Tabea Rebafka
October 2018

Correction de la Feuille N° 1

Exercice 1. *Représentation graphique*

(a) Les valeurs sont toutes positives. La loi est asymétrique (la moustache supérieure est plus longue que celle du bas et il y a pas mal de valeurs très élevées.) Ceci peut être une loi Gamma, plus précisément une loi exponentielle.

(b) L'histogramme montre que les observations sont distribuées relativement symétriquement autour de zéro. La loi est unimodale. On peut penser à une loi normale. En revanche, le coefficient d'aplatissement (indiqué en bas du graphique) vaut 3.906 ce qui est trop élevé pour que la loi soit une loi normale. Il est possible que la loi soit une loi de Cauchy (ou plus généralement une loi de Student), qui est symétrique, unimodale, à queues lourdes.

(c) Le graphique est un QQ-plot qui compare l'échantillon avec la loi normale standard. Les points sont bien alignés sur la première bissectrice, ce qui indique que la loi de l'échantillon est probablement la loi normale standard.

(d) Cette fonction de répartition empirique n'a que deux points de saut : en 0 et en 1. Donc, l'échantillon est composé que des 0 et des 1. Il s'agit alors très probablement d'une loi de Bernoulli. En plus, la hauteur des sauts correspond à la proportion des observations dont la valeur est égale au point de saut. Ici, l'hauteur du saut en 0 est de 0.7 environ. Donc, 70% des valeurs dans l'échantillons sont des 0, ce qui implique que le paramètre de la loi de Bernoulli est environ 0.3.

Exercice 2. 1. Soit $X \sim \mathcal{P}(\theta)$ avec $\theta > 0$. On a

$$\mathbb{P}_\theta(X = k) = \frac{\theta^k}{k!}\mathrm{e}^{-\theta}\ , \quad k = 0, 1, \ldots$$

Pour la méthode des moments on calcule le premier moment :

$$\mathbb{E}_\theta[X] = \sum_{k=0}^{n} k\mathbb{P}_\theta(X = k) = \mathrm{e}^{-\theta}\sum_{k=0}^{n} k\frac{\theta^k}{k!} = \mathrm{e}^{-\theta}\sum_{k=1}^{n}\frac{\theta^k}{(k-1)!} = \mathrm{e}^{-\theta}\sum_{k=0}^{n}\frac{\theta^{k+1}}{k!} = \mathrm{e}^{-\theta}\theta\mathrm{e}^{\theta} = \theta\ .$$

Par la méthode des moments, on pose

$$\mathbb{E}_\theta[X] = \theta \quad \Longleftrightarrow \quad \bar{x}_n = \theta\ .$$

L'EMM est alors $\hat{\theta}^{MM} = \bar{x}_n$.

En revanche, pour les valeurs du paramètre θ on a la contrainte $\theta \in]0, +\infty[$. En effet, on a $\bar{x}_n = 0$ si et seulement si $x_i = 0$, pour tout $i = 1, \ldots, n$. Dans ce cas, l'EMM n'existe pas. Ce cas se produit avec une probabilité positive, plus précisément,

$$\mathbb{P}_\theta(\bar{X}_n = 0) = \mathbb{P}_\theta(X_i = 0, i = 1, \ldots, n) = \mathbb{P}_\theta(X_1 = 0)^n = \mathrm{e}^{-n\theta} > 0\ .$$

Pour trouver l'EMV, on écrit la fonction de vraisemblance associée à $\mathbf{x} = (x_1, \ldots, x_n)$:

$$\mathcal{L}(\mathbf{x}; \theta) = \prod_{i=1}^{n}\mathbb{P}_\theta(X_i = x_i) = \prod_{i=1}^{n}\frac{\theta^{x_i}}{x_i!}\mathrm{e}^{-\theta} = \mathrm{e}^{-n\theta}\theta^{\sum_{i=1}^n x_i}\frac{1}{\prod_{i=1}^n (x_i!)}\ .$$

On peut passer à la fonction de log-vraisemblance

$$\ell(\theta) = \log\mathcal{L}(\mathbf{x}; \theta) = -n\theta + \log\theta\sum_{i=1}^{n} x_i - \sum_{i=1}^{n}\log(x_i!)\ .$$

Cette fonction est deux fois dérivable avec

$$\ell'(\theta) = -n + \frac{1}{\theta}\sum_{i=1}^{n} x_i \ , \qquad \ell''(\theta) = -\frac{1}{\theta^2}\sum_{i=1}^{n} x_i \leq 0 \ .$$

