Sorbonne Université
Master Mathématiques et Applications
MU4MA015 : Statistique

6 octobre 2023
Aucun document autorisé
Durée : 1 heure

Contrôle Continu

Rappels

— Il sera tenu compte du soin apporté à la présentation et à la rédaction.
— Si vous n'arrivez pas à démontrer un résultat, vous pouvez l'admettre pour la suite de l'exercice.

Exercice 1 Soient $\theta > 0$ et $X_1, \ldots, X_n$ des variables aléatoires i.i.d. de loi uniforme sur $[0, \theta]$. On note $X_{(n)} = \max_{1 \le i \le n} X_i$.

1. Déterminer la loi de $X_{(n)}$ et donner sa densité par rapport à la mesure de Lebesgue.
Correction. On a que pour tout $x \in [0, \theta]$,
$$\mathbb{P}\left[X_{(n)} \le x\right] = \mathbb{P}\left[X_1 \le x\right]^n = \left(\frac{x}{\theta}\right)^n$$
et la densité est donc définie pour tout x par
$$f_\theta(x) = \frac{n}{\theta}\left(\frac{x}{\theta}\right)^{n-1} \mathbf{1}_{[0,\theta]}(x).$$

2. Déterminer la loi limite de $n(\theta - X_{(n)})$ lorsque n tend vers l'infini.
Correction. Pour $t < 0$, on a $\mathbb{P}(n(\theta - X_{(n)}) > t) = 1$, et, pour $t \ge 0$,
$$\mathbb{P}\left(n(\theta - X_{(n)}) > t\right) = \mathbb{P}\left(X_{(n)} < \theta - \frac{t}{n}\right) = \mathbf{1}_{0 \le t \le \theta n}\left(1 - \frac{t}{\theta n}\right)^n \underset{n \to +\infty}{\longrightarrow} e^{-t/\theta}.$$
Ainsi $n(\theta - X_{(n)}) \xrightarrow{\mathcal{L}} \mathcal{E}(1/\theta)$.

3. Calculer $\mathbb{E}[X_{(n)}]$ et en déduire un estimateur sans biais de θ. On le notera $\widehat{\theta}_n$.
Correction. On trouve $\mathbb{E}[X_{(n)}] = \frac{n}{n+1}\theta$, ce qui conduit à proposer $\widehat{\theta}_n = \frac{n+1}{n} X_{(n)}$.

4. Quelle est la loi limite de $n(\theta - \widehat{\theta}_n)$?
Correction. On a
$$n(\theta - \widehat{\theta}_n) = n(\theta - X_{(n)}) - X_{(n)}.$$
Or $n(\theta - X_{(n)}) \xrightarrow{\mathcal{L}} \mathcal{E}(1/\theta)$, ce qui implique en particulier $X_{(n)} \xrightarrow{\mathbb{P}} \theta$, donc, par le lemme de Slutsky,
$$n(\theta - \widehat{\theta}_n) \xrightarrow{\mathcal{L}} \mathcal{E}(1/\theta) - \theta.$$

Exercice 2 Soient $\theta > 0$ et $X_1, \dots, X_n$ des variables aléatoires i.i.d. d'espérance $\sqrt{\theta}$ et de variance 1.

