Correction du CC 2023/2024
Proba - VA - Intervalles de confiance

Septembre 2024

QCM 9

Le volume expiratoire maximal par seconde (VEMS) correspond au volume d'air expiré pendant la première seconde d'une expiration forcée suite à une inspiration profonde. À partir d'un échantillon aléatoire de 250 hommes âgés de 20 à 29 ans, on a pu estimer le VEMS moyen à 3,4 L et son écart-type à 0,6 L. Vous ne ferez aucun arrondi dans les calculs intermédiaires. Les résultats finaux seront donnés avec 2 chiffres après la virgule. Parmi les propositions ci-dessous, cochez celle(s) qui est/sont vraie(s)

A. la borne sup. de l'intervalle de confiance à la confiance 0,95 du VEMS moyen dans la population des hommes âgés de 20 à 29 ans vaut 3,47 L

B. la borne inf. de l'intervalle de confiance à la confiance 0,968 du VEMS moyen dans la population des hommes âgés de 10 à 29 ans vaut 3,31 L

C. la borne inférieure de l'intervalle de confiance à la confiance 0,95 du VEMS moyen dans l'échantillon des hommes âgés de 10 à 29 ans vaut 3,32 L

D. la borne supérieure de l'intervalle de confiance à la confiance 0,90 du VEMS moyen l'échantillon des hommes âgés de 10 à 29 ans vaut 3,46 L

E. toutes choses égales par ailleurs, pour diminuer la largeur d'un intervalle de confiance, on peut augmenter la taille de l'échantillon étudié

QCM 9 - Réponses

Item A : bs_1=borne sup $ic_{0,95}(\mu)$ (VEMS moyen dans la population)

Intervalle de confiance d'une moyenne théorique, $n = 250$:
$ic_{0,95}(\mu) = m \pm z_{1-\frac{\alpha}{2}} \times \frac{s}{\sqrt{n}}$
$m = 3{,}4 \quad s = 0{,}6 \quad z_{1-\frac{0{,}05}{2}} = 1{,}96$
Application numérique : $bs_1 = 3.474376...$
On majore la borne supérieure $\rightarrow bs_1 = 3{,}48$
Item A : Faux

Item B : bi_1=borne inf $ic_{0,968}(\mu)$ (VEMS moyen dans la population)

Même formule que dans l'item A mais $\alpha = 1 - 0{,}968 = 0{,}032$.
On a donc $z_{1-\frac{\alpha}{2}} = z_{1-\frac{0{,}032}{2}} = z_{0{,}984} = 2{,}1444$ (lecture dans la table 2)
Application numérique : $bi_1 = 3.318625...$
On minore la borne inférieure $\rightarrow bi_1 = 3{,}31$
Item B : Vrai

QCM 9 - Réponses

Item C : bi_2= borne inf $ic_{0,95}(m)$ (VEMS moyen dans l'échantillon)

On calcule un intervalle de confiance d'une valeur théorique (=de la population), inconnue. Ici, m est connue, cela n'a pas de sens de chercher un intervalle de confiance de m

Item C : Faux

Item D : bs_2c= borne sup $ic_{0,90}(m)$ (VEMS moyen dans l'échantillon)

Même justification que pour l'item C

Item D : Faux

toutes choses égales par ailleurs, pour diminuer la largeur d'un intervalle de confiance, on peut augmenter la taille de l'échantillon étudié

cf cours

Item E : Vrai

QCM 10 - Énoncé

Le BNP (Brain Natriuretic Peptide) est un utilisé comme biomarqueur de l'insuffisance cardiaque. On définit la variable aléatoire X modélisant la concentration en BNP dans une population d'hommes âgés a priori sain. On considère que cette variable X suit une loi de Gauss d'espérance μ_X=80ng/mL et d'écart-type σ_X=30 ng/mL. On considère qu'un patient ayant une concentration de BNP supérieure à 100 a un risque d'être atteint d'insuffisance cardiaque. On s'intéresse à un échantillon aléatoire simple constitué de 50 hommes âgés a priori sain. On définit les variables aléatoires suivantes :

- Y : « nombre de patients de l'échantillon ayant un risque d'être atteint d'insuffisance cardiaque »
- Z : « proportion de patients ayant un risque d'être atteint d'insuffisance cardiaque dans l'échantillon »

Les résultats des calculs seront donnés avec 2 chiffres après la virgule.

