CORRIGES DES EXERCICES : Tests de Wilcoxon "signes et rangs" – Echantillons appariés

Exercice 4.1

$\mathcal{P}$ ={salariés d'une entreprise} de taille N inconnue,
X = score au test de perception d'autrui avant le programme,
Y = score au test de perception d'autrui après le programme, variables quantitatives définies sur $\mathcal{P}$

échantillon de X issu de $\mathcal{P}$ de taille n=11
échantillon de Y issu de $\mathcal{P}$ de taille n=11 $\Rightarrow$ deux échantillons appariés

Le programme est-il efficace, au risque 5% ?

Le programme est efficace se traduit par le fait que les valeurs de X sont globalement inférieures à celles de Y, c'est à dire que la loi de X est à gauche de celle de Y : " $X \prec Y$ "; c'est l'hypothèse que l'on veut accepter à tort au maximum dans 5% des cas, donc l'hypothèse alternative

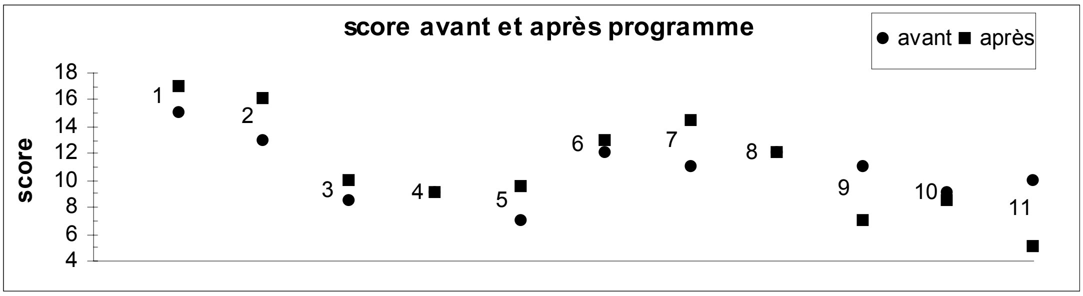
Le programme n'est pas efficace se traduit par le fait que les valeurs de X sont globalement égales à celles de Y, c'est à dire que la loi de X est identique à celle de Y : " $X \equiv Y$ "; c'est l'hypothèse nulle

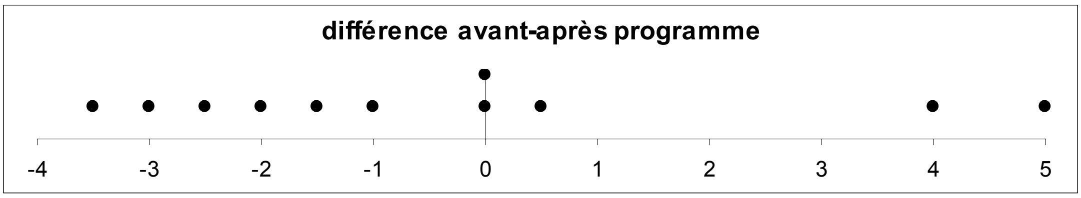
On considère D=X−Y la différence des scores au test de perception d'autrui avant et après le programme, de médiane M inconnue dans $\mathcal{P}$

hypothèse nulle	H_0 : "le programme est inefficace"	H_0 : " $X \equiv Y$ "	$\Leftrightarrow$	$H_0 : M = 0$
hypothèse alternative	H_1 : "le programme est efficace"	H_1 : " $X \prec Y$ "	$\Leftrightarrow$	$H_1 : M < 0$

unilatérale gauche au risque α=5%

salarié	1	2	3	4	5	6	7	8	9	10	11
score avant programme X	15	13	8,5	9	7	12	11	12	11	9	10
score après programme Y	17	16	10	9	9,5	13	14,5	12	7	8,5	5
différence D=X−Y	−2,0	−3,0	−1,5	0,0	−2,5	−1,0	−3,5	0,0	4,0	0,5	5,0
rang(|D|)	4	6	3		5	2	7		8	1	9
signe(D)	−	−	−	0	−	−	−	0	+	+	+





❶ **Comparaison de deux distributions sur échantillons appariés : test de Wilcoxon**

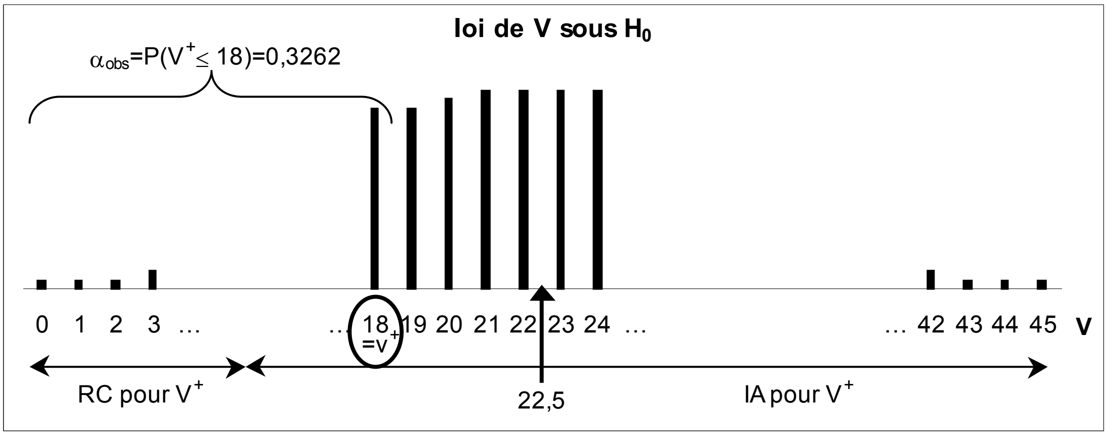
Echantillon de D issu de $\mathcal{P}$ de taille n=11−2=9 (nombre de valeurs de D non nulles)

Statistiques de test V^+ somme des rangs de |D| associés aux valeurs positives de D, et V^- :

somme $V^+ + V^- = 9\times10/2=45$
valeur observée de V^+ : $v^+=8+1+9=18$
valeur observée de V^- : $v^-=45-18=4+6+3+5+2+7=27$

loi exacte de la statistique de test V^+ sous l'hypothèse nulle (justifiée car n= 9 <20 et absence d'ex æquo)

sous H_0 : $X \equiv Y$ la statistique V^+ est quantitative discrète sur {0, …, 45} symétrique, de moyenne 45/2=22,5



- degré de signification du test unilatéral gauche (région de rejet à gauche du domaine de variation de V^+)

le degré de signification est la probabilité sous H_0 d'observer une valeur de la statistique de test V^+ au plus égale à la valeur v^+=18 observée car sous l'hypothèse H_1 on attend "peu" de valeurs positives de D avec des rangs faibles, donc de "petites" valeurs de V^+

$\alpha_{obs} = P_{H_0}(V^+ \leq v^+) = P_{H_0}(V^+ \leq 18) = 0{,}3262$ car $v^+=18>n=9$ table 3 colonne n=9 et ligne $v-n=18-9=9$

- décision

α_{obs} = 0,3262 > α=5% donc on ne rejette pas H_0 au seuil α=5% et au risque β inconnu

