

Majeure Science des données : UP Apprentissage statistique

Liseth Pasaguayo :

liseth.pasaguayo@emse.fr, Bureau 529

Thomas Galtier :

thomas.galtier@emse.fr, Bureau 532

- Pagès, J., & Husson, F. (2016). *Analyse factorielle des correspondances (AFC)* Laboratoire de mathématiques appliquées, Agrocampus Rennes.
- Greenacre, M. (2017). *Correspondence Analysis in Practice* (3.rd ed.). Chapman & Hall/CRC.
<https://doi.org/10.1201/9781315369983>

TP NOTÉ

❖ **Partie 1 : Interprétation et analyse des résultats issus de quelques cas d'AFC.**

- Durée : 1H00
- Date de remise : le 22 octobre à la fin du cours
- Consignes : l'utilisation de l'ordinateur n'est pas autorisée

❖ **Partie 2 : Réalisation et étude détaillée d'un cas d'AFC, avec tous les calculs effectués manuellement(sur papier).**

- Durée : 1H30
- Date de remise : le 23 octobre à la fin du cours
- Consignes : l'utilisation de l'ordinateur n'est pas autorisée

❖ **Partie 3 : Développement et implémentation d'un cas d'AFC, à l'aide des logiciels R ou Python.**

- Durée : 1H30
- Date de remise : le 7 de novembre de 2025

L'approche

1. Les données
2. Notion d'Indépendances entre les variables
3. Tableau de contingence et fréquence
4. Liaison entre deux variables qualitatives
5. Profils *ligne* et *colonne* - profils *moyens*
6. Réduction de dimension
7. Représentations simultanées *ligne et colonne*

Analyse Factorielle des Correspondances (AFC)

- ❖ Son objectif principal est de **réduire la dimension des données** tout en conservant le maximum d'information, afin de représenter les liaisons entre les variables sous forme de graphiques faciles à interpréter.
- ❖ L'AC est particulièrement utilisée dans les enquêtes, le marketing, la sociologie et d'autres domaines où l'on analyse des données catégorielles.

Exemples:

	Américain	Européen	Japonais	Total
Marié	37	14	51	102
Marié avec enfants	52	15	44	111
Célibataire	33	15	63	111
Célibataire avec enfants	6	1	8	15
Total	128	45	166	339

- Quels types de musique sont globalement les plus populaires, et lesquels sont les moins choisis ?
- Existe-t-il une relation entre le groupe d'âge (jeunes, âge moyen, personnes âgées) et le type de musique préféré ?

- Le choix de l'origine de l'automobile dépend-il de l'état familial ou semble-t-il indépendant ?
- Parmi les personnes *mariées*, quel type d'automobile est le plus courant ?

	Jeunes	Âge moyen	Âgées	Total
Musique disco	70	0	0	70
Rock'n'roll et musique américaine	45	45	0	90
Pop et musique anglaise	30	30	30	90
Jazz et musique autochtone	0	80	20	100
Musique classique	35	5	10	50
Total	180	160	60	400

1. Les données

Soit deux **variables qualitatives** observées simultanément sur n individus affectés de poids identiques $1/n$.

La première variable, notée X , possède I **modalités** notées x_1, \dots, x_I ,

la seconde variable, notée Y , possède J modalités notées y_1, \dots, y_J .

- La table de contingence (des correspondances)
associée à ces observations, de dimension $I \times J$:

tableau de contingence

=

tableau des effectifs croisés

		Ensemble J				
		1	j			J
Ensemble I	1					
	i		x_{ij}			
	I					

Où x_{ij} est le nombre d'individus appartenant à la modalité i de la première variable et à j la modalité de la deuxième variable

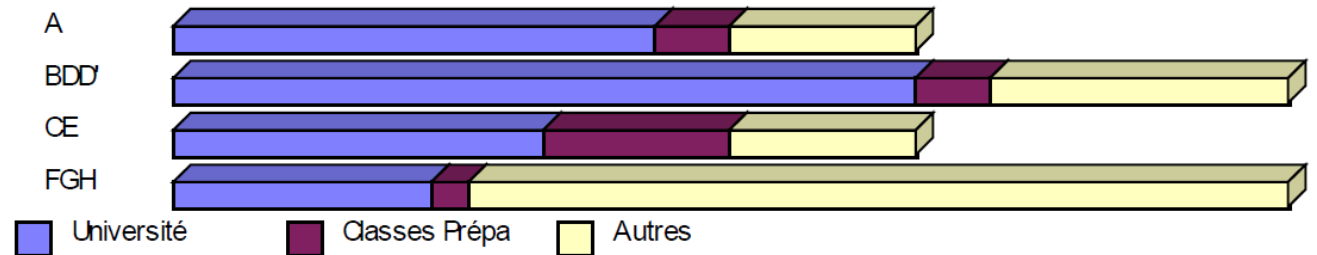
1. Les données

Un exemple : AFC - Pour faire quoi ?

Question : Que deviennent les bacheliers ?

Recensement par discipline/type études.

	destination			total
	université	classes prépa	autres	
A	13	2	5	20
BDD'	20	2	8	30
CE	10	5	5	20
FGH	7	1	22	30
total	50	10	40	100



Questions :

- S'il n'y avait pas de relation préférentielle entre un choix d'études supérieures et le bac obtenu : qu'aurait-on ?
- En d'autres termes, s'il y avait indépendance entre les deux choix consécutifs, qu'aurait-on dans la matrice T_0 des correspondances ?

→ Reconstruire les données en cas d'indépendance (matrice T_0).

Distribution des n individus dans les $I \times J$ cases du tableau de contingence T :

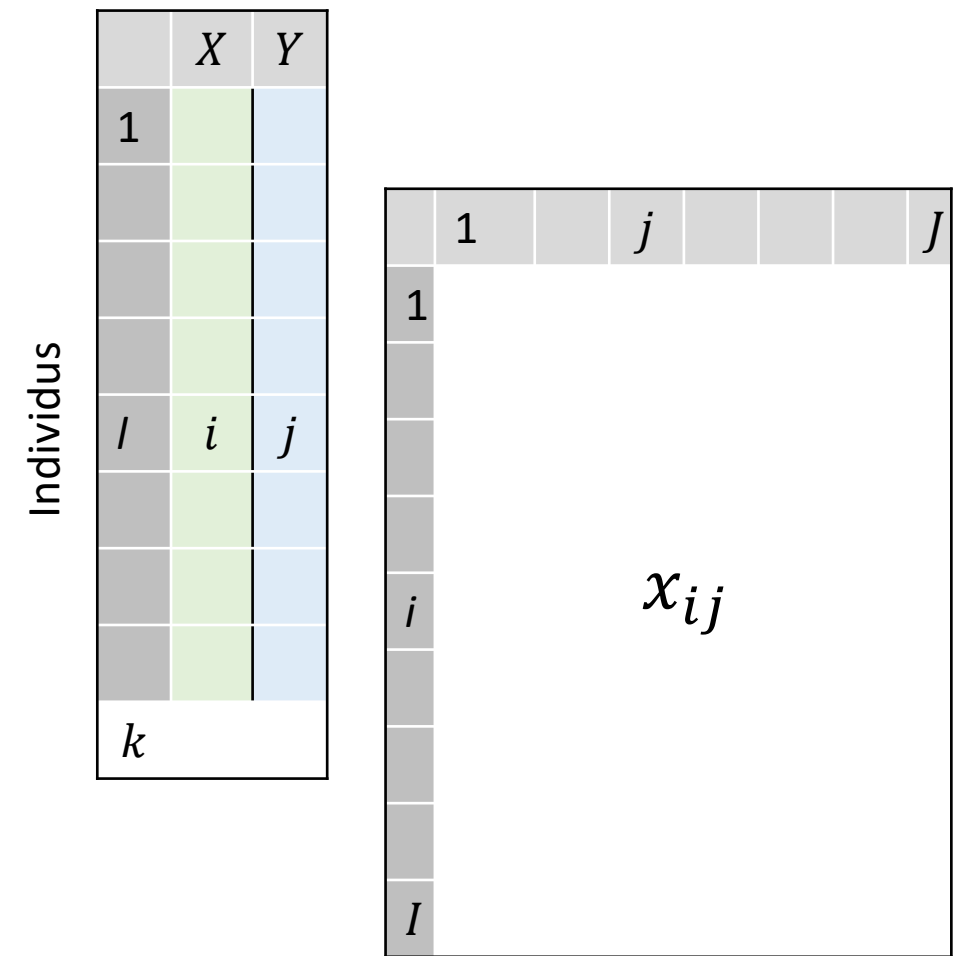


Tableau de contingence : $T = x_{ij}$

	Y1	Y2	Y3	Totaux ligne : $x_{i.}$
X1	13	2	5	20
X2	20	2	8	30
X3	10	5	5	20
X4	7	1	22	30
Totaux col. : $x_{.j}$	50	10	40	100

