Estimation et Identification statistiques
Contrôle : Partie théorique - Correction

Exercice 1 Estimateur du maximum de vraisemblance

1. Par le fait que les variables sont indépendantes et identiquement distribuées, la vraisemblance s'écrit :

$$L(x_1,...,x_n;\theta) = f(X_1 = x_1,...,X_n = x_n;\theta) = \prod_{i=1}^{n} f(X_i = x_i;\theta) = \prod_{i=1}^{n} \tfrac{x_i}{\theta} exp(\tfrac{-x_i}{2\theta}) I_{\mathbb{R}_+}(x_i) \quad (1)$$

Ainsi, la Log-vraisemblance s'écrit :

$$\ln L(x_1,...,x_n;\theta) = \sum_{i=1}^{n} \left(\ln(x_i) - \ln(\theta) - \tfrac{x_i^2}{2\theta}\right) = -n\ln(\theta) + \sum_{i=1}^{n} \left(\ln(x_i)\right) - \tfrac{1}{2\theta}\sum_{i=1}^{n} \left(x_i^2\right) \quad (2)$$

Calculons les dérivées d'ordre 1 et 2 par rapport à θ de cette expression. On trouve :

$$\begin{cases} \frac{\partial \ln L(x_1,...,x_n;\theta)}{\partial\theta} & = -\frac{n}{\theta} + \frac{1}{2\theta^2}\sum_{i=1}^{n} x_i^2 \\ \frac{\partial^2 \ln L(x_1,...,x_n;\theta)}{\partial\theta^2} & = \frac{n}{\theta^2} - \frac{1}{\theta^3}\sum_{i=1}^{n} x_i^2 \end{cases} \quad (3)$$

Ainsi, $\frac{\partial \ln L(x_1,...,x_n;\theta)}{\partial\theta} = 0$ équivaut à :

$$\hat{\theta} = \frac{1}{2n}\sum_{i=1}^{n} x_i^2, \quad (4)$$

et pour cette valeur de θ, on trouve :

$$\frac{\partial^2 \ln L(x_1,...,x_n;\theta)}{\partial\theta^2} = \frac{-2n^3}{[\sum_{i=1}^{n} x_i^2]^2} \quad (5)$$

Ainsi, la fonction de vraisemblance admet un unique extremum, et il s'agit d'un maximum global. D'où l'estimateur du maximum de vraisemblance suivant :

$$\hat{\theta} = \frac{1}{2n}\sum_{i=1}^{n} x_i^2. \quad (6)$$

2. Moyenne :

$$E[\hat{\theta}] = E[\frac{1}{2n}\sum_{i=1}^{n} x_i^2] = \frac{1}{2n}\sum_{i=1}^{n} E[x_i^2] = \frac{1}{2n}\sum_{i=1}^{n} 2\theta = \theta \quad (7)$$

Variance (la troisième égalité provient de l'indépendance des x_i) :

$$\begin{aligned} &Var(\theta) = \tfrac{1}{4n^2} Var(\sum_{i=1}^{n} x_i^2) = \tfrac{1}{4n^2}\sum_{i=1}^{n} Var(x_i^2) + \tfrac{1}{4n^2}\sum_{i=1}^{n}\sum_{j=1, j\neq i}^{n} Covar(x_i^2, x_j^2) = \tfrac{1}{4n} Var(x_1^2) + 0 \\ &= \tfrac{1}{4n}\left(E[x_1^4] - E[x_1^2]^2\right) = \tfrac{1}{4n}\left(8\theta^2 - 4\theta^2\right) = \tfrac{\theta^2}{n} \underset{n\to\infty}{\longrightarrow} 0 \end{aligned} \quad (8)$$

Ainsi, l'estimateur est non biaisé et converge vers θ en moyenne quadratique.

3. On a :
$$\begin{cases} E[\frac{\partial \ln L(x_1,\dots,x_n;\theta)}{\partial\theta}] & = -\frac{n}{\theta} + \frac{1}{2\theta^2}\sum_{i=1}^{n} E[x_i^2] = -\frac{n}{\theta} + \frac{1}{2\theta^2}\sum_{i=1}^{n} 2\theta = 0 \\ E[\frac{\partial^2 \ln L(x_1,\dots,x_n;\theta)}{\partial\theta^2}] & = \frac{n}{\theta^2} - \frac{1}{\theta^3}\sum_{i=1}^{n} E[x_i^2] = \frac{n}{\theta^2} - \frac{1}{\theta^3}\sum_{i=1}^{n} 2\theta = -\frac{n}{\theta^2} \end{cases} \tag{9}$$
Ainsi, la condition de régularité du théorème de Cramer-Rao est vérifiée, et la matrice d'information de Fisher vaut :
$$F(\theta) = -E[\frac{\partial^2 \ln L(x_1,...,x_n;\theta)}{\partial\theta^2}] = \frac{n}{\theta^2}. \tag{10}$$
Finalement, la borne de Cramer-Rao vaut :
$$BCR(\theta) = F(\theta)^{-1} = \frac{\theta^2}{n}, \tag{11}$$
et $\hat{\theta}$ est un estimateur sans biais atteignant la borne de Cramer-Rao, c'est donc un estimateur efficace.

