EXAMEN (ESTIMATION) - CORRIGÉ SUCCINCT

I.Nikiforov

29 novembre 2001

Documents autorisés : Polycopiés distribués, formulaires et notes de cours.

Exercice 1 *Soit $\xi_1, \xi_2, \ldots, \xi_n$ une suite de variables aléatoires indépendantes, $\xi_i \sim U(a,b)$, où $U(a,b)$ est la distribution uniforme sur $[a;b]$, $a < b$.*

1. *Calculer la densité et l'espérance mathématique de la variable aléatoire $\max(\xi_1, \ldots, \xi_n)$.*
2. *Calculer la densité et l'espérance mathématique de la variable aléatoire $\min(\xi_1, \ldots, \xi_n)$.*

Réponses :

1. La fonction de répartition d'une loi uniforme est définie par :

$$F(x) = \begin{cases} 0 & \text{si} \quad x < a \\ \dfrac{x-a}{b-a} & \text{si} \quad x \in [a;b] \\ 1 & \text{si} \quad x > b \end{cases} \quad , a < b.$$

La fonction de répartition de la variable $\xi_{(n)} = \max(\xi_1, \ldots, \xi_n)$ est (pour les $x \in [a;b]$)

$$F_{\xi_{(n)}}(x) = \mathbb{P}(\xi_{(n)} < x) = \prod_{i=1}^{n} \mathbb{P}(\xi_i < x) = [\mathbb{P}(\xi_i < x)]^n = \left(\frac{x-a}{b-a}\right)^n.$$

La densité de $\xi_{(n)}$ est

$$f_{\xi_{(n)}}(x) = \begin{cases} \dfrac{d}{dx}\left[\dfrac{x-a}{b-a}\right]^n = n\dfrac{(x-a)^{n-1}}{(b-a)^n} & \text{si} \quad x \in [a;b] \\ 0 & \text{si} \quad x \notin [a;b] \end{cases} \quad , a < b.$$

L'espérance mathématique de la variable aléatoire $\xi_{(n)}$ est

$$\mathbb{E}(\xi_{(n)}) = \int_a^b x f_{\xi_{(n)}}(x) dx = \int_a^b xn\frac{(x-a)^{n-1}}{(b-a)^n}\, dx = b - \frac{b-a}{n+1}.$$

2. La fonction de répartition de la variable $\xi_{(1)} = \min(\xi_1, \ldots, \xi_n)$ est (pour les $x \in [a;b]$)

$$F_{\xi_{(1)}}(x) = \mathbb{P}(\xi_{(1)} < x) = 1 - \prod_{i=1}^{n} \mathbb{P}(\xi_i \geq x) = 1 - \left(\frac{b-x}{b-a}\right)^n.$$

La densité de $\xi_{(1)}$ est

$$f_{\xi_{(1)}}(x) = \begin{cases} \dfrac{d}{dx}\left[1 - \left(\dfrac{b-x}{b-a}\right)^n\right] = n\dfrac{(b-x)^{n-1}}{(b-a)^n} & \text{si} \quad x \in [a;b] \\ 0 & \text{si} \quad x \notin [a;b] \end{cases} \quad , a < b.$$

L'espérance mathématique de la variable aléatoire $\xi_{(1)}$ est

$$\mathbb{E}(\xi_{(1)}) = \int_a^b x f_{\xi_{(1)}}(x) dx = \int_a^b xn\frac{(b-x)^{n-1}}{(b-a)^n}\, dx = a + \frac{b-a}{n+1}.$$

Exercice 2 *Soit $\xi_1, \xi_2, \ldots, \xi_n$ une suite de variables aléatoires indépendantes, $\xi_i \sim U(a,b)$, où $U(a,b)$ est la distribution uniforme sur $[a;b]$, $a < b$.*

On cherche à estimer le paramètre $\theta = b - a$. Soit

$$\hat{\theta} = \xi_{(n)} - \xi_{(1)},$$

où $\xi_{(n)} = \max(\xi_1, \ldots, \xi_n)$ et $\xi_{(1)} = \min(\xi_1, \ldots, \xi_n)$, un estimateur de θ.

