

Análisis de Correspondencias y estudio de historias de vida: Una aplicación a la Encuesta de Transición Educativo-Formativa e Inserción Laboral¹

Julio Abad González

julio.abad@unileon.es

Pilar Blanco Alonso

pilar.blanco@unileon.es

Ana García Gallego

ana.ggallego@unileon.es

Universidad de León

Fac. de Ciencias Económicas y Empresariales

Departamento de Economía y Estadística

Campus de Vegazana, s/n

24071 León (España)

Recibido: Junio 2008

Aceptado: Junio 2008

En las últimas décadas, muchos son los trabajos que han puesto de manifiesto la importancia de adoptar una perspectiva dinámica a la hora de abordar el estudio de los procesos sociales. Por esta razón, de forma creciente, la información que se recoge para realizar este tipo de estudios refiere para cada individuo los tiempos transcurridos en

In the last decades, many papers have outlined the importance of adopting a dynamic perspective to analyze social processes. That is the reason why, for these studies, the collected data increasingly report durations of episodes and changes in time.

This kind of information, the so-called

¹ Parte de este artículo ha sido aceptada en las *XVII Jornadas de la Asociación de Economía de la Educación* (AEDE), a celebrar los días 10 y 11 de julio de 2008 en Santiago de Compostela.

diferentes estados o entre distintos eventos.

Este tipo de información dinámica, que se suele denominar "datos de historias de vida" (*event history data*), habitualmente se analiza utilizando procedimientos relacionados con el análisis de supervivencia o los modelos de regresión para datos de duración. Sin embargo, en este artículo se adopta un enfoque distinto al utilizar un método multivariante exploratorio, como es el Análisis de Correspondencias, y aplicarlo a la *super-matriz indicador* que recoge las historias de vida de los individuos como secuencias de eventos.

Para ilustrar la utilización de esta metodología, presentaremos una aplicación de la misma al análisis de datos procedentes de la *Encuesta de Transición Educativo-Formativa e Inserción Laboral* elaborada por el Instituto Nacional de Estadística en 2005. Dicha encuesta tiene por objetivo conocer las diferentes formas de transición desde la educación y la formación al mercado laboral mediante el estudio de las trayectorias formativas y laborales de 45.000 jóvenes entre los años 2001 y 2005.

Palabras clave: Análisis de Correspondencias, datos de historias de vida, transición al empleo.

'event history data', is usually analyzed by means of statistical methods such as survival analysis or regression models for transition data. However, in this paper a different approach is outlined applying correspondence analysis, an exploratory multivariate method, to a super indicator matrix that records event history data as sequences of events.

In order to illustrate its use, in this paper we apply this methodology to analyze data from an education-to-job transition survey (*Encuesta de Transición Educativo-Formativa e Inserción Laboral*) conducted by the Spanish Statistics Institute (Instituto Nacional de Estadística) in 2005. This survey is intended to study the different ways of transition from education to labour market for 45.000 young people during the period 2001-2005.

Key words: Correspondence Analysis, event history data, job transition.

1. INTRODUCCIÓN

Aunque desarrollado inicialmente para el estudio de tablas bidimensionales de contingencia, el Análisis de Correspondencias es una técnica estadística cuya aplicación puede ser útil en todas aquellas situaciones en que sea posible construir una matriz de datos positivos donde sea de interés comparar perfiles-fila, perfiles-columna o ambas cosas. Este hecho legitima su uso en otros ámbitos, y en particular, en el marco del estudio de datos de historias de vida o acontecimientos (*event history data*).

Este tipo de datos están siendo crecientemente utilizados en el ámbito de las ciencias económicas y sociales puesto que permiten adoptar una perspectiva dinámica a la hora de estudiar procesos de muy diversa naturaleza, tal y como queda de manifiesto en la revisión elaborada por Blossfeld *et al.* (1989: 14-17). Sin embargo, no es menos cierto que en el tratamiento de este tipo de información suelen emplearse métodos estadísticos como el análisis de supervivencia o los modelos de

regresión para datos de duración y no técnicas de reducción de datos como es el Análisis de Correspondencias.

Por todo ello, con el fin de mostrar la utilización de esta metodología en este ámbito, presentamos una aplicación de la misma al análisis de datos procedentes de la *Encuesta de Transición Educativo-Formativa e Inserción Laboral* (ETEFIL) elaborada por el Instituto Nacional de Estadística en 2005.

De acuerdo con este planteamiento, la estructura de este artículo es la siguiente:

- En el primer apartado se presentan sucintamente las principales características teóricas del Análisis de Correspondencias para, a continuación, mostrar los pasos a seguir en su aplicación cuando se analizan datos de historias de acontecimientos.
- El segundo apartado se refiere a los datos: en primer lugar, se hace una descripción general de los objetivos y características de la Encuesta de Transición Educativo-Formativa e Inserción Laboral (ETEFIL) y, seguidamente, se hace referencia a los datos que de ella se han extraído y que serán objeto de análisis.
- El tercer apartado muestra las distintas aplicaciones de esta técnica a los datos extraídos de la ETEFIL y describe los principales resultados obtenidos.
- Por último, se exponen las conclusiones más relevantes del trabajo realizado hasta la fecha ya que este estudio se inscribe en un proyecto más amplio cuyo desarrollo está aún en curso.

2. METODOLOGÍA

En este apartado presentaremos los fundamentos teóricos del Análisis de Correspondencias y mostraremos diferentes procedimientos que permiten su aplicación al estudio de historias de vida.

2.1. *Análisis de Correspondencias*

Comenzamos abordando de forma breve en qué consiste esta técnica de análisis estadístico y cuáles son sus objetivos y características.

No obstante, un desarrollo más exhaustivo, se puede encontrar en Benzécri (1992), Greenacre (1984 y 1993) o Jambu (1989: 193-328).

Como ya se ha señalado, el Análisis de Correspondencias es un método inicialmente adaptado para el tratamiento de tablas de contingencia -tablas de frecuencias bidimensionales- que permite estudiar las eventuales relaciones existentes entre las filas y columnas de dicha tabla a través de la representación gráfica simultánea de las mismas. No obstante, tal y como señala Jambu (1989: 252), su aplicación se puede extender (y de hecho, se ha extendido) al estudio de otros tipos de tablas de datos siempre que las nociones de perfiles-fila y perfiles-columna, a los que nos referiremos a continuación, tengan una interpretación natural.

Partimos, por tanto, de una tabla de contingencia K que se obtiene clasificando una población o muestra según dos variables: la variable I , con n modalidades: $I = \{1, 2 \dots i \dots n\}$, y la variable J , con p modalidades: $J = \{1, 2 \dots j \dots p\}$, de modo que cada elemento de la tabla k_{ij} indica la frecuencia conjunta de la fila i y de la columna j . Las frecuencias marginales son designadas por $k_{i.} = \sum_j k_{ij}$ para las filas y por $k_{.j} = \sum_i k_{ij}$ para las columnas, siendo $k = \sum_i \sum_j k_{ij}$ la frecuencia total de la tabla. Todas estas frecuencias, que son absolutas, pueden ser transformadas en frecuencias relativas f_{ij} a través de la relación $f_{ij} = k_{ij}/k$.

Tablas de perfiles-fila y de perfiles-columna

Dado que las distancias entre filas (o entre columnas) están condicionadas por sus masas, es necesario hacerlas comparables mediante su transformación en perfiles-fila y en perfiles-columna.

La tabla de perfiles-fila recoge las probabilidades condicionadas de presentar una determinada modalidad j de la variable J sabiendo que se presenta la modalidad i de la variable I ($f_{ij}/f_{i.}$). El conjunto de perfiles-fila forma una nube de n puntos en el espacio de las p columnas. Cada punto i tiene por coordenadas en \mathbb{R}^p : $\{f_{ij}/f_{i.}; j = 1, 2 \dots p\}$, y su masa $f_{i.}$ es igual a su frecuencia relativa. El centro de gravedad G_I de esta nube es la media de los perfiles-fila ponderados por sus masas, y se corresponde con el perfil-fila medio, es decir, con las frecuencias marginales de las columnas ($f_{.j}$).

Análogamente, la tabla de perfiles-columna recoge las probabilidades condicionadas de presentar una determinada modalidad i de la variable I sabiendo que se presenta la modalidad j de la variable J

(f_{ij}/f_j) . El conjunto de perfiles-columna constituye una nube de p puntos en el espacio de las n filas. Las coordenadas en \mathbb{R}^n del punto j son: $\{f_{ij}/f_j; i = 1, 2 \dots n\}$, y cada punto está afectado de una masa f_j . El centro de gravedad G_J de la nube de perfiles-columna coincide con el perfil-columna medio, y sus componentes son las frecuencias marginales de las filas (f_i).

Distancia χ^2 (ji-cuadrado)

Se define la distancia entre dos puntos-fila y entre dos puntos-columna como:

$$d_{\chi^2}^2(i, i') = \sum_{j=1}^p \frac{1}{f_j} \left(\frac{f_{ij}}{f_i} - \frac{f_{i'j}}{f_i} \right)^2 \quad \text{y} \quad d_{\chi^2}^2(j, j') = \sum_{i=1}^n \frac{1}{f_i} \left(\frac{f_{ij}}{f_j} - \frac{f_{i'j}}{f_j} \right)^2$$

Ésta es una distancia euclídea ponderada respectivamente por el inverso de la masa de la columna o de la fila para no favorecer a aquellas columnas o filas que tienen más masa. Así, la distancia χ^2 va a otorgar la misma importancia a las columnas en el cálculo de la distancia entre dos perfiles-fila, y a las filas, en el cálculo de la distancia entre perfiles-columna, independientemente de cuáles sean sus frecuencias relativas.

