UNIVERSITÉ DE NANTES

Tests de Student

Comparaison de 2 groupes

Module HPS3-32
Année 2021-22

GALHARRET J-M,
Laboratoire de Mathématiques Jean Leray
Faculté de Psychologie.

Introduction

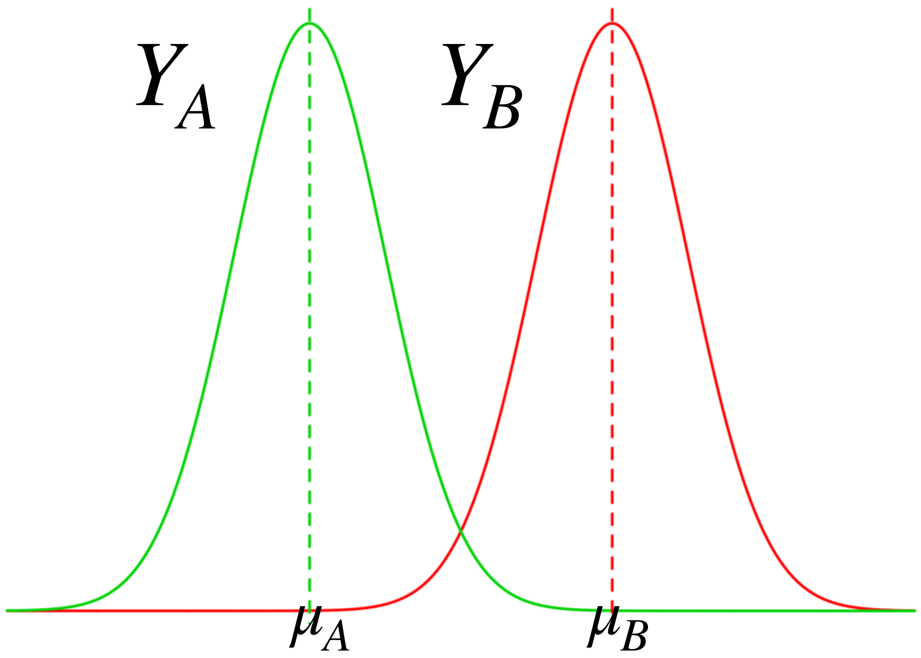
Cadre d'application

- On va étudier deux situations :
  - Les groupes sont indépendants : les deux séries de données correspondent à des individus différents (ex : jeune versus âgé).
  - Les groupes sont appariés : les deux séries de données correspondent aux mêmes individus (ex : résultat au test avant et après un traitement).

Comparaison d'une variable quantitative entre 2 groupes indépendants

Généralités

Conditions d'application du t test



- On considère une variable quantitative Y définie sur deux populations notées A et B.
- On suppose que la distribution de Y sur A et B est connue et que:

$$Y_A \sim \mathcal{N}(\mu_A, \sigma^2)$$
$$Y_B \sim \mathcal{N}(\mu_B, \sigma^2)$$

