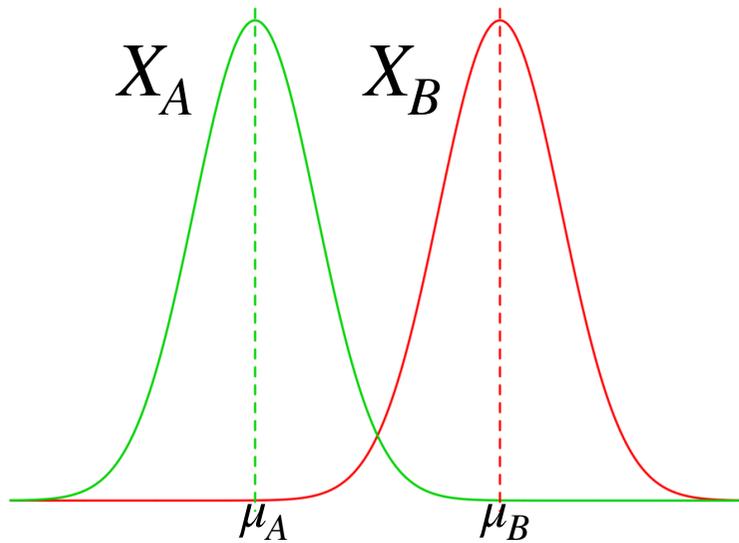


Tests non paramétriques basés sur les rangs

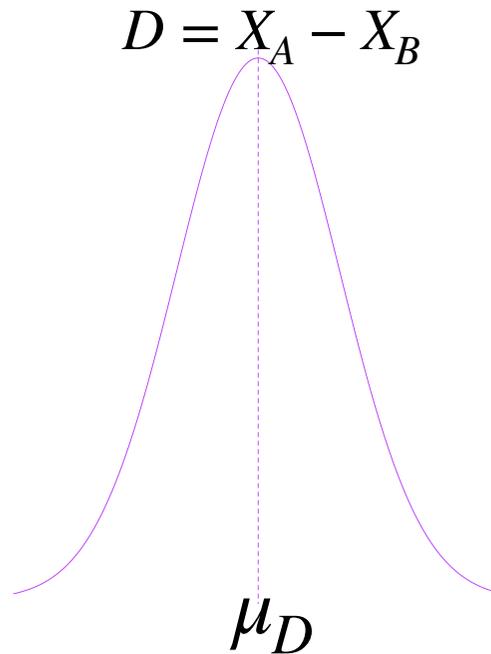
Module HPS3-3A

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

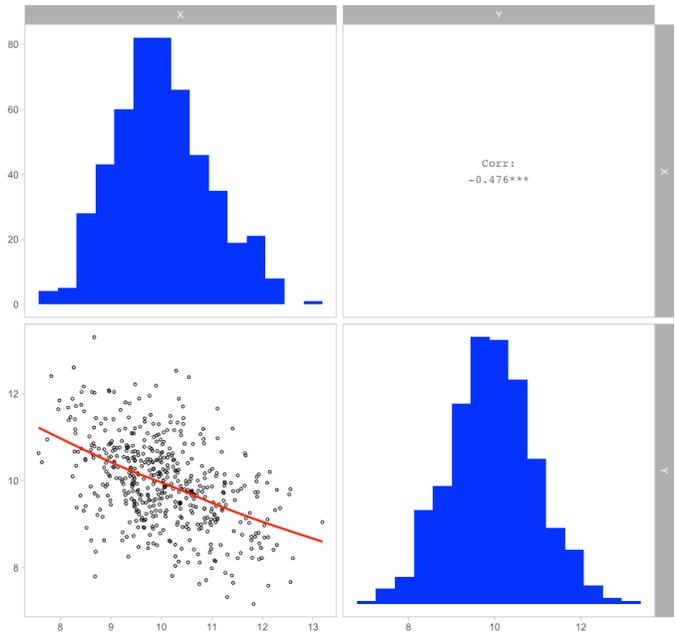
Tests paramétriques



$$H_0 : \mu_A = \mu_B$$



$$H_0 : \mu_D = 0$$



$$H_0 : r = 0$$

Question : Que faire lorsque la normalité des données peut être remise en cause ?

