

Visualización conjunta del uso del tiempo en días laborables y en festivos en la Comunidad Autónoma Vasca

M. Jesús Bárcena-Ruiz & Karmele Fernández-Aguirre
Dpto de Estadística y Econometría, Facultad de CC EE y Empresariales
Avda. Lehendakari Aguirre, 83, 48015 Bilbao, España
E-mail: KF@alcib.bs.ehu.es

ABSTRACT

In recent years researchers have focused on investigating the different ways of linking surveys (M. Lejeune, 1995) (L. Lebart & M. Lejeune, 1995 a). This is a complex problem that relates the information obtained in two different surveys on the same population. Our main interest is to compare two alternative ways which display jointly the specific information of two sets of variables Y and Z , each corresponding to one of the two surveys considered. The method uses the information from a set of variables X common to the two surveys. We test the robustness of the results using a non-parametric bootstrap, adapting the partial bootstrapping method to the case of linking surveys.

Palabras clave: Inserción de Encuestas, Análisis de Correspondencias Múltiples, Bootstrap Parcial
Proyección en suplementario.

1. OBJETIVO

El objetivo de la inserción consiste en visualizar conjuntamente la información específica de dos encuestas A y B realizadas, respectivamente, a N_A y N_B individuos de la misma población. Las dos encuestas tienen un conjunto \mathbf{X} de variables en común, además cada una de ellas tiene un conjunto de variables específicas \mathbf{Y} ó \mathbf{Z} , tal y como ilustra el siguiente diagrama:

Encuesta A			Encuesta B		
X_1, \dots, X_p	Y_1, \dots, Y_q	No observado	X_1, \dots, X_p	No observado	Z_1, \dots, Z_r

El objetivo del trabajo es doble: Se trata en primer lugar de obtener, mediante técnicas factoriales, representaciones gráficas en las que se proyecten las variables específicas \mathbf{Y} , del uso del tiempo en días laborables, y las variables específicas \mathbf{Z} , del uso del tiempo en festivos, sobre un subespacio de representación estable obtenido del análisis de las variables \mathbf{X} , comunes a las dos submuestras contenidas en la Encuesta de Presupuestos de Tiempo (EPT-93). Así se ha efectuado un análisis exploratorio multivariante que nos permite señalar distintas tipologías de población, cuyo uso del tiempo varía en función del tipo de jornada. Por otra parte, se trata de comparar, desde el punto de vista de la

AGRADECIMIENTOS: Este trabajo ha sido parcialmente subvencionado por la Dirección General de Enseñanza Superior del Ministerio Español de Educación y Ciencia y la Universidad del País Vasco/Euskal Herriko Unibertsitatea mediante el proyecto de investigación: PB95-0346.

aplicación de dos maneras distintas de efectuar la inserción. Ambas han sido consideradas anteriormente desde el punto de vista teórico pero, sobre todo, en cuanto a las aplicaciones, estas se han venido centrando en una de ellas, la presentada en la Sección 3.1. Posteriormente, se ha efectuado una validación mediante bootstrap parcial que muestra la estabilidad de los resultados obtenidos al aplicar las dos maneras distintas de inserción.

En el apartado siguiente vamos a presentar los datos objeto de experimentación y a continuación dos métodos alternativos para la inserción de las variables **Y** y **Z**, así como la adaptación del método de validación “partial bootstrapping” al caso de inserción de encuestas que nos ocupa, posteriormente presentaremos la aplicación a la Encuesta de Presupuestos de Tiempo (EPT-93)

2. LOS DATOS OBJETO DE EXPERIMENTACION

La EPT-93 fue realizada por el EUSTAT (Instituto de Estadística de la Comunidad Autónoma del País Vasco) entrevistando a 5040 individuos. Para los mismos se observaron sus características personales (variables comunes **X** cuya descripción y modalidades aparecen en el Anexo1) y el tiempo en minutos dedicado a diferentes actividades (como dormir, aseo, trabajo, etc.: variables específicas **Y** ó **Z** dependiendo del tipo de jornada y que figuran en el Anexo2) . La información se recogió mediante un diario que cada encuestado rellenó en un solo día previamente establecido. Contamos con dos submuestras *E1* con 2521 observaciones obtenidas en día laborable y *E2* con 2519 observaciones obtenidas en día festivo. Es necesario analizar por separado los resultados obtenidos para cada submuestra ya que el uso del tiempo es diferente en días laborables y en festivos.

La situación ideal sería la obtención de un diario semanal donde se recogiese, para cada encuestado, información sobre los tiempos dedicados a las distintas actividades durante todos los días de la semana, sin embargo, el diario semanal tiene muchos inconvenientes: "....su calidad disminuye a medida que avanza la semana, los diarios incompletos se multiplican y, además, es imposible realizarlo sin un incentivo económico alto." (EUSTAT, 1997 p.XII).

