

## INTRODUCTION NUMÉRIQUE À L'AFD

On considère dix clients d'une banque répartis en trois groupes de risque financier : haut risque (1), risque moyen (2) et risque faible ou nul (3). On connaît pour chaque client l'âge, le revenu familial annuel, le patrimoine, et le montant de l'emprunt obtenu :

	Age	Revenu	Patrimoine	Emprunt	Groupe
1	45	250kF	1300kF	600kF	3
2	47	160kF	1150kF	450kF	2
3	38	165kF	850kF	370kF	1
4	36	175kF	770kF	250kF	1
5	29	99kF	450kF	400kF	1
6	39	170kF	1400kF	120kF	3
7	27	120kF	1400kF	160kF	2
8	51	160kF	1300kF	320kF	3
9	32	155kF	1500kF	350kF	2
10	35	170kF	1400kF	180kF	2

Tableau 1 : unités statistiques *actives*

### 1. ANALYSE STATISTIQUE DESCRIPTIVE DES VARIABLES INITIALES.

#### 1.1 Moyennes et écarts types des variables observées sur la totalité des individus:

Variables	Moyennes	Variances	Écarts types
âge	37.9	55.09	7.422264
revenu	162.4	1383.84	37.2
patrimoine	1152	107936	328.5361
emprunt	320	19520	139.714

Moyennes et écarts types

	âge	revenu	patrimoine	emprunt
âge	1.000			
revenu	0.580	1.000		
patrimoine	0.193	0.345	1.000	
emprunt	0.419	0.430	-0.255	1.000

Matrice des corrélations

#### 1.2 Moyennes des variables observées sur chaque groupe d'individus:

	Groupe 1	Groupe 2	Groupe 3
âge	34.33333	35.25	45
revenu	146.3333	151.25	193.3333
patrimoine	690	1362.5	1333.333
emprunt	340	285	346.6667

Moyennes par groupe

### 1.3 Décomposition de la variance et rapports de corrélation

variance	totale	expliquée	résiduelle	rapport de corrélation
âge	55.09	21.74817	33.34183	0.395
revenu	1383.84	414.2324	969.6075	0.299
patrimoine	107936	91621.75	16314.25	0.849
emprunt	19520	823.3281	18696.67	0.042

#### Décomposition de la variance

Le rapport de corrélation le plus élevé est celui du patrimoine, dont le banquier tient compte de façon prépondérante dans l'évaluation du risque financier.

## 2. ANALYSE FACTORIELLE DISCRIMINANTE.

### 2.1 Résultats numériques

Les résultats de l'analyse discriminante sur les individus du tableau 1 sont les suivants :

n°	Groupe	axe 1	axe 2	n°	Groupe	axe 1	axe 2
1	3	1.3541	-1.5318	6	3	1.5447	-0.7005
2	2	1.4703	-0.6229	7	2	2.2553	2.2396
3	1	-3.1979	-0.4295	8	3	2.7347	-1.4256
4	1	-5.2125	-1.0479	9	2	3.8648	1.7423
5	1	-6.3944	1.6874	10	2	1.5809	0.0889

On vérifiera que les composantes discriminantes sont de moyenne nulle, qu'elles sont non corrélées et on calculera les variances totales.

### 2.2 Centres de gravité

Pour calculer les coordonnées des centres de gravité des groupes sur les axes 1 et 2, il suffit de calculer les moyennes des coordonnées des unités statistiques qui leur appartiennent.

On trouve :

Groupe	Axe 1	Axe 2
1	-4.9349	0.0700
2	2.2928	0.8620
3	1.8778	-1.2193

On vérifiera que les moyennes pondérées des coordonnées des centres de gravité sur chaque axe sont nulles et on calculera les variances expliquées des composantes discriminante.

### 2.3 Pouvoirs discriminants

Les pouvoirs discriminants sont les rapports de corrélation des composantes discriminantes:

	Pouvoirs discriminants
Première composante discriminante	$\eta^2 = 0.9128$

Seconde composante discriminante	$\eta^2 = 0.4268$
----------------------------------	-------------------

### 2.3 Plan discriminant

On représente les individus sur le plan discriminants 1x2 :

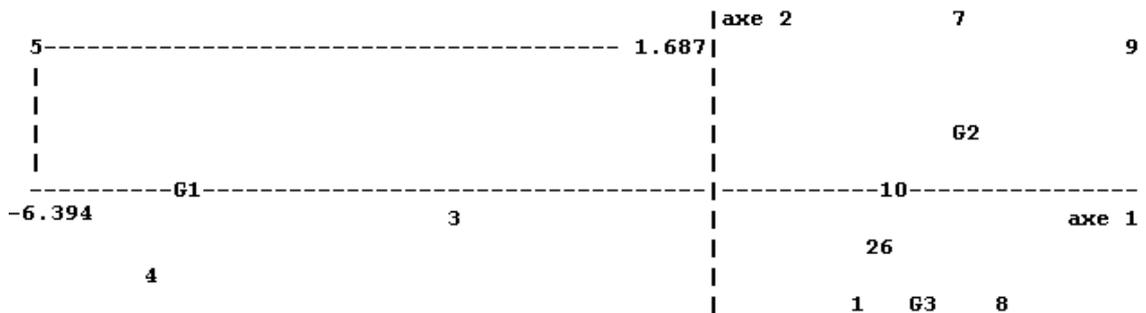


Figure1 : Plan discriminant 1x2

### 2.4 Corrélation des composantes discriminantes et des variables initiales.

Ces coefficients de corrélation donnent l'interprétation des composantes discriminantes :