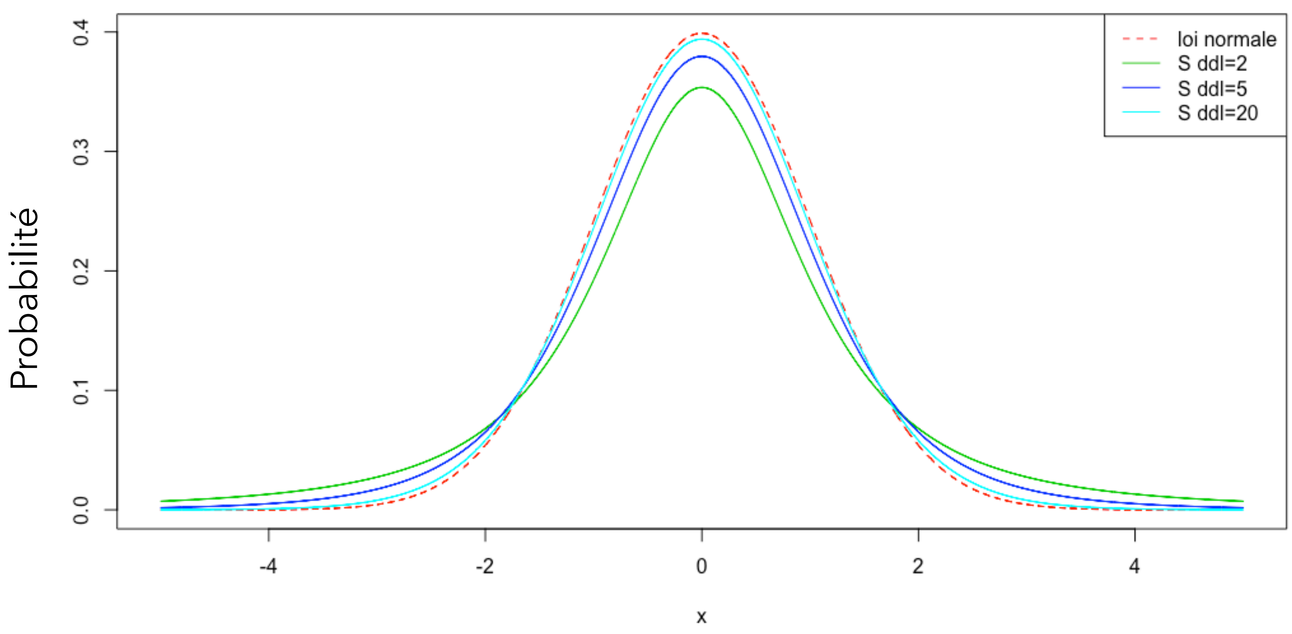
Procédure de test

t test de Student sur groupes indépendants

- On teste l'hypothèse : $H_0 : \mu_A = \mu_B$ versus $H_1 : \mu_A \neq \mu_B$
- On considère deux échantillons extraits respectivement de A et de B, ils seront de tailles respectives n_A, n_B. On calcule les moyennes et écarts types corrigés notés respectivement $\bar{y}_A, s_A$ et $\bar{y}_B, s_B$.
- On calcule $t = \dfrac{\bar{y}_A - \bar{y}_B}{s\sqrt{\frac{1}{n_A} + \frac{1}{n_B}}}$ où $s = \sqrt{\dfrac{(n_A - 1)s_A^2 + (n_B - 1)s_B^2}{n_A + n_B - 2}}$
- Sous Ho, on sait que T (la variable correspondant à t) suit une loi de Student à $n_A + n_B - 2$ degrés de liberté (ddl).
- La p-value du test est donnée par $p = \mathbb{P}_{H_0}(|T| > |t|)$ où $\mathbb{P}_{H_0}$ est la probabilité sachant que Ho est vraie.

Loi de Student

Généralités



- La loi de Student est symétrique de moyenne 0.
- Plus ddl augmente, plus elle se rapproche de la distribution normale

Lecture de la Table de Student

Table pour le test de Student

	p_value (test bilatéral)						
ddl	**0,2**	**0,1**	**0,05**	**0,02**	**0,01**	**0,002**	**0,001**
1	3,078	6,314	12,706	31,821	63,657	318,309	636,619
2	1,886	2,920	4,303	6,965	9,925	22,327	31,599
3	1,638	2,353	3,182	4,541	5,841	10,215	12,924
4	1,533	2,132	2,776	3,747	4,604	7,173	8,610
5	1,476	2,015	2,571	3,365	4,032	5,893	6,869
6	1,440	1,943	2,447	3,143	3,707	5,208	5,959
7	1,415	1,895	2,365	2,998	3,499	4,785	5,408
8	1,397	1,860	2,306	2,896	3,355	4,501	5,041
9	1,383	1,833	2,262	2,821	3,250	4,297	4,781
10	1,372	1,812	2,228	2,764	3,169	4,144	4,587
11	1,363	1,796	2,201	2,718	3,106	4,025	4,437
12	1,356	1,782	2,179	2,681	3,055	3,930	4,318
13	1,350	1,771	2,160	2,650	3,012	3,852	4,221
14	1,345	1,761	2,145	2,624	2,977	3,787	4,140
15	1,341	1,753	2,131	2,602	2,947	3,733	4,073
16	1,337	1,746	2,120	2,583	2,921	3,686	4,015
17	1,333	1,740	2,110	2,567	2,898	3,646	3,965
18	1,330	1,734	2,101	2,552	2,878	3,610	3,922
19	1,328	1,729	2,093	2,539	2,861	3,579	3,883
20	1,325	1,725	2,086	2,528	2,845	3,552	3,850

Pour ddl=14 on a par exemple :