La fonction $\ell(\theta)$ est alors concave. Pour la maximiser il suffit de déterminer un point où sa dérivée s'annule. On a

$$\ell'(\theta) = 0 \Longleftrightarrow -n + \frac{1}{\theta}\sum_{i=1}^{n} x_i = 0 \Longleftrightarrow \theta = \bar{x}_n \ .$$

L'EMV est alors $\hat{\theta}^{MV} = \hat{\theta}^{MM} = \bar{x}_n$.
La même remarque que pour l'EMM s'applique : si $\bar{x}_n = 0$, alors l'EMV n'existe pas.

2. Calculons le premier moment de la loi géométrique Geo(θ) avec $0 < \theta < 1$:

$$\mathbb{E}_\theta[X] = \sum_{\ell=0}^{\infty} \ell \mathbb{P}_\theta(X = \ell) = \theta \sum_{\ell=0}^{\infty} \ell(1-\theta)^\ell = \theta \frac{1-\theta}{(1-(1-\theta))^2} = \frac{1-\theta}{\theta} \ .$$

Posons

$$\mathbb{E}_\theta[X] = \bar{x}_n \quad \Longleftrightarrow \quad \frac{1-\theta}{\theta} = \bar{x}_n \quad \Longleftrightarrow \quad \theta = \frac{1}{\bar{x}_n + 1} \ .$$

Seulement si tous les $x_i = 0, i = 1, \dots, n$, on a $\frac{1}{\bar{x}_n+1} = 1 \notin]0,1[$. Dans tous les autres cas l'EMM est bien défini et donné par $\theta^{MM} = \frac{1}{\bar{x}_n+1}$.
Pour la fonction de vraisemblance on a

$$\mathcal{L}(\mathbf{x};\theta) = \prod_{i=1}^{n} \mathbb{P}_\theta(X_i = x_i) = \prod_{i=1}^{n} \theta(1-\theta)^{x_i} = \theta^n (1-\theta)^{\sum_{i=1}^n x_i} \ .$$

Passons à la fonction de log-vraisemblance :

$$\ell(\theta) = \log \mathcal{L}(\mathbf{x};\theta) = n \log\theta + \sum_{i=1}^{n} x_i \log(1-\theta) \ .$$

Ses dérivées sont données par

$$\ell'(\theta) = \frac{n}{\theta} - \sum_{i=1}^{n} x_i \frac{1}{1-\theta} \ , \qquad \ell''(\theta) = -\frac{n}{\theta^2} - \frac{\sum_{i=1}^n x_i}{(1-\theta)^2} < 0 \ , \forall \theta \in]0,1[\ .$$

La fonction $\ell(\theta)$ est donc concave, et tout point critique est alors maximum global de $\ell(\theta)$. Pour le point critique de $\ell(\theta)$ on trouve

$$\ell'(\theta) = 0 \Leftrightarrow \frac{n}{\theta} - \sum_{i=1}^{n} x_i \frac{1}{1-\theta} = 0 \Leftrightarrow \theta = \frac{1}{\bar{x}_n + 1} \ .$$

On en déduit que $\ell(\theta)$ est maximal en $\frac{1}{\bar{x}_n+1}$. Donc, l'EMV est donné par $\hat{\theta}^{MV} = \frac{1}{\bar{x}_n+1}$, et il coïncide avec l'EMM. Seulement, si $\bar{x}_n = 0$, l'EMV n'existe pas.

3. Soit $X \sim \mathcal{N}(\mu, \sigma^2)$. Alors

$$\mathbb{E}_{\mu,\sigma^2}[X] = \mu \ , \qquad \mathbb{E}^2_{\mu,\sigma}[X^2] = \sigma^2 + \mu^2 \ .$$

Par la méthode des moments on pose le système de deux équations

$$\mathbb{E}_{\mu,\sigma^2}[X] = \bar{x}_n \ , \quad \mathbb{E}_{\mu,\sigma^2}[X^2] = \frac{1}{n}\sum_{i=1}^{n} x_i^2 \Longleftrightarrow \mu = \bar{x}_n \ , \quad \sigma^2 + \mu^2 = \frac{1}{n}\sum_{i=1}^{n} x_i^2$$

$$\Longleftrightarrow \mu = \bar{x}_n \ , \quad \sigma^2 = \frac{1}{n}\sum_{i=1}^{n} x_i^2 - \bar{x}_n^2 = s_{\mathbf{x}}^2 \ .$$