- conclusion

les valeurs de X (score de perception d'autrui avant programme) ne sont pas globalement inférieures à celles de Y (score de perception d'autrui après programme) donc le programme n'est pas significativement efficace (est inefficace) au seuil α=5% et au risque β inconnu

❷ Comparaison de deux distributions sur échantillons appariés : test du signe pour deux échantillons appariés

La variable *signe*(D) est qualitative dichotomique définie sur E={positif, négatif}, p proportion de signes strictement positifs **inconnue** dans $\mathcal{P}$,

échantillon de D issu de $\mathcal{P}$ de taille n=11−2=9 (nombre de valeurs de D non nulles) : estimation ponctuelle de p par f fréquence observée de signes strictement positifs dans l'échantillon avec s=3 donc f=3/9=0,333=33,3%

Comparaison de la proportion p *inconnue* à la valeur théorique p_0=50%=0,5

hypothèse nulle $H_0 : M = M_0=0$ s'écrit $H_0 : p = p_0= 0{,}5$

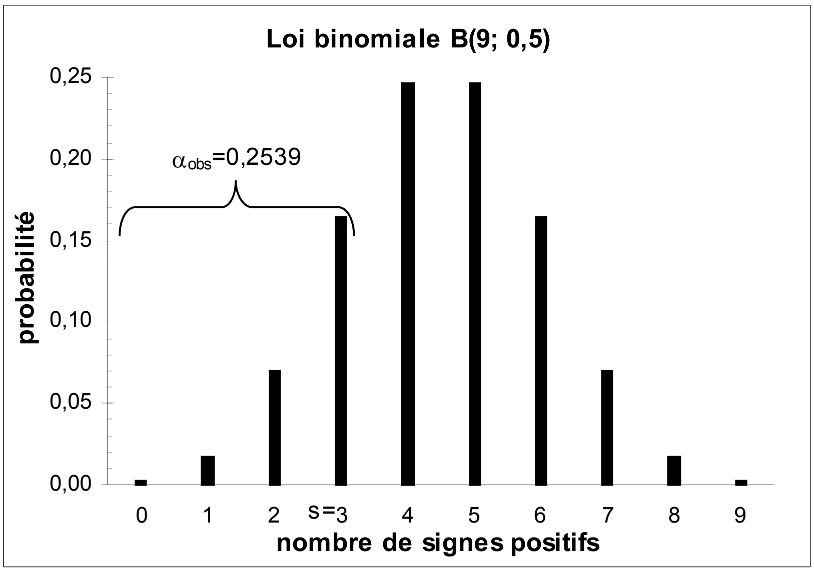
hypothèse alternative $H_1 : M < M_0=0$ s'écrit $H_1 : p < p_0 =0{,}5$ unilatérale gauche, au risque α=5%

- loi de la statistique de test sous l'hypothèse nulle

sous $H_0 : p = p_0 = 0{,}5$

la statistique de test S_n effectif empirique de signes strictement positifs, quantitative discrète définie sur $\{0, …, n\} = \{0, …, 9\}$ suit une loi binomiale $\mathcal{B}(n, p_0) = \mathcal{B}(9; 0{,}5)$ symétrique, de moyenne $np_0 = 9\times0{,}5 = 4{,}5$

	$S_n \sim\mathcal{B}$ (9; 0,5)	
s	$P(S_n=s)$	$P(S_n\leq s)$
0	0,0020	0,0020
1	0,0176	0,0195
2	0,0703	0,0898
3	0,1641	0,2539
4	0,2461	0,5000
5	0,2461	0,7461
6	0,1641	0,9102
7	0,0703	0,9805
8	0,0176	0,9980
9	0,0020	1,0000



- degré de signification du test unilatéral gauche (région de rejet à gauche du domaine de variation de S_n***)***

le degré de signification est la probabilité sous H_0 d'observer une valeur de la statistique de test S_n au plus égale à la valeur s=3 observée car sous H_1 on attend "peu" de valeurs de D positives, donc on prendra le risque de conclure H_1 (rejeter H_0 en faveur de H_1) à tort pour les "petites" valeurs de S_n

loi exacte $\mathcal{B}(9;\, 0{,}5)$: $\alpha_{obs} = P_{H_0}(S_n \leq s) = P_{H_0}(S_n \leq 3) = F_{\mathcal{B}}(3) = 0{,}2539$

où $F_{\mathcal{B}}$ est la fonction de répartition de la loi $\mathcal{B}(9;\, 0{,}5)$ (table 1 ligne n=9 colonne k=3)

- décision

$\alpha_{obs} = 0{,}2539 > \alpha = 5\%$ donc on ne rejette pas H_0 au seuil α=5% et au risque β inconnu

- conclusion

la proportion p de signes positifs de la différence D n'est pas significativement inférieure à 50%, au seuil α=5% et au risque β inconnu ; donc la médiane de la différence des scores de perception d'autrui avant et après le programme n'est pas significativement inférieure à 0, c'est à dire que le programme n'est pas significativement efficace (est inefficace) au seuil α=5% et au risque β inconnu

remarque

dans ce cas le test de Wilcoxon et celui du signe donnent des résultats proches et des conclusions identiques

Exercice 4.2

$\mathcal{P}$ ={paires de jumeaux} de taille N inconnue,
X = score au test de perception sociale à la garderie, Y = score au test de perception sociale à la maison,
variables quantitatives définies sur $\mathcal{P}$

échantillon de X issu de $\mathcal{P}$ de taille n=8
échantillon de Y issu de $\mathcal{P}$ de taille n=8 $\Rightarrow$ deux échantillons appariés

Les perceptions sociales à la garderie et à la maison sont-elles différentes, au risque 5% ?

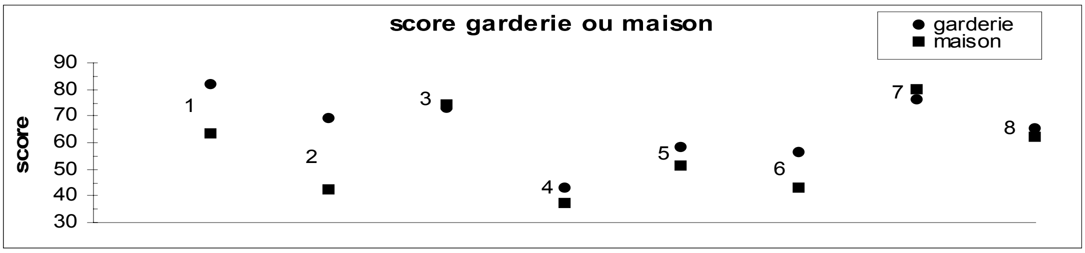
Les perceptions sociales sont différentes se traduit par le fait que les valeurs de X sont globalement différentes de celles de Y, c'est à dire que la loi de X est différente de celle de Y : " $X \not\equiv Y$ ", hypothèse que l'on veut accepter à tort au maximum dans 5% des cas, donc hypothèse alternative