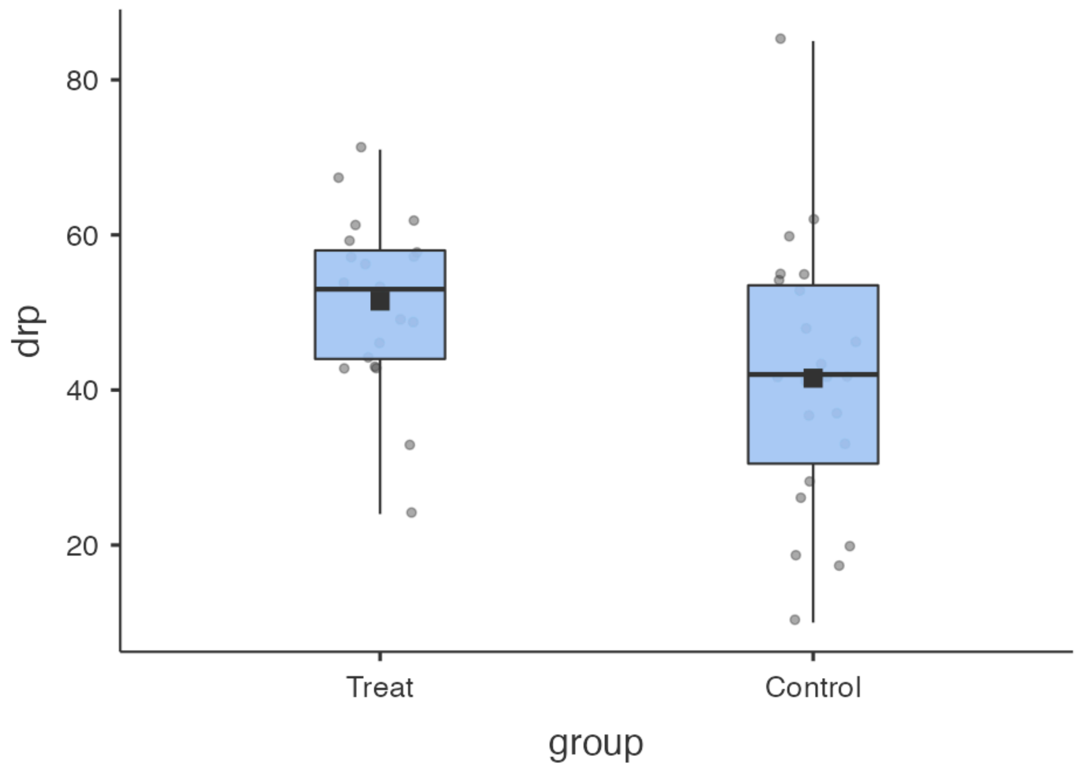
$$\mathbb{P}_{H_0}(|T| > 1.345) = .20$$
$$\mathbb{P}_{H_0}(|T| > 2.264) = .02$$

Donc si pour ce ddl=14 :

- on a obtenu $t = 2$ alors on sait que la p-value est comprise entre .05 et .10 (donc non significative).
- on a obtenu $t = -10$ alors on sait que p-value est inférieure à .001 (donc significative).

Retour sur l'exemple

	group	drp
N	Treat	21
	Control	23
Missing	Treat	0
	Control	0
Mean	Treat	51.5
	Control	41.5
Median	Treat	53.0
	Control	42.0
Standard deviation	Treat	11.0
	Control	17.1
Minimum	Treat	24.0
	Control	10.0
Maximum	Treat	71.0
	Control	85.0



On peut exprimer Ho de plusieurs façons (non exhaustives) :

- La différence de résultats de lecture entre les deux groupes n'est due qu'au hasard.
- Le fait d'avoir effectué des activités n'a pas permis d'améliorer significativement les performances des enfants.
- L'entrainement en lecture n'a pas d'impact significatif sur les résultats des enfants.
- Il n'y a pas de différence significative du niveau moyen de lecture entre les enfants du groupe contrôle et du groupe ayant effectué des activités de lecture.

Calcul de l'écart type commun (s)

- Dans la procédure du test on doit calculer $s = \sqrt{\dfrac{(n_A - 1)s_A^2 + (n_B - 1)s_B^2}{n_A + n_B - 2}}$ puisque l'on a supposé que les deux populations ont le même écart type.
- Il s'agit (pour la variance) d'un calcul de moyenne pondérée : c'est la moyenne des variances corrigées de chaque échantillon pondérées par leur taille -1.
- Lorsque les échantillons ont la même taille on a $s = \sqrt{\dfrac{s_A^2 + s_B^2}{2}}$
- Pour vérifier que l'on ne s'est pas trompé dans le calcul on peut utiliser que s doit toujours être compris entre les deux écarts types corrigés s_A, s_B.

	group	drp
N	Treat	21
	Control	23
Missing	Treat	0
	Control	0
Mean	Treat	51.5
	Control	41.5
Median	Treat	53.0
	Control	42.0
Standard deviation	Treat	11.0
	Control	17.1

$$s = \sqrt{\frac{20 \times 11^2 + 22 \times 17.1^2}{42}} = 14.55$$

$$t = \frac{51.5 - 41.5}{14.55 \times \sqrt{\frac{1}{21} + \frac{1}{23}}} = 2.28$$

$$ddl = 21 + 23 - 2 = 42$$

$$t = \frac{\bar{y}_A - \bar{y}_B}{s\sqrt{\frac{1}{n_A} + \frac{1}{n_B}}} = 2.28 \quad \text{on a } ddl = n_A + n_B - 2 = 42.$$

	p_value (test bilatéral)						
ddl	**0,2**	**0,1**	**0,05**	**0,02**	**0,01**	**0,002**	**0,001**
40	1,303	1,684	2,021	2,423	2,704	3,307	3,551

La p-value calculée par JAMOVI vaut p=.029.