Inercia de la nube de puntos

La inercia de la nube de puntos-fila respecto a su centro de gravedad es una medida de la dispersión de la nube, y se calcula como suma ponderada de las distancias entre los puntos-fila y su centro de gravedad usando como ponderación la masa de cada punto-fila, y como métrica la de la distancia χ^2 :

$$H(I, G_I) = \sum_{i=1}^n f_i d_{\chi^2}^2(i, G_I) = \sum_{i=1}^n f_i \sum_{j=1}^p \frac{1}{f_j} \left(\frac{f_{ij}}{f_i} - f_j \right)^2 = \sum_{i=1}^n \sum_{j=1}^p \frac{(f_{ij} - f_i f_j)^2}{f_i f_j}$$

Análogamente se obtiene la inercia de la nube de puntos-columna respecto a su centro de gravedad:

$$H(J, G_J) = \sum_{j=1}^p f_j d_{\chi^2}^2(j, G_J) = \sum_{j=1}^p f_j \sum_{i=1}^n \frac{1}{f_i} \left(\frac{f_{ij}}{f_j} - f_i \right)^2 = \sum_{i=1}^n \sum_{j=1}^p \frac{(f_{ij} - f_i f_j)^2}{f_i f_j}$$

Como vemos, la inercia de la nube de puntos-fila coincide con la de la nube de puntos columna, y resulta ser igual al estadístico χ^2 de Pearson dividido entre la frecuencia total de la tabla k .

Por lo tanto, la inercia de la nube será tanto mayor cuanto mayor sea la distancia entre los puntos y su centro de gravedad o, planteado de otro modo, cuanto más se alejen las frecuencias de la condición de independencia.

Obtención de las dimensiones y coordenadas factoriales

El objetivo del Análisis de Correspondencias será determinar sendos subespacios que maximicen las inercias de las proyecciones de las respectivas nubes de puntos-fila y puntos-columna.

Para el ajuste de la nube de puntos-fila, si denotamos por $\psi_{\alpha i}$ la proyección del punto-fila i sobre el eje factorial o dimensión α , la dimensión 1 será aquella que maximice la inercia o varianza ponderada de las proyecciones de los puntos-fila sobre dicha dimensión ($\lambda_1 = \sum_i f_i \cdot \psi_{1i}^2$). La dimensión 2 también será aquella que maximice la inercia de las proyecciones de los puntos-fila sobre dicha dimensión ($\lambda_2 = \sum_i f_i \cdot \psi_{2i}^2$) pero con la restricción de que las coordenadas de los proyecciones de los puntos-fila sobre la dimensión 2 deben ser ortogonales a las obtenidas para la dimensión 1 ($\sum_i f_i \cdot \psi_{1i} \psi_{2i} = 0$). Las dimensiones sucesivas serán también aquellas que maximicen la inercia de las proyecciones con la restricción de ser ortogonales a las dimensiones anteriores.

El ajuste de la nube de puntos-columna se realizará de forma análoga. Si se denota por $\varphi_{\alpha j}$ la proyección del punto-columna j sobre el eje factorial o dimensión α , la dimensión 1 será aquella que maximice la inercia ponderada de las proyecciones de los puntos-columna sobre dicha dimensión ($\lambda_1 = \sum_j f_j \cdot \varphi_{1j}^2$) y las dimensiones siguientes serán esas que maximicen la inercia de las proyecciones pero con la restricción de ser ortogonales a todas las dimensiones obtenidas anteriormente.

Relaciones de transición o cuasi-baricéntricas

Aunque, como se acaba de señalar, el Análisis de Correspondencias da lugar a sendas soluciones para las nubes de los puntos-fila y de los puntos-columna, una característica importante de esta técnica es que ambas soluciones están estrechamente ligadas a través de

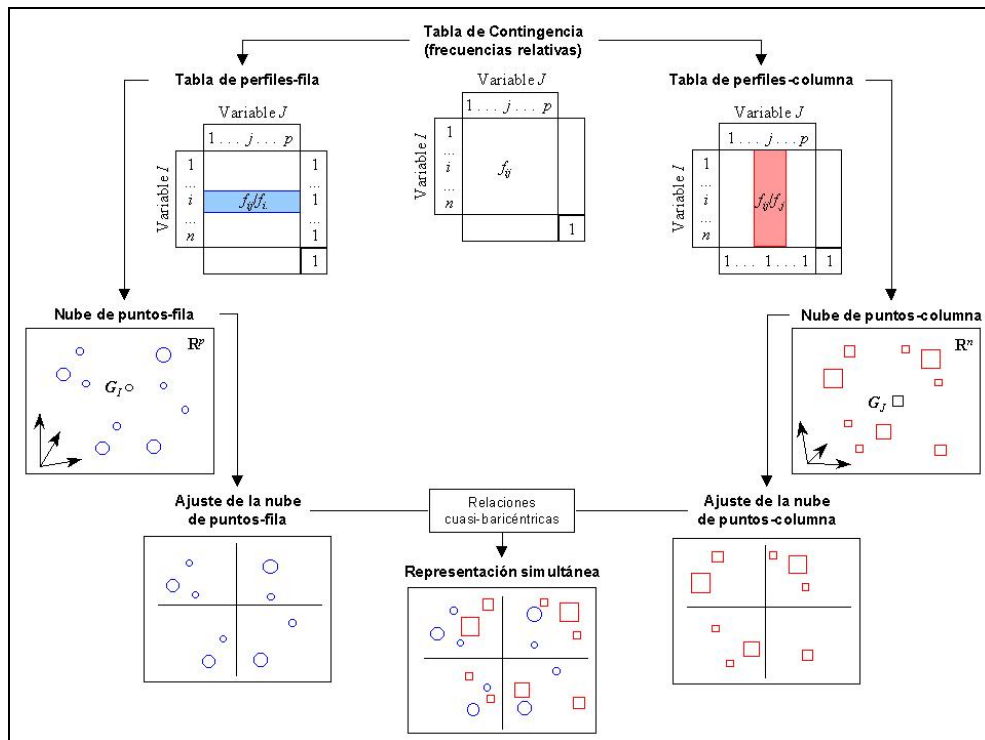
las denominadas relaciones de transición o cuasi-baricéntricas, que vienen dadas por las expresiones siguientes:

$$\Psi_{ai} = \frac{1}{\sqrt{\lambda_\alpha}} \sum_{j=1}^p \frac{f_{ij}}{f_i} \Phi_{aj} \quad \text{y} \quad \Phi_{aj} = \frac{1}{\sqrt{\lambda_\alpha}} \sum_{i=1}^n \frac{f_{ij}}{f_j} \Psi_{ai}$$

Estas relaciones permiten obtener, para cada dimensión, las coordenadas de los puntos-fila a partir de las coordenadas de los puntos-columna, y viceversa. Las relaciones de transición prueban también la igualdad de las dimensiones del ajuste de la nube de puntos-fila y de la nube de puntos-columna, siendo el número total de dimensiones extraídas igual al mínimo de $n-1$ y $p-1$.

Por último, hay que destacar que la existencia de estas relaciones cuasi-baricéntricas, justifican la representación gráfica simultánea de las filas y de las columnas de la tabla, y constituyen una de las propiedades más destacables del Análisis de Correspondencias.

Figura 1: Esquema del Análisis de Correspondencias



La Figura 1 muestra un esquema del proceso que sigue el Análisis de Correspondencias, desde la tabla de contingencia hasta llegar a la representación simultánea de puntos-fila y puntos-columna.

Reglas de interpretación

Las nubes de puntos-fila y de puntos-columna se representan en los planos formados por los primeros ejes factoriales tomados de dos en dos, y para su interpretación es necesario tener en cuenta los siguientes elementos:

- *Tasas de inercia*. Dado que la suma de las inercias asociadas a cada una de las dimensiones extraídas es igual a la inercia total de la nube ($\sum_{\alpha} \lambda_{\alpha} = H = \chi^2/k$), se utiliza la tasa de inercia $\tau_{\alpha} = \lambda_{\alpha} / \sum_{\alpha} \lambda_{\alpha}$ para medir la proporción de la inercia total de la nube que queda explicada por la dimensión α y, por tanto, para valorar la importancia de la misma.
- *Contribuciones absolutas*. Expresan en qué medida contribuye una fila i o una columna j a la inercia proyectada sobre cada dimensión. Se define como la proporción de la inercia explicada por el factor α que corresponde a la fila i ($C_{\alpha}(i) = f_{i.} \psi_{\alpha i}^2 / \lambda_{\alpha}$) o a la columna j ($C_{\alpha}(j) = f_{.j} \phi_{\alpha j}^2 / \lambda_{\alpha}$).
- *Contribuciones relativas o cosenos cuadrados*. Informan de si un punto está bien representado sobre un eje factorial. La contribución relativa de la dimensión α al punto-fila i (o al punto-columna j) es el porcentaje de la inercia de a la fila i (o la columna j) que queda explicado por la dimensión α ($\text{Cos}^2_{\alpha}(i) = \psi_{\alpha i}^2 / \sum_{\alpha} \psi_{\alpha i}^2$ y $\text{Cos}^2_{\alpha}(j) = \phi_{\alpha j}^2 / \sum_{\alpha} \phi_{\alpha j}^2$, respectivamente).

