1 Exercices

Exercice 1. Soit $\{X_n,\ n \in \mathbb{N}\}$ un échantillon de loi exponentielle :

$$(\mathbb{R}, \mathscr{B}(\mathbb{R}), \{\theta \exp(-\theta x)\mathbb{1}_{x \geq 0} \cdot \lambda_{\mathrm{Leb}} \,:\, \theta \in \Theta = \mathbb{R}_+^*\}) \,.$$

1. Construire les estimateurs de moments $\hat{\theta}_n^{(1)}$ et $\hat{\theta}_n^{(2)}$ associés aux fonctions $T^{(1)}(x) = x$ et $T^{(2)}(x) = x^2$.
2. Montrer que, pour $i = 1, 2$, les suites d'estimateurs $\{\hat{\theta}_n^{(i)}, n \in \mathbb{N}\}$ sont asymptotiquement normales.
3. Quel estimateur semble préférable d'un point de vue asymptotique ?

Exercice 2. Soit $\{X_n,\ n \in \mathbb{N}\}$ un échantillon d'une loi de Poisson de paramètre $\theta \in \Theta := \mathbb{R}_+^*$. On pose $\bar{X}_n := n^{-1} \sum_{i=1}^n X_i$.

1. Montrer que la suite $\{\bar{X}_n,\ n \in \mathbb{N}\}$ est une suite asymptotiquement normale d'estimateurs de θ.
2. On cherche à estimer la probabilité $g(\theta) = \mathbb{P}_\theta(X_1 = 0)$. Montrer que $\{\exp\{-\bar{X}_n\}, n \geq 0\}$ est une suite asymptotiquement normale d'estimateurs de g.
3. Construire un intervalle de confiance asymptotique pour g.

Exercice 3. Soit $\{X_k,\ k \in \mathbb{N}\}$ un échantillon d'une loi $\mathrm{N}(\theta, 1)$. Pour $n \in \mathbb{N}$, on pose $\bar{X}_n = n^{-1} \sum_{i=1}^n X_i$.

1. Montrer que l'estimateur du maximum de vraisemblance est asymptotiquement normal pour θ.
2. Soit $\{b_n,\ n \in \mathbb{N}\}$ une suite de nombre positifs tels que $\lim_{n\to\infty} b_n = 0$ et $\lim_{n\to\infty} \sqrt{n} b_n = \infty$. On considère la suite d'estimateurs :

   $$\tilde{\theta}_n = \bar{X}_n \mathbb{1}_{\{|\bar{X}_n \geq b_n\}} + a\bar{X}_n \mathbb{1}_{\{|\bar{X}_n| < b_n\}}$$

   Montrer que cette suite d'estimateurs est asymptotiquement normale.
3. Commenter.

Exercice 4. Soit $(\mathsf{X}, \mathscr{X}, \mu)$ un espace mesuré. Soit $\{X_k,\ k \in \mathbb{N}\}$ un échantillon du modèle statistique

$$(\mathsf{X}, \mathscr{X}, \{f_\theta \cdot \mu \,:\, \theta \in \Theta\})$$

où Θ est un intervalle ouvert de $\mathbb{R}$. On suppose que

— pour tout $x \in \mathsf{X}$, $\theta \mapsto f_\theta(x) = f(\theta, x)$ est continûment différentiable en θ et pour tout $x \in \mathsf{X}$ et $\theta \in \Theta$, $f(\theta, x) > 0$.

— pour tout $\theta, \theta_0 \in \Theta$ et,

$$\int |\log f(\theta, x)| f(\theta_0, x) \mathrm{d}x < \infty$$

— le modèle est identifiable.

Pour tout $n \in \mathbb{N}$ et $(x_1, \dots, x_n) \in \mathsf{X}^n$, la log-vraisemblance est donnée par :

$$\ell_n(\theta; x_1, \dots, x_n) := n^{-1} \sum_{i=1}^{n} \log f(\theta, x_i) \ .$$

Les équations de vraisemblance associées sont données par

$$\frac{\partial \ell_n}{\partial \theta}(\theta; x_1, \dots, x_n) = 0 \ . \tag{1}$$

On note $\mathscr{R}_n(x_1, \dots, x_n)$ l'ensemble des racines de l'équation de (1) L'objectif de cet exercice est d'établir l'existence d'une suite $\{\tilde{\theta}_n,\ n \in \mathbb{N}\}$ consistante d'estimateurs de θ telle que, pour tout $\theta \in \Theta$, $\lim_{n\to\infty} \mathbb{P}_\theta(\tilde{\theta}_n \in \mathscr{R}_n(X_1, \dots, X_n)) = 1$.