Au risque de 5%, on peut penser que la différence entre les deux groupes est significative.

Taille d'effet

Quantifier l'écart des populations.

- La significativité permet de savoir si la différence est due au hasard ou si elle est attribuable au fait d'une vraie différence entre les deux groupes.
- Elle ne permet pas de <u>**quantifier**</u> l'écart entre les groupes. Pour ce faire on définit la taille d'effet.
- Pour le test de Student il s'agit du δ de cohen défini par : $\delta = \dfrac{\bar{y}_A - \bar{y}_B}{s}$, c'est à dire que c'est la différence des moyennes rapporté à la dispersion de la population.
- Il existe des standards (Cohen, 1988) :

$\lvert\delta\rvert$	[0.2,0.5[	[0.5,0.8[	[0.8,2[	Plus de 2
Effet	Faible	Modéré	Fort	Très fort

Rédaction de la conclusion aux normes APA

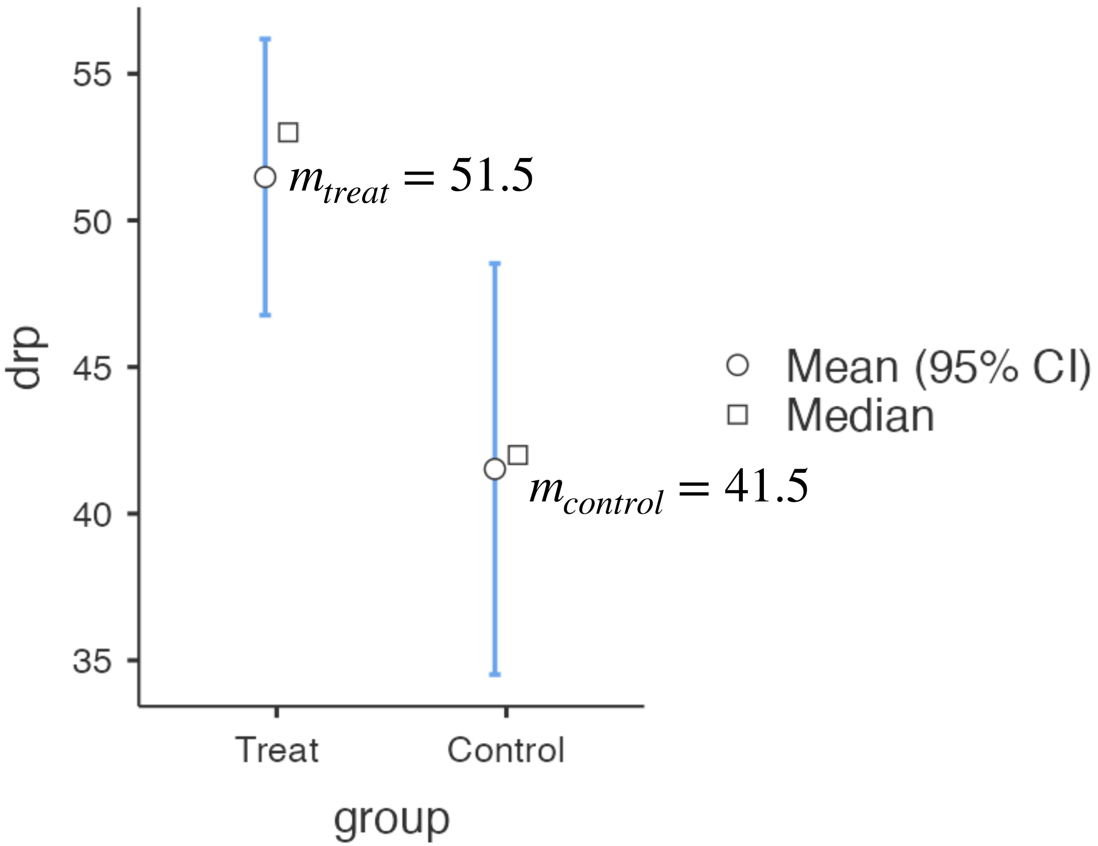
Test de Student sur groupes indépendants

On reprend l'ensemble des résultats précédents :

Un test de Student sur groupes indépendants a permis de mettre en évidence une différence moyenne et significative en terme de résultats en lecture en faveur du groupe des enfants ayant effectués un entrainement ($M = 41.5$, $SD = 17.15$) par rapport aux enfants du groupe contrôle ($M = 51.5$, $SD = 11.00$), $t(42) = 2.28$, $p = .029$, $\delta = 0.68$.