L'EMM de (μ, σ^2) est alors donné par $(\bar{x}_n, s^2_{\mathbf{x}})$.
Pour l'EMV, on calcule la fonction de vraisemblance

$$\mathcal{L}(x_1, \ldots, x_n; \mu, \sigma^2) = \prod_{i=1}^{n} f_{\mu,\sigma^2}(x_i) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{(x_i-\mu)^2}{2\sigma^2}\right\}$$
$$= \frac{1}{(2\pi\sigma^2)^{n/2}} \exp\left\{-\frac{\sum_{i=1}^{n}(x_i-\mu)^2}{2\sigma^2}\right\} .$$

La fonction de log-vraisemblance est donnée par

$$\ell(\mu, \sigma^2) = \log(\mathcal{L}(x_1, \ldots, x_n; \mu, \sigma^2)) = -\frac{n}{2}\log(2\pi) - \frac{n}{2}\log(\sigma^2) - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2 .$$

La dérivée première et seconde par rapport à μ sont

$$\frac{\partial}{\partial\mu}\ell(\mu,\sigma^2) = \frac{1}{\sigma^2}\sum_{i=1}^{n}(x_i-\mu) , \qquad \frac{\partial^2}{\partial^2\mu}\ell(\mu,\sigma^2) = -\frac{n}{\sigma^2} < 0 .$$

Donc, pour tout σ^2 fixé, $\mu \mapsto \ell(\mu, \sigma^2)$ est concave avec maximum global $\hat{\mu}$ donné par

$$\frac{\partial}{\partial\mu}\ell(\hat{\mu},\sigma^2) = 0 \Leftrightarrow \hat{\mu} = \bar{x}_n .$$

Quant à σ, on a

$$\frac{\partial}{\partial\sigma^2}\ell(\hat{\mu},\sigma^2) = -\frac{n}{2\sigma^2} + \frac{1}{2(\sigma^2)^2}\sum_{i=1}^{n}(x_i-\hat{\mu})^2 , \qquad \frac{\partial^2}{\partial^2\sigma^2}\ell(\hat{\mu},\sigma^2) = \frac{n}{2(\sigma^2)^2} - \frac{1}{(\sigma^2)^3}\sum_{i=1}^{n}(x_i-\hat{\mu})^2 .$$

On trouve alors

$$\frac{\partial}{\partial\sigma^2}\ell(\hat{\mu},\hat{\sigma}^2) = 0 \Leftrightarrow \hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i-\hat{\mu})^2 ,$$

et

$$\frac{\partial^2}{\partial^2\sigma^2}\ell(\hat{\mu},\hat{\sigma}^2) = \frac{n}{2\hat{\sigma}^4} - \frac{n\hat{\sigma}^2}{\hat{\sigma}^6} = -\frac{n}{2\hat{\sigma}^4} < 0 .$$

Donc, $\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x}_n)^2$ est un maximum local. Étant l'unique point critique, $\hat{\sigma}^2$ est bien le maximum global.
Donc, $\hat{\mu} = \bar{x}_n$ et $\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x}_n)^2$ sont les EMV de μ et σ^2. L'EMV coïncide alors avec l'EMM.

4. Pour l'EMM on calcule le premier moment. On a

$$\mathbb{E}_\theta[X_1] = \int_{\mathbb{R}} x f_\theta(x)\mathrm{d}x = \int_\theta^\infty x\mathrm{e}^{-(x-\theta)}\mathrm{d}x = \int_0^\infty (u+\theta)\mathrm{e}^{-u}\mathrm{d}u = \int_0^\infty u\mathrm{e}^{-u}\mathrm{d}u + \theta\int_0^\infty \mathrm{e}^{-u}\mathrm{d}u = 1+\theta .$$

