Correction du TD 2

Exercice 1

1. On a

$$\mathbb{E}\big(\overline{X}_n\big) = \frac{1}{n}\sum_{i=1}^{n}\mathbb{E}(X_i) = m \ \text{ et } \ \mathbb{E}\big(Z_n\big) = \frac{1}{2}\big(\mathbb{E}(X_n) + \mathbb{E}(X_{n-1})\big) = m,$$

   les deux estimateurs sont donc sans biais pour l'estimation de m.

2. Le risque quadratique d'un estimateur $\widehat{m}$ de m est donné par

$$R(\widehat{m}, m) = (\mathbb{E}(\widehat{m}) - m)^2 + \mathrm{Var}(\widehat{m}).$$

   On compare les risques quadratiques des deux estimateurs. On a d'une part

$$\begin{aligned} R(\overline{X}_n, m) &= 0 + \mathrm{Var}(\overline{X}_n) = \mathrm{Var}\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \frac{1}{n^2}\sum_{i=1}^{n}\mathrm{Var}(X_i) \text{ (par indépendance)} \\ &= \frac{1}{n}\mathrm{Var}(X_1) = \frac{\sigma^2}{n}, \end{aligned}$$

   et d'autre part

$$R(Z_n, m) = 0 + \frac{1}{4}\mathrm{Var}(X_n + X_{n-1}) = \frac{\sigma^2}{2}.$$

   Pour tout $n \geq 3$, pour tout $m \in \mathbb{R}$, on a $R(\overline{X}_n, m) < R(Z_n, m)$, l'estimateur $\overline{X}_n$ est donc préférable à Z_n au sens du risque quadratique. Ne pas oublier de comparer les risques quadratiques pour toutes les valeurs possibles du paramètre inconnu, ici $m \in \mathbb{R}$.

3. Le risque de l'estimateur (stupide) W_n est égal à m^2 donc il est préférable à $\overline{X}_n$ pour des valeurs de paramètre extrêmes $|m| \leq \sigma/\sqrt{n}$. Ainsi, $\overline{X}_n$ n'est pas toujours préférable à $W_n = 0$, ce qui montre bien la limite de la méthode consistant à comparer les risques des estimateurs pour toutes les valeurs possibles du paramètre (c'est-à-dire uniformément). Remarquons tout de même que cet effet s'estompe avec n tendant vers l'infini.

Exercice 2

1. Comme $\mathbb{E}X_1 = \theta$ et $\mathrm{Var}\, X_1 = \theta$, le théorème de la limite centrale appliqué aux variables X_i i.i.d. de carré intégrable donne

$$\sqrt{n}(\widehat{\theta}_n - \theta) \xrightarrow[n\to\infty]{d} \mathcal{N}(0, \theta).$$

   Par le théorème de continuité et le lemme de Slutsky, on a donc

$$\sqrt{n}\frac{\widehat{\theta}_n - \theta}{\sqrt{\widehat{\theta}_n}} \xrightarrow[n\to\infty]{d} \mathcal{N}(0, 1).$$

En notant $q := \Phi^{-1}(0.975) \approx 2$, un intervalle de confiance asymptotique à 95% pour θ est donc

$$\left[\widehat{\theta}_n - q\frac{\sqrt{\widehat{\theta}_n}}{\sqrt{n}} \ ; \ \widehat{\theta}_n + q\frac{\sqrt{\widehat{\theta}_n}}{\sqrt{n}}\right].$$

2. Pour toute fonction g dérivable en tout $\theta > 0$, nous pouvons appliquer la méthode Delta :

$$\sqrt{n}(g(\widehat{\theta}_n) - g(\theta)) \xrightarrow[n\to\infty]{d} \mathcal{N}(0, \theta(g'(\theta))^2).$$

Pour "stabiliser la variance", choisissons g de sorte que, pour tout $\theta > 0$,

$$\theta(g'(\theta))^2 = 1.$$

Ceci incite à choisir g comme une primitive de $\theta \mapsto 1/\sqrt{\theta}$, ainsi $g(\theta) = 2\sqrt{\theta}$ convient. Avec ce choix, nous trouvons

$$\sqrt{n}\left(2\sqrt{\widehat{\theta}_n} - 2\sqrt{\theta}\right) \xrightarrow[n\to\infty]{d} \mathcal{N}(0,1),$$

ce qui « stabilise » bien la variance asymptotique, toujours égale à 1 quel que soit $\theta > 0$. On en déduit que

$$\mathbb{P}\left(-q \leq \sqrt{n}\left(2\sqrt{\widehat{\theta}_n} - 2\sqrt{\theta}\right) \leq q\right) \xrightarrow[n\to\infty]{} 0.95,$$

