

7. ANÁLISIS DE FUNCIONES DISCRIMINANTES (AFD)

AMARN 2018 - IMFIA.FI.UDELAR -
Ing. Luis Silveira, Ph.D.

OBJETIVO

Separar dos o más grupos de observaciones, en base a las medidas de éstas sobre varias variables.

Ejemplos.

Tabla 1.1. Gorriones (21 s – 28 no-s, 5 variables)

¿Es posible utilizar las medidas del cuerpo para separar los gorriones s y no-s?

Tabla 1.2 Cráneos egipcios (5 períodos de tiempo, 4 variables)

¿Es razonable considerar las medidas de los cráneos para su “datación”?

AMARN 2018 - IMFIA.FI.UDELAR -
Ing. Luis Silveira, Ph.D.

OBJETIVO

Caso general

Sean m **muestras** aleatorias, de diferentes grupos de **tamaño** n_1, n_2, \dots, n_m , para las que se cuenta con datos de las observaciones sobre p **variables** X_1, X_2, \dots, X_p . Por tanto, los datos para el AFD toman la forma:

Observación	X_1	X_2	...	X_p	
1	x_{111}	x_{112}	...	x_{11p}	
2	x_{211}	x_{212}	...	x_{21p}	
.	Grupo 1
.	
n_1	x_{n_11}	x_{n_12}	...	x_{n_1p}	
1	x_{121}	x_{122}	...	x_{12p}	
2	x_{221}	x_{222}	...	x_{22p}	
.	Grupo 2
.	
n_2	x_{n_21}	x_{n_22}	...	x_{n_2p}	
1	x_{1m1}	x_{1m2}	...	x_{1mp}	
2	x_{2m1}	x_{2m2}	...	x_{2mp}	
.	Grupo m
.	
n_m	x_{n_m1}	x_{n_m2}	...	$x_{n_m p}$	

AMARN 2018 - IMFIA.FI.UDELAR -
Ing. Luis Silveira, Ph.D.

OBJETIVO

□ Datos en el AFD

No es necesario estandarizarlos, puesto que la escala de las variables no afecta el resultado.

□ Formas de discriminar / separar

1. Distancias de Mahalanobis
2. Funciones canónicas discriminantes

AMARN 2018 - IMFIA.FI.UDELAR -
Ing. Luis Silveira, Ph.D.

DISCRIMINACIÓN UTILIZANDO LAS DISTANCIAS DE MAHALANOBIS

- Los vectores de medias para las m muestras pueden considerarse como estimaciones de los verdaderos vectores de medias para los grupos.
- Se calculan las distancias de Mahalanobis, entre las observaciones individuales y el centroide de los grupos, y cada observación puede re-ubicarse en el grupo más próximo. Éste puede ser, o no, el grupo del que procede la observación.
- El porcentaje de re-ubicaciones correctas puede interpretarse como un indicador de cuán bien pueden separarse los grupos utilizando las variables disponibles.

AMARN 2018 - IMFIA.FI.UDELAR -
Ing. Luis Silveira, Ph.D.

DISCRIMINACIÓN UTILIZANDO LAS DISTANCIAS DE MAHALANOBIS

Sea: $\bar{x}'_i = (\bar{x}_{1i}, \bar{x}_{2i}, \dots, \bar{x}_{pi})'$ = vector de valores medios de la muestra del i -ésimo grupo.

C_i = matriz de covarianza de la misma muestra

$C = \sum_{i=1}^m (n_i - 1) C_i / \sum_{i=1}^m (n_i - 1)$ = matriz "media" de covarianza, en base a todas las muestras disponibles

La distancia de Mahalanobis entre una observación $x' = (x_1, x_2, \dots, x_p)'$

y el centroide del grupo, se estima según: Elemento en la r -ésima fila y s -ésima columna de la matriz inversa de C

$$D_i^2 = (x - \bar{x}_i)' C^{-1} (x - \bar{x}_i) = \sum_{r=1}^p \sum_{s=1}^p (x_r - \bar{x}_{ri}) c^{rs} (x_s - \bar{x}_{si})$$

La observación x se reubica en el grupo que produce el menor valor de D_i^2

AMARN 2018 - IMFIA.FI.UDELAR -
Ing. Luis Silveira, Ph.D.