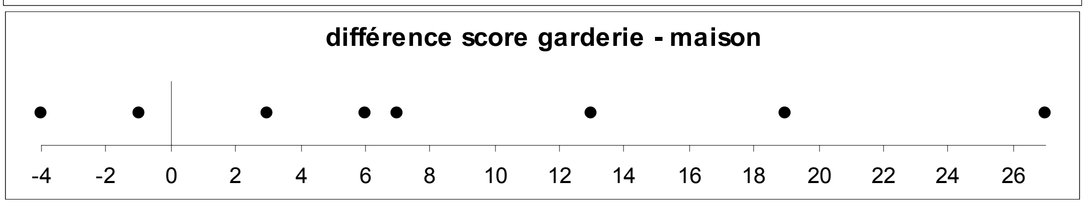
Les perceptions sociales ne sont pas différentes se traduit par le fait que les valeurs de X sont globalement égales à celles de Y, c'est à dire que la loi de X est identique à celle de Y, hypothèse nulle " $X \equiv Y$ "

On considère D=X−Y la différence du score au test de perception sociale du jumeau à la garderie et du score du jumeau à la maison, de médiane M inconnue dans $\mathcal{P}$

hypothèse nulle H_0 : "les perceptions sociales sont identiques" H_0 : " $X \equiv Y$ " $\Leftrightarrow$ $H_0 : M = 0$
hypothèse alternative H_1 : "les perceptions sociales sont différentes" H_1 : " $X \not\equiv Y$ " $\Leftrightarrow$ $H_1 : M \neq 0$
bilatérale au risque α=5%

paire	1	2	3	4	5	6	7	8
score à la garderie X	82	69	73	43	58	56	76	65
score à la maison Y	63	42	74	37	51	43	80	62
différence D=X−Y	19	27	−1	6	7	13	–4	3
signe(D)	+	+	–	+	+	+	–	+
rang(|D|)	7	8	1	4	5	6	3	2





❶ Comparaison de deux distributions sur échantillons appariés : test de Wilcoxon

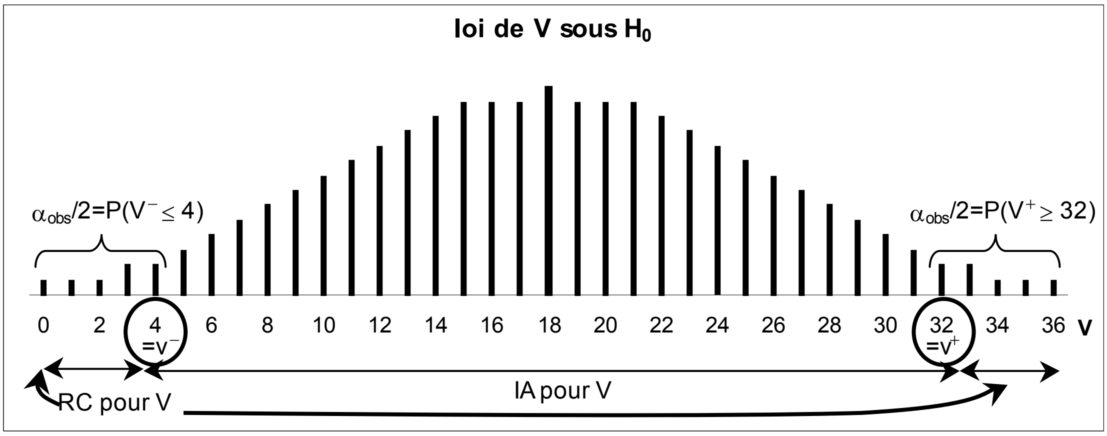
Echantillon de D issu de $\mathcal{P}$ de taille n=8 (nombre de valeurs de D non nulles)

Statistiques de test V^+ somme des rangs de |D| associés aux valeurs positives de D, et V^- :

somme $V^+ + V^- = 8\times9/2=36$
valeur observée de V^- : $v^-=1+3=4$
valeur observée de V^+ : $v^+=36-4=7+8+4+5+6+2=32$

loi exacte de la statistique de test V^+ sous l'hypothèse nulle (justifiée car n= 8 <20 et absence d'ex æquo)

sous H_0 : X ≡ Y la statistique V^+ est quantitative discrète sur {0, …, 36} symétrique, de moyenne 36/2=18



- degré de signification du test bilatéral (région de rejet bilatérale, petites ou grandes valeurs de V^+ et de V^-)

puisque $v^- = 4 < v^+ = 32$

$\alpha_{obs} = P_{H_0}(V^- \le v^-) + P_{H_0}(V^+ \ge v^+) = 2\times P_{H_0}(V^+ \ge v^+) = 2\times P_{H_0}(V^+ \ge 32) = 2\times(1-P_{H_0}(V^+ \le 31))$

$= 2\times(1-0{,}9727)) = 2\times0{,}0273 = 0{,}0546$ car v=31>n=8 table 3 colonne n=8 et ligne v–n=31–8=23

ou $\alpha_{obs} = P_{H_0}(V^- \le v^-) + P_{H_0}(V^+ \ge v^+) = 2\times P_{H_0}(V^- \le v^-) = 2\times P_{H_0}(V^- \le 4) = 2\times0{,}0273 = 0{,}0546$

car $v^-=4\le n=8$ table 2 colonne n=8 et ligne v=4

ou pour $v = \min(v^+, v^-) = v^- = 4$

$\alpha_{obs} = P_{H_0}(V \le v) + P_{H_0}(V \ge v) = 2\times P_{H_0}(V \le v) = 2\times P_{H_0}(V \le 4) = 2\times0{,}0273 = 0{,}0546$

car $v^-=4\le n=8$ table 2 colonne n=8 et ligne v=4

- décision

$\alpha_{obs} = 0{,}0546 > \alpha=5\%$ donc on ne rejette pas H_0 au seuil α=5% et au risque β inconnu

- conclusion

les perceptions sociales à la garderie et à la maison ne sont pas significativement différentes au seuil α=5% et au risque β inconnu

❷ Comparaison de deux distributions sur échantillons appariés : test du signe pour deux échantillons appariés

La variable *signe*(D) est qualitative dichotomique définie sur E={positif, négatif}, p proportion de signes strictement positifs **inconnue** dans $\mathcal{P}$,

échantillon de D issu de $\mathcal{P}$ de taille n=8 (nombre de valeurs de D non nulles) : estimation ponctuelle de p par f fréquence observée de signes strictement positifs dans l'échantillon avec s=6 donc f=6/8=0,75=75%

Comparaison de la proportion *p inconnue* à la valeur théorique p_0=50%=0,5

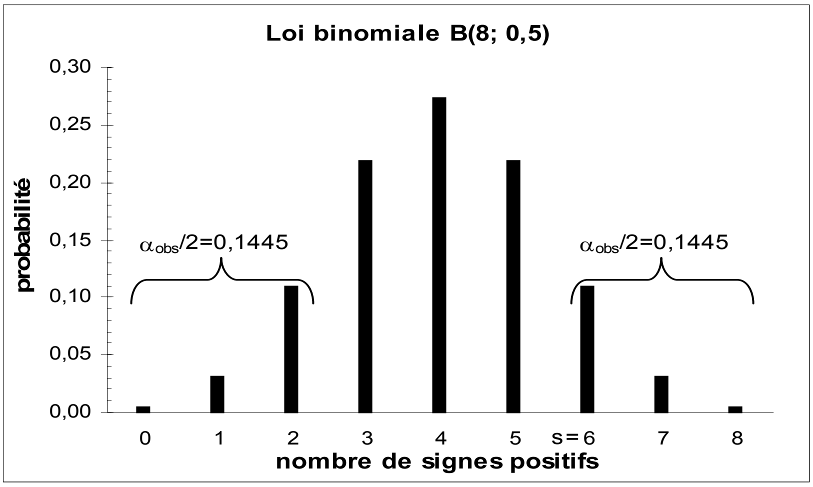
hypothèse nulle	$H_0 : M = M_0=0$	s'écrit	$H_0 : p = p_0 = 0{,}5$	
hypothèse alternative	$H_1 : M \ne M_0=0$	s'écrit	$H_1 : p \ne p_0 = 0{,}5$	bilatérale, au risque α=5%