Soit $\varepsilon > 0$. Considérons les suites de variables $\{Y_i^\varepsilon,\ i \in \mathbb{N}\}$ et $\{Z_i^\varepsilon,\ i \in \mathbb{N}\}$ où

$$Y_i^\varepsilon := \log\{f(\theta_0, X_i)/f(\theta_0 - \varepsilon, X_i)\} \ , \qquad Z_i^\varepsilon := \log\{f(\theta_0, X_i)/f(\theta_0 + \varepsilon, X_i)\} \ .$$

Soit $\theta_0 \in \Theta$ et $\varepsilon > 0$ tel que $]\theta_0 - \varepsilon, \theta_0 + \varepsilon[\subset \Theta$. On pose

$$S_n(\theta_0, \varepsilon) = \{\mathbf{x} = (x_1, \dots, x_n) \in \mathsf{X}^n \ : \ \ell_n(\theta_0, \mathbf{x}) > \ell_n(\theta_0 - \varepsilon, \mathbf{x}) \quad \text{et} \quad \ell_n(\theta_0, \mathbf{x}) > \ell_n(\theta_0 + \varepsilon, \mathbf{x})\} \ .$$

1. Réinterpréter $(X_1, \dots, X_n) \in S_n(\theta_0, \varepsilon)$ en fonction de $(Y_1^\varepsilon, \dots, Y_n^\varepsilon)$ et $(Z_1^\varepsilon, \dots, Z_n^\varepsilon)$. Montrer que
$$\lim_{n\to\infty} \mathbb{P}_{\theta_0}((X_1, \dots, X_n) \in S_n(\theta_0, \varepsilon)) = 1 \ .$$
2. Montrer que si $(x_1, \dots, x_n) \in S_n(\theta_0, \varepsilon)$,
$$\mathscr{R}_n(x_1, \dots, x_n) \cap]\theta_0 - \varepsilon, \theta_0 + \varepsilon[\neq \emptyset \ .$$
3. Supposons que pour tout $n \in \mathbb{N}$, il existe un ensemble X_n tel que pour tout $\theta_0 \in \Theta$,
$$\lim_{n\to\infty} \mathbb{P}_\theta((X_1, \dots, X_n) \in \mathsf{X}_n) = 1$$
et pour tout $(x_1, \dots, x_n) \in \mathsf{X}_n$, $\mathscr{R}_n(x_1, \dots, x_n))$ est un singleton. Soit $\theta_* \in \Theta$ un point arbitraire. On définit
$$\tilde{\theta}_n := \begin{cases} \text{la racine de l'équation de vraisemblance} & \text{si } \mathscr{R}_n(X_1, \dots, X_n) \neq \emptyset \\ \theta_* & \text{sinon} \end{cases}$$
Montrer que la suite d'estimateurs $\{\tilde{\theta}_n,\ n \in \mathbb{N}\}$ est consistante.
4. On s'intéresse maintenant au cas où l'ensemble des racines n'est pas réduit à un singleton. Soit $\{\hat{\theta}_n,\ n \in \mathbb{N}\}$ une suite consistante d'estimateurs de θ. On définit la suite $\{\tilde{\theta}_n,\ n \in \mathbb{N}\}$ de la façon suivante :
$$\tilde{\theta}_n := \begin{cases} \inf\{\theta \in \mathscr{R}_n(X_1, \dots, X_n) \ : \ |\theta - \hat{\theta}_n|\} & \text{si } \mathscr{R}_n(X_1, \dots, X_n) \neq \emptyset \\ \hat{\theta}_n & \text{sinon} \end{cases}$$
Montrer que $\tilde{\theta}_n \overset{\mathbb{P}_{\theta_0}-\text{prob}}{\longrightarrow} \theta_0$.

Prenons $\mathsf{X} \subseteq \mathbb{R}$ et $\Theta :=]0,1[$. Soient p_0 et p_1 deux densités par rapport à la mesure de Lebesgue sur $\mathbb{R}$ telles que $\int |p_0(x) - p_1(x)| \mathrm{d}x > 0$. On considère

$$f(\theta,x) := (1-\theta)p_0(x) + \theta p_1(x)$$

5. Interpréter ce modèle.
6. Écrire les équations de vraisemblance.
7. Montrer l'existence d'une suite $(\tilde{\theta}_n)_n$ qui converge en probabilités vers θ_0 et qui, $\mathbb{P}_{\theta_0}$-p.s., pour n assez grand, est une racine des eqs. de vraisemblance.
8. (Question subsidiaire !) un algorithme de recherche de racines ?

Exercice 5. Soit $g : \mathbb{R} \to \mathbb{R}$ une fonction continue et bornée. Montrer que

$$\mathrm{e}^{-n\lambda} \sum_{k=0}^{\infty} g\left(\frac{k}{n}\right) \frac{(n\lambda)^k}{k!} \to_{n\to\infty} g(\lambda)\ .$$