Théorie de l'information et de la décision
Partie 2, bayésienne (10 pts)

Exercice 1 (6 pts)

Le nombre de requêtes r arrivant sur un serveur suit une loi de Poisson de paramètre $\theta > 0$ et de densité

$$f(r|\theta) = \frac{\theta^r}{r!}\exp(-\theta)\mathbf{1}_{r\in\mathbb{N}}\,.$$

On se place dans le paradigme bayésien et on suppose que θ est une variable aléatoire.

1) (2 pts) Donner la loi a priori de Jeffreys pour θ. Est-elle propre ? Discuter de l'existence de la loi a posteriori en fonction de la valeur observée de r.

2) (2 pts) Montrer que la famille des lois Gamma est conjuguée. On rappelle que si θ suit une loi Gamma de paramètres $a > 0$ et $b > 0$, elle admet comme densité

$$\pi(\theta) = \frac{b^a}{\Gamma(a)}\theta^{a-1}\exp\left(-b\theta\right)\mathbf{1}_{\theta>0}\,,$$

$\mathbb{E}(\theta) = a/b$ et $\mathbb{V}(\theta) = a/b^2$

3) (2 pts) Pour une loi a priori Gamma de paramètres $a = 1$ et $b = 1$ sur θ et $r = 10$, donner les estimateurs bayésiens associés aux fonctions de perte

$$L_1(\theta, d) = (\theta - d)^2 \text{ et } L_2(\theta, d) = \frac{(\theta - d)^2}{\theta^2}.$$

Exercice 2 (4 pts)

Lucas et Romain jouent aux boules et s'entrainent aux tirs. Romain prétend pouvoir toucher à 10 mètres 2 fois sur 3. Lucas pense lui que Romain ne touche qu'une boule sur trois. Romain effectue 10 essais et touche cinq boules.

1) (3 pts) Modéliser le problème et donner la réponse bayésienne. Expliquer en quoi le résultat obtenu est extrêmement intuitif.

2) (1 pt) Même question que précédemment dans le cas où Romain touche 4 boules sur 10.

Correction examen
HMMA 101 25/11/2016

Exercice 1

1) $$\log\left[f(n|\theta)\right] = -\log(n!) + n\log(\theta) - \theta$$

égalité presque sûre pour θ fixé et n variable aléatoire.

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

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

Ainsi, $$I_n(\theta) = \frac{\mathbb{E}(n)}{\theta^2} = \left(\frac{1}{\theta}\right)$$

$$\Longrightarrow \pi^J(\theta) \propto \theta^{-1/2}\,\mathbb{1}\{\theta > 0\}$$

$\pi^J(\theta)$ est une loi impropre (2)

en effet $\int_0^{+\infty} \theta^{-1/2} \, d\theta$ n'est pas finie.

Nous avons

$$\pi^J(\theta|n) \propto \pi^J(\theta) f(n|\theta)$$

$$\pi^J(\theta|n) \propto \theta^{-1/2} \theta^n e^{-\theta} \mathbb{1}_{\{\theta > 0\}}$$

$$\pi^J(\theta|n) \propto \theta^{n-1/2} e^{-\theta} \mathbb{1}_{\{\theta > 0\}}$$

$$\Rightarrow \theta|n \sim \Gamma\text{amma}(n+\tfrac{1}{2}, 1)$$

La loi a posteriori existe quelle que soit la valeur où $n \in \mathbb{N}$.

2] $$\pi(\theta|n) \propto \theta^{a-1} e^{-b\theta} \theta^n e^{-\theta} \mathbb{1}_{\{\theta > 0\}}$$

$$\Rightarrow \theta|n \sim \Gamma\text{amma}(a+n, b+1)$$

La famille des lois Gamma est ③
donc conjuguée pour le modèle de Poisson.

3] $\pi(\theta|n) \propto e^{-\theta}\theta^{n}e^{-\theta}\mathbb{1}_{\{\theta>0\}}$