Introduction aux tests non paramétriques

- Ils s'utilisent couramment dans les cas suivants :
 - Présence de valeur(s) extrême(s) dans l'un au moins des échantillons.
 - Tailles d'échantillons faibles ($n_A < 10$ ou $n_B < 10$) \rightsquigarrow ce n'est pas une bonne idée, à éviter !!!
 - Très forte asymétrie des distributions (effet plancher ou effet plafond).
 - Les variables X_A, X_B sont ordinales (c'est à dire nominales mais pour lesquelles on peut classer les modalités).
- Ne pas utiliser les tests non paramétriques lorsque les variables n'ont pas la même dispersion !

Test de Mann-Whitney

(Comparaison de deux groupes indépendants)

Généralités

Conditions d'application

- On considère une variable mesurée X définie sur deux populations A et B. On note comme précédemment X_A, X_B les distributions de X sur A et sur B.
- On teste H_0 : « Les distributions X_A, X_B sont les mêmes » contre H_1 : « L'une des distribution a des valeurs systématiquement plus grandes que l'autre ».
- En termes mathématiques cela se traduit par

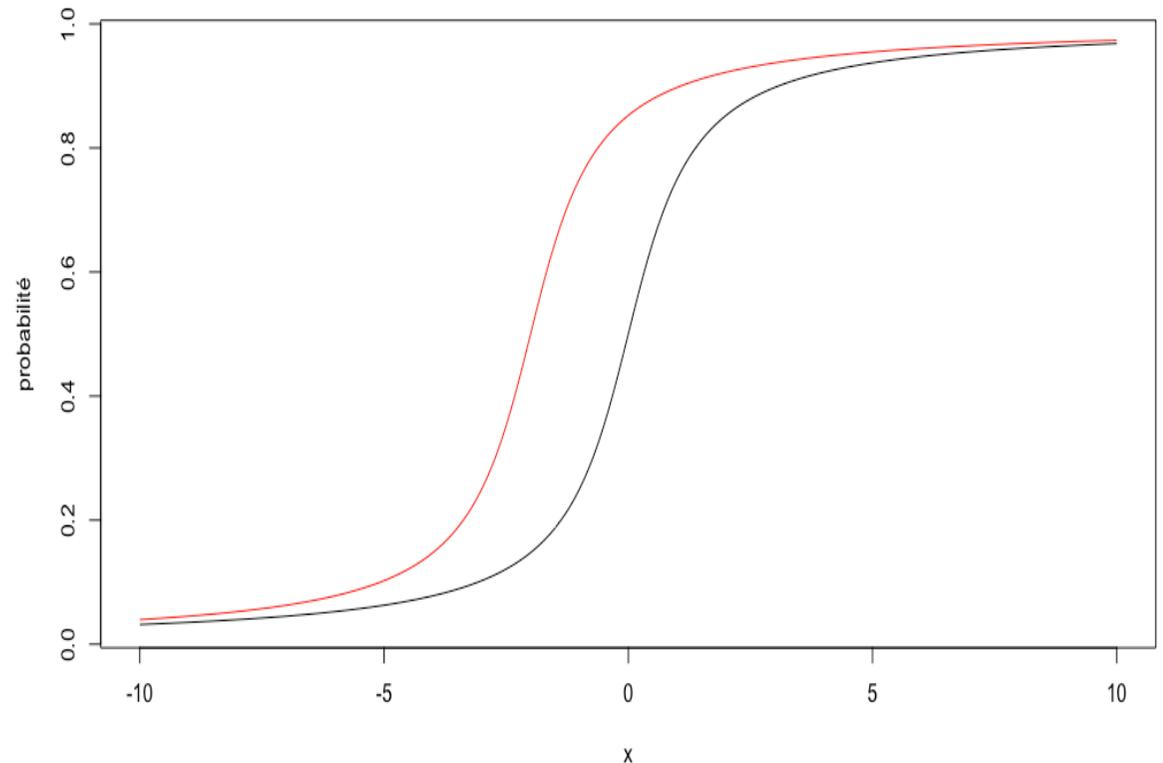
$$H_0 : F_A(x) = F_B(x), \text{ pour tout } x \in \mathbb{R}$$

versus

$$H_1 : F_A(x) > F_B(x), \text{ pour tout } x \in \mathbb{R}$$

Rappel : La fonction de répartition d'une variable X est définie par

$$F(x) = \mathbb{P}(X < x)$$



Généralités

Interprétation

- Lorsque les distributions de X sur A et B sont normales cette question de « décalage » des distributions est ramenée à la question des différences de moyennes.
- Souvent le test de M-W est décrit comme le test permettant de tester l'égalité des médianes. Ceci n'est vrai que si l'on suppose que les distributions ont exactement la même forme (ceci est en pratique invérifiable). Donc on évitera ce raccourci.

