Variables qualitatives : analyse des correspondances.

Jean-Marc Lasgouttes

http://ana-donnees.lasgouttes.net/

L'analyse factorielle des correspondances

But On cherche à décrire la liaison entre *deux* variables qualitatives.

Exemple on peut regarder la répartition de la couleur des yeux en fonction de la couleur des cheveux.

Différence avec l'ACP l'ACP se fait dans un cadre différent ; les variables sont quantitatives et donc

— il est possible de faire des opérations mathématiques sur les valeurs des variables ;

— par contre, il n'est en général pas possible de compter les individus qui ont une caractéristique donnée (taille=$1,83$m)

Pourquoi deux variables ? le cas de plus de deux variables est l'analyse de correspondance multiples, traité plus tard dans le cours.

Partie I. Les données qualitatives

Variables qualitatives

Soit $\mathcal{X}$ une variable qualitative. On dispose d'un échantillon de n individus sur lesquels la variable est mesurée.

Modalités (ou catégories) les valeurs que peut prendre une variable qualitative ; si la variable a m modalités (valeurs possibles), on note x_i, $1 \leq i \leq m$, ces modalités, ou plus simplement i.

Effectif le nombre d'occurrence de la modalité i dans l'échantillon ; on le note n_i, et on a $\sum_{i=1}^m n_i = n$.

Profil c'est l'ensemble des valeurs n_i/n ; la somme du profil sur les modalités est 1.

Tableau de contingence

Soient $\mathcal{X}_1$ et $\mathcal{X}_2$ deux variables qualitatives à m_1 et m_2 modalités respectivement décrivant un ensemble de n individus.

Définition le tableau de contingence est une matrice à m_1 lignes et m_2 colonnes renfermant les effectifs n_{ij} d'individus tels que $\mathcal{X}_1 = i$ et $\mathcal{X}_2 = j$.

$$\mathbf{N} = \begin{pmatrix} n_{11} & n_{12} & & \cdots & & n_{1m_2} \\ n_{21} & n_{22} & & & & \\ & & \ddots & & & \\ \vdots & & & n_{ij} & & \vdots \\ & & & & \ddots & \\ n_{m_1 1} & & & & & n_{m_1 m_2} \end{pmatrix}$$

La constitution de ce tableau est aussi appelé un « tri croisé ».

Marges et profils

Marge en ligne c'est la somme $n_{i\cdot} = \sum_{j=1}^{m_2} n_{ij}$, c'est-à-dire l'effectif total de la modalité i de $\mathcal{X}_1$.

On définit aussi le profil marginal des lignes $n_{i\cdot}/n$.

Marge en colonne c'est la somme $n_{\cdot j} = \sum_{i=1}^{m_1} n_{ij}$, c'est-à-dire l'effectif total de la modalité j de $\mathcal{X}_2$.

On définit aussi le profil marginal des colonnes $n_{\cdot j}/n$.

Deux lectures possibles selon la variable que l'on privilégie, on peut définir

— le tableau des *profils-lignes* $n_{ij}/n_{i\cdot}$, qui représente la fréquence de la modalité j conditionnellement à $\mathcal{X}_1 = i$; la somme de chaque ligne est ramenée à 100%.

— le tableau des *profils-colonnes* $n_{ij}/n_{\cdot j}$, qui représente la fréquence de la modalité i conditionnellement à $\mathcal{X} = j$; la somme de chaque colonne est ramenée à 100%.

Propriétés des profils

Moyenne la moyenne des profils-lignes (avec poids correspondant aux profils marginaux des lignes) est le profil marginal des colonnes :

$$\sum_{i=1}^{m_1} \frac{n_{i\cdot}}{n} \times \frac{n_{ij}}{n_{i\cdot}} = \frac{n_{\cdot j}}{n},$$

et de même pour les colonnes $\sum_{j=1}^{m_2} \frac{n_{\cdot j}}{n} \times \frac{n_{ij}}{n_{\cdot j}} = \frac{n_{i\cdot}}{n}$.

Indépendance empirique lorsque tous les profils lignes sont identiques, il y a indépendance entre $\mathcal{X}_1$ et $\mathcal{X}_2$, puisque la connaissance de $\mathcal{X}_1$ ne change pas la répartition de $\mathcal{X}_2$. On a pour tout j

$$\frac{n_{1j}}{n_{1\cdot}} = \frac{n_{2j}}{n_{2\cdot}} = \cdots = \frac{n_{m_1 j}}{n_{m_1\cdot}} = \frac{n_{1j} + \cdots + n_{m_1 j}}{n_{1\cdot} + \cdots + n_{m_1\cdot}} = \frac{n_{\cdot j}}{n}$$

et donc $n_{ij} = \frac{n_{i\cdot} n_{\cdot j}}{n}$.

Le χ^2 d'écart à l'indépendance

Définition c'est la grandeur suivante, aussi notée χ^2 ou X^2

$$d^2 = \sum_{i=1}^{m_1}\sum_{j=1}^{m_2} \frac{\left(n_{ij} - \frac{n_{i\cdot}n_{\cdot j}}{n}\right)^2}{\frac{n_{i\cdot}n_{\cdot j}}{n}} = n\left[\sum_{i=1}^{m_1}\sum_{j=1}^{m_2} \frac{n_{ij}^2}{n_{i\cdot}n_{\cdot j}} - 1\right].$$

$d^2 = 0 \iff$ les variables sont indépendantes.

Contribution au χ^2 c'est le terme

$$\frac{\left(n_{ij} - \frac{n_{i\cdot}n_{\cdot j}}{n}\right)^2}{\frac{n_{i\cdot}n_{\cdot j}}{n}}$$

qui permet de mettre en évidence les associations significatives entre modalités de deux variables.

