

## Cours 10

Test de comparaison de moyennes

---

---

---

---

---

---

---

---

## Plan de la rencontre

- ⦿ Le test de moyenne
  - > Utilité et postulats d'utilisation
  - > Les tests paramétriques
  - > Les tests non-paramétriques

☰ Test de moyenne avec SPSS

2

---

---

---

---

---

---

---

---

## Retour sur les analyses bivariées

- ⦿ Le type de test utilisé dépend du type de variables utilisées.
- ⦿ Trois possibilités avec deux familles de variables :

Une variable de type ...		En lien avec une variable...	On utilise :
QUALITATIVE	en lien avec	QUALITATIVE =	Tableau croisé
QUALITATIVE	en lien avec	QUANTITATIVE = (ou ordinale)	Test de moyennes
QUANTITATIVE	en lien avec	QUANTITATIVE = (ou ordinale)	Corrélation

---

---

---

---

---

---

---

---



## Le critère paramétrique

### Les postulats d'utilisation

- La distribution de la variable quantitative est normale. ( vérifiez avec l'histogramme)
- Le nombre de sujets est supérieur à 30 (au total, pas par groupe)
- Homogénéité des variances (comparables) (*Levene non significatif*)
- Échantillon aléatoire indépendant (inutile dans le cas de vos TP)

---

---

---

---

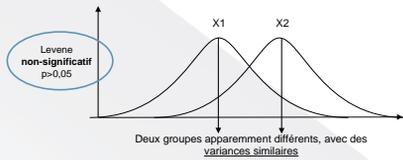
---

---

---

---

### L'homogénéité de la variance et le test de Levene



Tout va bien !

---

---

---

---

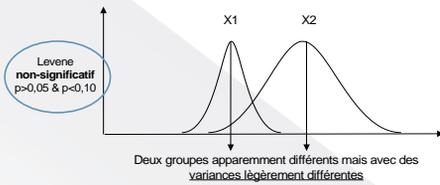
---

---

---

---

### L'homogénéité de la variance et le test de Levene




---

---

---

---

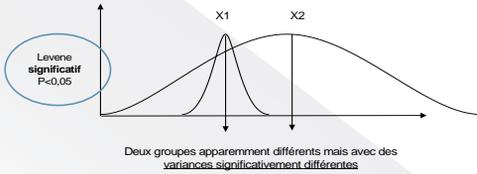
---

---

---

---

### L'homogénéité de la variance et le test de Levene



Ici, le test de Levene est **significatif**, donc les variances ne sont pas comparables. On brise un postulat donc on doit faire un test non-paramétrique.

---

---

---

---

---

---

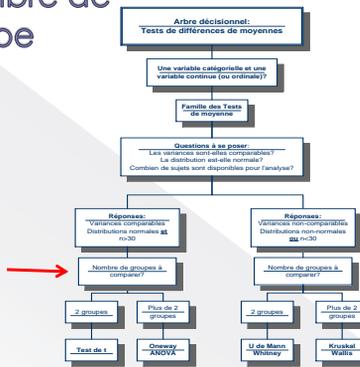
---

---

---

---

### Le nombre de groupe




---

---

---

---

---

---

---

---

---

---

### Le Test de T

Comparer les moyennes de deux groupes

---

---

---

---

---

---

---

---

---

---

## Le test de moyenne

Test T pour des échantillons indépendants

### Objectif:

- Le test de moyenne sert à évaluer s'il existe une différence statistiquement significative entre deux moyennes
- Concrètement, est-ce que deux groupes (ex : hommes vs. Femmes) ont des moyennes différentes à une échelle de troubles du comportement ?

13

---

---

---

---

---

---

---

---

## Trois questions incontournables

- Existe-t-il un lien entre mes variables?  
La question de la **Signification**
- Ce lien est-il important?  
La question de la **Force**
- Que veut dire ce lien en termes concrets? La question du **Sens** (ou de la nature) de la relation

---

---

---

---

---

---

---

---

## La signification

- Le test de T sert à répondre à la question suivante : Existe-t-il un lien entre une variable dichotomique et une variable quantitative? Il sert donc à tester des hypothèses.
- Il est utile pour savoir si oui ou non il existe un lien entre deux variables : il nous indique la signification de la relation.
- Lorsque c'est significatif à  $p < 0.05$ , nous rejetons  $H_0$ . Lorsque c'est supérieur, nous ne pouvons rejeter l'hypothèse nulle.

