Problèmes inverses — Estimation par maximum de vraisemblance

Corrigé de la feuille d'exercices — Exercices 9, 10 et 11

Exercice 9

1. La vraisemblance des observations $(x_1, \ldots, x_n) \in (\mathbb{R}_+^*)^n$ est

$$L_n(x_1, \ldots, x_n; \theta) = \prod_{i=1}^{n} \theta^2 x_i e^{-\theta x_i} = \theta^{2n} \left(\prod_{i=1}^{n} x_i \right) \exp\left(-\theta \sum_{i=1}^{n} x_i \right).$$

Donc

$$\frac{d}{d\theta} \ln L_n(x_1, \ldots, x_n; \theta) = \frac{2n}{\theta} - \sum_{i=1}^{n} x_i,$$

qui s'annule en

$$\frac{2n}{\sum_{i=1}^{n} x_i},$$

est strictement négatif pour θ plus grand et strictement positif pour θ plus petit. Donc

$$\hat{\theta}_n = \frac{2n}{\sum_{i=1}^{n} X_i} = \frac{2}{\bar{X}_n}$$

est l'estimateur du maximum de vraisemblance de θ.

2. Les v.a. X_i suivent la loi gamma de paramètre $(2, 1/\theta)$. D'après les propriétés des sommes de v.a. indépendantes de loi gamma, on en déduit que $Y = X_1 + X_2 + \ldots + X_n$ suit la loi gamma de paramètre $(2n, 1/\theta)$, c'est-à-dire de densité

$$\frac{\theta^{2n}}{(2n-1)!} x^{2n-1} e^{-\theta x}.$$

Donc

$$\mathbb{E}_\theta \hat{\theta}_n = 2n \mathbb{E}_\theta \frac{1}{Y} = 2n \int_0^\infty \frac{1}{x} \frac{\theta^{2n}}{(2n-1)!} x^{2n-1} e^{-\theta x} dx = \frac{2n\theta^{2n}}{(2n-1)!} \int_0^\infty x^{2n-2} e^{-\theta x} dx.$$

Puisque $\int_0^\infty x^{k-1} e^{-x} dx = (k-1)!$, on en déduit que

$$\mathbb{E}_\theta \hat{\theta}_n = \frac{2n\theta^{2n}}{(2n-1)!} \frac{(2n-2)!}{\theta^{2n-1}} = \frac{2n}{2n-1} \theta,$$

et le biais vaut

$$\mathbb{E}_\theta \hat{\theta}_n - \theta = \frac{\theta}{2n-1}.$$

En particulier, l'EMV est asymptotiquement sans biais.

Pour le risque quadratique,

$$\begin{aligned}
\mathcal{R}(\hat{\theta}_n, \theta) &= \mathbb{E}_\theta \hat{\theta}_n^2 - 2\theta \mathbb{E}_\theta \hat{\theta}_n + \theta^2 \\
&= \frac{4n^2\theta^{2n}}{(2n-1)!} \int_0^\infty x^{2n-3} e^{-\theta x} dx - 2\frac{2n}{2n-1}\theta^2 + \theta^2 \\
&= \frac{\theta^2(n+1)}{(2n-1)(n-1)},
\end{aligned}$$

qui tend vers 0 lorsque $n \to +\infty$.

3. D'après la LGN, $\bar{X}_n \to 2/\theta$ p.s. De plus, d'après le TCL,

$$\sqrt{n}(\bar{X}_n - \frac{2}{\theta}) \xrightarrow[n\to+\infty]{\text{en loi sous } \mathbb{P}_\theta} \mathcal{N}(0, \mathrm{Var}_\theta(X_1)) = \mathcal{N}(0, \tfrac{5\theta^2}{2}).$$

Afin d'étudier la loi limite de $\hat{\theta}_n$, on calcule

$$\sqrt{n}(\hat{\theta}_n - \theta) = \frac{\sqrt{n}(2 - \theta\bar{X}_n)}{\bar{X}_n} = -\theta\frac{\sqrt{n}(\bar{X}_n - 2/\theta)}{\bar{X}_n}.$$

Le numerateur converge en loi d'après le TCL et le denominateur converge p.s. d'après la LGN, donc converge en probabilité.

