Statistique

Examen du 11 juin 2003

Sans document (durée deux heures).
Calculatrice non programmable autorisée. Exercices indépendants.

Exercice 1 [9 points]

Soit $(X_1, \ldots, X_n)$ un échantillon id de loi de densité

$$f_\theta(x) = \theta x^{\theta-1}\,, \quad 0 < x < 1\,,$$

avec $\theta > 0$.

1. [1] Montrer que la vraisemblance associée à l'échantillon $(x_1, \ldots, x_n)$ ne dépend de l'échantillon qu'au travers de
$$\sum_{i=1}^{n} \log x_i\,.$$

2. [1] Montrer que $-\log(X_i)$ est de loi exponentielle de paramètre θ.

3. [1] Donner l'estimateur du maximum de vraisemblance de θ, $\widehat{\theta}(X_1, \ldots, X_n)$.

4. [1] Montrer que $n/\widehat{\theta}(X_1, \ldots, X_n)$ est une variable aléatoire de densité sur $\mathbb{R}^+$
$$\frac{\theta^n}{(n-1)!}\, y^{n-1} \exp\{-\theta y\}\,.$$

5. [1] Déduire de la question 2 que $1/\widehat{\theta}(X_1, \ldots, X_n)$ est un estimateur sans biais de $1/\theta$.

6. [2] Déduire de la question 2 la variance de $1/\widehat{\theta}(X_1, \ldots, X_n)$ et montrer qu'elle atteint la borne de Cramér–Rao.

7. [2] Montrer que $n\,\widehat{\theta}(X_1, \ldots, X_n)/\theta$ a une loi indépendante de θ et en déduire *la forme* d'un intervalle de confiance sur θ.

Exercice 2 [6 points]

Soit $n > 0$ un entier. On définit la variable aléatoire

$$T_n = \sum_{i=1}^{n} Y_i, \qquad Y_i \sim \mathcal{G}(i/n)\,,$$

où les Y_i sont des variables aléatoires indépendantes géométriques. *On rappelle que, pour une variable géométrique $X \sim \mathcal{G}(p)$, $P(X = k) = p(1-p)^{k-1}$ ($k = 1, 2, \ldots$), $\mathbb{E}[X] = 1/p$, $\mathrm{var}(X) = (1-p)/p^2$, et sa fgm est*

$$\mathbb{E}[\exp(tX)] = pe^t/[1-(1-p)e^t] \qquad\qquad (t < -log(1-p)\,.$$

1. [1] Représenter $\mathbb{E}[T_n]$ et $\mathrm{var}(T_n)$ sous forme de sommes.

2. [1] En utilisant la fonction génératrice des moments, montrer que T_n *n'est pas* une variable aléatoire de loi géométrique.

3. [1] Montrer que

$$\sum_{i=1}^{n} \frac{1}{i} \geq \int_1^{n+1} \frac{1}{x}\, dx = \log(n+1)$$

et que

$$\sum_{i=1}^{n} \frac{1}{i} \leq 1 + \int_1^{n} \frac{1}{x}\, dx = 1 + \log(n)$$

[0.5] En déduire que

$$\lim_{n\to\infty} \mathbb{E}[T_n]/(n \log n) = 1\,.$$

4. [0.5] Montrer que, pour toute variable aléatoire X et tout $a \in \mathbb{R}$,

$$\mathbb{E}[(X-a)^2] = \mathrm{var}(X) + (\mathbb{E}[X]-a)^2\ .$$

5. [2] En déduire que

$$T_n/(n \log n)$$

converge en probabilité vers 1 quand n tend vers l'infini. (*Indication:* On pourra utiliser, *sans les redémontrer,* l'inégalité $P(|S| > \epsilon) \leq \mathbb{E}[S^2]/\epsilon^2$ et l'identité

$$\left.\sum_{i=1}^{\infty} i^{-2} = \pi^2/6.\right)$$

Numéro identifiant :

Questionnaire à choix multiples (5 points)

(Une seule réponse valable par question à cocher sur cette feuille; réponse positive : `+1`*, non réponse :* `0` *et réponse négative :* `-1`*)*

Attention: Reportez le détail des calculs sur votre copie

1. Si $(X_1, \dots, X_n)$ est un échantillon normal $\mathcal{N}(0, \sigma^2)$, un estimateur sans biais de σ^2 est $\alpha \sum_{i=1}^n X_i^2$ avec

   ○ $\alpha = 1/(n-1)$ ○ $\alpha = 1/n$ ○ $\alpha = 1/(n+1)$
   ○ $\alpha = 1/(n+2)$ ○ $\alpha = n$ ○ $\alpha = n-1$

2. L'intégrale $I = \displaystyle\int_0^1 \int_0^1 ye^{-xy} dxdy$ vaut

   ○ $I = 1/2$ ○ $I = e^{-1}$ ○ $I = e^{-1}/2$
   ○ $I = 1 - e^{-1}$ ○ $I = e^{-2}$ ○ $I = 1/2 - e^{-2}$

3. Si U et V sont des variables uniformes $\mathcal{U}([0,1])$ indépendantes, $Z = \log(U/V)$ a comme densité

   ○ $f(z) = \exp(-|z|/2)/4$ ○ $f(z) = 2\exp(|z|)$
   ○ $f(z) = 1/\pi(1+z^2)$ ○ $f(z) = \exp(-|z|)/2$

4. Si $X_1, \dots, X_n$ est un échantillon iid de loi exponentielle $\mathcal{E}xp(\theta)$, $n \min(X_1, \dots, X_n)$ converge en loi vers une loi

   ○ $\mathcal{N}(\theta, \theta^2)$ ○ $\mathcal{E}xp(\theta)$ ○ $\mathcal{G}a(\theta, \theta)$
   ○ $\mathcal{N}(\theta, \theta)$ ○ $\mathcal{E}xp(n\,\theta)$ ○ $\mathcal{G}a(n, n\theta)$

5. Si $\overline{X}_n$ est une variable aléatoire de loi $\mathcal{N}(\theta, 1/n)$, un intervalle de confiance de niveau $\alpha = 0.95$ sur θ est donné par

   ○ $[\overline{X}_n - 1.96/n, \overline{X}_n + 1.96/n]$ ○ $[\overline{X}_n - 1.96\sqrt{n}, \overline{X}_n + 1.96\sqrt{n}]$
   ○ $[\overline{X}_n - 1.96/n^2, \overline{X}_n + 1.96/n^2]$ ○ $[\overline{X}_n - 1.96/\sqrt{n}, \overline{X}_n + 1.96/\sqrt{n}]$
