TUTORAT UE 4 2015-2016 – Biostatistiques
CORRECTION Séance n°7 – Semaine du 02/10/2015

Séance de révisions générales

QCM n°1 : **A, B, C, D**

A. **Vrai**. Ici, on tire les bonbons en même temps, c'est donc une Combinaison : $\frac{C_{30}^2}{C_{150}^2}$= 0,0389 Attention : on ne peut pas faire 150! sur une calculette, (on utilise donc la simplification entre 150! et 148! (respectivement au numérateur et dénominateur) :$\frac{149 \times 150}{2!} = C_{150}^2$

B. **Vrai**. Ici, on tire les bonbons successivement : $\frac{A_{30}^1 \times A_{20}^1 \times A_{40}^1 \times A_{40}^1 \times A_{20}^1}{A_{150}^5}$= 2,7045.$10^{-4}$

C. **Vrai**. C'est une permutation, donc 5! = 120.

D. **Vrai**. $\frac{C_{50}^3}{C_{150}^3}$= 196/5513

E. Faux. Ce n'est pas parce que l'on garde les mêmes proportions en bonbons que les probabilités restent identiques. $\frac{C_{90}^2}{C_{450}^2}$=0,0396.

QCM n°2 : **A, C, D**

A. **Vrai**. P(S)=P(S∩T)+P(S∩ $\bar{T}$) = P(S/T) ×P(T) + P(S/ $\bar{T}$)×P($\bar{T}$) On isole P(S/T) et on trouve 0,01.
B. Faux. On vérifie si P(S∩T) = P(T) × P(S) ; P(T) × P(S)=0,16 et P(S∩T) =2,5×10^{-3}.
C. **Vrai**. On utilise le théorème de Bayes. P($\bar{T}$/$\bar{S}$) =$\frac{P(\bar{S}/\bar{T}) \times P(\bar{T})}{P(\bar{S}/\bar{T}) \times P(\bar{T}) + P(\bar{S}/T) \times P(T)}$=5/16.
D. **Vrai**. Le dénominateur dans le théorème de Bayes correspond à la formule de la probabilité totale.
E. Faux. P(T/$\bar{S}$) = 1 - P($\bar{T}$/$\bar{S}$) = 11/16.

QCM n°3 : **C, D**

A. Faux. Ce n'est pas 15% mais bien 75%.
B. Faux. Si on prend un patient au hasard, il y a 9 % de chances qu'il ne soit pas vacciné efficacement : Par exemple, si 100 personnes se font vaccinés, à la première session, 10 d'entre eux seront vaccinés efficacement, ensuite 60% des 90 restant seront vaccinés efficacement à la deuxième session, soit 54, puis 75% des 36 restant non vaccinés le seront à la 3eme session, soit 27. Il ne restera donc plus que 9 personnes vaccinées inefficacement.
C. **Vrai**. 0,09 × 200 = 18. (cf. item B)
D. **Vrai**. Cf raisonnement item B
E. Faux, Il a 60% de chance d'être vacciné efficacement à la 2ème session, il faut ajouter la chance d'être vacciné efficacement à la 3ème session pour ceux pour qui la 2eme session n'a pas été efficace. Soit 0,6+0,75×0,4=0,9

QCM n°4 : **A, B, D**

A. **Vrai**. On a n>20 et p<0.5, on approxime par une loi de Poisson de paramètre λ=np=105*0,07=7,35
B. **Vrai**. P(X=0)= $e^{-7,35}$= 6,4259.10^{-4}.
C. Faux. P(X≥2)= 1-P(X<2)= 1-P(X=0)-P(X=1)= 1 - $e^{-7,35}$ – 7,35× $e^{-7,35}$=0,9946.
D. **Vrai**. P(X≤1)= P(X=0) + P(X=1)= $e^{-7,35}$ + 7,35× $e^{-7,35}$=5,3656× 10^{-3}.
E. Faux. Attention, la loi de Poisson est discrète même si elle prend un nombre infini de valeurs, ainsi aucune correction de continuité lorsqu'on passe d'une loi discrète à une autre loi discrète.

