

Nuevo procedimiento metodológico para el análisis exploratorio de una tabla estructurada en diversos conjuntos de individuos

Karmele Fernández Aguirre

Departamento de Economía Aplicada III
Universidad del País Vasco

M. Isabel Landaluce Calvo

Departamento de Economía Aplicada
Universidad de Burgos

Juan I. Modroño Herrán

Departamento de Economía Aplicada III
Universidad del País Vasco

Resumen

El objetivo del artículo es la presentación de un procedimiento metodológico para el análisis exploratorio de la tipología existente en el seno de diferentes grupos de individuos que han sido objeto del mismo estudio cualitativo. El núcleo principal de la técnica está basado en los principios del Análisis Factorial Múltiple (AFM), que tiene en cuenta la estructura en grupos de los datos, equilibrando su influencia en los resultados globales. En una primera etapa, la técnica propuesta utiliza los resultados proporcionados por el Análisis de Correspondencias Múltiples (ACM), conocido por sus numerosas ventajas en el análisis exploratorio de datos cualitativos. La metodología propuesta proporciona todos los resultados necesarios para un exhaustivo estudio de este tipo de tablas múltiples.

Palabras y frases clave: variables cualitativas, tablas múltiples, grupos de individuos, ACM y AFM

Clasificación AMS: 62H25

New methodological procedure for the exploratory analysis of a data table structured into different groups of individuals

Abstract

The aim of this article is to present a methodological procedure for the exploratory analysis of different groups of individuals who have participated in a qualitative study. The core of the technique is based on the principles of Multiple Factor Analysis (MFA), which takes into account the structure of the data groups, balancing their influence on the overall results. In the first stage, the proposed technique uses the results provided by Multiple Correspondence Analysis (MCA), known for its many advantages in the exploratory analysis of qualitative data. The procedure provides all the results necessary for a comprehensive study of this type of multiple data tables.

Keywords: qualitative variables, multiple tables, groups of individuals, MCA and MFA

Clasificación AMS: 62H25

1. Contextualización. Objetivos

En frecuentes ocasiones se realizan investigaciones en las que la información se obtiene a través de encuestas en las que la muestra seleccionada puede presentar diferentes estructuras de comportamiento, asociadas principalmente a variables de caracterización (sexo, edad, nivel de estudios, categoría laboral, etc.). Estas diferentes estructuras pueden quedar eclipsadas en los análisis de la información en su conjunto. Por ello, esta información debe ser estudiada desde la óptica de tabla múltiple, esto es, teniendo en cuenta la existencia de diversos grupos, lo que requiere realizar, por un lado, análisis parciales de cada uno de ellos y, por otro lado, un análisis global en el que la influencia individual de cada uno de los grupos esté equilibrada.

En el extenso campo de investigación del Análisis de Datos existen numerosas y diversas técnicas desarrolladas para el estudio exploratorio de tablas susceptibles de ser estructuradas en subtablas (también denominadas tablas múltiples, tablas de tres dimensiones, tablas de tres entradas,...) Esta estructuración puede estar asociada tanto a la dimensión de las columnas (variables) como a la dimensión de las filas (individuos). Las investigaciones en las que se dispone de información de esta naturaleza múltiple tienen objetivos más ambiciosos, ya que no se limitan a la búsqueda de relaciones entre variables y tipologías de los individuos, sino que se amplían al análisis comparativo de las realidades presentes en el seno de cada una de las tablas. Esta riqueza interpretativa, junto con la gran casuística de datos, ha animado a muchos investigadores a desarrollar metodologías apropiadas a estos objetivos.

Este trabajo se enmarca en una línea de investigación iniciada en la década de los 80 por los profesores B Escofier y J. Pagès en el seno de la Escuela Francesa de Análisis de Datos, el Análisis Factorial Múltiple (AFM), (Escofier et Pagès, 1986, 1992, 1994, 1998). Desde que estos autores publicaron sus avances sobre el tratamiento de tablas

múltiples, son muchos los estudiosos que se han sumado a su filosofía y han contribuido a afianzar el AFM como una metodología con una gran versatilidad en el tratamiento de información de tres dimensiones. Así lo ponen de manifiesto los numerosos trabajos que en los últimos años han visto la luz y han consolidado esta escuela, poniendo de manifiesto su potencialidad, tanto desde el punto de vista teórico como empírico, en muy diversas áreas. De la vasta literatura científica relacionada con esta técnica merecen ser destacados los siguientes trabajos: Pagès (1996, 2005), Pagès et al. (2001, 2002), Husson et al. (2006-1, 2006-2) en los que se compara, reflejando similitudes y diferencias, el AFM con otros métodos como el modelo INDSCAL, el análisis Procusteano o la metodología STATIS; Pagès (2002, 2004), Pagès et al. (2008), Bécue et al. (2001, 2003, 2008), Abascal et al. (2006) que presentan una extensión del AFM para el tratamiento de tablas mixtas y de tablas de frecuencias; Le Dien et al. (2003) una adaptación del AFM para el tratamiento de encuestas en las que las preguntas están agrupadas en temas (estructuración jerárquica); Le Dien et al. (2010) extensión del AFM para el tratamiento simultáneo de variables cuantitativas medidas en varios grupos de individuos; Morand et al. (2007) extensión del AFM incorporando las rotaciones procustas; Abascal et al. (2001, 2004, 2008), Landaluce et al. (1995, 1999), García Lautre (2001), García Lautre et al. (2003) en los que se recogen diversas aplicaciones del AFM.

Todos estos trabajos ponen de manifiesto que el AFM con el tiempo y uso ha pasado de ser una técnica de análisis de tablas múltiples a toda una filosofía de análisis comparativo, tanto gráfico como a través de indicadores numéricos, de diferentes conjuntos de datos, sea cual sea su naturaleza y su estructura. Nuestra aportación se enmarca en este contexto de adaptaciones, extensiones y aplicaciones del AFM.