FUNCIONES CANÓNICAS DISCRIMINANTES

Procedimiento

Determinar funciones de las variables X_1, X_2, \dots, X_p que en algún sentido separan los m grupos tan bien como es posible. La aproximación más simple es una combinación lineal de las variables:

$$Z = a_1X_1 + a_2X_2 + \dots + a_pX_p$$

Los grupos se pueden separar utilizando Z , si los valores medios cambian considerablemente de grupo en grupo, y son aproximadamente constantes en el grupo.

Un modo de determinar los coeficientes a_1, a_2, \dots, a_p consiste en maximizar la razón F en un análisis de varianza unidireccional.

AMARN 2018 - IMFIA.FI.UDELAR -
Ing. Luis Silveira, Ph.D.

FUNCIONES CANÓNICAS DISCRIMINANTES

Si N = nro. total de observaciones considerando todos los grupos, un análisis de la varianza de los valores de Z toma la siguiente forma:

Fuente de variación	Grados de libertad	Media al cuadrado	Razón F
Entre grupos	$m-1$	M_B	M_B/M_W
Intra grupos	$N-m$	M_W	
	$N-1$		

Una función adecuada para separar los grupos puede definirse como la combinación para la cual $F = M_B/M_W$ sea tan grande como sea posible (Fisher, 1936)

AMARN 2018 - IMFIA.FI.UDELAR -
Ing. Luis Silveira, Ph.D.

FUNCIONES CANÓNICAS DISCRIMINANTES

Utilizando esta aproximación, es posible determinar varias combinaciones lineales para separar los grupos. En general, el número disponible es:

$$s = \text{Mín}(p, m-1)$$

Estas s combinaciones lineales se denominan **funciones canónicas discriminantes**.

La primer función:

$$Z_1 = a_{11}X_1 + a_{12}X_2 + \dots + a_{1p}X_p$$

produce la máxima razón F posible en un análisis de varianza unidireccional, para la variación dentro y entre grupos.

AMARN 2018 - IMFIA.FI.UDELAR -
Ing. Luis Silveira, Ph.D.

FUNCIONES CANÓNICAS DISCRIMINANTES

Si existe más de una función, la segunda:

$$Z_2 = a_{21}X_1 + a_{22}X_2 + \dots + a_{2p}X_p$$

Produce la máxima razón F posible, en un análisis de varianza unidireccional, sujeto a la **no-correlación entre Z_1 y Z_2** dentro de los grupos (o intra grupos).

Las subsiguientes funciones se definen de igual modo. La i -ésima función canónica discriminante:

$$Z_i = a_{i1}X_1 + a_{i2}X_2 + \dots + a_{ip}X_p$$

es la combinación lineal que maximiza la razón F en un análisis de varianza unidireccional, sujeto a que Z_i está **no-correlacionada con Z_1, Z_2, \dots, Z_{i-1}** dentro de los grupos (o intra grupos).

AMARN 2018 - IMFIA.FI.UDELAR -
Ing. Luis Silveira, Ph.D.

FUNCIONES CANÓNICAS DISCRIMINANTES

La determinación de los coeficientes de las funciones canónicas discriminantes (FCD) se transforma en un problema de **obtención de los valores propios**:

- Cálculo de la matriz **W** intra muestras y la matriz **T** de la muestra total (ver Pruebas de significancia con datos multiv).
- Cálculo de la matriz entre grupos: **B = T - W**
- Determinación de los valores y vectores propios de la matriz **W⁻¹ B**

AMARN 2018 - IMFIA.FI.UDELAR -
Ing. Luis Silveira, Ph.D.

FUNCIONES CANÓNICAS DISCRIMINANTES

Si $\lambda_1 > \lambda_2 > \dots > \lambda_s$ son los valores propios, λ_i es la razón entre la suma de los cuadrados **entre** grupos y la suma de los cuadrados **dentro** del grupo, para la i-ésima combinación lineal, Z_i , mientras que los elementos de los correspondientes vectores propios, $a'_i = (a_{i1}, a_{i2}, \dots, a_{ip})$ son los coeficientes de Z_i

AMARN 2018 - IMFIA.FI.UDELAR -
Ing. Luis Silveira, Ph.D.

