

## Analyse de données quantitatives (LPOLS1221)

### Séance 8 : Comparaison de moyennes à l'aide du Test-t

#### 1. Principe général

Le test-t est un test inférentiel visant à comparer l'égalité de deux moyennes. Il existe plusieurs types de test-t en fonction de la nature des données. Dans le cadre de ce cours, l'on utilisera le test-t à échantillon unique et les test-t à échantillons indépendants.

Dans le cadre du test-t à échantillon unique, il s'agit de tester  $H_0$ , la moyenne de la variable dépendante est égale à une moyenne de référence. Cette moyenne de référence est extérieure au jeu de donnée (moyenne d'une autre enquête, issue de la littérature, etc.).

Dans le cas du test-t à échantillons indépendants, il s'agit de tester  $H_0$  qui prévoit une absence de différence significative entre deux moyennes issues de l'échantillon. Dans ce dernier cas, l'obtention des deux moyennes s'obtient via une variable qualitative.

#### 2. Conditions d'applications

La variable dépendante doit être continue et normalement distribuée (si  $n < 50$ ). En cas de test-t à échantillons indépendants, les deux variables doivent être indépendantes et les variances des deux groupes doivent être égales (condition d'homoscédasticité). Cette dernière condition est vérifiable via le test de Lévène qui a pour hypothèse  $H_0$ , les variances sont égales. La procédure d'SPSS propose en réalité deux tests : un test pour variances homogènes et un test non-paramétrique pour variances hétérogènes. Le test de Lévène vous permet ainsi de savoir quel test-t lire.

Test des échantillons indépendants										
Test de Levene sur l'égalité des variances			Test t pour égalité des moyennes							
		F	Sig.	t	ddl	Sig. (bilatéral)	Différence moyenne	Différence erreur standard	Intervalle de confiance de la différence à 95 %	
									Inférieur	Supérieur
how important in your life: work (Q1A)	Hypothèse de variances égales	2,346	,126	-2,445	1498	,015	-,097	,040	-,175	-,019
	Hypothèse de variances inégales			-2,450	1497,846	,014	-,097	,040	-,175	-,019

#### 3. Mesure d'association

Enfin, pour mesurer la force de l'association, on utilisera l'éta-carré qui mesure la part de variance de la variable dépendante expliquée par la variable indépendante. Celui-ci est calculé de la manière suivante :  $t^2 / (t^2 + N_1 + N_2 - 2)$  et s'interprète de la manière suivante : autour de 0.01 (effet de petite taille), autour de 0.06 (effet de taille moyenne) et autour voire, au-delà de 0.14 (effet de grande taille). Cette mesure n'est pas calculée par SPSS, il faut donc la calculer manuellement.

#### 4. Manipulations SPSS

Test-t à échantillon unique : Analyze → Compare Means → One Sample T-Test

- Insérez la variable dépendante dans l'encadré « Test Variables »
- Insérez la valeur à tester dans l'encadré « Value Test »

Test-t à échantillons indépendants : Analyze → Compare Means → Independent Samples T-test

- Insérez la variable dépendante dans l'encadré « Test Variables »
- Insérez la variable indépendante dans l'encadré « grouping variable ». Spécifiez les modalités des deux groupes ou utilisez l'option « cutpoint » en référant une valeur. Cette option va créer deux groupes : un groupe d'observations dont les valeurs sont strictement inférieures à la valeur renseignée et un groupe d'observations dont les valeurs sont égales ou supérieures à la valeur renseignée.

## **Exercices**

1. Quelle est la relation entre le niveau d'éducation du père (**V355\_r**) et le revenu annuel du ménage (**V353YR**) des répondants belges ? Comparez les moyennes pour les individus dont le père a un niveau d'éducation moyen ou élevé.
2. Refaites le même exercice mais comparez l'effet pour les individus dont le père a un niveau éducation faible ou moyen et ceux dont le père a un niveau d'éducation élevé.
3. En Belgique, en 1986, il y avait en moyenne 2,01 enfants par foyer et 1,98 en 1995. Réalisez deux test-t consécutifs avec la base de données de 2008 (**V324b**) et ces données. Quelles conclusions pouvons-nous en tirer ?
4. Existe-t-il une relation de dépendance entre le fait d'avoir un niveau d'éducation ou non (**V336\_r**) et le nombre d'institutions politiques et civiles envers lesquelles les citoyens accordent une grande confiance (**Variables V205 à V222**)?
5. Créez une variable à partir de la variable **V66**. Cette variable que vous nommerez confiance, distinguera les individus insatisfaits (valeurs 1 à 5) des individus satisfaits (6 à 10) de leur vie. Regardez ensuite à partir de cette variable, comment la satisfaction des répondants varie en fonction qu'ils vivent en Belgique, Hongrie ou au Kosovo.
6. Comparez les moyennes de revenu (**V353M\_PPP**) entre les luxembourgeois et les kosovars. Commentez.