Parmi les propositions ci-dessous, cochez celle(s) qui est/sont vraie(s)

QCM 10 - Analyse de l'énoncé

Variables de l'énoncé

- X : [BNP], X$\rightarrow \mathcal{N}(\mu_X = 80; \sigma_X = 30)$
  Patient "à risque d'insuffisance cardiaque" si [BNP] $>$ 100 ng/mL
  On note U la variable centrée réduite associée à X (Z est déjà utilisé dans l'exercice)
- T : "Statut du patient vis à vis de l'insuffisance cardiaque"
  - T$=$0 si le patient n'est pas à risque
  - T$=$1 si le patient est à risque

  $\mathrm{P}(T = 1) = \mathrm{P}(X > 100)$
  On centre et on réduit :
  $\mathrm{P}(X > 100) = \mathrm{P}\left(U > \frac{100-80}{30}\right) = \mathrm{P}(U > 0{,}67) = 1 - \mathrm{P}(U \leq 0{,}67)$
  On lit dans la table 1 : $\mathrm{P}(U \leq 0{,}67) = 0{,}7486$
  Donc $\mathrm{P}(T = 1) = 0{,}2514 \simeq 0{,}25$ (arrondi avec 2 chiffres après la virgule)
  On a donc T$\rightarrow$ Bern(p$=$0,25)

QCM 10 - Analyse de l'énoncé

Variables de l'énoncé (suite)

- Y : « nb patients de l'échantillon à risque d'insuffisance cardiaque »
  $Y = \sum_{i=1}^{50} Ti$, où les T_i sont toutes indépendantes et de même loi de Bernoulli de paramètre p=0,25
  Y $\to \mathcal{B}(n = 50, p = 0{,}25)$
  Peut-on approximer le loi de Y par une loi normale ?
  - $n = 50$, donc $n \geq 50$
  - $n \times p = 50 \times 0{,}25 = 12{,}5$ donc $n \times p \geq 5$
  - $n \times (1 - p) = 50 \times 0{,}75 = 37{,}5$ donc $n \times (1 - p) \geq 5$

  Donc $Y \rightsquigarrow \mathcal{N}(\mu_Y = 12{,}5; \sigma_Y = \sqrt{9{,}375})$
- Z : « prop de patients de l'échantillon à risque d'insuffisance cardiaque »
  $Z = \frac{Y}{n}$ Comme Y suit approximativement une loi normale, Z suit aussi approximativement une loi normale.
  $Z \to \mathcal{N}(\mu_Z = 0{,}25; \sigma_Z = \sqrt{0{,}00375} \simeq 0{,}06)$

QCM 10 - Réponses

Item A : La borne supérieure de l'intervalle de fluctuation, à la confiance 0,95, de la concentration moyenne en BNP vaut 88,32

Ici, on travaille sur la concentration moyenne, c'est à dire la va M, où $M = \sum_{i=1}^{50} \frac{X_i}{50}$ et où $X_i \to \mathcal{N}(80; 30)$.
M suit donc une loi normale de paramètres $\mu_m = 80$ et $\sigma_M = \frac{30}{\sqrt{50}} = 4{,}24$

$$IF_{0,95}(X) = \mu \pm z_{1-\frac{\alpha}{2}} \times \sigma$$

La borne supérieure *bs* se calcule donc ainsi :
$bs = \mu + z_{1-\frac{\alpha}{2}} \times \sigma = 80 + 1{,}96 \times 4{,}24 = 88{,}3104$
Pour l'arrondi, on majore la borne supérieure et on obtient $bs = 88{,}32$
Item A : Vrai