Por tanto, el objetivo de este trabajo es insertar las submuestras *E1* y *E2* para poder visualizar conjuntamente la información obtenida en días laborables y en festivos y así aproximar los resultados que se hubiesen obtenido con un cuestionario semanal.

Hay que ser consciente de que la inserción de encuestas no aumenta la información disponible pero enriquece las interpretaciones de los datos parciales lo que conlleva una reducción del efecto de la no respuesta, en nuestro caso, la ausencia de información sobre el uso del tiempo en días festivos en la submuestra *E1* y en días laborables en la submuestra *E2*.

3. DOS METODOS ALTERNATIVOS PARA LA INSERCIÓN

Los dos métodos se basan en la utilización de la técnica de proyección en suplementario. Se trata de una técnica de probada utilidad ilustrativa, en análisis descriptivos multivariantes, consistente en proyectar sobre ejes y planos factoriales elementos (observaciones o individuos y variables) que no han tomado parte en la obtención de los mismos (M. Greenacre, 1984) (L. Lebart *et al.* 1984).

Teniendo en cuenta las relaciones de transición entre el espacio de los individuos y el dual de las variables, las variables suplementarias se proyectan sobre un eje factorial de orden α , correspondiente al α -ésimo valor propio I_a , utilizando las coordenadas de los individuos activos sobre este eje (que constituyen el factor α) con el mismo esquema que si se tratara de variables activas. Así, la proyección de una variable cuantitativa sobre un eje, se calcula como la correlación entre la variable (activa ó suplementaria) y el factor. Si es cualitativa o categórica, cada modalidad (activa ó suplementaria) se proyecta como el centro de gravedad de las proyecciones de los individuos que han elegido esa modalidad, corregido por el inverso de la raíz cuadrada del valor propio asociado al eje.

3.1 Subespacio de representación parcial

Este método de inserción consiste en analizar la matriz de datos \mathbf{X} de una encuesta, por ejemplo A, tomada como referencia (S. Bonnefous *et al.* 1986) (A. Iliakopoulos, 1988) a través de métodos factoriales y proyectar la información específica \mathbf{Y} de la propia encuesta como suplementaria y las variables específicas \mathbf{Z} de la otra B como puntos doblemente suplementarios (Aluja-Banet *et al.* 1997). Los pasos del procedimiento serían:

- I) Análisis factorial descriptivo de la matriz de variables comunes \mathbf{X} de la encuesta A y obtención de los valores propios I_a y vectores propios correspondientes \mathbf{m}_a , $\alpha=1,2,...,k$ donde k es la dimensión del subespacio de referencia, en el caso de inserción de encuestas, generalmente, $k \leq 3$.
- II) Obtención de los factores \mathbf{y}_a^A y proyección de las variables \mathbf{Y} como suplementarias.
- III) Proyección como suplementarios los N_B individuos de la encuesta B. Denotaremos \mathbf{y}_a^S a los vectores formados por esas proyecciones.
- IV) Proyección las variables específicas \mathbf{Z} de B, sobre los ejes y planos obtenidos del análisis de A. Estas proyecciones se obtienen, como suplementarias, mediante los vectores \mathbf{y}_a^S son, por tanto, proyecciones doblemente suplementarias.

Antes de proceder a la inserción es necesario verificar que el análisis factorial exploratorio de la matriz de variables comunes \mathbf{X} proporcione, en ambas encuestas, subespacios de representación similares. Si es así, las proyecciones de los individuos de B como activos en su propia encuesta (componentes \mathbf{y}_a^B) y como suplementarios en el

análisis de A (\mathbf{y}_a^s) serán muy parecidas y la posición de las variables específicas \mathbf{Z} en los planos de la encuesta A será similar a la que se obtiene en el análisis de su propia encuesta. Así, los resultados de la inserción se pueden interpretar aplicando las reglas de interpretación tradicionales propias del tipo de análisis factorial efectuado.

3.2 Subespacio de representación global

Este método se basa en el análisis de la tabla de datos \mathbf{X} de las dos encuestas A y B tomadas conjuntamente, a continuación se insertan las variables específicas \mathbf{Y} y \mathbf{Z} . Los pasos del procedimiento serían:

- I) Análisis factorial descriptivo de la matriz de variables comunes \mathbf{X} para todos los individuos $N_A + N_B$ y obtención de los valores propios y vectores propios correspondientes.
- II) Obtención de los factores \mathbf{y}_a proyección de las variables específicas \mathbf{Y} de la encuesta A como variables suplementarias utilizando las N_A primeras componentes de \mathbf{y}_a y las variables específicas \mathbf{Z} de B utilizando las N_B últimas.

