FACULTE
De
PHARMACIE

TUTORAT UE4 2011-2012 – Biostatistiques
CORRECTION Séance n°7 – Semaine du 28/11/2011

Séance de révision générale

Séance préparée par Nicolas Dubois

QCM n°1 : **a, b, d, e**

a) **Vrai** : $A^5_{32} = 24165120$
b) **Vrai** : $C^5_{32} = 201376$
c) Faux : Il n'y a que 4 dames dans un jeu de 32 cartes. S'il le tirage comporte 4 dames, il ne reste que 32-4 = 28 possibilités pour la 5ème carte qui n'est pas une dame.
d) **Vrai** : $C^2_4 \times C^3_{28} = 19656$
e) **Vrai** : $(C^2_4 \times C^3_4) / C^5_{32} = 1,19 \times 10^{-4}$
En effet : le nombre de combinaisons favorables de 2 dames et de 3 as = $C^2_4 \times C^3_4$
Le nombre de combinaisons possibles de 5 cartes = C^5_{32}

QCM n°2 : **a, c**

a) **Vrai** : ces deux paramètres sont dits « intrinsèques » car ils ne dépendent pas de la prévalence, à la différence de la VPP et la VPN (paramètres extrinsèques).
b) Faux : c'est l'inverse !
c) **Vrai** : la sensibilité est la probabilité d'avoir un test positif sachant que l'on est malade.

$$Se = \frac{69}{69+27} = 0,72$$

d) Faux : la VPN est la probabilité de ne pas être malade sachant que le test est négatif.

$$VPN = \frac{94}{94+27} = 0,78$$

e) Faux : $VPP = P(M/T) = \dfrac{P(T/M).P(M)}{P(T/M).P(M) + P(T/\overline{M}).P(\overline{M})}$

NB : on voit ici que la VPP dépend de P(M) qui est la prévalence de la maladie (cf : b)).

QCM n°3 : **b**

a) Faux : l'échantillon est bien représentatif car il a été tiré au sort et non parce que 100>30 (aucun rapport.)

b) **Vrai**

c) Faux : Moyenne= $\frac{800\times13+900\times29+1000\times35+1100\times8+1200\times15}{100}=983€$

d'où variance= $\frac{(800-983)^2\times13+(900-983)^2\times29+(1000-983)^2\times35+(1100-983)^2\times8+(1200-983)^2\times15}{99}=14759€$

/ ! Bien diviser par 99, soit n-1 car nous cherchons sur la population. Sinon, diviser par 100 pour trouver la variance de l'échantillon et appliquer le facteur $\frac{n}{n-1}$ pour étendre à la population. D'où écart-type=$\sqrt{14759}$=121,48 €, attention aux unités !

d) Faux : Le mode vaut 1000 ici, soit la valeur ayant l'effectif le plus élevé. Il peut y en avoir plusieurs, sur un histogramme le mode est la valeur d'une colonne étant encadrée par deux colonnes plus petites. On comprend aisément comment il peut y en avoir plusieurs ainsi.

e) Faux : Il faut faire une ligne supplémentaire au tableau pour trouver les effectifs cumulés croissants :

Salaire en €	800	900	1 000	1 100	1 200
Nombre de personnes	13	29	35	8	15
Effectif cumulé croissant	13	42 (=13+29)	77 (=42+35)	85 (=77+8)	100 (=85+15)

Comment trouver Q1 ? on sait que par définition Q1=Q.0,25 donc 0,25.100 (effectif total)=25 => 42 (on prend la valeur d'effectif cumulé croissant supérieur à 25) donc Q1= 900€
Pour Q3, même démarche : Q3=Q.0,75 donc 0,75.100=75 =>77 donc Q3= 1000€
De même pour la médiane= Q.0,5 donc 0,5.100=50 =>77 donc médiane=1000€

QCM n°4 : **a, d**

a) **Vrai**
b) Faux : On a B(n ;p) => B (60 ; 0,04). On cherche ici P(X=7)= $C^7_{60}\,0{,}04^{7}(1-0{,}04)^{60-7}$=0,0073
c) Faux : E(X)=np=60x0,04=2,4 et Var(X)=np(1-p)=2,304
d) **Vrai :** On a bien n>20 et p<0,5 donc B(60 ;0,04) => P(lambda=np=2,4)
e) Faux : Elles sont toutes les deux des lois discrètes.

