TUTORAT UE4 2012-2013 – Biostatistiques
CORRECTION Séance n°3 – Semaine du 08/10/2012

Lois de probabilité- Estimations

Séance préparée par Chayma IGHIDI, Guilhaume GARRIGUES, Laura MILHAU
et Grégoire SARTHOULET

QCM n°1 : **A, B, C**

A. Vrai. C'est la définition

B. Vrai. la **variance estimée dans la population** (on ne connaît pas la valeur de la moyenne) se calcule ainsi :S² = $\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\mu)^2$

Et la **variance observée de l'échantillon** (on connaît la valeur de la moyenne) est :

$$s^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i-\mu)^2$$

C. Vrai.

D. Faux. $\mu = \frac{\sum x_i}{n} = \frac{3000}{100} = 30$

E. Faux. On calcule la variance observée dans l'échantillon : $s^2 = \frac{1}{n} \times (\sum x_i{}^2) - \bar{x}^2 = \frac{1}{100} \times (300\,000) - 30^2 = 2100$. Mais on nous demande de calculer s donc $s = \sqrt{s^2} = \sqrt{2100} = 45.8258 = 45.8 = 46$ (tous ces arrondis seraient justes).

QCM n°2 : **A, D, E**

A. **Vrai.** $\mu = \frac{140}{20} = 7$

B. Faux. On cherche la variance observée dans l'échantillon : s²= $\frac{1}{20} \times 1040 - 7^2 = 3$

C. Faux. L'écart type est donc s=$\sqrt{3} = 1{,}7321$

D. **Vrai.** La formule de l'intervalle de confiance de la variance est : $\left[(n-1)\frac{S^2}{b};(n-1)\frac{S^2}{a}\right]$ avec S² la variance estimée dans la population.

- Calcul de la variance estimée dans la population. On a calculé dans l'item b la variance observée dans l'échantillon donc on peut calculer la variance estimée dans la population: S²=$\frac{n}{n-1}s^2 = \frac{20}{19} \times 3 = 3{,}1579$.
- Calcul de a et de b : a tel que $P\left(\boldsymbol{\chi^2}_{\mathbf{19}} \leq \boldsymbol{a}\right) = {}^{\boldsymbol{\alpha}}/_{\mathbf{2}}$ et b tel que P($\chi^2{}_{19} \leq b) = 1 - {}^{\alpha}/_{2}$
  Pour a : on lit dans la table du Chi deux à 19 ddl et $\alpha = 0{,}90$ et on trouve a=11,651
  Pour b : on lit dans la table du Chi deux à 19 ddl et $\alpha = 0{,}10$ et on trouve b=27,204
- Calcul de l'intervalle : $\left[19 \times \frac{3{,}1579}{27{,}204}; 19 \times \frac{3{,}1579}{11{,}651}\right] = [2.2056; 5.1498] = [2.21; 5{,}15]$

E. **Vrai.** On fait la racine carrée de l'intervalle de la variance : $\left[\sqrt{2.2056};\sqrt{5.1498}\right] = [1.4851; 2.2693] = [1.49; 2.27]$.

QCM n°3 : **F**

A. Faux. $\lambda \geq 20$, on peut donc faire l'approximation par la loi normale mais on obtient $X \sim N(\lambda = 20; \sqrt{\lambda} = 4{,}4721) \rightarrow P(X = 30) = P(29{,}5 \leq X \leq 30{,}5) = \pi(2{,}35) - \pi(2{,}12) = 0{,}9906 - 0{,}9830 = 7{,}6.10^{-3}$.
B. Faux. $\lambda < 20$ on n'approxime donc pas par la loi normale.
C. Faux. $\sqrt{2X} \sim N(\sqrt{69}; 1)$ L'approximation est possible car n>30.
D. Faux. On ne peut faire l'approximation de la loi de Student par la loi Normale centrée réduite si n>30.
E. Faux. On ne fait la correction de continuité que lorsqu'on approxime une loi discrète par la loi continue.
F. Vrai.