---

---

---

---

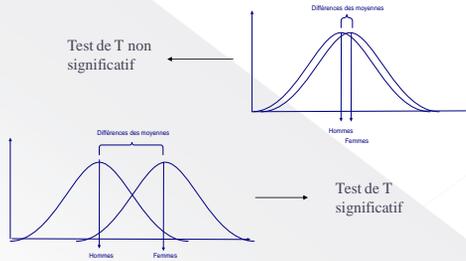
---

---

---

---

## La représentation graphique de ce qu'est en théorie un test de T




---

---

---

---

---

---

---

---

## La force de la relation

- ⦿ Le test de T donne la signification mais non la force de la relation...
- ⦿ Le  $\eta^2$  est le coefficient d'association (ou de force) des tests de moyennes paramétriques.

---

---

---

---

---

---

---

---

## Le coefficient de force des tests de moyenne paramétriques: le $\eta^2$

- ⦿ Il mesure la force de l'association entre la variable qualitative et la variable quantitative
- ⦿ Il varie entre 0 et 1
- ⦿ Il donne une indication de la proportion de la variance expliquée
- ⦿ On l'obtient en utilisant la commande « comparer des moyennes » (voir vignettes SPSS)

---

---

---

---

---

---

---

---

## Le sens de la relation

- ⦿ Pour les tests de moyennes (comme pour les tableaux de contingence) on ne parle pas de relation positive ou négative.
- ⦿ Le sens de la relation analysé à l'aide des moyennes de vos groupes.
- ⦿ Par exemple : les garçons commettent plus de violence physique que les filles.

---

---

---

---

---

---

---

---

---

---

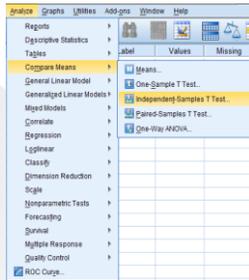
---

---

## Le test de T dans SPSS

- En Français :
- 1- Analyse
  - 2- Comparer les moyennes
  - 3- Test T pour échantillons indépendants

Les individus qui croient que les femmes victimes de viol le méritaient ont-ils une moyenne différente à une échelle de sadisme que ceux qui n'ont pas une telle croyance ?




---

---

---

---

---

---

---

---

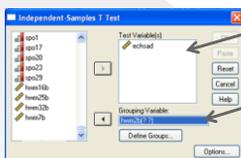
---

---

---

---

## Je sélectionne les variables...



- Variable quantitative (dépendante)
- Variable(s) à tester
- Variable dichotomique
- Critère de regroupement...

---

---

---

---

---

---

---

---

---

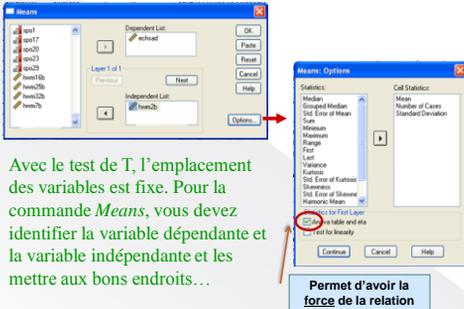
---

---

---



### On demande le Eta<sup>2</sup> dans les Options



Avec le test de T, l'emplacement des variables est fixe. Pour la commande Means, vous devez identifier la variable dépendante et la variable indépendante et les mettre aux bons endroits...

---

---

---

---

---

---

---

---

---

---

### Le Eta<sup>2</sup> est dans le dernier tableau

Means

ANOVA Table						
		Sums of Squares	df	Mean Square	F	Sig.
ehead * fwm2b	Between Groups (Combined)	6,629	1	6,629	14,218	,000
women get raped	Within Groups	123,943	216	,466		
probably deserve it, 0-4	Total	130,571	217			

Measures of Association		
	Eta	Eta Squared
ehead * fwm2b	<del>X</del>	.050
women get raped		
probably deserve it, 0-4		

La force

N.B. On doit prendre le Eta Squared (Eta<sup>2</sup>) et pas le Eta

---

---

---

---

---

---

---

---

---

---

### Présentation des résultats

Tableau X. Comparaisons de moyenne à l'échelle de sadisme des prisonniers selon l'opinion quant au sort des victimes de viol.

Les victimes de viol méritent ce qui leur arrive

	Moyennes		T	n
	Non	Oui		
Score à l'échelle de sadisme	0,36	0,74	-3,77 ***	273

\* = p ≤ 0,05      \*\* = p ≤ 0,01      \*\*\* = p ≤ 0,001

---

---

---

---

---

---

---

---

---

---

## Interprétation des résultats

- ⊙ Les individus qui croient que les femmes victimes de viol méritent ce qui leur arrive ont une moyenne de 0,74 à l'échelle de sadisme, alors que les individus qui n'ont pas cette croyance ont une moyenne de 0,36. Ces moyennes sont significativement différentes ( $p \leq 0.001$ ). La force de cette relation est très faible comme en témoigne un  $\eta^2$  de 0,050.