- loi de la statistique de test sous l'hypothèse nulle

sous $H_0 : p = p_0 = 0{,}5$

la statistique de test S_n effectif empirique de signes strictement positifs, quantitative discrète définie sur $\{0, \ldots, n\} = \{0, \ldots, 8\}$ suit une loi binomiale $\mathcal{B}(n, p_0) = \mathcal{B}(8; 0{,}5)$ symétrique, de moyenne $np_0 = 8\times0{,}5 = 4$

	$S_n \sim \mathcal{B}$ (8; 0,5)	
s	$P(S_n = s)$	$P(S_n \le s)$
0	0,0039	0,0039
1	0,0313	0,0352
2	0,1094	0,1445
3	0,2188	0,3633
4	0,2734	0,6367
5	0,2188	0,8555
6	0,1094	0,9648
7	0,0313	0,9961
8	0,0039	1,0000



- degré de signification du test bilatéral (région de rejet bilatérale, petites ou grandes valeurs de S_n***)***

loi exacte $\mathcal{B}(8; 0,5)$: $\alpha_{obs} = P_{H_0}(S_n \ge s) + P_{H_0}(S_n \le n-s) = P_{H_0}(S_n \ge 6) + P_{H_0}(S_n \le 2)$

car 2 (= nombre de signes négatifs) est symétrique de 6 par rapport à $np_0 = n/2 = 4$

$\alpha_{obs} = 2 \times P_{H_0}(S_n \le 2) = 2 \times F_{\mathcal{B}}(2) = 2 \times 0,1445 = 0,289$

où $F_{\mathcal{B}}($ est la fonction de répartition de la loi $\mathcal{B}(8; 0,5)$ (table 1 ligne n=8 colonne k=2)

- décision

$\alpha_{obs} = 0,289 > \alpha = 5\%$ donc on ne rejette pas H_0 au seuil α=5% et au risque β inconnu

- conclusion

la proportion p de signes positifs de la différence de D n'est pas significativement différente de 50%, au seuil α=5% et au risque β inconnu ; donc la médiane de la différence de scores de perception sociale entre la garderie et la maison n'est pas significativement différente de la valeur théorique 0, au seuil α=5% et au risque β inconnu ; les perceptions sociales entre la garderie et la maison ne sont pas significativement différentes au seuil α=5% et au risque β inconnu

remarque

dans ce cas, le test de Wilcoxon et celui du signe donnent des résultats très différents bien que les conclusions soient semblables, ce qui s'explique par le fait que le test de Wilcoxon est plus puissant que celui du signe : il a un degré de signification beaucoup plus faible (α_{obs} = 0,0546) que celui du signe (α_{obs} = 0,289)

Exercice 4.3

$\mathcal{P}$ ={paires de jumeaux} de taille N inconnue,
X = score au test d'agressivité du premier né, Y = score au test d'agressivité du puîné,
variables quantitatives définies sur $\mathcal{P}$

échantillon de X issu de $\mathcal{P}$ de taille n=12
échantillon de Y issu de $\mathcal{P}$ de taille n=12 $\Rightarrow$ deux échantillons appariés

Le premier né est-il plus agressif que le puîné, au risque 2% ?

Le premier né est plus agressif que le puîné se traduit par le fait que les valeurs de X sont globalement supérieures à celles de Y, c'est à dire que la loi de X est à droite de celle de Y : " X $\succ$ Y ", hypothèse que l'on veut accepter à tort au maximum dans 2% des cas, donc hypothèse alternative

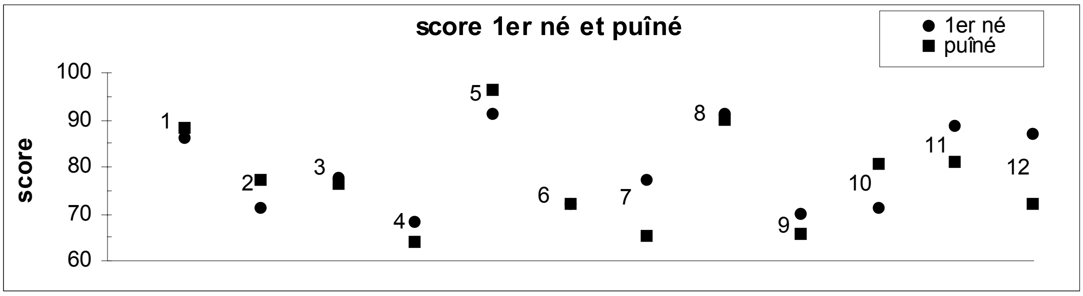
Le premier né n'est pas plus agressif que le puîné se traduit par le fait que les valeurs de X sont globalement égales à celles de Y, c'est à dire que la loi de X est identique à celle de Y : " X ≡ Y " hypothèse nulle

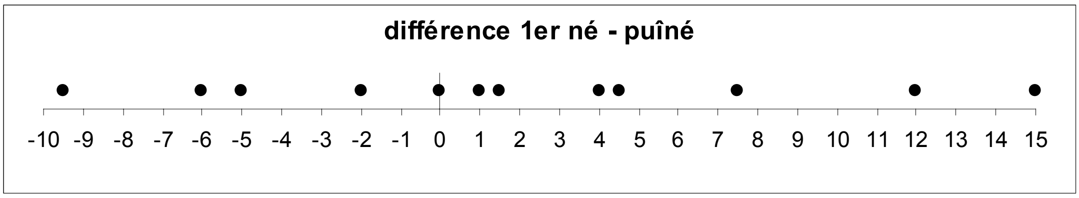
On considère D=X−Y la différence du score au test d'agressivité du jumeau né en premier et du score du jumeau né en second, de médiane M inconnue dans $\mathcal{P}$

hypothèse nulle H_0 : "les agressivités sont identiques" H_0 : " X ≡ Y " ⇔ $H_0 : M = 0$
hypothèse alternative H_1 : "le premier né est plus agressif que le puîné" H_1 : " X $\succ$ Y " ⇔ $H_1 : M > 0$
unilatérale droite au risque α=2%

paire	1	2	3	4	5	6	7	8	9	10	11	12
score d'agressivité du 1er né X	86	71	77,5	68	91	72	77	91	70	71	88,5	87
score d'agressivité du puîné Y	88	77	76	64	96	72	65	90	65,5	80,5	81	72
différence D=X−Y	−2,0	−6,0	1,5	4,0	−5,0	0,0	12,0	1,0	4,5	−9,5	7,5	15,0
rang(D)	4	2	7	8	3	5	11	6	9	1	10	12
signe(D)	–	–	+	+	–	0	+	+	+	–	+	+
rang(|D|)	3	7	2	4	6		10	1	5	9	8	11

pour n=12 la médiane observée m de D est égale à une valeur comprise entre les valeurs de rang (n/2)=6 et (n/2+1)=7 de D donc m se situe entre 1 et 1,5