Si no hay diferencias relevantes en la estructura de las variables comunes en ambas encuestas, el subespacio de representación obtenido considerando los $N_A + N_B$ individuos es similar a los que obtendríamos considerando N_A y N_B por separado.

4. EL BOOTSTRAP PARCIAL COMO METODO DE VALIDACION

El uso de técnicas de remuestreo sobre una distribución libre (Efron, 1979) sirve para valorar la variabilidad muestral basándose en el remuestreo de una distribución empírica. Las técnicas de remuestreo bootstrap proporcionan un marco apropiado para investigar la estabilidad de los resultados obtenidos mediante análisis factoriales descriptivos de una matriz de datos. El bootstrap de los resultados de un Análisis de Correspondencias Múltiples (ACM) denominado también Homogeneity Analysis (Gifi, 1990) o de un Análisis de Componentes Principales (ACP) se lleva a cabo extrayendo réplicas con reemplazamiento de la matriz de datos analizada. En este trabajo, analizamos la estabilidad de los resultados de la inserción mediante bootstrap parcial.

En un bootstrap parcial no se realiza el análisis factorial de cada réplica (Chateau and Lebart, 1996), se considera que los ejes obtenidos para los datos originales son más fiables que los obtenidos para cualquier réplica y se evitan reflexiones de los ejes debidas a la arbitrariedad de signo de los vectores propios. Las condiciones de ortogonalidad asociadas a una descomposición en valores singulares, que depende de una realización muestral, implican múltiples distorsiones difíciles de evitar (Milan and Whittaker, 1995) proponen técnicas para identificar los factores obtenidos de los análisis factoriales de cada réplica. En un bootstrap parcial, para cada réplica se obtienen la tabla de BURT (caso de ACM) o la matriz de covarianzas/correlación (caso de ACP) y se calculan las proyecciones de las variables como elementos suplementarios (Greenacre, 1984).

En el caso de un ACP normado, teniendo en cuenta las relaciones de transición (Lebart et al. 1997b) la proyección de una variable j sobre un eje α es igual a la correlación entre la variable y el eje, $\mathbf{j}_a(j) = \text{corr}(X_j, \mathbf{y}_a)$, así, una forma de calcular la proyección de la variable j de una réplica k sobre un eje α del análisis original, $\mathbf{j}_a^k(j)$, es como coeficiente de correlación entre los valores que la variable a proyectar y el factor toman en los individuos seleccionados en la réplica, X_j^k y \mathbf{y}_a^k respectivamente:

$$\mathbf{j}_a^k(j) = \text{corr}(X_j^k, \mathbf{y}_a^k)$$

En el caso de inserción de encuestas se obtienen muestras aleatorias con reemplazamiento de las observaciones de las variables a insertar y se calculan las proyecciones para cada réplica en el subespacio de referencia de las variables comunes. Si las variables a insertar son continuas, como en el caso de nuestro trabajo (Sección 5.3) estas proyecciones se calculan como coeficientes de correlación entre los valores que toman esas variables y el factor para los individuos seleccionados en la réplica.

5. APLICACION A LA ENCUESTA DE PRESUPUESTOS DE TIEMPO

La inserción de **Y** (tiempos de actividades en día laborable) y de **Z** (tiempos de actividades en día festivo) se efectúa sobre un subespacio de representación obtenido del conjunto de variables de caracterización **X** (edad, sexo, estado civil, nivel de instrucción y relación con la actividad). El caso que estudiamos es privilegiado dado que son dos partes de una misma encuesta y la comprobación de que las variables comunes **X** proporcionen un subespacio de representación similar no hará más que confirmar que el diseño de la encuesta y el reparto entre días laborables y festivos es correcto.

Esta comprobación puede hacerse aplicando el procedimiento ACM del paquete de programas (SPADN, 1997) a las variables **X**, que no presentan casos de no respuesta, de $E1 + E2$ (Tabla 1). En los ejes factoriales obtenidos se proyectan las categorías activas y como suplementarias las mismas categorías observadas por separado en $E1$ y en $E2$ completando con ceros sus respectivas columnas; así, cada categoría de **X** tiene tres proyecciones: una como activa y dos como suplementarias, las tres coincidirían en un punto si las dos submuestras fueran exactamente iguales. Al ser iguales los individuos de las dos submuestras, la única diferencia sería que en la tabla activa las modalidades tendrían peso doble pero el centro de gravedad no cambiaría.

$E1+E2$	$E1$	$E2$
X_1	X_1	0
X_2	0	X_2

Tabla 1: Tabla de datos objeto de ACM

Conviene recordar que en un ACM la coordenada de la proyección de una modalidad j sobre un eje α es el punto medio, corregido por $I_a^{-1/2}$, de las coordenadas de los individuos que la han escogido:

$$j_{aj} = \frac{1}{z_{.j} \sqrt{I_a}} \sum_{i \in I(j)} y_{ai}$$

Donde $I(j)$ designa el conjunto de individuos que han escogido la modalidad j y $z_{.j}$ el número de individuos que toman dicha modalidad.

