

CLASIFICACIÓN DE LAS ZONAS TURÍSTICAS ESPAÑOLAS EN FUNCIÓN DE LAS CARACTERÍSTICAS ESTRUCTURALES DE LA OFERTA Y LA DEMANDA

Una aplicación mediante un modelo de clases latentes

José Manuel Ramírez Hurtado*
Universidad Pablo de Olavide - España
Jesús Manuel López Bonilla**
Universidad de Sevilla - España

Resumen: La identificación de los segmentos turísticos por parte de los agentes y de los organismos implicados en el turismo permite una asignación más eficiente de los recursos. El objetivo de este trabajo es hacer una tipificación de las zonas turísticas españolas mediante el uso de un modelo de clases latentes en función de las características estructurales de la oferta y demanda asociadas a ellas. Los resultados muestran la existencia de tres grupos muy diferenciados de zonas turísticas. Se ofrece una clasificación de dichas zonas turísticas.

PALABRAS CLAVE: turismo, zona turística, clasificación, segmentación, clases latentes.

Abstract: A Classification of Spanish Tourist Zones Based on Structural Characteristics of Supply and Demand. An Application Using a Latent Class Model. Tourist agents and organisms can make an efficient assignation of resources by identifying tourist segments. The goal of this work is to classifying the Spanish tourism zones by means of supply and demand, using a latent class model. Results show the existing of three different groups. A classification of these tourism zones is listed in the paper.

KEY WORDS: tourism, tourism zones, classification, segmentation, latent class.

INTRODUCCIÓN

El éxito de un destino turístico depende de su mayor o menor habilidad para transformar sus atractivos recursos naturales y humanos en recursos turísticos (Ravelo, Moreno, Oreja & Deoula, 2007). Esto determinará en gran medida la capacidad que tiene el destino turístico para atraer a los potenciales clientes y lograr su satisfacción.

El entorno turístico en el que se mueven los destinos españoles es complejo y cambiante (Trujillo, López & Alarcón, 2002). La aparición de nuevos destinos, los cambios en la distribución afectados por las nuevas tecnologías, las alianzas estratégicas e integradoras de empresas, etc., hacen que el turismo sea un sistema complejo que goza de una característica muy importante como es la heterogeneidad. De esta forma, alcanzar óptimos niveles de competitividad se convierte en un factor absolutamente estratégico para las zonas turísticas (Alonso, 2010).

* Doctor por la Universidad Pablo de Olavide, de Sevilla, España. Departamento de Economía, Métodos Cuantitativos e Historia Económica, Universidad Pablo de Olavide, España. E-mail: jmramhur@upo.es.

** Doctor por la Universidad de Sevilla, España. Departamento de Administración de Empresas y Comercialización e Investigación de Mercados, Universidad de Sevilla, España. E-mail: lopezbon@us.es.

El conocimiento de los segmentos de la demanda turística por parte de las empresas y de los organismos públicos trae consigo un uso más eficiente de los recursos ya que se pueden seleccionar aquellos grupos estratégicamente más importantes y determinar las políticas de productos, precio y distribución más adecuadas (Palacio & McCool, 1997). En el mismo sentido, la segmentación de la demanda turística permite la elaboración de planes de acción y marketing apropiados para cada fracción del mercado (Raya & Caparrós, 2008).

Así pues, el objetivo de este trabajo es el de clasificar las zonas turísticas españolas en diversos cluster, con el fin de analizar y optimizar la gestión estratégica y comercial de la oferta y demanda de dichas zonas.

La literatura económica ha abordado la segmentación del mercado turístico desde diversas perspectivas (Eymann & Ronning, 1997). La gran heterogeneidad de trabajos existentes se debe principalmente a los criterios y técnicas de segmentación empleados (Raya & Caparrós, 2008). Los criterios hacen referencia a las variables que posteriormente son utilizadas para construir los segmentos turísticos (Picón, Varela & Real, 2003). En cuanto a las técnicas de segmentación, hacen alusión a la metodología utilizada para la obtención de los diversos segmentos. La mayoría de estas técnicas persiguen la tipificación de turistas a partir de un conjunto de variables (Fodnees & Murray, 1998; Jurowski & Reich, 2000; McKercher, 2002). En España destacan los trabajos de González (1999); Gutiérrez & Montero (2000); Dávila, Rodríguez & Rodríguez (2002); Roig, Rodríguez & Buitrago (2005); Albadalejo & Morera (2006); González, Sánchez & Sanz (2006); Rodríguez (2006); y Ferreria, Rial & Varela (2010); entre otros, para la identificación de perfiles turísticos.

En este estudio se pretende tipificar las zonas turísticas españolas en función de una serie de variables asociadas a la oferta y demanda que presenta cada una de ellas. A diferencia de los anteriores, en este trabajo se emplea como unidad la zona turística en vez del turista. También se utiliza una metodología diferente como es un modelo de clases latentes.

La justificación de la realización de esta investigación es la importante repercusión que tiene una adaptación eficiente de la oferta y demanda turística tanto para los agentes implicados en el turismo como para la elaboración de las políticas económicas.

El artículo está estructurado en varias partes. A continuación se analiza el concepto de zona turística, diferenciándolo del de destino turístico. Seguidamente se hace una revisión de la importancia de la segmentación en el turismo. Posteriormente, se desarrolla la metodología a utilizar en este estudio, como es el análisis de clases latentes. En el siguiente epígrafe se describe todo lo referente a la obtención de los datos para el análisis. Posteriormente, se estima el modelo y se obtienen los resultados. Finalmente se comentan las conclusiones más importantes.

