Université de Montpellier
M2 Biostat - HAX918X

Année 2023-2024

Examen partiel - 18 octobre 2023
Durée 1h - Documents interdits

Exercice 1 (6 pts) Nous considérons un n-échantillon $(x_1, \ldots, x_n)$ distribué suivant une loi $\mathcal{N}(\theta, 1)$. Pour $n = 100$, nous observons $\bar{x} = 1$. Nous nous plaçons dans le paradigme bayésien.

1) (3 pts) Expliciter la loi a priori de Jeffreys pour le paramètre θ. Est-elle propre ?

2) (3 pts) En utilisant le fait que $F^{-1}_{N(0,1)}(0.975) \approx 2$, donner un intervalle de crédibilité de niveau 95% pour le paramètre θ .

Exercice 2 (8 pts) Nous considérons la variable aléatoire X admettant pour densité

$$f_X(x) = \frac{\exp(-|x|)}{2(1-\exp(-2))}\mathbb{I}_{[-2,2]}(x)$$

1) (4 pts) Proposer une méthode de simulation exacte et directe de X (directe dans le sens où elle ne nécessite pas le recours à un algorithme itératif).

2) (4 pts) Proposer une méthode de simulation de X basée sur l'utilisation d'un algorithme acceptation rejet.

Exercice 3 (6 pts) Nous considérons $x|\theta \sim \mathcal{U}([0,\theta])$ avec $\theta > 0$ et supposons que θ est distribué a priori suivant une loi de Pareto de paramètre (θ_0, α) :

$$\pi(\theta; \theta_0, \alpha) = \alpha\,(\theta_0)^\alpha\,\theta^{-\alpha-1}\mathbf{1}_{[\theta_0,+\infty[}(\theta)$$

avec $\theta_0 > 0$ et $\alpha > 0$.

1) (3 pts) Donner la loi a posteriori de θ.

2) (3 pts) Calculer l'estimateur bayésien associé à la fonction de perte

$$L(\theta, d) = (\theta - \delta)^2\,.$$

Correction Examen Partiel
18 octobre 2023
HAX918X

Exercice 1

1) $$f(x|\theta) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}(x-\theta)^2\right)$$

$$\log(f(x|\theta)) = -\frac{1}{2}\log(2\pi) - \frac{1}{2}(x-\theta)^2$$

$$\frac{d\log}{d\theta} f(x|\theta) = (x-\theta)$$

$$\frac{d^2\log f}{(d\theta)^2}(x|\theta) = -1$$

$$\Rightarrow I_x(\theta) = 1$$

$$\Rightarrow I_n(\theta) = n$$

Nous en déduisons que

$$\pi^J(\theta) \propto 1$$

la loi a priori de Jeffreys est impropre.

②

2]

$$\pi^J(\theta | x_1, \dots, x_n)$$

$$\propto \pi^J(\theta) f(x_1, \dots, x_n | \theta)$$

$$\Rightarrow \pi^J(\theta | x_1, \dots, x_n)$$

$$\propto \exp\left(-\frac{1}{2} \sum_{i=1}^n (x_i - \theta)^2\right)$$

$$\Rightarrow \pi^J(\theta | x_1, \dots, x_n)$$

$$\propto \exp\left(-\frac{n\theta^2}{2} + \theta \sum_{i=1}^n x_i\right)$$

$$\Rightarrow \pi^J(\theta | x_1, \dots, x_n)$$

$$\propto \exp\left(-\frac{n}{2} (\theta - \overline{x})^2\right)$$

$$\Rightarrow \theta | x_1, \dots, x_n \sim N\left(\overline{x}, \frac{1}{n}\right)$$

Un intervalle de crédibilité pour le paramètre θ est un intervalle $[a, b]$ $(a < b)$ tel que

$$\mathbb{P}^{\pi^J}(\theta \in [a, b] | x_1, \dots, x_n) = 1 - \alpha$$

où $1-\alpha$ est le niveau (x le risque).

