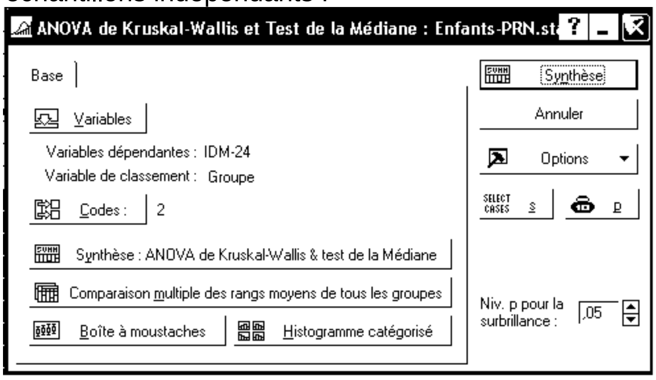
Master de Psychologie sociale des représentations
PSR73C : Informatique - TD N°4

Tests non paramétriques

1 Tests non paramétriques sur des groupes indépendants

1.1 Test de la médiane

Ouvrez le classeur Statistica Enfants-PRN.stw.
On veut comparer l'IDM à 24 mois dans le groupe témoin et dans le groupe expérimental à l'aide d'un test de la médiane.
Rappel de la méthode : on contruit un tableau de contingence en croisant les variables "Groupe" et "Position par rapport à la médiane" et on réalise un test du khi-deux sur le tableau de contingence obtenu.
En utilisant, par exemple, le menu Statistiques - Tests non paramétriques - Statistiques ordinales, vérifiez que la médiane des IDM à 24 mois est égale à 111,5.
Dans le cours, le test de la médiane a été présenté avec une variable "Groupe" à deux modalités. Cependant, la méthode peut s'étendre sans difficultés au cas où la variable "Groupe" comporte plus de deux modalités. C'est pourquoi Statistica range ce test dans le menu : Statistiques - Tests non paramétriques - Comparaison de plusieurs échantillons indépendants :



Spécifiez la variable dépendante et la variable de classement, puis cliquez sur le bouton "Synthèse : ANOVA de Kruskal-Wallis & test de la Médiane". On obtient le résultat suivant :

Dépendant : IDM-24	Test Médiane, Méd. Globale = 111,500; IDM-24 (Enfants-PRN.sta) Var. indépendante (classement) : Groupe Chi-Deux = 3,540645 dl = 1 p = ,0599		
	Témoin	Expérimental	Total
<= Médiane : observ.	19,00	9,00	28,00
théorique	15,50	12,50	
obs.-thé.	3,50	-3,50	
> Médiane : observée	12,00	16,00	28,00
théorique	15,50	12,50	
obs.-thé.	-3,50	3,50	
Total : observé	31,00	25,00	56,00

Remarque : Le test de la médiane ne met pas en évidence de différence entre les deux groupes. En revanche, un test unilatéral de comparaison de moyennes établit une différence au bénéfice du groupe expérimental. Mais le test de la médiane est moins puissant, et c'est nécessairement un test bilatéral.

1.2 Test bilatéral de Kolmogorov-Smirnov

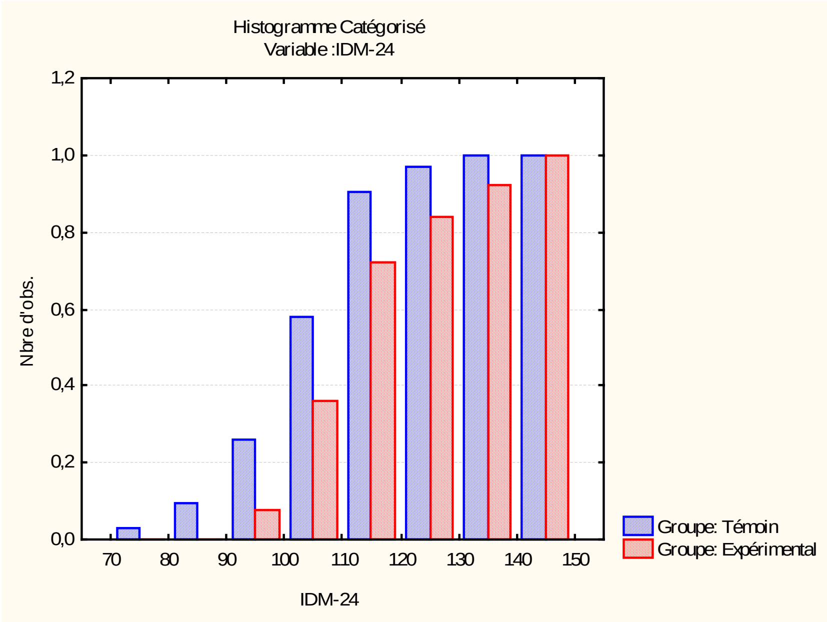
On reprend la comparaison des deux groupes à l'aide du test de Kolmogorov-Smirnov.
Reprenez le menu Statistiques - Tests non paramétriques. Sélectionnez l'item "Comparaison de deux échantillons indépendants". Si nécessaire, spécifiez de nouveau la variable dépendante et la variable de classement, puis cliquez sur le bouton "Test de Kolmogorov-S. de deux échant.".
Vous devriez obtenir le résultat suivant :