$x_{.j}$: Marge ligne (profil empirique de Y)

$x_{i.}$: Marge colonne (profil empirique de X)

Ce qui permet de construire
la Matrice de distribution des correspondances si les deux variables qualitatives sont indépendantes

k nombre de réponses

$$k = \sum_{i=1}^I \sum_{j=1}^J x_{ij}$$

Matrice T à la matrice T_0

Tableau de contingence : $T = x_{ij}$

	Y1	Y2	Y3	Marge ligne : $x_{i.}$
X1	13	2	5	20
X2	20	2	8	30
X3	10	5	5	20
X4	7	1	22	30
Marge col. : $x_{.j}$	50	10	40	100

Matrice T_0

En effectifs :

$$T_0 = \hat{x}_{ij} = \left[\frac{x_{i.} \cdot x_{.j}}{k} \right]$$

$$k = \sum_{i=1}^I \sum_{j=1}^J x_{ij} = 100$$

Soit :

$$\hat{x}_{11} = \left[\frac{20 \cdot 50}{100} \right] = 10 \quad \hat{x}_{12} = \left[\frac{20 \cdot 10}{100} \right] = 2 \quad \hat{x}_{13} = \left[\frac{20 \cdot 40}{100} \right] = 8$$

Écart entre les données de la table de contingence T et la table de contingence sous hyp. d'indépendance T_0 :

$$R = T - T_0$$

$$T = \begin{bmatrix} 13 & 2 & 5 \\ 20 & 2 & 8 \\ 10 & 5 & 5 \\ 7 & 1 & 22 \end{bmatrix} \quad T_0 = \begin{bmatrix} 10 & 2 & 8 \\ 15 & 3 & 12 \\ 10 & 2 & 8 \\ 15 & 3 & 12 \end{bmatrix}$$

$$R = \begin{bmatrix} 13 & 2 & 5 \\ 20 & 2 & 8 \\ 10 & 5 & 5 \\ 7 & 1 & 22 \end{bmatrix} - \begin{bmatrix} 10 & 2 & 8 \\ 15 & 3 & 12 \\ 10 & 2 & 8 \\ 15 & 3 & 12 \end{bmatrix} = \begin{bmatrix} 3 & 0 & -3 \\ 5 & -1 & -4 \\ 0 & 3 & -3 \\ -8 & -2 & 10 \end{bmatrix}$$

Si R matrice de zéros : indépendance

2. Indépendance entre deux variables (a)

Modèle d'indépendance

Les événements A et B sont indépendants si:

$$P(A \text{ et } B) = P(A) \times P(B)$$

1) Passer de données du tableau de contingence à des probabilités par calcul de fréquence et marge (col. ou lig.)
table de fréquence.

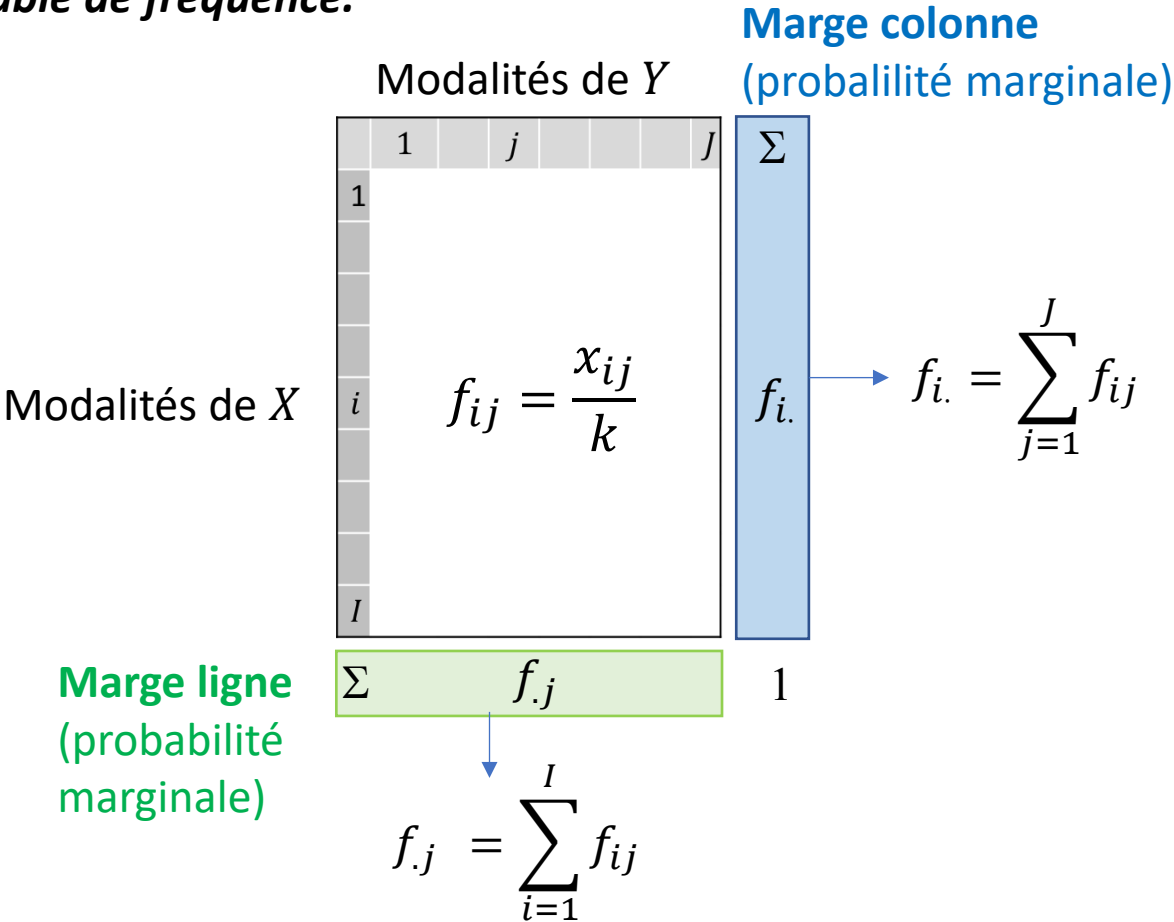


Tableau de fréquences

	Y1	Y2	Y3	Marge col.: $f_{i.}$
X1	0,13	0,02	0,05	0,20
X2	0,20	0,02	0,08	0,30
X3	0,10	0,05	0,05	0,20
X4	0,07	0,01	0,22	0,30
Marge ligne. : $f_{.j}$	0,50	0,10	0,40	1

$$f_{11} = \left[\frac{13}{100} \right] = 0.13 \quad f_{13} = \left[\frac{5}{100} \right] = 0.05$$
$$f_{12} = \left[\frac{2}{100} \right] = 0.02 \quad f_{21} = \left[\frac{20}{100} \right] = 0.2 \quad \dots\dots\dots$$

2. Indépendance entre deux variables (b)

Modèle d'indépendance

Les événements A et B sont indépendants si:

$$P(A \text{ et } B) = P(A) \times P(B)$$

Considérons deux variables qualitatives :

- X avec I modalités \rightarrow lignes,
- Y avec J modalités \rightarrow colonnes.

On définit :

- $A = \{X = i\}$,
- $B = \{Y = j\}$.

$$P(X = i, Y = j) = P(X = i) \cdot P(Y = j)$$

Variables qualitatives indépendantes :

Probabilité conjointe = produit des probabilités marginales, soit :

$$\forall i, \forall j, f_{ij} = f_{i.} \times f_{.j}$$

$$\frac{f_{ij}}{f_{i.}} = f_{.j}$$

$$\frac{f_{ij}}{f_{.j}} = f_{i.}$$

Indépendance si
on a :

\rightarrow Probabilité conditionnelle = probabilité marginale

3. Liaison entre deux variables qualitatives (a)

Écart (distance) entre les données observées

(f_{ij}) et le modèle d'indépendance qui est $(f_{i.} \ f_{.j})$

$$\chi^2_{obs} = \sum_{i=1}^I \sum_{j=1}^J \frac{(eff.obs - eff.théo)^2}{eff\ théo} = \sum_{i=1}^I \sum_{j=1}^J \frac{(k f_{ij} - k f_{i.} f_{.j})^2}{k f_{i.} f_{.j}}$$

$$\chi^2_{obs} = \sum_{i=1}^I \sum_{j=1}^J \frac{(probabilité\ obs - probabilité\ théo)^2}{probabilité\ théo} = k \Phi^2$$

Φ^2 Est l'écart entre probabilité théorique et observée = intensité de liaison (nature de la liaison entre deux var.)