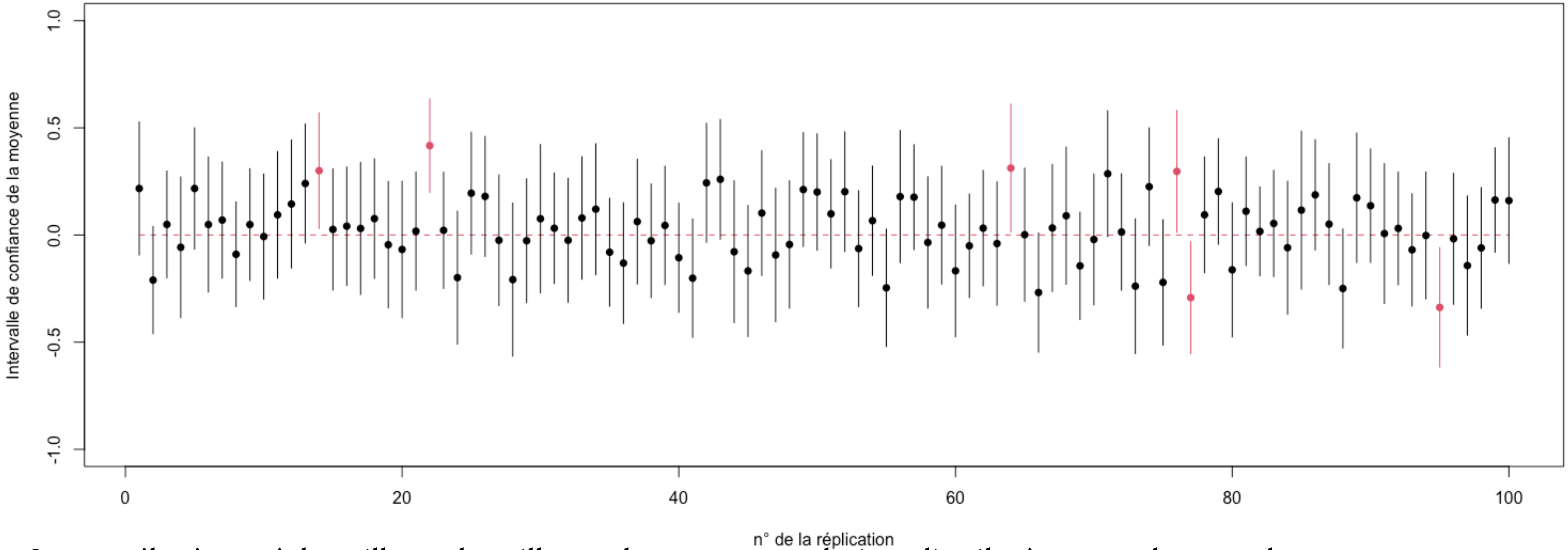
Intervalles de confiance

95%-CI



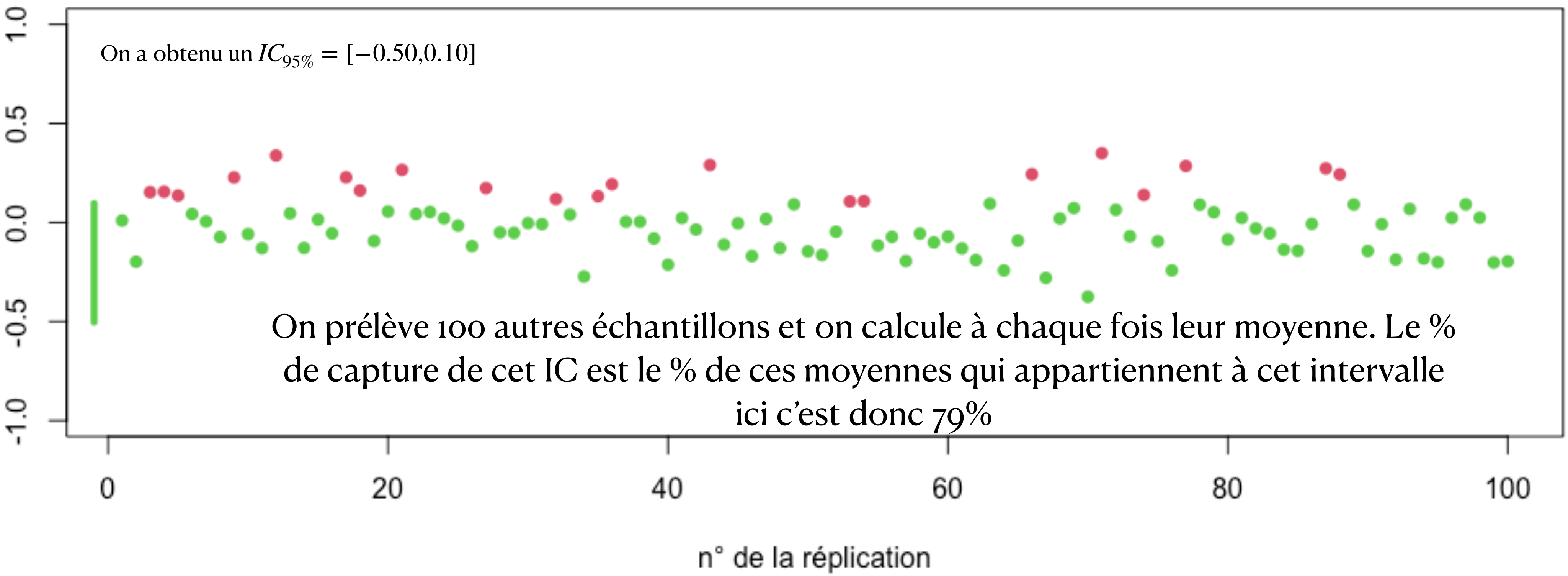
La notion d'intervalle de confiance est également difficile à appréhender :

