00321

UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

76

FACULTAD DE CIENCIAS

"ANÁLISIS DE CORRESPONDENCIAS PARA EL ESTUDIO DE ASOCIACIÓN ENTRE VARIABLES CATEGÓRICAS: UN ENFOQUE APLICADO"

T E S I S

QUE PARA OBTENER EL TÍTULO DE

A C T U A R I O

P R E S E N T A :

IVAN ALEJANDRO PINEDA SANTA CRUZ

DIRECTOR: M. EN C. INOCENCIO RAFAEL MADRID RÍOS

MÉXICO, D.F. 2003

PAGINACION DISCONTINUA

UNIVERSIDAD NACIONAL
AUTÓNOMA DE
MEXICO

DRA. MARÍA DE LOURDES ESTEVA PERALTA
Jefa de la División de Estudios Profesionales de la
Facultad de Ciencias
Presente

Comunicamos a usted que hemos revisado el trabajo escrito: Análisis de correspondencias para el estudio de asociación entre variables categóricas: un enfoque aplicado.

realizado por Iván Alejandro Pineda Santa Cruz

con número de cuenta 09220414-2 , quién cubrió los créditos de la carrera de actuaría.

Dicho trabajo cuenta con nuestro voto aprobatorio.

Atentamente

Director de Tesis Propietario	M. en C. Inocecnio Rafael Madrid Ríos.
Propietario	Dr. Rubén Hernández Cid.
Propietario	Mat. Margarita Elvira Chávez Cano.
Suplente	M. en A.P. María Del Pilar Alonso Reyes.
Suplente	Act. Jaime Vázquez Alamilla.

Consejo Departamental de Matemáticas

M. en C. José Antonio Flores Díaz

FACULTAD DE CIENCIAS
CONSEJO DEPARTAMENTAL
DE
MATEMÁTICAS

Doy gracias:

Al corazón del cielo y al corazón de la tierra,
por el infinito amor con que me crearon.
Ome teotl

A mis padres: Constantino y Esther , por su amor que transformado en paciencia y esmero lograron pulirme como una piedra preciosa.
A mis hermanos: Hugo, Ricardo, Lilia, Edgar, Oswaldo y Yolotzin.

A todos mis maestros, y en especial a:

M. en C. Inocencio Rafael Madrid Ríos.
Dr. Rubén Hernández Cid.
Mat. Margarita Elvira Chávez Cano.
M. en A.P. María Del Pilar Alonso Reyes.
Act. Jaime Vázquez Alamilla.

Por ser amigos y guías en esta aventura que se llama vida. Porque ellos todo lo sacan de su corazón; obran con deleite, todo lo obran con calma, con cuidado, crean, inventan hábilmente lo que piensan, arreglan las cosas, dialogan con su corazón, encuentran las cosas en su mente...

A todos mis amigos que han contribuido con sus trazos a dibujar la historia de mi vida.

A todas las personas que me ayudaron en la realización de este trabajo.

Recuerda:

"Eres águila, eres ocelote, es tu don, es tu merecimiento... que no desfallezcan , ni tu venerable mano, ni tu venerable pie..."

Huehuetlahtolli

"Sólo quien se tome tiempo alcanzará la sabiduría".
Mexicah Tlamatiliztli.

"Y todos somos estrellas cuando descubrimos nuestra esencia".

Celsa Xolalohco

Capitulo III
Análisis factorial de correspondencias múltiples
(AFC-M)

Conclusiones

Anexos

Prefacio

Dada la demanda de profesionistas calificados en el manejo práctico de los métodos estadísticos es indudable la importancia que tiene la incorporación a los cursos que se imparten en la Facultad de Ciencias de la UNAM, la enseñanza de otras técnicas estadísticas como el análisis factorial de correspondencias, debido a su diversidad de aplicaciones prácticas.

El análisis de correspondencias es una técnica estadística que pemite describir de manera gráfica la asociación de un conjunto de datos cualitativos.

Antes de que apareciera el análisis de correspondencias ya habían otros métodos estadísticos que permitían estudiar esta asociación por ejemplo la prueba χ^2 de Pearson o los modelos log lineales; sin embargo esta técnica empezó a generar interés debido a la relativa facilidad con que se puede interpretar una gráfica en lugar de una estadística de prueba, sobre todo para personas cuya formación no es matemática o estadística. Sin embargo, esta facilidad puede ser engañosa pues pueden darse por ciertas, conclusiones erróneas cuando existe falta de entendimiento de la técnica en sus supuestos, sus alcances y en la calidad de sus resultados.

La motivación para escribir este trabajo es mostrar los aspectos que deben tomarse en cuenta para una adecuada interpretación del análisis de correspondencias tanto en sus aspectos geométricos, algebráicos y prácticos.

Introducción

La estadística puede definirse como el desarrollo y aplicación de técnicas para colectar, analizar e interpretar datos cuantitativos o cualitativos. De tal manera que cuando se tiene la información o los datos acerca de un determinado fenómeno que se pretenda estudiar los métodos estadísticos pueden ser utilizados para tratar de describirlos o explicarlos.

Los métodos estadísticos se han desarrollado para estudiar la realidad tomando en cuenta la incertidumbre que un fenómeno aleatorio tiene. Una corriente de pensamiento que ha desarrollado algunos métodos para analizar datos, es la llamada escuela francesa cuyo enfoque se basa en la geometría y el álgebra lineal lo que permite realizar el estudio global de un gran número de variables, con un mínimo de supuestos. El análisis de correspondencias pertenece a esta escuela de pensamiento y permite estudiar la dependencia o asociación por ejemplo entre dos variables cualitativas o factores (variables que no toman valores numéricos), que generalmente son presentadas en forma de tablas de contingencia. En dichas tablas, los renglones y las columnas representan dos particiones de la misma población. Mediante el análisis de correspondencias es posible estudiar la "proximidad" entre las categorías de cada factor, y también se puede estudiar la proximidad entre las categorías de dos o más factores.

H.O.Hartley publicó en 1935 un artículo en el cual da una formulación algebráica de la correlación entre los renglones y columnas de una tabla de contingencia, de tal manera que se le puede atribuir el origen matemático del análisis de correspondencias. Louis Guttman (1941) trató el caso general de más de dos variables cualitativas desarrollando lo que hoy se conoce como análisis de correspondencias múltiples. La forma algebráica geométrica del análisis de correspondencias fue desarrollada por el francés Jean-Paul-Benzecri (1976) dentro de la llamada escuela francesa del análisis de datos. El término francés *correspondence*, fue usado para denotar *sistemas de asociaciones* entre los elementos de dos conjuntos: los renglones y las columnas de una tabla de contingencia.

El análisis de correspondencias puede ser utilizado para estudiar la asociación de más de dos factores generalizándose las ideas desarrolladas para el caso de dos factores además la interpretación de los factores sigue teniendo las mismas reglas que en el análisis de correspondencias simple. Sin embargo, existen algunos inconvenientes en el momento de interpretar los resultados ya que como se aumenta la dimensión del espacio para que las categorías de los factores puedan ser representados en un espacio de menor dimensión se tiene una pérdida de información que se traduce en que los porcentajes de inercia en los ejes son bajos y la descripción de la gráfica no es del todo satisfactoria, es decir, la parte de la inercia explicada pierde aquí su interés ya que los valores propios no representan más que una pequeña parte de la inercia total.

Las formas de realizar el análisis de correspondencias son las siguientes: utilizando la información por individuo que se encuentra en una matriz indicadora denominada por Z o bien utilizando la matriz de Burt; la cual recoge la clasificación de la muestra de acuerdo a cada factor y es en si una generalización de la tabla de contingencia. Cuando se trabaja con la matriz indicadora Z, se puede hacer uso del, análisis de correlación canónico.

En el capitulo I se introducen conceptos básicos que son requeridos para la comprensión del análisis de correspondecias.

En el capitulo II. Análisis Factorial de Correspondencias simples. Se estudiarán los elementos que permiten comprender el tratamiento geométrico y algebráico de los datos, y se desarrollará el análisis de correspondencias para una tabla de contingencia de $I \times J$.

En el capitulo III. Análisis de Correspondencias Múltiples. Se presenta el análisis de correspondencias múltiles, como una generalización del análisis de correspondencias simple, además presentarán las formas equivalentes entre el análisis de correspondencias llevado a cabo con los datos de la matriz indicadora Z o bien mediante el uso de la matriz de Burt. Así mismo, se presentarán ejemplos que ayuden a comprender la interpretación del análisis.

Finalmente se presentarán las conclusiones que se obtuvieron con el estudio del análisis de correspondencias aplicado a los ejemplos propuestos.

Para la lectura de esta tesis, se requieren conocimientos de álgebra lineal y de geometría analítica por lo que se han incluido varios anexos y bibliografía recomendada de modo que el lector interesado puede consultarlos al final del presente trabajo.

El paquete estadístico computacional que se utilizó fue Xlstat V.4.

Capítulo I
Elementos básicos.

En este capítulo, se mencionan los elementos que se necesitarán para la comprensión del análisis factorial de correspondencias simple (AFC-simple).

Conceptos tales como espacio vectorial, distancia entre los elementos que lo conforman y la representación gráfica de los elementos dentro del espacio, permiten el manejo de la información contenida en la tabla de contingencia de la cual hace uso el AFC .

1.1 Tabla de contingencia de dos dimensiones

Las poblaciones se describen mediante medidas numéricas llamadas parámetros. Desafortunadamente la mayoría de las veces la obtención de estos parámetros resulta demasiado costosa y es de difícil obtención. Uno de los objetivos de la estadística es inferir acerca de alguna o varias características de la población con base en la información contenida en una muestra haciendo uso de estimadores para esos parámetros. Un estimador puede ser entendido como una fórmula para calcular un valor estimado o representativo de un parámetro de la población basada en las mediciones contenidas en una muestra.

Existe toda una metodología para obtener muestras representativas de una población, para los propósitos del presente trabajo supondremos que se cuenta con una muestra de tamaño T de una distribución multinomial.

En forma general podemos decir que una tabla de contingencia de $I \times J$ es un arreglo en el cual una muestra de T observaciones es clasificada con respecto a dos factores o variables cualitativas.

Una variable cualitativa es una transformación que va de un conjunto de unidades de estudio a un conjunto de valores.

X: Unidades de estudio→ *Valores*

De tal manera que si V es un subconjunto del conjunto de valores, entonces

i) Si V sólo tiene estructura de orden, la variable X es ordinal, es decir todas aquellas variables que establecen cierta jerarquía. Por ejemplo: grado de estudios, nivel económico, preferencias, etc.

ii) Si V no tiene ninguna estructura en particular, la variable X es nominal, es decir son aquellas a las que se les asigna un nombre o etiqueta. Por ejemplo: estado civil, religión que profesa una persona, tipo de enfermedad, sexo, etc.

A estos dos tipos de variables comunmente se les conoce como variables cualitativas,

categóricas o factores. En lo que sigue se hará referencia a las variables cualitativas como factores.

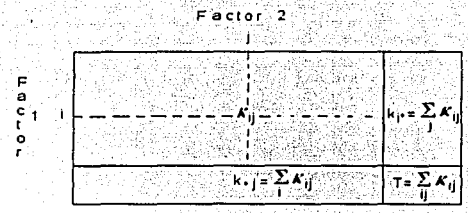


Fig. 1 Tabla de contingencia de dos factores

En la figura 1 se muestra esquemáticamente la forma de la tabla de contingencia en la que el primer factor tiene I categorías y el segundo factor tiene J categorías que corresponden a los renglones y columnas respectivamente.

La frecuencia observada en la i-ésima categoría del factor I y la j-ésima categoría del factor J se define como k_{ij}. El término k_{ij} debe cumplir con lo siguiente: $k_{ij} \geq 0$ para i=1,2,3,...,I ; j=1,2,3,...,J. Y, además para los renglones y columnas en la tabla se debe cumplir que $\sum_{j=1}^{J} k_{ij} > 0$ para i=1,2,3,...,I y $\sum_{i=1}^{I} k_{ij} > 0$ para j=1,2,3,...,J. Lo que significa que debe de existir para cada columna o renglón al menos algún valor distinto de cero.

Si suponemos que la muestra de tamaño T se distribuye como una variable aleatoria multinomial. Es decir se tiene la probabilidad

$$P(e_{11} = k_{11}, e_{12} = k_{12}, \ldots, e_{IJ} = k_{IJ}) = \frac{T! \prod_{i=1}^{I} \prod_{j=1}^{J} p_{ij}^{k_{ij}}}{\prod_{i=1}^{I} \prod_{j=1}^{J} k_{ij}!}$$

con

$$\sum_{i=1}^{I} \sum_{j=1}^{J} k_{ij} = T$$

$$\sum_{i=1}^{I} \sum_{j=1}^{J} p_{ij} = 1$$

Donde p_{ij} es la probabilidad de que de los T elementos de la población k_{ij} estén en la categoría i del factor I y la categoría j del factor J (celda (i,j)) de la tabla de contingencia. De tal manera que se tiene que estimar el valor de p_{ij}.

Mediante el método de máxima verosimilitud se sabe que un estimador de p_{ij} es $\hat{p}_{ij} = \frac{k_{ij}}{T}$ además por este mismo método un estimador para $p_{i\cdot}$ es $\hat{p}_{i\cdot} = \frac{k_{i\cdot}}{T}$ y para $p_{\cdot j}$ es $\hat{p}_{\cdot j} = \frac{k_{\cdot j}}{T}$ (Christensen,pág 41-43).

Como se verá más adelante estos estimadores son los que definen los pesos o masas de las

categorías de los factores y además definen a la matriz de correspondencias.

Dentro de la tabla de contingencia se tienen los siguientes conceptos: *Renglón marginal* o renglón de totales por columna: es aquel que se obtiene al sumar los elementos de cada una de las j columnas de la tabla de contingencia y cuyo j-ésimo término es $k_{\cdot j} = \sum_{i=1}^{I} k_{ij}$

Columna marginal; de manera análoga se tiene la columna de totales por renglón la cual se obtiene al sumar los elementos de cada uno de los i renglones de la tabla de contingencia y cuyo i-ésimo término es $k_{i\cdot} = \sum_{j=1}^{J} k_{ij}$. La suma de todos los elementos de la tabla de contingencia corresponde al tamaño de la muestra llamado también total o gran total que denotaremos como $T = \sum_{j=1}^{J} \sum_{i=1}^{I} k_{ij}$. Véase Fig 1.

1.2 Conceptos geométricos y algebráicos obtenidos de la tabla de contingencia, necesarios para el desarrollo del AFC-simple.

A continuación se expone el desarrollo algebráico del AFC en el cual la información de las categorías de la tabla de contingencia es tratada como vectores en diferentes espacios ; con esto se obtiene la ventaja de la aplicación del algebra lineal y la geometría analítica.

Con lo mencionado hasta el momento podemos relacionar la tabla de contingencia con una matriz, que denominaremos como $K_{I\times J}$ que tiene la siguiente forma:

$$
\begin{array}{c}
 \\
 \\
 \\
1 \\
\cdots \\
\textit{Factor 1}\ \ i \\
\cdots \\
I
\end{array}
\begin{array}{c}
\textit{Factor 2} \\
1 \quad 2 \quad \cdots\cdots \quad j \quad \cdots\cdots \quad J \\
\left[\begin{array}{ccccccc}
k_{11} & k_{12} & \cdots & k_{1j} & \cdots & k_{1J} \\
\cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\
k_{i1} & k_{i2} & \cdots & k_{ij} & \cdots & k_{iJ} \\
\cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\
k_{I1} & k_{I2} & \cdots & k_{Ij} & \cdots & k_{IJ}
\end{array}\right]
\end{array}
= K_{I\times J} \qquad (I)
$$

En este trabajo se usarán las siguientes notaciones:la notación punto, en la cual los puntos indican las sumas sobre los indices particulares donde aparecen puntos; por ejemplo $k_{i\cdot}$ indicará la suma de los elementos que forman el renglón i de la tabla de contingencia, ($k_{i\cdot} = k_{i1} + k_{i2} + k_{i3} + \ldots + k_{iJ} = \sum_{j=1}^{J} k_{ij}$), mientras que $k_{\cdot j}$, indicará la suma de los elementos que foman la columna j de la tabla de contingencia
($k_{\cdot j} = k_{1j} + k_{2j} + \ldots + k_{Ij} = \sum_{i=1}^{I} k_{ij}$); notación suma y la notación desarrollada

$k_{i\cdot}$	Notación punto.
$k_{i1} + k_{i2} + k_{i3} + \ldots + k_{iJ}$	Notación desarrollada.
$\sum_{j=1}^{J} k_{ij}$	Notación suma.

De esta manera se puede pensar que la tabla de contingencia puede ser vista como una matriz que está formada por I vectores renglón, cada uno con J componentes; o, formada por J vectores columna, cada uno con I componentes. Es decir en una matriz se encuentran implícitos dos espacios vectoriales: R^J para los renglones y R^I para las columnas.

La ventaja de trabajar con matrices, es desde luego el uso matemático de los espacios vectoriales el cual permite desarrollar un proceso de análisis de datos para estudiar la asociación entre variables categóricas con los datos presentados en la tabla de contingencia.

Los términos de la matriz $K_{I\times J}$ representan las frecuencias absolutas que aparecen en la tabla de contingencia.

Se puede obtener la matriz de frecuencias relativas (proporciones) con respecto al total de observaciones T multiplicando a la matriz K por el reciproco de la suma total T (recordemos que el estimador máximo verosimil de p_{ij} es $\hat{p}_{ij} = \frac{k_{ij}}{T}$ al cual de ahora en adelante lo denotaremos por f_{ij}) La matriz F representa la probabilidad conjunta estimada para cada celda , es decir la probabilidad de que la información esté clasificada en la categoría i del factor 1 y la categoria j del factor 2

$$F_{I\times J} = \frac{1}{T} K_{I\times J}$$

$$F_{I\times J} = \begin{bmatrix} f_{11} & f_{12} & \dots & f_{1j} & \dots & f_{1J} \\ f_{21} & f_{22} & \dots & f_{2j} & \dots & f_{2J} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ f_{i1} & f_{i2} & \dots & f_{ij} & \dots & f_{iJ} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ f_{I1} & f_{I2} & \dots & f_{Ij} & \dots & f_{IJ} \end{bmatrix} \quad \text{con } f_{ij} = \hat{p}_{ij} = \frac{k_{ij}}{T} \qquad \text{(II)}$$

1.2.1 Peso o relevancia del renglón i, dado por los datos.

El peso del renglón i puede entenderse como la importancia que tiene el $i-$ *ésimo* elemento del primer factor dentro del conjunto de los renglones que es determinado en términos de proporción con respecto a los demás renglones, y se calcula como el cociente entre la suma de los elementos del i-ésimo renglón de la matriz K (k_{i*}) y el total (T),

$$\frac{k_{i*}}{T} = \frac{\sum_{j=1}^{J} k_{ij}}{T}.$$

Con la matriz F el peso del renglón i es $f_{i*} = \sum_{j=1}^{J} f_{ij}$.

Así se tienen las siguientes equivalencias:

$$f_{i\bullet} = \sum_{j=1}^{J} f_{ij} = \frac{k_{i\bullet}}{T} = \frac{\sum_{j=1}^{J} k_{ij}}{T}$$

Estos i términos forman un vector, el vector de pesos por renglones:

$$r = \begin{bmatrix} f_{1\bullet} \\ f_{2\bullet} \\ \cdots \\ f_{I\bullet} \end{bmatrix}$$ Nótese que $\sum_{i=1}^{I} f_{i\bullet} = 1$ (III)

1.2.2 Peso o relevancia de la columna j, dado por los datos.

De manera análoga, el peso de la columna j puede entenderse como la importancia que tiene el *j – ésimo* elemento del segundo factor dentro del conjunto de las columnas el cual es determinado por la tabla de contingencia en términos de proporción con respecto a las demás columnas , y se calcula como el cociente entre el total de la j-ésima columna de la matriz K y la suma total.

$$\frac{k_{\bullet j}}{T} = \frac{\sum_{i=1}^{I} k_{ij}}{T}$$

Con la matriz F el peso de la columna *j* es $f_{\bullet j} = \sum_{i=1}^{I} f_{ij}$.

Teniendose las siguientes equivalencias:

$$f_{\bullet j} = \sum_{i=1}^{I} f_{ij} = \frac{k_{\bullet j}}{T} = \frac{\sum_{i=1}^{I} k_{ij}}{T}$$

Estos j términos forman el vector $c_{1\times J}$; el vector, el vector de pesos por columnas que denotaremos por:

$$c = \begin{bmatrix} f_{\bullet 1} & f_{\bullet 2} & \cdots & f_{\bullet p} \end{bmatrix}$$ (IV)

Observación: $\sum_{i=1}^{I}\sum_{j=1}^{J} \frac{k_{ij}}{T} = \sum_{i=1}^{I}\sum_{j=1}^{J} f_{ij} = \sum_{i=1}^{I} f_{\bullet i} = \sum_{j=1}^{J} f_{i\bullet} = 1$

1.2.3 Perfiles por renglón y por columna.

A los elementos del i-ésimo renglón de la matriz K dividido entre su total ($k_{i\bullet}$) se le conoce como perfil renglón.

Esto es: $f_{ij} = [\frac{k_{i1}}{k_{i\bullet}}, \frac{k_{i2}}{k_{i\bullet}}, \ldots, \frac{k_{iJ}}{k_{i\bullet}}]$ (V)

Observación: La notación f_{ij} quiere decir que el i-ésimo perfil tiene *J* elementos, es decir es un vector del espacio vectorial R^J.

A la j-ésima columna de la matriz K dividida entre el total de la columna se le conoce como perfil columna.

Esto es: $f_J = \begin{bmatrix} \frac{k_{1J}}{k_{\cdot J}} \\ \frac{k_{2J}}{k_{\cdot J}} \\ \cdots \\ \frac{k_{IJ}}{k_{\cdot J}} \end{bmatrix}$ (VI)

Observación: La notación f_J quiere decir que el j-ésimo perfil tiene I elementos, es decir es un vector del espacio vectorial R^I.

1.2.4 Centroide.

Hasta ahora hemos caracterizado algunos elementos que se pueden obtener de la tabla de contingencia en cada uno de los dos espacios generados por los dos factores que intervienen. En la estadística descriptiva, por ejemplo, se habla de medidas de tendencia central dentro de las cuales el promedio es una de las más utilizadas y que en términos generales pone de manifiesto alrededor de que valor oscilan los valores de una muestra, es decir, es un centro de gravedad; nótese que cuando se obtiene este promedio cada elemento considerado tiene igual importancia dentro de la muestra es decir tienen masa uno. Hasta ahora, hemos trabajado con vectores, por lo que podemos preguntarnos ¿existirá un vector promedio alrededor del cual se encuentran los vectores renglón y columna de la matriz K, y que además tome en cuenta su peso o masa?

Por ejemplo tomemos el caso de los renglones o sea el espacio R^J. Como hemos visto tenemos I vectores $\begin{bmatrix} f_1 \\ f_2 \\ \cdots \\ f_J \end{bmatrix}$ y cada uno de estos vectores tiene por peso $\begin{pmatrix} f_{1\cdot} \\ f_{2\cdot} \\ \cdots \\ f_{I\cdot} \end{pmatrix}$ el centro de gravedad o centroide o vector promedio es un vector $\in R^J$ al cual lo denotaremos por G_{R^J} y es tal que:

$$G_{R^J} = \frac{\sum_{i=1}^{I} f_{i\cdot} f_{ij}}{\sum_{i=1}^{I} f_{i\cdot}} = \sum_{i=1}^{I} f_{i\cdot} f_{ij} = (f_{\cdot 1}, f_{\cdot 2}, \ldots f_{\cdot J}) \qquad (VII)$$

El centroide para el espacio de los renglones representa el vector promedio de los I perfiles por renglón ponderados.

De manera análoga se tiene que el centroide para el espacio de las columnas es:

$$G_{IR'}=\frac{\sum_{j=1}^{J} f_{ij}f_i}{\sum_{j=1}^{J} f_{ij}} = \sum_{j=1}^{J} f_{ij}f_i = (f_{1\bullet}, f_{2\bullet}, \ldots, f_{I\bullet})' \qquad (VIII)$$

El cual representa el vector promedio de los perfiles por columna ponderados

1.2.5 Matriz perfil por renglones y columnas.

Al hacer para cada uno de los renglones o columnas la división respectiva para obtener los perfiles resulta tedioso. Una solución es el manejo matricial de la información.

Denotaremos por $R_{I\times J}$ a la matriz perfil por renglones

En términos algebráicos tenemos que:

$$R_{I\times J} = D_{r_{I\times I}}^{-1} F_{I\times J} =$$

$$\begin{bmatrix} \frac{f_{11}}{f_{1\bullet}} & \frac{f_{12}}{f_{1\bullet}} & \cdots & \frac{f_{1j}}{f_{1\bullet}} & \cdots & \frac{f_{1J}}{f_{1\bullet}} \\ \frac{f_{21}}{f_{2\bullet}} & \frac{f_{22}}{f_{2\bullet}} & \cdots & \frac{f_{2j}}{f_{2\bullet}} & \cdots & \frac{f_{2J}}{f_{2\bullet}} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \frac{f_{i1}}{f_{i\bullet}} & \frac{f_{i2}}{f_{i\bullet}} & \cdots & \frac{f_{ij}}{f_{i\bullet}} & \cdots & \frac{f_{iJ}}{f_{i\bullet}} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \frac{f_{I1}}{f_{I\bullet}} & \frac{f_{I2}}{f_{I\bullet}} & \cdots & \frac{f_{Ij}}{f_{I\bullet}} & \cdots & \frac{f_{IJ}}{f_{I\bullet}} \end{bmatrix} = \begin{bmatrix} f_{1j} \\ f_{2j} \\ \cdots \\ f_{ij} \\ \cdots \\ f_{Ij} \end{bmatrix} \qquad (IX)$$

y cuyo término general es $\frac{k_{ij}}{k_{i\bullet}} = \frac{f_{ij}}{f_{i\bullet}}$;

donde $D_r^{-1}{}_{I\times I}$ es la matriz diagonal asociada al vector de pesos por renglón es decir

$$D_r^{-1}{}_{I\times I} = \begin{bmatrix} \frac{1}{f_{1\bullet}} & 0 & 0 & 0 \\ 0 & \frac{1}{f_{2\bullet}} & 0 & 0 \\ 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & \frac{1}{f_{I\bullet}} \end{bmatrix}$$

"Lo que se logra al construir la matriz $R_{I\times J}$ es expresar las probabilidades condicionales de la forma:

Pr(categoría del factor (2)| categoría del factor(1))".

De manera análoga denotaremos por $C_{I\times J}$ la matriz perfil por columnas.

$$C_{I\times J} = F_{I\times J} D_c^{-1}{}_{J\times J}$$

$$= \begin{bmatrix} \frac{f_{11}}{f_{.1}} & \frac{f_{12}}{f_{.2}} & \cdots & \frac{f_{1j}}{f_{.j}} & \cdots & \frac{f_{1J}}{f_{.J}} \\ \frac{f_{21}}{f_{.1}} & \frac{f_{22}}{f_{.2}} & \cdots & \frac{f_{2j}}{f_{.j}} & \cdots & \frac{f_{2J}}{f_{.J}} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \frac{f_{i1}}{f_{.1}} & \frac{f_{i2}}{f_{.2}} & \cdots & \frac{f_{ij}}{f_{.j}} & \cdots & \frac{f_{iJ}}{f_{.J}} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \frac{f_{I1}}{f_{.1}} & \frac{f_{I2}}{f_{.2}} & \cdots & \frac{f_{Ij}}{f_{.j}} & \cdots & \frac{f_{IJ}}{f_{.J}} \end{bmatrix} = \begin{bmatrix} f_1^J & f_2^J & \cdots & f_i^J & \cdots & f_I^J \end{bmatrix} \qquad (X)$$

cuyo término general es $\frac{k_{ij}}{k_{.j}} = \frac{f_{ij}}{f_{.j}}$

donde $D^{-1}_{J\times J}$ es la matriz diagonal asociada al vector de sumas por columna de F es decir.

$$D_C^{-1}{}_{J\times J} = \begin{bmatrix} \frac{1}{f_{.1}} & 0 & 0 & 0 \\ 0 & \frac{1}{f_{.2}} & 0 & 0 \\ 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & \frac{1}{f_{.J}} \end{bmatrix}$$

De manera análoga las celdas que forman la matriz *C* representan la probabilidad condicional de la forma:

Pr(categoria del factor(1)| categoria del factor(2))"

Estas probabilidades condicionales nos van a permitir presentar de una manera intuitiva el concepto de asociación, y nos permitirá introducir el análisis de correspondencias.

Ya que la matriz de datos puede verse como un arreglo de vectores por renglones o un arreglo de vectores por columnas, entonces quedan determinados dos espacios vectoriales R^I y R^J. Dentro de esos espacios vectoriales los renglones en R^J y las columnas en R^I, permiten determinar la matriz perfil por renglones y la matriz perfil por columnas respectivamente. Cada uno de estos renglones o columnas tienen asignado un peso o masa en función de la frecuencia de los datos.

El hecho de que se trabaje con vectores que pertenezcan a un espacio vectorial en el cual se puedan medir distancias entre vectores nos permite por ejemplo tratar de encontrar vectores parecidos en términos de sus proximidades, es decir podemos trabajar en un espacio métrico, por lo que tenemos una manera de medir distancias: métrica. Usualmente se ha trabajado en un espacio euclidiano en el cual se tiene la distancia usual:

Un espacio euclidiano por ejemplo de n dimensiones es el conjunto de todos los vectores con n componentes para los cuales las operaciones de suma vectorial y multiplicación por escalar son permitidas, y además, para cualquiera dos vectores en el espacio, existe un numero no negativo

llamado la distancia Euclidiana (métrica). La distancia Euclidiana para $\underline{a}$ y $\underline{b} \in R^n$ está definida como:

$$d^2(\underline{a},\underline{b}) = \|\underline{a} - \underline{b}\|^2 = (a_1 - b_1)^2 + (a_2 - b_2)^2 + \ldots + (a_n - b_n)^2$$

De manera alternativa podemos definir $\|\underline{a} - \underline{b}\|^2$ en términos del producto escalar de $(\underline{a} - \underline{b})$ el cual es

$d^2(\underline{a},\underline{b}) = \|\underline{a} - \underline{b}\|^2 = \langle(\underline{a} - \underline{b}),(\underline{a} - \underline{b})\rangle = [(\underline{a} - \underline{b})'(\underline{a} - \underline{b})]$ donde $(\underline{a} - \underline{b})$ es el vector de diferencia.

Nótese que esta métrica no hace diferencia con respecto a la importancia que tienen los vectores, es decir todos tienen la misma masa. Sin embargo, en el análisis de datos, los métodos deben ser adaptados a situaciones prácticas donde los individuos o las categorías no necesariamente tengan la misma "importancia". Así, por ejemplo si lo que se quiere es determinar la proximidad o cercanía entre elementos que se consideren tengan el mismos peso o sean igualmente importantes se podría pensar en utilizar la distancia euclidiana. Pero, también se da el caso en que el investigador pueda privilegiar a algunos elementos más que otros en función de su importancia con respecto a algún criterio o que la estructura interna de los datos así lo establezca, de tal manera que cada elemento tiene asociado una masa f_w tal que $f_w > 0$ y $\sum_w f_w = 1$. Así, si se quiere determinar la proximidad o cercanía entre estos elementos ponderados, entonces tendremos que utilizar alguna métrica que tome en cuenta ese peso o masa; una métrica que satisface lo anterior es la llamada : métrica Ji cuadrada, la cual define un espacio métrico ponderado.

1.2.6 Distancia Ji cuadrada

En un espacio ponderado.en el que a los vectores de este espacio se les asigna una masa o un peso. Algunos autores denominan a este espacio como un espacio euclidiano ponderado en el cual se define el producto punto o producto escalar de la siguiente forma: para dos vectores $\underline{v}, \underline{w}$, con masa q_1 y q_2 respectivamente tenemos:

$\langle \underline{v}, \underline{w} \rangle = \underline{v}' D_q \underline{w} = \sum_j q_j v_j w_j$, Donde D_q es la matriz diagonal asociada a los pesos de los puntos.

De tal manera que la distancia al cuadrado de los dos puntos en este espacio euclidiano ponderado es:

$$d^2(\underline{v}, \underline{w}) = \langle \underline{v} - \underline{w}, \underline{v} - \underline{w} \rangle = (\underline{v} - \underline{w})' D_q (\underline{v} - \underline{w}) = \sum_j q_j (v_j - w_j)^2.$$

En particular, cuando se trabaja con una tabla de contingencia se tiene la llamada distancia Ji cuadrada, la cual está definida para dos perfiles por renglones i e i' como

$$d^2(i,i') = \sum_{j=1}^{J} \frac{1}{f_{.j}} \left(\frac{f_{ij}}{f_{i.}} - \frac{f_{i'j}}{f_{i'.}} \right)^2$$ y para perfiles por columna j y j' como:

$$d^2(j,j') = \sum_{i=1}^{I} \frac{1}{f_{i.}} \left(\frac{f_{ij}}{f_{.j}} - \frac{f_{ij'}}{f_{.j'}} \right)^2$$

Esta distancia difiere de la distancia usual Euclidiana en que cada término al cuadrado está ponderado por el inverso del peso correspondiente a cada término del otro espacio.

Hasta el momento lo que se ha hecho es lo siguiente:

1. Se ha visto a la tabla de contingencia como una matriz, la cual define de manera natural el espacio de los renglones y el espacio de las columnas R^J y R^I respectivamente.