3. Liaison entre deux variables qualitatives (b)

$$\chi^2_{obs} == \sum_{i=1}^I \sum_{j=1}^J \frac{(eff. obs - eff. théo)^2}{eff théo} = \sum_{i=1}^I \sum_{j=1}^J \frac{(k f_{ij} - k f_{i.} f_{.j})^2}{k f_{i.} f_{.j}}$$

	Y1	Y2	Y3
X1	$\frac{(13 - 10)^2}{10} = 0,9$	$\frac{(2 - 2)^2}{2} = 0$	$\frac{(5 - 8)^2}{8} = 1,1250$
X2	$\frac{(20 - 15)^2}{15} = 1,6667$	$\frac{(2 - 3)^2}{3} = 0,3333$	$\frac{(8 - 12)^2}{12} = 1,3333$
X3	$\frac{(10 - 10)^2}{10} = 0$	$\frac{(5 - 2)^2}{2} = 4,5$	$\frac{(5 - 8)^2}{8} = 1,1250$
X4	$\frac{(7 - 15)^2}{15} = 4,2667$	$\frac{(1 - 3)^2}{3} = 1,3333$	$\frac{(22 - 12)^2}{12} = 8,3333$

$$T = \begin{bmatrix} 13 & 2 & 5 \\ 20 & 2 & 8 \\ 10 & 5 & 5 \\ 7 & 1 & 22 \end{bmatrix} \text{ Donées observées}$$

$$T_0 = \begin{bmatrix} 10 & 2 & 8 \\ 15 & 3 & 12 \\ 10 & 2 & 8 \\ 15 & 3 & 12 \end{bmatrix} \text{ Donées théoriques}$$

$$\chi^2_{obs} = 0,9 + 0 + 1,125 + 1,6667 + 0,3333 + 1,333 + 0 + 4,5 + 1,125 + 4,2667 + 1,333 + 8,333 = 24,92$$

$$\chi^2_{obs} = \sum_{i=1}^I \sum_{j=1}^J k \frac{(\text{probabilité obs} - \text{probabilité théo})^2}{\text{probabilité théo}} = k \Phi^2 \quad \Phi^2 = \frac{\chi^2}{k} = \frac{24,917}{100} = 0,2492$$

4. Profil-lignes et profil-colonnes

Passage à la répartition des pourcentages à l'intérieur d'une ligne ou d'une colonne : tableau des profil-lignes et profil-colonnes.

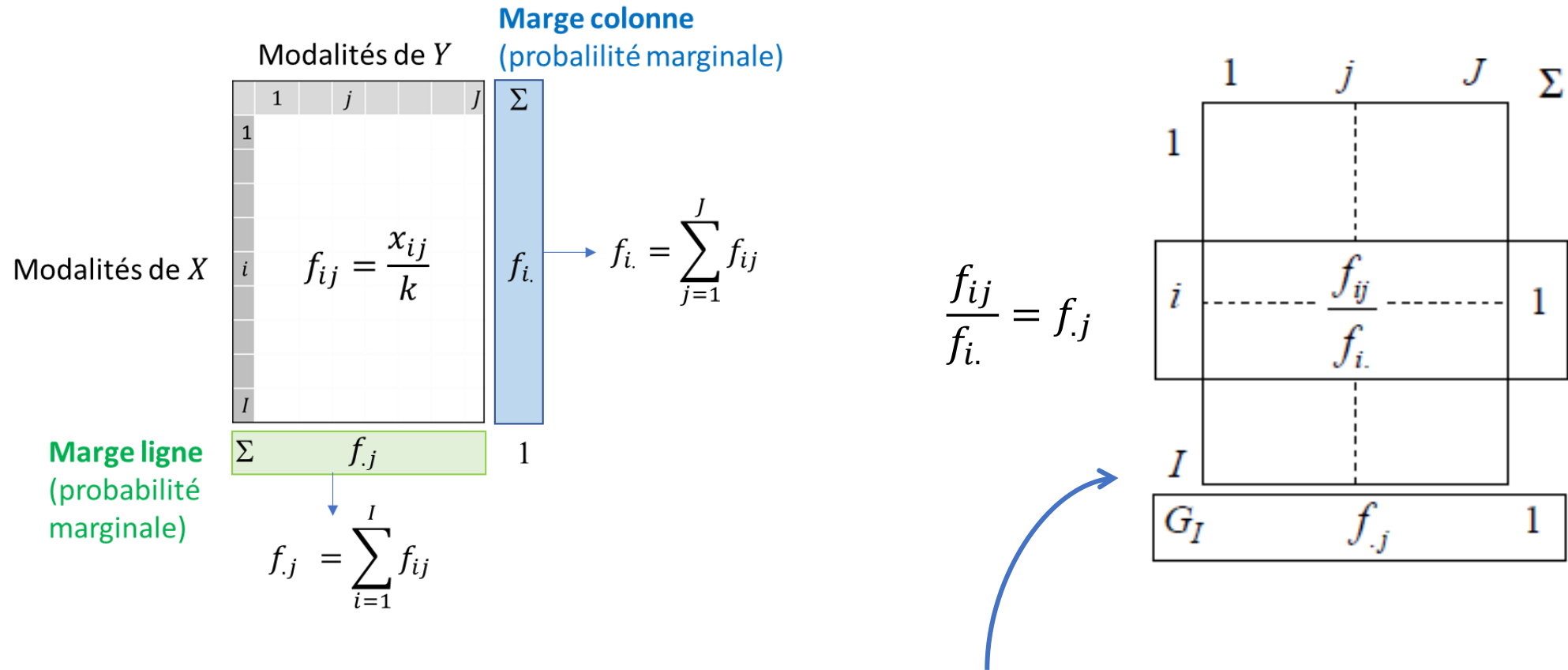
Profils-lignes → décrivent la répartition de chaque ligne (modalité de la variable X) entre les colonnes (modalités de la variable Y).

❖ Question : Pour une modalité de X , comment se répartissent les individus entre les modalités de Y ?

Profils-colonnes → décrivent la répartition de chaque colonne entre les lignes.

❖ Question : Pour une modalité de Y , comment se répartissent les individus entre les modalités de X ?

4.1. Profil-ligne et profil-moyen

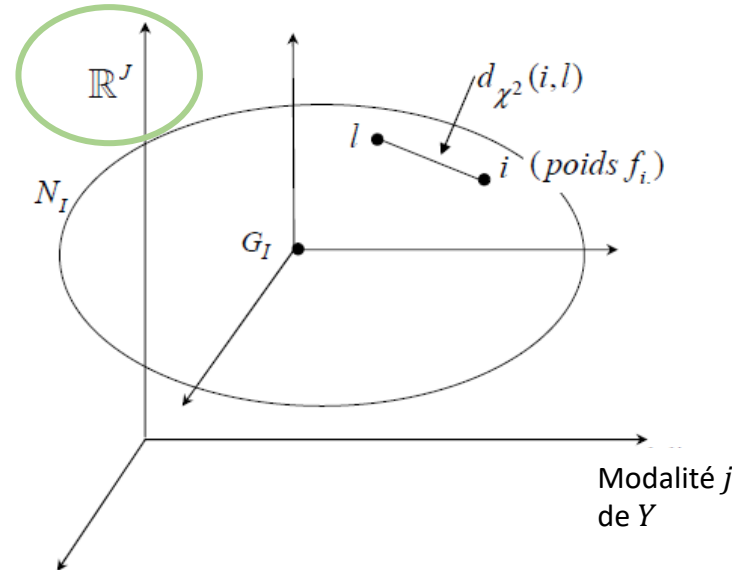
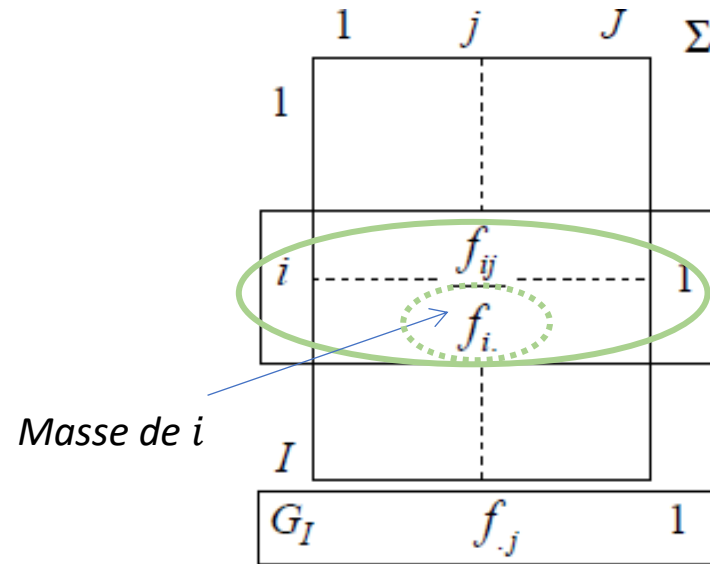


Profil ligne i = distribution conditionnelle : répartition de la var Y par rapport à la modalité i de la var X ; la somme des probabilités conditionnelles d'avoir $Y = j$ sachant que $X = i$, est = 1 (somme de ces probabilités = 1 en ligne)

Profil moyen ligne G_I : répartition sur l'ensemble de la population de la variable Y .

