

- Analyse de données -  
**TP 3 Analyse en composantes principales**

---

## 1 Introduction

La revue "60 Millions de consommateurs" a présenté dans son numéro de novembre 2001 une étude sur les téléphones mobiles aboutissant à un tableau comportant un nombre important de caractéristiques relatives à l'utilisation de ces appareils. Une part de ces caractéristiques représentent des observations correspondant à des mesures objectives, associées le plus souvent à des jugements de valeur exprimés sur une échelle comportant cinq degrés d'appréciation. Dans le but de fournir une cartographie de ces appareils pouvant servir à les différencier selon le maximum de ces critères il a été construit un tableau de données ci-dessous, retenant les variables jugées les plus pertinentes. Créer le à partir du fichier de données mobiles.dat en supprimant les 2 premières colonnes.

1) Les appareils testés sont donc au nombre de 15 ; le numéro de modèle n'a été précisé que lorsque plusieurs téléphones d'une même marque sont présents. On ne donne volontairement pas d'explication sur le sens des variables, même si certaines sont simples à deviner : l'analyse sera faite sur les propriétés numériques des variables, pas leur interprétation. Pour pouvoir utiliser l'ACP centrée réduite les degrés de l'échelle de jugement ont été transformés en notes allant de 1 à 5. Pour toutes ces données, les valeurs les plus grandes sont considérées comme meilleures.

Calculer la matrice de corrélation des 8 variables :

Quelles sont les variables les plus corrélées entre elles ? Les moins corrélées entre elles ?  
Quelles sont les variables les plus opposées ?

## 2 Première analyse en composantes principales

2) Faire une analyse en composantes principales des données centrées-réduites en utilisant la fonction `prcomp`. Quelles sont les valeurs propres obtenues ?

3) Faire une représentation en histogramme des valeurs propres. Combien de composantes principales faut-il retenir ? Quel est le pourcentage d'inertie totale expliquée par le sous-espace principal correspondant ? Si on ne retient que le premier plan principal, quel est le ratio expliqué ? La qualité de l'ACP est-elle bonne si on ne conserve que le premier plan principal ?

4) Représenter le cercle des corrélations du premier plan principal (axe 1 en abscisse, axe 2 en ordonnée), accompagné du tableau des corrélations entre les variables et les 2 (ou 3) premiers facteurs, la projection des individus sur le premier plan principal et les contributions (en 10000ièmes) de ces individus aux 2 (ou 3) premiers facteurs.

Quelles sont les variables qui déterminent la première composante principale ? La proximité de ces variables sur le cercle est-elle cohérente avec ce que l'on avait remarqué en 1) ?

Quels sont les 2 individus qui déterminent le plus la première composante principale ? Quelle est leur contribution aux axes ? Comparer ces contributions au poids respectif de ces individus. Quel est le lien entre les variables de la question précédente et ces 2 individus ?

Quels sont les 2 individus qui déterminent le plus la première composante principale ? Quelle est leur contribution aux axes ? Comparer ces contributions au poids respectif de ces individus. Quel est le lien entre les variables de la question précédente et ces 2 individus ?

### 3 Nouvelle ACP

On décide de faire une nouvelle analyse en composantes principales sur les variables centrées-réduites en passant les 2 téléphones mobiles perturbateurs en éléments supplémentaires.

5) Calculer les mêmes caractéristiques que dans la section précédente (valeurs propres, corrélations entre variables et facteurs, projection des individus sur le premier plan principal et contributions des individus aux axes, qualité de représentation des individus) :

6) Expliquer en quoi la nouvelle analyse est différente de la précédente et pourquoi elle est utile. Combien de valeurs propres faut-il retenir, et quelle proportion d'inertie totale cela représente-t-il ?

7) Représenter les individus supplémentaires :

Les 2 individus supplémentaires sont-ils bien représentés sur le plan principal ? Pour cela, on calcule leur "cosinus carré".