Corrigé : Feuille de travaux dirigés 4

Solution Exercice 1

Dans cet exercice, $\mu_0, \sigma^2, \tau_0^2$ sont des « hyper-paramètres » et supposés connus à l'avance.

1. L'espace des paramètres est : $\Theta = \mathbb{R}$.

Pour déterminer la loi a posteriori, on utilise la formule de la distribution a posteriori : $\pi(\theta|X=x) = \frac{\pi(\theta)p_{\theta}(x)}{m^X(x)}$, où $m^X(x)$ est la marginale de X en x et joue uniquement le rôle d'un facteur de normalisation pour $\pi(\theta|x)$ (fonction de θ).

On « voit » alors que $\theta|x$ suit une loi gaussienne. Nous calculons donc sa densité à une constante multiplicative (une quantité ne dépendant pas de θ) près : Dans toute la suite, la notation $\pi(\theta|x) \propto g(\theta)$ signifie « $\exists \lambda > 0$: $\forall \theta, \pi(\theta|x) = \lambda g(\theta)$ », où λ peut dépendre de x mais pas de θ . La constante λ est une constante de normalisation qui pourrait être calculée explicitement grâce au fait que $\int_{\theta} \pi(\theta|x) d\theta = 1$, mais dont on n'a pas besoin dans notre cas pour identifier la loi a posteriori.

$$\pi(\theta|x) \propto \pi(\theta)p_{\theta}(x)$$

$$\propto e^{-\frac{(\theta-\mu_{0})^{2}}{2\tau_{0}^{2}}}e^{-\frac{(\theta-x)^{2}}{2\sigma^{2}}}$$

$$\propto e^{-\frac{1}{2}\left(\left(\frac{1}{\tau_{0}^{2}} + \frac{1}{\sigma^{2}}\right)\theta^{2} - 2\left(\frac{\mu_{0}}{\tau_{0}^{2}} + \frac{x}{\sigma^{2}}\right)\theta\right)}$$

$$\propto e^{-\frac{1}{2}\frac{\tau_{0}^{2} + \sigma^{2}}{\tau_{0}^{2}\sigma^{2}}\left(\theta^{2} - 2\frac{\mu_{0}\sigma^{2} + x\tau_{0}^{2}}{\tau_{0}^{2} + \sigma^{2}}\theta\right)}$$

On identifie l'espérance et la variance de la loi grâce à la relation de proportionnalité $e^{-\frac{(\theta-m)^2}{2v^2}} \propto e^{-\frac{1}{2v^2}(\theta^2-2m\theta)}$ et on obtient :

$$\boldsymbol{\theta}|x \sim \mathcal{N}\left(\frac{\mu_0 \sigma^2 + x \tau_0^2}{\tau_0^2 + \sigma^2}, \frac{\tau_0^2 \sigma^2}{\tau_0^2 + \sigma^2}\right).$$

Si on désire faire tous les calculs (y compris celui de la marginale), on peut utiliser l'identité suivante :

$$\int_{-\infty}^{+\infty} e^{-ax^2 + bx + c} dx = \sqrt{\frac{\pi}{a}} e^{\frac{b^2}{4a} + c}.$$

2. Puisque l'échantillon est i.i.d., Y suit une loi produit. Le modèle bayésien pour Y est donc :

$$\begin{cases} \boldsymbol{\theta} \sim \mathcal{N}(\mu_0, \tau_0^2), \\ Y | \boldsymbol{\theta} \sim \mathcal{N}(\boldsymbol{\theta}, \sigma^2)^{\otimes n}. \end{cases}$$