2. Se han obtenido los elementos por perfiles dependiendo en que espacio estemos trabajando. A cada uno de estos perfiles se les ha asignado una masa que está en relación a su importancia dentro del espacio definido por los elementos que conforman la tabla de contingencia.

3. Se ha encontrado el centroide de cada uno de los espacios definidos en la tabla de contingencia.

5. La masa que tienen los perfiles establecen la necesidad de trabajar en un espacio métrico ponderado, donde la métrica es la distancia ji-cuadrada.

Ejemplo1 : Relación entre los espacios R^I y R^J con los perfiles por columna y por renglón

Tabla 1.1. Tabla de contingencia (K): número de muebles producidos con defectos clasificados de acuerdo a tipo de defecto y turno de producción

Turno de producción	Tipo de defecto A	B	C	D	Total
1	15	21	45	13	94
2	28	31	34	5	98
3	33	17	49	20	119
Total	74	69	128	38	309

⇐ Matriz de frecuencias absolutas K_{ij}

Tabla 1.2 Tabla de frecuencias relativas (F): porcentaje de muebles producidos con defectos clasificados de acuerdo a tipo de defecto y turno de producción

Turno de producción	Tipo de defecto A	B	C	D	Total
1	0.049	0.068	0.146	0.042	0.304
2	0.084	0.1	0.11	0.016	0.311
3	0.107	0.055	0.159	0.065	0.385
Total	0.239	0.223	0.414	0.123	1

⇐ Matriz de correspondencias

Cuadro 1: Relación entre los espacios R^I y R^J con los perfiles por columna y por renglón

Espacio R^I para turno de producción, hay 3 renglones son los "perfiles por renglón" (puntos* en R^4)

- Las coordenadas del renglon i=1,2,3 son $\frac{f_{ij}}{f_{i.}}$ perfil para el renglón i: $(\frac{f_{i1}}{f_{i.}}, \frac{f_{i2}}{f_{i.}}, \ldots, \frac{f_{ip}}{f_{i.}})$ son

Perfiles por turno (renglones)

Turno de producción	Tipo de defecto A	B	C	D
1	0.16957	0.2234	0.47873	0.1283
2	0.37083	0.32292	0.25417	0.05208
3	0.27731	0.14286	0.41176	0.12907

- El peso del i-ésimo renglón es $f_{i.}$: $\begin{bmatrix} 0.304 \\ 0.311 \\ 0.385 \end{bmatrix}$

-Los I perfiles por renglón corresponden a los I renglones de la matriz: $D_I^{-1}F$

-La distancia entre dos renglones i e i' es:

$d^2(i,i')=\sum_{j=1}^{J} \frac{1}{f_{.j}}\left[\frac{f_{ij}}{f_{i.}} - \frac{f_{i'j}}{f_{i'.}}\right]^2$

$d^2(1,2)=\sum_{j=1}^{4} \frac{1}{f_{.j}}\left[\frac{f_{1j}}{f_{1.}} - \frac{f_{2j}}{f_{2.}}\right]^2=0.194$

Adicionalmente a lo revisado se verá en el capitulo siguiente:

- El criterio de ajuste está determinado por una forma cuadrática en la cual intervienen los pesos de los perfiles los cuales aparecen en la matriz D_r
- La matriz de distancias entre renglones está determinada por una forma cuadrática en la cual interviene D_J^{-1}.

Espacio R^J para tipo de defecto, hay 4 columnas son los "perfiles por columna"(puntos en R^3)

- Las coordenadas de la columna j=1,2,3,4, son $\frac{f_{ij}}{f_{.j}}$ y los perfiles por columna j: $(\frac{f_{1j}}{f_{.j}}, \frac{f_{2j}}{f_{.j}}, \ldots, \frac{f_{Ij}}{f_{.j}})$ son

Perfiles por tipo de defecto (columnas)

Turno de producción	Tipo de defecto A	B	C	D
1	1.2007	0.30435	0.35156	0.34211
2	0.35135	0.44928	0.29663	0.13158
3	0.44595	0.24638	0.35181	0.52632

- El peso de la j-ésima columna es $f_{.j}$: $\begin{bmatrix} 0.239 \\ 0.223 \\ 0.414 \\ 0.123 \end{bmatrix}'$

- Los J perfiles por columna corresponden a las J columnas de la matriz FD_C^{-1}

- La distancia entre dos columnas j y j' es:

$d^2(j,j')=\sum_{i=1}^{I} \frac{1}{f_{i.}}\left[\frac{f_{ij}}{f_{.j}} - \frac{f_{ij'}}{f_{.j'}}\right]^2$

$d^2(A,B)=\sum_{i=1}^{3} \frac{1}{f_{i.}}\left[\frac{f_{iA}}{f_{A.}} - \frac{f_{iB}}{f_{B.}}\right]^2=0.168$

Adicionalmente a lo revisado se verá en el capitulo siguiente:

- El criterio de ajuste está determinado por una forma cuadrática en la cual intervienen los pesos de los perfiles los cuales aparecen en la matriz D_c
- La matriz de distancias entre columnas está determinada por la forma cuadrática asociada con: D_I^{-1}

En el siguiente capítulo mencionaremos lo que es asociación entre las categorías haciendo referencia a la prueba Ji-cuadrada de Pearson para independencia y mostraremos el desarrollo del AFC-simple para representar gráficamente de manera óptima la asociación entre los perfiles renglones y columnas a través de la proximidad entre ellos ya sea entre un mismo espacio o entre ambos espacios mediante un escalamiento.

Capítulo II
Analisis factorial de correspondencias simple.

El análisis factorial de correspondencias (AFC), desarrollado por un grupo de estadísticos franceses desde principios de 1960, se reporta que es teóricamente equivalente a casos especiales de algunas técnicas clásicas que han aparecido en diferentes contextos desde la mitad de los años 30's, como por ejemplo el análisis de correlación canónico y el análisis de componentes principales. El nombre de análisis factorial tiene dos acepciones la primera referida al estudio de la asociación entre factores o variables cualitativas y el segundo a la técnica del análisis multivariado llamada análisis de factores en la que se busca resumir en menos dimensiones la información contenida en una matriz de datos.

El AFC, es una herramienta estadística exploratoria para estudiar la asociación entre las categorías de factores presentadas en una tabla de contingencia, en este sentido puede ser utilizado como una técnica complementaria a otros métodos estadísticos como los modelos log-lineales.

El AFC determina gráficamente la posición de una serie de unidades de una población bajo estudio, que pueden referirse a objetos o sujetos (segmentos del mercado, grupos de individuos o personas físicas, sectores, productos, etc.) con relación a una serie de factores de interés (atributos, características, escalas de valoración, etc.) clasificados por dos factores de interés y de acuerdo a la proximidad que tienen las categorías dentro de cada factor y entre factores, reporta o refleja en dos o en a lo mucho tres dimensiones la estructura de asociación tanto de las categorías dentro de cada factor como entre las categorías de los factores, determinando además un escalamiento de las categorías de los factores.

Para introducir las ideas referentes al AFC-S utilizaremos el siguiente ejemplo donde se estudia la asociación entre dos factores:

Supongamos que se quiere clasificar los defectos encontrados en los muebles producidos en una fábrica, de acuerdo a 1) el tipo de defecto y 2) el turno en que fue producido el mueble. De tal manera que un total de 309 muebles defectuosos fueron registrados y clasificados de acuerdo a uno de los siguientes tipos de defecto: A, B, C o D. Al mismo tiempo, cada mueble fue identificado de acuerdo al turno de producción en el cual fue fabricado. La información es presentada en la siguiente tabla de contingencia:

Tabla 2.1 Tabla de contingencia número de muebles con defectos Por tipo de defecto y turno de producción

Turno de producción	Tipo de defecto				
	A	B	C	D	Total
1	15	21	45	13	94
2	26	31	34	5	96
3	33	17	49	20	119
Total	74	69	128	38	309

2.1 Definición de asociación o dependencia para una tabla de contingencia de I×J.

De la Tabla 2.1, para el caso del espacio R^J -que es el espacio de los renglones- se refiere al turno en que fueron fabricados los muebles, de tal forma que para cada turno se tiene un vector con cuatro entradas es decir nos referimos al espacio R^4, para el caso del espacio R^J -que es el espacio de las columnas- se refiere al tipo de defecto que tienen los muebles,de tal manera que para cada tipo de defecto se tiene un vector con tres entradas es decir nos referimos al espacio R^3.

La probabilidad de que los elementos de una población pertenezcan a una categoría j del factor tipo de defecto, sea mayor o menor que en una categoría i del factor turno que en otra categoría i' del mismo factor se representa como:

$$\Pr(x \in j | x \in i) \neq \Pr(x \in j | x \in i') \qquad \text{para alguna } i \neq i'.$$

Si esta desigualdad se cumple para alguna categoría j del factor tipo de defecto, entonces existe una asociación entre los factores turno de producción y tipo de defecto.Esta ecuación indica que la probabilidad de que un mueble x pertenezca a la categoría j del factor tipo de defecto depende la categoría del factor turno que se considere.

Como una explicación hipotética, en el caso de que se presentara la asociación pudiera ser que si pensamos que los turnos de fabricación corresponden al matutino, vespertino y nocturno; puede darse el caso por ejemplo que para trabajadores que no estén muy habituados a trabajar en la noche su capacidad productiva se ve disminuida por efecto del sueño, de tal manera que se esperaría un mayor número de muebles defectuosos en el turno nocturno presentándose de esta manera la posible asociación.

Para hablar del término de independencia se tiene lo siguiente:

$$\text{Si } \Pr(x \in j | x \in i) = \Pr(x \in j | x \in i') = \Pr(x \in j) \qquad i=1,2,3 \text{ con } i \neq i'$$

Es decir, si estas igualdades se cumplen para toda categoría j del factor tipo de defecto $j=1,2,3,4$ entonces las variables no están asociadas.Esto quiere decir que la probabilidad de que un mueble x pertenezca a la categoría j del factor tipo de defecto es independiente de las categorías del factor turno de producción.

Existen métodos para probar la independencia entre las categorías de los factores, dentro de los cuales la prueba Ji-cuadrada es una de las más usadas.Aplicaremos esta prueba estadística a los datos de la tabla de contingencia 2.1

2.2 Prueba de independencia χ^2 para una tabla de contingencia de $I \times J$.

Datos: Una muestra aleatoria de tamaño T es obtenida. Estas observaciones son clasificadas de acuerdo a dos criterios (factores): Para el ejemplo de la tabla 2.1 tenemos que T=309.

k_{ij} representa la frecuencia de muebles en la categoría i del primer factor turno de producción y la categoría j del segundo factor tipo de defecto.

Supuestos:

1. La muestra T es aleatoria.

2. Cada observación es clasificada en exáctamente una de las I categorías del primer factor y exáctamente en una de las J categorías del segundo factor.

Hipótesis:

H_0 : La aparición del defecto j en los muebles producidos, es independiente del turno de producción i para todo $i = 1,2,3$ y $j = 1,2,3,4$.

Es decir

$H_0 : p_{ij} = p_{i\bullet}p_{\bullet j} \quad \forall$ i,j $\quad Vs \quad H_1 : p_{ij} \neq p_{i\bullet}p_{\bullet j}$ para alguna i,j

Estadístico de prueba:

Sea $E_{ij} = \frac{k_{i\bullet}k_{\bullet j}}{T}$ el valor esperado para la celda ij suponiendo cierta la hipótesis H_0 de independencia. El estadístico esta dado por: $\chi_c^2 = \sum_{i=1}^{I}\sum_{j=1}^{J}\frac{(k_{ij}-E_{ij})^2}{E_{ij}}$.

Para los datos de la tabla 2.1 χ_c^2 =19.177

Distribución nula del estadístico:

La distribución nula del estadístico está dada aproximadamente por la distribución Ji-cuadrada con $(I-1)(J-1)$ grados de libertad ($\chi_{(I-1)(J-1)}$). La aproximación a la Ji-cuadrada es satisfactoria si todas las E_{ij} son mayores o iguales que .5 y que al menos la mitad son más grandes que 1.

Para el ejemplo se tiene que $\chi_c^2 \sim \chi_{(2)(3)} = \chi_{(6)}$

Regla de decisión:

Rachazar H_0 si $\chi_c^2 > \chi^2_{((1-\alpha),(I-1)(J-1))}$, donde $\chi^2_{((1-\alpha),(I-1)(J-1))}$ es el cuantil de orden 1-α de una distribución χ^2 con $((I-1)(J-1))$ grados de libertad.

Por lo tanto para el ejemplo se tiene que para un nivel de significancia $\alpha = 0.05$:

$\chi^{2}_{(6)}$ =12.591, de tal manera que se rechaza la hipótesis H_0 ya que χ^2_e =19.177>$\chi^2_{(.95,6)}$ =12.591

Por lo que se concluye que existe asociación entre el turno de fabricación del mueble y el tipo de defecto que presenta.

Una pregunta que quizás interese a la persona encargada del control de calidad de la fábrica es el poder determinar cuales categorías de los factores son los que están asociados; ya que la prueba Ji-cuadrada únicamente establece si existe asociación o no, pero no especifica cuales son las categorías causantes de ésta.

Un camino para descubrir que factores causan la asociación entre los factores es haciendo uso de la expresión que nos proporcionaba la definición de asociación:

$$\Pr(x \in j | x \in i) \neq \Pr(x \in j | x \in i') \qquad \text{para alguna } i \neq i' \qquad (1)$$

la cual tiene dos posibilidades:

$$\Pr(x \in j | x \in i) > \Pr(x \in j | x \in i') \text{ para alguna } i \neq i' \qquad (2)$$

ó

$$\Pr(x \in j | x \in i) < \Pr(x \in j | x \in i') \text{ para alguna } i \neq i' \qquad (3)$$

De manera intuitiva observamos de 2 que es más probable que la categoría j del segundo factor ocurra con la categoría i del primer factor que con la categoría i' .

De manera análoga de 3 observamos que es más probable que la categoría j del segundo factor ocurra con la categoría i' del primer factor que con la categoría i .

Ahora bien, si lográramos representar estas posibles asociaciones en un gráfico que muestre las similitudes entre las categorías, esperaríamos para 2 que la proximidad de la categoría j a la categoría i, sea menor que la proximidad de la categoía j a la categoría i', dado que es más probable que lo primero ocurra.

Como se recordará la matriz perfil por renglones y por columnas definidas en el capítulo uno - tablas 2.2 y 2.3 -, contine las probabilidades condicionales

Tabla 2.2 Perfiles por turno (renglones)

Turno de producción	Tipo de defecto				Total
	A	B	C	D	
1	0.15957	0.2234	0.47872	0.1383	1
2	0.27083	0.3292	0.35417	0.0508	1
3	0.27731	0.1486	0.41176	0.1267	1

Tabla 2.3 Perfiles por tipo de defecto (columnas)

Turno de producción	Tipo de defecto			
	A	B	C	D
1	0.2027	0.30435	0.35156	0.34211
2	0.35135	0.44928	0.26563	0.13158
3	0.44595	0.24638	0.38281	0.52632
Total	1	1	1	1

Por ejemplo:

1) Pr(x∈ C | $x \in 3$) es la Pr(los muebles tengan el defecto C cuando se fabricaron en el tercer turno)= 0.41176

2) Pr(x∈ 3 | $x \in C$) es la Pr(los muebles se hayan fabricado en el tercer turno cuando presentan el defecto C)= 0.38281

3) Pr(x∈ A | $x \in 2$) = 0.27083

4) =Pr(x∈ 2 | $x \in A$) = 0.35135

Una interpretación gráfica de la información presentada en las tablas 2.2 y 2.3, es la siguiente:

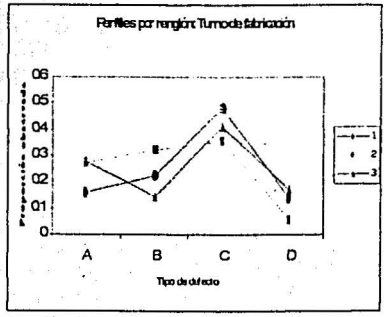


Fig 2.1

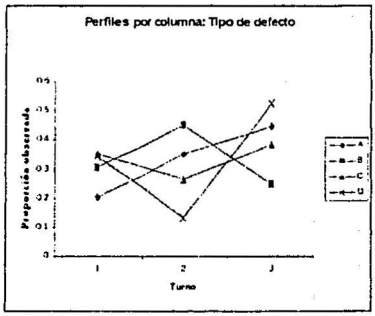


Fig 2.2

Así, en la figura 2.1 se tiene que el tipo de defecto A es más probable de ocurrir en las turnos 2 y 3 que en el primer turno. Para el defecto B se observa que en el segundo turno es más probable que este ocurra que en el tercer turno. Para los tipos de defectos C y D es casi igual de probable que ocurran en todos los turnos.

Si nos fijamos en la figura 2.2 vemos que en el turno 1 es más probable la ocurrencia de los defectos B,C y D que el defecto A.Para el turno 2, es más probable que ocurra el defecto B que el efecto D y para el turno 3 es más probable que ocurra el defecto D que el defecto B.

Estas proporciones observadas nos dan una idea acerca de cómo es la asociación; para conseguir una mejor apreciación de ésta se construyen las tablas 2.4 y 2.5 que representan las desviaciones de los perfiles renglón y columna respectivamente con respecto al centroide del espacio vectorial respectivo.

Tabla 2.4. Desviación de los perfiles renglón (Turno de producción) con respecto al centroide.

Turno de producción	Tipo de defecto A	B	C	D
1	-0.07991	0.0001	0.08448	0.015322
2	0.03135	0.09962	-0.09007	-0.07089
3	0.03783	-0.08044	-0.00247	0.04509

Tabla 2.5. Desviación de los perfiles columna (Tipo de defecto) con respecto al centroide.

Turno de producción	Tipo de defecto A	B	C	D
1	-0.1015	0.00014	0.04736	0.0379
2	0.04057	0.1386	-0.04505	-0.1791
3	0.06083	-0.13874	-0.0023	0.1412

Una presentación gráfica de la información contenida en las tablas 2.4 y 2.5 se presenta en las figuras 2.3 y 2.4.

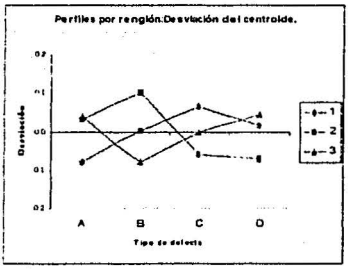


Fig 2.3

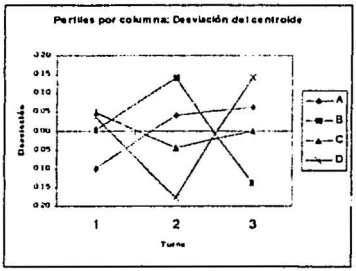


Fig 2.4

De la figura 2.3 interpretamos que el defecto A se encuentra más alejado en un sentido negativo del centroide para el turno 1 en el espacio de los renglones (R^4), lo cual nos habla de una asociación entre el turno 1 y el defecto A. Para el caso del defecto B, tenemos que el turno 2 y 3 se encuantran alejados del centroide en un sentido positivo y negativo respectivamente lo cual nos hace suponer una asociación entre el tipo de defecto B y los turnos 2 y 3. El defecto C por otra parte parece que tiene una asociación no tan marcada con respecto a los turnos 1 y 2 ,mientras que con el tercer turno no parece existir asociación. Por último, con respecto al defecto D se tiene que presenta una posible asociación con respecto al turno 2.

De la figura 2.4 se tiene que el turno 1 está más asociado con respecto al tipo de defecto A en un sentido negativo, mientras que existe una asociación en un sentido positivo entre el defecto B y el turno 2 y una asociación en un sentido negativo entre el defecto D y el turno 2. Se observa que, además existe una asociación en un sentido positivo entre el turno 3 y el defecto D y una asociación en un sentido negativo entre el turno 3 y el defecto B.

La forma de describir la asociación presentada arriba es un mejor acercamiento para conocerla, esto en el sentido de que únicamente se están describiendo los perfiles tal cual se obtienen de la tabla de contingencia, pero sin tomar en cuenta el peso que cada uno de ellos tiene en su respectivo

espacio, para lograr introducir este elemento en el análisis es necesario ponderar cada uno de los elementos. Recuérdese que precisamente en el capítulo I se hizo referencia a un espacio vectorial ponderado.

Así, el problema de descubrir la asociación de manera gráfica entre las categorías de los factores se puede tratar de la siguiente manera:

Para cada uno de los elementos que forman la matriz perfil por renglón y por columnas que son vectores en los espacios R^J y R^I respectivamente, se pretende encontrar una representación gráfica en un espacio de menor dimensión prefiriéndose una línea recta, enseguida un plano o bien un hiperplano, en el cual se tomen en cuenta la ponderación de los elementos y que presente las proximidades entre elementos en cada espacio y de manera simultanea se consideren los dos espacios para tener una representación óptima. **Categorías muy próximas entre si, con un alejamiento considerable del centroide implicarán una asociación fuerte entre ellas.**

Como se puede observar el problema arriba presentado equivale a encontrar vectores que generen el subespacio óptimo buscado.

2.3 Representación gráfica de la asociación entre factores con el AFC.

En el capítulo anterior se mencionó que los elementos de la tabla de contingencia ya sea por renglon o por columna podían ser vistos como vectores. Ahora bien, esos vectores pueden ser representados en un sistema de coordenadas, a través de unos vectores que forman una base para el espacio vectorial. Si pudieramos ver esa representación veríamos esos vectores como una nube de puntos.

Dentro de esa nube se esperaría que aquellos vectores parecidos estarán más próximos y aquellos vectores que sean distintos estáran más alejados. Por lo que el problema de encontrar un subespacio óptimo para esos vectores, será el encontrar un nuevo sistema de referencia -base- que conserve la estructura de la posición de los vectores y que además sea posible representarlos visualmente. Para encontrar esos vectores que forman la base que buscamos existen algunos métodos de ajuste como el de mínimos cuadrados.

En las siguientes secciones se muestra de manera gráfica la idea de este ajuste para el AFC y también se muestra la parte algebráica.

2.3.1. Formulación gráfica de ajuste de un subespacio vectorial para dos factores de la tabla de contingencia de I×J.

Se tienen dos nubes de puntos N(I) y N(J)

Un elemento de la nube N(I) es un par formado por un perfil renglón y su peso respectivo, es decir:

N(I)=$\{f_i, f_{i*}\}$, donde f_i representa al perfil renglón i-ésimo

Tabla 2.6. Nube de puntos N(I): Perfiles por renglón

Turno	A	B	C	D	Peso
1	0.160	0.223	0.479	0.138	0.304
2	0.271	0.323	0.354	0.052	0.311
3	0.277	0.143	0.412	0.168	0.385

Un elemento de la nube N(J) es un par formado por un perfil columna y su masa respectiva, es decir:

N(J)=$\{f_j, f_{*j}\}$, donde f_j representa al perfil columna j-ésimo.

Tabla 2.7.Nube de puntos N(J): perfiles por columna.

Turno	A	B	C	D
1	0.203	0.304	0.352	0.342
2	0.351	0.449	0.266	0.132
3	0.446	0.246	0.383	0.526
Peso	0.239	0.223	0.414	0.123

Gráficamente la nube de puntos N(J) se ve así:

Nube de puntos N(J): Perfiles por columna

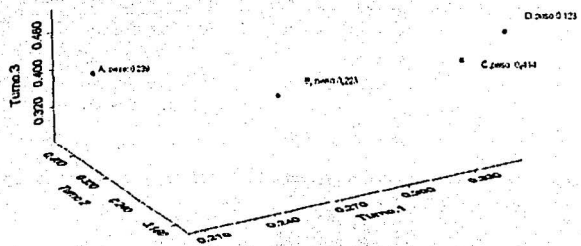


Fig. 2.5

Para estos dos conjuntos buscamos una representación que esté en un subespacio de la menor dimensión posible el cual puede ser R o R^2 y además que sea lo más confiable en el sentido de que se conserve la relación de asociación entre las categorías.

Para lograr esto se hará la proyección de los elementos de la nube de puntos en un espacio lineal S que puede ser una línea, un plano o un hiperplano, etc.; el cual contendrá al centro de gavedad G, si se consigue que la distancia de los puntos a su proyección sea mínima entonces el subespacio S será la aproximación a la representación deseada.

La razón por la que el centroide debe estar contenido en el subespacio, es porque éste es el punto más cercano a la nube de puntos en el sentido de minimiza la suma de cuadrados de las distancias entre los puntos de la nube y dicho punto. Por lo que la búsqueda de subespacios óptimos se restringirá a aquellos que contengan al centroide.

Para encontrar el subespacio de menor dimensión se tienen dos alternativas:

1. Se define una función llamada de cercanía entre la nube de puntos N(I) o N(J) con el subespacio S.

Por ejemplo para la nube de puntos N(I)

$\Psi(S, N(I)) = \sum_{i=1}^{I} f_{i.} d^2(f_i, S)$

donde

$d^2(f_i, S)$ es la distancia mínima al cuadrado entre el perfil f_i y el subespacio S.

$f_{i.}$ es el peso o masa del perfil i-ésimo.

Esta función establece que el subespacio S es el que mejor ajusta a los perfiles en términos de minimizar la distancia que existe entre estos y el subespacio, por lo que el problema se reduce a: minimizar la suma de las distancias elevadas al cuadrado ponderadas de los elementos de la nube de puntos N(I) al subespacio S.

2. La otra forma es maximizar la suma de distancias ponderadas del origen (el centroide) a las proyecciones de los puntos de la nube sobre el subespacio S.

Desde luego este subespacio S es generado por algunos vectores por lo que a fin de cuentas lo que importa es encontrar los vectores que lo generan.

Para presentar gráficamente esta idea y sin perder generalidad trabajemos con la nube de puntos por renglón N(I) y además supongamos que los elementos de la nube tienen pesos unitarios. Tomemos un perfil cualquiera f_i de la nube, y proyectémoslo de manera ortogonal sobre el subespacio S generado por el vector u -hasta ahora desconocido- denotemos a esta proyección como $Proy_u f_i$.

Observemos los segmentos $\overline{Gf_i}$, $\overline{G\,Proy_u f_i}$, $\overline{f_i\,Proy_u f_i}$. Véase figura 2.6 .

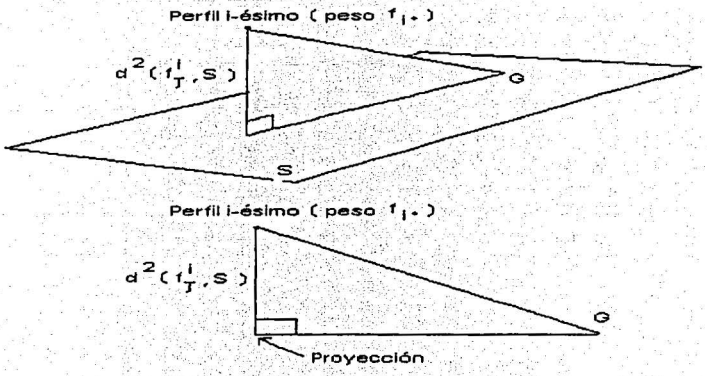


Fig 2.6

Lo que se quiere es minimizar la longitud o distancia del segmento $\overline{f_j \text{Pr}oy_u f_j}$, para todos los I elementos de la nube, de tal manera que la suma de las distancias de todos los puntos con respecto al subespacio generado por el vector u sea mínima. Tomemos la distancia al cuadrado $\overline{f_j \text{Pr}oy_u f_j}^2$. Aplicando el Teorema de Pitágoras tenemos que $\overline{f_j \text{Pr}oy_u f_j}^2 + \overline{G \text{Pr}oy_u f_j}^2 = \overline{Gf_j}^2$, sumando sobre los I elementos $\sum_{i=1}^{I} \overline{f_j \text{Pr}oy_u f_j}^2 + \sum_{i=1}^{I} \overline{G \text{Pr}oy_u f_j}^2 = \sum_{i=1}^{I} \overline{Gf_j}^2$;como el problema es minimizar $\sum_{i=1}^{I} \overline{f_j \text{Pr}oy_u f_j}^2$, esto implica que $\sum_{i=1}^{I} \overline{f_j \text{Pr}oy_u f_j}^2 = \sum_{i=1}^{I} \overline{Gf_j}^2 - \sum_{i=1}^{I} \overline{G \text{Pr}oy_u f_j}^2$.

Nótese que $\sum_{i=1}^{I} \overline{f_j \text{Pr}oy_u f_j}^2 = \sum_{i=1}^{I} f_{i\bullet} d^2(f_j, S)$ para $f_{i\bullet} = 1$ i=1,2,3...,I y para el subespacio generado por el vector u. Lo que implica que minimizar la suma de las distacias al cuadrado entre los perfiles y el subespacio generado por el vector u es equivalente a maximizar $\sum_{i=1}^{I} \overline{G \text{Pr}oy_u f_j}^2$, que es la suma de las proyecciones al cuadrado sobre el subespacio generado por el vector u (ver Fig. 2.6)

Para encontrar la solución al problema de encontrar el mejor subespacio que ajuste a los puntos algunos autores hacen uso de la técnica de descomposición en valores singulares (SVD) y otros lo plantean como un problema de optimización cuya solución corresponde a encontrar los eigenvectores y eigenvalores de la matriz que representa la suma ponderada de cuadrados de las proyecciones de los puntos de la nube. En este trabajo se presenta la segunda forma.

2.3.2 Presentación algebráica del análisis de correspondencias.

Algunos resultados importantes que se refieren al subespacio óptimo S son los siguientes:

1. El centroide de la nube de puntos N(I), es el punto que minimiza la función que describe la distancia entre la nube de puntos N(I) y el subespacio S.

2. El subespacio óptimo S debe contener al centroide.

Para garantizar que el centroide esté contenido en el subespacio S se lleva a cabo una traslación a un nuevo sistema de referencia en el cual el origen sea precisamente el centroide (el origen del sistema de referencia para el subespacio lineal). Para fines ilustrativos se trabajará en esta parte de la exposición con las matrices originales sin considerar por el momento esta transformación lineal, sin embargo al momento de considerar los ejemplos se tomará en cuenta.

Definición 1

Sea A una matriz simétrica, se dice que A es positiva definida si $A\,x > 0, \forall\, x \neq 0$.

Sea $X_{I\times J}$ una matriz de datos, que se refiere a una tabla de contingencia de I categorías para el primer factor y J categorías para el segundo factor, sea $M_{J\times J}$ una matriz simétrica, positiva definida; que define la distancia en R^J Sea $N_{I\times I}$ una matriz diagonal cuyos elementos son los pesos de los I puntos (si se considera como matriz de distancias a la matriz diagonal de pesos por renglón entonces la matriz $N_{I\times I}$ que representará la distancia para el espacio R^I) Véase cuadro 1 Capítulo I.

Sea u un vector unitario en el espacio R^J que satisfaga la ecuación $u^t\, M\, u = 1$. Las proyecciones de los I puntos sobre el eje que define el vector u son los elementos del vector $\hat{v}_{I\times 1}$ tal que $\hat{v} = XM\,u$. La suma de las distancias del centroide a la proyección ponderadas al cuadrado es $\hat{v}^t\, N\, \hat{v}$, a este término se le conoce como ajuste ya que es precisamente lo que corresponde a $\sum_{i=1}^{I} \overline{G\,\mathrm{Proy}_u f_i}^2$ y de acuerdo a lo expresado con anterioridad en la medida que esta cantidad sea maximizada ocurrirá que la distancia entre los perfiles y el subespacio generado por el vector u es menor.

En términos matriciales esto queda expresado de la siguiente manera:

$$\hat{v}^t\, N\, \hat{v} = u^t\, M X^t\, N X M\, u \qquad (4)$$

esta cantidad debe ser maximizada con la restricción:

$$u^t\, M\, u = 1 \qquad (5)$$

Es decir el problema es maximizar la suma del cuadrado de las proyecciones ponderadas sobre el subespacio generado por el vector u con la métrica M y además con la restricción de que el

vector u tenga norma 1 para la misma métrica definida por M.

De la expresión anterior observamos que lo único que desconocemos es el vector u, por lo que el problema ahora es su determinación. Para resolver este problema veamos la siguiente proposición:

Proposición 1

Sea A y M matrices simétricas, además M es positiva definida, un forma cuadrática para la matriz A es $u'Au$, maximizar esta forma cuadrática bajo la restricción $u'Mu = 1$. Se escribe:

$\max u'Au$

s.a $u'Mu = 1$

y u_a es el eigenvector de la matriz $M^{-1}A$ que corresponde al eigenvalor más grande λ_a.($a = 1, 2, ...,$ rango de A), el cual maximiza la forma cuadrática $u_a' A u_a$. Esta proposición además establece que el vector unitario u_a, para la distancia definida por M ($u_a' M u_a = 1$) es M-ortogonal a u_β(es decir $u_a' M u_\beta = 0$ para $\alpha > \beta$) †.

Utilizando la proposición anterior para (4), se tiene

$\max \hat{v}' N \hat{v} = u' M X' N X M u$

s.a

$u' M u = 1$

Sea $A = MX'NXM$

$\Rightarrow$ $\max \hat{v}' M \hat{v} = u' M X' N X M u = u' A u$

s.a

$u' M u = 1$

que es de la misma forma que la proporsición 1 además tenemos que

1) $A_{J\times J} = MX'NXM$ es simétrica.