QCM n°5 : **F**

A. Faux. On a X~B(200;3 10^{-2}). En effet, X compte le nombre total de succès après n expériences identiques et indépendantes.
B. Faux. En effet, nous pouvons approximer par une loi normale puisqu'on a n≥30, np=6≥5, et nq=194≥5. Cependant μ=np=6, et σ=√npq=2,41.
C. Faux. En utilisant la correction de continuité, on a P(X<5) en binomiale =P(X≤4,5) en normale =$\pi(\frac{4,5-6}{2,41})$= π(-0,62)=0,2676. On trouve 0,3409 sans la correction de continuité.
D. Faux. En utilisant la correction de continuité, on a P(4,5<X≤6,5) =$\pi(\frac{6,5-6}{2,41})$ −$\pi(\frac{4,5-6}{2,41})$= π(0,21) – π(–0,62)= 0,5832 – 0,2676= 0,3156.
E. Faux. Il ne faut pas oublier la correction de continuité ! P(1,5≤X≤2,5) =$\pi(\frac{2,5-6}{2,41})$ -$\pi(\frac{1,5-6}{2,41})$= π(-1,45) - π(-1,87) =0,0735 – 0,0307= 0,0428.

QCM n°6 : **A, B, C, D, E**

A. **Vrai**.
B. **Vrai**. On vérifie les conditions d'application : np=30>5 et nq=60>5. Pour 5%, c=1,96. Les bornes de cet intervalle sont $p_0 \pm$ c$\sqrt{\frac{p_0(1-p_0)}{n}}$ = $1/3 \pm$ 1,96$\sqrt{\frac{1/3(1-1/3)}{90}}$. On trouve bien l'intervalle [0,2359 ; 0,4307].
C. **Vrai**. Il suffit de multiplier notre intervalle de confiance précédent par n=90 pour trouver celui du nombre de moustachus.
D. **Vrai**. Avec un risque α plus petit, on a plus de chance d'être dans l'intervalle, ainsi celui-ci est plus grand.
E. **Vrai**. Pour 2%, c=2,326. Les bornes de cet intervalle sont $p_0 \pm$ c$\sqrt{\frac{p_0(1-p_0)}{n}}$ = $1/3 \pm$ 2,326$\sqrt{\frac{1/3(1-1/3)}{90}}$. On trouve bien l'intervalle [0,2178 ; 0,4489].

QCM n°7 : **B, E**

A. Faux. Pour une densité de probabilité, il faut qu'elle soit continue sur intervalle de définition et qu'elle soit positive. Ensuite, il faut que ça somme de –∞ à +∞ soit égale à 1 :

$\int_{-\infty}^{+\infty} f(x)dx = 1 \; soit \; en \; simplifiant: \int_0^2 f(x)dx = 1 = \int_0^2 (tx^3 + k)dx = \left[\frac{tx^4}{4} + kx\right]_0^2 = 4t + 2k$=1

$De \; plus \; on \; sait \; que: P(X < 1) = 0{,}325 \; Donc \; \int_0^1 f(x)dx = 0{,}325 = \left[\frac{tx^4}{4} + kx\right]_0^1 = 0{,}25t + k = 0{,}325$

On a donc un système à deux équations à deux inconnus qu'il faut résoudre :

$\begin{cases} 4t + 2k = 1 \\ 0{,}25t + k = 0{,}325 \end{cases} Donc \; k = 0{,}325 - 0{,}25t \; Ainsi \; on \; a \begin{cases} 4t + 2(0{,}325 - 0{,}25t) = 1 \\ k = 0{,}325 - 0{,}25t \end{cases} [\ldots] \; Donc \begin{cases} t = 0{,}1 \\ k = 0{,}3 \end{cases}$

