Troisième chapitre: Les tests non paramétriques :

Homogénéité entre échantillons appariés

1. Principe de l'appariement

Ce cas se présente chaque fois que l'on compare deux méthodes de mesures (ou deux séries de mesures différées dans le temps) en **soumettant les mêmes individus** à ces 2 méthodes. A chacune des méthodes correspond alors une population de mesures, mais ces populations et les échantillons que l'on peut en extraire ne sont pas indépendants. Il est aussi possible de soumettre les mêmes sujets à deux traitements différents. Chaque sujet est alors utilisé comme son propre contrôle.

La méthode paramétrique usuelle pour analyser deux échantillons non indépendants (appariés) est le test du t de Student. Si les conditions d'application de ce test ne sont pas réunies, il est alors possible d'utiliser différents tests non paramétriques : test du signe, test de rang de Wilcoxon, test de McNemar (Si les variables étudiées sont de natures dichotomiques, analyse de variance à deux facteurs de Friedman, test Q de Cochran (si la variable est catégorielle, par exemple, « succès » ou « échec »).

L'appariement, que l'on retrouve sous différentes appellation (mesures répétées, échantillons dépendants, paired samples ou matched pairs samples en anglais,) est une procédure très populaire en statistique. Elle permet une analyse fine des différences entre les populations.

2. Test des signes

2.1 Test d'hypothèses, statistique de test et région critique

Définition du test :

Nous considérons maintenant que nous disposons d'un échantillon de n observations. Chaque observation étant constituée d'une paire de valeurs. Nous formons une nouvelle variable aléatoire $D = X_1 - X_2$ dont les valeurs d_i sont obtenues par différences des paires de valeurs, c.-à-d.,

$$d_i = x_{i1} - x_{i2}$$

Le test des signes consiste à s'intéresser uniquement au sens de l'écart + = ($D > 0$) ou - = ($D < 0$), et non à son importance. Il est adapté à tout type de variables (quantitative, ordinale, même binaire) dès lors qu'il est possible de déterminer si une valeur est plus importante qu'une autre pour chaque paire d'observation. Ce gain en champ d'application entraîne en contrepartie une perte de puissance du test si d'aventure l'amplitude de l'écart était une information exploitable dans l'étude. Son homologue paramétrique est le test de Student pour échantillons appariés.

Soit π= Pr($D > 0$), le test d'hypothèses, s'il est bilatéral s'écrit

$$H_0 : \pi = \frac{1}{2} \quad \text{vs} \quad H_1 : \pi \neq \frac{1}{2}$$

Pour un test unilatéral, nous pouvons former l'hypothèse alternative $H_1 : \pi > \frac{1}{2}$ (resp. $H_1 : \pi < \frac{1}{2}$) lorsqu'on souhaite savoir si X_1 est stochastiquement plus grande (resp. plus petite) que X_2.

Statistique de test et région critique

Schématiquement, le test de signes consiste simplement à tester si le nombre des écarts positifs est significativement différent du nombre des écarts négatifs (plus élevée ou plus rare concernant les tests unilatéraux).

Pour un échantillon de taille n, une statistique naturelle pour le test des signes est le nombre d'occurrence S des écarts positifs, c.-à-d.,

$$S = \sum_{i=1}^{n} \delta_i .$$

Où $\delta_i = 1\ si\ d_i > 0, \delta_i = 0\ si\ d_i < 0.$

Sous H_0, $S \sim \mathcal{B}(n, \frac{1}{2})$. Comme ce dernier est symétrique, il est possible d'obtenir directement la probabilité critique (p-valeur) du test bilatéral en calculant la quantité

$$p = 2 \times \Pr[\mathcal{B}\left(n, \frac{1}{2}\right) \geq \max(S, n - S)]$$

Pour un test unilatéral $H_1 : \pi > \frac{1}{2}$, la probabilité critique sera $p = \Pr[\mathcal{B}\left(n, \frac{1}{2}\right) \geq \max(S, n - S)]$

Exemple : (Effet du glucose sur la mémoire).

Il s'agit d'identifier l'effet du glucose sur la mémoire de n = 16 patients âgés. L'expérimentation mise en place est la suivante : au réveil, on leur donne une boisson sucrée au moyen de glucose, on leur narre une histoire. Quelque temps après, on leur demande de la retracer. La qualité de la restitution est notée par un juge. Après un délai raisonnable, on recommence la même expérimentation mais en leur donnant une boisson à la saccharine (Remarque : on peut aussi imaginer que l'ordre d'administration des sucreries est défini aléatoirement selon les sujets). L'objectif de l'étude est de vérifier que l'ingestion de la boisson au glucose entraîne une qualité de mémorisation différente. Les données sont recensées dans le tableau suivant:

X_1(glucose)	X_2(saccharine	Signe(écart)
0	1	--
10	9	+
9	6	+
4	2	+
8	5	+
6	5	+
9	7	+
3	2	+
12	8	+
10	8	+
15	11	+
9	3	+
5	6	--
6	8	--
10	8	+
6	4	+

Nous disposons de n = 16 observations.

1. Nous formons une nouvelle colonne qui prend la valeur "+" ou "-" selon le signe de $d_i = x_{i1} - x_{i2}$.

2. Nous pouvons former la statistique $S = \sum_{i=1}^{n} \delta_i = 13$ correspondant au nombre de signe "+".

3. Calculons la quantité $p = Pr\left[B\left(n,\frac{1}{2}\right) \geq 13\right]$

$$p = Pr\left[B\left(n,\frac{1}{2}\right) = 13\right] + Pr\left[B\left(n,\frac{1}{2}\right) = 14\right] + Pr\left[B\left(n,\frac{1}{2}\right) = 15\right] + Pr\left[B\left(n,\frac{1}{2}\right) = 16\right]$$

$$= C_{16}^{13}(0.5)^{13}(0.5)^{16-13} + \cdots + C_{16}^{16}(0.5)^{16}(0.5)^{16-16} = 0.00854 + \cdots + 0.00002 = 0.01064$$

4. Dans le cadre d'un test bilatéral, la probabilité critique du test est égal à $p = 2 \times 0.01064 = 0.02127$. Au risque 5%, nous pouvons rejeter l'hypothèse nulle d'égalité des effets du glucose et de la saccharine sur la qualité de la mémorisation.

5. Pour un test unilatéral $H_1: \pi > \frac{1}{2}$, la probabilité critique serait directement

$p = Pr\left[B\left(n,\frac{1}{2}\right) \geq 13\right] = 0.01064$. On conclurait alors que l'effet du glucose sur la mémorisation est significativement meilleur.

