{w}_j + z(t)(\mathbf{x}_n - \mathbf{w}_j)$$

para aquellas neuronas  $j$  tales que  $|j - g_n| \leq h(t)$ , con  $h(t)$  la función que determina el rango de excitación lateral en el entorno de la neurona ganadora,

$$h(t) = \mathbf{H}\left(1 - \frac{t}{1 + NI}\right),$$

$$\text{y } z(t) = a\left(1 - \frac{t}{1 + NI}\right)$$

5. Evaluación: Calcular el error medio y volver a la etapa 2,

$$ErrorMedio = \frac{1}{N} \sum_{n=1}^N \|\mathbf{x}_n - \mathbf{w}_{g_n}\|.$$

---

<sup>2</sup> Som.prg es software de libre distribución que se puede encontrar en la siguiente dirección: [http://peabody.vanderbilt.edu/depts/psych\\_and\\_hd/faculty/wallern/SOM/Som.prg](http://peabody.vanderbilt.edu/depts/psych_and_hd/faculty/wallern/SOM/Som.prg)

Se han llevado a cabo distintos estudios de simulación intentando comparar la capacidad de las redes neuronales para realizar análisis cluster con la de diversos métodos de análisis de conglomerados tradicionales. Así, Balakrishnan, Cooper, Jacob y Lewis (1994) comparan las redes de Kohonen unidimensionales para análisis cluster con el tradicional *K*-means (Hartigan, 1975) mediante datos generados artificialmente, concluyendo que estas redes funcionan peor que *K*-means. También trabajan con otra red neuronal, concretamente Frequency-Sensitive Competitive Learning Algorithm (Krishnamurthi et al., 1990), esta vez con el objetivo particular de hacer una segmentación de mercado. Como los resultados obtenidos sobre un conjunto de datos reales son diferentes a los proporcionados por *K*-means, los autores finalmente combinan ambos enfoques puesto que así obtienen resultados con mejor interpretación. Otros investigadores que comparan mediante datos simulados técnicas cluster tradicionales (en este caso de cluster jerárquico) con SOM bidimensional son Mangiameli, Chen y West (1996), en cuyo caso las redes neuronales se revelan superiores. Más tarde, Waller, Kaiser, Illian y Manry (1998) llevan a cabo otro estudio de Monte Carlo con datos más generales que los del primer estudio de Balakrishnan et al. y obtienen que las redes tipo SOM unidimensional funcionan igual o mejor que algunos de los métodos de análisis cluster tradicionales (jerárquicos y no jerárquicos). Finalmente, Kuo, Ho y Hu (2002a, 2002b) utilizan en un ejemplo práctico de segmentación de mercado un método en dos pasos que es una combinación de SOM bidimensional y *K*-means, y realizan estudios de simulación que indican que el esquema propuesto es ligeramente mejor que el método convencional de análisis cluster en dos etapas sugerido por Punj y Stewart (1983) que consta de la integración de un método no jerárquico (como, por ejemplo, Ward que es de los que mejor funcionan en simulaciones de Monte Carlo) con uno no jerárquico como *K*-means. El motivo de plantear un paso previo a la aplicación del método de *K*-means se debe, principalmente, al problema de determinar puntos semillas iniciales para los centroides de los grupos, ya que este método no jerárquico se comportado bastante bien en ejercicios de simulación en los que se asignan

valores iniciales para los centroides próximos a los verdaderos valores de éstos, pero puede presentar problemas si se eligen semillas aleatorias (Milligan, 1980).

#### ***2.4. Segmentación del mercado de turismo rural.***

Teniendo en cuenta lo anteriormente expuesto, en este trabajo se emplea una combinación de redes neuronales autoorganizadas y  $K$ -means, donde los resultados obtenidos por el SOM se utilizan como centroides iniciales para el algoritmo  $K$ -means. Además, la red autoorganizada que se considera es unidimensional, con una capa de entrada que consta de seis neuronas (correspondientes a las seis componentes de la solución factorial ya estudiada) que reciben como patrones de entrada las puntuaciones factoriales de los 335 individuos encuestados, y en cuya capa de salida se dispone un número de neuronas  $K$  igual al número de conglomerados que se pretenden detectar. El motivo de utilizar una red de Kohonen unidimensional se debe precisamente a la simplicidad en la relación entre el número de clusters y la estructura de la capa de salida.