Pour déterminer la loi a posteriori, on utilise la formule de la distribution a posteriori : $\pi(\theta|y) = \frac{\pi(\theta)p_{\theta}(y)}{m^{Y}(y)}$, où m^{Y} est la marginale de Y. À une constante près cela

donne:

$$\begin{split} \pi(\theta|y) &\propto \pi(\theta) p_{\theta}(y) \\ &\propto e^{-\frac{(\theta - \mu_{0})^{2}}{2\tau_{0}^{2}}} e^{-\frac{\sum_{i=1}^{n} (\theta - x_{i})^{2}}{2\sigma^{2}}} \\ &\propto e^{-\frac{1}{2} \left((\frac{1}{\tau_{0}^{2}} + \frac{n}{\sigma^{2}}) \theta^{2} - 2(\frac{\mu_{0}}{\tau_{0}^{2}} + \frac{\sum_{i} x_{i}}{\sigma^{2}}) \theta \right)} \\ &\propto e^{-\frac{1}{2} \frac{n\tau_{0}^{2} + \sigma^{2}}{\tau_{0}^{2}\sigma^{2}} \left(\theta^{2} - 2\frac{\mu_{0}\sigma^{2} + \tau_{0}^{2} \sum_{i} x_{i}}{n\tau_{0}^{2} + \sigma^{2}} \theta \right)} \end{split}$$

On reconnaît une loi gaussienne:

$$\boldsymbol{\theta}|y \sim \mathcal{N}\left(\frac{\mu_0 \sigma^2 + \tau_0^2 \sum_i x_i}{n\tau_0^2 + \sigma^2}, \frac{\tau_0^2 \sigma^2}{n\tau_0^2 + \sigma^2}\right).$$

Pour étudier le comportement lorsque $n \to +\infty$, on réécrit l'espérance et la variance de la loi en introduisant la moyenne des x_i , $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$:

$$\boldsymbol{\theta}|y \sim \mathcal{N}\left(\frac{\mu_0 \sigma^2 + n\tau_0^2 \bar{x}}{n\tau_0^2 + \sigma^2}, \frac{\tau_0^2 \sigma^2}{n\tau_0^2 + \sigma^2}\right).$$

Lorsque
$$n \to +\infty$$
, $\mathbb{E}(\boldsymbol{\theta}|y) \sim \frac{n\tau_0^2 \bar{x}}{n\tau_0^2} = \bar{x} \to \mathbb{E}(X_1|\theta) = \theta$ et $\mathbb{V}\operatorname{ar}(\boldsymbol{\theta}|y) \sim \frac{\tau_0^2 \sigma^2}{n\tau_0^2} = \frac{\sigma^2}{n} \to 0$.

Solution Exercice 2

- 1. L'espace des paramètres est : $\Theta = \{\theta_1, \theta_2\}$. Le modèle statistique est : $\mathcal{P} = \{\mathcal{B}(\theta), \theta \in \Theta\}$, où $\mathcal{B}(\theta)$ est une loi de Bernoulli de paramètre θ .
- 2. Ici on définit un modèle bayésien pour X:

$$\begin{cases} \boldsymbol{\theta} \sim \pi, \\ X | \boldsymbol{\theta} \sim \mathcal{B}(\boldsymbol{\theta}). \end{cases}$$

De manière générale, on a :

$$\pi(\theta|x) = \frac{\pi(\theta)p_{\theta}(x)}{\pi(\theta_1)p_{\theta_1}(x) + \pi(\theta_2)p_{\theta_2}(x)} = \frac{\pi(\theta)\theta^x(1-\theta)^{1-x}}{\pi(\theta_1)\theta_1^x(1-\theta_1)^{1-x} + \pi(\theta_2)\theta_2^x(1-\theta_2)^{1-x}}.$$

Pour $\pi = \pi_1$ ($\pi_1(\theta_1) = \pi_2(\theta_2) = 0.5$), on obtient donc les valeurs $\pi_1(\theta|x)$:

$$\begin{array}{c|ccccc} \theta/x & 0 & 1 \\ \hline \theta_1 & \frac{1-\theta_1}{1-\theta_1+1-\theta_2} = \frac{2}{3} & \frac{\theta_1}{\theta_1+\theta_2} = \frac{1}{4} \\ \hline \theta_2 & \frac{1-\theta_2}{1-\theta_1+1-\theta_2} = \frac{1}{3} & \frac{\theta_2}{\theta_1+\theta_2} = \frac{3}{4} \\ \end{array}$$

