

## TD Analyse discriminante

Une firme qui vend des systèmes d'alarmes pour entreprises a mené une analyse discriminante sur un échantillon de 20 entreprises. Cette discrimination de l'achat ou non d'un système d'alarme (*achalarm*) a été menée selon les variables suivantes :

- Nombre d'ouvriers (*nbouv*)
- Le capital de la société en milliers de dinars (*capital*)
- La valeur moyenne de son stock en milliers de dinars (*valstock*)
- Ancienneté ou nombre d'année d'existence de la société (*nbannée*)
- L'âge du PDG de la société (*agepdg*)

Les données qui ont été utilisées pour cette analyse discriminante sont les suivantes :

<i>Code entreprise</i>	<i>achalarm</i>	<i>nbouv</i>	<i>capital</i>	<i>valstock</i>	<i>nbannée</i>	<i>agepdg</i>
1	OUI	350	100	40	10	50
2	OUI	230	150	50	9	40
3	OUI	60	200	60	5	60
4	OUI	200	250	30	12	50
5	OUI	100	200	35	10	45
6	OUI	300	300	50	10	50
7	OUI	140	90	30	2	41
8	OUI	80	80	20	7	42
9	NON	100	20	5	15	55
10	NON	150	10	3	20	60
11	NON	200	30	3	30	70
12	NON	300	50	6	25	65
13	NON	130	25	7	29	63
14	NON	50	10	1	32	67
15	NON	320	100	3	25	56
16	NON	100	25	15	24	59
17	NON	190	50	1	22	35
18	NON	200	60	2	20	60
19	NON	120	70	4	30	62
20	NON	140	50	5	15	64

1- Quelle est l'utilité de cette analyse discriminante pour cette entreprise qui vend des systèmes d'alarmes ?

CERNER LE PROFIL DES ACHETEURS POTENTIELS DU SYSTEME D'ALARME  
MIEUX CIBLER NOTRE ACTION MARKETING ET DE PROSPECTION

2- Donner une interprétation exhaustive et détaillée, aux résultats de cette analyse discriminante qui figurent en annexe.

### Test de signification globale de la fonction discriminante

Test pour savoir (H0) si la moyenne de toutes les fonctions discriminantes pour toutes les classes sont égales ou différente (H1).

Ce test sur SPSS s'appuie sur le Lambda de Wilks. On remarque dans ce cas que  $\text{sig} < 0,05$  (on accepte H1) donc il existe une différence statistiquement significative entre les moyennes des fonctions discriminantes du groupe oui (ACHAT ALARME) et le groupe non (NON ACHAT ALARME)

Wilks' Lambda

Test of Function(s)	Wilks' Lambda	Chi-square	df	Sig.
1	,102	35,445	5	,000

### Examen de la corrélation canonique

Comme il n'y a que deux groupes, une seule fonction discriminante est estimée (pour séparer entre deux groupes on a besoin d'une seule frontière). La valeur propre qui lui est associée est égale à 8,843. La corrélation canonique associée à cette fonction est de 0,948. Mais on considère souvent son carrée  $(0,948)^2 = 89,9\%$  pour dire que 89,9% de la variable dépendante achat d'alarme est expliqué par ce modèle

Eigenvalues

Function	Eigenvalue	% of Variance	Cumulative %	Canonical Correlation
1	8,843 <sup>a</sup>	100,0	100,0	,948

a. First 1 canonical discriminant functions were used in the analysis.

### Analyse de la corrélation entre les variables explicatives et aussi le pouvoir discriminant de chaque variable

La matrice de corrélation intra classes ou intra-groupes indique de faibles corrélations entre les variables (entre -0,076 et 0,369) Ainsi le problème de multicolinéarité ne se pose pas.

**Pooled Within-Groups Matrices**

		NBOUVR	CAPITAL	VALSTOCK	NBANNÉE	AGEPDG
Correlation	NBOUVR	1,000	,301	,090	,132	-,076
	CAPITAL	,301	1,000	,369	,197	,158
	VALSTOCK	,090	,369	1,000	-,017	,328
	NBANNÉE	,132	,197	-,017	1,000	,303
	AGEPDG	-,076	,158	,328	,303	1,000

On analyse ensuite la significativité des variables considérées individuellement à travers le test de Fisher des Lambdas de Wilks (principe du (H0) si la moyenne selon cette variable pour toutes les classes sont égales ou (H1) différentes). Ainsi on remarque que toutes les variables ont un fort pouvoir discriminant entre les deux groupes ( $\text{sig} < 0,000$ ) à l'exception de la variable nombre d'ouvriers (car  $\text{sig} = 0,708$  est  $> 0,05$ )

**Tests of Equality of Group Means**

	Wilks' Lambda	F	df1	df2	Sig.
NBOUVR	,992	,145	1	18	,708
CAPITAL	,396	27,475	1	18	,000
VALSTOCK	,192	75,852	1	18	,000
NBANNÉE	,267	49,441	1	18	,000
AGEPDG	,614	11,306	1	18	,003

Le pouvoir discriminant des variables est plus détaillé dans la matrice de structure et par ordre décroissant : Val stock, nb année, capital, age PDG et on trouve en fin de liste la variable nb ouvriers

**Structure Matrix**

	Function
	1
VALSTOCK	,690
NBANNÉE	-,557
CAPITAL	,415
AGEPDG	-,267
NBOUVR	,030

Pooled within-groups correlations between discriminating variables and standardized canonical discriminant functions  
Variables ordered by absolute size of correlation within function.