Borne supérieure comme $n_{ij} \leq n_{i\cdot}$, on a

$$\sum_{i=1}^{m_1}\sum_{j=1}^{m_2} \frac{n_{ij}^2}{n_{i\cdot}n_{\cdot j}} \leq \sum_{i=1}^{m_1}\sum_{j=1}^{m_2} \frac{n_{ij}}{n_{\cdot j}} = \sum_{j=1}^{m_2} \frac{\sum_{i=1}^{m_1} n_{ij}}{n_{\cdot j}} = \sum_{j=1}^{m_2} \frac{n_{\cdot j}}{n_{\cdot j}} = m_2,$$

et donc $d^2 \leq n(m_2 - 1)$. On fait de même pour m_1 et

$$\varphi^2 = \frac{d^2}{n} \leq \min(m_1 - 1, m_2 - 1).$$

Dépendance fonctionnelle si $\varphi^2 = m_2 - 1$, alors pour chaque i soit $n_{ij} = n_{i\cdot}$, soit $n_{ij} = 0$: il existe une unique case non nulle par ligne. $\mathcal{X}_2$ est donc fonctionnellement liée à $\mathcal{X}_1$.

Dépendance inverse cette relation ne signifie pas que $\mathcal{X}_1$ est fonctionnellement liée à $\mathcal{X}_2$, sauf si $m_1 = m_2$. On peut alors représenter le tableau comme une matrice diagonale.

Caractère significatif du χ^2

Problème à partir de quelle valeur de d^2 doit-on considérer que les variables $\mathcal{X}_1$ et $\mathcal{X}_2$ sont dépendantes ?

Méthode on suppose que $\mathcal{X}_1$ et $\mathcal{X}_2$ sont issus de tirages de deux variables aléatoires indépendantes. On peut alors montrer que d^2 est une réalisation d'une variable aléatoire D^2 qui suit une loi $\chi^2_{(m_1-1)(m_2-1)}$.

Définition Loi du khi-deux à p degrés de libertés χ^2_p est la loi de la variable $\sum_{i=1}^p U_i^2$, où les U_i sont des variables gaussiennes réduites indépendantes.

Le test du χ^2 on se fixe un risque d'erreur α (0.01 ou 0.05 en général) et on calcule la valeur d_c^2 telle que $\mathrm{P}\left(\chi^2_{(m_1-1)(m_2-1)} > d_c^2\right) = \alpha$. Si $d^2 > d_c^2$ on considère que l'événement est trop improbable et que donc que l'hypothèse originale d'indépendance doit être rejetée. On trouvera en général ces valeurs dans une table précalculée.

Cas p grand quand $p > 30$, on considère que $\sqrt{2\chi_p^2} - \sqrt{2p-1}$ est distribué comme une variable gaussienne centrée réduite $N(0,1)$.

Partie II. Géométrie de nuages de profils

Analyse des correspondances de deux variables : les données

Effectifs on a un tableau de contingence $\mathbf{N}$ à m_1 lignes et m_2 colonnes résultant du croisement de deux variables qualitatives $\mathcal{X}_1$ et $\mathcal{X}_2$ à m_1 et m_2 modalités respectivement. On note $\mathbf{D}_1$ et $\mathbf{D}_2$ les matrices diagonales des effectifs marginaux

$$\mathbf{D}_1 = \begin{pmatrix} n_{1\cdot} & & & 0 \\ & n_{2\cdot} & & \\ & & \ddots & \\ 0 & & & n_{m_1\cdot} \end{pmatrix}$$

$$\mathbf{D}_2 = \begin{pmatrix} n_{\cdot 1} & & & 0 \\ & n_{\cdot 2} & & \\ & & \ddots & \\ 0 & & & n_{\cdot m_2} \end{pmatrix}$$

Profils le tableau des profils des lignes $n_{ij}/n_{i\cdot}$ est donné par $\mathbf{D}_1^{-1}\mathbf{N}$ et celui des profils des colonnes $n_{ij}/n_{\cdot j}$ par $\mathbf{N}\mathbf{D}_2^{-1}$.

Représentation géométrique des profils

Nuage de points les profils-lignes forment un nuage de m_1 points de $\mathbb{R}^{m_2}$. Chaque point est affecté d'un poids égal à sa fréquence marginale $n_{i\cdot}/n$, et la matrice des poids est donc $\frac{1}{n}\mathbf{D}_1$.

Centre de gravité c'est le profil marginal des colonnes car

$$\mathbf{g}_\ell = \frac{1}{n}(\mathbf{D}_1^{-1}\mathbf{N})'\mathbf{D}_1\mathbf{1}_{m_1} = \left(\frac{n_{\cdot 1}}{n}, \cdots, \frac{n_{\cdot m_2}}{n}\right)'$$

Profils-colonnes les lignes du tableau $\mathbf{D}_2^{-1}\mathbf{N}'$ forment un nuage de m_2 points de $\mathbb{R}^{m_1}$, avec matrice de poids $\frac{1}{n}\mathbf{D}_2$ et centre de gravité

$$\mathbf{g}_c = \left(\frac{n_{1\cdot}}{n}, \cdots, \frac{n_{m_1\cdot}}{n}\right)'$$

Comment étudier ces données

Cas indépendant en cas d'indépendance empirique, on a

$$\frac{n_{ij}}{n_{i\cdot}} = \frac{n_{\cdot j}}{n} \text{ et } \frac{n_{ij}}{n_{\cdot j}} = \frac{n_{i\cdot}}{n}.$$

Les deux nuages sont donc réduits à leurs centres de gravité.

Dimension des nuages comme les profils somment à 1, les m_1 profils-lignes sont situés dans le sous-espace W_1 de dimension $m_2 - 1$ défini par $\sum_{j=1}^{m_2} x_j = 1$ et $x_j \geq 0$.

ACP l'étude de la forme des nuages au moyen de l'analyse en composantes principales permettra de rendre compte de la structure des écarts à l'indépendance.

Partie III. L'AFC : une ACP sur un nuage de profils

La métrique du χ^2

Profils-lignes la distance du χ^2 entre les profils-lignes $\mathbf{e}_i$ et $\mathbf{e}_{i'}$ est définie par

$$d^2_{\chi^2_\ell}(\mathbf{e}_i, \mathbf{e}_{i'}) = \sum_{j=1}^{m_2} \frac{n}{n_{\cdot j}} \left(\frac{n_{ij}}{n_{i\cdot}} - \frac{n_{i'j}}{n_{i'\cdot}} \right)^2,$$

ce qui revient à utiliser la métrique diagonale $n\mathbf{D}_2^{-1}$.