	Test de Kolmogorov-Smirnov (Enfants-PRN.sta) Par var. Groupe Tests significatifs marqués à p <,05000								
variable	Max Nég Différnc	Max Pos Différnc	niv. p	Moyenne Témoin	Moyenne Expérimental	Ec-Type Témoin	Ec-Type Expérimental	N Actif Témoin	N Actif Expérimental
IDM-24	-0,403871	0,00	p < .025	106,7097	117,2000	12,95426	12,68201	31	25

On sait que la mise en oeuvre du test de Kolmogorov-Smirnov repose sur le choix d'un découpage en classes, puis la détermination des fonctions de répartition (fréquences cumulées) des deux distributions observées. Il est légitime de se demander quelles sont les bornes de classes utilisées par Statistica.
Quelques manipulations sous Excel montrent que Statistica prend en fait l'ensemble des modalités observées comme bornes de classes, et construit donc le tableau de fréquences cumulées suivant :

	Groupe témoin		Groupe expérimental		
Classes	*Fréquence*	*% cumulé*	*Fréquence*	*% cumulé*	Différence
80	1	3,23%	0	,00%	-3,23%
81	1	6,45%	0	,00%	-6,45%
88	1	9,68%	0	,00%	-9,68%
91	3	19,35%	0	,00%	-19,35%
96	1	22,58%	1	4,00%	-18,58%
98	0	22,58%	1	8,00%	-14,58%
100	1	25,81%	0	8,00%	-17,81%
102	3	35,48%	0	8,00%	-27,48%
104	4	48,39%	0	8,00%	**-40,39%**
106	1	51,61%	1	12,00%	-39,61%
106	0	51,61%	0	12,00%	-39,61%
109	2	58,06%	6	36,00%	-22,06%
111	1	61,29%	0	36,00%	-25,29%
112	0	61,29%	3	48,00%	-13,29%
114	4	74,19%	2	56,00%	-18,19%
116	0	74,19%	1	60,00%	-14,19%
117	0	74,19%	1	64,00%	-10,19%
119	5	90,32%	2	72,00%	-18,32%
123	1	93,55%	0	72,00%	-21,55%
127	1	96,77%	3	84,00%	-12,77%
132	1	100,00%	0	84,00%	-16,00%
137	0	100,00%	2	92,00%	-8,00%
143	0	100,00%	2	100,00%	,00%
Total	31		25		

De plus, il semble que Statistica utilise des tables spécifiques à ce test, et non une approximation par un khi-2.
Il peut être intéressant de visualiser la "distance" entre les deux courbes cumulatives à l'aide d'un graphique. Par exemple, utilisez le bouton "Histogramme catégorisé par groupe" du dialogue obtenu par le menu Statistiques - Tests non paramétriques - Comparaison de deux échantillons indépendants. Avec quelques modifications du graphique, on peut obtenir la représentation suivante :



Remarque.
Le test de Kolmogorov-Smirnov peut être utilisé pour tester soit une hypothèse unilatérale (la VD a une intensité plus grande dans l'un des groupes), soit une hypothèse bilatérale (la distribution de la VD n'est pas la même dans les deux groupes). Comme pour les autres tests, Statistica ne fournit que le test bilatéral.

1.3 Test de Wald-Wolfowitz

Ainsi que nous l'avons vu en cours, le test de Wald-Wolfowitz s'applique à une variable continue, ne comportant pas d'ex-aequo. Son application à des données telles que celles de Enfants-PRN.stw risque donc de réserver quelques surprises... Nous utiliserons donc un autre exemple pour présenter ce test.

Exemple :
Des mesures de pollution organique dans deux rivières ont donné les résultats suivants :

Riv. A	34	12	36	31	43	16	15	10			
Riv. B	65	76	18	27	21	49	20	45	41	17	58

Au vu des valeurs rencontrées, est-il possible que ces pollutions soient dues à une origine commune (c'est-à-dire, est-il possible que ces valeurs soient obtenues par un tirage au hasard dans une même population) ?

Saisissez (sous une forme convenable) ces données dans une feuille de données Statistica.
Réalisez ensuite un test de Wald-Wolfowitz, à l'aide du menu : Statistiques - Tests non paramétriques - Comparaison de deux échantillons indépendants - Test des suites de Wald-Wolfowitz.

On obtient le résultat suivant :

	Test des Suites de Wald-Wolfowitz (Feuille de données25 dans Classeur2) Par var. Lieu Tests significatifs marqués à p <,05000									
Variable	N Actif Riv. A	N Actif Riv. B	Moyenne Riv. A	Moyenne Riv. B	Z	niv. p	Z ajusté	niv. p	Nbe de Suites	Nbe d' ex-aequo
Pollution	8	11	24,6250	39,7273	-2,0674	0,0387	1,8249	0,0680	6	0

Remarquez que la valeur de Z indiquerait un test significatif à 5%, alors que la valeur de "Z ajusté" indique un résultat non significatif. Vu la faible taille des échantillons, c'est ce dernier résultat qui doit être préféré. Les tables spécifiques pour ce test indiquent justement u=6 comme "valeur critique", c'est-à-dire plus grande valeur rendant le test significatif à 5%.