2.2. Le cas particulier des écarts nuls

Lorsqu'il y a des écarts nuls dans les données, c.-à-d. $d_i = x_{i1} - x_{i2} = 0$, la modélisation des signes via la loi binomiale pose problème. En effet, il y aurait à ce moment là, 3 résultats possibles (d_i <0; d_i = 0; d_i > 0). Plutôt que de s'embarquer dans des schémas compliqués, on conseille généralement de supprimer ces. On réduit d'autant l'effectif n dans les calculs. Voyons ce mécanisme sur un exemple.

Exemple : (La prise de décision dans un couple).

On a demandé à des couples d'évaluer le poids de leur conjoint dans la prise de décision lors des achats importants.

L'homme note l'influence de sa femme, et vice versa. L'étude porte sur 17 couples. Les données sont regroupés dans le tableau ci-dessous, on cherche à savoir si l'homme donne plus d'importance à sa femme qu'inversement, nous sommes dans le cadre d'un test unilatéral :

Numéro	Mari	Femme	Ecart	Signe
A	5	3	2	+
B	4	3	1	+
C	6	4	2	+
D	6	5	1	+
E	3	3	0	0
F	2	3	-1	-
G	5	2	3	+
H	3	3	0	0
I	1	2	-1	-
J	4	3	1	+
K	5	2	3	+
L	4	2	2	+
M	4	5	-1	-
N	7	2	5	+
O	5	5	0	0
P	5	3	2	+
Q	5	1	4	+

Nbre de couples	17
N (+ ou --)	14
S	11

Le nombre de couples dans l'étude est 17.

- Les hypothèses à tester sont : $\begin{cases} H_0 : \dots \\ H_1 : \dots \end{cases}$
- Mais lorsque nous comptabilisons les écarts, nous nous rendons compte qu'il est nul (égal à 0) pour les couples *{E;H;O}*. On décide donc de les exclure des calculs. Nous disposons finalement de n = 14 observations exploitables.
- Nous comptons les observations telles que $d_i > 0$, nous obtenons $S = 11$.
- Nous calculons alors la probabilité $p = Pr\left[B\left(n, \frac{1}{2}\right) \geq 11\right] = 0.02869$

j	Pr
11	0.02222
12	0.00555
13	0.00085
14	0.00006
Somme	0.02669

- Qui est égal à la probabilité critique du test aussi puisque nous réalisons un test unilatéral.
- Pour un niveau de signification de 5%, les hommes pensent que leur épouses ont une place plus importante dans la prise de décision lors des achats importants.

2.3. L'approximation normale pour les grands effectifs

Pour n assez grand, ($n > 35$) en pratique, la distribution de S, qui est binomiale, peut être approchée par la loi normale qui, sous $H_0 : \pi = \frac{1}{2}$ prend les paramètres :

$$E(S) = n\pi = \frac{n}{2} \quad et \quad V(S) = n\pi(1-\pi) = \frac{n}{4}$$

La statistique centrée et réduite du test s'écrit : $Z = \frac{S-\frac{n}{2}}{\sqrt{\frac{n}{4}}} = \frac{2S-n}{\sqrt{n}} \sim \mathcal{N}(0,1)$

Pour un test bilatéral au risque α, la région critique du test est définie par :

$$R.C.: |Z| \geq u_{1-\alpha/2}$$

Remarque : (Correction de continuité).

Pour les effectifs modérés, nous pouvons améliorer la précision en introduisant une correction de continuité. Pour le test bilatéral, il s'écrit

$$|Z| = \frac{|2S-n|-1.0}{\sqrt{n}}.$$

Pour un test unilatéral $Z = \frac{2S-n\pm1.0}{\sqrt{n}}$.

Nous rajoutons +1 pour un test unilatéral à gauche ($H_1 : \pi < \frac{1}{2}$) ; nous retranchons pour un test unilatéral à droite.

Exemple (Shooté à l'hélium).

On a demandé à 39 footballeurs de shooter dans des ballons gonflés à l'hélium et à l'air. On cherche à savoir si dans le premier cas, ils arrivent à envoyer le ballon plus loin.

Nous sommes dans le cadre d'un test unilatéral à droite $H_1 : \pi > \frac{1}{2}$. Les données sont présentées comme suit :

Essai	X_1 (Hélium)	X_2 (Air)	Signe de l'ecart
1	25	25	0
2	16	23	--
3	25	18	+
4	14	16	--
5	23	35	--
6	29	15	+
7	25	26	--
8	26	24	+
9	22	24	--
10	26	28	--
11	12	25	--
12	28	19	+
13	28	27	+
14	31	25	+
15	22	34	--
16	29	26	+
17	23	20	+
18	26	22	+
19	35	33	+
20	24	29	+
21	31	31	0
22	34	27	+
23	39	22	+
24	32	29	+
25	14	28	--
26	28	29	--
27	30	22	+
28	27	31	--
29	33	25	+
30	11	20	--
31	26	27	--
32	32	26	+
33	30	28	+
34	29	32	--
35	30	28	+
36	29	25	+
37	29	31	--
38	30	28	+
39	26	28	--

- Puisqu'on a deux « d_i =0 », $n = 37$ observations exploitables.
- La statistique du test

  $S = \sum_{i=1}^{n} \delta_i$. Où $\delta_i = 1\ si\ d_i > 0, \delta_i = 0\ si\ d_i < 0$. Sous H_0, $S \sim \mathcal{B}(n, \frac{1}{2})$. Mais puisque $n = 37 > 35$, On fait l'approximation par la loi normale et

  $$Z = \frac{S - \frac{n}{2}}{\sqrt{\frac{n}{4}}} = \frac{2S - n}{\sqrt{n}} \sim \mathcal{N}(0,1)$$

  La région de rejet : $Z_c > u_{1-\alpha}$.
- Le nombre de différences positives $S = \sum_{i=1}^{n} \delta_i = 20$. Avec la correction de continuité

$$Z_c = \frac{2S - n - 1}{\sqrt{n}} = \frac{2(20) - 37 - 1}{\sqrt{37}} = 0.33$$

- Pour un niveau de signification 5% , on a $u_{1-\alpha} = 1.645$ et $Z_c < u_{1-\alpha}$. Donc on ne peut pas rejeter l'hypothèse nulle : gonfler les ballons à l'hélium ne permet pas de les envoyer plus loin.

3. Test des rangs signés de Wilcoxon

3.1. Test d'hypothèses, statistique de test et région critique

Le test des rangs signés de Wilcoxon traite la comparaison d'échantillons appariés. Il répond donc à la même catégorie de problèmes que le test des signes. Nous ne devons pas le confondre avec le test des rangs de Wilcoxon-Mann-Whitney pour échantillons indépendants, même si le mécanisme du test repose sur une somme de rangs.