ce qui implique que

$$\left[\left(\max\left(0, \sqrt{\widehat{\theta}_n} - \frac{q}{2\sqrt{n}}\right)\right)^2 \ ; \ \left(\sqrt{\widehat{\theta}_n} + \frac{q}{2\sqrt{n}}\right)^2\right]$$

est un intervalle de confiance asymptotique à 95% pour θ. On peut en fait simplifier la borne de gauche puisque

$$\max\left(0, \sqrt{\widehat{\theta}_n} - \frac{q}{2\sqrt{n}}\right) = 0 \iff n\widehat{\theta}_n \leq \frac{q^2}{4}$$

or par la loi des grands nombres $\widehat{\theta}_n$ tend p.s. vers $\theta > 0$ donc $n\widehat{\theta}_n$ tend vers l'infini p.s., donc aussi en probabilité [1], d'où l'on déduit en particulier que

$$\mathbb{P}\left(n\widehat{\theta}_n \leq \frac{q^2}{4}\right) \xrightarrow[n\to\infty]{} 0.$$

Par conséquent

$$\left[\left(\sqrt{\widehat{\theta}_n} - \frac{q}{2\sqrt{n}}\right)^2 \ ; \ \left(\sqrt{\widehat{\theta}_n} + \frac{q}{2\sqrt{n}}\right)^2\right]$$

est un intervalle de confiance asymptotique à 95% pour θ.

3. Les intervalles précédents ne sont valables qu'asymptotiquement, c'est-à-dire lorsque n tend vers l'infini, or dans ce cas, on a pour la borne supérieure du dernier intervalle :

$$\left(\sqrt{\widehat{\theta}_n} + \frac{q}{2\sqrt{n}}\right)^2 = \widehat{\theta}_n + q\frac{\sqrt{\widehat{\theta}_n}}{\sqrt{n}} + \frac{q^2}{4n} \approx \widehat{\theta}_n + q\frac{\sqrt{\widehat{\theta}_n}}{\sqrt{n}},$$

1. Une suite de variables (X_n) tend en probabilité (respectivement p.s.) vers l'infini si $\forall M \in \mathbb{R}$ on a $\mathbb{P}(X_n \leq M) \xrightarrow[n\to\infty]{} 0$ (respectivement $\mathbb{P}(\{\omega \in \Omega,\ X_n(\omega) \xrightarrow[n\to\infty]{} \infty\}) = 1$), et comme d'habitude la convergence p.s. implique la convergence en probabilité.

qui correspond à la borne supérieure du second intervalle. Le fait que $\widehat{\theta}_n$ tende presque sûrement vers $\theta > 0$ assure en effet que $1/n$ est asymptotiquement négligeable devant $\sqrt{\widehat{\theta}_n}/\sqrt{n}$:

$$\frac{1/n}{\sqrt{\widehat{\theta}_n}/\sqrt{n}} = \frac{1}{\sqrt{n}} \times \frac{1}{\sqrt{\widehat{\theta}_n}} \xrightarrow[n\to\infty]{p.s.} 0.$$

Bref, la méthode de stabilisation de la variance n'apporte rien de plus qu'un simple plug-in, lequel ne requiert aucun calcul de primitive.

Remarque : Lorsqu'une suite U_n converge en loi vers une variable U de fonction de répartition F_U **continue** sur $\mathbb{R}$, nous avons pour $-\infty \leq a \leq b \leq +\infty$, avec les conventions usuelles $F(-\infty) = 0$ et $F(+\infty) = 1$ pour toute fonction de répartition :

$$\mathbb{P}\left(a \leq U_n \leq b\right) = F_{U_n}(b) - F_{U_n}(a) \xrightarrow[n\to\infty]{} F_U(b) - F_U(a).$$

Lorsque $U \sim \mathcal{N}(0,1)$ (ce qui est souvent le cas), il suffit de prendre pour $\beta \in [0,\alpha]$, $b = \Phi^{-1}(1-\beta)$ et $a = \Phi^{-1}(\alpha - \beta)$ pour en déduire que

$$\mathbb{P}\left(a \leq U_n \leq b\right) \xrightarrow[n\to\infty]{} 1 - \beta - (\alpha - \beta) = 1 - \alpha.$$

On a donc une infinité de possibilités, conduisant à une infinité d'intervalles de confiance. L'option la plus classique est de choisir $\beta = \alpha/2$, autrement dit $b = \Phi^{-1}(1-\alpha/2)$ et $a = -\Phi^{-1}(1-\alpha/2) = \Phi^{-1}(\alpha/2)$ car par symétrie de la loi gaussienne centrée, c'est ce qui donne l'intervalle de confiance le plus court. A contrario, les choix $\beta = \alpha$ (i.e. $b = \Phi^{-1}(1-\alpha)$ et $a = \Phi^{-1}(0) = -\infty$) ou $\beta = 0$ (i.e. $b = \Phi^{-1}(1) = \infty$ et $a = \Phi^{-1}(\alpha)$) donnent des intervalles de longueur infinie, non pertinents en tant que tels comme intervalles de confiance, mais qui peuvent l'être pour la construction de tests d'hypothèse unilatéraux.