EL CONCEPTO DE ZONA TURÍSTICA

El concepto de zona turística ha sido poco utilizado en la investigación sobre el turismo, quizá debido a la complejidad de su definición. Uno de los pocos autores que ha analizado el concepto de zona turística es Esteve Secall (2002: 2). Según este investigador, se entiende por zona turística *un territorio, normalmente de carácter supramunicipal, caracterizado por una gran concentración de actividades turísticas (o en el que el turismo es la actividad económica preferente), donde se ofertan diversos productos turísticos, -aunque predomine un tipo de producto concreto-, que integra diferentes destinos turísticos, y que se manifiesta como un conglomerado de tres tipos de agrupamientos de empresas o entidades turísticas: a) de carácter empresarial; b) de carácter territorial; y c) de carácter institucional.*

La aparición de las zonas turísticas se debe a la evolución expansiva de uno o varios destinos turísticos próximos, o de uno o varios productos turísticos concretos ofertados en un entorno geográfico, asimismo, próxima. En cualquier caso, la oferta de una zona turística debe tener una dimensión relevante, al igual que su demanda (Esteve Secall, 2002). Esto permite diferenciar la zona turística de lo que es un destino turístico. En opinión de Esteve Secall (2002: 3), *el destino turístico es un ámbito geográfico inferior en tamaño a la zona, normalmente de carácter municipal, caracterizado por la especialización en un producto turístico, por atender a un segmento turístico determinado o a un tipo de turista atraído por una sola motivación.* Por ejemplo, la Alhambra, Lourdes o La Meca, por citar algunos, son visitados al año por un gran número de personas pero tienen un carácter monotemático que los excluye como zona turística.

Una vez visto el concepto de zona turística hay que resaltar la importancia que tiene para los diversos agentes que intervienen en la actividad turística alcanzar ventajas competitivas en la zona. Ello va a permitir obtener objetivos por encima de la media del sector, lo cual puede traducirse en rentabilidades financieras, sociales y ambientales superiores, consiguiendo la máxima satisfacción para los turistas (Esteve Secall, 2002). Por tanto, el estudio de las características estructurales de la oferta y de la demanda de las zonas turísticas es un aspecto muy importante a analizar ya que la existencia de diferencias en estas medidas puede provocar una competitividad diferente entre las diversas zonas turísticas.

LA SEGMENTACIÓN EN TURISMO

Una importante característica del turismo es su heterogeneidad. Esto es debido a la existencia de muchos tipos de zonas turísticas con diferentes ofertas y demandas, y de muchos tipos de turistas con distintos comportamientos relacionados con un amplio abanico de motivaciones (González & Bello, 2002).

La segmentación del mercado puede ser definida como un proceso a través del cual la gente (tanto proveedores de turismo como usuarios) con similares necesidades, deseos y características son agrupados de modo que cualquier organización turística puede usar una mayor precisión en servir y comunicarse con dichos grupos. Desde el punto de vista de la oferta y demanda de las zonas turísticas, la segmentación de zonas turísticas puede ser definida como una agrupación de zonas que comparten las mismas características estructurales de oferta y demanda, donde los elementos de cada grupo tienen el mismo potencial de estructura. En este sentido, las zonas turísticas de un mismo grupo compartirán el mismo atractivo para los turistas. De hecho, el territorio o zona turística supone un importante recurso turístico tanto desde el punto de vista político como desde el punto de vista empresarial (Fortunato, 2005).

Según Moreira, Nascimento & Segre (2010), en turismo es importante resaltar el potencial de las características locales en detrimento de la hegemonía global valorando las particularidades culturales y regionales. Así pues, tiene sentido clasificar a las zonas turísticas en diversos grupos para evitar un análisis global que puede provocar resultados erróneos.

La segmentación del mercado en turismo permite obtener algunas ventajas competitivas. Entre otros factores, esas ventajas son obtenidas porque los efectos de marketing pueden estar dirigidos a un segmento concreto de demanda (Dolnicar, 2007). Asimismo, la segmentación del mercado permite operaciones empresariales más viables.

En definitiva, en este trabajo se pretende realizar una segmentación de las zonas turísticas españolas para proporcionar a los agentes implicados una visión efectiva y concreta de las características estructurales de oferta y demanda de dichas zonas, con el fin de llevar a cabo las políticas adecuadas.

MODELOS DE CLASES LATENTES

Tradicionalmente la clasificación de objetos se ha llevado a cabo mediante el análisis cluster. Un problema importante de esta metodología es la ausencia de un criterio objetivo que guíe a los investigadores en la determinación de un número óptimo de grupos (Kaufman & Rousseeuw, 1990). Por ello, en los últimos años muchos investigadores han seleccionado el análisis de clases latentes como una metodología alternativa para la clasificación de datos, en virtud de que exige menos criterios arbitrarios para determinar el número de grupos de la población.

El análisis de clases latentes (ACL) es una herramienta estadística que permite modelar las relaciones entre las variables observadas, suponiendo que la estructura de relaciones subyacentes es explicada por una variable latente categórica (no observada). Esta metodología clasificatoria se basa en la estimación de probabilidades condicionadas, lo que permite analizar variables medidas en diferentes métricas, especialmente datos categóricos (Magidson & Vermunt, 2001 y 2004).

El ACL fue introducido por primera vez por Lazarsfeld (1950) como herramienta para construir una tipología en el análisis de un conjunto de variables dicotómicas. Más tarde, Lazarsfeld & Henry (1968) continuaron con el desarrollo de la técnica. Posteriormente, otros autores han desarrollado esta metodología, entre los que cabe citar a Goodman (1974), McCutcheon (1987), Hagenaars (1990), Andersen (1993), Agresti (2002), Bartholomew, Steele, Moustaki & Galbraith (2002) y Vermunt & Magidson (2002, 2003 y 2005), entre otros.

De acuerdo con Bartholomew, Steele, Moustaki & Galbraith (2002), en el análisis de clases latentes se asume que:

- ✓ Cada individuo de una muestra aleatoria pertenece sólo a una de las clases latentes detectadas.
- ✓ La probabilidad de dar una respuesta a un ítem particular es la misma para todos los individuos que se encuentran en la misma clase, pero diferente a la de los individuos que pertenecen a una clase diferente.
- ✓ Dada la pertenencia de un individuo a una clase latente, sus respuestas a cada uno de los ítems son condicionalmente independientes.

Supóngase que existe un conjunto de p variables manifiestas $X_1, \dots, X_p$ que se consideran indicadoras de una variable latente Y , y que estas variables conforman un modelo de clases latentes con C clases o categorías. Sea $\pi_X(x)$ la densidad conjunta de las variables manifiestas $X' = (X_1, \dots, X_p)$, donde $x' = (x_1, \dots, x_p)$ representa un determinado patrón de respuesta, en el cual cada x_i toma diferentes valores dependiendo de las categorías de la correspondiente variable manifiesta. Consideremos que cada X_i tiene I_i categorías, $i = 1, \dots, p$, por lo que las variables manifiestas conforman una tabla de contingencia de p -vías con $\prod_{i=1}^p I_i$ patrones de respuesta.