Par rapport au test des signes, le test de Wilcoxon utilise l'importance relative des écarts lors de la définition de la statistique de test. Il est donc plus riche, plus puissant pour peu que l'on puisse les ordonner (les écarts).

Construction de la statistique de test

De nouveau, nous travaillons sur les écarts $d_i = x_{i1} - x_{i2}$, mais nous les exploitons de manière différente:

1. Nous construisons la valeur absolue des écarts $|d_i|$.

2. A partir de $|d_i|$, la valeur absolue des écarts, nous définissons les rangs des observations r_i. Ainsi la plus petite valeur des écarts, en valeur absolue, reçoit le rang 1 ; le plus grand écart, toujours en valeur absolue, se voit attribué le rang n.

3. A partir des rangs, nous pouvons calculer la statistique T^+, elle correspond à la somme des rangs pour les observations présentant un écart positif ($d_i > 0$)

$$T^+ = \sum_{i, d_i > 0} d_i$$

4. De la même manière, nous pourrions définir T^- pour les écarts négatifs. Mais sachant que la somme totale des rangs est égale à $\frac{n(n+1)}{2}$, nous pouvons déduire T^- à partir de la relation suivante :

$$T^- = \frac{n(n+1)}{2} - T^+ .$$

Sous H_0, les deux variables X_1 et X_2 ont une fonction de répartition identique, ou plus précisément, leurs paramètres de localisation sont équivalents, nous aurons l'égalité

$$T^+ = T^- = \frac{1}{2}[\frac{n(n+1)}{2}]$$

Plus T^+ sera grand (resp. petit) par rapport à T^-, plus nous serons amenés à conclure que les valeurs prises par X_1 sont stochastiquement plus élevés (resp. plus faibles) par rapport à celles de X_2.

Pour un test bilatéral, la région critique du test au risque α s'écrit

$$R.C.: (T^+ \leq \frac{n(n+1)}{2} - T_\alpha \ ) \quad ou \quad (T^+ \geq T_\alpha \ )$$

Où le seuil critique T_α est lue dans une table spécifique due a Wilcoxon

Exemple : (Effets de l'entraînement sur la tension artérielle).

On cherche à savoir si un entraînement régulier modifie la tension artérielle des personnes (test bilatéral). On a recueilli la tension systolique de *n* = 8 personnes, on leur a fait suivre un programme d'entraînement spécifique pendant 6 mois, puis on leur a de nouveau mesuré la tension.

Les données et les calculs sont réunis dans la table suivante :

N°	Avant	après	Ecart	|Ecart|	Rang|Ecart|	Rang signé
1	130	120	10	10	5	5
2	170	163	7	7	4	4
3	125	120	5	5	2	2
4	170	135	35	35	7	7
5	130	143	-13	13	6	-6
6	130	136	-6	6	3	-3
7	145	144	1	1	1	1
8	160	120	40	40	8	8

On a $n = 8$ paires d'observations et $T^+ = 27$.

- Dans la table des seuils critiques, pour un test bilatéral à 5%, nous lisons $T_{0.05} = 4$ pour un échantillon de taille *n* = 8. Manifestement, nous sommes dans la région critique. La tension artérielle n'est plus la même après un entraînement de 6 mois.

Remarques :

1. Les écarts nuls $d_i = 0$ posent des problèmes. La solution usuelle consiste tout simplement à supprimer les observations correspondantes.

2. La situation est différente lorsque nous avons des ex-aequo. Dans ce cas, plusieurs observations présentent une valeur identique de $|d_i|$, nous devons leur attribuer un rang moyen. La statistique de test T^+ n'est pas modifiée. Sa variance en revanche le sera. Ce qui aura un impact lorsque nous passons à l'approximation normale pour les grands échantillons, et qu'il faudra utiliser la statistique centrée et réduite *Z* pour définir la région critique.

Exemple (Influence des époux lors de la décision d'achat).

Nous reprenons l'exemple des époux qui jugent leurs influences respectives lors des importantes décisions d'achat. Nous souhaitons illustrer 2 configurations : comment exclure les observations à écarts nuls $d_i = 0$ des calculs ; comment mettre en place le principe des rangs moyens pour les ex-aequos.

Le tableau des opérations sera comme suit :

N°	mari	Femme	di	$\|d_i\|$	ri	r'i	Rang signé
E	3	3	0	0	-	-	-
H	3	3	0	0	-	-	-
O	5	5	0	0	-	-	-
B	4	3	1	1	1	3.5	3.5
D	6	5	1	1	2	3.5	3.5
F	2	3	-1	1	3	3.5	-3.5
I	1	2	-1	1	4	3.5	-3.5
J	4	3	1	1	5	3.5	3.5
M	4	5	-1	1	6	3.5	-3.5
A	5	3	2	2	7	8.5	8.5
C	6	4	2	2	8	8.5	8.5
L	4	2	2	2	9	8.5	8.5
P	5	3	2	2	10	8.5	8.5
G	5	2	3	3	11	11.5	11.5
K	5	2	3	3	12	11.5	11.5
Q	5	1	4	4	13	13	13
N	7	2	5	5	14	14	14

- Nous excluons les observations pour lesquelles $|d_i| = 0$, d'où $n = 14$,

La statistique de test

$$T^+ = \sum_{i, d_i > 0} d_i$$

- Sous H_0, $T^+ \sim \mathcal{W}ilcoxon$. D'après la table $T^+ = 94.5$.
- Pour un test bilatéral à 5%, le seuil critique du test pour $n = 14$ est égal à $T_{0.05} = 21$.
- R.C. : $\left(T^+ \leq \frac{14(15)}{2} - 21 = 84\right) \quad ou \quad \left(T^+ \geq 21\right)$
- Conclusion : Puisque $T^+ \geq 21$, Nous somme dans la région critique, très largement même, manifestement hommes et femmes n'évaluent pas de la même manière leurs influences respectives lors des décisions d'achat.