❶ Comparaison de deux distributions sur échantillons appariés : test de Wilcoxon

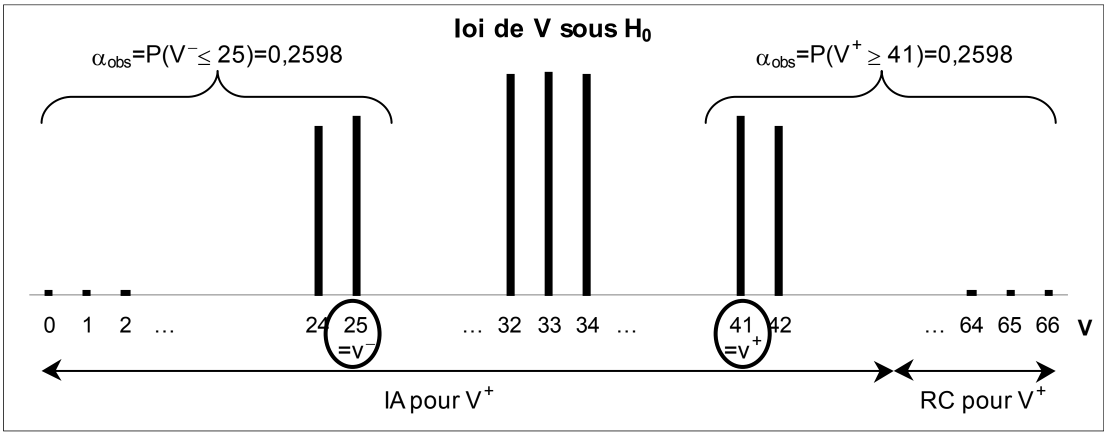
Echantillon de D issu de $\mathcal{P}$ de taille n=12−1=11 (nombre de valeurs de D non nulles)

Statistiques de test V^+ somme des rangs de |D| pour les différences positives, et V^- :

somme $V^+ + V^- = 11\times12/2=66$
valeur observée de V^- : $v^-=3+7+6+9=25$
valeur observée de V^+ : $v^+=66–25=2+4+10+1+5+8+11=41$

① **loi exacte de la statistique de test V^+ sous l'hypothèse nulle** (justifiée car n= 11 <20 et absence d'ex æquo)

sous H_0 : $X \equiv Y$ la statistique V^+ est quantitative discrète sur $\{0, \ldots, 66\}$ symétrique, de moyenne 66/2=33



- degré de signification du test unilatéral droit (région de rejet à droite du domaine de variation de V^+)

le degré de signification est la probabilité sous H_0 d'observer une valeur de la statistique de test V^+ au moins égale à la valeur v^+=41 observée car sous l'hypothèse H_1 on attend "beaucoup" de valeurs positives de D avec des rangs élevés, donc de "grandes" valeurs de V^+

$\alpha_{obs} = P_{H_0}(V^+ \geq v^+) = P_{H_0}(V^+ \geq 41) = 1 - P_{H_0}(V^+ \leq 40) = 1 - 0{,}7402 = 0{,}2598$

car v=40>n=11 table 3 colonne n=11 et ligne v−n=40−11=29

ou $\alpha_{obs} = P_{H_0}(V^+ \geq v^+) = P_{H_0}(V^- \leq v^-) = P_{H_0}(V^- \leq 25) = 0{,}2598$

car v^-=25>n=11 table 3 colonne n=11 et ligne $v^- - n$=25−11=14

- décision : $\alpha_{obs} = 0{,}2598 > \alpha = 2\%$ donc on ne rejette pas H_0 au seuil α=2% et au risque β inconnu

② **approximation normale** (justifiée car n =11 ≥10 mais peu utilisée en l'absence d'ex æquo pour n =11 ≤ 20)

- loi de la statistique de test sous l'hypothèse nulle

sous H_0 : " $X \equiv Y$ " les statistiques V^+ et V^- suivent approximativement une loi normale

de moyenne $\frac{n(n+1)}{4}$=66/2=33 et de variance $\frac{n(n+1)(2n+1)}{24}$=11×12×23/24=3036/24=126,5

donc la statistique de test $Z = \dfrac{V^+ - \frac{n(n+1)}{4}}{\sqrt{\frac{n(n+1)(2n+1)}{24}}}$ suit approximativement une loi normale $\mathcal{N}(0,1)$

- valeur observée de la statistique de test : $z = \dfrac{v^+ - \frac{n(n+1)}{4}}{\sqrt{\frac{n(n+1)(2n+1)}{24}}} = \dfrac{41-33}{\sqrt{126{,}5}} = 0{,}71$

- degré de signification du test unilatéral droit (région de rejet à droite)

$\alpha_{obs} = P_{H_0}(Z \geq z) = P_{H_0}(Z \geq 0{,}71) = 1 - P_{H_0}(Z \leq 0{,}71) = 1 - F_Z(0{,}71) = 1 - 0{,}7611 = 0{,}2389$

où F_Z est la fonction de répartition de la loi normale $\mathcal{N}(0,1)$

- décision : $\alpha_{obs} = 0{,}2389 > \alpha = 2\%$ donc on ne rejette pas H_0 au seuil α=2% et au risque β inconnu

remarque : le résultat avec l'approximation normale est légèrement différent de celui obtenu avec le test exact mais la conclusion reste identique

conclusion

le premier né n'est pas significativement plus agressif que le puîné, au seuil α=2% et au risque β inconnu

❷ **Comparaison de deux distributions sur échantillons appariés : test du signe pour deux échantillons appariés**

La variable *signe*(D) est qualitative dichotomique définie sur E={positif, négatif}, p proportion de signes strictement positifs **inconnue** dans $\mathcal{P}$,

échantillon de D issu de $\mathcal{P}$ de taille n=12−1=11 (nombre de valeurs de D non nulles) : estimation ponctuelle de p par f fréquence observée de signes strictement positifs dans l'échantillon avec s=7 donc f =7/11=0,636=63,6%