En posant $\mathbb{E}_\theta[X_1] = \bar{x}_n$, on obtient pour l'EMM $\hat{\theta}^{MM} = \bar{x}_n - 1$. En fait, cet estimateur peut poser des problèmes, car par la forme de la densité f_θ il est clair que $\theta < X_i$ *p.s.*. Or, il arrive qu'une observation x_j est inférieur à $\bar{x}_n - 1$ avec probabilité positive. Autrement dit, $\mathbb{P}_\theta(\bar{X}_n - 1 < \theta) > 0$.
Pour l'EMV on calcule la fonction de vraisemblance

$$\mathcal{L}(x_1, \ldots, x_n; \theta) = \prod_{i=1}^{n} f_\theta(x_i) = \prod_{i=1}^{n} \mathrm{e}^{-x_i+\theta}\mathbb{1}\{x_i \geq \theta\} = \mathrm{e}^{-\sum_{i=1}^{n} x_i + n\theta}\mathbb{1}\{x_{(1)} \geq \theta\} ,$$

où $x_{(1)} = \min\{x_1, \ldots, x_n\}$. D'une part, la fonction $\theta \mapsto \mathbb{1}\{x_{(1)} \geq \theta\}$ est maximale pour tout $\theta \leq x_{(1)}$. D'autre part, la fonction $\theta \mapsto \mathrm{e}^{-\sum_{i=1}^{n} x_i + n\theta}$ est croissante en θ. Cela implique que le maximum est atteint en $\theta = x_{(1)}$. Donc, l'EMV est $\hat{\theta}_n = x_{(1)}$.

Exercice 3. *Loi de Bernoulli*

1. (a) Pour calculer l'estimateur par la méthode des moments, il faut d'abord calculer le premier moment de X_1 en fonction du paramètre inconnu. Clairement, on a

$$\mathbb{E}_p[X_1] = p \ .$$

Pour obtenir l'estimateur par la méthode des moments, il suffit de remplacer le moment théorique $\mathbb{E}_p[X_1]$ par le moment empirique $\bar{x}_n = \frac{1}{n}\sum_{i=1}^n x_i$: par définition, l'EMM $\hat{p}^{MM}$ est tel que

$$\mathbb{E}_p[X_1] = \bar{x}_n \ ,$$

c'est-à-dire $\hat{p}^{MM} = \bar{x}_n$.

Une remarque importante est que $\bar{X}_n$ n'est pas forcément à valeurs $p.s.$ dans l'espace $p =]0,1[$, en effet, comme le modèle est discret, on peut avoir avec probabilité non nulle $\bar{X}_n \in \{0,1\}$. Ceci génère une difficulté quant à l'utilisation de l'EMM ; en toute rigueur, la méthode des moments ne fournit pas de solution pour le cas $\bar{X}_n \in \{0,1\}$. Ainsi, la réponse rigoureuse est $\hat{p}^{MM} = \bar{x}_n$ si $\bar{X}_n \in]0,1[$ et $\hat{p}^{MM}$ non défini si $\bar{x}_n \in \{0,1\}$.

Pour calculer l'estimateur du maximum de vraisemblance, on remarque que pour tout $x \in \{0,1\}$

$$\mathbb{P}_p(X_i = x) = p^x(1-p)^{1-x} \ .$$

D'où, pour tout $(x_1, \ldots, x_n) \in \{0,1\}^n$,

$$\mathcal{L}(x_1, \ldots, x_n; p) = \prod_{i=1}^n \mathbb{P}_p(X_i = x_i) = \prod_{i=1}^n p^{x_i}(1-p)^{1-x_i} = p^{\sum_{i=1}^n x_i}(1-p)^{n-\sum_{i=1}^n x_i} \ .$$

Par suite,

$$\ell(p) = \log(\mathcal{L}(x1, \ldots, x_n, p)) = \log(p)\sum_{i=1}^n x_i + \log(1-p)(n - \sum_{i=1}^n x_i) \ .$$

Il s'agit à présent de minimiser la fonction $\ell(p) : p \in]0,1[\mapsto \ell(p)$. Pour tout $p \in]0,1[$,

$$\ell'(p) = \frac{\sum_{i=1}^n x_i}{p} - \frac{n - \sum_{i=1}^n x_i}{1-p} \ .$$

On en déduit que pour tout $p \in]0,1[$, $\ell'(p) \geq 0 \Leftrightarrow p \geq \bar{x}_n$. Ainsi, lorsque $\bar{x}_n \in]0,1[$, le minimum de $p \mapsto \ell(p)$ est atteint en $\bar{x}_n$. Dans le cas où $\bar{x}_n \in \{0,1\}$, par contre, la fonction $\ell(p) : p \in]0,1[\mapsto \ell(p)$ n'a pas de minimum.

Finalement, l'EMV est $\hat{p}^{MV} = \bar{x}_n$ lorsque $\bar{x}_n \in]0,1[$ et l'EMV est non défini si $\bar{x}_n \in \{0,1\}$.