On peut donc appliquer le Lemme de Slutsky pour en déduire que le couple converge en loi, et donc leur quotient aussi. Finalement,

$$\sqrt{n}(\hat{\theta}_n - \theta) \to -\frac{\theta}{2/\theta}\mathcal{N}\left(0, \tfrac{5\theta^2}{2}\right) = \mathcal{N}\left(0, \tfrac{5\theta^2}{2} \times \left(-\tfrac{\theta}{2/\theta}\right)^2\right) = \mathcal{N}\left(0, \tfrac{5\theta^6}{8}\right).$$

Ainsi, l'estimateur $\hat{\theta}_n$ est de vitesse $\sqrt{n}$ et de loi limite $\mathcal{N}\left(0, \frac{5\theta^6}{8}\right)$.

Exercice 10

1. Soit $(x_1, \ldots, x_n) \in \mathbb{R}^n$.

$$L_n(x_1, \ldots, x_n; \theta) = \prod_{i=1}^n e^{x_i - \theta}\mathbb{1}_{]-\infty,\theta]}(x_i) = \begin{cases} e^{x_1+\ldots+x_n-n\theta} & \text{si } \max_{1\leqslant i\leqslant n} x_i \leqslant \theta, \\ 0 & \text{sinon.} \end{cases}$$

Comme fonction de θ, cette quantité est clairement maximale pour $\theta = \max_{1\leqslant i\leqslant n} x_i$, donc l'estimateur du maximum de vraisemblance est

$$\hat{\theta} = \max_{1\leqslant i\leqslant n} X_i.$$

2. Pour tout $z \in \mathbb{R}$, on a

$$\mathbb{P}_\theta(n(\hat{\theta} - \theta) \leqslant z) = \prod_{i=1}^n \mathbb{P}_\theta(X_i \leqslant \theta + \tfrac{z}{n}) = \begin{cases} \left(\int_{-\infty}^{\theta+\frac{z}{n}} e^{x-\theta}dx\right)^n & \text{si } z \leqslant 0, \\ 1 & \text{sinon} \end{cases} = \begin{cases} e^z & \text{si } z \leqslant 0, \\ 1 & \text{sinon.} \end{cases}$$

Donc $n(\hat{\theta} - \theta) = -Z$ où $Z \sim \mathrm{Exp}(1)$.

3. On en déduit que $\mathbb{E}_\theta\hat{\theta} = \mathbb{E}_\theta(\theta - \frac{Z}{n}) = \theta - \frac{1}{n}$.

4. Soit $\alpha \in]0,1[$ fixé. Remarquons que $\mathbb{P}(Z \in [0, \ln\frac{1}{\alpha}]) = \int_0^{\ln(1/\alpha)} e^{-x}dx = 1 - \alpha$. Puisque $\theta = \hat{\theta} + \frac{Z}{n}$, on en déduit que, pour tout $\theta \in \mathbb{R}$,

$$\mathbb{P}_\theta\left(\theta \in \left[\hat{\theta}, \hat{\theta} + \frac{\ln(1/\alpha)}{n}\right]\right) = 1 - \alpha,$$

et donc $[\hat{\theta}, \hat{\theta} + \frac{\ln(1/\alpha)}{n}]$ est un intervalle de confiance pour θ de niveau $1 - \alpha$.

Exercice 11

1. Soit $(x_1,\dots,x_n)\in(\mathbb{R}_+^*)^n$ des observations. On note $\theta=(a,b)$ le vecteur des paramètres. La vraisemblance

$$L_n(x_1,\dots,x_n;\theta)=\prod_{i=1}^n\frac{1}{b-a}\mathbb{1}_{[a,b]}(x_i)=\begin{cases}\frac{1}{(b-a)^n} & \text{si } \min_{1\leqslant i\leqslant n}x_i\geqslant a \text{ et } \max_{1\leqslant i\leqslant n}x_i\leqslant b,\\ 0 & \text{sinon.}\end{cases}$$

est maximale pour $b-a$ minimal sous la contrainte $\min_{1\leqslant i\leqslant n}x_i\geqslant a$ et $\max_{1\leqslant i\leqslant n}x_i\leqslant b$, donc l'estimateur du maximum de vraisemblance est

$$\hat a=\min_{1\leqslant i\leqslant n}X_i\quad\text{et}\quad \hat b=\max_{1\leqslant i\leqslant n}X_i.$$

2. On a d'une part $a\leqslant\hat a\leqslant\hat b\leqslant b$ presque sûrement. D'autre part, pour tout $\varepsilon\in]0,b-a[$,

$$\mathbb{P}_\theta(\hat a\geqslant a+\varepsilon)=\mathbb{P}_\theta(X_i\geqslant a+\varepsilon,\ \forall i\in\{1,\dots,n\})=\left(\frac{b-a-\varepsilon}{b-a}\right)^n\xrightarrow[n\to+\infty]{}0.$$

De même, $\mathbb{P}_\theta(\hat b\leqslant b-\varepsilon)=(\frac{b-a-\varepsilon}{b-a})^n\to0$ quand $n\to+\infty$. Donc $\hat a\to a$ et $\hat b\to b$ en probabilité sous $\mathbb{P}_\theta$ pour tout θ.