Quels sont les calculs faits par Statistica ?
On peut vérifier que Z= -2.0674 correspond à la formule donnée dans le cours, sans la correction de continuité, tandis que Z= -1,8249 correspond à cette même formule, correction de continuité comprise. En effet :

$$\mu = \frac{2\times8\times11}{8+1} + 1 = 10{,}2632 \quad \text{et} \quad \sigma = \sqrt{\frac{2\times8\times11(2\times8\times11-8-11)}{(8+11)^2(8+11-1)}} = 2{,}0621$$

D'où : $Z = \dfrac{6-10{,}2632}{2{,}0621} = -2{,}0674$ et $Z_{ajusté} = \dfrac{6-10{,}2632+0{,}5}{2{,}0621} = -1{,}8249$.

On peut remarquer également que Statistica ne prend aucune précaution particulière pour traiter les petits échantillons, et que c'est donc à l'utilisateur qu'il appartient d'apprécier si l'approximation par la loi normale est ou non légitime.
Notons enfin que, comme pour tous les autres tests, les niveaux de significativité indiqués correspondent à un test bilatéral.

Les résultats fournis par Statistica comportent une cellule "Nombre d'ex-aequo". En principe, le test des suites s'applique dans des situations où il n'y a pas d'ex-aequo. Il faut également remarquer que Statistica détecte très mal la présence d'ex-aequo, comme le montre le fichier Pollution.stw.

1.4 Protocoles de rangs et test de Wilcoxon Mann Whitney

1.4.1 Le test de Wilcoxon Mann Whitney - Groupes indépendants

On reprend le fichier Enfants-PRN.stw.
La comparaison précédente peut être reprise à l'aide d'un test de Wilcoxon Mann Whitney.
Reprenez le menu Statistiques - Tests non paramétriques. Sélectionnez l'item "Comparaison de deux échantillons indépendants". Si nécessaire, spécifiez de nouveau la variable dépendante et la variable de classement, puis cliquez sur le bouton "Test U de Mann-Whitney". Vous devriez obtenir comme résultat :

	Test U de Mann-Whitney (Enfants-PRN.sta dans Enfants-PRN.stw) Par var. Groupe Tests significatifs marqués à p <,05000									
variable	SommeRgs Témoin	SommeRgs Expérimental	U	Z	niv. p	Z ajusté	niv. p	N Actif Témoin	N Actif)ériment	2*(1-p) p exact
IDM-24	731,5	864,5	235,5	-2,50522	0,012238	-2,51430	0,011927	31	25	0,011429

Statistica nous indique ici trois niveaux de significativité différents : 1,22%, 1,19% et 1,14%. A quoi correspondent ces résultats ?

La première valeur indiquée pour Z, et le premier niveau de significativité indiqué correspondent à la statistique pour "grands échantillons" donnée dans le cours, pour un test bilatéral.

La valeur "Z ajusté" correspond à une statistique Z pour grands échantillons, avec la prise en compte d'une correction pour les ex-aequo.

Le troisième niveau de significativité (0,011429) correspond à l'utilisation de la "vraie" distribution des rangs, sans approximation par une loi normale, mais aussi sans tenir compte des ex-aequo.

1.4.2 Comparaison de la première valeur Z et de la valeur obtenue par la statistique du cours

La statistique calculée par Statistica est-elle la même statistique que celle indiquée en cours ?
Statistica calcule les sommes des rangs W_1 et W_2. On peut vérifier que la valeur Z indiquée (-2,505) correspond bien à la formule du cours :

$$Z = \frac{\overline{R}_1 - \overline{R}_2}{E} \quad \text{avec :} \quad E^2 = \frac{(n_1+n_2+1)(n_1+n_2)^2}{12n_1n_2}$$

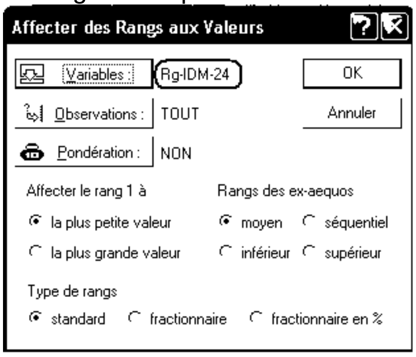
En effet, on a ici : $\overline{R}_1 = \dfrac{731{,}5}{31} = 23{,}60$ et $\overline{R}_2 = \dfrac{864{,}5}{25} = 34{,}58$;

$E^2 = \dfrac{(31+25+1)(31+25)^2}{12\times31\times25} = 19{,}22$, $E = 4{,}38$ et enfin : $Z = \dfrac{23{,}60-34{,}58}{4{,}38} = -2{,}505$

En revanche, Statistica calcule aussi une autre statistique : le U de Mann-Whitney.

1.4.3 Détermination du protocole des rangs et prise en compte des ex aequo

Pour la mise en oeuvre du test de Mann Whitney, la détermination préalable du protocole des rangs n'est pas nécessaire. Cependant, il peut être intéressant de le déterminer pour contrôler, par exemple, que les ex-aequo ne sont pas trop nombreux...
Le menu Données - Affecter les rangs... permet de déterminer le protocole des rangs. Mais, le protocole obtenu remplace le protocole observé à partir duquel il a été déterminé. Si nous voulons conserver à la fois le protocole des rangs et le protocole observé, nous devons au préalable faire une copie de ce dernier.
Insérez une nouvelle variable après la variable IDM-24. Cette nouvelle variable sera nommée Rg-IDM-24.
Recopiez les données de la colonne IDM-24 dans la colonne Rg-IDM-24.
Utilisez le menu Données - Affecter les rangs... en spécifiant comme variable : Rg-IDM-24 :



Au besoin, modifiez les caractéristiques de la variable Rg-IDM-24 de façon que les données s'affichent avec au moins une décimale.
On observe la présence d'assez nombreux ex-aequo dans ce protocole.