4.1. Profil-ligne et profil-moyen

- Profil ligne : répartition de la variable Y en fonction de la modalité i de la var X
- La proximité entre les points lignes (**espace \mathbb{R}^J**) donc une distance entre les points lignes (profils lignes)



- Un point i est affecté de la **masse** f_i : fréquence relative de i ème modalité de la var X
- Puisque $\sum_{j=1}^J \frac{f_{ij}}{f_{i.}} = 1$, les k points sont situés dans un espace de dim $(J - 1)$
- Le centre de gravité de ce nuage = moyenne des profils lignes affectés de leurs masses correspondantes: $f_{.j}$ est sa j ème composante vaut

4.1. Profil-ligne et profil-moyen-exemple $\left(\frac{f_{ij}}{f_{i.}}\right)_{j=1,\dots,J}$

Tableau de fréquences

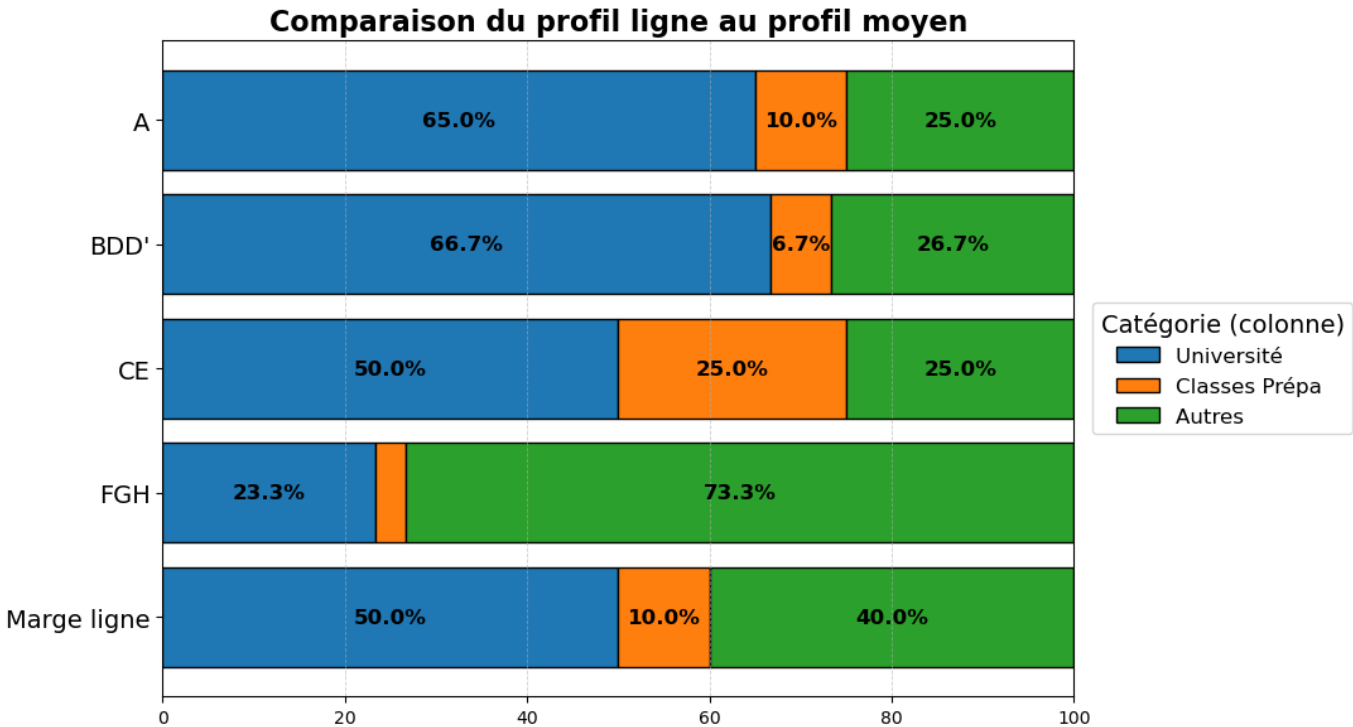
	Y1	Y2	Y3	Marge col.: $f_{i.}$
X1	0,13	0,02	0,05	0,20
X2	0,20	0,02	0,08	0,30
X3	0,10	0,05	0,05	0,20
X4	0,07	0,01	0,22	0,30
Marge ligne. : $f_{.j}$	0,50	0,10	0,40	1

$$f_{.j} = \frac{f_{1j}}{0,20} = \frac{0,13}{0,20} = 0,65$$

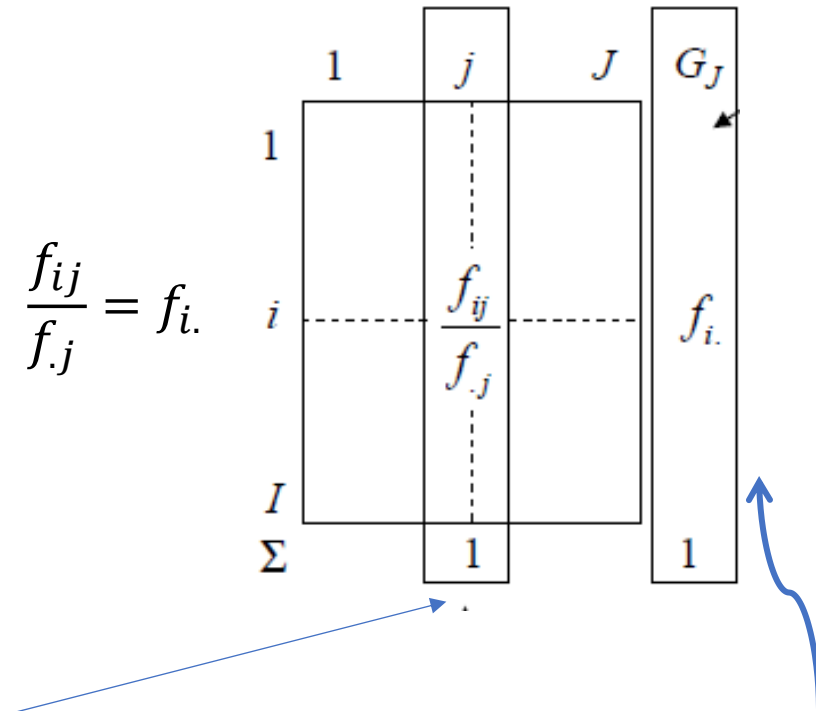
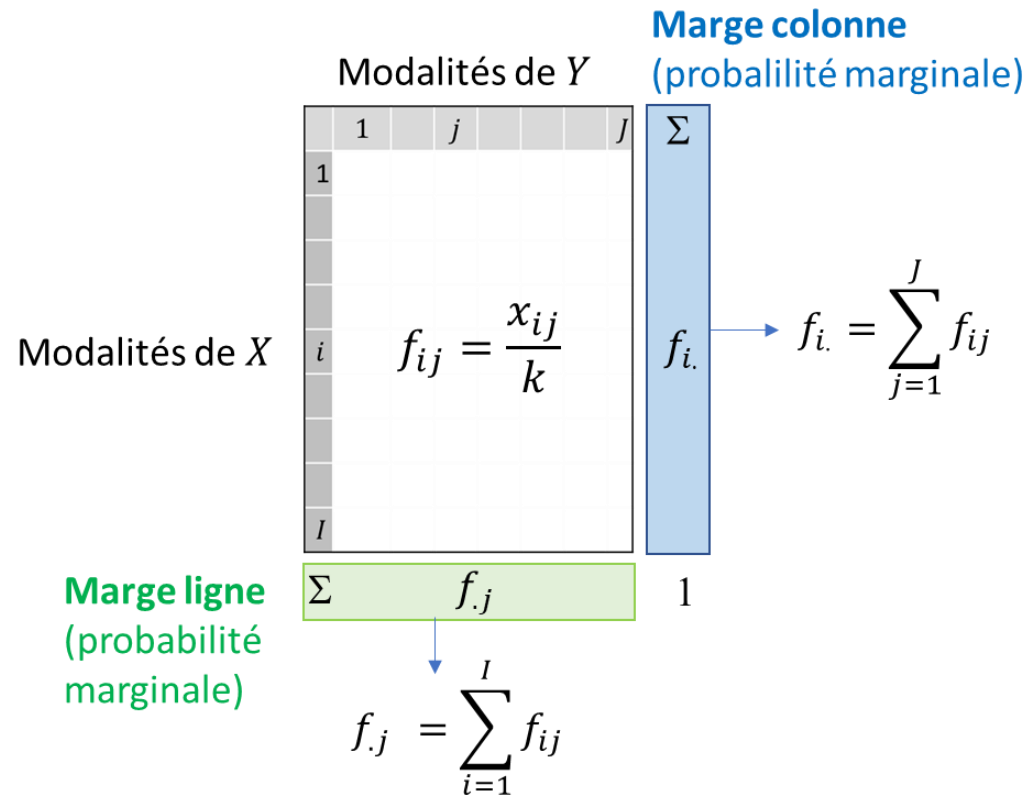
$$f_{.j} = \frac{f_{2j}}{0,20} = \frac{0,02}{0,20} = 0,10$$