3.2 Approximation normale pour les grands effectifs

Approximation normale

Lorsque n est assez grand, ($n > 15$) en pratique, on peut approximer la distribution de T^+ sous H_0 par une loi normale de paramètres :

$$E(T^+) = \frac{1}{4}n(n+1) \quad et \quad V(T^+) = \frac{1}{24}n(n+1)(2n+1)$$

La statistique de test : $Z = \frac{T^+ - \frac{1}{4}n(n+1)}{\sqrt{\frac{1}{24}n(n+1)(2n+1)}} \sim \mathcal{N}(0,1)$

La région critique pour un test bilatéral au risque α s'écrit :

$$R.C.: |Z| \geq u_{1-\alpha/2}$$

Remarque : (Correction de continuité). Pour les effectifs modérés, nous pouvons améliorer la précision en introduisant une correction de continuité. Pour le test bilatéral, il s'écrit

$$|Z| = \frac{\left|T^+ - \frac{1}{4}n(n+1)\right| - 0.5}{\sqrt{\frac{1}{24}n(n+1)(2n+1)}}$$

Correction de la variance pour les ex-aequo

Lorsque les $|d_i|$ comportent des ex-aequo, nous devons corriger la variance de la statistique de test. Soit G le nombre de valeurs distinctes de $|d_i|$, pour la valeur $n^o g$, nous observons t_g observations, la variance corrigée s'écrit alors

$$\tilde{V}(T^+) = \frac{1}{24}n(n+1)(2n+1) - \frac{1}{48}\sum_{g=1}^{G} t_g\left(t_g - 1\right)(t_g + 1)$$

La statistique centrée réduite destinée à définir la région critique est également modifiée

$$\bar{Z} = \frac{T^+ - \frac{1}{4}n(n+1)}{\sqrt{\tilde{V}(T^+)}}$$

Exercice 1. Nous reprenons l'exemple "Shooté à l'hélium" étudié lors de la présentation de l'approximation normale de la statistique du test des signes. On cherche à savoir si les shoots vont plus loin lorsque l'on gonfle des ballons avec de l'hélium plutôt qu'avec de l'air.

- Trouver ce que nous propose le test des rangs signés.

4. Test de McNemar

Le test de McNemar s'applique très bien aux mesures de type "avant-après". La problématique de ce test non paramétrique serait celle d’un test de fréquence sur deux échantillons appariés. Prenons l’exemple d’un panel de consommateurs, avant et après une action de marketing. Ces braves gens répondent à une question binaire (oui ou non, j’achète ou pas, satisfait ou insatisfait…). Intéressons-nous aux consommateurs qui ont changé d’avis entre les deux enquêtes.

Pour cela, il faut comparer l'effectif qui est passé de « oui » à « non » à celui qui est passé de « non » à « oui ». L'hypothèse nulle, que l'on va tester, est que ces deux sous-populations se compen
sent peu ou prou.

4.1 Principe, statistique de test et région critique

Si π_1 (resp. π_2) est la probabilité $\Pr(X = 1)$ lors de la 1^{ere} (resp. 2^{nde}) mesure. Le test de McNemar peut s'écrire

$$H_0 : \pi_1 = \pi_2 \text{ vs } H_1 : \pi_1 \neq \pi_2$$

Il suffit donc d'un tableau de contingence à quatre cases pour procéder au test du type :

		Après		
		0	1	Total
Avant	0	a	b	a+b
	1	c	d	c+d
	Total	a+c	b+d	n

A la marge, nous avons les informations sur les proportions de "1" avant les modifications $\hat{\pi}_1 = \frac{c+d}{n}$ et après $\hat{\pi}_2 = \frac{b+d}{n}$. Nous disposons des informations sur les flux de $0 \rightarrow 1$ (b) et de $1 \rightarrow 0$ (c). Les marges sont identiques si ($a + b = a + c$) c.-à-d. ($b = c$) pour les "0". De la même manière pour les "1", identité des marges ($c + d = b + d$) entraîne ($c = d$). Le test sera donc essentiellement fondé sur la confrontation entre les valeurs b et c.
Le test d'hypothèses peut s'écrire comme une comparaison de probabilités sous cet angle, on oppose :

$$H_0 : \Pr(0 \rightarrow 1) = \Pr(1 \rightarrow 0) \quad \text{vs} \quad H_1 : \Pr(0 \rightarrow 1) \neq \Pr(1 \rightarrow 0)$$

La statistique du test d'écrit alors :

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

Sous H_0, elle est distribuée selon une loi du χ^2 à 1 degré de liberté. La région critique au risque α est tout naturellement :

$$R.C : \chi^2 \geq {\chi^2}_{1-\alpha}(1)$$

Remarque 1 (Correction de continuité). Sur les petits effectifs, surtout lorsque ($b + c$) est faible puisque la statistique repose essentiellement dessus, certains conseillent d'introduire la correction de continuité de Yates. Elle améliore l'approximation à l'aide de la loi du χ^2. La statistique de test s'écrit alors

$$\chi^2 = \frac{(|b-c|-1)^2}{(b+c)}$$

Nous avons aussi la possibilité, comme nous le verrons plus loin, de passer par les calculs exacts basés sur la loi binomiale.

Remarque 2 (D'autres applications). Le schéma "avant - après" est certainement le plus commode pour exposer le test de McNemar. Mais les applications sont plus larges. Il faut tout simplement que la variable d'intérêt soit binaire et que nous travaillons sur des échantillons appariés.

Exemple lors d'un audit social, un consultant interroge 200 salariés sur l'organisation du travail. La question est fermée : satisfait ou non. Trois mois après quelques réaménagements, la même question est posée. Peut-on considérer que les salariés perçoivent un réel changement ?
Voici les résultats :

	Oui avant	Non avant
Oui après	55	25
Non après	38	82

En regardant le tableau, on peut préciser le sens du changement et reformuler la question : les choses ont-elles empiré ?

Soit à tester les hypothèses suivantes : $H_0 : \pi_1 = \pi_2$ vs $H_1 : \pi_1 < \pi_2$

Où π_1 (resp. π_2) est la probabilité $\Pr(X = 1)$ lors de la 1^{ere} (resp. 2^{nde}) mesure de la satisfaction des salariés.

- Nous constatons maintenant que b = 25 individus qui étaient non satisfaits, sont maintenant satisfaits. A l'inverse, c = 38 qui étaient satisfait sont maintenant opposés.
- Formons la statistique du test en introduisant la correction de continuité, nous obtenons

$$\chi^2 = \frac{(|b-c|-1)^2}{(b+c)} = \frac{(|25-38|-1)^2}{(25+38)} = 2{,}286$$

- Au risque 5%, le seuil critique est défini par ${\chi^2}_{0.95}(1) = 3.8415$.
- Les données sont compatibles avec l'hypothèse nulle au niveau de signification 5%. Les réponses des salariés ne sont pas significativement différentes aux deux dates étudiées, c-à-d, ils ne perçoivent aucun changement.

4.2. Un approche non symétrique du test de McNemar

Calcul exact de la probabilité critique

On remarque que la statistique du test repose uniquement sur les valeurs de b et c. Les quantités a et b n'interviennent pas, nous pouvons les considérer comme fixes.