Inertie l'inertie totale du nuage des profils-lignes par rapport à $\mathbf{g}_\ell$ est

$$I_{\mathbf{g}_\ell} = \sum_{i=1}^{m_1} \frac{n_{i\cdot}}{n} d^2_{\chi^2_\ell}(\mathbf{e}_i, \mathbf{g}_\ell) = \sum_{i=1}^{m_1} \sum_{j=1}^{m_2} \frac{n_{i\cdot}}{n_{\cdot j}} \left(\frac{n_{ij}}{n_{i\cdot}} - \frac{n_{\cdot j}}{n} \right)^2$$
$$= \sum_{i=1}^{m_1} \sum_{j=1}^{m_2} \frac{1}{n_{i\cdot} n_{\cdot j}} \left(n_{ij} - \frac{n_{i\cdot} n_{\cdot j}}{n} \right)^2 = \varphi^2$$

Cette inertie mesure donc l'écart à l'indépendance.

Pourquoi la métrique du χ^2 ?

Pondération la pondération $n/n_{\cdot j}$ permet de donner des importances comparables aux différentes « variables ».

Équivalence distributionnelle si deux colonnes j et j' de $\mathbf{N}$ ont le même profil, il est logique de les regrouper en une seule d'effectif $n_{ij}+n_{ij'}$; on a alors quand $n_{ij}/n_{\cdot j} = n_{ij'}/n_{\cdot j'}$

$$\frac{n}{n_{\cdot j}} \left[\frac{n_{ij}}{n_{i\cdot}} - \frac{n_{\cdot j}}{n} \right]^2 + \frac{n}{n_{\cdot j'}} \left[\frac{n_{ij'}}{n_{i\cdot}} - \frac{n_{\cdot j'}}{n} \right]^2$$
$$= \frac{n}{n_{\cdot j} + n_{\cdot j'}} \left[\frac{n_{ij} + n_{ij'}}{n_{i\cdot}} - \frac{n_{\cdot j} + n_{\cdot j'}}{n} \right]^2$$

La distance entre les profils-ligne est donc inchangée.

Profils-colonnes on définit la distance entre deux profils-colonnes $\mathbf{e}_j$ et $\mathbf{e}_{j'}$ comme

$$d^2_{\chi^2_c}(\mathbf{e}_j, \mathbf{e}_{j'}) = \sum_{i=1}^{m_1} \frac{n}{n_{i\cdot}} \left(\frac{n_{ij}}{n_{\cdot j}} - \frac{n_{ij'}}{n_{\cdot j'}} \right)^2,$$

soit une métrique de matrice $n\mathbf{D}_1^{-1}$. Les propriétés sont les mêmes que les profils lignes.

ACP des deux nuages de profils

Comment ? deux possibilités en dualité exacte

	données	métrique	poids
Profils-lignes	$\mathbf{X} = \mathbf{D}_1^{-1}\mathbf{N}$	$\mathbf{M} = n\mathbf{D}_2^{-1}$	$\mathbf{D} = \frac{\mathbf{D}_1}{n}$
Profils-colonnes	$\mathbf{X} = \mathbf{D}_2^{-1}\mathbf{N}'$	$\mathbf{M} = n\mathbf{D}_1^{-1}$	$\mathbf{D} = \frac{\mathbf{D}_2}{n}$

Centrage Il n'est pas très utile, puisque la différence des inerties est 1 :

$$I_{\mathbf{g}} = \frac{d^2}{n} = \sum_{i=1}^{m_1} \sum_{j=1}^{m_2} \frac{\left(n_{ij} - \frac{n_{i\cdot} n_{\cdot j}}{n}\right)^2}{n_{i\cdot} n_{\cdot j}}$$
$$= \sum_{i=1}^{m_1} \sum_{j=1}^{m_2} \frac{n_{ij}^2}{n_{i\cdot} n_{\cdot j}} - 1 = I_{\mathbf{0}} - 1,$$

Approche On effectue donc une ACP non centrée

- on remplace la variance par $\mathbf{X}'\mathbf{D}\mathbf{X}$.
- on élimine la valeur propre 1 associée à l'axe principal $\mathbf{g}$ (admis)

Calcul de l'ACP (profils-lignes)

Facteurs principaux ils sont vecteurs propres de

$$\mathbf{M}\mathbf{X}'\mathbf{D}\mathbf{X} = (n\mathbf{D}_2^{-1})(\mathbf{D}_1^{-1}\mathbf{N})'\frac{\mathbf{D}_1}{n}(\mathbf{D}_1^{-1}\mathbf{N}) = \mathbf{D}_2^{-1}\mathbf{N}'\mathbf{D}_1^{-1}\mathbf{N}.$$

et on a donc pour chaque axe principal k

$$\mathbf{D}_2^{-1}\mathbf{N}'\mathbf{D}_1^{-1}\mathbf{N}\mathbf{u}_k = \lambda_k \mathbf{u}_k$$

Composantes principales la composante principale associée au facteur $\mathbf{u}_k$ est $\mathbf{a}_k = \mathbf{X}\mathbf{u}_k = \mathbf{D}_1^{-1}\mathbf{N}\mathbf{u}_k$; elle est vecteur propre de la matrice $\mathbf{D}_1^{-1}\mathbf{N}\mathbf{D}_2^{-1}\mathbf{N}'$ car

$$\mathbf{D}_1^{-1}\mathbf{N}\mathbf{D}_2^{-1}\mathbf{N}'\mathbf{a}_k = \mathbf{D}_1^{-1}\mathbf{N}\mathbf{D}_2^{-1}\mathbf{N}'\mathbf{D}_1^{-1}\mathbf{N}\mathbf{u}_k$$
$$= \lambda_k \mathbf{D}_1^{-1}\mathbf{N}\mathbf{u}_k = \lambda_k \mathbf{a}_{\mathbf{k}}$$

Analyse des profils-colonnes on échange les indices 1 et 2 et on transpose $\mathbf{N}$.