3. On a maintenant comme distribution a priori : $\pi_2(\theta_1) = 1/4$, $\pi_2(\theta_2) = 3/4$. On obtient donc les valeurs $\pi_2(\theta|x)$:

$$\begin{array}{c|ccccc}
\theta/x & 0 & 1 \\
\hline
\theta_1 & \frac{1 \times (1 - \theta_1)}{1 \times (1 - \theta_1) + 3 \times (1 - \theta_2)} = \frac{2}{5} & \frac{1 \times \theta_1}{1 \times \theta_1 + 3 \times \theta_2} = \frac{1}{10} \\
\hline
\theta_2 & \frac{3}{5} & \frac{9}{10}
\end{array}$$

4. La densité de la loi prédictive a posteriori est donnée par :

$$p(y) = \pi_2(\theta_1|x)p_{\theta_1}(y) + \pi_2(\theta_2|x)p_{\theta_2}(y).$$

Pour x = 1 on obtient : $p(0) = \frac{1}{10} \times (1 - 0.2) + \frac{9}{10} \times (1 - 0.6) = \frac{11}{25}$ et $p(1) = \frac{14}{25}$.

5. Pour $X = (X_1, \dots, X_n)$ échantillon *i.i.d.* :

$$\begin{split} \pi(\theta|x) &= \frac{\pi(\theta)p_{\theta}(x)}{\pi(\theta_{1})p_{\theta_{1}}(x) + \pi(\theta_{2})p_{\theta_{2}}(x)} \\ &= \frac{\pi(\theta)\theta^{\sum_{i}x_{i}}(1-\theta)^{n-\sum_{i}x_{i}}}{\pi(\theta_{1})\theta_{1}^{\sum_{i}x_{i}}(1-\theta_{1})^{n-\sum_{i}x_{i}} + \pi(\theta_{2})\theta_{2}^{\sum_{i}x_{i}}(1-\theta_{2})^{n-\sum_{i}x_{i}}} \\ &= \frac{\pi(\theta)\theta^{s}(1-\theta)^{n-s}}{\pi(\theta_{1})\theta_{1}^{s}(1-\theta_{1})^{n-s} + \pi(\theta_{2})\theta_{2}^{s}(1-\theta_{2})^{n-s}}. \end{split}$$

Ainsi $\pi(\theta|x)$ ne dépend de x qu'à travers s.

6. On pose $\alpha = \frac{\theta_2(1-\theta_2)}{\theta_1(1-\theta_1)} = \frac{3}{2}$. Alors:

$$\pi(\theta_1|x) = \frac{\pi(\theta_1)}{\pi(\theta_1) + \alpha^{n/2}\pi(\theta_2)}$$

et

$$\pi(\theta_2|x) = \frac{\pi(\theta_2)}{\alpha^{-n/2}\pi(\theta_1) + \pi(\theta_2)}.$$

D'où, quand $n \to +\infty$, $\pi(\theta_1|x) \to 0$ et $\pi(\theta_2|x) \to 1$, indépendamment de la distribution a priori de $\boldsymbol{\theta}$. La loi conditionnelle a posteriori tend à privilégier le θ le plus proche de 0.5 (en effet $\theta_2 = 0.6 \approx 0.5$ et $\theta_1 = 0.2 \neq 0.5$).