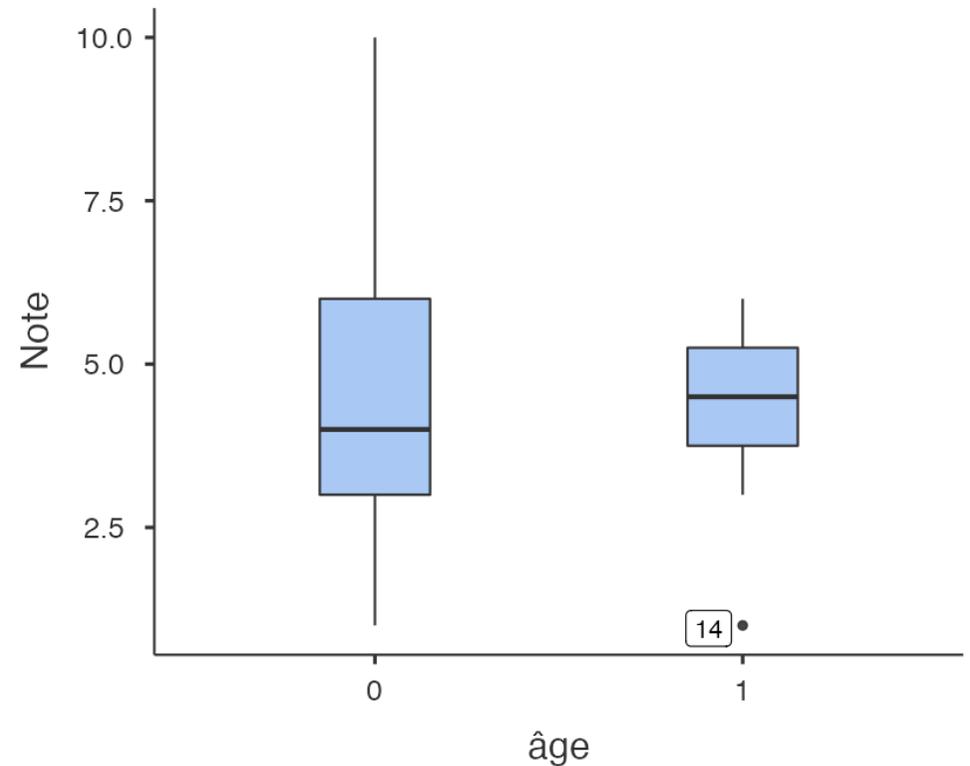
Exemple

Notes attribuées en fonction de l'âge

0=Jeunes	3	1	5	4	6	4	6	2	10
1=Âgés	4	3	6	6	1	5	5	4	

Question : Dans une pré-étude on a demandé à deux groupes d'individus qui sont caractérisés par leur âge (0 = moins de 40ans et 1 = plus de 40 ans) de noter un scénario qui leur a été proposé.

On ne va pas utiliser un test de Student car on a peu de valeurs (n=9 et n=8) et on a de plus une valeur extrême qui risque d'impacter la moyenne.



Intuition du résultat

Calcul des fonctions de répartition des deux échantillons

x	1	2	3	4	5	6	7	8	9	10
$F_0(x)$	11,1 %	22,2 %	33,3 %	55,6 %	66,7 %	88,9 %	88,9 %	88,9 %	88,9 %	100,0 %
$F_1(x)$	12,5 %	12,5 %	25,0 %	50,0 %	75,0 %	100,0 %	100,0 %	100,0 %	100,0 %	100,0 %

Notes attribuées en fonction de l'âge

0=Jeunes	3	1	5	4	6	4	6	2	10
1=Âgés	4	3	6	6	1	5	5	4	

Calcul de rangs

Mise en oeuvre du test

- Pour effectuer le test de M-W **on va remplacer les valeurs des échantillons par leur rang.**
- Le calcul des rangs s'effectuera de la façon suivante :
 - Les valeurs sont rangées par ordre croissant.
 - En cas d'ex-aequo on attribuera à chaque valeur un rang moyen.