Comparaison lignes-colonnes

	ACP profils-lignes	ACP profils-colonnes
Facteurs principaux	Vecteurs propres de $\mathbf{D}_2^{-1}\mathbf{N}'\mathbf{D}_1^{-1}\mathbf{N}$	Vecteurs propres de $\mathbf{D}_1^{-1}\mathbf{N}\mathbf{D}_2^{-1}\mathbf{N}'$
Composantes principales	Vecteurs propres de $\mathbf{D}_1^{-1}\mathbf{N}\mathbf{D}_2^{-1}\mathbf{N}'$ normalisés par $\operatorname{var}\mathbf{a}_k = \mathbf{a}_k'\frac{\mathbf{D}_1}{n}\mathbf{a}_k = \lambda_k$	Vecteurs propres de $\mathbf{D}_2^{-1}\mathbf{N}'\mathbf{D}_1^{-1}\mathbf{N}$ normalisés par $\operatorname{var}\mathbf{b}_k = \mathbf{b}_k'\frac{\mathbf{D}_2}{n}\mathbf{b}_k = \lambda_k$

Comparaison les deux analyses conduisent aux mêmes valeurs propres et les facteurs principaux de l'une sont les composantes principales de l'autre (à un facteur près).

Partie IV. Aspects pratiques

Interprétation des résultats

Coordonnées des points Les coordonnées des points-lignes et points-colonnes s'obtiennent en cherchant les vecteurs propres des produits des deux tableaux de profils. Ce sont les grandeurs principales à obtenir.

Projection des nuages il est possible de projeter les deux nuages de points sur le même représentations. On justifiera plus tard le sens de cette représentation et son interprétation.

Cercle des corrélations il n'a aucun intérêt ici, puisque les véritables variables sont qualitatives.

(non) effet de taille comme les composantes variables sont centrées ($\sum_{i=1}^{m_1} n_{i\cdot} a_{ik} = \sum_{j=1}^{m_2} n_{\cdot j} b_{jk} = 0$), ont sait que les coordonnées des $\mathbf{a}_k$ et $\mathbf{b}_k$ ne peuvent être toutes de même signe ; il n'y a donc jamais d'effet de « taille ».

Contributions à l'inertie

Contribution des profils-lignes On sait que $\lambda_k = \sum_{i=1}^{m_1} \frac{n_{i\cdot}}{n}(a_{ik})^2$, où a_{ik} est la coordonnée du profil-ligne i sur la k-ième composante principale de l'ACP sur les profils-lignes. On définit donc la contribution de la modalité i à l'axe principal k comme

$$\frac{n_{i\cdot}}{n} \cdot \frac{(a_{ik})^2}{\lambda_k}.$$

On considérera les modalités ayant l'influence la plus importante (typiquement $> \alpha n_{i\cdot}/n$, $\alpha = 2$ ou 3) comme constitutives des axes ; on regardera aussi le signe de la coordonnée.

Il n'y a pas ici de modalités sur-représentées, puisqu'on ne peut pas les retirer.

Contribution des profils-colonnes pour les mêmes raisons, la contribution du de la modalité j de $\mathcal{X}_2$ à l'axe k est

$$\frac{n_{\cdot j}}{n} \cdot \frac{(b_{jk})^2}{\lambda_k}.$$

Qualité de la représentation

Profils-lignes l'AFC est une ACP, et on peut donc mesurer la qualité de la représentation de la modalité i (son profil-ligne) par un sous-espace factoriel. La qualité (le $\cos^2$ de l'angle entre le point et sa projection) s'écrit encore, pour le plan formé des q premiers axes :

$$\frac{\sum_{k=1}^{q}(a_{ik})^2}{\sum_{k=1}^{m_2}(a_{ik})^2}.$$

Comme pour l'ACP, > 0.8 signifie « bien représenté » et < 0.5 veut dire « mal représenté ». Les valeurs sont souvent données en 10000$^{\text{e}}$ou en pourcents.

Profils-colonne Le principe est le même, mais la formule devient, pour la modalité j :

$$\frac{\sum_{k=1}^{q}(b_{jk})^2}{\sum_{k=1}^{m_1}(b_{jk})^2}.$$

Formules de transition

But on cherche une relation entre les vecteurs $\mathbf{a}_k$ et $\mathbf{b}_k$ pour éviter de faire deux diagonalisation de matrice. Par exemple, si $m_1 < m_2$, on diagonalisera la matrice $\mathbf{D}_1^{-1}\mathbf{N}\mathbf{D}_2^{-1}\mathbf{N}'$.

Formules un calcul simple donne les formules suivantes

$$\mathbf{b}_k = \frac{1}{\sqrt{\lambda_k}}\mathbf{D}_2^{-1}\mathbf{N}'\mathbf{a}_k, \text{ soit } b_{jk} = \frac{1}{\sqrt{\lambda_k}}\sum_{i=1}^{m_1}\frac{n_{ij}}{n_{\cdot j}}a_{ik},$$

$$\mathbf{a}_k = \frac{1}{\sqrt{\lambda_k}}\mathbf{D}_1^{-1}\mathbf{N}\mathbf{b}_k, \text{ soit } a_{ik} = \frac{1}{\sqrt{\lambda_k}}\sum_{j=1}^{m_2}\frac{n_{ij}}{n_{i\cdot}}b_{jk}.$$

Méthode comme $\mathbf{a}_k$ est (à une normalisation près) le facteur principal associé à $\mathbf{b}_k$, on sait que $\mathbf{b}_k = \alpha\mathbf{D}_2^{-1}\mathbf{N}'\mathbf{a}_k$. Pour déterminer α, il suffit d'écrire que $\mathbf{b}_k'\frac{\mathbf{D}_2}{n}\mathbf{b}_k = \lambda_k$.

Décomposition de l'inertie

Nombre de valeurs propres Comme $\mathbf{D}_1^{-1}\mathbf{N}\mathbf{D}_2^{-1}\mathbf{N}'$ est carrée de dimension m_1 et que $\mathbf{D}_2^{-1}\mathbf{N}'\mathbf{D}_1^{-1}\mathbf{N}$ est de dimension m_2, le nombre de valeurs propres non nulles est $\min(m_1, m_2)$. Mais comme une des valeurs propres est 1 (associée à $\mathbf{g}$) et n'est pas intéressante :

Il y a au plus $\min(m_1 - 1, m_2 - 1)$ *valeurs propres non nulles*

φ^2 et valeurs propres l'inertie totale (et donc la somme des valeurs propres) est égale à φ^2. Donc si si $m_1 < m_2$, on obtient $\varphi^2 = \sum_{k=1}^{m_1-1}\lambda_k$.