$$f_{.j} = \frac{f_{3j}}{0,20} = \frac{0,05}{0,20} = 0,25$$

	Y1	Y2	Y3
X1	0,65	0,10	0,25
X2	0,6667	0,0667	0,2667
X3	0,50	0,25	0,25
X4	0,2333	0,0333	0,7333
Profil moyen G_I	0,50	0,10	0,40



4.2. Profil-colonne et profil-moyen

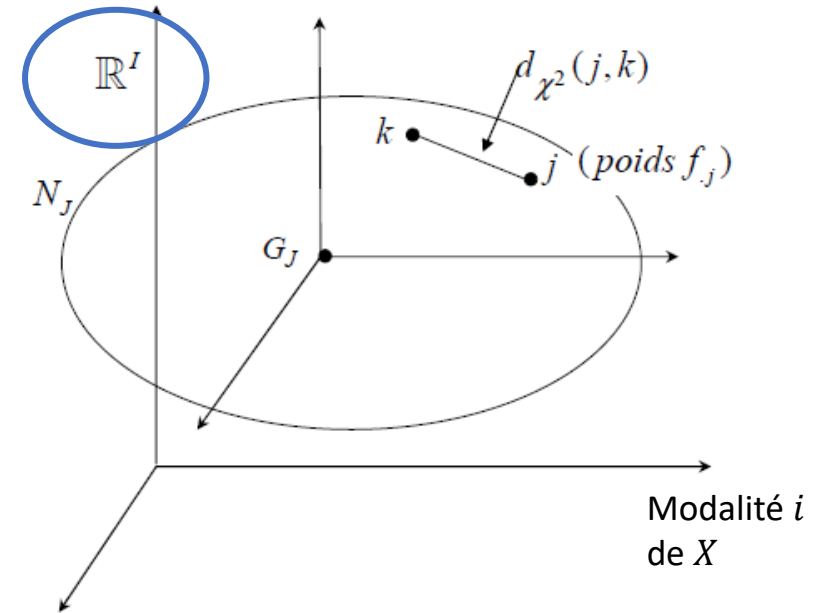
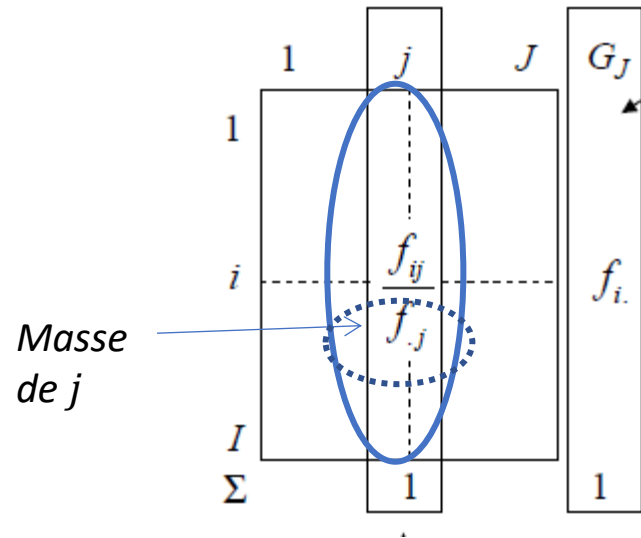


Profil colonne j = distribution conditionnelle : répartition de la var X par rapport à la modalité j de la var Y ; la somme des probabilités conditionnelles d'avoir $X = i$ sachant que $Y = j$ est = 1

Profil moyen de la variable X : en colonne G_J

4.3. Profil-colonne et profil moyen

- Nuage des J profils colonnes dans \mathbb{R}^I



- Chaque point j est affecté de la masse f_j dans un espace de dim $(I - 1)$
- G_J : le centre de gravité du nuage des profils-colonnes est le profil moyen de la var X ; sa i ème composante est f_i .
- C'est la fréquence marginale de lignes

5. Distance entre deux points profils lignes

- Traduits les différences d'effectifs sur les deux modalités de la var X
- Distance euclidienne usuelle entre **deux profils lignes** : ressemblance ou différence entre les 2 modalités i et i' de la var X sans tenir compte des effectifs de j : (dim sur Y est p , dim sur X est n)

- $$d^2(i, i') = \sum_{j=1}^J \left(\frac{f_{ij}}{f_{i.}} - \frac{f_{i'j}}{f_{i'.}} \right)^2$$

- Nécessité de prendre en compte les **masses des colonnes** (chaque j est de $f_{.j}$ différent):

➤ **Profils lignes**: la distance devient

- *distance du χ^2* :
$$d_{\chi^2}^2(i, i') = \sum_{j=1}^J \frac{1}{f_{.j}} \left(\frac{f_{ij}}{f_{i.}} - \frac{f_{i'j}}{f_{i'.}} \right)^2$$

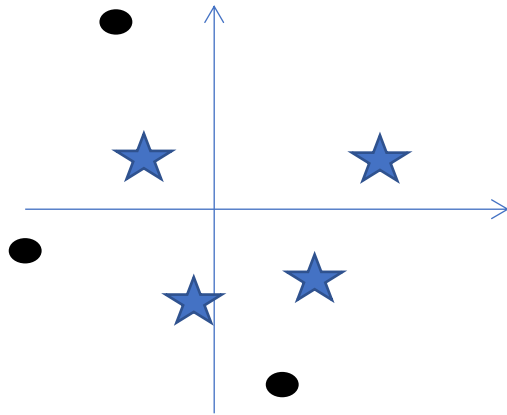
➤ Pour **profils colonne** : la distance devient

- *distance du χ^2* :
$$d_{\chi^2}^2(j, j') = \sum_{i=1}^I \frac{1}{f_{i.}} \left(\frac{f_{ij}}{f_{.j}} - \frac{f_{ij'}}{f_{.j'}} \right)^2$$

La **distance pondérée du Chi2** possède des propriétés remarquables, en particulier la symétrie entre les lignes et les colonnes. Grâce à cette symétrie, on peut fusionner des modalités ayant le même profil sans modifier les distances ni la géométrie des données. De plus, la propriété **quasi-barycentrique** est préservée : les coordonnées des lignes sont des barycentres pondérés des colonnes et réciproquement.

6. Visualisation des nuages

- Nuage de I profils lignes dans un espace à deux dimensions : réduction de R^J à R^Q (2 axes factoriels) – Nuage des I profils dans un espace à 2 dim .



distance du χ^2 :
$$d_{\chi^2}^2(i, i') = \sum_{j=1}^J \frac{1}{f_{.j}} \left(\frac{f_{ij}}{f_{i.}} - \frac{f_{i'j}}{f_{i'.}} \right)^2$$

- Représentation des deux nuages dans un espace réduit
Analyse du nuage de points pondérés dans un espace muni de la métrique du χ^2 → réduction de dimension.