Dem:

Sabemos que para dos matrices A y B $(AB)' = B'A'$. y que $(A')' = A$

$\Rightarrow A' = (MX'NXM)'$

$= (M(X'NXM))' = (X'(NXM))'M' = (N(XM))'XM' = (XM)'N'XM'$

$= M'X'N'XM'$

pero sabemos por hipótesis que M es simétrica y que N es una matriz diagonal

$\Rightarrow A^t = M X^t N X M$

$\therefore A = A^t$ es decir A es simétrica.†

2) Por hipótesis sabemos que la matriz de distancia definida por M es simétrica y positiva definida.

Así u_a es el eigenvector de la matriz $M^{-1}M X^t N X M = X^t N X M$ correspondiente al eigenvalor más grande $\lambda_a.(a = 1, 2, .., J)$

Al vector u_a se le conoce como el a eje principal. Este vector satisface la ecuación $X^t N X M u_a = \lambda_a u_a$.

El operador "proyección " sobre el eje u_a, definido como $\varphi_a = M u_a$, es llamado o se le conoce como "factor". Y satisface la ecuación:

$$M X^t N X \varphi_a = \lambda \varphi_a,$$

ya que $X^t N X M u_a = \lambda u_a$ y si premultiplicamos por M tenemos $M X^t N X (M u) = \lambda(M u)$ que da la relación anterior.

Observación: Los factores definidos como φ_a son α ($\alpha = 1,...,$ rango de A) para cada vector u_a definido por su respectivo eigenvalor. Además la norma de estos factores bajo la distancia definida por M^{-1} es uno:

$$\varphi_a^t M^{-1} \varphi_a = u_a^t M M^{-1} M u_a = u_a^t M u_a = 1$$

Al resolver el problema de maximización hemos encontrado los vectores u que generan al mejor espacio vectorial que ajusta a los perfiles.

En el análisis de correspondencias se muestran los perfiles por renglón y los perfiles por columna de una tabla de contingencia como puntos de una espacio vectorial de menor dimensión. Ambas representaciones pueden ser presentados en una sola gráfica que muestra los perfiles de ambos factores de manera simultánea.

Sin embargo, todavia tenemos un problema importante: ¿cuántos vectores de los u's que fueron encontrados son necesarios para generar el subespacio adecuado S y dar una buena representación de los perfiles en la gráfica?. Dejaremos por un momento la respuesta a esta pregunta, y obtengamos la representación de los perfiles para el ejemplo de la tabla 2.1. Los resultados fueron obtenidos mediante el uso del paquete mathematica V 4.0.

2.3.2.1 Análisis en R^J.Cálculo de factores .

Los elementos que definen a este espacio para los datos del ejemplo son:

1. I- perfiles por renglón, los cuales están centrados en el origen. Cada uno de estos vectores o puntos se encuentran en los renglones de la matriz R* referida a los renglones de la matriz perfil por renglón tomando su desviación con respecto al centroide, en el espacio R^J.

$R^* = R_{I*J} - 1r$

donde

$R = D_r^{-1}F$ es la matriz perfil por renglón

$$1^*r=\begin{pmatrix}1\\1\\1\\1\end{pmatrix}\begin{pmatrix}0.239 & 0.223 & 0.144 & 0.123\end{pmatrix}=\begin{pmatrix}0.239 & 0.223 & 0.144 & 0.123\\0.239 & 0.223 & 0.144 & 0.123\\0.239 & 0.223 & 0.144 & 0.123\\0.239 & 0.223 & 0.144 & 0.123\end{pmatrix}$$

así, los elementos de las celdas de la matriz R* para el ejemplo de la tabla 2.1 estan dados en la tabla 2.4.

Tabla 2.4 Perfiles Renglón: Desviación del centroide

	Tipo de defecto			
Turno	A	B	C	D
1	-0.07991	0.00010	0.05448	0.01532
2	0.03135	0.09952	-0.05007	-0.07089
3	0.03783	-0.08044	-0.00247	0.04509

Para fines del desarrollo en la explicación la matriz R* que se refiere a la matriz perfil por renglones considerando su desviación con respecto al centroide, será considerada conceptualmente como la matriz $D_r^{-1}F$ que es la forma matricial para calcular la matriz perfil por renglones.

2. La distancia definida entre dos perfiles de R* es

$$d^2(i,i') = \sum_{j=1}^{J} \frac{1}{f_{.j}}\left(\frac{f_{ij}}{f_{i.}} - \frac{f_{ij}}{f_{i.}}\right)^2$$

3. La distancia está caracterizada por la la matriz diagonal del recíproco de los pesos por columna.

$$D_c^{-1} = \begin{pmatrix} \frac{1}{0.239} & 0 & 0 & 0 \\ 0 & \frac{1}{0.223} & 0 & 0 \\ 0 & 0 & \frac{1}{0.414} & 0 \\ 0 & 0 & 0 & \frac{1}{0.123} \end{pmatrix}$$

4. Los pesos de los vectores para este espacio están dados por la la matriz diagonal de pesos por renglón

$$D_r = \begin{pmatrix} 0.304 & 0 & 0 \\ 0 & 0.311 & 0 \\ 0 & 0 & 0.385 \end{pmatrix}$$

5. La matriz transpuesta de correspondencias es:

$$F^t = \begin{pmatrix} 0.048 & 0.084 & 0.106 \\ 0.067 & 0.100 & 0.055 \\ 0.145 & 0.110 & 0.158 \\ 0.042 & 0.016 & 0.064 \end{pmatrix}$$

Así, de acuerdo al problema de encontrar un subespacio de menor dimensión que R^4 tenemos que encontrar los vectores u_a que generarán este nuevo subespacio; estos vectores son unitarios para la distancia definida en R^J la cual está dada por D_c^{-1} y además deberán satisfacer la siguiente restricción $u_a^t D_c^{-1} u_a = 1$. Es decir estos vectores serán ortonormales para la distancia definida.

Una vez encontrados esos vectores u_a podremos proyectar los perfiles de la matriz R^* en cada eje generado por el vector u_a conservando la distancia de cada perfil afectada por su peso. La proyección está dada por:

$R^* D_c^{-1} u_a = \hat{v}_{I\times 1}$., lo que se traduce como que la matriz perfil por renglones es proyectada a sobre el vetor u_a. lo cual da por resultado un vector en el cual cada entrada representa la proyección del elemento i-ésimo sobre el vector α.

La expresión anterior puede presentarse de manera matricial como $D_r^{-1} F D_c^{-1} u_a$.

Recordando que la cantidad que debe ser maximizada es la suma de cuadrados de las proyecciones de la nube de puntos al centroide (en la que se considera el peso de cada elemento dentro de la nube) esta suma de cuadrados está dada por la siguiente forma cuadrática $\hat{v}^t D_r \hat{v}$.

Estableciéndose el siguiente problema de maximización:

max $\hat{v}^t\, Dr\, \hat{v} = u_a^t\, D_c^{-1}\, F^t\, D_r^{-1}\, F\, D_c^{-1}\, u_a$

s a.

$u_a^t\, D_c^{-1}\, u_a = 1$.

Sea

$$A = D_c^{-1}\, F^t\, D_r^{-1}\, F\, D_c^{-1}$$

$$M = D_c^{-1}$$

Por lo tanto de acuerdo a la proposición 1 se tiene que: u_a es el eigenvector de la matriz

$M^{-1} A = (D_c^{-1})^{-1}\, D_c^{-1}\, F^t\, D_r^{-1}\, F\, D_c^{-1} = F^t\, D_r^{-1}\, F\, D_c^{-1}$ es decir u_a cumple con

$F^t\, D_r^{-1}\, F\, D_c^{-1}\, u_a = \lambda_a u_a$

Para el ejemplo tenemos que

S=$F^t\, D_r^{-1}\, F\, D_c^{-1}$

$$= \begin{bmatrix} 0.048 & 0.084 & 0.106 \\ 0.067 & 0.100 & 0.055 \\ 0.145 & 0.110 & 0.158 \\ 0.042 & 0.016 & 0.064 \end{bmatrix} \begin{bmatrix} -0.079 & 0.0001 & 0.044 & 0.015 \\ 0.031 & 0.0946 & -0.060 & -0.070 \\ 0.037 & -0.0804 & -0.002 & 0.045 \end{bmatrix} \begin{bmatrix} \frac{1}{0.239} & 0 & 0 & 0 \\ 0 & \frac{1}{0.223} & 0 & 0 \\ 0 & 0 & \frac{1}{0.414} & 0 \\ 0 & 0 & 0 & \frac{1}{0.123} \end{bmatrix}$$

$$= \begin{bmatrix} 0.0116 & -0.0009 & -0.0052 & -0.0033 \\ -0.0008 & 0.0249 & -0.0042 & -0.0291 \\ -0.0091 & -0.0079 & 0.0057 & 0.0128 \\ -0.0016 & -0.0160 & 0.0038 & 0.0196 \end{bmatrix}$$

Los eigenvalores de esta matriz determinan los vectores u_a que son conocidos como los α ejes principales. Al vector $\varphi_a = D_c^{-1}\, u_a$, se le conoce como primer factor, y corresponde al operador proyección. Además φ_a es un eigenvector de la matriz

$D_c^{-1}\, F^t\, D_r^{-1}\, F$, ya que si premultiplicamos (1) por D_c^{-1}, tenemos que $D_c^{-1}\, F^t\, D_r^{-1}\, F\, (D_c^{-1}\, u_a) = \lambda_a (D_c^{-1}\, u_a)$.

Las proyecciones de los I puntos sobre el eje principal u_a quedan determinados como:

$R\, D_c^{-1}\, u_a = D_r^{-1}\, F\, D_c^{-1}\, u_a = D_r^{-1}\, F\, \varphi_a$, ya que $\varphi_a = D_c^{-1}\, u_a$

En general la matriz S no es simétrica y por lo tanto no es posible aplicar el resultado de la

proposición 1. Para encontrar la solución a nuestro problema haremos lo siguiente:

Sea

$S=F^{-1} D_r F D_c^{-1}$ y llamemos

$\hat{A}=F' D_r^{-1} F$, la cual es simétrica.

Dem.

$$\hat{A}' = \left(F' D_r^{-1} F\right)' = \left(D_r^{-1} F\right)'(F')' = F' (D_r^{-1})' F = F' D_r^{-1} F = \hat{A} \dagger$$

Además la matriz D_c^{-1} es una matriz diagonal. Por lo anterior se puede escribir como $D_c^{-1} = D_c^{-\frac{1}{2}} D_c^{-\frac{1}{2}}$. Asi, S puede ser reescrita como: $S=\hat{A} D_c^{-\frac{1}{2}} D_c^{-\frac{1}{2}}$.

Como nuestro problema era encontrar los vectores que cumplieran con lo siguiente: $S\ u=\lambda u$, entonces tenemos: $\hat{A} D_c^{-\frac{1}{2}} D_c^{-\frac{1}{2}} cu=\lambda u$.

Si premultiplicamos en los dos lados de la igualdad por $D_c^{-\frac{1}{2}}$, se tiene

$$D_c^{-\frac{1}{2}} \hat{A} D_c^{-\frac{1}{2}} D_c^{-\frac{1}{2}} u=\lambda D_c^{-\frac{1}{2}} u. \qquad (6)$$

Haciendo $w=D_c^{-\frac{1}{2}} u$ en (6) tenemos lo siguiente: $D_c^{-\frac{1}{2}} \hat{A} D_c^{-\frac{1}{2}} w=\lambda w$.

La matriz $S_1 = D_c^{-\frac{1}{2}} \hat{A} D_c^{-\frac{1}{2}}$ es simétrica

Dem.

$$S_1' = \left(D_c^{-\frac{1}{2}} \hat{A} D_c^{-\frac{1}{2}}\right)' = \left(\hat{A} D_c^{-\frac{1}{2}}\right)'\left(D_c^{-\frac{1}{2}}\right)' = \left(D_c^{-\frac{1}{2}}\right)' \hat{A}' \left(D_c^{-\frac{1}{2}}\right)' = D_c^{-\frac{1}{2}} \hat{A} D_c^{-\frac{1}{2}} \dagger$$

además la matriz S_1 tiene los mismos eigenvalores λ que S.

La matriz S_1 para los datos del ejemplo de la tabla 2.1 es:

$$S_1 = \begin{pmatrix} 0.0116 & -0.0008 & -0.0069 & -0.0023 \\ -0.0008 & 0.0249 & -0.0058 & -0.0216 \\ -0.0069 & -0.0058 & 0.0057 & 0.0070 \\ -0.0023 & -0.0216 & 0.0070 & 0.0196 \end{pmatrix}$$

Los eigenvalores son:

$\lambda_1 = 0.0463$

$\lambda_2 = 0.0157$

$\lambda_3 = -2.72\times10^{-11}$

$\lambda_4 = 9.7911\times10^{-19}$

Los eigenvectores son:

$$W_1 = \begin{pmatrix} -0.0719, & -0.7200, & 0.2287, & 0.6511 \end{pmatrix}$$

$$W_2 = \begin{pmatrix} -0.8528, & 0.2447, & 0.4610, & 0.0144 \end{pmatrix}$$

$$W_3 = \begin{pmatrix} -0.4894, & -0.4725, & -0.6435, & -0.3505 \end{pmatrix}$$

$$W_4 = \begin{pmatrix} 0.1670, & -0.4453, & 0.5665, & -0.6729 \end{pmatrix}$$

Los eigenvectores de la matriz original se obtiene como: $\mathbf{y}=D_c^{\frac{1}{2}}\mathbf{w}$, y los factores son obtenidos de la siguiente manera:

$\varphi = D_c^{-1}\,\mathbf{y} = D_c^{-1}\mathbf{w}$

los cuales son:

La matriz de factores es:
$$\begin{pmatrix} -0.147087 & -1.74231 & -1.00004 & 0.341292 \\ -1.52358 & 0.517967 & -0.999958 & -0.942305 \\ 0.355445 & 0.716465 & -1.00007 & 0.9804 \\ 1.95722 & 0.0411931 & -0.999896 & -1.91915 \end{pmatrix}$$

Así, la proyección sobre los ejes principales son:

La matriz de proyecciones es:
$$\begin{pmatrix} 0.059701 & 0.186109 & 2.71061\times10^{-11} & 5.55112\times10^{-17} \\ -0.30402 & -0.048942 & 2.72575\times10^{-11} & 8.32667\times10^{-17} \\ 0.1986 & -0.107494 & 2.72267\times10^{-11} & -9.71445\times10^{-17} \end{pmatrix}$$

2.3.2.2 Análisis en R^J. Cálculo de factores.

Se hará el mismo análisis para encontrar los factores y las proyecciones para el espacio R^J cuyos elementos se encuentran en la matriz perfil por columnas.

Los elementos que definen a este espacio son:

1. J-perfiles por columna, los cuales están centrados en el origen. Cada uno de estos vectores o puntos se encuentran en los renglones de la matriz C^* referida a los renglones de la matriz perfil por columnas tomando su desviación con respecto al centroide en el espacio R^J.

Tabla 2.3.Perfiles por columna

Turno	Tipo de defecto A	B	C	D
1	0.2027	0.30435	0.35156	0.34211
2	0.35135	0.44928	0.26563	0.13158
3	0.44595	0.24638	0.38281	0.52632
Total	1	1	1	1

$C^* = C_{I\times J} - c1$

donde $C_{I\times J} = F_{I\times J}\, D_c^{-1}{}_{J\times J}$ es la matriz perfil por columnas.

$$c^*1 = \begin{pmatrix} 0.304 \\ 0.311 \\ 0.385 \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 & 1 \end{pmatrix} = \begin{pmatrix} 0.304 & 0.304 & 0.304 & 0.304 \\ 0.311 & 0.311 & 0.311 & 0.311 \\ 0.385 & 0.385 & 0.385 & 0.385 \end{pmatrix}$$

De manera similar a lo establecido en la sección anterior, la matriz C* será referida conceptualmente como la matriz perfil por columnas considerdando su desviación con respecto al centroide: $F\, D_c^{-1}$.Otra forma equivalente de trabajar esta matriz es consedarando su transpuesta $D_c^{-1}\, F^T$, y esto es precisamente lo que se hará en esta sección.de tal manera que C* esta dada por la tabla 2.5.

Tabla 2.5.Perfiles columna: Desviación del centroide

Turno	Tipo de defecto A	B	C	D
1	-0.10150	0.00014	0.04736	0.03790
2	0.04057	0.13860	-0.04505	-0.17910
3	0.06083	-0.13874	-0.00230	0.14120

2. La distancia definida entre dos puntos o columnas es:

$$d^2(j,j') = \sum_{i=1}^{I} \frac{1}{f_{i\cdot}} \left(\frac{f_{ij}}{f_{\cdot j}} - \frac{f_{ij'}}{f_{\cdot j'}} \right)^2$$

3. La distancia está caracterizada por la forma cuadrática D_r^{-1}: la matriz diagonal de pesos por renglón.

$$D_r^{-1} = \begin{pmatrix} \frac{1}{0.304} & 0 & 0 \\ 0 & \frac{1}{0.311} & 0 \\ 0 & 0 & \frac{1}{0.385} \end{pmatrix}$$

4. La matriz F es

$$F = \begin{pmatrix} 0.0485 & 0.0679 & 0.14563 & 0.0420 \\ 0.0841 & 0.1003 & 0.1100 & 0.0161 \\ 0.1067 & 0.0550 & 0.1585 & 0.0647 \end{pmatrix}$$

De manera análoga al análisis en R^J, queremos encontrar vectores v_a ortonormales de tal manera que para la distancia definida por la matriz D_r^{-1} se tenga $v_a' D_r^{-1} v_a = 1$.

Así, el vector de las "J" proyecciones sobre el eje definido por v es $\tilde{w}$ tal que

$\tilde{w} = D_c^{-1} F' D_r^{-1} v$, donde el factor de proyección es $\psi = D_r^{-1} v$

De acuerdo a lo establecido en la interpretación gráfica queremos maximizar la suma ponderada de cuadrados de las proyecciones tomando en cuenta los pesos de cada uno de los elementos, los cuales se encuentran en la matriz diagonal D_c.

Así el problema queda representado como:

Max $\tilde{w}' D_c \tilde{w} = v' D_r^{-1} F D_c^{-1} D_c D_c^{-1} F' D_r^{-1} v = v' D_r^{-1} F D_c^{-1} F' D_r^{-1} v$

s.a

$v_a' D_r^{-1} v_a = 1$

Sea $A = D_r^{-1} F D_c^{-1} F' D_r^{-1}$

$M = D_r^{-1}$

de manera semejante que al análisis en R^J, por la proposición 1 tenemos que:

v_a es el eigen vector de la matriz $(D_r^{-1})^{-1} D_r^{-1} F D_c^{-1} F' D_r^{-1} = F D_c^{-1} F' D_r^{-1}$ de manera que se satisface la ecuación:

$$F D_c^{-1} F' D_r^{-1} = \lambda_a v_a \qquad (7)$$

Para los datos del ejemplo tenemos:

$S^* = F D_c^{-1} F' D_r^{-1}$

$$= \begin{pmatrix} 0.0485 & 0.0679 & 0.14561 & 0.0420 \\ 0.6841 & 0.1003 & 0.1100 & 0.0161 \\ 0.1067 & 0.0350 & 0.1595 & 0.0647 \end{pmatrix} \begin{pmatrix} -0.1015 & 0.04067 & 0.06083 \\ 0.00014 & 0.1386 & -0.13874 \\ 0.04736 & -0.04505 & -0.0023 \\ 0.0379 & -0.1791 & 0.1412 \end{pmatrix} \begin{pmatrix} \frac{1}{0.304} & 0 & 0 \\ 0 & \frac{1}{0.311} & 0 \\ 0 & 0 & \frac{1}{0.385} \end{pmatrix}$$

$$= \begin{pmatrix} 0.0119 & -0.0090 & -0.0021 \\ -0.0086 & 0.0301 & -0.0174 \\ -0.0025 & -0.0221 & 0.0199 \end{pmatrix}$$

Los eigenvalores de esta matriz determinan a los vectores v_a que es conocido como el primer α eje principal, al operador $\psi_a = D_r^{-1} v_a$ se le conoce como el factor α-ésimo. Además ψ_a satisface la ecuación $D_r^{-1} F D_c^{-1} F' \psi_a = \lambda_a \psi_a$, es decir ψ_a (premultiplicando (7) por D_r^{-1}) es un eigenvector de la matriz $D_r^{-1} F D_c^{-1} F'$.

Las coordenadas de los "J"puntos sobre el eje generado por v_a son: $D_c^{-1} F' D_r^{-1} v = D_c^{-1} F' \psi_a$. Donde $\psi_a = D_r^{-1} v_a$

Infortunadamente, de manera general la matriz S' no es simétrica por lo que no se puede aplicar la proposición 1. Para poder resolver el problema, haremos lo siguiente:

Sea

$S' = F D_c^{-1} F' D_r^{-1}$

La matriz $\hat{A}' = F D_c^{-1} F'$ es simétrica y además $D_r^{-1} = D_r^{-\frac{1}{2}} D_r^{-\frac{1}{2}}$. De tal manera que $S' = \hat{A}' D_r^{-\frac{1}{2}} D_r^{-\frac{1}{2}}$. Si premultiplicamos por $D_r^{-\frac{1}{2}}$ y hacemos $w = D_r^{-\frac{1}{2}} v$, entonces el problema original tiene la siguiente forma:

Encontrar los eigenvalores λ y eigenvectores w de:

$D_r^{-\frac{1}{2}} \hat{A}' D_r^{-\frac{1}{2}} w = \lambda w$

La matriz $S_2 = D_r^{-\frac{1}{2}} \hat{A}' D_r^{-\frac{1}{2}}$ es simétrica y tiene los mismos eigenvalores de S'. Con esta nueva matriz los factores se calculan como: $\psi_a = D_r^{-\frac{1}{2}} w$.

$$S_2 = \begin{pmatrix} 0.0119 & -0.0091 & -0.0024 \\ -0.0085 & 0.0301 & -0.0194 \\ -0.0023 & -0.0199 & 0.0199 \end{pmatrix}$$

cuyos eigenvalores son:

$\lambda_1 = 0.0463$

$\lambda_2 = 0.0157$

$\lambda_3 = 1.2845 \times 10^{-6}$

asociados a estos eigen_valores, sus eigenvectores son:

w1=(0.1688, −0.7933, 0.5849)

w2=(0.8298, −0.2060, −0.5185)

w3=(−0.5516, −0.5576, −0.6202)

ψ_a se le conoce como el factor correspondiente al eigenvalor λ_a.

Así, la matriz de factores es:

$$\text{La matriz de factores es:} \begin{pmatrix} 0.306265 & 1.50514 & -1.00052 \\ -1.42254 & -0.369434 & -1.00002 \\ 0.942701 & -0.835667 & -0.999574 \end{pmatrix}$$

Y, las proyecciones de los perfiles sobre los ejes principlaes

$$\text{La matriz de proyecciones es:} \begin{pmatrix} -0.0309726 & -0.218319 & -2.41289 \cdot 10^{-7} \\ -0.327257 & 0.0652777 & -4.17077 \cdot 10^{-7} \\ 0.0770498 & 0.0501763 & 1.40693 \cdot 10^{-6} \\ 0.400143 & 0.00554634 & -7.14046 \cdot 10^{-6} \end{pmatrix}$$

2.3.2.3 Relación del análisis de R^I con el análisis en R^J : fórmulas de transición.

En R^J tenemos que $F^t D_r^{-1} F D_c^{-1} u_a = \lambda_a u_a$. (8)

es decir u_a es un eigenvector de la matriz $F^t D_r^{-1} F D_c^{-1}$ correspondiente al eigenvalor λ_a .

Si premultiplicamos (8) por $F D_c^{-1}$ tenemos que

$(F D_c^{-1} F^t D_r^{-1}) F D_c^{-1} u_a = \lambda_a (F D_c^{-1} u_a)$, lo que significa que $F D_c^{-1} u_a$ es un eigenvector de la matriz $F D_c^{-1} F^t D_r^{-1}$ para el mismo eigenvalor λ_a en R^J lo cual implica que v_a en R^I es proporcional a $F D_c^{-1} u_a$ en R^J.

Dándose las fórmulas de transición:

$$v_a = \frac{1}{\sqrt{\lambda_a}} F D_c^{-1} u_a$$

$$u_a = \frac{1}{\sqrt{\lambda_a}} F^t D_r^{-1} v_a$$

Nótese que

$$\psi_a = D_r^{-1} v_a = D_r^{-1} \frac{1}{\sqrt{\lambda_a}} F D_c^{-1} u_a = \frac{1}{\sqrt{\lambda_a}} D_r^{-1} F D_c^{-1} u_a = \frac{1}{\sqrt{\lambda_a}} D_r^{-1} F \varphi_a$$

$$\varphi_a = D_c^{-1} u_a = D_c^{-1} \frac{1}{\sqrt{\lambda_a}} F^T D_r^{-1} v_a = \frac{1}{\sqrt{\lambda_a}} D_c^{-1} F^T D_r^{-1} v_a = \frac{1}{\sqrt{\lambda_a}} D_c^{-1} F^T \psi_a$$

Así, por ejemplo una vez hecho el análisis para R^J y si queremos la representación de los datos en R^I se tiene:

$$\psi_1 = \frac{1}{\sqrt{0.01463}} \begin{pmatrix} \frac{1}{0.304} & 0 & 0 \\ 0 & \frac{1}{0.311} & 0 \\ 0 & 0 & \frac{1}{0.385} \end{pmatrix} \begin{pmatrix} 0.0485 & 0.0679 & 0.11563 & 0.0420 \\ 0.0841 & 0.1003 & 0.1100 & 0.0161 \\ 0.1067 & 0.0550 & 0.1585 & 0.0647 \end{pmatrix} \begin{pmatrix} -0.1470 \\ -1.5235 \\ 0.3554 \\ 1.8572 \end{pmatrix}$$

$$= \begin{pmatrix} 0.2926 \\ -1.4378 \\ 0.9287 \end{pmatrix}$$

$$\psi_2 = \frac{1}{\sqrt{0.01374}} \begin{pmatrix} \frac{1}{0.304} & 0 & 0 \\ 0 & \frac{1}{0.311} & 0 \\ 0 & 0 & \frac{1}{0.385} \end{pmatrix} \begin{pmatrix} 0.0485 & 0.0679 & 0.11563 & 0.0420 \\ 0.0841 & 0.1003 & 0.1100 & 0.0161 \\ 0.1067 & 0.0550 & 0.1585 & 0.0647 \end{pmatrix} \begin{pmatrix} -1.7423 \\ 0.5179 \\ 0.7164 \\ 0.0411 \end{pmatrix}$$

$$= \begin{pmatrix} 1.4832 \\ -0.3903 \\ -0.8566 \end{pmatrix}$$

Las relaciones anteriores muestran que las coordenadas de los puntos sobre un eje principal en el espacio R^I son proporcionales a las coordenadas de los puntos en el otro espacio R^J correspondientes al mismo eigenvalor .

Las relaciones anteriores pueden ser escritas en forma explícita es decir:

$$\psi_{ai} = \frac{1}{\sqrt{\lambda_a}} \sum_{j=1}^{J} \frac{f_{ij}}{f_{i\bullet}} \varphi_{aj}$$

..... (9)

$$\varphi_{aj} = \frac{1}{\sqrt{\lambda_a}} \sum_{i=1}^{I} \frac{f_{ij}}{f_{\bullet j}} \psi_{ai}$$

Estas fórmulas de transición son las que permiten representar graficamente de manera simultánea los perfiles de los dos espacios, sobre un subespacio (plano) generado por los dos primeros ejes principales. Y son las fórmulas que nos permitirán en el próximo capítulo establecer

las relaciones entre varios factores.

En la Fig 2.7 se muestran los perfiles para tipo de defecto y turno, representados en un subespacio de dimensión dos.

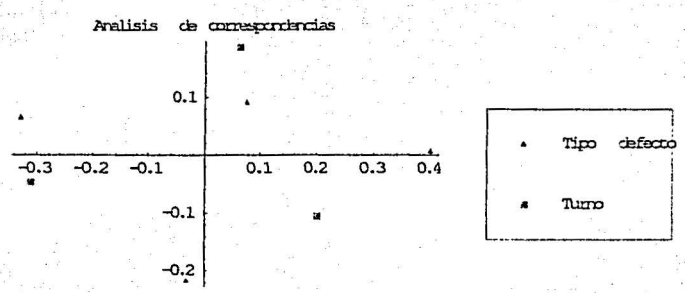


Fig. 2.7 Perfiles por tipo de defecto y turno en un espacio de dos dimensiones.

Sin embargo, existen algunas preguntas que se tienen que hacer:

1) ¿Porqué nada más se eligieron dos ejes?

2) ¿Qué calidad de representación tienen los perfiles en esta gráfica?

3) ¿Puede hacerse una interpretación en terminos de proximidad de manera simultánea entre los perfiles por renglón y los perfiles por columna?

Estas preguntas serán respondidad en las siguientes secciones.

2.4 Calidad del análisis: contribuciones.

En el análisis de correspondencias, cuando se obtienen las proyecciones sobre los nuevos ejes de referencia o ejes factoriales se presenta una gráfica de los perfiles, generalmente en un espacio de menor dimensión (dos o tres a lo mucho).

Para poder interpretar está gráfica de una manera adecuada es necesario conocer con que certeza se están describiendo las relaciones entre las categorías analizadas, es decir, que tan buena es la representación de los perfiles en la gráfica.

Algunos elementos que nos ayudarán a conocer la calidad de la representación se presentan a continuación.

2.4.1 Inercia: descomposición sobre los ejes factoriales ¿Qué significado geométrico y estadístico tiene la inercia?

Regresando a las preguntas finales de la sección 2.3.2.3

Se define la inercia de un punto i con masa $f_{i\bullet}$ con respecto al centroide G como la distancia al cuadrado del punto i al centroide G ponderada por su masa, es decir :

$$I_G(i, f_{i\bullet}) = f_{i\bullet} d^2(G, i).$$

Y, se define la inercia de una nube de puntos N(I) o N(J) como:

$$I_G N(I) = \sum_{i=1}^{I} f_{i\bullet} d^2(G, i)$$

$$I_G N(J) = \sum_{j=1}^{J} f_{\bullet j} d^2(G, j)$$

Otra manera de representar la inercia del punto j-ésimo es:

$$I(j) = f_{\bullet j}(\mathrm{Proy}_1^2 M_1 + \mathrm{Proy}_2^2 M_2 + \ldots + \mathrm{Proy}_n^2 M_J)$$

La inercia de la nube en la dirección del eje α es el eigenvalor λ_α correspondiente a este eje. La inercia es la suma extendida a varios puntos de la nube, de las cantidades:

$$\lambda_\alpha = \sum_{i=1}^{I} f_{i\bullet} \mathrm{Proy}_\alpha^2(i)$$

de tal manera que λ_α aparece como la media ponderada sobre los elementos de la nube N(I), con su sistema asociado de masas $f_{i\bullet}$ $i = 1,..I,$ de Las proyecciones al cuadrado generadas por el α -ésimo factor.

Regresando a la idea que se desarrolló para encontrar, para la nube de puntos N(I) o N(J) el subespacio de menor dimensión que ajustara de manera óptima a los puntos, se obuvo un subespacio de dos dimensiones (plano), el cual generado por dos vectores que son ortogonales y que generan a los elementos que son proyectados (perfiles), es decir se puede descomponer la inercia en dos distancias debido al teroema de pitágoras, por lo que tenemos lo siguiente:

La inercia total = Inercia correspondiente al eje 1 + Inercia correspondiente al eje 2.

Y, demás la inercia correspondiente al primer eje factorial está registrada con el primer eigenvalor λ_1, y la inercia correspondiente al segundo eje está registrada por el eigenvalor λ_2. De manera adicional se cumple la siguiente relación:

$$\lambda_1 \leq \lambda_2 \leq \lambda_3 \leq \lambda_4.$$

Como un comentario adicional, se tiene que el número de ejes factoriales de la nube de puntos N(I) no puede exceder al $\min\{(J-1),(I-1)\}$.

Para el centroide G de la nube de puntos N(I) siendo el origen del sistema de ejes coordenadas se tiene que todos sus componentes son ceros. En lenguaje probabilistico, la media ponderada del cuadrado de una función centrada en cero es llamada varianza. Ahora como $Proy_a$ es de media cero sobre I con el sistema de masas $f_{i.}$ asociado para cada perfil, la media ponderada por los masas o pesos $f_{i.}$ del cuadrado de la proyección $Proy_a^2$ es su varianza, así, diremos que el factor α tiene λ_α por varianza, de tal manera que el factor 1 tiene λ_1 de varianza, el segundo factor tiene λ_2 y así sucesivamente. Un aspecto importante para mencionar es el hecho de que en análisis de correspondencias λ_1 no puede ser mayor que 1.

2.4.2- Inercia relativa y contribución absoluta.

1) ¿Porqué nada más se eligieron dos ejes?

La inercia relativa y la contribución absoluta son algunas medidas que ayudan a saber que tan buena es la representación de los datos por mediante el análisis.

La inercia relativa es la proporción de la inercia total correspondiente a cada eje factorial. Es decir, que porcentaje de la inercia está siendo representado por cada eje factorial.

$$\tau_\alpha = \frac{\lambda_\alpha}{\sum_{i=1}^{r} \lambda_i} = \frac{\lambda_\alpha}{Traza}$$

En nuestro ejemplo tenemos que

$$\tau_1 = \frac{\lambda 1}{\sum_{i=1}^{2} \lambda_i} = \frac{\lambda 1}{Traza} = \frac{0.0463}{0.0463+0.0157} = 74.6\%$$ de la inercia es explicada por el primer eje.

$$\tau_2 = \frac{\lambda 2}{\sum_{i=1}^{2} \lambda_i} = \frac{\lambda 2}{Traza} = \frac{0.0157}{0.0463+0.0157} = 25.3\%$$ de la inercia es explicada por el segundo eje.

