

Mise en œuvre sous le logiciel PSPP

Pour faire exécuter une commande sur des variables il y a deux possibilités : utiliser les menus déroulants de la fenêtre « Editeur de données » ou utiliser une suite de commandes dans la fenêtre « Editeur de syntaxe ».

Pour ouvrir une fenêtre « Editeur de syntaxe », une solution est d'utiliser le menu déroulant (de la fenêtre « Editeur de données ») « fichier » (puis « nouveau » et enfin « syntaxe »).

Pour plus de détails consulter la documentation du logiciel : PSPP Users' Guide.

Les exemples de ce document sont réalisés à partir du fichier texte **tp.txt** qu'il faut préalablement importer sous PSPP ou alors il faut utiliser directement le fichier **tp.sav** qui est au bon format. Ces 2 fichiers sont disponibles. Les variables utilisées dans ce document sont :

- Le sexe : 2 pour les femmes et 1 pour les hommes
- Le résultat au test du MMS (score croissant avec les performances : max = 30)
- VISIO : déficiences visuelles codée 0 non et 1 oui
- L'âge en année

1 Variable quantitative

1.1 Décrire

Il y a plusieurs solutions.

Une solution simple est d'utiliser la commande « descriptives », soit en ligne de commande, soit dans le menu déroulant (« Statistiques », puis « Statistiques Descriptives », puis « Descriptives »).

Pour décrire la variable « age » la ligne de commande est :

descriptives / variables = age

Si on souhaite plus de statistiques on peut compléter en précisant :

descriptives / variables = age / statistics = all

ou en faisant la liste de ce que l'on souhaite, exemple :

descriptives / variables = age / statistics = MEAN,STDDEV,VARIANCE,MINIMUM,MAXIMUM

On peut aussi utiliser la commande « frequencies », soit en ligne de commande, soit dans le menu déroulant (« Statistiques », puis « Statistiques Descriptives », puis « Fréquences ») mais en enlevant le tableau des fréquences.

1.2 Graphique

Pour les graphiques on pourra faire un histogramme ou une boîte à moustaches si on croise la variable quantitative avec une variable qualitative ou un nuage de points si on croise la variable quantitative avec une autre variable quantitative.

Pour faire un histogramme, par menu déroulant choisir « Graphiques » puis « Histogramme ». Sinon taper la ligne de commande :

```
graph / histogram = age
```

Pour une boîte à moustaches (ici l'âge en fonction du sexe) :

```
examine age by sexe / plot = boxplot
```

et pour un nuage de points, par menu déroulant choisir « Graphiques » puis « Nuage de points ». Sinon taper la ligne de commande (ici le MMS en fonction de l'âge) :

```
graph / scatterplot = age with mms
```

ou si on veut différencier les femmes des hommes :

```
graph / scatterplot = age with mms by sexe
```

1.3 Comparaison de distributions

1.3.1 Test de Student pour tester des moyennes

Pour faire un test de comparaison de moyennes de Student par menu déroulant choisir « Statistiques » puis « Comparaison des moyennes » puis « Test T pour Echantillons Indépendants ». IL faudra préciser les valeurs de la variable définissant les groupes. Sinon taper la ligne de commande :

```
T-test / variable = age / groups=sexe(1,2)
```

Ici on teste une différence de moyenne de l'âge en fonction du sexe (codé 1 ou 2).

Pour tester une moyenne observée par rapport à une moyenne théorique, par menu déroulant, choisir « Statistiques » puis « Comparaison des moyennes » puis « Test T pour Un-Echantillon ». Sinon taper la ligne de commande :

```
T-test / testval= 75 / variables = age
```

Ici on teste une moyenne de l'âge égale à 75.

1.3.2 Test de Mann-Whitney / Wilcoxon pour tester des distributions

Pour un test non paramétrique de Mann et Withney / Wilcoxon il faut utiliser les commandes suivantes :

`npar test / mann-whitney = age by sexe(1,2)`

Ici on teste une différence de distribution de l'âge en fonction du sexe (codé 1 ou 2).

Pour pouvoir le faire avec les menus déroulants il faut faire un test de Kruskal-Wallis qui est équivalent à un Mann et Withney dans le cas de 2 échantillons. Il faut choisir « Statistiques » puis « Statistiques non-paramétriques » puis « Echantillons K Indépendants ». Il faudra préciser les valeurs de la variable qui définissent les groupes.

1.4 Régression linéaire

Pour faire un modèle de régression linéaire, par menu déroulant choisir « Statistiques » puis « Régression » puis « Linéaire ». Sinon taper la ligne de commande :

`regression / variables=age / dependent=mms`

Dans ce cas on fait une régression linéaire du MMS en fonction de l'âge. Si on souhaite ajuster sur la variable sexe on fait :

`regression / variables=age sexe / dependent=mms`

1.5 Corrélation

Pour étudier la corrélation entre deux variables quantitatives (ici l'âge et le MMS), par menu déroulant choisir « Statistiques » puis « Corrélation bivariée ». Sinon taper la ligne de commande :

`correlation / variables=age mms`

2 Variable qualitative

2.1 Décrire

Pour décrire une variable qualitative avec les menus déroulants choisir « Statistiques », puis « Statistiques Descriptives », puis « Fréquences ». Il faudra penser à ne pas demander les statistiques (qui sont pour des variables quantitatives) mais demander le tableau des fréquences.

Pour décrire la variable « sexe » la ligne de commande est :

`frequencies / variables = sexe / statistics = none`

La commande « statistics=none » est là pour enlever les affichages de statistiques (moyenne, ...) qui n'ont pas lieu d'être calculées avec une variable qualitative.

2.2 Graphique

Pour faire un diagramme en barres, par menu déroulant choisir « Graphiques » puis « Diagramme en barres ». Sinon taper la ligne de commande :

`graph / bar = count by sexe`

ou si on souhaite des pourcentages

`graph / bar = pct by sexe`

et si on souhaite croiser les variables sexe et déficiences visuels (visio)

`graph / bar = pct by sexe by visio`

2.3 Comparaison de fréquences

2.3.1 Test de comparaison de fréquences observées

Pour faire un test du χ^2 d'indépendance grâce au menu déroulant choisir « Statistiques » puis « Statistiques Descriptives », puis « Tableaux croisés ». On vérifiera que l'on a précisé dans l'onglet « Statistiques » que la case Khi-deux est bien cochée. Sinon taper la ligne de commande :

`crosstab / tables = sexe by visio / statistics=chisq / cells = count expected`

La commande « expected » nous permet d'afficher dans le tableau croisé les effectifs attendus (théoriques) pour vérifier les conditions de validité du test. Pour un tableau 2x2 sont aussi fournis le test avec correction de Yates (appelé correction de continuité) et le test de Fisher exact. Mais ce n'est plus le cas pour des tableaux plus grands.

2.3.2 Test de comparaison de fréquences observées et théoriques

Pour un test du χ^2 d'ajustement il faut utiliser les commandes suivantes :

`npar test / chisquare=sexe / expected=0.4 0.6`

On teste ici par rapport à une population avec 40% d'hommes et 60% de femmes. Pour pouvoir le faire avec les menus déroulants il faut choisir « Statistiques » puis « Statistiques non-paramétriques » puis « Khi deux ».

1.4 Régression logistique

Pour faire un modèle de régression logistique, par menu déroulant choisir « Statistiques » puis « Régression » puis « Logistique binaire ». Sinon taper la ligne de commande :

`logistic regression visio with age sexe`

Dans ce cas on fait une régression logistique pour étudier la probabilité d'avoir une déficience visuelle en fonction de l'âge et ajustée sur le sexe.