Tests d'adéquation, d'indépendance et non paramétriques

Partie 1 Tests d' adéquation et d' indépendance

A] Test du χ^2 d' adéquation à une loi théorique

Illustration: §4.3.3, p. 123. Adéquation à une loi binomiale dans le cas où π est estimée

H0: le nombre de réponses correctes pour un sujet quelconque suit une loi binomiale $\mathcal{B}$(8; 0,4612)

H1: ce n'est pas le cas

Il existe ici une fonction excel très rustre pour calculer la p-value.

Etape 1:

Entrer les données sous la forme de deux colonnes effectifs observés/effectifs attendus. Demander à Excel de réaliser le test sur la base de ces informations.

	A	B
1	effectifs observés	effectifs attendus
2	32	11,14
3	24	29,14
4	38	49,9
5	42	53,38
6	29	36,58
7	15	15,64
8	20	4,22
9		
10		=TEST.KHIDEUX(A2:A8;B2:B8)
11		

Etape 2: En sortie Excel donne la p-value du test:

	A	B
1	effectifs observés	effectifs attendus
2	32	11,14
3	24	29,14
4	38	49,9
5	42	53,38
6	29	36,58
7	15	15,64
8	20	4,22
9		
10		0,0000

Ici, on voit qu'il y a rejet de H0 au seuil 0,001 et donc que les observations ne sont (absolument) pas compatibles avec l'idée qu'elles proviennent d'une loi binomiale $\mathcal{B}$(8; 0,4612).

Remarques:

a) les conditions d'application ne sont pas vérifiées par Excel. A vous de le faire (c'est facile puisque vous devez entrer les effectifs attendus)!

b) Si l'on souhaite un peu plus détailler les choses, on pourra intégrer des petites formules pour calculer les termes de t, par exemple:

	A	B	C
1	effectifs observés	effectifs attendus	
2	32	11,14	=PUISSANCE(A2-B2;2)/B2
3	24	29,14	=PUISSANCE(A3-B3;2)/B3
4	38	49,9	=PUISSANCE(A4-B4;2)/B4
5	42	53,38	=PUISSANCE(A5-B5;2)/B5
6	29	36,58	=PUISSANCE(A6-B6;2)/B6
7	15	15,64	=PUISSANCE(A7-B7;2)/B7
8	20	4,22	=PUISSANCE(A8-B8;2)/B8
9			
10		=TEST.KHIDEUX(A2:A8;B2:B8)	=SOMME(C2:C8)

ce qui donne le détail des calculs et la valeur de t:

	A	B	C
1	effectifs observés	effectifs attendus	
2	32	11,14	39,0610
3	24	29,14	0,9066
4	38	49,9	2,8379
5	42	53,38	2,4261
6	29	36,58	1,5707
7	15	15,64	0,0262
8	20	4,22	59,0067
9			
10		0,0000	105,8352

où l'on voit que l'écart au modèle stipulée en H0 vient surtout des extrêmes; il y a beaucoup plus d'observés que d'attendus pour les nombres de réponses correctes faibles ou élevés.

B] Test du χ^2 d' indépendance - test du χ^2 d' homogénéïté de plusieurs distributions

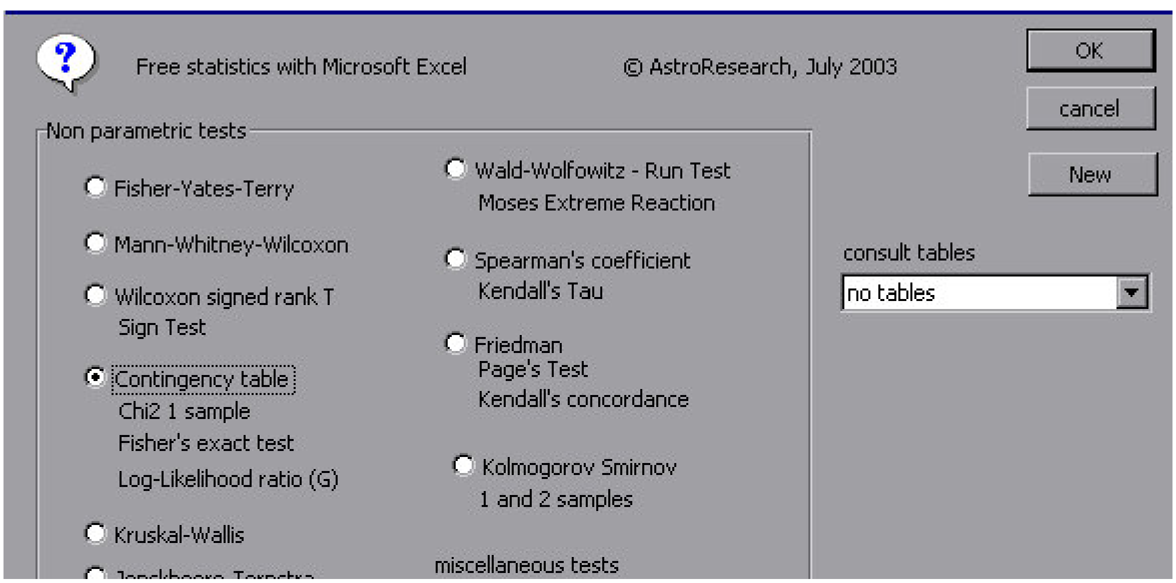
On reprendra l' ensemble de macros Excel "tests non paramétriques" gracieusement fournis par le concepteur du site "Croire Savoir Imaginer" et que l' on peut récupérer à l' adresse internet suivante: http://perso.club-internet.fr/hdelboy/Nonparam.htm et à partir du lien fourni sur notre page dédié à Excel. Il existe sur le même site aussi une documentation pdf, regroupant à la fois des éléments théoriques et statistiques, pour utiliser ces tests.

Pour utiliser ces macros, il suffira de les "unzipper" puis dans le dossier obtenu d' ouvrir le programme MANNWHIT.xlm . Les données pourront être entrées au cas par cas et sauvegardées dans des classeurs différents. La consultation des pages et fichiers du concepteur des macros donne aussi tous les renseignements nécessaires pour ce faire.

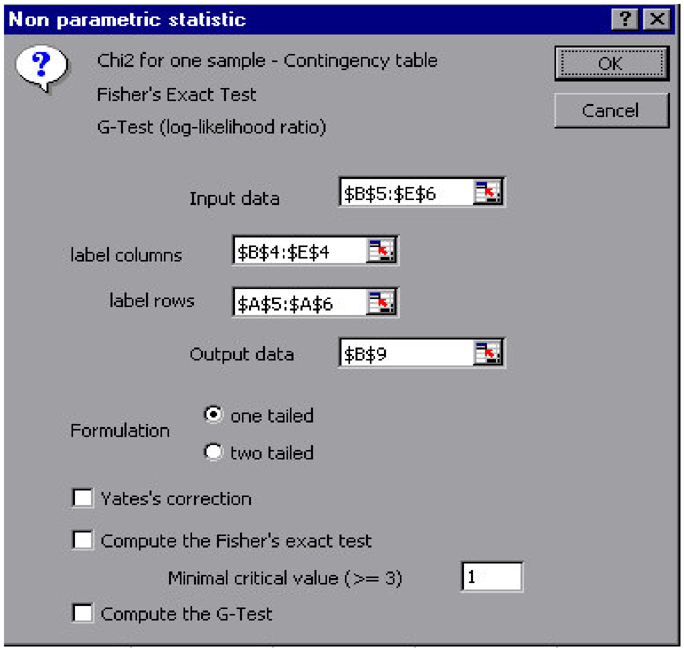
Illustration: §4.4.1, p. 127.
H0: couleurs préférées et genres sont indépendants
H1: le contraire
Etape 1: Entrer le tableau de contingence:

	A	B	C	D	E
1					
2	***Click the button to appear tests or type the sequence ctrl - a***				
3					
4		rose	blanche	bleue	rouge
5	Garçons	30	35	82	53
6	Filles	52	50	48	33

Etape 2: Cliquez sur la case A1 (bouton de cette macro) pour lancer les macros non paramétriques, puis choisir dans la menu proposé:



Etape 3: Dans le menu suivant, entrer: la plage de valeurs; la plage contenant les noms de colonnes et lignes et enfin -dans output data- la cellule où l' affichage des résultats doit débuter.