---

---

---

---

---

---

---

---

## Récapitulons...

1. S'assurer que l'on respecte les postulats.
  - > Critères paramétriques (normalité,  $N > 30$ , égalité variances)
  - > Nombre de groupes (2)
2. Si la relation n'est pas significative, je ne traite pas des deux autres étapes.
3. Analyser le sens de la relation. Quel groupe a la moyenne la plus élevée ?
4. Analyser la force de la relation ( $\eta^2$ ).

---

---

---

---

---

---

---

---

## Le Oneway ANOVA

Comparer les moyennes de trois groupes et plus

---

---

---

---

---

---

---

---

## Le Oneway ANOVA (test F)

- Aussi appelé analyse de variance simple ou à un facteur
- Il vise à comparer les moyennes non pas deux groupes mais bien trois groupes ou plus
- Concrètement, est-ce que trois groupes (ex : mineurs, jeunes adultes, aînés) ont des moyennes différentes à une échelle de troubles du comportement ?

---

---

---

---

---

---

---

---

## Trois questions incontournables

1. Existe-t-il un lien entre mes variables?  
La question de la **Signification**
2. Ce lien est-il important?  
La question de la **Force**
3. Que veut dire ce lien en termes concrets?  
La question du **Sens** (ou de la nature) de la relation

---

---

---

---

---

---

---

---

## La signification

- Le Ratio F (plutôt que le t en T de Student) teste l'hypothèse d'une différence de moyennes entre les groupes.
- Il est utile pour savoir si oui ou non il existe un lien entre deux variables : il nous indique la signification de la relation.
- Lorsque c'est significatif à  $p < 0.05$ , nous rejetons  $H_0$ . Lorsque c'est supérieur, nous ne pouvons rejeter l'hypothèse nulle.

---

---

---

---

---

---

---

---

## La signification et les tests à posteriori (post hoc)

- Lorsque le test F est significatif, il nous indique que la moyenne d'au moins 1 groupe se distingue significativement des autres. Mais il ne permet pas d'établir si toutes les moyennes sont différentes.
- Nous devons utiliser une autre série d'analyses afin de connaître plus en détail quels groupes sont significativement différents desquels.
- Il faut alors réaliser des tests post hoc.

---

---

---

---

---

---

---

---

---

---

## Les tests Post hoc: pour mieux comprendre les différences entre les groupes

- Les différents tests post hoc visent à comparer les différents sous-groupes entre eux afin de connaître où se trouvent les différences parmi les groupes.
- Il existe plusieurs tests post hoc. Pour l'illustration, nous nous attarderons principalement à deux d'entre eux : le Scheffe et le Tukey.

---

---

---

---

---

---

---

---

---

---

## Le Scheffe et le Tukey

- **Le Scheffe** : Le Scheffe est un test commun utilisé dans le cas d'échantillons (i.e. la taille des groupes à comparer) de différentes tailles. Il pourrait cependant aussi être utilisé dans le cas d'échantillon de taille semblable.
- **Le Tukey** : Le Tukey est similaire au Scheffe, mais ne doit être utilisé que dans les cas où les échantillons (i.e. la taille des groupes à comparer) sont de taille semblable.

---

---

---

---

---

---

---

---

---

---

## La force et le sens de la relation

- ⊙ La force et le sens de la relation pour le Oneway Anova sont déterminés de manière similaire au test de T :
  - > Force : Regarder le Eta<sup>2</sup>
  - > Sens : Regarder les moyennes des groupes

---

---

---

---

---

---

---

---

---

---

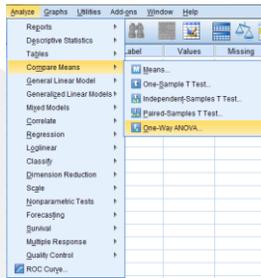
---

---

## Le Oneway ANOVA dans SPSS

Existe-t-il un lien entre niveau de précipitation en pluie (3 catégories : pas de pluie, faible pluie, forte pluie) et le nombre d'accidents avec blessés ?