FUNCIONES CANÓNICAS DISCRIMINANTES

Z_1 refleja, tanto como es posible, las diferencias del grupo
 Z_2 captura, tanto como es posible, las diferencias del grupo no reflejadas por Z_1
 Z_3 captura, tanto como es posible, las diferencias del grupo no reflejadas por Z_1 y Z_2
etc.

Es de esperar que las primeras FCD sean suficientes para reflejar la mayor parte de las importantes diferencias entre grupos. En particular, si sólo se necesita Z_1 o Z_1 y Z_2 , es posible representar gráficamente la relación entre los diversos grupos graficando los valores de estas funciones para las observaciones de la muestra.

AMARN 2018 - IMFIA.FI.UDELAR -
Ing. Luis Silveira, Ph.D.

PRUEBAS DE SIGNIFICANCIA

Pruebas de significancia de utilidad en conjunción con el análisis de funciones discriminantes:

- El estadígrafo T^2 , para analizar si existe una diferencia significativa entre las medias, para cualquier par de grupos.
- **La prueba Φ_j^2** para analizar si la media las FCD Z_j difieren significativamente de grupo en grupo, basada en los valores propios de la matriz $\mathbf{W}^{-1}\mathbf{B}$:

$$\phi_j^2 = \{N - 1 - (p + m) / 2\} \ln(1 + \lambda_j)$$

donde N es el número total de observaciones considerando todos los grupos. Este estadígrafo se compara con la distribución χ^2 con $p+m-2j$ g.d.l.

AMARN 2018 - IMFIA.FI.UDELAR -
Ing. Luis Silveira, Ph.D.

PRUEBAS DE SIGNIFICANCIA

Un valor significativamente grande de ϕ_j^2 indica que Z_j varía de grupo en grupo.

- Examinar las distancias de Mahalanobis, desde las observaciones al centroide de sus grupos. Estas distancias deben aproximarse a una distribución χ^2 con p g.d.l.
Si, en base a la distribución χ^2 , se encuentra que una observación está significativamente alejada del centroide de su grupo, conviene preguntarse si la observación realmente proviene de este grupo.

AMARN 2018 - IMFIA.FI.UDELAR -
Ing. Luis Silveira, Ph.D.

SUPOSICIONES DEL AFD

Los métodos discutidos se basan en dos supuestos:

- La matriz de covarianza de la población, dentro del grupo, debe ser la misma para todos los grupos. (Si esta suposición no es correcta, el AFD no funciona muy bien).
- Para las pruebas de significancia, los datos dentro de los grupos deben seguir una distribución multinormal.

La falla de una o más suposiciones no necesariamente significa que el AFD constituya una pérdida de tiempo. Una discriminación excelente puede ser posible con datos que no siguen una distribución multinormal, aunque puede que no sea fácil establecer la significación estadística de las diferencias grupales.

AMARN 2018 - IMFIA.FI.UDELAR -
Ing. Luis Silveira, Ph.D.

APLICACIONES DEL AFD

Los valores propios de $\mathbf{W}^{-1}\mathbf{B}$ son: $\lambda_1 = 0,437$; $\lambda_2 = 0,035$; $\lambda_3 = 0,015$ y $\lambda_4 = 0,002$. Las correspondientes funciones canónicas

discriminantes son: $Z_1 = -0,0107X_1 + 0,0040X_2 + 0,0119X_3 - 0,0068X_4$

$$Z_2 = 0,0031X_1 + 0,0168X_2 - 0,0046X_3 - 0,0022X_4$$

$$Z_3 = -0,0068X_1 + 0,0010X_2 + 0,0000X_3 + 0,0247X_4$$

$$Z_4 = 0,0126X_1 - 0,0001X_2 + 0,0112X_3 + 0,0054X_4$$

Debido a que λ_1 es mucho mayor que los otros valores propios, es evidente que la mayoría de las diferencias de la muestra se describen sólo con Z_1 .

Las variables X_i ($i=1$ a 4) son los valores mostrados en la Tabla 1.2 sin estandarización.

AMARN 2018 - IMFIA.FI.UDELAR -
Ing. Luis Silveira, Ph.D.