<u>Remarque:</u> pour ce test, peu importe de demander du one ou two tailed (uni ou bilatéral). N' ayant pas utilisé la correction de Yates dans l' ouvrage, on s' en abstiendra ici.
<u>Résultats:</u>

On trouve dans chaque case (i,j) du premier tableau résultat le nombre $\dfrac{(n_{ij} - \frac{n_{i.}n_{.j}}{N})^2}{\frac{n_{i.}n_{.j}}{N}}$

($\dfrac{(\text{effectif observé - effectif attendu})^2}{\text{effectif attendu}}$. Les marges de ce tableau, intitulées nij ou nji, donnent les sommes en lignes ou en colonnes des termes précédents. Enfin le total du tableau est donné dans la case en bas à droite du tableau.

C' est la valeur de l' indice Khi2.

	rose	blanche	bleue	rouge	nij
Garçons	3,84	1,98	2,93	1,46	10,21
Filles	4,19	2,17	3,21	1,59	11,16
nji	8,03	4,15	6,14	3,05	21,37

$$\frac{(n_{23} - \frac{n_{2.}n_{.3}}{N})^2}{\frac{n_{2.}n_{.3}}{N}}$$

Le second tableau résultat donne le nombre de ddl; la valeur de l' indice t et la p-value:

Contingency table	
ddl	3
chi2	21,381
p	0

p<0,001 indique un confortable rejet de l' hypothèse d' indépendance.

Remarque: conditions d'application:

La condition d'application retenue dans cette macro est que les effectifs théoriques soient ≥3. Cette règle ne correspond pas aux règles de "Cochran" (aucun effectif plus petit que 1 et moins de 20% d'entre eux inférieurs à 5) qui sont plus habituelles:

La première des conditions (aucun effectif théorique < 1)pourra être testée en entrant dans le menu où on règle les plages de valeurs à utiliser (cf. étape 3) la commande suivante

Minimal critical value (>= 3)	1

.

La seconde condition ne peut par contre être vérifiée qu'en faisant soi-même les calculs des effectifs théoriques, ce qu'on pourra par exemple faire en entrant les formules suivantes dans deux cellules Excel puis en les appliquant (par Glisser-Déposer) aux 3 colonnes suivantes (pour les données de l'exemple ci-dessus):

```
=SOMME($B$5:$E$5)*SOMME(B5:B6)/SOMME($B$5:$E$6)
=SOMME($B$6:$E$6)*SOMME(B5:B6)/SOMME($B$5:$E$6)
```

Ce qui donne:

19		42,8198	44,3864	67,8851	44,9086
20		39,1802	40,6136	62,1149	41,0914

C] Test d'adéquation de Kolmogorov-Smirnov

N'ayant pas trouvé de macros freeware pour réaliser ce test, la possibilité offerte ici est d'adapter à vos observations la procédure ci-dessous - dont le classeur excel est récupérable sur notre site.

Notons que seuls le cas où $\mathcal{N}$ et σ sont inconnus est programmé ici (test de Lilliefors).

Illustration: à l'ouverture, le fichier traité correspond à la question 1c du problème 3 du chapitre 8 (§8.3, p. 235)

H0: Le temps mis à résoudre la tâche a une distribution $\mathcal{N}(494; \sqrt{54,2222})$

H0: l'inverse

On notera que les résultats diffèrent un peu de ceux figurant dans l'ouvrage. En effet, ces derniers ayant été faits "à la main" utilisaient des arrondis beaucoup plus grossier dans les étapes intermédiaires.

L'exemple qui suit consiste à tester s'il est plausible que la suite d'observations *54 - 67 - 45 - 60 - 55 - 63 - 64,5 - 49 - 53 - 62 - 62 - 77 - 52 - 56 - 60 - 54 - 61 - 68 - 52 - 54* provienne d'une loi normale.

Etape 1:

Si vous avez moins d'observation qu'il y en a dans le fichier proposé, Lors supprimer entre la ligne correspondant à la première valeur et celle correspondant à la dernière valeur autant de lignes qu'il en faut pour qu'il en reste exactement le nombre d'observations correspondant à vos mesures. **Prendre garde de ne pas toucher à la première ligne de notre fichier** (c'est là où figurent nos commandes)

Si vous avez plus d' observation qu' il y en a dans le fichier proposé, insérer entre la ligne correspondant à la première valeur et celle correspondant à la deuxième valeur autant de lignes qu' il en faut pour faire figurer toutes vos données.

Sur l'exemple:
Vous avez 20 observations. Il faut donc insérer 10 lignes entre les première et deuxième lignes:

	A	B
1	Valeurs triées	Effectifs valeurs
2		
3	482	1
4	489	1
5	490	1
6	491	2
7	491	0
8	493	1
9	497	1
10	499	2
11	499	0
12	509	1
13		
14		

10 Valeurs de la feuille initiale

	A	B
1	Valeurs triées	Effectifs valeurs
2		
3	482	1
4		
5		
6		
7		
8		
9		
10		
11		
12		
13		
14	489	1
15	490	1
16	491	2
17	491	0
18	493	1
19	497	1
20	499	2
21	499	0
22	509	1

20 Valeurs après insertion des 10 lignes

Etape 2:

Entrer vos observations **triées au préalable** à la place de nos données en prenant garde de **bien démarrer à notre première donnée**.

Exemple:

	A	B
1	Valeurs triées	Effectifs valeurs
2		
3	482	1
4		
5		
6		
7		
8		
9		
10		
11		
12		
13		
14	489	1
15	490	1
16	491	2
17	491	0
18	493	1
19	497	1
20	499	2
21	499	0
22	509	1

Feuille initiale avec lignes vierges insérées

	A	B
1	Valeurs triées	Effectifs valeurs
2		
3	45	1
4	49	
5	52	
6	52	
7	53	
8	54	
9	54	
10	54	
11	55	
12	56	
13	60	
14	60	2
15	61	1
16	62	2
17	62	0
18	63	1
19	64,5	1
20	67	1
21	68	1
22	77	1

"Vos 20 valeurs préalablement triées"

Etape 3:

Pour les colonnes autres que les données triées et celle donnant la valeur de t, t' et p, recopier par un "Glisser-déposer" les formules se trouvant sur la première ligne non vide de chaque colonne vers l' ensemble des cellules situées au-dessous. Sur l' exemple ci-dessus, vous devez obtenir le tableau (la première colonne ne faisant pas l' objet de la recopie de formule est grisée):

