

# Análisis Multivariado I

## EJERCICIOS DISCRIMINANTE

### Ejercicio 1

El siguiente ejercicio consiste en la aplicación del análisis discriminante a un juego de datos de tipo financiero. La base consiste en 46 empresas de las cuales se extrajo información contable y la condición actual de la empresa, es decir si la misma quebró o no quebró.

Las variables utilizadas para el análisis son:

- FLUJO DE CAJA/DEUDA TOTAL (FC\_DT),
- INGRESO NETO/ACTIVO TOTAL(IN\_AT),
- ACTIVO CORRIENTE/PASIVO CORRIENTE(AC\_PC),
- ACTIVO CORRIENTE/VENTAS NETAS(AC\_VN),
- GRUPO=1, QUEBRÓ MPRESA, GRUPO=2 NO QUEBRÓLA EMPRESA

Se presenta a continuación el output del Análisis discriminante lineal sobre el cual debe responder preguntas.

#### TEST DE MULTINORMALIDAD

```
grupos:0
beta1hat  kappa1      pvalsim beta2hat  kappa2  pvalkurt  n
9.505884  39.60785  0.05597422  28.88862  1.764028  0.07772735  25
-----
grupos:1
beta1hat  kappa1      pvalsim beta2hat  kappa2  pvalkurt  n
10.29878  36.04575  0.1519176  28.81927  1.593825  0.1109752  21
```

#### Tests for the Equality of Means:

Group Variable: GRUPO

	Statistics	F	df1	df2	Pr
Wilks Lambda	0.51248	9.7509	4	41	0.000012289
Pillai Trace	0.48752	9.7509	4	41	0.000012289
Hotelling-Lawley Trace	0.95130	9.7509	4	41	0.000012289
Roy Greatest Root	0.95130	9.7509	4	41	0.000012289

\* Tests assume covariance homoscedasticity.

F Statistic for Wilks' Lambda is exact.

F Statistic for Roy's Greatest Root is an upper bound.

**Constants:**

0 1  
-7.394962 -5.202876

**Linear Coefficients:**

0 1  
Fc.dt 5.36265 4.08154  
In.AT -9.98536 -18.59032  
Ac.pc 3.30209 1.62405  
AC.VN 10.06639 12.24185

**Classification Results<sup>b,c</sup>**

			Predicted Group Membership		Total
			,000	1,000	
Original	Count	,000	23	2	25
		1,000	3	18	21
	%	,000	92,0	8,0	100,0
		1,000	14,3	85,7	100,0
Cross-validated <sup>a</sup>	Count	,000	23	2	25
		1,000	4	17	21
	%	,000	92,0	8,0	100,0
		1,000	19,0	81,0	100,0

a. Cross validation is done only for those cases in the analysis. In cross validation, each case is classified by the functions derived from all cases other than that case.

b. 89,1% of original grouped cases correctly classified.

c. 87,0% of cross-validated grouped cases correctly classified.

### Responda las siguientes preguntas:

1. ¿Qué indica el test de multinormalidad? Plantee cual es la hipótesis nula, cual es el resultado del test y como consecuencia del mismo cual sería el método discriminante más apropiado. Con la información existente que análisis discriminante recomendaría?
2. ¿En qué caso es aplicable la regla discriminante lineal?
3. ¿Qué información aporta el siguiente cuadro? ¿Cómo puede ser utilizada la misma?

Discriminant Analysis      Linear Discriminant Function

#### Constants:

	0	1
	-7.394962	-5.202876

#### Linear Coefficients:

	0	1
Fc.dt	5.36265	4.08154
In.AT	-9.98536	-18.59032
Ac.pc	3.30209	1.62405
AC.VN	10.06639	12.24185

4. ¿En qué consiste el error de **crossvalidation**? ¿Cuál es el nivel de error en este caso?
5. ¿Cuántas empresas fueron reclasificadas? Plantee la regla que se utilizó para determinar la clasificación de las distintas empresas.
6. ¿Cómo podrían usarse los resultados obtenidos si se tienen los datos de una nueva empresa y se quiere determinar a que grupo pertenece: quiebra o no quiebra?

## Ejercicio 2

**Se desea hacer un análisis discriminante. Con la información que se presenta a continuación determine que Análisis Discriminante recomendaría. Fundamente.**

### Test de Igualdad de Medias

	Statistics	F	df1	df2	Pr
Wilks Lambda	0.4307737	32.53959	8	197	0
Pillai Trace	0.5692263	32.53959	8	197	0
Hotelling-Lawley Trace	1.3214046	32.53959	8	197	0
Roy Greatest Root	1.3214046	32.53959	8	197	0

### Test de Mardia de Multinormalidad por grupos

grupos:1

beta1hat	kappa1	pvalsim	beta2hat	kappa2	pvalkurt	n
83.7802	1424.263	0.00005	137.5079	22.9582	0.0003	102

-----  
grupos:2

beta1hat	kappa1	pvalsim	beta2hat	kappa2	pvalkurt	n
163.0215	2825.706	0.00004	243.7311	66.00213	0.00004	104

### Test de Box de Homogeneidad de Varianzas

	Statistic	df	Pr
Box.M	1392.345	36	0
adj.M	1335.082	36	0

### Ejercicio 3 (Tomado de año 2015)

- Se desea discriminar entre dos poblaciones normales con vector de medias  $(0, 0)$  y  $(1, 1)$  y varianzas  $(2, 4)$  con coeficiente de correlación lineal  $\rho = 0,8$ . Construir la función discriminante e interpretarla.
- Las probabilidades a priori son 0.7 para la primera población y 0.3 para la segunda. Calcular la función discriminante en cada caso.
- Si los costes de equivocarse no son los mismos, de manera que el costo de clasificar en la segunda población cuando viene de la primera es el doble, calcular las funciones discriminantes