Partiel de Statistique

Mardi 25 Janvier 2000, 10h00-12h00

Des études préalables ont permis de considérer que le pourcentage de fer contenu dans des éprouvettes de type E était une variable aléatoire X de densité de probabilité :

$$f_\theta(x) = \frac{1}{\theta} x^{\frac{1}{\theta}-1} I_{[0,1]}(x)$$

où $I_{[0,1]}(x)$ est la fonction indicatrice de l'intervalle $[0,1]$ et θ est un paramètre inconnu > 0. On prélève au hasard $n = 30$ éprouvettes de type E et on observe le pourcentage de fer x_i contenu dans chaque éprouvette i :

0.90	0.13	0.78	0.23	0.72	0.75	0.37	0.78	0.47	0.70
0.53	0.82	0.95	0.57	0.92	0.89	0.69	0.92	0.91	0.80
0.76	0.66	0.32	0.84	0.09	0.26	0.70	0.43	0.76	0.65

Pour les diverses applications numériques, on pourra utiliser les résultats suivants :

$$\sum_{i=1}^{30} \ln x_i \approx -16.92 \text{ et } \sum_{i=1}^{30} x_i \approx 19.30$$

$\boxed{1^{ère}\ \textbf{Partie : Estimation}}$

1) Déterminer la loi de la variable aléatoire $Y = -\ln X$ et en déduire $E[Y]$ et $Var[Y]$.

2) Déterminer l'estimateur de θ noté $\widehat{\theta}_{MV}$ obtenu à l'aide de la méthode du maximum de vraisemblance. Faire l'application numérique. A l'aide des résultats de la question 1), dire si $\widehat{\theta}_{MV}$ est un estimateur sans biais, convergent et efficace de θ.

3) Déterminer l'estimateur de θ noté $\widehat{\theta}_{MOM}$ obtenu à l'aide de la méthode des moments. Faire l'application numérique.

4) De nombreuses études en laboratoire ont montré que le paramètre θ pouvait être muni de la loi a priori

$$g(\theta) = \frac{1}{\theta} I_{[\frac{1}{4},\frac{e}{4}]}(\theta)$$

a) Déterminer l'estimateur du maximum a Posteriori du paramètre θ noté $\widehat{\theta}_{MAP}$. Faire l'application numérique et comparer $\widehat{\theta}_{MV}$ et $\widehat{\theta}_{MAP}$.

b) On pose $T = -\sum_{i=1}^{n} \ln X_i$ et $G_k(a,b) = \int_a^b u^{k-1} e^{-u} du$. Montrer que l'estimateur de la moyenne a posteriori de θ noté $\widehat{\theta}_{MOAP}$ peut s'écrire :

$$\widehat{\theta}_{MOAP} = T \frac{G_{n-1}(\frac{4}{e}T, 4T)}{G_n(\frac{4}{e}T, 4T)}$$

$2^{ème}$ Partie : Tests d'Hypothèses

1) On désire tester les deux hypothèses

$$\begin{cases} H_0 : \theta = \theta_0 \\ H_1 : \theta = \theta_1 \text{ (avec } \theta_1 > \theta_0) \end{cases}$$

a) Déterminer la région critique du test de Neyman-Pearson.
b) Quelle est la loi approchée de la statistique de test pour n grand ?
c) En utilisant la loi approchée du b), montrer que pour $\alpha = 0.05$ et $\theta_0 = 0.4$, le test de Neyman-Pearson est de la forme :

$$\text{Rejet de } H_0 \text{ si } \frac{-1}{30} \sum_{i=1}^{30} \ln x_i > 0.52$$

d) Effectuer le test et calculer le risque β pour $\theta_1 = 1$.

2) On désire vérifier si les données x_i sont issues d'un échantillon de variables aléatoires X_i de densité $f_\theta(x)$. Effectuer un test du chi2 avec un risque $\alpha = 0.05$ en répartissant les données dans trois classes équiprobables et conclure.

3) Afin de tester si les données x_i sont issues d'un échantillon de variables aléatoires de densité

$$g(x) = 3x^2 I_{[0,1]}(x)$$

(correspondant à $\theta = 1/3$), effectuer un test du chi2 dans les mêmes conditions qu'à la question 2) et commenter les résultats obtenus.