	A	B	C	D	E	F	G	H
1	Valeurs triées	Effectifs valeurs	Effectifs cumulés	z	$F_N(j)$	$p_{H0}(j)$	$p_{H0}(j)-F_N(j-1)$	$p_{H0}(j)-F_N(j)$
2								
3	45	1	1	-1,80	0,0500	0,0363	0,0363	0,0137
4	49	1	2	-1,26	0,1000	0,1038	0,0538	0,0038
5	52	2	4	-0,86	0,2000	0,1951	0,0951	0,0049
6	52	0	4	-0,86	0,2000	0,1951	0,0000	0,0000
7	53	1	5	-0,73	0,2500	0,2341	0,0341	0,0159
8	54	3	8	-0,59	0,4000	0,2770	0,0270	0,1230
9	54	0	8	-0,59	0,4000	0,2770	0,0000	0,0000
10	54	0	8	-0,59	0,4000	0,2770	0,0000	0,0000
11	55	1	9	-0,46	0,4500	0,3235	0,0765	0,1265
12	56	1	10	-0,32	0,5000	0,3729	0,0771	0,1271
13	60	2	12	0,21	0,6000	0,5834	0,0834	0,0166
14	60	0	12	0,21	0,6000	0,5834	0,0000	0,0000
15	61	1	13	0,34	0,6500	0,6347	0,0347	0,0153
16	62	2	15	0,48	0,7500	0,6837	0,0337	0,0663
17	62	0	15	0,48	0,7500	0,6837	0,0000	0,0000
18	63	1	16	0,61	0,8000	0,7296	0,0204	0,0704
19	64,5	1	17	0,81	0,8500	0,7917	0,0083	0,0583
20	67	1	18	1,15	0,9000	0,8742	0,0242	0,0258
21	68	1	19	1,28	0,9500	0,8998	0,0002	0,0502
22	77	1	20	2,48	1,0000	0,9935	0,0435	0,0065

Et les résultats du test:

t	0,1271
t'	0,5914
p	>0,15

Partie 2: Tests non paramétriques

On reprendra l' ensemble de macros Excel "tests non paramétriques" gracieusement fournis par le concepteur du site "Croire Savoir Imaginer" et que l' on peut récupérer à l' adresse internet suivante: http://perso.club-internet.fr/hdelboy/Nonparam.htm et à partir du lien fourni sur notre page dédié à Excel.
Pour utiliser ces macros, il suffira de les "unzipper" puis dans le dossier obtenu d' ouvrir le programme MANNWHIT.xlm . Les données pourront être entrées au cas par cas et sauvegardées dans des classeurs différents. La consultation des pages et fichiers du concepteur des macros donne aussi tous les renseignements nécessaires pour ce faire.

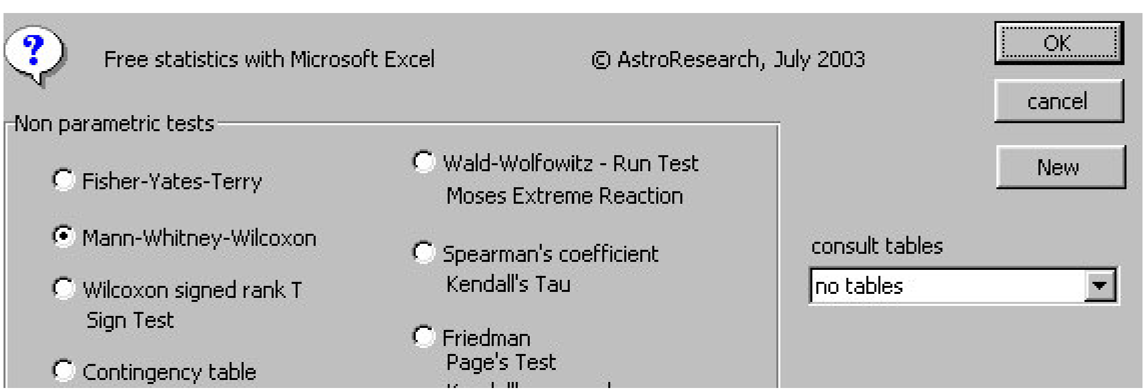
A] Test de Wilcoxon pour échantillons indépendants

Illustration: §9.2.4, p. 267
H1: le rappel est globalement meilleur lorsque le contexte dans lequel il a lieu est similaire à celui d' apprentissage que lorsque le contexte est différent.
H0: l' inverse

Etape 1: entrer les données n' importe où sur la page.

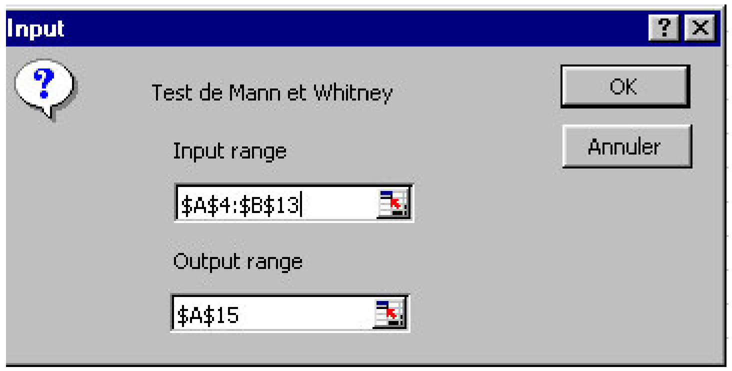
	A	B
1		
2	*Click the button to appear tes*	
3	***Similaire***	***Différent***
4	16	12
5	20	22
6	19	10
7	22	7
8	25	8
9	13	15
10	14	12
11	25	6
12		9
13		19

Etape 2: Cliquez en A1 puis demander dans le menu résultat de cette action un test de Mann-Whithney-Wilcoxon :



Etape 3: dans input range, on rentrera la plage sur laquelle les valeurs des deux conditions - et seulement ces valeurs- sont contenues.

Dans Ouput range, indiquez la cellule à partir de laquelle les résultats doivent être affichés.



Résultats:

A droite, on trouve les rangs dans les deux conditions:

rank x	rank y
11	6,5
14	15,5
12,5	5
15,5	2
17,5	3
8	10
9	6,5
17,5	1
	4
	12,5

Les deux premières lignes du tableau suivant sont les sommes des rangs dans chaque condition; les deux suivantes sont les moyennes des rangs.

rank S xy [Wx]	105
rank S yx [Wy]	66
M rank xy	13,13
M rank yx	6,6

Le programme utilise à partir de là une approximation normale valable dès que la taille du plus petit des deux échantillons est supérieure ou égale à 8, ce qui est le cas ici (cette approximation n' est pas utilisée dans le cas contraire).

La valeur de t est donnée dans la ligne intitulé e qui suit la dernière ligne du tableau précédent:

27	M rank yx	6,6
28	e	2,577

La p-value bilatérale est donnée trois lignes plus tard:

31	p	0,01

$p^b/2$ (=0,005) étant inférieure à 0,01 indique qu' une des hypothèses unilatérales associées au test bilatéral peut être rejetée au seuil 0,01 (rejet de H0: "le rappel est au moins aussi bon lorsque le contexte dans lequel il a lieu est différent de celui d' apprentissage que lorsque le contexte est similaire ou de H0 : "le rappel est au moins aussi bon lorsque le contexte dans lequel il a lieu est similaire à celui d' apprentissage que lorsque le contexte est différent" au seuil 0,01). Pour vérifier que c' est bien H0: "le rappel est au moins aussi bon lorsque le contexte dans lequel il a lieu est différent de celui d' apprentissage que lorsque le contexte est similaire" qui est rejetée, il suffit de jeter un coup d' œil sur les moyennes des rangs pour vérifier qu' elles sont bien contradictoires avec cette hypothèse. Moyenne des rangs "similaire" = 13,13 > Moyenne des rangs "différent" = 6,6 indique qu' en moyenne les rangs sont plus élevés dans le contexte similaire, c' est-à-dire que les valeurs sont elles aussi plus élevées. Nos observations sont contradictoires avec H0, que l' on peut donc rejeter pour tout seuil plus grand que $p^b/2$ (la p-value associée à ce test est égale à $p^b/2$, soit étant donné les approximations réalisées à peu près 0,005).

