

M1 - Statistiques bayésiennes

Examen, le 10/05/2016

Durée 3h. Les documents ne sont pas autorisés.

On rappelle que, si l'on note $\Gamma(a) = \int_0^\infty x^{a-1} e^{-x} dx$ pour $a > 0$,

- la densité de $Y \sim \text{Gamma}(a, b)$ est $\frac{b^a x^{a-1} e^{-bx}}{\Gamma(a)} \mathbb{1}_{x>0}$ et $E[Y] = \frac{a}{b}$, $\text{Var}[Y] = \frac{a}{b^2}$
- la densité d'une loi inverse gamma $\text{IG}(a, b)$ est $\frac{b^a}{\Gamma(a)} x^{-a-1} e^{-\frac{b}{x}} \mathbb{1}_{x>0}$.

Exercice 1

A. On s'intéresse à l'estimation de μ dans le cadre bayésien

$$\begin{aligned} X_1, \dots, X_n \mid \mu &\sim \mathcal{N}(\mu, 1)^{\otimes n} \\ \mu &\sim \mathcal{N}(0, 1) = \Pi. \end{aligned}$$

1. Déterminer la loi a posteriori $\Pi[\cdot \mid X_1, \dots, X_n]$, sa moyenne μ_X et sa variance σ_X^2 .
2. Construire un intervalle J_X de crédibilité $1 - \alpha$ pour la loi a posteriori. Un tel intervalle est-il unique? Votre intervalle J_X est-il de taille minimale?

B. On s'intéresse à l'estimation de σ^2 dans le cadre bayésien

$$\begin{aligned} X_1, \dots, X_n \mid \sigma^2 &\sim \mathcal{N}(0, \sigma^2)^{\otimes n} \\ \sigma^2 &\sim \Pi, \end{aligned}$$

avec $\Pi = \text{IG}(a, b)$ la loi inverse gamma définie dans le préambule.

1. Justifier mathématiquement l'appellation *inverse gamma*.
2. Déterminer la loi a posteriori $\Pi[\cdot \mid X_1, \dots, X_n]$.

C. Pour $\Theta = \mathbb{R}$, on souhaite faire un test de deux hypothèses $H_0 = \{\theta : \theta \in \Theta_0\}$ et $H_1 = \{\theta : \theta \in \Theta_1\}$, avec Θ_0, Θ_1 des sous-ensembles de Θ disjoints et $\Theta = \Theta_0 \cup \Theta_1$. On observe $X = (X_1, \dots, X_n)$ suivant le modèle $\{\mathcal{N}(\theta, 1)^{\otimes n}, \theta \in \mathbb{R}\}$. On définit la fonction de perte équilibrée

$$L(\theta, \varphi) = \mathbb{1}_{\theta \in \Theta_0} \mathbb{1}_{\varphi=1} + \mathbb{1}_{\theta \in \Theta_1} \mathbb{1}_{\varphi=0}.$$

On rappelle que le test bayésien associé pour une loi a priori Π sur Θ est l'estimateur de Bayes pour L et Π .

1. Justifier le choix de L dans le but de faire un test de H_0 contre H_1 .
2. Démontrer que le test bayésien s'écrit $\varphi(X) = \mathbb{1}_{\Pi(\Theta_0 \mid X) \leq \Pi(\Theta_1 \mid X)}$.
3. Déterminer le test bayésien pour $H_0 = \{\theta > 0\}$, $H_1 = \{\theta \leq 0\}$ et $\Pi = \mathcal{N}(0, 1)$.

4. On veut tester $H_0 = \{\theta = 0\}$ contre $H_1 = \{\theta \neq 0\}$. Proposer une modification de la loi a priori Π pour faire ce test [seulement si vous avez le temps en fin d'examen, déterminer le test correspondant].

Exercice 2

On souhaite estimer $\theta \in \Theta = (0, +\infty)$ dans le cadre bayésien

$$\begin{aligned} X_1, \dots, X_n | \theta &\sim \text{Unif}[0, \theta]^{\otimes n} \\ \theta &\sim \mathcal{P}(a, b), \end{aligned}$$

où la loi de Pareto $\mathcal{P}(a, b)$ de paramètres $a > 0, b > 0$ est la loi de densité sur \mathbb{R}

$$f_{a,b}(\theta) = \frac{ab^a}{\theta^{a+1}} \mathbb{1}_{\theta > b}.$$

On rappelle que la loi uniforme sur $[0, \theta]$ a pour densité $x \rightarrow \frac{1}{\theta} \mathbb{1}_{[0, \theta]}(x)$.

