

TUTORAT UE4 2011-2012 – Biostatistiques

CORRECTION Séance n°4 – Semaine du 17/10/2011

Tests Statistiques – Molinari

Séance préparée par Zahir AKKARI et Valentin DUFFIEUX

QCM n°1 : A, C, D :

- a) **Vrai**
- b) Faux : voir c)
- c) **Vrai** : on fait un test en unilatéral.
- d) **Vrai** : car on a un a priori sur le résultat.
- e) Faux
- f) Faux

QCM n°2 : A, B

- a) **Vrai**
- b) **Vrai** : On fait un test bilatéral.
- c) Faux : $t_{obs} = 4.19$
Les deux échantillons sont de taille supérieure à 30. On peut donc faire un test de l'écart-réduit.

$$T_{obs} = \frac{(mp - ml)}{\sqrt{\frac{sp^2}{np} + \frac{sl^2}{nl}}} = 4.19$$

- d) Faux : pvalue < 0.0001, par lecture inverse de la table de l'écart-réduit.
- e) Faux : $t_{obs} > t_{\alpha}$ (ou p-value < α) donc on rejette H_0
On conclut que la consommation de café en PACES est plus forte qu'en fac de lettres, le test nous montre la différence entre les deux moyennes, pas que le facteur causal est le stress.
- f) Faux

QCM n°3 : A, C, D

- a) **Vrai**
- b) Faux : $Z = 5,06$
 $N > 30$, on peut utiliser un test de l'écart-réduit.

$$Z = \frac{\bar{X} - \mu_0}{\frac{S}{\sqrt{n}}} = \frac{140 - 120}{\frac{25}{\sqrt{40}}} = 5,06$$

- c) **Vrai** car $Z = 5,06 > 1.96 = t_{\alpha}$ (ou pvalue < α)
- d) **Vrai**, c'est demander un intervalle de confiance de niveau 95 % de la moyenne
$$IC = \left[\bar{x} - C_{\alpha/2} \frac{S}{\sqrt{n}} ; \bar{x} + C_{\alpha/2} \frac{S}{\sqrt{n}} \right] = \left[140 - 1.96 \frac{25}{\sqrt{40}} ; 140 + 1.96 \frac{25}{\sqrt{40}} \right] = [132,3 ; 147,7]$$
- e) Faux, il n'y a pas besoin de faire de calcul. Avec un intervalle de confiance de niveau de 90 %, on doit obtenir un intervalle moins large que le précédent, car l'on prend plus de risques de ne pas avoir la valeur de la moyenne dans l'intervalle. Or l'intervalle proposé dans l'item est plus large que celui calculé dans le d). Pour info, l'intervalle de confiance est [133,5 ; 146,5] avec un $C_{\alpha/2} = 1,645$.

QCM n°4 : B, E

- a) Faux : H_0 : Le fluor ne modifie pas le pourcentage d'enfants indemnes de caries dentaires.
b) **Vrai** : On prend un a priori, on fait un test en unilatéral
c) Faux : On cherche dans la table de la loi normale centrée réduite la valeur de u correspondant à $P(X \leq u) = 1 - 0,05$ OU On cherche dans la table de l'écart-réduit le u correspondant à un α de 10 % (en effet, la table de l'écart-réduit est en bilatéral, et 5 % en unilatéral correspond à 10% en bilatéral)
On obtient dans les deux méthodes, $u = 1,65$ (c'est un « t_α »)

- d) Faux : $t_{obs} = 28,15$

$$p = \frac{n_1 \times p_2 + n_2 \times p_1}{n_1 + n_2} = \frac{0,056 \times 12000 + 0,178 \times 9000}{12000 + 9000} = 0.108$$

à 10^{-3} près. Mais pour le calcul du t_{obs} , il faudra utiliser la valeur précise de p (La touche ANS de la calculatrice est mon amie !)

on a bien $n_1 p = 12000 \times 0,11 = 1320 > 5$; $n_1 q = 12000 \times (1 - 0,11) = 10680 > 5$;

$n_2 p = 990 > 5$ et $n_2 q = 8010 > 5$

On peut appliquer le test de l'écart-réduit.

$$t_{obs} = \frac{p_1 - p_2}{p \times (1 - p) \times \left(\frac{1}{n_1} + \frac{1}{n_2} \right)} = 28,15$$

- e) **Vrai** : $t_{obs} = 28.15 > 1,65 = t_\alpha$, donc on rejette H_0 , et on accepte H_1 .
Attention, on ne peut pas pour autant en conclure que cette amélioration est due à la mise sur le marché du sel fluoré. D'autres facteurs causals sont peut être responsables.
f) Faux

QCM n°5 : A, B, D

- a) **Vrai**

b) **Vrai** car tous les effectifs théoriques obtenus sous H_0 sont supérieurs ou égaux à 5.

$$\frac{500 \cdot 200}{2000} = 50 \quad ; \quad \frac{500 \cdot 1800}{2000} = 450 \quad ; \quad \frac{1500 \cdot 200}{2000} = 150 \quad ; \quad \frac{1800 \cdot 1500}{2000} = 1350$$

c) Faux, si la statistique de test est égale à 9,6, $0,001 < p\text{-value} < 0,01$

d) **Vrai** : $p\text{-value} < 0,05$, donc on rejette H_0

e) Faux : puisqu'en rejetant H_0 , on accepte H_1 : les deux échantillons ne sont pas issus de la même population. Mais on ne peut pas conclure cliniquement, juste statistiquement !

