Université de Montpellier
M2 SSD-Biostat - HAX912X

Année 2021-2022

Examen partiel - 8 octobre 2021
Durée 1h30 - Documents interdits

Exercice 1 (4 pts) L'entraineur d'une équipe de football souhaite classer ses joueurs par rapport à leurs performances aux tirs aux buts (penaltys). Il cherche donc à estimer la probabilité intrinsèque de chacun de marquer un penalty. Pour ce faire, pendant 30 jours, il enregistre le nombre de tirs consécutifs qu'il leur a été nécessaire pour marquer un penalty et ce face à différents gardiens.
Nous nous concentrons sur le cas d'un joueur et nous supposons que tous les essais sont indépendants. Nous disposons alors de 30 réalisations indépendantes $x_1, \ldots, x_{30}$ suivant une loi géométrique de paramètre $\theta \in [0,1]$ où θ est la probabilité intrinsèque que le joueur marque un penalty. Nous rappelons que, pour tout $i \in \{1, \ldots, 30\}$,

$$\mathbb{P}(X_i = x_i) = (1-\theta)^{x_i - 1}\theta \times \mathbb{I}_{x_i \in \mathbb{N}_*}\,.$$

Nous rappelons que l'espérance d'une loi géométrique de paramètre θ est égale à $1/\theta$.

Nous nous plaçons dans le paradigme bayésien.

1 (2 pts) Donner la loi a priori de Jeffreys $\pi^J(\theta)$ pour le paramètre θ.

2 (2 pts) Donner la loi a posteriori de θ associé à $\pi^J(\theta)$.

Exercice 2 (6 pts) Nous considérons un n-échantillon $(x_1, \ldots, x_n)$ iid suivant une loi normale centrée de variance $\theta > 0$. Nous considérons que $\theta \in \Theta = \{1, 2\}$. Nous nous plaçons dans le paradigme bayésien et supposons que $\mathbb{P}(\theta = 1) = \mathbb{P}(\theta = 2) = \dfrac{1}{2}$.

1) (2 pts) Donner la loi a posteriori de θ. Est-elle propre ?

2) (2 pts) Étudier le comportement asymptotique des probabilités a posteriori lorsque n tend vers l'infini.

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

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

Exercice 3 (6 pts) Nous considérons le modèle de régression gaussien tel que

$$y|\theta, x \sim \mathcal{N}(\theta x, 1)$$

où $x \in \mathbb{R}$ est fixé et $\theta \in \mathbb{R}$ est un paramètre inconnu. Nous supposons que nous observons un n réalisations indépendantes $(y_1, \ldots, y_n)$ de la variable à expliquer y associées aux valeurs $(x_1 \ldots, x_n)$ pour la variable explicative x.

1 (2 pts) Donner la loi a priori de Jeffreys $\pi^J(\theta)$ pour le paramètre θ.

2 (2 pts) Donner l'estimateur bayésien de θ associé à $\pi^J(\theta)$ et la fonction de perte quadratique.

3 (2 pts) Donner la famille de lois conjuguées pour le paramètre θ.

Exercice 4 (4 pts) Nous considérons un n-échantillon $x_1, \ldots, x_n$ distribué suivant la loi uniforme $[\theta, 2\theta]$ avec $\theta > 0$. Nous supposons que $\pi(\theta) \propto \mathbb{I}_{\theta>0}$.

1 (2 pts) Montrer que la loi a posteriori de θ est propre et l'expliciter. Pour le modèle considéré, nous savons que presque sûrement $2\min(x_1, \ldots, x_n) \geq \max(x_1, \ldots, x_n)$.

2 (2 pts) Donner un intervalle de crédibilité de niveau $1 - \alpha$ pour le paramètre θ.

