

VIDAL DÍAZ DE RADA IGÚZQUIZA (*)

Diseño de tipologías de consumidores mediante la utilización conjunta del Análisis Cluster y otras técnicas multivariantes

1. INTRODUCCIÓN

Este trabajo analiza los resultados proporcionados por la técnica estadística *Análisis de Cluster* cuando ésta se aplica no a datos directos sino a puntuaciones factoriales. Entre otras razones, una de las ventajas de analizar las puntuaciones factoriales es su facilidad para conseguir que los datos cumplan determinados requisitos *necesarios* para utilizar el *Análisis de Cluster*: requisitos como la no existencia de un gran número de variables, igualdad de medida, y la existencia de incorrelación entre ellas; aspectos que serán expuestos con más detalle en las páginas siguientes. Asimismo, la utilización de otras técnicas estadísticas como el Análisis de Varianza permitirá comprobar si todas las variables utilizadas en la definición de los clusters producen diferencias significativas entre ellos, al tiempo que conoceremos –mediante el Análisis Discriminante– en qué medida los individuos se inscriben adecuadamente en cada uno de los *clusters* seleccionados.

La exposición va a estar dividida en 4 partes que comienzan con la realización de una pequeña explicación del procedimiento utilizado para reducir la información proporcionada por 28 variables a un conjunto menor de factores, y continúa

(*) Dpto. de Sociología. Universidad Pública de Navarra.

con una exposición más detallada sobre la técnica del Análisis de Cluster. Posteriormente se procederá a validar los resultados obtenidos, para realizar a continuación una interpretación de los mismos.

Con objeto de realizar las comprobaciones señaladas en el párrafo anterior utilizaremos datos obtenidos de un estudio más amplio sobre hábitos de consumo en Navarra basado en una muestra aleatoria estratificada por afijación proporcional de 552 entrevistas (1). En concreto, nos centraremos en tres preguntas que tratan de conocer comportamientos generales de compra y que hacen referencia a los elementos que más influyen en la decisión de comprar, los motivos por los cuales «los otros» compren productos de marca, y el grado de acuerdo con diversos comportamientos de compra (2). Así, la clasificación de los consumidores navarros en cinco agrupaciones proporciona una adecuada interpretación de la realidad del consumidor navarro, realidad que podría ser extrapolable a otras Comunidades Autónomas de características similares, al tiempo que nos permite explicar con mayor facilidad la combinación de técnicas estadísticas realizada en este trabajo.

1. LA UTILIZACIÓN DEL ANÁLISIS FACTORIAL DE COMPONENTES PRINCIPALES

El Análisis Factorial es un método de análisis multivariante que intenta reducir, a través de un modelo lineal, un conjunto extenso de variables observables a número menor de variables hipotéticas llamadas Factores (Kim y Mueller, 1978: p. 191), tratando de extraer la mayor cantidad posible de la variabilidad total. Este método fue planteado en primer lugar por Pearson (1901) y posteriormente desarrollado por Hotelling (1933). El modelo de Componentes Principales se refleja en la siguiente expresión:

$$Z_{ik} = F_{1k} * A_{i1} + F_{2k} * A_{i2} + \dots + F_{mk} * A_{in} \quad (k = 1, \dots, n).$$

(1) A un nivel de confianza del 95 por ciento y un valor «p» de 0,5 corresponde un error muestral para datos globales del 4 por ciento.

(2) La formulación exacta de las preguntas analizadas puede consultarse en el anexo a este trabajo.

Donde Z_{ik} representa la puntuación del individuo k en la variable i , F_{1k} la puntuación del sujeto k en el factor 1 y A_{i1} el peso factorial de la variable i en el factor 1. Eludiremos la explicación técnica del Análisis Factorial –remitiendo al lector interesado la bibliografía expuesta al final de este trabajo– y únicamente analizaremos algunas pautas utilizadas en la elección del número de factores, considerando para ello los criterios enunciados por E. Diday (1982: pp. 186-187):

- 1) Fijar un porcentaje de varianza y tomar los factores que lleguen a explicarlo. En una línea similar hay autores que sugieren observar el porcentaje de varianza explicado por cada factor y seleccionar aquellos cuyo porcentaje acumulado sea relativamente alto (Bisquerra, 1989: p. 308).
- 2) Tomar los factores que expliquen una proporción de varianza por encima de la expresión:

$$K * (100/p)$$
donde k es una constante que suele tomar el valor de 2 ó 3, y p el número total de variables.

- 3) Representar en ordenadas los autovalores y en abcisas el número de Componentes Principales e ir seleccionando los factores hasta que se obtenga un «codo» en la representación gráfica.

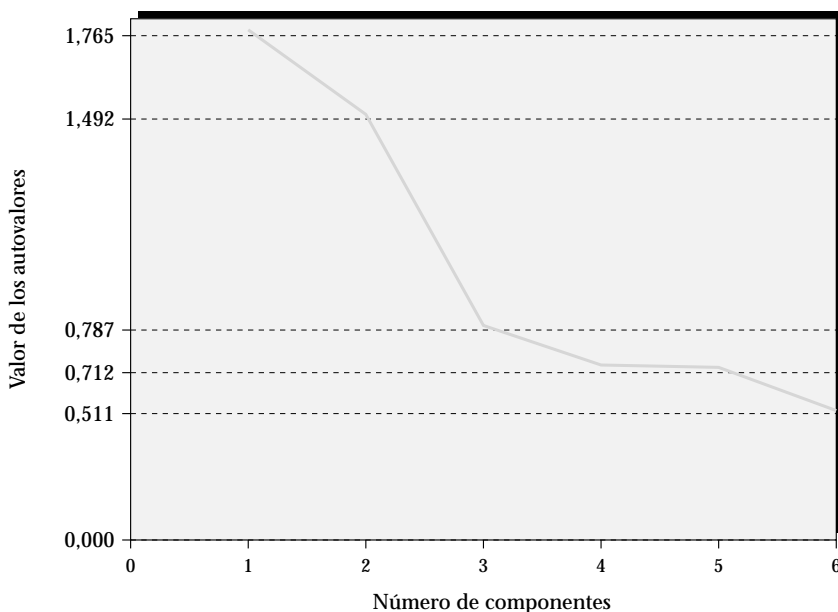
Con objeto de explicar mejor este criterio hemos representado en el gráfico 1 el análisis factorial realizado con la pregunta 2 (cuadro 2). De los seis autovalores se han seleccionado los dos primeros ya que la línea que los une sufre una gran desviación cuando llega al tercer factor.

Seleccionado el número de factores, y después de aplicarles una rotación Varimax, se ha procedido a calcular la puntuación de los entrevistados en cada uno de ellos. Expuestos brevemente los principios metodológicos de la técnica del Análisis Factorial, procederemos a interpretar los resultados obtenidos mediante su aplicación.

1.1. Conductas de compra

En la primera pregunta (consúltese Anexo I) se han recogido distintos aspectos que forman parte de los procesos deci-

Gráfico 1



sorios utilizados en la elección de los productos. Debe considerarse que en ningún momento se ha hecho referencia a productos concretos y, por lo tanto, las respuestas de los entrevistados no distinguen entre compras realizadas bajo procesos de decisión simple o compleja, sino que son respuestas espontáneas surgidas del recuerdo de los últimos productos comprados, y que reflejan un hábito de compra «asentado en la costumbre».

Tras aplicar el Análisis Factorial y extraer tres factores que explican el 51,2 por ciento de la varianza total, procedemos a interpretar el primer factor cuyas variables muestran la existencia de una «*Racionalidad culta*» caracterizada por un alto grado de información y reflexión ante de la compra («Seguir los consejos del dependiente, Comparar precios.... Preguntar a los amigos..., y relación precio-calidad»).

El factor dos refleja una «*racionalidad de vida cotidiana*», más espontánea, más forjada y fundamentada en el hábito diario: Si bien hay una preocupación y un análisis del producto comprado, este análisis es menos reflexivo que en el caso

Cuadro 1

PREGUNTA 1:
CUANDO REALIZA COMPRAS ¿CUÁNTO INFLUYE EN SU DECISIÓN...?

