

## Analyse des données

### TP 2

#### Objectif

Le but de ce TP est de réaliser une ACP normée, il s'agit de déterminer les composantes principales permettant de résumer l'information des données initiales, de visualiser les individus dans le nouveau espace (espace de projection) et d'analyser les résultats obtenus ainsi que les graphes obtenus.

#### Exercice 1

Dans ce premier exercice, nous reprenons les données de l'exercice 1 du TP 1 ainsi que les quelques résultats utiles et fondamentaux pour réaliser une ACP.

| $I \backslash J$ | Moyenne | Age | Taille | Poids |
|------------------|---------|-----|--------|-------|
| Enfant 1         | 14      | 13  | 1.50   | 45    |
| Enfant 2         | 16      | 13  | 1.60   | 50    |
| Enfant 3         | 15      | 13  | 1.65   | 50    |
| Enfant 4         | 9       | 15  | 1.75   | 60    |
| Enfant 5         | 10      | 14  | 1.70   | 60    |
| Enfant 6         | 7       | 14  | 1.70   | 60    |
| Enfant 7         | 8       | 14  | 1.60   | 70    |
| Enfant 8         | 13      | 13  | 1.60   | 65    |
| Enfant 9         | 17      | 15  | 1.55   | 60    |
| Enfant 10        | 11      | 14  | 1.70   | 65    |

Ainsi,

- 1) Afficher la matrice centrée réduite de la matrice des données.
- 2) Afficher une matrice à **trois colonnes** qui nous retourne les **informations suivantes** :
  - Les valeurs propres de la matrice des corrélations  $R$  des données de la matrice  $X$  (déjà calculées) sur la première colonne.
  - Le **Taux** d'inertie (par rapport à l'inertie totale) projeté sur chaque **axe principal** sur la 2<sup>ème</sup> colonne.
  - Les **Taux cumulés** de l'information retenue sur la dernière colonne.
  - Commenter les résultats obtenus.
- 3) Déterminer tous les axes principaux associés aux valeurs propres déjà calculées.
- 4) Calculer la somme des inerties de tous les axes. Que constatez-vous ?
- 5) Déterminer les coordonnées de tous les individus sur tous les axes principaux. Afficher les résultats dans une matrice.

### *Analyse des données*

- 6) Déterminer le **meilleur plan ajustant** le nuage des points des individus. Préciser les **composantes principales** c'est-à-dire les nouvelles variables.
- 7) Exprimer les **nouvelles variables** c'est-à-dire les **composantes principales** en fonction des **variables initiales**.
- 8) Calculer la variance de chaque composante principale. Que constatez-vous ?
- 9) En déduire la covariance entre les deux composantes principales. Que constatez-vous ?
- 10) Représenter graphiquement tous les individus dans le nouveau plan. Nommer vos axes ainsi que les individus dans votre graphe. Préciser (afficher) le taux d'inertie porté par chaque axe au-dessous de l'axe correspondant.
- 11) Déterminer les individus homogènes. Autrement dit ceux qui forment des groupes homogènes. Afficher le résultat dans des listes séparées.

Justifier votre réponse en précisant le critère de regroupement.

#### **Indication :**

Vous pouvez exploiter la matrice demandée à déterminer dans la question 5).

- 12) Dans une autre fenêtre de présentation graphique, encadrer les groupes homogènes.
- 13) Déterminer les individus qui sont proches de l'origine. Que signifient ça ?
- 14) Déterminer les individus qui sont éloignés. Par quoi, pouvez-vous interpréter ça ?