Donc $f(x) = 0{,}1x^3 + 0.3$ sur 0<x<2

B. **Vrai**. *Cf* A
C. Faux. On a $F(u) = \int_{-\infty}^{u} f(x)dx \; soit \; en \; simplifiant: \int_0^u f(x)dx = F(u) = \left[\frac{0{,}1x^4}{4} + 0{,}3x\right]_0^u = \frac{0{,}1x^4}{4} + 0{,}3u$
D. Faux. On utilise la relation trouvée ci-dessus :

$$F(u) = \frac{0{,}1x^4}{4} + 0{,}3u \; donc \; P(x < 1{,}5) = F(1{,}5) = \frac{0{,}1 \times 1{,}5^4}{4} + 0{,}3 \times 1{,}5 = 0{,}58$$

E. **Vrai**. On a $E(x) = \int_{-\infty}^{+\infty} xf(x)dx \; Donc \; en \; simplifiant: E(x) = \int_0^2 0{,}1x^4 + 0{,}3x \, dx = \left[\frac{0{,}1x^5}{5} + \frac{0{,}3x^2}{2}\right]_0^2 =$ 1,24

QCM n°8 : **B, D, E**

A. Faux. On réalise un test du chi 2 pour deux variables qualitatives, si α augmente, tα diminue.
B. **Vrai**. Si β augmente, la puissance (1-β) = probabilité de rejeter H0 alors qu'elle est fausse, diminue
C. Faux. La probabilité de rejeter H0 alors qu'elle est vraie correspond au risque α. Celui-ci est indépendant du risque β, et donc de la puissance.
D. **Vrai**. α correspond à la « probabilité de rejeter H0 alors qu'elle est vraie », donc « La probabilité de ne pas rejeter H0 alors qu'elle est vraie » correspond à 1-α. Ces deux élements varient donc en sens inverse.
E. **Vrai**.

QCM n°9 : **C, D, E**

A. Faux. On utilise bien un test de l'écart-réduit en unilatéral car on cherche à montrer que les étudiants descendent plus de pistes que la référence, mais on n'a besoin que de la condition n>30. Les autres conditions sont nécessaires pour le test de Student.
B. Faux. $t_{obs} = \frac{\bar{x}-\mu}{\frac{S}{\sqrt{n}}} = \frac{117-114}{\frac{\sqrt{81}}{\sqrt{50}}} = 2{,}357$.
C. **Vrai**. On est en <u>unilatéral</u> donc on lit le t_α à α×2=0,1.
D. **Vrai**. $t_{obs} > t_\alpha$ → Rejet H0.
E. **Vrai**. On cherche le t_{obs} dans la table de l'écart-réduit et on trouve 0,01 < p-value < 0,02.
Mais on est en unilatéral, il faut donc diviser la p-value par deux. On a donc : 0,005 < p-value < 0,01.

QCM n°10 : **E**

	Stévia	**Glucose**	**Fructose**	**Totaux**
Pas d'obésité	**13** $\frac{57\times38}{150} = 14{,}44$	**26** $\frac{57\times62}{150} = 23{,}56$	**18** $\frac{57\times50}{150} = 19$	57
Obésité	**25** $\frac{93\times38}{150} = 23{,}56$	**36** $\frac{93\times62}{150} = 38{,}44$	**32** $\frac{93\times50}{150} = 31$	93
Totaux	38	62	50	150

A. Faux**.** Le test de l'écart réduit ne permet pas de comparer plus de deux proportions.
B. Faux. $X^2_{obs} = \sum \frac{(O_{ij}-E_{ij})^2}{E_{ij}}$.
C. Faux. $X^2_{obs} = \frac{(13-14,44)^2}{14,44} + \frac{(26-23,56)^2}{23,56} + \frac{(18-19)^2}{19} + \frac{(25-23,56)^2}{23,56} + \frac{(36-38,44)^2}{38,44} + \frac{(32-31)^2}{31} = 0{,}724$.
Mais X^2_α n'est pas lu à 6 ddl : (r-1)×(k-1) = (3-1)×(2-1) = 2×1 = 2 ddl.
(Attention de ne pas compter les colonnes et lignes des totaux)
D. On a X^2_{obs}= 0,724 et X^2_α= 5,991 à 5% et 2ddl. $X^2_{obs} < X^2_\alpha$ → Non rejet de H0.
E. **Vrai**. On cherche la p-value dans la table du X² à 2 ddl.