	Factor 1	Factor 2	Factor 3
Comp. precios	0,75726	0,06492	-0,10468
Infor. Dependien	0,73687	-0,10739	0,12024
Preg. amigos	0,53228	0,17255	0,15443
Precio-calidad	0,51372	0,24255	0,12773
Prod. Conocid	0,06046	0,75313	-0,04585
Análisis etiq	-0,01457	0,64655	0,29700
Ofertas-rebaja	0,20096	0,60100	0,00042
Prod. moda	-0,01722	0,15056	0,77459
Prod. anuncian	0,26398	-0,02857	0,77649
Valor Popio	2,22	1,27	1,11
Porcen Varianza	24,70	14,20	12,30
Porcen Acum.	24,70	38,90	51,20

anterior y se realiza frecuentemente en el mismo lugar de compra.

La diferenciación fundamental entre ambos factores se localiza en las diferentes predisposiciones (activa o pasiva) hacia la recogida de información. Frente al análisis de la etiqueta, comprar productos conocidos y aprovechar ofertas y rebajas (predisposición pasiva), otras conductas como «acudir a los amigos para conocer los productos que compran», realizar «análisis comparativos entre los diferentes precios», etc. implican una actitud distinta, mucho más comprometida y con un mayor tiempo invertido en ellas. (Shaughnessy, 1989: pp. 39-57 y Vazquez Casielles, 1989: pp. 55-72).

El tercer factor, con un poder explicativo menor, hace referencia a una «conducta no racional» que considera los anuncios de los medios de comunicación y la influencia de la moda como patrones condicionantes en la realización de sus compras.

1.2. Los motivos de comprar «marcas prestigiosas»

Analizadas las actitudes ante ciertos comportamientos de compra, el objetivo siguiente es conocer cuáles son las razones que cada entrevistado cree que conducen a los demás a comprar estas marcas. Realizando el mismo proceso que en la pre-

Cuadro 2

PREGUNTA 2:
MOTIVOS POR LOS CUALES LA GENTE COMPRO «MARCAS» PRESTIGIOSAS

	Factor 1	Factor 2
Comodidad	0,70784	0,12165
Compra menos tiempo	0,66021	-0,16197
«Individualidad»	0,63588	0,41428
Se fían más	0,61041	-0,21976
Diferenciarse otros	0,06119	0,83188
Presumir	-0,21704	0,73425
Valor Popio.....	1,76471	1,49210
Porcen Varianza	29,4	24,9
Porcen Acum.	29,4	54,3

gunta anterior extraemos dos factores que explican el 54,3 por ciento de la varianza.

El primero de ellos destaca la comodidad como uno de los criterios más importantes en la elección de las marcas, y está asociado a otras conductas como realizar la compra en menos tiempo, la individualidad que proporciona la marca y la confianza en marcas conocidas. En este sentido, este factor está reflejando los aspectos utilitarios de la marca, unido a una cierta individualidad e identificación (Bello Acebrón y otros, 1994: pp. 18-129). El segundo factor hace referencia a comprar «marcas» para diferenciarse del resto, la posesión de objetos de «marca» a fin de identificar socialmente a su portador y excluirlo de otros grupos, es un aspecto que hace referencia a un consumo conspicuo; la marca como distinción social, identificando y excluyendo a su portador respecto a otros grupos sociales (Blas, 1990: pp. 19-20).

1.3. Actitudes de los consumidores

En la pregunta tres se solicita del entrevistado que manifieste su opinión sobre un conjunto de comportamientos sobre distintos «estadios» del proceso de consumo: Se muestran comportamientos de compra (decido sobre la marcha, me gusta conocer tiendas nuevas, etc.), hay items que hacen referencia a la consideración de la moda y al sentimiento de exclusividad de la ropa, otro grupo cuestiona el status que se adquiere al poseer ciertos productos, etc.

Extraídos cuatro factores, este análisis explica el 53,5 por ciento de la varianza total. El primer factor muestra el *placer* experimentado en la *realización de las compras*, la compra como un *acto lúdico y placentero* (Vázquez Casielles, 1989: pp. 56-59). (Me gusta conocer tiendas nuevas, me gusta acompañar a alguien que va de compras, es muy frecuente que compre cosas que no tenía pensadas). Muy relacionado con este placer proporcionado por el acto de compra está la consideración e importancia de la moda a la hora de elegir las prendas de vestir, así como la idea de exclusividad e individualidad que confiere la ropa con la que cada uno viste.

El factor dos refleja un comportamiento basado y fundamentado en la *calidad de las compras*, la calidad adquiere más importancia que los precios, las marcas son una garantía; las diferencias entre las marcas se notan mucho.

El tercer factor muestra el *status que los productos poseídos* confieren al poseedor; una casa elegante dice mucho de los que en ella viven, gusto por ir bien vestido puesto que es muy importante la impresión que doy a los demás, el coche puede mostrar el status de quien lo lleva. Analizando la importancia que los españoles conceden al modo de vestir, Amando de Miguel expone que en España «nadie quiere ser menos que otro en las apariencias», y relaciona este hecho con «la idea de la representación escénica que de continuo muestran los españoles»: el análisis de la elevada proporción de ingresos que se gasta en ropa, así como la enorme importancia que se le da al vestir, «revela el alto valor que concede la cultura española a las apariencias» (1990: p. 77). A la vez, estas conductas están asociadas con una compra basada en la *reflexión* y en el análisis de los productos al no decidir sobre la marcha los objetos que se necesitan.

El factor número cuatro hace referencia a la idea que la gente compra en *consonancia con el ambiente social* en el que vive, a la vez que se enfatiza la *calidad de vida* como el motivo principal para realizar compras.

2. ANÁLISIS DE CLUSTER

El término «Análisis de cluster» (o conglomerados) define una gran diversidad de técnicas y métodos cuyo fin es lograr

Cuadro 3

PREGUNTA 3:
COMPORTAMIENTOS Y ACTITUDES DE COMPRA

	Factor 1	Factor 2	Factor 3	Factor 4
Conocer tiendas nuevas ...	0,774	-0,60	-0,150	-0,081
Acompañar a alguien	0,701	-0,026	-0,073	-0,111
Vestir a la moda	0,609	0,442	-0,284	-0,102
Cosas no pensadas	0,607	-0,007	0,278	-0,327
Molesta ver ropa igual	0,545	0,254	-0,159	0,191
Calidad más que precio ...	-0,056	0,753	0,153	0,027
Diferencias Marcas	0,087	0,641	-0,116	-0,186
Invertir en una casa.....	0,152	0,086	-0,689	-0,206
Decido marcha compras ..	0,231	0,206	0,564	-0,445
Bien vestido impresión	0,376	0,414	-0,471	-0,155
Coche indica prestigio	0,303	-0,064	-0,460	-0,329
Compra según ambiente .	0,111	-0,038	-0,070	-0,661
Vida más agradable	-0,099	0,241	-0,190	-0,656
Valor Propio	2,47	1,53	1,52	1,45
Porcen Varianza	18,98	11,74	11,67	11,14
Porcen Acum.	18,98	30,72	42,39	53,53

Cuadro 4

DENOMINACIÓN DE CADA UNO DE LOS FACTORES

Racionalidad culta
Racionalidad de vida cotidiana
Conducta no racional

Elección de marcas atendiendo a aspectos utilitarios
Elección de marcas porque indican status

Placer en el acto de compra
Calidad en las compras
Consumo conspicuo
Importancia del ambiente social y la calidad de vida

una clasificación o agrupamiento de n individuos en g grupos según el comportamiento de éstos en una serie de variables. A diferencia de otras técnicas multivariantes basadas en criterios de dependencia e independencia entre las variables, las técnicas de cluster consideran a todas las variables como intervinientes y conceden a todas ellas la misma función en la configuración de los grupos. De modo que el fin del Análisis de Cluster no es explicar estadísticamente ningún fenómeno

concreto sino clasificar una población –definida por unas características determinadas– en un reducido número de grupos mutuamente exclusivos y exhaustivos (Sánchez Cuenca, 1990: pp. 524-539).