Par conséquent, l'EMV coïncide avec l'EMM.

(b) Pour le risque quadratique de $\hat{p}^{MV} = \bar{X}_n$ on a

$$\begin{aligned}\mathcal{R}(\hat{p}^{MV}, p) &= \mathbf{Var}_p(\hat{p}^{MV}) + \left(\mathbb{E}_p[\hat{p}^{MV}] - p\right)^2 = \mathbf{Var}(\bar{X}_n) + \left(\mathbb{E}_p[\bar{X}_n] - p\right)^2 \\ &= \frac{1}{n}\mathbf{Var}(X_1) + (p-p)^2 = \frac{1}{n}p(1-p) \ .\end{aligned}$$

On voit que le risque quadratique tend vers 0 lorsque n tend vers l'infini. Cela implique la consistance de l'estimateur. De plus, d'après le TCL, l'estimateur $\hat{p}^{MV} = \bar{X}_n$ est asymptotiquement normal :

$$\sqrt{n}(\hat{p}^{MV} - p) = \sqrt{n}(\bar{X}_n - \mathbb{E}_p[X_1]) \xrightarrow{\mathcal{L}} \mathcal{N}(0, \mathbf{Var}(X_1)) = \mathcal{N}(0, p(1-p)) \ , \quad n \to \infty \ .$$

2. (a) On a

$$v = p(1-p) = \mathbb{E}_p[X_1](1 - \mathbb{E}_p[X_1]) \ .$$

Par la méthode de substitution on estime $\mathbb{E}_p[X_1]$ par $\bar{x}_n$. Donc, $\hat{v}_n = \bar{x}_n(1 - \hat{x}_n)$ est un estimateur par la méthode de substitution de v.

Par la loi des grands nombres, on a $\bar{X}_n \xrightarrow{P} p$ lorsque $n \to \infty$. Par le 1er théorème de continuité on en déduit que

$$\hat{v}_n = \bar{X}_n(1-\bar{X}_n) \xrightarrow{P} p(1-p) = v \;, \quad n \to \infty \;.$$

D'où la consistence de $\hat{v}_n$.

(b) On vérifie que

$$\begin{aligned}
\mathbb{E}_p[\hat{v}_n] &= \mathbb{E}_p[\bar{X}_n(1-\bar{X}_n)] = \mathbb{E}_p[\bar{X}_n] - \mathbb{E}_p[\bar{X}_n^2] = p - \frac{1}{n^2}\mathbb{E}_p\left[\sum_{i=1}^n X_i^2 + \sum_{i\neq j} X_i X_j\right] \\
&= p - \frac{1}{n}\mathbb{E}_p\left[X_1^2\right] - \frac{n-1}{n}\left(\mathbb{E}_p\left[X_1\right]\right)^2 = p - \frac{1}{n}p - \frac{n-1}{n}p^2 \\
&= \frac{n-1}{n}p(1-p) = \frac{n-1}{n}v \neq v \;.
\end{aligned}$$

On en déduit que

$$\tilde{v}_n = \frac{n}{n-1}\hat{v}_n = \frac{n}{n-1}\bar{X}_n(1-\bar{X}_n)$$

est un estimateur sans biais de v.

(c) Pour la loi limite de $\tilde{v}_n$ on trouve

$$\sqrt{n}(\tilde{v}_n - v) = \sqrt{n}\left(\frac{n}{n-1}\hat{v}_n - v\right) = \underbrace{\frac{n}{n-1}}_{\to 1}\sqrt{n}\,(\hat{v}_n - v) + \underbrace{\frac{\sqrt{n}}{n-1}}_{\to 0}v \;.$$

Donc, par le lemme de Slutsky, la loi limite de $\tilde{v}_n$ coïncide avec la loi limite de $\hat{v}_n$. Déterminons la loi limite de $\hat{v}_n$. Notons $h(u) = u(1-u)$. Clairement, h est continue et dérivable avec $h'(u) = 1-2u$. Par la delta méthode, on a

$$\begin{aligned}
\sqrt{n}(\hat{v}_n - v) &= \sqrt{n}\left(\bar{X}_n(1-\bar{X}_n) - p(1-p)\right) \\
&= \sqrt{n}\left(h(\bar{X}_n) - h(p)\right) \\
&\xrightarrow{\mathcal{L}} \mathcal{N}(0, (h'(p))^2\mathbf{Var}(X_1)) = \mathcal{N}(0, (1-2p)^2 p(1-p)) \;, \quad n \to \infty \;.
\end{aligned}$$