- En français :
- 1- Analyse
  - 2- Comparer les moyennes
  - 3- ANOVA à 1 facteur




---

---

---

---

---

---

---

---

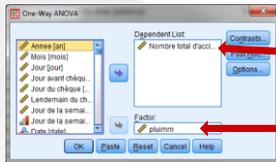
---

---

---

---

## Je sélectionne les variables...



- Variable quantitative  
\* Liste variables dépendantes
- Variable qualitative (3 groupes ou +)  
\* Critère

---

---

---

---

---

---

---

---

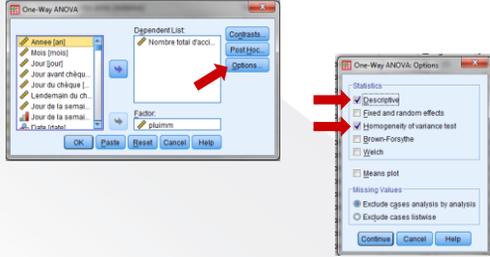
---

---

---

---

## Je demande les statistiques descriptives et le test de Levene




---

---

---

---

---

---

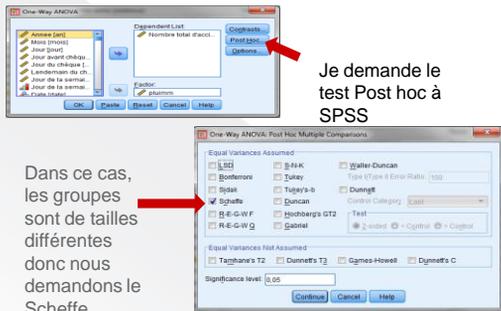
---

---

---

---

## La comparaison a posteriori des groupes



Je demande le test Post hoc à SPSS

Dans ce cas, les groupes sont de tailles différentes donc nous demandons le Scheffe

---

---

---

---

---

---

---

---

---

---

## Les résultats du test F (output)

La première étape est de vérifier si les variances de nos groupes sont comparables à l'aide des résultats du test de Levene.

Test of Homogeneity of Variances

Nombre total d'accidents avec blessés

Levene Statistic	df1	df2	Sig.
3,528	2	1442	,050

Il importe de se rappeler que le test de Levene ne doit **PAS** être significatif... Dans notre exemple, ce postulat est respecté.

---

---

---

---

---

---

---

---

---

---





## Présentation des résultats

Tableau X. Titre du tableau

	Précipitations en pluie			F	n
	Moyennes				
	Aucune	Faible	Forte		
Accidents avec blessés	72,38	74,08	89,05	28,30 ***	1445

\* =  $p \leq 0,05$     \*\* =  $p \leq 0,01$     \*\*\* =  $p \leq 0,001$

Afin de simplifier la présentation des résultats, on suggère, dans le cadre du cours, de faire mention dans le texte des résultats obtenus à l'aide des tests post-hoc.

---

---

---

---

---

---

---

---

---

---

---

---

## Interprétation des résultats

- Il existe une relation statistiquement significative entre les précipitations en pluie et le nombre d'accidents avec blessés ( $F = 28,30$ ;  $p \leq 0,001$ ). Le test de Scheffe montre qu'il y a significativement plus d'accidents avec blessés les jours de fortes pluies (89,05 accidents) que les jours sans pluie (72,38 accidents) ou de faibles pluies (74,08 accidents). Bien que la moyenne des jours de faibles pluies soit plus élevée que celle des jours sans pluie, cette différence n'est pas significative. La force de cette relation est toutefois faible ( $\text{Eta}^2 = 0,04$ ).

---

---

---

---

---

---

---

---

---

---

---

---

## Récapitulons...

- S'assurer que l'on respecte les postulats.
  - Critères paramétriques (normalité,  $N > 30$ , égalité variances)
  - Nombre de groupes (3 ou plus)
- Regarder mon seuil de tolérance (test F). Si la relation n'est pas significative, je ne traite pas des autres étapes.
- Identifier quelles moyennes sont différentes à l'aide des tests Post hoc
- Analyser le sens de la relation. Quel groupe a la moyenne la plus élevée ?
- Analyser la force de la relation ( $\text{Eta}^2$ ).