Así, se eligieron nada más los dos primeros ejes para los resultados del ejemplo de la tabla 2.1 porque entre ambos recogen el 100% de la inercia.

Para interpretar la información que proporcionan los ejes principales calcularemos las contribuciones absolutas para cada elemento j de la nube de puntos N(J) que representa los defectos de los muebles.

Estas contribuciones se representan por (($ca_\alpha(j)$.)) e indican la proporción de la varianza explicada por cada perfil del factor tipo de defecto en relación a cada eje princial.

Sabemos que $u_a' D_c^{-1} u_a = 1$, para la distancia definida en R^J. Además sabemos que $\varphi_a = D_c^{-1} u_a$. lo cual implica que $u_a = D_c \varphi_a$, dado que D_c^{-1} es positiva definida. Así, tenemos que $u_a' D_c^{-1} u_a = (D_c \varphi_a)' D_c^{-1} D_c \varphi_a = 1$

$= \varphi_a' D_c' D_c^{-1} D_c \varphi_a = \varphi_a' D_c D_c^{-1} D_c \varphi_a = 1$ (debido a que D_c^{-1} es simétrica)

$=\varphi_a^t D_c \varphi_a = 1$,

lo cual de manera explícita queda expresado como: $\sum_{j=1}^{J} f_j \varphi_{aj}^2 = 1$

Recordemos que la proyección del elemento j-ésimo en R^I es

$\varphi_{aj} = \frac{1}{\sqrt{\lambda_a}} \sum_{i=1}^{I} \frac{f_{ij}}{f_j} \psi_{ai}$, desarrollando tenemos que:

$\sqrt{\lambda_a}\, \varphi_{aj} = \sum_{i=1}^{I} \frac{f_{ij}}{f_j} \psi_{ai}$, sea $\hat{\varphi}_{aj} = \sqrt{\lambda_a}\, \varphi_{aj}$,

por lo tanto tenemos que: $\hat{\varphi}_{aj} = \sum_{i=1}^{I} \frac{f_{ij}}{f_j} \psi_{ai}$.

Así, la varianza de el conjunto de puntos proyectado sobre el eje α con respecto a G es:

$\sum_{j=1}^{J} f_j \hat{\varphi}_{aj}^2 = \sum_{j=1}^{J} f_j \left(\sqrt{\lambda_a}\, \hat{\varphi}_{aj}\right)^2 = \sum_{j=1}^{J} f_j \lambda_a \hat{\varphi}_{aj}^2 = \lambda_a$

donde f_j es la masa asignada a la j-ésima columna y $\hat{\varphi}_{aj}^2$ representa la distancia de la proyección del perfil j-ésimo con respecto al centroide definido por las columnas.

Dado que la varianza está representada por el eigenvalor λ_a , pero además este valor representa la inercia del conjunto de puntos en R^I con respecto al centro de gravedad. Es decir $\lambda_1 + \lambda_2 + \lambda_3 + ... + \lambda_I$ es igual a la inercia de todos los puntos con respecto a los ejes principales y además esta suma representa la varianza que cada uno de los ejes recoge.

Además $\lambda_1 + \lambda_2 + \lambda_3 + ... + \lambda_I = \sum_{i=1}^{I} \lambda_i$, que corresponde a la inercia total de ambas nubes es igual al coeficiente de contingencia en media cuadrática $\phi^2 = \frac{\chi^2}{n}$ la cual en principio no es acotada y su interpretación para valores cercanos a cero se puede afirmar que los factores no están asociadas entre sí.

El cociente $\frac{\lambda_a \varphi_{aj}^2 f_j}{\lambda_a} = f_j \varphi_{aj}^2 = ca_a(j)$ Representa la contribución absoluta del elemento j-ésimo al eje principal α.

J	*Masa*	*Contribución* primer eje	*Contribución* Segundo eje
A	0.2394	0.00022	0.01141
B	0.2233	0.02391	0.00095
C	0.4142	0.00245	0.00336
D	0.12297	0.01969	3.78×10^{-6}

Observación

$\sum_{j=1}^{J} ca_a(j)=\lambda_a$.

Si ahora, dividimos la contribución de cada perfil entre su respectivo eigenvalor tendremos la siguiente tabla

J	*Masa*	*Contribución %* primer eje	*Contribución %* Segundo eje
A	0.2394	0.00496	0.7250
B	0.2233	0.51648	0.00095
C	0.4142	0.05311	0.21395
D	0.12297	0.4252	0.00024

Que representa la contribución en términos de porcentaje de cada perfil tipo de defecto a la inercia que recoge el eje respectivo. En este ejemplo vemos que para el primer eje la categoría Tipo de defecto B y D son las que aportan una mayor inercia; mientras que para el segundo eje las categorías que mayor cantidad de inercia aportan son el Tipo de defecto A y C.

$\sum_{j=1}^{J} \frac{ca_a(j)}{\lambda_a}=1$

De manera análoga en el espacio de los renglones o R^J tenemos que

$f_{i*}\psi_{ai}^2 = ca_a(i)$

I	Masa	Contribución primer eje	Contribución Segundo eje
1	0.304	0.00120	0.01052
2	0.311	0.02977	0.00074
3	0.385	0.01537	0.00444

donde $\sum_{i=1}^{I} ca_a(i)=\lambda_a$

Si ahora, dividimos la contribucón de cada perfil entre su respectivo eigen_valor tendremos la siguiente tabla

I	*Masa*	*Contribución %* primer eje	*Contribución %* Segundo eje
1	0.304	0.02603	0.66879
2	0.311	0.64297	0.04739
3	0.385	0.33212	0.28256

Que representa la contribución en términos de porcentaje de la contribución de cada perfil Turno a la inercia que recoge el eje respectivo. En este ejemplo vemos que para el primer eje la categoría "segundo" y "tercer" turno son las que aportan una mayor inercia; mientras que para el segundo eje se tienen las categorías con mayor cantidad de inercia son el "primer" y "tercer" turno.

2.4.3 Correlaciones al cuadrado: cosenos cuadrados.($Cor^2(j)$)

2) ¿Qué calidad de representación tienen los perfiles en esta gráfica?

Explica la parte de la varianza de una variable explicada por un eje principal.

Para entender este punto supongamos que estamos en el espacio definido por las columnas, es decir, estamos en R^I. Como estamos trabajando en un espacio ponderado hemos visto que la distancia entre dos puntos se calcula de la siguiente manera

$$d^2(j,j') = \sum_{i=1}^{I} \frac{1}{f_{i\cdot}} \left(\frac{f_{ij}}{f_{\cdot j}} - \frac{f_{ij'}}{f_{\cdot j'}} \right)^2$$

Pero veamos que pasa con un elemento especial del espacio R^I: el centroide. Por lo que para calcular la distancia entre un columna j y el centroide tendremos

$d^2(j,G) = \sum_{i=1}^{I} \frac{1}{f_{i\cdot}} \left(\frac{f_{ij}}{f_{\cdot j}} - f_{i\cdot} \right)^2$, dado que el centroide se puede calular como: $G_j = \sum_{i=1}^{I} f_{i\cdot} \frac{f_{ij}}{f_{\cdot j}} = \sum_{i=1}^{I} f_{ij} = f_{i\cdot}$

El cuadrado de la proyección de la variable j sobre el eje α es igual a:

$$d_\alpha^2(j,G) = \left(\sqrt{\lambda_\alpha}\, \varphi_{\alpha j} \right)^2$$

Observación

$$\sum_a d_a^2(j,G) = d^2(j,G)$$

Así, tenemos que $\cos^2\omega = \frac{d_a^2(j,G)}{d^2(j,G)} = Cor(j)$

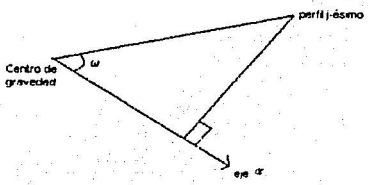


Fig 2.8 $\cos^2\omega = \frac{d_a^2(j,G)}{d^2(j,G)} = Cor(j)$

En particular para el ejemplo tenemos para las columnas:

J	Dist.1erEje	Dist.2doEje	Dist.Total	cos²1er eje	cos²2do eje	C.R*.
A	0.00095	0.04766	0.04862	0.01972	0.98027	0.9999
B	0.10709	0.00426	0.11135	0.96173	0.03826	1
C	0.00593	0.00813	0.01406	0.42195	0.57801	1
D	0.16011	0.00003	0.16014	0.9998	0.00019	0.9999

*calidad de representación

Ahora, para los renglones:

I	Dist.1erEje	Dist.2doEje	Dist.Total	cos²1er eje	cos²2do eje	C.R*.
1	0.00396	0.0316	0.03560	0.10272	0.8972	1
2	0.09572	0.00239	0.09812	0.96173	0.0244	1
3	0.03994	0.01155	0.05149	0.77562	0.2243	1

*calidad de representación

2.4.4 Calidad de representación de un perfil.

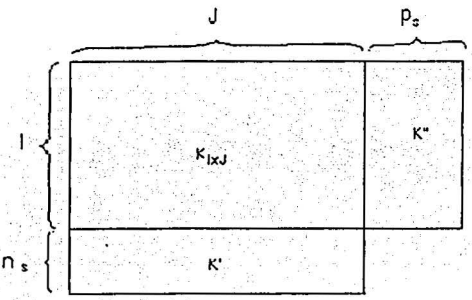
Explica la varianza explicada por los primeros α ejes principales

$QLT=Cor(1)+Cor(2)+\ldots+Cor(\alpha)$

2.5. Tratamiento de los elementos suplementarios en la tabla de contingencia.

Para fines prácticos, dentro del análisis de correspondencias existe la posibilidad representar categorías suplementarias también denominadas ilustrativas, las cuales son otras categorías de la tabla de contingencia que se quieran añadir para proyectarlas en el subespacio previamente encontrado por el análisis de correspondencias y sin que la inserción de tales categorías interfiera en el resultado de la determinación del subespacio óptimo. De este modo se pueden crear varios escenarios, los cuales representan posibles hipótesis de investigación .Lo que se espera es que el nuevo elemento quede situado en la proximidad de los elementos que se le parecen.

Retornemos a la matriz de datos original $K_{I\times J}$, la cual quede ser incrementada por p_s categorías del factor que define las columnas o por n_s categorías del factor que define los renglones.



Estamos interesados en proyectar los perfiles de los p_s nuevos puntos con respecto a los p puntos analizados en el espacio R^I. Sea k''_{ij} la i-ésima coordenada de la j-ésima columna suplementaria. El perfil de esta columna suplementaria es el vector cuyo i-esimo componente es $\frac{k''_{ij}}{k''_{.j}}$ es decir

$$j'' = \begin{bmatrix} \frac{k''_{1j}}{k''_{.j}} \\ \frac{k''_{2j}}{k''_{.j}} \\ \frac{k''_{3j}}{k''_{.j}} \\ \cdots \\ \frac{k''_{ij}}{k''_{.j}} \end{bmatrix} \qquad \text{con } k''_{.j} = \sum_{i=1}^{I} k''_{ij}$$

Dado que ya se obtuvieron las fórmulas de transición para las proyecciones sobre los ejes principales (4), lo que se hace es proyectar el punto j'' sobre el eje α , usando la formula:

$$\varphi_{\alpha j''} = \frac{1}{\sqrt{\lambda_\alpha}} \sum_{i=1}^{I} \left(\frac{k''_{ij}}{k''_{.j}} \right) \psi_{\alpha i}$$

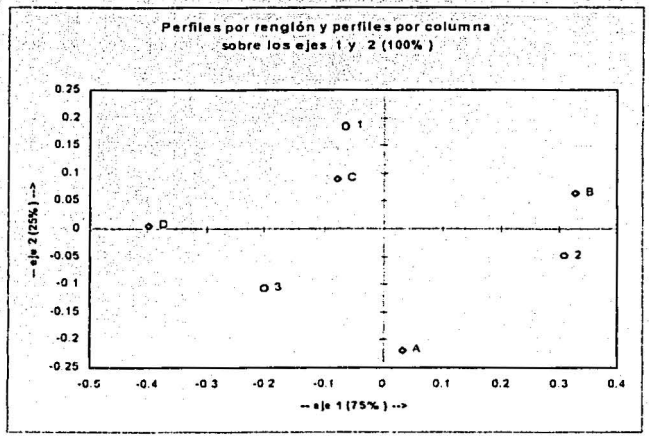
Para el caso de un categoría por renglón complementaria tenemos que los perfiles de las n_s categorías suplementarias con respecto a los I puntos analizados en el espacio R^J. Sea k'_{ij} la j-ésima cooredanda del i-ésimo renglón suplementario. El pérfil de este renglón es el vector cuyo j-ésimo elemnto es $\frac{k'_{ij}}{k'_{i.}}$ es decir:

$$i' = \begin{bmatrix} \frac{k'_{i1}}{k'_{i.}} & \frac{k'_{i2}}{k'_{i.}} & \frac{k'_{i3}}{k'_{i.}} & \cdots\cdots & \frac{k'_{ip}}{k'_{i.}} \end{bmatrix} \qquad \text{donde } k'_{i.} = \sum_{j=1}^{J} k'_{ij}$$

De nuevo lo que se hace es proyectar el punto i' sobre el eje α , usando la fórmula:

$$\psi_{\alpha i'} = \frac{1}{\sqrt{\lambda_\alpha}} \sum_{j=1}^{J} \left(\frac{k'_{ij}}{k'_{i.}} \right) \varphi_{\alpha j}$$

2.6 Descripción de la gráfica del Analisis de correspondencias para los datos de la tabla 2.1.



En base a las consideraciones hechas con las contribuciones observamos que con los dos primeros ejes principales se explica el 100% de la inercia, y además la calidad de respresentación también es del 100% para cada perfil, tanto por renglón o por columna. Se observa que en el primer eje el cual recoge el 75% de la inercia ésta es generada principalmente por el tipo de defecto D y B, en el caso del turno se tiene que el tercero y el segundo son los que más aportan a la inercia del primer eje. Con lo anterior se puede establecer que debido al alejamiento del origen (centroide) el tipo de defecto D se encuentra asociado mayoritariamente con el turno de fabricación 3, en contraparte con el tipo de defecto B que está más asociado con el turno de producción 2; es importante señalar que tomando en cuenta únicamente el primer eje no se muestra alguna relación con respecto a los tipos de defecto A y C y turno de producción 1.

Con el segundo eje, se puede ver que el tipo de defecto A es opuesto al B y C y además se aprecia que existe una asociación no tan marcada entre el turno 3 y el defecto A la cual es opuesta a la asociación entre el tipo de defecto C y el primer turno.

Observación:

En vista de lo anteriormente expuesto, es necesario poner atención en lo siguiente:

1. Inercia acumulada por los ejes que queremos retener.

2. La calidad de representación de los puntos.

3. Tener en cuenta que u y $-u$ son solución al problema de optimización, lo que ocasiona que se puedan obtener gráficas "aparentemente" diferentes. Las discrepancias son debidas a que en ocasiones lo que aparecía a la "derecha" ahora aparece a la izquierda y lo de "arriba" aparece ahora "abajo".

Resúmen del capítulo

A manera de resumen se plantean en forma esquemática los pasos realizados hasta el momento para el desarrollo del análisis de correspondencias.

1. Se tienen los datos de un fenómeno o problema de investigación en una tabla de contingencia de dos factores con I y J categorías respectivamente.

2. El número de categorías en las variable determina los espacios correspondientes a los renglones (R^J) y a las columnas (R^I).

3. La tabla de contingencia puede ser vista como un matriz de $I \times J$.

4. Se busca en el análisis indices o factores a partir de los datos originales de tal manera que conserven la mayor cantidad de varianza original y además generen un espacio óptimo en el cual la distancia de los puntos a dicho espacio sea mínima.

5. Se encontraron para el espacio R^J esos factores o índices por medio de la relación $\varphi_\alpha = D_c^{-1} u_\alpha$, donde el vector u_α maximiza la forma cuadrática asociada a suma ponderada de cuadrados de las proyecciones sobre el espacio generado (este vector se conoce como el α eje principal)con la restricción $u_\alpha' D_c^{-1} = 1$ (norma uno). Y donde se utilza la distancia Ji-cuadrada.

6. Encontrados los factores se procedió a encontrar las proyecciones sobre los ejes principales.

7. Para saber que tan buena es esta representación en los ejes se caculan las contribuciones respectivas de cada eje.

8. Todo este desarrollo se debe hacer con origen en el centroide debido a que el subespacio óptimo debe contener al centroide.

9. Todo esto es aplicable de la misma manera al espacio en R^I y para el cual se establece una espacio de dualidad con respecto al espacio R^J mediante las fórmulas de transición encontradas.

Capítulo III
Análisis de correspondencias múltiples.

El análisis de correspondencias simple, el cual ha sido descrito en el capítulo II, puede ser generalizado para poder estudiar la asociación de más de dos factores de manera simultánea, sin embargo como se aumenta la dimensión del espacio, para que las categorías de los factores puedan ser representados en un espacio de menor dimensión existe una pérdida de información en los factores por lo que los porcentajes de inercia en los ejes son bajos y la descripción de la gráfica no es del todo satisfactoria. En este caso la interpretación de los factores sigue las mismas reglas que en el análisis de correspondencias simple, la única peculiaridad reside en que la parte de la inercia explicada pierde aquí su interés ya que los valores propios no representan más que una pequeña parte de la inercia total.

Existen dos formas de presentar el análisis de correspondencias, una utilizando la información por individuos que se encuentra en una matriz indicadora Z o bien utilizando la matriz de Burt; la cual recoge la clasificación de la muestra de acuerdo a cada factor y es en si una generalización de la tabla de contingencia. Ambas formas son equivalentes en el sentido de la descripción entre las categorías y difieren unicamente en los valores de las inercias que cada uno arroja.

Cuando se trabaja con la matriz indicadora Z, lo que se hace es una extensión del análisis de correlación canónico. El propósito de presentar al Análisis de Correlación Canónico es el de ejemplificar una forma alternativa de trabajar con el AFC, recordemos que en el capítulo anterior el desarrollo del AFC se hizo con base a una tabla de contingencia, pero también se puede comenzar el análisis a través de los datos obtenidos de manera directa es decir, la información codificada. Esta información se presenta en una matriz indicadora Z en la que aparece un 1 para la categoría del factor que se haya elegido. El tener esta matriz, en la cual se le asigna un valor numérico a una característica cualitativa nos permitirá el uso de otras técnicas como el análisis de correlación canónico y además nos permitirá estudiar el problema de la generalización del análisis de correspondencias para el caso en que se tengan más de dos factores.

Es importante mencionar que la forma en que se quiera empezar con el análisis depende de las características particulares del problema que esté siendo atacado, así por ejemplo, si se quiere hacer un estudio de discriminación con respecto a un grupo de variables que puedan describir de manera adecuada el perfil de un trabajador (Pérez Soto Claudia: Análisis de correspondencias múltiples como una técnica para el estudio de datos cualitativos,1990) lo acertado sería utilizar la matriz indicadora Z , para hacer un análisis por individuos y por variables. Pero en dado caso de que se quiera estudiar la asociación entre las categorías sería recomendable el uso de la matriz de Burt.

3.1 El análisis de correlación canónico (ACC).

El Análisis de correlación canónico (ACC) es una técnica multivariada que investiga la relación entre dos conjuntos de variables que pueden ser un conjunto de variables de respuesta y un conjunto de variables explicactivas, ambos conjuntos de variables continuas.

Desde el punto de vista del análisis de datos el ACC consiste en estudiar las relaciones lineales existentes entre dos grupos de variables cuantitativas observadas sobre un mismo conjunto de individuos.

Para la utilización del análisis de correlación canónico con propósitos descriptivos no se requieren supuestos distribucionales. En estos casos, las variables de predicción y las variables de respuesta pueden ser medidas en una escala nominal u ordinal. Sin embargo para probar la significancia de la relación entre las variables canónicas , los datos deben presentar los requerimientos de normalidad multivariada y homogeneidad de varianza; es decir deben ser cuantitativas.

Como el ACC pretende poner en evidencia las relaciones lineales que existen entre los dos conjuntos de variables, la pregunta de interés es ¿cuál es el criterio para decir que las relaciones lineales entre el conjunto de variables explicativas y de respuesta son próximas?.

Así, en el ACC se tienen los siguientes elementos:

$\{x^1, x^2, x^3, ..., x^p\}$ y $\{y^1, y^2, y^3, ..., y^q\}$ dos grupos de variables cuantitativas medidas sobre sobre un conjunto I de individuos.

Estos grupos de variables definen matrices así, por ejemplo la matriz $X=[x^1\ x^2\ x^3 \dots\ x^p]$ está asociada a la aplicación X y es la matriz de datos con respecto al primer paquetes de variables, de manera análoga $Y=[y^1\ y^2\ y^3 \dots\ y^q]$ es la matriz de datos para el segundo grupo de variables.

Como un caso particular se tienen el ACC clásico en el cual se considera que cada individuo tiene asociado un peso p_i, positivo y además $\sum_{i=1}^{n} p_i = 1$. Para cada paquete de variables se considera que cada variable está centrada, por lo que el centro de gravedad de la nube de individuos coincide con el origen.

El espacio de variables tiene asociada la métrica de pesos M

$$M=D_p=\begin{pmatrix} p_1 & 0 & 0 \\ 0 & \dots & 0 \\ 0 & 0 & p_n \end{pmatrix}$$

para este caso las matrices $V_{XX} = X^t D_p X$, $V_{YY} = Y^t D_p Y$, $V_{XY} = X^t D_p Y$, representan matrices de varianzas y covarianzas asociadas respectivamente a los paquetes de variables X y Y y

entre ambos grupos.

Así, los factores canónicos pueden ser obtenidos a partir de las ecuaciones:

$$V_{XX}^{-1} V_{XY} V_{YY}^{-1} V_{YX} a = \lambda a$$

$$V_{YY}^{-1} V_{YX} V_{XX}^{-1} V_{XY} b = \lambda b$$

con $\|a\|_{V_{XX}} = 1$ y $\|b\|_{V_{YY}} = 1$

de tal manera que las variables canónicas ξ^1 y η^1 verifican que $\xi^1 = X a_1$ y $\eta^1 = Y b_1$ y además si $\lambda_i \neq 0$, las ecuaciones de transición pueden ser expresadas como:

$$a_i = \frac{1}{\sqrt{\lambda_i}} V_{XX}^{-1} V_{XY} b_i$$

$$b_i = \frac{1}{\sqrt{\lambda_i}} V_{YY}^{-1} V_{YX} a_i$$

Además, como las variables canónicas son centradas y reducidas se tiene que $corr^2(\xi^i, \eta^i) = \lambda_i$, de lo que se deduce que el i-ésimo valor propio λ_i es el cuadrado del i-ésimo coeficiente de correlación canónica.

La interpretación geométrica del análisis de correlación canónico es la siguiente:

De manera general se tiene

$$a = \frac{1}{\sqrt{\lambda_i}} V_{XX}^{-1} V_{XY} b$$

$$b = \frac{1}{\sqrt{\lambda_i}} V_{YY}^{-1} V_{YX} a$$

Si se premultiplican los elementos de cada ecuación por X y Y respectivamente, se tiene:

$$Xa = \frac{1}{\sqrt{\lambda_i}} X (X^t X)^{-1} X^t Y b \qquad \text{donde } V_{XX}^{-1} = (X^t X)^{-1} \text{ y } V_{XY} = X^t Y$$

$$Yb = \frac{1}{\sqrt{\lambda_i}} Y (Y^t Y)^{-1} Y^t X a \qquad \text{donde } V_{YY}^{-1} = (Y^t Y)^{-1} \text{ y } V_{YX} = Y^t X$$

Llamemos $\wp_X, \wp_Y$ los subespacios lineales de R^n, que son generados por las columnas de X y Y respectivamente. Las combinaciones lineales de a y b definen los puntos de $\wp_X$ y de $\wp_Y$ respectivamente, cuyas coordenadas son los componentes de los vectores Xa y Yb respectivamente. Por lo que las matrices

$$P_x = X (X^t X)^{-1} X^t$$

$$P_Y = Y (Y^t Y)^{-1} Y^t$$

son los operadores de proyección sobre $\wp_X$ y $\wp_Y$ respectivamente. Es decir cada vector es

colineal a la proyección del otro.

Como los vetores Xa y Yb son unitarios, se tiene que:

$\sqrt{\lambda} = \cos\omega = \cos(Xa, Yb)$

Así, se tiene que la primera raíz canónica λ es el coseno al cuadrado de el ángulo más pequeño entre los subespacios $\wp_X$ y $\wp_Y$

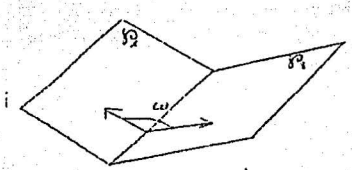


Fig 3.1. Interpretación geométrica del ACC

Hasta ahora se han mostrado en que consiste en ACC de manera general. Las ideas aquí presentadas nos van a permitir entender parte de la generzalización de la técnica del análisis de correspondencias múltiples.

Para introducir la generalización del análisis de correspondencias múltiples (AFC-múltiples) presentamos algunas consideraciones

3.2 Consideraciones para el análisis de correspondencias múltiple.

- Se tienen Q factores . Para un factor q en particular se tiene un conjunto de p_q categorías.

Así entonces, el número total de categorías que se tienen es $p=\sum_{q=1}^{Q} p_q$

- El tamaño de muestra que es considerada para los Q factores es n.

Lo anterior queda ejemplificado a través de lo siguiente: cuando en algunos estudios estadísticos se tienen una gran cantidad de preguntas para las cuales las respuestas son categorías y además se tiene que para cada pregunta las categorías son mutuamente excluyentes; entonces para cada pregunta o factor que tiene k categorías se tienen a lo más k grupos en las cuales la muestra de tamaño n puede ser repartida. Es decir si preguntamos a 10 personas acerca de su religión (suponiendo que sólo consideramos católica, protestante, musulmana) entonces a lo más se pueden formar con la muestra los tres grupos referentes a cada religión o pudiera darse el caso de que los 10 fueran católicos o los 10 fueran protestantes o qu los 10 fueran musulmanes.

3.2.1 Matriz indicadora $Z=[Z_1, Z_2,..., Z_q,...,Z_Q]$.

Como se ha dicho si tenemos Q factores, entonces la persona únicamente puede escoger una sola categoría para cada factor q q=1,2,...,Q. Esta información se recoge en la matriz Z, es decir, se denota por Z a la matriz con I renglones (número de individuos) y p (número de categorías) columnas que describe para cada individuo i de la muestra de tamaño I su elección o clasificación de acuerdo a las categorías de los factores.

La matriz Z es la yuxtaposición de las Q submatrices tales que

$Z=[Z_1, Z_2,..., Z_q,...,Z_Q]$

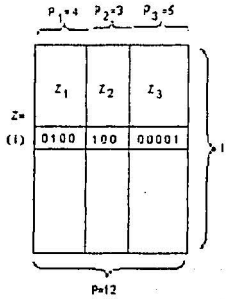


Fig 3.2. Matriz indicadora

La submatriz Z_q (la cual tiene I renglones y P_q columnas) es tal que su i-ésimo renglón tiene $P_q - 1$ veces el valor cero y una vez el valor uno, en la columna correspondiente a la categoría del factor q en el cual fue clasificado el sujeto i.

Esta matriz representa la información por individuo, pero si quisieramos trabajar con una matriz con datos condensados entonces podemos hacer uso de la llamada tabla de contingencia asociada a la matriz de Burt.

3.2.2 Tabla de Burt asociada con Z.

La tabla de Burt condensa la información generada por la matriz indicadora Z, y puede ser vista como bloques de matrices, donde los bloques son tablas de contingencia tomadas de dos en dos. Los bloques que contienen a los elementos de la diagonal son matrices diagonales debido a que corresponden a la tabla de contingencia de una variable en ella misma, por lo que los elementos de la diagonal son las frecuencias de dicha variable.

A la matriz $B=Z'Z$ se le conoce como tabla de contingencia de Burt asociada con Z y tiene la siguiente forma:

$$B=\begin{bmatrix} Z_1'Z_1 & \dots & Z_1'Z_q & \dots & Z_1'Z_Q \\ \dots & \dots & \dots & \dots & \dots \\ Z_q'Z_1 & \dots & Z_q'Z_q & \dots & Z_q'Z_Q \\ \dots & \dots & \dots & \dots & \dots \\ Z_Q'Z_1 & \dots & Z_Q'Z_q & \dots & Z_Q'Z_Q \end{bmatrix}$$

Nótese que la matriz B está formada por Q^2 bloques. La q-ésima submatriz $Z_q'Z_q$ es una matriz diagonal cuyos elementos son las frecuencias para cada categoría del factor q. La submatriz $Z_q'Z_{q'}$ es la tabla de contingencia que clasifica la muestra de tamaño n en los factores q y q'.

La matriz diagonal D es una matriz cuadrada que tiene los mismos elementos en la diagonal que B; esos elementos en la diagonal son las frecuencias para cada una de las categorías. La matriz D también tiene Q^2 bloques. Y además solamente las Q submatrices diagonales son diferentes a la matriz cero. De tal manera que el q-ésimo bloque $D_q=Z_q'Z_q$ es una matriz diagonal cuyos elementos son las frecuencias correspondientes a las categorías del factor q.

Ejemplo1 : Supongamos que tenemos 10 personas a las cuales las clasificamos de acuerdo a los siguientes factores: Z_1 = sexo (masculino , femenino), Z_2 =estado civil (soltero, casado, viudo) y Z_3 = partido político de preferencia (PRI,PAN,PRD). La matriz Z que representa la información es la siguiente:

$$Z = \begin{array}{c} \begin{matrix} Z_1 & Z_2 & Z_3 \end{matrix} \\ \left[\begin{array}{ccc||c} 10 & 100 & 100 & 1 \\ 01 & 100 & 010 & 2 \\ 01 & 010 & 100 & 3 \\ 10 & 100 & 100 & 4 \\ 10 & 010 & 001 & 5 \\ 10 & 001 & 001 & 6 \\ 01 & 001 & 010 & 7 \\ 01 & 010 & 100 & 8 \\ 10 & 100 & 010 & 9 \\ 01 & 100 & 010 & 10 \end{array}\right] \end{array} \left.\begin{matrix} \\ \\ \\ \\ \\ \\ \\ \\ \\ \\ \end{matrix}\right\} \text{individuo}$$

En esta tabla se lee por ejemplo que para el individuo 4: es hombre, soltero y tiene preferencia por el PRI. Para el décimo individuo: es mujer, soltera y con preferencia por el PAN. Se observa además que de los 10 individuos hay 5 hombres y 5 mujeres; 5 solteros, 3 casados y 2 viudos; 4 con preferencias por el PRI, 4 con preferencias por el PAN y 2 que tienen prefemcias por el PRD.

La matriz de Burt se obtiene de la siguiente manera:

$$Z' = \begin{bmatrix} Z_1' \\ Z_2' \\ Z_3' \end{bmatrix}$$

$$B = \underbrace{\begin{bmatrix} 1 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 & 1 & 1 & 0 & 1 \\ 1 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \end{bmatrix}}_{Z'} \underbrace{\begin{bmatrix} 10 & 100 & 100 \\ 01 & 100 & 010 \\ 01 & 010 & 100 \\ 10 & 100 & 100 \\ 10 & 010 & 001 \\ 10 & 001 & 001 \\ 01 & 001 & 010 \\ 01 & 010 & 100 \\ 10 & 100 & 010 \\ 01 & 100 & 010 \end{bmatrix}}_{Z}$$

Haciendo las operaciones se tiene que

$$B = \begin{bmatrix} 5 & 0 & 3 & 1 & 1 & 2 & 1 & 2 \\ 0 & 5 & 2 & 2 & 1 & 2 & 3 & 0 \\ 3 & 2 & 5 & 0 & 0 & 2 & 3 & 0 \\ 1 & 2 & 0 & 3 & 0 & 2 & 0 & 1 \\ 1 & 1 & 0 & 0 & 2 & 0 & 1 & 1 \\ 2 & 2 & 2 & 2 & 0 & 4 & 0 & 0 \\ 1 & 3 & 3 & 0 & 1 & 0 & 4 & 0 \\ 2 & 0 & 0 & 1 & 0 & 0 & 0 & 2 \end{bmatrix} \Rightarrow$$

Tabla 1. Matriz de Burt.

	Hombre	Mujer	Soltero	Casado	Viudo	PRI	PAN	PRD
Hombre	5	0	3	1	1	2	1	2
Mujer	0	5	2	2	1	2	3	0
Soltero	3	2	5	0	0	2	3	0
Casado	1	2	0	3	0	2	0	1
Viudo	1	1	0	0	2	0	1	1
PRI	2	2	2	2	0	4	0	0
PAN	1	3	3	0	1	0	4	0
PRD	2	0	0	1	0	0	0	2

Nótese que la matriz B está formada por Q^2 bloques, es decir el cruce del factor sexo consigo mismo y con los otros dos, el cruce del factor estado civil consigo mismo y con los otros dos y el

cruce del factor preferencia política, igual, consigo mismo y los otros dos, en total se tiene 9 bloques. La matriz $Z_1'Z_1$ (en rojo) es una matriz diagonal que representa la frecuencia de cada categoría del factor sexo (recuérdese que se tenían 5 hombres y 5 mujeres); la matriz $Z_2'Z_2$ (en azul) es una matriz diagonal que representa la frecuencia de cada categoría del factor estado civil (recuérdese que se tenían 5 solteros, 3 casados y 2 viudos), la matriz $Z_3'Z_3$ (en verde) es una matriz diagonal que representa la frecuencia de cada categoría del factor preferencia por el partido político(recuérdese que se tenían 4 personas con preferencias por el PRI, 4 con preferencias por el PAN y 2 que tienen preferencias por el PRD). El bloque $Z_1'Z_{2'}$ es la tabla de contingencia que clasifica a la muestra de 10 personas con respecto a los factores sexo y estado civil; de tal manera que se lee en la tabla que de los 5 hombres en la muestra 3 eran solteros, 1 era casado y 1 viudo. De manera semejante se lee para el caso de las mujeres.