Choix du nombre de valeurs propres On se contente souvent de regarder le premier plan principal car

— la règle de Kaiser $\lambda_k > \varphi^2/(m_1 - 1)$ s'applique mal ;
— la règle du coude reste valide, mais est subjective ;
— il existe un test sur de la part d'inertie non expliquée, mais il est un peu compliqué.

Résumé des notations

Notation	taille	description
m_1 et m_2	entiers	nombre de modalités des variables 1 et 2
$\mathbf{N} = (n_{ij})$	$m_1 \times m_2$	table de contingence
$\mathbf{D}_1 = \text{diag}(n_{i\cdot})$	$m_1 \times m_1$	effectifs (marges) de lignes
$\mathbf{D}_2 = \text{diag}(n_{\cdot j})$	$m_2 \times m_2$	effectifs (marges) de colonnes
$n_{ij}/n_{i\cdot}$ et $n_{ij}/n_{\cdot j}$	$m_1 \times m_2$	profils lignes et colonnes
$d^2 = n\varphi^2$	réel > 0	χ^2 d'écart à l'indépendance
$\mathbf{a}_k$ et $\mathbf{b}_k$	m_1 et m_2	coordonnées des lignes et colonnes sur l'axe k
λ_k	réel > 0	Valeur propre associée à l'axe k

Partie V. Analyse des correspondances multiples

Analyse des correspondances multiples

But on veut étendre l'AFC au cas de $p \geq 2$ variables $\mathcal{X}_1, \mathcal{X}_2 \ldots, \mathcal{X}_p$ à $m_1, m_2, \ldots, m_p$ modalités. Ceci est particulièrement utile pour l'exploration d'enquêtes où les questions sont à réponses multiples.

Problème l'analyse des correspondances utilise une table de contingence qui est difficilement généralisable au cas $p > 2$.

Méthode on cherche un moyen différent d'analyser $p > 2$ variables et on vérifie que les résultats sont comparables à l'AFC pour $p = 2$.

Les données

Données brutes chaque individu est décrit par les numéros des modalités qu'il possède pour chacune des p variables. Il n'est pas possible de faire des calculs sur ce tableau, où les valeurs sont arbitraires.

Tableau disjonctif on remplace la v-ième colonne par m_v colonnes d'indicatrices : on met un zéro dans chaque colonne, sauf celle correspondant à la modalité de l'individu i qui reçoit 1.

Exemple On interroge 6 personnes sur la couleur de leurs cheveux (`CB`, `CC` et `CR` pour blond, châtain et roux), la couleur de leurs yeux (`YB`, `YV` et `YM` pour bleu, vert et marron) et leur sexe (`H`/`F`). On a donc trois variables (avec respectivement 3, 3 et 2 modalités) mesurées sur 6 individus. Les tableaux brut (ci-dessous à gauche) sont équivalents aux tableaux disjonctifs (à droite).

$$\begin{pmatrix} \texttt{CB} \\ \texttt{CB} \\ \texttt{CC} \\ \texttt{CC} \\ \texttt{CR} \\ \texttt{CB} \end{pmatrix} \begin{pmatrix} \texttt{YB} \\ \texttt{YV} \\ \texttt{YB} \\ \texttt{YM} \\ \texttt{YV} \\ \texttt{YB} \end{pmatrix} \begin{pmatrix} \texttt{H} \\ \texttt{H} \\ \texttt{F} \\ \texttt{H} \\ \texttt{F} \\ \texttt{F} \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \end{pmatrix}$$

Tableau disjonctif et tableau de contingence

Tableau disjonctif à la variable $\mathcal{X}_v$ on associe le tableau disjonctif $\mathbf{X}_v$ à n lignes et m_v colonnes.

Tableau de contingence on vérifie facilement que le tableau de contingence des variables $\mathcal{X}_v$ et $\mathcal{X}_w$ est donné par

$$\mathbf{N}_{vw} = \mathbf{X}_v' \mathbf{X}_w.$$

Effectifs marginaux la matrice diagonale des effectifs marginaux de la variable $\mathcal{X}_v$ est donnée par

$$\mathbf{D}_v = \mathbf{X}_v' \mathbf{X}_v.$$

Exemple (suite) Table de contingence Cheveux/Yeux et matrice d'effectif marginaux de la couleur de cheveux

$$\mathbf{N}_{12} = \begin{pmatrix} 2 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \qquad \mathbf{D}_1 = \begin{pmatrix} 3 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Tableau disjonctif joint

Définition c'est la matrice $\mathbf{X} = (\mathbf{X}_1 | \mathbf{X}_2 | \cdots | \mathbf{X}_p)$, qui possède n lignes et $m_1 + \cdots + m_p$ colonnes. Chaque colonne représente une *catégorie*, c'est-à-dire une modalité d'une variable.

Exemple pour l'exemple de variables précédentes, on a le tableau disjonctif joint suivant

$$\mathbf{X} = \left(\begin{array}{ccc|ccc|cc} 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \end{array} \right)$$

Chaque somme de lignes vaut 3. Les sommes de colonnes valent

$$\left(\begin{array}{ccc|ccc|cc} 3 & 2 & 1 & 3 & 2 & 1 & 3 & 3 \end{array} \right)$$

Le tableau de Burt

Définition c'est un super-tableau de contingence des variables $\mathcal{X}_1, \ldots, \mathcal{X}_p$, formé de tableaux de contingence et de matrices d'effectifs marginaux. :

$$\mathbf{B} = \mathbf{X}'\mathbf{X} = \begin{bmatrix} \mathbf{X}_1'\mathbf{X}_1 & \mathbf{X}_1'\mathbf{X}_2 & \cdots & \mathbf{X}_1'\mathbf{X}_p \\ \mathbf{X}_2'\mathbf{X}_1 & \mathbf{X}_2'\mathbf{X}_2 & & \\ \vdots & & \ddots & \vdots \\ \mathbf{X}_p'\mathbf{X}_1 & \cdots & & \mathbf{X}_p'\mathbf{X}_p \end{bmatrix}$$

$$= \begin{bmatrix} \mathbf{D}_1 & \mathbf{N}_{12} & \cdots & \mathbf{N}_{1p} \\ \mathbf{N}_{21} & \mathbf{D}_2 & & \\ \vdots & & \ddots & \vdots \\ \mathbf{N}_{p1} & \cdots & & \mathbf{D}_p \end{bmatrix}$$