---

---

---

---

---

---

---

---

---

---

---

---

Pour savoir comment faire avec SPSS,  
consulter le site de Jean-Pierre Guay :

<http://www.mapageweb.umontreal.ca/quayjea/>

52

---

---

---

---

---

---

---

---

## L'UTILITÉ ET LA LOGIQUE DES TESTS NON- PARAMÉTRIQUES

---

---

---

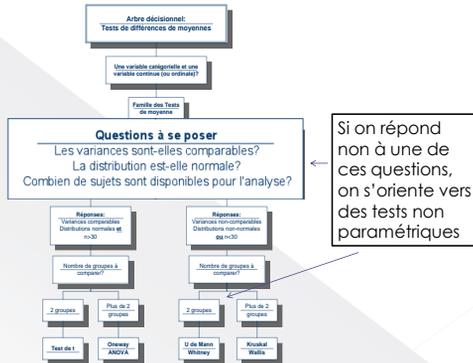
---

---

---

---

---




---

---

---

---

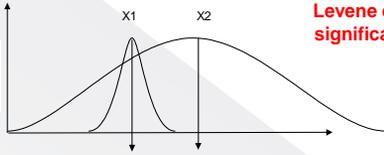
---

---

---

---

## DES VARIANCES NON COMPARABLES, C'EST QUOI DÉJÀ ?



Le test de Levene est significatif

Deux groupes apparemment différents mais avec des variances significativement différentes.

---

---

---

---

---

---

---

---

## Les tests non-paramétriques

### U de Mann-Whitney

- ⦿ Détermine si les médianes de 2 groupes sont différentes.

### Kruskal Wallace

- ⦿ Détermine si les médianes de 3 ou + groupes sont différentes.

56

---

---

---

---

---

---

---

---

## ET LES TROIS GRANDES QUESTIONS ...

- Existe-t-il un lien ?
  - On regarde la signification du U de Mann Whitney ou celle du Kruskal-Wallis.
- Si oui, quel est la force de ce lien ?
  - Il n'y a pas de coefficient de force avec les tests non paramétriques (chanceux).
- Quel est le sens ? (ou qu'est-ce que ça veut dire en langage clair)
  - On regarde les médianes.

---

---

---

---

---

---

---

---

## LES TESTS NON-PARAMÉTRIQUES DANS SPSS

---

---

---

---

---

---

---

---

## LES TESTS NON-PARAMÉTRIQUES DANS SPSS

On veut savoir s'il existe une différence significative entre les types d'agresseurs sexuels (variable dichotomique: pédophiles vs violeurs) sur une échelle d'anxiété hétérosexuelle (variable continue).

2 groupes à comparer donc j'utilise le U de Mann-Whitney

---

---

---

---

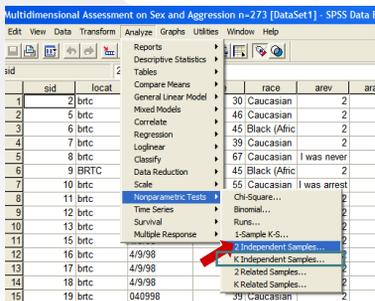
---

---

---

---

## LES TESTS NON-PARAMÉTRIQUES DANS SPSS



On clique *K-Independent samples* si on a plus de deux groupes à comparer avec le Kruskal-Wallis

Tests non paramétriques  
2 échantillons indépendants (U de Mann Whitney)

K échantillons indépendants (Kruskal Wallace)

---

---

---

---

---

---

---

---





## COMMENT ON FAIT DÉJÀ POUR ALLER CHERCHER LES MÉDIANES ?

The screenshot shows the SPSS 'Frequencies' dialog box. The 'Statistics' sub-dialog is open, and the 'Median' checkbox is checked for both the 'Voleurs' and 'Pédophiles' groups. The 'Anxiété hétérosexuelle' variable is selected in the main dialog.

Il existe une différence significative entre les pédophiles et les voleurs sur l'échelle d'anxiété hétérosexuelle ( $Z=-4,53$ ,  $p<0,001$ ). En effet, les pédophiles ont une médiane plus élevée (médiane=1,9) que les voleurs (médiane=1,1).

---

---

---

---

---

---

---

---

---

---

## Et le Kruskal-Wallis ?

- ⊙ Lorsque le Kruskal-Wallis est significatif, ça indique que la médiane d'au moins 1 groupe est différente des autres.
- ⊙ Il n'existe pas de Post Hoc en non paramétrique. Pour voir s'il y a des différences entre nos groupes, il faut faire plusieurs U de Mann Whitney.

---

---

---

---

---

---

---

---

---

---

## Présentation des résultats

Tableau X. Titre du tableau

Score sur échelle de consommation	Médianes (rang moyen)		Z	n
	Oui	Non		
	5(168,1)	2(120,6)		

\* =  $p \leq 0,05$       \*\* =  $p \leq 0,01$       \*\*\* =  $p \leq 0,001$

---

---

---

---

---

---

---

---

---

---

Pour savoir comment faire avec SPSS,  
consulter le site de Jean-Pierre Guay :

<http://www.mapageweb.umontreal.ca/quayjea/>

70

---

---

---

---

---

---

---

---