Elementos suplementarios

Una última característica interesante del Análisis de Correspondencias es la posibilidad de representar elementos suplementarios en los planos factoriales obtenidos del análisis. Se trata de añadir filas o columnas a la tabla base para, una vez obtenidos los ejes factoriales, obtener sus coordenadas factoriales y representarlas junto al resto de elementos. A estos elementos se les denomina también ilustrativos, dado que, a diferencia del resto de elementos -denominados activos-, no intervienen en la determinación de los ejes factoriales, sino que se

introducen a posteriori con el fin de "ilustrar" o facilitar la interpretación de las representaciones gráficas.

Las coordenadas factoriales de la fila suplementaria i^+ y de la columna suplementaria j^+ se obtienen a partir de las relaciones de transición antes comentadas, resultando, respectivamente, las expresiones siguientes:

$$\Psi_{ai^+} = \frac{1}{\sqrt{\lambda_\alpha}} \sum_{j=1}^p \frac{f_{i^+j}}{f_{i^+}} \varphi_{aj} \quad \text{y} \quad \varphi_{aj^+} = \frac{1}{\sqrt{\lambda_\alpha}} \sum_{i=1}^n \frac{f_{ij^+}}{f_{j^+}} \Psi_{ai}$$

Donde f_{i^+j}/f_{i^+} y f_{ij^+}/f_{j^+} son, respectivamente, el perfil-fila de la fila suplementaria i^+ y el perfil-columna de la columna suplementaria j^+ .

2.2. Extensiones del Análisis de Correspondencias: aplicación al estudio de historias de vida

Como ya hemos apuntado, el Análisis de Correspondencias puede extenderse a otros tipos de tablas que no son *strictu sensu* de contingencia, siempre y cuando siga teniendo sentido estudiar las diferencias entre los perfiles-fila, entre los perfiles-columna, o entre los dos.

En este sentido, una de las generalizaciones más simples y, probablemente la más utilizada, es el ANÁLISIS DE CORRESPONDENCIAS MÚLTIPLES (ACM) -a la versión a la que nos hemos referido anteriormente se le suele denominar Análisis de Correspondencias Simples (ACS)-. Esta técnica consiste en aplicar el Análisis de Correspondencias a una *tabla disyuntiva completa*, es decir, una tabla cuyas filas corresponden a individuos, y que tiene por columnas las distintas modalidades -codificadas como indicadores de presencia/ausencia (unos y ceros)- de un conjunto de variables cualitativas.

Asimismo, también tiene sentido la utilización del Análisis de Correspondencias en el estudio de datos de historias de vida. Los datos de historias de vida o datos de historias de acontecimientos (*event history data*) son registros biográficos en los que, para cada individuo de la muestra, se recoge no sólo una secuencia con los estados en que se encuentra éste a lo largo del periodo de estudio, sino también cuánto tiempo se encuentra en cada uno de ellos.

Cada una de estas historias individuales puede ser codificada en términos de frecuencias del siguiente modo: cada frecuencia indicaría el tiempo que cada individuo ha estado en cada uno de los distintos

estados mutuamente excluyentes considerados durante el periodo de estudio (por ejemplo, cuántos meses ha estado un individuo estudiando, trabajando o desempleado durante un año).

Si no sólo es de interés el tiempo transcurrido en cada estado, sino también el orden en que los diferentes estados se van sucediendo en el tiempo, se construye una nueva tabla yuxtaponiendo las m subtablas obtenidas de la división del periodo de estudio en m subperiodos. En último término, si la unidad de medida del tiempo coincide con el subperiodo definido, se llegaría a una "supermatriz de indicadores" de presencia/ausencia que indicaría si el individuo estaba en ese estado durante ese subperiodo o no.

La Figura 2 recoge ambos tipos de codificación para un ejemplo con tres individuos: la "tabla agregada" muestra como el individuo 1 permaneció 6 meses estudiando (E), 3 trabajando (T) y 3 desempleado (D); esa misma información aparece desglosada por meses de modo que cada patrón 100, 010 ó 001 indica en qué situación está el individuo ese mes (estudiando, trabajando o desempleado, respectivamente). Así, el individuo 1 entre enero y junio estudiaba, de julio a septiembre estaba desempleado, y entre octubre y diciembre tenía trabajo.

Figura 2: Tabulación de historias de vida para Análisis de Correspondencias

Tabla agregada				Supermatriz de Indicadores												
				ene	feb	mar	abr	may	jun	jul	ago	sep	oct	nov	dic	
i	E	T	D	i	ETD	ETD	ETD	ETD	ETD	ETD	ETD	ETD	ETD	ETD	ETD	
1	6	3	3	1	100	100	100	100	100	100	001	001	001	010	010	010
2	9	0	3	2	100	100	100	100	100	100	100	100	100	001	001	001
3	0	6	6	3	001	001	001	001	001	001	010	010	010	010	010	010

El tratamiento de estas tablas nos permitirá determinar qué individuos y qué estados tienen perfiles semejantes (o diferentes) en función de qué estado han ocupado durante cada subperiodo (mediante un Análisis de Correspondencias Múltiples de la supermatriz de indicadores) o en función de cuánto tiempo han pasado en cada estado (mediante un Análisis de Correspondencias Simples de la tabla agregada).

Además y previamente al análisis de esta tabla, van der Heijden y de Leeuw (1989) proponen que se aplique el Análisis de Correspondencias Simples para estudiar la matriz subperiodos×estados,

tabla de contingencia cuyas frecuencias indican cuántos individuos se encuentran en cada estado durante cada uno de los subperiodos de tiempo definidos (en la Figura 3 se muestra la tabla subperiodos×estados asociada a los datos del ejemplo de la Figura 2).

Figura 3: Tabla subperiodos×estados para Análisis de Correspondencias

	ene	feb	mar	abr	may	jun	jul	ago	sep	oct	nov	dic
<i>E</i>	2	2	2	2	2	2	1	1	1	0	0	0
<i>T</i>	0	0	0	0	0	0	1	1	1	2	2	2
<i>D</i>	1	1	1	1	1	1	1	1	1	1	1	1

Por lo tanto, de acuerdo con estos autores, tres son los tratamientos a realizar cuando se utiliza el Análisis de Correspondencias para estudiar datos de historias de vida:

- a) Un ACS de la matriz subperiodos×estados que permitirá definir el "perfil medio" de la nube.
- b) Un ACM de la supermatriz de indicadores que revelará qué individuos y qué estados se alejan del "perfil medio" establecido por el análisis anterior.
- c) Un ACS de la "tabla agregada" individuos×estados que puede ser considerado como una versión con restricciones del análisis del punto anterior.

Si las configuraciones de los perfiles obtenidas en los análisis b) y c) son semejantes, concluiremos que la información proporcionada por la supermatriz de indicadores utilizada en el procedimiento b) no aporta una información significativamente diferente de la que recoge la "tabla agregada" analizada en c). En esa situación, el procedimiento c) resultará preferible al b) puesto que, proporcionando unos resultados similares, se gana en la estabilidad de los mismos (van der Heijden *et al.* 2000: 41).

3. DATOS: ETEFIL-2005

La aplicación que presentamos se ha realizado a partir de micro-datos proporcionados por el Instituto Nacional de Estadística (INE)

correspondientes a la *Encuesta de Transición Educativo-Formativa e Inserción Laboral* (ETEFIL-2005)². De acuerdo con su resumen metodológico, "es la primera investigación estadística de esta magnitud desarrollada en España" dedicada al "estudio del conjunto de itinerarios seguidos dentro del sistema educativo y las transiciones entre el estudio y el trabajo". Esta encuesta es fruto de la colaboración entre el Ministerio de Educación y Ciencia, el Ministerio de Trabajo y Asuntos Sociales y el propio INE, que fue el encargado de llevarla a cabo en el año 2005.

3.1. *Población, muestra y periodo de referencia*

La población objeto de la encuesta está constituida por todas aquellas personas que, siendo menores de 25 años a 31/12/2001, terminaron estudios del sistema educativo no universitario durante el curso 2000/01, abandonaron la educación secundaria obligatoria sin obtener la titulación en ese mismo curso, o completaron programas del sistema de formación ocupacional en 2001.

La muestra final está formada por 45.620 individuos y es el resultado de la agregación de las siete muestras correspondientes a cada uno de los siete colectivos siguientes: alumnos que abandonaron el sistema educativo sin el título de graduado en Educación Secundaria Obligatoria (AESO), graduados en Educación Secundaria Obligatoria (ESO), graduados en Bachillerato (BACH), graduados en Ciclos Formativos de Formación Profesional de Grado Medio (CFGM), y de Grado Superior (CFGS), alumnos que finalizaron un curso del Plan Nacional de Formación e Inserción Profesional (FIP), y alumnos que finalizaron un programa de Escuelas Taller y Casas de Oficios (ETCO).