En el caso de las variables de $E1$ y de $E2$ tenemos una modalidad adicional que corresponde a valores no observados (ceros en la segunda y tercera parte de la Tabla 1), la proyección de esta modalidad para todas las variables se sitúa en el centro de gravedad corregido de los individuos de la otra submuestra $E2$ y $E1$ respectivamente. Si las dos submuestras fueran exactamente iguales coincidirían en el origen de coordenadas (centro de gravedad de las dos submuestras). Si no son iguales, como en nuestro caso los dos centros de gravedad corregidos son: (-0.17, 0.10) y (-0.12, 0.06) y nos dan una idea de la proximidad en el plano principal de las dos nubes.

Para las categorías bien representadas en el plano principal (Fig. 1) se observan tan sólo pequeñas diferencias entre cada proyección activa: LAH (labores de hogar), PRI (estudios primarios), PAR (parados) y sus suplementarias correspondientes: LAH1, LAH2 (de $E1$ ó $E2$), PRI1, PRI2 (de $E1$ ó $E2$), PAR1, PAR2 (de $E1$ ó $E2$)... Las mayores diferencias se observan en las categorías: ED1 (edad hasta 34 años), MED (estudios medios), SUP (estudios superiores), SOL (soltero) y las categorías de la variable sexo: VAR (varón) y MUJ (mujer) cambian bastante en el segundo eje, sin embargo, es en el tercero donde están mucho mejor representadas y en él apenas presentan cambios. Por

tanto, los rasgos principales de \mathbf{X} son similares en $E1$ y $E2$ y tiene sentido realizar su inserción.

Al analizar mediante ACM las tablas \mathbf{X} de cada submuestra observamos que la variable sexo es la que presenta mayores diferencias en ambas, en el caso de $E1$ no contribuye prácticamente a la creación del segundo eje y contribuye en un 44.7% al tercero, mientras que en el caso de $E2$ contribuye en un 12.6% al segundo y en un 34.3% al tercero. Este hecho va a influir ligeramente (Sección 5.2) en los resultados de la aplicación de los dos métodos de inserción.

El plano principal de este análisis (Fig. 1) recoge un 40.17% de la inercia total y los tres primeros ejes recogen un 53.79%. El análisis de estos 3 primeros ejes nos proporciona el conocimiento de los rasgos principales de \mathbf{X} comunes a ambas submuestras. El primer eje contrapone la población “joven”: ED1 (edad hasta 34 años), EST (estudiante), MED (estudios medios) y SUP (estudios superiores), frente a la población “mayor”: ED3 (>60 años), CAS (casados) y RES (resto), PRI (estudios primarios), JUB (jubilados) y LAH (labores del hogar). El segundo eje complementa y matiza la información del primero. En este eje, destaca la población “ocupada”: ED2 (edad 34-59 años), OCU (ocupados) y CAS (casados) frente a la “no ocupada”: ED3 (>60 años), JUB (jubilados), EST (estudiantes), SOL (solteros) y RES (resto). En el tercer eje se recoge la contraposición entre mujeres y labores del hogar frente a varones ocupados y jubilados o rentistas.

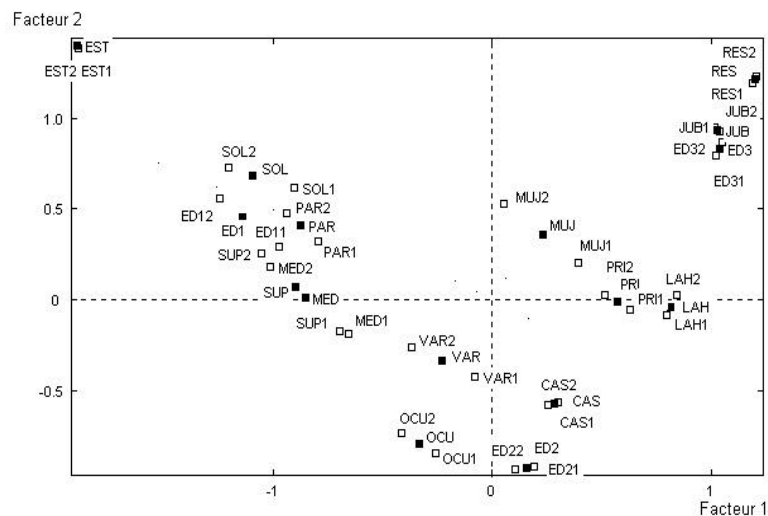


Fig. 1 Plano principal del análisis de \mathbf{X} global y proyección de $E1$ y $E2$

5.1 Inserción de EPT-93 en un subespacio de representación global

Siguiendo los pasos indicados en la Sección 3.2, proyectamos como suplementarias las variables \mathbf{Y} y \mathbf{Z} de $E1$ y $E2$ en los tres primeros ejes factoriales del análisis de las variables activas \mathbf{X} de $E1+E2$. Estas proyecciones se calculan como coeficientes de correlación entre los valores que toman las variables \mathbf{Y} (\mathbf{Z}) y el factor y_a en los individuos de $E1(E2)$.