Este artículo está integrado por 3 secciones, además de esta primera que ha servido para situar la metodología propuesta en la precisa línea de investigación. En la sección 2 aparece reflejada la técnica que los autores han desarrollado, poniendo de manifiesto las ventajas que ofrece dado el problema y los objetivos planteados. Esta sección se complementa con la tercera en la que se presentan los resultados de una aplicación empírica de la técnica en un estudio real. Las conclusiones alcanzadas, así como las referencias bibliográficas consultadas para su realización completan el trabajo.

2. Procedimiento Metodológico

En análisis de encuestas es muy habitual que la información de partida (el cuestionario) esté referida a una muestra de individuos (anónimos) que pueden ser clasificados en distintos grupos atendiendo a una serie de variables de caracterización (algunas propias del muestreo). Frecuentemente las preguntas (variables) que integran el cuestionario son cerradas y ofrecen diferentes opciones de respuesta (modalidades). El análisis exploratorio de este tipo de datos requiere una codificación previa, siendo la codificación condensada la más utilizada. La codificación condensada consiste en asignar a cada variable tantos valores enteros consecutivos como modalidades tenga la misma. Esta codificación es necesariamente transformada para el procesamiento estadístico. El Análisis de Correspondencias Múltiples (ACM) es el método más apropiado para el análisis de la tabla única resultante de la codificada en forma

disyuntiva completa (yuxtaposición de matrices binarias con tantas columnas como modalidades de respuesta tenga la cuestión planteada); o bien, de la tabla de contingencias múltiples de Burt obtenida al cruzar más de dos variables categóricas (ver, por ejemplo, Escofier & Pagès 1992, 2008 o Greenacre 2007). El ACM trata por igual variables categóricas nominales y ordinales, siendo el carácter ordinal del etiquetado de las modalidades elegido en la codificación condensada irrelevante. Sin embargo, esta técnica no permite contemplar la posibilidad de la estructuración en varias subtablas del conjunto de individuos encuestados y de su estudio comparativo. Únicamente ofrece la posibilidad de analizar una de las subtablas como activa y el resto como suplementarias o apilarlas y utilizar una variable suplementaria como indicadora de la tabla correspondiente.

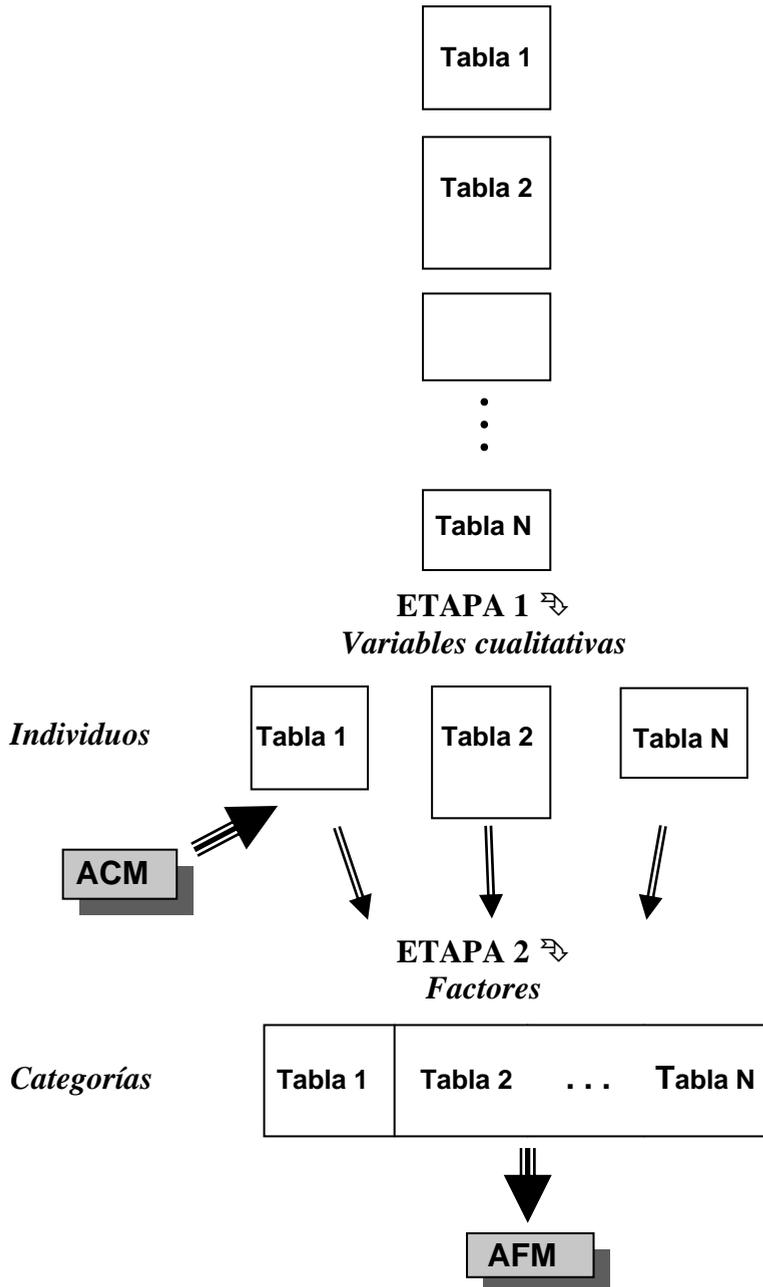
El tratamiento de tablas múltiples supone un enfoque mucho más completo, en el que cada subtabla tiene identidad propia, esto es, tiene un papel activo en los resultados globales, proporcionando, además, indicadores apropiados para medir las semejanzas y (en su caso) las diferencias entre las estructuras internas de cada uno de los grupos considerados.

En la sección previa ya se ha señalado que el análisis exploratorio de tablas múltiples es un campo que ha evolucionado de una forma muy fructífera en los últimos años, existiendo hoy en día métodos adaptados a casi la totalidad de alternativas de tablas de tres dimensiones. Sin embargo, el tratamiento simultáneo de varios conjuntos de individuos en los que se haya medido la misma información de naturaleza cualitativa es un campo que está aún en su etapa experimental.