QCM n°6 B, C

- a) Faux, on doit poser cette condition (on doit aussi poser comme condition , si on compare deux moyennes observées sur deux échantillons, que les deux ont la même variance.)
b) **Vrai**
c) **Vrai**
d) Faux : Pour le test de Fisher, il n'y a pas de condition de taille des effectifs.
e) Faux : ce sont le test de Wilcoxon et le test des signes.
f) Faux

QCM n°7 : F

- a) Faux : c'est H_0
- b) Faux : on réalise un test de Mann-Whitney car les données ne sont pas appariées
- c) Faux : On construit le tableau suivant :

Valeurs	12	20	28	32	34	39	48	65	65	84
Rangs	1	2	3	4	5	6	7	8,5	8,5	10
R_F		2	3		5	6	7		8,5	
R_G	1			4				8,5		10

$$S_f = 2+3+5+6+7+8,5 = 31,5$$

$$S_g = 1+4+8,5+10 = 23,5$$

$$U_f = 4 \times 6 + \frac{6 \times 7}{2} - 31,5 = 13,5$$

$$U_g = 4 \times 6 + \frac{4 \times 5}{2} - 23,5 = 10,5$$

$$U = \min (U_f ; U_g) = 10,5$$

$n_1 > 10$ et $n_2 > 10$, donc on lit Useuil dans la table de Mann et Whitney.

Useuil = 2 ($n_1 = 4$; $n_2 = 6$ et $n_1 - n_2 = 2$)

Useuil < U , donc on ne rejette pas H_0 .

- d) Faux : on ne rejette pas H_0 (Attention, règle de décision différente par rapport aux tests précédents)
- e) Faux : On n'a pas pu rejeter H_0 , on ne peut conclure à la différence des deux distributions.
- f) **Vrai**

QCM n°8 : A, B, D

- a) **Vrai**
- b) **Vrai**
- c) Faux : la condition d'application du test de Mac Nemar est que le nombre de paires discordantes doit être supérieur ou égal à 10.
- d) **Vrai** : il est égal à 9.14

Il faut d'abord retranscrire l'énoncé en tableau :

	Goût (avant)	Équilibre (avant)	total
Goût (après)	22	6	28
Equilibre (après)	22	18	40
total	44	24	68

La somme des paires discordantes vaut $6 + 22 = 28 > 10$

On peut donc faire ce test.

$$\chi^2_{obs} = \frac{(22 - 6)^2}{22 + 6} = 9,14$$

Par lecture inverse dans la table de X^2 , on obtient $0,001 < pvalue < 0,01$

A 1 ddl, avec un risque de 5 % , $t_\alpha = 3,841$

$X^2_{obs} > t_\alpha$, donc on rejette H_0

- e) Faux : On rejette H_0 car $pvalue < \alpha$
- f) Faux

QCM n°9 : C, D

- a) Faux : car on prend toujours un risque d'erreur (c'est de la statistique) : α
- b) Faux : on peut comparer Pvalue à α .
- c) **VRAI**
- d) **VRAI**
- e) Faux : on peut comparer un échantillon à lui-même, après un traitement par exemple.

QCM n°10 : A, D

- a) **Vrai**
- b) Faux : car on fait un test bilatéral pour cette raison mais on aura besoin d'un Tobs plus grand pour rejeter H_0 ce sera plus dur à mettre en évidence.
- c) Faux : les échantillons sont indépendants.
- d) **Vrai** $Tobs = (m1 - m2) / ((s1^2/n) + (s2^2/n))^{1/2} = 1.015$ ensuite on lit sa Pvalue qui vaut 0.3
- e) Faux : car $Tobs = 1.015 < 1.96$, d'où on ne peut pas rejeter H_0 à 5%.

QCM n°11 : A, B, C, D, E

- a) **Vrai**
- b) **Vrai** la comparaison peut se faire aussi par le test de l'écart réduit.
- c) **Vrai**
- d) **Vrai** $tobs = (p1 - p2) / (p(1-p) * (2/n))^{1/2}$ avec $p = (p1 * n1 + p2 * n2) / (2n) = 5,769$
- e) **Vrai**

QCM n°12 : A, C, D

- a) **VRAI**
- b) Faux : on peut aussi faire un test exact de Fischer.
- c) **Vrai**
- d) **Vrai**
- e) Faux à $(k-1) * (r-1)$ ddl.

QCM n°13 : A, C, E

- a) **Vrai**
- b) Faux c'est à 1.74% en unilatéral !
- c) **Vrai**
- d) Faux, on manque d'informations (le signe de la différence $m - \mu$)
- e) **Vrai**

QCM n°14 : B, C, E

- a) Faux on retire de l'effectif de base le nombre de paires égales donc $m = 25 - 6 = 19$
- b) **Vrai**
- c) **Vrai**
- d) Faux : réfuter H_0 revient à dire que le traitement est efficace !! De plus on prendra en Bilatéral un risque de $2 * 0.01 = 0.02$ (par lecture dans la table).
- e) **Vrai**

QCM n°15 : A, D

- a) **Vrai**
- b) Faux à 6 ddl mais les conditions d'effectif sont respectées.
- c) Faux

O	12	89	111	77	80	31	0
E	19,2	43,2	87,6	100	87,6	43,2	19,2
(O-E) ²	51,84	2097,64	547,56	529	57,76	148,84	368,64
(O-E) ² /E	2,7	48,5564815	6,25068493	5,29	0,65936073	3,44537037	19,2
$\Sigma(O-E)^2/E$	86,1018975						

- d) **Vrai**
- e) Faux ça peut être une explication (on n'est pas sur).