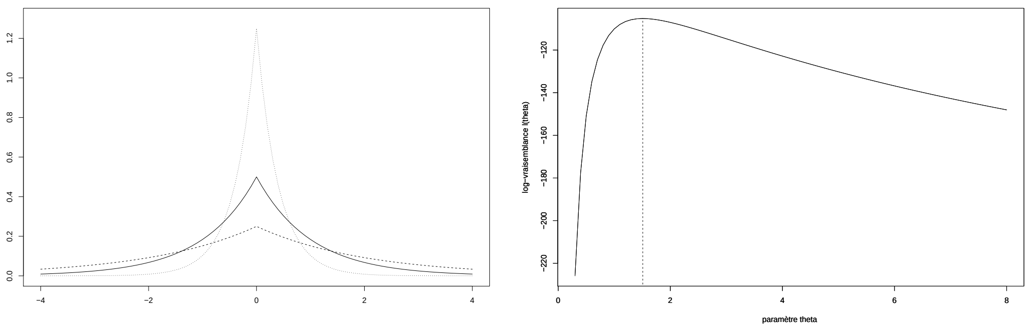
Alternativement, on peut obtenir ce résultat de la façon suivante :

$$\begin{aligned}
\sqrt{n}(\hat{v}_n - v) &= \sqrt{n}\left(\bar{X}_n(1-\bar{X}_n) - p(1-p)\right) \\
&= \sqrt{n}(\bar{X}_n - p - \bar{X}_n^2 + p^2) \\
&= \sqrt{n}(\bar{X}_n - p) - \sqrt{n}(\bar{X}_n - p + p)^2 + \sqrt{n}p^2 \\
&= \sqrt{n}(\bar{X}_n - p) - \sqrt{n}(\bar{X}_n - p)^2 - 2p\sqrt{n}(\bar{X}_n - p) - \sqrt{n}p^2 + \sqrt{n}p^2 \\
&= \underbrace{(1-2p)\underbrace{\sqrt{n}(\bar{X}_n - p)}_{\xrightarrow{\mathcal{L}}\mathcal{N}(0,p(1-p))}}_{\xrightarrow{\mathcal{L}}\mathcal{N}(0,(1-2p)^2p(1-p))} - \underbrace{\underbrace{\frac{1}{\sqrt{n}}}_{\to 0}\underbrace{\left(\underbrace{\sqrt{n}(\bar{X}_n - p)}_{\xrightarrow{\mathcal{L}}\xi\sim\mathcal{N}(0,p(1-p))}\right)^2}_{\xrightarrow{\mathcal{L}}\xi^2}}_{\xrightarrow{P}0} \\
&\xrightarrow{\mathcal{L}} \mathcal{N}(0, (1-2p)^2 p(1-p)) \;, \quad n \to \infty \;,
\end{aligned}$$

par le TCL, le 1er théorème de continuité et le lemme de Slutsky.

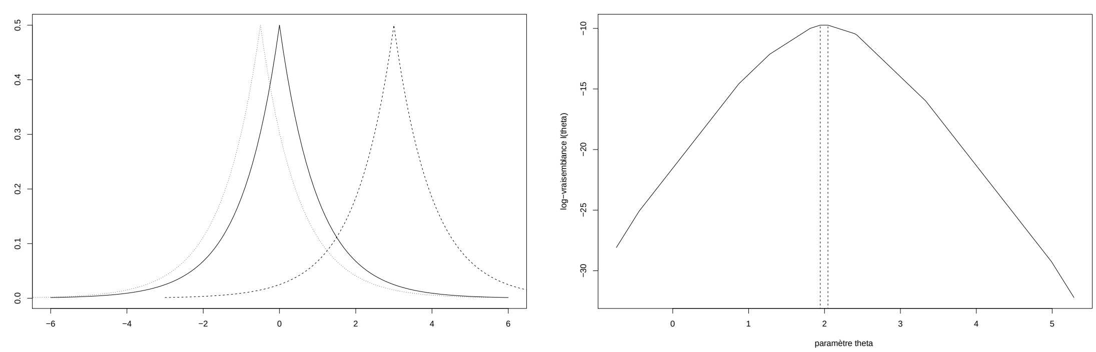
Exercice 4. *Loi double exponentielle*

1. (a) Voir Figure 1 (a) pour l'allure de la densité.