Notons B la variable aléatoire associée à l'évènement $(0 \to 1)$, que l'on mesure avec la quantité b dans le tableau de contingence. Les hypothèses à confronter dans le test de McNemar peuvent s'écrire de la manière suivante :

$$H_0 : \Pr(B) = \frac{1}{2} \quad \text{vs} \quad H_1 : \Pr(B) \neq \frac{1}{2}$$

Sous H_0, B est distribuée selon une loi binomiale de paramètres $B(m;\ 1/2\ )$, où $m = b{+}c$. La probabilité critique du test peut être calculée à l'aide de l'expression :

$$p = 2 \times \Pr[B \geq \max(b, m-b)]$$

Où

$$\Pr[\,B \geq \max(b, m-b)] = \sum_{j=\max(b,m-b)}^{m} C_m^j (0.5)^m$$

Exemple (Retour sur l'audit social). Dans notre exemple, *m* = 63 et *b* = 25. La probabilité critique basée sur la distribution binomiale s'écrit

$p = 2 \times \sum_{j=\max(25,38)}^{m} C_{63}^{j} (0.5)^{63} = 2 \times \sum_{j=38}^{63} C_{m}^{j} (0.5)^{63} = 2 \times 0.03846292 = 0.07692583$

qui est supérieur à 0.05.

Définir un test unilatéral

On se rend compte que la présentation usuelle du test de McNemar, un peu globalisée, masque les possibilités d'analyse fine des données. La formulation du test dans la section précédente nous éclaire mieux sur les informations que l'on peut extraire du tableau de contingence. Rien ne nous empêche en réalité de définir un test unilatéral du type

$$H_0 : \Pr(B) = \frac{1}{2} \quad \text{vs} \quad H_1 : \Pr(B) \geq \frac{1}{2}$$

La probabilité critique du test peut être alors calculée en se focalisation sur l'aile droite de la distribution de la loi binomiale, soit

$$p = \Pr[B \geq b]$$

Remarque (Approximation lorsque *m* est grand). Lorsque *m* est suffisamment élevé (en pratique *m* >10), nous pouvons utiliser l'approximation normale. Nous formons la quantité *Z*, en introduisant la correction de continuité

$$Z = \frac{\left(b - \frac{m}{2}\right) - 0.5}{\sqrt{\frac{m}{4}}} = \frac{(2b - m) - 1.0}{\sqrt{m}}$$

La probabilité critique du test peut être obtenue avec

$$p' = 1 - \Phi(Z) = 1 - \Phi\left(\frac{(2b - m) - 1.0}{\sqrt{m}}\right)$$

Où $\Phi(.)$ est la fonction de répartition de la loi normale centrée et réduite.

Exercice 2 :

Un directeur d'école souhaite sensibiliser les élèves de primaire à l'équilibre de l'alimentation. Il réalise une enquête dont il ressort qu'à la cantine, 18 élèves sur 52 (35% des élèves) choisissent leur menu en fonction de l'équilibre alimentaire et pas du seul critère de goût. Pendant deux semaines, une information est proposée dans les classes et pendant la demi-pension. Une seconde enquête est alors réalisée qui montre une augmentation de 23 points du pourcentage des élèves qui choisissent leur menu en fonction de l'équilibre alimentaire. Il y a en effet 30 élèves (58% des élèves) qui choisissent maintenant leur menu en tenant compte de l'équilibre alimentaire.

- Que penser de l'efficacité d'une telle campagne d'information sachant que 19 élèves, qui ne prenaient pas en compte l'équilibre alimentaire composant leurs repas, ont changé d'avis ?

Exercice 3 :

Nous disposons des notes données par deux correcteurs différents sur des copies de bac de mathématiques. On se demande si l'effectif d'élèves qui ont une note supérieure à la moyenne est dépendant du correcteur.

Le nombre d'élèves qui ont eu une note supérieur à 10 est donné dans la table (2×2) suivant deux variables sup2 et sup3 qui valent 1 lorsque la note de la copie est supérieure à 10, et 0 sinon.

Table de sup2 par sup3			
	Sup3		
Sup2	0	1	Total
0	13	8	21
1	0	9	9
Total	13	17	30

- Faire un test non paramétrique pour répondre à cette question ? (coursNP)

5. ANOVA de Friedman

5.1 Principe, statistique de test et région critique

L'analyse de variance (ANOVA) de Friedman consiste à comparer K paramètres de localisation sur K échantillons liés. Le tableau de données comporte donc n lignes, et K colonnes (pour nous ce sont des variables, on parle aussi de "traitements" dans la terminologie des plans d'expériences). Dans sa construction, chaque ligne correspond à un "K-uplet" de mesures. Ces dernières peuvent être continues ou ordinales, l'essentiel est que l'on puisse les exploiter de manière à produire un classement, c.-à-d., affecter des rangs aux traitements. Par hypothèse, les formes des distributions des K échantillons sont identiques, quand bien même elles seraient décalées, en particulier elles doivent présenter une dispersion identique.

Le test de Friedman est une généralisation du test des signes

Le test d'hypothèses s'écrit :

$$H_0 : \theta_1 = \cdots = \theta_K$$

$$H_1 : \exists k, k' \ \ tel\ que\ \theta_k \neq \theta_{k'}$$

Il s'agit bien d'un test de comparaison de K paramètres de localisation. Mais, à la différence du test de Kruskal-Wallis par exemple, les valeurs seront uniquement comparables à l'intérieur de chaque groupe.

Ainsi, nous travaillerons bien sur des rangs, mais calculés à l'intérieur des blocs. C'est ce qui différencie le test de Friedman de tous les tests pour échantillons indépendants où les rangs étaient toujours construits à partir de la totalité de observations mesurées.

Notre tableau des rangs qui servira au calcul de la statistique de test se présente comme suit

Bloc vs. Traitement	X_1	…	X_k	…	X_K	Somme
1	r_{11}	…	r_{1k}	…	r_{1K}	$R_{1.} = K(K+1)/2$
…						…
i	r_{i1}	…	r_{ik}	…	r_{iK}	$R_{i.} = K(K+1)/2$
…						…
n	r_{n1}	…	r_{nk}	…	r_{nK}	$R_{n.} = K(K+1)/2$
Somme	$S_{.1}$	…	$S_{.k}$	…	$S_{.K}$	-
moyenne	$\bar{R}_{.1}$	…	$\bar{R}_{.k}$	…	$\bar{R}_{.K}$	-

Première information importante, les rangs étant calculés à l'intérieur de chaque bloc, la somme des rangs d'un bloc est toujours égal à $R_{i.} = K(K+1)/2$.

