TUTORAT UE 4 2014-2015 – Biostatistiques
CORRECTION Séance n°7 – Semaine du 24/11/2014

Séance de révisions générales

<u>QCM n°1 :</u> **A, D, E**

A. **Vrai**. $\frac{C_5^2 \times C_7^2}{C_{20}^4} = \frac{10 \times 21}{4845}$ = 0,043.

B. Faux. $\frac{C_8^4}{C_{20}^4} + \frac{C_7^4}{C_{20}^4} + \frac{C_5^4}{C_{20}^4} = \frac{70+35+5}{4845}$ = 0,023.

C. Faux. $\frac{C_7^3 * C_{13}^1}{C_{20}^4} + \frac{C_7^4}{C_{20}^4} = 0{,}1011.$

D. **Vrai**. A_{20}^3= 6840.

E. **Vrai**. $\frac{A_8^1 \times A_7^1 \times A_5^1}{A_{20}^3} = \frac{8 \times 7 \times 5}{6840}$ = 0,041.

<u>QCM n°2 :</u> **E**

A. Faux. P(E) = P(E/A)xP(A) + P(E/B)xP(B) = 0,05x0,25+0,15x0,75 = 0,125.

B. Faux. P(A/E)= $\frac{P(E/A) \times P(A)}{P(E)} = \frac{0{,}05 \times 0{,}25}{0{,}125}$ = 0,1.

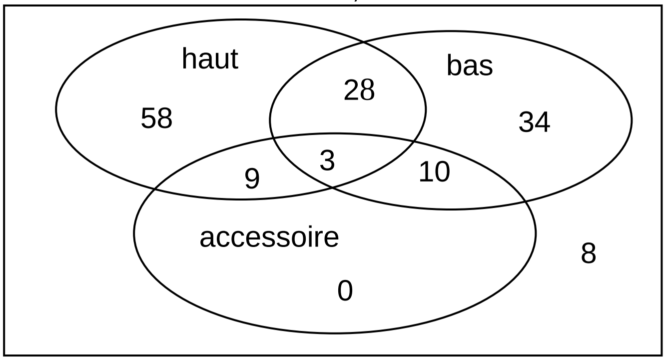
C. Faux. P(B/E) = P($\bar{A}$/E) = 1-P(A/E) = 1-0,1 = 0,9.

D. Faux. P($\bar{E}$/B) = 1-P(E/B) = 1-0,15 = 0,85.

E. **Vrai**. P($\bar{E}/\bar{A}$) = P($\bar{E}$/B) = 0,85.

<u>QCM n°3 :</u> **B, D**

Soit les événements A « acheter un accessoire », B « acheter un bas » et H « acheter un haut ».



A. Faux. $\frac{3}{150}$=0,02.

B. **Vrai**.

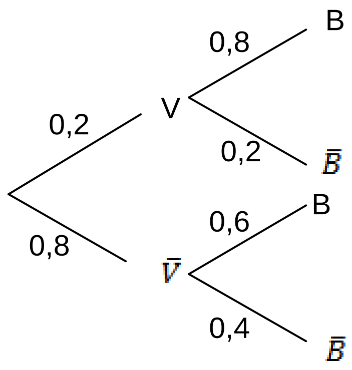
C. Faux. $\frac{8}{150}$=0,05.

D. **Vrai**.

E. Faux. P(H∩A) = $\frac{12}{150}$=0,08 est différent de P(H)xP(A)=0,096, avec P(H)=$\frac{98}{150}$ et P(A)=$\frac{22}{150}$ (selon les données de l'énoncé).

QCM n°4 : **A, B, C**

Soit les événements B « avoir les cheveux bruns » et V « avoir les yeux verts ».



A. **Vrai**. P(B)=0,2x0,8+0,8x0,6=0,64.
B. **Vrai**. P($\bar{B}$/V)=0,2.
C. **Vrai**. P(B∩V)=P(B/V)xP(V)=0,2x0,8=0,16.
D. Faux. P($\bar{B} \cap \bar{V}$)= P($\bar{B}/\bar{V}$)xP($\bar{V}$)=0,4x0,8=0,32.
E. Faux. P(B)xP(V)=0,64x0,2=0,128 est différent de P(B∩V)=0,16.

QCM n°5 : **B, D, E**

A. Faux. n=44 et p=1/5
B. **Vrai**. P(X=44)=$(1/5)^{44}$ =1,76 x 10^{-31}.
C. Faux. P(X=9)=0,147>P(X=6)=0,0938.
D. **Vrai**. n>20 et p<0,5 : les conditions sont vérifiées donc on peut réaliser une approximation par une loi de Poisson de paramètre λ=np=8,8.
E. **Vrai**. P(X=0)=$e^{-\lambda}$=$e^{-8,8}$=1,5x10^{-4}.