Calcul de la correction pour ex-aequo.
La présence d'ex-aequo a pour effet de diminuer la dispersion des données. L'écart type est ajusté à l'aide d'un facteur correctif donné par la formule suivante :

$$\sigma'^2 = \sigma^2\left(1 - \frac{1}{N\left(N^2-1\right)}\sum_j \left(t_j^3 - t_j\right)\right)$$

où N désigne le nombre total d'observations, la somme comporte autant de termes que de "paquets" d'ex aequo, et pour un paquet donné, t_j désigne le nombre d'observations rassemblées dans le paquet.

Dans notre exemple, on dénombre 4 paquets de 2 ex aequo, 3 paquets de 3 ex aequo, 2 paquets de 4 ex aequo, et un paquet de 6, un paquet de 7 et un paquet de 8 ex aequo. Le calcul du facteur correctif donnera donc :

$$\sigma'^2 = \sigma^2\left(1 - \frac{1}{56\left(56^2-1\right)}\left(4\times\left(2^3-2\right)+3\times\left(3^3-3\right)+2\times\left(4^3-4\right)+\left(6^3-6\right)+\left(7^3-7\right)+\left(8^3-8\right)\right)\right)$$

Calcul fait, on obtient : $\frac{\sigma'^2}{\sigma^2} = 0{,}992789$ d'où $\frac{\sigma'}{\sigma} = 0{,}996388$ et $Z_{ajusté} = -\frac{2{,}505220}{0{,}996388} = -2{,}514302$.

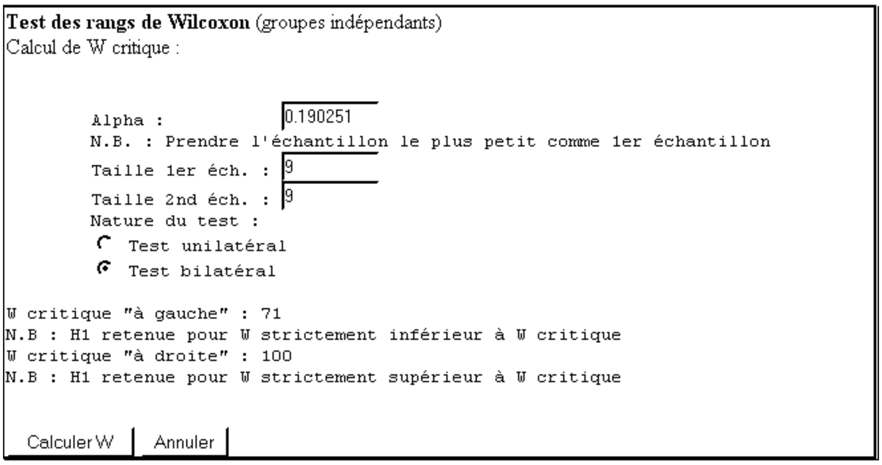
On constate que la valeur trouvée correspond bien à celle indiquée par Statistica. Ce calcul montre également que, si les échantillons sont de taille suffisante, l'effet des ex aequo est assez limité, même dans le cas où ceux-ci sont assez nombreux.

1.4.4 Le test de Mann Whitney sur de petits échantillons

On considère les données contenues dans le classeur : Delinquance-Juvenile.stw
On compare les deux groupes "Maison des parents" et "Foyer".
Réalisez un test de Mann-Whitney : vous devriez obtenir le résultat suivant :

variable	Test U de Mann-Whitney (Délinquance dans Delinquance-Juvenile.stw) Par var. Domicile Tests significatifs marqués à p <,05000									
	SommeRgs Maison des parents	SommeRgs Foyer	U	Z	niv. p	Z ajusté	niv. p	N Actif Maison des parents	N Actif Foyer	2*(1-p) p exact
Absentéisme	101	70	25	1,37	0,17	1,37	0,17	9	9	0,190251

Compte tenu de la faible taille des échantillons, ce n'est pas le niveau de la statistique Z qu'il faut ici prendre en compte, mais la valeur indiquée dans la dernière colonne : 2*(1-p) - p exact.
Pour vérifier cette valeur, on peut se servir des "tables statistiques en ligne" accessibles à l'adresse http://geai.univ-brest.fr/~carpenti/statistiques/table1.php :


Test des rangs de Wilcoxon (groupes indépendants)
Calcul de W critique :

Alpha : 0.190251
N.B. : Prendre l'échantillon le plus petit comme 1er échantillon
Taille 1er éch. : 9
Taille 2nd éch. : 9
Nature du test :
Test unilatéral
Test bilatéral

W critique "à gauche" : 71
N.B : H1 retenue pour W strictement inférieur à W critique
W critique "à droite" : 100
N.B : H1 retenue pour W strictement supérieur à W critique

Calculer W | Annuler

Pour le niveau de significativité calculé par Statistica, les valeurs indiquées par les tables en ligne (calculées par le logiciel de Statistiques R) sont compatibles avec les sommes de rangs observées.
En revanche, on remarquera que Statistica ne fait pas de correction pour tenir compte des ex-aequo. D'autres logiciels (Statgraphics, Minitab) font cette correction, et affichent W=56, avec un niveau de significativité de 0,1844.