El supuesto fundamental en todo modelo de variables latentes es el principio de *Independencia Condicional o Local*, el cual establece que dado un valor fijo de la variable latente, las variables manifiestas son mutuamente independientes. Básicamente, este supuesto viene a señalar que toda la asociación observada entre las variables manifiestas está medida o explicada por las variables latentes (Sepúlveda, 2004). Por este principio se tiene que la probabilidad $P(X = x / Y = c)$ viene dada por:

$$\pi_{X/Y(c)}(x) = \prod_{i=1}^p \pi_{X_i/Y(c)}(x_i)$$

donde $\pi_{X_i/Y(c)}(x_i) = P(X_i = x_i / Y = c)$, $x_i = 1, \dots, I_i$; $c = 1, \dots, C$.

La distribución conjunta de X e Y , denotada por $\pi_{X,Y}(x,c)$ puede ser obtenida como $\pi_{X,Y}(x,c) = \pi_Y(c)\pi_{X/Y(c)}(x)$, donde $\pi_Y(c) = P(Y=c)$ representa la proporción de individuos que se encuentran en la clase latente c , $c = 1, \dots, C$, también conocida como probabilidad a priori.

A partir de aquí, el modelo de clases latentes puede ser escrito como:

$$\pi_X(x) = \sum_{c=1}^C \pi_{X,Y}(x,c)$$

o equivalentemente,

$$\pi_X(x) = \sum_{c=1}^C \pi_Y(c) \pi_{X/Y(c)}(x) = \sum_{c=1}^C \pi_Y(c) \prod_{i=1}^p \pi_{X_i/Y(c)}(x_i)$$

donde los parámetros están sujetos a las siguientes restricciones:

$$\sum_{c=1}^C \pi_Y(c) = 1$$

$$\sum_{x_i=1}^{I_i} \pi_{X_i/Y(c)}(x_i) = 1, i = 1, \dots, p; c = 1, \dots, C$$

Como se puede observar, el análisis de clases latentes es un modelo paramétrico que utiliza los datos observados para estimar los parámetros del modelo. En este caso, los parámetros del modelo son:

- a) La probabilidad de cada una de las clases latentes, $\pi_Y(c)$, $c = 1, \dots, C$.
- b) Las probabilidades de respuesta condicional de cada una de las variables manifiestas dentro de cada clase latente, $\pi_{X_i/Y(c)}(x_i)$, $i = 1, \dots, p$; $c = 1, \dots, C$; $x_i = 1, \dots, I_i$.

El análisis posterior del modelo de clases latentes está relacionado con lo que se puede decir acerca de los individuos que pertenecen a una clase determinada. Para ello se utiliza la distribución de Y condicionada a X , $\pi_{Y/X(x)}(c) = P(Y=c / X=x)$, denominada distribución a posteriori. En la práctica, cada individuo se clasifica en la clase latente en la cual esta probabilidad sea mayor.

La estimación de los parámetros en los modelos de clases latentes depende de la escala de medición de las variables observadas, ya que se asumen diferentes distribuciones para las variables nominales, ordinales y continuas. La estimación se hace mediante estimadores de máxima verosimilitud. El más conocido es el denominado algoritmo *EM* (Goodman, 1974; Hagenaars, 1990).

La evaluación del ajuste del modelo se realiza usualmente mediante el criterio de información bayesiana (*BIC*). Este estadístico pondera según el número de parámetros, la bondad del ajuste obtenido. Como regla para la selección del modelo que mejor se ajusta a los datos, se elige el modelo que presente un menor valor de *BIC* (Gill, 2002).

Hoy día, el desarrollo de software informático con la implementación de importantes algoritmos ha hecho que se haya extendido el uso del análisis de clases latentes.

En este estudio se ha utilizado el software Latent Gold 4.5 de la compañía Statistical Innovations (2010).

DATOS DEL ESTUDIO

Los datos de este trabajo han sido obtenidos de La Encuesta de Ocupación Hotelera, elaborada por el Instituto Nacional de Estadística (INE, 2010). Dicha encuesta sustituye desde enero de 1999 a la antigua Encuesta de Movimiento de Viajeros en Establecimientos Hoteleros (MVEH), ampliando la investigación a la categoría de una estrella y similares puesto que además de suponer más del 50 por ciento del total de establecimientos, representan más del 5 por ciento de entrada de viajeros. En concreto, se utilizan los datos detallados correspondientes al año 2008.

Las unidades de análisis son todos los establecimientos hoteleros inscriptos como tales en el correspondiente registro de las Consejerías de Turismo de cada Comunidad Autónoma. Son establecimientos hoteleros aquellos establecimientos que prestan servicios de alojamiento colectivo mediante precio con o sin otros servicios complementarios (hotel, hotel-apartamento o apartahotel, motel, hostel, pensión,...).

En este caso se va a utilizar como objeto para definir los conglomerados la zona turística. La zona turística se define como el conjunto de municipios en los que se localiza de forma específica la afluencia turística.

Las variables a utilizar en el estudio son aquéllas que están relacionadas con la estructura de la oferta y demanda de las zonas turísticas. En concreto, se han utilizado las siguientes variables:

- V1: Número medio de establecimientos abiertos al año.
- V2: Número medio de plazas al año.
- V3: Grado de ocupación medio al año.
- V4: Grado de ocupación medio por fines de semana al año.
- V5: Número medio de personal ocupado por meses.
- V6: Viajeros entrados al año.
- V7: Viajeros residentes en España.

V8: Viajeros no residentes en España.

V9: Primer país con más procedencia.

V10: Segundo país con más procedencia.

V11: Tercer país con más procedencia.

V12: Número de pernoctaciones al año.

Las variables V1 a V5 se han obtenido como medias al año dividiendo la suma de los valores mensuales por 12. De esta forma se ha evitado el efecto de la estacionalización. Para todas las variables, salvo V9, V10 y V11, se han definido tres modalidades (bajo, medio y alto), calculado el percentil 25 y el percentil 75, con el objetivo de tener una discriminación acentuada en sus valores.