Comparaison de la proportion p *inconnue* à la valeur théorique p_0=50%=0,5

hypothèse nulle $H_0 : M = M_0 = 0$ s'écrit $H_0 : p = p_0 = 0{,}5$

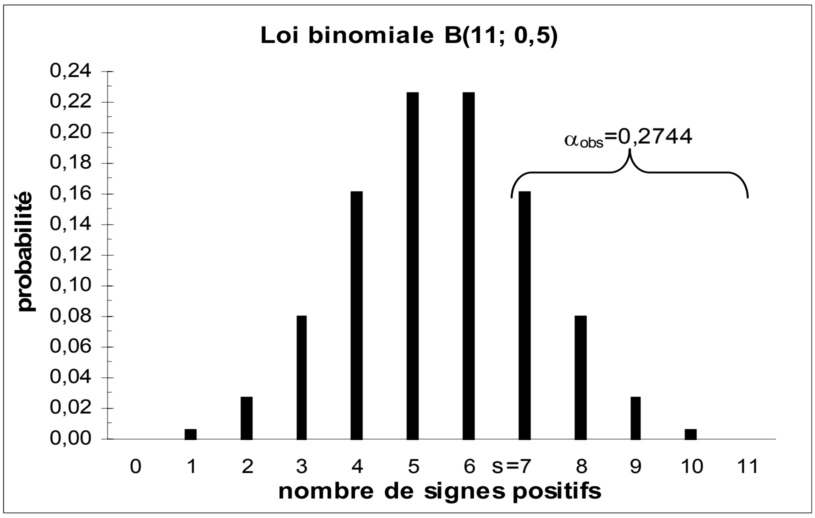
hypothèse alternative $H_1 : M > M_0 = 0$ s'écrit $H_1 : p > p_0 = 0{,}5$ unilatérale droite, au risque α=2%

- loi de la statistique de test sous l'hypothèse nulle

sous $H_0 : p = p_0 = 0{,}5$

la statistique de test S_n effectif empirique de signes strictement positifs, quantitative discrète définie sur $\{0, \ldots, n\} = \{0, \ldots, 11\}$ suit une loi binomiale $\mathcal{B}(n, p_0) = \mathcal{B}(11; 0{,}5)$ symétrique, de moyenne $np_0 = 11\times0{,}5 = 5{,}5$

	$S_n \sim\mathcal{B}$ (11; 0,5)	
s	P(S_n =*s*)	P($S_n \leq s$)
0	0,0005	0,0005
1	0,0054	0,0059
2	0,0269	0,0327
3	0,0806	0,1133
4	0,1611	0,2744
5	0,2256	0,5000
6	0,2256	0,7256
7	0,1611	0,8867
8	0,0806	0,9673
9	0,0269	0,9941
10	0,0054	0,9995
11	0,0005	1,0000



- degré de signification du test unilatéral droit (région de rejet à droite du domaine de variation de* S_n*)

le degré de signification est la probabilité sous H_0 d'observer une valeur de la statistique de test S_n au moins égale à la valeur s=7 observée car sous H_1 on attend "beaucoup" de valeurs de D positives, donc on prendra le risque de conclure H_1 (rejeter H_0 en faveur de H_1) à tort pour les "grandes" valeurs de S_n

loi exacte $\mathcal{B}(11; 0{,}5)$: $\alpha_{obs} = P_{H_0}(S_n \geq s) = P_{H_0}(S_n \geq 7) = 1-P_{H_0}(S_n < 7) = 1-P_{H_0}(S_n \leq 6)$

$\alpha_{obs} = 1-F_{\mathcal{B}}(6) = 1-0{,}7256 = 0{,}2744$ (table 1 ligne n=11 colonne k=6)

ou $\alpha_{obs} = P_{H_0}(S_n \geq s) = P_{H_0}(S_n \leq n-s) = P_{H_0}(S_n \leq 4)$

car 4 (= nombre de signes négatifs) est symétrique de 7 par rapport à np_0=n/2=5,5

$\alpha_{obs} = F_{\mathcal{B}}(4) = 0{,}2744$

où $F_{\mathcal{B}}$ est la fonction de répartition de la loi $\mathcal{B}(11; 0{,}5)$ (table 1 ligne n=11 colonne k=4)

- décision

$\alpha_{obs} = 0{,}2744 > \alpha$=2% donc on ne rejette pas H_0 au seuil α=2% et au risque β inconnu

- conclusion

la proportion p de signes positifs de la différence D n'est pas significativement supérieure à 50%, au seuil α=2% et au risque β inconnu ; donc la médiane de la différence des scores d'agressivité entre premier né et puîné n'est pas significativement supérieure à 0, c'est à dire que le premier né n'est pas significativement plus agressif que le puîné, au seuil α=2% et au risque β inconnu

remarque

dans ce cas le test de Wilcoxon et celui du signe donnent des résultats proches et des conclusions identiques

Exercice 4.4

$\mathcal{P}$ ={personnes} de taille N inconnue, X = poids avant entraînement, Y = poids après les 3 mois d'entraînement, variables quantitatives définies sur $\mathcal{P}$

échantillon de X issu de $\mathcal{P}$ de taille n=10
échantillon de Y issu de $\mathcal{P}$ de taille n=10 $\Rightarrow$ deux échantillons appariés

L'entraînement intensif aide-t-il à perdre du poids, au risque 5% ?

On considère D=X−Y la différence de poids entre avant et après entraînement, c'est à dire la perte de poids

L'entraînement aide à perdre du poids se traduit par le fait que les valeurs de X sont globalement supérieures à celles de Y, c'est à dire que la loi de X est à droite de celle de Y, c'est à dire que la perte de poids D a des valeurs globalement positives : " X $\succ$ Y ", hypothèse que l'on veut accepter à tort au maximum dans 5% des cas, donc hypothèse alternative

L'entraînement n'aide pas à perdre du poids se traduit par le fait que les valeurs de X sont globalement égales à celles de Y, c'est à dire que la loi de X est identique à celle de Y : " X $\equiv$ Y " hypothèse nulle

On considère D=X−Y la différence de poids entre avant et après entraînement, c'est à dire la perte de poids, de médiane M inconnue dans $\mathcal{P}$

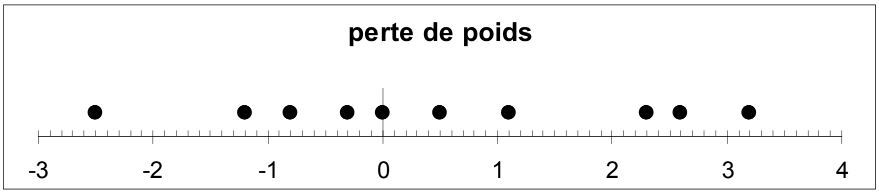
hypothèse nulle — H_0 : "l'entraînement n'aide pas à perdre du poids" — H_0 : " X $\equiv$ Y " $\Leftrightarrow$ $H_0 : M = 0$

hypothèse alternative — H_1 : "l'entraînement aide à perdre du poids" — H_1 : " X $\succ$ Y " $\Leftrightarrow$ $H_1 : M > 0$

unilatérale droite au risque α=5%

personne	1	2	3	4	5	6	7	8	9	10
différence D=X−Y	2,6	2,3	0,0	3,2	0,5	1,1	−1,2	−2,5	−0,8	−0,3
rang(D)	9	8	5	10	6	7	2	1	3	4
signe(D)	+	+	0	+	+	+	–	–	–	–
rang(|D|)	8	6		9	2	4	5	7	3	1

pour n=10 la médiane observée m de D est égale à une valeur comprise entre les valeurs de rang (n/2)=5 et (n/2+1)=6 de D donc m se situe entre 0 et 0,5