Exercice : Procéder de même pour effectuer les deux autres comparaisons de groupes pris deux à deux. La seule comparaison qui nous conduit à accepter l'hypothèse alternative est la troisième : les enfants placés en foyer sont moins souvent absents que les enfants placés en famille adoptive.

1.5 Test de Kruskal-Wallis

1.5.1 Exemple 1

Les données contenues dans le classeur Delinquance-Juvenile.stw concernent trois groupes indépendants. La comparaison globale de ces trois groupes peut être réalisée à l'aide d'un test de Kruskal-Wallis ou d'un test de la médiane.

Utilisez le menu Statistiques - Tests non paramétriques - Comparaison de plusieurs échantillons indépendants.
Sélectionnez Absentéisme comme variable dépendante, et Domiciel comme variable de classement.
Vous devriez obtenir les résultats suivants :

Dépend. : Absentéisme	ANOVA de Kruskal-Wallis par Rangs; Absentéisme (Délinquance dans Delinqu Var. indépendante (classement) : Domicile Test de Kruskal-Wallis : H (2, N= 27) =6,784418 p =,033		
	Code	N Actifs	Somme Rangs
Maison des parents	101	9	124,5000
Foyer	102	9	83,0000
Famille adoptive	103	9	170,5000

Dépendant : Absentéisme	Test Médiane, Méd. Globale = 13,0000; Absentéisme (Délinquance dans D Var. indépendante (classement) : Domicile Chi-Deux = 5,637363 dl = 2 p = ,0597			
	Maison des parents	Foyer	Famille adoptive	Total
<= Médiane : observ.	5,000000	7,00000	2,00000	14,00000
théorique	4,666667	4,66667	4,66667	
obs.-thé.	0,333333	2,33333	-2,66667	
> Médiane : observée	4,000000	2,00000	7,00000	13,00000
théorique	4,333333	4,33333	4,33333	
obs.-thé.	-0,333333	-2,33333	2,66667	
Total : observé	9,000000	9,00000	9,00000	27,00000

On voit que le test de Kruskal-Wallis conduit à un résultat significatif au seuil de 5%, alors que le test de la médiane ne met pas en évidence de différence entre les groupes. En effet, le test de la médiane est moins puissant que celui de Kruskal-Wallis.

1.5.2 Exemple 2

Ouvrez la feuille de données Kruskal.sta.
Cette feuille de données est l'un des exemples fournis avec Statistica. La présentation de ces données est la suivante :

Cet exemple est basé sur un ensemble de données (fictives) reprises de (Hays, 1981, p. 592).

De jeunes enfants ont été affectés au hasard dans trois groupes expérimentaux. On montre à chaque enfant une série de paires de stimuli. Sa tâche consiste à choisir l'un de ces stimuli et si le choix est "correct", il reçoit une récompense. Dans l'un des groupes, le critère déterminant le choix correct est la forme (groupe 1 - Forme), dans le second groupe, le critère pertinent est la couleur (groupe 2 - Couleur) et dans le troisième groupe, le critère pertinent est la taille (groupe 3 - Taille). La variable dépendante est le nombre d'essais réalisés par l'enfant pour détecter le choix qui sera récompensé.

Réalisez un test de Kruskal-Wallis sur ces données. Vous obtenez le résultat suivant :

Dépend. : PERFRMNC	ANOVA de Kruskal-Wallis par Rangs; PERFRMNC (Kruskal.sta) Var. indépendante (classement) : CONDITN Test de Kruskal-Wallis : H (2, N= 36) =13,84438 p =,001		
	Code	N Actifs	Somme Rangs
FORME	1	12	139,0000
COULEUR	2	12	200,0000
TAILLE	3	12	327,0000

Dépendant : PERFRMNC	Test Médiane, Méd. Globale = 19,5000; PERFRMNC (Kruskal.sta) Var. indépendante (classement) : CONDITN Chi-Deux = 8,666667 dl = 2 p = ,0131			
	FORME	COULEUR	TAILLE	Total
<= Médiane : observ.	9,00000	7,00000	2,00000	18,00000
théorique	6,00000	6,00000	6,00000	
obs.-thé.	3,00000	1,00000	-4,00000	
> Médiane : observée	3,00000	5,00000	10,00000	18,00000
théorique	6,00000	6,00000	6,00000	
obs.-thé.	-3,00000	-1,00000	4,00000	
Total : observé	12,00000	12,00000	12,00000	36,00000

Autrement dit, les résultats des deux tests sont significatifs. On constate encore que le test de la médiane est moins puissant que le test de Kruskal-Wallis.