QCM 10 - Réponses

Item B : La probabilité qu'un individu de l'échantillon ait un risque d'être atteint d'insuffisance cardiaque est inférieure à 2×10^{-5}

On travaille ici sur la va X. Cette question revient à calculer $\mathrm{P}(X > 100)$ (cf diapo 6)

On avait trouvé $\mathrm{P}(X > 100) = 0{,}25$

Item B : Faux

Item C : la variable aléatoire Y suit une loi de Bernoulli de paramètre p environ égal à 0,25

Dans la diapo 7, nous avons montré que Y suit une loi binomiale de paramètres n=50 et p=0,25

Item C : Faux

QCM 10 - Réponses

Item D : la variable aléatoire Z suit approximativement une loi normale d'espérance $\mu_Z \simeq 0{,}25$ et d'écart-type $\sigma_Z \simeq 0{,}06$

Cet item est vrai, voir les explications sur la diapo 7

Item D : Vrai

Item E : la probabilité d'avoir exactement un patient à risque d'insuffisance cardiaque dans l'échantillon vaut environ $50 \times 0{,}25 \times 0{,}75^{49}$

Ici, on travaille sur la variable Y qui modélise le nombre de patients à risque d'insuffisance cardiaque dans un échantillon de 50 hommes

On cherche $\mathrm{P}(Y = 1)$ avec $Y \to \mathcal{B}(n = 50, p = 0{,}25)$

$\mathrm{P}(Y = 1) = \mathrm{C}_{50}^{1}(0{,}25)^1 \times (0{,}75)^{49}$

Or,

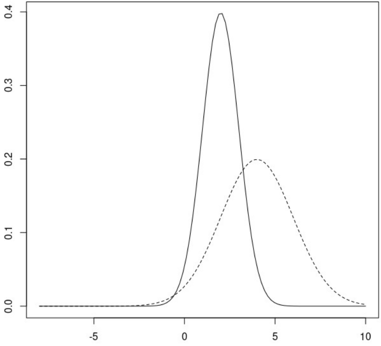
$$\mathrm{C}_{50}^{1} = \frac{50!}{1! \times 49!} = \frac{50 \times 49 \times 48 \times \ldots \times 1}{1 \times 49 \times 48 \times \ldots \times 1} = 50$$

Donc $\mathrm{P}(Y = 1) = 50 \times 0{,}25 \times 0{,}75^{49}$

Item E : Vrai

QCM 11 - Enoncé et réponses

Soient U et V, 2 variables aléatoires Gaussiennes dont les densités de probabilité sont représentées sur la figure ci-dessous, en traits pleins pour V et en pointillés pour U. Parmi les propositions suivantes relatives aux variables aléatoires et aux intervalles de confiances, cochez celle(s) qui est/sont vraie(s).



Item A : on voit sur la figure que $E(U) > E(V)$

Pour une variable Gaussienne, l'espérance est égale à la médiane et et égale au mode

Item A : Vrai

Item B : on voit sur la figure que $var(U) > var(V)$

Item B : Vrai

QCM 11 - Suite

Item C : les bornes de l'intervalle de confiance dépendent de l'échantillon considéré

Item C : Vrai (cf cours)

Item D : un intervalle de confiance est forcément bilatéral

Un intervalle de confiance peut-être unilatéral ou bilatéral

Item D : Faux

Item E : soit T l'estimateur de θ, le biais de T se calcule ainsi : $E(T) - \theta$