❶ Comparaison de deux distributions sur échantillons appariés : test de Wilcoxon

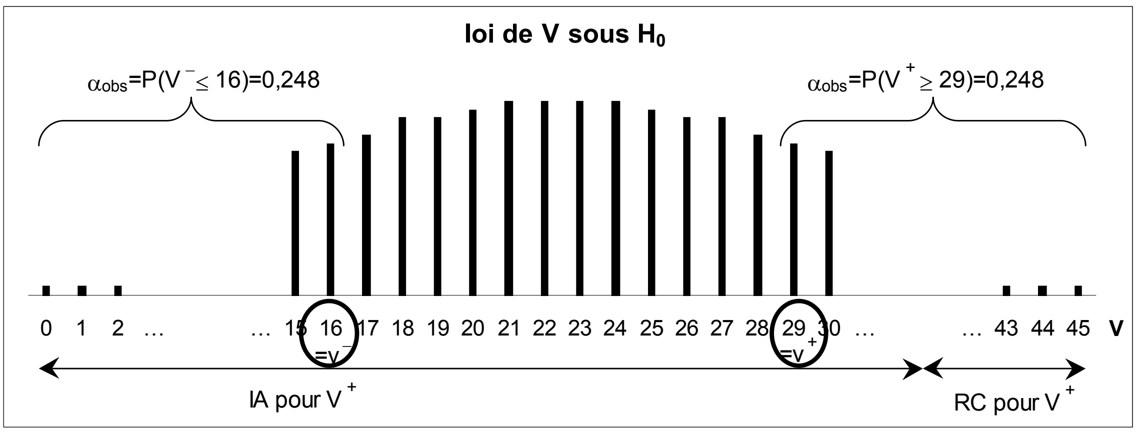
Echantillon de D issu de $\mathcal{P}$ de taille n=10−1=9 (nombre de valeurs de D non nulles)

Statistiques de test V^+ somme des rangs de |D| associés aux valeurs positives de D, et V^- :

somme $V^+ + V^- = 9\times10/2=45$
valeur observée de V^- : v^-=5+7+3+1=16
valeur observée de V^+ : v^+=45−16=8+6+9+2+4=29

loi exacte de la statistique de test V^+ sous l'hypothèse nulle (justifiée car n= 10 <20 et absence d'ex æquo)

sous H_0 : X $\equiv$ Y la statistique V^+ est quantitative discrète sur {0, …, 45} symétrique, de moyenne 45/2=22,5



- degré de signification du test unilatéral droit (région de rejet à droite du domaine de variation de V^+)

le degré de signification est la probabilité sous H_0 d'observer une valeur de la statistique de test V^+ au moins égale à la valeur $v^+=29$ observée car sous l'hypothèse H_1 on attend "beaucoup" de valeurs positives de D avec des rangs élevés, donc de "grandes" valeurs de V^+

$\alpha_{obs} = P_{H_0}(V^+ \geq v^+) = P_{H_0}(V^+ \geq 29) = 1 - P_{H_0}(V^+ \leq 28) = 1 - 0{,}752 = 0{,}248$

car $v=28>n=9$ table 3 colonne $n=9$ et ligne $v-n=28-9=19$

ou $\alpha_{obs} = P_{H_0}(V^+ \geq v^+) = P_{H_0}(V^- \leq v^-) = P_{H_0}(V^- \leq 16) = 0{,}248$

car $v^-=16>n=9$ table 3 colonne $n=9$ et ligne $v^- - n=16-9=7$

- décision

$\alpha_{obs} = 0{,}248 > \alpha=5\%$ donc on ne rejette pas H_0 au seuil $\alpha=5\%$ et au risque β inconnu

- conclusion

la perte de poids n'est pas significativement positive au seuil $\alpha=5\%$ et au risque β inconnu donc l'entraînement intensif n'est pas significativement efficace (est inefficace) sur la perte de poids au seuil $\alpha=5\%$ et au risque β inconnu

❷ Comparaison de deux distributions sur échantillons appariés : test du signe pour deux échantillons appariés

La variable *signe*(D) est qualitative dichotomique définie sur E={positif, négatif}, p proportion de signes strictement positifs **inconnue** dans $\mathcal{P}$,

échantillon de D issu de $\mathcal{P}$ de taille $n=10-1=9$ (nombre de valeurs de D non nulles) : estimation ponctuelle de p par f fréquence observée de signes strictement positifs dans l'échantillon avec $s=5$ donc $f=5/9=0{,}556=55{,}6\%$

Comparaison de la proportion p *inconnue* à la valeur théorique $p_0=50\%=0{,}5$

hypothèse nulle $H_0 : M = M_0=0$ s'écrit $H_0 : p = p_0 = 0{,}5$

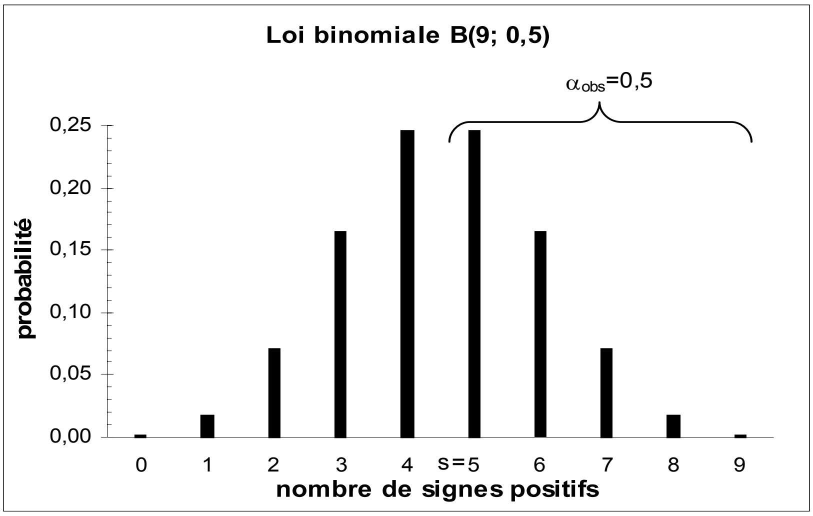
hypothèse alternative $H_1 : M > M_0=0$ s'écrit $H_1 : p > p_0 = 0{,}5$ unilatérale droite, au risque $\alpha=5\%$

- loi de la statistique de test sous l'hypothèse nulle

sous $H_0 : p = p_0 = 0{,}5$

la statistique de test S_n effectif empirique de signes strictement positifs, quantitative discrète définie sur $\{0, \ldots, n\} = \{0, \ldots, 9\}$ suit une loi binomiale $\mathcal{B}(n, p_0) = \mathcal{B}(9; 0{,}5)$ symétrique, de moyenne $np_0 = 9\times0{,}5 = 4{,}5$

	$S_n \sim \mathcal{B}$ (9; 0,5)	
s	$P(S_n=s)$	$P(S_n \leq s)$
0	0,0020	0,0020
1	0,0176	0,0195
2	0,0703	0,0898
3	0,1641	0,2539
4	0,2461	0,5000
5	0,2461	0,7461
6	0,1641	0,9102
7	0,0703	0,9805
8	0,0176	0,9980
9	0,0020	1,0000



- degré de signification du test unilatéral droit (région de rejet à droite du domaine de variation de S_n)

