TUTORAT UE 4 2015-2016 – Biostatistiques
CORRECTION Séance n°4 – Semaine du 12/10/2015

Tests statistiques
Professeur Molinari

<u>QCM n°1 :</u> **B, D, E**

A. Faux. On n'accepte jamais H0, on la rejette ou non.
B. **Vrai**.
C. Faux. Il faut inverser les parenthèses.
D. **Vrai**.
E. **Vrai**.

<u>QCM n°2 :</u> **B, D, E**

A. Faux. C'est le test unilatéral qui a plus de puissance, puisqu'il prend en compte le sens de la différence, mais on l'utilise que très rarement.
B. **Vrai**.
C. Faux. Seulement β. Puissance = 1 - β .
D. **Vrai**.
E. **Vrai**.

<u>QCM n°3 :</u> **A, E**

A. **Vrai**. Car n>30.
B. Faux. L'hypothèse nulle est que les deux moyennes sont égales.
C. Faux. t_{obs}=|m-µ|/(S/√n)=0,5/(2/√100)=2,5. On trouverait 1,25 si l'on prenait S²=4 au lieu de S.
D. Faux. A 5%, t_α=1,96, ainsi on a bien $t_{obs}>t_\alpha$, on rejette H0 au risque de 5%. Mais à 1%, t_α=2,576, ainsi $t_{obs}<t_\alpha$, on ne peut pas rejeter H0 au risque de 1%.
E. **Vrai**. on cherche dans la table de l'écart réduit la valeur de α pour que $t_\alpha=t_{obs}$.

<u>QCM n°4 :</u> **A, C**

A. **Vrai**. t_{obs}=(m-µ)/(s/√n)=0,03/(0,1/√35)=1,775
B. Faux. Le risque en unilatéral reste de 5%, c'est simplement la lecture de la table qui se fera à 10%.
C. **Vrai**. A 5%, t_α=1,645, ainsi $t_\alpha<t_{obs}$, on peut rejeter H0 d'une manière unilatérale.
D. Faux. A 2%, t_α=2,054, ainsi $t_\alpha>t_{obs}$, on ne peut pas rejeter H0 d'une manière unilatérale.
E. Faux. Les tests statistiques mettent en évidence un lien statistique, et non une relation causale, qui est démontrée par un faisceau d'arguments.

<u>QCM n°5 :</u> **B**

A. Faux. Ici, n<30, et la variable suit une loi Normale. On utilise donc un test de Student pour comparer 2 moyennes sur échantillons indépendants.
B. **Vrai**. Appliquer la formule du Formulaire: $\frac{17,78-14,69}{\sqrt{(\frac{8,333}{9}+\frac{8,333}{17})}}$ =2,6 avec S²=8,333.
C. Faux. On lit t_α dans la table de Student à n_1+n_2-2 ddl.
D. Faux. Pour α=5%, t_α=2,064, donc inférieur à t_{obs}. On rejette donc H0 à 5%.
E. Faux. Pour α=1%, t_α=2,797, on <u>ne rejette pas</u> H0. On n'accepte jamais H0.

QCM n°6 : **B, D, E**

A. Faux. Pour un test de Student, l'hypothèse de normalité est requise.
B. **Vrai**.
C. Faux. $t_{obs} = \frac{\bar{x}-\mu}{S/\sqrt{n}}$ =9,613, attention de ne pas utiliser la variance à la place de l'écart-type.
D. **Vrai**. t_α=2,508, inférieur à t_{obs}, donc on rejette H0 à 2%.
E. **Vrai**. On calcule le nouveau t_{obs}, avec n =4 on trouve 4,01. Puis on lit t_α = 5,841 à 3 ddl.

QCM n°7 : **B, C, D**

A. Faux. Le test d'Aspin-Welch est utilisé uniquement après que le test de comparaison de deux variances (test F) ait conduit à l'inégalité des deux variances. Or ici les variances sont égales.
B. **Vrai**. $t_{obs}=\frac{|m_1-m_2|}{\sqrt{\frac{S_1^2}{n_1}+\frac{S_2^2}{n_2}}}=\frac{|m_2-8|}{\sqrt{\frac{100}{120}+\frac{100}{70}}}$.=3,32. D'où m_2=13 ou m_2=3.
C. **Vrai**. t_α= 2,576, inférieur à t_{obs} donc on rejette H0.
D. **Vrai**.
E. Faux. On fait un test de l'écart réduit, qui ne nécessite pas que les variances soient égales.

QCM n°8 : **F**

A. Faux. Ici, on ne sait pas si les variances sont égales ou non, il faut donc réaliser un test F afin de vérifier. On fait donc le rapport des deux variances : $\frac{20,25}{7}$ = 2,89. On lit ensuite $t_\alpha \in [2,33; 2,40]$ dans la table de Fisher à 17 et 15 ddl. On rejette donc H0, les variances sont significativement différentes.
B. Faux. cf: A
C. Faux. Comme les variances sont différentes, on va utiliser un test d'Aspin-Welsh, on lira donc t_α dans la table de Student à m ddl. On calcul tout d'abord c= 0,72 (cf formulaire) puis m=28ddl (cf formulaire)
D. Faux. Lecture de la table de Student à 28 ddl, t_α=2,048 pour α=5%.
E. Faux. On calcul notre $t_{obs} = \frac{45-35}{\sqrt{\frac{20,25}{18}+\frac{7}{16}}}$ = 8, supérieur à t_α, on rejette donc H0 au risque de 5%.

