Master de Psychologie
PSY73B : Informatique : traitement des données - TD N°3
Tests non paramétriques

15. Tests non paramétriques sur des groupes indépendants

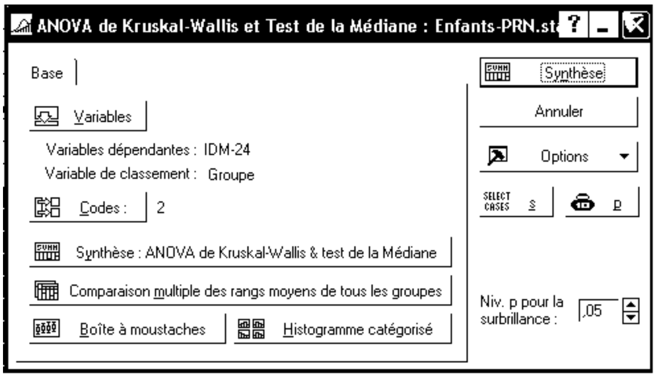
15.1. Test de la médiane

Objectif du test : comparer les médianes dans deux ou plusieurs groupes indépendants, lorsque la variable dépendante est ordinale ou numérique.

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 construit 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 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.

15.2. Test bilatéral de Kolmogorov-Smirnov

Objectif du test : comparer les distributions de la variable dépendante dans deux ou plusieurs groupes indépendants, lorsque la variable dépendante est ordinale ou numérique.

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 :

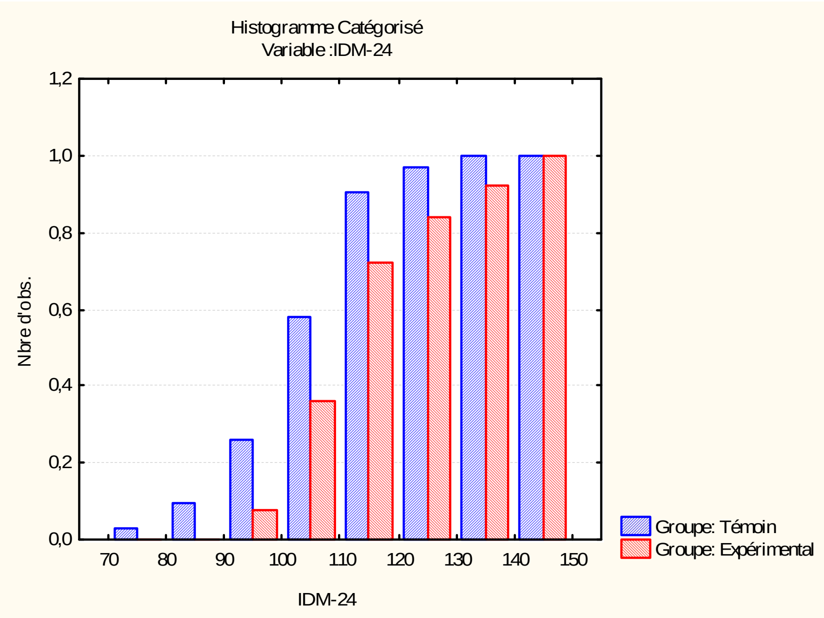
	Test de Kolmogorov-Smirnov (Enfants-PRN.sta) Par var. Groupe Tests significatifs marqués à p <,05000								
variable	Max Nég Différnc	Max Pos Différnc	niv. p	Moyenne Témoin	Moyenne Expérimental	Ec-Type Témoin	Ec-Type Expérimental	N Actif Témoin	N Actif Expérimental
IDM-24	-0,403871	0,00	p < .025	106,7097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



Modifications à faire à partir du graphique produit avec les réglages par défaut de Statistica : à l'aide du bouton droit de la souris, sélectionnez l'item de menu Propriétés du graphique (Toutes les options)... puis :
- Sous l'onglet Tracé -- Histogramme, sélectionnez Représentation de l'histogramme : cumulé
- Sous l'onglet Tracé -- Histogramme, dans la zone Propriétés, cochez la boîte Effectifs relatifs cumulés
- Sous l'onglet Catégorisation, dans la zone Mise en forme des catégories, sélectionnez Superposées
- Sous l'onglet Tracé -- Ajustement, cliquez sur le bouton Supprimer de la zone Type d'ajustement.

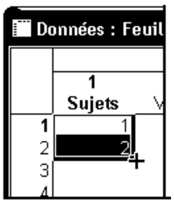
Remarque.
Le test de Kolmogorov-Smirnov peut être utilisé pour tester soit une hypothèse unilatérale (la VD a une intensité plus grande dans l'un des groupes), soit une hypothèse bilatérale (la distribution de la VD n'est pas la même dans les deux groupes). Comme pour les autres tests, Statistica ne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

15.3. Test de Wald-Wolfowitz

Objectif du test : comparer les distributions de la variable dépendante dans deux ou plusieurs groupes indépendants, lorsque la variable dépendante, ordinale ou numérique, ne comporte pas d'ex aequo. Ce test étudie notamment si l'interclassement des valeurs issues des deux groupes peut être dû au hasard.

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.