Una definición de las variables que marcan el comportamiento de la oferta y demanda de las zonas turísticas se puede obtener en el Sitio Web del INE.

RESULTADOS

Una vez ejecutado el procedimiento mediante el software Latent Gold 4.5, se obtuvo en primer lugar el resumen de modelos (Tabla 1).

Tabla 1: Resumen de modelos (I)

	LL	BIC(LL)	Nº de parámetros	L ²	Grados de libertad	p-valor	Error
Modelo 1 1-Cluster	-537,7765	1199,231	34	801,8691	4	3,00E-172	0,0000
Modelo 2 2-Cluster	-419,9924	1058,240	60	566,3009	-22	0,0000	0,0003
Modelo 3 3-Cluster	-329,1088	971,050	86	384,5336	-48	0,0000	0,0000
Modelo 4 4-Cluster	-306,1432	1019,696	112	338,6024	-74	0,0000	0,0001
Modelo 5 5-Cluster	-289,8776	1081,7421	138	306,0712	-100	0,0000	0,0007

Fuente: Elaboración propia

El estadístico L^2 muestra la cantidad de asociación entre las variables que permanece no explicada después de la estimación del modelo. Por tanto, el valor más bajo de este estadístico es el que proporciona un mejor ajuste del modelo a los datos. En este caso se puede observar que el modelo 5 es el que presenta un valor del estadístico L^2 más bajo, sin embargo, el número de parámetros a estimar es el más elevado.

Un criterio para determinar el número de clusters es mirar el p-valor de cada modelo, bajo el supuesto de que el estadístico L^2 sigue una distribución Chi-cuadrado. Para ello se elige aquel modelo cuyo p-valor sea superior a 0,05 y con un menor número de parámetros a estimar. En este caso, todos los p-valores son inferiores a 0,05, por lo que no es conveniente utilizar este criterio.

Otra forma para determinar el modelo más adecuado es mirar el criterio de información bayesiana (BIC). Como ya se ha visto anteriormente, se elige aquel modelo que presente un menor BIC. Así pues, utilizando este estadístico, se selecciona el modelo 3, que es el que presenta el valor más bajo, además de que el número de parámetros estimados es mucho menor que en el modelo 5.

El programa Latent Gold ofrece una opción alternativa para determinar el modelo, utilizando un procedimiento bootstrap del estadístico L^2 para estimar el p-valor. Este procedimiento proporciona una estimación más precisa, relajando el supuesto de que el estadístico L^2 sigue una distribución Chi-cuadrado (Tabla 2).

Tabla 2: Ajuste del modelo utilizando el procedimiento bootstrap

Estadísticos Chi-cuadrado		Bootstrap		
Grados de libertad	-48	p-valor	p-valor	Error estándar
L-cuadrado (L^2)	384,5336	0,0000	0,2220	0,0186
X-cuadrado	163890,842	0,0000		
Cressie-Read	9664,8863	0,0000		
BIC (basado en L^2)	559,1377	0,0000		
AIC (basado en L^2)	480,5336	0,0000		
AIC3 (basado en L^2)	528,5336	0,0000		
CAIC (basado en L^2)	607,1377	0,0000		
Índice de disimilaridad	0,9799	0,0000		

Fuente: Elaboración propia

Se puede observar en la Tabla 2 que el p-valor asociado al estadístico L^2 vale 0,2220 con un error estándar de 0,0186, por lo que las estimaciones obtenidas con este procedimiento son mucho mejores que las anteriores. Así pues, se selecciona el modelo 3, que presenta 3 clases latentes.

Otra información que se puede obtener es la de los parámetros del modelo (Tabla 3). Los parámetros del modelo miden la influencia de cada variable predictora en la definición de cada cluster. Un p-valor inferior a 0,05 nos indica que la hipótesis de que todos los efectos asociados con la variable predictora debe ser rechazada. Se puede observar que en todas las variables predictoras se rechaza la hipótesis nula, salvo en las variables *primer país con más viajeros*, *segundo país con más viajeros* y *tercer país con más viajeros*.

La última columna señala el valor del coeficiente R^2 . Este coeficiente indica la proporción de varianza de cada variable predictora que es explicada por el modelo de 3 clusters.

Toda esta información exige el planteamiento de nuevo de la estimación de un modelo sin las tres variables anteriormente señaladas. Si se hace de nuevo una estimación de los parámetros, sin tener en cuenta las tres variables anteriores, se obtienen los resultados que aparecen en la Tabla 4. Se puede observar en este caso que el valor del estadístico L^2 ha disminuido con respecto a la primera estimación que se hizo. Sin embargo, al igual que ocurrió con la primera estimación, el p-valor es

inferior a 0,05 por lo que se rechaza la hipótesis de que el estadístico L^2 se distribuye según un modelo Chi-cuadrado.

