

Analyse des données avancée

Chapitre 16 : L'analyse factorielle discriminante



Jaouad Madkour

Faculté d'Économie et de Gestion de Tanger · 27 août 2026

Un changement de situation

Le critère

Ce que l'axe discriminant révèle

Ce qu'il faut retenir

Un changement de situation

Les groupes sont désormais connus

Jusqu'ici, aucune méthode ne savait à quel groupe appartenait un individu — la classification devait les construire.

L'**analyse factorielle discriminante** part de groupes **donnés**, et cherche ce qui les distingue.

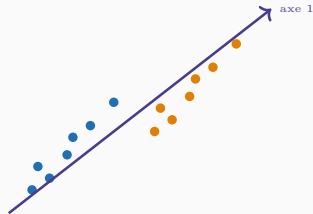
La question posée à nos données

Les clients se répartissent en trois groupes selon leur fréquence de visite : **112** rares, **87** mensuels, **101** hebdomadaires.

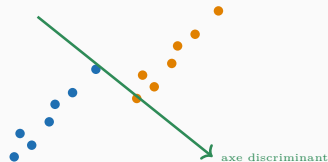
Qu'est-ce qui, dans les huit notes de satisfaction, sépare le client qui revient de celui qui ne revient pas? L'ACP suggérerait une réponse; l'AFD va la mesurer.

ACP et AFD : la différence en une image

ACP : l'axe de plus grande dispersion



AFD : l'axe qui sépare le mieux



Même nuage, même données — deux axes presque perpendiculaires.

L'ACP ignore les couleurs : elle cherche la direction où le nuage s'étale le plus, et la trouve *le long* des deux groupes.

L'AFD connaît les couleurs : elle cherche la direction qui *sépare* les groupes, et la trouve perpendiculaire à la précédente.

Formellement, elle maximise

$$\frac{\text{inertie interclasse}}{\text{inertie intraclasse}}$$

L'ACP est aveugle aux groupes ; l'AFD ne voit qu'eux. Chercher une séparation sur un plan d'ACP est une erreur courante — et la figure de gauche montre pourquoi.

Le critère

Les deux matrices

$$T = B + W$$

B : la dispersion **entre** les centres des groupes. W : la dispersion **à l'intérieur** des groupes. T : la dispersion totale.

Ce que l'AFD maximise

$$\max_u \frac{u' Bu}{u' Wu}$$

Une séparation nette suppose deux choses à la fois : des centres éloignés — grand numérateur — et des groupes resserrés autour de leur centre — petit dénominateur.

Un axe où les centres sont loin mais les groupes très étalés ne sépare rien : les nuages se recouvrent.

Encore un problème de vecteurs propres

Les axes discriminants sont les vecteurs propres de $W^{-1}B$, et les valeurs propres μ_k mesurent le **pouvoir discriminant** de chaque axe.

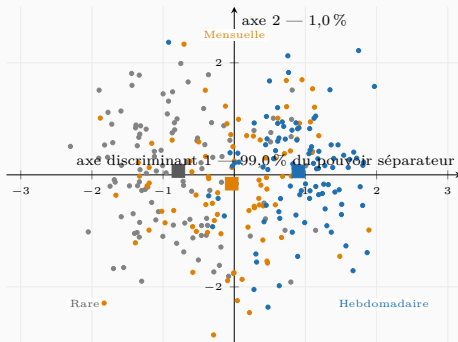
Leur nombre est au plus $\min(K - 1, p)$: avec trois groupes, **deux axes au maximum**.

Sur nos données

$$\mu_1 = 1,0161 \quad \mu_2 = 0,0103$$

Le premier axe porte **99,0 %** du pouvoir séparateur. **La distinction entre les trois groupes est unidimensionnelle** : un seul nombre suffit à situer un client sur l'échelle de la fidélité.

Ce que l'axe discriminant révèle



Un seul axe fait tout le travail.

Le pouvoir discriminant du premier axe vaut 1,016 ; celui du second 0,010. La séparation est **unidimensionnelle**, et les trois groupes s'y ordonnent : $-0,79$, $-0,04$, $+0,90$.

Et cet axe a un sens : ses poids sont 1,00 pour le conseil, 0,90 pour la rapidité, 0,85 pour l'accueil — et $-0,20$ pour le prix.

Ce qui fait revenir un client, c'est le service, pas l'offre. L'ACP suggérerait ce résultat ; l'AFD le mesure.

Les groupes se chevauchent, en revanche : la frontière n'est pas nette, et le taux d'erreur le dira.

Corrélation de chaque note avec l'axe discriminant

Note	r	Note	r
conseil	0,907	agencement	0,446
accueil	0,899	qualité	0,397
rapidité	0,867	choix	0,357
propreté	0,603	prix	0,110

Le classement est sans appel

Les quatre premières places sont occupées par les quatre critères de **service**. Le **prix** arrive dernier, à 0,110 — autant dire rien.

Le prix n'a aucun pouvoir de distinguer un client fidèle d'un client occasionnel. C'est le résultat du chapitre 15, désormais chiffré variable par variable.

Le lambda de Wilks

$$\Lambda = \prod_k \frac{1}{1 + \mu_k} = 0,4909$$

Il varie de 0 — séparation parfaite — à 1 — aucune séparation.

Une exception dans ce cours

L'AFD est la seule méthode du programme qui admette un **cadre inférentiel** : sous hypothèse de normalité multivariée, Λ se transforme en une statistique de Fisher.

On peut donc, ici seulement, tester la séparation. Nous n'en ferons pas usage : l'hypothèse de normalité est douteuse sur des notes discrètes, et $\Lambda = 0,49$ parle assez de lui-même.

Ce qu'il faut retenir

Six points

1. L'AFD part de groupes **connus** et cherche ce qui les sépare.
2. L'ACP maximise l'inertie totale; l'AFD maximise le rapport B/W .
3. Les axes sont les vecteurs propres de $W^{-1}B$; il y en a au plus $\min(K - 1, p)$.
4. Chez nous : $\mu_1 = 1,016$ et $\mu_2 = 0,010$ — la séparation est **unidimensionnelle**.
5. L'axe est fait de **conseil**, **accueil** et **rapidité**; le prix y pèse 0,110.
6. Le **lambda de Wilks** vaut 0,491 : c'est la seule méthode du cours qui admette un test.

La suite

Nous savons ce qui distingue les trois groupes. Reste à s'en servir : classer un client dont on ignore la fréquence de visite.

C'est le passage de la description à la **règle de décision** — et il faudra en mesurer honnêtement les erreurs.