La metodología que se propone en este trabajo para esta situación se puede esquematizar de la siguiente manera:

Figura 1

Esquema de la información analizada y de la metodología propuesta



Esto es, partiendo de una encuesta realizada a varios grupos de individuos (que integran en su conjunto una muestra):

Etapla 1: se realiza un Análisis de Correspondencias Múltiples sobre cada una de las subtablas asociadas a estos grupos, X_1, \dots, X_J , donde J es el número de subtablas. Cada una de estas tablas tiene como filas los N_j individuos encuestados, $j = 1, \dots, J$ y por columnas las K modalidades de las variables objeto de la encuesta que han sido elegidas para el análisis. No es necesario que el número de individuos que aparecen en cada tabla sea el mismo, pero sí que hayan contestado a las mismas preguntas, que además deben contener el mismo número de respuestas. De cada uno de los ACM correspondientes, se extrae un número reducido de factores r_j y se computan las coordenadas de las modalidades en esos factores, $G_i^j, i = 1, \dots, r_j, j = 1, \dots, J$, de dimensión, $K \times 1$ en un proceso de reducción de la dimensionalidad de la tabla de datos. No es formalmente necesario extraer el mismo número r_j de factores en cada tabla, aunque creemos que hacerlo ($r_i = r_i, \forall i \neq i'$) puede ser una buena práctica de cara a la interpretación futura de los elementos del análisis.

Etapla 2: se construyen J subtablas $G^{(j)}, j = 1, \dots, J$ de dimensión $K \times r_j$ cada una de ellas conteniendo las coordenadas de las modalidades de los ACM de cada subtabla, $G^{(j)} = [G_1^j, \dots, G_{r_j}^j]$. Después, se yuxtaponen las tablas de los factores $G^{(j)}$, formando una gran tabla G de dimensión $K \times \sum_{j=1}^J r_j$ que, por filas, tiene las modalidades de las variables elegidas para el análisis. Las columnas corresponden a los factores elegidos en el paso anterior para todas las subtablas. Esta tabla múltiple así obtenida, de naturaleza cuantitativa, será analizada a través de un AFM, que permitirá obtener una visión completa y simplificada de diferencias y similitudes entre las subtablas/grupos de individuos estudiados. En este caso, el AFM consiste un ACP de la tabla G^* donde las subtablas $G^{(j)}$ han sido previamente ponderadas por el inverso de la raíz cuadrada del primer valor propio de cada ACP separado (o ACP parcial) de cada subtabla:

$$G^* = \left[\frac{1}{\sqrt{\lambda_1^1}} G^{(1)}, \frac{1}{\sqrt{\lambda_1^2}} G^{(2)}, \dots, \frac{1}{\sqrt{\lambda_1^j}} G^{(j)}, \dots, \frac{1}{\sqrt{\lambda_1^J}} G^{(J)} \right]$$

A continuación, se presentan brevemente, y de manera simplificada (prescindiendo del aparato matemático) las metodologías seleccionadas en cada una de las etapas propuestas, poniendo de manifiesto la idoneidad de cada una de ellas dada la naturaleza de los datos y los objetivos planteados en el estudio de los mismos. Son muchas las referencias bibliográficas en las que se puede encontrar una detallada presentación de ambas técnicas. Los autores recomiendan, por la proximidad al trabajo aquí expuesto, la obra de Escofier & Pagés (1992).

2.1 Análisis de Correspondencias Múltiples (ACM)

Es una técnica factorial desarrollada para estudiar una población de individuos descritos por un conjunto de variables categóricas, con un determinado número de categorías cada una de ellas. Una de las aplicaciones más habituales del ACM es el tratamiento del conjunto de respuestas a una encuesta. En este caso, la metodología se basa en que cada pregunta constituye una variable cuyas categorías son las respuestas propuestas, entre las cuales cada encuestado debe de elegir una. Los datos así definidos son de tres tipos: individuos, variables y categorías. Sin embargo, integran una tabla única que se estudia en su conjunto.

En el ACM se establecen varios objetivos, proporcionando un análisis muy completo de este tipo de información. Por un lado, obtener una tipología de individuos a través de las clases definidas por las categorías. Se considera que dos clases se parecen tanto más cuanto más próximos son sus perfiles sobre el conjunto de las categorías. Asimismo, se obtiene una tipología de categorías, que permite estudiar la asociación mutua entre las categorías. Por otro lado, resumir el conjunto de variables por un pequeño número de variables numéricas, denominadas factores. Estos factores están relacionados con el conjunto de variables estudiadas.

2.2 Análisis Factorial Múltiple

Es un método factorial adaptado al tratamiento de tablas de datos en las que un mismo conjunto de individuos se describe a través de varios grupos de variables. En esta técnica los objetivos perseguidos no se limitan a la obtención de una tipología de los individuos definida a través del conjunto de variables, sino que se amplían a la búsqueda de posibles relaciones entre las estructuras obtenidas en el seno de cada uno de los grupos.

Respecto a los métodos factoriales clásicos, el AFM presenta algunas ventajas. Quizás la más destacable sea la de equilibrar la influencia de los grupos en el análisis global, ofreciendo además resultados típicos de la estructura en grupos del conjunto de variables.

El objetivo principal de este método es analizar la estructura común de las distintas tablas de datos objeto de estudio, poniendo de manifiesto cuáles son los elementos heterogéneos, es decir, aquéllos que tienen un comportamiento diferente al resto.

El AFM inicialmente realiza Análisis de Componentes Principales de cada grupo por separado, denominados análisis parciales. De estos análisis se estudian los valores propios, que ponen de manifiesto la dimensionalidad de la estructura interna de cada uno de ellos, comprobando, a la vez, el número de dimensiones de cada grupo que intervendrán con cierto peso en el análisis general.

Posteriormente, realiza el análisis simultáneo de los grupos previamente ponderados, que se denomina análisis global, tal como se ha indicado en la sección 2.

La ponderación de todas las variables que integran un grupo es idéntica, por lo que no se ve alterada la estructura interna asociada a cada uno de los grupos. Esta ponderación

no equilibra la inercia total, pero normaliza las inercias parciales. Al multiplicar la inercia de cada una de las subtablas por esa ponderación, en cada dirección del espacio (eje factorial), la inercia de la primera componente principal de cada grupo se iguala a 1.