Procédure de test

Statistique de test

- Soient $n_A < n_B$ les tailles des échantillons de valeurs de X sur les groupes A et B
- On fusionne les deux échantillons et on calcule le rang de chaque valeurs dans l'ensemble obtenu. On appelle W_A, W_B la somme des rangs des valeurs du groupe A et B.
- On calcule $U_A = W_A - \frac{n_A(n_A + 1)}{2}$ et $U_B = W_B - \frac{n_B(n_B + 1)}{2}$.
- La statistique du test de Mann-Whitney est donnée par :

$$U = \min(U_A, U_B)$$

Remarque : Lorsque l'hypothèse H_0 est vraie alors on a $U = U_A = U_B$

Calcul des rangs des valeurs considérées

0=Jeunes	3	1	5	4	6	4	6	2	10	Somme
1=Âgés	4	3	6	6	1	5	5	4		
Rang (0)	4,5	1,5	11	7,5	14,5	7,5	14,5	3	17	81
Rang (1)	7,5	4,5	14,5	14,5	1,5	11	11	7,5		72

$$U_0 = 81 - \frac{9 \times 10}{2} = 36$$

$$U_1 = 72 - \frac{8 \times 9}{2} = 36$$

$$U = 36$$

Lecture de la table de M-W

p_value (test bilatéral)	n ₂ =3		4		5		6		7		8		9		10		11		12		13		14		15	
	0,05	0,01	0,05	0,01	0,05	0,01	0,05	0,01	0,05	0,01	0,05	0,01	0,05	0,01	0,05	0,01	0,05	0,01	0,05	0,01	0,05	0,01	0,05	0,01	0,05	0,01
n ₁ =3	-	-	-	-	0	-	1	-	1	-	2	-	2	0	3	0	3	0	4	1	4	1	5	1	5	2
4			0	-	1	-	2	0	3	0	4	1	4	1	5	2	6	2	7	3	8	3	9	4	10	5
5					2	0	3	1	5	1	6	2	7	3	8	4	9	5	11	6	12	7	13	7	14	8
6							5	2	6	3	8	4	10	5	11	6	13	7	14	9	16	10	17	11	19	12
7									8	4	10	6	12	7	14	9	16	10	18	12	20	13	22	15	24	16
8											13	7	15	9	17	11	19	13	22	15	24	17	26	18	29	20
9													17	11	21	13	23	16	26	18	28	20	31	22	34	24
10															23	16	26	18	29	21	33	24	36	26	39	29
11																	30	21	33	24	37	27	40	30	44	33

On regarde dans la table de M-W pour $n_A = 9, n_B = 8$ on constate que pour :

$$U = 15, p = .05$$

$$U = 9, p = .01$$

Donc plus U augmente et plus la p-value augmente également.

On a $U = 28,5$ donc $p > .05$.

Rédaction de la conclusion aux normes APA

Test de Mann-Whitney

Un test de Mann-Whitney n'a pas permis de mettre en évidence qu'il y a une différence significative au niveau de la distribution des notes selon l'âge du juge, $U = 36$, $p > .05$.

Remarque : il existe également des tailles d'effet pour les tests non paramétriques mais le calcul est trop complexe pour être réalisé à la main.

Quelques remarques

Test de M-W

- Si on fait le test de M-W, pour les statistiques descriptives, on n'utilise plus les moyennes ni les écarts types des groupes mais les médianes et les IQR (intervalles interquartiles).
- Lorsque $n_A > 10$ et $n_B > 10$, plutôt qu'utiliser la table de M-W on peut utiliser l'approximation

$$\frac{U - \mathbb{E}(U)}{\sqrt{\text{Var}(U)}} \sim \mathcal{N}(0,1)$$

Avec si H_0 est vraie : $\mathbb{E}(U) = \frac{n_A n_B}{2}$ et $\text{Var}(U) = \frac{n_A n_B (n_A + n_B + 1)}{12}$.

Ceci nous permet d'obtenir une p-value approchée du test. Pour que $p < .05$ il suffit que le rapport précédent soit supérieur à $z = 1.96$.