Il s'agit bien de la définition donnée en cours

Item E : Vrai

QCM 12 - Énoncé

Un médicament pouvant être pris par les enfants, les adultes ou les seniors, existe sous 2 formes : poudre à diluer ou comprimé à avaler. Parmi les patients qui prennent ce médicament, 20% sont des enfants et 50% sont des adultes de moins de 65 ans. 90% des enfants qui prennent ce médicament utilisent la forme « poudre à diluer » alors que 80% des adultes de moins de 65 ans préfèrent la forme « comprimé ». On sait aussi que 15% des personnes prenant le médicament sont des seniors utilisant la forme « comprimé ».
Parmi les propositions ci-dessous, cochez celle(s) qui est/sont vraie(s) :

A. Parmi les individus prenant le médicament, 40% sont des adultes ayant choisi la forme « comprimé »

B. Parmi les individus prenant le médicament, 10% sont des enfants ayant choisi la forme « comprimé »

C. Parmi les individus prenant le médicament, 57% ont choisi la forme « comprimé »

D. On choisit aléatoirement un individu ayant choisi la forme « comprimé », la probabilité qu'il s'agisse d'un adulte de moins de 65 ans vaut environ 0,7

E. La moitié des seniors ont choisi la forme « comprimé »

QCM 12 - Informations de l'énoncé

Énoncé

Un médicament pouvant être pris par les enfants, les adultes ou les seniors, existe sous 2 formes : poudre à diluer ou comprimé à avaler. Parmi les patients qui prennent ce médicament, 20% sont des enfants et 50% sont des adultes de moins de 65 ans. 90% des enfants qui prennent ce médicament utilisent la forme « poudre à diluer » alors que 80% des adultes de moins de 65 ans préfèrent la forme « comprimé ». On sait aussi que 15% des personnes prenant le médicament sont des seniors utilisant la forme « comprimé ».

Infos de l'énoncé :

On définit les évènements suivants :

C : « prendre la forme comprimé » et $\bar{C}$: « prendre la forme poudre »

E : « être un enfant », A : « être un adulte » et S : « être un sénior »

$\mathrm{P}(E) = 0{,}2 \quad \mathrm{P}(A) = 0{,}5 \quad \mathrm{P}\left(\bar{C}|E\right) = 0{,}9 \quad \mathrm{P}(C|A) = 0{,}8$

et $\mathrm{P}(S \cap C) = 0{,}15$

QCM 12 - Correction

Item A : Parmi les individus prenant le médicament, 40% sont des adultes ayant choisi la forme comprimé

On cherche $\mathrm{P}(A \cap C)$. Pour cela, on utilise la formule des probabilités composées :

$\mathrm{P}(A \cap C) = \mathrm{P}(C|A) \times \mathrm{P}(A)$

$\mathrm{P}(A \cap C) = 0{,}8 \times 0{,}5 = 0{,}4$

Item A : Vrai

Item B : Parmi les individus prenant le médicament, 10% sont des enfants ayant choisi la forme comprimé

On cherche $\mathrm{P}(E \cap C)$. On utilise la formule des probabilités composées :

$\mathrm{P}(E \cap C) = \mathrm{P}(C|E) \times \mathrm{P}(E)$

$\mathrm{P}(E \cap C) = (1 - \mathrm{P}(\bar{C}|E)) \times \mathrm{P}(E)$

$\mathrm{P}(E \cap C) = 0{,}1 \times 0{,}2 = 0{,}02$

Item B : Faux

QCM 12 - Correction

Item C - Parmi les individus prenant le médicament, 57% ont choisi la forme « comprimé »

On cherche $\mathrm{P}(C)$. Pour cela, on va utiliser la formule des probabilité totales avec le système complet d'évènements $\{A, E, S\}$.