QCM n°4 : **B, C, E**

A. Faux. Lecture directe de la table du Chi2 pour 25 d.d.l pour une probabilité de 0,5 on lit une valeur de t=24,337
B. **Vrai**. Lecture directe de la table du Chi2 pour 13 d.d.l on cherche la valeur 12,340 qui correspond à une probabilité de 0,5.
C. **Vrai.** Définition.
D. Faux. D'après la table : $P(|T| > 1{,}069) = 0.30$ or $P(|T| > 1{,}069) = P(T > 1{,}069) + P(T < -1{,}069)$=$2P(T > 1{,}069)(car\ la\ loi\ est\ symétrique) = 0{,}3 \rightarrow P(T > 1{,}069) = 0{,}15$.
E. Vrai. D'après la table, $P(|T| > 1.341) = 0{,}10$

QCM n°5 : **A, C, E**

A. **Vrai** $\bar{x} = \frac{1}{n}\sum n_i x_i = 30{,}785$.
B. Faux. Puisque n>30, $S^2 = s^2 = \frac{1}{n}.\sum n_i {x_i}^2 - \bar{x}^2$=21,5kg² (C'est l'unité qui est fausse)
C. **Vrai**. car n>30 donc S²=s².
D. Faux. Cf réponse E
E. **Vrai**. $\mu\ \epsilon\ [\bar{x} - c_{\alpha/2}\frac{S}{\sqrt{n}};\ \bar{x} + c_{\alpha/2}\frac{S}{\sqrt{n}}]$ on trouve c=1.96 dans la table de la loi normale d'où [30,26 ; 31,31]

QCM n°6 : **A, D, E**

A. **Vrai** $\bar{x} = \frac{49{,}5}{25} = 1{,}98\ kg$

$$s^2 = \frac{1}{n}.\sum n_i {x_i}^2 - \bar{x}^2 = \frac{98{,}3}{25} - 1{,}98^2 = 0{,}0116\ kg^2$$

Il faut S² car production par échantillon $S^2 = \frac{25}{24}.0{,}0116 = 0{,}0121$

B. Faux $X \rightarrow N(\mu\,;\sigma)$ car n>30.

Il faut $S^2 = \frac{25}{24}.0{,}0116 = 0{,}0121$

$\sigma^2\ \epsilon\ [(n-1)\frac{S^2}{b}; (n_1)\frac{S^2}{a}]$ on cherche a et b dans table Chi 2 à 24 ddl .

a : ligne n-1 et colonne 1 - $\frac{\alpha}{2}$; a =15,659

b : ligne n-1 et colonne $\frac{\alpha}{2}$; b=33,196

donc $\sigma^2\ \epsilon\ [0{,}0087; 0{,}0185]$

C. Faux $\sigma\ \epsilon\ [0{,}093;\ 0{,}136]$
D. **Vrai**.$\mu\ \epsilon\ \left[\bar{x} - \frac{t_{n-1,\alpha}S}{\sqrt{n}}\ ;\ \bar{x} - \frac{t_{n-1,\alpha}S}{\sqrt{n}}\right] on\ cherche\ t\ dans\ student\ a\ ligne\ n-1 = 24\ et\ colonne\ \alpha = 0{,}05$

d'où $t_{n-1,\alpha} = 2.064$. On obtient $\mu\ \epsilon\ [1.9346; 2.0254] = [1.93; 2.03]$

E. **Vrai.** σ étant connue, $\mu\ \epsilon\ [\bar{x} - c_{\alpha/2}\frac{\sigma}{\sqrt{n}};\ \bar{x} + c_{\alpha/2}\frac{\sigma}{\sqrt{n}}]$=[1.9408 ;2.0192]=[1.94 ;2.02].

QCM n°7 : **A, B, E**

A. **Vrai** $p_0 = \frac{167}{220} = 0{,}7591 = 0.759$

B. **Vrai.** On a bien un contexte de loi binomiale avec n=220>30, np=167>5 et nq=53>5 donc $\pi \in \left[p_0 - c_{\frac{\alpha}{2}}\sqrt{\frac{p_0(1-p_0)}{n}}\ ;\ p_0 + c_{\frac{\alpha}{2}}\sqrt{\frac{p_0(1-p_0)}{n}}\right]$ = [0,7591$\mp$1.96$\sqrt{\frac{0,7591.0,2409}{220}}$] $= [0.7026; 0.8156] = [0,70; 0,82]$.

C. Faux. n>30, on peut donc appliquer le théorème central limite. Donc X suit une loi normale de paramètre $N\left(\mu; \frac{\sigma}{\sqrt{n}}\right)$ que l'on approche par $N\left(\bar{x}; \frac{S}{\sqrt{n}}\right)$ avec $\bar{x} = \sum \frac{x_i}{n} = \frac{420}{167} = 2,5150$
et comme n>30 $S^2 = s^2 = \sum \frac{{x_i}^2}{n} - \bar{x}^2 = 2,0580$ d'où $s = \sqrt{s^2} = 1,4346$ donc $\frac{s}{\sqrt{n}} = \frac{1,4346}{\sqrt{167}}$=0,1110.

D. Faux $\mu \in \left[\bar{x} - c_{\alpha/2}\frac{S}{\sqrt{n}};\ \bar{x} + c_{\alpha/2}\frac{S}{\sqrt{n}}\right]$ = [2.3324 ;2.6976]=[2,33 ; 2,70] (on cherche $c_{\alpha/2}$ dans la table pour α=0.10 et il vaut 1.645.