APLICACIONES DEL AFD

Medias y desviación estándar para la función canónica discriminante Z_1 para las cinco muestras de cráneos egipcios

Muestra	Media	Desv. Std.
Predinástico temprano	-0,029	0,097
Predinástico tardío	-0,043	0,071
12 y 13 dinastías	-0,099	0,075
Ptolomeico	-0,143	0,08
Romano	-0,167	0,095

Se puede ver claramente que la media de Z_1 disminuye con el paso del tiempo.

AMARN 2018 - IMFIA.FI.UDELAR -
Ing. Luis Silveira, Ph.D.

APLICACIONES DEL AFD

Los valores de la **prueba de significancia** son:

$$\begin{aligned}\phi_1^2 &= 52,40 & 7 & \text{ g.d.l.} \\ \phi_2^2 &= 5,03 & 5 & \text{ g.d.l.} \\ \phi_3^2 &= 2,16 & 3 & \text{ g.d.l.} \\ \phi_4^2 &= 0,30 & 1 & \text{ g.d.l.}\end{aligned}$$

g.d.l.: $p+m-2j = 4+5-2j$

Solamente ϕ_1^2 es significativo para un nivel de significancia del 5%.

Conclusión: Las diferencias entre las 5 muestras se describen adecuadamente con solamente Z_1 .

AMARN 2018 - IMFIA.FI.UDELAR -
Ing. Luis Silveira, Ph.D.

APLICACIONES DEL AFD

Reclasificación según las distancias de Mahalanobis:

Grupo fuente	Reclasificación					Total
	1	2	3	4	5	
1	12	8	4	4	2	30
2	10	8	5	4	3	30
3	4	4	15	2	5	30
4	3	3	7	5	12	30
5	2	4	4	9	11	30

Solamente 51 de 150 cráneos (34%) se reclasifica en el grupo al que realmente pertenece. **Conclusión: Si bien el AFD es exitoso para señalar los cambios que se producen con el tiempo en las dimensiones de los cráneos, no es un método satisfactorio para la "datación" de los cráneos.**

AMARN 2018 - IMFIA.FI.UDELAR -
Ing. Luis Silveira, Ph.D.

APLICACIONES DEL AFD

Ej. 7.3 Discriminación entre grupos de países Europeos

Datos: Tabla 1.5

Ejemplo 5.2: A.C.P.

Ejemplo 6.1: A.F.

AFD : ¿En que medida es posible discriminar entre grupos de países en base a los patrones de empleo?

AMARN 2018 - IMFIA.FI.UDELAR -
Ing. Luis Silveira, Ph.D.

APLICACIONES DEL AFD

Cuando se recolectaron los datos existían 4 grupos naturales:

1) Unión Europea (UE): Bélgica, Dinamarca, Francia, Alemania, Grecia, Irlanda, Italia, Luxemburgo, Países Bajos, Portugal, España y el Reino Unido.

2) Zona de Libre Comercio Europea (AELC): Austria, Finlandia, Islandia, Noruega, Suecia y Suiza.

3) Los países del Este: Albania, Bulgaria, Republicas Checo/Eslovaca, Hungría, Polonia, Rumania, la ex URSS y la ex Yugoslavia.

4) Los restantes países: Chipre, Gibraltar, Malta y Turquía.

Estos cuatro agrupamientos pueden utilizarse como base en un AFD.

AMARN 2018 - IMFIA.FI.UDELAR -
Ing. Luis Silveira, Ph.D.

APLICACIONES DEL AFD

El porcentaje de los 9 grupos industriales suma 100% para cada uno de los 30 países. Por lo tanto, cualquiera de los 9 variables porcentuales puede expresarse como 100 menos las restantes variables.

Por consiguiente, es necesario omitir una de las variables para poder calcular las funciones canónicas discriminantes.

La última variable, el porcentaje de empleados en Transporte y Comunicaciones se omite en el análisis que se describe a continuación:

AMARN 2018 - IMFIA.FI.UDELAR -
Ing. Luis Silveira, Ph.D.