El origen de estas técnicas hay que situarlo a principios de los años 60 en el campo de las ciencias naturales. La publicación del libro de R. Sokal y P. Sneath *Principles of Numerical Taxonomy* (1963) es considerado como el inicio formal de la materia, aunque muy pronto traspasaron el campo de las ciencias naturales y fueron aplicadas con éxito a otros campos como la sociología empírica, la medicina, la psiquiatría, la psicología, etc. Debido a la gran cantidad de publicaciones aparecidas en los últimos 25 años es muy difícil realizar un análisis exhaustivo de todos los métodos y procedimientos desarrollados, de modo que nos limitaremos a la revisión de los métodos más utilizados en las Ciencias Sociales.

3. ETAPAS EN UN ANÁLISIS DE CLUSTER

Mark S. Aldenderfer y Roger K. Blashfield (1980: p. 12) establecen cuatro etapas básicas en la realización del Análisis de Cluster:

- Estudio pormenorizado de las variables que van a ser incluidas en el análisis.
- Uso de un método de clasificación de cluster con objeto de crear grupos similares.
- Cálculo de las distancias o similitudes entre los casos.
- Validación de los resultados del análisis.

3.1. Selección de las variables que van a ser incluidas en el análisis

El Análisis de Cluster requiere prestar una atención fundamental a los datos seleccionados ya que las clasificaciones resultantes van a estar totalmente condicionadas por éstos. En este sentido, tres son los aspectos fundamentales que hay que considerar en el tratamiento de los datos (Martínez Ramos, 1984: p. 167);

- Que las variables estén incorrelacionadas.
- Que la unidad de medida sea la misma para todas las variables analizadas.
- Que el número de variables no sea demasiado grande.

Los datos utilizados en este estudio cumplen todos estos requisitos ya que la correlación entre los factores es muy baja (3), el número de variables no es demasiado grande, y la unidad de medida es la misma para todas al tratarse de puntuaciones factoriales (4).

3.2. Elección de la técnica de agrupamiento

Una vez formada la matriz de similaridad o distancias entre los individuos, el proceso siguiente es seleccionar la técnica de agrupamiento a utilizar con objeto de establecer las agrupaciones entre los individuos. Éstas se dividen en dos; Jerárquicas y No jerárquicas. A su vez, las técnicas jerárquicas pueden dividirse en aglomerativas o ascendentes y disociativas o descendentes. Las primeras *parten de los casos individuales* y siguiendo un determinado criterio los clasifican en distintos grupos hasta conseguir un único conglomerado; mientras que las técnicas disociativas *parten del conjunto general y van desgredando los individuos hasta la unidad*. Una vez realizada la agrupación, es el propio investigador el que deberá decidir el número de grupos que desea (5).

Por otro lado, las *técnicas no Jerárquicas* (también llamadas de partición) parten de un número de grupos determinado y

(3) Los factores extraídos dentro de cada pregunta son ortogonales, y por lo tanto la correlación entre ellos es nula. No obstante, existe relación entre los factores de las diversas preguntas, relación que puede ser desestimada al ser muy baja y no alcanzar valores significativos.

(4) No obstante, puede realizarse Análisis de Cluster con variables de distintas unidades empleando para ello distancias que contemplen esta particularidad, como la distancia de Gower. Consultar Everitt, 1993: pp. 39 y sigs.

(5) Las técnicas Jerárquicas Aglomerativas más utilizadas son la de las *Distancias Mínimas* (o vecino más próximo), las *Distancias Máximas* (o vecino más lejano) y la *Promedio entre Grupos*. De las Técnicas de Agrupamiento Jerárquicas Disociativas el método más utilizado es el de William Lambert; que está indicado principalmente cuando se trabaja con variables dicotómicas y esta basado en la distancia Chi-Cuadrado. Otro de los métodos más utilizados es el «*Automatic Interaction Detector*». Aunque no es propiamente un método de Análisis de Cluster, el A.I.D. trata de combinar las categorías de las distintas variables a fin de generar grupos que difieran al máximo entre sí respecto de los valores de una variable dependiente generalmente cuantitativa.

van agrupando a los sujetos en cada fase según una determinada medida de similaridad o distancia (6). La gran diferencia entre ambas técnicas es que en éstas el número de grupos suele ser elegido a priori por el investigador, mientras que en las técnicas Jerárquicas los grupos son elegidos al final del proceso.

Atendiendo a su procedimiento, las técnicas No-Jerárquicas tratan de formar grupos bajo la premisa de maximizar la varianza inter-grupos y minimizar la varianza dentro del grupo (intra-grupos). De hecho, tal y como es expuesto por M. D. Paz Caballero (1989: pp. 277-279), estas técnicas se diferencian de las jerárquicas en:

- Parten de la matriz original de puntuaciones; y no de la matriz de proximidades.
- El investigador debe decidir cuántos clusters desea obtener antes de comenzar con todo el proceso de cálculo.
- Los clusters resultantes no están anidados unos en otros, sino que son independientes.

La elección de que técnica va a ser utilizada es algo que depende del propio investigador, y tiene que ser compatible con su deseo de clasificación, la métrica de las variables utilizadas y la medida de similaridad seleccionada (Aldenderfer y Blashfield, 1980: p. 34). En este caso concreto, de todas las técnicas disponibles vamos a utilizar las No-Jerárquicas ya que los métodos jerárquicos tienen el inconveniente de ser muy inestables y poco fiables en sus resultados, y es muy difícil decidir cuántos grupos deben seleccionarse tras el análisis (Martínez Ramos 1984: pp. 165-208). Otros autores aconsejan utilizar métodos jerárquicos cuando se trabaja con muestras pequeñas (Aaker y Day, 1989: p. 510).

Desde la perspectiva de Francisco Alcantud (1985: p. 81) las técnicas No-Jerárquicas son las que mejor se adaptan a los estudios psicológicos y sociológicos caracterizados por el empleo de grandes conjuntos de datos. Una de las ventajas de este enfoque es su gran fiabilidad, ya que una incorrecta asignación de un individuo a un grupo puede ser modificada y corregida posteriormente (Bigne 1990: pp. 368-390).

(6) Dentro de estas técnicas, las más utilizadas son el método K-Means, el Hill-Climbing y el Q-Technique.

De todas las técnicas No-Jerárquicas expuestas anteriormente hemos seleccionado el K-Means; que comienza con una división del conjunto de los datos en x grupos configurados al azar y posteriormente trata de mejorar esta primera clasificación reasignando aquellos elementos que tienen una gran distancia al centroide del cluster que pertenecen, tratando de este modo de reducir la distancia (7) media entre cada elemento de un grupo y su centroide. Como señala M. Anderberg (1973), el proceso de funcionamiento de este método es el siguiente:

- Se comienza con una partición inicial de los datos en un específico número de agrupamientos para calcular posteriormente el centroide de cada uno.
- El siguiente paso trata de reasignar cada caso al agrupamiento más cercano (cuya distancia al centroide del cluster sea menor).
- Calcula los nuevos centroides de los clusters cada vez que se incorpora un nuevo caso.
- Repite alternativamente el segundo y tercer paso hasta que ninguna reasignación de un caso a un nuevo cluster permita reducir más la distancia entre los individuos dentro de cada agrupamiento, ni aumentar la distancia entre los distintos clusters.

Los métodos no jerárquicos requieren que el investigador fije de antemano el número de agrupaciones o clusters que desea obtener; pudiéndose verificar la exactitud de esta clasificación por las técnicas de validación que serán expuestas en el apartado 3.4. Por otro lado, diversos autores aconsejan su utilización cuando se desea analizar no tanto cual es la estructura jerárquica de los individuos; sino cuando el objetivo es conocer el número de grupos construidos y las características de cada uno (Paz Caballero, 1989: p. 282).