①

Correction examen contrôle continu 8 octobre 2021 HAX912X

Exercice 1

1) $\underline{x} = (x_1, \dots, x_{30})$, $x_i \in \mathbb{N}_*$

$$f(\underline{x}|\theta) = \theta^n (1-\theta)^{\sum x_i - n}, \quad n = 30$$

$$\log f(\underline{x}|\theta) = n \log(\theta) + (\sum x_i - n) \log(1-\theta)$$

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

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

2

$$I_{\underline{x}}(\theta) = \frac{n}{\theta^2} + \frac{n/\theta - n}{(1-\theta)^2}$$

$$I_{\underline{x}}(\theta) = \frac{n}{\theta^2} + \frac{n}{\theta(1-\theta)} = \frac{n(1-\theta) + n\theta}{\theta^2(1-\theta)}$$

$$I_{\underline{x}}(\theta) = \frac{n}{\theta^2(1-\theta)}$$

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

C'est une loi a priori impropre.

2) $\pi^J(\theta|\underline{x}) \propto \theta^{n}(1-\theta)^{\sum x_i - n} \theta^{-1}(1-\theta)^{-\frac{1}{2}} \mathbb{1}_{]0,1[}(\theta)$

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

$n = 30$

$\theta|\underline{x} \sim \mathcal{B}(n, \sum x_i - n + 1/2)$

Exercice 2

③

1) $\underline{x} = (x_1, \dots, x_n)$

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

$\theta \in \Theta = \{1, 2\}$

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

$$\pi(\theta=2|\underline{x}) \propto \frac{1}{2^{n/2}} e^{-\frac{1}{4}\sum x_i^2} \left(\frac{1}{2}\right)$$

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

$$\pi(\theta=2|\underline{x}) = \frac{\frac{1}{2^{n/2}} e^{-\frac{1}{4}\sum x_i^2}}{e^{-\frac{1}{2}\sum x_i^2} + \frac{1}{2^{n/2}} e^{-\frac{1}{4}\sum x_i^2}}$$

La loi a posteriori est parfaitement définie : elle est propre.

2]

$$\mathbb{P}(\theta=1\,|\,\underline{X}) = \frac{1}{1+\frac{1}{2^{n/2}}\, e^{\frac{1}{4}\sum x_i^2}}$$

$$\frac{1}{2^{n/2}}\, e^{\frac{1}{4}\sum x_i^2} = e^{-\frac{n}{2}\log(2)+\frac{n}{4}\,\frac{\sum x_i^2}{n}}$$

$$= e^{-n\left[\frac{\log(2)}{2} - \frac{1}{4}\,\frac{\sum x_i^2}{n}\right]} = u_n$$

Si $\theta = 1$, alors

$$\frac{\sum x_i^2}{n} \xrightarrow[n\to\infty]{ps} 1$$

$$\Rightarrow u_n \xrightarrow[n\to\infty]{ps} 0 \quad \left(\frac{\log(2)}{2} - \frac{1}{4} > 0\right)$$

$$\Rightarrow \mathbb{P}(\theta=1\,|\,\underline{X}) \xrightarrow[n\to\infty]{ps} 1$$

Si $\theta =$, alors

$$\sum x_i^2 \xrightarrow{ps} 2$$

$$\Rightarrow u_n \xrightarrow[n\to\infty]{ps} \infty \quad \left(\frac{\log(2)}{2} - \frac{1}{2} < 0\right)$$

5

$$\Rightarrow \mathbb{P}(\theta = 1 \mid \underline{x}) \xrightarrow[n\to\infty]{PS} 0$$

$$\Rightarrow \mathbb{P}(\theta = 2 \mid \underline{x}) \xrightarrow[n\to\infty]{PS} 1$$

$\boxed{3}$

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

$$\hat{\theta}_n \in \arg\min_{d \in \{1,2\}} E^{\pi}(L(\theta, d) \mid \underline{x})$$

$$\hat{\theta}_n = \begin{cases} 1 & \text{si } \mathbb{P}(\theta = 1 \mid \underline{x}) \geq 2\,\mathbb{P}(\theta = 2 \mid \underline{x}) \\ 2 & \text{si } \mathbb{P}(\theta = 1 \mid \underline{x}) < 2\,\mathbb{P}(\theta = 2 \mid \underline{x}) \end{cases}$$

$$\hat{\theta}_n = \begin{cases} 1 & \text{si } \mathbb{P}(\theta = 1 \mid \underline{x}) \geq \frac{2}{3} \\ 2 & \text{si } \mathbb{P}(\theta = 1 \mid \underline{x}) < \frac{2}{3} \end{cases}$$

On peut éventuellement continuer le calcul.