El bloque $Z_1'Z_{3'}$ es la tabla de contingencia que clasifica a la muestra de 10 personas con respecto a los factores sexo y preferencia política; de tal manera que se lee en la tabla que de los 5 hombres en la muestra 2 preferían al PRI, 1 al PAN y 2 al PRD. De manera semejante se lee para el caso de las mujeres.

El bloque $Z_2'Z_{3'}$ es la tabla de contingencia que clasifica a la muestra de 10 personas con respecto a los factores estado civil y preferencia política; de tal manera que se lee en la tabla que de las 5 personas solteras en la muestra 2 preferían al PRI, 3 al PAN y ninguna al PRD. De las 3 personas casadas 2 preferían al PRI, ninguna al PAN y 1 al PRD; y, de las dos personas viudas ninguna prefería al PRI, una al PAN y una al PRD.

Hasta aquí, se ha mencionado la forma en que se obtiene la información, resulta importante el hecho de que es posible hacer el análisis con los datos de la matriz indicadora (datos crudos, sin procesar) y haciendo uso de la matriz de Burt (datos procesados).

En lo que sigue se expondrán las equivalencias para el caso de dos factores del análisis de correspondencias utilizando las matrices que hemos descrito hasta el momento:

- Tabla de contingencia entre dos factores.

- Matriz Indicadora Z.

- Matriz de Burt.

Esto para mostrar las formas usuales en que se reporta el análisis de correspondencias para el caso de más de dos factores.

3.3 Análisis de correspondencias para una tabla de dos factores considerando diversas matrices.

La matriz de respuestas contenidas en la matriz Z es $Z = [Z_1|Z_2]$

Es equivalente, desde el punto de vista de describir la relación entre categorías, llevar a cabo alguno de los siguientes análisis:

1. Un análisis de correspondencias de la matriz $Z_{I\times p}$

2. Un análisis de correspondencias de la matriz de Burt $B_{p\times p}$

3. Un análisis de correspondencias de la tabla de contingencia $K_{I\times J} = Z_1' Z_2$

4. Un análisis canónico de los dos conjuntos de bloques Z_1 y Z_2.

Para mostrar lo anterior a excepción del punto 4 sea lo siguiente:

Equivalencia entre el análisis 1 y 2.

El α-ésimo factor ϕ_a extraído del análisis 1 es el eigenvector de la siguiente ecuación:

$$\frac{1}{Q} D^{-1} Z' Z \phi_a = \mu_a \phi_a$$

Lo anterior se muestra de lo siguiente; recordando la notación utilizada para el análisis de correspondencias utilizado en el capítulo II en el cual se estableció que el factor φ era el eigenvector de la matriz $A=D_c^{-1}F'D_r^{-1}F$

Haciendo

$$F=\frac{1}{nQ}Z$$

$$D_c = \frac{1}{nQ}D$$

$D_r = \frac{1}{n} I_n$ con I_n la matriz identidad.

sustituyendo se tiene que:

$$A=D_J^{-1}F'D_I^{-1}F = \left(\frac{1}{nQ}D\right)^{-1}\frac{1}{nQ}Z'\left(\frac{1}{n}I_n\right)^{-1}\frac{1}{nQ}Z = \frac{1}{Q}D^{-1}Z'Z.$$

Así, los factores son :

$$\frac{1}{Q} D^{-1} Z' Z \phi_a = \mu_a \phi_a \qquad (3.1)$$

Lo anterior se ejemplifica con los datos del ejemplo 1:

Se tiene la matriz Z que es la siguiente:

$Z =$

	Sexo		Estado civil			Preferencia política			
	Hombre	Mujer	Soltero	Viudo	Casado	PRI	PAN	PRD	Total
	1	0	1	0	0	1	0	0	3
	0	1	1	0	0	0	1	0	3
	0	1	0	1	0	1	0	0	3
	1	0	1	0	0	1	0	0	3
	1	0	0	1	0	0	0	1	3
	1	0	0	0	1	0	0	1	3
	0	1	0	0	1	0	1	0	3
	0	1	0	1	0	1	0	0	3
	1	0	1	0	0	0	1	0	3
	0	1	1	0	0	0	1	0	3
Total	5	5	5	3	2	4	4	2	30

entonces la matriz F está dada la expresión siguiente donde Q es el número de factores.

$F = \frac{1}{10*3} Z = \frac{1}{30} Z$ = matriz de correspondencias

$$D_c = \frac{1}{10*3} D = \begin{pmatrix} 5 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 5 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 5 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 3 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 4 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 2 \end{pmatrix}$$

la matriz diagonal de pesos por columnas

$$D_r = \frac{1}{10} I_{10} = \begin{pmatrix} \frac{1}{10} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \frac{1}{10} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{10} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{10} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{10} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \frac{1}{10} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{10} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{10} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{10} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{10} \end{pmatrix}$$

Ahora, consideremos la matriz B tal que la matriz B=Z'Z. B es simétrica. Además los renglones y las columnas marginales son los elementos de la matriz QD.

En el análisis de la matriz B

Sea la matriz

$F=\frac{1}{nQ^2}B$

las matrices

$D_c = D_r = \frac{1}{nQ}D$

Así la matriz a ser diagonalizada es:

$A=D_c^{-1} F^t D_r^{-1} F = \left(\frac{1}{nQ}D\right)^{-1}\frac{1}{nQ^2}B^t\left(\frac{1}{nQ}D\right)^{-1}\frac{1}{nQ^2}B = \frac{1}{Q^2}D^{-1}B^t D^{-1}B$

si premultiplicamos $\frac{1}{Q}D^{-1}Z^tZ\phi_a = \mu_a\phi_a$ por $\frac{1}{Q}D^{-1}B$ obtenemos lo siguiente:

$\left(\frac{1}{Q}D^{-1}B\right)\frac{1}{Q}D^{-1}Z^tZ\phi_a = \left(\frac{1}{Q}D^{-1}B\right)\mu_a\phi_a$

$\Leftrightarrow$

$\frac{1}{Q^2}D^{-1}BD^{-1}B\phi_a = \left(\frac{1}{Q}D^{-1}B\right)\mu_a\phi_a$, pero $\left(\frac{1}{Q}D^{-1}B\right)$ es un eigenvector para (3.1)

por lo que

$\frac{1}{Q^2}D^{-1}BD^{-1}B\phi_a = \mu_a^2\phi_a$

por lo que los factores son idénticos para ambos análisis.

La matriz de Burt para los datos del ejemplo 1 son:

B=

	Hombre	Mujer	Soltero	Casado	Viudo	PRI	PAN	PRD	Total
Hombre	5	0	3	1	1	2	1	2	15
Mujer	0	5	2	2	1	2	3	0	15
Soltero	3	2	5	0	0	2	3	0	15
Casado	1	2	0	3	0	2	0	1	9
Viudo	1	1	0	0	2	0	1	1	6
PRI	2	2	2	2	0	4	0	0	12
PAN	1	3	3	0	1	0	4	0	12
PRD	2	0	0	1	1	0	0	2	6
Total	15	15	15	9	6	12	12	6	90

por lo que las matrices son

$F=\frac{1}{10\times3^2}B=\frac{1}{90}B$ ⇐ matriz de correspondencias

$$D_c = D_l = \frac{1}{10\times3}D = \frac{1}{30}D = \begin{pmatrix} \frac{5}{30} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \frac{5}{30} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{5}{30} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{3}{30} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{2}{30} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \frac{4}{30} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \frac{4}{30} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{2}{30} \end{pmatrix}$$

Nótese que la matriz anterior representa la matriz diagonal de pesos por renglón y a la matriz diagonal de pesos por columna. Y se obtiene si a la columna y renglón marginal se divide entre la suma total. Igual que se hacía en el análisis de correspondencias simple expuesto en el capítulo II.

Las equivalencias anteriores fueron presentadas utilizando 3 factores como caso particular cuando se tienen dos factores está la siguiente equivalencia.

Ahora, la equivalencia entre el análisis 1 y el análisis 3

Se mostrará que para cada par de factores (φ_a, ψ_a) relativos al mismo eigenvalor λ_a el cual es extraido del análisis de la tabla de contingencia $K=Z_1' Z_2$, hay un factor ϕ_a del análisis Z (o B) tal que:

$$\phi_a = \begin{bmatrix} \varphi_a \\ \psi_a \end{bmatrix}$$

Recordemos que la matriz $D_1 = Z_1' Z_1$ y la matriz $D_2 = Z_2' Z_2$ son las matrices diagonales que representan el renglón y columna marginales de la tabla de contingencia.

$$D=\begin{bmatrix} D_1 & 0 \\ 0 & D_2 \end{bmatrix}$$

Recordemos que las formulas de transición fueron establecidas de la siguiente manera:

$$\varphi_a = \frac{1}{\sqrt{\lambda_a}} D_1^{-1} Z_1' Z_2 \psi \qquad (3.2)$$

$$\psi_a = \frac{1}{\sqrt{\lambda_a}} D_2^{-1} Z_2' Z_1 \varphi_a$$

De (3.2) tenemos que $\varphi_a\sqrt{\lambda_a} = D_1^{-1}Z_1'Z_2\psi_a$ y sumando a ambos lados φ_a y agrupando tenemos lo siguiente:

$(1+\sqrt{\lambda_a})\varphi_a = D_1^{-1}(D_1\varphi_a + Z_1'Z_2\psi_a)$, haciendo lo mismo con la otra ecuación se tiene lo siguiente:

$(1+\sqrt{\lambda_a})\psi_a = D_2^{-1}(D_2\psi_a + Z_2'Z_1\varphi_a)$

Este sistema de ecuaciones en forma matricial se reescribe de la siguiente manera

$$\begin{bmatrix} D_1 & 0 \\ 0 & D_2 \end{bmatrix}^{-1}\begin{bmatrix} D_1 & Z_1'Z_2 \\ Z_2'Z_1 & D_2 \end{bmatrix}\begin{bmatrix} \varphi_a \\ \psi_a \end{bmatrix} = (1+\sqrt{\lambda_a})\begin{bmatrix} \varphi_a \\ \psi_a \end{bmatrix}$$

esto a su vez representa lo siguiente:

$$D^{-1}Z'Z\begin{bmatrix} \varphi_a \\ \psi_a \end{bmatrix} = (1+\sqrt{\lambda_a})\begin{bmatrix} \varphi_a \\ \psi_a \end{bmatrix}$$

$$\Leftrightarrow$$

$$D^{-1}Z'Z\phi_a = (1+\sqrt{\lambda_a})\phi_a$$

Como estamos tratando con dos factores, y como se mostró en la ecuación (3.1) para el análisis de la matriz Z, entonces haciendo Q=2 se tiene finalmente:

$$\frac{1}{Q}D^{-1}Z'Z\phi_a = \left(\frac{1+\sqrt{\lambda_a}}{2}\right)\phi_a$$

lo cual tiene los mismos factores que el análisis 1, sin embargo el eigenvalor ahora es:

$$\mu_a = \frac{1+\sqrt{\lambda_a}}{2}$$

Así, si λ_a es el eigenvalor más grande obtenido del análisis de la tabla de contingencia K entonces la ecuación anterior muestra el $\alpha - \acute{e}simo$ eigenvalor más grande del análisis de la matriz Z.

Por transitividad los tres análisis son equivalentes, por lo tanto hacer un análisis de correspondencias de la tabla de contingencias es equivalente a realizar un análisis de correspondencias de la matriz indicadora y es equivalente también a efectuar un análisis de correspondencias de la tabla de Burt.

En la sección siguiente se hará una generalización para resolver el problema de encontrar los componentes y los ejes para el caso de más de dos factores, la idea que se pretende aprovechar es la generalización del análisis de correlación canónico a través del cual se concluye que los eigenvalores se obtienen de la utiliación de la matriz de Burt B y como se mostró en el apartado

anterior son proporcionales a los que se obtienen utilizando la matriz Z.

3.4 Generalización del análisis de correspondencias utilizando la matriz Z, para más de dos factores.

De la figura 1 se observa que $Z=[Z_1, Z_2, ..., Z_Q]$ tiene P columnas que es el número de categorías totales y cada columna corresponde a un puntos en R^n. Consideremos el espacio R^n. Cada submatriz Z_q genera un subespacio lineal $\wp_q$ con p_q dimensiones. El rango de la matriz Z es a lo más igual a P-(Q-1) ya que todos esos subespacios tienen en común al menos el primer bisector (el vector cuyos componentes son iguales a uno).

Sea φ_q el vector cuyas p_q componentes son las coordenadas de un punto m_q de $\wp_q$ en la base definida por las columnas de Z_q.

Las coordenadas de m_q en R^n son los componentes de $m_q = Z_q\varphi_q$

La distancia al cuadrado de este punto m_q a el origen es :

$$\varphi_q' Z_q' Z_q \varphi_q = \varphi_q' D_q \varphi_q$$

Así, el análisis de correspondencias de la tabla de contingencia para dos preguntas q y q' se reduce a estudiar las posiciones relativas de los subespacios $\wp_q$ y $\wp_{q'}$. Y esto precisamente corresponde al análisis canónico de la matriz $[Z_q|Z_{q'}]$.

Por lo que las ecuaciones de transición quedan de la siguiente forma:

$$\varphi_q = \frac{1}{\sqrt{x}}(Z_q' Z_q)^{-1} Z_q' Z_{q'} \varphi_{q'} = \frac{1}{\sqrt{x}} D_q^{-1} Z_q' Z_{q'} \varphi_{q'}$$

$$\varphi_{q'} = \frac{1}{\sqrt{x}}(Z_{q'}' Z_{q'})^{-1} Z_{q'}' Z_q \varphi_q = \frac{1}{\sqrt{x}} D_{q'}^{-1} Z_{q'}' Z_q \varphi_q$$

De manera análoga a lo hecho para el análisis canónico si premultiplicamos por Z_q y $Z_{q'}$ respectivamente tenemos lo siguiente:

$$Z_q\varphi_q = \frac{1}{\sqrt{x}} Z_q (Z_q' Z_q)^{-1} Z_q' Z_{q'} \varphi_{q'}$$

$$Z_{q'}\varphi_{q'} = \frac{1}{\sqrt{x}} Z_{q'} (Z_{q'}' Z_{q'})^{-1} Z_{q'}' Z_q \varphi_q$$

Donde se obtiene que los operadores de proyección sobre $\wp_q$ y $\wp_{q'}$ son:

$$P_q = Z_q (Z_q' Z_q)^{-1} Z_q'$$

$$P_{q'} = Z_{q'} (Z_{q'}' Z_{q'})^{-1} Z_{q'}'$$

lo que finalmente proporciona las coordenadas de los puntos m_q y $m_{q'}$ para ambos

subespacios:

$$m_q = \frac{1}{\sqrt{\lambda}} P_q m_{q'}$$

$$m_{q'} = \frac{1}{\sqrt{\lambda}} P_{q'} m_q$$

De manera análoga a lo establecido en el análisis canónico la proyección ortogonal de m_q sobre $\varphi_{q'}$ es colineal a $m_{q'}$ y viceversa.

Una generalización para el problema del análisis canónico cuando se tienen más de dos factores.

Sean $\varphi_1, \varphi_2, ..., \varphi_Q$, respectivamente, los vectores de los componentes de los Q puntos, $m_1, m_2, ..., m_Q$ en las bases $Z_1, Z_2, ..., Z_q$, y sea $m=m_1 + m_2 + ... + m_Q$

La cantidad a ser maximizada es

$$\|m\|^2 = \sum \{\varphi_q' Z_q' Z_{q'} \varphi_{q'} | q \in Q \text{ y } q' \in Q\}$$

con la restricción

$$\sum \{\varphi_q' D \varphi_q \ | q \in Q\} = Qn$$

Si ϕ es el vector con p componentes definido por

$$\phi' = \{\varphi_1', \varphi_2', ..., \varphi_Q'\}$$

El problema llega a ser el

$$\max \phi' B \phi$$

$$\text{s.a } \phi' D \phi = Qn$$

Los factores requeridos ϕ son los eigenvectores de la matriz $D^{-1}B$, relativos a los eigenvalores más grandes. Ellos son proporcionales a aquellos que son extraídos del análisis de correspondencias de la matriz Z y coinciden, con los extraídos del análisis de la matriz B.

3.5 Resultados del análisis de correspondencias multiples utilizando la matriz Z.

Recordemos que los ϕ factores extraídos del análisis de la matriz Z son tales que:

$$\frac{1}{Q} D^{-1} B \phi = \mu \phi,$$

lo cual rearreglando los términos de esta ecuación para mostrar los componentes φ_q de ϕ relativos al factor q, de acuerdo a los bloques de las matrices D y B, se tiene:

$\frac{1}{Q}\sum_{q=1}^{Q} D_q^{-1} Z_q' Z_q \varphi_q = \mu \varphi_q$

Los Q subconjuntos de puntos correspondientes a las p_q categorías del factor q tienen el mismo centro de gravedad, el cual es también el centro de gravedad del conjunto entero de puntos. J_q denota el subconjunto de los p valores del índice j correspondeintes al factor q (J_q tiene p_q elementos)

Las coordenadas del subconjunto de puntos relativos al factor q son las columnas de $Z_q D_q^{-1} = Z_q (Z_q' Z_q)^{-1}$ y los elementos de la diagonal de $\frac{1}{I} D_q$ son las masas relativas de los p_q puntos del subconjunto q.

La i-ésima coordenada del centro de gravedad G_q es

$g_{qi} = \sum_{j \in J_q} \frac{d_{jj}}{I} \frac{z_{ij}}{d_{jj}} = \frac{1}{I}$ ya que $\sum_{j \in J_q} z_{ij} = 1$

De tal manera que g_{qi} es independiente de q, es decir $g_{qi} = g_i$.

Los φ_q factores correspondientes a los factores no triviales están centrados porque esos factores corresponden a un análisis del conjunto de puntos después de la traslación del origen a G.

Algunos resultados del análisis de correspondencias múltiples.

1. La suma de los pesos de las columnas $Z_q = \frac{1}{Q}$.

Es decir cada factor recibe el mismo peso, el cual es distribuido sobre todas las categorías de acuerdo a la frecuencia de cada catregoría del factor

Por ejemplo:

El peso para la submatriz $Z_1 = \frac{1}{3} = 0.16 + .016 = 0.33333$

El peso para la submatriz $Z_2 = \frac{1}{3} = 0.16 + 0.1 + .06 = 0.33333$

El peso para la submatriz $Z_3 = \frac{1}{3} = 0.13 + .013 + .06 = 0.33333$

2. El centroide de los perfiles columna de Z_q es el centro de la gráfica que es el de todos los perfiles columna, de tal manera que la nube de los perfiles para cada factor está balanceada en el origen.

3. La inercia total de los perfiles columna y renglón es:

$In(P) = \frac{J}{Q-1}$

4. La inercia de los perfiles columna de Z_q es:

$in(J_q) = \frac{(J_q - 1)}{Q}$

De esto se deduce que la inercia con la que contribuye un factor se incrementa linealmente con el número de categorías del factor.

5. La inercia de una categoría en particular (j) es:

$in(j) = \frac{1}{Q} - c_j^Z$

De esto se deduce que la inercia con la que contribuye una categoría incrementa en la medidad que decrece la frecuencia de esta categoría, con un límite superior de $\frac{1}{Q}$

6. El número de dimensiones no triviales con inercia positiva es a lo más J-Q, es decir el número de dimensiones de los perfiles por renglón y por columnas es a lo más J-Q.

7. Los perfiles renglón están en el centroide de las perfiles por columna, al reescalar por el inversa de la raíz cuadrada de las inercias principales a través de los ejes respéctivos.

Al igual que en el capítulo II nuestro problema consiste en encontrar una representación gráfica que nos permita apreciar en un espacio de menor dimensión, los perfiles de las categorías para poder estudiar posibles asociaciones entre los factores, por tal razón surge de nuevo la pregunta de ¿cuántas gráficas puedo obtener y cuál es la calidad de cada una?

3.6 Dimensionalidad de la configuración de las p categorías en R^n

Las coordenadas de las categorías en R^n son las columnas de ZD^{-1}. Ellas generan un subespacio cuya dimensión es el rango de ZD^{-1} y por lo tanto el rango de $Z=[Z_1, Z_2, ..., Z_q, ..., Z_Q]$

Todos los $\wp_q$ subespacios generados por las columnas de las Z_q matrices tienen en común el primer bisector Δ, por lo tanto el rango máximo de Z es

$$p_1 + (p_2 - 1) + ... + (p_Q - 1) = p - Q + 1$$

Por lo tanto el rango máximo de $D^{-1}Z'Z$, la matriz a ser diagonalizada es p-Q+1. pero el análisis con respecto al orígen O, el primer biserctor Δ es el eigenvector correspondiente al eigenvalor 1, en el análisis con respecto al centro de gravedad G , (p-Q) se encuentran (p-Q) eigenvalores diferentes de cero.

Así, escogiendo una base en el subespacio del conjunto de puntos, el problema puede ser reducido a encontrar eigenvalores y eigenvectores para una matriz de rango (p-Q).

3.7 La mejor representación simultánea para los individuos de la matriz Z

Nosostros buscamos las abscisas de los I individuos y las p categorías en el mismo eje tales que:

1. La abscisa de un individuo particular i es la media aritmética de su respuesta.

2. La abscisa de una categoría j es la media aritmética de las abcisas de los individuos que escogieron esto.

Como se ha mencionado las formulas de transición obtenidas para el análisis de la matriz Z son:

$$\Psi = \frac{1}{\sqrt{\mu}} \frac{1}{Q} Z\phi$$

$$\phi = \frac{1}{\sqrt{\mu}} \frac{1}{Q} D^{-1} Z'\Psi$$

donde Ψ_i es la abscisa del individuo i, y ϕ_j de la categoría j.

3.8 Ejemplo introductorio.Una forma alternativa de trabajar con más de dos factores.

El siguiente ejemplo aparece en el libro Applied Multivariate Data Analysis, Jobson. y establece lo siguiente:

Los datos que se presentan en la tabla 3.1 fueron obtenidos de los archivos de un estudiante de la carrera de servicios de advertencias legales para pobres y establece las frecuencias observadas para el número de personas a las que se les imputaba por un delito (cargo) y el resultado de la querella es decir si fueron sentenciados o absueltos por el delito por el que se les acusaba. Esta información estaba disponible tanto para hombres como para mujeres. El objetivo era examinar la asociación entre el cargo por el que se acusa a una persona y el resultado penal para hombres y mujeres. Véase Tabla 3.1

Tabla 3.1 Tabla de contingencia. Frecuencias observadas para personas acusadas por un cargo penal y el veredicto del juez, tanto para hombres como para mujeres

Veredicto del juez	Sexo	Cargo penal					
		Daños impulsivos	Robo < $1000	Desorden público	Posesión de narcóticos	Otros	Total
Absuelto	Hombre	8	11	5	7	12	43
	Mujer	5	15	3	1	6	30
Sentenciado	Hombre	105	32	11	23	37	208
	Mujer	32	57	6	2	25	122
Total		150	115	25	33	80	403

Una forma de averiguar esa relación sería el estudio de manera simultánea de los tres factores,

lo cual puede ser hecho con el análisis de correspondencias múltiples. Otra forma es la siguiente, se pueden concatenar algunas variables, por ejemplo el veredicto del juez y el sexo de la persona, de tal manera que se tiene la siguiente tabla de contingencia Tabla 3.2.

Tabla 3.2. Tabla de contingencia.Frecuencias observadas para personas acusadas por un cargo penal y concatenación de sexo y veredicto del juez.

Sexo/Veredicto del juez	Cargo Penal					
	Daños impulsivos	Robo <$1000	Desorden público	Posesión de narcóticos	Otros	Total
Hombre/Absuelto	8	11	5	7	12	43
Mujer/Absuelta	5	15	3	1	6	30
Hombre/Condenado	105	32	11	23	37	208
Mujer/Condenada	32	57	6	2	25	122
Total	150	115	25	33	80	403

De la tabla anterior se observa que para el factor Sexo/veredicto del juez y para el factor cargo penal se tienen cuatro y cinco categorías respectivamente. Los resultados de la prueba Ji cuadrada de independencia son:

$\chi^2_E = 70.9167$ que se distribuye $\chi^2_{(12)}$ y el cuantil de orden 0.95 de una distribución Ji-cuadrada con 12 grados de libertad es 21.0261 por lo que se rechaza la hipótesis de independencia entre los renglones y las columnas de la tabla de contingencia; además el nivel de significancia descriptivo p-value= $\Pr(\chi^2 > 70.9167) \cong 0.0000$.Por lo que se concluye que existe asociación entre las categorías.

Como se realizó en el capitulo II, un estudio gráfico exploratorio podría ser de utilidad. Así se tienen las siguientes gráficas con respecto a los perfiles con desviación del origen.

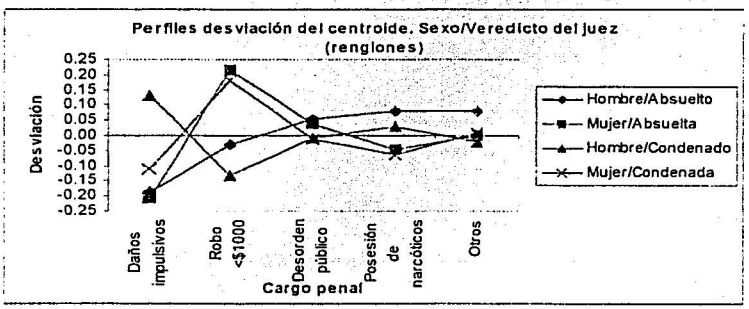


Fig 3.1

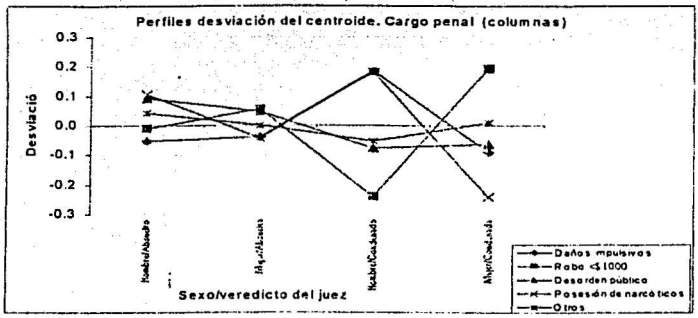


Fig 3.2

En la Fig 3.1 se observa que para el caso en que los hombres son condenados la desviación mayor se encuentra en la categoría de daños impulsivos, en contraparte con robo < $1000. Para el caso de las mujeres que si son sentenciadas se observa que la desviación mayor aparece en la categoría de robo<$1000 en contraparte con posesión de narcóticos.

De la Fig. 3.2 se observa que para el caso de posesión de narcóticos y daños impulsivos los hombres si son condenados, en contraparte a con robo < $1000. De manera análoga para robo abajo < $1000 las mujeres son mayoritariamente sentenciadas mientras que ocurre lo contrario para las mujeres son condenadas cuando se les enfrenta por "otros" tipos de cargos.

Estas dos gráficas dan una idea general de cómo es la relación entre las categorías de los factores.

Haciendo uso del análisis de correspondencias (AFC) se tiene lo siguiente:

Tabla 3.3 Valores propios y porcentaje de la inercia.

	Valores propios		
	1	2	3
Valor	0.1419	0.0329	0.0012
% de la inercia	81%	19%	1%
% acumulado	81%	99%	100%

De la tabla 3.3 se tiene que los tres primeros ejes factoriales recogen el 100% de la inercia total, y si nos fijamos en los dos primeros ejes estos recogen el 99%. Por lo que de acuerdo a lo establecido en el capítulo II, la gráfica con los dos primeros ejes explica el 99% de la dispersión de los datos, la cual está repartida en el primer eje con un 81% y con el segundo eje con un 19%.

La gráfica del AFC es la siguiente:

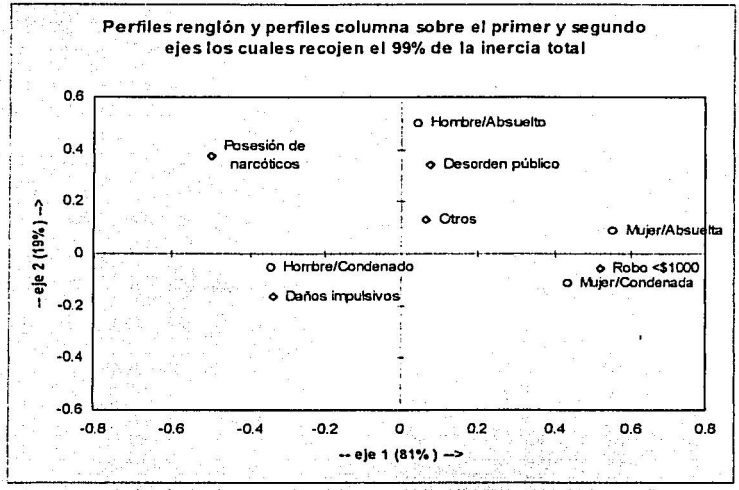


Fig 3.3

Recordemos que la interpretación de la gráfica se hace a partir de la posición de los perfiles en los ejes y la interpretación de los ejes tiene que ver con qué tan dispersos están los perfiles sobre éste.

Por lo que la interpretación de la gráfica puede ser la siguiente:

1) Tomando en cuenta el primer eje:

a) Considerando los perfiles por renglón se observan dos grupos: por una parte los hombres que son condenados y por otro lado las mujeres para las cuales no existe gran diferencia entre ser condenadas o ser absueltas.

b) En cuanto a los perfiles por columna se observa también que existen dos grupos que contrastan de manera fuerte, por un lado los cargos de posesión de narcóticos y daños impulsivos y por el otro el robo < $1000.

2) Tomando en cuenta el segundo eje:

a') Se observa para los hombres que son condenados, las mujeres que son condenadas y las

mujeres que son absueltas, se encuentran muy cerca del centroide (origen), mientras que los hombres que son absueltos presentan un alejamiento considerable del centroide.

b') Se observa que robo abajo <$1000 y otros se encuentran muy cerca del centroide, mientras que posesión de narcóticos y desorden público .

3) Tomando en cuenta los perfiles por renglón y por columnas de manera simultánea, se observa una relación entre los hombres que son condenados y los cargos de daños impulsivos y posesión de narcóticos en el sentido de alejamiento del origen, esta relación está en contraste con las mujeres que son condenadas por robo < $1000 y que también son sentenciadas por el mismo cargo. Con respecto al segundo eje parece que los hombres que son sentenciados están relacionados con posesión de narcóticos y daños públicos en contraste con daños impulsivos.

Lo anterior sugiere lo siguiente: los hombres tienden a tener una fuerte asociación en las tasas de condena por daños impulsivos, mientras que las mujeres están relacionadas con el robo <$1000.

Esta interpretación debe de ser sostenida por las contribuciones y la calidad de la representación de los perfiles.

Tabla 3.4 Componentes de los perfiles renglón sobre los ejes factoriales

				Coordenadas en los ejes		
Perfiles renglón	Peso	Inercia	% inercia	1	2	3
Hombre/Absuelto	0.1067	0.0273	0.1554	0.0443	0.5035	-0.0278
Mujer/Absuelta	0.0744	0.0243	0.1383	0.5541	0.0892	0.1097
Hombre/Condenado	0.5161	0.0627	0.3561	-0.3447	-0.0503	0.0053
Mujer/Condenada	0.3027	0.0616	0.3503	0.4359	-0.1137	-0.0262

De la tabla 3.4 se observa que los perfiles Hombre/condenado y mujer/condenada, tienen mayor peso y además esos perfiles aportan la mayor cantidad de inercia.

Tabla 3.5 Componentes de los perfiles columna sobre los ejes factoriales

				Coordenadas en los ejes		
Perfiles columna	Peso	Inercia	% inercia	1	2	3
Daños impulsivos	0.3722	0.0528	0.2989	-0.3384	-0.1634	0.0082
Robo <$1000	0.2854	0.0787	0.4473	0.5219	-0.0582	0.0041
Desorden público	0.0620	0.0083	0.0471	0.0751	0.3420	0.1053
Posesión de narcót	0.0819	0.0017	0.1802	-0.4982	0.3728	-0.0138
Otros	0.1985	0.0047	0.0265	0.0663	0.1293	-0.0485

Para la tabla 3.5 se observa que los perfiles que más peso tienen son Daños impulsivos, Robo<$1000 y otros. Aquí mientras Daños impulsivos , Robo<$1000 aportan una gran cantidad de inercia, el perfil Otros no aporta mucho.

Ahora analicemos las contribuciones que cada perfil aportan a cada eje, recuérdese que el primer eje aportaba la mayor cantidad de inercia y es interesante saber como está distribuida esa inercia entre los perfiles.