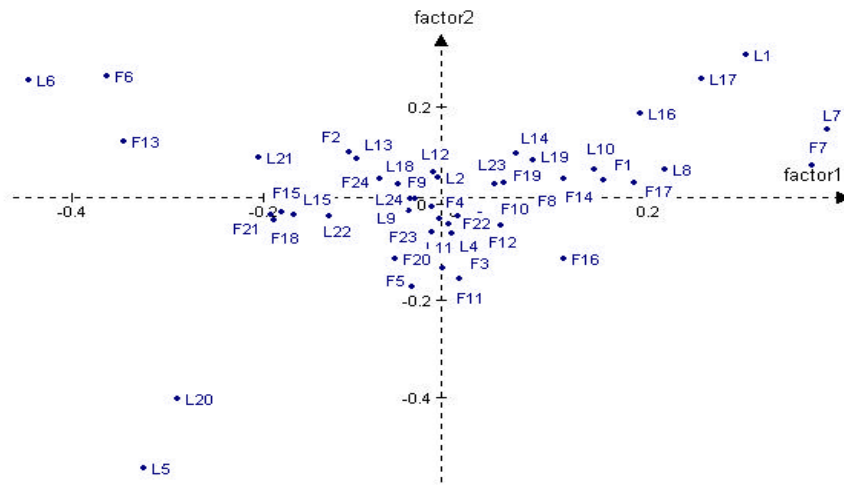


Fig. 2 Inserción de tiempos en Laborable y Festivo en el plano principal

En la Fig. 2 presentamos el plano principal de la inserción de las variables Y , tiempos de las 24 actividades consideradas en día laborable (que en el plano aparecen indicadas mediante la letra L) y de Z , tiempos en festivo (indicados mediante la letra F). Podemos observar que algunas actividades se sitúan bastante próximas: L7, F7 (tareas domésticas), L6, F6 (formación), L15, F15 (gimnasia y deporte) y otras (más cerca del origen, peor representadas en el plano); esto indica que el tipo de personas que más tiempo dedica a estas actividades es el mismo en días laborables que en festivos. Otras actividades se alejan bastante según sea el tipo de jornada y relacionándolas con las modalidades de las variables de caracterización (Fig. 1) puede decirse que las reuniones con amigos, fiestas etc. en festivos: F13 está mucho más claramente relacionada con los jóvenes, solteros y estudiantes que la misma actividad en día laborable: L13. Las actividades dormir: L1, ocio en el hogar: L17 y excursiones y paseos: L16 están mucho más relacionadas con las personas mayores jubiladas y resto (ni solteras ni casadas) en día laborable. El trabajo: L5, F5 y el trayecto al trabajo o formación: L20, F20 también están bien diferenciados. Estas dos últimas son mucho más características de personas ocupadas, jóvenes y de mediana edad en día laborable.

5.2 Inserción de EPT-93 en un subespacio de representación parcial

En la Fig. 3 presentamos el plano principal del ACM de la tabla X de EI y la inserción de las variables Y como suplementarias y las Z como doblemente suplementarios (método de inserción descrito en la Sección 3.1).