1. Proposer un estimateur consistant de θ par la méthode des moments. On le notera $\widehat{\theta}_n$.
   Correction. Puisque $\mathbb{E}_\theta[X_1] = \sqrt{\theta} = \varphi(\theta)$, on propose $\widehat{\theta}_n = \varphi^{-1}(\overline{X}_n) = (\overline{X}_n)^2$. La consistance vient de la loi des grands nombres et du théorème de continuité.
2. Déterminer la loi limite de $\sqrt{n}(\widehat{\theta}_n - \theta)$ et en déduire, pour tout $\alpha \in]0,1[$, un intervalle de confiance de niveau asymptotique $1-\alpha$ pour θ.
   Correction. Par le TCL, on sait que $\sqrt{n}(\overline{X}_n - \sqrt{\theta}) \xrightarrow{\mathcal{L}} \mathcal{N}(0,1)$, et la méthode Delta appliquée à la fonction $x \mapsto x^2$ donne
   $$\sqrt{n}(\widehat{\theta}_n - \theta) \xrightarrow{\mathcal{L}} \mathcal{N}(0, 4\theta)\,.$$
   De plus, le lemme de Slutsky garantit que
   $$\frac{\sqrt{n}(\widehat{\theta}_n - \theta)}{2\sqrt{\widehat{\theta}_n}} \xrightarrow{\mathcal{L}} \mathcal{N}(0,1)\,.$$
   Ainsi, en notant $\Phi^{-1}(1-\alpha/2)$ le quantile d'ordre $1-\alpha/2$ de la loi $\mathcal{N}(0,1)$, l'intervalle
   $$\left[\widehat{\theta}_n - \frac{2\sqrt{\widehat{\theta}_n}\Phi^{-1}(1-\alpha/2)}{\sqrt{n}}\,,\, \widehat{\theta}_n + \frac{2\sqrt{\widehat{\theta}_n}\Phi^{-1}(1-\alpha/2)}{\sqrt{n}}\right]$$
   est un intervalle de confiance de niveau asymptotique $1-\alpha$ pour θ.
3. On souhaite tester
   $$H_0 \,:\, \theta = 1 \qquad \text{contre} \qquad H_1 \,:\, \theta \neq 1\,.$$
   Proposer un test T_n de niveau asymptotique $\alpha \in]0,1[$.
   Correction. Le lien entre intervalles de confiance et tests assure que le test T_n donné par
   $$T_n(X_1, \dots, X_n) = \mathbf{1}_{\left\{|\widehat{\theta}_n - 1| > \frac{2\sqrt{\widehat{\theta}_n}\Phi^{-1}(1-\alpha/2)}{\sqrt{n}}\right\}}$$
   est de niveau asymptotique α.
4. Soit $\overline{x}_n$ une réalisation de $\overline{X}_n = \frac{1}{n}\sum_{i=1}^n X_i$. Exprimer la p-valeur α_0 du test T_n. Pour $n = 100$, et $\overline{x}_n = 0.8$, on trouve $\alpha_0 \simeq 2.4\%$. Rejette-t-on H_0 au niveau 5% ? Au niveau 1% ?

   Correction. Il est clair que
   $$\alpha_0 = 2\left(1 - \Phi\left(\frac{\sqrt{n}}{2}\left||\overline{x}_n| - \frac{1}{|\overline{x}_n|}\right|\right)\right)\,.$$
   Comme $\alpha_0 \simeq 2.4\%$, on rejette H_0 au niveau 5%, mais pas au niveau 1%.

Exercice 3 Soit X une variable aléatoire de densité f définie par
$$\forall x \in \mathbb{R}\,,\; f(x) = |x|\mathbf{1}_{[-1,1]}(x)\,,$$
et, pour $\theta \in \mathbb{R}$, soient $X_1, \dots, X_n$ des variables aléatoires i.i.d. de même loi que $X + \theta$. Pour $p \in]0,1[$, on note x_p le quantile d'ordre p de la loi des X_i, et $x_p(n)$ le quantile empirique.

1. Donner la fonction de répartition de X, en déduire celle de X_1.
Correction. On a

$$F_X(x) = \frac{1}{2}(1 - x^2)\mathbf{1}_{[-1,0[}(x) + \frac{1}{2}(1 + x^2)\mathbf{1}_{[0,+1[}(x) + \mathbf{1}_{[1,+\infty[}(x),$$

donc, en notant F_θ la fonction de répartition de X_1,

$$F_\theta(x) = \frac{1}{2}(1 - (x-\theta)^2)\mathbf{1}_{[\theta-1,\theta[}(x) + \frac{1}{2}(1 + (x-\theta)^2)\mathbf{1}_{[\theta,\theta+1[}(x) + \mathbf{1}_{[\theta+1,+\infty[}(x).$$

2. Déterminer $x_{1/2}$. Pourquoi n'est-il pas judicieux d'estimer θ par $x_{1/2}(n)$?
Correction. La médiane est donc $x_{1/2} = \theta$. Cependant, ce n'est pas une très bonne idée d'estimer θ par $x_{1/2}(n)$ car la densité est nulle en θ. En particulier, le résultat de normalité asymptotique ne s'applique pas.
3. Déterminer $x_{3/4}$ et en déduire un estimateur de θ, que l'on notera $\widehat{\theta}_n$.
Correction. On vérifie que $x_{3/4} = \theta + 1/\sqrt{2}$, ce qui conduit à proposer $\widehat{\theta}_n = x_{3/4}(n) - 1/\sqrt{2}$.
4. Quelle est la loi limite de $\sqrt{n}(\widehat{\theta}_n - \theta)$?
Correction. Par le théorème de normalité asymptotique des quantiles empiriques, puisque la densité en $x_{3/4}$ vaut $1/\sqrt{2} > 0$, on a

$$\sqrt{n}(\widehat{\theta}_n - \theta) = \sqrt{n}(x_{3/4}(n) - x_{3/4}) \xrightarrow{\mathcal{L}} \mathcal{N}\left(0, \frac{\frac{1}{4} \times \frac{3}{4}}{\left(\frac{1}{\sqrt{2}}\right)^2}\right) = \mathcal{N}\left(0, \frac{3}{8}\right).$$