Exercice 2 Estimation Bayésienne

1. La loi a posteriori s'écrit :
$$f(\theta|x_1,\dots,x_n) = \frac{f(x_1,\dots,x_n|\theta)f(\theta)}{f(x_1,\dots,x_n)} = \frac{\left(\prod_{i=1}^{n} \frac{x_i}{\theta} exp(\frac{-x_i^2}{2\theta}) I_{\mathbb{R}_+}(x_i)\right)\left(\frac{\beta^\alpha}{\Gamma(\alpha)}\frac{1}{\theta^{\alpha+1}}\exp\left(\frac{-\beta}{\theta}\right)\mathcal{I}_{\mathbb{R}^+}(\theta)\right)}{f(x_1,\dots,x_n)} \tag{12}$$
2. La loi peut se réécrire :
$$\begin{aligned} & f(\theta|x_1,\dots,x_n) \\ & = \frac{\beta^\alpha}{\Gamma(\alpha)}\frac{\left(\prod_{i=1}^{n} x_i I_{\mathbb{R}_+}(x_i)\right)}{f(x_1,\dots,x_n)} \times \left(\frac{1}{\theta^n}\exp\left(\frac{-1}{2\theta}\sum_{i=1}^{n} x_i^2\right)\right)\left(\frac{1}{\theta^{\alpha+1}}\exp\left(\frac{-\beta}{\theta}\right)\mathcal{I}_{\mathbb{R}^+}(\theta)\right) \\ & = \frac{\beta^\alpha}{\Gamma(\alpha)}\frac{\left(\prod_{i=1}^{n} x_i I_{\mathbb{R}_+}(x_i)\right)}{f(x_1,\dots,x_n)} \times \left(\frac{1}{\theta^{n+\alpha+1}}\exp\left(\frac{-1}{\theta}\left(\beta + \frac{1}{2}\sum_{i=1}^{n} x_i^2\right)\right)\mathcal{I}_{\mathbb{R}^+}(\theta)\right), \end{aligned} \tag{13}$$
ce qui peut se mettre sous la forme d'une loi inverse-gamma avec $\alpha' = n+\alpha$, et $\beta' = \beta + \frac{1}{2}\sum_{i=1}^{n} x_i^2$.
De plus, comme $\alpha > 0$ et $\beta > 0$, on a également $\alpha' > 0$ et $\beta' > 0$.
3. La log Vraisemblance s'écrit :
$$\ln f(\theta|x_1,\dots,x_n) = C - (\alpha'+1)\ln(\theta) - \frac{\beta'}{\theta}, \tag{14}$$
où C est une constante indépendante de θ.
Ainsi :
$$\frac{\partial \ln f(\theta|x_1,\dots,x_n)}{\partial\theta} = -\frac{\alpha'+1}{\theta} + \frac{\beta'}{\theta^2} \tag{15}$$
Cette dérivée s'annule en l'unique valeur $\theta = \frac{\beta'}{\alpha'+1}$, valeur pour laquelle on a :
$$\ln f(\theta|x_1,\dots,x_n) = C - (\alpha'+1)\left(\ln\left(\frac{\beta'}{\alpha'+1}\right)+1\right). \tag{16}$$
De plus, sur l'intervalle $[0;\infty]$ la fonction log-vraisemblance est continue, et $\lim_{\theta\to 0} \ln f(\theta|x_1,\dots,x_n) = -\infty$, et $\lim_{\theta\to+\infty} \ln f(\theta|x_1,\dots,x_n) = -\infty$.
Ainsi, la fonction croit sur l'intervalle $]-\infty;\frac{\beta'}{\alpha'+1}[$, et décroit sur l'intervalle $[\frac{\beta'}{\alpha'+1};\infty[$. En particulier, l'extremum calculé est un maximum.

4. Compte-tenu de ce qui précède, l'expression de l'estimateur au sens du maximum a posteriori s'écrit :

$$\hat{\theta}_{MAP} = \frac{\beta'}{\alpha' + 1} = \frac{\beta + \frac{1}{2}\sum_{i=1}^{n} x_i^2}{\alpha + n + 1}. \tag{17}$$

5. Comme $\hat{\theta}_{MV} = \frac{1}{2n}\sum_{i=1}^{n} x_i^2$ (cf. Exercice 1), on a :

$$\hat{\theta}_{MAP} = \frac{\beta + n\hat{\theta}_{MV}}{\alpha + n + 1} \tag{18}$$

De plus, comme $\lim_{n\to\infty} \theta_{MV} = \theta$ (cf. Exercice 1), on a :

$$\lim_{n\to\infty} \hat{\theta}_{MAP} = \theta. \tag{19}$$

L'estimateur $\hat{\theta}_{MAP}$ se comporte donc comme $\hat{\theta}_{MV}$ lorsque $n \to \infty$. Lorsque le nombre d'observations devient très important, on fait confiance à ces observations et l'influence de la loi a priori devient négligeable.