QCM n°6 : **A, B, C, D, E**

A. **Vrai**.
B. **Vrai**. n≥30, np≥5 et n(1-p)≥5 donc on peut approximer par N(np ; $\sqrt{npq}$) soit N(22 ; 4,28).
C. **Vrai**. */!\ pensez à utiliser l'écart type sans arrondir et à employer la correction de continuité*.

$$P(X<10)=P(U<\frac{9,5-22}{\sqrt{\frac{22*5}{6}}})\rightarrow \pi(-2,92)=0,0017\,.$$

D. **Vrai**. $P(X\leq 23)=P(U<\frac{23,5-22}{\sqrt{\frac{22*5}{6}}})\rightarrow \pi(0,35)=0,6368$

E. **Vrai**. P $P(X\geq 26)=1-P(X<26)=1-P(U<\frac{25,5-22}{\sqrt{\frac{22*5}{6}}})\rightarrow 1-\pi(0,82)=1-0,7939=0,2061.$

QCM n°7 : **A, C, E**

A. **Vrai**. $S^2 = 4 * \frac{27}{26} = \frac{54}{13} \approx 4{,}154$.

B. Faux. n<30, on utilise la table de Student.

C. **Vrai**. $[10{,}4 - 2{,}056 * \frac{\sqrt{\frac{54}{13}}}{\sqrt{27}} ; 10{,}4 + 2{,}056 * \frac{\sqrt{\frac{54}{13}}}{\sqrt{27}}] = [9{,}594 ; 11{,}206]$.

Remarque : on lit t_α dans la table de Student à 26 ddl.

D. Faux. On lit t_α dans la table de l'écart-réduit et on n'a pas besoin d'estimer S^2 car n>30.

E. **Vrai**. IC à 95% est de $[11{,}2 - 1{,}96 * \frac{1{,}8}{\sqrt{36}} ; 11{,}2 + 1{,}96 * \frac{1{,}8}{\sqrt{36}}] = [10{,}612 ; 11{,}788]$.

QCM n°8 : **C, E**

A. Faux.

B. Faux.

C. **Vrai**.

D. Faux. C'est le résultat que l'on trouve si l'on n'avait pas employé la variance <u>estimée</u>.

E. **Vrai**. On calcule la variance estimée $S^2 = \frac{n}{n-1} s^2 = 9{,}375$.

On cherche a et b tels que : $P(khi^2 \leq a) = 0{,}10$
$P(khi^2 \leq b) = 0{,}90$

D'où a=15,659 et b=33,196 lus dans la table du Chi-deux à 24 ddl.

Puis on remplace dans la formule de notre intervalle : $[(n-1)\frac{S^2}{b} ; (n-1)\frac{S^2}{a}] = [24 * \frac{9{,}375}{33{,}196} ; 24 * \frac{9{,}375}{15{,}659}] = [6{,}78 ; 14{,}37]$.

QCM n°9 : **C, D, E**

A. Faux. H0 : l'espérance de vie des sportifs n'est pas différente de celle de la population.

B. Faux. n>30, on utilise le test de l'écart réduit.

C. **Vrai**.

D. **Vrai**. $t_{obs} = \frac{84{,}3 - 82{,}5}{1/\sqrt{200}} = 25{,}46$ et $t_\alpha = 1{,}645$, t_{obs} est supérieur à t_α donc on rejette H0.

E. **Vrai**. $t_\alpha = 1{,}96$, t_{obs} est supérieur à t_α donc on rejette H0.

QCM n°10 : **A, C, D, E**

A. **Vrai**.

B. Faux. Les échantillons sont indépendants.

C. **Vrai**. Le test de Student permet de comparer deux moyennes sur deux échantillons indépendants. Il nécessite des conditions d'applications : normalité des valeurs (car petits échantillons) et égalité des variances.

D. **Vrai**. $t_{obs} = \frac{|142 - 139|}{\sqrt{\frac{S^2}{8} + \frac{S^2}{5}}}$ avec $S^2 = \frac{7 * 1{,}5^2 + 4 * 1{,}5^2}{8 + 5 - 2} = 2{,}25$ donc $t_{obs} = 3{,}508$.

On lit dans la table de Student à 8+5-2=11 ddl un $t_\alpha = 1{,}796$.

E. **Vrai**. $t_\alpha = 3{,}106$, t_{obs} est supérieur à t_α donc on rejette H0.

QCM n°11 : C

A. Faux. Le test du X² de Mac Nemar s'emploie pour des échantillons appariés (avant/après), ce qui n'est pas le cas de l'échantillon considéré.
B. Faux. Les effectifs théoriques sont inférieurs à 5.
C. **Vrai**. On utilise le test exact de Fischer car les conditions d'application du test du chi-deux et de l'écart réduit ne sont pas vérifiées.
D. Faux. p-value>α, pour α=5% et pour α=1% donc on ne rejette pas l'hypothèse H0 dans les deux cas.
E. Faux. *Cf. item D.*

QCM n°12 : A, B, C, D, E

A. **Vrai**. r = $\frac{\sum(x_i-m_x)(y_i-m_y)}{\sqrt{\sum(x_i-m_x)^2\sum(y_i-m_y)^2}} = \frac{18-1-2+21}{\sqrt{(9+1+1+9)(36+1+4+49)}} = \frac{3\sqrt{2}}{5}$.