$\mathrm{P}(C) = \mathrm{P}(C|E) \times \mathrm{P}(E) + \mathrm{P}(C|A) \times \mathrm{P}(A) + \mathrm{P}(C|S) \times \mathrm{P}(S)$

On écrit l'égalité en utilisant les termes connus dans l'énoncé :

$\mathrm{P}(C) = (1 - \mathrm{P}(\bar{C}|E)) \times \mathrm{P}(E) + \mathrm{P}(C|A) \times \mathrm{P}(A) + \mathrm{P}(S \cap C)$

$\mathrm{P}(C) = 0{,}1 \times 0{,}2 + 0{,}8 \times 0{,}5 + 0{,}15 = 0{,}57$

Item C : Vrai

QCM 12 - Correction

Item D : on choisit aléatoirement un individu ayant choisi la forme « comprimé », la probabilité qu'il s'agisse d'un adulte de moins de 65 ans vaut environ 0,7

On cherche $\mathrm{P}(A|C)$. On utilise la définition de la probabilité conditionnelle :

$$\mathrm{P}(A|C) = \frac{\mathrm{P}(A \cap C)}{\mathrm{P}(C)} = \frac{\mathrm{P}(C|A) \times \mathrm{P}(A)}{\mathrm{P}(C)}$$

$$\mathrm{P}(A|C) = \frac{0{,}8 \times 0{,}5}{0{,}57} \simeq 0{,}7$$

Item D : Vrai

Remarque : Cela revient à utiliser la formule de Bayes dont le dénominateur a été calculé lors de l'item C

QCM 12 - Correction

Item E : la moitié des seniors ont choisi la forme « comprimé »

On cherche $\mathrm{P}(C|S)$. On utilise la définition de la probabilité conditionnelle :

$$\mathrm{P}(C|S) = \frac{\mathrm{P}(C \cap S)}{\mathrm{P}(S)}$$

Or, $\mathrm{P}(S) = 1 - \mathrm{P}(A) - \mathrm{P}(E) = 1 - 0{,}5 - 0{,}2 = 0{,}3$
D'où :

$$\mathrm{P}(C|S) = \frac{0{,}15}{0{,}3} = 0{,}5$$

Item E : Vrai

Remarque : pour résoudre cet exercice, on peut tout à fait faire un arbre en mettant au 1er niveau $\{A, E, S\}$ et au second niveau $\{C, \bar{C}\}$

Correction du CT 2023/2024
Proba - VA - Intervalles de confiance

Énoncé et définition des événements

Énoncé

On réalise une étude sur une pathologie P pour laquelle il existe différents niveaux de gravité allant de G1 (le moins grave) à G4 (le plus grave). Dans une population de patients atteints de la pathologie P, on a constaté que 40% des individus sont au niveau de gravité G1, 30% au niveau G2, 20% au niveau G3 et 10% au niveau G4. Des études précédentes ont montré que, tout niveaux de gravité confondus, la survie des patients 3 ans après le diagnostic est de 65%. On sait par ailleurs que parmi les personnes atteintes de la pathologie P de niveau de gravité G3, la probabilité de survie 3 ans après le diagnostic vaut 50% et que la probabilité qu'un patient atteint de la pathologie P soit vivant 3 ans après le diagnostic et de niveau de gravité G4 vaut 0,02.

Définition des évènements et infos de l'énoncé

G_i : « être atteint au niveau de gravité G_i » et S : « être vivant 3 ans après le diagnostic »

$\mathrm{P}(G_1) = 0{,}4 \quad \mathrm{P}(G_2) = 0{,}3 \quad \mathrm{P}(G_3) = 0{,}2 \quad \mathrm{P}(G_4) = 0{,}1 \quad \mathrm{P}(S) = 0{,}65$

$\mathrm{P}(S|G_3) = 0{,}5 \quad \mathrm{P}(S \cap G_4) = 0{,}02$

Correction

Item A : 80% des personnes atteintes de la pathologie P au niveau G4 décèdent dans les 3 ans suivant le diagnostic

On cherche $\mathrm{P}\left(\bar{S}|G_4\right)$. On passe par l'évènement complémentaire.
$\mathrm{P}\left(\bar{S}|G_4\right) = 1 - \mathrm{P}(S|G_4)$
Pour calculer $\mathrm{P}(S|G_4)$, on utilise la définition d'une probabilité conditionnelle :

$$\mathrm{P}(S|G_4) = \frac{\mathrm{P}(S \cap G_4)}{\mathrm{P}(G4)} = \frac{0{,}02}{0{,}1} = 0{,}2$$

