

6 TD6

La loi *inverse-Gamma* $IG(a, b)$ est la loi sur \mathbb{R}^+ de densité, pour $a, b > 0$,

$$x \rightarrow x^{-(a+1)} e^{-b/x} \frac{b^a}{\Gamma(a)}.$$

La loi de *Student* à p degrés de liberté a une densité $f_p(u)$ sur \mathbb{R} proportionnelle à $\left(1 + \frac{u^2}{p}\right)^{-\frac{p+1}{2}}$.

6.1 Famille gaussienne conjuguée dans \mathbb{R}^d

Soit $\mathcal{P} = \{P_\theta = \mathcal{N}(\theta, \Sigma), \theta \in \mathbb{R}^d\}$, où Σ est une matrice de variance-covariance inversible connue. Soit Π la loi a priori $\mathcal{N}(0, \Lambda)$ sur θ , avec Λ inversible.

1. Montrer que si

$$(X_1, \dots, X_n) | \theta \sim P_{\theta, \Sigma}^{\otimes n},$$

la loi a posteriori s'écrit $\theta | (X_1, \dots, X_n) \sim \mathcal{N}(\theta_X, \Sigma_X)$, avec,

$$\begin{aligned}\Sigma_X &= (n\Sigma^{-1} + \Lambda^{-1})^{-1} \\ \theta_X &= n\Sigma_X \Sigma^{-1} \bar{X}.\end{aligned}$$

2. La classe de lois a priori $\{\mathcal{N}(0, \Lambda), \Lambda \text{ inversible}\}$ est-elle conjuguée ?

6.2 Famille gaussienne à moyenne et variance inconnues

On considère le modèle $\mathcal{P} = \{P_\theta = P_{\mu, \sigma^2} = \mathcal{N}(\mu, \sigma^2), \mu \in \mathbb{R}, \sigma^2 > 0\}$ où l'on a posé $\theta = (\mu, \sigma^2)$. On dispose d'observations $X = (X_1, \dots, X_n)$, de loi $P_\theta^{\otimes n}$ sachant θ . On définit la mesure Π

$$d\Pi(\mu, \sigma^2) = \frac{1}{\sigma^2} d\mu d\sigma^2,$$

où $d\mu d\sigma^2$ s'interprète comme $d\text{Leb}_{\mathbb{R}}(\mu) d\text{Leb}_{\mathbb{R}^+}(\sigma^2)$.

1. Vérifier que Π est une "loi" a priori impropre.
2. Montrer que $\mathcal{L}(\sigma^2 | X)$ est une loi $IG(\frac{n-1}{2}, \frac{s}{2})$, où $s = \sum_{i=1}^n (X_i - \bar{X}_n)^2$. Pour ce faire
 - (a) Montrer que l'intégrale $\int p_{\mu, \sigma^2}(X) \sigma^{-2} d\mu$ est finie. Elle permet donc de définir à constante de proportionnalité près une "densité jointe" de (σ^2, X) .
 - (b) En déduire $\mathcal{L}(\sigma^2 | X)$.
 - (c) Vérifier également en passant que la loi a posteriori $\mathcal{L}((\mu, \sigma^2) | X)$ est bien définie.
3. Caractériser la loi a posteriori $\mathcal{L}(\mu, \sigma^2 | X)$ à l'aide de $\mathcal{L}(\mu | \sigma^2, X)$ et $\mathcal{L}(\sigma^2 | X)$.
4. Construire une région de crédibilité pour μ au niveau $1 - \alpha$.

6.3 Empirical Bayes et lois normales

On se place dans le cadre du modèle fondamental $\mathcal{P} = \{P_\theta = \mathcal{N}(\theta, 1), \theta \in \mathbb{R}\}$. On dispose de n observations X_1, \dots, X_n i.i.d. sachant θ de loi P_θ . Soit $\Pi = \Pi_\mu = \mathcal{N}(\mu, 1)$ loi a priori sur θ . On se propose de déterminer μ par une méthode bayésienne empirique.

1. En quoi consiste cette méthode ?

2. On construit $\hat{\mu}$ par la méthode du maximum de vraisemblance marginal. Rappeler le principe de cette méthode en deux lignes maximum.
3. Montrer que la loi marginale de (X_1, \dots, X_n) est celle d'un vecteur gaussien. On pourra s'inspirer de l'exercice 2 du TD5.
4. En déduire que $\hat{\mu} = \overline{X}$. Quelle est la loi a posteriori finale suggérée par la méthode bayésienne empirique ?

6.4 Empirical Bayes et lois de Poisson

Soit $\mathcal{P} = \{P_\theta = \mathcal{P}(\theta), \theta > 0\}$. On dispose d'observations X_1, \dots, X_n i.i.d. de loi P_θ sachant θ . Soit $\Pi = \Pi_\lambda = \mathcal{E}(\lambda)$ loi a priori sur θ . On se propose de déterminer λ par une méthode bayésienne empirique.

1. Montrer que la loi marginale de X_1 est une loi géométrique de paramètre $\lambda/(\lambda + 1)$
2. Calculer la densité marginale de (X_1, \dots, X_n) en fonction de λ .
3. En déduire que $\hat{\lambda}^{EB} = 1/\overline{X}$ puis la loi a posteriori finale suggérée par la méthode bayésienne empirique

6.5 Un exemple non-paramétrique (*)

On observe $X_1, \dots, X_n \in [0, 1]$ i.i.d. de densité inconnue f sur l'intervalle $[0, 1]$. On suppose que f appartient à l'intersection \mathcal{C} de l'ensemble de toutes les densités avec la classe de Hölder $\Sigma(1, L)$ définie par, pour $L > 0$,

$$\Sigma(1, L) = \{f : [0, 1] \rightarrow \mathbb{R}, \|f\|_\infty + \|f'\|_\infty \leq L\}.$$

1. Ecrire le modèle statistique. On se propose de minorer le risque minimax pour l'estimation de f au point $x_0 \in (0, 1)$ soit, pour $X = (X_1, \dots, X_n)$,

$$R_M = \inf_T \sup_{f \in \mathcal{C}} E_f(T(X) - f(x_0))^2.$$

On pose $f_0 = 1$ la densité sur $[0, 1]$ constante égale à 1, et $f_1 = 1 + \varphi_n$, où

$$\varphi_n(x) = d_n K\left(\frac{x - x_0}{h_n}\right),$$

avec d_n, h_n à choisir tendant vers 0 et K une fonction de $\Sigma(1, L)$ nulle à l'extérieur de $[-1/2, 1/2]$ et telle que $\int_{\mathbb{R}} K = 0$ et $K(0) > 0$ (une telle fonction existe bien).

2. A l'aide de la méthode de minoration de Le Cam, minorer le risque minimax R_M en fonction de $(f_1 - f_0)(x_0)^2$ et de $\|P_1^n - P_0^n\|_1$, où

$$P_0^n = P_{f_0}^{\otimes n}, \quad P_1^n = P_{f_1}^{\otimes n}.$$

3. En utilisant la majoration $\|P - Q\|_1^2 \leq \int (dP/dQ - 1)^2 dQ = \int (dP/dQ)^2 dQ - 1$, majorer $\|P_1^n - P_0^n\|_1$ en fonction de d_n et h_n .
4. En choisissant bien d_n et h_n , en conclure qu'il existe une constante $C_0 > 0$ telle que pour n assez grand,

$$R_M \geq C_0 n^{-2/3}.$$