Analyser cette quantité n'a aucun sens. Pour comparer l'efficacité des traitements, nous devrons donc nous intéresser à la somme des rangs par traitements $S_{.k}$. Ainsi, la statistique du test de Friedman s'écrit :

$$Fr = \frac{12n}{K(K+1)} \sum_{k=1}^{K} (\bar{R}_{.k} - \bar{\bar{R}})^2$$

Où $\bar{\bar{R}}$ est la moyenne globale des rangs.

Soit

$$\bar{\bar{R}} = \frac{\sum_i \sum_k r_{ik}}{nK} = \frac{nK(K+1)/2}{nK} = (K+1)/2$$

D'où

$$Fr = \frac{12n}{K(K+1)} \sum_{k=1}^{K} (\bar{R}_{.k} - \frac{\bar{R}+1}{2})^2$$

$$= \frac{12n}{K(K+1)} \sum_{k=1}^{K} \bar{R}_{.k}^2 - 3\mathrm{n}(\mathrm{K}+1) \quad \text{(Démontrer cette formule)}$$

Remarques:

1. Une autre écriture courante consiste à exprimer la statistique à partir de la somme des rangs comme suit :

$$Fr = \frac{12}{nK(K+1)} \sum_{k=1}^{K} S^2{}_{.k} - 3\mathrm{n}(\mathrm{K}+1)$$

2. Si $\bar{R}_{.k} = 1$, cela veut dire que le traitement est systématiquement le moins bon dans tous les blocs ; à l'inverse, si $\bar{R}_{.k} = K$, il est systématiquement le meilleur.

3. De fait, avec la première formulation, on comprend que *Fr* traduit l'idée d'une variabilité inter-traitements, elle correspond à la dispersion des rangs moyens conditionnels autour de la moyenne globale. Si les traitements se valent tous, nous obtiendrons *Fr* = 0. Plus ils se démarqueront les uns des autres, plus *Fr* prendra une valeur élevée.

4. On peut construire de manière différente la statistique *Fr* qui est comme un rapport de variance inter-traitement et de variance intra-traitements sur les rangs.

La région critique du test correspond donc aux grandes valeurs de Fr, soit, au risque α:

$$R.C: Fr \geq Fr_{\alpha,k,n}$$

Où $Fr_{\alpha,k,n}$ est lue dans la table des valeurs critiques de la statistique de Friedman

Remarque: (Anova de Friedman, généralisation du test des signes). Le parallèle entre le test des signes et le test de Friedman est possible. Il a été montré en effet pour $K = 2$ que

$$Fr = \frac{4}{n}(S - \frac{n}{2})^2$$

Où S est la statistique du test des signes. De ce point de vue, on peut considérer que le test de Friedman est une généralisation du test des signes qui s'applique pour la comparaison de $K \geq 2$ populations.

Exemple (Comparaison de marques de pneumatiques). On souhaite comparer la longévité de 4 marques de pneumatiques (traitement). Nous leur faisons parcourir 10 scénarios de parcours routier (blocs). La variable d'intérêt est la distance parcourue jusqu'à une certaine limite d'usure. Les données sont présentées dans le tableau suivant :

		Pneumatiques (Traitements)			
		A	F	G	R
Blocs	1	38	29	41.5	39
	2	24.5	36	35.5	26
	3	37.5	38.5	31.5	29.5
	4	20.5	33.5	29.5	21.5
	5	29.5	35	34	11
	6	22	33	35	17
	7	29	37.5	38	18.5
	8	25	21.5	28	18
	9	26	35.5	28	17
	10	17	23.5	34	16.5

- Nous devons tester l'égalité de $K=4$ moyennes avec un test non paramétrique, c-à-d, tester

$$H_0: \theta_1 = \cdots = \theta_4$$
$$H_1: \exists k, k' \ tel\ que\ \theta_k \neq \theta_{k'}$$

- Nous disposons de $n = 10$ observations (blocs), et nous souhaitons comparer $K = 4$ marques de pneumatiques (traitements).
- Première étape indispensable, nous devons transformer les données. Les rangs sont calculés à l'intérieur des blocs.

Rangs			
A	F	G	R
2	1	4	3
1	4	3	2
3	4	2	1
1	4	3	2
2	4	3	1
2	3	4	1
2	3	4	1
3	2	4	1
2	4	3	1
2	3	4	1

$S_{.k}$	20	32	34	14
$\bar{R}_{.k}$	2.0	3.2	3.4	1.4
$S_{.k}^2$	400	1024	1156	196

$$\sum_{k=1}^{K} S^2{}_{.k} = 400 + 1024 + 1156 + 196 = 2776$$

- Nous pouvons maintenant produire la statistique du test en utilisant la deuxième formule :

$$Fr = \frac{12}{nK(K+1)} \sum_{k=1}^{K} S^2{}_{.k} - 3\text{n}(\text{K}+1) = \frac{12}{10 \times 4(4+1)} 2776 - 3 \times 10(4+1) = 16.56$$

(Trouver cette valeur en utilisant la première formule)

- Le seuil critique du test à 5% est égal à 7.68 (voir table statistique de Friedman). Les marques de pneumatiques ne présentent pas des performances identiques au risque 5%.
- Pour un risque à 1%, le seuil est à 10.68 dans la table, nous pouvons en déduire le même résultat. On en conclut qu'il y a une différence de performances très significative entre les marques.

5.2 Approximations pour les grands échantillons

Lorsque n ou K sont assez grands, en pratique (n > 15 ou K > 4), la distribution de Fr peut être approximée par une loi du χ^2 à (K - 1) degrés de liberté. La région critique sera ainsi :

$$R.C : Fr \geq \chi^2_{1-\alpha}(K-1)$$

où $\chi^2_{1-\alpha}(K-1)$ est le quantile d'ordre $1-\alpha$ de la loi du χ^2 à (K - 1) degrés de liberté.

Exemple : (Approximation χ^2- Tester la longévité des pneumatiques). Revenons sur notre exemple ci-dessus "Longévité des pneumatiques". Nous sommes à la lisière des conditions d'une bonne approximation avec n = 10 et K = 4. Essayons néanmoins, ne serait-ce que pour comparer les conclusions du test.

Revenons sur notre feuille de calcul.

Le nombre de degrés de liberté est ddl = K - 1 = 4 - 1 = 3. Le seuil critique au risque 1% est le quantile $\chi^2_{0.99}(3) = 11.34$. Nous sommes dans la région critique Fr = 16:56 > 11.34. Les résultats restent cohérents avec l'approche exacte.

