Université de Montpellier — Année 2022-2023

M2 SSD-Biostat - HAX912X

Examen partiel - 14 octobre 2022
Durée 1h30 - Documents interdits

Exercice 1 (7 pts) Un des premiers exemples d'utilisation de la statistique bayésienne remonte à Laplace en 1786. Celui-ci décida de répondre à la question suivante : au regard du nombre observé de naissances de filles entre 1745 et 1770 $x = 241945$ parmi les $n = 493472$ naissances à Paris, est-il credible que la probabilité de naissance d'une fille soit plus petite que 1/2 ? On note cette probabilité θ et il est clair que $x|\theta \sim \mathcal{B}(n, \theta)$.

1) (3 pts) Laplace munit le paramètre θ d'une loi uniforme sur $[0, 1]$. Répondre à la question qu'il s'est posé sous forme d'une intégrale.

2) (2 pts) Pour le modèle considéré, expliciter la loi a priori de Jeffreys.

3) (2 pts) Donner la loi a posteriori de θ pour la loi a priori de Jeffreys et l'estimateur bayésien associé à la fonction de perte quadratique.

On rappelle que si la variable aléatoire θ suit une loi Beta de paramètre (a, b), elle admet comme densité

$$\pi(\theta) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \theta^{a-1}(1-\theta)^{b-1}\mathbf{1}_{\theta \in [0,1]}$$

avec $a > 0$ et $b > 0$, $\mathbb{E}(\theta) = \dfrac{a}{a+b}$ et $\mathbb{V}(\theta) = \dfrac{ab}{(a+b)^2(a+b+1)}$.

Exercice 2 (7 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) (2 pts) Calculer l'estimateur bayésien associé à la fonction de perte

$$L_1(\theta, \delta) = (\theta - \delta)^2 .$$

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

$$L_2(\theta, \delta) = \frac{(\theta - \delta)^2}{\theta^2} .$$

Exercice 3 (6 pts) Nous considérons un n-échantillon $(x_1, \ldots, x_n)$ iid suivant une loi $\mathcal{N}(\theta, 1)$. Nous considérons que $\theta \in \Theta = \{1, 4\}$. Nous nous plaçons dans le paradigme bayésien et supposons que $\mathbb{P}(\theta = 1) = \mathbb{P}(\theta = 4) = \dfrac{1}{2}$.

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

2) (3 pts) Donner un estimateur bayésien associé à la fonction de perte suivante

$$L(\theta, d) = \begin{cases} 0 & \text{si} \quad \theta = d \\ 1 & \text{si} \quad \theta = 1 \quad \text{et} \quad d = 4 \\ 2 & \text{si} \quad \theta = 4 \quad \text{et} \quad d = 1 \end{cases}.$$

Correction examen partiel HAX918X 14/10/2022

(1)

Exercice 1

1) $\theta \sim U_{[0,1]}$

$x|\theta \sim \mathcal{B}(n,\theta)$

Laplace cherche à calculer

$\mathbb{P}^{\pi}(\theta < 1/2 | x)$.

Donnons la loi a posteriori de θ

$\pi(\theta|x) \propto C_n^x \, \theta^x (1-\theta)^{n-x} \, \mathbb{I}_{[0,1]}(\theta)$

$\Rightarrow \theta|x \sim \text{Beta}(x+1, n-x+1)$

On en déduit que

$$\mathbb{P}^{\pi}(\theta < 1/2 \mid x) = \int_0^{1/2} \frac{\Gamma(n+2)}{\Gamma(x+1)\Gamma(n-x+1)} \theta^{x} (1-\theta)^{n-x} \, d\theta$$

$$\mathbb{P}^{\pi}(\theta < 1/2 \mid x) = \frac{\Gamma(n+2)}{\Gamma(x+1)\Gamma(n-x+1)} \int_0^{1/2} \theta^{x} (1-\theta)^{n-x} \, d\theta$$

2] $L(\theta; x) = ct + x \log(\theta) + (n-x)\log(1-\theta)$

$\forall x \in \{0, ----, n\}$

$$\frac{dL}{d\theta}(\theta; x) = +\frac{x}{\theta} - \frac{n-x}{1-\theta}$$

$$\frac{d^2L}{(d\theta)^2}(\theta; x) = -\frac{x}{\theta^2} - \frac{(n-x)}{(1-\theta)^2}$$

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

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

$$\Rightarrow \theta \sim \mathrm{Beta}(1/2, 1/2)$$

3]

$$\pi^J(\theta|x) \propto f(x|\theta)\,\pi^J(\theta)$$

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

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

$$\Rightarrow \theta|x \sim \text{Beta}\left(x+\tfrac{1}{2},\, n-x+\tfrac{1}{2}\right)$$

L'estimateur bayésien pour la fonction de perte quadratique est la moyenne de la loi a posteriori

$$\hat{\theta} = E^\pi(\theta|x) = \frac{x+1/2}{n+1}$$

4

Exercice 2

1] $\pi(\theta|x) \propto f(x|\theta)\pi(\theta)$

$$\pi(\theta|x) \propto \mathbb{1}_{[0,\theta]}(x) \; \theta^{-\alpha-2} \; \mathbb{1}_{[\theta_0;+\infty[}(\theta)$$

$$\Rightarrow \pi(\theta|x) \propto \theta^{-\alpha-2} \; \mathbb{1}_{[\max(x,\theta_0);+\infty[}(\theta)$$

$$\Rightarrow \theta|x \sim \text{Pareto}(\max(x,\theta_0), \alpha+1)$$

2] L'estimateur bayésien associé à la fonction de perte $L_1(\theta,\delta)$ est la moyenne de la loi a posteriori

