Théorie de l'information et de la décision
Partie 2, bayésienne (10 pts)
Documents interdits (sauf formulaire de probabilités)

Une variable aléatoire u distribuée suivant une loi Inverse Gamma de paramètres $\alpha > 0$ et $\beta > 0$ admet pour densité de probabilité

$$f(u) = \frac{\beta^\alpha}{\Gamma(\alpha)} u^{-\alpha-1} \exp\left(-\frac{\beta}{u}\right) \mathbb{I}_{u>0}\,.$$

Nous admettons que $\mathbb{E}(u) = \dfrac{\beta}{\alpha - 1}$ et $\mathbb{V}(u) = \dfrac{\beta}{(\alpha-1)^2(\alpha-2)}$.

Exercice 1 (6 pts) Nous modélisons la distance y à laquelle se trouve une particule de son point de départ (après un temps fixé) par une loi de Rayleigh de paramètre $\theta > 0$

$$f(y|\theta) = \frac{y}{\theta} \exp\left(-\frac{y^2}{2\theta}\right) \mathbb{I}_{y\geq 0}\,.$$

Nous admettons que $\mathbb{E}(y|\theta) = \sqrt{\dfrac{\theta\pi}{2}}$ et $\mathbb{E}(y^2|\theta) = 2\theta$.
Nous observons un n-échantillon $y_1, \ldots, y_n$ et souhaitons estimer θ. Nous nous plaçons dans le paradigme bayésien et supposons que θ est une variable aléatoire.

1) (1 pts) Montrer que la famille des lois Inverse Gamma est conjuguée.

2) (2 pts) Donner la loi a priori de Jeffreys pour θ. Est-elle propre ?

3) (3 pts) Pour la loi a priori de Jeffreys, donner l'estimateur associé à la fonction de perte $L(\theta, d) = (\theta - d)^2$ et celui du Maximum A Posteriori (MAP). Comparer.

Exercice 2 (4 pts) La durée de vie y d'ampoules d'un certain stock est modélisée par une loi exponentielle de paramètre $\left(\dfrac{1}{\theta}\right)$; $\mathbb{E}(y|\theta) = \theta > 0$. Nous tirons au hasard 100 ampoules dans le stock et nous mesurons leurs durées de vie, nous observons $\sum_{i=1}^{100} y_i = 450$. Le stock étant de grande taille, nous supposons que les durées de vie sont indépendantes.

Un expert indique que, selon lui, la durée de vie moyenne des ampoules est de 5 années avec un écart-type d'une année. Proposer un estimateur bayésien pour le paramètre θ tenant compte de cette information a priori.

1

Correction examen final
HΠΠA 101
Janvier 2018

Exercice 1

1] $\pi(\theta) \propto \theta^{-\alpha-1} e^{-\beta/\theta} \mathbb{1}_{\{\theta > 0\}}$

$f(y|\theta) = \left(\frac{y}{\theta}\right) e^{-y^2/2\theta} \mathbb{1}_{\{y \geq 0\}}$

$\pi(\theta|y) \propto \theta^{-\alpha-2} e^{-\frac{1}{\theta}\left(\beta + \frac{y^2}{2}\right)} \mathbb{1}_{\{\theta > 0\}}$

$\Rightarrow \theta | y \sim \mathcal{IG}\left(\alpha+1, \beta + \frac{y^2}{2}\right)$

La famille des lois Inverse Gamma est conjuguée.

2] $\log(f(y|\theta)) = \log(y) - \log(\theta) - \frac{y^2}{2\theta}$

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

$$\frac{d^2 \log f}{d\theta^2}(y|\theta) = \frac{1}{\theta^2} - \frac{y^2}{2}\frac{2\theta}{\theta^4}$$

$$\frac{d^2 \log f}{d\theta^2}(y|\theta) = \frac{1}{\theta^2} - \frac{y^2}{\theta^3}$$

$$I(\theta) = -\frac{1}{\theta^2} + \frac{2\theta}{\theta^3}$$

$$\Rightarrow I(\theta) = \left(\frac{1}{\theta^2}\right)$$

Ainsi, $I_n(\theta) = \frac{n}{\theta^2}$

$$\Rightarrow \pi^J(\theta) \propto \left(\theta^{-2}\right)^{1/2}$$

$$\Rightarrow \pi^J(\theta) \propto \theta^{-1}$$

Cette loi a priori est impropre car $\int_0^{+\infty} \theta^{-1} d\theta$ est divergente.