3.3. Selección de las medidas de proximidad y distancia

Llegados a este punto en el que ya se conocen los datos a utilizar y el proceso de «encadenamiento», procederemos a

(7) El concepto de distancia y las distintas distancias disponibles serán explicadas en el siguiente apartado.

medir «qué grado de similitud» o «de diferencia» tienen los individuos seleccionados. Con objeto de solucionar esta cuestión se han desarrollado una gran cantidad de índices de distancia y de proximidad que permiten medir el grado de diferencia o semejanza entre dos individuos. No debe olvidarse que la distancia va a variar en función de las variables seleccionadas y el coeficiente de ponderación (si es que existe) asignado a cada una. Algunas de las medidas utilizadas más frecuentemente se exponen en el cuadro 5.

De todas las medidas disponibles, en este trabajo vamos a utilizar la Distancia Euclídea puesto que consideramos que es la más adecuada para nuestros datos (Jain y Dubes, 1988: p. 17): La distancia euclídea entre dos individuos, caracterizados por dos variables, se define como «la raíz cuadrada de la suma de las diferencias entre los elementos al cuadrado».

De hecho, la distancia entre los individuos 1 y 2, considerando dos variables (X e Y) es:

$$d_{12}^2 = (X_2 - X_1)^2 + (Y_2 - Y_1)^2$$

Si en vez de dos variables tenemos m, la distancia entre los dos sujetos se calcula generalizando la fórmula anterior:

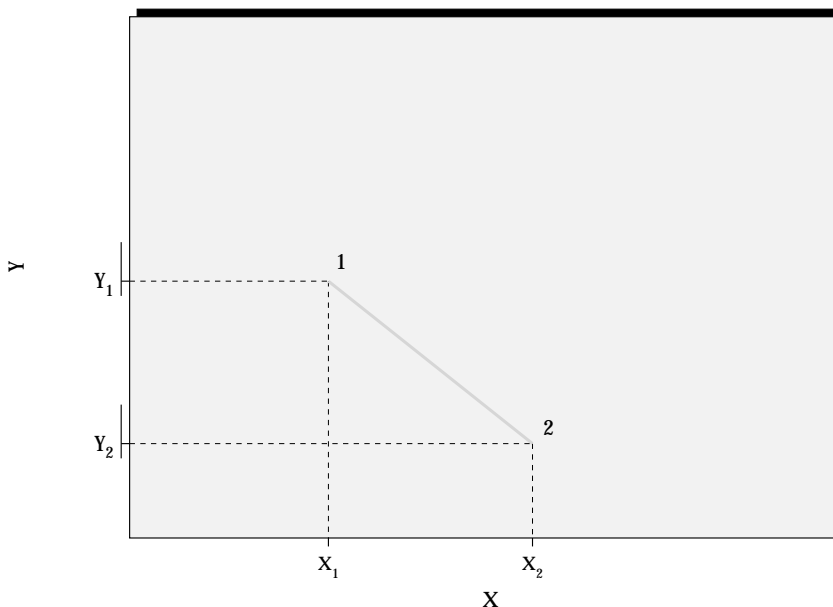
$$d_{12} = \sqrt{\sum_{i=1}^n (X_{2k} - X_{1k})^2}$$

Cuadro 5

CLASIFICACIÓN DE LAS PRINCIPALES MEDIDAS DE DISTANCIA

Datos binarios: Jaccard Russel y Rao Sokal y Sneath Rogers y Tanimoto
Variables discretas: Chi-Cuadrado Phi-Cuadrado
Variables continuas: Distancia euclídea, euclídea al cuadrado Coef de correlación entre vectores (similitud) Coseno de vectores (similitud) Distancia métrica de Chebynev Distancia Manhattan o City-block Distancia de Minkowski

Gráfico 2



donde k indica el número de variables.

Una vez realizada esta exposición, y concluidos los cálculos pertinentes, han sido seleccionado cinco clusters que oscilan entre un tamaño mínimo de 77 personas y un máximo de 188 (cuadro 6). Las distancias entre los centros de cada cluster nos indican el grado de diferenciación entre cada uno, de modo que puede decirse que son los grupos 4 y 1, así como el 3 y 2 los que presentan más diferencias entre ellos.

No obstante, antes de comenzar con la definición y caracterización pormenorizada de los rasgos característicos de cada uno de éstos procederemos a validar los resultados obtenidos. Debe quedar claro que el fin del próximo apartado no es realizar una exposición exhaustiva de estas técnicas –ya que la literatura sobre validación de los clusters es casi tan grande como la literatura dedicada al Análisis de Cluster– sino exponer someramente algunas de las técnicas más importantes, a la vez que aconsejamos al lector interesado la lectura del cuarto capítulo del trabajo de Anil K. Jain y Richard C. Dubes (1988: pp. 143-122).

Cuadro 6

CARACTERÍSTICAS DE LOS CLUSTERS SELECCIONADOS

Número de casos en cada agrupamiento		
Cluster	Casos	% respecto al total
1	87,0	15,8%
2	122,0	12,2%
3	78,0	14,1%
4	77,0	13,9%
5	188,0	34,0%
Total	552,0	100,0

Distancias finales entre los centros de los clusters					
Cluster	1	2	3	4	5
1	0,0000				
2	2,3755	0,0000			
3	2,5339	3,2325	0,0000		
4	3,3303	2,7607	2,9001	0,0000	
5	2,0617	2,3215	2,2537	2,5212	0,0000

3.4. Comprobación v validación de los resultados obtenidos

El último paso en la realización de un Análisis de Cluster es evaluar la clasificación obtenida.

Las técnicas más utilizadas en la validación de los cluster han sido el Análisis de Varianza a fin de conocer las diferencias que presentan las variables (factores en este caso) en cada uno de los clusters, y el Análisis Discriminante con objeto de descubrir el porcentaje de sujetos que están asignados correctamente.

3.3.1. Análisis de Varianza

La técnica estadística del Análisis de Varianza trata de dividir la varianza de la variable dependiente –en este caso cada uno de los factores– en varios componentes, cada uno de los cuales puede ser atribuido a una fuente (variable o factor) identificable. En este apartado se utilizará para conocer si los factores extraídos permiten localizar diferencias significativas entre cada una de las agrupaciones realizadas (8). De no ser

(8) Una exposición detallada de esta técnica puede encontrarse en M. C. Cuadras (1981), F. J. Tejedor (1985) y F. Calvo (1993).

así, se debería eliminar esa variable en el proceso de formación de los clusters, o proceder a una nueva reconfiguración de los grupos extrayendo un número distinto de clusters. Se trata, en definitiva, de conocer hasta qué punto cada uno de los factores nos permite diferenciar significativamente los individuos de cada agrupación.

En el cuadro 7 se muestra el nombre de cada factor, la F de Fisher resultado de aplicar el Análisis de Varianza y la significación de ésta, siendo esta última columna la que nos indica que todos los factores producen diferencias significativas en la definición de los cinco clusters.

3.4.2. Análisis Discriminante

El segundo criterio utilizado para validar los resultados se fundamenta en evaluar la clasificación y asignación de un individuo –cuyas características se conocen– a un determinado grupo (Paz, 1989: p. 284); o como afirma F. Alcantud «analizar la bondad de la clasificación utilizando la asignación de cada sujeto al cluster que pertenece como una nueva variable, y mediante el Análisis Discriminante estudiar la bondad de la clasificación obtenida» (1985: p. 96). Recordaremos que el Análisis Discriminante analiza cómo un conjunto de individuos con unas determinadas características pueden clasificarse en una serie de grupos definidos a priori; tratando de loca-

Cuadro 7

DIFERENCIAS SIGNIFICATIVAS DE LOS FACTORES EN CADA CLUSTER

Factor	F	Signif.
Racionalidad culta	28,9962	0,000
Racionalidad vida cotidiana .	136,6972	0,000
Conducta no racional	58,3213	0,000
Comprar marcas por aspectos utilitarios	48,9154	0,000
Comprar marcas porque indican status	94,6832	0,000
Placer en el acto de compra .	59,0817	0,000
Calidad en las compras	23,7810	0,000
Consumo conspicuo	39,1691	0,000
Import. ambiente social y calidad de vida	45,4478	0,000

lizar cuáles son las variables que mejor contribuyen a la clasificación de cada individuo. Por las propias pretensiones de este trabajo se eludirá realizar una exposición pormenorizada de esta técnica y nos centraremos únicamente en analizar el número de individuos que han sido clasificados correctamente, una vez comprobada que la función discriminante es significativa (Klecka, 1980).