Solution Exercice 3 1. Comme la loi a priori est $\pi = \mathcal{G}amma(a, \lambda)$, la densité a priori est donnée par

$$\pi(\theta) = \mathbb{1}_{\theta > 0} \frac{\lambda^a}{\Gamma(a)} \theta^{a-1} e^{-\lambda \theta}, \quad \forall \theta \in \Theta.$$

Pour $\theta \in \Theta$, le modèle statistique des observations est décrit par la densité, calculée dans la question 2.1.2 :

$$p_{\theta}^{\otimes n}(\mathbf{x}) = \frac{\theta^{(\sum_{i=1}^{n} x_i)}}{\prod_{i=1}^{n} x_i!} e^{-n\theta}, \quad \forall \mathbf{x} \in \mathbb{N}^n.$$

La formule de la distribution a posteriori nous donne

$$\pi(\theta|\mathbf{x}) = \frac{\pi(\theta)p_{\theta}^{\otimes n}(\mathbf{x})}{m^{X}(\mathbf{x})} = C(\mathbf{x})\mathbb{1}_{\theta>0}\theta^{a-1}e^{-\lambda\theta}\theta^{(\sum_{i=1}^{n}x_{i})}e^{-n\theta}$$

où $C(\mathbf{x})$ est une grandeur qui ne dépend pas de θ . En réorganisant la formule, on trouve

$$\pi(\theta|\mathbf{x}) = C(\mathbf{x}) \mathbb{1}_{\theta > 0} \theta^{(a-1 + \sum_{i=1}^{n} x_i)} e^{-(\lambda + n)\theta}$$

On reconnaît à une constante près la densité d'une loi gamma de paramètres $a + \sum_{i=1}^{n} x_i$ et $\lambda + n$ et on en déduit la valeur de la constante $C(\mathbf{x})$:

$$\pi(\theta|\mathbf{x}) = \mathbb{1}_{\theta>0} \frac{(\lambda+n)^{(a+\sum_{i=1}^{n} x_i)}}{\Gamma(a+\sum_{i=1}^{n} x_i)} \theta^{(a-1+\sum_{i=1}^{n} x_i)} e^{-(\lambda+n)\theta}$$

Cela montre que la loi a posteriori est $\pi(\cdot|\mathbf{x}) = \mathcal{G}amma(a + \sum_{i=1}^{n} x_i, \lambda + n)$.

2. En utilisant la formule de l'espérance pour la loi Gamma, on obtient

$$\mathbb{E}(\boldsymbol{\theta}|\mathbf{x}) = \int_0^{+\infty} \theta \pi(d\theta|\mathbf{x}) = \frac{a + \sum_{i=1}^n x_i}{\lambda + n}$$

3. D'après la loi des grands nombres, la moyenne empirique des observations d'une suite de variables aléatoires i.i.d. converge presque sûrement vers leur espérance. On a donc

$$\mathbb{E}(\boldsymbol{\theta}|\mathbf{X}) = \frac{a + \sum_{i=1}^{n} X_i}{\lambda + n} = \frac{n^{-1}a + n^{-1} \sum_{i=1}^{n} X_i}{n^{-1}\lambda + 1} \sim \frac{1}{n} \sum_{i=1}^{n} X_i \xrightarrow{\text{p.s.}} \mathbb{E}(X_1) = \theta_0$$

lorsque $X_1 \sim \mathcal{P}oiss(\theta_0)$.

4. Soit Y une variable aléatoire qui suit la loi a posteriori $\pi(\cdot|\mathbf{X}) = \mathcal{G}amma(a + \sum_{i=1}^{n} X_i, \lambda + n)$. D'après les indications de l'énoncé,

$$\operatorname{Var}(\boldsymbol{\theta}|\mathbf{X}) = \operatorname{Var}(Y) = \frac{a + \sum_{i=1}^{n} X_i}{(\lambda + n)^2}$$

Comme dans la question 3, on utilise la loi des grands nombres :

$$\operatorname{Var}(\boldsymbol{\theta}|\mathbf{X}) = \frac{a + \sum_{i=1}^{n} X_i}{(\lambda + n)^2} \sim \frac{1}{n^2} \sum_{i=1}^{n} X_i \overset{\text{p.s.}}{\sim} \frac{\mathbb{E}(X_1)}{n} \to 0$$