Asimismo, se estudian, por un lado, los valores propios obtenidos en el análisis global. Estos valores proporcionan una primera información sobre las relaciones existentes entre los grupos analizados. La existencia de factores comunes puede ser detectada, igualmente, a través del cálculo de las correlaciones entre los factores globales y los factores del mismo orden correspondientes a cada uno de los grupos analizados, obtenidos en los respectivos análisis parciales.

El AFM es un método que permite el estudio tanto de la intra-estructura, esto es, estudio de la inercia de los puntos de cada nube (grupo) con respecto a su centro de gravedad, como de la inter-estructura o estudio comparativo de la proximidad entre los diferentes grupos. La suma de estas dos inercias es lo que se denomina inercia total. La razón inercia inter/inercia total es un índice que refuerza las conclusiones alcanzadas a través del análisis de las correlaciones antes mencionadas. En caso de estar próxima a 1 confirma el carácter de común para el factor en cuestión. Además, los puntos que representan al mismo individuo, desde los diferentes puntos de vista referentes a los grupos, estarán próximos entre sí.

Por último, señalar que el AFM proporciona resultados de gran interés desde el punto de vista de la interpretación. En un mismo gráfico pueden aparecer conjuntamente individuos, variables y, en su caso, modalidades. La riqueza interpretativa alcanza su máximo en los planos factoriales de las representaciones superpuestas de los individuos vistos a través de cada una de las tablas analizadas. Un estudio minucioso de estos planos permite detectar la tendencia general de los individuos, los grupos que definen en mayor medida esta tendencia, así como los individuos raros, es decir, aquéllos cuyo comportamiento es diferente según el grupo que los describe. Esta tendencia queda definida a partir de todos los grupos de variables, sin que haya grupos que queden eclipsados por la presencia de otros con mayor peso inicial.

3. Aplicación empírica

Con el objetivo de ilustrar la idoneidad de la metodología propuesta para el análisis de varios grupos de individuos en los que se ha medido información de naturaleza cualitativa, a continuación se presenta e interpretan los principales resultados obtenidos en una aplicación empírica a partir de una encuesta real.

Este análisis se enmarca en un estudio más ambicioso sobre la “Imagen de Marca y Viabilidad de una tienda de regalos con el logotipo de la UPV/EHU”. Los gestores de dicha universidad, como parte de un amplio proyecto de relanzamiento de su imagen, pensaron en una tienda, abierta al público, de productos corporativos con el logotipo de la misma. Antes de embarcarse en dicho proyecto, consideraron necesario conocer la aceptación que entre los actuales integrantes de la universidad tendría la tienda, tanto en general, como en cuanto a los productos (o, al menos, una parte de ellos) que en ella se van a vender.

Para ello, se realiz3 un muestreo estratificado con afijaci3n proporcional por campus, sexo y edad, tanto para el personal de administraci3n y servicios (PAS) como para el profesorado y personal investigador (PDI) y, por campus, g3nero y ciclo para los alumnos. El universo estaba compuesto por 48.995 alumnos, 1.128 PAS y 3.982 PDI. En concreto, se enviaron 2.289 invitaciones a realizar la encuesta a los alumnos, 768 al personal de administraci3n y servicios y 1.499 a los profesores. Las respuestas obtenidas fueron 547 para los alumnos (23.9% de tasa de respuesta), 444 para el PAS (57,81%) y 754 para el PDI (50.30%), tasas ciertamente elevadas para un encuesta de esta naturaleza. Los datos fueron obtenidos mediante una encuesta on-line que estuvo disponible para su realizaci3n durante un mes, entre febrero y marzo de 2005.

En la encuesta, adem3s de las preguntas de identificaci3n: Sexo: hombre/mujer; Edad: 18-22/23-29/30-44/>45; Estamento: Profesor/PAS/Estudiante; Campus: Bizkaia/Araba/Gipuzkoa, se preguntaba, de forma individual para cada uno de los 21 productos seleccionados (una vez enseada una foto del mismo y su precio de venta), la valoraci3n, en una escala de muy poco probable (1), poco probable (2), probable (3) o muy probable (4) la intenci3n de adquirir alguno de estos productos, ya sea para uso personal o para regalar.

La relevancia de estudiar la actitud respecto de la tienda corporativa, de los miembros de la instituci3n universitaria, en cualquiera de los estamentos considerados, (PAS, PDI y alumnos), es evidente; dado que dicha actitud ser3 reflejo de la identificaci3n que los mismos tienen con los principios, valores y objetivos de la propia instituci3n, medida a trav3s de la intenci3n que tales miembros tienen de poseer productos con el logotipo que la representa.

Este trabajo tiene como objetivo aportar resultados que ayuden a los gestores de la UPV/EHU en su elecci3n entre el establecimiento de una 3nica estrategia de marketing orientada a todos los estamentos internos de la universidad, dado su similar perfil respecto a la tienda de productos con el logotipo o, por el contrario, la necesidad de establecer diferentes estrategias dirigidas, respectivamente, a cada uno de los estamentos, dado su particular perfil.

El estudio de las posibles similitudes y/o diferencias entre los tres estamentos internos, en lo que a la intenci3n de comprar productos con el logo de la UPV/EHU se refiere, as3 como a su caracterizaci3n, se puede realizar a trav3s de t3cnicas factoriales. No obstante, se debe seleccionar un m3todo que permita tener en cuenta la estructura en tres tablas de naturaleza cualitativa y distinto tamao, en cuanto al n3mero de integrantes, que poseen los datos iniciales. Adem3s, debe permitir la ponderaci3n de la influencia de cada uno de los grupos de individuos, para que ninguno tenga mayor peso en los resultados y proporcionar medidas de las relaciones entre los mismos. La t3cnica de an3lisis propuesta en este trabajo resulta muy adecuada para estos objetivos.

El programa estadístico utilizado para la obtenci3n de resultados ha sido Syst3me Pour L'Analyse des Donn3es (SPAD).