Exercice 3

1. Les X_i $(1 \leq i \leq n)$ suivent une loi normale $\mathcal{N}(0,\theta)$. On sait donc que $\mathrm{Var}(X_i) = \mathbb{E}(X_i^2) = \theta$. Pour estimer θ, on propose l'estimateur suivant :

$$\widehat{\theta}_n = \frac{1}{n}\sum_{i=1}^{n} X_i^2.$$

On a alors $\frac{n\widehat{\theta}_n}{\theta} = \sum_{i=1}^{n}\left(\frac{X_i}{\sqrt{\theta}}\right)^2$. Or comme pour tout $1 \leq i \leq n$, $\frac{X_i}{\sqrt{\theta}} \sim \mathcal{N}(0,1)$ et comme les v.a. X_i sont i.i.d., on en déduit que

$$\frac{n\widehat{\theta}_n}{\theta} \sim \chi_n^2.$$

Cette caractérisation de la loi de $\widehat{\theta}_n$ suffit, cependant on sait aussi qu'une loi du χ_n^2 est une loi $\gamma(\frac{n}{2}, \frac{1}{2})$. Ainsi,

$$\widehat{\theta}_n = \frac{\theta}{n}\frac{n\widehat{\theta}_n}{\theta} \sim \gamma\left(\frac{n}{2}, \frac{n}{2\theta}\right).$$

2. On note $F_{\chi_n^2}$ la f.d.r. de la loi du χ_n^2 et $F_{\chi_n^2}^{-1}$ la fonction quantile, qui est en fait la réciproque au sens usuel de $F_{\chi_n^2}$ car cette dernière est une bijection de $[0,\infty)$ vers $[0,1)$. On peut alors écrire :

$$\mathbb{P}\left(F_{\chi_n^2}^{-1}(\alpha/2) \leq \frac{n\widehat{\theta}_n}{\theta} \leq F_{\chi_n^2}^{-1}(1-\alpha/2)\right) = 1 - \alpha,$$

donc

$$\mathbb{P}\left(\frac{n\widehat{\theta}_n}{F^{-1}_{\chi^2_n}(1-\alpha/2)} \leq \theta \leq \frac{n\widehat{\theta}_n}{F^{-1}_{\chi^2_n}(\alpha/2)}\right) = 1-\alpha.$$

En posant

$$A_1 = \frac{n\widehat{\theta}_n}{F^{-1}_{\chi^2_n}(1-\alpha/2)} \quad \text{et} \quad B_1 = \frac{n\widehat{\theta}_n}{F^{-1}_{\chi^2_n}(\alpha/2)},$$

on en déduit que $\mathbb{P}(\theta < A_1) = \mathbb{P}(\theta > B_1) = \alpha/2$:

$$\begin{aligned}
\mathbb{P}\big(\theta < A_1\big) &= \mathbb{P}\left(\theta < \frac{n\widehat{\theta}_n}{F^{-1}_{\chi^2_n}(1-\alpha/2)}\right) = \mathbb{P}\left(F^{-1}_{\chi^2_n}(1-\alpha/2) < \frac{n\widehat{\theta}_n}{\theta}\right) \\
&= 1 - F_{\chi^2_n}\big(F^{-1}_{\chi^2_n}(1-\alpha/2)\big) = \alpha/2. \\
\mathbb{P}\big(\theta > B_1\big) &= \mathbb{P}\left(\theta > \frac{n\widehat{\theta}_n}{F^{-1}_{\chi^2_n}(\alpha/2)}\right) = \mathbb{P}\left(F^{-1}_{\chi^2_n}(\alpha/2) > \frac{n\widehat{\theta}_n}{\theta}\right) = F_{\chi^2_n}\big(F^{-1}_{\chi^2_n}(\alpha/2)\big) = \alpha/2.
\end{aligned}$$