Note: Une deuxième statistique est proposée ainsi qu' une seconde p-value. Il s' agit d' une correction par les ex-aequos qui devrait normalement être utilisée, notamment lorsque les échantillons sont de très faible tailles. Il vaut mieux consulter les deux informations.

14	***Mann Whitney test***		
15	N	18	
16	nx	8	
17	ny	10	
18	Uxy	69	
19	Uyx	11	
20	Uo	40	
21	s2U	126,666667	
22	mxy	19,25	
23	myx	12	
24	rank S xy [Wx]	105	somme des rangs condition 1
25	rank S yx [Wy]	66	somme des rangs condition 2
26	M rank xy	13,13	moyenne des rangs condition 1
27	M rank yx	6,6	moyenne des rangs condition 2
28	e	2,577	t observé
29	critical value of U 5%	17	
30	critical value of U 1%	11	
31	p	0,01	p-value bilatérale
32	corrections for ties	yes	
33	ties	8	
34	s2U	126,143791	
35	e	2,58205244	t corrigé par les ex-aequos
36	p	0,0098	p-value bilatérale pour t corrigé par les ex-aequos

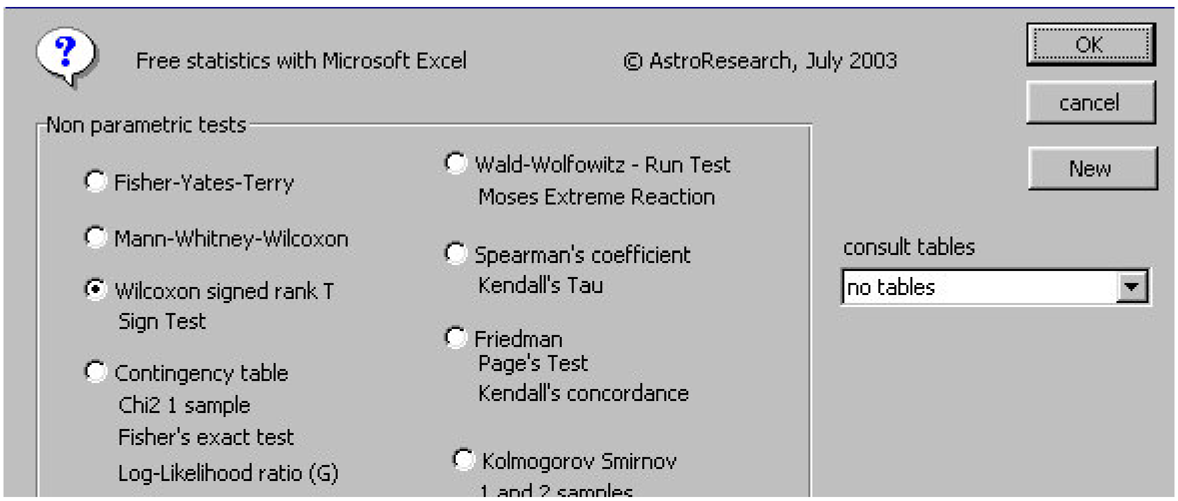
B] Test de Wilcoxon et test du signe pour échantillons appariés

Illustration: §9.4.3, p. 272
H1: le degré explicatif des origines internes est plus élevé que celui des origines externes.
H0: le degré explicatif des origines externes est au moins aussi élevé que celui des origines internes.

Etape 1: entrer les données n' importe où sur la page.

4	Externe	Interne
5	85	91
6	78	75
7	80	80
8	84	87
9	86	90
10	88	96
11	83	82
12	80	85
13	90	95
14	88	90

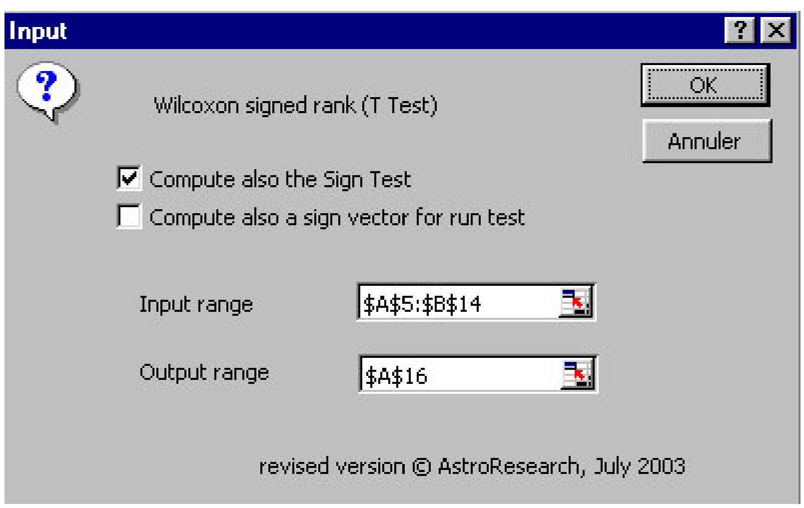
Etape 2: Cliquez en A1 puis demander dans le menu résultat de cette action un test Wilcoxon signed rank T :



Etape 3:

Indiquez, si vous le souhaitez comme c' est le cas ici, de calculer aussi le "Sign Test". Dans input range, on rentrera la plage sur laquelle les valeurs des deux conditions - et seulement ces valeurs- sont contenues.

Dans Ouput range, indiquez la cellule à partir de laquelle les résultats doivent être affichés.



Résultats:

Un tableau situé à droite donne sur une première colonne les rangs des différences x-y (externe-interne ici). Une deuxième colonne donne les rangs associés à des différences positives et une troisième colonne les rangs associés à des différences négatives.

rank diff.	rank +	rank -
-6		8
3	3,5	
-3		3,5
-4		5
-8		9
1	1	
-5		6,5
-5		6,5
-2		2

Le tableau intitulé "Wilcoxon rank test" donne les résultats du test de Wilcoxon:

16	***Wilcoxon rank test***	
17	Np	9
18	Mx	84,2
19	My	87,1
20	P+	4,5
21	P-	40,5
22	P0	1
23	T the.	22,5
24	varT	71,25
25	e	-2,132
26	p	p < 0.05
27	ties	4
28	critical value l	6

somme des rangs associés à des différences positives

somme des rangs associés à des différences négatives

valeur de la statistique utilisant une approximation normale

p-value bilatérale.

Il utilise une approximation normale à partir de 20 observations. Etant en-dessous ici, la p-value est obtenue directement sur la base des tables.

$p^{b}/2$ (<0,025) étant inférieure à 0,05 (par exemple si on a décidé de mener notre test au seuil 0,05) indique qu' une des hypothèses unilatérales associées au test bilatéral peut être rejetée au seuil 0,05 (rejet de H0: "le degré explicatif des origines externes est au moins aussi élevé que celui des origines internes" ou de H0 : "le degré explicatif des origines internes est au moins aussi élevé que celui des origines externes" au seuil 0,05). Pour vérifier que c' est bien H0: "le degré explicatif des origines externes est au moins aussi élevé que celui des origines internes" qui est rejetée, il suffit de jeter un coup d' œil sur les sommes des rangs associés aux différences positives et négatives. Somme des rangs "différences positives" = 4,5 < somme des rangs "différences négatives" = 45,5 indique qu' en général les différences sont associées à des différences négatives, c' est-à-dire qu' en général les différences "externe - interne" sont négatives donc qu' en général "externe < interne". Nos observations sont donc bien contradictoires avec H0, que l' on peut donc rejeter pour tout seuil plus grand que $p^{b}/2$ (la p-value associée à ce test est égale à $p^{b}/2$, soit <0,025).