Le degré de signification est la probabilité sous H_0 d'observer une valeur de la statistique de test S_n au moins égale à la valeur $s=5$ observée car sous H_1 on attend "beaucoup" de valeurs de D positives, donc on prendra le risque de conclure H_1 (rejeter H_0 en faveur de H_1) à tort pour les "grandes" valeurs de S_n

loi exacte $\mathcal{B}(9; 0{,}5)$: $\alpha_{obs} = P_{H_0}(S_n \geq s) = P_{H_0}(S_n \geq 5) = P_{H_0}(S_n \leq n-s) = P_{H_0}(S_n \leq 4)$

$= F_{\mathcal{B}}(4) = 0{,}5$

où $F_{\mathcal{B}}$ est la fonction de répartition de la loi $\mathcal{B}(9; 0{,}5)$ (table 1 ligne $n=9$ colonne $k=4$)

- décision

$\alpha_{obs} = 0{,}5 > \alpha=5\%$ donc on ne rejette pas H_0 au seuil $\alpha=5\%$ et au risque β inconnu

- conclusion

la proportion p de signes positifs de la différence D n'est pas significativement supérieure à 50%, au seuil α=5% et au risque β inconnu ;
donc la médiane de la différence des poids avant et après l'entraînement n'est pas significativement supérieure à 0, c'est à dire que l'entraînement intensif n'est pas significativement efficace (est inefficace) sur la perte de poids au seuil α=5% et au risque β inconnu

remarque

dans ce cas, le test de Wilcoxon et celui du signe donnent des résultats très différents bien que les conclusions soient semblables, ce qui s'explique par le fait que le test de Wilcoxon est plus puissant que celui du signe : il a un degré de signification deux fois plus faible (α_{obs} = 0,248) que celui du signe (α_{obs} = 0,5)

Exercice 4.5

$\mathcal{P}$ ={sujets atteints de la maladie de Parkinson} de taille N inconnue,
X = vitesse de marche avant le médicament (en m/s), Y = vitesse de marche après quelques jours de traitement (en m/s), variables quantitatives définies sur $\mathcal{P}$

échantillon de X issu de $\mathcal{P}$ de taille n=21
échantillon de Y issu de $\mathcal{P}$ de taille n=21 $\Rightarrow$ deux échantillons appariés

Le nouveau médicament a-t-il un effet bénéfique sur la marche, au risque α =1% ?

Le nouveau médicament a un effet bénéfique sur la marche se traduit par le fait que les valeurs de X sont globalement inférieures à celles de Y, c'est à dire que la loi de X est à gauche de celle de Y : " $X \prec Y$ "; c'est l'hypothèse que l'on veut accepter à tort au maximum dans 1% des cas, donc l'hypothèse alternative

Le nouveau médicament n'a pas un effet bénéfique sur la marche se traduit par le fait que les valeurs de X sont globalement égales à celles de Y, c'est à dire que la loi de X est identique à celle de Y : " $X \equiv Y$ "; c'est l'hypothèse nulle

On considère D=X−Y la différence des vitesses avant et après traitement, de médiane M inconnue dans $\mathcal{P}$

hypothèse nulle	H_0 : "le traitement n'est pas bénéfique"	H_0 : " $X \equiv Y$ " $\Leftrightarrow$	$H_0 : M = 0$
hypothèse alternative	H_1 : "le traitement est bénéfique"	H_1 : " $X \prec Y$ " $\Leftrightarrow$	$H_1 : M < 0$

unilatérale gauche au risque α=1%

❶ Comparaison de deux distributions sur échantillons appariés : test de Wilcoxon

Echantillon de D issu de $\mathcal{P}$ de taille n=21 (nombre de valeurs de D non nulles ?)

Statistiques de test V^+ somme des rangs de |D| associés aux valeurs positives de D, et V^- :

amélioration des performances physiques c-à-d Y > X d'où D=X−Y < 0 correspond à valeur observée de V^- : v^-=200
dégradation des performances physiques c-à-d Y < X d'où D=X−Y > 0 correspond à valeur observée de V^+ : v^+=31
somme $V^+ + V^-$= 21×22/2=231

① loi exacte de la statistique de test V^+ sous l'hypothèse nulle

sous H_0 : $X \equiv Y$ la statistique V^+ est quantitative discrète sur {0, …, 231} symétrique, de moyenne 231/2=115,5

- degré de signification du test unilatéral gauche (région de rejet à gauche du domaine de variation de V^+)

le degré de signification est la probabilité sous H_0 d'observer une valeur de la statistique de test V^+ au plus égale à la valeur v^+=31 observée car sous l'hypothèse H_1 on attend "peu" de valeurs positives de D avec des rangs faibles, donc de "petites" valeurs de V^+

$\alpha_{obs} = P_{H_0}(V^+ \leq v^+) = P_{H_0}(V^+ \leq 31)$ = ? car les tables 2 et 3 ne donnent pas de valeurs pour n=21>20

② approximation normale (justifiée car n =21 > 20)

- loi de la statistique de test sous l'hypothèse nulle

sous H_0 : " $X \equiv Y$ " les statistiques V^+ et V^- suivent approximativement une loi normale

de moyenne $\frac{n(n+1)}{4}$=231/2=115,5 et de variance $\frac{n(n+1)(2n+1)}{24}$=21×22×43/24=19 866/24=827,75 donc la statistique de test $Z = \dfrac{V^+ - \frac{n(n+1)}{4}}{\sqrt{\frac{n(n+1)(2n+1)}{24}}}$ suit approximativement une loi normale $\mathcal{N}(0,1)$

- valeur observée de la statistique de test : $z = \dfrac{v^+ - \frac{n(n+1)}{4}}{\sqrt{\frac{n(n+1)(2n+1)}{24}}} = \dfrac{31-115{,}5}{\sqrt{827{,}75}} = \dfrac{-84{,}5}{28{,}77} = -2{,}937$

- degré de signification du test unilatéral gauche (région de rejet à gauche)

$\alpha_{obs} = P_{H_0}(Z \le z) = P_{H_0}(Z \le -2{,}937) \approx F_Z(-2{,}94) = (1-F_Z(2{,}94)) = 1-0{,}9984 = 0{,}0016$

où F_Z est la fonction de répartition de la loi normale $\mathcal{N}(0,1)$

- ***décision*** : $\alpha_{obs} = 0{,}0016 < \alpha=1\%$ donc on rejette H_0 en faveur de H_1 au risque $\alpha=1\%$; le test est significatif au degré de signification $\alpha_{obs} = 0{,}0016$

- conclusion

les valeurs de X (vitesse de marche avant traitement) sont globalement inférieures à celles de Y (vitesse de marche après traitement) donc le nouveau médicament a un effet significativement bénéfique sur la marche au risque $\alpha=1\%$ et au degré de signification $\alpha_{obs} = 0{,}0016$

❷ Comparaison de deux distributions sur échantillons appariés : test du signe pour deux échantillons appariés

La variable *signe*(D) est qualitative dichotomique définie sur E={positif, négatif}, *p* proportion de signes strictement positifs **inconnue** dans $\mathcal{P}$,

échantillon de D issu de $\mathcal{P}$ de taille $n=21$: estimation ponctuelle de *p* par *f* fréquence observée de signes strictement positifs dans l'échantillon que l'on ne connaît pas ici : $s=?$

Il est impossible de faire le test du signe avec ces données