	CD <sub>1</sub>	CD <sub>2</sub>
âge	0.299	-0.857
revenu	0.269	-0.739
patrimoine	0.964	-0.019
emprunt	-0.108	-0.272

L'importance du patrimoine dans la première composante discriminante est évidente. L'âge et le revenu expliquent la seconde composante discriminante. L'emprunt n'intervient guère : on retrouve ici plus ou moins les rapports de corrélation ; ce n'est pas nécessairement vrai dans le cas général.

### 3. RÈGLE D'AFFECTION.

On décide d'affecter les individus au groupe dont le centre de gravité leur est le plus proche. On peut effectuer cette affectation en examinant le plan discriminant 1x2 ci-dessus ou en effectuant les calculs :

- Les individus 3, 4 et 5 sont affectés au groupe 1.
- Les individus 7, 9 et 10 sont affectés au groupe 2.
- Les individus 1, 2, 6 et 8 sont affectés au groupe 3.

Le calcul des distances est élémentaire : les axes discriminants étant orthonormés, le carré de la distance entre deux points est la somme des carrés des différences de leurs coordonnées.

individu 1	1.3541	-1.5318
centre de gravité de G <sub>3</sub>	1.8778	-1.2193

$$d^2(1, G_3) = (1.3541 - 1.8778)^2 + (-1.5318 - (-1.2193))^2 = 0.3719$$

Un seul individu sur les dix est mal classé. Le pourcentage de bien classés est donc de 90%. Comme il s'agit du fichier de calcul, ce pourcentage doit être interprété avec prudence, et il est préférable d'effectuer ce calcul sur les individus du fichier test ci-dessous :

	Age	Revenu	Patrimoine	Emprunt	Groupe
1	35	170kF	1100kF	400kF	2
2	37	160kF	750kF	350kF	2
3	48	195kF	1150kF	570kF	3
4	26	95kF	470kF	250kF	1
5	39	99kF	350kF	350kF	1
6	29	120kF	300kF	220kF	3
7	47	140kF	1400kF	360kF	3
8	51	160kF	1900kF	320kF	3
9	42	255kF	1200kF	350kF	1
10	35	270kF	1150kF	280kF	2

Tableau 2 : unités statistiques tests

La démarche est exactement la même. Les résultats donnés par le logiciel sont les suivants :

n°	Groupe	affectation	axe 1	axe 2	n°	Groupe	affectation	axe 1	axe 2
1	2	2	-0.5674	0.4519	6	3	1	-9.6954	0.1616
2	2	1	-4.3785	-0.4508	7	3	2	4.3465	-0.0685
3	3	3	1.2491	-1.0650	8	3	2	9.5212	-0.3009
4	1	1	-7.1331	1.5896	9	1	3	-1.5589	-2.4841
5	1	1	-7.0499	-0.0528	10	2	1	-3.4699	-2.2807

Les individus 2, 6, 7, 8, 9 et 10 du fichier test sont mal classés. Le pourcentage de bien classés tombe à 40%.

Il est bien évident que les effectifs sont insuffisants pour que l'analyse discriminante puisse produire des résultats significatifs.

La représentation graphique des individus tests sur le plan discriminant 1x2 est donnée ci-dessous :

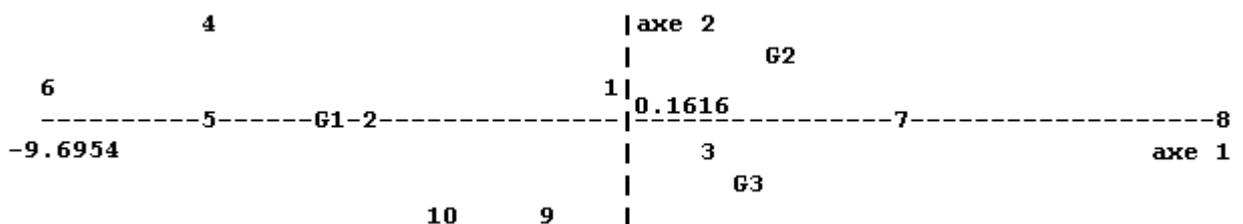


Figure 2 : plan discriminant 1 x 2 (unités tests)

Cette représentation graphique permet de contrôler la validité de la règle de classification en examinant les différentes notions de proximité entre un point et un groupe.