Aunque el trabajo de campo se realizó entre abril y septiembre de 2005, el periodo de referencia es único para cada individuo, dado que se trata de un estudio longitudinal y retrospectivo acerca de lo realizado por el encuestado desde que finalizó o abandonó el sistema educativo hasta la fecha de la entrevista. No obstante, y con el fin de homogeneizar todas las observaciones, nos hemos centrado en los años 2002 a 2004, periodo durante el cual todos los individuos se encuentran en observación.

² http://www.ine.es/prodyser/micro_etefil.htm. Consultado el 17/10/2007.

3.2. Variables estudiadas en el Análisis de Correspondencias

Como ya se ha indicado, uno de los resultados más interesantes del Análisis de Correspondencias son las representaciones gráficas a que da lugar. Sin embargo, este aspecto gráfico de la técnica no permite incluir una gran cantidad de variables en el análisis, puesto que la profusión de puntos podría dar lugar a una mayor confusión en vez de a una clarificación de los resultados. Por ello, hemos optado inicialmente por restringir el estudio a las siguientes variables por ser las de mayor interés: el mes de observación (36 meses entre enero de 2002 y diciembre de 2004), el colectivo de pertenencia (los 7 colectivos considerados a los que nos referimos al hablar de la muestra), y la situación básica del encuestado (ver Tabla 1), que recoge de forma exhaustiva los 12 estados mutuamente excluyentes en que un individuo puede estar en cada mes del periodo de referencia.

Tabla 1: Categorías de la variable "situación básica del encuestado"

ETIQUETA	CATEGORÍA
SE T>20	Estudia en el sistema educativo (SE) y trabaja más o igual de 20 h. semanales
SE T<20 BE	Estudia en el SE, trabaja menos de 20 h. semanales y busca empleo
SE T<20 NBE	Estudia en el SE, trabaja menos de 20 h. semanales y no busca empleo
SE NT BE	Estudia en el SE, no trabaja y busca empleo
SE NT NBE	Estudia en el SE, no trabaja y no busca empleo
NSE T>20	No estudia en el SE y trabaja más o igual de 20 h. semanales
NSE T<20 BE	No estudia en el SE, busca empleo y trabaja menos de 20 h. semanales.
NSE T<20 NBE OE	No estudia en el SE, trabaja menos de 20 h. semanales, no busca empleo y recibe algún tipo de educación/formación fuera del SE
NSE T<20 NBE NOE	No estudia en el SE, trabaja menos de 20 h. semanales, no busca empleo y no recibe ningún tipo de educación/formación fuera del SE
NSE NT BE	No estudia en el SE, no trabaja y busca empleo
NSE NT NBE OE	No estudia en el SE, no trabaja, no busca empleo y recibe algún tipo de educación/formación fuera del SE
NSE NT NBE NOE	No estudia en el SE, no trabaja, no busca empleo y no recibe ningún tipo de educación/formación fuera del SE

Asimismo, como ilustrativas se incluyen, además del sexo del individuo, otras variables que surgen al agrupar algunas de las situaciones descritas en la Tabla 1: si estudia en el sistema educativo o no, si trabaja o no, y en su caso, si trabaja menos de 20 h./semana o no, y si busca empleo o no.

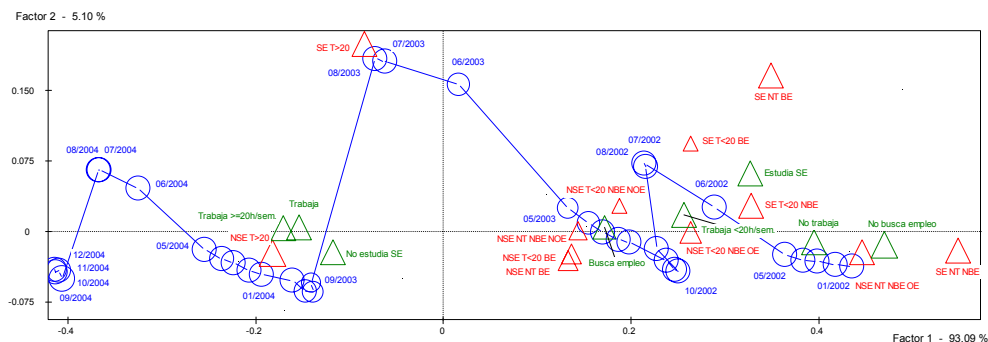
4. APLICACIONES Y RESULTADOS

En este apartado se muestran los principales resultados de la aplicación del Análisis de Correspondencias a las distintas tablas en que se organizan los datos extraídos de la ETEFIL-2005 de acuerdo con el procedimiento descrito en el Apartado 2.2.

4.1. ACS de la tabla subperiodos×estados

El primer análisis se efectúa sobre la tabla de contingencia que cruza subperiodos y estados y cuyas frecuencias recogen cuántos individuos se encuentran en cada una de las 12 situaciones básicas en cada uno de los 36 meses considerados. A esta tabla, hemos añadido como columnas suplementarias las agrupaciones de las situaciones básicas. Los resultados del análisis, que se detallan en el Anexo A1, se sintetizan en el Gráfico 1, que recoge la representación simultánea de filas y columnas sobre el primer plano factorial (que explica un 98,19% de la inercia total de la nube). Al igual que en los gráficos siguientes, el tamaño del icono de cada punto -salvo cuando éste corresponde a un individuo- es proporcional a su contribución relativa al plano, por lo que las modalidades con iconos pequeños estarán mal representadas en el plano y no han de ser tenidas en cuenta en la interpretación del mismo.

Gráfico 1: Primer plano factorial del ACS de la tabla subperidos×estados



La primera dimensión, a la que corresponde la mayor parte de la inercia (93,09%), queda muy identificada con la secuencia temporal ya que los meses se ordenan prácticamente de forma cronológica en el sentido siguiente: desde enero de 2002, en el extremo positivo, hasta

diciembre de 2004, en el extremo negativo. Desde el punto de vista de las columnas, las situaciones con mayores contribuciones absolutas son "SE NT NBE" (*estudia en el SE, no trabaja y no busca empleo*), con un 59,8%, en el extremo positivo, y "NSE T>20" (*no estudia en el SE y trabaja más o igual de 20h. semanales*), con un 26,5% en el negativo. Por lo tanto, este primer eje muestra la transición de una etapa de continuación de estudios (año 2002, sobre todo el primer semestre) a otra de trabajo a tiempo completo (año 2004, sobre todo el segundo semestre). El segundo factor es mucho menos importante que el primero pero aporta una información complementaria. Desde el punto de vista de las filas, separa en positivo los meses de junio, julio y agosto de los tres años (sobre todo de 2003) del resto de meses. Respecto de las columnas, los estados que más contribuyen son "SE T>20" (*estudia en el SE y trabaja más o igual de 20h. semanales*), con un 78,3% y, en menor medida, "SE NT BE" (*estudia en el SE, no trabaja y busca empleo*), con un 10,2%. Estas dos situaciones se ubican en el extremo positivo de esta dimensión, frente al resto de situaciones cuyas coordenadas en este factor son negativas o próximas al origen. Esta dimensión refleja la actividad laboral de los meses de verano en los que los encuestados conjugan estudio y trabajo (2003) o estudio y búsqueda de empleo (2002).

4.2. ACM de la matriz de indicadores

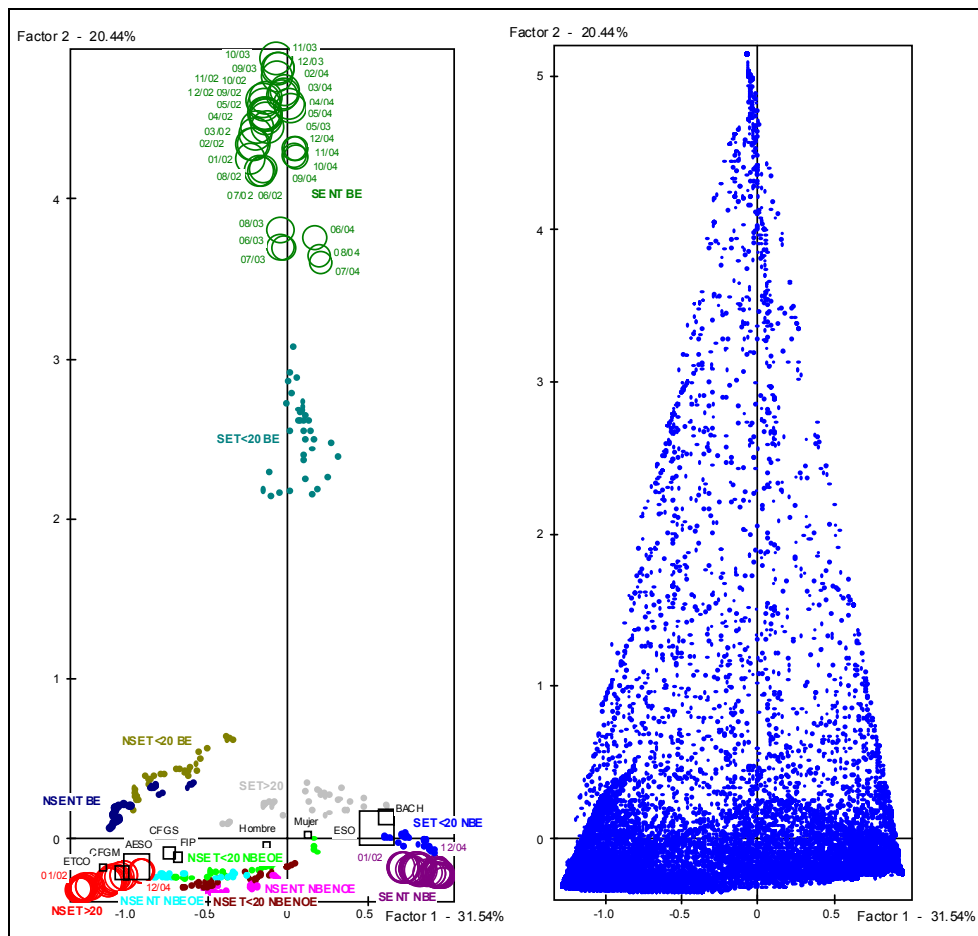
El segundo paso consiste en aplicar un ACM a la supermatriz de indicadores -incluyendo las variables *colectivo* y *sexo* como ilustrativas- para encontrar las situaciones que se alejan del "perfil medio" descrito por el análisis anterior. Los cuatro primeros factores explican³ un 76,02% de la inercia total, correspondiendo un 51,98% al primer plano factorial, que es el que se muestra en el Gráfico 2 (en el plano de la izquierda se representan las columnas o situaciones-periodo, y en el de la derecha, las filas o individuos). El plano con los factores 3 y 4, que explica un 24,04% adicional, se muestra de forma análoga en el Gráfico 3.