QCM n°9 : **A**

A. **Vrai**.
B. Faux. H_1 est l'hypothèse selon laquelle au moins deux moyennes sont différentes l'une de l'autre.
C. Faux. Il teste la liaison entre une variable quantitative (la quantité d'endorphine libérée) et une variable qualitative (le type d'activité).
D. Faux. Il faut aussi supposer l'égalité des variances dans chacun des groupes.
E. Faux. Les tests de l'écart réduit et de Student ne permettent de comparer des moyennes que deux par deux or le chercheur veut les comparer globalement.

QCM n°10 : **A**

A. **Vrai**.
B. Faux. Il faut ici utiliser un test de Student pour données appariées car n≤30 (en supposant la normalité).
C. Faux. On le lit bien dans cette table, mais il faut le lire à (n-1)=5 ddl.
D. Faux. On ne peut attribuer de causalité à ce régime, la perte de poids peut être due à pleins d'autres facteurs comme par exemple une maladie ou autre.
E. Faux. On cherche la p-value, il faut donc calculer le t_{obs} : voici les différentes étapes :

1-Puisque l'on a un échantillon apparié, il faut faire la différence entre avant et après (ou après et avant) :

Différence : Poids avant – Poids après (=Poids perdu)	+4	0	+6	+8	+11	+1

2-Maintenant, il faut calculer la moyenne et l'écart-type estimé:

$$\bar{x} = \frac{\sum x_i}{n} = \frac{4+0+6+8+11+1}{6} = 5$$

$s^2 = \frac{1}{n}(\sum {x_i}^2 - n\bar{x}^2) = \frac{1}{6} * (4^2 + 0^2 + 6^2 + 8^2 + 11^2 + 1^2 - 6 \times 5^2) = \frac{44}{3}$ Donc $S^2 = \frac{44}{3} \times \frac{6}{5} = 17,6$ donc $S = 4,195$.

3-On peut à présent utiliser la formule en comparant à la valeur de référence qui est alors 0:

$$t_{obs} = \frac{\bar{x} - 0}{S/\sqrt{n}} = \frac{5}{4,195/\sqrt{6}} = 2,92$$

Donc, en lisant dans la table de Student à 5ddl on trouve un risque entre 0,02 et 0,05. Le test est en unilatéral on divise donc par deux le risque : 0,01<p-value<0,025.

QCM n°11 : **A, C, D**

	Nombre de cas avec complications	Nombre de cas sans complications	Pourcentage de complications	Totaux
Nouvelle Approche	Observé=18 Théorique=14,4	Observé=27 Théorique=30,6	40%	45
Ancienne Approche	Observé=6 Théorique=9,6	Observé=24 Théorique=20,4	20%	30
Totaux	24	51		75

A. **Vrai**. Les conditions sont réunies

$$On\ a\ p = \frac{p_1 n_1 + p_2 n_2}{n_1 + n_2} = 0,32$$

$$Donc\ n_1 p = 0,32 \times 45 = 14,4 \geq 5\ et\ n_2 p = 0,32 \times 30 = 9,6 \geq 5\ et$$
$$n_2 q = 0,68 \times 45 = 30,6 \geq 5\ et\ n_2 q = 0,68 \times 30 = 20,4 \geq 5$$

B. Faux. p=32%

C. **Vrai**. Cf tableau ci-dessus

D. **Vrai**.

$$\chi^2{}_{obs} = \sum \frac{(O_{ij} - E_{ij})^2}{E_{ij}} = \frac{(18 - 14,4)^2}{14,4} + \frac{(27 - 30,6)^2}{30,6} + \frac{(6 - 9,6)^2}{9,6} + \frac{(24 - 20,4)^2}{20,4} = 3,31$$

E. Faux. A 1ddl, la valeur seuil est de 3,841 à 5% donc on ne rejette pas H_0.

QCM n°12 : **A, B, C, D, E**

A. **Vrai**.

B. **Vrai**. Pour calculer les effectifs espérés, il suffit de compter le nombre total de thons et de les répartir équitablement entre les horaires puisque sous H_0, il s'agit d'une loi uniforme. Ainsi on trouve qu'il y a 80 thons donc on devrait espérer en vendre 20 par tranche d'horaires.

C. **Vrai**. Tous les effectifs espérés sont supérieur à 5 (ils sont tous égaux à 20).

D. **Vrai**.

$$\chi^2{}_{obs} = \sum \frac{(O_{ij} - E_{ij})^2}{E_{ij}} = \frac{(15 - 20)^2}{20} + \frac{(30 - 20)^2}{20} + \frac{(25 - 20)^2}{20} + \frac{(10 - 20)^2}{20} = 12,5$$

E. **Vrai**. Il suffit de lire dans la table du $\chi^2{}_{obs}$ à 3ddl, et on voit que la p-value est compris entre 0,01 et 0,001.