- Ici on considère un intervalle de confiance de la moyenne (cela peut concerner un autre paramètre tel que l'écart type la médiane ...)
- L'intervalle de confiance de la moyenne est centré sur la moyenne de l'échantillon.
- Signification : Si on répète l'expérience un très grand nombre de fois alors 95 % des IC que l'on obtiendra contiendront la valeur du vrai paramètre.



On a prélevé 100 échantillons de taille 50 dans une population distribuée normalement de moyenne 0 et d'écart type 1 (on parle de loi normale centrée réduite). Pour chaque réplication on calcule l'IC de niveau 95% correspondant. On voit que :
Parmi les 100 IC seuls 6 ne contiennent pas 0. (On est bien proche du niveau de 95% si on avait eu plus de réplication alors on aurait exactement 95%)

Pourcentage de capture d'un IC



On a obtenu un $IC_{95\%} = [-0.50, 0.10]$

On prélève 100 autres échantillons et on calcule à chaque fois leur moyenne. Le % de capture de cet IC est le % de ces moyennes qui appartiennent à cet intervalle ici c'est donc 79%

Lien entre la p-value et l'IC de la différence

On veut comparer un variable quantitative sur deux groupes A et B. On calcule la différence des moyennes des deux échantillons et l'IC à 95% (noté $\mathscr{I}$) de cette différence de moyennes. On calcule également la p-value du test de Student entre ces deux groupes. Alors on a la propriété suivante :

$$p < .05 \text{ si et seulement si } 0 \text{ n'appartient pas à } \mathscr{I}$$

Retour sur l'exemple :

Mean différence	SE Différence	Lower 95%-CI	Upper 95%-CI
9.95	4.39	1.09	18.8

Significativité de l'écart entre les 2 groupes
(p-value)

Quantification de l'écart entre les 2 groupes
(Taille d'effet)

Evolution d’une variable quantitative entre deux mesures (Comparaison de groupes appariés)

Introduction

Position du problème

- On considère une variable quantitative Y qui va être mesurée à deux reprises sur les mêmes individus (on parle de groupes appariés). Par exemple :
  - La moyenne des étudiants au semestre 1 et au semestre 2.
  - Le niveau de stress avant et après un examen.
  - Le nombre de situations correctement rappelées lorsqu'ils sont décrites par une phrase et lorsqu'elles sont décrites par un film.
- On veut tester si l'évolution de X entre les deux mesures est significative.

Conditions d'application

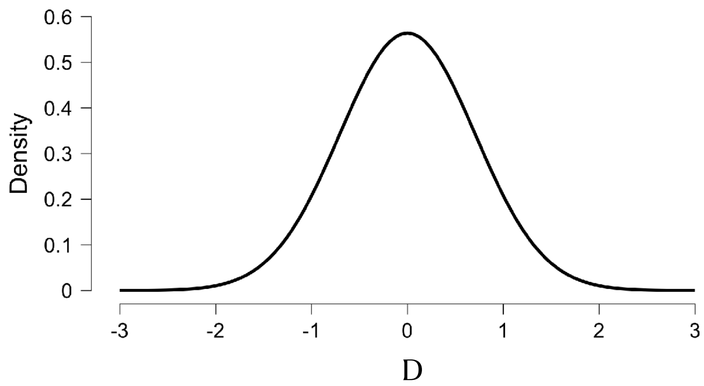
- Soient Y1 et Y2 les deux mesures considérées de la variable X. On définit

$$D = Y_2 - Y_1 \text{ (Différence entre les deux mesures)}$$

On note μ_D, σ_D la moyenne et l'écart type de D.