En los planos de la izquierda de los Gráficos 2 y 3 (situaciones-periodo), se han representado de un mismo color los puntos correspondientes a los 36 meses de cada una de las 12 situaciones básicas. A

³ Dado que los autovalores del ACM dan una idea muy pesimista de la inercia extraída en el análisis, las tasas de inercia se han recalculado de acuerdo con la fórmula propuesta por Benzécri (1992: 411-412).

diferencia del resto de gráficos en los que el tamaño del icono correspondiente a cada punto es proporcional a su contribución relativa al plano, en los planos de la derecha de ambos gráficos (individuos), el tamaño de los iconos es fijo.

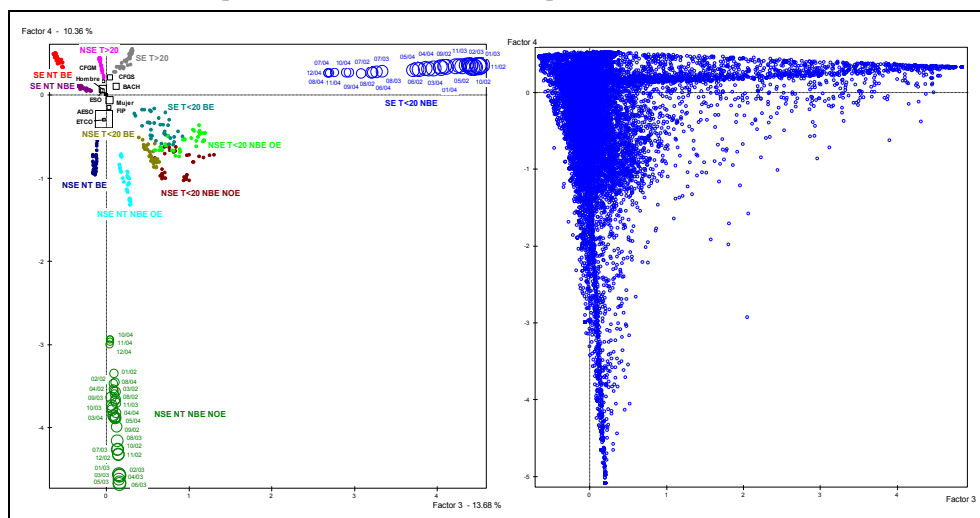
Gráfico 2: Representación de situaciones-periodo e individuos sobre el 1º plano factorial del ACS de la supermatriz de indicadores



El Gráfico 2 muestra como el primer plano factorial tiene una disposición "triangular": los individuos se distribuyen dentro de un triangulo cuyos vértices se sitúan en el extremo positivo del primer factor -caracterizado por "SE NT NBE" (*estudia en el SE, no trabaja y no busca empleo*)-, en el extremo negativo del primer factor -identificado con "NSE

T>20" (*no estudia en el SE y trabaja más o igual de 20 h. semanales*)-, y en el extremo positivo del segundo factor -definido por "SE NT BE" (*estudia en el SE, no trabaja y busca empleo*)-.

Gráfico 3: Representación de situaciones-periodo e individuos sobre el 2º plano factorial del ACS de la supermatriz de indicadores



Por su parte, el segundo plano queda definido por las situaciones "SE T<20 NBE" (*estudia en el SE, trabaja menos de 20 h. semanales y no busca empleo*), en el extremo positivo del tercer factor, y "NSE NT NBE NOE" (*no estudia en el SE, no trabaja, no busca empleo, y no recibe ningún tipo de educación/formación fuera del SE*), en el extremo negativo del cuarto factor.

Por lo tanto, estas cinco situaciones básicas son aquellas cuyo perfil difiere más del "perfil medio" descrito en el Apartado 4.1, dado que, en conjunto, son las que más se alejan del origen. El que aparezcan próximos entre sí los puntos correspondientes a los meses de una misma situación indica que una mayoría de los individuos que están en cada una de esas situaciones permanecen en ella durante gran parte del periodo de observación (por ejemplo, *estudia en el SE, no trabaja y no busca empleo* -"SE NT NBE"- y *no estudia en el SE y trabaja más o igual de 20 h. semanales* -"NSE T>20"-). Sin embargo, en las otras tres situaciones se observa que los valores más próximos al origen corresponden a los últimos

meses de 2004 (*no estudia en el SE, no trabaja, no busca empleo y no recibe ningún tipo de educación/formación fuera del SE* -"NSE NT NBE NOE"-), a los meses de verano (*estudia en el SE, no trabaja y busca empleo* -"SE NT BE"-) o a ambos (*estudia en el SE, trabaja menos de 20 h. semanales y no busca empleo* -"SE T<20 NBE"), lo cual indica que ocupar durante estos meses esas situaciones se aleja menos del perfil medio que hacerlo en el resto de meses.

4.3. ACS de la tabla agregada individuos×situaciones

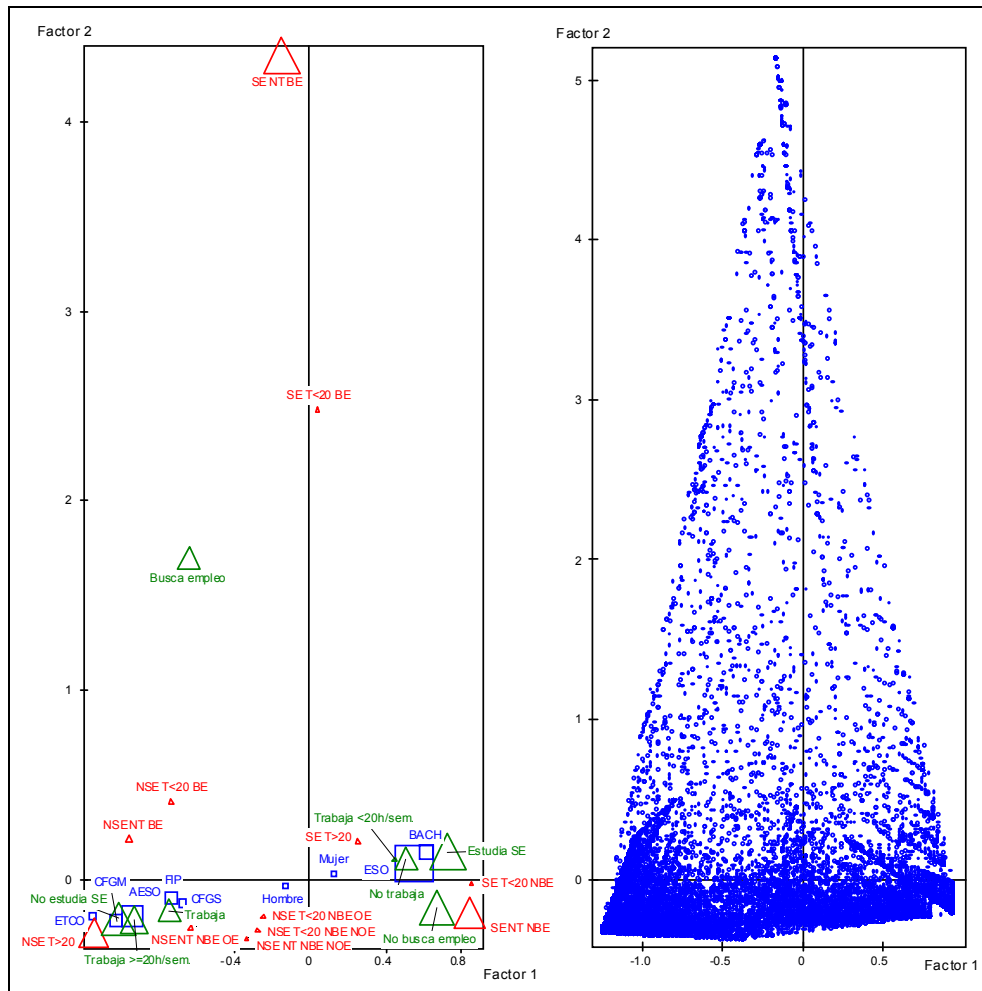
El tercer tratamiento realizado es un Análisis de Correspondencias Simples sobre la tabla que cruza los individuos (en filas) y las 12 situaciones básicas (en columnas) y que recoge el número de subperiodos (meses) que cada individuo ha pasado en cada una de las 12 situaciones (también se incluyen como ilustrativas las agrupaciones de las situaciones básicas así como las variables *colectivo* y *sexo*). Tal y como se ha señalado en el Apartado 2.2, este análisis viene a ser equivalente al comentado en el Subapartado 4.2 con la restricción de que, en lugar de proyectar por cada situación 36 puntos (uno por mes), sólo se proyectará uno. Los resultados de este análisis se recogen en el Anexo A2.