7. Réduction de dimension dans R^J et R^I : transformation des données

1) Tableau de contingence X avec x_{ij} de $\dim(I, J)$

	Y1	Y2	Y3	Totaux ligne : $x_{i.}$
X1	13	2	5	20
X2	20	2	8	30
X3	10	5	5	20
X4	7	1	22	30
Totaux col. : $x_{.j}$	50	10	40	100

2) Fréquences relatives F d'élément f_{ij} de $\dim(I, J)$

	Y1	Y2	Y3	Marge col.: $f_{.j}$
X1	0,13	0,02	0,05	0,20
X2	0,20	0,02	0,08	0,30
X3	0,10	0,05	0,05	0,20
X4	0,07	0,01	0,22	0,30
Marge ligne. : $f_{i.}$	0,50	0,10	0,40	1

3) Table des profils lignes et profils colonnes

Profils-lignes

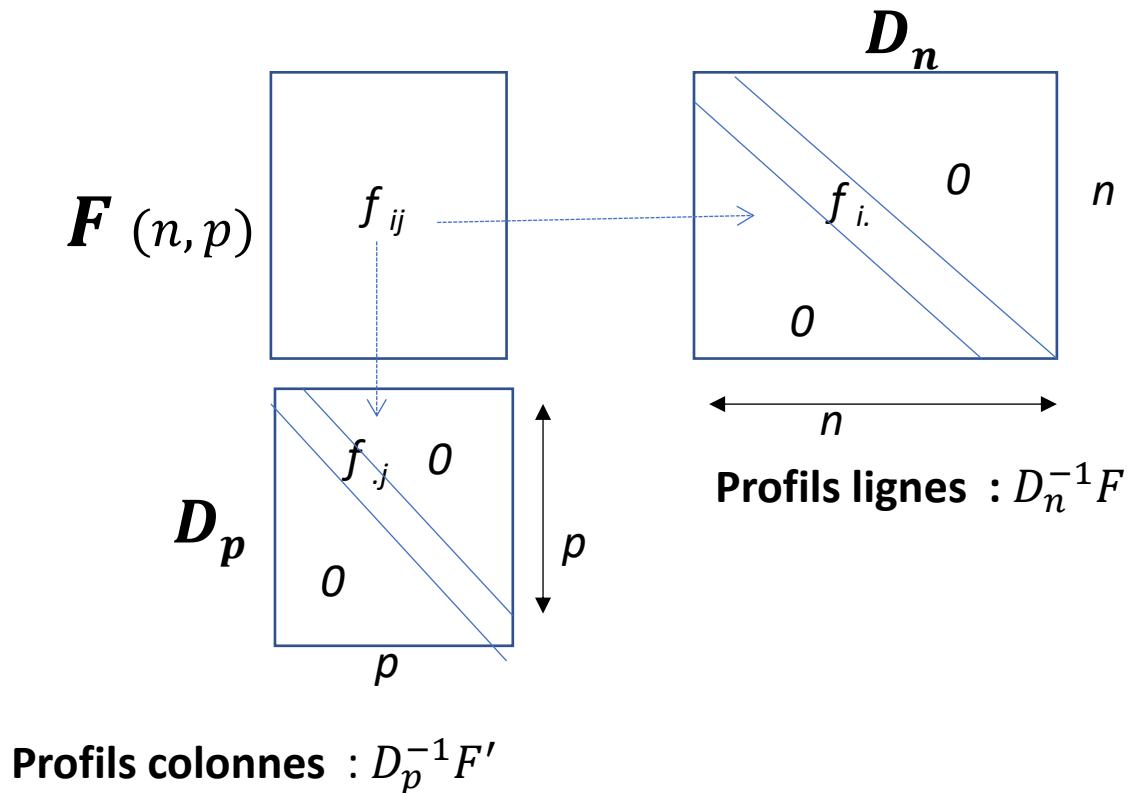
	Y1	Y2	Y3
X1	0,65	0,10	0,25
X2	0,6667	0,0667	0,2667
X3	0,50	0,25	0,25
X4	0,2333	0,0333	0,7333
Profil moyen G_I	0,50	0,10	0,40

Profils-colonnes

	Y1	Y2	Y3	Profil moyen G_J
X1	0,26	0,2	0,125	0,20
X2	0,40	0,2	0,2	0,30
X3	0,2	0,5	0,125	0,20
X4	0,14	0,1	0,55	0,30

8. Matrices de transformation

- F : Matrice des fréquences relatives (n,p)
- D_n : Matrice diagonale des marges en ligne $f_{.j}$ de dim (n,n)
- D_p : Matrice diagonale des marges en colonne $f_{i.}$ de dim (p,p)



$$D_n = \begin{bmatrix} 0.20 & 0 & 0 & 0 \\ 0 & 0.30 & 0 & 0 \\ 0 & 0 & 0.20 & 0 \\ 0 & 0 & 0 & 0.30 \end{bmatrix}$$

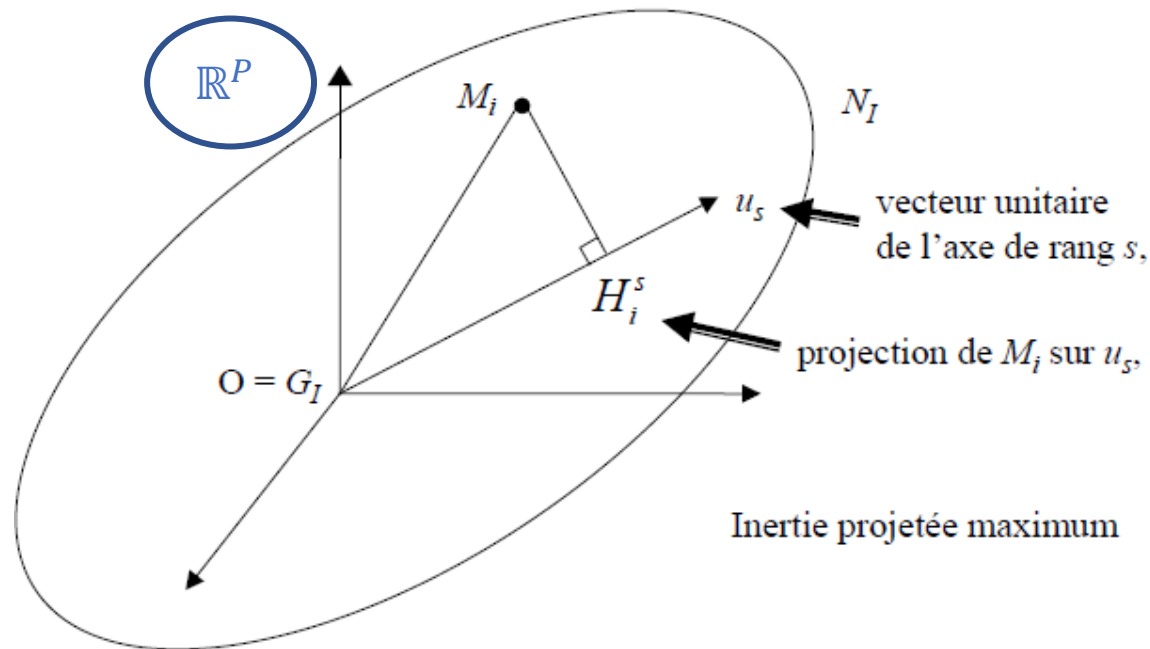
$$D_p = \begin{bmatrix} 0.50 & 0 & 0 \\ 0 & 0.10 & 0 \\ 0 & 0 & 0.40 \end{bmatrix}$$

$$D_n^{-1}F = \begin{bmatrix} 0,65 & 0,10 & 0,25 \\ 0,6667 & 0,0667 & 0,2667 \\ 0,2333 & 0,333 & 0,333 \end{bmatrix}$$

$$D_p^{-1}F' = \begin{bmatrix} 0,26 & 0,40 & 0,20 & 0,14 \\ 0,20 & 0,20 & 0,50 & 0,10 \\ 0,125 & 0,20 & 0,125 & 0,55 \end{bmatrix}$$