- On suppose que la distribution des différences est connue et que

$$D \sim \mathcal{N}(\mu_D, \sigma_D)$$



Procédure de test

t test de Student sur groupes appariés

- On teste l'hypothèse : $H_0 : \mu_D = 0$ versus $H_1 : \mu_D \neq 0$.
- On considère un échantillon de différences de taille n. On calcule respectivement $\bar{d}$, s_d la moyenne et l'écart type corrigé de D sur cet échantillon.
- On calcule $t = \dfrac{\bar{d}}{\frac{s_d}{\sqrt{n}}}$
- Sous Ho, T (la variable correspondant à t) suit une loi de Student à $n-1$ degrés de liberté (ddl).
- La p-value du test est donnée par $p = \mathbb{P}_{H_0}(|T| > |t|)$ où $\mathbb{P}_{H_0}$ est la probabilité sachant que Ho est vraie.
- La taille d'effet (Cohen) est $\delta = \dfrac{\bar{d}}{s_d}$. Les standards vus pour les groupes indépendants restent les mêmes.

Exemple

Data base de JASP

Description:

This data set, "Moon & Aggression", provides the number of disruptive behaviors by dementia patients during two different phases of the lunar cycle (Moore et al, 2012, p. 410). Each row corresponds to one participant.

Variables:

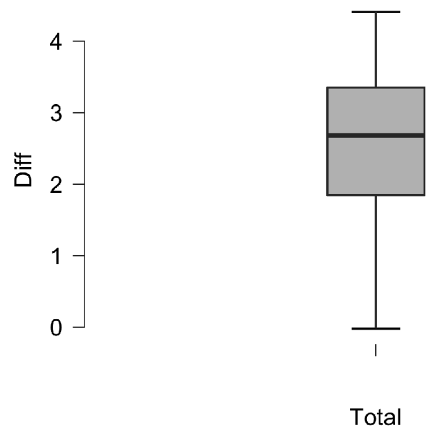
- **Moon** - The average number of disruptive behaviors during full moon days.

- **Other -** The average number of disruptive behaviors during other days.

Moon	Other	Diff
3,33	0,27	3,06
3,67	0,59	3,08
2,67	0,32	2,35
3,33	0,19	3,14
3,33	1,26	2,07
3,67	0,11	3,56
4,67	0,3	4,37
2,67	0,4	2,27
6	1,59	4,41
4,33	0,6	3,73
3,33	0,65	2,68
0,67	0,69	-0,02
1,33	1,26	0,07
0,33	0,23	0,1
2	0,38	1,62

Mise en oeuvre du test

L'échantillon des différences peut être considéré comme extrait d'une distribution normale car il n'y a aucune valeur extrême et sa distribution est relativement symétrique (voir boxplot ci-contre).



- On a sur l'échantillon : $n = 15,\ \bar{d} = 2.433,\ s_d = 1.460.$
- $t = \dfrac{2.433}{\frac{1.46}{\sqrt{15}}} = 6.45.$
- $\delta = \dfrac{2.433}{1.46} = 1.66$

	p_value (test bilatéral)							
ddl	**0,2**	**0,1**	**0,05**	**0,02**	**0,01**	**0,002**	**0,001**	
14	1,345	1,761	2,145	2,624	2,977	3,787	4,140	$\longrightarrow$ 6.45