2 Tests non paramétriques sur des groupes appareillés

2.1 Test du khi-2 de Mac Nemar

On reprend l'exemple traité en cours : reconnaissance d'une série de portraits à deux semaines et à 1 an. Les observations sont résumées par le tableau suivant :

		1 an	
		Reconnu	Non reconnu
Deux semaines	Reconnu	81	46
	Non reconnu	8	49

Utilisez le menu Statistiques - Tests non paramétriques - Tables 2x2.
Indiquez les effectifs ci-dessus dans la fenêtre de dialogue et cliquez sur le bouton "Synthèse".
Statistica nous sert en vrac différents résultats : khi-deux "classique", phi-deux, khi-deux de Mac Nemar... A nous de savoir choisir le résultat qui nous intéresse (et qui a un sens par rapport à nos données) :

	Table 2 x 2 (Feuille de données59)		
	Colon. 1	Colon. 2	Totaux Bruts
Effectifs, ligne 1	81	46	127
%age du total	44,022%	25,000%	69,022%
Effectifs, ligne 2	8	49	57
%age du total	4,348%	26,630%	30,978%
Totaux colonne	89	95	184
%age du total	48,370%	51,630%	
Chi-deux (dl=1)	38,98	p= ,0000	
V-deux (dl=1)	38,77	p= ,0000	
Chi ² corrigé de Yates	37,02	p= ,0000	
Phi-deux	,21186		
p exact Fisher, unilatéral		p= ,0000	
bilatéral		p= ,0000	
Chi ² de McNemar (A/D)	**7,39**	**p= ,0066**	
Chi-deux (B/C)	25,35	p= ,0000	

2.2 Test du signe - Groupes appareillés

On reprend le classeur Enfants-PRN.stw et on se propose de comparer l'IDM à 6 mois et l'IDM à 24 mois dans le groupe témoin.
On veut essayer de montrer que le nombre de différences négatives est significativement grand, ou, de manière symétrique, que le nombre de différences positives est suffisamment faible pour montrer une baisse de l'IDM entre 6 et 24 mois, dans la population dont est tiré l'échantillon.
On va donc utiliser un test du signe pour comparer les scores des enfants du groupe témoin à 6 mois et à 24 mois.
Utilisez le menu Statistiques - Tests non paramétriques - Comparaison de deux échantillons appariés.
Indiquez IDM-6 et IDM-24 comme variables et cliquez sur le bouton "Test des signes".
Vous devriez obtenir le résultat suivant :

	Test des Signes (Enfants-PRN.sta) Tests significatifs marqués à p <,05000			
Couple de variables	Nbe Non ex-aequo	%age v < V	Z	niv. p
IDM-6 & IDM-24	31	38,70968	1,077632	0,281198

Statistica nous indique que 38,71% des paires sont telles que IDM-6 est inférieur à IDM-24. Il calcule l'approximation par une loi normale donnée par :

$$Z = \frac{2D - 1 - N}{\sqrt{N}} \quad \text{où} \quad D = Max(D_+, D_-)$$

et indique que le niveau de significativité de cette statistique est de 28% pour un test bilatéral.

Conclusion : on n'a pas démontré de différence significative entre l'IDM à 6 mois et l'IDM à 24 mois pour la population d'où a été tiré l'échantillon d'enfants du groupe témoin.

Remarques.
1. Nos données comprennent 31 observations pour IDM-6 (le groupe témoin seul), mais 56 pour IDM-24 (groupe témoin et groupe expérimental). Remarquez que Statistica réalise le test en ne considérant que les 31 paires "complètes" : les valeurs manquantes sont ignorées.
2. Statistica ne prévoit ici aucune procédure pour traiter le cas des petits échantillons, et l'aide renvoie à l'ouvrage de Siegel et Castellan pour traiter les cas où n < 20...

2.3 Le test de Wilcoxon - Groupes appareillés

2.3.1 Le test des rangs signés de Wilcoxon

La comparaison des scores IDM-6 et IDM-24 peut également être effectuée à l'aide d'un test de Wilcoxon (test des rangs signés).
Utilisez le menu Statistiques - Tests non paramétriques - Comparaison de deux échantillons appariés.
Indiquez IDM-6 et IDM-24 comme variables et cliquez sur le bouton "Test de Wilcoxon, échantillons appariés".
Vous devriez obtenir le résultat suivant :

	Test de Wilcoxon pour Ech. Appariés (Enfants-PRN.sta) Tests significatifs marqués à p <,05000			
Couples de variables	N Actifs	T	Z	niv. p
IDM-6 & IDM-24	31	177,5000	1,381556	0,167109

On vérifie que la statistique calculée par Statistica est :

$$Z = \frac{T - \frac{N(N+1)}{4}}{E} \quad \text{avec} \quad E^2 = \frac{N(N+1)(2N+1)}{24}$$

où T est le maximum des deux sommes de rangs des différences positives et négatives. Cette statistique est pratiquement celle donnée en cours.
Il faut également remarquer que, comme précédemment :
- Il n'est tenu compte des ex-aequo : Statgraphics, qui introduit une correction pour les ex-aequo, trouve Z=1,48944 au lieu de 1,3815.
- Il n'est pas prévu de procédure pour traiter le cas des petits échantillons

Remarque.
Le traitement des différences nulles est particulièrement mal pris en compte par le test de Wilcoxon mis en oeuvre dans Statistica. On pourra, par exemple, reprendre le fichier de données précédent et modifier les données en introduisant de plus en plus de différences nulles. Celles-ci ne seront éliminées (N actifs inférieur à 31) que dans de rares occasions.

