TUTORAT UE 4 2013-2014 – Biostatistiques
Séance n°4 – Semaine du 14/10/2013

Lois de Probabilités - Tests statistiques
Sabatier - Molinari

Séance préparée par les tuteurs du TSN

QCM n°1 : **Dans un service de MME, les infirmiers réalisent tous les matins le contrôle de la glycémie X de 10 patients diabétiques de type II. Les valeurs relevées sont : 1,50 ; 1,96 ; 1,10 ; 2,05 ; 2,30 ; 1,75 ; 2,10 ; 1,05 ; 0,90 ; 2,80.**
On veut estimer la glycémie du matin de tous les diabétiques de type II présents dans l'hôpital. On pose l'hypothèse selon laquelle X suit une loi normale. Choisir la ou les propositions exactes.

A. La moyenne de l'échantillon est de 1,751 et l'écart type observé est de 0 ,5796.
B. On peut estimer la variance de la population à 0,373.
C. L'IC au seuil de 95% de la glycémie moyenne de la population est de [1,314; 2,188].
D. L'IC au seuil de 90% de la glycémie moyenne de la population est de [1,41; 2,09].
E. Les hypothèses auraient été strictement les mêmes si le nombre de patients avait été de 45.
F. Toutes les propositions précédentes sont fausses.

QCM n°2 : **On s'intéresse à la distance séparant 40 étudiants en PACES de la fac, on a mesuré une valeur moyenne de 9,4 km et une variance de 5,76 km². Choisir la ou les propositions exactes.**

A. On peut estimer la variance de la population S^2 =5,91.
B. Des hypothèses supplémentaires sont nécessaires afin de calculer l'IC dans lequel se trouve la distance moyenne pour tous les PACES.
C. Pour un risque alpha=5%, on a $c_{\alpha/2}$=1,960.
D. L'IC au seuil de 95% est de [8,65; 10,15].
E. L'IC au seuil de 90% est de [8,77; 10,03].
F. Toutes les propositions précédentes sont fausses.

QCM n°3 : **On choisit parmi la population un échantillon, tiré de manière aléatoire, de 19 personnes. On cherche à connaître leur pratique sportive hebdomadaire (variable X). Sur cet échantillon on trouve qu'en moyenne, ils pratiquent 2h de sport avec un écart type de 0,5h. Choisir la ou les propositions exactes.**

A. On doit poser l'hypothèse selon laquelle la variable X suit une loi normale afin de pouvoir calculer des IC.
B. En posant toutes les hypothèses nécessaires, l'IC au seuil de 95% de l'espérance moyenne est de [1,752 ; 2,248].
C. L'IC de la variance se calcule avec la loi de Student à n-1 ddl.
D. L'IC au seuil de 80% pour la variance est de [0,4276 ; 0,6612].
E. L'IC au seuil de 80% pour la variance est de [0,18 ; 0,44].
F. Toutes les propositions précédentes sont fausses.

QCM n°4: **On réalise un sondage auprès de 80 étudiants en PACES choisis au hasard dans l'UM1. A la question :"appréhendez-vous le concours?", 30 courageux répondent négativement. On s'intéresse à ceux qui se disent craintifs. Choisir la ou les propositions exactes.**

A. Le nombre de réponses positives dans l'échantillon suit une loi binomiale.
B. Pour calculer un intervalle de confiance, on peut utiliser comme approximation la loi de Poisson.
C. Un intervalle de confiance de la proportion p_0 de craintifs au risque α=5 % est [0,508 ; 0,742].
D. Si on raisonne en nombre de PACES, l'intervalle de confiance au risque α=5 % est [41,5 ; 58,5].
E. Si on prend un risque α plus faible, à 3%, l'intervalle de confiance sera plus grand.
F. Toutes les propositions précédentes sont fausses.