Nous remarquons que

$$\sqrt{n}(\theta - \bar{x}) \mid x_1, \dots, x_n \sim \mathcal{N}(0,1)$$

La densité de la loi $\mathcal{N}(0,1)$ étant symétrique par rapport à l'axe des ordonnées, il est rationnel de prendre des risques symétriques

Comme $F^{-1}_{\mathcal{N}(0,1)}(0{,}975) = 2$, $F^{-1}_{\mathcal{N}(0,1)}(0{,}025) = -2$

et $\mathbb{P}^{\pi^J}\left(-2 \leq \sqrt{n}(\theta - \bar{x}) \leq 2 \mid x_1, \dots, x_n\right) = 0{,}95$

Comme $n = 100$ et $\bar{x} = 1$, nous avons

$$\mathbb{P}^{\pi^J}\left(1 - 0{,}2 \leq \theta \leq 1 + 0{,}2 \mid x_1, \dots, x_n\right) = 0{,}95$$

Un intervalle de crédibilité de niveau 95% pour θ est $[0{,}8 ; 1{,}2]$

Exercice 2

1) $$f_X(x) = \frac{\exp(-|x|)}{2(1 - \exp(-2))} \mathbb{1}_{[-2,2]}(x)$$

4

Calculons la fonction de répartition de X

$$\mathbb{P}(X \leq x) = \begin{cases} 0 \text{ si } x < -2 \\ 1 \text{ si } x > 2 \\ \int_{-2}^{x} f_X(u)\,du \text{ si } x \in [-2,2] \end{cases}$$

Soit $x \in [-2,0]$,

$$F_X(x) = \int_{-2}^{x} \frac{\exp(-|u|)}{2(1-\exp(-2))}\,du$$

$$F_X(x) = \frac{1}{2(1-\exp(-2))}\left[\int_{-2}^{x} e^{u}\,du\right]$$

$$= \frac{1}{2(1-\exp(-2))}\left(e^{x} - e^{-2}\right)$$

Soit $x \in [0,2]$,

$$F_X(x) = \frac{1}{2(1-\exp(-2))}\left[\int_{-2}^{0} e^{u}\,du + \int_{0}^{x} e^{-u}\,du\right]$$

Comme $\forall_{[0,1]} |(1+x)^{-1} e^{-x}| < \frac{1}{2}$, (5)
nous cherchons la plus petite valeur de N telle que

$$N \geq \frac{10^6}{2 I^2}.$$

Par ailleurs,

$$I = \int_0^1 (1+x)^{-1/2} e^{-x} dx$$

$$\frac{1}{e^x \sqrt{1+x}} \geq \frac{1}{\sqrt{2}\, e} \quad \forall x \in [0,1]$$

$$I \geq (2)^{-1/2} e^{-1}$$

$$I^2 \geq 2^{-1} e^{-2}$$

$$\frac{1}{I^2} \leq 2e^2$$

Nous cherchons donc la plus petite valeur de N telle que

$$N \geq 10^6 e^2$$

Il faut donc que $N \geq 7389057$

⑥

Exercice 3

Nous proposons d'utiliser un algorithme acceptation-rejet avec comme loi de proposition une loi uniforme sur $[0,1]$

$$\frac{f_X(x)}{g(x)} \leq 2 \quad \forall x \in [0,1]$$

Algorithme

Simuler $y \sim \mathcal{U}_{[0,1]}$ (*)

Calculer $\frac{f_X(y)}{2g(y)} = r$

Simuler $u \sim \mathcal{U}_{[0,1]}$, si $r \leq u$ alors $x = y$

sinon retourner à (*)

Exercice 4

1) $\log(f(x|\theta)) = x\log(\theta) + \log(1-\theta)$

$\forall x \in \mathbb{N}$

$$\frac{d \log f}{d\theta}(x|\theta) = \frac{x}{\theta} - \frac{1}{1-\theta}$$

$$\frac{d^2 \log f}{(d\theta)^2}(x|\theta) = -\frac{x}{\theta^2} - \frac{1}{(1-\theta)^2}$$

$$I_x(\theta) = \frac{1}{\theta(1-\theta)} + \frac{1}{(1-\theta)^2}$$

$$I_x(\theta) = \frac{1}{\theta(1-\theta)^2}$$

$$\Longrightarrow \pi^J(\theta) \propto \theta^{-1/2}(1-\theta)^{-1}\mathbb{1}_{]0,1[}(\theta)$$

loi impropre

$$2) \quad \pi^J(\theta|x_1,\dots,x_n) \propto \theta^{\sum_{i=1}^n x_i - 1/2}(1-\theta)^n \mathbb{1}_{]0,1[}(\theta)$$

$$\Longrightarrow \theta|x_1,\dots,x_n \sim \text{Beta}\left(\sum_{i=1}^n x_i + \tfrac{1}{2}, n\right)$$