TUTORAT UE 4 2012-2013 – Biostatistiques
CORRECTION Colle n°2 – Semaine du 05/11/2012

Probabilités – Lois de probabilités – Tests statistiques

Dujols – Sabatier - Molinari

QCM n°1 : **A, E**

A. **Vrai**.
B. Faux. P(A/B)=P(A∩B) que si les évènements sont compatibles. C'est la probabilité conditionnelle.
C. Faux. C'est la proportion de faux négatifs.
D. Faux. C'est l'arrangement qui tient compte de l'ordre.
E. **Vrai**.

QCM n°2 : **ANNULE**

QCM n°3 : **C, E**

A. Faux. P(S) x P(C)=0,35. Alors que P(S∩C)=0,3 (calcul en C). Les évènements ne sont pas indépendants, ils sont compatibles.
B. Faux. P(S/$\bar{C}$)=P(S∩$\bar{C}$)/P($\bar{C}$)=0,2/0,3=0,67.
C. **Vrai**. P(S∩C)=P(S) - P(S∩$\bar{C}$)= 0,5 - 0,2 = 0,3
D. Faux. P(S/C)= P(S∩C)/P(C)=0,3/0,7=0,43.
E. **Vrai**. P(C/S)=P(S/C)xP(C) / P(S/C)xP(C) + P(S/$\bar{C}$)P($\bar{C}$) = 0,6

QCM n°4 : **A, C, E**

A. **Vrai**. On peut considérer l'échantillon comme représentatif car il a été tiré au sort.
B. Faux. La classe modale est [8 ;12[car effectif le plus élevé.
C. **Vrai**. Moyenne=$\sum n_i x_i/n$ = 7,78
D. Faux.
E. **Vrai**. La classe médiane est [8 ;12[. Car on a la moitié de l'effectif au dessous, et l'autre moitié au dessus. La moitié de l'effectif est de 50,5.

QCM n°5 : **A, C, E**

A. **Vrai**
B. Faux. P(X=2)=C^k_nxp^kx$(1-p)^{n-k}$=(26!/(24!x2!))x0,15²x$0,85^{24}$=0,14795 à 10^{-5} près.
C. **Vrai.** Pour réaliser une approximation d'une loi Binomiale par une loi de Poisson, il faut remplir les 2 conditions suivantes : n>20 et p<0,5, ce qui est le cas ici. On approxime alors la loi Binomiale par une loi de Poisson de paramètre λ=np=26x0,15=3,9.
D. Faux. Afin de réaliser une approximation de la loi Binomiale par une loi Normale, il faut remplir les 3 conditions suivantes : n≥30, np>5 et n(1-p)>5. Ici, n<30 donc l'approximation par une loi Normale n'est pas possible.
E. **Vrai**. Par approximation par une loi de Poisson : P(X=2)=$e^{-\lambda} \times \frac{\lambda^k}{k!} = e^{-3,9} \times \frac{3,9^2}{2!}$ =0,15394 à 10^{-5} près.

QCM n°6 : **A**

A. Vrai. P(X=k)=$e^{-\lambda} \times \frac{\lambda^k}{k!}$
P(X=0)=0,014=$e^{-\lambda}$
λ=-ln(0,014)=4,27
B. Faux. L'approximation est impossible car λ <20.
C. Faux. $P(X > 3) = 1 - P(X \leq 3) = 1 - P(X = 0) - P(X = 1) - P(X = 2) - P(X = 3)$

$$= 1 - e^{-4,27} - e^{-4,27} \times \frac{4,27^1}{1!} - e^{-4,27} \times \frac{4,27^2}{2!} - e^{-4,27} \times \frac{4,27^3}{3!} = 0,617$$

D. Faux. $P(X < 2) = P(X = 0) + P(X = 1) = e^{-4,27} + e^{-4,27} \times \frac{4,27^1}{1!} = 0,074$

E. Faux. V(X)=λ donc σ(X)=$\sqrt{4,27} = 2,07$.

QCM n°7 : **C, E**

A. Faux. P(X<90)=P(U<(90-96)/10) d'où π(-0,6)=0,2743, que l'on multiplie par l'effectif total, afin d'obtenir l'effectif recherché : 0,2743x1200=329,16.

B. Faux. P(100<X<110)=P($\frac{100-960}{10}$<U<$\frac{110-96}{10}$) d'où π(1,4)-π(0,4)=0,9192-0,6557=0,2635, soit un effectif égal à 316,2.

C. **Vrai**. Cf item B.

D. Faux. Nous sommes ici dans le cadre d'une loi continue, donc P(X=98)=0.

E. **Vrai.** P(95<X<105)=P($\frac{95-96}{10}$<U<$\frac{105-96}{10}$) d'où π(0,9)-π(-0,1)=0,8159-0,4602=0,3557 soit 35,57%.

