

Analyse des données

Projet 1 (Groupe 1)

Considérons le tableau des données ci-dessous résumant les résultats d'une enquête effectuée sur quelques étudiants de certaines universités afin d'étudier les relations entre le choix de la **Discipline** (**10 modalités c'est-à-dire 10 disciplines**) et **LMD (H, F)** le niveau de la discipline par étudiants (Hommes ou Femmes) avec **6 modalités**.

Avec :

L-H : Licence pour hommes, L-F : Licence pour femmes.

M-H : Master pour hommes, M-F : Master pour femmes.

D-H : Doctorant pour hommes, D-F : Doctorant pour femme.

Par conséquent, les données sont les effectifs par niveau pour hommes et femmes.

| LMD (H, F) Discipline | L-H | L-F | M-H | M-F | D-H | D-F |
|--|------------|------------|------------|------------|------------|------------|
| Informatique | 54861 | 22559 | 48293 | 17078 | 11491 | 4407 |
| Science et vie | 37317 | 69373 | 21693 | 42371 | 4342 | 4029 |
| Langues | 21291 | 62736 | 3874 | 13186 | 907 | 1839 |
| Droit | 17850 | 48691 | 5853 | 17672 | 2401 | 4531 |
| Génie civile | 17253 | 8248 | 4172 | 1963 | 328 | 188 |
| Administration | 12388 | 18574 | 2884 | 4183 | 0 | 0 |
| Sciences humaines | 41050 | 94346 | 20447 | 43016 | 6972 | 7787 |
| Sciences humaines | 726 | 1779 | 811 | 2356 | 15 | 13 |
| Gestion | 37157 | 38387 | 26929 | 29466 | 2552 | 1983 |
| Biologie | 15004 | 24318 | 8457 | 11090 | 5232 | 5641 |

Pour atteindre le but de cette étude, nous réalisons une AFC sur les données de ce tableau de contingence.

- 1) Déterminer et afficher la matrice des fréquences relatives ainsi que les fréquences marginales.
- 2) Déterminer et afficher les matrices des profils lignes et colonnes.
- 3) Afficher un tableau résumant les informations suivantes liées au nuage des profils-lignes :

Analyse des données

- Les valeurs propres et représenter les graphiquement.
 - Le **Taux** d'inertie (par rapport à l'inertie totale) porté par chaque **axe principal**.
 - Les **Taux cumulés** de l'information retenue.
- 4) Donner les axes principaux de l'ACP des profils-lignes et ceux de l'ACP des profils-colonnes.
 - 5) Déterminer les **Facteurs lignes** (Composantes principales) correspondant aux axes du plan principal. On les note par : ϕ_i . Présenter les résultats dans un tableau.
 - 6) Déterminer les **contributions relatives et absolues** des lignes sur chaque axe. Présenter les résultats obtenus dans deux tableaux. Un tableau pour chaque type de contribution.
 - 7) Déterminer les **Facteurs colonnes** du nuage des profils-colonnes. On les notera par ψ_i . Afficher les résultats obtenus dans un tableau à deux colonnes.
 - 8) Reprenez la question 5) pour les profils colonnes.
 - 9) **Visualiser** les deux nuages dans le nouveau plan principal (présentation simultanée). **Etiqueter** les profils-lignes et les profils-colonnes et **Utiliser** deux couleurs différentes afin de distinguer entre les deux nuages.
 - 10) Analyser et interpréter les résultats obtenus.
 - 11) Que signifie chaque axe ?