1. *L'estimateur $\hat{\theta}$ est-il biaisé ?*
2. *Supposons que $n \to \infty$. L'estimateur $\hat{\theta}$ est-il asymptotiquement biaisé ?*

Réponses :

1. En utilisant les résultats de l'exercice 1, on obtient
$$\mathbb{E}_{a,b}(\hat{\theta}) = \mathbb{E}_{a,b}(\xi_{(n)} - \xi_{(1)}) = \frac{n-1}{n+1}(b-a) \neq \theta = b-a.$$
Donc, l'estimateur $\hat{\theta} = \xi_{(n)} - \xi_{(1)}$ est biaisé.
2. L'estimateur $\hat{\theta} = \xi_{(n)} - \xi_{(1)}$ est asymptotiquement non biaisé, car
$$\lim_{n\to\infty} \mathbb{E}_{a,b}(\xi_{(n)} - \xi_{(1)}) = \theta = b-a.$$

Exercice 3 *La loi binomiale $B(1,p)$ est définie par :*
$$\mathbb{P}(\xi = k) = p^k(1-p)^{1-k},$$
où $k \in \{0,1\}$ sont entiers et $0 < p < 1$. Soit $(\xi_1, \ldots, \xi_n)$ un vecteur aléatoire, où $\xi_i \sim B(1,p)$. On cherche à estimer le paramètre θ tel que
$$\theta(p) = \arcsin \sqrt{p}.$$
Soit
$$\hat{\theta} = \arcsin \sqrt{\overline{\xi}} \;\; \textit{avec} \;\; \overline{\xi} = \frac{1}{n}\sum_{i=1}^{n} \xi_i$$
un estimateur de θ. Supposons que $n \to \infty$.

1. *L'estimateur $\hat{\theta}$ est-il convergent ?*
2. *L'estimateur $\hat{\theta}$ est-il asymptotiquement normal ?*

Réponses :

1. En utilisant la loi des grands nombres, on trouve que la moyenne $\overline{\xi} = \frac{1}{n}\sum_{i=1}^{n} \xi_i$ converge en probabilité vers $\mathbb{E}_p(\xi) = p : \overline{\xi} \underset{\mathbb{P}}{\to} p$. En utilisant le théorème de continuité, on obtient que $\hat{\theta} = \arcsin \sqrt{\overline{\xi}} \underset{\mathbb{P}}{\to} \theta(p) = \arcsin \sqrt{p}$.
2. L'estimateur $\hat{\theta} = h(\overline{\xi}) = \arcsin \sqrt{\overline{\xi}}$ est asymptotiquement normal car la fonction $h \;: u \mapsto h(u) = \arcsin \sqrt{u}$ est dérivable, $0 < \left|\frac{dh}{du}(u_0)\right| < \infty$, où $u_0 = \mathbb{E}_p(\xi) = p$, et $\mathbb{E}_p(\xi^2) < \infty$.

Exercice 4 *À partir des observations $y_1, y_2, \ldots, y_n$ d'une certaine grandeur économique y, on retient le modèle suivant*
$$y_i = a y_{i-1} + \xi_i, \quad i = 1, 2, \ldots, n.$$
On considère que $\xi_i \sim \mathcal{N}(0, \sigma^2)$, $\sigma^2 > 0$, et $y_0 = 0$.

1. *Donner l'expression de la log-vraisemblance pour le modèle retenu.*
2. *Déterminer l'estimateur du maximum de vraisemblance $\hat{a}$ de a.*

Réponses :

1. La densité conditionnelle de y_i sachant y_{i-1} est
$$f(y_i|y_{i-1}) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{-\frac{(y_i - a y_{i-1})^2}{2\sigma^2}\right\}.$$
On obtient la densité des n premières observations $y_1, \ldots, y_n$ sachant y_0 :
$$f_a(y_1, \ldots, y_n \mid y_0) = \prod_{i=1}^{n} f(y_i|y_{i-1}).$$
D'où on trouve la log-vraisemblance pour le modèle
$$\ln f_a(y_1, \ldots, y_n \mid y_0) = -\frac{n}{2}\ln 2\pi\sigma^2 - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(y_i - a y_{i-1})^2.$$
2. On calcule la dérive partielle
$$\frac{\partial}{\partial a}\left[-\frac{n}{2}\ln 2\pi\sigma^2 - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(y_i - a y_{i-1})^2\right] = \frac{1}{\sigma^2}\sum_{i=1}^{n}(y_i - a y_{i-1})y_{i-1} = 0.$$
D'où on trouve l'estimateur du maximum de vraisemblance de a :
$$\hat{a} = \frac{\sum_{i=1}^{n} y_i y_{i-1}}{\sum_{i=1}^{n} y_{i-1}^2} = \frac{\sum_{i=2}^{n} y_i y_{i-1}}{\sum_{i=2}^{n} y_{i-1}^2}.$$