QCM n°5: **En moyenne, un plan de tomate cultivé sans engrais donne 42,4 tomates avec un écart type de 4,7. Sur un échantillon de 80 pieds de tomates avec engrais, on a observé une moyenne de production de 43,2 tomates avec un écart type de 3,3. On se demande si l'ajout d'engrais a une influence sur le nombre de tomates par plan. Choisir la ou les propositions exactes.**

A. L'hypothèse H1 est μ=m0 pour un test bilatéral.
B. On utilise un test de l'écart réduit.
C. t_{obs}=1,522.
D. Au risque α=5%, on rejette H0 mais on accepte H0 pour un risque α=2 %.
E. La p-value est 0,03.
F. Toutes les propositions précédentes sont fausses.

QCM n°6: **Pour pouvoir bronzer au soleil sans complexe cet été, un groupe de 10 personnes a décidé de faire un régime en privilégiant les fruits et les légumes. On note leur poids avant et après leur régime, en kg. On suppose que la variable poids suit une loi normale. Choisir la ou les propositions exactes.**

Avant	61	82	74	130	90	53	79	103	72	66
Après	56	80	78	112	93	53	80	97	75	67

A. On utilise le test de student avec 18 degrés de liberté.
B. La moyenne des différences est m=1,9.
C. L'écart type estimé des différences est de S=47,2.
D. Pour un risque α=5%, la perte de poids est significative.
E. Pour avoir un plus grand nombre de données on refait la même étude sur 40 personnes. On utilisera le test de l'écart réduit.
F. Toutes les propositions précédentes sont fausses.

QCM n°7: **Concernant les estimations, choisir la ou les propositions exactes.**

A. L'inférence statistique consiste à déduire les paramètres d'une population à partir d'un échantillon tiré de cette population.
B. Pour cela, on utilise uniquement des estimateurs par intervalle.
C. Un intervalle de confiance permet d'estimer un paramètre avec un risque beta connu de se tromper, beta étant la probabilité de ne pas rejeter l'hypothèse émise alors qu'elle est fausse.
D. On peut estimer le paramètre p de la loi binomiale en utilisant une approximation par la loi normale dès lors que la taille de l'échantillon est supérieure à 30.
E. On peut estimer le paramètre de la loi de Poisson dès lors que λ est supérieur à 20.
F. Toutes les propositions précédentes sont fausses.

QCM n°8: **A propos de la comparaison de 2 moyennes observées sur des échantillons indépendants, choisir la ou les propositions exactes.**

A. Dans le test de Student, son utilisation nécessite 2 populations de moyennes m_1 et m_2 et de variances σ_1 et σ_2.
B. Dans le test de l'écart réduit; t_α sera lu dans la table de la loi Normale centrée réduite N(0,1).
C. Pour pouvoir utiliser le test de Student, on doit juste vérifier que les 2 populations ont la même variance.
D. Dans le test de l'écart réduit, la statistique du test se lira à n1+n2 -2 ddl.
E. Les tests de Student et de l'écart-réduit ne sont utilisés que pour des échantillons indépendants.
F. Toutes les propositions précédentes sont fausses.

QCM n°9: **Concernant la comparaison de 2 variances, choisir la ou les propositions exactes.**

A. Ce test se nomme aussi test de Fisher ou test F ou test exact de Fischer.
B. Par ce test, on peut vérifier les conditions d'application de certains tests non paramétriques pour lesquels on n'a aucune information sur la distribution des variables.
C. t_{obs}=S1/S2 avec S1>S2.
D. Le test F permet parfois d'utiliser le test de Student.
E. Le test d'ANOVA permet également de vérifier les conditions du test de Student pour comparer des échantillons indépendants.
F. Toutes les propositions précédentes sont fausses.

QCM n°10 : **On se rend en amphi PACES et on divise la population en deux groupes : les filles et les garçons. On compte dans chaque groupe le nombre d'individus portant des lunettes. On obtient le tableau suivant :**

	Avec lunettes	Sans lunettes	Total	Pourcentage
Filles	18	109	127	p1 =14,17%
Garçons	16	57	73	p2 =21,91%
	34	166	200	

On cherche à savoir si ces deux pourcentages sont significativement différents. On définit H0 : $\pi 1=\pi 2$.