Como se puede observar, las representaciones, tanto de columnas como de filas, en el primer plano factorial, son muy semejantes a las homólogas resultantes del análisis anterior (Gráfico 2).

El primer plano factorial tiene la misma disposición triangular puesto que los individuos se distribuyen dentro de un triángulo cuyos vértices se sitúan:

- En el extremo positivo del primer factor, caracterizado por "SE NT NBE" (*estudia en el SE, no trabaja y no busca empleo*), con una contribución absoluta al mismo del 40,1%.
- En el extremo negativo del primer factor, donde se sitúa "NSE T>20" (*no estudia en el SE y trabaja más o igual de 20 h. semanales*), que contribuye al primer factor un 47,7%.
- En el extremo positivo del segundo factor, definido por la situación "SE NT BE" (*estudia en el SE, no trabaja y busca empleo*) que genera prácticamente todo el segundo factor dado que presenta una contribución absoluta al mismo del 90,8%.

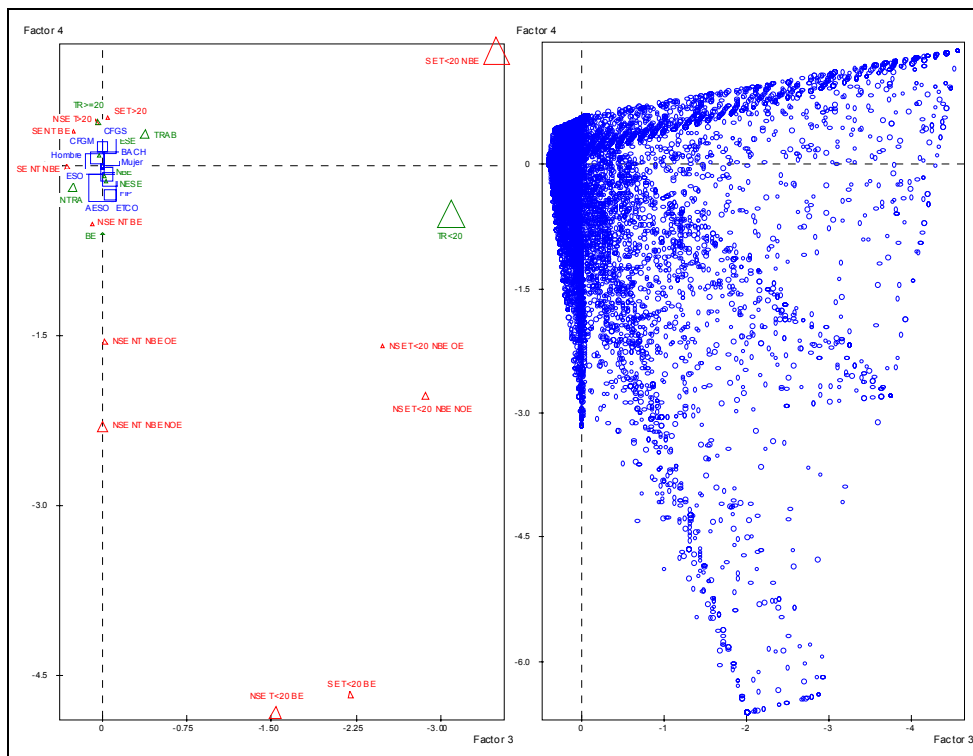
Gráfico 4: Representación de las columnas (situaciones) y filas (individuos) sobre el 1º plano factorial del ACS de la tabla agregada individuos×situaciones



El segundo plano factorial, que se muestra en el Gráfico 5, también guarda una gran similitud con los representados en el Gráfico 3 en lo que al tercer factor se refiere. Este factor queda básicamente caracterizado por la situación "SE T<20 NBE" (*estudia en el SE, trabaja menos de 20 h. semanales y no busca empleo*), con una contribución del 75,3%. En cuanto al cuarto factor, esta similitud no es tan acusada puesto que, además de la situación "NSE NT NBE NOE" (*no estudia en el SE, no trabaja, no busca empleo, y no recibe ningún tipo de educación/formación fuera del SE*) -con una contribución al factor del 26,3%- , intervienen también en su formación

los estados "NSE T<20 BE" (*no estudia en el SE, trabaja menos de 20 h. semanales y busca empleo*), con un 26,0%, y "SE T<20 BE" (*estudia en el SE, trabaja menos de 20 h. semanales y busca empleo*), con un 11,1%, todos ellos en el extremo negativo del factor.

Gráfico 5: Representación de columnas (situaciones) y filas (individuos) sobre el 2º plano factorial del ACS de la tabla agregada individuos×situaciones



Por lo tanto, cabe señalar que, aunque la matriz tratada por el ACM en el Apartado 4.2 ofrece una información más desagregada que la tabla analizada mediante ACS en 4.3, sin embargo, este mayor detalle no aporta diferencias significativas en lo que a los resultados se refiere. Como ya se indicó anteriormente, esto hace que sea preferible estudiar los resultados obtenidos en el procedimiento explicado en 4.3 puesto que éstos estarán dotados de una mayor estabilidad.

Así, al perfil medio definido por el primer análisis (el expuesto en 4.1 y representado en el plano factorial del Gráfico 1), queda matizado

por los perfiles diferenciales que determinan los planos factoriales representados en los Gráficos 4 y 5. De acuerdo con ellos, se pueden definir cuatro tipologías de individuos que se alejan de ese perfil medio:

- Los individuos que se sitúan en la esquina inferior derecha del primer plano factorial, que son aquellos que continúan en el sistema educativo, no trabajan y no buscan empleo durante los 36 meses (la proyección de los colectivos como variables ilustrativas nos indica que pertenecen fundamentalmente a los colectivos de graduados en ESO y, en menor medida, en Bachillerato).
- Los individuos situados en la esquina inferior izquierda del primer plano factorial, que son aquellos que ya no estudian en el sistema educativo y trabajan al menos 20 h. semanales durante los 36 meses (las variables ilustrativas muestran que pertenecen al resto de colectivos, sobre todo al de los que abandonaron la ESO).
- Los individuos que se aproximan al extremo superior del primer plano factorial, que son aquellos que continúan estudios en el sistema educativo, no trabajan pero buscan empleo durante la mayoría de los 36 meses (ya hemos comentado que esta situación se aproxima al perfil medio durante los meses de verano).
- Los individuos situados en la esquina superior derecha del segundo plano factorial, que son aquellos que continúan estudios en el SE, trabajan menos de 20 h. semanales y no buscan empleo durante la mayoría de los 36 meses (en este caso, esta situación se aproxima al perfil medio durante el último semestre de 2004 y los meses de verano de 2003).

5. CONCLUSIONES

En este artículo hemos pretendido presentar el Análisis de Correspondencias como una técnica multivariante de reducción de datos que puede ser útil en el estudio de historias de vida. A tal fin, una vez descritas las características generales de la técnica, se detallan los aspectos propios de su aplicación cuando la información de partida son datos de historias de vida. En este sentido, se plantea la idoneidad de realizar tres análisis complementarios:

- Un ACS de la tabla subperiodos×estados que permita describir el perfil medio de la nube de puntos.
- Un ACM de la matriz de superindicadores que revelará qué situaciones, periodos e individuos se alejan del patrón trazado por el análisis anterior.
- Un ACS de la tabla individuos×estados que se podrá interpretar como una síntesis del análisis anterior puesto que en él no interviene la dimensión temporal. Además, en el caso de que ambos ofrezcan resultados similares, éste último será preferible al anterior dada la mayor estabilidad de sus resultados.

Con el fin de ilustrar todo este procedimiento, se emplean datos extraídos de la Encuesta de Transición Educativo-Formativa e Inserción Laboral (ETEFIL-2005). Concretamente, se analizan, mes a mes durante el periodo 2002-2004, las diferentes situaciones por las que pasan los 45.620 individuos de la muestra (agrupados en 7 colectivos) en lo que a su vinculación con el sistema educativo, el empleo y la búsqueda del mismo se refiere.

