Problèmes inverses — Estimation par maximum de vraisemblance

Corrigé de la feuille d'exercices — Exercice 5

Exercice 5

1. Si $x < 0$, alors $\mathbb{P}_\theta(\hat{\theta}_1 \leqslant x) = 0$. Soit $x \geqslant 0$.

$$\mathbb{P}_\theta(\hat{\theta}_1 \leqslant x) = \prod_{i=1}^{n} \mathbb{P}_\theta(X_i \leqslant x) = \prod_{i=1}^{n} \int_0^{\min\{\theta, x\}} \frac{dy}{2\sqrt{y\theta}} = \begin{cases} \left(\frac{x}{\theta}\right)^{\frac{n}{2}} & \text{si } x < \theta, \\ 1 & \text{si } x \geqslant \theta. \end{cases}$$

Si on note $g_\theta(x)$ la densité de $\hat{\theta}_1$, alors

$$g_\theta(x) = \frac{d}{dx}\mathbb{P}_\theta(\hat{\theta}_1 \leqslant x) = \frac{nx^{\frac{n}{2}-1}}{2\theta^{\frac{n}{2}}}\mathbb{1}_{[0,\theta]}(x).$$

2. $\mathbb{E}_\theta\hat{\theta}_1 = \int_0^\theta x g_\theta(x)dx = \theta\frac{n}{n+2}$. L'estimateur $\hat{\theta}_1$ de θ est donc biaisé, mais il est asymptotiquement sans biais puisque $\lim_{n\to\infty} \mathbb{E}_\theta\hat{\theta}_1 = \theta$.

3. On a $\mathbb{E}_\theta X_1 = \int_0^\theta x f_\theta(x)dx = \frac{\theta}{3}$, donc $\theta = 3\mathbb{E}_\theta X_1$. La méthode des moments consiste à remplacer dans l'expression précédente $f_\theta(x)dx$ (c'est-à-dire la loi de X_1) par la mesure empirique des observations $\hat{\mathbb{Q}}_n = \frac{1}{n}\sum_{i=1}^n \delta_{X_i}$. On obtient

$$\hat{\theta}_2 = 3\int_0^\theta x\hat{\mathbb{Q}}_n(dx) = \frac{3}{n}\sum_{i=1}^{n} X_i = 3\bar{X}_n.$$

On a $\mathbb{E}_\theta\hat{\theta}_2 = \frac{3}{n}\sum_{i=1}^n \mathbb{E}_\theta X_i = 3\mathbb{E}_\theta X_1 = \theta$, donc $\hat{\theta}_2$ est un estimateur sans biais de θ.

4. Le calcul donne

$$\mathcal{R}(\hat{\theta}_1, \theta) = \frac{8\theta^2}{(n+2)(n+4)} \quad \text{et} \quad \mathcal{R}(\hat{\theta}_2, \theta) = \frac{4\theta^2}{5n}.$$

Asymptotiquement, $\mathcal{R}(\hat{\theta}_1, \theta) = o(\mathcal{R}(\hat{\theta}_2, \theta))$, donc pour n assez grand, l'estimateur $\hat{\theta}_1$ est préférable à $\hat{\theta}_2$. Les deux sont consistants puisque $\mathcal{R}(\hat{\theta}_1, \theta) \to 0$ et $\mathcal{R}(\hat{\theta}_2, \theta) \to 0$ lorsque $n \to +\infty$.

5. Soit $\theta > 0$ fixé. Pour tout $x \in \mathbb{R}$, on a

$$\mathbb{P}_\theta(n(\hat{\theta}_1 - \theta) \leqslant x) = \mathbb{P}_\theta(\hat{\theta}_1 \leqslant \theta + \frac{x}{n}) = \begin{cases} \left(1 + \frac{x}{\theta n}\right)^{n/2} & \text{si } x < 0, \\ 1 & \text{sinon} \end{cases} \xrightarrow[n\to+\infty]{} \begin{cases} e^{x/2\theta} & \text{si } x < 0, \\ 1 & \text{sinon.} \end{cases}$$

Donc $n(\hat{\theta}_1 - \theta)$ converge en loi sous $\mathbb{P}_\theta$ vers $-Z$ où $Z \sim \text{Exp}(1/2\theta)$. L'estimateur $\hat{\theta}_1$ a donc pour vitesse n et loi limite $-Z$.

Pour $\hat{\theta}_2$, on utilise le TCL : pour tout $\theta > 0$,

$$\sqrt{n}(\hat{\theta}_2 - \theta) \xrightarrow[n\to+\infty]{\text{en loi sous } \mathbb{P}_\theta} \mathcal{N}(0, \text{Var}_\theta(3X_1)) = \mathcal{N}(0, \frac{4}{5}\theta^2).$$

L'estimateur $\hat{\theta}_2$ a donc pour vitesse $\sqrt{n}$ et loi limite $\mathcal{N}(0, \frac{4}{5}\theta^2)$.