## Examen de la qualité de la discrimination

L'examen de la qualité de la discrimination se fait à travers l'examen des individus correctement classés par cette fonction. Pour cela on calcule le hit ratio = (nb d'individus bien classés/Nb total d'individus). Dans notre cas =  $(12+8)/20 = 100\%$  des individus sont correctement classés ce qui est une excellente performance pour cette fonction discriminante.

**Classification Results<sup>a</sup>**

			Predicted Group Membership		Total
ACHALARM			NON	OUI	
Original	Count	NON	12	0	12
		OUI	0	8	8
	%	NON	100,0	,0	100,0
		OUI	,0	100,0	100,0

a. 100,0% of original grouped cases correctly classified.

## Analyse des profils des groupes

Groupe 1 (oui) : ceux qui ont acheté une alarme

Valeur moyenne du stock : 39 375 dinars

Ancienneté moyenne de la sté : 8,12 ans

Capital moyen : 171 250

Age moyen du PDG : 47,25 ans

Nombre moyen d'ouvriers : 182,5 ouvriers

Groupe 2 (non) : ceux qui n'ont pas acheté d'alarme

Valeur moyenne du stock : 4 583 dinars

Ancienneté moyenne de la sté : 23,9 ans

Capital moyen : 41 666 dinars

Age moyen du PDG : 59,6 ans

Nombre moyen d'ouvriers : 166,66 ouvriers

Les sociétés ayant une valeur de stock importante, assez récentes, avec un capital important et un PDG plutôt jeune (moins de 50 ans) sont celles les plus susceptibles d'acheter un système d'alarme.

Group Statistics

ACHALARM		Mean	Std. Deviation	Valid N (listwise)	
				Unweighted	Weighted
NON	NBOUVR	166,6667	80,37790	12	12,000
	CAPITAL	41,6667	26,82717	12	12,000
	VALSTOCK	4,5833	3,77692	12	12,000
	NBANNÉE	23,9167	5,72805	12	12,000
	AGEPDG	59,6667	8,88649	12	12,000
OUI	NBOUVR	182,5000	105,93124	8	8,000
	CAPITAL	171,2500	80,07809	8	8,000
	VALSTOCK	39,3750	13,21187	8	8,000
	NBANNÉE	8,1250	3,27054	8	8,000
	AGEPDG	47,2500	6,64938	8	8,000
Total	NBOUVR	173,0000	89,09486	20	20,000
	CAPITAL	93,5000	83,79329	20	20,000
	VALSTOCK	18,5000	19,45169	20	20,000
	NBANNÉE	17,6000	9,27021	20	20,000
	AGEPDG	54,7000	10,04778	20	20,000

L'équation de la fonction discriminante :

$$Z = 2,627 - 0,001 * nbouvr + 0,006 * capital + 0,080 * Valstock - 0,096 * nbannée - 0,051 * agePDG$$

Canonical Discriminant Fur

	Function
	1
NBOUVR	-,001
CAPITAL	,006
VALSTOCK	,080
NBANNÉE	-,096
AGEPDG	-,051
(Constant)	2,627

Unstandardized coefficients

Le barycentre des scores des individus qui ont acheté une alarme est de 3,455 alors que celui de ceux qui n'ont pas acheté est de -2,303. La frontière de séparation entre les deux groupes :  $(3,455 + (-2,303))/2 = 0,576$

Functions at Group Centroids

	Function
ACHALARM	1
NON	-2,303
OUI	3,455

Unstandardized canonical discriminant functions evaluated at group means

3- On dispose des informations suivantes relatives à une société :

<i>nbouvr</i>	<i>capital</i>	<i>valstock</i>	<i>nbannée</i>	<i>agepdg</i>
100	50	50	8	45

Pour mieux cibler l'action commerciale et marketing, on voudrait savoir si cette société serait un acheteur potentiel d'un système alarme

$$Z = 2,627 - 0,001 \cdot \text{nbouvr} + 0,006 \cdot \text{capital} + 0,080 \cdot \text{Valstock} - 0,096 \cdot \text{nbannée} - 0,051 \cdot \text{agePDG}$$

$$Z = 7,795 \text{ (plus proche du barycentre 3,455 (oui))}$$

Donc on peut conclure que cette société constitue un acheteur potentiel d'un système d'alarme

4- Peut-on proposer une amélioration pour cette analyse discriminante ? Expliquer.

Oui on peut refaire cette analyse discriminante en supprimant la variable n'ayant pas un pouvoir discriminant significatif à savoir la variable nombre d'ouvriers.