Tabla 3.6 Contribuciones (en porcentaje) de los perfiles renglón a la inercia asociada a cada eje factorial

	Ejes		
Perfiles renglón	1	2	3
Hombre/Absuelto	0.0015	0.8232	0.0686
Mujer/Absuelta	0.1610	0.0180	0.7465
Hombre/Condenad	0.4322	0.0397	0.0120
Mujer/Condenada	0.4053	0.1191	0.1729

Tabla 3.7 Contribuciones (en porcentaje) de los perfiles columna a la inercia asociada a cada eje factorial

	Ejes		
Perfiles columna	1	2	3
Daños impulsivos	0.3004	0.3024	0.0208
Robo <$1000	0.5478	0.0294	0.0041
Desorden público	0.0025	0.2209	0.5727
Posesión de narcót	0.1432	0.3464	0.0130
Otros	0.0061	0.1010	0.3894

De la tabla 3.6 para el primer eje principal se tiene que los perfiles Hombre/condenado y Mujeres/condenada son los que más aportan al primer eje (un 43.22% y un 40.53% respectivamente), y además son los que determinan su orientación. Para el segundo eje se tiene Hombre/ absuelto (82.32%) es el perfil que más contribuye y por lo mismo el que determina la orientación del segundo eje (Véase Fig 3.3).

De manera análoga para la tabla 3.7 en el primer eje se tiene que Daños impulsivos, Robo< $1000 contribuyen con mayor cantidad de inercia (30.04% y 54.78% respectivamente), mientras que para el segundo eje se tiene que Daños impulsivos, Posesión de narcóticos y Desorden público (30.24%, 22.09% y 34.64% respectivamente) son los que contribuyen a la orientación del eje. (Véase Fig 3.3).

Ahora bien, la calidad de representación de los perfiles en la gráfica es determinada por los $\cos^2\theta$ o las correlaciones al cuadrado, recuérdese que si $\cos^2\theta$ es grande , entonces θ es pequeño y puede decirse que el perfil va en dirección del eje o que está correlacionado con el eje factorial, estos elementos se presentan en las siguientes tablas.

Tabla 3.8 Cosenos cuadradados de los ángulos de los perfiles renglón con respecto a los ejes factoriales

	Ejes		
Perfiles renglón	1	2	3
Hombre/Absuelto	0.0077	0.9893	0.0030
Mujer/Absuelta	0.9389	0.0243	0.0368
Hombre/Condenad	0.9790	0.0208	0.0002
Mujer/Condenada	0.9331	0.0635	0.0034

Tabla 3.9 Cosenos cuadradados de los ángulos de los perfiles columna con respecto a los ejes factoriales

	Ejes		
Perfiles columna	1	2	3
Daños impulsivos	0.8106	0.1889	0.0005
Robo <$1000	0.9877	0.0123	0.0001
Desorden público	0.0421	0.8750	0.0829
Posesión de narcót	0.6407	0.3588	0.0005
Otros	0.1873	0.7124	0.1003

A partir de la tabla 3.8 se puede decir que el primer eje factorial guarda una fuerte correlación con respecto a los perfiles Mujer/condenada, Hombre/condenado y Mujer/absuelta; mientras que el segundo eje se encuentra fuertemente correlacionado con el perfil Hombre/absuelto; esto corrobora lo que ya se había dicho en cuanto a la dirección de los ejes

Similarmente, para la tabla 3.9 se observa que Daños impulsivos, Robo< $1000 están altamente correlacionados con el primer eje factorial, mientras que Desorden público se encuentra muy correlacionado con el segundo eje.

Además, como se recordará en el capítulo II, la calidad de representación de cualquier perfil en la gráfica es la suma de las correlaciones al cuadrado, por lo que se puede decir que la calidad de representación de cualquier perfil tanto para los renglones como para las columnas es muy buena con los dos primeros ejes.

Este ejemplo muestra una forma particular de hacer un estudio de la relación concatenando dos factores, sin embargo el problema que ahora surge es el siguiente, ¿por qué no concatenar, por ejemplo, el sexo de la persona y el cargo penal y estudiar la asociación con el factor veredicto del juez, o concatenar el veredicto del juez con el cargo penal y estudiar la asociación con respecto al sexo de la persona?. Este tipo de estudios nos llevaría al estudio gráfico de la asociación condicional y parcial, pero que pasaría con la asociación total, en la que se utilicen de manera simultánea los tres factores.

Una forma alternativa para llevar a cabo el estudio de la asociación en el que se involucren los tres factores de manera simultánea es mediante el uso del análisis de correspondencias múltiples.

Del cual haremos su introduccción a continuación:

Utilización del análisis de correspondencias multiples (AFCM)

El objetivo es determinar si existe asociación o no entre las categorías de los tres factores, sexo, resultado de la sentencia y el cargo por el que se acusa.

Para iniciar la interpretación del análisis de correpondencias multiple se hacen las siguientes interrogantes que se tratarán de contestar mediante el análisis:

1. ¿Quiénes son más sentenciados los hombres o las mujeres?.

2. Para cada tipo de delito,¿ son igualmente sentenciados hombres y mujeres?.

3. ¿Qué delito es más sentenciado?

El procedimiento para la interpretación del Análisis Factorial de Correspondencias Multiples (AFCM).principalemente considera:

1. El número de ejes que se utilicen para la construcción del gráfico.

De acuerdo a lo establecido hasta el momento se tiene que el número total de ejes que recuperan el 100% de la estructura de la información es 6, por lo que debemos escoger entre estos seis ejes aquellos que recuperen la mayor cantidad de inercia. De manera análoga a lo establecido en el análisis de correspondencias simple se escogen aquellos ejes que conserven la mayor cantidad de inercia y además se considera que las categorías de los factores estén bien representadas en la gráfica. Así, para hacer la interpretación utilizaremos el subespacio generado por un solo eje (una linea recta) tomando en cuenta la proporción de inercia que se representa en los eigenvalores.

Valeurs pr.	1	2	3	4	5	6
Valeur	0.4626	0.3858	0.3333	0.3333	0.2745	0.2105
% de vari.	23%	19%	17%	17%	14%	11%
% cumulé	23%	42%	59%	76%	89%	100%

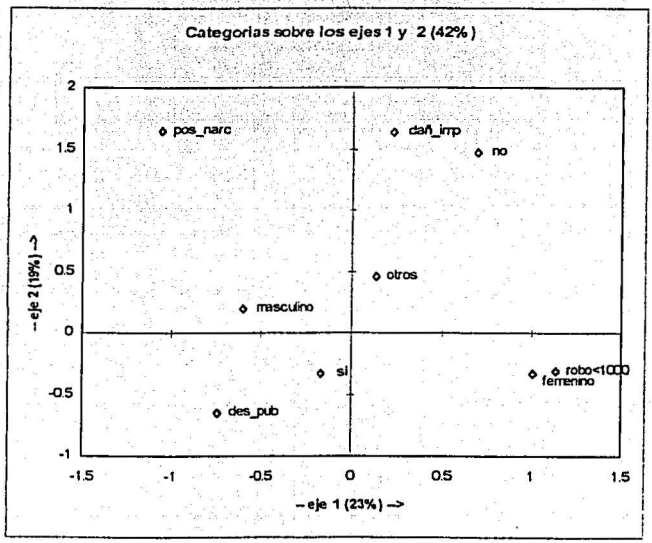
El subespacio generado por el primer eje recupera el 23% de la inercia total, el segundo el 19% y así sucesivamente. Considerar cualquiera de estos seis posibles subespacios no ayuda a establecer la relación de asociación que pudiera existir debido a la poca información que recogen los ejes.

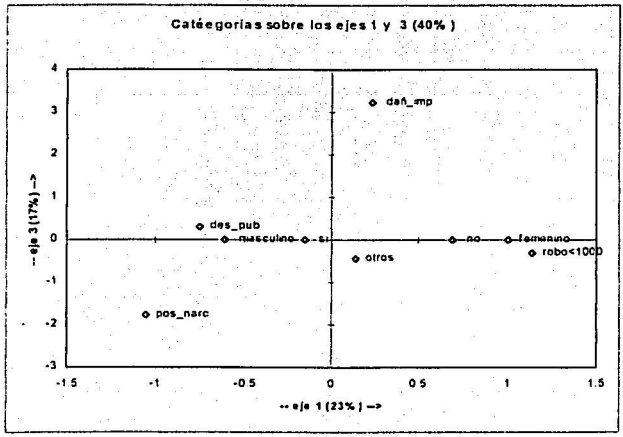
Si consideramos como el mejor subespacio el formado ahora por los pares de ejes, tendríamos los siguientes:

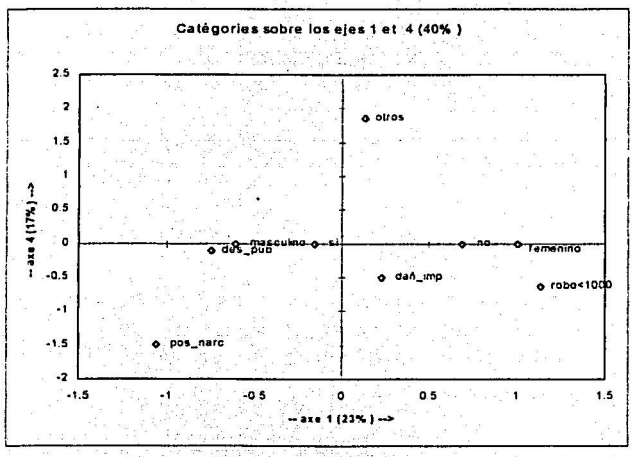
Gráfica	% de la inercia
eje 1 Vs eje 2	42*
eje1 Vs eje 3	40*
eje1 Vs eje 4	40*
eje 1 Vs eje 5	37
eje 1 Vs eje 6	34
eje 2 Vs eje 3	36
eje 2 Vs eje 4	36

eje 2 Vs eje 5	33
eje 2 Vs eje 6	30
eje 3 Vs eje 4	34
eje 3 Vs eje 5	31
eje 3 Vs eje 6	28
eje 4 Vs eje 5	31
eje 4 Vs eje 6	28
eje 5 Vs eje 6	25

De tal manera que si queremos conservar la mayor cantidad de inercia escogeriamos la interpretación de las gráficas: eje1 Vs eje 2,eje1 Vs eje 3, eje1 Vs eje 4.







Primera conclusión a partir de los gráficos.

1. Las mujeres están asociadas con los delitos de robo<1000. Este delito no es penalizado.

2. Los hombres están asociado con los delitos de posesión de narcótico y con daños impulsivos. Estos delitos **son penalizados**.

Si empleamos las restricciones debidas a la calidad de la representación de las categorias debemos tener en cuenta lo siguiente: la representación gráfica debe ser nitida, como una foto: para ello se requiere buena cantidad de inercia explicada, buena calidad de representación de la foto, buena correlación de los puntos con los ejes; se requieren que en el peor de los casos se pueda consultar simultáneamente todas las fotos, aunque ello propicie una actitud anti_parsimoniosa dentro del análisis de datos.

En base a lo anterior y consultando la salida del paquete Xlstat V4. para los datos de este problema tenemos lo siguiente Véase anexo 6.

a) Ninguna de las gráficas con dos ejes explica más del 42%. Por lo que la cantidad de información que se pierde es de más de la mitad.

b) La calidad de representación utilizando los dos primeros ejes (aquellos que nos proporcionan una mayor cantidad de inercia) es a lo más del 62% por lo que no podemos hablar de una buena representación con estos dos ejes. Si tomamos el primer y tercer ejes (40% de la inercia total) no hay una mejora, únicamente la categoria de daños impulsivos aumenta a un 68% de calidad en la representación. Utilizando el primer y cuarto ejes (40 % de la inercia total) tampoco hay un aumento en la calidad de la representación unicamente la categoría cargo "otros" aumenta a un 86%. En base a estos resultados podemos concluir que con una gráfica formada por pares de ejes la calidad de representación no es adecuada y por lo tanto no podemos decir todavía nada de manera segura.

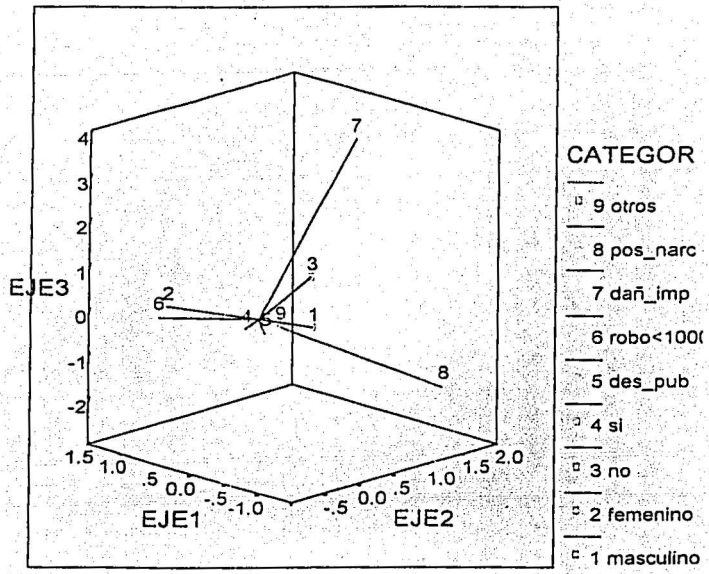
Es importante notar en este momento que la representación de la asociación será más sencilla gráficamente cuando en realidad ésta se encuentre bien estructurada, cuando no es así entonces se tiene que hacer una intensa búsqueda de con cuales y cuántos ejes se puede representar la asociación y que las categorías de los factores quedan bien representadas, considerando al mismo tiempo que la inercia recuperada sea alta.

Una alternativa sería estudiar la representación con los tres primeros ejes que corresponde a 59% de inercia recuperada y en cuanto a la calidad por arriba de 59% para todas las categorias, con excepción de el cargo de otros, como puede verse en el apéndice relativo a la salida del paquete Xlstat V4.

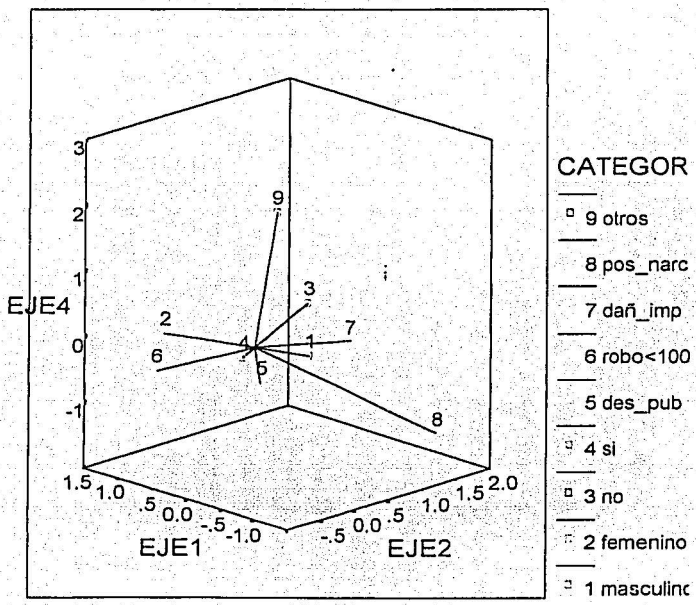
Ahora bien, si aumentamos una dimensión a la gráfica tenemos las siguientes posibles gráficas.

Gráfica	% de inercia
eje1 Vs eje2 Vs eje 3	59*
eje1 Vs eje2 Vs eje 4	59*
eje1 Vs eje2 Vs eje 5	56
eje1 Vs eje2 Vs eje 6	53
eje1 Vs eje3 Vs eje 4	57*
eje1 Vs eje3 Vs eje 5	54
eje1 Vs eje3 Vs eje 6	51
eje1 Vs eje4 Vs eje 5	54
eje1 Vs eje4 Vs eje 6	51
eje2 Vs eje3 Vs eje 4	53
eje2 Vs eje3 Vs eje 5	50
eje2 Vs eje3 Vs eje 6	47
eje2 Vs eje4 Vs eje 5	50
eje2 Vs eje4 Vs eje 6	47
eje2 Vs eje5 Vs eje 6	44
eje3 Vs eje4 Vs eje 5	48
eje3 Vs eje4 Vs eje 6	45
eje3 Vs eje5 Vs eje 6	42

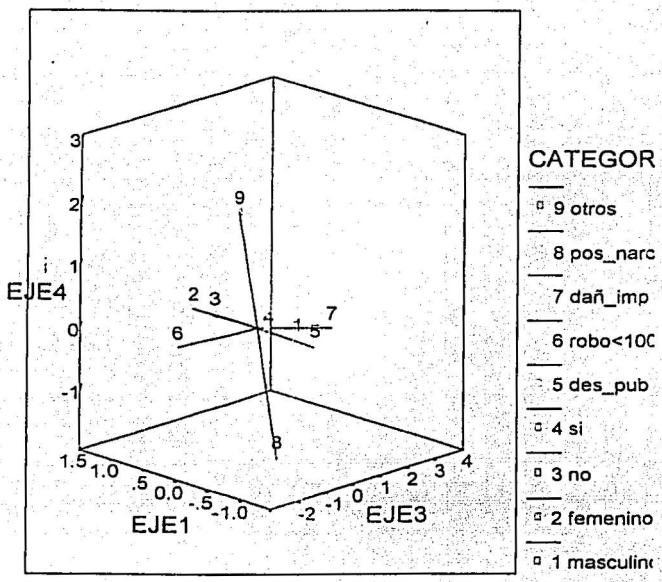
De manera análoga las gráficas que más inercia aportan son: eje1 Vs eje2 Vs eje3, eje1 Vs eje2 Vs eje4 y eje1 Vs eje3 Vs eje 4. Las cuales se presentan a continuación:



De esta gráfica que representa el 59% de la inercia total se puede notar que los hombres (1) están relacionados con daños impulsivos (7) y con posesión de narcóticos (8) y además esta asociación apunta en la dirección de no penalizados (3). A diferencia de las mujeres (2) que están asociadas con robo<1000 (6) y esta asociación está dada también con el hecho de que si son sentenciadas (4). Nótese que la calidad de representación es del 59% para todas las categorías a excepción del cargo otros.



De esta gráfica que también recoge el 59% de la inercia total se puede notar que los hombres (1) están relacionados con posesión de narcóticos (8) y no son sentenciados (3), nótese que aunque también parece que está asociado con daños impulsivos (7) la calidad de representación de esta categoría es del 20% lo cual es muy bajo. Por otro lado, las mujeres (2) están asociadas con robo<1000 (6) y además si son sentenciadas (4).



De esta gráfica lo que más se puede decir es que los daños impulsivos (7) son opuestos a robo<1000 (6) y además la relación con el sexo sigue siendo la misma que en las gráficas anteriores, sólo que ahora posesión de narcóticos (8) no tiene una buena calidad de representación 58% y esta gráfica no es informativa con respecto a la asocación con el sexo ya que para hombres (1) y mujeres (2) la calidad de representación es apenas del 11%.

Del ejemplo anterior se aprecia la dificultad para observar la estructura de la asociación de los datos, desde luego que no es sencilla la interprteación de una gráfica de mayor dimensión y que además tiene una calidad de representación modesta o baja . Este problema muy probablemente se deba a los datos. En estos casos se sugiere hacer uso de otras técnicas que pueden complementar las interprteaciones que se han hecho, por ejemplo,se puede hacer uso de los modelos log-lineales.

Bajo las consideraciones anteriores se tiene que: el delito más frecuente en las mujeres es el (robo<1000) y además es sancionado, a diferencia del daño impulsivo asociado con el hombre y que no es sancionado.

Ahora bien, veamos un ejemplo que se ha manejado mucho en la literatura estadística y que se refiere a las flores de Iris, trabajada por Fisher. Se reporta (Dallas E. Johnson) que aplicando la técnica del análisis discriminante a la base de datos compuesta por 50 observaciones para el tipo de iris setosa, 50 para el tipo de iris versicolor, 50 para el tio de iris virginica; con las siguientes variables:

1. Sepallen : longitud del sépalo.

2. Sepalwid: ancho del sépalo.

3. Petallen: longitud del pétalo.

4. Petalwid: ancho del pétalo.

Se obtiene tres grupos diferenciados para cada especie de flor de iris. En el análisis reportado el porcentaje de flores que fueron clasificadas correctamente fue es del 98%, y lo que se concluye es.

a) Las variables explicativas individualmente diferencían en promedio a las especies, por lo que son importantes para la clasificación de estas.

b) La variable que más aporta a la discriminación es el largo del pétalo.

c) Con las variables explicativas estudiadas la especie setosa se identifica claramente, pero las especies versicolor y virginica podrían confundirse.

Ante este conocimiento a priori, seleccionamos dos de las tres variedades de iris: la setosa y la virginica ya que son los grupos que más se diferencian entre si. Lo que nos interesa conocer es si existe alguna característica que determine a que variedad pertenece, es decir, si por ejemplo medidas pequeñas en las variables están relacionadas con alguna especie en particular.

Para llevar a cabo lo anterior, dado que la escala de medida es de razón, se propusieron las siguientes categorías, que se formaron de la siguiente manera: 1) se obtuvieron los valores mínimo y máximo de las variables estudiadas. 2) como se deseaba que fueran tres categorías: corto, mediano, largo, entonces se tomo $\Delta E=\frac{max-min}{3}$ como la longitud del intervalo, y así las categorías fueron:

corto: (min,min+ΔE)

mediano: (min+ΔE,min+2ΔE)

largo:(min+2ΔE,max)

Con estas categorías tenemos una medida cualitativa que nos relaciona las longitudes de las variables.

La nueva base de datos se presenta en el anexo 7, y a ésta se le aplicó un análisis de correspondencias multiples, para las variables: variedad de flor, sepallen, sepalwid, petallen, petalwid. Por lo que en total se tienen catorce categorías y cinco factores.

Los resultados se muestran a continuación:

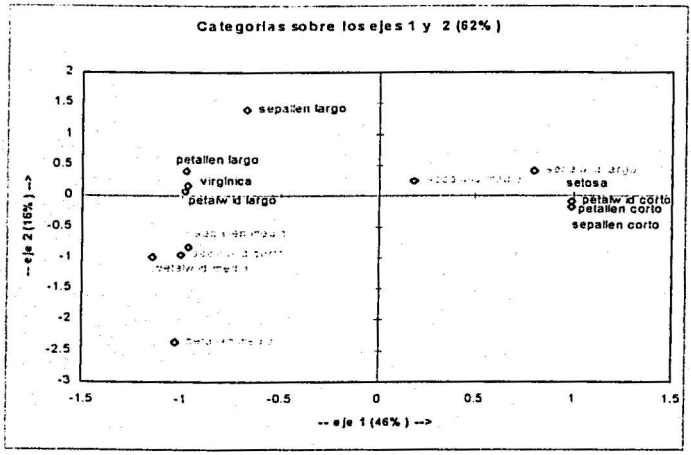
1. El número de ejes que se tienen es: 14-5=9.

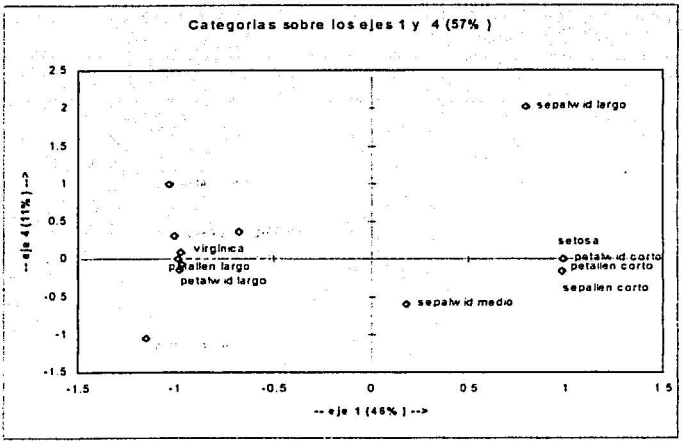
2. Dentro de esos 9 ejes debemos escoger aquellos que conserven la mayor cantidad de inercia

Valeurs pr.	1	2	3	4	5	6	7	8	9
Valeur	0.8223	0.2653	0.2227	0.2006	0.1343	0.1053	0.0275	0.0000	0.0000
% de var.	46%	16%	12%	11%	7%	6%	2%	0%	0%
% cumulé	46%	62%	74%	86%	93%	98%	100%	100%	100%

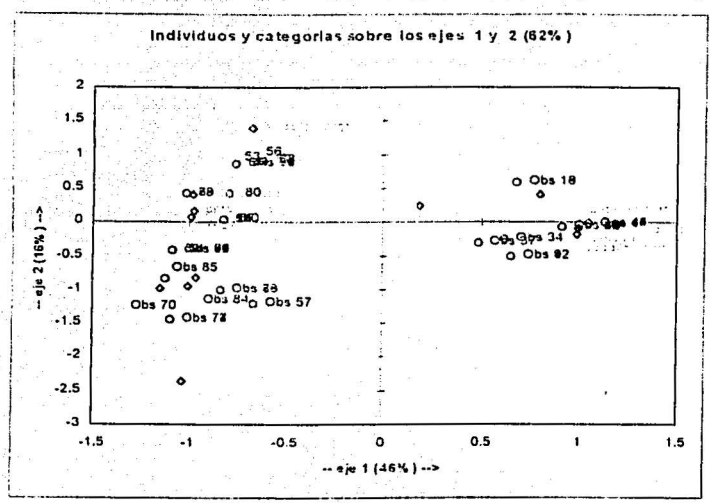
de la tabla anterior se observa que el primer eje recoge el 48% de la inercia total, el segundo el 16% y el tercero y cuarto el 12% y 11% respectivamente. Lo anterior sugiere que si se está captando la estructura de asociación en los datos y además de una manera fuerte, con los tres primero ejes se recoge casi el 75 % de la inercia total.

3. La gráfica de correspondencias con los dos primero ejes es:



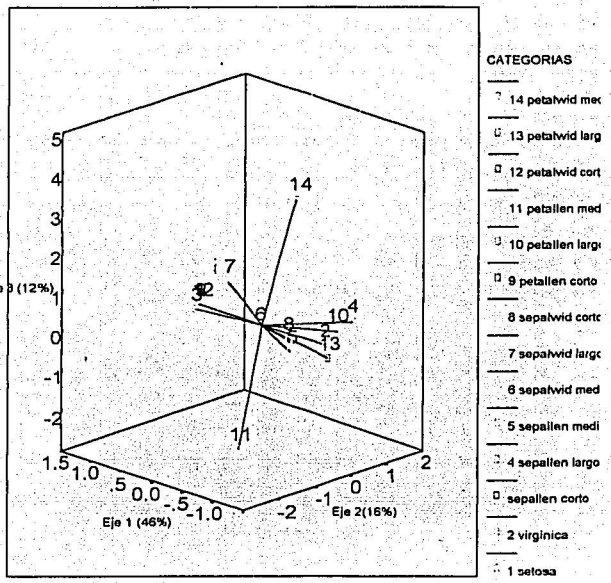


De las gráficas anteriores se concluye que las variedades de iris son opuestas totalmente, esto muestra como se sabía una clara diferenciación entre estas variedades, ahora bien, otro punto interesante era determinar que características se relacionaban con estas variedades. Así, de las gráficas se concluye que para el caso de la iris virginica la asociación está mayoritariamente dada por longitudes grandes tanto del sépalo como del pétalo y una anchura en el pétalo grande. Pero para el caso de la iris setosa se tiene que está asociada mayoritariamente a longitudes pequeñas tanto del pétalo, el sépalo y el ancho del pétalo. Hay que hacer notar que en las gráficas anteriores las categorías en rojo representan aquellas cuya calidad de representación era baja (abajo del 50%).



Si quisieramos hacer un análisis sobre los individuos entonces vemos que estos se encuentran también en dos grupos bastante definidos y además como era de esperarse, las observaciones del 1 al 50 se encuentran relacionadas con la variedad de iris setosa, y los datos del 50 al 100 se encuentran relacionados con la variedad virginica, además las calidades de representación de los individuos es bastante buena considerando los dos primero ejes. Veáse datos anexo 7.

A pesar de que con estas gráficas se tiene una corroboración acerca del comportamiento que tienen los datos obtenido a través del análisis discriminante, si seguimos con la idea de recuperar la mayor cantidad de información posible entonces agreguemos una dimensión más.



Se observa que efectivamente la iris setosa (1) es opuesta a la iris virginica (2) ambas con una calidad de representación del 99%. Además se aprecia que la iris virginica está asociada a petallen largo, (10) petalwid largo(13) y sepallen largo (4) tal como se había concluido en las gráficas anteriores. Por otro lado la iris setosa está asociada a petalwid corto (12), petallen medio (11), sepallen corto (3). Lo cual concuerda con lo establecido en las conclusiones de las gráficas con dos ejes. Sin embargo, con esta tres dimensiones se recupera el 75% de la inercia total, lo cual es muy bueno y se garantiza una calidad de representación buena para las categorías.

Así, podemos concluir con base al análisis que medidas cortas en el largo y ancho del pétalo y del sépalo están asociadas fuertemente con la iris setosa, mientras que medidas grandes de estas variables están asociadas a las iris virginica.

Como conclusión para este capítulo tenemos

1) En el capítulo anterior se mostró que existia una relación entre la cantidad total de inercia y

el coeficiente de contengencia en media cuadrática ϕ^2 que nos permitia interpretar para cada gráfica el grado de asociación entre las categorías; en el AFC-multiple no existe esta relación, por lo que la interpretación de la asociación se vuelve más heuristica al momento de interpretar la gráfica, por lo que esta interpretación debe estar basada en buena medida con el conocimiento que se tenga de la posible asociación.

2) A veces es posible reducir la dimensión de la tabla de contingencia para más de dos factores y convertirla en una tabla de contingencia de dos dimensiones por medio de la concatenación de algunas categorías o a través de procedimientos acorde a los objetivos del estudio y que buscan generar nuevas variables para trabajar con ellas.

3) A veces resulta difícil determinar la estructura de asociación entre las categorías de los factores estudiados a través del análisis de correspondencias múltiple, sin embargo ante estas problemáticas se recomienda recurrir a otras técnicas estadísticas que ayuden a interprtear los resultados.

4) Cuando la estructura de asociación en los datos es clara, entonces el análisis de correspondencias resulta muy útil, ya que presenta en una gráfica esa relación tomando en cuenta los elementos que permiten determinar la calidad de las gráficas.

Conclusiones.

La posible asociación entre las categorías de los factores puede ser estudiada a través de un AFC-S o un AFC-M según sea el caso, mediante representaciones gráficas que pueden ser presentadas como una línea recta, un plano, un plano en tres dimensiones o un hiperplano. Como una similitud podemos decir que estas representaciones son semejantes a un conjunto de "fotografías", en las que con un grado de nitidez aceptable en los elementos que la conforman se puede estudiar la posible asociación entre las categorías de los factores de la tabla de contingencia ya sea entre las categorías de los factores o bien por categorías entre factores;dicha asociación puede establecerse en la medida que:

i) Los perfiles de las categorías se alejen del origen (centroide) y muestren un proximidad entre sí, si esto ocurre se podrá decir que existe una asociación fuerte.

ii) La cantidad de inercia que sea recogida por los ejes que definen la(s) representación(es) gráfica(s) sea alta. A medidad que se aumenta el numero de dimensiones en la gráfica aumenta la cantidad de inercia representada.

iii) La calidad de representación de los perfiles dentro de la(s) representacion(es) gráfica(s) sea buena, lo que da la confianza para poder hablar de la posible asociación.

Para el caso del AFC-S, se observó que cuando existen pocas categorías la cantidad de inercia que recogen las gráficas es alta, sin embargo en la práctica puede suceder que el número de categorías de los factores en la tabla de contingencia sean muchos por lo que también el número de ejes lo serán , decidir cuales representaciones se conservarán para el análisis requerirá un conocimeinto muy particular acerca de la asociación que se quiere estudiar ya que la interpretación va más allá de las consideraciones geométricas por que puede darse el caso de que las gráficas seleccionadas conserven una cantidad de inercia alta pero tengan una calidad de representación baja o al contrario la cantidad de representación sea alta pero la cantidad de inercia sea baja. Así, se sugiere que se conserven los ejes que cumplan con: cantidad de inercia alta y calidad de representación alta; considerándose las posibles combinaciones de gráficas de dos dimensiones o las posibles combinaciones de gráficas en tres dimensiones etc.

Para el caso del AFC-M con la introducción de la matriz indicadora Z, el análisis presenta el problema de que la cantidad de inercia explicada disminuye, por lo que hacer una interpretación semejante al AFC-S resulta complicado debido a que la inercia se distribuye en un mayor número de ejes lo que ocasiona que la interpretación esté basada en gráficas con mayor dimensión o bien la interpretación de gráficas de dos dimensiones que recogen poca inercia; esto evidencia la necesidad de contar con un conocimiento general del contexto del problema para el que se utiliza el AFC.

Por otra parte puede darse el caso en que los datos presenten una estructura interna de asociación marcada por ejemplo en los datos para el estudio de las "iris de fisher". En el cual sin bien en el AFC-M no se recogía una cantidad de inercia muy alta con los tres primeros ejes, la interpretación esta de acuerdo a lo reportado en la literatura estadística.

Ante estos problemas de interpretación del análisis para tablas de contingencia con muchas categorías o con más de dos factores se han propuesto para el primer caso reducir el número de categorías colapsando éstas de tal manera que se reduzca el número de ejes en la gráfica y para el segundo caso reducir la dimensión de una tabla de contingencia a una de dos dimensiones mediante la concatenación de las categorías de un factor sobre otro, de tal manera que se pueda aplicar un AFC-S. Esto desde luego lleva el riesgo en la metodología que se utilice para abordar el problema bajo estudio, ya que colapsar categorías o concatenar factores puede ser incongruente con la investigación en que se utilice el AFC.

Ante esta situación se recomienda un estudio a cerca de la metodología del uso del AFC-M y la utilización de técnicas complementarias como pueden ser los modelos log-lineales.

