

L'objet de ce T.P. est de manipuler les procédures SAS permettant de mettre en œuvre les principales méthodes de statistique descriptive multidimensionnelle : Analyse en Composantes Principales (A.C.P.), Analyse Factorielle des Correspondances (A.F.C.) et Analyse des Correspondances Multiples (A.C.M.). Pour l'A.C.P., on va utiliser des macros SAS archivées et accessibles dans le répertoire : `~besse/sas/macros`.

On se reportera au chapitre 8 du cours photocopié pour plus de détails sur les différentes procédures.

### Exercice 08.1

Pour se familiariser avec les macros d'A.C.P., taper les instructions suivantes :

```
%acp(notes,ind,MECA ALIN ALGB ANLS STAT,red=cov,q=4);
```

cette macro réalise l'A.C.P. sans produire de graphique.

*Nota* : La structure des paramètres dans la macro `acp` est la suivante : nom de la table SAS permanente à analyser (`notes` pour `sasuser.notes`); nom de la variable contenant les étiquettes des individus; liste des noms des variables à prendre en compte (les séparateurs sont des blancs); `red=cov` : paramètre facultatif à mettre si l'on désire une A.C.P. non réduite (par défaut, l'A.C.P. est réduite); `q=4` : paramètre facultatif pour indiquer le nombre de colonnes à produire pour les coordonnées des individus sur les axes, les contributions des individus à l'inertie des axes et les cosinus carrés entre les vecteurs représentant les individus et leurs projections sur les axes (valeur par défaut : 3).

Faire ensuite :

```
%gacpix;
```

cette macro réalise le graphique des individus sur les 2 premiers axes. Pour un graphique sur les axes 2 et 3, faire :

```
%gacpix(x=2,y=3);
```

Continuer avec :

```
%gacpvx;
```

réalise le graphique des variables sur les 2 premiers axes (même chose que la macro précédente pour d'autres axes).

```
%gacpsx;
```

réalise le *scree-graph*, ou éboulis des valeurs propres.

```
%gacpbx;
```

réalise les boîtes parallèles des composantes principales.

### Exercice 08.2

Recommencer maintenant avec les instructions suivantes :

```
%acp(notes,ind,MECA ALIN ALGB ANLS STAT,red=cov);
```

```
%epsf(notesind);
```

```
%gacpix;
```

```
%epsf(notesvar);
```

```
%gacpvx;
```

```
%epsf(notesebo);
```

```
%gacpsx;
```

```
%epsf(notesbox);
```

```
%gacpbx;
```

Commenter.

### Exercice 08.3

On s'intéresse maintenant à l'A.F.C. Copier les fichiers suivants :

```
~baccini/tpsas/exolim/data/agr.txt
~baccini/tpsas/exolim/data/agr.don
~baccini/tpsas/exolim/exo08_3.sas
```

Lire le premier, lancer l'exécution du troisième et commenter.

Lorsqu'on réalise une A.F.C. avec la procédure `corresp`, il convient de noter les points suivants.

- Lorsque le fichier des données est constitué de la table de contingence à analyser (c'est le cas ici), la commande `var` est nécessaire pour déclarer les variables intervenant en colonnes (elles doivent être quantitatives puisqu'elles contiennent des effectifs).
- Toujours dans ce cas, la commande `id` permet de déclarer la variable intervenant en lignes (elle est nécessaire pour pouvoir représenter ces lignes dans le graphique ; elle est qualitative puisqu'elle contient des étiquettes).
- Si le fichier des données comporte en lignes les individus et en colonnes les 2 variables qualitatives codées, la commande `var` doit être remplacée par la commande `tables` (voir l'exercice 08.4).

### Exercice 08.4

On s'intéresse enfin à l'A.C.M. Copier les fichiers suivants :

```
~baccini/tpsas/exolim/data/deug.txt
~baccini/tpsas/exolim/data/deug.don
~baccini/tpsas/exolim/exo08_4.sas
```

Lire le premier, lancer l'exécution du troisième et commenter.

Pour réaliser une A.C.M. avec SAS, on retiendra les points suivants.

- C'est encore la procédure `corresp` que l'on utilise.
- L'option `mca` (*multiple correspondence analysis*) est nécessaire pour réaliser l'A.C.M.
- Dans une A.C.M., le fichier des données se présente en général sous la même forme que dans le fichier `deug.don` : les individus en lignes et les variables codées en colonnes. Une phase de recodification est alors nécessaire pour définir les nouvelles variables contenant les étiquettes qui apparaîtront sur le graphique.
- La commande `tables` permet de déclarer les variables prises en compte dans l'A.C.M. et donc de construire le tableau de Burt correspondant.
- L'option `observed` permet de faire afficher, dans les sorties, la table analysée, c'est-à-dire ici le tableau de Burt.