Resultados Primera etapa-ACM

En los ACM previos, correspondientes a cada uno de los estratos, se han obtenido resultados similares en los ejes de mayor inercia, especialmente en los dos primeros. Los planos principales reflejan un efecto Guttman, tan habitual en este tipo de datos, que opone, por un lado, las valoraciones extremas entre sí (1= muy poco probable frente a 4= muy probable) y, por otro lado, éstas frente a las valoraciones medias (2= poco probable y 3= probable). Estos resultados parciales ponen de manifiesto las similitudes de los comportamientos de los tres grupos analizados y, por tanto, su idoneidad para ser analizados conjuntamente a través de técnicas factoriales de tablas múltiples.

Sin embargo, no es conveniente tomar decisiones exclusivamente a partir de estos resultados, ya que se trata de análisis individuales de cada grupo por separado, por lo que no hay medidas concretas de las similitudes existentes entre los mismos. Es necesario completar el análisis con una segunda etapa que permita confirmar, desde la óptica del análisis simultáneo de varias tablas, a través de indicadores apropiados, hasta qué punto la similaridad, que estos gráficos parecen vislumbrar en las estructuras de los estamentos analizados, existe.

Resultados Segunda etapa-AFM

En lo que respecta al número de factores que elegimos en la segunda etapa, señalar que dado que nuestros objetivos se centran en la búsqueda de posibles grupos de clientes en el seno de los colectivos universitarios, tanto en el ámbito puramente académico (profesores y alumnos) como en el ámbito de la administración (personal de administración y servicios), seleccionamos aquéllos ejes de dispersión que reflejen comunalidades entre los tres estamentos considerados, utilizando el criterio de ejes asociados a valores propios superiores a la unidad (ya que en el AFM, las tablas son tratadas desde la óptica cuantitativa).

La extensión del AFM que aquí se propone, es un método de análisis factorial de tablas múltiples que tiene la peculiaridad de equilibrar la influencia de cada una de las matrices analizadas en los resultados globales. Esto es, en los resultados presentados a continuación (sobre todo los correspondientes al primer factor) el peso de los tres estamentos es equivalente. Además, se ha elegido el análisis normado que equilibra la influencia de las variables. El grupo 1 corresponde a los profesores, el grupo 2 a los alumnos y el grupo 3 al personal de administración y servicios.

Los siguientes indicadores permiten confirmar y medir la gran similitud existente entre las tres tablas analizadas. Esto es, los tres estamentos analizados presentan un comportamiento muy similar en cuanto a la intención de compra de sus integrantes respecto de los productos con el logo de la UPV/EHU. Este resultado se mantiene sobre todo en los primeros factores globales, a los que contribuyen de manera bastante equilibrada los tres grupos y que reflejan una realidad presente en todos los estamentos analizados.

- Histograma de los valores propios de la tabla global : la proximidad a 3 (máximo valor posible) del primer valor propio es un claro indicador de las importantes

semejanzas que mantienen los tres grupos analizados (hay que recordar que debido a la ponderación propia del AFM, la inercia de la primera componente de cada una de las subtablas se iguala a 1)

Tabla 1

Valores propios del análisis global (5 primeros)

	<i>Valor propio</i>	<i>Porcentaje de inercia</i>	<i>Porcentaje acumulado</i>
1	2,865	26,57	26,57
2	1,752	16,25	42,82
3	1,570	14,56	57,38
4	1,560	14,47	71,85
5	1,195	11,14	82,99

- Los coeficientes RV constituyen otro indicador de la relación entre los tres estamentos, pudiendo comprobar que todos ellos son de una magnitud similar. Lógicamente la relación es mayor entre cada uno de los grupos y la tabla global.

Tabla 2

Coefficientes RV de relación entre los grupos

	<i>Grupo 1</i>	<i>Grupo 2</i>	<i>Grupo 3</i>	<i>AFM</i>
<i>Grupo 1</i>	1,000			
<i>Grupo 2</i>	0,627	1,000		
<i>Grupo 3</i>	0,679	0,593	1,000	
<i>AFM</i>	0,887	0,846	0,875	1,000

- La correlación entre los factores parciales de cada una de las tablas y los correspondientes a la tabla global muestran que ninguno de los grupos tiene mayor protagonismo en el análisis global. Además, el análisis global refleja, con gran fidelidad, en sus primeros ejes realidades que están presentes en los estamentos estudiados.

Tabla 3

Correlación entre los factores parciales y los factores del análisis global

	<i>Factor 1</i>	<i>Factor 2</i>	<i>Factor 3</i>
<i>Grupo 1</i>	0,99	0,96	0,86
<i>Grupo 2</i>	0,98	0,93	0,88
<i>Grupo 3</i>	0,98	0,96	0,81

- La razón inercia inter e inercia total señala, de forma resumida, la similitud o diferencia que presentan los grupos en los diversos ejes de variabilidad globales. Se

comprueba que en los dos primeros ejes los tres estamentos considerados manifiestan casi una misma estructura

Tabla 4

Razón inercia inter/inercia total en los tres primeros factores

<i>Factor 1</i>	<i>Factor 2</i>	<i>Factor 3</i>
0,96	0,90	0,72

Los indicadores numéricos analizados ponen de manifiesto que los resultados de mayor interés, dada su posterior utilización, corresponden a los dos primeros factores globales. Por ello, a continuación se presentan e interpretan los principales planos factoriales proporcionados por el AFM. Previamente, se presentan en la siguiente tabla las etiquetas utilizadas en los gráficos para su correcta interpretación.