Le tableau intitulé "Sign test" donne les résultats du test du signe:

La p-value bilatérale est manifestement fausse: H1bil doit être multipliée par 2.

$p^b/2$ (=0,09) n' étant pas inférieure à 0,05 (par exemple si on a décidé de mener notre test au seuil 0,05) indique que notre hypothèse unilatérale H0 ne peut pas être rejetée à ce seuil. Par contre $p^b/2$ étant inférieure à 0,1 indique qu' une des hypothèses unilatérales associées au test bilatéral peut être rejetée au seuil 0,1 (rejet de H0: "le degré explicatif des origines externes est au moins aussi élevé que celui des origines internes" ou de H0 : "le degré explicatif des origines internes est au moins aussi élevé que celui des origines externes" au seuil 0,1). Pour vérifier que c' est bien H0: "le degré explicatif des origines externes est au moins aussi élevé que celui des origines internes" qui est rejetée, il suffit de jeter un coup d' œil sur les nombres de différences positives et négatives. nombre de "différences positives" = 2 < nombre de "différences négatives" = 7 indique qu' en général les différences sont négatives, c' est-à-dire qu' en général les différences "externe - interne" sont négatives donc qu' en général "externe < interne". Nos observations sont donc bien contradictoires avec H0, que l' on peut donc rejeter pour tout seuil plus grand que $p^b/2$ (la p-value associée à ce test est égale à $p^b/2$, soit 0,09).

29	***Sign Test***	**2**
30	S+	2
31	S-	7
32	S min	2
33	S max	7
34	H1bil.	0,09
35	H'1 uni	0,09
36	H"1	0,98

Nombre de différences positives

Nombre de différences négatives

p-value bilatérale/2

C] Test de corrélation des rangs de Spearman

Illustration: §9.5.3, p. 277

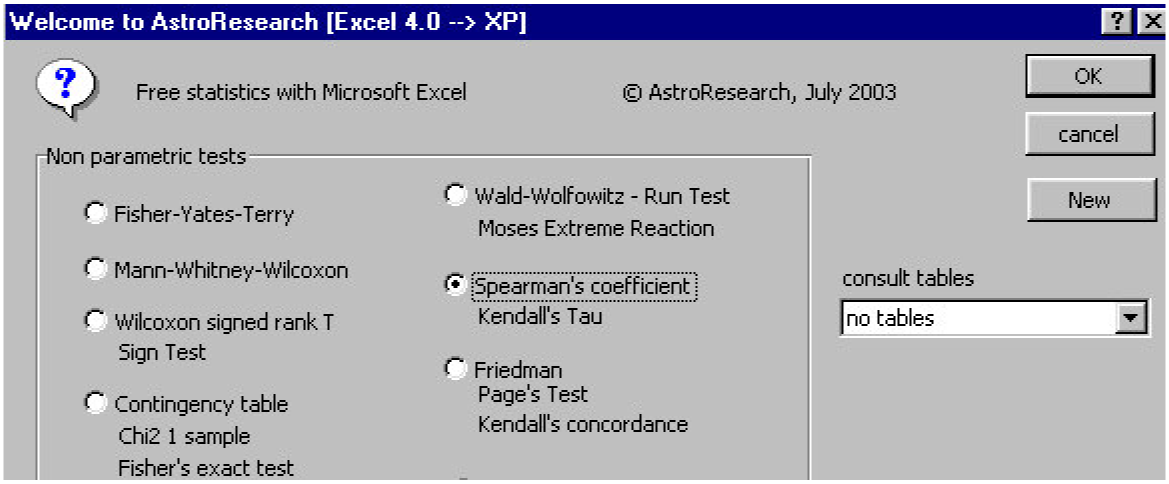
H1: la corrélation entre les degrés explicatifs donnés aux origines internes et externes est positive ⌠ $\varphi_{xy}>0$.

H0: la corrélation entre les degrés explicatifs donnés aux origines internes et externes est négative ou nulle ⌠ $\varphi_{xy}\leq 0$.

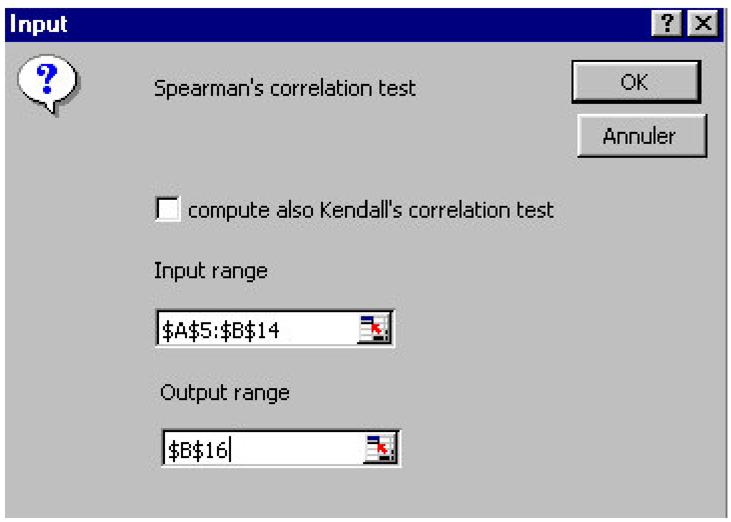
Etape 1: entrer les données n' importe où sur la page.

	A	B	C
1			
2	***Click the button to appear tests or type***		
3			
4	Externe	Interne	
5	85	91	
6	78	75	
7	80	80	
8	84	87	
9	86	90	
10	88	96	
11	83	82	
12	80	85	
13	90	95	
14	88	90	

Etape 2: Cliquez en A1 puis demander dans le menu résultat de cette action "Spearman' s coefficient" :



Etape 3: Entrer la plage de valeurs pour les observations en input range et la cellule où doit commencer l' affichage des résultats en Output range.



Résultats:

Un tableau situé à droite donne les rangs des valeurs pour les deux variables. Il est à noter qu' avant l' établisssement des rangs, les couples de valeurs ont été classés par ordre croissant de la première variable.

rank x	rank y	LS for ranked data
1	1	1,4
2,5	2	2,77
2,5	4	2,77
4	3	4,13
5	5	5,04
6	8	5,96
7	6,5	6,87
8,5	10	8,23
8,5	6,5	8,23
10	9	9,6

La valeur du coefficient est donnée dans la ligne **r'** du second tableau. Le programme utilise une approximation par une loi de Student à n-2 ddl pour déterminer la p-value, approximation utilisable dès que $n \geq 10$ (dans le cas contraire le logiciel utilise une table). La p-value bilatérale est donnée dans la ligne **p**.

Notons que l' approximation par une loi de Student entraîne une correction à réaliser lorsqu' il existe des ex-aequos. Cette correction est notée **r' corr.** pour le coefficient de corrélation corrigée et la p-value correspondante est notée **p corr**.