QCM n°8 : **A, B, E**

A. **Vrai.**

B. **Vrai.** D'après les données de l'énoncé : P(X>86)=0,3050 → P(X<86)=1-0,3050=0,6950 → π(0,51).
P(X<132)=0,9850 → π(2,17).
Il nous suffit alors de résoudre un système à deux équations, tel que :
(86-μ)=0,51σ et 132-μ=2,17σ
D'où μ=86-0,51σ, que l'on intègre dans l'autre équation : 132-(86-0,51σ)=2,17σ → 46=1,66σ → σ=(46/1,66)=27,71 à 10^{-2} près.
Ne reste plus qu'à calculer l'espérance μ, telle que : μ=86-0,51x27,71=71,87 à 10^{-2} près.

C. Faux.

D. Faux.

E. **Vrai.**

QCM n°9 : **C**

A. Faux. $\bar{x} = \frac{\Sigma\text{xi}}{n} = \frac{17,55}{15} = 1,17$

Calcul de la variance observée dans l'échantillon : $s^2 = \frac{\Sigma xi^2}{n} - \bar{x}^2 = \frac{24,13}{15} - 1,17^2 = 0,2398$.

Calcul de la variance estimée dans la population : $S^2 = \frac{n}{n-1} \times s^2 = \frac{15}{14} \times 0,24 = 0,2569$.

B. Faux. (Cf item A)

C. **Vrai**. (Cours)

D. Faux. Si n≥30, l'hypothèse de normalité n'est pas obligatoire.

E. Faux. On lit les valeurs de a et b dans la table du chi².

QCM n°10 : **B, E**

A. Faux. $\sigma^2 \epsilon [(n-1)\frac{S^2}{b}; (n-1)\frac{S^2}{a})]$. On lit a et b dans la table du chi² à ddl=n-1=14.
Pour a on lit dans la colonne $1 - \frac{\alpha}{2} = 0,90$ donc a=7,790.
Pour b on lit dans la colonne $\frac{\alpha}{2} = 0,10$ donc b=21,064.
S² est la variance estimée dans la population.
On a donc : $\sigma^2 \epsilon [0,1707; 0,4617]$

B. **Vrai**. (Cf item A)

C. Faux. $\mu \, \epsilon \, [\bar{x} - t\frac{S}{\sqrt{n}}; \bar{x} + t\frac{S}{\sqrt{n}}]$
On utilise la table de student à ddl = n-1=14 car n<30 et σ est inconnu.
t=1,761 pour α=0,10.
Ecart type estimé dans la population : S=$\sqrt{0,2569} = 0,5069$
Donc $\mu \, \epsilon \, [0,9395; 1,4005]$

D. Faux.
idem item C avec t=2,145 pour α=0,05.
donc $\mu \, \epsilon \, [0,8893; 1,4507]$

E. **Vrai**. Plus le seuil (1- α) est élevé, plus l'intervalle est étendu et moins il est précis.

QCM n°11 : **A, B, C, E**

A. **Vrai**.
B. **Vrai**.
C. Vrai.
D. Faux. C'est justement dans un test unilatéral qu'on s'occupe de savoir lequel est supérieur à l'autre.
E. **Vrai**.

QCM n°12 : **B**

A. Faux. Les hypothèses H1 comme H0 ne portent que sur la population. Donc H1 : $\mu 1 > \mu 2$ et H0 :$\mu 1 = \mu 2$
B. **Vrai**.
C. **Vrai**. n1=20<30 et n2=18<30. On pense donc directement à utiliser le test de student. Sauf que pour appliquer ce test, il faut en plus respecter certaines conditions d'application.
   1- Respecter l'hypothèse de normalité : le poids des filles aussi bien dans la population PACES que DFGSM2 doit suivre la loi normale
   2- Les variances des 2 populations sont considérées comme égales
D. Faux. tα se lira à (n1+n2–2) ddl soit 36 ddl
E. Faux. Dans son cours le professeur Molinari vous spécifie qu'il y a certaines formules à connaitre et d'autres pas du tout. $T_{obs}=\frac{\bar{X}-\mu_o}{S/\sqrt{n}}$ est l'une des formules à connaitre. Il est utilisé pour la comparaison d'une moyenne observée à une moyenne théorique. Bien sûr la formule de Tobs pour la comparaison de deux moyennes observées n'est pas à connaitre et est donné dans le formulaire.
(La formule étant: $t_{obs}=\frac{|m1-m2|}{\sqrt{\frac{S^2}{n1}+\frac{S^2}{n2}}}$ avec $s^2=\frac{(n1-1)\times s1^2+(n2-1)\times s2^2}{n1+n2-2}$)