Bibliografía

1. J.P. Benzecri. *Correspondence Analysis Handbook.* Marcel Dekker, Inc, New York, 1992.

2. Ludovic Lebart, Alain Marineau, kenneth M Warwick. *Multivariate Descriptive Analysis* (*Correspondence Analysis and Related Techniques For Largue Matrices*), John Wiley & Sons , USA, 1984.

3. Michael J. Greenacre. *Theory and Applications of Correspondence Analysis,* Academic Press, London, 1984.

4. J.M. Cornejo.*Técnicas de Investigación social. El Análisis de Correspondencias* (*Teoria y Práctica*). Biblioteca Universitaria de Ciencias Sociales, Serie Medium, Barcelona España, 1988.

5. Brian.Sidney Everit. *The Analysis of Contingence Tables.* London, Chapman and Hall, 1977.

6. Francisco Casanova del Ángel. *These de doctorat de 3e Cycle. Mathematiques; Traitement des donnees statistiques.* Paris, 1980.

7. Rubén Hernández. *Apuntes: Análisis de datos.*

8. K.V. Mardia. *Multivariate Analysis* . Academic Press, USA, 1979.

9. Alvin C. Rencher. *Metods of Multivariate Analysis,* Wiley Series in Probability and Mathematical Statistics, USA, 1995.

10. Bryan F.J. Manly. *Multivariate Satistical Methods. A Primer.* Chapman and Hall, Great Britain, 1990.

11. Eduardo Solar González, *Apuntes de Álgebra Lineal,* UNAM, 1985

12. Brian,S. Everitt and Graham Dunn. *Applied Multivariate Data Analysis.* Ed. Edward Arnold A. division of Hodder &Stoughton, Londres 1991.

13. Journal of Econometrics 22 (1983) 139-167. North-Holland Publishing Company*Analyzing rectangular tables by joint and constrained multidimensional scaling.* Willenm.J.Heiser and Jaqueline Meulman.

14. Alan,Agresti. *Analysis of Ordinal Categorical Data.* Wiley Series in Probability and Mathematical Statistics, USA, 1984

15. *Interpreting Multivariate Data.* Edited by Vic Barnett.Wiley Series in probability and mathematical statistics. 1981.

16. *A users guide of principal Components Analysis*.Wiley Series in Probability and Mathematical Statistics, USA

18. W.J Conover. *Practical Nonparametric Statistics*.John Wiley & Sons, Inc. 3rd. Edition, USA, 1999.

19. Visauta Vinacua Bienvenido. *Análisis Estadistico con SPSS para Windows. Estadistica Multivariante*. Mc Graw Hill, España, 1999.

20. J.D.Jobson. *Applied Multivariate Data Analysis*.Vol II. ,Springler Verlag.

21. Luis Joaristi Olariaga, Luis Lizasoain Hérnández, Análisis de correspondencias. Cuadernos de esttadistica, Madrid, España, Ed. La muralla, 1999.

22 Dallas E. Johnson, *Métodos multivariados aplicados al análisis de datos*, México, Ed. International Thomson,2000.

23. Lara Pérez Soto, Claudia. *Análisis de Correspondencias Múltiples como una técnica para el estudio de datos cualitativos*.Tesis de licenciatura,Facultad de Ciencias, UNAM, 1990.

24. Humberto Soto de la Rosa, *Algunas Técnicas para el análisis de datos categóricos*, Tesis de licenciatura, Instituto Tecnológico Autónomo de México.1999.

25. Nuria Guerra Vargas, *Análisis de correspondencias simples y múltiples (tres enfoques teóricos, aplicaciones y programación*), Tesis de licenciatura, Instituto Tecnológico Autónomo de México.1996.

Anexos

Anexo 1. Resultados de álgebra lineal y álgebra de matrices

1. Espacio Vectorial

Sea V un conjunto no vacío y sea $(K,+,\bullet)$ un campo. Se dice que V es un espacio vectorial sobre K si están definidas dos leyes de composición llamadas adición y multiplicación por un escalar, tal que:

a) La adición asigna a cada pareja ordenada $(\bar{u},\bar{v})$ de elementos de V un único elemento $(\bar{u}+\bar{v}) \in V$, llamado la suma de V.

b) $\forall\ \bar{u},\bar{v},\bar{w} \in V$ se tiene que $\bar{u}+(\bar{v}+\bar{w}) = (\bar{u}+\bar{v})+\bar{w}$

c) $\exists\ \bar{0} \in V$ tal que $\bar{0}+\bar{v} = \bar{v}$, $\forall\ \bar{v} \in V$.

d) $\forall\ \bar{v} \in V\ \exists\ -\bar{v} \in V$ tal que $-\bar{v}+\bar{v} = \bar{0}$

e) $\forall\ \bar{u},\bar{v} \in V\ \ \bar{u}+\bar{v} = \bar{v}+\bar{u}$

f) La mltiplicación por un escalar asigna a cada pareja ordenada $(\alpha,\bar{v})$ de lelementos $\alpha \in K$ y $\bar{v} \in V$ un único elemento $\alpha\bar{v} \in V$, llamado producto de α por $\bar{v}$.

g) $\forall\ \alpha \in K$; $\bar{u},\bar{v} \in V$ se tiene que $\alpha(\bar{v}+\bar{w}) = \alpha\bar{u}+\alpha\bar{w}$

h) $\forall\ \alpha,\beta \in K$; $\bar{v} \in V$ se tiene que $(\alpha+\beta)\bar{v} = \alpha\bar{v}+\beta\bar{v}$

i) $\forall\ \alpha,\beta \in K$; $\bar{v} \in V$ se tiene que $\alpha(\beta\bar{v}) = (\alpha\beta)\bar{v}$

j) Si 1 es la unidad en K se tiene que $\forall\ \bar{v} \in V$, $1\bar{v} = \bar{v}$

A los elementos de V se les llama vectores y a los elementos de K se les llama escalares.

2. Combinación lineal

Un vector $\bar{w}$ es una combinación lineal de los vectores $\bar{v}_1,\bar{v}_2,....,\bar{v}_n$ si puede ser expresado en la forma:

$$\bar{w} = \alpha_1\bar{v}_1+\alpha_2\bar{v}_2+....+\alpha_n\bar{v}_n$$

donde $\alpha_1,\alpha_2,...,\alpha_n$ son escalares.

3 Vectores linealmente independientes y vectores linealmente dependientes

Sea $S = \{\bar{v}_1,\bar{v}_2,....,\bar{v}_n\}$ un conjunto de vectores:

i) S es linealmente dependiente si existen escalares $\alpha_1,\alpha_2,...,\alpha_n$, no todos iguales a cero, tales

que

$$\bar{0} = \alpha_1\bar{v}_1 + \alpha_2\bar{v}_2 + ... + \alpha_n\bar{v}_n$$

ii) S es linealmente independiente si

$$\bar{0} = \alpha_1\bar{v}_1 + \alpha_2\bar{v}_2 + ... + \alpha_n\bar{v}_n$$

$$\text{con } \alpha_1 = \alpha_2 = ... = \alpha_n = 0.$$

4. Conjunto generador.

Sea V un espacio vectorial sobre K, y sea $G = \{\bar{v}_1, \bar{v}_2, ..., \bar{v}_m\}$ un conjunto de vectores de V.

Se dice que G es un generador de V si para todo vector $\bar{x} \in V$ existen escalares $\alpha_1, \alpha_2, ..., \alpha_m$ tales que $\bar{x} = \alpha_1\bar{v}_1 + \alpha_2\bar{v}_2 + ... + \alpha_m\bar{v}_m$. es decir todos los vectores de V son combinaciones lineales de los elementos

de G.

5. Base de un espacio vectorial V

Se llama base de un espacio vectorial V a un conjunto generador de V que es linealmente

independiente.

5. Dimensión de un espacio vectorial

Sea V un espacio vectorial sobre K. Si $B = \{\bar{v}_1, \bar{v}_2, ..., \bar{v}_n\}$ es una base de V se dice que V es de dimensión n, lo cual se denota con

$$dim\ V = n$$

En particular, si $V = \{\bar{0}\}$, $\qquad dim\ V = 0$

6. Transformaciones :

Entenderemos por transformaciones lineales a las funciones vectoriales de variable vectorial, es decir funciones del tipo $\bar{w} = f(\bar{v})$. En erminos formales tenemos: Si V y W son espacios vectoriales una función $T : V \longrightarrow W$ recibe el nombre de transformación . Los espacios V y W se llaman, respectivamente dominio y codominio de la transformación.

Operaciones elementales de matrices.

7. Para multiplicar dos matrices, el niumero de columnas en la primera debe ser igual al número de renglones en la segunda . De ese modo, la multiplicación de matrices está dada de la siguiente manera:

$$A_{m\times n}B_{n\times p} = C_{n\times p}$$

8. Definición de Producto:

Sean $A = [a_{ij}]$ y $B = [b_{ij}]$ dos matrices con elementos en C (numeros complejos), de $m \times n$ y $n \times q$ respectivamente . El producto AB es una matriz $P = [p_{ij}]$, de $m \times q$, definida por:

$p_{ij} = \sum_{k=1}^{n} a_{ik}b_{kj}$; para $i = 1,2,3,...,m$ y $j = 1,2,3,...,q$.

9.Cuando todos los productos de matrices están definidos, se cumple la ley asociativa para la multiplicación de las mismas. Por Tanto $A(BC) = (AB)C$.

10. La multiplicación de matrices no siempre es conmutativa. Es decir AB no siempre es igual a BA.

11. Definición (Suma)

Sean $A = [a_{ij}]$ y $B = [b_{ij}]$ dos matrices de $m \times n$ con elementos en C (numeros complejos). La suma de $A + B$ es una matriz $S = [s_{ij}]$, de $m \times n$, definida por

$s_{ij} = a_{ij} + b_{ij}$, para $i = 1,2,3,...,m$ y $j = 1,2,3,...,n$.

12. Teorema.

Si A, B son matrices de $m \times n$ cuyos elementos son números complejos, y sea $\alpha \in R$ entonces,

i). Existe una matriz $-A$ de $m \times n$ tal que $A + (-A) = 0$

ii). $\alpha(A + B) = \alpha A + \alpha B$

Álgebra de matrices.

13. Si

$$A = \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1p} \\ a_{21} & a_{22} & \ldots & a_{2p} \\ \ldots & \ldots & \ldots & \ldots \\ a_{n1} & a_{n2} & \ldots & a_{np} \end{bmatrix}$$

entonces A^t se conoce como la transpuesta de la matriz A. Ésta se define por

$$A^t = \begin{bmatrix} a_{11} & a_{21} & \ldots & a_{n1} \\ a_{12} & a_{22} & \ldots & a_{n2} \\ \ldots & \ldots & \ldots & \ldots \\ a_{1p} & a_{2p} & \ldots & a_{np} \end{bmatrix}$$

14. $(AB)' = B'A'$

15. Se dice que una matriz A es simétrica si $A = A'$

16. La traza de una matriz cuadrada $A_{p\times p}$, se define como la suma de sus elementos en la diagonal. Se escribe:

$$tr(A) = \sum_{i=1}^{p} a_{ii}$$

17. Se dice que una matriz cuadrada es una matriz diagonal si todos sus elemntos que no están en la diagonal son ceros.

18. El rango de una matriz se define como el número máximo de renglones (columnas) en A que son linealmente independientes. De modo equivalente , el rango de A es la dimensión del subespacio vectorial generado por los renglones (columnas) de la matriz A.

19. Proposiciones:

a) El Rango de $A_{m\times n} \leq min(m.n)$

b) Rango de A =rango de A'

20. Dos vectores X y Y , son ortogonales si $XY' = 0$. Además se dice que dos vetores ortogonales son ortonormales si $X'X = 1$ y $Y'Y = 1$.

21. Se dice que un vector X está normalizado si $X'X = 1$

Formas cuadráticas

22. Una forma cuadrática en p variables $x_1, x_2, ..., x_p$, es una función de la forma $\sum_{i=1}^{p} \sum_{j=1}^{p} a_{ij}x_ix_j$ .Una forma cuadrática de este tipo siempre se puede escribir como:

$X'AX$, para alguna matriz simétrica $A_{p\times p}$, en donde $X = \begin{bmatrix} x_1 \\ x_2 \\ \cdots \\ x_p \end{bmatrix}$

23. se dice que una matriz simétrica A es

a) Positiva definida si $X'AX > 0$, para toda $X \neq 0$.

b) Positiva semidefinida si $X'AX \geq 0$, para toda X y si $X'AX = 0$ para algún X diferente de cero.

Eigenvalores y Eigenvectores.

24.Los eigenvalores de una matriz simétrica $A_{p\times p}$, son las raíces de una ecuación polinomial dad por

$$|A - \lambda I| = 0,$$

la cual es una ecuación polinomial en λ, de p-ésimo grado

25. A cada eigenvalor de A le corresponde un vector diferente de cero llamado eigenvector, que satisface

$$Ac_i = \lambda_i c_i, \text{ para } i = 1, 2, 3, ..., p$$

26. Si A es una matriz simétrica de numeros reales , entonces sus eigenvalores y eigenvectores también consistirán en numeros reales.

27. Los eigenvectores no son únicos, de modo que a menudo se normalizan de tal manera que $c_i'c_i = 1$.

28. La traza de una matriz simétrica es igual a la suma de sus eigenvalores, es decir, $tr(A) = \sum_{i=1}^{p} \lambda_i$

Ángulos

29. El ángulo θ entre X y Y se expresa por:

$$\cos\theta = \frac{X'Y}{\|X\| * \|Y\|}$$

Anexo 2. Espacio métrico: determinación del cálculo de la distancia Ji cuadrada

Una distancia o una métrica en un conjunto S es una función d:S×$S \rightarrow R$ que satisface las siguientes propiedades:

a) d(x,y) ≥ 0 para toda x,y ∈ S

b) d(x,y) = 0 ⇔ x=y

c) d(x,y) = d(y,x) para toda x,y ∈ S

d) d(x,y) ≤ d(x,z)+d(z,y) para toda x,y,z ∈ S

Un espacio m'terico (S,d) es un conjunto S con una métrica d en S.

Para ilustrar un poco como se ha llegado a la formula de distancia para dos renglones tomemos la matriz perfil por renglones $R_{n\times p}=D_n^{-1}F$, cuya forma general es:

$$\begin{bmatrix} \frac{f_{11}}{f_{1\cdot}} & \frac{f_{12}}{f_{1\cdot}} & \frac{f_{13}}{f_{1\cdot}} & \cdots & \frac{f_{1J}}{f_{1\cdot}} \\ \frac{f_{21}}{f_{2\cdot}} & \frac{f_{22}}{f_{2\cdot}} & \frac{f_{23}}{f_{2\cdot}} & \cdots & \frac{f_{2J}}{f_{2\cdot}} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \frac{f_{n1}}{f_{n\cdot}} & \frac{f_{n2}}{f_{n\cdot}} & \frac{f_{n3}}{f_{n\cdot}} & \cdots & \frac{f_{nJ}}{f_{n\cdot}} \end{bmatrix}$$

Ahora calculemos la distancia entre dos renglones de acuerdo a la métrica definida por Dc^{-1} (matriz diagonal inversa de los pesos de las columnas). Los renglones se encuentran en un espacio R^J ponderado por los pesos asociados a la matriz Dc^{-1}, .para cada elemento del renglón, es decir

$$f_j=\left(\frac{f_{i1}}{f_{\cdot\cdot}}, \frac{f_{i2}}{f_{\cdot\cdot}}, \frac{f_{i3}}{f_{\cdot\cdot}}, \ \cdots \ , \frac{f_{iJ}}{f_{\cdot\cdot}} \right)$$

$$f_j'=\left(\frac{f_{i'1}}{f_{i'\cdot}}, \frac{f_{i'2}}{f_{i'\cdot}}, \frac{f_{i'3}}{f_{i'\cdot}}, \ \cdots \ , \frac{f_{i'J}}{f_{i'\cdot}} \right)$$

lo cual implica que $(f_j-f_j')=\left(\frac{f_{i1}}{f_{\cdot\cdot}}, \frac{f_{i2}}{f_{\cdot\cdot}}, \frac{f_{i3}}{f_{\cdot\cdot}}, \ \cdots \ , \frac{f_{iJ}}{f_{\cdot\cdot}} \right) - \left(\frac{f_{i'1}}{f_{i'\cdot}}, \frac{f_{i'2}}{f_{i'\cdot}}, \frac{f_{i'3}}{f_{i'\cdot}}, \ \cdots \ , \frac{f_{i'J}}{f_{i'\cdot}} \right)$

$$=\left(\frac{f_{i1}}{f_{\cdot\cdot}} - \frac{f_{i'1}}{f_{i'\cdot}}, \frac{f_{i2}}{f_{i'\cdot}} - \frac{f_{i'2}}{f_{i'\cdot}}, \frac{f_{i3}}{f_{\cdot\cdot}} - \frac{f_{i'3}}{f_{i'\cdot}}, \ldots, \frac{f_{iJ}}{f_{\cdot\cdot}} - \frac{f_{i'J}}{f_{i'\cdot}} \right)$$

De acuerdo a la definición de producto punto en este espacio ponderado tenemos.

$$d^2(i,i') = \langle i - i', i - i' \rangle$$

$$= \left(\frac{f_{i1}}{f_{i\cdot}} - \frac{f_{i'1}}{f_{i'\cdot}}, \frac{f_{i2}}{f_{i\cdot}} - \frac{f_{i'2}}{f_{i'\cdot}}, \ldots, \frac{f_{iJ}}{f_{i\cdot}} - \frac{f_{i'J}}{f_{i'\cdot}} \right) Dc^{-1} \left(\frac{f_{i1}}{f_{i\cdot}} - \frac{f_{i'1}}{f_{i'\cdot}}, \frac{f_{i2}}{f_{i\cdot}} - \frac{f_{i'2}}{f_{i'\cdot}}, \ldots, \frac{f_{iJ}}{f_{i\cdot}} - \frac{f_{i'J}}{f_{i'\cdot}} \right)^t$$

donde

$$Dc^{-1} = \begin{bmatrix} \frac{1}{f_{\cdot 1}} & 0 & 0 & 0 \\ 0 & \frac{1}{f_{\cdot 2}} & 0 & 0 \\ 0 & 0 & \ldots & 0 \\ 0 & 0 & 0 & \frac{1}{f_{\cdot J}} \end{bmatrix}$$

y por lo tanto

$$d^2(i,i') = \frac{1}{f_{\cdot 1}} \left[\frac{f_{i1}}{f_{i\cdot}} - \frac{f_{i'1}}{f_{i'\cdot}} \right]^2 + \frac{1}{f_{\cdot 2}} \left[\frac{f_{i2}}{f_{i\cdot}} - \frac{f_{i'2}}{f_{i'\cdot}} \right]^2 + \ldots + \frac{1}{f_{\cdot J}} \left[\frac{f_{iJ}}{f_{i\cdot}} - \frac{f_{i'J}}{f_{i'\cdot}} \right]^2$$

$$= \sum_{j=1}^{J} \frac{1}{f_{\cdot j}} \left[\frac{f_{ij}}{f_{i\cdot}} - \frac{f_{i'j}}{f_{i'\cdot}} \right]^2$$

Que es la misma expresión que se planteó en el capítulo I. Ponderar las diferencias con respecto a los elementos de los espacios respectivos implica introducir las matrices diagonales de pesos $Dc_{J\times J}^{-1}$ para el caso de R^J : $Dr_{I\times I}^{-1}$ y algunos autores llaman a estas matrices "metricas" o la distancia. De tal manera que se habla de la distancia definida por la matriz Dc^{-1}, para el espacio R^J de los renglones.

Habría que probar que la distancia así definida en el espacio ponderado es efectivamente una distancia. Por lo pronto dejaremos eso de lado y supondremos que la distancia está bien definida.

Una de las razones para escoger la distancis Ji-cuadrada, es que en esta se verifica la propiedad de equivalencia distribucional, que se expresa para el caso de perfiles por renglon y columna como:

"Si dos renglones de la tabla de contingencia original, son proporcionales y, se reemplazan por un sólo renglón el cual es la suma de los dos renglones, entonces las distancias entre las columnas no cambian. Esta propiedad es importante porque garantiza invarianza en resultados sin importar como fueron recogidos.

Anexo 3

Maximización de una forma cuadrática bajo un restricción cuadrática.

Encontremos un vetor u tal que maximice la cantidad $u'Au$ con la restricción $u'Mu=1$, donde A y M son matrices simétricas , y adicionalmente M es positiva definida.

Es decir:

Max $u'Au$

s.a. $u'Mu = 1$

Desarrollo:

La forma cuadrática $u'Au$ puede ser escrita como:

$$u'Au = \sum_i \sum_j a_{ij} u_i u_j \qquad (1)$$

Derivando esta cantidad para para las p componentes del vector u sucesivamente, se observa que el vector de las derivadas parciales de $u'Au$ se escriben en forma matricial como:

$$\frac{\partial(u'Au)}{\partial u} = 2Au \qquad (2)$$

y

$$\frac{\partial(u'Mu)}{\partial u} = 2Mu \qquad (3)$$

El lagrangiano es:

$$\pounds = u'Au - \lambda(u'Mu - 1) \qquad (4)$$

donde λ es un multiplicador de Lagrange

Para encontrar un máximo se requiere igualar las derivadas de la ecuación de Lagrange a cero:

$$\frac{\partial \pounds}{\partial u} = 2Au - 2\lambda Mu = 0 \qquad (5)$$

De la ecuación anterior se deduce que:

$$Au = \lambda Mu \qquad (6)$$

Cuando premultiplicamos a ambos lados de la ecuación por u', tomando en cuenta que el hecho de que $u'Mu = 1$ se tiene que

$$u'Au = \lambda u'Mu = \lambda, \tag{7}$$

el valor del parámetro λ es máximo.

Como la matriz M es positiva definida, y por lo tanto no singular, la relación (6) se escribe como:

$$M^{-1}Au = \lambda u \tag{8}$$

u es el eigenvector de la matriz $M^{-1}A$ que corresponde al eigenvalor más grande λ

De aquí en adelante llamaremos u_1 al vector u que corresponde al valor más grande λ_1 tal que la ecuación (6) es cierta. Encontraremos un vector u_2 tal que es M-ortogonal a u_1, es decir $u_1Mu_2 = 0$; y tenga norma 1, es decir $u_2'Mu_2 = 1$.

Igualando a cero las derivadas parciales de la ecuación de Lagrange

$$\pounds = u_2'Au_2 - \lambda_2(u_2'Mu_2 - 1) - \mu_2 u_2'Mu_1 \tag{9}$$

donde λ_2 y μ_2 son los dos multiplicadores de Lagrange.

La condición para un máximo o un mínimo es escrita para u_2 como:

$$\frac{\partial \pounds}{\partial u} = 2Au_2 - 2\lambda_2 Mu_2 - \mu_2 Mu_1 = 0 \tag{10}$$

Si multiplicamos los elementos de la ecuación (10) por u_1', se tiene que $\mu_2 = 0$ porque

$$u_1'Au_2 = u_1'Mu_2 = 0 \Rightarrow \mu_2 u_1'Mu_1 = 0 \Rightarrow \mu_2 = 0$$

Así, de acuerdo a la restricción de ortogonalidad se tiene que

$$2Au_2 - 2\lambda_2 Mu_2 = 0 \tag{11}$$

es decir:

$$Au_2 = \lambda_2 Mu_2 \tag{12}$$

Como M es nosingular, u_2 es el segundo eigenvector de $M^{-1}A$,relativo al segundo eigenvalor más grande λ_2,. si este es único.

En forma general podemos escribir que para encontrar el $\alpha - ésimo$ vector se resuleve el sistema:

$$Au_\alpha = \lambda_\alpha Mu_\alpha$$

y si M es no singular, $M^{-1}Au_\alpha = \lambda_\alpha u_\alpha$, α no puede ser más grande que el orden de la matriz A.

Anexo 4. Análisis de correspondencias como un caso particular del análisis de correlación canónico.

Los datos con los que trabaja el AFc, cuando hace uso de la información por individuos es $X=(x^1\ x^2\ x^J)$ y $Y=(y^1\ y^2\ y^J)$ donde x^J y y^J representan las variables indicadoras asociadas respectivamente a las categorías de los factores X y Y.

Esta forma de presentar el problema, permite estudiar la relación entre las características cualitativas a través del análisis de la dependencia entre dos grupos de variables cuantitativas muy particlares (las variables indicadoras).

De acuerdo a lo planteado el análisis canónico de las matrices X y Y consiste en encontrar parejas de variables canónicas (ξ^J,η^J) lo más correlacionadamente posible. De tal manera que las variables canónicas ξ^J,η^J transforman una característica cualitativa en una característica cuantitativa. A esta operación de cuantificar se conoce en estadistica como "escalamiento de una variable cualitativa".

En esas condiciones, el AFC de la tabla original permite obtener los mejores escalamientos (en el sentido de máxima correlación) de las variables cualitativas X y Y.

Se identifican las matrices V_{XX} y V_{YY} por:

$$V_{XX} = \frac{1}{n}\begin{pmatrix} n_{1\bullet} & 0 & 0 & 0 & 0 \\ 0 & \dots & 0 & 0 & 0 \\ 0 & 0 & n_{i\bullet} & 0 & 0 \\ 0 & 0 & 0 & \dots & 0 \\ 0 & 0 & 0 & 0 & n_{I\bullet} \end{pmatrix} = Dp_X$$

$$V_{YY} = \frac{1}{n}\begin{pmatrix} n_{\bullet 1} & 0 & 0 & 0 & 0 \\ 0 & \dots & 0 & 0 & 0 \\ 0 & 0 & n_{\bullet j} & 0 & 0 \\ 0 & 0 & 0 & \dots & 0 \\ 0 & 0 & 0 & 0 & n_{\bullet J} \end{pmatrix} = Dp_Y$$

Estas matrices tienen en sus diagonales los pesos relativos de cada renglón y de cada columna respectivamente.

Y, además la matriz V_{XY} está dada por:

$$V_{XY} = \frac{1}{n}X^tY = \frac{1}{n}\begin{pmatrix} n_{11} & \dots & n_{1j} & \dots & n_{1J} \\ \dots & \dots & \dots & \dots & \dots \\ n_{i1} & \dots & n_{ij} & \dots & n_{iJ} \\ \dots & \dots & \dots & \dots & \dots \\ n_{I1} & \dots & n_{Ij} & \dots & n_{IJ} \end{pmatrix} = \frac{1}{c}C = C^*$$

de acerdo al problema planteado para el análisis canónico clásico, los factores canónicos son solución de las ecuaciones:

$$D_{p_X}^{-\frac{1}{2}}C^*D_{p_Y}^{-\frac{1}{2}}C^{*t}a_k = \lambda_k a_k$$

$$D_{p_Y}^{-\frac{1}{2}}C^{*t}D_{p_X}^{-\frac{1}{2}}C^*b_k = \lambda_k b_k$$

Nótese que la matriz $D_{p_X}^{-\frac{1}{2}}C^*$ contiene los perfiles por englón para cada categoría i de X. Así mismo, la matriz $D_{p_Y}^{-\frac{1}{2}}C^{*t}$ contiene los perfiles por columna para cada categoría j de Y.

Los factores del AFC son entonces los vectores propios del producto de dos matrices de perfiles:

$$P_r P_c^t a_k = \lambda_k a_k$$

$$P_c^t P_r b_k = \lambda_k b_k$$

Anexo 5. Salida del programa Xlstat V 4.0 para los datos de la tabla 1. capítulo II..

Interpretación de la salida del paquete:

XLSTAT - Analyse Factorielle des Correspondances
Plage de données :

Test de l'hypothèse d'indépendance entre les lignes et les colonnes :
Distance du Chi2 sur les données = 19.1780 ~ Nombre de degrés de liberté = 6
Probabilité associée : 0.0039
Chi2 limite pour l'intervalle de confiance choisi = 12.5916
Sur la base de ce test on doit rejeter l'hypothèse d'indépendance entre lignes et colonnes.

Prueba Ji cuadrada de indepe

Valeurs propres et pourcentage d'inertie correspondant :

Valeurs pr.	1	2
Valeur	0.0463	0.0157
% d'inertie	75%	25%
% cumulé	75%	100%

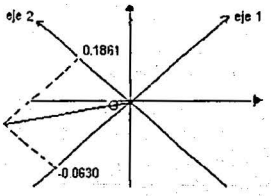
Valores propios, que definen a los vectores propios que generan la base para el nuevo sistema de referencia: los ejes factoriales, solución al problema de maximización.

Traza=0.0463+0.0157 = 0.062

Composantes des profils lignes sur les axes factoriels :

	Poids	Inertie	Inertie normée	1	2
1	0.3042	0.0117	0.1892	-0.0630	0.1861
2	0.3107	0.0305	0.4912	0.3094	-0.0490
3	0.3851	0.0198	0.3196	-0.1999	-0.1075
Perfil renglón	Peso	0.062	Inertie/traza	proyección eje 1	proyección eje 2

Componentes de los perfile por renglón: turno



Inercia (Turno 1) = $0.3042*(-0.0630)^2 + 0.3042*(0.1861)^2$
$= 0.0117427$
primer turno

Composantes des profils colonnes sur les axes factoriels :

	Poids	Inertie	Inertie normée	1	2
A	0.2395	0.0117	0.1883	0.0317	-0.2186
B	0.2233	0.0250	0.4023	0.3280	0.0650
C	0.4142	0.0058	0.0929	-0.0765	0.0898
D	0.1230	0.0196	0.3165	-0.3996	0.0052
▼ Perfil columna	▼ Peso	▼ 0.062	▼ Inerbe/traza	▼ proyección eje 1	▼ proyección eje 2

Componentes de los perfiles por columna: Tipo de defecto

Contributions des lignes aux inerties associées aux axes factoriels

	1	2
1	0.0260	0.6698
2	0.6419	0.0474
3	0.3321	0.2828

Con tribuciones relativas de los perfiles renglón a la inercia que recoge cada eje

$$= \frac{\text{Contribución absoluta}}{\lambda_\alpha} = \frac{0.3042 * (0.083)^2}{0.0463} = 0.026$$

primer turno

Contributions des colonnes aux inerties associées aux axes factoriels

	1	2
A	0.0052	0.7274
B	0.5186	0.0599
C	0.0523	0.2124
D	0.4240	0.0002

Contribuciones relativas de los perfiles columna a la inercia que recoge cada eje

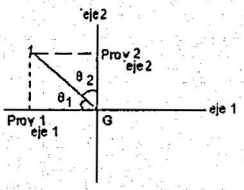
$$= \frac{\text{Contribución absoluta}}{\lambda_\alpha} = \frac{0.2395*(-0.2186)^2}{0.0157} = 0.72$$

Defecto A

Cosinus carrés des angles des vecteurs-lignes avec les axes

	1	2
1	0.1027	0.8973
2	0.9755	0.0245
3	0.7756	0.2244

Cosenos al cuadrado: correlaciones al cuadrado de los perfiles renglón con respecto a cada eje



$$d^2(G, \text{Turno1}) = \text{Prov}^2_{\text{eje 1}} + \text{Prov}^2_{\text{eje2}}$$

$$\text{Cos}^2\,\theta_1 = \frac{d_1^2(G, \text{Turno 1})}{d^2(G, \text{Turno1})} = \frac{-0.0630^2}{-0.063^2 + 0.1861^2} = 0.102817$$

correlación con el primer eje primer turno

$$\text{Cos}^2\,\theta_2 = \frac{d_2^2(G, \text{Turno 1})}{d^2(G, \text{Turno1})} = \frac{0.1861^2}{-0.063^2 + 0.1861^2} = 0.897182$$

correlación con el segundo eje primer turno

Cosinus carrés des angles des vecteurs-colonnes avec les axes

	1	2
A	0.0205	0.9795
B	0.9622	0.0378
C	0.4202	0.5798
D	0.9998	0.0002

Cosenos al cuadrado; correlaciones al cuadrado de los perfiles por columna con respecto a cada eje

Anexo 6. Base de datos y resultados del AFCM. Para el ejemplo de tres factores : Sexo, sentencia, y cargo. Capítulo III

Sexo	Sentenci	Cargo	Frecuencia
HOMBRE	SI	DAÑ_IMP	8
HOMBRE	SI	ROBO<10000	11
HOMBRE	SI	DES_PUB	5
HOMBRE	SI	NARCOT	7
HOMBRE	SI	OTROS	12
MUJER	SI	DAÑ_IMP	5
MUJER	SI	ROBO<10000	15
MUJER	SI	DES_PUB	3
MUJER	SI	NARCOT	1
MUJER	SI	OTROS	6
HOMBRE	NO	DAÑ_IMP	105
HOMBRE	NO	ROBO<10000	32
HOMBRE	NO	DES_PUB	11
HOMBRE	NO	NARCOT	23
HOMBRE	NO	OTROS	37
MUJER	NO	DAÑ_IMP	32
MUJER	NO	ROBO<10000	57
MUJER	NO	DES_PUB	6
MUJER	NO	NARCOT	2
MUJER	NO	OTROS	25

XLSTAT - Analyse Factorielle des Correspondances Multiples / Début le 03/02/2003 à 09:43:02 a.m.
Plage de données : classeur = Libro2 / feuille = Hoja1 / plage = Hoja1!$A$2:$C$21
Variables supplémentaires : 0
Observations supplémentaires : 0

Tableau de Burt :

	HOMBRE	MUJER	SI	NO	DAÑ_IMP	ROBO<10000	DES_PUB	NARCOT	OTROS
HOMBRE	251	0	43	208	113	43	13	30	49
MUJER	0	152	30	122	37	72	9	3	31
SI	43	30	73	0	13	26	8	8	18
NO	208	122	0	330	137	89	17	25	62
DAÑ_IMP	113	37	13	137	150	0	0	0	0
ROBO<10000	43	72	26	89	0	115	0	0	0
DES_PUB	16	9	8	17	0	0	25	0	0
NARCOT	30	3	8	25	0	0	0	33	0
OTROS	49	31	18	62	0	0	0	0	80