En definitiva, la utilización que hacemos de esta técnica es similar a la que realizan S. Shunglu y M. Sarkar (1995: pp. 123-131) cuando aplican el Análisis Discriminante a los resultados de un análisis de cluster formado por una escala de 74 ítems sobre Actitudes, Intereses y Opiniones: Una vez formados 8 clusters aplican un Análisis Discriminante «step-wise» y deciden eliminar 35 ítems que aportan poca capacidad discriminante al modelo. Posteriormente realizan de nuevo un Análisis de Cluster con los ítems restantes, comprobando que –con la mitad de variables– apenas se alteran los resultados obtenidos, consiguiendo un notable ahorro de tiempo y dinero en la realización de sucesivos trabajos de campo. No obstante, y debido a la alta capacidad discriminante de nuestro modelo, en este trabajo no hemos procedido a la eliminación de variables.

Así, el cuadro 8 muestra un cruce de tablas ente los individuos pertenecientes a cada cluster y aquellos que están correctamente especificados según el Análisis Discriminante. Así, en el Cluster UNO los 87 individuos están correctamente incluidos, mientras que en el Cluster DOS hay 5 individuos que el Análisis Discriminante los coloca en los cluster UNO y CUATRO. No obstante, la clasificación en 5 grupos es válida ya que el 94 por ciento de los individuos está correctamente clasificado.

Estas técnicas, que analizan fundamentalmente el poder explicativo de determinadas variables, apenas informan del número de agrupaciones que es necesario formar ni de la consistencia interna de cada uno. Por este motivo se han desarrollado otras fórmulas como es el proceder a la extracción de un diferente número de grupos y comparar sucesivamente las distancias entre los clusters y las distancias dentro de cada agrupación, tratando de seleccionar aquella división que maximice la primera (distancia inter-grupos) y que minimice la segunda (intra-grupos). Llegados a este nivel, hay autores que

Cuadro 8

TABLA DE CLASIFICACIONES

Cluster	N.º total de casos	Miembros predichos en cada cluster				
		1	2	3	4	5
UNO	87	87 100,0%	0 0,0%	0 0,0%	0 0,0%	0 0,0%
DOS	122	1 0,8%	117 95,9%	0 0,0%	0 0,0%	4 3,3%
TRES	78	2 2,6%	0 0,0%	71 91,0%	3 3,8%	2 2,6%
CUATRO	77	0 0,0%	3 3,9%	0 0,0%	74 96,1%	0 0,0%
CINCO	188	4 2,1%	4 2,1%	5 2,7%	5 2,7%	170 90,4%

Número de casos correctamente clasificados: 518 (94 por ciento).

proponen repetir el análisis de cluster final en submuestras aleatorias extraídas de la muestra objeto de estudio, técnica conocida como el «grado de replicabilidad de un cluster» (Aldenderfer y Blashfield, 1985: p. 65).

4. CONCLUSIÓN: INTERPRETACIÓN DE LOS RESULTADOS OBTENIDOS

Pese a la importancia de los criterios expuestos con anterioridad ninguno de ellos es adecuado si la clasificación resultante carece de sentido teórico o es muy difícil su interpretación (Fernández Santana, 1991: pp. 65-76). En última instancia, éste es el criterio más importante de validación; una correcta interpretabilidad. En esta línea, Francisco Alcantud propone un método de validación para los métodos no jerárquicos que consiste en extraer paulatinamente diferentes clusters y comprobar la solidez teórica de cada uno de ellos, al tiempo que se realiza un «recorrido» de cómo se distribuyen los sujetos de cada partición: Es decir, cuando se han seleccionado cuatro clusters y se hace de nuevo el análisis con cinco clusters; se trata de comprobar el comportamiento de cada uno de éstos en ambas particiones a fin de localizar cómo se distribuyen los sujetos en cada uno de los distintos agrupa-

mientos, así como conocer el origen de los sujetos del nuevo cluster (1985: pp. 85-88).

Aplicado conjuntamente estos criterios, y una vez seleccionados cinco clusters, se realizará una interpretación analizando las puntuaciones factoriales en cada uno de los cinco clusters, ayudados por el análisis de cómo se comportan las diversas variables sociodemográficas en cada uno de ellos (cuadro 9). Debe quedar claro que el objetivo de esta interpretación es puramente técnico y tiene como objeto comprobar la homogeneidad interna de cada uno y la diferencia con el resto, de modo que eludiremos realizar un comentario sociológico de los hallazgos encontrados.

4.1. TIPOLOGÍA 1: «Consumo conspicuo»: 87 personas, 15,8 por ciento del total

El cluster número uno presenta la composición del primer cluster con las puntuaciones de los diferentes factores utilizados en su configuración. Este colectivo, formado por 87 personas y que representa el 16 por ciento del total de los entrevistados, se caracteriza por la importancia que conceden a factores como el «Consumo conspicuo», «las marcas indican el status de quien las lleva», al tiempo que consideran el ir de compras como un acto lúdico y placentero.

El análisis del cuadro 9 nos informa que un 54,3 por ciento de este colectivo deja de usar las prendas cuando ya no les gustan, son los que más frecuentemente compran cosas que no llegan a utilizar y realizan sus compras de ropa y calzado cuando ven algo que les gusta, frente a otros colectivos que adquieren los productos en temporada de rebajas o cuando necesitan algo. En cuanto a sus rasgos sociodemográficos, el 68,4 por ciento son personas menores de 35 años, hay un predominio de mujeres frente a hombres y de solteros frente a casados.

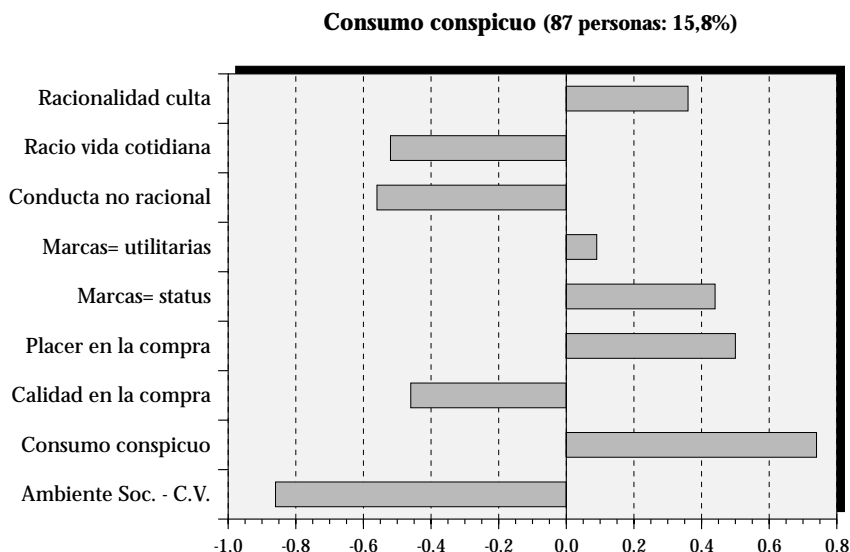
4.2. TIPO 2: «Jóvenes consumistas, placer en la compra»: 122 personas, 22,2 por ciento del total

En términos numéricos esta es la segunda agrupación ya que representa al 22,2 por ciento de los entrevistados, caracterizan-

