Corrigé TD 5
Février 2011

Exercice 1. Soit X une v.a. de densité :

$$f(x) = \frac{1}{2}e^{-|x-\theta|}, x \in \mathbb{R}$$

où θ est un paramètre réel inconnu.

1.1. Calculer $E_\theta[X]$ et Var$_\theta[X]$. En déduire un estimateur T_n de θ.
Calcul de l'espérance :

$$\begin{aligned} E_\theta[X] &= \frac{1}{2}\int_{-\infty}^{+\infty} xe^{-|x-\theta|}dx \\ &= \frac{1}{2}\int_{-\infty}^{+\infty} (u+\theta)e^{-|u|}du \\ &= \frac{\theta}{2}\int_{-\infty}^{+\infty} e^{-|u|}du \\ &= \theta \end{aligned}$$

De même, on calcule le moment d'ordre 2 :

$$E_\theta[X^2] = 2 + \theta^2,$$

puis on en déduit la variance de X : Var$_\theta(X) = 2$.

D'après la méthode des moments, on en déduit que la moyenne empirique $\bar{X}_n$ est un estimateur de θ. De plus, c'est un estimateur convergent par la loi forte des grands nombres.

1.2. Construire un intervalle de confiance de niveau asymptotique 95% pour θ dans le cas où $n = 200$.

D'après le Théorème Central Limite, on obtient :

$$\sqrt{n}\left(\frac{\bar{X}_n - E_\theta[X]}{\sqrt{Var_\theta(X)}}\right) \Rightarrow^{\text{loi}} \mathcal{N}(0,1).$$

on obtient l'intervalle de confiance asymptotique suivant :

$$\text{IC}(\theta) = \left[\bar{X}_n - u_{\frac{\alpha}{2}}\sqrt{\frac{2}{n}}\ ,\ \bar{X}_n + u_{\frac{\alpha}{2}}\sqrt{\frac{2}{n}}\right]$$

Exercice 2. On rappelle que dans le modèle uniforme $\mathcal{P} = \{\mathcal{U}[0,\theta], \theta > 0\}$, l'Estimateur du Maximum de Vraisemblance de θ est $\hat{\theta}_n = \max(X_1, ..., X_n)$.

2.1. Pour $x \in \mathbb{R}$, calculer $\mathbb{P}_\theta(\frac{\hat{\theta}_n}{\theta} \leq x)$ et en déduire que la loi de $\frac{\hat{\theta}_n}{\theta}$ ne dépend pas de θ.

On a déjà vu en cours que $X_{(n)}$ admet comme densité $y \mapsto \frac{n}{\theta^n} y^{n-1} \mathbf{1}_{[0,\theta]}(y)$. Comme θ est un paramètre d'échelle, $\frac{X_{(n)}}{\theta}$ a pour densité $z \mapsto n z^{n-1} \mathbf{1}_{[0,1]}(z)$.

Donc $\frac{\theta}{X_{(n)}}$ a pour densité $f : t \mapsto \frac{n}{t^{n+1}} \mathbf{1}_{[1,\infty]}(t)$. Ainsi, $\frac{\theta}{X_{(n)}}$ est bien une statistique libre pour θ. Sa densité est bien unimodale puisque constante (et égale à 0 avant 1 et strictement décroissant après 1. La valeur 1 est l'unique mode et point de discontinuité.

2.2. Construire un intervalle de confiance de niveau $1 - \alpha$ pour θ.

On cherche le plus court intervalle $[a, b]$ tel que

$$P(a \leq \frac{\theta}{X_{(n)}} \leq b) = \int_a^b f(t)dt = 1 - \alpha$$

Comme f est une fonction strictement décroissante sur $]1, \infty[$, nécessairement $a = 1$, alors

$$\int_1^b \frac{n}{t^{n+1}} dt = 1 - \alpha$$

soit

$$b = \alpha^{-1/n}$$

$P(1 \leq \frac{\theta}{X_{(n)}} \leq \alpha^{-1/n}) = 1 - \alpha$ s'écrit

$$P(X_{(n)} \leq \theta \leq \alpha^{-1/n} X_{(n)}) = 1 - \alpha$$

autrement, l'IC s'écrit

$$IC =]X_{(n)} \,;\, \alpha^{-1/n} X_{(n)}[$$

Exercice 3. 3.1. La vraisemblance associée à l'échantillon est définie par :

$$L(x, \theta) \;=\; \exp\left(-\sum_{i=1}^{n}(x_i - \theta)\right) \Pi_{i=1}^{n} \mathbf{1}_{[\theta,\infty[}(x_i)$$

3.2. $L(x, \theta)$ est maximum en θ ssi $\Pi_{i=1}^{n} \mathbf{1}_{[\theta,\infty[}(x_i) = 1$. Autrement dit,

$\forall i = 1, ..., n,\ \theta \leq x_i$, donc $\inf_{1 \leq i \leq n} x_i \geq \theta$. D'où $\hat{\theta}_n = \inf_{1 \leq i \leq n} X_i$ est l'EMV.