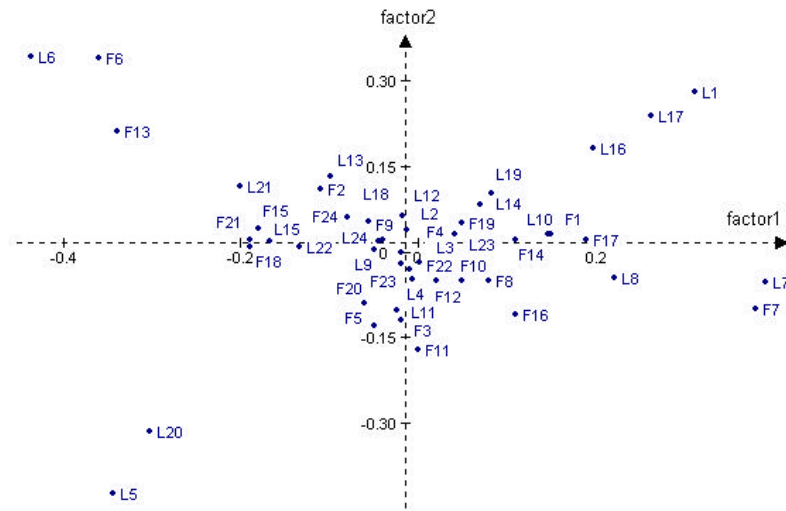


Fig. 3 Plano principal del análisis de X de EI e inserción de Y y de Z

Comparando la Fig.2 con la Fig. 3 podemos apreciar que el resultado de los dos métodos de inserción es prácticamente igual en cuanto a las proyecciones en el primer eje factorial, en el segundo las diferencias más apreciables se dan en L7, F7 (tareas domésticas) y L8, F8 (compras). Estas pequeñas diferencias se explican por la distinta contribución de la variable sexo a la formación del segundo eje factorial en las dos submuestras (Sección 5)

5.3 Validación de la inserción mediante remuestreo parcial

Usando el paquete estadístico Splus, hemos construido y apilado 30, 150 y 250 réplicas bootstrap de los datos a insertar: variables **Y** y **Z** en el subespacio de representación global y variables **Z** en el parcial, hemos insertado esas réplicas proyectándolas en el subespacio correspondiente (como se ha indicado al final de la Sección 4). Hay que hacer notar que al tomar muestras con reemplazamiento, del mismo tamaño que la muestra original, estamos dando de hecho nuevos pesos a las filas de la tabla, pesos múltiples de $1/n$ (donde n es el total de observaciones de *E1* ó *E2*) a las que intervienen en la muestra bootstrap y pesos 0 a las que no intervienen.

Hemos calculado también las medias y varianzas de estas proyecciones replicadas. En las Fig.4 y Fig. 5 podemos observar los resultados de 30 y 150 réplicas para el caso de las variables específicas **Z**, actividades en festivo (suprimidas las centrales por estar peor representadas) y la inserción en el subespacio global.

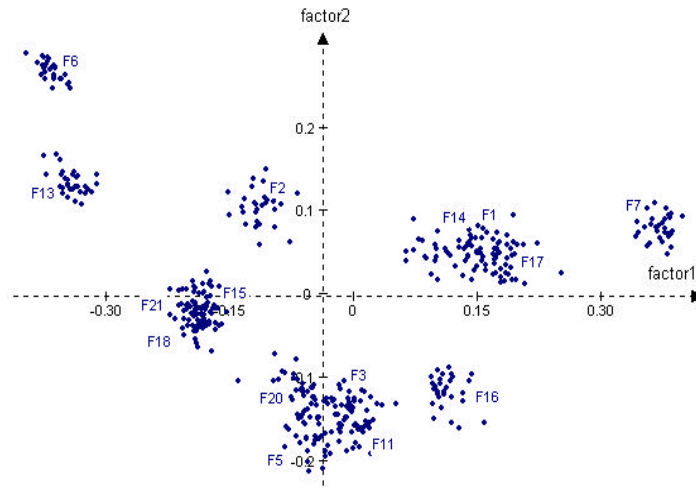


Fig. 4 Inserción en el plano de representación global y 30 réplicas

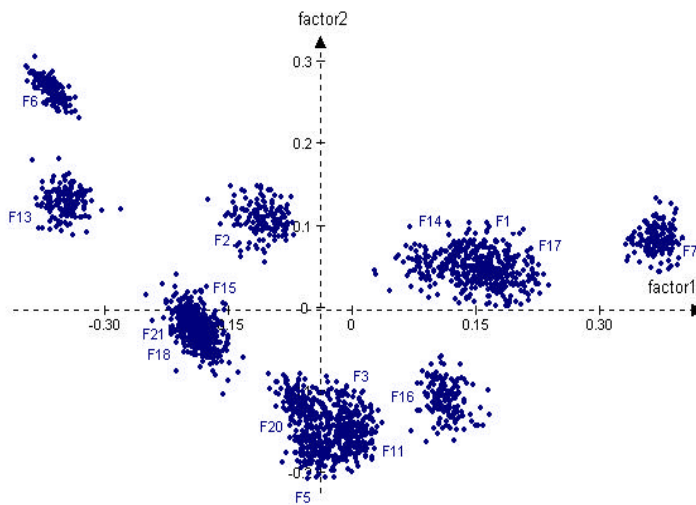


Fig. 5 Inserción en el plano de representación global y 150 réplicas

Las gráficas muestran que los resultados de la inserción son muy estables, las proyecciones de las réplicas forman una nube alrededor de la proyección primera. En las Tablas 2 y 3 podemos observar que las medias de 30 y 150 réplicas son prácticamente iguales. Las varianzas en el caso de 150 réplicas son menores (la forma de las nubes es semejante, sólo que para 150 réplicas son aproximadamente cinco veces más densas).