Cuadro 9

RASGOS DE LOS TIPOS DE CONSUMIDORES

	Consum conspicuo	Jóvenes consumi- stas	Desinte- resado	Decisión del exterior	Utiliza- dores
Hace personalmente la compra de ropa y calzado					
Siempre y casi	88,2%	88,2%	79,8%	88,1%	74,5%
Alguna vez, rara vez, Nunca ..	11,8%	11,8%	20,2%	11,9%	25,5%
Razones por las cuales deja de usar prendas					
Mal estado	36,2%	34,6%	66,3%	47,6%	68,5%
Ya no gustan	54,3%	43,3%	24,4%	23,8%	21,3%
Pasadas de moda	9,6%	22,0%	9,3%	28,6%	10,2%
¿Suele comprar cosas que después no usa?					
Frecuentemente	18,1%	10,2%	9,0%	2,4%	4,9%
Alguna vez	59,6%	57,5%	32,6%	46,4%	33,8%
Nunca	22,3%	32,3%	58,4%	51,2%	61,3%
¿Cuándo suele comprar la ropa y el calzado?					
Comienz temporada-rebaja	9,7%	18,1%	12,2%	3,6%	9,8%
Necesito algo	66,7%	59,1%	73,2%	72,6%	81,9%
Le gusta algo	23,7%	22,8%	14,6%	23,8%	8,3%
Utilización de la misma marca en determinados productos					
Fidelidad	23,2%	44,1%	30,3%	47,6%	27,9%
No me fijo en las marcas ..	52,6%	34,6%	52,8%	31,0%	63,2%
Cambio de marca con frecuencia	24,2%	21,3%	16,9%	21,4%	8,8%
Tipo de familia					
2 núcleos (F. extensa)	14,7%	13,4%	34,8%	21,4%	15,2%
Fam. nuclear	62,1%	73,2%	40,4%	60,7%	71,1%
Solo, compartido, etc.	10,5%	3,1%	15,7%	11,9%	4,4%
Pareja sin hijos	12,6%	10,2%	9,0%	6,0%	9,3%
Estado Civil					
Soltero	53,8%	53,7%	52,9%	30,3%	33,5%
Casado/convive pareja	46,2%	46,3%	47,1%	69,7%	66,5%
Sexo					
Hombre	29,5%	38,6%	62,9%	57,1%	61,3%
Mujer	70,5%	61,4%	37,1%	42,9%	38,7%
Edad					
16-25	35,8%	37,8%	16,9%	13,1%	17,2%
26-35	32,6%	23,6%	24,7%	26,2%	17,6%
36-45	15,8%	15,0%	21,3%	27,4%	21,1%
46-55	6,3%	11,8%	16,9%	19,0%	21,6%
56-65	9,5%	11,8%	20,2%	14,3%	22,5%
Ingresos económicos					
Menos de 100.000 ptas.	8,2%	2,5%	9,5%	10,8%	7,0%
100.001-150.000 ptas.	23,5%	17,4%	29,7%	16,2%	21,0%
150.001-200.000 ptas.	25,9%	29,8%	25,7%	21,6%	26,9%
200.001-275.000 ptas.	15,3%	33,1%	12,2%	23,0%	14,5%
Más de 275.000 ptas.	27,1%	17,4%	23,0%	28,4%	30,6%

Clúster núm. 1

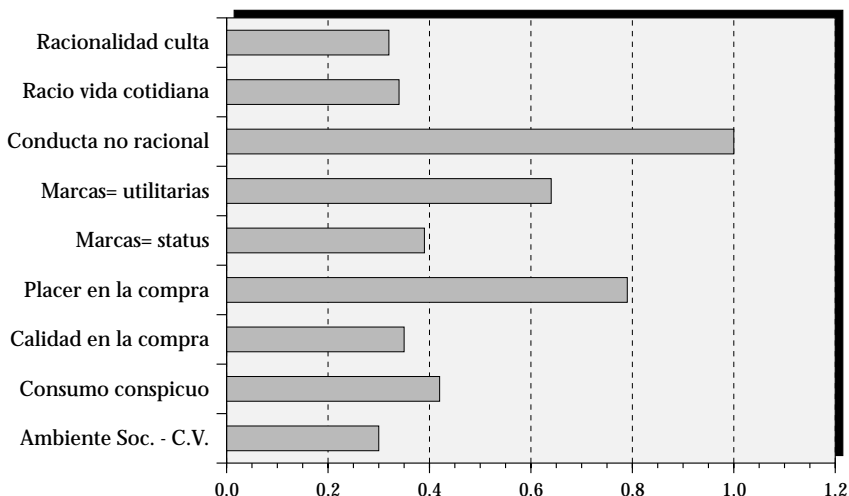
dose por un predominio de jóvenes menores de 25 años, mujeres, ingresos entre 150.000 y 275.000 ptas., y familias nucleares.

La característica esencial y definitoria de este colectivo es la escasa racionalidad que muestra en el acto de compra, el gran placer que experimentan en el mismo y su consideración que la marca proporciona una comodidad y un ahorro de tiempo (aspectos utilitarios), aunque sin desdeñar que ésta indica el status social del que la lleva. Tal y como se muestra en el cuadro 9 este colectivo se caracteriza por considerar la marca de los productos (únicamente un 34,6 por ciento no se fija en la marca), destacando sobre todo su elevada fidelidad a determinadas marcas (el 44 por ciento suele comprar siempre las mismas marcas). Por otro lado, el momento de realización de sus compras de ropa y calzado suele ser a principio de temporada y cuando ven algo que les gusta.

Esta elevada consideración de la marca de los productos, según J. Bilas (1990: pp. 19-20), debe interpretarse considerando que la función de la marca es «identificar socialmente» y permitir a su portador se igual «al nosotros» y diferente a los

Clúster núm. 2

Jóvenes consumistas (122 personas: 22,2%)



«otros», de modo que estas prácticas sociales quedan convertidas en «comportamientos-signo» con objeto de manifestar la pertenencia del individuo a grupos sociales que poseen más o menos «clase», más o menos «distinción-pretensión-necesidad» (Bourdieu, 1984: pp. 28, 53-66 y 397-403). Desde esta perspectiva, la «compra de marcas» es muy frecuente entre los jóvenes debido a la gran necesidad de identificación con un grupo que se experimenta en la juventud y adolescencia. Desde una perspectiva más amplia, Jean Baudrillard considera que una de las características definitorias de la sociedad actual es que los otros sistemas de reconocimiento se retraen y se considera el «standing» –producido por la posesión de ciertos productos– como reflejo del status social (1988: p. 217).

**4.3. TIPO 3: «Desinteresados por el acto de compra»:
78 personas. 14,1 por ciento del total**

Este tercer cluster, que agrupa al 14 por ciento de los entrevistados, se caracteriza por puntuar negativamente todos

los factores a excepción del que hace referencia a «la gente compra en consonancia con el ambiente social en el que vive». Analizando estas puntuaciones destaca sobre todo la no consideración de la racionalidad de vida cotidiana, el rechazo a la calidad en las compras y la escasa importancia concedida al factor «consumo conspicuo».

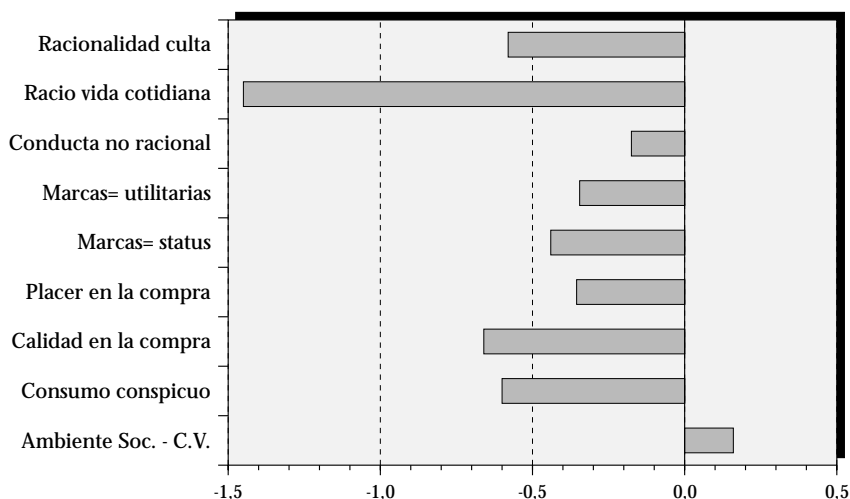
Por otro lado, el análisis del cuadro 9 deja entrever que este colectivo está formado mayormente por hombres, con predominio de familias extensas. El análisis de su conducta de compra permite comprobar que nunca compran productos que no utilizan, dejan de usar las prendas cuando se encuentran en mal estado y no prestan atención a la la marca de los productos.