Así, los resultados obtenidos en cada etapa permiten identificar el "perfil medio" de la nube así como los estados e individuos que más se alejan de este perfil medio a través de las representaciones gráficas de los planos factoriales característicos del Análisis de Correspondencias.

En este sentido, el perfil medio se concreta en una primera dimensión que traza el tránsito de una etapa de continuación de estudios en el sistema educativo -sin trabajar ni buscar empleo- (año 2002, sobre todo el primer semestre), a otra en la que se han abandonado los estudios y se trabaja a tiempo completo (año 2004, sobre todo el segundo semestre). Esta transición, reflejada en este primer factor, se ve matizada por una segunda dimensión que refleja la actividad laboral de los meses de verano en los que los encuestados conjugan estudio y trabajo (2003) o estudio y búsqueda de empleo (2002).

Una vez descrito este perfil medio, los dos análisis siguientes, que ofrecen resultados muy semejantes, muestran aquellas otras pautas que difieren del patrón general, como son los individuos que

durante los 36 meses continúan en el sistema educativo sin trabajar ni buscar empleo (sobre todo, graduados en ESO); los que abandonaron el sistema educativo y trabajan al menos 20 h. semanales durante todo el periodo (fundamentalmente, pertenecientes al colectivo de los que abandonaron la ESO); los que continúan estudios, no trabajan pero buscan empleo (esta situación se aproxima al perfil medio durante los meses de verano); y los que continúan estudios en el SE, trabajan menos de 20 h. semanales y no buscan empleo (en este caso, esta situación se asemeja al perfil medio durante el último semestre de 2004 y los meses de verano de 2003).

Consideramos que, con la aplicación desarrollada en este artículo, se cubre el objetivo de mostrar las posibilidades que ofrece el Análisis de Correspondencias en el estudio de datos de historias de acontecimientos. Asimismo, se ilustra cómo esta técnica puede proporcionar resultados de interés complementarios a los que ofrecen otros métodos estadísticos cuyo uso está más extendido cuando se analiza este tipo de información.

BIBLIOGRAFÍA

- BENZÉCRI, J.-P. (1992) *Correspondence Analysis Handbook*. New York: Marcel Dekker.
- BLOSSFELD, H.-P.; A. HAMERLE & K.-U. MAYER (1989) *Event History Analysis: Statistical Theory and Application in the Social Sciences*. Hillsdale (New Jersey): Lawrence Erlbaum Associates.
- DE LEEUW, J.; P.G.M. VAN DER HEIJDEN & I. KREFT (1985) "Homogeneity Analysis of Event History Data", *Methods of Operations Research*, 50, pp. 299-316.
- FRANCIS, B. & M. FULLER (1996) "Visualization of Event Histories", *Journal of the Royal Statistical Society, Series A*, 159, pp. 301-308.
- GREENACRE, M. (1984) *Theory and Applications of Correspondence Analysis*. London: Academic Press.
- (1993) *Correspondence Analysis in Practice*. London: Academic Press.

- INSTITUTO NACIONAL DE ESTADÍSTICA (2007) *Encuesta de transición educativo-formativa e inserción laboral. Resumen Metodológico*. (www.ine.es/daco/daco42/etefil/notaetefil.pdf). Consultado el 17/10/2007).
- JAMBU, M. (1989) *Exploration Informatique et Statistique des Données*. Paris: Dunod.
- LEBART, L.; A. MORINEAU et M. PIRON (1995) *Statistique Exploratoire Multidimensionnelle*. Paris: Dunod.
- MARTENS, B. (1994) "Analyzing Event History Data by Cluster Analysis and Multiple Correspondence Analysis: An Example Using Data about Work and Occupations of Scientists and Engineers". M. GREENACRE & J. BLASIUS (eds.) *Correspondence Analysis in the Social Sciences*. San Diego: Academic Press.
- VAN DER HEIJDEN, P.G.M. (1987) *Correspondence Analysis of Longitudinal Categorical Data*. Leiden: DSWO.
- & J. DE LEEUW (1989) "Correspondence Analysis with Special Attention to the Analysis of Panel Data and Event History Data". C.C. CLOGG (ed.) *Sociological Methodology*. Oxford: Basil Blackwell.
- ; J. TEUNISSEN et C. VAN ORLÉ (2000) "L'analyse des correspondances multiples: un outil pour la classification de données de cursus". J. MOREAU, P.-A. DOUDIN et P. CAZES (eds.) *L'analyse des correspondances et les techniques connexes. Approches nouvelles pour l'analyse statistique des données*. Berlin: Springer.

ANEXO

A1. Resultados del ACS de la tabla subperiodos×estados

Number	Eigenvalue	Percentage	Cumulated Percentage
1	0,0765	93,09	93,09
2	0,0042	5,10	98,19
3	0,0010	1,16	99,35
4	0,0002	0,20	99,56
5	0,0002	0,19	99,75
6	0,0001	0,10	99,85
7	0,0001	0,07	99,92
8	0,0000	0,03	99,95
9	0,0000	0,02	99,97
10	0,0000	0,02	99,99
11	0,0000	0,01	100,00
Trace of matrix:	0,0821		

Active columns	Relative Weight	Distance to origin	Coordinates		Squared cosines			Contributions	
			Axis 1	Axis 2	Axis 1	Axis 2	Sum	Axis 1	Axis 2
SE T>20	8,28	0,04685	-0,08	0,20	0,15	0,85	1,00	0,76	78,33
SE T<20 BE	0,10	0,13930	0,26	0,09	0,50	0,06	0,56	0,09	0,22
SE T<20 NBE	1,39	0,11648	0,33	0,03	0,92	0,01	0,93	1,95	0,24
SE NT BE	1,55	0,15587	0,35	0,17	0,78	0,18	0,96	2,46	10,19
SE NT NBE	15,20	0,30224	0,55	-0,02	1,00	0,00	1,00	59,79	1,49
NSE T>20	60,93	0,03378	-0,18	-0,02	0,98	0,02	1,00	26,46	7,51
NSE T<20 BE	0,54	0,02572	0,14	-0,02	0,72	0,02	0,74	0,13	0,07
NSE T<20 NBE OE	0,18	0,08701	0,26	0,00	0,80	0,00	0,80	0,16	0,00
NSE T<20 NBE NOE	0,53	0,06321	0,19	0,03	0,56	0,01	0,57	0,24	0,09
NSE NT BE	6,82	0,02444	0,13	-0,03	0,72	0,04	0,76	1,58	1,59
NSE NT NBE OE	2,21	0,20707	0,45	-0,02	0,96	0,00	0,96	5,74	0,27
NSE NT NBE NOE	2,29	0,03091	0,14	0,00	0,67	0,00	0,67	0,62	0,00

Active rows	Relative Weight	Distance to origin	Coordinates		Squared cosines			Contributions	
			Axis 1	Axis 2	Axis 1	Axis 2	Sum	Axis 1	Axis 2
ene-02	2,78	0,19315	0,44	-0,04	0,98	0,01	0,99	6,89	0,91
feb-02	2,78	0,17710	0,42	-0,04	0,99	0,01	1,00	6,36	0,84
mar-02	2,78	0,16034	0,40	-0,03	0,99	0,01	1,00	5,77	0,69
abr-02	2,78	0,14885	0,38	-0,03	0,99	0,01	1,00	5,35	0,64
may-02	2,78	0,13366	0,36	-0,03	0,99	0,00	1,00	4,81	0,42
jun-02	2,78	0,08530	0,29	0,03	0,98	0,01	0,99	3,03	0,43
jul-02	2,78	0,05309	0,21	0,07	0,86	0,10	0,96	1,66	3,47
ago-02	2,78	0,05298	0,22	0,07	0,88	0,09	0,97	1,70	3,16
sep-02	2,78	0,05324	0,23	-0,02	0,97	0,01	0,98	1,88	0,23
oct-02	2,78	0,06500	0,25	-0,04	0,97	0,03	1,00	2,28	1,22
nov-02	2,78	0,06280	0,25	-0,04	0,97	0,03	0,99	2,21	1,09
dic-02	2,78	0,05776	0,24	-0,03	0,98	0,02	0,99	2,05	0,65
ene-03	2,78	0,04135	0,20	-0,01	0,95	0,00	0,95	1,42	0,09
feb-03	2,78	0,03731	0,19	-0,01	0,93	0,00	0,93	1,26	0,05
mar-03	2,78	0,03101	0,17	0,00	0,93	0,00	0,93	1,05	0,00
abr-03	2,78	0,02646	0,15	0,01	0,90	0,00	0,91	0,87	0,05
may-03	2,78	0,02142	0,13	0,02	0,82	0,03	0,85	0,64	0,39
jun-03	2,78	0,02673	0,02	0,16	0,01	0,91	0,92	0,01	16,20
jul-03	2,78	0,03744	-0,06	0,18	0,10	0,87	0,97	0,14	21,58
ago-03	2,78	0,03996	-0,07	0,18	0,13	0,84	0,98	0,19	22,35
sep-03	2,78	0,03085	-0,14	-0,05	0,64	0,10	0,74	0,72	1,99
oct-03	2,78	0,02767	-0,14	-0,06	0,70	0,15	0,85	0,70	2,73
nov-03	2,78	0,02799	-0,15	-0,06	0,77	0,15	0,92	0,78	2,70
dic-03	2,78	0,03047	-0,16	-0,05	0,85	0,09	0,93	0,94	1,79
ene-04	2,78	0,04017	-0,19	-0,05	0,93	0,05	0,98	1,36	1,40
feb-04	2,78	0,04495	-0,21	-0,04	0,95	0,04	0,99	1,55	1,13
mar-04	2,78	0,05166	-0,22	-0,03	0,97	0,02	0,99	1,82	0,77
abr-04	2,78	0,05731	-0,24	-0,03	0,97	0,01	0,99	2,03	0,55
may-04	2,78	0,06595	-0,25	-0,02	0,98	0,01	0,98	2,35	0,25
jun-04	2,78	0,10831	-0,32	0,05	0,97	0,02	0,99	3,83	1,38
jul-04	2,78	0,13883	-0,37	0,07	0,97	0,03	1,00	4,87	2,85
ago-04	2,78	0,13991	-0,37	0,06	0,96	0,03	1,00	4,90	2,79
sep-04	2,78	0,16794	-0,41	-0,05	0,98	0,01	1,00	6,00	1,67
oct-04	2,78	0,17072	-0,41	-0,04	0,98	0,01	1,00	6,11	1,17
nov-04	2,78	0,17406	-0,41	-0,04	0,98	0,01	0,99	6,21	1,26
dic-04	2,78	0,17488	-0,41	-0,04	6,25	1,11	0,98	0,01	1,26

