TUTORAT UE4 2012-2013 – Biostatistiques
CORRECTION Séance n°5 – Semaine du 22/10/2012

Tests pour variables qualitatives- Tests non paramétriques

Séance préparée par Chayma IGHIDI, Guilhaume GARRIGUES, Laura MILHAU
et Grégoire SARTHOULET

__QCM n°1:__ **A, C, D**

A. Vrai.
B. Faux, uniquement si les conditions d'application du test paramétriques sont validés.
C. Vrai.
D. Vrai.
E. Faux. Plutôt quand l'échantillon est petit.

__QCM n°2:__ **C, D**

A. Faux. Les données sont appariées.
B. Faux. On fait un test de Wilcoxon sur données appariées.
C. Vrai.
D. Vrai.
E. Faux. H_0: l'étudiant en terminale travaille autant qu'en PACES
H_1 : L'étudiant en PACES étudie plus qu'en terminale

	1	2	3	4	5	6	7	8	9
Terminale	1	2	1,5	1	2,5	3	4	1,5	0,5
PACES	6	7	5	1	2	9	3	5,5	6,5
Différences	-5	-5	-3,5	0	0,5	-6	1	-4	-6
Rang	5,5	5,5	3		1	7,5	2	4	7,5

- On calcule tout d'abord les différences x-y puis on ordonne en ordre croissant sans tenir compte du signe
- Puis on affecte un rang sans compter les 0
- Calcul de la somme des rangs + et des rangs - :
  $S+= 1+2=3$ et S-=5,5+5,5+3+7,5+4+7,5=33 donc on a $S_{min}=3$
- On a 8 différences non nulles, on lit dans la table de Wilcoxon à 5% que $T_8=4>S_{min}$ donc on rejette H_0.

__QCM n°3 :__ **A**

A. **Vrai,** définition du cours (c'est la seule réponse possible pour ce qcm)
B. Faux. Mann-Whitney
C. Faux. Test des signes
D. Faux. Test de l'analyse de la variance ANOVA.
E. Faux. Test F

__QCM n°4 :__ **C, D**

A. Faux
B. Faux. Au risque 5% $S_{seuil}=11>4=S_+$(le plus petit des deux S) donc on rejette H_0 (Il y a une différence).
C. **Vrai**. A 2% $S_{seuil}=7> S_+$ donc on rejette H_0.
D. **Vrai**. A 1% $S_{seuil}=5> S_+$ donc on rejette H_0.
E. Faux c'est le principe des tests non paramétriques de travailler sur de petits effectifs.

QCM n°5 : **D**

A. Faux. Rf= 32 et RNf= 59.
B. Faux : Useuil = 6 (lu dans la Table de Mann Whitney à 5% pour n1=6 et n2-n1=1)
C. Faux : cf. B
D. **Vrai**. Uf= U_1= $n_1 . n_2 + \frac{n1(n1+1)}{2}$ – Srg_1 $U_{nf} = U_2 = n_1.n_2 + \frac{n2(n2+1)}{2}$ – Srg_2

U_f= 38 et Unf= 4 donc U= U_{min} (U_1, U_2)= 4 <Useuil.

Donc U<U_{seuil}, On rejette alors Ho : il ya une différence significative entre le groupe des fumeurs et le groupe des non-fumeurs.

E. Faux. Cf. item D.

QCM n°6 : **B, C, E**

A. Faux. On interroge le même échantillon de P2 avant et après leur WEI. On compare ici plusieurs distributions observées sur séries appariées. Les groupes ne sont donc pas indépendants.
B. **Vrai**. On compare deux variables qualitatives à 2 modalités sur des séries appariées, on utilise donc le test de Mac Nemar.
C. **Vrai**. Donc condition validée ici.
D. Faux. z=$\frac{(f-g)^2}{f+g}$
E. **Vrai**. z=$\frac{(16-5)^2}{16+5} = 5{,}7619$ On a $t_\alpha = 3{,}841$ (lu à 5% dans la table du chi-deux à 1 ddl) donc z>t_α donc on rejette H_0 .
<u>Remarque</u> : on ne prend en compte que les paires discordantes pour le calcul de z.
Avec H_0: les 2 sondages montrent que les P2 ont la même idée du WEI avant et après l'avoir fait.
Et H_1: les 2 sondages montrent que les p2 ont vu leur idée sur le WEI évoluer avant et après l'avoir fait.