Las medias de las réplicas están, asimismo, muy próximas de las proyecciones primeras, en la figura siguiente (Fig. 6) podemos observar esta proximidad con las medias de 30 réplicas.

Los resultados de la inserción sobre el subespacio de representación parcial han sido sometidos al mismo estudio de estabilidad y los resultados obtenidos son similares. Los dos métodos de inserción aplicados son igualmente estables frente a la prueba de parcial bootstrap efectuada.

Tabla 2. Medias de las proyecciones de 30 réplicas en los tres primeros factores

m1	m2	m3	m5	m6	m7	m11	m13	m14	m15	m16	m17	m18	m20	m21
0.16	-0.11	-0.01	-0.05	-0.37	0.37	0.00	-0.34	0.12	-0.18	0.11	0.19	-0.19	-0.06	-0.19
0.05	0.10	-0.13	-0.17	0.27	0.08	-0.16	0.13	0.05	-0.01	-0.12	0.04	-0.03	-0.12	-0.02
0.08	0.08	0.06	0.14	-0.07	-0.53	-0.08	0.02	-0.02	0.07	0.11	0.14	0.01	0.07	0.02

Tabla 3. Medias de las proyecciones de 150 réplicas en los tres primeros factores

m1	m2	m3	m5	m6	m7	m11	m13	m14	m15	m16	m17	m18	m20	m21
0.15	-0.11	-0.01	-0.05	-0.36	0.37	0.00	-0.35	0.12	-0.18	0.11	0.19	-0.19	-0.07	-0.19
0.05	0.11	-0.13	-0.17	0.26	0.08	-0.16	0.13	0.05	-0.02	-0.11	0.04	-0.03	-0.11	-0.02
0.08	-0.11	0.06	0.14	-0.07	-0.53	-0.08	0.02	-0.02	0.08	0.12	0.14	0.01	0.07	0.02

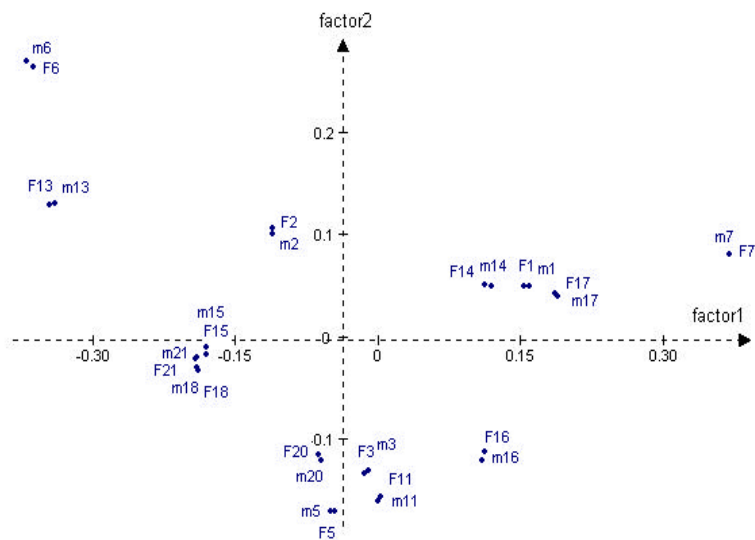


Fig. 6 Inserción de tiempos en festivo y medias de 30 réplica

6. CONSIDERACIONES FINALES

I) La inserción en referencia (Sección 3.1) puede parecer más natural si el tamaño de las dos muestras es muy distinto, la filosofía con la que se plantea es la de una gran encuesta que se toma como referencia y una o sucesivas encuestas con muchos menos individuos que se insertan en el subespacio de referencia obtenido de la primera. Pensamos, sin embargo, que la inserción en un subespacio de representación global es preferible a la basada en un subespacio de representación parcial. La proyección de las variables del conjunto Z se efectúa, en el análisis global, sobre un subespacio de representación más próximo al de la propia muestra y en nuestro caso de la submuestra $E2$. Por otra parte la inserción en referencia no presenta ninguna ventaja desde el punto de vista computacional y es mas delicada en cuanto a las condiciones de similitud en la estructura de las variables comunes de las dos encuestas, véase Rius, R., (1998, p. 142).

II) Observamos la suficiencia de 30 réplicas para este tipo de aplicaciones bootstrap tal como sugieren L. Lebart y M. Lejeune (1995a, p. 15) para el caso de fusión de encuestas.

III) Consideramos que sería interesante la inclusión de estos métodos en los paquetes estadísticos habituales, para su posible aplicación de manera compacta y automática.