On en déduit l'intervalle de confiance :

$$I^{(1)}_{1-\alpha} = \left[\frac{n\widehat{\theta}_n}{F^{-1}_{\chi^2_n}(1-\alpha/2)}, \frac{n\widehat{\theta}_n}{F^{-1}_{\chi^2_n}(\alpha/2)}\right].$$

En appliquant un raisonnement similaire et en remarquant que

$$\mathbb{P}\left(0 \leq \frac{n\widehat{\theta}_n}{\theta} \leq F^{-1}_{\chi^2_n}(1-\alpha)\right) = 1-\alpha,$$

on peut aussi choisir

$$A_2 = \frac{n\widehat{\theta}_n}{F^{-1}_{\chi^2_n}(1-\alpha)} \quad \text{et} \quad B_2 = \infty.$$

Il vient alors $\mathbb{P}(\theta < A_2) = \mathbb{P}\left(F^{-1}_{\chi^2_n}(1-\alpha) < \frac{n\widehat{\theta}_n}{\theta}\right) = \alpha$, d'où le nouvel intervalle de confiance :

$$I^{(2)}_{1-\alpha} = \left[\frac{n\widehat{\theta}_n}{F^{-1}_{\chi^2_n}(1-\alpha)}, \infty\right[.$$

De même, en prenant

$$A_3 = 0 \quad \text{et} \quad B_3 = \frac{n\widehat{\theta}_n}{F^{-1}_{\chi^2_n}(\alpha)},$$

on obtient $\mathbb{P}(\theta > B_3) = \mathbb{P}\left(\frac{n\widehat{\theta}_n}{\theta} < F^{-1}_{\chi^2_n}(\alpha)\right) = \alpha$, puis le troisième intervalle :

$$I^{(3)}_{1-\alpha} = \left[0, \frac{n\widehat{\theta}_n}{F^{-1}_{\chi^2_n}(\alpha)}\right].$$

Si, pour un niveau de confiance donné, on mesure la qualité d'un intervalle de confiance par sa longueur, le deuxième est disqualifié d'emblée puisque de longueur infinie. Le premier a pour longueur

$$L_1 = n\widehat{\theta}_n\left(\frac{1}{F^{-1}_{\chi^2_n}(\alpha/2)} - \frac{1}{F^{-1}_{\chi^2_n}(1-\alpha/2)}\right)$$

et le troisième $L_3 = (n\widehat{\theta}_n)/F^{-1}_{\chi^2_n}(\alpha)$. Ainsi, $I^{(1)}_{1-\alpha}$ est plus court que $I^{(3)}_{1-\alpha}$ si

$$\frac{1}{F^{-1}_{\chi^2_n}(\alpha)} + \frac{1}{F^{-1}_{\chi^2_n}(1-\alpha/2)} - \frac{1}{F^{-1}_{\chi^2_n}(\alpha/2)} \geq 0. \tag{1}$$

Dans le cas général, on ne peut rien dire : selon les valeurs de n et α, l'inégalité (1) est vérifiée ou non (voir aussi plus bas sur l'application numérique). Cependant, pour $\alpha \in]0,1[$ fixé et lorsque n tend vers l'infini, on s'attend à ce que $I^{(1)}_{1-\alpha}$ se concentre autour de la vraie valeur inconnue $\theta > 0$, et ce à vitesse $1/\sqrt{n}$ puisque $\widehat{\theta}_n$ est obtenu par moyennisation (voir question suivante), donc il semble raisonnable de supposer que L_1 tende vers 0. En revanche, ce ne sera pas le cas pour L_3 puisque, asymptotiquement, la borne de droite est au moins de l'ordre de θ donc L_3 aussi. Par ailleurs, comme précisé dans la remarque ci-dessus, en dépit de leur longueur, $I^{(2)}_{1-\alpha}$ et $I^{(3)}_{1-\alpha}$ peuvent présenter un intérêt pour la construction de tests unilatéraux (voir ci-dessous).

3. Les X_i suivent une loi normale et admettent donc des moments à tous les ordres, en particulier à l'ordre 4. Les v.a. X_i^2 sont donc dans $\mathbb{L}^2$ et satisfont le TCL :

$$\sqrt{n}\frac{\widehat{\theta}_n - \theta}{\sqrt{\mathrm{Var}(X_1^2)}} = \sqrt{n}\frac{\frac{1}{n}\sum_{i=1}^n X_i^2 - \theta^2}{\sqrt{\mathrm{Var}(X_1^2)}} \xrightarrow[n\to\infty]{d} \mathcal{N}(0,1).$$