QCM n°5 : **b, c, e**

a) Faux

b) Vrai

c) **Vrai,** on suit une loi binomiale X~B (n ; 1/6) car le dé est non truqué et on effectue n fois la même expérience de façon indépendante.

d) Faux

e) **Vrai :** $P(X \geq 1) > 0.9$

$P(X \geq 1) = 1 - P(X=0) = 1-(5/6)^n$ car P(X=0) = $C^0_n(1/6)^0(1-1/6)^n=(5/6)^n$

$1-(5/6)^n > 0.9 \rightarrow (5/6)^n < 0.1$

$n\,ln(\frac{5}{6}) < \ln(0.1)$

$n > \frac{\ln(0.1)}{\ln(\frac{5}{6})} = 12.629\ soit\ n \geq 13$

QCM n°6 : **a, b, c, e**

a) **Vrai**
b) **Vrai**
c) **Vrai** : on a $p_o - c_{\alpha/2} \times \sqrt{p_o(1-p_o)/n}$ =0.462

Et $p_o + c_{\alpha/2} \times \sqrt{\frac{p_o \times (1-p_o)}{n}}$ = 0.656

⇨ $P_o = \frac{0.462 + 0.656}{2}$ = 0.559

d) Faux : on cherche $c_{\alpha/2}$ pour n=40, puis par lecture inverse sur la table de l'écart réduit on lira α.
On veut le même intervalle :
Donc on a $p_o + c_{\alpha/2} \times \sqrt{\frac{p_o \times (1-p_o)}{n}}$ = 0.656
Donc $c_{\alpha/2}$=(0.656-0.559) $\times \sqrt{\frac{40}{0.559 \times 0.441}}$ = 1.236
Par lecture inverse, $\alpha \approx 0.22$
e) **Vrai**

QCM n°7 : **c, e**

a) Faux : $\bar{X}$ = **3,48**
b) Faux : La variance de la moyenne se calcule ainsi : Var($\bar{X}$) = $\frac{\sigma^2}{n} = \frac{0,27739^2}{7}$ = **0,01099**
c) **Vrai** : Ici on aura bien $t_{n-1,\alpha/2}\frac{S}{\sqrt{n}} = 2,447\frac{0,27739}{\sqrt{7}}$ = **0,257.**
d) Faux : Si on multiplie l'effectif par 3, on aura un effectif de 21. Refaisons le calcul et vérifions simplement : $t_{n-1,\alpha/2}\frac{S}{\sqrt{n}} = 2,447\frac{0,27739}{\sqrt{21}}$ = **0,148.** Ainsi cela aura été divisé par $\sqrt{3}$ et non 3.
e) **Vrai** : voir a)

QCM n°8 : **b, e**

a) Faux : X suit en fait une loi Binomiale : $X \sim B(300; 0,1)$
Avec alors : n=300>30 ; np= 30>5 et nq= 270>5. On peut donc approximer X par une loi Normale telle que $X \sim N(np = 30; \sqrt{npq} = \sqrt{27})$
Donc $\sigma(x) = \sqrt{27}$
b) **Vrai :**P(30≤X≤40)=P(29,5≤X≤40,5)=$\Pi(\frac{40,5-30}{\sqrt{27}}) - \Pi(\frac{29,5-30}{\sqrt{27}}) = \Pi(2,02) - \Pi(-0,1)$
=0,9783-0,4602 = 0,5181

c) Faux
d) Faux : P(X≤35)= P(X≤35,5)= $\pi(\frac{35,5-30}{\sqrt{27}}) = \pi(1,06) = 0,8554$
e) **Vrai**

<u>QCM n°9 :</u> **b, d**

a) Faux : on vérifie que tous les effectifs théoriques soient supérieurs ou égaux à 5.
b) **Vrai** : on utilisera un test de l'écart-réduit pour des échantillons de taille supérieure à 30, tandis que l'on pourra utiliser le test de Student pour tout échantillon dont la distribution suivra une loi normale.
c) Faux : les distributions doivent être normales dans les deux populations d'où proviennent les échantillons.
d) **Vrai**
e) Faux : pour un test de l'écart-réduit, on rejette H_0 si $t_{obs} > t_\alpha$ ou si p-value < α.

<u>QCM n°10</u> : **a, b, c**

a) **Vrai**
b) **Vrai** : par lecture dans la table de Chi ² à 2 ddl.
c) **Vrai :** le tableau théorique :

	peu essoufflé	très essoufflé	sur les rotules
non-fumeur	15,875	63,5	47,625
fumeur	9	36	27

$\sum \frac{(E-O)^2}{E}$ = 5,27

d) Faux : ça n'a rien à voir, on augmentera la faculté du test à montrer une différence. Le α est fixé a priori.
e) Faux : on choisi en fonction de α on prend ce risque de se tromper mais la deuxième partie est fausse !!