$\theta|n \sim \text{Gamma}(n+1, 2)$

Pour la fonction de perte
$L_1(\theta, d) = (\theta - d)^2$, l'estimateur
bayesien est la moyenne de loi
a posteriori et donc

$$\hat{\theta}^1 = \mathbb{E}^{\pi}(\theta|n) = \frac{n+1}{2} = 5{,}5$$

Pour la fonction de perte
$L_2(\theta, d) = \frac{(\theta - d)^2}{\theta^2}$, nous devons
minimiser en d

$$\mathbb{E}^{\pi}(L_2(\theta, d) \mid n).$$

$$\mathbb{E}^{\pi}\left[L_2(\theta,d)\mid x\right]$$

$$= \int \frac{(\theta-d)^2}{\theta^2}\,\pi(\theta\mid x)\,d\theta$$

$$= 1 - 2d\int \theta^{-1}\pi(\theta\mid x)\,d\theta$$

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

Soit d^* tel que $\frac{dh}{dd}(d^*) = 0$

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

Par ailleurs, $\frac{d^2h}{(dd)^2}(d) = 2\int \theta^{-2}\,\pi(\theta\mid x)\,d\theta > 0$

Ainsi,
$$\widehat{\theta^2} = \frac{\int \theta^{-1}\,\pi(\theta\mid x)\,d\theta}{\int \theta^{-2}\,\pi(\theta\mid x)\,d\theta}$$

5

$$\int_0^{+\infty} \theta^{-1} \frac{2^{n+1}}{\Gamma(n+1)} \theta^{n} e^{-2\theta} d\theta$$

$$= \frac{2^{n+1}}{\Gamma(n+1)} \int_0^{+\infty} \theta^{n-1} e^{-2\theta} d\theta$$

$$= \frac{2^{n+1}}{\Gamma(n+1)} \frac{\Gamma(n)}{2^{n}} = \frac{2}{n}$$

$$\int_0^{+\infty} \theta^{-2} \frac{2^{n+1}}{\Gamma(n+1)} \theta^{n} e^{-2\theta} d\theta$$

$$= \frac{2^{n+1}}{\Gamma(n+1)} \frac{\Gamma(n-1)}{2^{n-1}} = \frac{4}{n(n-1)}$$

Ainsi,

$$\widehat{\theta^2} = \frac{n-1}{2} = 4,5$$

Exercice 2

Soit x le nombre de boules touchées par Romain et θ la probabilité que Romain touche une boule à 10 mètres.

1)

$$x \mid \theta \sim \mathcal{B}(10, \theta)$$

$$\theta \in \Theta = \left\{ \frac{1}{3}, \frac{2}{3} \right\}$$

$$\mathbb{P}\left(\theta = \frac{1}{3}\right) = \mathbb{P}\left(\theta = \frac{2}{3}\right) = \frac{1}{2}$$

$$\mathbb{P}\left(\theta = \frac{1}{3} \mid x\right) \propto \left(\frac{1}{2}\right)\left(\frac{1}{3}\right)^{x}\left(\frac{2}{3}\right)^{n-x}$$

$$\mathbb{P}\left(\theta = \frac{2}{3} \mid x\right) \propto \left(\frac{1}{2}\right)\left(\frac{2}{3}\right)^{x}\left(\frac{1}{3}\right)^{n-x}$$

Si $x = 5$, alors

$$\mathbb{P}\left(\theta = \frac{1}{3} \mid x\right) = \mathbb{P}\left(\theta = \frac{2}{3} \mid x\right) = \frac{1}{2}$$

C'est tout à fait intuitif car $\frac{5}{10} = \frac{1}{2}$ et équivalent au $\frac{1}{3}$ et $\frac{2}{3}$.
On ne peut pas décider.

2) Si $x = 4$,

$$\mathbb{P}\left(\theta = \frac{1}{3} \middle| x\right) \propto \left(\frac{1}{2}\right)\left(\frac{1}{3}\right)^4\left(\frac{2}{3}\right)^6$$

$$\mathbb{P}\left(\theta = \frac{2}{3} \middle| x\right) \propto \left(\frac{1}{2}\right)\left(\frac{1}{3}\right)^6\left(\frac{2}{3}\right)^4$$

$$\Rightarrow \mathbb{P}\left(\theta = \frac{1}{3} \middle| x\right) = \frac{\left(\frac{2}{3}\right)^2}{\left(\frac{2}{3}\right)^2 + \left(\frac{1}{3}\right)^2}$$

$$\mathbb{P}\left(\theta = \frac{2}{3} \middle| x\right) = \frac{\left(\frac{1}{3}\right)^2}{\left(\frac{2}{3}\right)^2 + \left(\frac{1}{3}\right)^2}$$

$$\mathbb{P}\left(\theta = \frac{1}{3} \middle| x\right) = \frac{4}{5} \text{ et } \mathbb{P}\left(\theta = \frac{2}{3} \middle| x\right) = \frac{1}{5}$$