Como no hay contrastes estadísticos que permitan determinar el número más adecuado de clusters, en nuestro caso particular se ensayó con distintos valores, en principio pequeños (de 2 a 5), puesto que conceptualmente parece lógico esperar un número reducido de grupos debido a que el mercado está todavía en fase de desarrollo, y los recursos disponibles no permiten una amplia variedad de opciones (Kastenholz, Davis y Paul, 1999). Finalmente, se consideró la mejor solución cluster como aquella que permitía la identificación de los segmentos más distinguibles, interpretables y significativos desde el punto de vista de asociación con variables externas (demográficas, socioeconómicas, de comportamiento, etc.) no utilizadas en la construcción de los clusters.

La solución elegida es la correspondiente a cuatro conglomerados, y la interpretación de éstos se lleva a cabo teniendo en cuenta tanto los valores medios en los mismos de las valoraciones para las distintas motivaciones (tabla 3, en la que se recogen únicamente los casos en los que el contraste ANOVA es significativo al 5%, aunque estos contrastes deben interpretarse más bien desde un punto de vista descriptivo, ya

que no todos los supuestos se verifican) como las variables exógenas que muestran asociación significativa con la agrupación a nivel del 10%.

MOTIVACIONES PARA PRÁCTICAR TURISMO RURAL	MEDIA GLOBAL	MEDIAS			
		Cluster 1	Cluster 2	Cluster 3	Cluster 4
Atmósfera relajada	4,12	2,64	4,55	4,48	3,64
Aire fresco y naturaleza	4,26	2,78	4,58	4,73	3,70
Belleza del entorno	3,65	2,36	4,09	3,93	3,15
Lugar poco aglomerado	4,01	2,91	4,25	4,33	3,70
Realizar actividades al aire libre	3,49	2,55	3,83	3,76	2,76
Ver monumentos o arq. típica	2,49	1,85	2,71	2,54	2,61
Actividades culturales	2,06	1,56	2,16	2,17	2,06
Comida típica o tradicional	3,09	2,51	2,82	3,48	3,30
Convivir con los amigos	4,64	4,85	4,87	4,87	2,55
Convivir con la familia	3,07	1,95	1,44	4,74	3,21
Relacionarse con gente del lugar	3,18	3,15	2,97	3,52	2,42
Conocer gente nueva	3,16	3,16	3,02	3,41	2,52
Viajar con niños	2,48	1,29	1,26	3,76	2,94
Posibilidad de descansar	4,31	3,42	4,55	4,47	4,30
Independencia y flexibilidad	4,19	3,71	4,44	4,34	3,61
Distancia de viaje razonable	3,85	3,53	3,59	4,29	3,36

Tabla 3

En primer lugar, en cuanto a las motivaciones, y siempre hablando en términos de los valores medios, el grupo 1 (16,42% de la muestra) está formado por los individuos que menos valoran las variables que componen el primer factor (naturaleza, belleza del entorno, posibilidad de descansar, ...) y las actividades de tipo cultural. También son, junto con los del grupo 4 (9,85% de los entrevistados), los que menos valoran las actividades al aire libre y la independencia y flexibilidad, y junto con los del grupo 2 (31,94% de los encuestados) los que menos interés muestran por la comida

típica y la posibilidad de convivir con la familia y viajar con niños. Por el contrario, valoran mucho la posibilidad de convivir con los amigos (al igual que todos los demás grupos, excepto el 4) y son, junto con el grupo 3 (41,79% de la muestra) los que más valoran las posibilidades de hacer nuevas amistades.

Los que más valoran las variables correladas con el primer factor así como la posibilidad de practicar actividades de todo tipo y la independencia y flexibilidad que este tipo de turismo suele proporcionar son los individuos de los grupos 2 y 3, diferenciándose entre sí por el hecho de que los turistas del cluster 3 valoran más la distancia de viaje razonable, la posibilidad de viajar con niños y de convivir con la familia, así como la gastronomía tradicional.

Finalmente, el cuarto conglomerado se caracteriza principalmente por su valoración relativamente baja para la posibilidad de convivir con los amigos, en contraste con los otros tres grupos. Son también los que presentan puntuaciones más bajas en la posibilidad de realizar actividades al aire libre (junto con el grupo 1) y de realizar nuevas amistades. Valoran un poco menos que los grupos 2 y 3 la naturaleza, la belleza, la poca aglomeración, etc., pero más que los del grupo 1. También muestran cierto interés por la familia y los niños pero menos que los del grupo 3.