En vue d'une expérience, vous avez recruté 50 sujets que vous devez affecter au hasard à l'une ou l'autre de deux conditions expérimentales. Pour cela :
- Vous saisissez les identifiants des 50 sujets dans une colonne d'une feuille de données Statistica (par exemple, vous utilisez comme identifiants les nombres de 1 à 50).
Définissez un nouveau classeur Statistica, insérez dans ce classeur une feuille de données comportant 50 lignes. Dans la première variable, appelée Sujets, saisissez les valeurs 1 et 2, puis utilisez la souris pour réaliser une copie incrémentée :



- Vous générez dans une deuxième colonne une suite de nombres aléatoires compris entre 0 et 1 (menu Edition - Remplir Centrer-réduire le bloc - Remplir de valeurs aléatoires.

- Vous affectez ensuite chacun des sujets à l'une ou l'autre des conditions expérimentales A et B en comparant à 0,5 le nombre aléatoire correspondant.
Pour cela, créez une troisième variable, calculée à l'aide de la formule :

=iif(v2<0,5;"A";"B")

Vous obtenez ainsi une feuille ayant l'allure suivante :

	1 Sujets	2 Aleas	3 Condition
1	1	0,948245	B
2	2	0,900394	B
3	3	0,465691	A
4	4	0,167943	A
5	5	0,990672	B
6	6	0,984044	B

N.B. Les résultats qui suivent dépendent des valeurs aléatoires générées par le logiciel. Ils ne sont donc pas strictement reproductibles.

Vous vous demandez alors si l'affectation des sujets aux 2 conditions s'est bien faite de façon aléatoire (le générateur de nombres aléatoires utilisé est-il correct ?). Deux questions peuvent se poser :

- Les effectifs des groupes A et B sont-ils compatibles avec l'hypothèse d'une affectation "au hasard" (p=0,5) des sujets dans les groupes ?

Il n'est pas si simple de traiter cette question à l'aide de Statistica. Par exemple, on pourra :

- Remarquer que les étiquettes de texte A et B sont associées aux valeurs numériques 101 et 102 (dans un ordre quelconque) ;
- Définir une 4è variable, contenant la constante 101,5
- Réaliser un test des signes entre la variable "Condition" et la variable "Constante". (menu Statistiques - Tests non paramétriques - Comparaison de deux échantillons appariés)

	Test des Signes (Affectation-Sujets) Tests significatifs marqués à p <,05000			
Couple de variables	Nbe Non ex-aequo	%age v < V	Z	niv. p
Condition & Constante	50	38,00	1,56	0,12

On constate qu'ici, le déséquilibre entre les groupes (19 sujets dans l'un des groupes pour 31 dans l'autre) est compatible avec l'hypothèse d'une affectation au hasard, au seuil de 5%.

Remarque : On peut aussi recenser les valeurs de la variable "Condition" à l'aide du menu Statistiques - Statistiques élémentaires - Tables de fréquences :

	Table de fréquences : Condition: =iif(v2<0,5;"A";"B'			
Catégorie	Effectif	Effectifs Cumulés	%age	%age Cumulé
B	19	19	38,00000	38,0000
A	31	50	62,00000	100,0000
VM	0	50	0,00000	100,0000

puis utiliser le menu Statistiques - Tests non paramétriques - Tables 2x2, en indiquant les effectifs 19 et 31 dans les cases A/D ou B/C du tableau de contingence. On lit alors le résultat comme valeur du Chi-deux de Mac Nemar :

	Table 2 x 2 (Affectation-Sujets dans Affectation-Sujets.stw		
	Colon. 1	Colon. 2	Totaux Bruts
Effectifs, ligne 1	19	0	19
%age du total	38,000%	0,000%	38,000%
Effectifs, ligne 2	0	31	31
%age du total	0,000%	62,000%	62,000%
Totaux colonne	19	31	50
%age du total	38,000%	62,000%	
Chi ² de McNemar (A/D)	**2,42**	**p= ,1198**	
Chi-deux (B/C)	----	----	

- Les valeurs aléatoires générées sont-elles indépendantes entre elles : l'alternance de l'affectation à l'un ou l'autre groupe est-elle due au hasard ?

Là, c'est le test de Wald-Wolfowitz qui nous permet de répondre à cette interrogation. Utilisez le menu Statistiques - Tests non paramétriques - Comparaison de deux échantillons indépendants - Test des suites de Wald-Wolfowitz en spécifiant "Sujets" comme variable dépendante et "Condition" comme variable de classement. Vous obtenez, par exemple :

	Test des Suites de Wald-Wolfowitz (Affectation-Sujets dans Classeur8) Par var. Condition Tests significatifs marqués à p <,05000									
Variable	N Actif B	N Actif A	Moyenne B	Moyenne A	Z	niv. p	Z ajusté	niv. p	Nbe de Suites	Nbe d' ex-aequo
Sujets	19	31	26,32	25,00	-0,78	0,44	0,63	0,53	22	0

Quels sont les calculs faits par Statistica ?