2.3.2 Calcul du protocole des rangs signés

Il est possible de calculer le protocole des rangs signés à l'aide de Statistica. Pour cela :
- Ajoutez 4 colonnes supplémentaires en fin de tableau de données. Ces colonnes pourront être nommées : Diff, RgDiffAbs, RgPlus, RgMoins.
- Pour la colonne Diff, introduisez la formule : = V3-V4
- Pour la colonne RgDiffAbs, introduisez la formule =abs(Diff), puis transformez les valeurs en rangs.
- Pour la colonne RgPlus, introduisez la formule : = iif(Diff >0; RgDiffAbs; -9999)
- Pour la colonne RgMoins, introduisez la formule : = iif(Diff <0; RgDiffAbs; -9999).

Remarques.
1. La formule de calcul de la colonne RgDiffAbs reste mémorisée avec la feuille de données. Si on doit demander un recalcul des autres colonnes, il faudra éviter que le recalcul concerne cette colonne, ou refaire le calcul des rangs.
2. Remarquez l'utilisation de la valeur -9999 comme code pour les valeurs manquantes.
3. Le calcul précédent est correct en l'absence de différences nulles. S'il y avait des différences nulles, la formule de la colonne RgDiffAbs devrait être remplacée par : = iif(Diff<>0;abs(Diff);-9999)

Exercice
Ouvrez le fichier Wilcoxon.xls à l'aide d'Excel.

Recopiez les données dans une nouvelle feuille de données Statistica et faites un test de Wilcoxon pour étudier s'il existe une différence significative entre l'ainé et le cadet du point de vue de la variable étudiée.
Vous devriez obtenir une statistique Zobs égale à 2,18, et donc conclure à une différence significative entre les deux conditions.

2.4 Test Q de Cochran

Nous allons illustrer la mise en oeuvre de ce test à l'aide de l'exemple figurant dans la fiche de TD de statistiques. Cet exemple, donné par Siegel et Castellan, est également fourni avec Statistica. On peut en donner l'énoncé suivant :

Supposez que vous souhaitez connaître l'effet du style d'entretien sur le nombre de sondés qui acceptent de répondre à des questions personnelles lors d'enquêtes en face-à-face. Par exemple, il est toujours plus difficile d'obtenir des réponses concernant les finances personnelles ou la santé dans ce type de sondage, et vous pouvez vouloir apprendre si le style d'entretien améliore le taux de réponses (non-refus) de telles questions personnelles. Dans ce but, vous pouvez apprendre aux sondeurs à diriger l'entretien soit (1) de manière enthousiaste, en étant amical et en montrant de l'intérêt (Enquêteur 1), (2) d'une façon très réservée et formelle (Enquêteur 2), ou (3) de façon désintéressée et abrupte (Enquêteur 3). Vous pouvez ensuite sélectionner 18 jeux de trois chefs de famille dont les réponses à de précédents sondages répondent à certains critères, puis assigner de façon aléatoire les trois chefs de famille à l'un des types d'entretien. Comme variable dépendante, vous enregistrez si le sondé respectif répond (1-Oui) ou non (0-Non) aux questions personnelles. Les résultats de cette étude sont enregistrés dans le fichier de données Interview.sta.

Vérifiez les codes attribués aux modalités OUI et NON dans le fichier de données.
Utilisez le menu Statistiques - Tests non paramétriques - Comparaison de plusieurs échantillons appariés.
Spécifiez ENQUETEUR_1, ENQUETEUR_2 et ENQUETEUR_3 comme variables et cliquez sur le bouton "Synthèse : Test Q de Cochran".
Vous obtenez le résultat suivant :

	Test Q de Cochran (Interview.sta) Nbre d'obs. actives :18 Q = 16,66667, dl = 2, p < ,000240		
Variable	Somme	%age 0	%age 1
ENQUETEUR_1	13,00000	27,77778	72,22222
ENQUETEUR_2	13,00000	27,77778	72,22222
ENQUETEUR_3	3,00000	83,33333	16,66667

Le style d'interview a donc un effet sur le consentement de la personne interviewée à répondre à des questions d'ordre personnel.

2.5 Test de Friedman

Comme précédemment, nous travaillerons sur l'exemple fourni par Statistica. Il s'agit ici des données contenues dans le fichier Mothers.sta.

L'énoncé est le suivant :

L'exemple suivant se base sur un fichier de données reporté par Siegel (1956, page 233). Vingt mères et leurs enfants mal-entendants ont participé à un séminaire pour apprendre à s'occuper de leur enfant. À la fin du séminaire, les 13 personnes de l'équipe organisatrice ont noté les 20 mères en fonction de la probabilité qu'avait la mère respective d'élever son enfant d'une façon qui serait nocive à son développement. Les données sont contenues dans le fichier de données Mothers.sta.

But de l'analyse. Cette analyse poursuit deux buts :

(1) Selon les jugements portés, y a-t-il une différence significative entre les mères du point de vue de leur capacité à élever leur enfant ? Cette question relève de l'ANOVA par les rangs de Friedman.

(2) Faut-il faire confiance aux jugements portés ? Autrement dit, y a-t-il une corrélation entre les scores attribués par les juges. Si ce n'est pas le cas, il n'est guère possible d'accorder une grande confiance dans leurs affirmations. Cette question peut être traitée en utilisant le coefficient de concordance de Kendall.

