

CC2 de statistique

Exercice 1 : On considère des variables aléatoires i.i.d $X_1, \dots, X_n$ avec $X_1 \sim U[0, \sqrt{\theta}]$ où $\theta > 0$.

1. Donner l'espérance de X_1 .
2. En déduire un estimateur de θ , que l'on notera θ_n .
3. Est-il convergent ?
4. Donner sa normalité asymptotique.
5. Soit $\alpha \in (0, 1)$, construire un intervalle de confiance asymptotique de niveau $1 - \alpha$.
6. Soit $\theta_0 > 0$, en déduire un test asymptotique de niveau α pour tester

$$H_0 : \theta = \theta_0 \quad \text{contre} \quad H_1 : \theta \neq \theta_0$$

7. Montrer que l'estimateur du maximum de vraisemblance est $\hat{\theta}_n = X_{(n)}^2$, où $X_{(n)} = \max X_i$.
8. Donner la convergence en loi de $n(\theta - \hat{\theta}_n)$.
9. Construire un intervalle de confiance de niveau $1 - \alpha$ de la forme

$$[\hat{\theta}_n; \hat{\theta}_n c_\alpha].$$

10. Soit $\theta_0 > 0$, construire un test de taille α pour tester

$$H_0 : \theta \geq \theta_0 \quad \text{contre} \quad \theta < \theta_0.$$

1. **0.5 pts** On a $\mathbb{E}[X_1] = \sqrt{\theta}/2$.
2. **0.5 point** On a $\theta = (2\mathbb{E}[X_1])^2 =: g(\mathbb{E}[X_1])$. On considère donc l'estimateur

$$\theta_n = g(\bar{X}_n) = (2\bar{X}_n)^2.$$

3. **1pt** Par la loi des grands nombres, on a

$$\bar{X}_n \xrightarrow[n \rightarrow +\infty]{\mathbb{P}} \sqrt{\theta}/2$$

et comme la fonction g est continue en $\sqrt{\theta}/2$, on a par le théorème de continuité

$$\theta_n = g(\bar{X}_n) \xrightarrow[n \rightarrow +\infty]{\mathbb{P}} g(\sqrt{\theta}/2) = \theta.$$

4. **2pts** Comme $\mathbb{V}[X_1] = \theta/12$, le TCL nous donne

$$\sqrt{n} \left(\bar{X}_n - \sqrt{\theta}/2 \right) \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} \mathcal{N} \left(0, \frac{\theta}{12} \right).$$

De plus, pour tout $x > 0$, on a $g'(x) = 8x$ et en particulier g est dérivable en $\sqrt{\theta}/2$ avec $g'(\sqrt{\theta}/2) = 4\sqrt{\theta}$. Par la méthode delta, on obtient donc

$$\sqrt{n} \left(g(\bar{X}_n) - g\left(\frac{\sqrt{\theta}}{2}\right) \right) = \sqrt{n} (\theta_n - \theta) \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} \mathcal{N} \left(0, \frac{4\theta^2}{3} \right).$$

5. **3 pts** On peut réécrire la normalité asymptotique comme

$$\frac{\sqrt{3n}}{2} \frac{\theta_n - \theta}{\theta} \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} \mathcal{N}(0, 1).$$

Option 1 : On a

$$\frac{\sqrt{3n}}{2} \left(\frac{\theta_n}{\theta} - 1 \right) \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} \mathcal{N}(0, 1)$$

et donc, en notant $q_{1-\alpha/2}$ le quantile d'ordre $1 - \alpha/2$ de la loi normale centrée réduite, on obtient

$$\begin{aligned} 1 - \alpha &= \mathbb{P} \left[-q_{1-\alpha/2} \leq \frac{\sqrt{3n}}{2} \left(\frac{\theta_n}{\theta} - 1 \right) \leq q_{1-\alpha/2} \right] \\ &= \mathbb{P} \left[\frac{1 - 2q_{1-\alpha/2}/\sqrt{3n}}{\theta_n} \leq \frac{1}{\theta} \leq \frac{1 + 2q_{1-\alpha/2}/\sqrt{3n}}{\theta_n} \right] \\ &= \mathbb{P} \left[\frac{\theta_n}{1 - 2q_{1-\alpha/2}/\sqrt{3n}} \geq \theta \geq \frac{\theta_n}{1 + 2q_{1-\alpha/2}/\sqrt{3n}} \right] \end{aligned}$$

la dernière égalité n'étant vraie que pour n suffisamment grand.

Option 2 : On a

$$\frac{\sqrt{3n}}{2} \frac{\theta_n - \theta}{\theta_n} = \frac{\theta}{\theta_n} \frac{\sqrt{3n}}{2} \frac{\theta_n - \theta}{\theta}.$$

L'application $h : x \mapsto x^{-1}$ est continue en θ , et donc, par le théorème de continuité,

$$\frac{1}{\theta_n} \xrightarrow[n \rightarrow +\infty]{\mathbb{P}} \frac{1}{\theta}$$

et donc en particulier

$$\frac{\theta}{\theta_n} \xrightarrow[n \rightarrow +\infty]{\mathbb{P}} 1.$$

Par le théorème de Slutsky, on obtient donc

$$\frac{\sqrt{3n}}{2} \frac{\theta_n - \theta}{\theta_n} = \frac{\theta}{\theta_n} \frac{\sqrt{3n}}{2} \frac{\theta_n - \theta}{\theta} \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} \mathcal{N}(0, 1)$$

et on obtient donc l'intervalle de confiance

$$\left[\theta_n \pm 2q_{1-\alpha/2} \frac{\theta_n}{\sqrt{3n}} \right]$$

où $q_{1-\alpha/2}$ est le quantile d'ordre $1 - \alpha/2$ de la loi normale centrée réduite.