En relación a las variables externas, las variables socioeconómicas y demográficas significativas para diferenciar entre los clusters al 10% son el estado civil, la edad, la tipología familiar, los estudios y la situación laboral, siendo las asociaciones entre modalidades más destacadas las siguientes:

- En los grupos 1 y 2 apenas hay casados (5,5% y 6,5% respectivamente), mientras que en los grupos 3 y 4 más de la mitad están casados (65% y 57,6%).
- En los grupos 1 y 2 la mayoría son menores de 30 años (87,3% y 83,2%, respectivamente); en cambio, aproximadamente la mitad de los individuos de los grupos 3 y 4 tienen más de esa edad (57,1% y 51,5%).

- En los grupos 3 y 4 la mayoría de los entrevistados viven en pareja y/o tienen hijos a su cargo (70,7% y 63,6%) y en los clusters 1 y 2 la mayoría viven con sus padres (81,8% y 84,1%).
- En el grupo 1 sólo el 12,7% tiene estudios universitarios frente a los 36,4% en el grupo 4.
- El grupo con mayor porcentaje de no ocupados es el primero (36,4%), siendo todos ellos estudiantes. Los grupos 3 y 4 son los de mayor porcentaje de ocupados (82,1% y 81,8%).

Por otra parte, las asociaciones significativas para diferenciar entre los clusters al 10% en el caso de las variables de comportamiento en los viajes se resumen en los siguientes puntos:

- El mayor porcentaje de los que no había practicado turismo rural antes corresponde al grupo 2 (37,4%).
- En todos los grupos, excepto el 4, casi la totalidad de los individuos viajan con algún amigo, mientras que en este último el 30,3% viaja sólo con la pareja o la familia.
- En los grupos 3 y 4 es bastante frecuente viajar con algún miembro de la familia (72,1% y 63,6%), mientras que esta situación es más rara en los grupos 1 y 2 (9,1% y 6,5%).
- La mayoría de los individuos del cluster 3 viajan con algún niño en el grupo (72,9%) y aproximadamente la mitad de los del cluster 4 lo hacen (57,6%). Además, el número medio de niños con los que se viaja en el grupo 3 es mayor que en el 4, puesto los turistas de ese cluster suelen viajar en grupos de familias. En cambio, los de los grupos 1 y 2 viajan casi todos sin ningún menor (96,4% y 92,5%).
- Los turistas de los clusters 1 y 3 suelen viajar en grupos más grandes (el 52,7% y el 45,7%, respectivamente, lo hacen en grupos de más de 6 individuos), mientras que los grupos más pequeños se dan en el cluster 4 (el 45,5% se desplazan en grupos de menos de 6 personas).
- Los que más gastan por término medio son los del grupo 4 (el 30,3% gasta más de 30 € por persona y día, frente a porcentajes en dicho tramo inferiores

al 20% en los restantes clusters, y el 33,3% gasta menos de 15 € frente a porcentajes en dicho tramo superiores al 40% en los otros grupos).

- Los del cluster 2 son los que más interés tienen por practicar senderismo (62,6%), montar a caballo (56,1%) y hacer la bajada del río Segura (42,1%). No olvidemos que es el grupo que más valoraba las actividades al aire libre, junto con el 3. Además, el número medio de actividades que piensan realizar en estos grupos 2 y 3 es mayor que en los demás.

En conclusión, los perfiles definitivos de los grupos serán los siguientes:

- **CLUSTER 1** (16,42% de la muestra): Son generalmente jóvenes y solteros, la mayoría vive con los padres, y muchos de ellos son estudiantes, aunque pocos tienen título universitario (es posible que algunos sean tan jóvenes que todavía no les haya dado tiempo a conseguirlo, pero eso no lo podemos determinar al estar la variable edad dividida en tramos y ser “menor de 30 años” el primero de los tramos). Viajan en grupos grandes de amigos, valoran mucho la posibilidad de convivir con éstos y también bastante la posibilidad de hacer nuevas amistades, pero muy poco la posibilidad de convivir con la familia. Son los que menos valoran el entorno natural, la tranquilidad y la ausencia de contaminación, y también los que menos interés muestran por la práctica de actividades.
- **CLUSTER 2** (31,94% de la muestra): Está formado principalmente por jóvenes solteros que viven con sus padres. Viajan en grupos de amigos, pero más pequeños que los del cluster 1, siendo asimismo en este caso muy valorada la posibilidad de convivir con los amigos, a diferencia de la convivencia con la familia. Son, junto con los del grupo 3 los que más importancia conceden al entorno natural, la tranquilidad y la ausencia de contaminación, así como a la posibilidad de practicar actividades de todo tipo.
- **CLUSTER 3** (41,79% de la muestra): Lo componen en su mayoría grupos de familias con niños, aunque también hay grupos de parejas y matrimonios, por

lo que son los que más destacan la posibilidad de viajar en familia, niños incluidos, valorando también mucho la compañía de los amigos. Son, junto con los del cluster anterior, los que más valoran el ambiente tranquilo, el contacto con la naturaleza y la práctica de actividades.