On peut vérifier que Z= -0,78 correspond à la formule donnée dans le cours, sans la correction de continuité, tandis que Z ajusté = 0,63 correspond à cette même formule, correction de continuité comprise. En effet :

$$\mu = \frac{2\times 19\times 31}{19+31}+1 = 24{,}56 \text{ et } \sigma = \sqrt{\frac{2\times 19\times 31\times(2\times 19\times 31-19-31)}{(19+31)^2(19+31-1)}} = 3{,}2935$$

D'où : $Z = \frac{22-24{,}56}{3{,}2935} = -0{,}7772$ et $Z_{ajusté} = \frac{24{,}56-22-0{,}5}{3{,}2935} = 0{,}6254$.

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.

15.4. Protocoles de rangs et test de Wilcoxon Mann Whitney

Objectif du test : comparer les médianes de la variable dépendante dans deux groupes indépendants, lorsque la variable dépendante est ordinale ou numérique.

15.4.1 Le test de Wilcoxon Mann Whitney - Groupes indépendants

On reprend le fichier Enfants-PRN.stw.

La comparaison faite précédemment à l'aide d'un test de la médiane 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". En utilisant Statistica 10, 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	valeur p	Z ajusté	valeur p	N Actif Témoin	N Actif Expérimental	2*(1-p) p exact
IDM-24	731,5	864,5	235,5	-2,49698	0,012526	-2,50603	0,012210	31	25	0,011429

Statistica nous indique ici trois niveaux de significativité différents : 1,25%, 1,22% et 1,14%. A quoi correspondent ces résultats ?

La première valeur indiquée pour Z, et le premier niveau de significativité indiqué correspondant (presque exactement) 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.

15.4.2 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 (-2,4993) correspond à la formule du cours à laquelle on a appliqué une correction de continuité au numérateur :

$$Z = \frac{\overline{R}_1 - \overline{R}_2 \pm 0{,}5\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}{E} \quad \text{avec :} \quad E^2 = \frac{(n_1 + n_2 + 1)(n_1 + n_2)^2}{12 n_1 n_2}$$

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

$E^2 = \frac{(31+25+1)(31+25)^2}{12 \times 31 \times 25} = 19{,}2206$, $E = 4{,}3841$ et enfin :

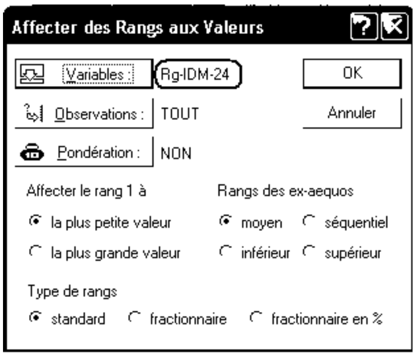
$$Z = \frac{23{,}60 - 34{,}58 + 0{,}5\left(\frac{1}{31} + \frac{1}{25}\right)}{4{,}3841} = -2{,}4970$$

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

15.4.3 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 traditionnellement 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 $Z_{ajusté} = -\frac{2{,}4970}{0{,}996388} = -2{,}5061$.

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.

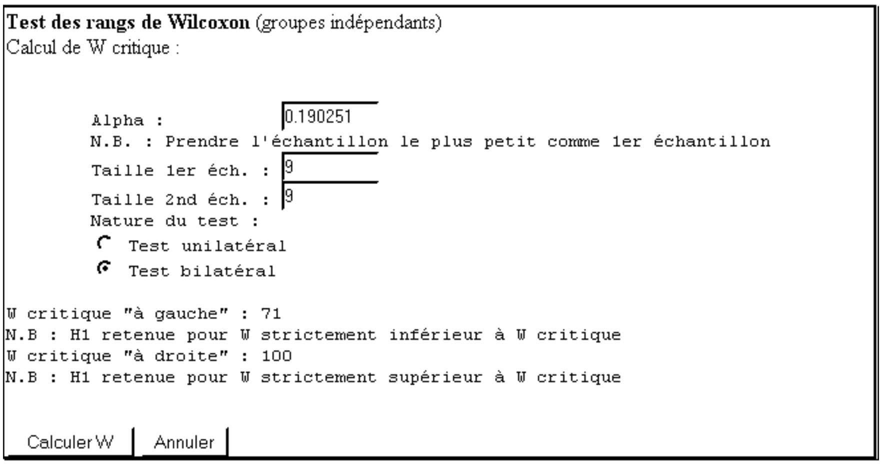
Remarque : Ce facteur de correction peut également être calculé en faisant le quotient de la variance du protocole des rangs effectivement observé sur les données expérimentales par la variance du protocole des rangs sans ex aequo, c'est-à-dire du protocole (1, 2, ..., n_1+n_2)

15.4.4 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 :

	Test U de Mann-Whitney (Délinquance dans Delinquance-Juvenile.stw) Par var. Domicile Tests significatifs marqués à p <,05000									
variable	SommeRgs Maison des parents	SommeRgs Foyer	U	Z	valeur p	Z ajusté	valeur p	N Actif Maison des parents	N Actif Foyer	2*(1-p) p exact
Absentéisme	101	70	25	1,32	0,19	1,33	0,18	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.

15.5. Test de Kruskal-Wallis

Objectif du test : comparer les médianes de la variable dépendante dans plusieurs groupes indépendants, lorsque la variable dépendante est ordinale ou numérique.