Exemple Toujours pour les mêmes variables

$$\mathbf{B} = \left(\begin{array}{ccc|ccc|cc} 3 & 0 & 0 & 2 & 1 & 0 & 2 & 1 \\ 0 & 2 & 0 & 1 & 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 & 0 & 0 & 1 \\ \hline 2 & 1 & 0 & 3 & 0 & 0 & 1 & 2 \\ 1 & 0 & 1 & 0 & 2 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 & 0 & 1 & 1 & 0 \\ \hline 2 & 1 & 0 & 1 & 1 & 1 & 3 & 0 \\ 1 & 1 & 1 & 2 & 1 & 0 & 0 & 3 \end{array} \right)$$

Partie VI. L'ACM : une AFC sur tableau disjonctif

Comment utiliser l'AFC pour analyser p variables

But on cherche à faire une représentation des $m_1 + \cdots + m_p$ catégories comme points d'un espace de faible dimension.

Méthode on fait une AFC sur le tableau disjonctif joint $\mathbf{X} = (\mathbf{X}_1 | \mathbf{X}_2 | \cdots | \mathbf{X}_p)$.

Les lignes la somme des éléments de chaque ligne de $\mathbf{X}$ est égale à p. Le tableau des profils-lignes est donc $\frac{1}{p}\mathbf{X}$.

Les colonnes la somme des éléments de chaque colonne de $\mathbf{X}$ est égale à l'effectif marginal de la catégorie correspondante. Le tableau des profils colonnes est donc $\mathbf{XD}^{-1}$, où $\mathbf{D}$ est la matrice diagonale par blocs

$$\mathbf{D} = \begin{pmatrix} \mathbf{D}_1 & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & \mathbf{D}_p \end{pmatrix}$$

Les coordonnées factorielles des catégories

Notation On note $\mathbf{a}_k = (\mathbf{a}_{1k}, \dots, \mathbf{a}_{1k})'$ le vecteur à $m_1 + \cdots + m_p$ composantes des coordonnées factorielles des catégories sur l'axe k.

Calcul de l'AFC sur $\mathbf{X}$ comme la matrice des profils lignes est $\frac{1}{p}\mathbf{X}$ et celle des profils colonnes $\mathbf{XD}^{-1}$, $\mathbf{a}_k$ est vecteur propre de

$$(\mathbf{XD}^{-1})'\frac{1}{p}\mathbf{X} = \frac{1}{p}\mathbf{D}^{-1}\mathbf{X}'\mathbf{X} = \frac{1}{p}\mathbf{D}^{-1}\mathbf{B}$$

et donc l'équation des coordonnées des catégories est

$$\frac{1}{p}\mathbf{D}^{-1}\mathbf{B}\mathbf{a}_k = \mu_k\mathbf{a}_k$$

avec la convention de normalisation

$$\frac{1}{np}\mathbf{a}_k'\mathbf{D}\mathbf{a}_k = \mu_k.$$

Résolution dans le cas $p = 2$

On note $\mathbf{a}_k$ (resp. $\mathbf{b}_k$) les m_1 premières (resp. m_2 dernières) coordonnées de la composante principale k et μ_k la valeur propre correspondante :

$$\frac{1}{2}\mathbf{D}^{-1}\mathbf{B}\begin{bmatrix}\mathbf{a}_k\\ \mathbf{b}_k\end{bmatrix} = \frac{1}{2}\begin{bmatrix}\mathbf{I}_{m_1} & \mathbf{D}_1^{-1}\mathbf{N}\\ \mathbf{D}_2^{-1}\mathbf{N}' & \mathbf{I}_{m_2}\end{bmatrix}\begin{bmatrix}\mathbf{a}_k\\ \mathbf{b}_k\end{bmatrix} = \mu_k\begin{bmatrix}\mathbf{a}_k\\ \mathbf{b}_k\end{bmatrix}.$$

On obtient les équations

$$\begin{cases}\mathbf{D}_1^{-1}\mathbf{N}\mathbf{b}_k = (2\mu_k - 1)\mathbf{a}_k\\ \mathbf{D}_2^{-1}\mathbf{N}'\mathbf{a}_k = (2\mu_k - 1)\mathbf{b}_k\end{cases},$$

et donc on retrouve les coordonnées des modalités de lignes et de colonnes dans l'AFC classique (avec $\lambda_k = (2\mu_k - 1)^2$) :

$$\begin{cases}\mathbf{D}_2^{-1}\mathbf{N}'\mathbf{D}_1^{-1}\mathbf{N}\mathbf{b}_k = (2\mu_k - 1)^2\mathbf{b}_k\\ \mathbf{D}_1^{-1}\mathbf{N}\mathbf{D}_2^{-1}\mathbf{N}'\mathbf{a}_k = (2\mu_k - 1)^2\mathbf{a}_k\end{cases}.$$

Différences ACM/AFC pour $p = 2$

Nombre de valeurs propres on a *a priori* $m_1 + m_2 - 2$ valeurs propres non nulles, ce qui est plus important que dans le cas classique. En particulier pour chaque λ_k, on a deux μ_k possibles

$$\begin{cases}\mu_k = \frac{1+\sqrt{\lambda_k}}{2} & \text{associée à } \begin{bmatrix}\mathbf{a}_k\\ \mathbf{b}_k\end{bmatrix}\\ \mu_k' = \frac{1-\sqrt{\lambda_k}}{2} & \text{associée à } \begin{bmatrix}\mathbf{a}_k\\ -\mathbf{b}_k\end{bmatrix}\end{cases}$$

On ne garde donc que les valeurs $\mu_k > \frac{1}{2}$. On peut montrer qu'il y en a $\min(m_1 - 1, m_2 - 1)$.

Inertie L'interprétation de la part d'inertie expliquée par les valeurs propres est maintenant très différente. En particulier les valeurs propres qui étaient très séparées dans l'AFC de $\mathbf{N}$ le sont beaucoup moins dans celle de $\mathbf{X}$.

