EPIDATA Analysis (« classic ») 2.2 – Fiches mémo

→ Les commandes essentielles de base

Pour PC

Les conseils donnés ici concernent l'écriture dans la fenêtre des commandes.

F1 : Aide en ligne (fichier *.pdf) - F2 : liste des commandes possibles - F3 : Liste des variables

NB : `varqual` est une variable qualitative (sexe par exemple), `varquant` est une variable quantitative (âge par exemple). La barre / pour des commandes à suivre sont précédées d'un espace.

→ Pour lire un fichier

- `Read` ⌨ (touche « entrée ») puis choisir un fichier dans l'ordinateur (en *.rec)

→ Pour décrire

- `freq varqual` pour la fréquence des données codées (qualitatives) ; `freq *` trie toutes les variables qualitatives. Pour ajouter les pourcentages en colonne, faire (après un espace) `/c`, et pour avoir les pourcentages **cumulés** `/cum`, et pour avoir aussi les manquants `/m`
- `means varquant` pour les données quantitatives seules : donne, entre autres, le nombre d'observations, la moyenne (mean), la médiane, le minimum, le maximum, l'écart-type (Dev Stand)
- `describe (ou des) varquant1 varquant2 varquant3` pour décrire toutes les variables quantitatives. Voir F1 pour d'autres modalités.

→ Pour comparer

- `means varquant varqual` pour croiser une variable quantitative selon deux modalités qualitatives ou plus. A signaler que EPIDATA demande que la `varqual` soit de type numérique (c'est-à-dire # dans Epidata Entry !) → c'est pour cela que les variables doivent être plutôt de type #. `/t` pour avoir le test statistique (test de t de Student pour deux moyennes, ou ANOVA et test de F pour trois moyennes ou plus).
- `tables varqual1 (en colonne) varqual2 (en ligne)` pour croiser deux données qualitatives → `/r` pour les pourcentages en lignes (pour row, et suite à un espace), ou `/c` pour les pourcentages en colonnes (pour column), `/t` pour les tests statistiques (c'est le χ^2 de Pearson pour deux pourcentages ou plus), `/ex` pour le test de Fisher avec tableau 2x2 (pour deux pourcentages), `/rr` pour le risque relatif, `/o` pour l'odds ratio (attention à l'ordre des codages : ajouter `/sa` pour avoir les codages dans l'ordre croissant (le 2 après le 1)).

→ Pour sélectionner

- Un critère : `select var=1` ou, par exemple, `select age > 30`, ou `select sexe=1`
- Plusieurs critères : par exemple avec le « ou » ou le « and »
 - `select (var1=1) or (var2=1)`, ou `select (var1=1) and (var2=1)`
- Pour ôter la sélection : `select` ⌨

EPIDATA Analysis récent (2021) – Fiches mémo

→ Les commandes essentielles de base

Pour Mac et PC

Les conseils donnés ici concernent l'écriture dans la fenêtre des commandes.

F1 : Aide en ligne (fichier *.pdf) - F2 : liste des commandes possibles - F3 : Liste des variables

NB : `varqual` est une variable qualitative (sexe par exemple), `varquant` est une variable quantitative (âge par exemple). La barre / pour des commandes à suivre sont précédées d'un espace.

→ Pour lire un fichier

- `Read` ↵ puis choisir un fichier dans l'ordinateur (en *.rec, ou en *epx)

→ Pour décrire

- `freq varqual` pour la fréquence des données codées (qualitatives). Ajouter les pourcentages en colonne, faire (après un espace) `!pr`, et pour avoir les pourcentages cumulés `!cum`, et pour avoir aussi les manquants `!m`. Possibilité de cliquer sur *onglet Analysis et Frequency*
- `means varquant` pour les données quantitatives seules : donne, entre autres, le nombre d'observations, la moyenne (mean), la médiane, le minimum, le maximum, l'écart-type (Dev Stand)
- `describe (ou des) varquant1 varquant2 varquant3` pour décrire toutes les variables quantitatives. Voir F1 pour d'autres modalités.

→ Pour comparer

- `means varquant !by:=varqual` pour croiser une variable quantitative selon une variable qualitative codée en deux modalités ou plus. A signaler que EPIDATA demande que la `varqual` soit de type numérique. `!t` pour avoir le test statistique (test de t de Student pour deux moyennes, ou ANOVA et test de F pour trois moyennes ou plus).
- `tables varqual1 (en colonne) varqual2 (en ligne)` pour croiser deux données qualitatives → `!pr` pour les pourcentages en lignes (pour row, et suite à un espace), ou `!pc` pour les pourcentages en colonnes (pour column), `!t` pour les tests statistiques (c'est le χ^2 de Pearson pour deux pourcentages ou plus), `!ex` pour le test de Fisher avec tableau 2x2 (pour deux pourcentages), `!rr` pour le risque relatif, `!odds` pour l'odds ratio. Possibilité de cliquer sur *onglet Analysis et Tables*. Choisir les variables, voir *Options et Statistics*

→ Pour sélectionner

- Un critère pour éliminer un codage : `select (var=1) do drop data`. Si la variable est codée 1, 2 et 3, ne restera que les codages 2 et 3.
- Pour lever la sélection : ???

Retour pour PC

Autres commandes fréquentes

Pour voir toutes les commandes qui ont été effectuées

Les commandes réalisées au cours d'une « session » de travail sont sauvées dans un fichier *temp.pgm* (jusqu'à *temp2.pgm*). Pour le trouver, voir dans *C:\Programmes\Epidata\Testdata*, ou à l'endroit où on a décidé de les mettre à l'installation ; attention sous Vista et quelques Windows 7, le chercher dans *Fichier de compatibilité* dans le même endroit. Pour l'ouvrir, clic droit et *ouvrir avec...* Bloc Notes ou Word ou Open Office. On peut aussi le charger dans ANALYSIS avec **Editor** et ne prendre que les lignes intéressantes.

Pour visualiser les données comme dans un tableau

- Sans modifier les données : **browse** var1 var2 ou cliquer sur le bouton **Browse** ou faire F6. Les données peuvent être ordonnées, et copiées (clic droit sur une colonne)
- Pour modifier quelques données sans passer par la saisie : la commande **update** ou **update** var1 permet de modifier des valeurs. Les données sont modifiées dans le fichier d'origine si on fait **savedata** à la fin.

Changer, créer des variables avec if... then...

- Pour définir un seuil d'âge : **define** old30 _
 if age<30 **then** old30=1 **if** age>= 30 **then** old30=2
 ou bien **if** age<30 **then** old30=1 **else** old30=2
- Pour redéfinir le sexe d'une lettre en chiffres) : **define** sexe #
 if (var1 ="M") **then** sexe=1 **else** sexe=2
- Pour exclure un codage, utiliser les deux caractères "<" et ">" : tous les sexes codés « pas 1 » seront 2
 if sexe <> 1 **then** sexe =2
- Pour définir des codages : **labelvalue** sexe /1="Homme" /2="Femme"

Autres analyses

Voir à *Further analysis* dans la boîte des commandes

- **regress** varquant1 varquant2 : Régression linéaire de Y en X
- **correlate** varquant1 varquant2 : Coefficient de corrélation de Pearson
- **kwallis** varquant varqual : Test de Kruskal-Wallis pour comparer des moyennes entre elles (test non paramétrique au même objectif que le test t commandé *par means*).

Pour sauver le fichier modifié (attention enlever le select)

savedata : sauver au même endroit, et (préférable) donner un nouveau nom comme *fichier_2.rec*. Ne sauve que la sélection si on a fait *select* avant **+++**.