Supplementary columns	Relative Weight	Distance to origin	Coordinates		Squared cosines		
			Axis 1	Axis 2	Axis 1	Axis 2	Sum
No estudia SE	73,49	0,01443	-0,12	-0,02	0,96	0,04	1,00
Trabaja	71,94	0,02384	-0,15	0,00	1,00	0,00	1,00
No trabaja	28,06	0,15670	0,40	-0,01	1,00	0,00	1,00
No busca empleo	21,78	0,22163	0,47	-0,01	1,00	0,00	1,00
Estudia SE	26,51	0,11091	0,33	0,06	0,96	0,03	1,00
Trabaja <20h/sem.	2,74	0,07096	0,26	0,02	0,93	0,00	0,93
Trabaja >=20h/sem.	69,20	0,02912	-0,17	0,00	1,00	0,00	1,00
Busca empleo	9,01	0,03345	0,17	0,00	0,88	0,00	0,88

Supplementary rows	Relative Weight	Distance to origin	Coordinates		Squared cosines		
			Axis 1	Axis 2	Axis 1	Axis 2	Sum
ESO	29,72	0,32913	0,49	0,15	0,72	0,07	0,79
BACH	13,25	0,79877	0,66	0,48	0,55	0,29	0,83
CFGM	14,41	0,16395	-0,33	-0,21	0,68	0,27	0,95
CFGS	16,90	0,05356	-0,22	-0,03	0,89	0,02	0,91
AESO	33,63	0,16218	-0,23	-0,19	0,33	0,22	0,55
FIP	16,35	0,08043	-0,17	-0,08	0,36	0,09	0,45
ETCO	3,61	0,30662	-0,35	-0,28	0,40	0,26	0,65
Hombre	72,53	0,00632	-0,04	-0,05	0,30	0,40	0,71
Mujer	55,34	0,02095	0,11	0,04	0,58	0,07	0,66

A2. Resultados del ACS de la tabla agregada individuos×situaciones

Number	Eigenvalue	Percentage	Cumulated Percentage
1	0,8136	14,14	14,14
2	0,7136	12,40	26,54
3	0,5847	10,16	36,70
4	0,5330	9,26	45,96
5	0,5228	9,08	55,04
6	0,4857	8,44	63,48
7	0,4645	8,07	71,55
8	0,4535	7,88	79,43
9	0,4290	7,45	86,89
10	0,3816	6,63	93,52
11	0,3731	6,48	100,00
Trace of matrix	5,7551		

Active columns	Relative Weight	Distance to origin	Coordinates				Squared cosines					Contributions			
			Axis 1	Axis 2	Axis 3	Axis 4	Axis 1	Axis 2	Axis 3	Axis 4	Sum	Axis 1	Axis 2	Axis 3	Axis 4
SE T>20	5,31	6,81	0,25	0,20	-0,05	0,43	0,01	0,01	0,00	0,03	0,04	0,41	0,31	0,03	1,87
SE T<20 BE	0,27	177,76	0,04	2,48	-2,20	-4,67	0,00	0,03	0,03	0,12	0,18	0,00	2,35	2,25	11,12
SE T<20 NBE	3,60	15,35	0,85	-0,02	-3,49	1,01	0,05	0,00	0,80	0,07	0,91	3,23	0,00	75,26	6,86
SE NT BE	3,43	19,70	-0,15	4,35	0,25	0,31	0,00	0,96	0,00	0,00	0,97	0,10	90,83	0,38	0,60
SE NT NBE	45,52	0,89	0,85	-0,18	0,31	0,00	0,81	0,04	0,11	0,00	0,95	40,14	2,01	7,39	0,00
NSE T>20	29,77	1,69	-1,14	-0,28	0,05	0,38	0,77	0,05	0,00	0,09	0,90	47,74	3,32	0,14	8,16
NSE T<20 BE	0,59	81,96	-0,73	0,41	-1,54	-4,83	0,01	0,00	0,03	0,28	0,32	0,39	0,14	2,39	25,97
NSE T<20 NBE OE	0,21	183,98	-0,24	-0,19	-2,48	-1,59	0,00	0,00	0,03	0,01	0,05	0,02	0,01	2,26	1,02
NSE T<20 NBE NOE	0,70	68,02	-0,28	-0,27	-2,86	-2,04	0,00	0,00	0,12	0,06	0,18	0,07	0,07	9,81	5,47
NSE NT BE	5,76	7,78	-0,95	0,22	0,09	-0,51	0,12	0,01	0,00	0,03	0,16	6,45	0,39	0,08	2,78
NSE NT NBE OE	2,17	21,73	-0,63	-0,25	-0,03	-1,55	0,02	0,00	0,00	0,11	0,13	1,07	0,19	0,00	9,81
NSE NT NBE NOE	2,65	18,71	-0,34	-0,32	0,00	-2,30	0,01	0,01	0,00	0,28	0,37	0,37	0,00	0,00	26,33

Supplementary columns	Relative Weight	Distance to origin	Coordinates				Squared cosines				
			Axis 1	Axis 2	Axis 3	Axis 4	Axis 1	Axis 2	Axis 3	Axis 4	Sum
ESE	58,13	0,57	0,73	0,15	0,02	0,10	0,93	0,04	0,00	0,02	0,99
TRM20	5,38	10,25	0,45	0,11	-3,09	-0,42	0,02	0,00	0,93	0,02	0,97
TRMOI20	35,08	1,25	-0,93	-0,21	0,04	0,39	0,69	0,03	0,00	0,12	0,85
BE	10,05	5,35	-0,64	1,70	-0,01	-0,60	0,08	0,54	0,00	0,07	0,68
INOE	3,35	15,03	-0,33	-0,31	-0,60	-2,25	0,01	0,01	0,02	0,34	0,37
IEFSE	2,39	19,84	-0,60	-0,25	-0,25	-1,55	0,02	0,00	0,00	0,12	0,15
NESE	41,87	1,10	-1,01	-0,20	-0,03	-0,14	0,93	0,04	0,00	0,02	0,99
TRAB	40,47	0,97	-0,75	-0,17	-0,38	0,28	0,58	0,03	0,15	0,08	0,84
NTRA	59,54	0,45	0,51	0,11	0,26	-0,19	0,58	0,03	0,15	0,08	0,84
NBE	60,17	0,51	0,67	-0,14	-0,02	-0,09	0,89	0,04	0,00	0,02	0,94

Supplementary categories	Relative Weight	Distance to origin	Coordinates				Squared cosines				
			Axis 1	Axis 2	Axis 3	Axis 4	Axis 1	Axis 2	Axis 3	Axis 4	Sum
ESO	42,47	2,25679	0,55	0,08	0,06	0,02	0,13	0,00	0,00	0,00	0,14
BACH	17,82	8,30645	0,62	0,15	-0,09	0,06	0,05	0,00	0,00	0,00	0,05
CFGM	6,15	33,44140	-1,03	-0,22	0,00	0,17	0,03	0,00	0,00	0,00	0,03
CFGS	7,93	26,24970	-0,68	-0,12	-0,06	0,18	0,02	0,00	0,00	0,00	0,02
AESO	15,34	12,04580	-0,94	-0,19	0,00	-0,19	0,07	0,00	0,00	0,00	0,08
FIP	8,56	15,95280	-0,73	-0,09	-0,07	-0,12	0,03	0,00	0,00	0,00	0,04
ETCO	1,73	78,37210	-1,15	-0,19	-0,07	-0,26	0,02	0,00	0,00	0,00	0,02
Hombre	49,25	6,38256	-0,13	-0,03	0,04	0,08	0,00	0,00	0,00	0,00	0,00
Mujer	50,75	6,34032	0,13	0,03	-0,04	-0,07	0,00	0,00	0,00	0,00	0,00