

Analyse des correspondances binaires

Interprétation

Véronique Tremblay

- Analyse en composante principale
- Analyse des correspondances binaires
- Analyse des correspondances multiples
- Analyse factorielle de données mixtes
- Analyse factorielle multiple
- Analyse factorielle hiérarchique

Analyse des correspondances multiples (ACM)

Tableau disjonctif complet (tableau logique)

$$Z = [Z_1|Z_2|\dots|Z_Q]$$

| ID | Type d'employé | | | | | Type de fumeur | | | Total |
|-----|----------------|---|---|---|---|----------------|---|---|-------|
| | 1 | 2 | 3 | 4 | 5 | 1 | 2 | 3 | Q |
| 1 | 0 | 0 | 1 | 0 | 0 | 0 | 1 | 0 | 2 |
| 2 | 0 | 1 | 0 | 0 | 0 | 1 | 0 | 0 | 2 |
| 3 | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 2 |
| 4 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 1 | 2 |
| ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ |
| 193 | 0 | 1 | 0 | 0 | 0 | 1 | 0 | 0 | 2 |

Notation supplémentaire

Un peu de notation additionnelle:

Q = nombre de questions

n = nombre d'individus répondant au questionnaire

p_q = nombre de modalités de la question q

$p = p_1 + \dots + p_Q$

Matrice carrée $p \times p$ définie comme étant

$$B = ZZ^{\top}$$

$$B = \begin{bmatrix} Z_1^{\top} Z_1 & Z_1^{\top} Z_2 & \cdots & Z_1^{\top} Z_Q \\ Z_2^{\top} Z_1 & Z_2^{\top} Z_2 & \cdots & Z_2^{\top} Z_Q \\ \vdots & \vdots & \cdots & \vdots \\ Z_Q^{\top} Z_1 & Z_Q^{\top} Z_2 & \cdots & Z_Q^{\top} Z_Q \end{bmatrix}.$$

Tableau de Burt

| | AS | AJ | ES | EJ | SE | NON | MOY | GRO |
|-----|----|----|----|----|----|-----|-----|-----|
| AS | 11 | 0 | 0 | 0 | 0 | 4 | 5 | 2 |
| AJ | 0 | 18 | 0 | 0 | 0 | 4 | 10 | 4 |
| ES | 0 | 0 | 51 | 0 | 0 | 25 | 22 | 4 |
| EJ | 0 | 0 | 0 | 88 | 0 | 18 | 57 | 13 |
| SE | 0 | 0 | 0 | 0 | 25 | 10 | 13 | 2 |
| NON | 4 | 4 | 25 | 18 | 10 | 61 | 0 | 0 |
| MOY | 5 | 10 | 22 | 57 | 13 | 0 | 107 | 0 |
| GRO | 2 | 4 | 4 | 13 | 2 | 0 | 0 | 25 |

- L'ACM permet de projeter l'ensemble des modalités de toutes les variables dans le même plan.
- L'interprétation est similaire à celle de l'ACB:
 - Les modalités d'une même variable sont éloignées
 - Les modalités de deux variables différentes sont proches si elles sont partagées par les mêmes individus.
 - Une modalité est éloignée des autres si elle n'est pas partagée par les mêmes individus

Tableau disjonctif complet vs tableau de Burt

- L'analyse des correspondances multiple peut être effectuée sur l'un ou l'autre des tableaux.
- Les résultats seront identiques (voir les notes de cours)
- Quand privilégier l'un ou l'autre?

Méthodes plus avancées

Analyse factorielle de données mixtes (AFDM)

Permet de présenter des variables continues et catégorielles dans le même plan.

Analyse factorielle multiple (AFM)

Permet de grouper les variables ou les individus (AFMD).

Analyse factorielle multiple hiérarchique (AFMH)

Permet de grouper les variables de façon hiérarchique.

FactoMineR

| Analyse | Fonction |
|---------|----------|
| ACP | PCA |
| ACB | CA |
| ACM | MCA |
| AFDM | FAMD* |
| AFM | MFA |
| AFMH | HMFA* |