STATISTIQUE MATHÉMATIQUE - EXAMEN.

ENSAI - Mastères spécialisés - 2024/2025 -

École nationale de la statistique et de l'analyse de l'information

Le sujet comporte 2 pages. Durée de l'épreuve : 2h. Tous documents autorisés, calculatrices autorisées. Le barème est indicatif et sans engagement. La qualité de la rédaction entre pour une part importante dans la notation. Vous pouvez admettre le résultat d'une question pour passer à la suivante.

I. 14 points

On considère un n-échantillon $X = (X_1, ..., X_n)$ d'une variable aléatoire de loi de Bernoulli de paramètre $\theta \in]0,1[$. On note $S_n = \sum_{i=1}^n X_i$ et $\overline{X} = S_n/n$ la moyenne empirique de l'échantillon.

1°. Déterminer la vraisemblance du n-échantillon $X_1, ..., X_n$. Le modèle est-il exponentiel ?

2°. Déterminer un estimateur $\widehat{\theta}_n$ de θ par la méthode du maximum de vraisemblance, puis par la méthode des moments. Que constatez-vous ?

3°. Montrer que cet estimateur est fortement consistant et converge en moyenne quadratique. Démontrer que $\widehat{\theta}_n$ est asymptotiquement normal et déterminer la variance de la loi limite.

4°. Déterminer un intervalle de confiance asymptotique pour θ, de niveau $1-\alpha$.

5°. $\widehat{\theta}_n$ est-il exhaustif ? Minimal ? Complet ?

6°. Le modèle est-il régulier ?

7°. Démontrer que $\widehat{\theta}_n$ est un estimateur VUMSB (sans biais de variance minimale) de θ.

8°. Calculer l'information de Fisher relative à θ et montrer que $\widehat{\theta}_n$ est un estimateur efficace de θ.

On considère que la loi *a priori* de θ est une loi bêta $\mathscr{B}(a,b)$, avec $a, b \in [0,1]$, dont la densité est donnée par

$$f(\theta) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}\theta^{a-1}(1-\theta)^{b-1}\mathbb{1}_{[0,1]}(\theta) \tag{1}$$

où Γ est la fonction Gamma d'Euler donnée par

$$\Gamma(a) = \int_0^{+\infty} x^{a-1}e^{-x}dx \tag{2}$$

On rappelle que l'espérance d'une variable aléatoire H de loi bêta est donnée par

$$\mathbb{E}[H] = \frac{a}{a+b} \tag{3}$$

9°. Déterminer la loi *a posteriori* de θ sachant $[X = x]$. En déduire que l'estimateur de la moyenne *a posteriori* peut s'exprimer au travers de la statistique S_n. Montrer que cet estimateur est biaisé, mais asymptotiquement sans biais.

10°. On suppose $n > 2$. On souhaite estimer la valeur de θ^2. Démontrer que $Y = X_1X_2$ est un estimateur sans biais de θ^2. Démontrer que $T_n = \mathbb{E}[Y|S_n]$ est un estimateur sans biais de variance minimale pour θ^2.

11°. On se propose de calculer explicitement l'expression de T_n. Calculer $\mathbb{P}[Y=1|S_n = s]$ (on distinguera les cas $s < 2$ et $s \geq 2$). En déduire que

$$T_n = \frac{S_n(S_n-1)}{n(n-1)}\mathbb{1}_{[S_n \geq 2]}. \tag{4}$$

II. 8 points

On considère un n-échantillon $X = (X_1, ..., X_n)$ d'une variable aléatoire dont la loi a pour densité :

$$f(x) = \theta e^{-\theta x}\mathbb{1}_{[0,+\infty[}(x) \tag{5}$$

avec $\theta > 0$ paramètre inconnu. On suppose que les X_i ne sont pas observés directement. La seule information connue est le fait que X_i soit supérieur à 2 ou non. On pose $Y_i = \mathbb{1}_{[X_i>2]}$ pour $i = 1, \ldots, n$ et

$$\overline{Y}_n = \frac{1}{n} \sum_{i=1}^{n} Y_i. \tag{6}$$

1°. Préciser la loi de Y_1 ainsi que celle de $n\overline{Y}_n$. Calculer $\mathbb{E}[Y_1]$ en fonction de θ.

On souhaite estimer le paramètre θ. On pose $\lambda = e^{-2\theta}$ et

$$\widehat{\theta}_n = \begin{cases} -\frac{1}{2} \ln \overline{Y}_n & \text{si } \overline{Y}_n > 0 \\ 0 & \text{sinon.} \end{cases} \tag{7}$$

2°. Calculer $\mathbb{P}\left[\overline{Y}_n \neq 0\right]$ et en déduire

$$\lim_{n \to +\infty} \mathbb{P}\left[\overline{Y}_n \neq 0\right]. \tag{8}$$

3°. Démontrer que $\overline{Y}_n$ converge presque sûrement vers λ et que $\sqrt{n}(\overline{Y}_n - \lambda)$ est asymptotiquement normal. Préciser la variance de la loi limite en fonction de θ.

4°. Expliquer pourquoi $\overline{Y}_n$ est différent de 0, presque sûrement, lorsque n suffisamment grand.

5°. Démontrer que $\widehat{\theta}_n$ est fortement consistant et asymptotiquement normal. Préciser la variance de la loi limite en fonction de θ.