Partie VII. Aspects pratiques

Formules barycentriques

Les coordonnées des individus Soit $\mathbf{c}_k$ le vecteur à n composantes des coordonnées des n individus sur l'axe factoriel associé à la valeur propre μ_k. D'après les résultats sur l'AFC, on a

$$\mathbf{c}_k = \frac{1}{\sqrt{\mu_k}}\frac{1}{p}\mathbf{X}\mathbf{a}_k \quad \text{et donc} \quad c_{ik} = \frac{1}{\sqrt{\mu_k}}\frac{1}{p}\sum_{j \text{ catégorie de } i} a_{jk}$$

Les seuls termes non nuls dans le calcul de $\mathbf{X}\mathbf{a}_k$ sont les coordonnées de la catégorie de chaque variable possédée par l'individu. Comme on est est dans le cadre de l'AFC, la variance de $\mathbf{c}_k$ est toujours

$$\operatorname{var}\mathbf{c}_k = \frac{1}{n}\mathbf{c}_k'\mathbf{c}_k = \mu_k$$

Barycentre des catégories *À $1/\sqrt{\mu_k}$ près, la coordonnée d'un individu est égale à la moyenne arithmétique simple des coordonnées des catégories auxquelles il appartient.*

Les coordonnées des catégories On a de même la seconde formule

$$\mathbf{a}_k = \frac{1}{\sqrt{\mu_k}}\mathbf{D}^{-1}\mathbf{X}'\mathbf{c}_k \quad \text{c-à-d} \quad a_{jk} = \frac{1}{\sqrt{\mu_k}}\frac{1}{n_j}\sum_{i \text{ de catégorie } j} c_{ik}$$

Les seuls termes non nuls de $\mathbf{X}'\mathbf{c}_k$ sont les coordonnées des individus ayant une catégorie donnée.

Barycentre des individus *À $1/\sqrt{\mu_k}$ près, la coordonnée d'une catégorie est égale à la moyenne arithmétique des coordonnées des n_j individus de cette catégorie.*

Barycentres et représentation

Représentation commune Les points représentatifs des catégories sont barycentres des groupes d'individus. On peut donc représenter individus et catégories dans un même plan factoriel.

Moyennes Comme $\mathbf{c}_k$ est une variable de moyenne nulle, la formule de barycentre indique que pour chaque variable $\mathcal{X}_i$, les coordonnées de ses catégories (pondérés par les effectifs) sont de moyenne nulle. Aucun centrage n'est donc nécessaire

Échelle pour que les catégories se trouvent visuellement au barycentre des individus qui les représentent on peut remplacer $\mathbf{a}_k$ par

$$\boldsymbol{\alpha}_k = \mathbf{D}^{-1}\mathbf{X}'\mathbf{c}_k = \sqrt{\mu_k}\mathbf{a}_k$$

Propriétés des valeurs propres

Valeur propres triviales La valeur propre 1 est associée (comme en AFC) à la composante $\mathbf{z}^0 = (1, \ldots, 1)$ dans l'espace des individus. Les autres vecteurs propres lui sont orthogonaux, et donc de moyenne nulle.

Autres valeurs propres Si $n > \sum_{v=1}^{p} m_v$, le rang de $\mathbf{X}$ est $\sum_{v=1}^{p} m_v - p + 1$ et le nombre de valeurs propres non trivialement égales à 0 ou 1 est $q = \sum_{v=1}^{p} m_v - p$.

Somme La somme des valeurs propres non triviales est donc

$$\sum_{k=1}^{q} \mu_k = \mathrm{Tr}\left(\frac{1}{p}\mathbf{D}^{-1}\mathbf{B}\right) - 1 = \frac{1}{p}\sum_{v=1}^{p} m_v - 1 = \frac{q}{p}$$

La moyenne des q valeurs propres vaut $1/p$.

Sélection de variables et axes

Sélection des variables on décide souvent de ne garder qu'un nombre réduit de variables actives et de garder les autres comme variables supplémentaires.

Sélection des axes

— règle courante : garder les axes tels que $\mu_k > 1/p$ (la moyenne des valeurs propres est $1/p$).

— les axes intéressants sont ceux que l'on peut interpréter, en regardant les contributions des variables actives et les valeurs-tests associées aux variables supplémentaires (définies plus tard).

— En pratique on se contente souvent d'interpréter le premier plan principal.

Inertie expliquée elle est moins intéressante qu'en ACP.

Catégories et axes factoriels

Si n_j est l'effectif de la catégorie j et a_{jk} sa coordonnée sur l'axe factoriel k, alors

$$\mathrm{var}\,\mathbf{a}_k = \sum_{j \in \text{catégories}} \frac{n_j}{np}(a_{jk})^2 = \mu_k$$

Catégorie La contribution de la catégorie j à l'axe factoriel est

$$\frac{n_j}{np}\frac{(a_{jk})^2}{\mu_k},$$

intéressante si elle est supérieure au poids n_j/np (à un facteur près comme en ACP et AFC).

Variable la contribution totale de la variable $\mathcal{X}_v$ à l'axe factoriel est

$$\frac{1}{\mu_k}\frac{1}{np}\sum_{j \text{ modalité de } \mathcal{X}_v} n_j (a_{jk})^2$$

Individus et axes factoriels

La normalisation de $\mathbf{c}_k$ est $\sum_{i=1}^{n}(c_{ik})^2 = n\mu_k$, où c_{ik} est la coordonnée de l'individu i sur l'axe factoriel k associé à la valeur propre μ_k.

Contribution d'un individu elle est égale pour l'individu i à

$$\frac{1}{n}\frac{(c_{ik})^2}{\mu_k}$$

Cette contribution est jugée en la comparant au poids $1/n$ comme en ACP et AFC.

Qualité de la représentation pour le sous-espace formé par les ℓ premier axes, la qualité de la représentation de l'individu i est le cosinus carré habituel

$$\frac{\sum_{k=1}^{\ell}(c_{ik})^2}{\sum_{k=1}^{q}(c_{ik})^2}$$

On définit de même sur les a_{jk} la qualité de la représentation d'une catégorie j.