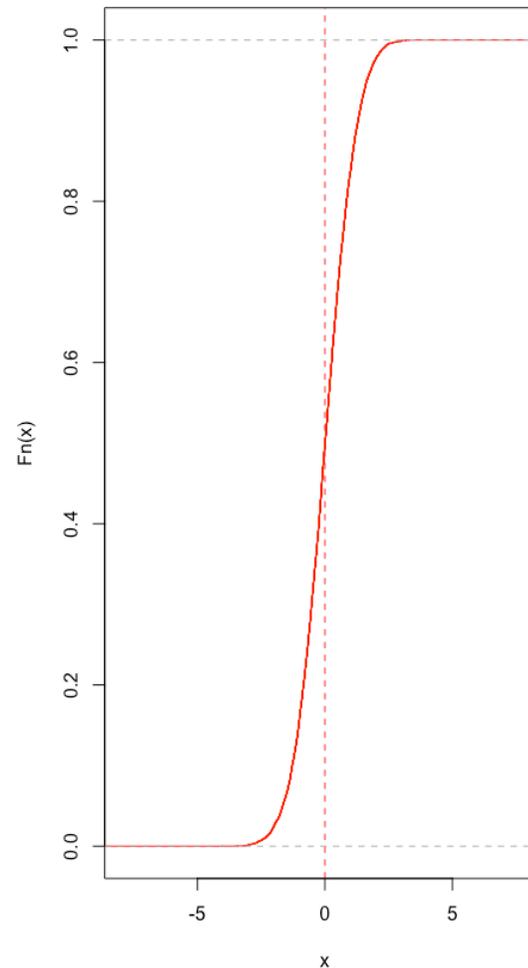
- Lorsque les deux distributions sont normales le test de M-W est 5% moins puissant que le test de Student. De plus le test de Student est plus facile à interpréter.

Test de Wilcoxon

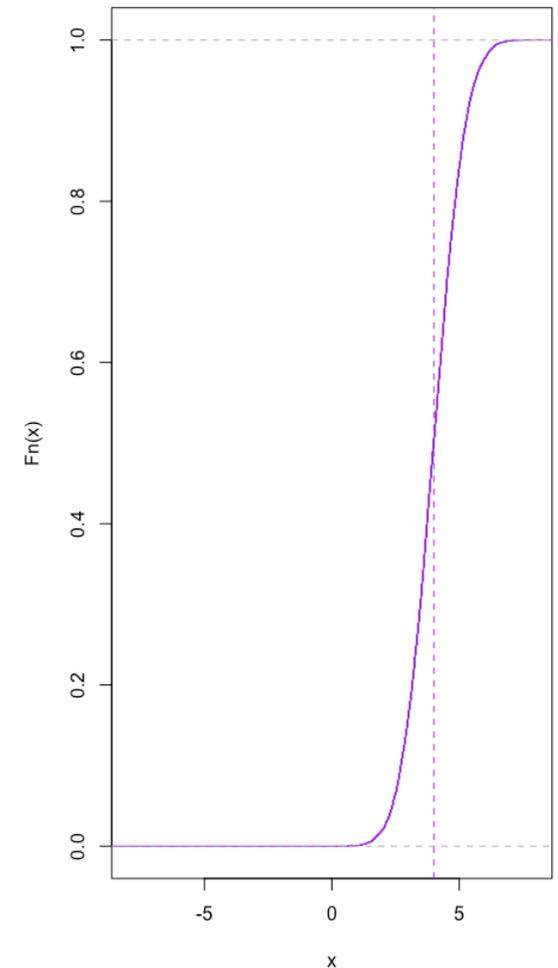
Test de Wilcoxon

- On considère une variable quantitative X mesurée à deux reprises. On note comme précédemment X_1, X_2 les distributions de X et $D = X_2 - X_1$ la différence entre les deux mesures. Soit F_D la fonction de répartition de la différence et \tilde{m} la médiane de D sur la population.
- On teste $H_0 : \tilde{m} = 0$ versus $H_1 : \tilde{m} > 0$. Autrement dit H_0 : la distribution des différences positives n'est pas significativement différente de celle des différences négatives

