M1 - Statistiques bayésiennes

Devoir à la maison

à rendre avant vendredi 3 avril 17h à anna.ben-hamou@upmc.fr

Exercice 1

On considère le cadre bayésien suivant : $\boldsymbol{\sigma}^2 \sim \Pi = \mathrm{IG}(a, b)$, pour $a, b > 0$ (loi inverse-Gamma), et $\mathbf{X} = (X_1, \ldots, X_n) \,|\, \boldsymbol{\sigma}^2 \sim \mathcal{N}(0, \boldsymbol{\sigma}^2)^{\otimes n}$. Dans tout l'exercice, la fonction de perte considérée est la fonction $\ell : \mathbb{R}_+^* \times \mathbb{R}_+^* \to \mathbb{R}_+$ donnée par

$$\forall \sigma^2, t \in \mathbb{R}_+^*,\ \ell(\sigma^2, t) = \left(\frac{t}{\sigma^2} - 1\right)^2 .$$

1. Quelle est la loi a posteriori $\Pi[\cdot \,|\, \mathbf{X}]$?
2. Déterminer un estimateur de Bayes pour Π et la perte ℓ. On le notera T^*.
3. Donner la fonction de risque $\sigma^2 \mapsto \mathbf{R}(\sigma^2, T^*)$ de T^*.
4. Que vaut le risque de Bayes $\mathbf{R}_B(\Pi)$?
5. Montrer que le risque minimax $\mathbf{R}_M$ vérifie $\mathbf{R}_M \geq \frac{2}{n+2}$.
6. Calculer la fonction de risque (toujours pour la perte ℓ) de l'estimateur $T(\mathbf{X}) = \frac{1}{n}\sum_{i=1}^n X_i^2$. Que vaut $\mathbf{R}_{\max}(T)$?

Exercice 2

On considère le cadre bayésien suivant : $\boldsymbol{\theta} \sim \Pi = \mathrm{Beta}(a, b)$, pour $a, b > 0$, et $\mathbf{X} = (X_1, \ldots, X_n) \,|\, \boldsymbol{\theta} \sim \mathrm{Geom}(\boldsymbol{\theta})^{\otimes n}$. On rappelle que pour $p \in]0, 1[$, $X \sim \mathrm{Geom}(p)$ (loi géométrique de paramètre p) si et seulement si pour tout $k \in \mathbb{N}^*$, $\mathbb{P}(X = k) = (1 - p)^{k-1} p$.

1. Quelle est la loi a posteriori $\Pi[\cdot \,|\, \mathbf{X}]$? Que valent la moyenne a posteriori, notée $m_{\mathbf{X}}$, et la variance a posteriori, notée $v_{\mathbf{X}}$?
2. Montrer que, pour tout $\theta_0 \in]0, 1[$, sous $P_{\theta_0}^{\otimes n}$, on a
$$m_{\mathbf{X}} \xrightarrow{\mathbb{P}} \frac{1}{\theta_0} \qquad \text{et } n v_{\mathbf{X}} \xrightarrow{\mathbb{P}} \theta_0^2 \,(1 - \theta_0) \,.$$
3. Montrer que, pour tout $\theta_0 \in]0, 1[$, la loi a posteriori est consistante en θ_0, à vitesse $\frac{1}{\sqrt{n}}$, c'est-à-dire que pour toute suite $(M_n)_{n \geq 1}$ telle que $M_n \to +\infty$, on a, sous $P_{\theta_0}^{\otimes n}$,
$$\mathbb{P}\left(|\boldsymbol{\theta} - \theta_0| > \frac{M_n}{\sqrt{n}} \,|\, \mathbf{X}\right) \xrightarrow{\mathbb{P}} 0 \,.$$
Pour cela on pourra d'abord montrer que $\sqrt{n}(m_{\mathbf{X}} - \theta_0)$ converge en loi sous $P_{\theta_0}^{\otimes n}$ vers une variable aléatoire que l'on précisera.

Exercice 3

On considère le modèle d'échantillonnage de Bernoulli $\mathcal{P} = \{\mathcal{B}(\theta)^{\otimes n},\ \theta \in [0, 1]\}$, et, pour $0 < q < p < 1$, on souhaite tester

$$H_0 : \theta = q \qquad \text{contre} \qquad H_1 : \theta = p \,.$$

Pour cela, on commence par choisir comme loi a priori $\Pi = \frac{1}{2}\delta_q + \frac{1}{2}\delta_p$, et l'on considère le modèle : $\boldsymbol{\theta} \sim \Pi$ et $\mathbf{X} = (X_1, \ldots, X_n) \,|\, \boldsymbol{\theta} \sim \mathcal{B}(\boldsymbol{\theta})^{\otimes n}$.

1. Déterminer le test de Bayes pour Π et la perte équilibrée, noté φ^*, et l'écrire sous la forme $\varphi^*(\mathbf{X}) = \mathbb{1}_{\{\overline{X}_n \geq c_{p,q}\}}$, pour une certaine constante $c_{p,q}$ à déterminer.
2. Montrer que le risque de Bayes est égal à
$$\mathbf{R}_B(\Pi) = \frac{1}{2}\left(1 - \|\mathcal{B}(p)^{\otimes n} - \mathcal{B}(q)^{\otimes n}\|_{\mathrm{VT}}\right) .$$
3. Montrer que
$$\|\mathcal{B}(p)^{\otimes n} - \mathcal{B}(q)^{\otimes n}\|_{\mathrm{VT}} \geq \mathbb{P}_p\left(\sum_{i=1}^n X_i \geq \frac{p+q}{2}n\right) - \mathbb{P}_q\left(\sum_{i=1}^n X_i \geq \frac{p+q}{2}n\right) .$$
4. En utilisant l'inégalité de Hoeffding, montrer que
$$1 - \|\mathcal{B}(p)^{\otimes n} - \mathcal{B}(q)^{\otimes n}\|_{\mathrm{VT}} \leq 2e^{-\frac{(p-q)^2 n}{2}} .$$
5. En déduire que le test de Bayes est consistant.