Spearman rank order cor	
N	20
Np	10
Mx	84,2
My	87,1
Sd	0
Sd2	15
df	8
r'	0,909
t	6,169
p	0,00027
Z	0,00319
ties	6
r' corr.	0,908
t corr.	6,14
p corr.	0,00028

$p^b/2$ (corrigée ou non) étant inférieure à 0,001 indique qu' une des hypothèses unilatérales associées au test bilatéral peut être rejetée au seuil 0,001 (rejet de H0: $\varphi_{xy} \leq 0$ ou de H0 : $\varphi_{xy} \geq 0$ au seuil 0,001). Pour vérifier que c' est bien H0 $\varphi_{xy} \leq 0$ qui est rejetée, il suffit de jeter un coup d' œil sur la corrélation des rangs observées. r' =0,909 (ou r' corr.=0,908) indique que nos observations sont bien contradictoires avec H0, que l' on peut donc rejeter pour tout seuil plus grand que $p^b/2$ (la p-value associée à ce test est égale à $p^b/2$, soit, étant donné l' approximation à peu près 0,0001).

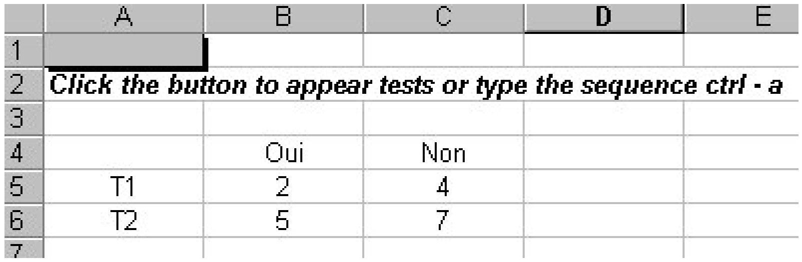
D] Test d' indépendance d' Irwin-Fisher

<u>Illustration: §9.6.1, p. 279</u>

H1: la proportion de guérisons est plus élevée avec le second traitement qu' avec le premier ($\pi_1 < \pi_2$).

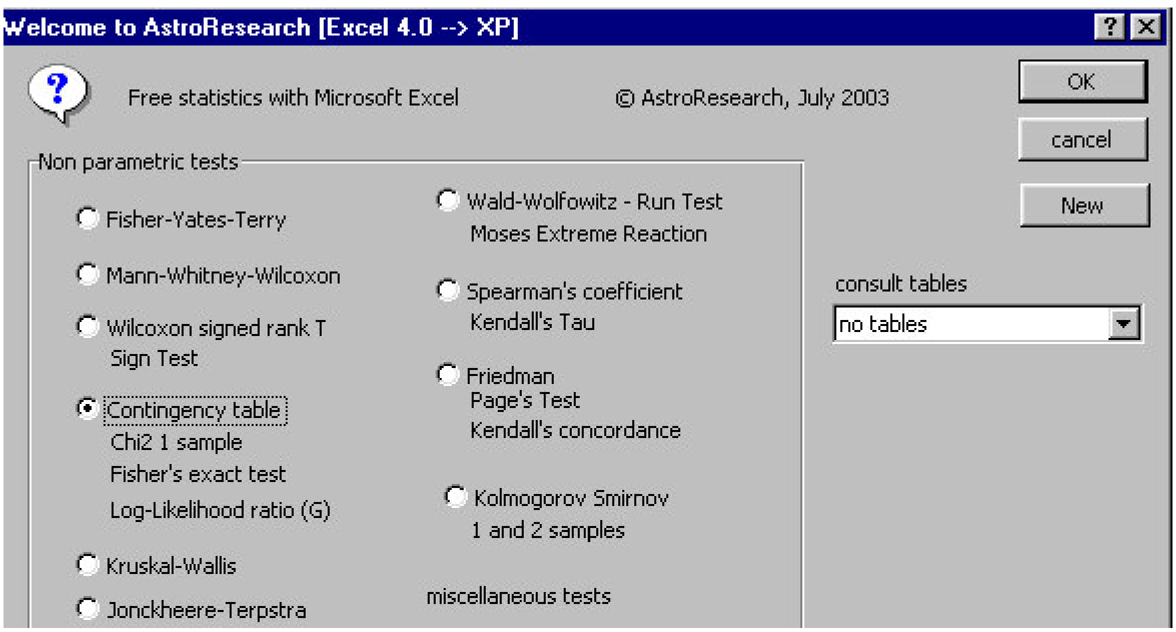
H0: la proportion de guérisons est au moins aussi élevée avec le premier traitement qu' avec le second ($\pi_1 \geq \pi_2$).

<u>Etape 1</u>: entrer les données n' importe où sur la page.

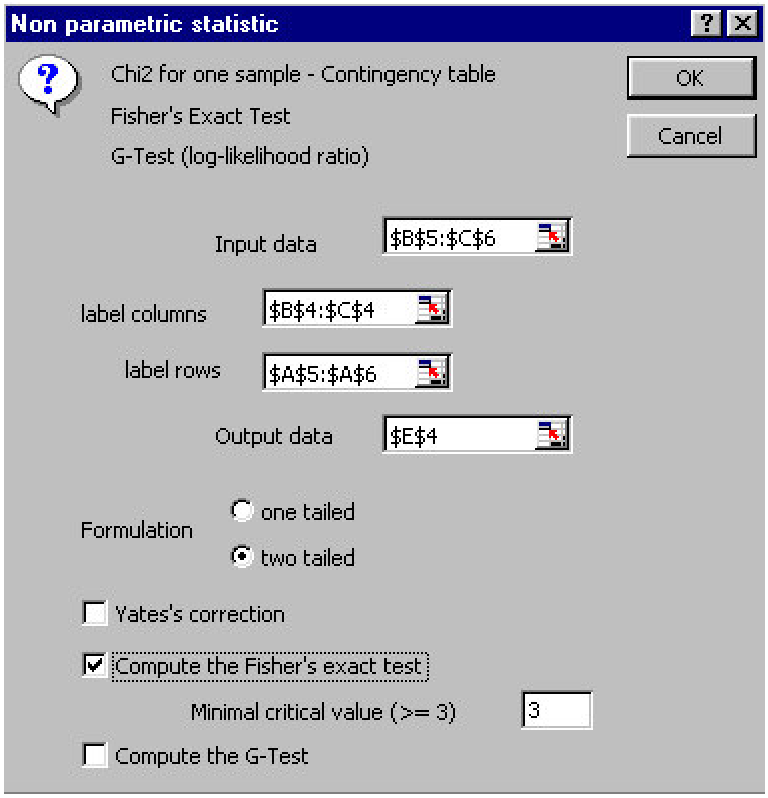


	A	B	C	D	E
1					
2	***Click the button to appear tests or type the sequence ctrl - a***				
3					
4		Oui	Non		
5	T1	2	4		
6	T2	5	7		
7					

<u>Etape 2:</u> Cliquez en A1 puis demander dans le menu résultat de cette action "Fisher' s exact test" :



<u>Etape 3:</u> Entrer la plage de valeurs pour les observations en "input data", les noms des colonnes en "label columns", les noms des lignes en "label rows" et la cellule à partir de laquelle les résultats seront affichés en "Output data". Demander "Compute Fisher' s exact test" et du "two tailed" (test bilatéral).



<u>Résultats:</u>

Le haut du tableau donne la distribution de probabilité des valeurs possibles de la statistique:

E	F
Possible matrices and their Ps	
p for 0 : [0,6,7,5]	
0,02488688	
p for 1 : [1,5,6,6]	
0,17420814	
p for 2 : [2,4,5,7]	
0,37330317	
p for 3 : [3,3,4,8]	
0,31108597	
p for 4 : [4,2,3,9]	
0,10369532	
p for 5 : [5,1,2,10]	
0,01244344	
p for 6 : [6,0,1,11]	
0,00037707	

Vient ensuite le tableau décisionnel proprement dit. Bien qu' y figure une p-value bilatérale (two tailed p), la logique de sa construction ne correspond pas à celle classiquement utilisée. Pour retrouver la p-value bilatérale, on prendra la valeur minimale de "one tailed p (0,5724)" et de "other (0,8009)", soit 0,5724, que l' on multipliera par 2, ce qui donne 1,1448.