Tabla 3: Parámetros estimados en el modelo

Modelos para los indicadores							
	Cluster1	Cluster2	Cluster3	Wald	p-valor	R²	
Nº medio de establecimientos abiertos al año	0,2202	1,7564	-1,9766	11,3899	0,0034	0,4034	
Nº medio de plazas al año	-1,0208	6,5412	-5,5204	19,707	5,30E-05	0,8561	
Grado de ocupación medio al año	-0,5069	2,5312	-2,0243	13,173	0,0014	0,4821	
Grado de ocupación medio por fines de semana al año	-0,7016	2,3667	-1,6651	11,9828	0,0025	0,4436	
Media de personal ocupado por meses	-0,001	7,6177	-7,6166	16,5234	0,00026	0,943	
Viajeros entrados al año	1,02	5,5189	-6,5389	19,7031	5,30E-05	0,856	
Viajeros residentes en España	0,9365	2,4498	-3,3863	14,566	0,00069	0,5756	
Viajeros no residentes en España	-0,001	7,6177	-7,6166	16,5234	0,00026	0,943	
Primer país con más procedencia							
Alemania	-1,0336	0,7052	0,3284	4,6931	0,58	0,0919	
Francia	-0,1425	-0,3992	0,5417				
Portugal	0,5095	-2,1011	1,5916				
Reino Unido	0,6666	1,7951	-2,4617				
Segundo país con más procedencia							
Alemania	-0,6820	-0,5023	1,1842	1,4182	0,96	0,028	
Francia	0,6652	0,6628	-1,328				
Italia	0,8349	0,4369	-1,2718				
Reino Unido	-0,8182	-0,5974	1,4156				
Tercer país con más procedencia							
Alemania	-0,1523	0,0091	0,1433	9,0289	0,98	0,0605	
Bélgica	-1,4430	1,5171	-0,0741				
Estados Unidos	0,2696	2,1596	-2,4292				
Francia	-1,4430	1,5171	-0,0741				
Irlanda	2,0866	-0,7828	-1,3038				
Italia	0,2674	-1,9037	1,6363				
Países Bajos	0,2697	-1,9082	1,6386				
Portugal	0,2697	-1,9082	1,6386				
Reino Unido	0,4592	-1,8891	1,4299				
Rusia	2,0866	-0,7828	-1,3038				
Suecia	-2,6704	3,9720	-1,3015				
Pernoctaciones	-1,0208	6,5412	-5,5204	19,707	5,30E-05	0,8561	
Modelo para los clusters							
Constante	Cluster1	Cluster2	Cluster3	Wald	p-value		
	0,3823	-0,1911	-0,1912	3,1942	0,2		

Fuente: Elaboración propia

Tabla 4: Resumen de modelos (II)

		LL	BIC(LL)	Nº de parámetros	L ²	Grados de libertad	p-valor	Error
Modelo 1	1-Cluster	-361,3489	788,1744	18	498,6078	20	5,70E-93	0,0000
Modelo 2	2-Cluster	-252,4615	606,7754	28	280,8329	10	1,70E-54	0,0000
Modelo 3	3-Cluster	-174,7949	487,8181	38	125,4997	0	0,0000	0,0000
Modelo 4	4-Cluster	-162,8241	500,2523	48	101,5581	-10	0,0000	0,0089
Modelo 5	5-Cluster	-153,6464	518,2728	58	83,2028	-20	0,0000	0,0202

Fuente: Elaboración propia

Al igual que se hizo con la primera estimación, si se utiliza el criterio BIC, se elige el modelo de 3 clases latentes. Toda esta información exige ejecutar de nuevo un procedimiento bootstrap para elegir el modelo, relajando la hipótesis de que el estadístico L^2 siga una distribución Chi-cuadrado. En este caso, el p-valor asociado al estadístico L^2 vale 0,2860 con un error estándar de 0,0202, por lo que obtenemos mejores estimaciones.

La Tabla 5 muestra en primer lugar el tamaño de cada cluster. Se puede observar que el cluster 1 está formado por un 47,01% del total de zonas turísticas, mientras que el cluster 2 está formado por un 26,50% de zonas turísticas y el cluster 3 por un 26,49%. También se muestran las probabilidades condicionadas de cada cluster en función de cada variable predictora. Así, se puede observar que el cluster 1 se caracteriza por ser zonas turísticas donde el número medio de establecimientos abiertos al año se encuentra en un nivel medio. El cluster 2 se caracteriza por tener un número medio de establecimientos abiertos al año alto, mientras que el 3 se caracteriza por tener un número bajo. En relación al número medio de plazas al año, el cluster 1 se caracteriza por tener un nivel medio, mientras que los clusters 2 y 3 se caracterizan por tener un nivel alto y bajo, respectivamente. De igual forma se pueden interpretar el resto de variables. En todas las variables analizadas, las probabilidades condicionadas son muy altas en la categoría nivel medio para el cluster 1, nivel alto para el cluster 2 y nivel bajo para el cluster 3, lo cual hace que existan diferencias significativas en los distintos grupos. La variable en la que existen menos diferencias es la correspondiente al grado de ocupación medio por fines de semana al año. Se puede observar que la probabilidad de tener un grado de ocupación bajo en esta variable es muy similar para los clusters 1 y 3.

Tabla 5: Probabilidades condicionadas

		Cluster1	Cluster2	Cluster3
Tamaño del cluster		0,4701	0,2650	0,2649
Indicadores				
Nº medio de establecimientos abiertos al año				
	Bajo	0,2971	0,0000	0,7029
	Medio	0,6077	0,2242	0,1681
	Alto	0,3965	0,6035	0,0000
Nº medio de plazas al año				
	Bajo	0,0987	0,0000	0,9013

	Medio	0,9437	0,0000	0,0563
	Alto	0,0000	1,0000	0,0000
Grado de ocupación medio al año				
	Bajo	0,3965	0,0000	0,6035
	Medio	0,6635	0,1122	0,2243
	Alto	0,1977	0,8023	0,0000
Grado de ocupación medio por fines de semana al año				
	Bajo	0,4964	0,0000	0,5036
	Medio	0,6077	0,1121	0,2802
	Alto	0,1977	0,8023	0,0000
Media de personal ocupado por meses				
	Bajo	0,0001	0,0000	0,9999
	Medio	1,0000	0,0000	0,0000
	Alto	0,0000	1,0000	0,0000
Viajeros entrados al año				
	Bajo	0,0001	0,0000	0,9999
	Medio	0,9437	0,0563	0,0000
	Alto	0,0987	0,9013	0,0000
Viajeros residentes en España				
	Bajo	0,0987	0,0000	0,9013
	Medio	0,7194	0,2245	0,0561
	Alto	0,3965	0,6035	0,0000
Viajeros no residentes en España				
	Bajo	0,0001	0,0000	0,9999
	Medio	1,0000	0,0000	0,0000
	Alto	0,0000	1,0000	0,0000
Pernoctaciones				
	Bajo	0,0987	0,0000	0,9013
	Medio	0,9437	0,0000	0,0563
	Alto	0,0000	1,0000	0,0000