9. Réduction de dimension : Critère à maximiser

- Proximités entre profils : analyse par rapport l'origine mais possible à partir des centres de gravité (F : Matrice des fréquences relatives (n, p))
- Dans espace \mathbb{R}^p avec analyse par rapport à l'origine : réduction de dimension
- Critère de projection orthogonale selon un axe où inertie maximale (variance) passant par O et engendré par un vecteur unitaire u et de métrique D_p



- AFC Revient à faire une ACP sur les profils ligne
- AFC : une ACP sur les profils colonne

10. Matrice à diagonaliser (espace des colonnes \mathbb{R}^p)

- Maximiser la somme pondérée des carrés des projections sur l'axe de vecteur unitaire u soit

$$\text{Max} \left\{ \sum_i f_{i.} d_{\chi^2}(i, O) \right\}$$

Soit maximiser la somme des distances au carré de OMi :

➤ Soit rendre max la quantité :

$$\mathbf{u}' \mathbf{D}_p^{-1} \mathbf{F}' \mathbf{D}_n^{-1} \mathbf{F} \mathbf{D}_p^{-1} \mathbf{u}$$

Avec la contrainte $\mathbf{u}' \mathbf{D}_p^{-1} \mathbf{u} = 1$

Alors \mathbf{u} est vecteur propre la matrice \mathbf{S} :

$$\mathbf{S} = \mathbf{F}' \mathbf{D}_n^{-1} \mathbf{F} \mathbf{D}_p^{-1} \text{ de terme général}$$

Profils lignes : $\mathbf{D}_n^{-1} \mathbf{F}$

Profils colonnes : $\mathbf{D}_p^{-1} \mathbf{F}'$

vecteur unitaire \mathbf{u} de métrique \mathbf{D}_p

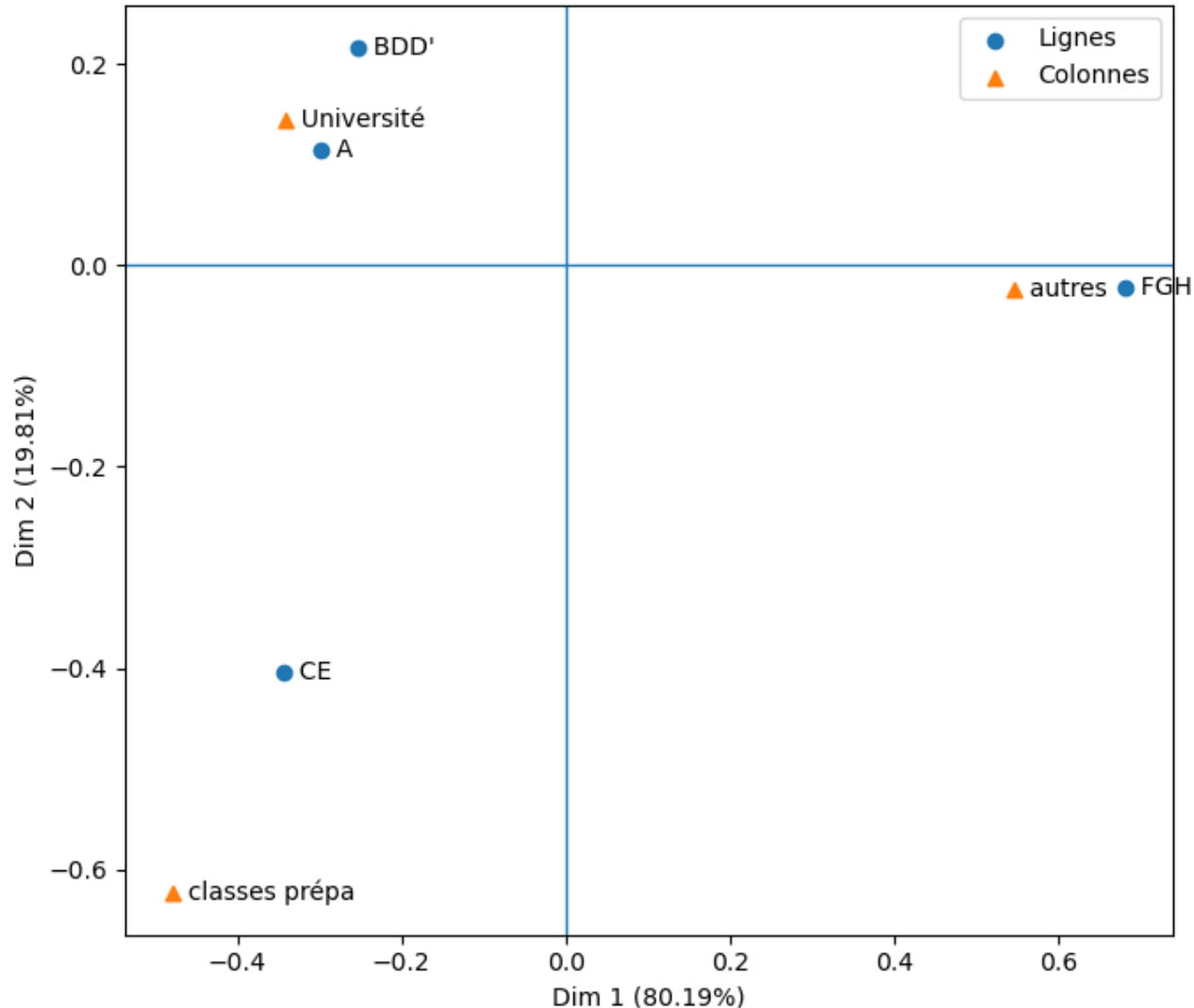
$$s_{jj'} = \sum_{i=1}^n \frac{f_{ij} f_{ij'}}{f_{i.} f_{.j'}}$$

11. Axes factoriels et coordonnées factorielles

Dans R^p		Dans R^n
$S = F' D_n^{-1} F D_p^{-1}$	Matrice à diagonaliser	$T = F D_p^{-1} F' D_n^{-1}$
$S u_\alpha = \lambda_\alpha u_\alpha$	Axe factoriel	$T v_\alpha = \lambda_\alpha v_\alpha$
$\psi_\alpha = D_n^{-1} F D_p^{-1} u_\alpha$	Coordonnées factorielles	$\varphi_\alpha = D_p^{-1} F' D_n^{-1} v_\alpha$
$\psi_{\alpha i} = \sum_{j=1}^p \frac{f_{ij}}{f_{i.} f_{.j}} u_{\alpha i}$	Coordonnées factorielles	$\varphi_{\alpha i} = \sum_{i=1}^n \frac{f_{ij}}{f_{i.} f_{.j}} v_{\alpha i}$

- AFC -> une ACP sur les profils lignes
- AFC -> une ACP sur les profils colonnes
- Analyse factorielle : réduction de dim + qualité

