Université de Montpellier Année 2016-2017
M1 BIOSTATISTIQUE & MIND - HMMA101

Examen partiel - 9 novembre 2016
Durée 1h - Documents interdits

Exercice 1 (10 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 x de naissances masculines parmi n naissances à Paris, peut-on dire si la probabilité θ qu'un enfant qui naisse soit un garçon est supérieure à $1/2$? Il est clair que $x|\theta \sim \mathcal{B}(n, \theta)$.

1) (3 pts) Laplace munit le paramètre θ d'une loi uniforme sur $[0, 1]$. Donner la loi a posteriori de θ. Ecrire la probabilité $\mathbb{P}(\theta > 1/2|x)$ sous forme d'intégrale.

2) (2 pts) Calculer l'espérance et la variance de la loi a posteriori calculée à la question 1). Exprimer leur limite presque sûre lorsque $n \to \infty$ (on pourra utiliser la loi des grands nombres).

3) (2 pts) Expliciter la loi a priori de Jeffreys.

4) (3 pts) Donner la loi a posteriori de θ pour la loi a priori de Jeffreys, calculer son espérance, sa variance et exprimer leur limite presque sûre lorsque $n \to \infty$ (on pourra utiliser la loi des grands nombres).

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 (10 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) (4 pts) Donner la loi a posteriori de θ (indication : la loi de Pareto est conjuguée pour la distribution uniforme).

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

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

3) (4 pts) Calculer l'estimateur bayésien associé à la fonction de perte

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

Correction examen de contrôle continu HMMA 101 - 09/11/2016

(1)

Exercice 1

1) $f(x|\theta) = C_n^x \theta^x (1-\theta)^{n-x} \mathbb{1}_{\{0,\dots,n\}}(x)$

$$\pi(\theta) = \mathbb{1}_{[0,1]}(\theta)$$

$$\pi(\theta|x) \propto \theta^x (1-\theta)^{n-x}$$

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

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

2) $\mathbb{E}(\theta|x) = \left[\frac{x+1}{n+2} \right]$

$$\mathbb{V}(\theta|x) = \frac{(x+1)(n-x+1)}{(n+2)^2(n+3)}$$

$$K = \sum_{i=1}^{n} y_i$$

avec $y_1, \ldots, y_n$
n-échantillon suivant une loi de Bernoulli de paramètre $\theta^* > 0$.
θ^* est la vraie valeur de θ

$$\mathbb{E}(\theta \mid K) = \frac{\sum_{i=1}^{n} y_i}{n} \times \left(\frac{n}{n+1}\right) \xrightarrow[n \to \infty]{PS} \theta^*$$

(loi forte des grands nombres)

$$\mathbb{V}(\theta \mid K) = \frac{\left(\sum_{i=1}^{n} y_i + 1\right)\left(n - \sum_{i=1}^{n} y_i + 1\right)}{(n+2)^2 (n+3)}$$

$$= \left(\frac{\sum_{i=1}^{n} y_i}{n} + \frac{1}{n}\right)\left(1 - \frac{\sum_{i=1}^{n} y_i}{n} + \frac{1}{n}\right)\left[\frac{n^2}{(n+2)^2}\right]\left[\frac{1}{(n+3)}\right]$$

$$\xrightarrow[n \to \infty]{PS} 0$$

3)

3

$$I_x(\theta) = -\mathbb{E}\left[\frac{d^2}{(d\theta)^2}\left[\log(C_n^x) + x\log(\theta) + (n-x)\log(1-\theta)\right]\right]$$

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

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

$$\Rightarrow \theta \sim Beta\left(\frac{1}{2}, \frac{1}{2}\right)$$

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

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

$$\mathbb{E}(\theta|x) = \frac{x+\frac{1}{2}}{n+1}$$

$$\mathbb{V}(\theta|x) = \frac{(x+\frac{1}{2})(n-x+\frac{1}{2})}{(n+1)^2(n+2)}$$

4

$$\mathbb{E}(\theta|\mathbf{1}) = \frac{\sum_{i=1}^{n} y_i + \frac{1}{2}}{n+1}$$

$$= \left(\frac{\sum y_i}{n} + \frac{1}{2n}\right)\left(\frac{n}{n+1}\right) \xrightarrow[n\to\infty]{PS} \theta^*$$

d'après la loi des grands nombres

$$\mathbb{V}(\theta|\mathbf{1}) = \left[\frac{\sum y_i}{n} + \frac{1}{2n}\right]\left[1 - \frac{\sum y_i}{n} + \frac{1}{2n}\right]$$

$$\left[\frac{n^2}{(n+1)^2}\right]\left[\frac{1}{n+2}\right] \xrightarrow[n\to\infty]{PS} 0$$

Exercice 2

1)

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

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

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

2) On sait que l'estimateur bayésien associé à la fonction de perte quadratique est la moyenne de la loi a posteriori.

$$\widehat{\theta}_1 = E(\theta/x)$$

$$= (\alpha+1)\int_{\max(x,\theta_0)}^{+\infty} (\max(x,\theta_0))^{\alpha+1}\,\theta^{-\alpha-1}\,d\theta$$

$$= (\alpha+1)\left[\max(x,\theta_0)\right]^{\alpha+1}\left[-\frac{\theta^{-\alpha}}{\alpha}\right]_{\max(x,\theta_0)}^{+\infty}$$

$$= \left[\max(x,\theta_0)\right]\left(\frac{\alpha+1}{\alpha}\right)$$

3]

$$h(\delta) = \int L_2(\theta, \delta) \; \pi(\theta | x) \, d\theta$$

$$= \int \theta^{-2}(\theta^2 - 2\theta\delta + \delta^2) \pi(\theta|x) \, d\theta$$

$$= 1 \quad - 2\delta \int \theta^{-} \; \pi(\theta|x) \, d\theta$$

$$+ \delta^2 \int \theta^{-2} \pi(\theta|x) \, d\theta$$

$$\frac{dh}{d\delta}(\delta^*) = 0$$

$$\Leftrightarrow \quad \delta^* = \frac{\int \theta^{-1} \pi(\theta|x) \, d\theta}{\int \theta^{-2} \pi(\theta|x) \, d\theta}$$

$$\frac{d^2 h}{(d\delta)^2}(\delta) = 2 \int \theta^{-2} \pi(\theta|x) \, d\theta > 0$$

Ainsi

$$\hat{\theta}_2 = \frac{\int \theta^{-1} \pi(\theta | x)\, d\theta}{\int \theta^{-2} \pi(\theta | x)\, d\theta}$$

$$= \left[\frac{\int_{\max(x,\theta_0)}^{+\infty} \theta^{-\alpha-3}\, d\theta}{\int_{\max(x,\theta_0)}^{+\infty} \theta^{-\alpha-4}\, d\theta} \right]$$

$$= \left(\frac{\alpha+3}{\alpha+2} \right) \max(x, \theta_0).$$