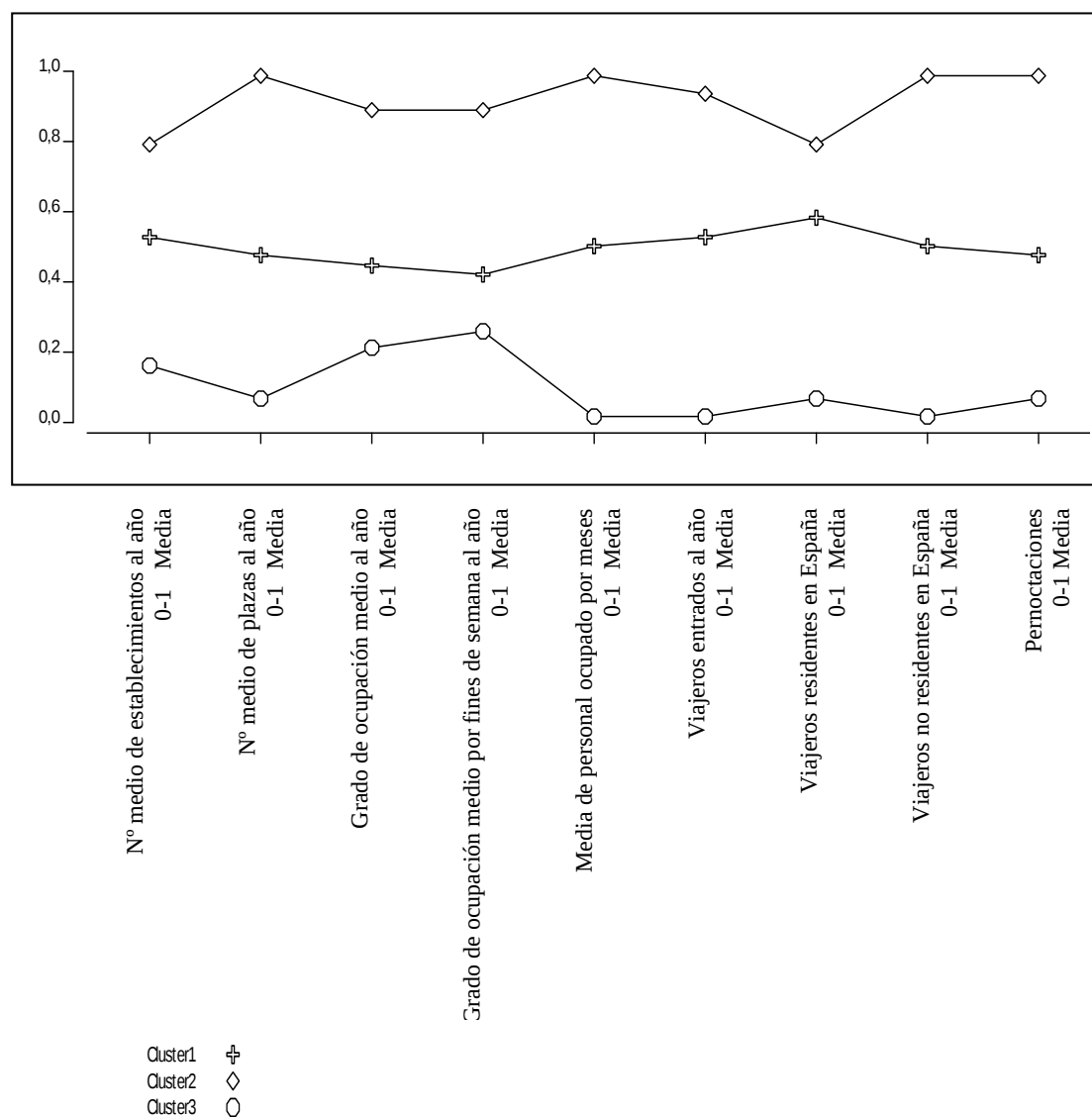
Fuente: Elaboración propia

Estas diferencias entre los grupos se pueden observar también de forma visual en el Gráfico 1. Se observa que la línea del cluster 1 se encuentra entre las líneas del cluster 2 (más elevada) y la línea del cluster 3 (más baja). El cluster 1 destaca fundamentalmente por tener valores elevados en el número medio de plazas al año, en la media de personal ocupado por meses, en los viajeros no residentes en España y en el número de pernoctaciones. En el lado opuesto se encuentra el cluster 3, con los valores más bajos en la media del personal ocupado por meses, en los viajeros entrados al año y en los viajeros no residentes en España.

De forma general, se observan grandes diferencias dentro de cada variable en función de cada cluster. Las diferencias más pequeñas se encuentran en el *grado de ocupación medio por fines de semana al año* entre los clusters 1 y 3, así como el *número de viajeros residentes en España* entre los clusters 1 y 2.

Finalmente, el programa permite obtener una clasificación de los destinos turísticos en cada uno de los grupos (Tabla 6).

Gráfico 1: Perfiles para el modelo de 3 clases latentes



Fuente: Elaboración propia

Tabla 6: Clasificación de las zonas turísticas

Clúster 1	Costa de Almería, Costa de la Luz de Cádiz, Costa de la Luz de Huelva, Costa Verde (Asturias), Isla de Menorca, Islas de Ibiza-Formentera, Isla de Lanzarote (Las Palmas), Isla de Fuerteventura (Las Palmas), Costa del Garraf (Barcelona), Costa del Maresme (Barcelona), Costa Daurada (Tarragona), Costa Azahar (Castellón), Rías Altas (A Coruña), Rías Baixas (Pontevedra y A Coruña), Costa Cálida (Murcia), Costa Guipuzcoana, Pirineo Aragonés y Pirineo Catalán.
-----------	--

Clúster 2	Costa del Sol (Málaga), Isla de Mallorca, Palma-Calviá, Isla de Gran Canaria (Las Palmas), Sur de Gran Canaria (Las Palmas), Isla de Tenerife (Tenerife), Sur de Tenerife (Tenerife), Costa de Barcelona (Barcelona), Costa Brava (Girona) y Costa Blanca (Alicante).
Clúster 3	Costa Tropical (Granada), Isla de la Gomera (Tenerife), Isla de la Palma (Tenerife), Isla de Hierro (Tenerife), Costa Valencia (Valencia), Costa Da Morte (A Coruña), Costa A Mariña Lucense (Lugo), Costa Vizcaína, Pirineo Navarro y Pirineo Vasco.

