EMV de la loi uniforme

Akita & Richi plaisir

ENS Rennes, 2013-2014

Référence : Cours de statistique de M1 de F.Malrieu.

Développement pour les leçons :
- 260. Espérance, variance et moments d'une variable aléatoire.
- 262. Modes de convergence d'une suite de variables aléatoires. Exemples et applications.
- 263. Variables aléatoires à densité. Exemples et applications.

Résultats :

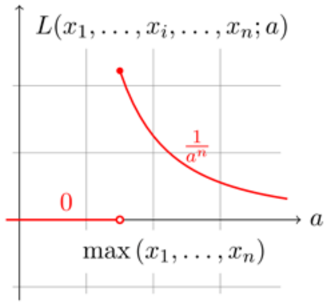
Soient $X_1, \ldots, X_n$ des variables aléatoires de même loi uniforme $\mathcal{U}_{[0,\theta]}$, avec $\theta \in \mathbb{R}_+^*$. Notons $\hat{\theta}_n$ l'estimateur du maximum de vraisemblance du paramètre θ. Nous allons montrer les résultats suivants :

1. $\hat{\theta}_n = X_{(n)} := \max\limits_{1\leq i\leq n} X_i$,
2. le biais est : $\mathbb{E}(\hat{\theta}_n) = \dfrac{n}{n+1}\theta$,
3. consistance : $\hat{\theta}_n \xrightarrow[n]{\text{ps}} \theta$,
4. vitesse et loi limite : $\dfrac{n}{\theta}(\hat{\theta}_n - \theta) \xrightarrow{\mathcal{L}} Y$, où $-Y \sim \mathcal{E}\left(\dfrac{1}{\theta}\right)$,
5. risque quadratique : $\mathbb{E}((\hat{\theta}_n - \theta)^2) = \dfrac{2\theta^2}{(n+1)(n+2)}$.

Démonstration :

1) La vraisemblance des observations $(x_1, \ldots x_n)$ est :

$$L(x1, \ldots, x_n; \theta) = \prod_{i=1}^n f_\theta(x_i) = \prod_{i=1}^n \frac{1}{\theta}\mathbb{1}_{[0,\theta]} = \begin{cases} 0, & \text{si } \theta \leq \max(x_1, \ldots, x_n) \\ \\ \dfrac{1}{\theta^n}, & \text{si } \theta \geq \max(x_1, \ldots, x_n) \end{cases}.$$



Vraisemblance d'un échantillon suivant la loi uniforme

D'où l'estimateur cherché : $\hat{\theta}_n = X_{(n)}$.

2) Calculons la loi de $\hat{\theta}_n$ en utilisant les fonctions de répartition et les densités :

$$\forall t \geq 0,\ F_{\hat{\theta}_n}(t) := \mathbb{P}(\hat{\theta}_n \leq t) = \mathbb{P}(X_i \leq t, \forall i \in [\![1,n]\!]) = \mathbb{P}(X_1 \leq t)^n = \begin{cases} \left(\dfrac{t}{\theta}\right)^n, & \text{si } t \leq \theta, \\ 1, & \text{sinon} \end{cases}.$$

D'où la densité de $\hat{\theta}_n$: $f_{\hat{\theta}_n}(t) = \dfrac{n}{\theta}\left(\dfrac{t}{\theta}\right)^n \mathbb{1}_{[0,\theta]}(t)$

Nous pouvons maintenant calculer le biais :

$$\mathbb{E}(\hat{\theta}_n) = \int_0^\theta t f_{\hat{\theta}_n}(t)\mathrm{d}t = n\int_0^\theta \left(\frac{t}{\theta}\right)^n \mathrm{d}t = \frac{n}{n+1}\theta.$$

3) Soit $\varepsilon > 0$. Calculons :

$$\mathbb{P}(|\hat{\theta}_n - \theta| \geq \epsilon) = 1 - \mathbb{P}(\theta - \varepsilon \leq \hat{\theta}_n \leq \theta + \epsilon) = 1 - (F_{\hat{\theta}_n}(\theta+\varepsilon) - F_{\hat{\theta}_n}(\theta-\varepsilon)) = \left(\frac{\theta-\varepsilon}{\theta}\right)^n = \left(1 - \frac{\varepsilon}{\theta}\right)^n.$$

$1 - \varepsilon/\theta < 1$ donc la série $\sum \mathbb{P}\left(|\hat{\theta}_n - \theta| \geq \varepsilon\right)$ converge et d'après le lemme de Borel-Cantelli :

$$\hat{\theta}_n \xrightarrow[n]{\text{ps}} \theta.$$

4) Soit $t \in \mathbb{R}$. Calculons en utilisant les fonctions de répartition :

$$\begin{aligned}
\mathbb{P}(n(\theta - \hat{\theta}_n) \geq t) &= \mathbb{P}\left(\hat{\theta}_n \leq \theta - \frac{t}{n}\right) \\
&= \left(\frac{\theta - t/n}{\theta}\right)^n \\
&= \left(1 - \frac{t}{n\theta}\right)^n \\
&= \exp\left(n\ln\left(1 - \frac{t}{n\theta}\right)\right) \underset{n}{\longrightarrow} \mathrm{e}^{-t/\theta}.
\end{aligned}$$

Donc : $n(\hat{\theta}_n - \theta) \xrightarrow{\mathcal{L}} -Y$, où $Y \sim \mathcal{E}\left(\frac{1}{\theta}\right)$.

5) Nous avons :

$$\begin{aligned}
\mathbb{E}(\hat{\theta}_n^2) &= \int_0^{+\infty} \mathbb{P}(\hat{\theta}_n^2 > t)\mathrm{d}t \\
&= \int_0^{+\infty} \mathbb{P}(\hat{\theta}_n > \sqrt{t})\mathrm{d}t \\
&= \int_0^{\theta^2} \left(1 - \left(\frac{\sqrt{t}}{\theta}\right)^n\right)\mathrm{d}t \\
&= \left[t - \frac{t^{n/2+1}}{\theta^n(n/2+1)}\right]_0^{\theta^2} \\
&= \theta^2 - \frac{\theta^2}{n/2+1}.
\end{aligned}$$

D'où le risque quadratique :

$$\begin{aligned}
\mathbb{E}((\hat{\theta}_n - \theta)^2) &= \mathbb{E}(\hat{\theta}_n^2 - 2\theta\hat{\theta}_n + \theta^2) \\
&= \theta^2 - \frac{\theta^2}{n/2+1} - 2\theta\frac{n}{n+1}\theta + \theta^2 \\
&= \theta^2\left(1 - \frac{1}{n/2+1} - \frac{2n}{n+1} + 1\right) \\
&= \frac{2\theta^2}{(n+1)(n+2)}.
\end{aligned}$$

Remarques :

- 2) : c'était prévisible que l'estimateur serait de moyenne plus petite que θ car $\hat{\theta}_n$ prend des valeurs plus petites que θ,
- 4) : l'estimateur n'est donc pas asymptotiquement normal,
- ici la log-vraisemblance n'est pas définie.