Tabla 5

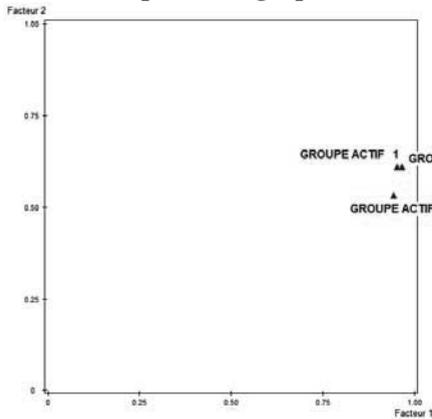
Etiquetas de las variables que contienen las valoraciones 1-4 de los productos

<i>Etiquetas</i>	<i>Artículo</i>	<i>Etiquetas</i>	<i>Artículo</i>	<i>Etiquetas</i>	<i>Artículo</i>
Backpack	Mochila	PenBlack	Bolígrafo negro	Trayplas	Vaciabolsillos de plástico
Bag	Bolso	PenBlue	Bolígrafo azul	TraySkin	Vaciabolsillos de piel
Cap	Gorra	PenSilve	Bolígrafo de plata	t-shirt	Camiseta
Cup	Taza	PenSWC	Bolígrafo en estuche	t-shirtV	Camiseta cuello en V
Keyring	Llavero	Pin	Pin de plata	Wallet	Monedero
Lighter	Mechero	Sculptur	Escultura	WatchSki	Reloj
MousePad	Alfombrilla ratón	Sweater	Sudadera	WatchMet	Reloj metálico

- El primer gráfico corresponde a una representación simplificada de los tres grupos a través de un solo punto asociado a cada uno de ellos. La posición de cada punto se corresponde con los pares correspondientes al coeficiente de ligazón \mathcal{L}_g (ver, p.ej., Escofier & Pagès, 1992) entre una tabla j y cada uno de los dos primeros factores globales del AFM, $(\mathcal{L}_g(K_j, F_1), \mathcal{L}_g(K_j, F_2))$, donde $j=1,2,3$. Su proximidad es una confirmación de la similitud en sus estructuras internas, a la vez que informa sobre la calidad de representación de cada grupo (etiquetas “GROUPE ACTIF” 1, 2 y 3) sobre los factores globales.

Figura 2

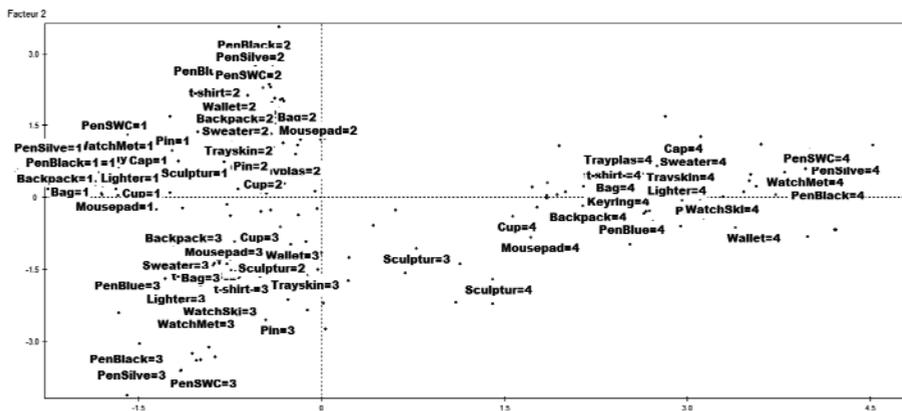
Plano Principal de los grupos



- El siguiente gráfico refleja la posición relativa de los individuos medios, esto es, las intenciones (desde muy poco probable hasta muy probable) media de compra de los distintos productos, desde el punto de vista del análisis múltiple. En él se puede observar que el análisis global, como media equilibrada de los tres estamentos, y dada la gran similitud entre los mismos, reproduce los planos obtenidos en los ACM parciales de cada uno de los grupos considerados. Así, el plano principal refleja claramente la existencia de tres grupos de respuesta frente a la posibilidad de comprar en la EHUDENDA productos con el logo de la institución universitaria: los “clientes” (cuadrantes 1 y 4) que manifiestan una clara intención de compra, los “posibles clientes” (cuadrante 3) que manifiestan una cierta intención de compra y los “dudosos clientes” (cuadrante 2), que manifiestan una poco o muy poco probable intención de compra.

Figura 3

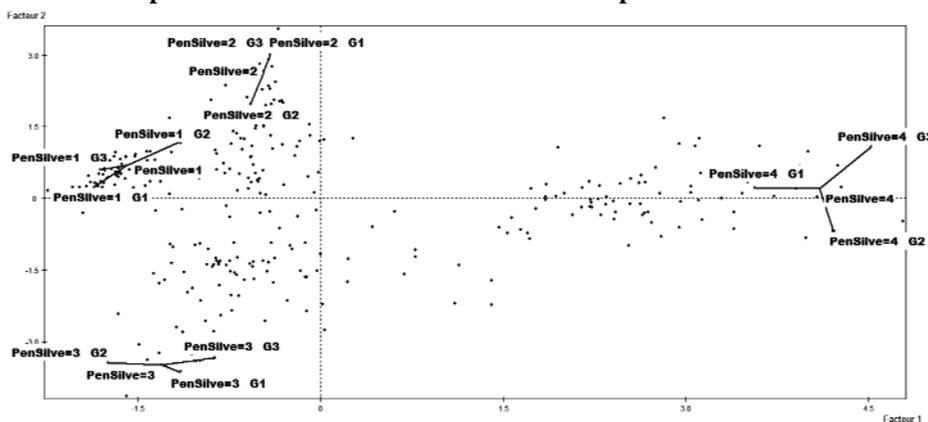
Plano Principal de los individuos medios



- El AFM también proporciona planos factoriales en los que se proyectan simultáneamente los llamados puntos parciales (asociados a cada uno de los grupos) y los puntos medios, de gran riqueza interpretativa ya que permite detectar los individuos responsables de las posibles diferencias entre los tablas analizadas. Sin embargo, en este estudio dicho gráfico constituye una confirmación más de las importantes similitudes que los tres estamentos presentan en lo que a intención de compra se refiere, proyectándose los tres puntos parciales y el punto medio muy próximos.