Fuente: Elaboración propia

CONCLUSIONES

En este trabajo se han analizado la oferta y la demanda turística de las zonas turísticas existentes en España. El objetivo que ha perseguido este trabajo es el de clasificar a las zonas turísticas en grupos homogéneos dentro de sí y heterogéneos entre ellos. Un mayor conocimiento sobre la existencia de distintos grupos de zonas turísticas permite una asignación de recursos más eficiente y eficaz, tanto por parte de las empresas integradas en el sector como por parte de las Administraciones Públicas responsables de las políticas en materia de turismo en dichas zonas, en concreto a las Comunidades Autónomas competentes en materia turística.

El estudio pone de relieve la existencia de tres grupos diferenciados de zonas turísticas, en función de la oferta y demanda que presentan cada una de ellas. Según los resultados obtenidos, se puede clasificar de forma diferenciada a las zonas turísticas en tres grupos, en función de aquellas zonas turísticas que presentan unos valores de oferta y demanda bajos, medios y altos. Para ello se ha utilizado de un modelo de clases latentes.

Las zonas turísticas que presentan unas características estructurales de la oferta y la demanda más elevadas son Costa del Sol (Málaga), Isla de Mallorca, Palma-Calviá, Isla de Gran Canaria (Las Palmas), Sur de Gran Canaria (Las Palmas), Isla de Tenerife (Tenerife), Sur de Tenerife (Tenerife), Costa de Barcelona (Barcelona), Costa Brava (Girona) y Costa Blanca (Alicante). Existe un amplio grupo de zonas turísticas que se encuentran en un nivel medio y, finalmente, existe un grupo con unos niveles bajos en relación a la oferta y demanda.

Una conclusión importante que podemos obtener de la clasificación obtenida es que no existe una uniformidad geográfica en la situación de las zonas turísticas similares, es decir, no existe área geográfica concreta que se diferencie del resto de áreas geográficas. Por ejemplo, en Andalucía hay zonas turísticas con características de oferta y de demanda superiores, como la Costa del Sol, con características medias, como la Costa de Almería, y con características inferiores, como la Costa Tropical de Granada. Por tanto, los agentes implicados deberían considerar a las zonas turísticas con niveles más bajos de oferta y demanda, no como una zona en la que no invertir, sino como una

oportunidad de mercado. También, relacionado con todo esto, hay que decir que todas las zonas turísticas que configuran el cluster 2 se encuentran en la costa mediterránea y en el archipiélago canario. Así pues, los agentes implicados deberían hacer un esfuerzo en otras áreas geográficas, como la zona del mar Cantábrico o la zona del Atlántico para conseguir zonas turísticas de alto nivel.

Otra conclusión importante es que para los clusters 2 y 3 el nivel medio de viajeros es similar tanto para los residentes en España como para los no residentes en España. Sin embargo, el cluster 2 se caracteriza por tener una media de viajeros no residentes en España superior a la de los viajeros residentes en España. En este sentido, los esfuerzos de marketing de los agentes implicados deben ir dirigidos hacia una mayor atracción de los turistas nacionales hacia las zonas turísticas que configuran el cluster 2.

En definitiva, este trabajo permite a los agentes implicados obtener un panorama clarificador de la realidad existente entre las zonas turísticas españolas en función de la oferta y demanda que presenta cada una de ellas.

Las variables utilizadas en este estudio contienen información facilitada por los establecimientos hoteleros al Instituto Nacional de Estadística de España. La justificación de su utilización está en la fácil accesibilidad hacia dicha información. Por tanto, la clasificación de las zonas turísticas está limitada por este conjunto de variables. En este sentido, los resultados deben ser mirados desde la óptica de la oferta y de la demanda, por lo que deben ser tenidos en cuenta de forma precavida, ya que, aparte de estas variables económicas, se podrían haber utilizado otras variables desde el punto de vista medioambiental y/o social, lo que podría haber dado como resultado otra clasificación totalmente diferente de la encontrada en este estudio.

La realización de este trabajo abre las puertas para la realización de otros estudios en los que se puedan analizar otras variables de las zonas turísticas, si bien es complicada la obtención de los datos. También sería interesante estudiar la oferta y la demanda turística para la obtención de una clasificación sobre el mapa de provincias españolas, en vez de zonas turísticas.

REFERENCIAS BIBLIOGRÁFICAS

- Agresti, A.** (2002) "Categorical Data Analysis". Wiley-Interscience, New York
- Albadalejo, I. y Morera, L.** (2006) "Profiling segments of tourists in rural areas of South-Eastern Spain". *Tourism Management*, 28: 757-767
- Alonso Ferreras, V. H.** (2010) "Factores críticos de éxito y evaluación de la competitividad de destinos turísticos". *Estudios y Perspectivas en Turismo*, 19 (2): 201-220
- Andersen, E. B.** (1993) "The analysis of categorical data". Springer-Verlag, Berlin
- Bartholomew, D. J.; Steele, F.; Moustaki, I. & Galbraith, J. I.** (2002) "The Analysis and Interpretation of Multivariate Data for Social Scientists". Chapman & Hall, New York