6

Exercice 3

$y \mid \theta, x \sim N(\theta x, 1)$

1] $f(\underline{y} \mid \theta, \underline{x}) = \frac{1}{(2\pi)^{n/2}} \exp\left\{-\frac{1}{2}\sum_{i=1}^{n}(y_i - \theta x_i)^2\right\}$

$$\log(f(\underline{y} \mid \theta, \underline{x})) = -\frac{n}{2}\log(2\pi) - \frac{1}{2}\sum_{i=1}^{n}(y_i - \theta x_i)^2$$

$$\frac{d\log f}{d\theta}(\underline{y} \mid \theta, \underline{x}) = \sum_{i=1}^{n} x_i (y_i - \theta x_i)$$

$$\frac{d^2\log f}{(d\theta)^2}(\underline{y} \mid \theta, \underline{x}) = -\sum_{i=1}^{n} x_i^2$$

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

C'est une loi impropre.

2] $\pi^J(\theta \mid \underline{y}, \underline{x})$

$$\propto \exp\left\{-\frac{1}{2}\left[\theta^2 \sum x_i^2 - 2\theta \sum y_i x_i\right]\right\}$$

$\pi^J(\theta \mid \underline{y}, \underline{x})$

$$\propto \exp\left\{-\frac{\sum x_i^2}{2}\left(\theta^2 - 2\theta\frac{\sum y_i x_i}{\sum x_i^2}\right)\right\}$$

$$\propto \exp\left\{-\frac{\sum x_i^2}{2}\left(\theta - \frac{\sum y_i x_i}{\sum x_i^2}\right)^2\right\}$$

$$\Rightarrow \theta \mid \underline{y}, \underline{x} \sim \mathcal{N}\left(\frac{\sum y_i x_i}{\sum x_i^2}, \frac{1}{\sum x_i^2}\right)$$

$$\mathbb{E}^{\pi^J}(\theta \mid \underline{y}, \underline{x}) = \frac{\sum y_i x_i}{\sum x_i^2} = \hat{\theta}$$

$\hat{\theta}$ est l'estimateur de θ sous la fonction de perte quadratique.

3) Il s'agit de la famille des lois gaussiennes

$\theta \sim \mathcal{N}(\mu, \sigma^2)$

Exercice 4 $\quad n \geq 1$

1] $\min(x_1, \ldots, x_n) \geq \theta$

$\max(x_1, \ldots, x_n) \leq 2\theta$

$$\min(x_1, \ldots, x_n) \geq \theta \geq \frac{1}{2}\max(x_1, \ldots, x_n)$$

$$f(\underline{x}|\theta) = \prod_{i=1}^{n} \mathbb{1}_{[\theta, 2\theta]}(x_i) \left(\frac{1}{\theta}\right)$$

$$f(\underline{x}|\theta) = \left(\frac{1}{\theta^n}\right) \mathbb{1}_{[0.5\max(\underline{x}), \min(\underline{x})]}(\theta)$$

$$\pi(\theta|\underline{x}) \propto \frac{1}{\theta^n} \mathbb{1}_{[0.5\max(\underline{x}), \min(\underline{x})]}(\theta)$$