15.5.1 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 <u>Statistiques - Tests non paramétriques - Comparaison de plusieurs échantillons indépendants</u>.
Sélectionnez Absentéisme comme variable dépendante, et Domicile comme variable de classement.
Vous devriez obtenir les résultats suivants :

	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		
Dépend. : Absentéisme	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.

15.5.2 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.

15.5.3 Test post hoc après un test de Kruskal-Wallis

Lorsque le test de Kruskal-Wallis conclut sur une différence significative entre les groupes, on peut compléter l'étude par des tests de comparaison par paires.

Statistica fournit un test analogue au test de Bonferroni-Dunn. Pour exécuter ce test, utilisez le bouton Comparaison multiple des rangs moyens de tous les groupes. du dialogue ANOVA de Kruskal-Wallis et test de la médiane.

Deux feuilles de résultats sont produites par Statistica :

	Valeurs z des comp. multiples ; PERFRMNC Var. indépendante (classement) : CONDITN Kruskal-Wallis : H (2, N= 36) =13,84 p =,001		
Dépend. : PERFRMNC	FORME R:11,583	COULEUR R:16,667	TAILLE R:27,250
FORME		1,1819	3,6424
COULEUR	1,1819		2,4606
TAILLE	3,6424	2,4606	

	Valeurs p des Comp. Multiples (bilatéral) Var. indépendante (classement) : CONDITN Kruskal-Wallis : H (2, N= 36) =13,84 p =,001		
Dépend. : PERFRMNC	FORME R:11,583	COULEUR R:16,667	TAILLE R:27,250
FORME		0,7118	0,0008
COULEUR	0,7118		0,0416
TAILLE	0,0008	0,0416	

Dans la première, Statistica nous indique la valeur observée de la statistique de test, pour chaque paire de groupes. Cette statistique suit une loi normale centrée réduite.

La deuxième feuille indique le niveau de significativité global (familywise) de chacune des comparaisons. Ce dernier est obtenue en multipliant par le nombre de comparaisons le niveau de significativité individuel. Par exemple, Z=1,1819 correspond au niveau de significativité bilatéral : P(|Z|>1,1819)= 0,237245. Le deuxième tableau indique donc : p-value = 0,237245 x 3 = 0,7117.

Sur l'exemple fourni, on voit que la condition "Taille" diffère significativement des conditions "Forme" et "Couleur", et que ces deux dernières conditions ne diffèrent pas significativement entre elles.

16. Tests non paramétriques sur des groupes appariés

16.1. Test du khi-2 de Mac Nemar

Objectif du test : comparer deux groupes appariés (deux conditions expérimentales) lorsque la variable dépendante est une variable dichotomique.

16.1.1 Le test du khi-2 de Mac Nemar sur un tableau d'effectifs

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 données1)		
	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**	

16.1.2 Le test de comparaison de deux proportions sur des groupes appariés, à partir d'un tableau protocole

Ouvrez le classeur Reconnaissance-Portraits.stw. Dans la feuille de données contenue dans ce classeur, le protocole des 184 photos a été saisi.

Utilisation du test du signe

On pourra remarquer que le test de comparaison de deux variables dichotomiques sur des groupes appariés vu en cours de statistiques est en fait un cas particulier du test du signe. On peut ainsi le mettre en oeuvre comme suit.

Utilisez le menu Statistiques - Tests non paramétriques - Comparaison de deux échantillons appariés.
Indiquez "Deux semaines" et "Un an" comme variables et cliquez sur le bouton "Test des signes". Vous devriez obtenir le résultat suivant :

Test des Signes (Feuille de données2) Tests significatifs marqués à p <,05000				
	Nbe Non ex-aequo	%age v < V	Z	niv. p
Deux semaines & Un an	54	14,81	5,035	0,000

La statistique Z indiquée par Statistica est presque identique à celle donnée dans le cours. La seule différence est l'introduction d'une correction de continuité ; la formule exacte est ainsi :

$$Z = \frac{|b - c| - 1}{\sqrt{b + c}}$$

Utilisation du test Q de Cochran ou du test deFriedman

On peut aussi remarquer que le test Q de Cochran (comparaison de plusieurs conditions expérimentales sur des groupes appariés, avec une variable dépendante dichotomique) est identique au test du khi-2 de Mac Nemar lorsque le nombre de conditions est égal à 2. On peut ainsi le mettre en oeuvre comme suit.

Utilisez le menu Statistiques - Tests non paramétriques - Test Q de Cochran.
Indiquez "Deux semaines" et "Un an" comme variables et cliquez sur le bouton "Synthèse". Vous devriez obtenir le résultat suivant :

Variable	Test Q de Cochran Nbre d'obs. actives :184 Q = 26,74074, dl = 1, p < ,000000 Somme	%age 0	%age 1
Deux semaines	127,0000	30,9783	69,0217
Un an	89,0000	51,6304	48,3696

Le test de Friedman (test non paramétrique de comparaison de plusieurs conditions expérimentales sur des groupes appariés, pour une variable ordinale ou numérique) est identique au test du khi-2 de MacNemar lorsque la variable est dichotomique et le nombre de conditions égal à 2.