Fisher's exact test	
table = [2,4,5,7]	
one tailed p	0,57239819
two tailed p	1
other	0,80090498
theor. value of chi2 < 3	2.33 for [2]

$p^b/2$ (=0,5724) n' étant inférieure à aucun des seuils classiques, il ne peut y avoir rejet de notre hypothèse nulle.

Partie 3 - Compléments tests non paramétriques: les tests de Kruskal-Wallis et de Friedman

Ces tests sont les extensions des tests de Wilcoxon indépendant et apparié lorsqu' il existe plus de deux conditions expérimentales.
Ce sont aussi les alternatives non paramétriques aux analyses de variance dans les plans d' expérience comportant une unique VI et des observations provenant soit d' échantillons indépendants, soit d' échantillons appariés.

A] Le test de Kruskal-Wallis

Exemple:
On souhaite vérifier l'existence de l'effet de l'absorption d'un tranquilisant sur la performance d'un individu. On construit une expérience en utilisant une VI à 4 modalités: absorption d'un placebo (groupe 1) ou d'une des formes du médicament (groupes 2, 3, 4). La performance d'un sujet est mesurée par un nombre compris entre 0 et 100. les résultats sont les suivants:

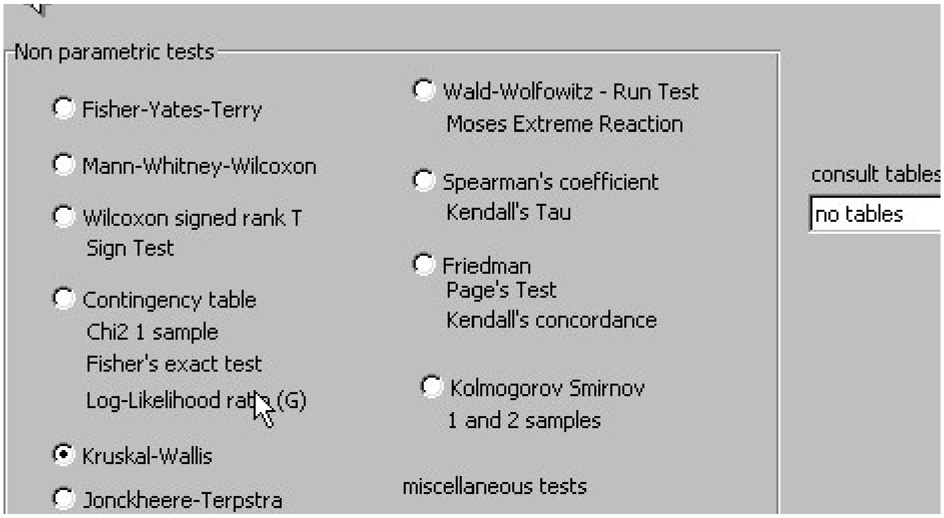
groupe 1: 38 - 42 - 55 - 47 - 84 - 53 - 45 - 73 - 81 - 54
groupe 2: 57 - 43 - 45 - 35 - 52 - 24 - 36 - 43 - 61
groupe 3: 33 - 36 - 58 - 47 - 39 - 24 - 18
groupe 4: 28 - 21 - 50 - 39 - 44 - 50 - 53 - 46 - 42 - 30

En utilisant un test non paramétrique, peut-on conclure et à quel seuil à une différence entre les 4 conditions expérimentales quant à la performance d'un individu ?

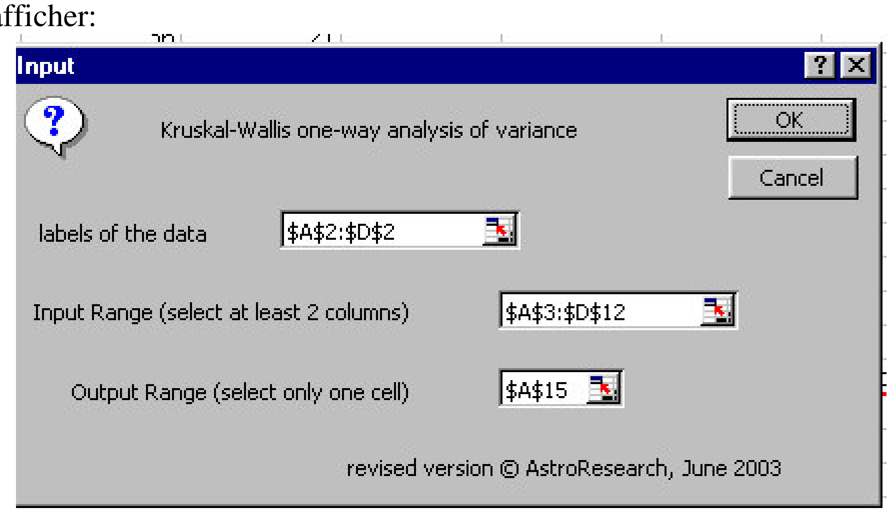
Entrée des données:
Autant de colonnes (lignes) qu' il y a de conditions (4). Une ligne comprend les noms des conditions.

	A	B	C	D
1				
2	groupe 1	groupe 2	groupe 3	groupe 4
3	38	57	33	28
4	42	43	36	21
5	55	45	58	50
6	47	35	47	39
7	84	52	39	44
8	53	24	24	50
9	45	36	18	53
10	73	43		46
11	81	61		42
12	54			30

Analyse:
Cliquez sur le bouton de la Macro, puis demander Kruska-Wallis:



Dans le menu qui suit, labels of data correspond aux noms des conditions, Input range à la plage de valeurs à utiliser et Output range est une cellule à partir de laquelle les résultats devront s' afficher:



Résultats:
Dans un premier tableau, on trouve les rangs des observations:

groupe 1	groupe 2	groupe 3	groupe 4
11	31	7	5
14,5	16,5	9,5	2
30	19,5	32	24,5
22,5	8	22,5	12,5
36	26	12,5	18
27,5	3,5	3,5	24,5
19,5	9,5	1	27,5
34	16,5		21
35	33		14,5
29			6

Un second tableau donne diverses statistiques: nombre total de mesures; dans chaque échantillon nombre de mesures, moyenne des scores bruts, somme et moyenne des rangs, nombre total d' ex-aequos, valeur de l' indice statistique, indice statistique corrigé par les ex-aequos.

	groupe 1	groupe 2	groupe 3	groupe 4
sample size	36			
count	10	9	7	10
average	57	44	36	40
rank sum	259	163,5	88	155,5
rank average	25,9	18,17	12,57	15,55
ties	18			
Q	7,943			
Q corr.	7,952			
p	NS			

p désigne comme d' habitude la p-value. Elle est cependant manifestement incorrecte ici. Pour la calculer correctement, on utilisera la fonction suivante:

=LOI.KHIDEUX(F32;COLONNES(A3:D3)-1)

où F32 correspond à la valeur de la statistique corrigée par les ex-aequos (valeur de Q corr.); COLONNES(A3:D3) donne le nombre de colonnes de la plage de données, c' est-à-dire le nombre de conditions expérimentales.