QCM n°7: **A, B**

Ho: Fa = Fb.
H1: Fa<Fb.

A. **Vrai**. les 2 groupes sont issus d'échantillons indépendants.
B. **Vrai**. RA= 64.5 et RB= 106.5
C. **Faux,** pas d'a priori sur le sens de la différence
D. Faux. Ua= 9*9 + $\frac{9*10}{2}$ - 64.5= 61.5 ; Ub= 9*9 + $\frac{9*10}{2}$ - 106.5 = 19.5 ; Useuil = 17
E. Faux. Umin (Ua, Ub) = 19.5. U> $Useuil$ donc on ne rejette pas Ho et on ne peut conclure que les différences de résultats influencent significativement les résultats.

QCM n°8 : **F**

A. Faux. Ils sont appariés.
B. Faux. On utilise un test des signes. Dans l'énoncé, on considère que m est petit donc on ne peut pas utiliser l'approximation de la loi binomiale par la loi normale. En revanche si m avait été grand, on aurait pu faire l'approximation par la loi normale de paramètres $\mu = \frac{1}{2}$ et $\sigma = \frac{1}{2}\sqrt{m}$.
C. Faux. la paramètre m de la table correspond au nombre de paires dont la différence est non nulle, on a donc ici $m = 9$
D. Faux. dans la table de la loi binomiale on trouve P$(T \leq 1) = 0{,}020$. Les résultats dans la table sont donnés en unilatéral, en bilatéral on doit donc multiplier par 2 la probabilité donc on obtient une probabilité de 0,040.
E. Faux. On comptabilise le nombre de « + » et le nombre de « -» et on désigne par T le plus petit des 2 $\sum(+) = 1$ $\sum(-) = 80$, donc T=1, on a vu dans la table que P$(T \leq 1) = 0{,}020$. Les résultats dans la table sont donnés en unilatéral, en bilatéral on doit donc multiplier par 2 la probabilité donc on obtient une probabilité de 0,04. 0.040<0,05 donc on rejette l'équivalence pour $\alpha = 0{,}05$.
F. **Vrai**.

QCM n°9 : **A, B, E.**

A. **Vrai**. On étudie le lien entre deux variables qualitatives sur échantillons indépendants.
B. **Vrai.** On peut reformuler la question comme la comparaison entre la proportion de diabétiques chez les gros consommateurs de fast-food et la proportion de diabétiques chez les autres personnes.
C. Faux, on calcule les effectifs théoriques sous l'hypothèse H0 (en italique dans le tableau suivant) :

	Diabète.	Pas de diabète	Total
Mc.	60 / *57.96*	32 / *34.04*	92
Non Mc.	3 / *5.04*	5 / *2.96*	8
Total	63	37	100

Il y a une valeur théorique qui est inférieure à 5 donc on ne peut pas utiliser ce test.
D. Faux. On estime la proportion globale de diabétiques : $p = \frac{60+3}{100} = 0.63$. On calcule ensuite $n_M p = 92 \times 0.63 = 57.96, n_{NM} p = 8 \times 0.63 = 5.04,\ n_M(1-p) = 92 \times 0.37 = 34.04\ et\ n_{NM}(1-p) = 8 \times 0.37 = 2.96$. On a 2.96<5 donc on ne peut utiliser ce test.
E. **Vrai**. On peut utiliser le test exact de Fisher quand on souhaite comparer deux proportions issues de deux populations indépendantes et que les conditions de validité du test du khi-2 et de l'écart-réduit ne sont pas remplies.
F. Faux.