5.3 Traitement des ex-aequo

Lorsqu'il y a des ex-aequo à l'intérieur d'un bloc, le principe des rangs moyens est utilisé, c.-à-d., on réalise la péréquation des rangs affectés aux traitements qui présentent la même valeur. La statistique de test doit être corrigée pour tenir compte de l'ajustement, avec les caractéristiques suivantes : la statistique corrigée est toujours supérieure ou égale à la statistique non corrigée ; s'il n'y a pas aucun ex-aequo, la statistique corrigée et non corrigée doivent être identiques ; la correction sera d'autant plus sensible que le nombre d'ex-aequo est élevé.

Pour le bloc $n^o i$: nous notons G_i le nombre de valeurs différentes, la valeur $n^o g$ étant répétée t_{ig} fois. La statistique de Friedman ajustée pour les ex-aequo s'écrit alors

$$\widetilde{Fr} = \frac{Fr}{1 - \frac{\sum_{i=1}^{n} \sum_{g=1}^{Gi} (t_{ig}{}^{3} - t_{ig})}{n(K^3 - K)}}$$

La quantité au dénominateur est toujours $\leq\ 1$, on comprend aisément que $\widetilde{Fr} \geq\ Fr$. Sous H_0, les lois asymptotiques de $\widetilde{Fr}$ et Fr sont identiques.

Exemple (Torches de soudage).

Six (*n* = 6) soudeurs avec différents niveaux de savoir faire ont été invités à souder deux tuyaux à l'aide cinq (*K* = 5) différentes torches de soudage. Les torches ont été affectées dans un ordre aléatoire pour chaque soudeur. La qualité de soudure a été évaluée par un juge (note allant de 1 à 10, 10 correspondant à la meilleure évaluation). On cherche à savoir si la qualité de soudure dépend du type de torche utilisé.

Ces données nous intéressent surtout parce qu'elles comportent des ex-aequo dans certaines lignes du tableau. Nous utilisons les rangs moyens dans ce cas, et nous devrons également corriger la statistique de test.

		Traitement				
	Soudeur	Torche1	Torche2	Torche3	Torche4	Torche5
Blocs	1	3.9	4.1	4.2	4.1	3.3
	2	9.4	9.5	9.4	9.0	8.6
	3	9.7	9.3	9.3	9.2	8.4
	4	8.3	8.0	7.9	8.6	7.4
	5	9.8	8.9	9.0	9.0	8.3
	6	9.9	10.0	9.7	9.6	9.1

- Testons les hypothèses :

$$H_0 : \theta_1 = \cdots = \theta_5$$
$$H_1 : \exists k \neq k' \ ;\ tel\ que\ \theta_k \neq \theta_{k'}$$

- Nous disposons de *n* = 6 observations (blocs), et nous souhaitons comparer *K* = 5 marques de torches (traitements).
- Les rangs sont calculés à l'intérieur des blocs.

		Rangs				
	Soudeur	Torche1	Torche2	Torche3	Torche4	Torche5
Blocs	1	2	3.5	5	3.5	1
	2	3.5	5	3.5	2	1
	3	5	3.5	3.5	2	1
	4	4	3	2	5	1
	5	5	2	3.5	3.5	1
	6	4	5	3	2	1
	$S_{.k}$	23.5	22	20.5	18	6
	$\bar{R}_{.k}$	3.92	3.67	3.42	3	1
	$S_{.k}^2$	552.25	484	420.25	324	36

On a $\sum_{k=1}^{K} S^2{}_{.k} = 1816.5$

- La statistique du test en utilisant la deuxième formule et puisque K $= 5 > 4$, *nous pouvons utiliser la loi asymptotique* :

$$Fr = \frac{12}{nK(K+1)}\sum_{k=1}^{K} S^2{}_{.k} - 3\mathrm{n}(\mathrm{K}+1) \sim \chi^2\,(K-1)$$

La région critique : $Fr \geq \chi^2_{0.95}(4)$

$$Fr = \frac{12}{nK(K+1)}\sum_{k=1}^{K} S^2{}_{.k} - 3\mathrm{n}(\mathrm{K}+1) = \frac{12}{6\times 5(5+1)}1816.5 - 3\times 6(5+1)$$
$$= 13.1$$

- On a des ex-aequo, donc il faut faire la correction. Par exemple détaillons les calculs pour le premier bloc. Nous lisons les différentes valeurs et leurs nombre d'apparition :$\{3.3(1), 3.9(1), 4.1(2), 4.2\,(1)\}$. Nous pouvons former :

$$\sum_{g=1}^{4}(t_{ig}{}^3 - t_{ig}) = (1^3-1)+(1^3-1)+(2^3-2)+(1^3-1) = 6$$

- Le facteur de correction est

$$C = 1 - \frac{\sum_{i=1}^{n}\sum_{g=1}^{Gi}(t_{ig}{}^3 - t_{ig})}{n(K^3-K)} = 1 - \frac{24}{6(5^3-5)} = 0.9667$$

	Calcul de $(t_{ig}{}^3 - t_{ig})$	
Blocs	1	6
	2	6
	3	6
	4	0
	5	6
	6	0
	Total	24

- La statistique corrigée :

$$\widetilde{Fr} = \frac{Fr}{C} = \frac{13.1}{0.9667} = 13.5517$$

Conclusion : Le seuil critique au risque 5% est $\chi^2_{0.95}(4) = 9.4877$. Nous nous situons dans la région critique,

$$\widetilde{Fr} = 13.5517 > 9.4877 = \chi^2_{0.95}(4)$$

Nous rejetons l'hypothèse nulle d'influence égale des torches sur la qualité de la soudure.

Exercice 4 :

Quatre membres d'une équipe d'athlétisme sont classés par l'entraîneur-chef en fonction de leur capacité sur six épreuves d'athlétisme. Pour chaque épreuve, l'entraîneur attribue un rang de 1 à l'athlète qui est le meilleur à l'épreuve et un rang de 4 à l'athlète qui est le pire à l'épreuve. Le tableau suivant résume les données de l'étude.

Événement	Athlète			
	Athlète 1	Athlète 2	Athlète 3	Athlète 4
Sprint	3	2	1	4
1500 mètres	3	2	1	4
Saut à la perche	3	2	1	4
Long saut	4	2	1	3
Lancer du poids	3	2	1	4
400 mètres	4	1	2	3

Y a-t-il un lien significatif entre les classements attribués aux athlètes sur les six épreuves?

6. Test Q de Cochran pour $K \geq 2$ populations

6.1 Principe, statistique de test et région critique

Le test Q de Cochran est une généralisation du test de McNemar où l'on traite $K \geq 2$ échantillons appariés, dans un plan d'expérience en bloc aléatoire complet. Il s'applique aussi au cas des mesures répétées c.-à-d. un seul échantillon dans laquelle une variable a été mesurée K fois. La variable d'intérêt est binaire. On peut le voir comme une variante du test de Friedman.