1. Ecrire le modèle statistique. Quelle est la loi a priori Π ?
2. Montrer que la loi a posteriori correspondante est une loi $\mathcal{P}(n + a, Y_n)$, où $Y_n = \max\{X_1, \dots, X_n, b\}$.
3. Quel est l'estimateur de Bayes $T_1(X_1, \dots, X_n)$ pour la fonction de perte $L_1(\theta, T) = (\theta - T)^2$ et une loi a priori Π quelconque ? (on ne demande pas de le redémontrer). Donner l'expression explicite de T_1 pour $\Pi = \mathcal{P}(a, b)$.
4. On définit la fonction de perte

$$L_2(\theta, T) = \left(\frac{T}{\theta} - 1 \right)^2.$$

Déterminer un estimateur de Bayes T_2 pour cette fonction de perte et une loi a priori Π quelconque. Donner son expression lorsque $\Pi = \mathcal{P}(a, b)$.

5. Rappeler la définition de la consistance au sens fréquentiste de la loi a posteriori $\Pi[\cdot | X_1, \dots, X_n]$ au point $\theta_0 \in \Theta$. La loi a posteriori peut-elle ici être consistante en tout point $\theta_0 \in \Theta$?
6. On suppose $b = 1$. Montrer que si $\theta_0 > 1$, la variable Y_n définie ci-dessus converge en probabilité sous P_{θ_0} vers θ_0 . En déduire qu'il en est de même pour les estimateurs T_1 et T_2 ci-dessus.
7. Montrer que la variance a posteriori vérifie, quand $n \rightarrow \infty$,

$$\text{Var}[\theta | X_1, \dots, X_n] = O(n^{-2}) Y_n^2.$$

Construire un intervalle de crédibilité au moins $1 - \alpha$ centré en la moyenne a posteriori. Comment varie son diamètre en fonction de n ? Quelle est, selon vous, la vitesse de convergence de la loi a posteriori ? (on ne demande pas de démontrer ce dernier point)

Exercice 3

Soit le modèle $\mathcal{P} = \{P_\theta = \text{Gamma}(c, c\theta), \theta > 0\}$ pour $c > 0$ une constante donnée.

1. Proposer une famille de lois a priori conjuguée pour l'estimation de θ et calculer la loi a posteriori correspondante à une de ces lois a priori pour n observations X_1, \dots, X_n i.i.d. de loi P_θ sachant θ .

On rappelle que si $dP_\theta(x) = p_\theta(x)dx$ et $\ell_\theta(X) = \log p_\theta(X)$, alors l'information de Fisher est définie par

$$I(\theta) = E_\theta \left[\left(\frac{\partial}{\partial \theta} \ell_\theta(X) \right)^2 \right].$$

On rappelle qu'en dimension 1 l'a priori de Jeffreys est proportionnel (ou égal) à $I(\theta)^{1/2}$.

2. Quel est l'intérêt de l'a priori de Jeffreys ?
3. Déterminer l'a priori de Jeffreys pour le modèle \mathcal{P} ci-dessus.
4. On veut évaluer numériquement l'intégrale suivante, pour $n \geq 1$ fixé,

$$I = I_n = \int_{\mathbb{R}} h(u) \frac{f_n(u)}{C_n} du,$$

avec h et f_n des fonctions positives simples à évaluer et $C_n = \int_{\mathbb{R}} f_n(u) du > 0$. Pour le calcul de I , on voudrait éviter de devoir calculer la constante de normalisation C_n . Soit Z_1, \dots, Z_N i.i.d. de loi de densité g sur \mathbb{R} . On suppose qu'il est facile de produire un tel échantillon par simulation. On pose pour tout réel u ,

$$\zeta(u) = \frac{f_n(u)}{g}.$$

- (a) Montrer, sous une condition d'intégrabilité à préciser, que

$$\frac{1}{N} \sum_{i=1}^N h(Z_i) \zeta(Z_i)$$

converge en probabilité (ou p.s.) quand $N \rightarrow \infty$ et déterminer sa limite. Répondre aux mêmes questions pour $\frac{1}{N} \sum_{i=1}^N \zeta(Z_i)$.

- (b) En déduire une quantité \hat{I}_N qui converge en probabilité vers I quand $N \rightarrow \infty$, et dont l'évaluation ne fait pas appel à C_n .
5. On souhaite appliquer le résultat de la question 5 au modèle \mathcal{P} et à la loi a priori $\Pi = \text{Gamma}(1, 1)$, afin de calculer

$$I = \int \theta^{1/2} d\Pi(\theta | X_1, \dots, X_n)$$

sans devoir évaluer la fonction Γ . Identifier une fonction f_n et proposer une densité g simple sur \mathbb{R}^+ pour appliquer la méthode en question. Quel est le nom de ce type de méthode ? Quel type de fonction g est-il préférable de prendre ?