6. **1pt** Le test consiste en

$$\text{"On rejette " } H_0 \iff \theta_0 \notin IC_{1-\alpha}(\theta) \iff |\theta_n - \theta_0| > 2q_{1-\alpha/2} \frac{\theta_n}{\sqrt{3n}}.$$

On a donc la zone de rejet

$$ZR = \left\{ |\theta_n - \theta_0| > 2q_{1-\alpha/2} \frac{\theta_n}{\sqrt{3n}} \right\}.$$

7. **2 pts** On a la vraisemblance

$$L_{\mathbf{X}}(\theta) = \prod_{i=1}^n \frac{1}{\sqrt{\theta}} \mathbf{1}_{X_i \leq \theta^2} = \frac{1}{\theta^{n/2}} \mathbf{1}_{X_{(n)} \leq \sqrt{\theta}}$$

et comme $X_{(n)}$ et θ^2 sont positifs, on obtient

$$L_{\mathbf{X}}(\theta) = \frac{1}{\theta^{n/2}} \mathbf{1}_{\theta \geq X_{(n)}^2},$$

qui est une fonction nulle sur $[0, X_{(n)}^2)$, et positive et strictement décroissante sur $[X_{(n)}^2, +\infty)$, et le maximum est donc atteint en $X_{(n)}^2$, i.e $\hat{\theta}_n = X_{(n)}^2$.

8. **2 pts** On a $\theta - \hat{\theta}_n$ qui est positif, et donc pour tout $t > 0$, on a

$$\mathbb{P} \left[n (\theta - \hat{\theta}_n) \leq t \right] = \mathbb{P} \left[\hat{\theta}_n \geq \theta - \frac{t}{n} \right]$$

Pour n suffisamment grand, i.e. tel que $t/n \leq \theta$, on a

$$\begin{aligned}
 \mathbb{P} \left[\hat{\theta}_n \geq \theta - \frac{t}{n} \right] &= \mathbb{P} \left[X_{(n)} \geq \sqrt{\theta - \frac{t}{n}} \right] \\
 &= 1 - \mathbb{P} \left[X_{(n)} \leq \sqrt{\theta - \frac{t}{n}} \right] \\
 &= 1 - \left(\mathbb{P} \left[X_1 \leq \sqrt{\theta - \frac{t}{n}} \right] \right)^n \\
 &= 1 - \left(\frac{\sqrt{\theta - t/n}}{\sqrt{\theta}} \right)^n \\
 &= 1 - \left(1 - \frac{t}{\theta n} \right)^{n/2}.
 \end{aligned}$$

On a donc

$$\mathbb{P} \left[n \left(\theta - \hat{\theta}_{(n)} \right) \leq t \right] = 1 - \exp \left(\frac{n}{2} \log \left(1 - \frac{t}{\theta n} \right) \right) \xrightarrow{n \rightarrow +\infty} 1 - e^{-t/2\theta}$$

et on reconnaît la fonction de répartition d'une loi exponentielle de paramètre $1/2\theta$, i.e. on a

$$n(\theta - \theta_n) \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} \mathcal{E} \left(\frac{1}{2\theta} \right).$$

9. **2pts** On cherche $c_\alpha \geq 1$ tel que

$$\begin{aligned}
 1 - \alpha &= \mathbb{P} [\theta \leq \theta_n c_\alpha] \iff 1 - \alpha = \mathbb{P} \left[\theta_n \geq \frac{\theta}{c_\alpha} \right] \\
 &\iff 1 - \alpha = \mathbb{P} \left[X_{(n)} \geq \sqrt{\frac{\theta}{c_\alpha}} \right] \\
 &\iff \alpha = \mathbb{P} \left[X_{(n)} \leq \sqrt{\frac{\theta}{c_\alpha}} \right] \\
 &\iff \alpha = \left(\mathbb{P} \left[X_1 \leq \sqrt{\frac{\theta}{c_\alpha}} \right] \right)^n = \frac{1}{c_\alpha^{n/2}}.
 \end{aligned}$$

On a donc $c_\alpha = \alpha^{-2/n}$.

10. **2 pts** Sous H_0 , comme θ_n est aussi grand que l'on veut, on a

$$\alpha = \sup_{\theta \geq \theta_0} \mathbb{P}_\theta [\text{On rejette } H_0] = \sup_{\theta \geq \theta_0} \mathbb{P}_\theta [\theta_n \leq c_\alpha] = \sup_{\theta \geq \theta_0} \mathbb{P}_\theta [X_{(n)} \leq \sqrt{c_\alpha}] = \left(\sqrt{\frac{c_\alpha}{\theta}} \right)^n$$

en prenant $c_\alpha \leq \theta_0$. Comme cette fonction est décroissante en θ , on a le sup qui est atteint

en θ_0 , et on obtient donc

$$\alpha = \left(\frac{c_\alpha}{\theta_0} \right)^{n/2} \iff c_\alpha = \theta_0 \alpha^{2/n}.$$

On obtient donc la zone de rejet

$$ZR = \left\{ \theta_n \leq \theta_0 \alpha^{2/n} \right\}.$$