ANEXO1

Variables comunes (de caracterización) junto a sus modalidades

Variable	Descripción	Identificador	Modalidad
X1	Edad	ED1 ED2 ED3	Hasta 34 años Entre 35 y 59 años 60 años o más
X2	Sexo	VAR MUJ	Varón Mujer
X3	Estado civil	SOL CAS RES	Soltero Casado Resto
X4	Nivel de instrucción	PRI MED SUP	Primarios Medios Superiores
X5	Relación con la actividad	SRM OCU PAR JUB EST LAH OTR	Servicio militar Ocupados Parados Jubilados Estudiantes Labores del hogar Otros

ANEXO2

Variables específicas de uso del tiempo. Las variables Y1,...,Y24 corresponden a usos del tiempo en días laborables, y Z1,...,Z24 son las variables homólogas para los días festivos.

Variables	Descripción
Y1,Z1	Dormir
Y2,Z2	Higiene y cuidado personal
Y3,Z3	Comer
Y4,Z4	Actividades privadas y actividades no descritas
Y5,Z5	Trabajo
Y6,Z6	Formación
Y7,Z7	Tareas domésticas (cocinar, fregar, limpiar la casa arreglo y cuidado ropa y trabajos diversos)
Y8,Z8	Compras
Y9,Z9	Gestiones
Y10,Z10	Actividades de semiocio (punto, costura, pintura, escultura, reparaciones, bricolaje, jardinería, cuidado de animales..)
Y11,Z11	Cuidado de niños y adultos
Y12,Z12	Reuniones de tipo familiar (comidas, defunciones, bodas, visitas hospitalarias..)
Y13,Z13	Reuniones con amigos, fiestas, ir de potes o copas...
Y14,Z14	Participación religiosa o política
Y15,Z15	Gimnasia y deporte
Y16,Z16	Excursiones y paseos
Y17,Z17	Ocio en el hogar (TV, vídeo, música, radio...)
Y18,Z18	Ocio fuera del hogar (cine, teatro, conciertos, museos y exposiciones, espectáculos deportivos..)
Y19,Z19	Otras actividades de ocio (micro informática, fotografía, cartas, juegos, crucigramas...)
Y20,Z20	Trayecto al trabajo o formación
Y21,Z21	Acompañar a otros
Y22,Z22	Esperas en el trabajo o formación
Y23,Z23	Esperas en cuidados médicos y gestiones administrativas
Y24,Z24	Otras esperas

Referencias

Aluja-Banet, T., Rius, R., Nonell, R., and Martínez-Abarca, M. J. (1997). Data fusion and file grafting. *Analyses Multidimensionnelles des Données*. IV-ème Congrès International NGUS'97 . K. Fernández-Aguirre et A. Morineau (Eds.), p. 7-13, CISIA-CERESTA.

Bonnefous S., Brenot J., et Pàges J. P. (1986). Methode de la greffe et communication entre enquêtes. *Data Analysis and Informatics IV*. Diday, p. 603-617. North Holland.

Chateau, F. and Lebart, L. (1996). Assessing Sample Variability in the Visualization Techniques Related to Principal Component Analysis: Bootstrap and Alternative Simulation Methods. *Proceedings of the COMPSTAT 1996*, p. 205-210: Physica-Verlag

Efron, B. (1979). Bootstrap methods: another look at the jackknife. *Ann. Statist.*, 7, n.1, p. 1-26

EUSTAT (1997). *Análisis de Tipologías de Jornadas Laborales*. Vitoria/Gasteiz: Instituto Vasco de Estadística (EUSTAT).

Gifi, A. (1990). *Nonlinear Multivariate Analysis*. Department of Data Theory, University of Leiden: John Wiley & Sons. Chichester

Greenacre, M. J. (1984). *Theory and Applications of Correspondence Analysis*: Academic Press, London

Iliakopoulos, A. (1988). *Etude de l'opinion et enquête de référence*. PhD thesis. Paris-Dauphine.

Lebart, L., Morineau, A. and Warwick, K. (1984). *Multivariate Descriptive Statistical Analysis*. J. Wiley, New York

Lebart, L., and Lejeune, M. (1995a). Assessment of Data Fusion and Injection. *Encuentro Internacional AIMC sobre Investigación de Medios*, p. 1-18, Madrid, 1995

Lebart, L., Morineau, A., Piron, M. (1997b). *Statistique exploratoire multidimensionnelle 2. édition*: Dunod, Paris

Lejeune, M. (1995). De L'usage des Fusions de Données dans les Etudes de Marché. *IASS Proceedings – Beijing 1995*, p. 137-149

Millan, L., and Whittaker, J. (1995). Application of the Parametric Bootstrap to Models that Incorporate a Singular Value Descomposition. *Appl. Statist.* 44, 1, p. 31-49

Rius, R. (1998). *Iserció de Dades d' Enquesta*. Tesis Doctoral. Universidad Politécnica de Catalunya

SPADN (1997) *Système Portable d'Analyse des Données Numériques*. CISIA, Paris.

S-PLUS 4.0 (1998) Mathsoft