D'où, $\mathrm{P}\left(\bar{S}|G_4\right) = 1 - 0{,}2 = 0{,}8$
Item A : Vrai

Correction

Item B : si on considère un patient décédé moins de 3 ans après le diagnostic, la probabilité qu'il soit de niveau de gravité G_4 vaut 0,8

On cherche $\mathrm{P}\left(G_4|\bar{S}\right)$. On utilise la définition d'une probabilité conditionnelle :

$$\mathrm{P}\left(G_4|\bar{S}\right) = \frac{\mathrm{P}\left(G_4 \cap \bar{S}\right)}{\mathrm{P}\left(\bar{S}\right)} = \frac{\mathrm{P}\left(\bar{S}|G_4\right) \times \mathrm{P}(G_4)}{\mathrm{P}\left(\bar{S}\right)}$$

$$\mathrm{P}\left(G_4|\bar{S}\right) = \frac{0{,}8 \times 0{,}1}{0{,}35} \simeq 0{,}23$$

Item B : Faux

Item C : 10% des personnes atteintes de la pathologie P sont de niveau de gravité G_3 et sont vivants 3 ans après le diagnostic

On cherche $\mathrm{P}(G_3 \cap S)$. On utilise la formule des probabilités composées :
$\mathrm{P}(G_3 \cap S) = \mathrm{P}(S|G_3) \times \mathrm{P}(G_3) = 0{,}5 \times 0{,}2 = 0{,}1$
Item C : Vrai

Correction

Item D : la probabilité qu'une personne atteinte de la maladie P soit de niveau de gravité G_3 ou G_4 et encore vivant 3 ans après le diagnostic vaut 0,12

On cherche $P((G_3 \cup G_4) \cap S)$.

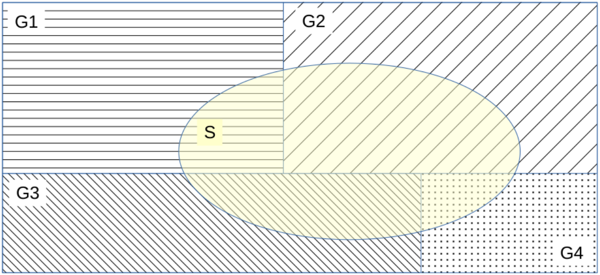
$P((G_3 \cup G_4) \cap S) = P((G_3 \cap S) \cup (G_4 \cap S))$

Les évènements G_3 et G_4 sont incompatibles donc

$P((G_3 \cup G_4) \cap S) = P(G_3 \cap S) + P(G_4 \cap S)$

$P((G_3 \cup G_4) \cap S) = 0{,}1 + 0{,}02 = 0{,}12$ (cf Item C et énoncé)

Item D : Vrai



Correction

Item E : 40% des personnes atteintes de la pathologie P de niveau de gravité G3 ou G4 sont encore vivantes 3 ans après le diagnostic

On cherche $\mathrm{P}(S|(G_3 \cup G_4))$. On utilise la définition de la probabilité conditionnelle :

$$\mathrm{P}(S|(G_3 \cup G_4)) = \frac{\mathrm{P}(S \cap (G_3 \cup G_4))}{\mathrm{P}(G_3 \cup G_4)}$$

G_3 et G_4 étant incompatibles,
$\mathrm{P}(G_3 \cup G_4) = \mathrm{P}(G_3) + \mathrm{P}(G_4) = 0{,}2 + 0{,}1 = 0{,}3$
$\mathrm{P}(S \cap (G_3 \cup G_4)) = 0{,}12$ (cf item D)
Donc $\mathrm{P}(S|(G_3 \cup G_4)) = \frac{0{,}12}{0{,}3} = 0{,}4$
Item E : Vrai