Conclusion normes APA

Test de Student sur groupes appariés

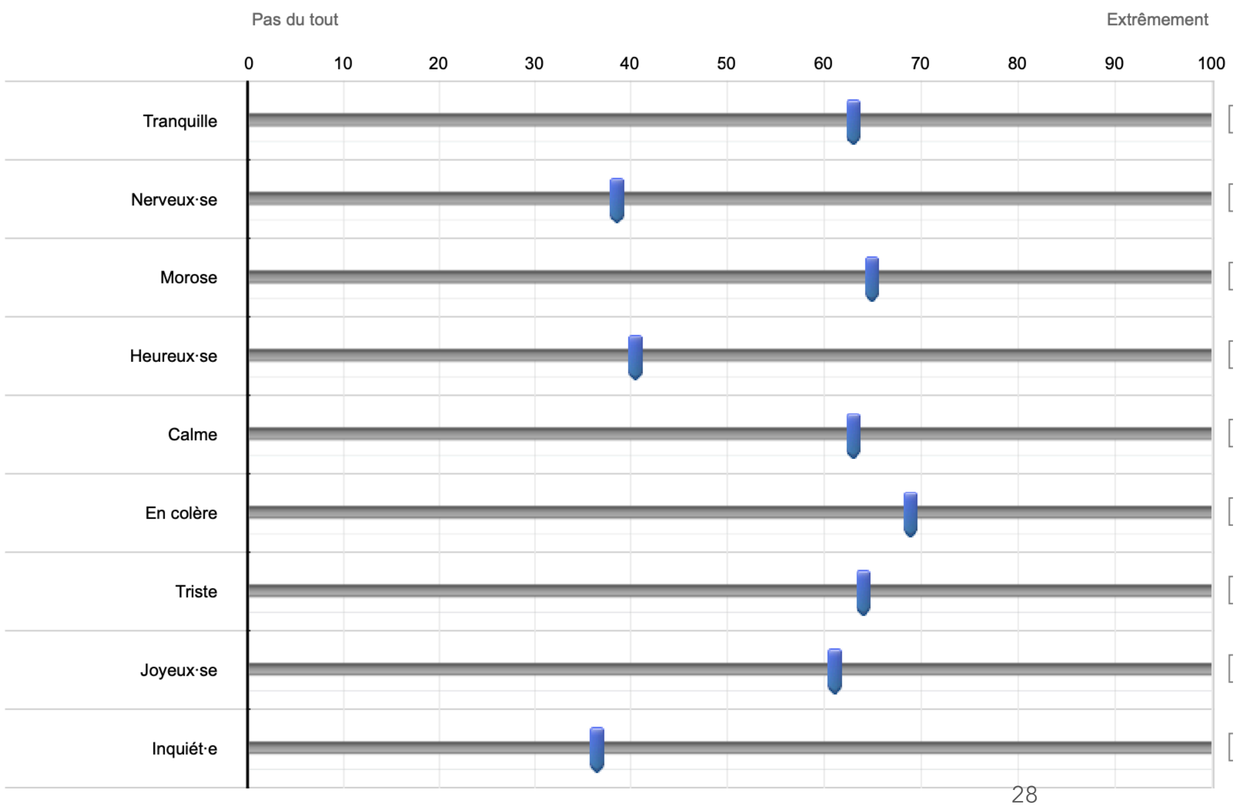
Le nombre moyen de comportements agressifs lors des jours de pleine lune ($M = 3.0$, $SD = 1.50$) des patients atteints de démence est significativement supérieur aux autres jours ($M = 0.6$, $SD = 0.45$), $t(14) = 6.45$, $p < .001$. Cette différence est grande d'après les standards de Cohen ($\delta = 1.66$)

Exercice d'application

Contexte de l'étude

COVID-19 2020-21

MAVA. Dans l'immédiat face à la crise du COVID-19, je me sens :



AP : affects positifs

API : Moyenne de Tranquille + Calme

APA : Moyenne de Heureux·se + Joyeux·se

AP : Moyenne de API et APA

AN : affects négatifs

ANI : Morose + Triste

ANA : Nerveux·se + Inquiét·e

AN : Moyenne de de ANI et ANA

Analyses différentielles selon le sexe

Confinement 1 (C1)

Tableau 1 Analyse descriptive

Levels	Counts	% of Total
1=Femme	102	62.2 %
2=Homme	62	37.8%

Tableau 2 Analyse descriptive des variables d’intérêt (m : moyenne, s : écart type)

	Sexe	API_C1	APA_C1	ANI_C1	ANA_C1	AP_C1	AN_C1
m	1	57.3	36.3	43.0	57.3	46.8	50.1
	2	62.8	38.6	39.8	47.2	50.7	43.5
s	1	23.5	25.6	26.6	23.3	20.9	22.6
	2	24.8	24.7	27.4	26.4	21.8	23.5

Tableau 3 : Résultats des Analyses différentielles selon le sexe

Variable	t	df	p	Mean difference	SE difference	Effect Size
API_C1	-1.419	162	0.158	-5.48	3.87	-0.2284
APA_C1	-0.570	162	0.569	-2.32	4.07	-0.0918
ANI_C1	0.739	162	0.461	3.20	4.33	0.1189
ANA_C1	2.555	162	0.012	10.09	3.95	0.4114
AP_C1	-1.140	162	0.256	-3.90	3.42	-0.1836
AN_C1	1.797	162	0.074	6.65	3.7	0.2894

Analyse différentielle entre les confinements

Évolution des affects négatifs entre C1 et C2

Tableau 1 Analyse différentielle des affects négatifs entre le confinement C1 et C2

		statistic	df	p	Effect Size
ANI_C1	ANI_C2	-2.82	163	0.005	-0.220
ANA_C1	ANA_C2	5.52	163	< .001	0.431