

TP RD6 Analyse des correspondances multiples

Le fichier sciences.csv contient des données issues d'une enquête réalisée en 1993 auprès de $n = 871$ enquêtés sur leur avis sur le rôle de la science. Les enquêtés doivent répondre s'ils sont d'accord ou non avec les déclarations suivantes :

- A : nous croyons trop souvent en la science et pas assez à l'intuition (aux sentiments et aux croyances).
- B : globalement les sciences font plus de mal que de bien.
- C : tous les changements que les hommes font à la nature, aussi scientifiques soient-ils, sont susceptibles d'empirer les choses.
- D : la science va résoudre les problèmes d'environnement sans affecter notre façon de vivre.

Il y a cinq réponses possibles :

- 1 tout à fait d'accord ;
- 2 d'accord ;
- 3 pas d'avis ;
- 4 pas d'accord ;
- 5 absolument pas d'accord.

Puis des questions de signalétique leur ont été posées :

- Sexe : Homme (sex1), Femme (sex2)
- Age (6 classes) : 16-24 (age1), 25-34, 35-44, 45-54, 55-64, 65 et plus (age6)
- Niveau d'éducation (6 niveaux) : enseignement primaire (edu1), niveau collège, niveau lycée, niveau Bac + 2, niveau Bac + 5, niveau Bac + 8 (edu6)

Le but de l'étude est de caractériser, de décrire la population de cette enquête, de faire une typologie des enquêtés en fonction de leur attitude vis-à-vis des sciences.

1. Importer le fichier sciences.csv. En utilisant la commande summary, expliquer pourquoi le tableau n'est pas adapté pour une ACM. Le transformer en conséquence.

```
library(FactoMineR)
library(factoextra)
data=read.csv("sciences.csv",sep=";")
summary(data)

for (i in 1:7) data[,i]=as.factor(data[,i])
summary(data)
```

2. Construire le tableau disjonctif et le tableau de Burt.

```
Z=tab.disjonctif(data)
t(Z)%*%Z
```

3. Réaliser l'ACM sans l'analyser de ce tableau avec FactoMineR.

```
acm=MCA(data)
```

4. Toutes les variables ont-elles le même poids? La variable sexe a-t-elle un rôle important? Reprendre l'analyse en mettant cette variable en supplémentaire.

```
acm$var$eta2
plotellipses(acm,keepvar=c(5),means=FALSE)
acm=MCA(data,quali.sup=5)
```

5. Les représentations graphiques suggèrent un effet particulier, lequel? Expliquez le.

```
table(data$A,data$C)
image(1:5,1:5,t(table(data$A, data$C)),xlab="C",ylab="A")
table(data$A,data$B,data$C)
```

5. Combien d'axes sont définis? Déterminer les axes pertinents

```
fviz_eig(acm, addlabels = TRUE)
E=acm$eig[,1]
q=length(acm$call$quali)#q = nombre de variables actives
E[acm$eig[,1]<1/q]=1/q
E=E[E-1/q>0]
E
E=(q/(q-1))^2*(E-1/q)^2
barplot(E)
title("Diagramme du critère de Benzecri")
```

6. Déterminer les variables et les modalités pertinentes pour expliquer le plan F1-F2.

```
acm$var$eta2
acm$var$v.test
plotellipses(acm,means=FALSE)
```

7. Comment peut-on expliquer le faible rapport de corrélation entre la 4ème variable D et la première dimension ?

8. Interpréter finement la proximité des modalités "C5" et "B5".

9. Y-a-t'il un lien entre le niveau d'éducation et le comportement face à la science ?

```
par(mfrow=c(3,2))
for (i in 1:5) barplot(table(data[data$A==i,"edu"]))
par(mfrow=c(1,1))
acm.edu = MCA(data[,c("A","D","edu")])

plot(acm.edu,invisible="ind")
```

10. Comment sont positionnées les modalités Homme et Femme ?

```
plotellipses(acm,keepvar=c(5),means=FALSE)
```

11. Comment peut-on faire pour étudier le comportement des femmes qui ont un haut niveau d'éducation et des hommes qui ont un haut niveau d'éducation ? Que concluez-vous ?

```
par(mfrow=c(2,1))
barplot(table(data[data$sex==1&data$edu==5,"A"]))
barplot(table(data[data$sex==2 & data$edu==5,"A"]))
```

12. Comment interpréter la distance entre deux individus ? Quelle distance est utilisée en AFCM ?