B. **Vrai**. On calcule les variances observées : s_y^2=22,5 et s_x^2=5 puis les écarts types estimés :
$S_y=\sqrt{\frac{4}{3}\times 22{,}5} = \sqrt{30}$ et $S_x=\sqrt{\frac{4}{3}\times 5} = \sqrt{\frac{20}{3}}$.
L'équation de la droite est : $y - m_y = r\frac{S_y}{S_x}(x - m_x) \rightarrow y - 11 = \frac{3\sqrt{2}}{5} \times \frac{\sqrt{30}}{\sqrt{\frac{20}{3}}}(x-5)$. Par simplification, on obtient bien une équation telle que $y = 2 + 1{,}8x$.

C. **Vrai**. $y = 2 + 1{,}8 * 4 = 9{,}2$.

D. **Vrai**. On calcule d'abord l'écart type lié $S'_y = \sqrt{(1-r^2)S_y^2} = \sqrt{(1-(\frac{3\sqrt{2}}{5})^2)\times 30} = 2{,}8983$.
L'intervalle est [9,2-1,96x2,8983 ; 9,2+1,96x2,8983] → [3,519 ; 14,881].

E. **Vrai**. On calcule la note moyenne en ayant suivi 6 séances : $y = 2 + 1{,}8 * 6 = 12{,}8$.
L'intervalle est [12,8-1,96x2,8983 ; 12,8+1,96x2,8983] → [7,119 ; 18,481].

QCM n°13 : A, E

A. **Vrai**. On sélectionne entre autres un échantillon parmi une population de malades.
B. Faux. Un fort biais de classement dû au recueil a posteriori de l'information.
C. Faux. C'est une étude rétrospective, elle ne donne pas l'incidence puisque l'on n'a pas la possibilité d'apprécier l'apparition des nouveaux cas, tous les cas étant déjà connus.
D. Faux. VPP et VPN ne sont pas des indices à calculer dans le cadre d'enquêtes ; leur application se réserve aux tests et signes diagnostiques.
E. **Vrai**.

QCM n°14 : **A, D**

Un tableau nous facilitera le raisonnement :

	Malades	Non malades	Total
Exposés	7500	2500	10000
Non Exposés	11000	29000	40000
Total	18500	31500	50000

/!\ on ne tient pas compte des gens déjà atteints dans une enquête prospective.

A. **Vrai**.

B. Faux. $RR = \frac{7500/10000}{11000/40000} = \frac{30}{11} \approx 2{,}727.$

C. Faux. $ER = \frac{7500}{10000} - \frac{11000}{40000} = \frac{19}{40} = 0{,}475.$

D. **Vrai**. Soit la proportion de personnes exposées au facteur $p = \frac{10000}{50000} = 0{,}2.$

On a : $PRA = \frac{p*(RR-1)}{p*RR+(1-p)} = \frac{19}{74} \approx 0{,}257..$

E. Faux. Le caractère prospectif limite les biais de classement. En revanche, on prêtera attention particulièrement au biais de perdus de vue et au biais de confusion.

QCM n°15 : **A, C, E**

Un tableau nous aidera :

	malades	Non malades	tot
Test +	160	60	220
Test -	40	140	180
tot	200	200	400

A. **Vrai**. P(T$^-$/M$^-$)= $\frac{140}{200} = 0{,}7.$

B. Faux. Jamais de VPP, car elle dépend de la prévalence et donc de notre échantillon. Ici, il est volontairement non représentatif de la population.

C. **Vrai**. $RV+ = \frac{Se}{1-Sp} = \frac{P(T/M)}{P(T/\overline{M})} = \frac{0{,}8}{1-0{,}7} = \frac{8}{3}.$

D. Faux. Il y a 220 positifs dont 60 faux positifs donc seulement 160 vrais positifs (il y a bien 40 faux négatifs).

E. **Vrai**. Plus RV$^-$ est proche de 0 et plus le résultat d'un test négatif est fiable. $RV- = \frac{P(\overline{T}/M)}{P(\overline{T}/\overline{M})} = \frac{40/200}{140/200} = \frac{2}{7}$

QCM n°16 : **D, E**

A. Faux. L'augmentation de l'incidence est toujours péjorative ; en revanche, la prévalence peut augmenter si on arrive à mieux prendre en charge une pathologie en diminuant sa létalité par exemple, ce qui n'est pas péjoratif.

B. Faux. Si l'incidence augmente, il y a plus de porteurs de la maladie donc plus de mortalité alors que la létalité reste la même.

C. Faux. Le NSN augmente quand la variabilité augmente ou que la différence attendue diminue.

D. **Vrai**.

E. **Vrai**.