Annale 2024/2025
Probabilités, variables aléatoires, intervalle de confiance

C. Bardel

CC, QCM 3

La mucoviscidose est une maladie génétique liée à une anomalie du gène codant pour la protéine CFTR. Afin d'étudier les mutations responsables de la maladie, le gène est séquencé dans un échantillon de 200 individus atteints de mucoviscidose. Parmi ces 200 individus, on en dénombre 140 qui sont porteurs de la mutation ΔF508. On souhaite calculer un intervalle de confiance de la proportion de porteurs de mutation ΔF508 dans la population des patients atteints de mucoviscidose (appelés dans la suite de l'exercice, les malades). Dans cet exercice, les notations utilisées sont celles vues dans le cours sur les estimations et intervalles de confiance. Indiquez la ou les réponse(s) juste(s)

A. la borne sup. de l'intervalle de confiance à la confiance 0,95 de la proportion de porteurs de ΔF508 dans la population de malades vaut 0,76
B. la borne inf. de l'intervalle de confiance à la confiance 0,95 de la proportion de porteurs de ΔF508 dans l'échantillon de malades étudié vaut 0,63
C. la borne inf. de l'intervalle de confiance à la confiance 0,87 de la proportion de porteurs de ΔF508 dans la population de malades vaut 0,65
D. les conditions de validité de cet intervalle de confiance sont : $n \geq 30$, $n \times f \geq 5$, $n \times (1 - f) \geq 5$, où f est l'estimation de la proportion de malades porteurs de la mutation ΔF508
E. si on augmente la taille de l'échantillon étudié, on diminuera la largeur de l'intervalle de confiance

CC, QCM 4

Une étude est réalisée au sein de la population âgée (plus de 80 ans). Dans cette population, 22,4% des individus sont atteints de la maladie d'Alzheimer et 24% sont porteurs d'au moins un allèle ϵ_4 du gène *ApoE*. Parmi les porteurs de l'allèle ϵ_4, on a constaté que 30% sont atteints de la maladie d'Alzheimer.
Indiquez la ou les réponses juste(s)

A. la probabilité de ne pas être atteint de la maladie d'Alzheimer et porteur de l'allèle ϵ_4 vaut 0,072
B. parmi les non porteurs de l'allèle ϵ_4, la probabilité d'être atteint de la maladie d'Alzheimer vaut 0,2
C. parmi les individus qui ne sont pas atteints de la maladie d'Alzheimer, la probabilité d'être porteur de l'allèle ϵ_4 vaut environ 0,09
D. la probabilité de ne pas être porteur de l'allèle ϵ_4 ou de ne pas être atteint de la maladie d'Alzheimer vaut 0,928
E. les évènements « être atteint de la maladie d'Alzheimer » et « être porteur de l'allèle ϵ_4 » ne sont pas indépendants

CC, QCM 5

Dans un pays, durant la saison estivale, on considère que la probabilité qu'il pleuve est identique chaque jour et vaut 0,8. Durant la saison hivernale, on considère que la probabilité qu'il pleuve est également identique chaque jour et vaut 0,3. On note X_E et X_H les variables aléatoires modélisant le nombre de jour de pluie pendant les saisons estivales et hivernales respectivement. On considère que les saisons estivales et hivernales durent toutes les deux 90 jours et qu'elles sont indépendantes. On définit la variable aléatoire D comme la différence $(X_H - X_E)$

A. la variable aléatoire X_E suit une loi de Bernoulli de paramètre p=0,8

B. en moyenne, sur l'ensemble des 2 saisons, on s'attend à avoir 99 jours de pluie

C. la variance de D est égale à 4,5

D. la fonction de répartition de X_H est une fonction F définie par $F(x) = P(XH \leq x)$, $\forall x \in \mathbb{R}$

E. on peut approximer la loi de X_H par une loi normale d'espérance égale à 0,8 et d'écart-type égal à $\sqrt{\frac{0{,}8 \times 0{,}2}{90}}$

CC, QCM 6

Soit T la variable aléatoire modélisant la concentration sanguine de la protéine A (en μg/mL). T suit approximativement une loi normale d'espérance $\mu_T = 3$ et d'écart-type $\sigma_T = 0{,}5$. Chez les individus atteints de la maladie M, la concentration sanguine moyenne de la protéine A est réduite et vaut 2, avec un écart-type toujours égal à 0,5. On considère un échantillon de 25 patients atteints de la maladie M, chez qui la concentration en protéine A est abaissée. On note :

- M l'évènement « le patient est malade »
- Z une variable aléatoire qui suit la loi normale centrée réduite
- Φ la fonction de répartition de la loi normale centrée réduite

Indiquez la ou les réponses juste(s)

A. $P(2 \leq T \leq 4) = P(2 \leq T) + P(T \leq 4)$

B. dans la population générale, l'intervalle de fluctuation à 0,95 de la concentration sanguine de la protéine A vaut [2,02 ; 3,98]

C. dans l'échantillon de 25 patients malades, la probabilité que la concentration sanguine en protéine A moyenne soit supérieure à 1,5 μg/mL est égale à $P(Z \leq 5)$

D. $P(4 \leq T \leq 5) = \Phi(5) - \Phi(4)$

E. $P(T \leq 1|M) = P(Z \leq -\frac{1}{5})$

CT, QCM 5

Une usine de production de comprimés comporte 2 lignes de production utilisant 2 technologies différentes. Chaque jour, la ligne 1 (L1) produit deux fois plus de comprimés que la ligne 2 (L2). On sait par ailleurs que les probabilités de produire des comprimés non conformes valent respectivement 0,03 et 0,04 sur les lignes L1 et L2.
On prélève aléatoirement un comprimé produit dans une journée.
Indiquez la ou les réponse(s) juste(s)

A. le comprimé a une chance sur 2 de provenir de la ligne L1

B. la probabilité que le comprimé provienne de la ligne L1 et soit non conforme vaut 0,015

C. la probabilité que le comprimé soit non conforme vaut environ 0,033

D. la probabilité que le comprimé provienne de la ligne L1 sachant qu'il est non conforme vaut environ 0,6

E. la variable aléatoire modélisant le nombre de comprimés non conformes dans un échantillon de 10 comprimés provenant de la ligne 1 suit une loi de Bernoulli de paramètre p=0,03

CT, QCM 6

Une étude a été réalisée pour estimer le temps que passent les lycéens français sur leur smartphone chaque jour. Un échantillon aléatoire de 100 lycéens, représentatif de la population des lycéens français, a été constitué. Cette étude a permis d'estimer qu'en moyenne, un lycéen passe 442 minutes sur son smartphone avec un écart-type estimé à 303 minutes.
On souhaite calculer un intervalle de confiance du temps moyen passé par un lycéen sur son smartphone. Indiquez la ou les réponse(s) juste(s).

A. les conditions d'application pour calculer cet intervalle sont remplies

B. il faut vérifier a posteriori que $n \geq 30$, $n \times f_1 \geq 5$, $n \times (1 - f_1) \geq 5$, $n \times f_2 \geq 5$ et $n \times (1 - f_2) \geq 5$ où f_1 et f_2 sont les bornes inférieures et supérieures de l'intervalle de confiance

C. la borne supérieure de l'unintervalle de confiance à la confiance 0,95 de l'estimation du temps moyen passé par un lycéen sur son smartphone est égale à 502 minutes

D. l'estimation du temps moyen d'utilisation du smartphone par les lycéens est une variable aléatoire quantitative

E. l'intervalle de confiance du temps moyen passé par un lycéen sur son smartphone contient de façon certaine la valeur de l'estimation de ce temps moyen