- **CLUSTER 4** (9,85% de la muestra): Está constituido por parejas, familias que viajan solas y grupos pequeños de amigos y/ parejas. Son los que menos interés muestran por convivir con los amigos y por practicar actividades al aire libre. Presentan el mayor porcentaje de estudios superiores de todos los grupos y son los que hacen un mayor gasto en la zona turística.

Finalmente, conviene destacar varios aspectos técnicos de la solución obtenida. Por un lado, el SOM se comportó de manera bastante estable, en el sentido de que las diferencias en los resultados obtenidos para distintas inicializaciones de los parámetros fueron mínimas, eligiéndose la solución que proporcionaba un menor error medio. No obstante, su interpretación era bastante peor que la de su combinación con *K*-means. Por otra parte, los segmentos obtenidos con este método combinado son prácticamente los mismos que los que resultan de aplicar una combinación del método jerárquico de Ward con *K*-means (valor  $-0,316$  para el coeficiente Kappa de Cohen, con un *p*-valor muy inferior a  $0,05$ ), aunque la clasificación de SOM junto con *K*-means parece ligeramente mejor, puesto que si comparamos el porcentaje de predicciones correctas de los respectivos análisis discriminantes en los que las puntuaciones factoriales son las variables independientes y la dependiente es la que indica la pertenencia a los distintos cluster, obtenemos  $96,4\%$  para SOM más *K*-means y  $94,6\%$  para Ward más *K*-means. Nótese que para *K*-means con los centroides iniciales considerados por defecto por SPSS 11.0 el porcentaje clasificaciones correctas baja a  $91\%$ .

### **3. Conclusiones.**

A partir de la información de la encuesta, y utilizando técnicas multivariantes tradicionales y redes neuronales autoorganizadas, se han logrado identificar cuatro grupos de turistas rurales diferenciados en algunas de las motivaciones para la

práctica de turismo rural consideradas en el estudio, así como en varias de las variables externas (socioeconómicas, demográficas, ...).

De todos modos, vistos los resultados, podría ser interesante realizar una nueva encuesta donde se integren otra serie de motivos (clima, ambiente nocturno, alojamiento con encanto rústico, etc.) para ver si permitan diferenciar con mayor claridad distintos segmentos en la demanda de turismo rural. Además, esto nos permitiría subsanar la limitación temporal ya comentada, y contrastar si existen o no diferencias sustanciales entre el turista rural de todo el año y el de Semana Santa.

Por último, aunque nuestros resultados parecen buenos desde el punto de vista de interpretación de los segmentos, sería conveniente realizar estudios de simulación comparativos del método de análisis cluster en dos pasos utilizado (SOM unidimensional más *K*-means) con otros alternativos, puesto que precisamente esta combinación específica no ha sido analizada todavía.

## **Bibliografía**

1. Allen, L.H., Long, P., Perdue, R., Kieselbach (1988): "The impact of tourism development on residents' perceptions of community life", *Journal of Travel Research*, **27**, (summer), pp. 16-21.
2. Allen, L.H., Hafer, H., Long, P., Perdue, R. (1993): "Rural residents' attitudes toward recreation and tourism development", *Journal of Travel Research*, **31**, (spring), pp. 27-33.
3. Arabie, P., Hubert, L. (1994): "Cluster analysis in market research", in R.P. Bagozzi (ed.), *Advanced methods in marketing research*, pp 160-189, Blackwell, Oxford.
4. Balakrishnan, P.V., Cooper, M.C., Jacob, V.S., Lewis, P.A. (1994): "A study of the classification capabilities of neural networks using unsupervised learning: A comparison with K-means clustering", *Psychometrika*, **59** (1), pp. 509-525.
5. Balakrishnan, P.V., Cooper, M.C., Jacob, V.S., Lewis, P.A. (1996): "Comparative performance of the FSCL neural network and K-means algorithm