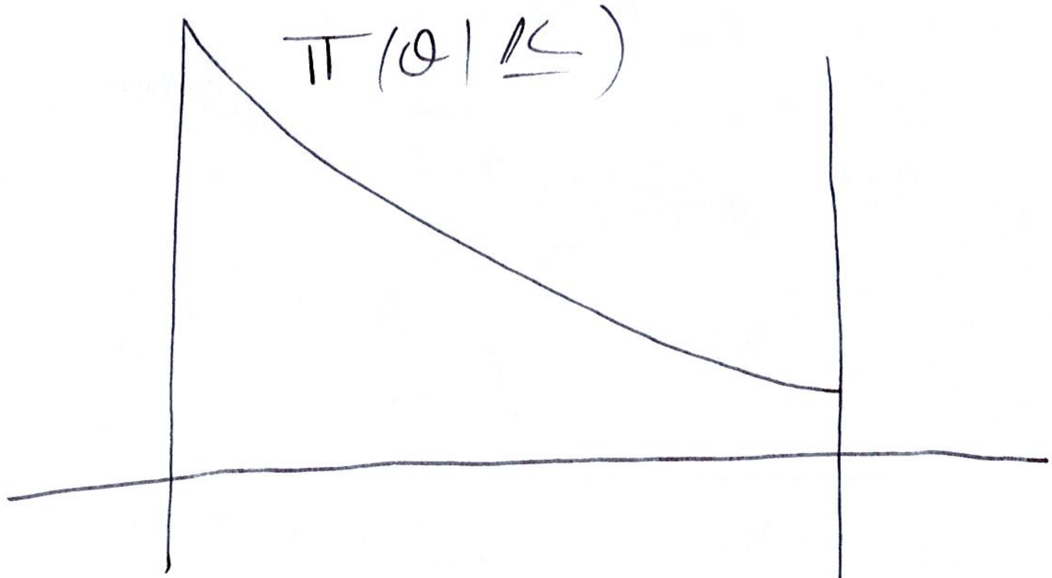
$$\int_{0.5\max(\underline{x})}^{\min(\underline{x})} \theta^{-n} d\theta = \left[\frac{\theta^{-n+1}}{-n+1}\right]_{0.5\max(\underline{x})}^{\min(\underline{x})}$$

(9)

$$\int_{0.5\,\max(\underline{x})}^{\min(\underline{x})} \theta^{-n}\,d\theta = \left[\frac{(0.5\max(\underline{x}))^{-n+1} - \min(\underline{x})^{-n+1}}{n-1}\right] = c(\underline{x})$$

La loi a posteriori est propre.

2) $$\pi(\theta \mid \underline{x}) = \frac{1}{\theta^{n}\, c(\underline{x})}\, \mathbb{1}_{[0.5\max(\underline{x}),\ \min(\underline{x})]}(\theta)$$



$$HPD = \{\theta : 0.5\max(\underline{x}) \le \theta \le k_\alpha\}$$

où k_α est tel que

$$\mathbb{P}(0.5\max(\underline{x}) \le \theta \le k_\alpha) = 1-\alpha$$

on peut faire le calcul.

$$\int_{0.5\max(\underline{x})}^{k_\alpha} \pi(\theta|\underline{x})\, d\theta = 1-\alpha$$

$$\int_{0.5\max(\underline{x})}^{k_\alpha} \frac{1}{\theta^n c(\underline{x})}\, d\theta$$

$$= \frac{(n-1)\left[\theta^{-n+1}/-n+1\right]_{0.5\max(\underline{x})}^{k_\alpha}}{(0.5\max(\underline{x}))^{-n+1} - \min(\underline{x})^{-n+1}}$$

$$= \frac{(0.5\max(\underline{x}))^{-n+1} - k_\alpha^{-n+1}}{\underbrace{(0.5\max(\underline{x}))^{-n+1} - \min(\underline{x})^{-n+1}}_{d(\underline{x})}} = 1-\alpha$$

$$\Rightarrow k_\alpha = \left[-(1-\alpha)\,d(\underline{x}) + (0.5\max(\underline{x}))^{-n+1}\right]^{\frac{1}{-n+1}}$$