Valeurs propres et coordonnées des catégories sur les axes correspondants :

Valeurs pr.	1	2	3	4	5	6
Valeur	0.4626	0.3858	0.3333	0.3333	0.2745	0.2105
% de vari.	23%	19%	17%	17%	14%	11%
% cumulé	23%	42%	59%	76%	89%	100%
Catégories	1	2	3	4	5	6
HOMBRE	-0.6075	0.1969	0.0000	0.0000	0.1378	-0.4228
MUJER	1.0032	-0.3251	0.0000	0.0000	-0.2275	0.6982
SI	0.6917	1.4704	0.0000	0.0000	1.3466	0.2577
NO	-0.1530	-0.3253	0.0000	0.0000	-0.2979	-0.0570
DAÑ_IMP	-0.7480	-0.6449	0.3174	-0.1023	0.6097	0.4777
ROBO<100(	1.1321	-0.3125	-0.3063	-0.6238	0.0963	-0.7956
DES_PUB	0.2312	1.6407	3.2181	-0.4996	-1.3283	-0.0664
NARCOT	-1.0558	1.6479	-1.7698	-1.4928	-1.1625	0.8194
OTROS	0.1385	0.4660	-0.4306	1.3605	-0.3870	-0.0689

Contributions, contributions centrales et positions centrales

Catégories / Axes	Effectif	Poids relatif	1	2	3	4	5	6	1	2	3	4	5	6
[illegible]	251	20.76	333.79	42.52	0.00	0.30	28.93	306.29	0.61	0.08	0.00	0.00	0.03	0.32
[illegible]	142	12.97	561.20	89.38	0.00	0.00	47.77	149.59	0.41	0.08	0.00	0.00	0.03	0.30
Sum for variable 1	433	33.33	894.99	131.90	0.00	0.30	76.69	541.82						
[illegible]	73	6.04	125.95	691.84	0.00	0.00	833.91	26.31	0.11	0.46	0.00	0.00	0.40	3.01
NO	330	27.30	27.84	150.83	0.00	0.05	177.53	5.46	0.11	0.46	0.00	0.00	0.40	0.01
Sum for variable 2	433	33.33	153.79	832.67	0.00	0.05	981.74	45.69						
[illegible]	150	12.41	322.42	956.51	75.58	7.65	336.65	270.84	0.23	0.25	0.06	0.01	3.23	0.1-
[illegible]	119	9.91	531.02	48.52	53.94	223.79	6.47	576.31	0.51	0.04	0.04	0.19	0.03	3.28
[illegible]	26	2.07	4.82	343.71	1294.52	31.21	267.97	0.87	0.03	0.18	0.68	0.02	0.12	0.00
[illegible]	33	2.75	132.53	367.14	518.81	87.68	270.62	175.03	0.10	0.24	0.26	0.30	0.12	3.08
[illegible]	105	8.70	9.53	75.35	74.16	1,264.52	72.75	3.01	0.00	0.05	0.05	0.48	0.04	0.00
Sum for variable 3	433	33.33	978.31	1070.43	2015.00	2015.00	954.57	1026.26						

Distances et valeurs-test

Catégories / Axes	Distances^2	1	2	3	4	5	6
[illegible]	0.81						
[illegible]	1.76						
M	4.52						
NO	0.22						
[illegible]	1.91						
[illegible]	2.76						
[illegible]	15.12						
[illegible]	11.21						
[illegible]	4.04						

[illegible] test values, absolutely [illegible] is significant for axis

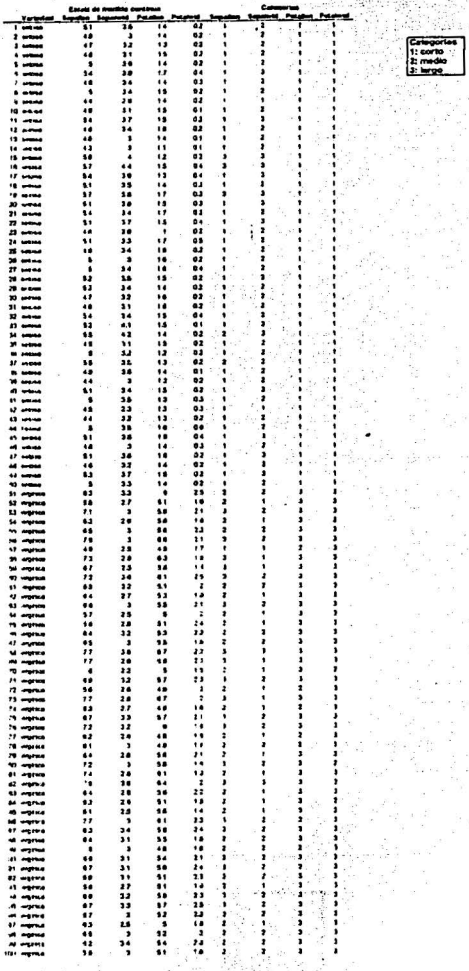
Coordonnées et contributions des observations

Observations	Poids relatif	Distances^2	Coordonnées en each axis						Contributions en each axis					
			1	2	3	4	5	6	1	2	3	4	5	6
Obs 1	1.16	16.17	-0.92	1.55	0.52	-0.17	3.77	0.64	0.46	1.55	0.20	0.02	12.94	0.46
Obs 2	2.73	27.91	1.85	3.41	-0.58	-1.19	3.34	-2.31	2.10	3.74	0.28	1.05	10.05	6.32
Obs 3	1.24	33.74	0.35	3.87	4.15	-0.85	0.22	-0.39	0.05	10.13	12.65	0.31	0.04	0.17
Obs 4	1.74	38.12	-1.28	4.71	-2.70	-2.28	0.54	1.28	0.85	14.28	5.44	3.87	0.27	1.68
Obs 5	2.36	39.66	0.38	3.97	-0.06	3.72	2.42	-0.56	0.09	10.12	0.05	10.21	4.59	0.41
Obs 6	1.24	13.10	1.04	0.65	0.41	-0.13	2.45	2.33	0.58	0.22	0.13	0.01	5.47	6.39
Obs 7	3.72	43.38	5.27	1.73	-0.65	-1.39	3.00	0.45	15.46	1.93	0.35	1.45	8.11	0.24
Obs 8	0.74	21.39	1.64	2.59	3.22	-0.80	-0.23	1.12	1.43	4.21	7.71	0.19	0.05	1.48
Obs 9	0.26	8.79	0.31	1.50	-1.02	-0.98	-0.03	1.29	0.05	1.46	0.78	0.96	0.00	1.91
Obs 10	1.49	30.42	2.30	3.12	-0.91	2.83	1.14	1.56	2.40	2.89	0.28	5.15	1.16	2.94
Obs 11	26.05	87.97	-7.58	-4.25	1.68	-0.61	2.43	-0.02	30.79	11.83	2.63	0.27	7.77	0.00
Obs 12	7.94	35.53	1.05	-1.34	-1.00	-2.04	-0.23	-5.24	0.97	1.15	0.74	3.09	0.05	32.39
Obs 13	2.75	58.47	-0.95	2.68	6.16	-0.95	-3.14	-1.32	0.42	4.56	26.27	0.98	1.42	2.04
Obs 14	5.71	62.30	-4.27	3.91	-4.30	4.13	-4.04	1.18	9.78	8.64	17.88	12.72	14.73	1.84
Obs 15	9.18	90.94	-1.85	1.10	-1.51	9.53	-2.12	-2.42	1.84	0.78	1.70	31.78	4.26	6.83
Obs 16	7.94	37.96	0.28	-3.43	1.04	-0.33	0.30	4.63	0.04	9.94	0.83	0.08	0.09	24.63
Obs 17	14.14	63.16	7.33	-3.93	-1.34	-2.72	-2.05	-0.76	26.48	9.79	1.35	5.50	3.64	0.65
Obs 18	1.49	33.94	1.30	1.35	4.56	-0.71	-2.79	1.02	0.91	1.09	18.42	0.37	7.55	1.23
Obs 19	0.50	6.72	0.14	0.79	-1.48	-1.22	-1.52	1.50	0.01	0.27	1.98	1.11	2.29	2.66
Obs 20	6.70	44.26	2.42	-0.49	-1.24	1.37	-2.40	2.78	3.16	0.15	1.15	21.47	7.42	5.09

Coordonnées centrales pour les observations

Observations	1	2	3	4	5	6
Obs 1	0.36	0.13	0.01	0.00	0.78	0.02
Obs 2	0.14	0.21	0.01	0.05	0.40	3.19
Obs 3	0.00	0.47	0.51	0.01	0.00	0.00
Obs 4	0.04	0.56	0.19	0.14	0.01	0.04
Obs 5	0.00	0.43	0.02	0.38	0.16	3.01
Obs 6	0.08	0.03	0.01	0.00	0.46	0.41
Obs 7	0.45	0.07	0.01	0.04	0.21	0.00
Obs 8	0.13	0.31	0.49	0.01	0.00	0.08
Obs 9	0.02	0.39	0.18	0.13	0.00	0.29
Obs 10	0.24	0.22	0.02	0.34	0.08	0.12
Obs 11	0.45	0.21	0.04	0.00	0.10	0.01
Obs 12	0.03	0.06	0.03	0.12	0.00	0.77
Obs 13	0.01	0.12	0.65	0.02	0.17	0.33
Obs 14	0.20	0.17	0.28	0.14	0.18	0.02
Obs 15	0.05	0.02	0.04	3.71	0.07	0.10
Obs 16	0.03	0.41	0.03	0.03	0.03	0.56
Obs 17	0.65	0.18	0.02	0.04	0.05	0.01
Obs 18	0.05	0.76	0.37	0.01	0.25	0.03
Obs 19	0.03	0.07	2.24	0.17	0.26	0.28
Obs 20	0.12	0.33	0.01	0.58	0.17	0.46

Anexo 7. Base de datos para las flores de Iris. Resultados del AFCM. Capitulo III



XLSTAT - Analyse Factorielle des Correspondances Multiples / Début le 03/02/2003 à 10:49:29 a.m.
Plage de données : classeur = anexo7.xls / feuille = Hoja2 / plage = $B$1:$F$101
Variables supplémentaires : 0
Observations supplémentaires : 0

Tableau de Burt :

	Variedad / setosa	Variedad / virgin	Sepalen / 1	Sepalen / 3	Sepalen / 2	Sepalwid / 2	Sepalwid / 3	Sepalwid / 1	Petalen / 1	Petalen / 3	Petalen / 2	Petalwid / 1	Petalwid / 3	Petalwid / 2
Variedad / se	50	0	45	3	2	35	13	2	50	0	0	50	0	0
Variedad / vi	0	50	1	21	28	27	2	21	0	44	6	0	46	4
Sepalen / 1	45	1	46	0	0	34	9	3	45	0	1	45	1	0
Sepalen / 3	3	21	0	24	0	15	5	4	3	21	0	3	20	1
Sepalen / 2	2	28	0	0	30	13	1	16	2	23	5	2	25	3
Sepalwid / 2	35	27	34	15	13	62	0	0	35	25	2	35	26	1
Sepalwid / 3	13	2	9	5	1	0	15	0	13	2	0	13	2	0
Sepalwid / 1	2	21	3	4	16	0	0	23	2	17	4	2	18	3
Petalen / 1	50	0	45	3	2	35	13	2	50	0	0	50	0	0
Petalen / 3	0	44	0	21	23	25	2	17	0	44	0	0	40	4
Petalen / 2	0	6	1	0	5	2	0	4	0	0	6	0	6	0
Petalwid / 1	50	0	45	3	2	35	13	2	50	0	0	50	0	0
Petalwid / 3	0	46	1	20	25	26	2	18	0	40	6	0	46	0
Petalwid / 2	0	4	0	1	3	1	0	3	0	4	0	0	0	4

Valeurs propres et coordonnées des catégories sur les axes correspondants

Valeurs pr.	1	2	3	4	5	6	7	8	9
Valeur	0.4223	0.2853	0.2227	0.2029	0.1343	0.1053	0.0275	0.0000	0.0000
% de vari	44%	16%	12%	11%	7%	6%	2%	0%	0%
% cumulé	46%	62%	74%	85%	93%	98%	100%	100%	100%
**Catégories **	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
Variedad / se	0.9881	-0.0808	0.0430	0.0018	-0.0369	0.0505	-0.1067	0.0000	0.0000
Variedad / vi	0.9881	0.0808	-0.0430	-0.0018	0.0369	-0.0505	0.1067	0.0000	0.0000
Sepalen / 1	0.9828	-0.1858	-0.0449	-0.1535	0.0219	0.1841	0.3368	0.0000	0.0000
Sepalen / 3	-0.6756	1.3867	0.0175	0.3743	0.6400	0.3775	-0.3077	0.0000	0.0000
Sepalen / 2	0.9665	-0.9245	0.0548	-0.0642	-0.5457	-0.5843	0.2703	0.0000	0.0000
Sepalwid / 2	0.1799	0.2503	-0.2941	0.5005	0.1438	-0.2235	0.0162	0.0000	0.0000
Sepalwid / 3	0.7960	0.4101	0.5568	2.0119	0.1508	-0.5734	0.1750	0.0000	0.0000
Sepalwid / 1	-1.0040	-0.9422	0.4297	0.3066	-0.2883	1.0418	0.0710	0.0000	0.0000
Petalen / 1	0.9881	-0.0808	0.0430	0.0018	0.0369	0.0505	0.1067	0.0000	0.0000
Petalen / 3	-0.9820	0.4143	0.2097	-0.1372	-0.2431	-0.0243	0.1196	0.0000	0.0000
Petalen / 2	-1.0329	-2.3645	-1.8957	0.9928	2.0906	-0.2421	0.0120	0.0000	0.0000
Petalwid / 1	0.9881	0.0808	0.0430	0.0018	-0.0369	0.0505	-0.1067	0.0000	0.0000
Petalwid / 3	-0.9742	0.1736	-0.3992	0.0893	0.1282	0.0015	0.1048	0.0000	0.0000
Petalwid / 2	-1.1480	-0.9865	4.0535	-1.0467	1.9376	-0.6479	0.1293	0.0000	0.0000

Contributions, contributions cumulées et cosinus carrés

Catégories \ Ax	Effectif	Poids relatif	1	2	3	4	5	6	7	8	9	1	2	3	4	5	6	7	8	9
Variedad / se	50	10.00	11.87	0.23	0.08	0.00	0.10	0.24	4.14	0.79	11.48	0.98	0.01	0.00	0.00	0.00	0.00	0.01	0.00	0.0[illegible]
Variedad / vi	50	10.00	11.87	0.23	0.08	0.00	0.10	0.24	4.14	3.03	3.22	0.98	0.01	0.00	0.00	0.00	0.00	0.01	0.00	0.0[illegible]
Sum for Var	100	20.00	23.75	0.46	0.17	0.00	0.20	0.48	8.27	3.82	11.70									
Sepalen / 1	46	9.20	10.81	1.11	0.08	1.07	0.03	2.96	37.93	1.66	3.25	0.82	0.03	0.00	0.02	0.00	0.03	0.10	0.00	0.00
Sepalen / 3	24	4.80	2.56	32.15	0.01	3.32	14.64	8.49	16.52	0.87	1.69	0.14	0.61	0.00	0.04	0.13	0.04	0.03	0.00	0.00
Sepalen / 2	30	6.00	6.92	14.30	0.09	0.12	13.30	19.45	15.93	1.08	2.12	0.40	0.29	0.00	0.00	0.13	0.15	0.03	0.00	0.0[illegible]
Sum for Sep	100	20.00	20.29	47.76	0.17	4.51	27.98	28.91	70.38	3.61	7.0[illegible]									
Sepalwid / 2	62	12.40	0.49	2.72	4.52	22.07	1.90	5.98	0.12	4.29	27.39	0.05	0.10	0.14	0.59	0.03	0.08	0.00	0.00	0.0[illegible]
Sepalwid / 3	15	3.00	2.31	1.77	4.19	59.94	0.51	12.92	3.38	1.03	6.53	0.11	0.03	0.05	0.71	0.00	0.08	0.01	0.00	0.0[illegible]
Sepalwid / 1	23	4.60	5.64	14.31	3.81	2.13	2.85	47.41	0.84	1.59	10.16	0.30	0.27	0.06	0.03	0.02	0.32	0.00	0.00	0.00
Sum for Sep	100	20.00	8.44	18.81	12.81	84.14	5.25	66.21	4.34	6.90	44.17									
Petalen / 1	50	10.00	11.87	0.23	0.08	0.00	0.10	0.24	4.14	2.59	23.78	0.98	0.01	0.00	0.00	0.00	0.00	0.01	0.00	0.0[illegible]
Petalen / 3	44	8.80	10.32	5.29	1.74	0.82	3.87	0.05	4.58	41.59	10.12	0.76	0.13	0.03	0.01	0.05	0.00	0.01	0.00	0.0[illegible]
Petalen / 2	6	1.20	1.56	23.52	19.36	5.84	39.05	0.67	0.01	5.67	1.38	0.07	0.35	0.23	0.06	0.29	0.00	0.00	0.00	0.00
Sum for Pet	100	20.00	23.75	29.04	21.18	6.66	43.02	0.96	8.72	49.85	35.28									
Petalwid / 1	50	10.00	11.87	0.23	0.08	0.00	0.10	0.24	4.14	3.93	1.79	0.98	0.01	0.00	0.00	0.00	0.00	0.01	0.00	0.0[illegible]
Petalwid / 3	46	9.20	10.62	0.97	6.58	0.36	1.13	0.00	3.87	29.34	0.01	0.81	0.03	0.14	0.01	0.01	0.00	0.01	0.00	0.0[illegible]
Petalwid / 2	4	0.80	1.28	2.73	59.01	4.33	22.32	3.19	0.48	2.55	0.00	0.05	0.04	0.88	0.05	0.15	0.02	0.00	0.00	0.0[illegible]
Sum for Pet	100	20.00	23.77	3.93	65.58	4.69	23.54	3.43	8.29	35.82	1.7[illegible]									

Distances et valeurs test

Catégories \	Distance^2	1	2	3	4	5	6	7	8	9
Variedad / se	1.00	**9.83**	-0.80	0.43	0.02	-0.37	0.50	-1.06	0.00	0.00
Variedad / vi	1.00	**-9.83**	0.80	-0.43	-0.02	0.37	-0.50	1.06	0.00	0.00
Sepalen / 1	1.17	**9.03**	-1.71	-0.41	-1.41	0.20	**1.69**	**3.09**	0.00	0.00
Sepalen / 3	3.17	-3.78	7.75	0.10	**2.09**	**3.58**	2.11	-1.72	0.00	0.00
Sepalen / 2	2.33	-6.39	5.37	0.36	-0.42	-3.55	3.81	-1.76	0.00	0.00
Sepalwid / 2	0.61	**2.29**	3.18	-3.74	7.63	**1.82**	-2.54	-0.21	0.00	0.00
Sepalwid / 3	5.67	**3.33**	1.71	**2.33**	**8.41**	-0.63	-2.81	0.74	0.00	0.00
Sepalwid / 1	3.35	-5.45	-5.12	**2.34**	**1.67**	-1.57	**5.67**	-0.39	0.00	0.00
Petalen / 1	1.00	**9.83**	-0.80	0.43	0.02	-0.37	0.50	-1.06	0.00	0.00
Petalen / 3	1.27	-8.68	**3.65**	**1.86**	-1.21	-2.14	-0.21	1.06	0.00	0.00
Petalen / 2	15.67	-2.60	-5.94	-4.77	**2.50**	**5.20**	-0.61	0.03	0.00	0.00
Petalwid / 1	1.00	**9.83**	-0.80	0.43	0.02	-0.37	0.50	-1.06	0.00	0.00
Petalwid / 3	1.17	-8.95	1.59	-3.67	0.82	-1.18	0.01	0.96	0.00	0.00
Petalwid / 2	24.00	-2.33	2.09	**8.23**	-2.13	**3.93**	-1.32	0.26	0.00	0.00

In bold, test values showing category is significant for axis

Courtes cartes pour les observations

Observation	1	2	3	4	5	6	7	8	9
Obs 1	0.87	0.00	0.01	0.12	0.00	0.00	0.00	0.00	0.00
Obs 2	0.87	0.00	0.01	0.12	0.00	0.00	0.00	0.00	0.00
Obs 3	0.87	0.00	0.01	0.12	0.00	0.00	0.00	0.00	0.00
Obs 4	0.87	0.00	0.01	0.12	0.00	0.00	0.00	0.00	0.00
Obs 5	0.87	0.00	0.01	0.12	0.00	0.00	0.00	0.00	0.00
Obs 6	0.56	0.00	0.04	0.35	0.01	0.02	0.03	0.00	0.00
Obs 7	0.87	0.00	0.01	0.12	0.00	0.00	0.00	0.00	0.00
Obs 8	0.87	0.00	0.01	0.12	0.00	0.00	0.00	0.00	0.00
Obs 9	0.28	0.18	0.03	0.00	0.03	0.46	0.00	0.00	0.00
Obs 10	0.87	0.00	0.01	0.12	0.00	0.00	0.00	0.00	0.00
Obs 11	0.56	0.00	0.04	0.35	0.01	0.02	0.03	0.00	0.00
Obs 12	0.87	0.00	0.01	0.12	0.00	0.00	0.00	0.00	0.00
Obs 13	0.87	0.00	0.01	0.12	0.00	0.00	0.00	0.00	0.00
Obs 14	0.87	0.00	0.01	0.12	0.00	0.00	0.00	0.00	0.00
Obs 15	0.20	0.14	0.04	0.48	0.02	0.00	0.13	0.00	0.00
Obs 16	0.20	0.14	0.04	0.48	0.02	0.00	0.13	0.00	0.00
Obs 17	0.56	0.00	0.04	0.35	0.01	0.02	0.03	0.00	0.00
Obs 18	0.87	0.00	0.01	0.12	0.00	0.00	0.00	0.00	0.00
Obs 19	0.20	0.14	0.04	0.48	0.02	0.00	0.13	0.00	0.00
Obs 20	0.56	0.00	0.04	0.35	0.01	0.02	0.03	0.00	0.00
Obs 21	0.87	0.00	0.01	0.12	0.00	0.00	0.00	0.00	0.00
Obs 22	0.56	0.00	0.04	0.35	0.01	0.02	0.03	0.00	0.00
Obs 23	0.87	0.00	0.01	0.12	0.00	0.00	0.00	0.00	0.00
Obs 24	0.87	0.00	0.01	0.12	0.00	0.00	0.00	0.00	0.00
Obs 25	0.87	0.00	0.01	0.12	0.00	0.00	0.00	0.00	0.00
Obs 26	0.87	0.00	0.01	0.12	0.00	0.00	0.00	0.00	0.00
Obs 27	0.87	0.00	0.01	0.12	0.00	0.00	0.00	0.00	0.00
Obs 28	0.87	0.00	0.01	0.12	0.00	0.00	0.00	0.00	0.00
Obs 29	0.87	0.00	0.01	0.12	0.00	0.00	0.00	0.00	0.00
Obs 30	0.87	0.00	0.01	0.12	0.00	0.00	0.00	0.00	0.00
Obs 31	0.87	0.00	0.01	0.12	0.00	0.00	0.00	0.00	0.00
Obs 32	0.87	0.00	0.01	0.12	0.00	0.00	0.00	0.00	0.00
Obs 33	0.56	0.00	0.04	0.35	0.01	0.02	0.03	0.00	0.00
Obs 34	0.17	0.03	0.04	0.34	0.09	0.21	0.11	0.00	0.00
Obs 35	0.87	0.00	0.01	0.12	0.00	0.00	0.00	0.00	0.00
Obs 36	0.87	0.00	0.01	0.12	0.00	0.00	0.00	0.00	0.00
Obs 37	0.15	0.08	0.00	0.07	0.07	0.14	0.45	0.00	0.00
Obs 38	0.87	0.00	0.01	0.12	0.00	0.00	0.00	0.00	0.00
Obs 39	0.87	0.00	0.01	0.12	0.00	0.00	0.00	0.00	0.00
Obs 40	0.87	0.00	0.01	0.12	0.00	0.00	0.00	0.00	0.00
Obs 41	0.87	0.00	0.01	0.12	0.00	0.00	0.00	0.00	0.00
Obs 42	0.28	0.18	0.03	0.00	0.03	0.46	0.00	0.00	0.00
Obs 43	0.87	0.00	0.01	0.12	0.00	0.00	0.00	0.00	0.00
Obs 44	0.87	0.00	0.01	0.12	0.00	0.00	0.00	0.00	0.00
Obs 45	0.56	0.00	0.04	0.35	0.01	0.02	0.03	0.00	0.00
Obs 46	0.87	0.00	0.01	0.12	0.00	0.00	0.00	0.00	0.00
Obs 47	0.56	0.00	0.04	0.35	0.01	0.02	0.03	0.00	0.00
Obs 48	0.87	0.00	0.01	0.12	0.00	0.00	0.00	0.00	0.00
Obs 49	0.56	0.00	0.04	0.35	0.01	0.02	0.03	0.00	0.00
Obs 50	0.87	0.00	0.01	0.12	0.00	0.00	0.00	0.00	0.00
Obs 51	0.53	0.00	0.03	0.08	0.13	0.23	0.00	0.00	0.00
Obs 52	0.64	0.09	0.01	0.00	0.22	0.03	0.00	0.00	0.00
Obs 53	0.40	0.52	0.03	0.01	0.04	0.00	0.00	0.00	0.00
Obs 54	0.64	0.09	0.01	0.00	0.22	0.03	0.00	0.00	0.00
Obs 55	0.53	0.00	0.03	0.08	0.13	0.23	0.00	0.00	0.00
Obs 56	0.40	0.52	0.03	0.01	0.04	0.00	0.00	0.00	0.00
Obs 57	0.10	0.33	0.15	0.07	0.29	0.07	0.04	0.00	0.00
Obs 58	0.52	0.09	0.00	0.04	0.00	0.35	0.00	0.00	0.00
Obs 59	0.52	0.09	0.00	0.04	0.00	0.35	0.00	0.00	0.00
Obs 60	0.40	0.52	0.03	0.01	0.04	0.00	0.00	0.00	0.00
Obs 61	0.53	0.00	0.03	0.08	0.13	0.23	0.00	0.00	0.00
Obs 62	0.64	0.09	0.01	0.00	0.22	0.03	0.00	0.00	0.00
Obs 63	0.40	0.52	0.03	0.01	0.04	0.00	0.00	0.00	0.00
Obs 64	0.64	0.09	0.01	0.00	0.22	0.03	0.00	0.00	0.00
Obs 65	0.64	0.09	0.01	0.00	0.22	0.03	0.00	0.00	0.00
Obs 66	0.53	0.00	0.03	0.08	0.13	0.23	0.00	0.00	0.00
Obs 67	0.53	0.00	0.03	0.08	0.13	0.23	0.00	0.00	0.00
Obs 68	0.16	0.35	0.01	0.44	0.00	0.02	0.02	0.00	0.00
Obs 69	0.52	0.09	0.00	0.04	0.00	0.35	0.00	0.00	0.00
Obs 70	0.20	0.11	0.62	0.03	0.04	0.00	0.00	0.00	0.00
Obs 71	0.40	0.52	0.03	0.01	0.04	0.00	0.00	0.00	0.00
Obs 72	0.25	0.45	0.13	0.07	0.25	0.00	0.00	0.00	0.00
Obs 73	0.52	0.09	0.00	0.04	0.00	0.35	0.00	0.00	0.00
Obs 74	0.25	0.45	0.13	0.07	0.09	0.00	0.00	0.00	0.00
Obs 75	0.40	0.52	0.03	0.01	0.04	0.00	0.00	0.00	0.00
Obs 76	0.40	0.52	0.03	0.01	0.04	0.00	0.00	0.00	0.00
Obs 77	0.25	0.45	0.13	0.07	0.09	0.00	0.00	0.00	0.00
Obs 78	0.17	0.24	0.29	0.01	0.18	0.11	0.00	0.00	0.00
Obs 79	0.64	0.09	0.01	0.00	0.22	0.03	0.00	0.00	0.00
Obs 80	0.11	0.03	0.46	0.07	0.31	0.02	0.00	0.00	0.00
Obs 81	0.64	0.09	0.01	0.00	0.22	0.03	0.00	0.00	0.00
Obs 82	0.18	0.35	0.01	0.44	0.00	0.02	0.02	0.00	0.00
Obs 83	0.64	0.09	0.01	0.00	0.22	0.03	0.00	0.00	0.00
Obs 84	0.20	0.11	0.62	0.03	0.04	0.00	0.00	0.00	0.00
Obs 85	0.20	0.11	0.62	0.03	0.04	0.00	0.00	0.00	0.00
Obs 86	0.40	0.52	0.03	0.01	0.04	0.00	0.00	0.00	0.00
Obs 87	0.53	0.00	0.03	0.08	0.13	0.23	0.00	0.00	0.00
Obs 88	0.53	0.00	0.03	0.08	0.13	0.23	0.00	0.00	0.00
Obs 89	0.17	0.24	0.29	0.01	0.18	0.11	0.00	0.00	0.00
Obs 90	0.40	0.52	0.03	0.01	0.04	0.00	0.00	0.00	0.00
Obs 91	0.40	0.52	0.03	0.01	0.04	0.00	0.00	0.00	0.00
Obs 92	0.40	0.52	0.03	0.01	0.04	0.00	0.00	0.00	0.00
Obs 93	0.64	0.09	0.01	0.00	0.22	0.03	0.00	0.00	0.00
Obs 94	0.40	0.52	0.03	0.01	0.04	0.00	0.00	0.00	0.00
Obs 95	0.40	0.52	0.03	0.01	0.04	0.00	0.00	0.00	0.00
Obs 96	0.40	0.52	0.03	0.01	0.04	0.00	0.00	0.00	0.00
Obs 97	0.64	0.09	0.01	0.00	0.22	0.03	0.00	0.00	0.00
Obs 98	0.53	0.00	0.03	0.08	0.13	0.23	0.00	0.00	0.00
Obs 99	0.53	0.00	0.03	0.08	0.13	0.23	0.00	0.00	0.00
Obs 100	0.53	0.00	0.03	0.08	0.13	0.23	0.00	0.00	0.00

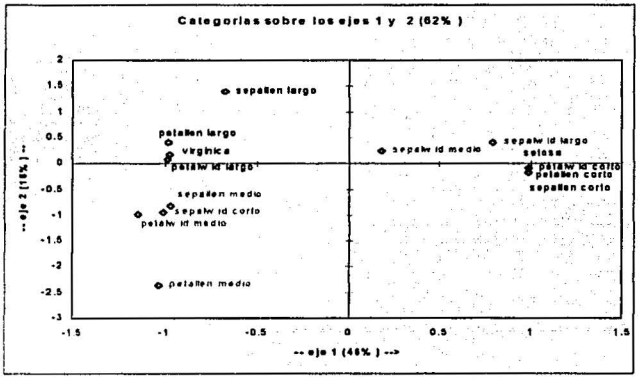
Fe de erratas.

Página 55

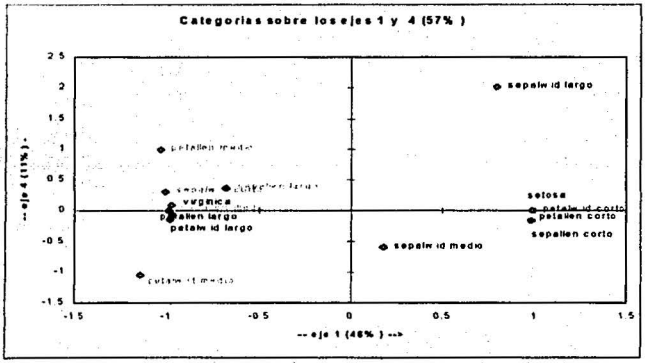
Tabla 1. Matriz de Burt.

	Hombre	Mujer	Soltero	Casado	Viudo	PRI	PAN	PRD
Hombre	5	0	3	1	1	2	1	2
Mujer	0	5	2	2	1	2	3	0
Soltero	3	2	5	0	0	2	3	0
Casado	1	2	0	3	0	2	0	1
Viudo	1	1	0	0	2	0	1	1
PRI	2	2	2	2	0	4	0	0
PAN	1	3	3	0	1	0	4	0
PRD	2	0	0	1	1	0	0	2

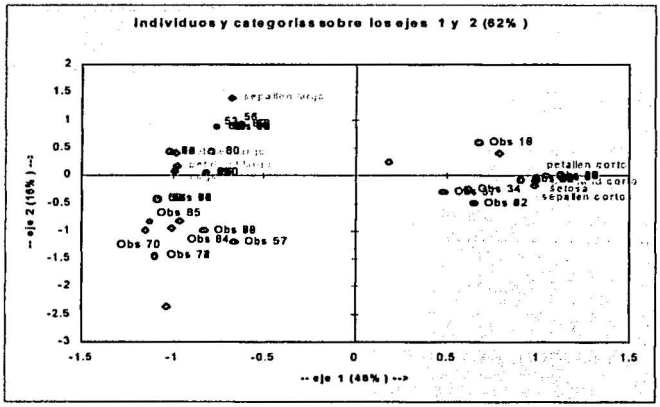
Página 82



Página 83



Página 84.



Página 90.

25. Ronald Christensen, *Log-Linear Models*, Springer-Verlag, New York, USA, 1990.