

Dans ce T.P., on va étudier la procédure `freq` de SAS qui permet de traiter une ou plusieurs variables qualitatives. On va également réaliser des graphiques, en particulier les premiers graphiques SAS “haute résolution”.

Le support est constitué des chapitres 3 et 5 du cours polycopié.

Exercice 03.1

Copier les fichiers suivants :

`~baccini/tpsas/exolim/data/quali.don`

`~baccini/tpsas/exolim/data/quali.txt`

Ils sont relatifs à un exemple fictif de données qualitatives.

1. Lire ces données, puis utiliser la procédure `freq` (page 34) pour faire un tri à plat de chacune des 2 variables (utiliser la commande `tables` — le `s` est facultatif — suivie du nom des variables).
2. Toujours avec la procédure `freq`, construire la table de contingence croisant le sexe en lignes et la couleur des cheveux en colonnes (toujours la commande `tables` ; le symbole du croisement de 2 variables est le signe `*`).
3. Dans la procédure ci-dessus, après la commande `tables`, ajouter l’option `/chisq` pour obtenir le khi-deux. Commenter.
4. Utiliser maintenant la procédure `chart` (pages 35–36) pour réaliser des graphiques relatifs aux 2 variables considérées. On utilisera la commande `vbar` pour obtenir un diagramme en colonnes (barres verticales), `hbar` pour obtenir un diagramme en barres horizontales et `pie` pour obtenir un diagramme en secteurs. Observer l’effet produit par les options `/type = percent`, `/type = cfreq` et `/type = cpercent` dans ces différents graphiques.

Exercice 03.2

Reprendre la dernière question de l’exercice précédent en réalisant 3 graphiques “haute résolution”.

1. Dans un premier temps, changer simplement la commande `chart` en la commande `gchart`. Remarquer que les graphiques sont ainsi obtenus dans une nouvelle fenêtre de SAS, spécialement ouverte par cette procédure : la fenêtre graphique (*graph1*).
2. Pour sauvegarder le premier graphique dans un fichier au format *postscript encapsulé*, ajouter les 2 commandes suivantes avant de relancer la procédure `gchart` (voir page 40) :

```
goptions device=psepsf gend='0a'x gaccess=gsasfile;
filename gsasfile 'exo03_2.eps';
```

Dans la fenêtre UNIX, visualiser le fichier obtenu en faisant la commande :
`gv exo03_2.eps` (*gv* signifie *ghostview*).
3. Essayer maintenant l’option `group` de la commande `hbar` en faisant :
`hbar COUL / group=SEXE`; Commenter les résultats obtenus.

Exercice 03.3

On va observer maintenant, à partir de différents exemples, l’effet de diverses options très utiles pour améliorer les graphiques haute résolution (avec `gchart` comme avec `gplot`).

Copier le fichier `~baccini/tpsas/exolim/exo03_3.sas` et l’exécuter, procédure par procédure, en commentant au fur et à mesure.

Remarques complémentaires

- En général, il est recommandé de faire la mise au point des graphiques (haute résolution) en utilisant la fenêtre graphique de SAS, puis de les archiver, une fois au point, dans un fichier au format postscript encapsulé.
- Tout programme SAS réalisant un graphique haute résolution (avec les procédures `gchart` ou `gplot`) doit s'achever par les commandes suivantes :
`quit;`
`goptions reset=all;`
`quit` termine la phase de réalisation d'un graphique; cette commande est également recommandée à l'issue d'un certain nombre de procédures statistiques complexes de SAS (`glm...`); `goptions reset=all` remet toutes les options graphiques à leurs valeurs par défaut.
- On notera les différents niveaux de la notion d'option dans SAS :
 - il existe des options générales, externes à toute procédure, qui se déclarent par les directives `option(s)` ou `goption(s)`
(exemples : `options linesize=78; goptions reset=all;`);
 - il existe des options de procédure, qui se déclarent dans la première ligne d'une procédure, juste après son nom (exemple : `proc corr vardef = n;`);
 - il existe des options de commande, qui se déclarent à la suite d'une commande; dans ce cas, la syntaxe n'est pas uniforme; on a déjà rencontré `by descending varX;`; on vient de voir `table SEXE*COUL/chisq;` (forme plus usuelle pour les options de commande).