M1 - Statistiques bayésiennes

Mini-test 1, le 13/02/2018

Durée 30mn. Les documents ne sont pas autorisés.

Exercice 1

Soient $X_1, \dots, X_n$ des variables aléatoires indépendantes et identiquement distribuées selon une loi P_θ sur $\mathbb{R}$, avec θ inconnu *fixé* dans $\Theta \subset \mathbb{R}$. On note $\mathcal{P} = \{P_\theta,\, \theta \in \Theta\}$.

1. (a) Quand dit-on que le modèle $\mathcal{P}$ est dominé ?
   (b) Quand dit-on que le modèle $\mathcal{P}$ est identifiable ?
2. Soit $\hat{\theta}_n$ un estimateur de θ.
   (a) Quand dit-on que l'estimateur $\hat{\theta}_n$ est consistant ?
   (b) Asymptotiquement normal ?
   (c) Quel est le risque quadratique de $\hat{\theta}_n$?
   (d) Exprimer ce risque comme une somme d'un terme de biais et d'un terme de variance.

Exercice 2

On se place dans un cadre bayésien.

1. On considère $\Theta \subset \mathbb{R}$ et l'on se donne une loi a priori Π sur Θ, ainsi qu'un ensemble de lois $\{P_\theta,\ \theta \in \Theta\}$ où les P_θ sont des lois sur $\mathbb{R}$. On suppose que P_θ ainsi que Π ont des densités $f_\theta(\cdot)$ et $\pi(\cdot)$ par rapport à la mesure de Lebesgue sur $\mathbb{R}$, soit

$$dP_\theta(x) = f_\theta(x)dx, \quad d\Pi(\theta) = \pi(\theta)d\theta.$$

   On dispose d'un échantillon $X = (X_1, \dots, X_n)$ dont la loi conditionnelle sachant θ est $P_\theta^{\otimes n}$. Donner
   (a) la densité de la loi a priori de θ,
   (b) la densité de la loi de $X \,|\, \theta$,
   (c) la densité de la loi a posteriori de θ sachant $X = x$.
2. On considère maintenant le scénario bayésien suivant :

$$\begin{aligned} \theta &\sim \text{Beta}(a,b) \text{ avec } a, b > 0 \\ X = (X_1, \dots, X_n) \,|\, \theta &\sim \text{Bernoulli}(\theta)^{\otimes n}. \end{aligned}$$

   On rappelle que la densité de la loi Beta(a, b) est donnée, pour tout θ de $[0, 1]$, par

$$\pi(\theta) = \frac{\theta^{a-1}(1-\theta)^{b-1}}{\mathcal{B}(a,b)}, \text{ avec } \mathcal{B}(a,b) = \int_0^1 u^{a-1}(1-u)^{b-1}du\,,$$

   et que $\mathbb{E}\left[\text{Beta}(a,b)\right] = \frac{a}{a+b}$ et $\text{Var}\left(\text{Beta}(a,b)\right) = \frac{ab}{(a+b)^2(a+b+1)}$.
   (a) Déterminer la loi a posteriori.
   (b) En déduire l'expression de la moyenne a posteriori, notée $\hat{\theta}(X)$. Interpréter brièvement.
   (c) À l'aide de l'inégalité de Tchebytchev, construire une région de crédibilité au moins $1 - \alpha$, pour $\alpha \in]0, 1[$, centrée autour de $\hat{\theta}(X)$.