

ISIMA

Première année

Probabilités et statistiques

Documents et calculatrices autorisées

Des vers microscopiques, qui naissent d'œufs, sont, à la naissance, tous de même longueur, prise ici comme unité. Puis leur longueur croît. On a mesuré la longueur de 200 vers prélevés au hasard dans la population (ayant un effectif inconnu, mais gigantesque) avec une précision de 0,1. Voici les résultats obtenus :

centre de classe : x_i	1,05	1,15	1,25	1,35	1,45	1,55	1,65	1,75	1,85	1,95	2,05	2,15	2,25
effectif : n_i	42	37	29	21	11	11	4	11	7	3	2	5	1
centre de classe : x_i	2,35	2,45	2,55	2,65	2,75	2,85	2,95	3,05	3,25	3,45	3,55	3,75	4,95
effectif : n_i	2	1	1	1	1	2	1	2	1	1	1	1	1

1) Déterminer, sur les données ci-dessus, le mode (valeur du caractère d'effectif maximal), la médiane (valeur m du caractère telle qu'au moins 50% de l'effectif ait une valeur supérieure ou égale à m et 50% une valeur inférieure ou égale à m .) et les quartiles (définitions analogues à celle de la médiane, mais avec respectivement 25% -75% , 50%-50% et 75% -25%) de cette série.

2) On pense que la taille d'un ver peut être modélisée par une variable aléatoire de densité $f_\alpha(x) = \frac{\alpha}{x^{1+\alpha}}$ pour $x \geq 1$, où α est un paramètre strictement positif à déterminer. Montrer que, pour $\alpha > 0$ fixé, f_α est bien une densité de probabilité sur $[1; +\infty[$.

3) Montrer que, pour des variables aléatoires indépendantes $(X_i)_{1 \leq i \leq n}$ représentant des tailles de vers et suivant toutes une loi de densité f_α , l'estimateur du maximum de vraisemblance de α vaut : $\frac{1}{\bar{Y}}$ avec $\bar{Y} = \frac{1}{n} \cdot \sum_{i=1}^n \ln(X_i)$.

4) Montrer que, si X suit une loi de densité f_α , alors $\mathbb{E}(\bar{Y}) = \frac{1}{\alpha}$.

5) Expliquer pourquoi, lorsque les x_i sont les réalisations de 200 variables aléatoires indépendantes X_i représentant chacune la longueur d'un ver tiré au hasard, on peut considérer que la variable aléatoire \bar{Y} suit une loi normale.

6) En posant, dans la série précédente : $y_i = \ln(x_i)$, on constate que la moyenne des y_i vaut environ $\bar{y} \simeq 0,335$ et que leur écart-type vaut environ $s \simeq 0,301$. Déterminer un intervalle de confiance au taux de 95% pour $\frac{1}{\alpha}$. En déduire un intervalle de confiance au même taux pour α .

7) Montrer que la fonction de répartition $F_\alpha(x) = \mathbb{P}\{X \leq x\}$ d'une variable aléatoire X de densité $f_\alpha(x) = \frac{\alpha}{x^{1+\alpha}}$ pour $x \geq 1$ vérifie : $F_\alpha(x) = 1 - \frac{1}{x^\alpha}$ pour $x \geq 1$.

8) On veut s'assurer que le modèle consistant à attribuer à la longueur X d'un ver la densité f_3 est réaliste. Pour cela, effectuer un test d'ajustement, au seuil de risque de 5%, pour déterminer si la série considérée est ou non issue d'une population suivant une loi de densité $f_3(x) = \frac{3}{x^4}$ (pour $x \geq 1$).

9) On a étudié également la longueur de 120 vers d'une espèce proche. On a trouvé, pour la série des logarithmes des longueurs, une moyenne d'environ 0,290 pour un écart-type d'environ 0,271. Peut-on admettre, au seuil de risque de 5%, que les coefficients α des deux populations de vers sont les mêmes ?