E. **Vrai** Cf réponse D

F. Faux.

QCM n°8 : **B, D, E**

A. Faux. Voir items suivants

B. **Vrai.** Loi binomiale car 2 réponses possibles : «Avoir ou pas des vomissements ». De plus approximation possible par la loi Normale ($n > 30$, $nq > 5$,$np > 5$). Avec $C\alpha/2$: fractile $\alpha/2$ de la loi Normale ; pour α= 5%, $C\alpha/2$= 1.96. Donc $\pi \in \left[p_0 - c_{\frac{\alpha}{2}}\sqrt{\frac{p_0(1-p_0)}{n}}\ ;\ p_0 + c_{\frac{\alpha}{2}}\sqrt{\frac{p_0(1-p_0)}{n}}\right]$ = [0.8 $\pm$ 1.96 *$\sqrt{\frac{0.8*0.2}{70}}$]= [0.7063 ; 0.8937]

C. Faux. Voir B

D. **Vrai.** L'amplitude de l'intervalle diminue quand n augmente, donc plus l'intervalle est petit et plus l'estimation est précise

E. **Vrai.** C'est la définition d'une estimation par intervalle, on a 95% de chances de trouver la valeur de p dans la population dans cet intervalle.

F. Faux.

QCM n°9 : **A, C, D, E**

A. **Vrai** La moyenne par le calcul nous donne $\bar{x} = \sum \frac{x_i}{n}$ = 0.8846.

La Variance est trouvée par la formule suivante : Var(x) = $\sum \frac{xi^2}{n} - \bar{x}^2$ = 0.0198.

Donc l'écart-type vaut : 0.1407.

B. Faux. Cf. ci-dessus

C. **Vraie**. On a besoin de supposer car $n<30$

On va calculer l'intervalle de confiance au seuil de 95%.

On lit dans la table de student à n-1 ddl la valeur de $t_{n-1,\ \alpha/2}$ pour α =5% et on trouve 2.179.

L'intervalle est définit par la formule suivante : [$\bar{x} \mp t_{n-1,\alpha/2}\ \frac{S}{\sqrt{n}}$]. Dans la table, on trouve $t_{n-1,\alpha/2}$ =2.179

On a $S = \sqrt{\frac{n}{n-1} s^2}$= 0.1465 ET n= 13.

Par remplacement on trouve : [0.796 ; 0.973]

D. **Vrai**. Cf. Ci-dessus

E. **Vrai**. Par raisonnement analogue (avec $t_{n-1,\alpha/2}$ =2.681)on trouve : [0.776;0.994]

f) Faux.

QCM n°10 : **B, C.**

A. Faux, cf. item B

B. **Vrai**, loi binomiale soit n=8 et p=$\frac{3}{8} = 0.375$

C. **Vrai**, X suit une loi B(8 ; 0.375) on cherche kmin tel que P(X≤kmin)= 0.025 (ou α/2).
P(X=0)=$C_8^0 0.375^0 0.625^8 = 0.0233$
P(X=1)=0.1118 →P(X$\leq 1) = 0.1351$
Donc kmin$\epsilon[0;1]$
Interpolation linéaire : $kmin = 0 + \frac{0.025-0.0233}{0.1351-0.0233} = 0.0152$
On cherche kmax tel que P(X≤kmax)= 0.975 (ou 1-α/2).
P(X=8)=0.0004
P(X=7)=0.0052 → P(X<7)=1-0.0056=0.9944
P(X=6)=0.0304 →P(X<6)=1-0.036=0.964
Donc kmax$\epsilon[6;7]$
Interpolation linéaire : $kmax = 6 + \frac{0.975-0.964}{0.9944-0.964} = 6.3618$
Soit l'intervalle de confiance est $\left[\frac{0.0152}{8};\frac{6.3618}{8}\right] = [0.0019;0.7952]$

D. Faux, cf item C.

E. Faux, cf. item C.

QCM n°11 : **B, C.**

A. Faux c'est une loi binomiale de paramètres, n=210 et p=45/210= 0.2143.

B. **Vrai**, soit np= 210x0.214 = 44.5> 5 et npq= 210x0.214x0.786= 35.32>5. Remarque : il faut aussi n≥ 30

C. **Vrai**, $\left[0.2143 - 1.96\sqrt{\frac{0.2143(1-0.2143)}{210}};0.2143 + 1.96\sqrt{\frac{0.2143(1-0.2143)}{210}}\right]$=$[0.1588;0.2698] = [0.159;0.270]$.

D. Faux, item C.

E. Faux, item A.