Nous disposons de K mesures binaires, prenant leurs valeurs dans $(1, 0)$. Nous les noterons X_k; $k = 1, \ldots, K$. Il s'agit de réaliser un test de comparaison de proportions :

$$H_0 : \pi_1 = \cdots = \pi_K \text{ vs } H_1 : Une\ au\ moins\ diff\grave{e}re\ des\ autres$$

où π_k représente la probabilité d'occurrence de la valeur "1" pour la variable X_k :

$$\pi_k = \Pr(X_k = 1).$$

Nous travaillons sur un tableau de données de la forme suivante :

.	Traitement					.
Bloc	X_1	…	X_k	…	X_K	Somme
1	x_{11}	…	x_{1k}	…	x_{1K}	L_1
$\vdots$	$\vdots$		$\vdots$	$\vdots$	$\vdots$	$\vdots$
i	x_{i1}	…	x_{ik}	…	x_{iK}	L_i
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
n	x_{n1}	…	x_{nk}	…	x_{nK}	L_n
Somme	C_1	…	C_k	…	C_K	S

Où

- x_{ik} prend la valeur 1 ou 0 ;
- $C_k = \sum_i x_{ik}$ est le nombre d'apparition de la valeur 1 dans la k^{eme} colonne ;
- $L_i = \sum_k x_{ik}$ est le nombre d'apparition de la valeur 1 dans la *ieme* ligne ;

- $S = \sum_k C_k = \sum_i L_i = \sum_i \sum_k x_{ik}$ est la somme totale des valeurs du tableau.
  La statistique du test de Cochran est définie de la manière suivante :

$$Q = K(K-1)\frac{\sum_{k=1}^{K}\left(C_k - \frac{S}{K}\right)^2}{\sum_{i=1}^{n} L_i(K - L_i)}$$

Remarques : Cette formule nous éclaire sur le mode de fonctionnement du test :

1- Si $\pi_k = \pi;\ \forall k$, c.-à-d., tous les traitements se valent, nous aurons, aux fluctuations d'échantillonnage près, $C_k = \frac{S}{K}$, alors $Q = 0$.

2- Plus les traitements induiront des effets différents, plus la disparité des C_k autour de $\frac{S}{K}$ sera grande, induisant une augmentation de Q.

3- On retrouve également l'écriture suivante dans la littérature :

$$Q = \frac{(K-1)[KC_k^2 \sum_{k=1}^{K} - (\sum_{k=1}^{K} C_k)^2]}{K\sum_{i=1}^{n} L_i - \sum_{i=1}^{n} L_i{}^2}$$

Sous H_0, Q suit asymptotiquement une loi du χ^2 à $(K - 1)$ degrés de liberté. L'approximation est de bonne qualité dès que pour ($n \geq 4$) et ($n \times K > 24$). La région critique du test au risque α s'écrit naturellement :

$$R.C : Q \geq \chi^2_{1-\alpha}(K-1)$$

Exemple : (Difficultés pour alpinistes).

$K = 3$ ascensions différentes sont tentées par $n = 5$ alpinistes. Les succès sont notés "1", les échecs "0". On cherche à savoir si les ascensions sont de difficultés égales. En d'autres termes, on veut comparer la probabilité de réussir les différents types d'ascensions.

individu	ascension	ascension	ascension	L_i	$L^2{}_i$	
1	1	1	0	2	4	
2	1	0	1	2	4	
3	0	0	1	1	1	
4	0	1	1	2	4	
5	1	0	1	2	4	
C_k	3	2	4	9	17	Somme
C_k^2	9	4	16			

On va comparer les difficultés de $K = 3$ ascensions en utilisant le Test Q de Cochran pour tester les hypothèses: $H_0 : \pi_1 = \pi_2 = \pi_3$ vs $H_1 : Une\ au\ moins\ diff\grave{e}re\ des\ autres$ où π_k représente la probabilité d'occurrence de la valeur "1" pour la variable X_k :

$$\pi_k = \Pr(X_k = 1).$$

- La statistique du test est

$$Q = \frac{(K-1)\left[K\sum_{k=1}^{K} C_k^2 - \left(\sum_{k=1}^{K} C_k\right)^2\right]}{K\sum_{i=1}^{n} L_i - \sum_{i=1}^{n} {L_i}^2} = \frac{(3-1)[3(9+4+16)-(3+2+4)^2]}{3\times 9-17} = \frac{12}{10} = 1.2$$

- Sous H_0, Q suit une loi du χ^2 à (K – 1=2) degrés de liberté. Le seuil critique du test au risque α = 5% est $\chi^2_{0.95}(2) = 5.9915$. Les données sont compatibles avec l'hypothèse nulle. Les ascensions sont de difficultés égales.
-

Exercice 5:

On a interrogé K = 3 échantillons de n = 18 familles, elles devaient répondre "oui" (1) ou "non" (0) à une question. Dans chaque échantillon, l'interview a été menée d'une manière différente (enthousiaste, courtoise, détachée). On veut vérifier que la proportion de "oui" est la même dans chaque échantillon, c.-à-d., on veut savoir si la manière de mener une interview influence les réponses. Les réponses sont données dans le tableau suivant :

Bloc	Interview1	Interview2	Interview3
1	0	0	0
2	1	1	0
3	0	1	0
4	0	0	0
5	1	0	0
6	1	1	0
7	1	1	0
8	0	1	0
9	1	0	0
10	0	0	0
11	1	1	1
12	1	1	1
13	1	1	0
14	1	1	0
15	1	1	0
16	1	1	1
17	1	1	0
18	1	1	0

Exercice 6

On a demandé à n = 12 femmes si elles étaient d'accord pour l'achat de véhicule d'une marque spécifique. K = 3 marques ont été proposées [Susuki (A), Mercedes (B) et Hyundai(C)], à chaque fois la personne interrogée devait répondre oui (1) ou non (0).
Les réponses étaient comme suit :

i	Susuki (A)	Mercedes (B)	Hyundai(C)
1	1	1	0
2	0	1	0
3	1	1	1
4	0	1	0
5	0	1	0
6	0	1	1
7	0	0	0
8	0	1	0
9	1	1	0
10	0	1	0
11	0	0	0
12	0	0	1

Peut-on conclure qu'il existe des différences en ce qui concerne la préférence de voiture en fonction des réponses des sujets?

Est-ce qu'on peut savoir si ces personnes sont plus attirées par une des marques en particulier ?