Utilisez le menu Statistiques - Tests non paramétriques - Comparaison de plusieurs échantillons appariés. Spécifiez MERE_1 à MERE_20 comme variables et cliquez sur le bouton "Synthèse : ANOVA de Friedman et concordance de Kendall".

N.B. Les 13 juges forment ici un échantillon de taille 13 dans la population à étudier. Les 20 mères, quant à elles, constituent 20 conditions différentes auxquelles les juges sont soumis.

On obtient les résultats suivants :

ANOVA de Friedman & Coef. de Concord. de Kendall (Mothers.sta)
ANOVA du Chi² (N = 13, dl = 19) = 139,9814 p <0,00000
Coeff. de Concordance = ,56673 Rang moy. r = ,53062

Variable	Rang Moyen	Somme Rangs	Moyenne	Ec-type
MERE_1	7,26923	94,5000	7,15385	4,597937
MERE_2	4,69231	61,0000	4,69231	5,266001
MERE_3	11,92308	155,0000	11,69231	4,230536
MERE_4	6,69231	87,0000	6,69231	3,351234
MERE_5	8,57692	111,5000	8,46154	3,755338
MERE_6	4,38462	57,0000	4,38462	3,150092
MERE_7	8,00000	104,0000	7,92308	3,988766
MERE_8	11,07692	144,0000	10,92308	3,475187
MERE_9	7,19231	93,5000	7,07692	3,174417
MERE_10	5,92308	77,0000	5,84615	3,436232
MERE_11	14,03846	182,5000	13,84615	3,262058
MERE_12	8,53846	111,0000	8,46154	3,017046
MERE_13	5,53846	72,0000	5,46154	5,767949
MERE_14	10,84615	141,0000	10,69231	5,344540
MERE_15	12,65385	164,5000	12,46154	4,858247
MERE_16	14,61538	190,0000	14,38462	2,256046
MERE_17	17,46154	227,0000	17,23077	3,192539
MERE_18	15,61538	203,0000	15,38462	3,042435
MERE_19	16,53846	215,0000	16,30769	2,626297
MERE_20	18,42308	239,5000	18,23077	4,085622

Comme on peut le voir, il existe des différences très significatives entre les mères. De plus, le coefficient de concordance de Kendall montre un accord satisfaisant entre les différents jugements portés (0,57, alors que la moyenne des coefficients obtenus en faisant varier le protocole est de 0,53).

3 Exercice à rendre par mail

Réf. Snehendu B. Kar, Catherine A. Pascual, Kirstin L. Chickering, Empowerment of women for health promotion: a meta-analysis, Social Science & Medecine, 49 (1999), pp. 1431-1460.
(accessible via le site de la BU, dans le bouquet de revues Science Direct d'Elsevier)

L'"empowerment" est défini comme un processus au cours duquel des individus, des communautés et des organisations prennent le contrôle de situations et de problèmes dans lesquels ils sont particulièrement impliqués. Dans l'article cité supra les auteurs ont analysé 40 cas d'étude de mouvements réussis d'"empowerment", pris en charge par des femmes, en provenant de pays développés et de pays moins développés. Parmi les cas étudiés : le mouvement des Mères de la Plaza de Mayo durant la dictature Argentine, le mouvement des femmes contre la violence par armes à feu aux Etats-Unis, etc.

Les auteurs ont identifié 7 méthodes utilisées dans ces processus : (1) entraînement à la prise de contrôle et au leadership, (2) utilisation des médias, (3) éducation publique et participation, (4) organisation de partenariats, associations, coopératives, (5) formation à un travail et micro-entreprises, (6) développement de services et d'assistance, (7) protection des droits et action sociale.

Pour chacun des 40 cas étudiés, les méthodes qui ont été utilisées sont indiquées dans le tableau suivant :

Méthode	Cas	
	Pays industrialisés	Pays moins industrialisés
1	8, 20, 23, 31, 33, 38	5, 6, 13, 15, 17, 19, 22, 24, 25, 27, 34, 35
2	8, 10, 11, 12, 18, 20, 23, 26, 30, 31, 32, 36, 38	1, 4, 5, 7, 16, 34, 35
3	8, 10, 11, 12, 18, 20, 23, 26, 31, 36, 37, 38, 39	3, 4, 5, 6, 7, 16, 19, 27, 34, 35
4	8, 10, 11, 18, 31, 32	2, 7, 13, 14, 15, 16, 17, 19, 21, 24, 25, 28, 29, 35
5	30	13, 14, 15, 17, 19, 24, 25
6	8, 10, 12, 20, 23, 26, 30, 33, 36, 37, 38, 39	1, 4, 5, 6, 7, 9, 13, 16, 17, 19, 21, 22, 24, 25, 27, 28, 29, 34, 35
7	8, 10, 11, 12, 18, 20, 23, 26, 31, 32	1, 2, 3, 5, 6, 7, 9, 13, 17, 22, 25, 29, 34, 40

Les auteurs utilisent un test Q de Cochran pour déterminer si la fréquence d'utilisation des sept méthodes varie significativement selon le cas étudié.
Saisissez les données sous une forme convenable dans Statistica et retrouvez le résultat indiqué par les auteurs.

Reprenez le même test pour les pays industrialisés d'une part, pour les pays moins industrialisés d'autre part.

Pour chacune des 7 méthodes, étudiez ensuite à l'aide du test du khi-2, si son usage est lié au niveau d'industrialisation.