Utilisez le menu Statistiques - Tests non paramétriques - Comparaison de plusieurs échantillons appariés (variables).
Indiquez "Deux semaines" et "Un an" comme variables et cliquez sur le bouton "Synthèse". Vous devriez obtenir le résultat suivant :

Variable	ANOVA de Friedman & Coef. de Concord. de Kendall ANOVA du Chi² (N = 184, dl = 1) = 26,74074 p = ,0000(Coeff. de Concordance = ,14533 Rang moy. r = ,14066 Rang Moyen	Somme Rangs	Moyenne	Ec-type
Deux semaines	1,603261	295,0000	0,690217	0,463666
Un an	1,396739	257,0000	0,483696	0,501098

Dans l'un et l'autre cas, la statistique Q qui est calculée est simplement :

$$Q = \frac{(b-c)^2}{b+c}$$

Utilisation de la méthode Tableaux et tris croisés

Le résultat du test du khi-2 de Mac Nemar peut également être obtenu à l'aide de la méthode "Tableaux et tris croisés".

Utilisez le menu Statistiques - Statistiques Elémentaires - Tableaux et tris croisés. Choisissez l'un ou l'autre des onglets Tableaux croisés et Tris croisés.
Indiquez "Deux semaines" comme première liste de variables et "Un an" comme seconde liste. Cliquez sur OK.
Sélectionnez ensuite l'onglet Options et cochez l'item "Exact de Fisher, Yates, MacNemar (2x2)".
Affichez l'onglet "Avancé" et cliquez sur le bouton "Tableaux détaillés à double entrée".
La feuille de résultats "Stats : Deux semaines(2) x Un an(2)" fournit alors différents résultats, dont le calcul du khi-2 de MacNemar (ligne (B/C)). Le calcul est fait ici avec la correction de Yates.

16.2. Test du signe - Groupes appariés

Objectif du test : comparer deux groupes appariés (deux conditions expérimentales) lorsque la variable dépendante est une variable ordinale ou numérique.

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 :

Couple de variables	Test des Signes (Enfants-PRN.sta) Tests significatifs marqués à p <,05000 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 :

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

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...

16.3. Le test de Wilcoxon - Groupes appariés

Objectif du test : comparer deux groupes appariés (deux conditions expérimentales) lorsque la variable dépendante est une variable ordinale ou numérique.

16.3.1 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 :

Couples de variables	Test de Wilcoxon pour Ech. Appariés (Enfants-PRN.sta) Tests significatifs marqués à p <,05000 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 :

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

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.

16.3.2 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; -999999998)
- Pour la colonne RgMoins, introduisez la formule : = iif(Diff <0; RgDiffAbs; -999999998).

N.B. Dans les versions plus anciennes de Statistica, le code des valeurs manquantes est : -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 -999999998 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);-999999998)

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.

16.4. Test Q de Cochran

Objectif du test : comparer plusieurs groupes appariés (plusieurs conditions expérimentales) lorsque la variable dépendante est une variable dichotomique.

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 :

Variable	Test Q de Cochran (Interview.sta) Nbre d'obs. actives :18 Q = 16,66667, dl = 2, p < ,000240 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.

16.5. Test de Friedman

Objectif du test : comparer plusieurs groupes appariés (plusieurs conditions expérimentales) lorsque la variable dépendante est une variable ordinale ou numérique.

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

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 du classeur Mothers.stw.

Buts 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.

Observez les données concernant les notes attribuées par les 20 juges. En fait, il semble que les juges aient opéré un classement des 20 mères (notes de 1 à 20), mais que certains (les juges 5 et 9) aient introduit des ex aequo, sans appliquer de règle de rang moyen ni de décalage dans les rangs. Les notes de ces deux juges vont ainsi de 1 à 19.

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 :

Variable	ANOVA de Friedman & Coef. de Concord. de Kendall (Mothers.stw ANOVA du Chi² (N = 13, dl = 19) = 146,5510 p =0,00000 Coeff. de Concordance = ,59332 Rang moy. r = ,55943			
	Rang Moyen	Somme Rangs	Moyenne	Ec-type
MERE_1	7,19	93,50	7,15	4,60
MERE_2	4,69	61,00	4,69	5,27
MERE_3	11,85	154,00	11,69	4,23
MERE_4	6,69	87,00	6,69	3,35
MERE_5	8,54	111,00	8,46	3,76
MERE_6	4,38	57,00	4,38	3,15
MERE_7	7,92	103,00	7,92	3,99
MERE_8	11,00	143,00	10,92	3,48
MERE_9	7,12	92,50	7,08	3,17
MERE_10	5,85	76,00	5,85	3,44
MERE_11	13,96	181,50	13,85	3,26
MERE_12	8,46	110,00	8,46	3,02
MERE_13	5,46	71,00	5,46	5,77
MERE_14	10,85	141,00	10,69	5,34
MERE_15	12,58	163,50	12,46	4,86
MERE_16	14,54	189,00	14,38	2,26
MERE_17	17,38	226,00	17,23	3,19
MERE_18	15,54	202,00	15,38	3,04
MERE_19	16,46	214,00	16,31	2,63
MERE_20	19,54	254,00	19,38	0,96

Statistica nous indique la valeur de la statistique F_r de Friedman : F_r=146,55 et sa p-value : p<10^{-5}. Comme on peut le voir, il existe des différences très significatives entre les mères.