On obtient une p-value alors égale à 0,047, ce qui est significatif au seuil 0,05. On peut donc conclure à ce seuil qu' il existe globalement des différences de performances selon le type de substance utilisée.

B] Le test de Friedman

Exemple:

Une étude avait pour objet de comparer l'importance que des adolescents donnent à diverses explications de la pauvreté et de la richesse. On a sélectionné au hasard parmi un grand nombre de lycéens de même profil (âge, origine ethnique, type d'établissement d'origine,…) 10 garçons. Ceux-ci ont eu pour tâche d'attribuer un score explicatif à 4 origines possibles de la pauvreté: (1) origines familiales (p.e. la famille ne met pas l'accent sur le succès) (2) chance (p.e. manque de chance) (3) origines externes (p.e. salaires trop bas) (4) origines internes (p.e. mauvaise gestion des ressources financières). Le score attribué à une explication consistait en une note pouvant aller de 70 (explication pas importante) à 100 (explication très importante). Les mesures sont consignées dans le tableau suivant:

Famille	Chance	Externe	interne
84	82	85	91
74	77	78	75
82	85	80	80
87	85	84	87
82	84	86	90
85	86	88	96
77	78	83	82
82	80	80	85
84	85	90	95
86	86	88	90

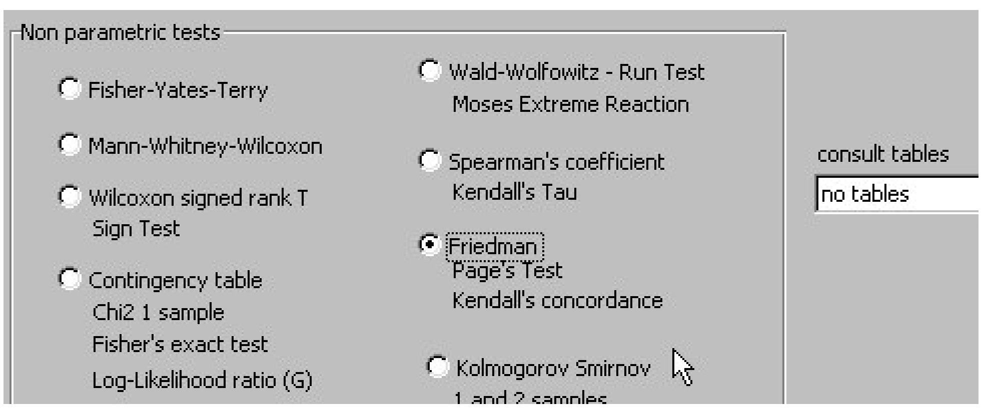
Peut-on conclure à des degrés explicatifs globalement différents des quatre types d' explication ?

Entrée des données:

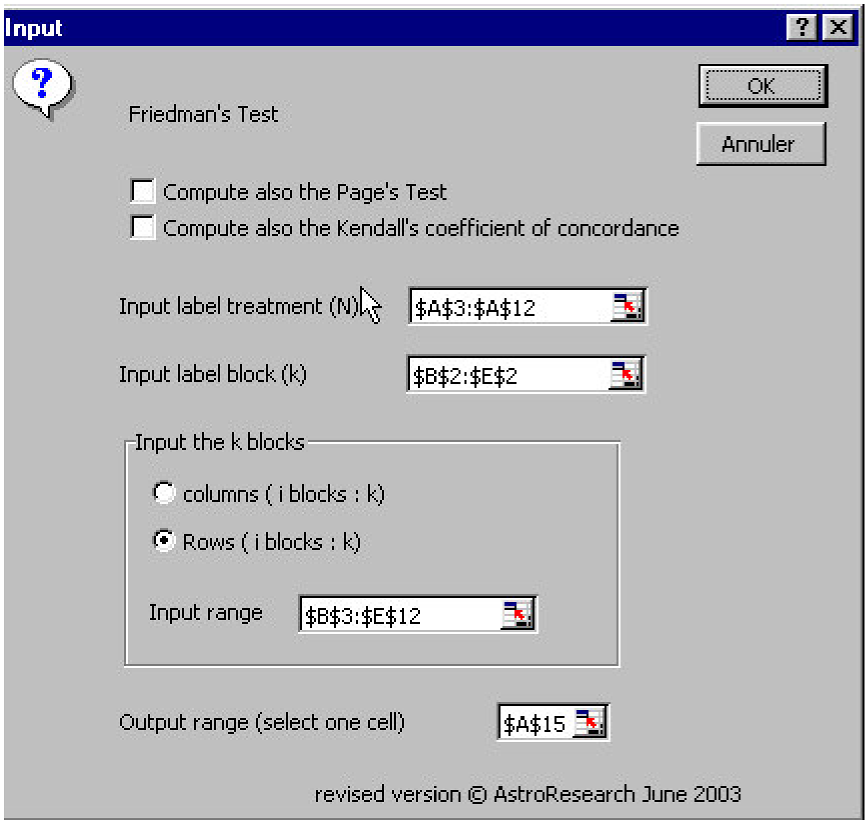
Autant de colonnes (lignes) qu' il y a de conditions (4). Une ligne comprend les noms des conditions et une colonne un repérage des sujets.

	A	B	C	D	E
1					
2		Famille	Chance	Externe	interne
3	s1	84	82	85	91
4	s2	74	77	78	75
5	s3	82	85	80	80
6	s4	87	85	84	87
7	s5	82	84	86	90
8	s6	85	86	88	96
9	s7	77	78	83	82
10	s8	82	80	80	85
11	s9	84	85	90	95
12	s10	86	86	88	90

Analyse:
Cliquez sur le bouton de la Macro, puis demander Friedman:



Le menu qui suit est assez difficile à utiliser. Ceci tient à ce que ce test peut aussi être conçu comme traitant un plan d' expérience à deux facteurs sans interaction, un facteur "colonne" et l' autre "ligne".
On entrera dans *Input label treatment* la plage de valeurs donnant les noms des sujets. Dans *Input label block*, on fera figurer la plage de valeurs donnant les noms des conditions expérimentales.
Enfin, dans *Input the k blocs*, a) on choisira l' option *rows* et b) on entrera la plage de valeurs contenant les données.



Résultats:

Dans un premier tableau, on trouve les rangs des observations déterminés par les sujets, et en fin de ligne la somme de ces rangs par conditions.

	s1	s2	s3	s4	s5	s6	s7	s8	s9	s10	total Ri
Famille	2	1	3	3,5	1	1	1	3	1	1,5	18
Chance	1	3	4	2	2	2	2	1,5	2	1,5	21
Externe	3	4	1,5	1	3	3	4	1,5	3	3	27
interne	4	2	1,5	3,5	4	4	3	4	4	4	34

Un second tableau donne le nombre total d' observations; le nombre de sujets; le nombre de conditions; la somme des carrés des sommes des rangs; la valeur F de l' analyse de variance à un facteur sur données répétées réalisée sur la base des rangs et la p-value associée; le nombre total d' ex-aequos et la valeur du coefficient correctif. Enfin, on trouve la valeur de la statistique à utiliser (correction par les ex-aequos) en T' . La p-value qui suit est ici signifcative, ce qui signifie que l' on peut conclure à des différences dans les degrés explicatifs moyens donnés aux 4 types d' explication. L' analyse des rangs semble dire que les explications internes sont considérées comme les plus explicatives et que la famille vient en dernière position.

Friedman's Test	
N	40
N (=t)	10
k (=b)	4
SRj2	2650
Fr (3, 27)	4,09
p	0,01622
critical value 5%	
ties	8
corr. for ties	0,96
T'	9,375
p'	0,0247