3.3. La loi densité de la loi de $\hat{\theta}_n$ est définie par :

$$f_{\hat{\theta}_n}(x, \theta) = n e^{-n(x-\theta)} \mathbf{1}_{[\theta,\infty[}(x)$$

D'où $E[\hat{\theta}_n] = \theta - \frac{1}{n}$.

3.4. $T_n = \hat{\theta}_n - \theta$ est de loi densité :

$$f_{T_n}(x, \theta) = n e^{-nx} \mathbf{1}_{[0,\infty[}(x)$$

C'est donc une statistique libre pour θ.

3.5. Pour un α donné, on cherche un $A = A(\hat{\theta}_n)$ tel que $P(A \leq \theta \leq \hat{\theta}_n) = 1 - \alpha$. On a :

$$P(A \leq \theta \leq \hat{\theta}_n) = P(0 \leq T_n \leq \hat{\theta}_n - A) = F_{T_n}(\hat{\theta}_n - A) = 1 - \alpha$$

Soit $A = \hat{\theta}_n - F_{T_n}^{-1}(1-\alpha)$ soit $A = \hat{\theta}_n + \frac{1}{n}\log\alpha$.

L'intervalle de confiance est :

$$\left[\hat{\theta}_n + \frac{1}{n}\log\alpha) \ ; \ \hat{\theta}_n\right]$$

3.6. et 3.7. A faire vous même !

6 Estimation par intervalle de confiance

1. (a) La moyenne empirique est $\overline{x} = 159$ et la médiane $x_{(4)} = 156$.

   (b) A partir de la fonction pivotale

$$\frac{\overline{X} - \mu}{\sigma/\sqrt{n}} \sim \mathcal{N}(0,1)$$

   on obtient l'intervalle bilatéral suivant :

$$1 - \alpha = P\left(\overline{X} - \frac{\sigma}{\sqrt{n}} u_{1-\alpha/2} < \mu < \overline{X} + \frac{\sigma}{\sqrt{n}} u_{1-\alpha/2}\right).$$

   A.N. : [151.59,166.41]

   (c) La longueur de l'intervalle est

$$\ell = \frac{2\sigma}{\sqrt{n}} u_{1-\alpha/2}.$$

   La condition $\ell \leq \epsilon$ équivaut donc à

$$n \geq \frac{4\sigma^2 u_{1-\alpha/2}^2}{\epsilon^2}.$$

   A.N. : $n \geq 62$.

   (d) Si σ^2 est inconnu, on utilise la fonction pivotale :

$$\frac{\overline{X} - \mu}{S^*/\sqrt{n}} \sim \mathcal{T}_{n-1}$$

   qui conduit à l'intervalle :

$$1 - \alpha = P\left(\overline{X} - \frac{S^*}{\sqrt{n}} t_{n-1,1-\alpha/2} < \mu < \overline{X} + \frac{S^*}{\sqrt{n}} t_{n-1,1-\alpha/2}\right).$$

   A.N. : [144.41,173.59]

   (e) Dans le cas où μ est connue, on utilise la fonction pivotale

$$\frac{n\widehat{\sigma}^2}{\sigma^2} \sim \chi^2_n$$

   avec

$$\widehat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \mu)^2.$$

   On a

$$1 - \alpha = P\left[\frac{n\widehat{\sigma}^2}{\sigma^2} > \chi^2_{n,\alpha}\right] = P\left[\sigma^2 < \frac{n\widehat{\sigma}^2}{\chi^2_{n,\alpha}}\right].$$

   A.N. : [0,690.8].

   Dans le cas où μ est inconnue, on utilise la fonction pivotale

$$\frac{(n-1)S^{*2}}{\sigma^2} \sim \chi^2_{n-1}.$$

   On a

$$1 - \alpha = P\left[\frac{(n-1)S^{*2}}{\sigma^2} > \chi^2_{n-1,\alpha}\right] = P\left[\sigma^2 < \frac{(n-1)S^{*2}}{\chi^2_{n-1,\alpha}}\right].$$

   A.N. : [0,909.8].

2. Il y a $n = 50$ enfants. On a donc $\widehat{p} = 15/50 = 0.3$.