La seconde valeur indiquée par Statistica est le coefficient de concordance de Kendall $\tilde{W}$. Ici, on a : $\tilde{W} = 0{,}59332$. Cette valeur peut être déterminée facilement à partir de F_r. En effet, on a : $\tilde{W} = \dfrac{F_r}{n(k-1)}$, où n désigne le nombre d'observations et k le nombre de groupes appariés. Sur notre exemple, on peut vérifier que $\tilde{W} = \dfrac{146{,}5510}{13 \times 19} = 0{,}59332$. L'interprétation statistique faite par Kendall est cependant légèrement différente : on calcule un protocole de rangs par individu statistique (par ligne), on fait des sommes de rangs par colonne, et on s'intéresse à la variance de la série des sommes de rangs. Si les n protocoles de rangs sont tous identiques, cette variance est maximale et $\tilde{W} = 1$. Sur notre exemple, cela correspondrait à un accord parfait entre les 13 juges, et d'une manière générale, cela correspond à un même classement des conditions expérimentales pour tous les sujets. Dans le cas contraire, $\tilde{W} < 1$.

La troisième valeur indiquée par Statistica (rang moyen r=0,55943) est la moyenne des $\dfrac{n(n-1)}{2}$ coefficients de corrélation obtenus en prenant deux à deux les protocoles de rangs des différentes lignes. Ces coefficients de corrélation (appelés coefficients de corrélation de Spearman, car ils sont calculés à partir des rangs et non pas directement à partir des valeurs de la variable dépendante) mesurent les degrés d'accord entre les juges. Cette moyenne est liée au coefficient de concordance par la relation : $r = \dfrac{n\tilde{W} - 1}{n-1}$ où n est le nombre d'individus statistiques et on vérifie bien sur notre exemple que $r = \dfrac{13 \times 0{,}59332 - 1}{12} = 0{,}55943$.

Remarque. Pour la plupart des tests sur les rangs, Statistica ne fait pas de correction pour les ex aequo. Bien que la documentation fournie avec le logiciel ne le mentionne pas, en traitant un mini-exemple comportant des ex aequo à la main et sous Statistica, on pourra constater qu'en revanche, cette correction est faite pour le test de Friedman.

16.6. Exercices

16.6.1 Exercice 1

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.

Ouvrez le classeur Statistica Empowerment.stw et vérifiez la saisie des données.

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.

16.6.2 Exercice 2

Dans un article publié en 2003 (Clark C.F., Kotchen M.J., Moore M.R., Internal and external influences on pro-environmental behavior: Participation in a green electricity program, Journal of Environmental Psychology, Vol. 23, pp. 237-246, 2003), des chercheurs ont analysé, d'un double point de vue à la fois économique et psychologique, le comportement "pro-environnement" des consommateurs. Une étude précédente avait montré que de nombreux foyers américains étaient prêts à payer leur électricité plus cher, à condition qu'elle soit produite à partir d'énergies renouvelables.

Dans le cadre d'un programme "d'électricité verte", une compagnie du Michigan propose à ses clients de payer une redevance additionnelle, en contrepartie de laquelle la compagnie connecte sur le réseau un bloc photovoltaïque (production d'électricité solaire) d'une puissance déterminée.

Afin d'étudier les motivations des souscripteurs de ce programme, une enquête, sous forme de questionnaire envoyé par courrier, a été adressée à 281 participants et à 619 non-participants au programme. Les taux de réponse ont été de 95% pour les participants et de 76% pour les non-participants.

On a notamment demandé aux participants au programme de classer par ordre de préférence les cinq motivations environnementales suivantes :

1) La réduction de la pollution de l'air améliorera la santé des écosystèmes naturels
2) La réduction de la pollution de l'air profitera aux personnes résidant dans ma région
3) Ma santé, et celle de ma famille pourront s'améliorer car le programme améliorera la qualité de l'air
4) La diminution des émissions de dioxyde de carbone liées à la production d'électricité ralentira la vitesse du réchauffement du climat
5) J'éprouve de la satisfaction à participer à ce programme, indépendamment de ses effets environnementaux.

Sur un échantillon de 20 réponses choisies au hasard, les classements des cinq motivations sont les suivants :

	Motivation1	Motivation2	Motivation3	Motivation4	Motivation5
s1	2	4	3	1	5
s2	1	2	4	3	5
s3	1	4	2	3	5
s4	1	3	2	4	5
s5	1	3	2	4	5
s6	2	4	3	1	5
s7	1	2	3	4	5
s8	1	4	2	3	5
s9	2	3	1	4	5
s10	4	3	5	1	2
s11	2	1	5	3	4
s12	3	2	1	5	4
s13	5	3	4	1	2
s14	3	2	1	5	4
s15	3	1	5	2	4
s16	2	3	1	4	5
s17	2	1	4	3	5
s18	1	2	4	3	5
s19	1	2	3	4	5
s20	3	1	4	2	5

Etudier à l'aide d'un test approprié si les rangs moyens des cinq motivations sont significativement différents.