- for market segmentation", *European Journal of Operational Research*, **93**, pp. 346-357.
6. Bardón Fernández, E. (1990): "Consideraciones sobre el turismo rural en España y medidas de desarrollo", *Estudios Turísticos*, 108, pp. 61-82.
  7. Blaine, T., Golan, M. (1993): "Demand for rural tourism: an exploratory study", *Annals of Tourism Research*, **20**, pp. 770-773.
  8. Bote, V. (1987): "Importancia de la demanda turística en espacio rural en España", *Estudios Turísticos*, **93**, pp. 79-88.
  9. Cavaco, C. (1995): "Rural tourism: the creation of new tourism spaces", in A. Montanari, A. Williams (ed.), *Regions, spaces and restructuring*, pp 129-149, Wiley, Chichester.
  10. Consejería de Turismo y Cultura (2001): *El turismo en la Región de Murcia. Resultados del 2000*. Murcia Turística.
  11. Davis, D., Allen, J., Cosenza, R. (1988). "Segmenting local residents by their attitudes, interests and opinions towards tourism", *Journal of Travel Research*, **27**, (fall), pp. 2-8.
  12. Fuentes García, R. (1995): "Análisis de las principales características de la demanda de turismo rural en España", *Estudios Turísticos*, **127**, pp. 19-52.
  13. Green, P.E., Krieger, A.M. (1995): "Alternative approaches to cluster-based market segmentation", *Journal of the Marketing Research Society*, **37** (3), pp. 221-239.
  14. Hair, J.F., Anderson, R.E., Tatham, R.L., Black, W.C. (1999): *Análisis multivariante*, Prentice Hall, Madrid.
  15. Hartigan, J.A. (1975): *Clustering algorithms*, Wiley, New York.

16. Kastenholz, E., Davis, D., Paul, G (1999): "Segmenting tourism in rural areas: the case of North and Central Portugal", *Journal of Travel Research*, **37** (may), pp. 353-363.
17. Kohonen, T. (1982): "Self-organized formation of topologically correct feature maps", *Biological Cybernetics*, **43**, pp. 59-69.
18. Kohonen, T. (2001): *Self-Organizing Maps*, third edition, Springer, Berlín, Heidelberg.
19. Kuo, R.J., Ho, L.M., Hu, C.M. (2002a): "Integration of self-organizing feature map and K-means algorithm for market segmentation", *Computers & Operations Research*, **29**, pp. 1475-1493.
20. Kuo, R.J., Ho, L.M., Hu, C.M. (2002b): "Cluster analysis in industrial market segmentation through artificial neural network", *Computers & Industrial Engineering*, **42**, pp. 391-399.
21. Mangiameli, P., Chen, S.K., West, D. (1996): "A comparison of SOM neural network and hierarchical clustering methods", *European Journal of Operational Research*, **93**, pp. 402-417.
22. Martín del Brío, B., Sanz Molina, A. (2001): *Redes neuronales y sistemas borrosos*, Ra-Ma, Madrid.
23. Martín Gil, F. (1994): "Nuevas formas de turismo en los espacios rurales españoles", *Estudios Turísticos*, **122**, pp. 15-39.
24. Milligan, G.W. (1980): "An examination of the effect of six types of error perturbations on fifteen clustering algorithms", *Psychometrika*, **45**, pp. 325-342.
25. Punj, G., Steward, D.W. (1983): "Cluster analysis in marketing research: review and suggestions for applications", *Journal of Marketing Research*, **20**, pp. 134-148.

26. Schaffer, C.M., Green, P.E. (1998): "Cluster-based market segmentation: some further comparisons of alternative approaches", *Journal of the Marketing Research Society*, **40** (2), pp. 155-163.
27. Sharpley, J., Sharpley, R. (1997): *Rural tourism. An introduction*. London: International Thomson Business Press.
28. Waller, N.G., Kaiser, H.A., Illian, J.B., Manry, M. (1998): "A comparison of the classification capabilities of the 1-dimensional Kohonen neural network with two partitioning and three hierarchical cluster analysis algorithms", *Psychometrika*, **63** (1), pp. 5-22.
29. Yagüe Perales, R.M. (1998): "El desarrollo del turismo rural a través del programa LEADER I: El caso de la comarca valenciana de Els Ports", *Estudios Turísticos*, **136**, pp. 5-26.
30. Yagüe Perales, R.M. (2002): "Rural tourism in Spain", *Annals of Tourism Research*, **29** (4), pp. 1101-1110.