4.4. TIPO 4: «Decisión de compra influida por el exterior»: 77 personas, 13,9 por ciento del total

Este agrupamiento, con una representación poblacional similar al anterior viene caracterizado por la importancia que se

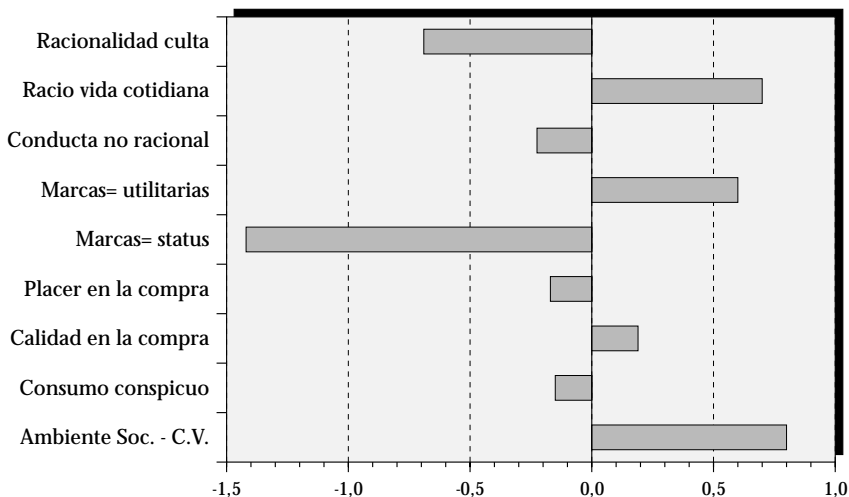
Clúster núm. 3

Desinteresados por comprar (78 personas: 14,0%)



Clúster núm. 4

Compra influida por otros (77 personas: 13,9%)



concede a la idea de que la gente compra en consonancia con el ambiente social en el que vive, a la vez que sus conductas de compra se caracterizan por una racionalidad de vida cotidiana y por la importancia que conceden a los aspectos funcionales que proporcionan la compra de «marcas prestigiosas» (Comodidad y un ahorro de tiempo).

En este grupo hay un predominio de hombres, casados, y edades entre los 36 y 45 años que nunca compran en rebajas, ya que el 50 por ciento de ellos consideran los productos rebajados como auténticos engaños, y manifiestan una elevada fidelidad a las marcas. Para A. del Pino la elevada fidelidad a las marcas que muestran los hombres hay que interpretarla considerando que el hombre es más deductivo que la mujer, más escéptico y más propenso a creer y confiar en argumentos lógicos. El hombre muestra fidelidad a determinados productos y marcas porque las ha elegido por convicción y ha elaborado fuertes razones para cada elección (1990: p. 137).

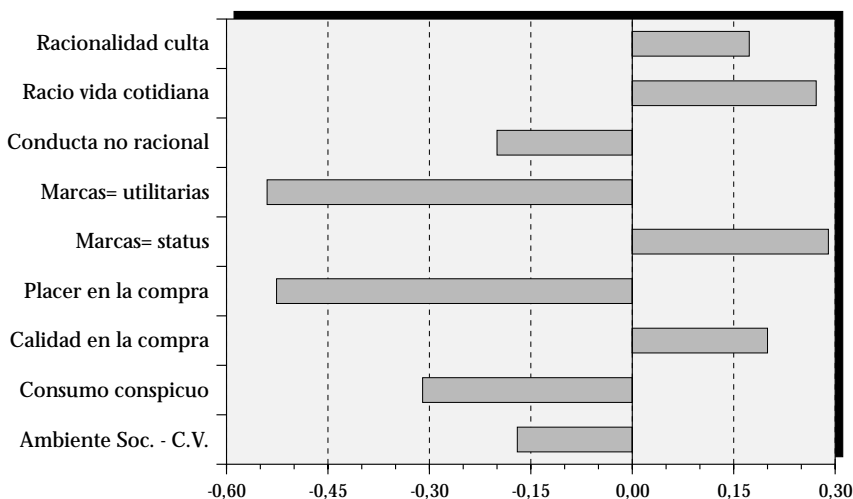
4.5. TIPO 5: «Utilizadores»: 188 personas, 34,0 por ciento del total

Esta última agrupación es la mayor de todas las analizadas y comprende al 34 por ciento de la muestra. Sus puntuaciones en los factores se caracterizan por una racionalidad de vida cotidiana, así como un comportamiento basado y fundamentado en la calidad de las compras ya que consideran que la calidad adquiere más importancia que los precios.

Los pertenecientes a este colectivo son los que menos realizan personalmente sus compras de ropa y calzado, y son los que más expresan que NUNCA compran productos que no utilizan. Por otro lado, suelen comprar ropa y calzado cuando necesitan los productos y no prestan atención a las marcas. Sus rasgos sociodemográficos indican que hay un predominio de familias nucleares, hombres, casados, y edades superiores a los 56 años. □

Clúster núm. 5

Utilizadores (188 personas: 34,0%)



BIBLIOGRAFÍA

- AAKER, D. A. y DAY, G. S. (1989): *Investigación de Mercados*, McGraw-Hill, México.
- ALCANTUD, F. (1985): *Técnicas de clasificación*, Ediciones Universitarias de Valencia, Valencia.
- ALDENDERFER, M. S. y BLASHFIELD, R. K. (1980): *Cluster Análisis*, Sage, London.
- ANDERBERG, M. (1973): *Cluster Análisis for Applications*, Academic Press, Nueva York.
- BAUDRILLARD, J. (1988): *El sistema de los objetos*. Siglo XXI, Madrid.
- BELLO ACEBRÓN, L. (1994): «El valor de la marca: Elementos y enfoques de medición», en *Esic Market*, julio, pp. 18- 129.
- BISQUERRA, R. (1989): *Análisis Multivariable*, PPU, Barcelona.
- BLAS, J. (1990): «Márcame», en *Revista Kontsumo*, 6, marzo-abril.
- BOURDIEU, P. (1988): *La distinción. Criterio y bases sociales del gusto*. Taurus Humanidades, Madrid.
- CALVO GÓMEZ, F. (1993): *Técnicas Estadísticas Multivariantes*, Univ. de Deusto, Bilbao.
- COMREY, A. L. (1985): *Manual de Análisis Factorial*, Cátedra, Madrid.
- CUADRAS, C. M. (1981): *Métodos de Análisis Multivariable*, Eubibar, Barcelona.
- DIDAY, E. y otros (1982): *Elements d'Analyse de Données*, Dunod, Paris.
- ENRIQUE BIGNE, J. (1990): «El Análisis de Cluster: Una aplicación a las áreas de interés de las Agencias de Publicidad», en *Resic-Market*, julio-septiembre, pp. 368-390.
- FERNÁNDEZ SANTANA, J. O. (1991): «El Análisis de Cluster: Aplicación, interpretación y validación», en *Papers, Revista de Sociología*, 37, Publicaciones de la Universidad Autónoma de Barcelona, pp. 65-76.
- HARMAN, H. (1980): *Análisis Factorial Moderno*, Saltés, Madrid.
- JAIN, A. K. y DUBES, R. C. (1988): *Algorithms for Clustering Data*, Prentice Hall, New Jersey.
- LAGO, C. (1990): «Enganchados al consumo», en *Revista Kontsumo*, 7, mayo, pp. 39-12.
- LOZARES, C. (1990): «La tipología en Sociología, más allá de la simple taxonomía: Conceptualización y cálculo» en *Papers, Revista de Sociología*, 34, Barcelona, pp. 139-163.