APLICACIONES DEL AFD

Número de variables canónicas: 3 {=Mín (nro. de variables, $p=8$; nro. de grupos - 1 = $m-1 = 4-1 = 3$)}

Las **variables canónicas** son:

$$Z_1 = 0,427AGR + 0,295MIN + 0,359MAN + 0,339PS + 0,222CON + 0,688SER + 0,464FIN + 0,514SPS$$

$$Z_2 = 0,674AGR + 0,579MIN + 0,550MAN + 1,576PS + 0,682CON + 0,658SER + 0,349FIN + 0,682SPS$$

$$Z_3 = 0,732AGR + 0,889MIN + 0,873MAN + 0,410PS + 0,524CON + 0,895SER + 0,714FIN + 0,764SPS$$

Diferentes paquetes computacionales pueden generar estas variables canónicas con todos los signos invertidos para una o más de las variables. También puede ser deseable invertir los signos.

Es importante notar que se utilizan los porcentajes originales en estas ecuaciones, y **NO** porcentajes estandarizados.

AMARN 2018 - IMFIA.FI.UDELAR -
Ing. Luis Silveira, Ph.D.

APLICACIONES DEL AFD

Los **valores propios** de $W^{-1}B$ son: $\lambda_1 = 5,349$; $\lambda_2 = 0,570$ y $\lambda_3 = 0,202$. Por consiguiente, la primer variable canónica es la más importante.

Debido a que todos los coeficientes son positivos para las tres variables canónicas, es difícil interpretar lo que exactamente significan en términos de las variables originales. En este sentido, es útil considerar las **correlaciones entre las variables originales y las canónicas**.

AMARN 2018 - IMFIA.FI.UDELAR -
Ing. Luis Silveira, Ph.D.

APLICACIONES DEL AFD

Correlaciones entre las variables originales y las canónicas

Grupo	Z_1	Z_2	Z_3
AGR	-0,50	0,37	0,09
MIN	-0,62	0,03	0,20
MAN	-0,02	-0,20	0,12
PS	0,17	0,18	-0,23
CON	0,14	0,26	-0,34
SER	0,82	-0,01	0,08
FIN	0,61	-0,36	-0,09
SPS	0,56	-0,19	-0,28
TC	-0,22	-0,47	-0,41

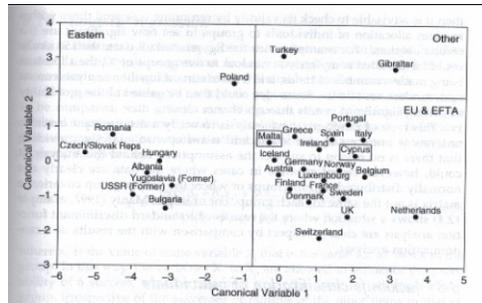
Z_1 tiene correlaciones $> 0,5$ para SER, FIN y SPS, y una correlación $< -0,5$ para AGR y MIN. Por lo tanto, Z_1 representa tipos de servicios de la industria en lugar de industrias tradicionales.

Z_2 No tiene correlaciones importantes, positivas o negativas, con las variables originales. Sin embargo, por las mayores correlaciones, parece representar AGR y CON, con una ausencia de TC y FIN.

Z_3 no muestra correlaciones importantes, pero representa, si algo, una ausencia de TC y CON.

AMARN 2018 - IMFIA.FI.UDELAR -
Ing. Luis Silveira, Ph.D.

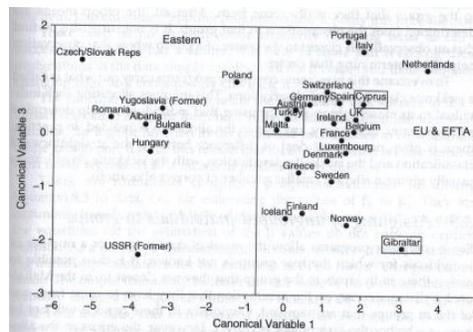
APLICACIONES DEL AFD



Distingue entre los países del este y los otros grupos. No hay separación clara entre los países de la UE y de la AELC, con Malta y Chipre en el mismo grupo. Turquía y Gibraltar del "otro" grupo de países aparecen en la parte superior a la derecha. La mayor separación ocurre con los valores horizontales de Z_1 . Sobre la base de la interpretación de Z_1 , parece que en los países del este predominan las industrias tradicionales en lugar de servicios, mientras que lo contrario tiende a ser cierto para los demás países. Del mismo modo, Turquía y Gibraltar se destacan por AGR y CON en lugar de TC y FIN. Para Gibraltar, al parecer no predomina AGR, pero un porcentaje muy alto en CON.