Fonction de répartition des différences sous H_0



Fonction de répartition des différences sous H_1



Procédure de test

Statistique de test

- On calcule les différences $D = X_2 - X_1$. On appelle n_d le nombre de différences non nulles.
- On range par ordre croissant les valeurs $|D| \neq 0$. On note r_i les rangs obtenus
- On calcule $W_+ = \sum_{i: D_i > 0} r_i$ la somme des rangs positifs et $W_- = \sum_{i: D_i < 0} r_i$ la somme des rangs négatifs.
- La statistique du test de Wilcoxon est donnée par :

$$W = \min(W_+, W_-)$$

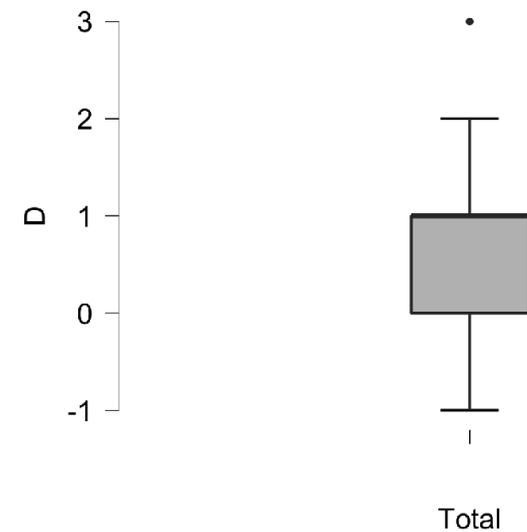
Remarque : Lorsque l'hypothèse H_0 est vraie alors on a $W = W_+ = W_-$

Exemple

Notations de restaurant

- Un journaliste d'une revue spécialisée à noter 10 restaurants (sur /10) selon deux critères : la qualité du service et de la nourriture.
- On veut prouver que la nourriture est systématiquement mieux notée que le service.

Restaurant	Service	Nourriture	D
1	9	9	0
2	8	9	1
3	5	8	3
4	7	6	-1
5	6	8	2
6	7	8	1
7	6	7	1
8	10	10	0
9	8	7	-1
10	5	6	1



Mise en oeuvre du test

Restaurant	D	D	rang D	rang(+)	rang(-)
1	0	-	-	-	-
2	1	1	3.5	3.5	-
3	3	3	8	8	-
4	-1	1	3.5	-	3.5
5	2	2	7	7	-
6	1	1	3.5	3.5	-
7	1	1	3.5	3.5	-
8	0	-	-	-	-
9	-1	1	3.5	-	3.5
10	1	1	3.5	3.5	-
Somme			29	7	

$$n_d = 8$$

$$W_+ = 29$$

$$W_- = 7$$

$$W = 7$$

19

n	p_value (test bilatéral)		
	0,05	0,02	0,01
6	0		
7	2	0	
8	4	2	0
9	6	3	2
10	8	5	3

On regarde la table de Wilcoxon à la ligne $n_d = 8$ et on a donc $p=.05$ pour $W=4$, $p=.02$ pour $W=2$.
Or $W=7$ donc $p>.05$ dans notre exemple.

Rédaction de la conclusion aux normes APA

Test de Wilcoxon

Un test de Wilcoxon n'a pas permis de mettre en évidence que les notes de restauration ($Mdn =$) sont systématiquement supérieures à celles du service ($Mdn =$), $W = 7$, $p > .05$.