Figura 4

Plano Principal de los individuos medios e individuos parciales



En este plano se ha resaltado únicamente, con el objetivo de ilustrar el comentario del párrafo anterior, el resultado obtenido para uno de los artículos considerados, el bolígrafo de plata (PenSilve). De los 16 puntos señalados, 4 corresponden a los llamados “puntos medios” (correspondientes a la tabla global), cada uno de ellos asociado a una de las modalidades de respuesta sobre la probabilidad de compra: (1) muy poco hasta (4) muy probable. A su alrededor se proyectan, respectivamente, otros 3 puntos que corresponden a los llamados puntos parciales (asociados a cada uno de los grupos analizados, profesores G1, estudiantes G2 y personal de administración y servicios G3). Se puede observar la existencia de una mínima inercia intra (distancia) entre el punto medio y los puntos parciales de cada una de las posibles opciones sobre la intención de compra. Las distancias observadas no reflejan, en ningún caso, un distinto comportamiento entre los tres estamentos internos de la universidad, ya que los puntos correspondientes quedan proyectados en la misma zona del plano: la de los clientes (cuadrantes 1 y 4), la de los posibles clientes (cuadrante 2) y la de los dudosos clientes (cuadrante 3). Únicamente se deben a una más clara intención (o no intención) de compra por parte de alguno de los grupos. Siendo este resultado la tónica general para todos y cada uno de los artículos de la tienda institucional.

- El resto de variables de caracterización medidas en la encuesta: sexo, edad y campus se han utilizado de manera ilustrativa, con el fin de encontrar diferentes perfiles entre las tres clases de clientes resultantes. Sin embargo, todas ellas resultaron no significativas.

4. Conclusiones

El trabajo aquí presentado permite establecer dos tipos de conclusiones. Por un lado, las referentes al procedimiento metodológico:

- Este procedimiento, combinación de técnicas factoriales de tabla única (ACM) y de tabla múltiple (AFM), permite el tratamiento simultáneo, desde un punto de vista descriptivo y comparativo, de grupos de individuos en los que se ha medido la misma información a través de variables nominales. La naturaleza exploratoria de la técnica tiene la ventaja de dejar que sean los propios datos los que sugieran cuales son los factores más relevantes para cada grupo de individuos, cuál es su importancia relativa de cara a un análisis global y cómo de próximos están unos de otros (en definitiva, unas tablas de otras).
- La metodología proporciona una visión global de los datos, sin que ninguno de los grupos considerados tenga protagonismo ni quede eclipsado por el resto. Además, proporciona unos resultados con gran riqueza interpretativa, entre los que destacan las distintas medidas de relación entre los grupos analizados. Estas medidas ayudan a tomar decisiones sobre los grupos a mantener en un análisis global, por su similitud, o a analizar por separado dado su comportamiento específico y distinto al resto.
- Los individuos originales pierden su identidad. Sin embargo, es bien sabido que, habitualmente, en las encuestas los individuos de la muestra no tienen interés en sí mismos, siendo verdaderamente importantes las clases definidas a través de ellos en su diferente caracterización. De hecho, las categorías, que constituyen los nuevos individuos en la tabla analizada en la segunda etapa de la metodología propuesta, definen estas clases.

Por otro lado, las conclusiones referentes al estudio empírico:

- Hay que recordar, que el estudio presentado forma parte de un análisis más ambicioso sobre la “Imagen de Marca y Viabilidad de una tienda de regalos (Ehudenda) con el logotipo de la UPV/EHU”. Los resultados aquí obtenidos permiten completar y/o confirmar los obtenidos con otras técnicas, que utilizan ópticas diversas (ACM de la muestra total, PLS, Logit...), Fernández & al. (2011).
- La metodología propuesta ha proporcionado indicadores (numéricos y gráficos) que miden la similaridad entre los perfiles de los tres estamentos internos de la UPV/EHU (profesores, estudiantes y personal de administración y servicios). Similaridad, tanto en lo que respecta a la opinión frente al posible interés/no interés por la adquisición de productos con el logo de la institución, como en lo que se refiere a su caracterización.

- Los resultados han permitido, por un lado, confirmar la existencia de un mercado potencial para los productos de la tienda de regalos institucionales de la UPV/EHU. Así, entre los miembros de la institución universitaria se ha detectado un primer grupo con una clara actitud positiva respecto de la tienda corporativa, medida a través de la intención que tales miembros tienen de poseer productos con el logotipo que la representa. Un segundo grupo que ha manifestado una cierta disposición a comprar artículos de la tienda institucional, existiendo un tercer grupo con muy poca disposición. Por tanto, son resultados que ayudan a determinar la viabilidad de la tienda universitaria, dada la aceptación de los productos que en ella se van a comercializar.
- Por otro lado, la realidad sobre el similar perfil de los actuales integrantes de la comunidad universitaria, respecto a los productos institucionales, aconseja a los gestores de la universidad a establecer una única estrategia de marketing orientada a todos los estamentos internos de la universidad, sin distinguir género, edad o campus.

Referencias

- ABASCAL E., FERNÁNDEZ K., LANDALUCE M. Y MODROÑO J. (2001). «Diferentes aplicaciones de las técnicas factoriales de análisis de tablas múltiples en las investigaciones mediante encuestas». *Metodología de Encuestas*, Vol. 3, Nº 2, pp. 251-280
- ABASCAL FERNÁNDEZ, I. GARCÍA LAUTRE Y LANDALUCE CALVO M.I. (2004). «Análisis de la evolución a través de encuestas. Trayectoria electoral de las comunidades autónomas españolas en el periodo 1977-2004». *Metodología de Encuestas*, Vol. 6, Nº 2, pp. 147-162
- ABASCAL, E., GARCÍA-LAUTRE, LANDALUCE, I. (2006) «Multiple factor analysis of mixed tables of metric and categorical data». *Multiple Correspondence Analysis and related Methods*. Edited by Jörg Blasius and Michael Greenacre. Chapman & Hall (TAYLOR & FRANCIS GROUP. Book Chapter 15, pp 351-367
- ABASCAL, E., DÍAZ DE RADA, V., GARCÍA-LAUTRE, I. LANDALUCE, I. «Three-way analysis of categorical tables of different populations. The case of face-to-face and telephone survey». *7th International Conference on Social Science Methodology rc33-logig and Methodology in Sociology*. September 1-5 2008. Naples. Italy
- BÉCUE M. & PAGÈS J. (2001). «Analyse simultanée de questions ouvertes et de questions fermées. Méthodologie, exemple». *Journal de la SFdS*. vol. 142 num. 4 pp. 91-104
- BÉCUE M. & PAGÈS J. (2003). «A principal axes method for comparing contingency tables: MFACT». *Computational Statistics and Data Analysis*. vol. 45 num. 3 pp. 481-503