Lors d'études précédentes, des auteurs ont identifié l'égoïsme, le social-altruisme et l'intérêt pour la biosphère comme prédicteurs des comportements pro-environnement. Cependant, lorsque le comportement pro-environnement nécessite la volonté de payer un surcoût pour la protection de l'environnement, ils ont trouvé que seul l'égoïsme est un prédicteur fiable de tels comportements.
Les résultats précédents sont-ils en accord avec ces conclusions ?

16.6.3 Exercice 3

Fichier contenant les données à traiter : Conso-crustaces.stw

Une étude de consommation a été menée auprès des pêcheurs à pied en Bretagne (Cyndie Picot, thèse soutenue à l'UBO en 2010). 510 pêcheurs à pied fréquentant les plages bretonnes ont rempli un questionnaire relatif à leurs habitudes de consommation. A partir des données récoltées, on a évalué leur consommation alimentaire de crustacés, et on souhaite étudier la variation du niveau de consommation selon différents facteurs socio-économiques.

On s'intéresse ici aux variables suivantes :

- Conso : consommation de crustacés en grammes par personne et par jour ;
- SD sexe : sexe de la personne interrogée
- SD Age Catégorisé : catégorie d'âge, avec la catégorisation suivante :
  - 1 : moins de 50 ans
  - 2 : de 50 ans à moins de 60 ans
  - 3 : de 60 ans à moins de 65 ans
  - 4 : 65 ans et plus
- SD Etudes : niveau d'études, catégorisé de 1 à 4
  - 1 : primaire ou sans diplôme
  - 2 : CAP/BEP
  - 3 : Etudes secondaires, Bac
  - 4 : études supérieures
- SD Résidence : catégorisé de 1 à 4
  - 1 : nord de la zone étudiée
  - 2 : centre de la zone
  - 3 : sud de la zone
  - 4 : autre zone
- SD CSP : catégorie socio-professionnelle
  - 1 : agriculteur
  - 2 : petit patron
  - 3 : cadre supérieur
  - 4 : profession intermédiaire
  - 5 : employé
  - 6 : ouvrier qualifié
  - 7 : retraité
  - 8 : inactif

On souhaite étudier l'effet de chacune des variables catégorisées sur la consommation de crustacés chez les pêcheurs à pied.

1) A l'aide de tests de normalité, justifier l'utilisation de tests non paramétriques pour étudier ces données.

2) Pour chacune des cinq variables catégorisées, étudier son effet sur la consommation, à l'aide d'un test non paramétrique.

3) Dans le cas où le test précédent conduit à un résultat significatif, compléter l'étude par des comparaisons des catégories prises 2 à 2.

16.6.4 Exercice 4

Ref. Costalat-Founeau A.-M. et al., Représentation du corps et de l'alimentation chez une population de femmes de plus de 75 ans, in Papers on Social Representations, Textes sur les représentations sociales Volume 11, pages 4.1 - 4.20 (2002).

Dans une étude publiée en 2002, les auteurs s'intéressent aux représentations sociales du corps et de l'alimentation chez les femmes de plus de 75 ans. L'étude porte sur un échantillon de 52 sujets. Les participantes sont d'abord été interviewées dans le cadre d'un entretien semi-directif. A partir des thèmes recueillis au cours des entretiens, cinq propositions ont été élaborées :

Mécanique: Le corps est avant tout une machine vivante qui obéit à une mécanique naturelle très précise. Chaque organe joue un rôle particulier et le moindre petit problème peut déboucher sur une maladie qui affectera tout le corps.

Séduction: Le corps est la première chose que l'on offre au regard des autres. C'est pour cela qu'il est important de l'entretenir, soit par des exercices, soit par des soins cosmétiques afin de conserver une apparence physique agréable.

Communication: Le corps est sans doute le premier moyen de communication de l'homme. Il nous permet d'abord d'établir des contacts physiques avec des personnes ou des animaux; nous les voyons, les touchons, les sentons. En fait, c'est grâce à nos cinq sens que nous percevons notre univers dans sa globalité.

Emotion: Le corps est avant tout le siège des émotions ; il nous permet de ressentir le plaisir ou bien la douleur, mais aussi d'extérioriser nos sentiments. Il est aussi le porte parole de l'amour et de l'amitié que celui de la haine.

Maternité: La fonction première du corps de la femme est sans doute celle de pouvoir concevoir, porter et donner le jour à un enfant. Le corps de la femme est avant toute autre chose, le corps nourricier de la mère.
Dans une seconde phase, chacun des sujets devait classer les cinq propositions en fonction de son degré d'accord (1 pour la proposition qui lui semble la plus juste, 5 pour celle qui semble la moins satisfaisante). Par ailleurs on a relevé également diverses caractéristiques socio-économiques des sujets. On s'intéresse ici aux 13 sujets dont les revenus sont compris entre 900 € et 1300 €. Les résultats ont été les suivants :

Sujet	Mécanique	Séduction	Communication	Emotion	Maternité
s1	3	1	4	5	2
s2	3	1	4	5	2
s3	3	1	4	5	2
s4	3	1	4	5	2
s5	4	2	5	1	3
s6	1	2	3	5	4
s7	4	2	3	5	1
s8	4	2	3	5	1
s9	5	2	4	3	1
s10	5	2	3	4	1
s11	5	3	2	1	4
s12	3	4	1	2	5
s13	5	2	3	1	4

1) On souhaite étudier si les 5 propositions obtiennent des scores significativement différents. Quel test statistique permet d'apporter une réponse à cette question ?