(a) (b)

FIGURE 1 – (a) Densité $f_\theta(x) = \frac{1}{2\theta}\mathrm{e}^{-\frac{|x|}{\theta}}$ de la loi double exponentielle pour des différentes valeurs de θ. (b) Fonction de log-vraisemblance $\ell(\theta)$ pour un échantillon de taille $n = 50$ avec un maximum global en $\hat{\theta}^{MV} = 1{,}51$.



(a) (b)

FIGURE 2 – (a) Densité $f_\mu(x) = \frac{1}{2}\mathrm{e}^{-|x-\mu|}$ de la loi double exponentielle pour des différentes valeurs de θ. (b) Fonction de log-vraisemblance $\ell(\theta)$ (moins une constante) pour un échantillon de taille $n = 10$ avec l'intervalle autour de 1,99 où la fonction est maximale.

(b) La fonction de vraisemblance s'écrit

$$\mathcal{L}(\mathbf{x};\theta)=\prod_{i=1}^n f_\theta(x_i)=\frac{1}{2^n\theta^n}\mathrm{e}^{-\frac{1}{\theta}\sum_{i=1}^n |x_i|}\ .$$

En passant au log on obtient la fonction de log-vraisemblance

$$\ell(\theta)=-n\log 2-n\log\theta-\frac{1}{\theta}\sum_{i=1}^n|x_i|\ .$$

La fonction $\ell(\theta)$ est deux fois dérivable et ses dérivées sont données par

$$\ell'(\theta)=-\frac{n}{\theta}+\frac{1}{\theta^2}\sum_{i=1}^n|x_i| \qquad \ell''(\theta)=\frac{n}{\theta^2}-\frac{2}{\theta^3}\sum_{i=1}^n|x_i|=\frac{1}{\theta^2}\left(n-\frac{2}{\theta}\sum_{i=1}^n|x_i|\right)\ .$$

On calcule

$$\ell'(\theta)=0\Leftrightarrow 0=-n\theta+\sum_{i=1}^n|x_i|\Leftrightarrow\theta=\frac{1}{n}\sum_{i=1}^n|x_i|\ .$$

On vérifie qu'il s'agit d'un maximum de $\ell(\theta)$ par la dérivée seconde :

$$\ell''\left(\frac{1}{n}\sum_i|x_i|\right)=\frac{n^2}{(\sum_{i=1}^n|x_i|)^2}(n-2n)<0\ .$$

La fonction de vraisemblance atteint donc son maximum en $\frac{1}{n}\sum_{i=1}^n|x_i|$. L'EMV est alors $\hat\theta^{MV}=\frac{1}{n}\sum_{i=1}^n|x_i|$. Il est bien défini et unique. Figure 1 (b) illustre la fonction $\ell(\theta)$ pour un échantillon i.i.d. de taille 50. On voit bien que la fonction est régulière et a un seul maximum. Ici le vrai paramètre est $\theta_0=1{,}5$, et l'EMV vaut 1,51.
On a par la loi des grands nombres,

$$\hat\theta_n^{MV}=\frac{1}{n}\sum_{i=1}^n|X_i|\xrightarrow{P}\mathbb{E}_\theta[|X_1|]\ ,\quad n\to\infty\ .$$

Par symétrie, on a

$$\mathbb{E}_\theta[|X_1|]=\int_{\mathbb{R}}|x|f_\theta(x)\mathrm{d}x=\frac{1}{2\theta}\int_{\mathbb{R}}|x|\mathrm{e}^{-|x|/\theta}\mathrm{d}x=\frac{1}{\theta}\int_0^\infty x\mathrm{e}^{-x/\theta}\mathrm{d}x=\theta\int_0^\infty u\mathrm{e}^{-u}\mathrm{d}u=\theta\ .$$

D'où la consistance de l'EMV.
De plus, l'EMV est asymptotiquement normal, car d'après le théorème central limite, on a

$$\sqrt{n}(\hat\theta_n^{MV}-\theta)=\sqrt{n}\left(\frac{1}{n}\sum_{i=1}^n|X_i|-\mathbb{E}_\theta[|X_1|]\right)\xrightarrow{\mathcal{L}}\mathcal{N}(0,\mathbf{Var}_\theta(|X_1|)\ ,\quad n\to\infty\ .$$