- KIM, J. O. y MUELLER, C. W. (1978): *Introduction to Factor Analysis*, Sage, London.
- KIM, J. O. y MUELLER, C. W. (1978): *Factor Analysis: Statistical Methods and Practical Issues*, Sage, London.
- KLECKA, J. (1980): *Discriminant Analysis*, SAGE, New York.
- MALLO, F. (1985): *Análisis de Componentes Principales*, Universidad de León, León.
- MARTÍNEZ RAMOS, E. (1984): «Aspectos teóricos del Análisis de Cluster y aplicación a la caracterización del electorado», en J. J. Sánchez Carrión (ed.), *Introducción a las técnicas de Análisis Multivariable*, Centro de Investigaciones Sociológicas, Madrid, pp. 165-208.
- PAZ CABALLERO, M. D. (1989): «Análisis de Cluster», en G. Vallejo, *Análisis Multivariante aplicado a las ciencias del comportamiento*, Servicio de Publicaciones de la Universidad de Oviedo, Oviedo, pp. 277-279.
- PINO, A. (1990): *Los nuevos consumidores españoles*. Deusto, Bilbao.
- SÁNCHEZ CUENCA, J. (1990): «La tipología», en E. Ortega Martínez *Manual de Investigación Comercial*, Pirámide, Madrid, pp. 524-539.
- SHUNGLU, S. y SARKAR, M. (1995): «Researching the Consumer» en *Marketing and Research*, today, vol. 23, 2.
- SOKAL, R. y SNEAT, P. (1963): *Principles of Numerical Taxonomy*, W. H. Freeman, S. Francisco.
- TABACHNICK, B. G y FIDELL, L. S. (1989): *Using Multivariate Statistics*, Harper & Row, New York.
- TEJEDOR, F. J. (1985): *Análisis de la Varianza aplicada a la investigación en Pedagogía y Psicología*, Anaya, Madrid.
- VÁZQUEZ CASIELLES, R. (1989): «Estrategias de decisión del consumidor», en *Revista IPMARK*, 330, mayo, pp. 55-73.
- VÁZQUEZ CASIELLES, R. (1989): «La imagen de la empresa detallista y su percepción por los consumidores (Factores que influyen en la conveniencia y en lo agradable que puede ser ir de compras)», en *Esic Market*, julio, pp. 49-75.

RESUMEN

Diseño de tipologías de consumidores mediante la utilización conjunta del Análisis Cluster y otras técnicas multivariantes

El objetivo del presente trabajo es analizar los resultados obtenidos en la clasificación de una determinada población cuando el Análisis de Cluster se aplica a un

conjunto de factores comunes obtenidos mediante el Análisis Factorial de Componentes Principales. Posteriormente, y con objeto de validar los resultados obtenidos se aplicará un criterio mixto de verificación consistente en el análisis de los resultados aportados por dos técnicas estadísticas (Análisis de Varianza y A. Discriminante) y, fundamentalmente, en una adecuada interpretabilidad.

De los diversos criterios de clasificación existentes (en base a las características sociodemográficas, según los estilos de vida, etc.) este trabajo agrupa la población Navarra atendiendo a su comportamiento como consumidores. No obstante, y con objeto de conocer mejor los rasgos de los distintos segmentos encontrados, se analizará también la caracterización sociodemográfica de cada uno de ellos.

PALABRAS CLAVE: Análisis multivariante, Análisis Cluster, segmentación, comportamiento del consumidor.

RÉSUMÉ

Établissement de typologies de consommateurs sur la base d'une utilisation conjointe de l'analyse cluster et d'autres techniques multivariantes

Le présent travail a pour objet d'analyser les résultats obtenus dans la classification d'une population déterminée quand l'analyse cluster est appliquée à un ensemble de facteurs communs obtenus par une analyse factorielle des composants principaux. Postérieurement, et afin de valider les résultats obtenus, est appliqué un critère mixte de vérification consistant dans une analyse des résultats apportés par deux techniques statistiques (analyse de variance et analyse discriminante) et fondé essentiellement sur une interprétabilité appropriée.

Choisissant parmi les différents critères de classification à retenir (caractéristiques sociodémographiques, modes de vie, etc.), ce travail regroupe la population de Navarre en fonction de son comportement en tant que consommateurs. Toutefois, et afin de mieux cerner les caractères des différents segments retrouvés, est étudiée de même la caractérisation sociodémographique de chacun de ces segments.

MOTS CLÉF: Analyse multivariante, Analyse Cluster, segmentation, comportement de consommateur.

SUMMARY

Consumer typing by means of the joint use of cluster analysis and other multivariate techniques

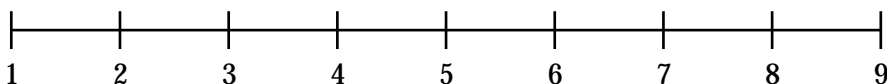
The purpose of this paper is to analyse the results obtained by classifying a particular population, applying cluster analysis to a series of common factors outputted by means of Main Component Factor Analysis. A hybrid verification criterion will then be applied, involving the analysis of the results outputted by two statistical techniques (Variance Analysis and Discriminant Analysis) and, fundamentally, satisfactory interpretability, for the purpose of validating the results obtained.

Of the different criteria of classification available (based on socio-demographic characteristics, by life styles, etc.), this paper groups the population of Navarre according to their behaviour as consumer. However, with a view to gaining a better understanding of the traits of the different segments identified, the socio-demographic characterisation of each one will also be analysed.

KEYWORDS: Analysis multivariant, Analysis Cluster, segmentation, behaviour of the consumers.

ANEXO: PREGUNTAS ANALIZADAS

1. Cuando realiza compras, cuánto cree Vd. que influye en su decisión cada uno de estos aspectos:



1 No influye nada; **5** Influye medianamente; **9** Influye mucho

- Comprar productos que se anuncian
- El análisis de la etiqueta y las características del producto ...
- Comprar productos de moda
- Aprovechar ofertas y rebajas
- La información o consejos del dependiente
- La relación precio-calidad
- Comprar productos conocidos que ha usado anteriormente
- Comparar precios en distintos establecimientos
- Preguntar a sus amigos sobre los productos que utilizan

2. Puntuando 1 si está totalmente en desacuerdo y 9 si está totalmente de acuerdo, ¿podría decirme cuáles cree usted que son los motivos que llevan a la gente a comprar «marcas» prestigiosas?

LA GENTE COMPRA MARCAS «PRESTIGIOSAS»...

- Para presumir
- Porque la gente se fía más de productos con marcas conocidas
- Porque hay marcas que proporcionan una cierta «individualidad», van con la forma de ser de cada uno
- Por que se encuentran más cómodos con esas marcas
- Por diferenciarse de otros
- Porque al comprar productos con marcas conocidas se puede realizar la compra en menos tiempo puesto que así evitan tener que mirar muchos productos

3. Le voy a leer una serie de frases sobre distintos comportamientos y actitudes de compra. ¿Podría decirme en qué medida está usted de acuerdo con cada una de estas frases, puntuando 1 si está totalmente en desacuerdo y 9 si está totalmente de acuerdo?

- A) Procuo ir siempre bien vestido, porque para mi es muy importante la impresión que doy a los demás

- B) Normalmente decido sobre la marcha lo que compro
- C) Me gusta conocer tiendas nuevas
- D) Me gusta acompañar a alguien que va de compras
- E) Cuando compro doy más importancia a la calidad
que a los precios
- F) Me gusta «invertir» en una casa elegante,
ya que dice mucho de los que en ella viven
- G) Las diferencias entre unas marcas y otras se notan
mucho, la marca es una garantía
- H) Me molesta ver a alguien con ropa igual que la mía
- I) El coche es un objeto que puede indicar el prestigio
de quien lo lleva
- J) Es muy frecuente que compre cosas
que no tenía pensadas
- K) Me gusta vestir a la moda
- L) La gente compra productos en consonancia
con el ambiente social en el que vive
- LL) Cuando compro algún producto para mi casa
lo hago pensando únicamente en que va a hacerme
la vida más agradable