

Master professionnel II : Ingénierie mathématique : Option Statistique

Statistique Bayésienne.

Anne PHILIPPE
Université de Nantes
Laboratoire de Mathématiques Jean Leray

Fiche 5. Construction de lois a priori

EXERCICE 1.

On dispose de n observations N_1, \dots, N_n . La i^{e} observation N_i est égale au nombre de pièces défectueuses dans le i^{e} lot. Un lot est constitué de 50 pièces. On veut estimer la probabilité p de produire une pièce défectueuse.

- 1) Décrire le modèle statistique.
- 2) On suppose que la loi a priori sur p est la loi beta de paramètres $(a, b) \in \mathbb{R}_+^* \times \mathbb{R}_+^*$. Quelle est la loi a posteriori ?
- 3) En déduire une famille de lois conjuguées pour le modèle binomial.
- 4) Le service qualité fournit les informations suivantes
 - la proportion de pièces défectueuses est en moyenne 0.15
 - la proportion de pièces X_1, \dots, X_n défectueuses appartient à l'intervalle $[0.1, 0.2]$ avec une probabilité de 95%

Utiliser ces informations pour fixer les paramètres (a, b) de la loi a priori.

EXERCICE 2. CALCUL DES LOIS A PRIORI CONJUGUÉES

Donner une famille de lois conjuguées pour les modèles $\{P_\theta, \theta \in \Theta\}$ suivants

- 1) Poisson : $\{\mathcal{P}(\theta), \theta \in \mathbb{R}_+^*\}$
- 2) Gaussien de variance connue : $\{\mathcal{N}(\theta, \sigma^2), \theta \in \mathbb{R}\}$

Rappel sur les lois .

On dit que X suit une loi inverse Gamma de paramètres $(a, b) \in \mathbb{R}_+^{*2}$, si X^{-1} suit une loi Gamma de paramètres (a, b) . Cette loi admet pour densité sur \mathbb{R}_+^* .

$$b^a \frac{1}{\Gamma(a)} e^{-b/x} x^{-a-1} \mathbb{I}_{\mathbb{R}_+^*}(x).$$

Si $a > 1$ alors X appartient à L^1 et $\mathbb{E}(X) = \frac{b}{a-1}$.

Si $a > 2$ alors X appartient à L^2 et $\text{Var}(X) = \frac{b^2}{(a-1)^2(a-2)}$.

- 3) Gaussien de moyenne connue : $\{\mathcal{N}(\mu, \theta), \theta \in \mathbb{R}_+^*\}$

- 4) Gaussien : $\{\mathcal{N}(\theta_1, \theta_2), \theta_1 \in \mathbb{R}, \theta_2 \in \mathbb{R}_+^*\}$
Indication : Considérer la famille des lois de paramètres $(\lambda, \tau, a, b) \in \mathbb{R} \times \mathbb{R}_+^{*3}$ définies par
 — la loi conditionnelle de θ_1 sachant θ_2 est la loi gaussienne de moyenne λ et de variance θ_2/τ
 — la loi de θ_2 est la loi inverse Gamma de paramètres (a, b) .

EXERCICE 3. CALCUL DES LOIS DE JEFFREY

Calculer la loi non informative de Jeffrey pour les modèles suivants :

- 1) Poisson : $\{\mathcal{P}(\theta), \theta \in \mathbb{R}^+\}$
- 2) Gaussien : $\{\mathcal{N}(\mu, \sigma^2), (\mu, \sigma^2) \in \mathbb{R} \times \mathbb{R}_+\}$ dans les deux situations suivantes
 - a. $\theta = \mu$ et σ^2 est connu.
 - b. $\theta = \sigma$ et μ est connu.
 - c. les deux paramètres sont inconnus : $\theta = (\mu, \sigma)$.

EXERCICE 4.

Soit $\mathbf{X} = (X_1, \dots, X_n)$ des variables aléatoires indépendantes et identiquement distribuées suivant la loi uniforme sur $[0 ; \theta]$ avec $\theta > 0$ inconnu.

On pose

$$M_n = \max(X_1, \dots, X_n).$$

Soit (a, b) deux réels tels que $a > 1$ et $b > 0$. On considère la densité de probabilité $\pi_{a,b}$ définie par

$$\pi_{a,b}(\theta) = ab^a \frac{1}{\theta^{a+1}} \mathbb{I}_{[b; +\infty[}(\theta).$$

- 1) Montrer que $\{\pi_{a,b} : a > 1, b > 0\}$ est une famille de lois conjuguées pour le modèle uniforme.
- 2) Montrer que l'estimateur de Bayes sous le coût quadratique est égal à

$$\delta_n^{a,b}(\mathbf{X}) = \frac{a+n}{a+n-1} \max(b, M_n).$$

- 3) On suppose que (X_1, \dots, X_n) sont iid suivant la loi uniforme sur $[0 ; \theta]$ et on fixe b tel que $P_\theta(X_1 > b) > 0$.
 - a) Montrer que les variables aléatoires M_n et $\max(b, M_n)$ sont presque sûrement égales à partir d'un certain rang.
 - b) En déduire que l'estimateur de Bayes $\delta_n^{a,b}(\mathbf{X})$ converge presque sûrement vers la vraie valeur du paramètre.
- 4) Que se passe-t-il lorsque $P_\theta(X_1 > b) = 0$?

EXERCICE 5.

On suppose que la densité de la loi a posteriori vérifie les conditions suivantes :

$\theta \mapsto \pi(\theta|X_1, \dots, X_n)$ et $\theta \mapsto \theta\pi(\theta|X_1, \dots, X_n)$ sont intégrables et admettent une primitive C^1

- (1) Montrer que pour tout $\delta \in \mathbb{R}$ on a

$$\mathbb{E}(|\delta - \theta| | X_1, \dots, X_n) = \delta(2P(\theta < \delta | X_1, \dots, X_n) - 1) + \mathbb{E}(\theta | X_1, \dots, X_n) - 2 \int_{-\infty}^{\delta} \theta \pi(\theta | X_1, \dots, X_n) d\theta$$

- (2) En déduire que la médiane de la loi a posteriori est un estimateur de Bayes sous coût L^1