Asie 2014. Enseignement spécifique

EXERCICE 2 (6 points) (commun à tous les candidats)

Le taux d'hématocrite est le pourcentage du volume de globules rouges par rapport au volume total du sang. On note X la variable aléatoire donnant le taux d'hématocrite d'un adulte choisi au hasard dans la population française. On admet que cette variable suit une loi normale de moyenne $\mu = 45,5$ et d'écart-type σ.

Partie A

On note Z la variable aléatoire $Z = \dfrac{X - \mu}{\sigma} = \dfrac{X - 45,5}{\sigma}$.

1) a) Quelle est la loi de la variable aléatoire Z ?

b) Déterminer $P(X \leqslant \mu)$.

2) En prenant $\sigma = 3,8$, déterminer $P(37,9 \leqslant X \leqslant 53,1)$. Arrondir le résultat au centième.

Partie B

Une certaine maladie V est présente dans la population française avec la fréquence 1 %. On sait d'autre part que 30 % de la population française a plus de 50 ans, et que 90 % des porteurs de la maladie V dans la population française ont plus de 50 ans.

On choisit au hasard un individu dans la population française.

On note α l'unique réel tel que $P(X \leqslant \alpha) = 0,995$, où X est la variable aléatoire définie au début de l'exercice. On ne cherchera pas à calculer α.

On définit les évènements :

- M « l'individu est porteur de la maladie V » ;
- S « l'individu a plus de 50 ans » ;
- H « l'individu a un taux d'hématocrite supérieur à α ».

Ainsi $P(M) = 0,01$, $P_M(S) = 0,9$ et $P(H) = P(X > \alpha)$.

D'autre part, une étude statistique a révélé que 60 % des individus ayant un taux d'hématocrite supérieur à α sont porteurs de la maladie V.

1) a) Déterminer $P(M \cap S)$.

b) On choisit au hasard un individu ayant plus de 50 ans. Montrer que la probabilité qu'il soit porteur de la maladie V est égale à $0,03$.

2) a) Calculer la probabilité $P(H)$.

b) L'individu choisi au hasard a un taux d'hématocrite inférieur ou égal à α. Calculer la probabilité qu'il soit porteur de la maladie V. Arrondir au millième.

Partie C

Le but de cette partie est d'étudier l'influence d'un gène sur la maladie V.

1) Déterminer l'intervalle de fluctuation asymptotique au seuil de 95 % de la fréquence de la maladie V dans les échantillons de taille 1 000, prélevés au hasard et avec remise dans l'ensemble de la population française. On arrondira les bornes de l'intervalle au millième.

2) Dans un échantillon aléatoire de 1 000 personnes possédant le gène, on a trouvé 14 personnes porteuses de la maladie V.
Au regard de ce résultat, peut-on décider, au seuil de 95 %, que le gène a une influence sur la maladie ?

Asie 2014. Enseignement spécifique

EXERCICE 2 : corrigé

Partie A

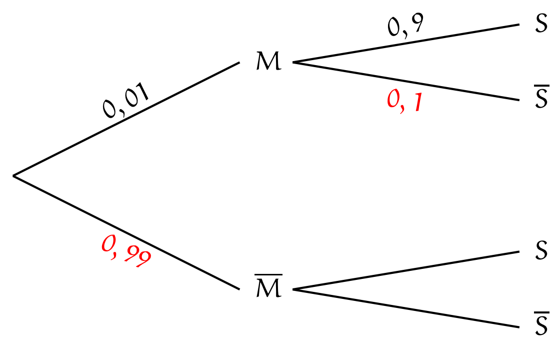
1) a) La variable aléatoire Z suit la loi normale centrée réduite c'est-à-dire la loi normale de moyenne 0 et d'écart-type 1.

b) On sait que $P(X \leqslant \mu) = \dfrac{1}{2}$.

2) La calculatrice fournit $P(37,9 \leqslant X \leqslant 53,1) = 0,95$ arrondi au centième.

Partie B

1) a) Représentons la situation par un arbre de probabilités.



$P(M \cap S) = P(M) \times P_M(S) = 0,01 \times 0,9 = 0,009$.

b) La probabilité demandée est $P_S(M)$. L'énoncé donne $p(S) = 0,3$ et d'autre part, d'après la question précédente, $P(M \cap S) = 0,009$.

$$P_S(M) = \frac{P(M \cap S)}{P(S)} = \frac{0,009}{0,3} = 0,03.$$

2) a) $P(H) = P(X > \alpha) = 1 - P(X \leqslant \alpha) = 1 - 0,995 = 0,005$.

b) La probabilité demandée est $P_{\overline{H}}(M)$. D'après la formule des probabilités totales,

$$P(M) = P(M \cap H) + P\left(M \cap \overline{H}\right) = P(H) \times P_H(M) + P\left(\overline{H}\right) \times P_{\overline{H}}(M)$$

et donc

$$0,01 = 0,005 \times 0,6 + 0,995 \times P_{\overline{H}}(M)$$

puis

$$P_{\overline{H}}(M) = \frac{0,01 - 0,005 \times 0,6}{0,995} = 0,007 \text{ arrondi au millième.}$$

Partie C

1) Ici, $n = 1\,000$ et $p = P(M) = 0,01$. On note que $n \geqslant 30$ puis $np = 10$ et donc $np \geqslant 5$ et $n(1-p) = 990$ et donc $n(1-p) \geqslant 5$.

Un intervalle de fluctuation asymptotique au seuil de $95\ \%$ de la fréquence de la maladie V dans les échantillons de taille $1\,000$ est

$$\left[p - 1,96\frac{\sqrt{p(1-p)}}{\sqrt{n}}, p + 1,96\frac{\sqrt{p(1-p)}}{\sqrt{n}}\right] = \left[0,01 - 1,96\frac{\sqrt{0,01 \times 0,99}}{\sqrt{1000}}, 0,01 + 1,96\frac{\sqrt{0,01 \times 0,99}}{\sqrt{1000}}\right].$$

En arrondissant au millième, on obtient l'intervalle $[0,004; 0,016]$.

2) La fréquence observée est $f = \dfrac{14}{1000} = 0,014$. Cette fréquence appartient à l'intervalle de fluctuation. Donc, on ne pas peut décider au seuil de $95\ \%$, que le gène a une influence sur la maladie.