- Dávila, D.; Rodríguez, S. & Rodríguez, A.** (2002) "La promoción turística espacial dentro de un destino vacacional homogéneo. Una aplicación a Canarias". *Estudios de Economía Aplicada*, 201(1): 45-60
- Dolnicar, S.** (2007) "Market segmentation in tourism". In A. Woodside y D. Martin (Eds.), *Tourism management, analysis, behaviour and strategy*. CABI, Cambridge, pp. 129-150
- Esteve Secall, R.** (2002) "La competitividad de las zonas turísticas". Ponencia en la Facultad de CC. EE. y EE. de la Universidad de Málaga, Málaga
- Eymann, A. & Ronning, G.** (1997) "Microeconomic models of tourists' destination choice". *Regional Science and Urban Economics*, 27(6): 735-761
- Ferrerira, S. D.; Rial, A. & Varela, J.** (2010) "Segmentación post hoc del mercado turístico español. Aplicación del análisis cluster en dos etapas". *Estudios y Perspectivas en Turismo*, 19 (5): 592-606
- Fodness, D. y Murray, B.** (1998) "A typology of tourist information search strategies". *Journal of Travel Research*, 37(2): 108-119
- Fortunato, N.** (2005) "El territorio y sus representaciones como recurso turístico. Valores fundacionales del concepto de parque nacional". *Estudios y Perspectivas en Turismo*, 14 (4): 314-348
- Gill, J.** (2002) "Bayesian methods: A social and behavioral sciences approach". CRC Press, Boca Raton
- González, A.** (1999) "Estilos de vida y segmentación del turismo". *Actas del Congreso de Turismo, Universidad y Empresa*. Benicassim, 1998: Organización administrativa, calidad de servicios y competitividad empresarial, Ed. Tirant lo Blanch, Valencia, pp. 621-636
- González, A.; Sánchez, I. y Sanz, S.** (2006) "Caracterización del turista de sol y playa: una comparativa con el turista rural y cultural". *Actas del IX Congreso de Turismo, Sol, Playa y Turismo Residencia*, Alicante
- González, A. y Bello, L.** (2002) "The construct lifestyle in market segmentation: The behaviour of tourist consumers". *European Journal of Marketing*, 36(1/2): 51-85
- Goodman, L. A.** (1974) "Exploratory latent structure analysis using both identifiable and unidentifiable models". *Biometrika*, 61(2): 215-231
- Gutiérrez, T. y Montero, I.** (2000) "La emergencia de la cultura en la demanda del producto turístico en los destinos naturales". *Actas del Congreso de Turismo, Universidad y Empresa*. Benicassim, 1999: Comercialización de productos, gestión de organizaciones, aeropuertos y protección de la naturaleza. Tirant Lo Blanch, Valencia, pp. 163-180
- Hagenaars, J. A.** (1990) "Categorical longitudinal data: loglinear analysis of panel, trend and cohort data". Sage Publications, Newbury Park
- INE** (2010) Instituto Nacional de Estadística. www.ine.es (Acceso el 18 de diciembre de 2010)
- Jurowski, C. y Reich, A.** (2000) "An explanation and illustration of cluster analysis for identifying hospitality market segments". *Journal of Hospitality and Tourism Research*, 24(1): 667-691
- Kaufman, L. y Rousseeuw, P. J.** (1990) "Finding groups in data: an introduction to cluster analysis". John Wiley and Sons, Inc, New York

- Lazarsfeld, P. F.** (1950) "The logical and mathematical foundation of latent structure analysis and the interpretation and mathematical foundation of latent structure analysis". En Stouffer, S.A. et al. (eds.), *Measurement and prediction*. Princeton University Press, pp. 362–472
- Lazarsfeld, P. F. y Henry, N. W.** (1968) "Latent Structure Analysis". Houghton Mill, Boston
- Magidson, J. y Vermunt, J. K.** (2001) "Latent class factor and cluster models, bi-plots and related graphical displays". *Sociological Methodology*, 31(1): 223–264
- Magidson, J. & Vermunt, J. K.** (2004) "Latent class models". En D. Kaplan (ed.), *The Sage handbook of quantitative methodology y for the social sciences*. Thousand Oakes: Sage Publications, pp. 175-198
- McKercher, B.** (2002) "Towards a classification of cultural tourists". *International Journal of Tourism Research*, 4(1): 29-38
- McCutcheon, A. L.** (1987) "Latent Class Analysis". Sage Publications, Newbury Park
- Moreira, M.; Nascimento, R. & Segre, L.** (2010) "¿Cuál es el papel del turismo en el desarrollo local? Un análisis crítico del cluster turístico de Santa Teresa – RJ, Brasil". *Estudios y Perspectivas en Turismo*, 19 (5): 812-834
- Palacio, V. & McCool, S. F.** (1997) "Identifying ecotourists in Belice through benedit segmentation: a preliminary analysis". *Journal of Sustainable Tourism*, 5: 234-243
- Picón, E.; Varela, J. y Real, E.** (2003) "Clasificación y segmentación post-hoc mediante el análisis de conglomerados". En Levy, J. P. y Varela, J., *Análisis multivariable para las ciencias sociales*. Pearson Educación, Madrid pp. 417-449
- Ravelo, T.; Moreno, M. C.; Oreja, J. R. & Deoula, M.** (2007) "Evaluación de la capacidad de atracción de las principales zonas turísticas de Tenerife. Una aplicación del modelo Rasch". IX Reunión de Economía Mundial, Madrid
- Raya, P. & Caparrós, A.** (2008) "La segmentación de la demanda turística en Andalucía". En *Informe Anual del Turismo en Andalucía 2008*. Consejería de Comercio, Turismo y Deporte de la Junta de Andalucía, pp. 383-445
- Rodríguez, P.** (2006) "La importancia del turismo deportivo en la naturaleza de España". En Moyano, E. & Moscoso, Y. (dir.): *Deporte y desarrollo rural*. Instituto Andaluz del Deporte, Málaga, pp. 181-194
- Roig, B.; Rodríguez, J. E. & Buitrago, J. M.** (2005) "Turismo rural: segmentación del mercado turístico español". En www.infoagro.com (Acceso el 22 de marzo de 2011)
- Sepúlveda, R.** (2004) "Contribuciones al análisis de clases latentes en presencia de dependencia local". Tesis Doctoral. Departamento de Estadística Aplicada. Universidad de Salamanca, Salamanca
- Statistical Innovations Inc.** (2010) *Latent Gold 4.5*
- Trujillo, F.; López, P. & Alarcón, P.** (2002) "Imagen turística de las Comunidades Autónomas españolas ante una demanda segmentada". *Estudios de Economía Aplicada*, 20(3): 627-649
- Vermunt, J. K. & Magidson, J.** (2002) "Latent class cluster analysis". En J. Hagenaars & A. McCutcheon (eds.). *Applied Latent Class Models*. Cambridge University Press, New York, pp. 89-106
- Vermunt, J. K. & Magidson, J.** (2003) "Addendum to Latent gold user's G guide: Upgrade for Version 3". Statistical Innovations Inc, Boston

Vermunt, J. K. & Magidson, J. (2005) "Technical guide for Latent gold 4.0: Basic and advanced". Statistical Innovations, Inc, Boston

Recibido el 03 de mayo de 2011

Correcciones recibidas el 17 de junio de 2011

Aceptado el 24 de junio de 2011

Arbitrado anónimamente