Contribution à l'inertie totale

Soit $\mathbf{x}^j = (x_i^j)$ le vecteur colonne de $\mathbf{X}$ correspondant à une catégorie j. On rappelle que l'inertie totale vaut

$$I_{\mathbf{g}} = \sum_{j \in \text{catégories}} \frac{n_j}{np} d^2(\mathbf{z}^j, \mathbf{g}) = \frac{1}{p}\sum_{v=1}^{p} m_v - 1,$$

où la distance du profil-colonne j au centre de gravité des profils-colonnes $\mathbf{g} = \mathbf{1}/n$ est

$$\begin{aligned} d^2(\mathbf{z}^j, \mathbf{g}) &= \sum_{i=1}^{n} \frac{np}{p}\left(\frac{x_i^j}{n_j} - \frac{1}{n}\right)^2 = n\sum_{i=1}^{n}\left(\frac{x_i^j}{n_j^2} + \frac{1}{n^2} - \frac{2x_i^j}{nn_j}\right) \\ &= n\left(\frac{n_j}{n_j^2} + \frac{n}{n^2} - \frac{2n_j}{nn_j}\right) = \frac{n}{n_j} - 1 \end{aligned}$$

Contribution d'une catégorie La contribution absolue de la catégorie à l'inertie est

$$\frac{n_j}{np} d^2(\mathbf{z}^j, \mathbf{g}) = \frac{1}{p}\left(1 - \frac{n_j}{n}\right),$$

qui est une fonction décroissante de l'effectif. Il faut donc éviter les catégories d'effectif trop faible, qui d'ailleurs se retrouveront dans les premiers axes

Contribution d'une variable La contribution de la variable $\mathcal{X}_v$ est

$$\sum_{j \text{ modalité de } \mathcal{X}_v} \frac{1}{p}\left(1 - \frac{n_j}{n}\right) = \frac{m_v - 1}{p}$$

Elle est d'autant plus grande que le nombre de modalités de $\mathcal{X}_i$ est élevé. Il faut donc éviter les disparités trop grandes entre les nombre de modalités (quand on a le choix du découpage...)

Partie VIII. Interprétation externe

Les variables supplémentaires

Leur usage est très courant en analyse des correspondances multiples.

Variables quantitatives on calcule « à la main » leur corrélation avec les axes factoriels et on les place sur un cercle de corrélations. Si $\hat{\mathbf{z}}$ est une version centrée-réduite de la variable, alors

$$\operatorname{cor}(\hat{\mathbf{z}}, \mathbf{c}_k) = \frac{1}{\sqrt{\mu_k}} \frac{1}{n} \sum_{i=1}^{n} \hat{z}_i c_{ik}$$

On peut aussi les découper en classes et les traiter comme des variables qualitatives.

Variables qualitatives on calcule directement les coordonnées de leurs modalités en utilisant la formule de barycentre des individus : la coordonnée la catégorie supplémentaire $\hat{\jmath}$ sur l'axe principal k est

$$a_{\hat{\jmath}k} = \frac{1}{\sqrt{\mu_k}} \frac{1}{n_{\hat{\jmath}}} \sum_{i \text{ de catégorie } \hat{\jmath}} c_{ik}$$

Valeurs-test pour les variables supplémentaires qualitatives

But on cherche à savoir si une catégorie $\hat{\jmath}$ d'effectif $n_{\hat{\jmath}}$ et de coordonnée $a_{\hat{\jmath}k}$ sur cet axe est liée à cet axe.

Idée du calcul si les $n_{\hat{\jmath}}$ individus d'une catégorie étaient pris au hasard, la moyenne de leurs coordonnées serait une variable aléatoire centrée (les $\mathbf{c}$ sont de moyenne nulle) et de variance $\frac{\mu_k}{n_{\hat{\jmath}}} \frac{n-n_{\hat{\jmath}}}{n-1}$. De plus, la moyenne des coordonnées est égale à $\sqrt{\mu_k} a_{\hat{\jmath}k}$.

Valeur-test c'est la version centrée et réduite de la moyenne des coordonnées

$$a_{\hat{\jmath}k} \sqrt{n_{\hat{\jmath}}} \sqrt{\frac{n-1}{n-n_{\hat{\jmath}}}}.$$

Quand $n_{\hat{\jmath}}$ et $n - n_{\hat{\jmath}} =$ sont assez grand (en général > 30), elle est significative si elle est supérieure à 2 ou 3 en valeur absolue. On ne doit pas l'utiliser sur les variables actives.

Partie IX. Récapitulatif

Notations

Notation	taille	description
$m_1, \dots, m_p$	entiers	nombres de modalités des variables
$\mathbf{X} = (x_i^j)$	$n \times$ (nb. cat.)	tableau disjonctif
$\mathbf{N}_{k\ell}$	$m_k \times m_\ell$	table de contingence des variables k et ℓ
$\mathbf{D} = \operatorname{diag}(n_i)$	nb. cat.	effectifs (marges) des catégories
$\mathbf{B}$	(nb. cat.)^2	matrice de Burt
μ_k	réel > 0	Valeur propre associée à l'axe k
$\mathbf{a}_k$	nb. cat.	coordonnées des catégories sur l'axe k
$\mathbf{c}_k$	n	coordonnées des individus sur l'axe k

nb. cat. $= m_1 + \cdots + m_p$.

Points communs entre AFC et ACM

But	décrire les liaisons entre plusieurs variables qualitatives
Cas $p = 2$	les coordonnées des modalités sont les mêmes pour les deux analyses
Représentation	toutes les modalités peuvent être représentées sur le même diagramme
Contribution d'une modalité à un axe	$\text{poids} \times \dfrac{(\text{coordonnée})^2}{\text{valeur propre}}$
Qualité de la représentation d'une modalité par un sous espace	$\cos^2\theta = \dfrac{\sum_{\text{axes du sous esp.}} (\text{coord sur l'axe})^2}{\sum_{\text{tous les axes}} (\text{coord sur l'axe})^2}$

Différences entre AFC et ACM

	AFC	ACM
Individus	non	oui
Données	tableau de contingence profils lignes/colonnes	tableau disjonctif tableau de Burt
Poids d'une modalité	$\frac{n_{i\cdot}}{n}$ (profil-ligne) $\frac{n_{\cdot j}}{n}$ (profil-colonne)	$\frac{n_j}{np}$
Nb de val. propres	$\min(m_1 - 1, m_2 - 1)$	$\sum_{v=1}^{p} m_v - p$
Axes à conserver	pas de règle Kaiser ; peut-être part d'inertie.	$\mu > \frac{1}{p}$
Variables supplémentaires	pas vraiment de sens	qualitatives et quantitatives