12. Interprétation-exemple (a)



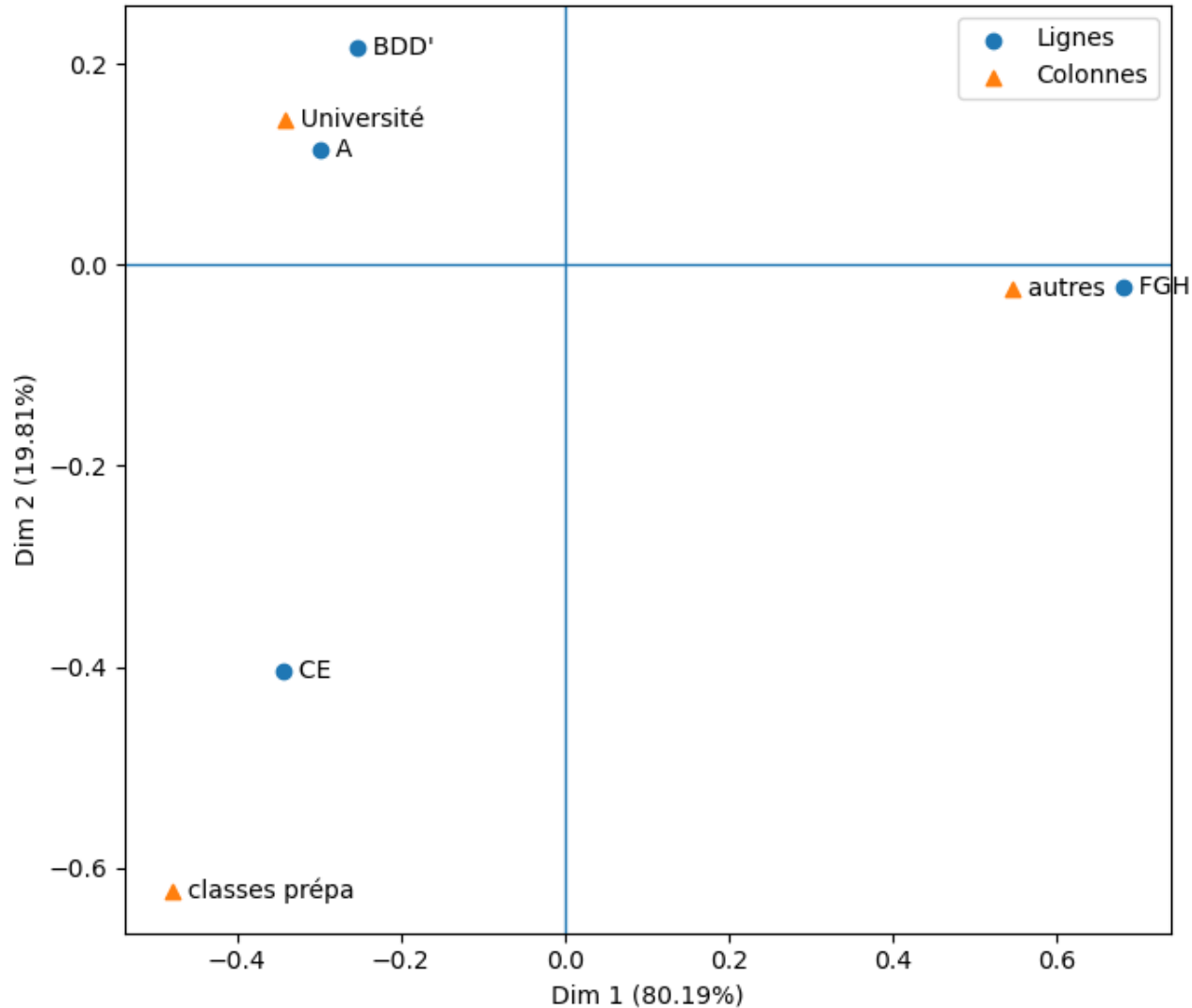
❖ Coordonnées factorielles lignes

	Dim1	Dim2
A	-0.297248	0.113549
BDD'	-0.253928	0.215943
CE	-0.343564	-0.403998
FGH	0.681136	-0.022311

❖ Coordonnées factorielles colonnes

	Dim1	Dim2
Université	-0.340516	0.143929
Classes Prépa	-0.478534	-0.622633
Autres	0.545278	-0.024254

12. Interprétation-exemple (b)



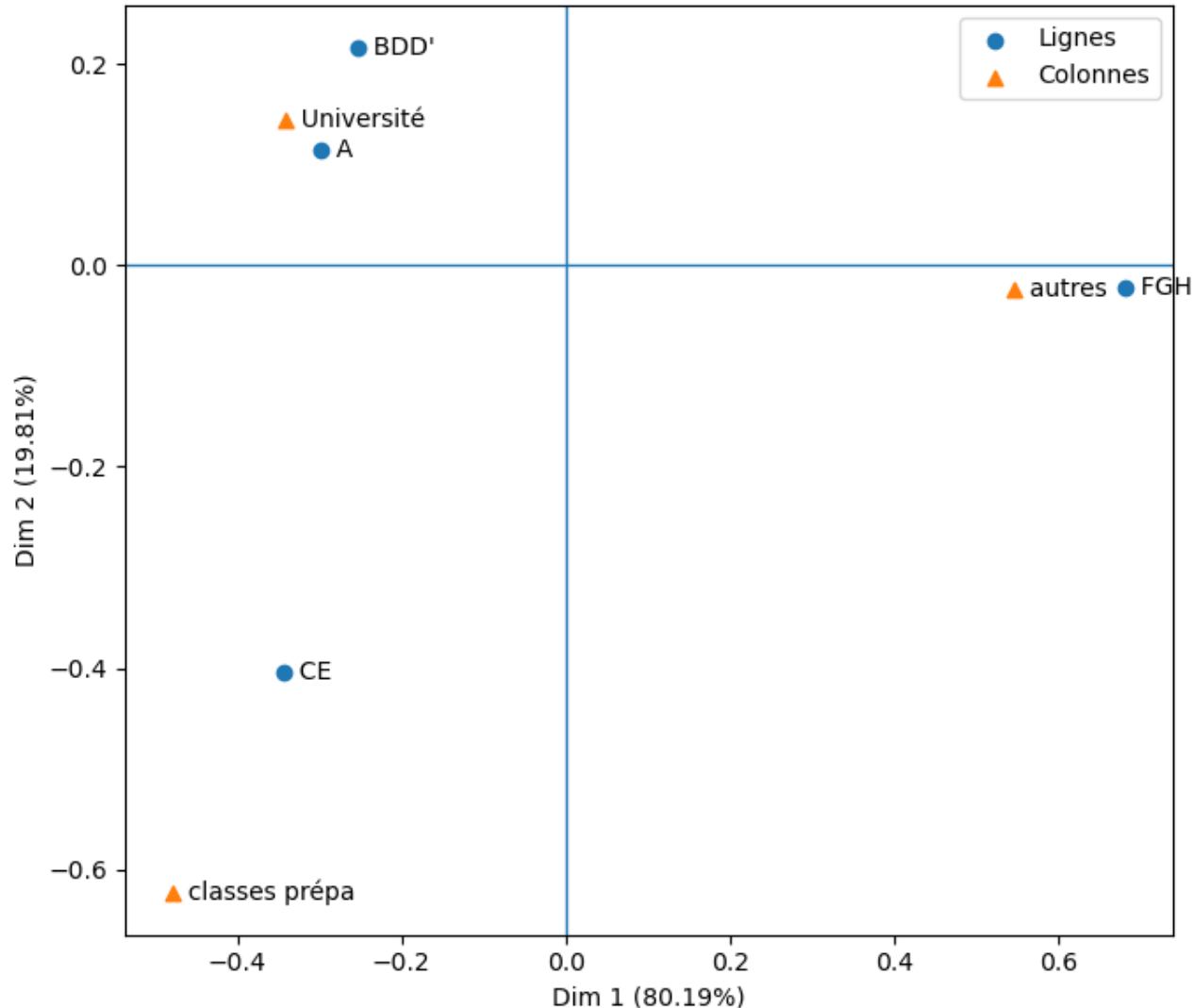
❖ Pourcentages d'inertie

En AFC: $0 \leq \lambda_s \leq 1$

$$\frac{\text{inertie projetée de } N_I \text{ sur } u_s}{\text{inertie totale de } N_I} = \frac{\sum_{i=1}^I f_{i.}(OH_i^s)^2}{\sum_{i=1}^I f_{i.}(OM_i)^2} = \frac{\lambda_s}{\sum_{k=1}^K \lambda_k}$$

	Inertie (= valeurs propres)	Inertie %
Dim 1	0,19980638	80,19
Dim 2	0,049936029	19,81
Somme	0,2492	100

12. Interprétation-exemple (c)



❖ Aides à l'interprétation: qualité de représentation

$$\frac{\text{inertie projetée de } M_i \text{ sur } u_s}{\text{inertie totale de } M_i} = \frac{f_i \cdot (OH_i^s)^2}{f_i \cdot (OM_i)^2} = \cos^2(\overrightarrow{OM_i}, u_s)$$

Profil ligne	Axe 1	Axe 2	Total
A	0.8727	0.1273	1.000
BDD'	0.5803	0.4197	1.000
CE	0.4197	0.5803	1.000
FGH	0.9989	0.0011	1.000

(calculés de la même façon sur la matrice transposée)

Profil colonne	Axe 1	Axe 2	Total
Université	0.8484	0.1516	1.000
classes prépa	0.3713	0.6287	1.000
autres	0.9980	0.0020	1.000

13. Aide à l'interprétation- contribution

- ❖ Contribution d'un point i à l'inertie de l'axe
 - Indicateur brut: inertie projetée de Mi sur $us = f_i.(OH_i^s)^2$
 - Indicateur relative : inertie projetée du point / inertie totale de l'axe : $f_i.(OH_i^s)^2/\lambda_s \times 100$
- ❖ On peut additionner la contribution de plusieurs éléments.
- ❖ Ceci indique alors dans quelle mesure on peut considérer qu'un axe est dû à un seul ou quelques éléments.
- ❖ Contributions réalisent un compromis entre distance à l'origine et poids
- ❖ Ceci peut être utilisé pour sélectionner un sous-ensemble d'éléments lors de l'interprétation (*Pagès 2008*)