3

3] Soit $\underline{y} = (y_1, ----, y_n)$

$$\pi^J(\theta|\underline{y}) \propto \left(\frac{1}{\theta^n}\right) e^{-\frac{1}{2\theta}\sum y_i^2} \theta^{-1} \mathbb{1}_{\{\theta>0\}}$$

$$\Rightarrow \pi^J(\theta|\underline{y}) \propto \theta^{-n-1} e^{-\frac{1}{\theta}\left(\frac{\sum y_i^2}{2}\right)}$$

$$\Rightarrow \theta|\underline{y} \sim I\Gamma\left(n, \frac{\sum y_i^2}{2}\right)$$

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

$$\hat{\theta}^1 = \frac{\sum y_i^2}{2(n-1)}$$

Par ailleurs, $\hat{\theta}^2 \in \arg\max_{\theta>0} \pi^J(\theta|\underline{y})$

$$\Rightarrow \hat{\theta}^2 \in \arg\max_{\theta>0} \log(\pi^J(\theta|\underline{y}))$$

$$\Rightarrow \hat{\theta}^2 \in \arg\max_{\theta>0} \Big\{ \underbrace{-(n+1)\log(\theta) - \frac{1}{2\theta}\sum_{i=1}^n y_i^2}_{h(\theta)} \Big\}$$

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

$$\iff -\frac{(n+1)}{\theta^*} + \frac{1}{2(\theta^*)^2}\sum y_i^2 = 0$$

$$\Rightarrow (n+1)\theta^* = \frac{\sum y_i^2}{2}$$

$$\Rightarrow \theta^* = \frac{\sum y_i^2}{2(n+1)}$$

$$\frac{d^2h}{(d\theta)^2}(\theta) = \frac{n+1}{\theta^2} - \frac{\sum y_i^2}{\theta^3}$$

$$\Rightarrow \frac{d^2h}{(d\theta)^2}(\theta^*) = \frac{4(n+1)^2}{(\sum y_i^2)^2} - \frac{8(n+1)^3}{(\sum y_i^2)^2} < 0$$

Ainsi $\widehat{\theta^2} = \left[\frac{\sum y_i^2}{2(n+1)}\right]$

Exercice 2

$y_1, \dots, y_n \overset{iid}{\sim} Exp\left(\frac{1}{\theta}\right)$

Pour tenir compte de l'information a priori donnée par l'expert sur le paramètre θ nous utilisons une loi Inverse Gamma (α, β) telle que

$$\frac{\beta}{(\alpha-1)} = 5 \text{ et } \frac{\beta}{(\alpha-1)^2(\alpha-2)} = 1$$

$$\Rightarrow \begin{cases} \dfrac{\beta}{\alpha-1} = 5 \\ \dfrac{1}{(\alpha-1)(\alpha-2)} = \dfrac{1}{5} \end{cases}$$

$$\Rightarrow \begin{cases} (\alpha-1)(\alpha-2) = 5 \\ \dfrac{\beta}{\alpha-1} = 5 \end{cases}$$

$$\Rightarrow \begin{cases} \alpha^2 - 2\alpha - \alpha + 2 = 5 \\ \beta = 5(\alpha-1) \end{cases}$$

$$\Rightarrow \begin{cases} \alpha^2 - 3\alpha - 3 = 0 \\ \beta = 5(\alpha-1) \end{cases}$$

$$\Rightarrow \begin{cases} \alpha = \dfrac{3+\sqrt{21}}{2} \\ \beta = \dfrac{5(1+\sqrt{21})}{2} \end{cases}$$

6

$$\text{Si } \theta \sim \mathcal{I}\Gamma\left(\frac{3+\sqrt{21}}{2}, \frac{5(1+\sqrt{21})}{2}\right)$$

$$\text{alors } \pi(\theta|\underline{y}) \propto \theta^{-\frac{3+\sqrt{21}}{2}-1} e^{-\frac{5(1+\sqrt{21})}{2\theta}}$$

$$\theta^{-n} e^{-\sum y_i/\theta} \mathbb{1}_{\{\theta>0\}}$$

$$\Rightarrow \pi(\theta|\underline{y}) \propto \theta^{-\frac{3+\sqrt{21}}{2}-n-1} \exp\left(-\frac{1}{\theta}\left(\frac{5(1+\sqrt{21})}{2}+\sum y_i\right)\right)$$

$$\Rightarrow \theta|\underline{y} \sim \mathcal{I}\Gamma\left(n+\frac{3+\sqrt{21}}{2}, \frac{5(1+\sqrt{21})}{2}+\sum y_i\right)\mathbb{1}_{\{\theta>0\}}$$

$$\hat{\theta} = \mathbb{E}^{\pi}(\theta|\underline{y}) = \left[\frac{\frac{5(1+\sqrt{21})}{2}+\sum y_i}{n+\frac{3+\sqrt{21}}{2}-1}\right]$$

Nous avons $n = 100$ et $\sum_{i=1}^{n} y_i = 450$

Ainsi $\hat{\theta} = \left[\frac{905+5\sqrt{21}}{201+\sqrt{21}}\right] \simeq 4,5$