3. a) Soit $a\leqslant x\leqslant y\leqslant b$. D'une part

$$\mathbb{P}_\theta(\hat a\geqslant x,\ \hat b\leqslant y)=\mathbb{P}_\theta(X_i\in[x,y],\ \forall i\in\{1,\dots,n\})=\left(\frac{y-x}{b-a}\right)^n.$$

D'autre part, si $f_\theta(\alpha,\beta)$ désigne la densité jointe du couple de variables aléatoires $(\hat a,\hat b)$, on a $f_\theta(\alpha,\beta)=0$ si $\alpha>\beta$ ou $\alpha<a$ ou $\beta>b$, et

$$\mathbb{P}_\theta(\hat a\geqslant x,\ \hat b\leqslant y)=\int_x^b d\alpha\int_\alpha^y d\beta\, f_\theta(\alpha,\beta),$$

donc, pour tout $a\leqslant\alpha\leqslant\beta\leqslant b$,

$$f_\theta(\alpha,\beta)=-\frac{\partial^2}{\partial x\partial y}\left(\frac{y-x}{b-a}\right)^n=\frac{n(n-1)(\beta-\alpha)^{n-2}}{(b-a)^n}.$$

b) On en déduit que, pour tout $x\in[0,b-a]$,

$$\begin{aligned}\mathbb{P}_\theta(\hat b-\hat a\leqslant x)&=\int_a^b d\alpha\int_\alpha^{(\alpha+x)\wedge b}d\beta\,\frac{n(n-1)(\beta-\alpha)^{n-2}}{(b-a)^n}\\&=\frac{n}{(b-a)^n}\int_a^b(x\wedge(b-\alpha))^{n-1}\,d\alpha\\&=\frac{nx^{n-1}}{(b-a)^n}\int_a^{b-x}d\alpha+\frac{n}{(b-a)^n}\int_{b-x}^b(b-\alpha)^{n-1}\,d\alpha\\&=n\left(\frac{x}{b-a}\right)^{n-1}-(n-1)\left(\frac{x}{b-a}\right)^n.\end{aligned}$$

Par ailleurs, on a clairement $\mathbb{P}_\theta(\hat b-\hat a\leqslant x)=0$ si $x<0$ et $\mathbb{P}_\theta(\hat b-\hat a\leqslant x)=1$ si $x>b-a$.

c) On déduit de la question précédente que pour tout $y\in[0,1]$,

$$\mathbb{P}_\theta(\tfrac{\hat b-\hat a}{b-a}\leqslant y)=ny^{n-1}-(n-1)y^n,$$

qui ne dépend pas de a ou b. Donc $\frac{\hat{b}-\hat{a}}{b-a}$ est une variable aléatoire pivotale.

d) Soit $\alpha \in]0,1[$ fixé. On note x_α l'unique $x \in [0,1]$ tel que $nx^{n-1} - (n-1)x^n = 1-\alpha$. On déduit de la question précédente que, pour tout θ,

$$\mathbb{P}_\theta(0 \leqslant \tfrac{\hat{b}-\hat{a}}{b-a} \leqslant x_\alpha) = 1-\alpha,$$

et donc

$$\mathbb{P}_\theta\left(b-a \in \left[\frac{\hat{b}-\hat{a}}{x_\alpha}, +\infty\right[\right) = 1-\alpha.$$

On a donc montré que $[\frac{\hat{b}-\hat{a}}{x_\alpha}, +\infty[$ est un intervalle de confiance pour le paramètre $b-a$ au niveau de confiance $1-\alpha$.

Remarque : on aurait également pu définir x'_α tel que $nx^{n-1} - (n-1)x^n = \alpha$ et utiliser que $\mathbb{P}_\theta(x'_\alpha \leqslant \frac{\hat{b}-\hat{a}}{b-a} \leqslant 1) = 1-\alpha$, ce qui aurait conduit à un autre intervalle de confiance de la forme $[\hat{b}-\hat{a}, \frac{\hat{b}-\hat{a}}{x'_\alpha}]$. Même si cet intervalle est beaucoup plus petit que le précédent, ces deux intervalles ont exactement le même niveau de confiance.