AMARN 2018 - IMFIA.FI.UDELAR -
Ing. Luis Silveira, Ph.D.

APLICACIONES DEL AFD



La representación gráfica de Z_3 contra Z_1 no muestra una separación vertical de la UE, la AELC y otros grupos de países. Sin embargo, hay algunos patrones obvios, como los países escandinavos que aparecen muy juntos.

AMARN 2018 - IMFIA.FI.UDELAR -
Ing. Luis Silveira, Ph.D.

APLICACIONES DEL AFD

Conclusión:

- El AFD ha tenido éxito en separar a los países del este de los otros, con menos éxito en la separación de los otros grupos.
- La separación es quizás más clara que la que se obtuvo usando ACP.

AMARN 2018 - IMFIA.FI.UDELAR -
Ing. Luis Silveira, Ph.D.

APLICACIONES DEL AFD

El mismo ejemplo. Discriminación entre grupos de países Europeos con XLSTAT:

Valores propios:

	F1	F2	F3
Valor propio	5,349	0,570	0,202
Discriminaci	87,391	9,310	3,299
% acumulad	87,391	96,701	100,000

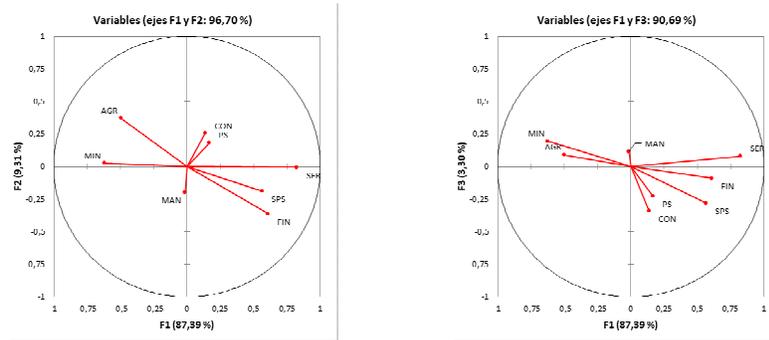
Correlaciones Variables/Factores:

	F1	F2	F3
AGR	-0,50	0,37	0,09
MIN	-0,62	0,03	0,20
MAN	-0,02	-0,20	0,12
PS	0,17	0,18	-0,23
CON	0,14	0,26	-0,34
SER	0,82	-0,01	0,08
FIN	0,61	-0,36	-0,09
SPS	0,56	-0,19	-0,28

AMARN 2018 - IMFIA.FI.UDELAR -
Ing. Luis Silveira, Ph.D.

APLICACIONES DEL AFD

El mismo ejemplo con XLSTAT:



AMARN 2018 - IMFIA.FI.UDELAR -
Ing. Luis Silveira, Ph.D.

APLICACIONES DEL AFD

El mismo ejemplo con XLSTAT:

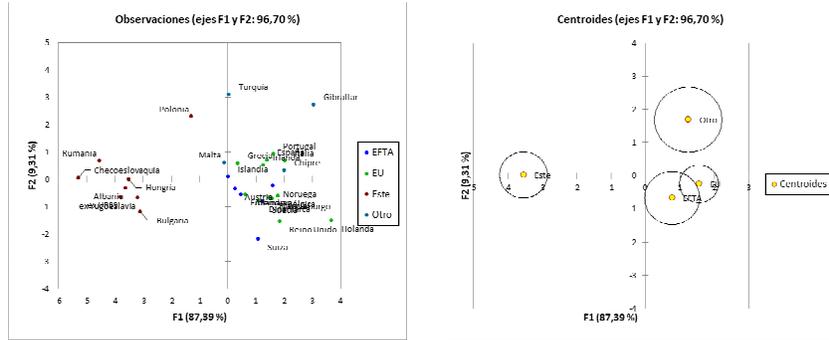
Coefficientes de las funciones discriminantes canónicas:

	F1	F2	F3
Intercepción	-43,156	-58,777	-73,318
AGR	0,427	0,674	0,732
MIN	0,295	0,579	0,889
MAN	0,359	0,550	0,873
PS	0,339	1,576	0,410
CON	0,222	0,682	0,524
SER	0,688	0,658	0,895
FIN	0,464	0,349	0,714
SPS	0,514	0,682	0,764

AMARN 2018 - IMFIA.FI.UDELAR -
Ing. Luis Silveira, Ph.D.

