

TP2 : enquête sur les clients d'une banque

On veut analyser des données relatives à une enquête réalisée sur 810 clients d'une banque. On s'intéresse tout d'abord aux variables suivantes

- **solde** : solde du compte courant (p4 (>20 k€), p3 (12–20 k€), p2 (4–12 k€), p1 (0–4k€), n1 (-4–0 k€), n2 (< - 4 k€)).
- **interdit** : interdit de chéquier (oui ou non)
- **eparlog** : plan d'épargne logement (nul (non), fai (<20 k€), for (>20 k€))
- **eparliv** : livret d'épargne (nul (non), fai (<20 k€), for (>20 k€))
- **credcon** : crédits à la consommation (nul (non), fai (<20 k€), for (>20 k€))
- **porttit** : portefeuille de titres (nul (non), fai (<20 k€), moy (20–100 k€), for (>100 k€))

Le code pour charger les données est le suivant :

```
> require(ade4)
> source("fonctions.R")
> data(banque)
> banque1=subset(banque,select=c(solde,interdit, eparlog, eparliv, credcon, porttit))
> # renomme solde en solde
> colnames(banque1)[1]="solde"
```

1 Les données

Question 1 Calculez le tableau de Burt des données ci-dessus avec la fonction `acm.burt`.

Question 2 Calculez avec la fonction `chisq.test` le χ^2 des variables `eparlog` et `credcon`. Que pouvez-vous en dire ? Même question avec les variables `eparlog` et `eparliv`.

2 Analyse des correspondances multiples

Question 3 Réaliser une ACM sur les données ci-dessus à l'aide de la fonction `dudi.acm`.

Question 4 Combien d'axes propres faut-il conserver d'après les règles habituelles ? Quelle est la proportion de l'inertie expliquée par le sous-espace propre correspondant (on pourra utiliser la fonction `inertia.dudi`) ?

Question 5 Faites une représentation des catégories. Quelles sont celles qui déterminent les deux premiers axes principaux ? (on détaillera les critères et on cherchera à être précis dans la réponse).

Question 6 Interprétez les axes et identifiez les difficultés.

2.1 Catégories supplémentaires

On cherche à préciser les caractéristiques des axes en termes de type de client. On s'intéresse donc aux variables supplémentaires suivantes :

- **age** : âge du client (ai25 [18, 25[, ai35 [25, 35[, ai45 [35, 45[, ai55 [45, 55[, ai75 [55, 75]);
- **sexe** : sexe du client (hom ou fem);
- **csp** : catégorie socio-professionnelle : agriculteur (`agric`), artisan (`artis`), cadre supérieur (`cadsu`), profession intermédiaire (`inter`), employé (`emplo`), ouvrier (`ouvri`), retraité (`retra`), inactif (`inact`) et étudiant (`etudi`).

Question 7 Utilisez `acm.suppl` du fichier `fonctions.R` pour calculer les effectifs et les valeurs tests correspondant aux individus supplémentaires.

Question 8 Vérifier si on peut utiliser les valeurs-test. Complétez l'interprétation des deux premiers axes grâce à ces valeurs.