<u>QCM n°11</u> : **f**

a) Faux : H_0 : les deux distributions sont identiques.
b) Faux : on utilise un test de Mann-Whitney, mais parce que les distributions ne sont pas appariées.
c) Faux : H_0 ne nous permet pas d'admettre que U suit une loi normale, pour cela, il faudrait que n_1 et/ou n_2 soient grands, ce qui n'est pas le cas ici, donc on ne peut pas l'admettre.
d) Faux : le test des signes est un test utilisé pour des distributions appariées.
e) Faux : On réalise le tableau suivant :

Valeurs	0.3	0.4	0.5	0.7	0.7	1	1.5	1.5	1.7	2	2	3
Rangs	1	2	3	4.5	4.5	6	7.5	7.5	9	10.5	10.5	12
R_{P1}	1	2	3	4.5		6	7.5			10.5		
R_{P2}					4.5			7.5	9		10.5	12

S_{P1} = 1+2+3+4.5+6+7.5+10.5 = 34.5
S_{P2} = 4.5+7.5+9+10.5+12 = 43.5

$$U_{P1} = 7 \times 5 + \frac{7 \times (7+1)}{2} - 34.5 = 28.5$$

$$U_{P2} = 7 \times 5 + \frac{5 \times (5+1)}{2} - 43.5 = 6.5$$

U = min (UP1 ; UP2) = 6.5
Par lecture de la table de Mann-Whitney, on a U_{seuil} = 5
$U > U_{seuil}$, donc on ne peut pas rejeter H_0

QCM n°12 **: a, c, e**

a) **Vrai**
b) Faux
c) **Vrai** : $T_{obs}= m_{\Delta}/S_{\Delta}$ *racine de n = 97.8
d) Faux : ho : $m_{\Delta}=0$
e) **Vrai** : bande d'incultes !

QCM n°13 **: a, c, d, e**

a) **Vrai**
b) Faux : On utilisera le test de l'écart réduit car n_1 et n_2 sont tous les deux supérieurs à 30.
c) **Vrai** : Tobs=$\frac{m1-m2}{\sqrt{(\frac{s1^2}{n1}+s2^2/n2)}}$=4,35
d) **Vrai** : 4,35>1,96=tα donc rejet de Ho
e) **Vrai** : Etant donné que nous avons rejeté Ho, on peut dire qu'il y a un effet de l'anticoagulant sur le corps humain.

QCM n°14 ***: a, b, d***

a) **Vrai** : grâce aux indicateurs obtenus par les enquêtes descriptives (prévalence, incidence, mortalité…), on peut faire des hypothèses étiologiques sur les éventuels facteurs de risque. On vérifiera ensuite ces hypothèses grâce aux indicateurs obtenus par les enquêtes analytiques (FR, OR…).
b) **Vrai** : les mesures d'association (facteur de risque, odds ratio) sont des indicateurs multiplicatifs.
c) Faux : Le biais de confusion est une erreur systématique et non aléatoire (le reste de la phrase est vrai). Attention ! les 3 conditions du facteur de confusion sont à connaître +++.
d) **Vrai**
e) Faux : la clause d'ambivalence est un critère de non-inclusion.

QCM n°15 : **a, d**

a) **Vrai** : la prévalence d'une maladie est bien le rapport entre le nombre de malades de la maladie considérée sur l'effectif total de la population considérée.
b) Faux : c'est une enquête de prévalence que l'on réalise.
c) Faux : l'enquête de prévalence est bien une enquête observationnelle mais elle est transversale (on voit les patients à un moment donné précis) et non pas longitudinale (où l'on suit les patients dans le temps).
d) **Vrai** : c'est bien une étude de la répartition d'un état de santé, sans recherche d'étiologie, ni évaluation de mesures prises contre l'épidémie.
e) Faux : on est bien dans de l'épidémiologie évaluative mais de population (action de prévention primaire, ici la vaccination) et non de recherche clinique (par exemple, évaluation de médicaments).

__QCM n°16 :__ **e**

a) Faux : c'est une enquête de cohorte exposés-non exposés, car elle est constituée de 2 groupes de personnes indemnes, suivies pendant 6 ans.

b) Faux : Excès de risque = P(M/A) - P(M / $\overline{A}$) = $\frac{190}{190+60} - \frac{20}{20+230}$ = 0,76-0,08 =0,68

c) Faux : RR = $\frac{\mathrm{P(M/A)}}{\mathrm{P(M/\overline{A})}}$ = $\frac{0{,}76}{0{,}08}$ = 9,5

d) Faux : PRA = $\frac{p \times (RR-1)}{p \times RR + (1-p)}$ = $\frac{0{,}3 \times (9{,}5-1)}{0{,}3 \times 9{,}5 + (1-0{,}3)}$ = 0,72

e) **Vrai** : car PRA = 0,72

__QCM n°17 :__ **a, b**

a) **Vrai**

b) **Vrai**

c) Faux : biais de sélection avec les perdus de vue, celui d'information est maitrisable par la mesure de l'expo …

d) Faux : dans l'exposé non exposé on cherche les maladies favorisées par l'exposition et non l'inverse.

e) Faux : dans le nombre de malades qui étaient exposés certains ont quand même développé la maladie, d'où le facteur rr-1 au numérateur.