QCM n°11 : **B, C**

A. Faux. Les échantillons sont bien indépendants, mais on utilise donc un test de Mann-Whitney, le test de Wilcoxon étant pour des échantillons appariés.
B. **Vrai**.
On range les différentes valeurs dans l'ordre croissant et on leur attribue un rang :

Valeur	0,8	0,85	0,9	1,05	1,1	1,2	1,2	2,4	3,2
Rang	1	2	3	4	5	6,5	6,5	8	9
Groupe	1	1	2	1	2	1	2	2	2

S_{rg1} = 1+2+4+6,5 = 13,5
S_{rg2} = 3+5+6,5+8+9 = 31,5
C. **Vrai**.
$U_1 = n_1n_2 + \frac{n1(n1+1)}{2} - S_{rg1} = 4\times5 + \frac{4(4+1)}{2} - 13{,}5 = 16{,}5$
$U_2 = n_1n_2 + \frac{n2(n2+1)}{2} - S_{rg2} = 4\times5 + \frac{5(5+1)}{2} - 31{,}5 = 3{,}5$
U = 3,5.
D. Faux. La valeur critique de la statistique de test est celle qui est lue dans la table. 3,5 est la statistique de test.

E. Faux. $U_\alpha = 1$ à 5%. On a $U > U_\alpha$ donc on ne rejette pas H0.

QCM n°12 : **F**

A. Faux. L'erreur systématique correspond aux biais, elle n'est pas liée à la fluctuation d'échantillonnage.
B. Faux. C'est le biais de sélection.
C. Faux. C'est l'inverse, les enquêtes prospectives possèdent en principe moins de bais que les rétrospectives.
D. Faux. L'analyse en ITT n'a de sens que dans les enquêtes expérimentales.
E. Faux. C'est le but de l'ETC.

QCM n°13 : **A, C, E**

A. **Vrai**. Cf. : cours
B. Faux. On suppose les « deux groupes représentatifs de leurs populations respectives » donc pas de biais important de sélection.
C. **Vrai**. On cherche donc P(Médecins/Cirrhose) ; elle est de 40/2000 soit de 2%.
D. Faux. On ne calcule pas de risque relatif dans les enquêtes cas/témoins.
E. **Vrai**. On peut approximer car la prévalence est inférieure à 1% ; $OR = \frac{ad}{bc} = \frac{40 \times 9549}{51 \times 1960} = 3{,}82.$

QCM n°14 : **B, C**

A. Faux. Etude longitudinale puisqu'elle s'étend dans le temps.
B. **Vrai**. RR=$\frac{5/_{5+7}}{20/_{20+3}} = 0{,}479 < 1$, et comme l'intervalle de confiance ne contient pas la valeur 1 (voir énoncé), il s'agit bien d'un facteur de protection.
C. **Vrai**. PRA= $\frac{p \times (RR-1)}{p \times RR + (1-p)}$ avec p comme proportion de personnes exposées dans la population.
ER= $P(M/E) - P(M/\bar{E})$
D. Faux. Bien qu'il existe des biais de perdus de vue, les biais de sélection sont quantitativement peu importants.
E. Faux. Une enquête de cohorte exposés/non exposés permet d'étudier UNE EXPOSITION RARE pour une OU PLUSIEURES maladies.

QCM n°15 : **A, D**

A. **Vrai**. sp= $P(T^-/M^-) = \frac{5}{5} = 1$
B. Faux. se= $P(T^+/M^+) = \frac{3}{5}$ De plus, la sensibilité est une probabilité conditionnelle, donc forcément ≤1.
C. Faux. $RV^- = \frac{1-Se}{Sp} = \frac{2/5}{1} = \frac{2}{5} \approx 0{,}4$ 0,714 correspond à la VPN.
D. **Vrai**. $RV^+ = \frac{Se}{1-Sp}$ $(1 - Sp) = 0$, donc RV^+ tend vers l'infini. Or la valeur diagnostique d'un test positive est d'autant plus grande que RV^+ est grand.
E. Faux. On ne peut pas calculer la VPP ici.