APLICACIONES DEL AFD

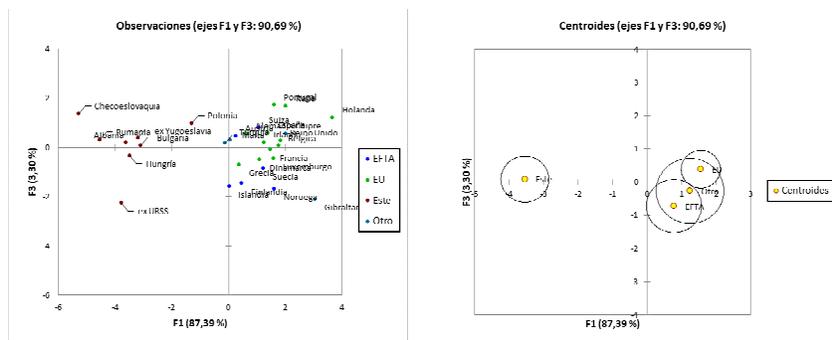
El mismo ejemplo con XLSTAT:



AMARN 2018 - IMFIA.FI.UDELAR -
Ing. Luis Silveira, Ph.D.

APLICACIONES DEL AFD

El mismo ejemplo con XLSTAT:



AMARN 2018 - IMFIA.FI.UDELAR -
Ing. Luis Silveira, Ph.D.

OPCIONES PARA VARIAR EL AFD

❖ **Permitiendo probabilidades a-priori de pertenecer a un grupo.**

Opción de paquetes computacionales, en caso que existe mayor probabilidad de pertenecer a un cierto grupo.

❖ **AFD paso a paso**

En este caso, las variables se adicionan a las FD una a una, hasta que se constata que la incorporación de una nueva variable no da lugar a una discriminación significativamente mejor.

Problema: Sesgo que el procedimiento introduce en las pruebas de significancia. Si se ejecuta un análisis **paso a paso**, es recomendable verificar su validez re-ejecutándolo varias veces, con una ubicación aleatoria de las observaciones en grupos, para verificar cuán significativos son los resultados obtenidos.

AMARN 2018 - IMFIA.FI.UDELAR -
Ing. Luis Silveira, Ph.D.

OPCIONES PARA VARIAR EL AFD

❖ **Clasificación de las observaciones denominada "cuchillo de Jack"**

Cada observación se ubica en el grupo más próximo, sin utilizar dicha observación para determinar el centro del grupo. Evita el sesgo en la ubicación.

AMARN 2018 - IMFIA.FI.UDELAR -
Ing. Luis Silveira, Ph.D.

MODELOS ALTERNATIVOS DEL AFD NO TRATADOS EN EL TEXTO

La matriz de covarianza, dentro del grupo, es la misma para todos los grupos. (Si no es así el AFD no funciona muy bien)

Modelo lineal o cuadrático

Dos modelos de AFD se utilizan en función de un supuesto básico: si las **matrices de covarianza son idénticas**, se utiliza el **análisis discriminante lineal**. Si, por el contrario, **las matrices de covarianza difieren al menos en dos grupos**, se utiliza el **modelo cuadrático**. La prueba de Box se utiliza para probar esta hipótesis. Se comienza con el análisis discriminante lineal y, en función de los resultados de la prueba de Box, se realiza un análisis cuadrático si es necesario.

AMARN 2018 - IMFIA.FI.UDELAR -
Ing. Luis Silveira, Ph.D.

AFD - RESUMEN

- ✓ Se describió el problema de **separar las observaciones en diferentes grupos** sobre la base de las mediciones que se tienen para p variables.
- ✓ Una **solución** a este problema implica asignar cada observación el grupo más cercano en base al **cálculo de distancias de Mahalanobis**, y luego calcular qué proporción de observaciones resultan mal clasificadas.
- ✓ Un enfoque alternativo – **funciones canónicas discriminantes** - atribuido a Fisher (1936) se basa en encontrar la combinación lineal de las p variables que maximiza las diferencias entre los grupos la razón F en un análisis de varianza unidireccional.

AMARN 2018 - IMFIA.FI.UDELAR -
Ing. Luis Silveira, Ph.D.