- BÉCUE, M. & PAGÈS, J. (2008). «Analysis of a mixture of quantitative, categorical and frequency data through an extension of multiple factor analysis. Application to survey data». *Computational Statistics and Data Analysis*. vol. 52 pp. 3255-3268
- ESCOFIER, B. & PAGES J. (1986). «Le traitement des variables qualitatives et tableaux mixtes par analyse factorielle multiple». *Data Analysis and Informatics*, IV(2), 179-191
- ESCOFIER, B. & PAGÈS, J. (1992) «Análisis factoriales simples y múltiples. Objetivos, métodos e interpretación». Servicio editorial de la Universidad de País Vasco
- ESCOFIER, B. & PAGÈS, J. (1994): «Multiple factor analysis (AFMULT package)». *Computational Statistics & Data Analysis* 18 121-140. North-Holland.
- ESCOFIER, B. & PAGÈS, J. (2008) «Analyses Factorielles Simples et Multiples. Objectifs, méthodes et interprétation». 4 ed. Dunod.
- ESCOUFIER, Y. (1973) «Le Traitement des Variables Vectorielles». *Biometrics* 24 (4), pp. 751-760
- FERNANDEZ-AGUIRRE, K., LANDALUCE, M.I., MARTIN, A. Y MODROÑO J.I. (2011). «Knowledge extraction from a large on-line survey: a case study for a higher education corporate marketing». *Journal of Applied Statistics*, Vol. 38, N° 11, pp. 2661-2679
- GARCÍA LAUTRE I. (2001) «Medición y análisis de las infraestructuras. Una nueva metodología basada en el Análisis Factorial Múltiple», Tesis doctoral, Departamento de Estadística e Investigación Operativa de la Universidad Pública de Navarra.
- GARCÍA LAUTRE, I. Y ABASCAL, E. (2003): «Una metodología para el estudio de la evolución de variables latentes. Análisis de las infraestructuras de carreteras de las comunidades autónomas (1975-2000)». *Estadística Española*, Vol. 45, N° 153, págs. 193 a 210.
- GREENACRE, M. (2007). *Correspondence Analysis in Practice, Second Edition*. London: Chapman & Hall/CRC.
- HUSSON F. & PAGÈS J. (2006). «Aspects méthodologiques du modèle Indscal». *Revue de Statistique Appliquée*. vol. LIV num. 2 pp. 83-100
- HUSSON F. & PAGÈS J. (2006). «Indscal Model: geometrical interpretation and methodology». *Computational Statistics and Data Analysis*. vol. 50 num. 2 pp. 358-378
- LANDALUCE, M^a I. (1995). «Estudio de la estructura de gasto medio de las Comunidades Autónomas españolas. Una aplicación del Análisis Factorial Múltiple». Tesis doctoral. Universidad del País Vasco
- LANDALUCE, M. I., FERNÁNDEZ, K., MODROÑO, J. I. (1999). «Reflexiones sobre el Uso Comparativo del Análisis Factorial Múltiple (AFM) y de la Metodología Statis para el Análisis de Tablas Múltiples». *Methodologica*, Vol. 7, pp. 37-65.

- LE DIEN S. & PAGÈS J. (2003). «Analyse Factorielle Multiple Hiérarchique». *Revue de Statistique Appliquée*. vol. LI num. 4 pp. 83-93
- LE DIEN S. & PAGÈS J. (2010). «DMFA: Dual Multiple Factor Analysis». *Communications in Statistics-Theory and Methods*. vol. 39:3 pp. 483-492
- MORAND E. & PAGÈS J. (2007). «L'analyse factorielle multiple procustéenne». *Journal de la Société Française de Statistique*. vol. 148 num. 2 pp. 65-97
- PAGÈS J. (1996). «Eléments de comparaison entre l'Analyse Factorielle Multiple et la méthode STATIS». *Revue de Statistique Appliquée*. vol. XLIV num. 4 pp. 81-95
- PAGÈS J. & TENENHAUS M. (2001). «Multiple factor analysis combined with PLS path modelling. Application to the analysis of relationships between physicochemical variables, sensory profiles and hedonic judgements». *Chemometrics and Intelligent Laboratory Systems*. vol. num. 58 pp. 261-273
- PAGÈS J. & TENENHAUS M. (2002). «Analyse factorielle multiple et approche PLS». *Revue de Statistique Appliquée*. vol. L num. 1 pp. 5-33
- PAGÈS J. (2002). «Analyse factorielle multiple appliquée aux variables qualitatives et aux données mixtes». *Revue de Statistique Appliquée*. vol. L num. 4 pp. 5-37
- PAGÈS J. (2004). «Analyse factorielle de données mixtes». *Revue Statistique Appliquée*. vol. LII num. 4 pp. 93-111
- PAGÈS J. (2005). «Analyse factorielle multiple et analyse procustéenne». *Revue de Statistique Appliquée*. vol. LIII num. 4 pp. 61-86
- PAGÈS, J. & CAMIZ S. (2008). «Analyse factorielle multiple de données mixtes : application à la comparaison de deux codages». *Revue de Modulad*. vol. 38 pp. 178-183
- PITMAN, T. (2000). «Perceptions of academics and students as customers: A survey of administrative staff in higher educations». *J. Higher Educ. Policy Manage*. 22 (2), pp. 165-176
- SPAD v5.5 (2000). «Système Portable d'Analyse des Données Numériques», CISIA, Montreuil, France.
- TRAVERSO, J. (2005). «Imagen interna de las Institución Universitaria. Modelo para el Personal de Administración y Servicios». *Revista de Economía y Empresa XXII* (54/55), pp. 95-112