A. Les conditions d'applications du χ^2 d'homogénéité sont remplies car les effectifs observés sont tous supérieurs à 5.
B. On compare 2 pourcentages observés sur échantillons indépendants, on peut donc utiliser le X^2 de McNemar.
C. En faisant un test du χ^2, la statistique du test vaut 2,97.
D. On lit dans la table du χ^2 à 1ddl et pour un α de 5% : χ^2=3,841.
E. Au risque de 5%, on ne rejette pas H0.
F. Toutes les propositions précédentes sont fausses.

QCM n°11: **Dans la population générale, la prévalence de la dépression est de 9%. On cherche à connaître la prévalence au sein des étudiants en PACES. On constitue un échantillon représentatif de 75 étudiants en PACES à qui l'on fait passer une visite médicale. 13 sont diagnostiqués dépressifs. On cherche à savoir si les étudiants en PACES sont plus susceptible de déprimer. Choisir la ou les propositions exactes.**

On définit H0 : P(pourcentage de la population des PACES) = 0,09.

A. Le pourcentage estimé d'étudiants en PACES dépressifs est de 17,33%.
B. On peut réaliser un test de l'écart-réduit pour savoir si ce pourcentage est significativement éloigné du pourcentage théorique.
C. La statistique de test vaut t_{obs}=1,91.
D. Avec un α de 5%, on rejette H0. On en déduit que les étudiants en PACES sont plus susceptibles de déprimer que la population générale.
E. Avec un échantillon de 80 étudiants en PACES et un même pourcentage observé, on n'aurait pas rejeté H0.
F. Toutes les propositions précédentes sont fausses.

QCM n°12: **On recense les habitants d'un petit village au sud de Nîmes selon leur groupe sanguin : A, B, AB ou O. On cherche à savoir si les fréquences par groupe sanguin sont significativement différentes de la distribution théorique, calculée sur toute la population française. Choisir la ou les propositions exactes.**

Groupe sanguin	Fréquence théorique	Distribution observée (nombre d'habitants)
A	44,00%	1018
O	45,00%	1113
B	8,00%	142
AB	3,00%	96
	100,00%	2369

A. Pour réaliser un χ² d'ajustement, tous les effectifs théoriques doivent être supérieurs à 10.
B. Le χ² observé suivra une loi à 4 ddl.
C. La statistique de test vaut 23,55.
D. Avec un risque de 5%, les distributions sont significativement différentes.
E. Avec un risque de 0,1%, les distributions sont significativement différentes.
F. Toutes les propositions précédentes sont fausses.

<u>QCM n°13 :</u> **Nous nous intéressons à la pratique sportive de 150 DFGSM2 par rapport à la PACES :**

- ✓ **30 DFGSM2 déclarent avoir fait du sport les 2 années**
- ✓ **50 DFGSM2 déclarent avoir fait du sport en DFGSM 2 et non en PACES**
- ✓ **30 DFGSM2 déclarent avoir fait du sport en PACES mais pas en DFGSM2**

On cherche à savoir si les étudiants changent leur activité sportive en DFGSM2. Choisir la ou les propositions exactes.

A. On utilise un test du χ^2 de MacNemar.
B. $X^2_{obs}=\frac{(f-g)^2}{f+g}=5$.
C. $X^2_\alpha = 3{,}841$ pour $\alpha=5\%$.
D. On ne rejette pas H_0 au risque 5%.
E. On ne rejette pas H_0 au risque 2%.
F. Toutes les propositions précédentes sont fausses.

<u>QCM n°14 :</u> **On a 3 échantillons différents de malades selon la nature de leur maladie. Parmi les groupes, on les range selon leur rhésus (rhésus+ et rhésus-). On souhaite étudier s'il y a un lien entre le système rhésus et les natures des maladies. Choisir la ou les propositions exactes.**

A. On utilise un test d'homogénéité entre les 2 distributions
B. $t_{obs} = \frac{\sum(Oij-Eij)^2}{Oij}$

En calculant le χ^2_{obs}, on obtient $\chi^2_{obs} = 8{,}25$.

C. On utilise un χ^2 à 2 ddl.
D. On conclut qu'il existe une liaison statistique entre la nature de la maladie et le système rhésus à 5%.
E. On conclut qu'il existe une liaison statistique entre la nature de la maladie et le système rhésus à 1%.
F. Toutes les propositions précédentes sont fausses.