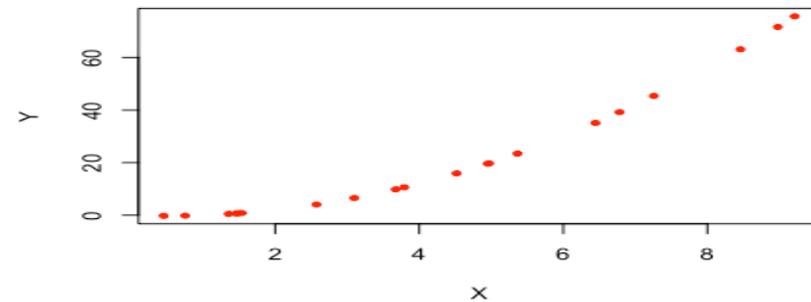
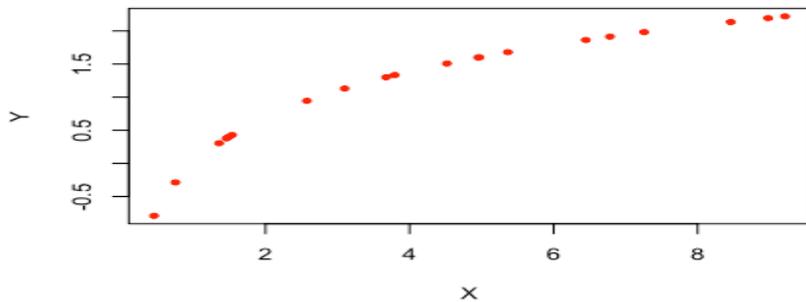
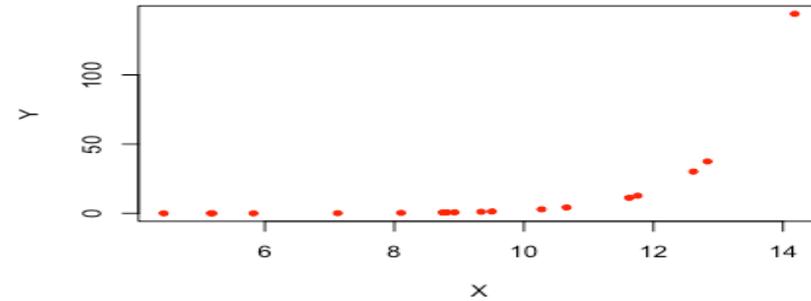
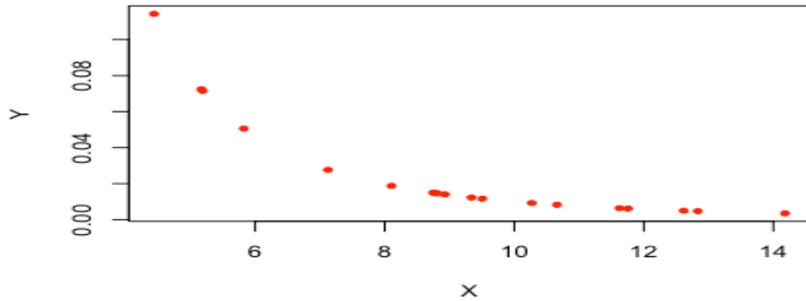
Remarque : Lorsque $n_d > 20$ plutôt qu'utiliser la table de Wilcoxon on peut utiliser l'approximation

$$\frac{W - \mathbb{E}(W)}{\sqrt{\text{Var}(W)}} \sim \mathcal{N}(0,1)$$

Avec, si H_0 est vraie : $\mathbb{E}(W) = \frac{n_d(n_d + 1)}{4}$ et $\text{Var}(W) = \frac{n_d(n_d + 1)(2n_d + 1)}{24}$.

Test de Spearman

Relations non linéaires



La relation entre X et Y n'est pas linéaire. Elle doit cependant être monotone pour pouvoir appliquer le test.

Procédure de test

Test de Spearman

- Les valeurs des variables sont remplacées par leur rangs c'est à dire (X_i, Y_i) est remplacé par (R_i, S_i) .
- On calcule le coefficient de corrélation ρ des rangs (R_i, S_i) (coefficient de corrélation de Spearman).
- La distribution de ρ est donnée dans la table de Spearman.

Mise en oeuvre du test

On a demandé à 10 individus d'évaluer leur niveau de stress sur une échelle de 1 à 5 et leur temps de sommeil (1 : moins de 4 h, 2 : entre 4h et 6h, 3 : de 6h à 8h, 4 : plus de 8h).

On pense que les personnes qui dorment le plus auront tendance à être moins stressées que les autres.

Participant	stress (X)	Temps_sommeil (Y)	Rang(X)	Rang(Y)
1	10	1	10	1,5
2	9	1	8,5	1,5
3	2	4	2	9
4	3	4	4	9
5	7	2	6	4,5
6	8	2	7	4,5
7	9	2	8,5	4,5
8	2	4	2	9
9	2	3	2	7
10	5	2	5	4,5


$$\rho = - .897$$

Conclusion normes APA

Test de Spearman

n	ddl	p_value (test bilatéral)			
		0,1	0,05	0,02	0,01
4	2	1,000	–	–	–
5	3	0,900	1,000	1,000	–
6	4	0,829	0,886	0,943	1,000
7	5	0,714	0,786	0,893	0,929
8	6	0,643	0,738	0,833	0,881
9	7	0,600	0,700	0,783	0,833
10	8	0,564	0,648	0,745	0,794
11	9	0,536	0,618	0,709	0,755
12	10	0,503	0,587	0,671	0,727
13	11	0,484	0,560	0,648	0,703

Un test de Spearman a permis de mettre en évidence qu'une diminution du stress est liée à une augmentation du temps de sommeil, $\rho(8) = - .897, p < .01$.