QCM n°13 : **F**

A. Faux.
Principe: Ho se pose tjrs sur la population. μ=moyenne théorique, μo=moyenne de référence=moyenne de la population dont est issue l'échantillon. m=moyenne de l'échantillon. Le test va permettre de montrer si μ est différent de μo et donc si l'échantillon appartient bien à la population. Par conséquent ho: $\mu = \mu$o. Si nous revenons à l'énoncé, μo=80, m= 150. Donc ho: $\mu = \mu$o=80
B. Faux. par lecture inverse de la table de l'écart réduit, on a p-value=0,03.
C. Faux. n>30. La comparaison se fait entre une valeur observée et une valeur de référence donc le test employé est le test de l'écart réduit.
D. Faux. La puissance du test est égale à 1–β or β=risque de 2eme espèce.
E. Faux.
<u>1ère Méthode simple</u> : p-value=0,03 > α=0,0038 → on ne rejette pas H0.

<u>2ème méthode</u> : pour comprendre le raisonnement, on cherchera d'abord à déterminer tα lorsque α=5%. En effet si α=5%, α/2=2,5%=0,025. P(X>tα)=0,025 → P(T≤tα)=1-0.025=0,9750. Par lecture inverse de la table de la loi normale centrée réduite, on a tα=1,96.

En utilisant un raisonnement analogue au précedent on a α=0,38% donc α/2=0,19%=0,0019 P(X> tα)=0,0019 donc P(X≤tα)=1-0,0019=0,9981. Par lecture inverse de la table de la loi normale centrée réduite, on tα=2,9.

Tobs<Tα donc ne rejette pas l'hypothèse H0 → il n'y a pas une différence significative de la pression artérielle diastolique.
F. Vrai.

QCM n°14 : **A, C, D**

A. **Vrai**. La problématique statistique est la comparaison de 2 moyennes observées sur 2 séries appariés et n=50>30. Donc le test à utiliser est le test de l'écart réduit
B. Faux. Le test employé est un test pour données appariées

C. **Vrai**.
D. **Vrai**.α=2% alors par lecture de la table de l'écart réduit Tα=2,326 < Tobs=6
On rejette Ho puisque Tobs>Tα. Par conséquent il y a une modification significative du cholestérol avant et après traitement. Le médicament M est donc efficace

Calcul de Tobs

$$Tobs = \frac{\delta d - 0}{S/\sqrt{n}} = \frac{/différence\ des\ moyennes\ /}{écart{-}type\ de\ la\ difference\ des\ moyennes} = \frac{0{,}18}{0{,}03} = 6$$

E. Faux. α=5%, Tα=1,96 <Tobs=6 donc le médicament M est efficace

QCM n°15 : **B, D, E**

A. Faux. Les échantillons sont indépendants, on utilise un test du Khi 2.
B. **Vrai**.
C. Faux. Tableau des effectifs théoriques :

	PACES	P2
Filles	1091 (1200X2100/2310)	109
Garçons	1009	101

X²obs=tobs= $\frac{(1100-1091)^2}{1091} + \frac{(100-109)^2}{109} + \frac{(1000-1009)^2}{1009} + \frac{(110-101)^2}{101}$ =1,6996.

On lit le tα pour α=0,05 dans la table du Khi 2 à 1 ddl ; tα=3,841.
tα > tobs, on ne rejette pas H0 pour alpha à 5%.
D. **Vrai**. Tα=1,642 pour α=0,20 donc on rejette H0 car tobs > tα ; il y a une différence significative pour un risque de 20%.
E. **Vrai**. Par lecture inverse de la table, notre tobs est compris entre 1,642 et 2,706.

QCM n°16 : **A**

A. Vrai.
B. Faux.
C. Faux.

Sujets	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
Avant	**22**	**25**	**20**	**34**	**20**	**47**	**30**	**41**	**31**
Après	**20**	**25**	**21**	**32**	**16**	**40**	**33**	**45**	**26**
Différence	**-2**	**0**	**+1**	**-2**	**-4**	**-7**	**+3**	**+4**	**-5**
Rang	**2,5**		**1**	**2,5**	**5,5**	**8**	**4**	**5,5**	**7**

S+ = somme des rangs + =25,5
S- = somme des rangs - = 10,5
S=10,5 (minimum entre S+ et S-)
D. Faux. On lit dans la table de Wilcoxon en unilatéral au risque alpha de 2,5% et pour N=8 (différences non nulles). Tα=4.
E. Faux. S>T donc on ne rejette pas H0 et il n'y a pas de réduction significative pour un risque de 2,5%.