On a

$$\mathbb{E}_\theta[|X_1|^2]=\int_{\mathbb{R}}x^2f_\theta(x)\mathrm{d}x=\frac{1}{2\theta}\int_{\mathbb{R}}x^2\mathrm{e}^{-|x|/\theta}\mathrm{d}x=\frac{1}{\theta}\int_0^\infty x^2\mathrm{e}^{-x/\theta}\mathrm{d}x=\theta^2\int_0^\infty u^2\mathrm{e}^{-u}\mathrm{d}u=\theta^2\Gamma(3)=2\theta^2\ .$$

D'où $\mathbf{Var}_\theta(|X_1|)=2\theta^2-\theta^2=\theta^2$.

2. 2.1 Figure 2 (a) montre l'allure de la densité.

   2.2 La fonction de vraisemblance est donnée par

$$\mathcal{L}(\mathbf{x};\mu)=\prod_{i=1}^n f_\mu(x_i)=\frac{1}{2^n}\mathrm{e}^{-\sum_{i=1}^n|x_i-\mu|}\ ,$$

avec fonction de log-vraisemblance

$$\ell(\mu) = -n \log 2 - \sum_{i=1}^{n} |x_i - \mu| \ .$$

On constate que

$$\arg\max_{\mu \in \mathbb{R}} \ell(\mu) = \arg\min_{\mu \in \mathbb{R}} \sum_{i=1}^{n} |x_i - \mu| \ .$$

Maximiser $\ell(\theta)$ est alors équivalent à chercher le point minimum de $g(\mu) = \sum_{i=1}^{n} |x_i - \mu|$. Clairement, la fonction $g(\mu)$ n'est pas dérivable en μ. On vérifie que g est continue et linéaire par morceaux. En effet, en utilisant la notation des statistiques d'ordre,

$$g(\mu) = \begin{cases} \sum_{i=1}^{n}(x_i - \mu) = c_1 - n\mu \ , & \text{si } \mu \leq x_{(1)} \\ (\mu - x_{(1)}) + \sum_{i=2}^{n}(x_{(i)} - \mu) = c_2 - (n-2)\mu \ , & \text{si } \mu \in [x_{(1)}, x_{(2)}] \\ \sum_{i=1}^{2}(\mu - x_{(i)}) + \sum_{i=3}^{n}(x_{(i)} - \mu) = c_3 - (n-4)\mu \ , & \text{si } \mu \in [x_{(2)}, x_{(3)}] \\ \vdots & \\ \sum_{i=1}^{n}(\mu - x_i) = c_n + n\mu \ , & \text{si } \mu \geq x_{(n)} \end{cases}$$

où les c_i sont des constantes. On observe que la pente de g passe progressivement d'une valeur négative à une valeur positive, ce qui implique que g est convexe.

Si n est impair, le signe de la pente de g change en $x_{((n+1)/2)}$. Elle est donc minimale en $x_{((n+1)/2)}$, et l'EMV vaut $\hat{\mu}^{MV} = x_{((n+1)/2)}$.

Si n est pair, la pente de g s'annule sur tout l'intervalle $]x_{(n/2)}, x_{(n/2+1)}[$. Alors, g est minimale sur $[x_{(n/2)}, x_{n/2+1}]$, et donc tout point de cet intervalle est EMV de μ. On a ici un exemple qui montre que l'EMV n'est pas toujours unique. Figure 2 (b) donne la fonction $\ell(\mu)$ pour un échantillon de taille 10. On voit que la fonction atteint son maximum sur un intervalle.

Afin d'étudier la consistance et la loi limite de l'EMV, il faut manipuler des statistiques d'ordre, ce qui n'est pas toujours évident. Les théorèmes classiques comme la loi des grands nombres ou le théorème central limite ne s'applique pas.

Néanmoins, concernant la consistance, on constate que μ est la médiane $q_{\alpha/2}$ de la loi f_μ. De plus, la médiane empirique $\hat{q}_{\alpha/2}^n = X_{(\lceil n/2 \rceil)}$ est l'EMV de μ. Par une proposition du cours sur les quantiles empiriques, on a $\hat{q}_{\alpha/2}^n \xrightarrow{P} q_{\alpha/2}$ lorsque n tend vers l'infini. D'où la consistance de l'EMV.

Quant à la loi limite, on peut noter qu'il y a un théorème de la limite centrale pour les quantiles, mais c'est hors programme.