⑤

$$\widehat{\theta}^1 = E^{\Pi}(\theta | x)$$

$$\Rightarrow \widehat{\theta}^1 = \int_{-\infty}^{+\infty} \theta (\alpha+1) \left[\max(\theta_0, x) \right]^{\alpha+1} \theta^{-\alpha-2} \mathbb{1}(\theta)\, d\theta$$

$[\max(\theta_0, x) ; +\infty[$

$$\Rightarrow \widehat{\theta}^1 = (\alpha+1) \left[\max(\theta_0, x) \right]^{\alpha+1} \int_{\max(\theta_0, x)}^{+\infty} \theta^{-\alpha-1}\, d\theta$$

$$\Rightarrow \widehat{\theta}^1 = \frac{(\alpha+1)(\max(\theta_0, x))^{\alpha+1}}{\alpha\, (\max(\theta_0, x))^{\alpha}}$$

$$\Rightarrow \widehat{\theta}^1 = \left(\frac{\alpha+1}{\alpha} \right) \max(\theta_0, x)$$

3)

$$\hat{\theta}^2 \in \arg\min_{d} \underbrace{E^{\Pi}\left[\frac{(\theta-d)^2}{\theta^2} \middle| x\right]}_{h(d)}$$

$$h(d) = E^{\Pi}\left[\frac{\theta^2 - 2\theta d + d^2}{\theta^2} \middle| x\right]$$

$$h(d) = 1 - 2d\, E^{\Pi}(\theta^{-1}/x) + d^2 E^{\Pi}(\theta^{-2}/x)$$

$$h'(d) = -2\, E^{\Pi}(\theta^{-1}/x) + 2d\, E^{\Pi}(\theta^{-2}/x)$$

$$h'(d^*) = 0 \iff d^* = \frac{E^{\Pi}(\theta^{-1}/x)}{E^{\Pi}(\theta^{-2}/x)}$$

$$h''(d) = 2\, E^{\Pi}(\theta^{-2}/x) > 0$$

$$\hat{\theta}^2 = \frac{\mathbb{E}^{\pi}(\theta^{-1}/x)}{\mathbb{E}^{\pi}(\theta^{-2}/x)}$$

$$\mathbb{E}^{\pi}(\theta^{-1}/x) = \frac{(\alpha+1)(\max(\theta_0,x))^{\alpha+1}}{(\alpha+2)(\max(\theta_0,x))^{\alpha+2}}$$

$$\mathbb{E}^{\pi}(\theta^{-2}/x) = \frac{(\alpha+1)(\max(\theta_0,x))^{\alpha+1}}{(\alpha+3)(\max(\theta_0,x))^{\alpha+3}}$$

$$\Rightarrow \hat{\theta}^2 = \frac{(\alpha+3)}{(\alpha+2)}\left[\max(\theta_0,x)\right]$$

Εργασία 3

$\underline{1}\rfloor$ $$\pi^{\pi}(\theta=1/x) \propto \frac{1}{2} e^{-\frac{1}{2}\sum_{i=1}^{n}(x_i-1)^2}$$

$$\pi^{\pi}(\theta=4/x) \propto \frac{1}{2} e^{-\frac{1}{2}\sum_{i=1}^{n}(x_i-4)^2}$$

$$\Rightarrow \mathbb{P}^{\pi}(\theta=1|x) = \frac{e^{-\frac{1}{2}\sum_{i=1}^{n}(x_i-1)^2}}{e^{-\frac{1}{2}\sum_{i=1}^{n}(x_i-1)^2} + e^{-\frac{1}{2}\sum_{i=1}^{n}(x_i-4)^2}}$$

et $\mathbb{P}^{\pi}(\theta=4|x) = \dfrac{e^{-\frac{1}{2}\sum_{i=1}^{n}(x_i-4)^2}}{e^{-\frac{1}{2}\sum_{i=1}^{n}(x_i-1)^2} + e^{-\frac{1}{2}\sum_{i=1}^{n}(x_i-4)^2}}$

2)
$$\mathbb{E}^{\pi}\left[L(\theta,\delta)|x\right] = \mathbb{1}_{\{\delta=4\}}\mathbb{P}^{\pi}(\theta=1|x) + 2\,\mathbb{1}_{\{\delta=1\}}\mathbb{P}^{\pi}(\theta=4|x)$$

$$\hat{\theta}=1 \iff 2\,\mathbb{P}^{\pi}(\theta=4|x) < \mathbb{P}^{\pi}(\theta=1|x)$$

$$\hat{\theta}=1 \iff \mathbb{P}^{\pi}(\theta=4|x) < \frac{1}{3}$$

$$\hat{\theta}=1 \iff \mathbb{P}^{\pi}(\theta=1|x) \geq \frac{2}{3}$$

$$\hat{\theta} = \begin{cases} 1 & \text{si } \mathbb{P}^{\pi}(\theta=1|x) \geq 2/3 \\ 4 & \text{sinon} \end{cases}$$