2) Réaliser le test et conclure au seuil de 5%.

17. Exercices à rendre par mail

Réalisez les études demandées dans les 2 exercices ci-dessous. Faites parvenir votre travail (classeur Statistica contenant les traitements demandés, commentaire saisi dans un rapport Statistica ou un fichier Word) par mail à votre enseignant (adresse : Francois.Carpentier@univ-brest.fr).

17.1. Exercice 1

Ref. Wagstaff, G., Equity Versus Equality in Allocations to Adults and Children, in The Journal of Social Psychology, Vol. 137 Issue 4, pp. 445 - 448 (1997).

Le respect de la règle d'équité impose de répartir les revenus entre des partenaires en fonction de leurs contributions. Cependant, une règle d'égalité de traitement peut s'imposer si les partenaires ne sont pas considérés comme pleinement responsables.

L'auteur mène une expérience mettant en jeu ces règles. Les participants sont 50 étudiants britanniques.

L'auteur propose aux participants 3 scénarios, qui, tous, mettent en scène des jumeaux.

Dans le premier, les jumelles ont 7 ans, l'une est paresseuse et désobéissante, l'autre est gentille et serviable. Les parents ont acheté 10 ensembles d'habits pour poupées. On demande aux participants comment ils répartiraient les ensembles entre les jumelles.

Dans le second, les jumeaux ont 16 ans. De nouveau, l'un est paresseux et source d'ennuis, l'autre est courageux et admiré. On demande aux participants comment ils répartiraient 10 disques de musique entre les jumeaux.

Dans le troisième, les jumeaux ont 40 ans. L'un est serviable et aide ses parents, l'autre ne s'occupe pas d'eux. Les parents ont économisé une somme de 10 000 £ pour les jumeaux. On demande aux participants comment ils répartiraient l'argent entre les jumeaux.

Pour chacune des répartitions faites par chaque participant, on note si celle-ci se fait sur la base de l'égalité (codée 1) ou l'équité (codée 0).

Le tableau des effectifs des différents types de réponses est le suivant :

Scén. 1	Scén. 2	Scén. 3	Effectif
1	1	1	15
1	1	0	27
1	0	0	7
0	0	0	1

L'auteur indique une différence significative (p<.0001) entre les répartitions proposées pour les 3 scénarios. Indiquer quel test peut être utilisé et réaliser ce test et retrouver ainsi les résultats de l'auteur.

17.2. Exercice 2

Enquête réalisée par des étudiants de L2 de STAPS de l'Université Paris X et publiée sur le site de Christophe Génolini : http://christophe.genolini.free.fr/enseignement/enqueteFraude2008.php
Les données utilisées ici ont été légèrement modifiées pour éviter les problèmes liés aux valeurs manquantes. Elles sont rassemblées dans le classeur *<u>Enquete-fraude.stw</u>*

Une enquête a été menée en 2007 auprès d'un échantillon de 286 étudiants d'une Université parisienne. Il s'agissait d'un questionnaire portant sur la fraude aux examens universitaires.

Les 286 personnes interrogées répondaient à quelques questions d'ordre général (âge, sexe, niveau d'études, UFR, redoublement éventuel, mention au Bac), puis à plusieurs questions relatives à la fraude. Dix techniques de fraude étaient envisagées : copier en regardant la copie ou les brouillons d’un autre (variable "Copier"), communiquer avec un voisin (variable "Communiquer"), échanger des brouillons (variable "EchangeBrouillon"), fabriquer ou utiliser des antisèches (papier, calculatrice, portable, MP3) (variable "Antiseche"), envoyer et recevoir des SMS (variable "SMS"), avoir le cours sur les genoux (variable "CoursGenoux"), ne pas rendre sa copie (variable "GarderCopie"), préparer la salle avant (écrire sur les tables, laisser des documents aux toilettes) (variable "PreparerSalle"), voler les sujets avant l’examen (fouiller le sac ou le bureau du prof) (variable "VolerSujet"), Autres (variable "Autres").
\medskip

Pour chaque technique, les sujets interrogés devaient répondre sur une échelle de Likert à 5 niveaux : jamais, rarement, parfois, souvent, toujours. Les réponses recueillies ont été codées de 0 (pour "jamais") à 4 (pour "toujours") et un score de synthèse (variable ScoreTricheTotal) est calculé en additionnant les valeurs correspondant à chaque méthode.

On souhaite étudier si les scores observés pour les dix techniques envisagées sont significativement différents, et, comme la distribution de chacune des variables s'écarte notablement d'une loi normale, on souhaite réaliser un test non paramétrique.

a) Quel test permettrait de faire cette étude ?

b) Réaliser le test et conclure.