<u>QCM n°10:</u> **A, B, C, D, E.**

A. **Vrai**.
B. **Vrai**. le nombre de filles dans une famille de 5 personnes suit une loi binomiale de paramètre n=5 (nombre de répétitions indépendantes de l'événement) et si on suppose que l'on a équiprobabilité la probabilité du succès est 0.5.
C. **Vrai.** cf. B.
D. **Vrai**. On calcule P(X=0) jusqu'à P(X=5) ou on les trouve grâce à la table de la loi binomiale (pour m=5) :

P(X=0)	P(X=1)	P(X=2)	P(X=3)	P(X=4)	P(X=5)
0.031	0.188-0.031=157	0.500-0.188=0.344	0.656	0.891	0.984

On en déduit les effectifs théoriques par multiplication des probabilités par l'effectif (320)

	A	**B**	**C**	**D**	**E**	**F**
Filles	**0**	**1**	**2**	**3**	**4**	**5**
Total	**18 / *9.92***	**56 / *50.24***	**110 / *99.84***	**88 / *99.84***	**46 / *50.24***	**8 / *9.92***

Donc le calcul de la statistique du X2 est :

$$\frac{(18-9.92)^2}{9.92} + \frac{(56-50.24)^2}{50.24} + \frac{(110-99.84)^2}{99.84} + \frac{(88-99.84)^2}{99.84} + \frac{(46-50.24)^2}{50.24} + \frac{(8-9.92)^2}{9.92} = 10.409$$

E. **Vrai**. le degré de liberté correspond au nombre de classes -1, ici 6-1=5, la valeur limite du X^2 est donc égale à 11.070 pour un ddl de 5 et un risque de 0.05%. On conserve donc H0 au risque de 5%.
F.

<u>QCM n°11 :</u> **A, B, E.**

A. **Vrai**. Le ddl est de 2=[(3-1)*(2-1)]
B. **Vrai**. Par le calcul on trouve bien que X2=9.348 de plus tous les Cij sont supérieurs à 5.

	Hypertension	**Tension Normale**	**Hypotension**	**Total**
Hommes	**72 / 66**	**192 / 186**	**36 / 48**	**300**
Femmes	**38 / 44**	**118 / 124**	**44 / 32**	**200**
Total	**110**	**310**	**80**	**500**

$$X2 = \frac{(72-66)^2}{66} + \frac{(38-44)^2}{44} + \frac{(192-186)^2}{186} + \frac{(118-124)^2}{124} + \frac{(36-48)^2}{48} + \frac{(44-32)^2}{32} = 9.348$$

X2 alpha se lit dans la table du Chi deux pour un ddl de 2 et un risque de 5% soit une valeur de 5.991.
C. Faux. Ho est rejetée car X2 est supérieure à X2 alpha
D. Faux. H1 est acceptée
E. **Vrai.** Par lecture dans la table X2 alpha=5.991 donc Ho est rejeté.

QCM n°12 : **A, C, D, E**

A. **Vrai**.
B. Faux. Nous avons à faire à des variables ordinales, on utilise donc le test des signes.
C. **Vrai**.
D. **Vrai**. $Soit\ H0 : le\ film\ ne\ change\ pas\ les\ opinions.$
$\sum + = 59\ et\ \sum - = 26. On\ en\ déduit\ T = 26 (le\ plus\ petit\ des\ deux)$
$8 + 7 = 15\ personnes\ qui\ ne\ changent\ pas\ d'avis.$
$m = 100 - 15 = 85$
n trop grand $\longrightarrow$ $pas\ dans\ la\ table : il\ faut\ approximer\ par\ une\ loi\ normale.$
T $\longrightarrow N\left(\frac{1m}{2} ; \frac{1\sqrt{m}}{2}\right) \longrightarrow N$ (42,5 ;4,61)
$On\ calcule\ Z = \frac{|(T+0,5)-\mu|}{|\sigma|} = 3,47$. (+0,5 pour la correction de continuité)
$Dans\ la\ table\ de\ l'écart\ réduit\ on\ trouve\ t\alpha = 1,96\ (risque\ 5\%)$.
$Z > t\alpha$ Donc on rejette H_0.
E. **Vrai**. Cf. D