De plus, $\mathrm{Var}(X_1^2) = \theta^2 \mathrm{Var}((X_1/\sqrt{\theta})^2) = 2\theta^2$ car la variance d'une χ^2_1 est égale à 2. On peut alors construire un IC asymptotique pour θ. La convergence en loi vue plus haut donne, en notant Φ la fonction de répartition de la $\mathcal{N}(0,1)$,

$$\mathbb{P}\left(-\Phi^{-1}(1-\alpha/2) \leq \sqrt{n}\frac{\widehat{\theta}_n - \theta}{\sqrt{2\theta^2}} \leq \Phi^{-1}(1-\alpha/2)\right) \xrightarrow[n\to\infty]{} 1-\alpha,$$

c'est-à-dire,

$$\mathbb{P}\left(1 - \Phi^{-1}(1-\alpha/2)/\sqrt{n/2} \leq \widehat{\theta}_n/\theta \leq 1 + \Phi^{-1}(1-\alpha/2)/\sqrt{n/2}\right) \xrightarrow[n\to\infty]{} 1-\alpha,$$

on encore

$$\mathbb{P}\left(\frac{\widehat{\theta}_n}{1+\Phi^{-1}(1-\alpha/2)/\sqrt{n/2}} \leq \theta \leq \frac{\widehat{\theta}_n}{1-\Phi^{-1}(1-\alpha/2)/\sqrt{n/2}}\right) \xrightarrow[n\to\infty]{} 1-\alpha.$$

Ainsi un intervalle de confiance asymptotique est

$$J_{1-\alpha} = \left[\frac{\widehat{\theta}_n}{1+\Phi^{-1}(1-\alpha/2)/\sqrt{n/2}}, \frac{\widehat{\theta}_n}{1-\Phi^{-1}(1-\alpha/2)/\sqrt{n/2}}\right],$$

en prenant la convention $[a,b] = \emptyset$ si $b < a$.

4. On a $F^{-1}_{\chi^2_n}(\alpha/2) \approx 3.25$, $F^{-1}_{\chi^2_n}(1-\alpha/2) \approx 20.48$ et $\Phi^{-1}(1-\alpha/2) \approx 2$. Les réalisations des intervalles de confiance sont donc

$$\begin{cases} I^{(1)}_{1-\alpha}(\omega) \approx [0.97, 6.16] & \text{(non-asymptotique)} \\ J_{1-\alpha}(\omega) \approx [1.05, 18.95] & \text{(asymptotique).} \end{cases}$$

On remarque que l'intervalle de confiance asymptotique est différent de l'intervalle de confiance non asymptotique. Ceci est normal car $n = 10$ est trop petit pour que l'approche asymptotique soit valide (la convergence n'a pas eu lieu). Puisque $n = 10$, on préférera bien sûr l'IC non-asymptotique.
Remarque : Pour ces valeurs numériques, on peut voir que $I^{(3)}_{1-\alpha}(\omega)$ est en fait plus court que $I^{(1)}_{1-\alpha}(\omega)$ puisque $I^{(3)}_{1-\alpha}(\omega) \approx [0, 5.08]$.

5. (a) Puisque le test $T_\alpha(\mathbb{X})$ est, par définition, une variable de Bernoulli telle que l'on accepte H_1 (ou que l'on rejette H_0) lorsque $T_\alpha(\mathbb{X}) = 1$ et que H_1 correspond à $\mathbb{1}_{\theta>3} = 1$, on choisit comme test $T_\alpha(\mathbb{X}) = \mathbb{1}_{\widehat{\theta}_n > c_\alpha}$, avec $c_\alpha \geq 0$ un réel à déterminer de sorte que le test soit de niveau α. Or la taille du test est :

$$\begin{aligned}
\sup_{\text{« sous } H_0 \text{ »}} \mathbb{P}_\theta(T_\alpha(\mathbb{X}) = 1) &= \sup_{\theta \leq 3} \mathbb{P}_\theta\big(\widehat{\theta}_n > c_\alpha\big) \\
&= \sup_{\theta \leq 3} \mathbb{P}_\theta\left(\frac{n\widehat{\theta}_n}{\theta} > \frac{nc_\alpha}{\theta}\right) \\
&= 1 - \inf_{\theta \leq 3} F_{\chi^2_n}\left(\frac{nc_\alpha}{\theta}\right) \\
&= 1 - F_{\chi^2_n}(nc_\alpha/3).
\end{aligned}$$

En choisissant $c_\alpha = 3n^{-1}F^{-1}_{\chi^2_n}(1-\alpha) \geq 0$, il vient $\sup_{\text{« sous } H_0 \text{ »}} \mathbb{P}_\theta(T_\alpha(\mathbb{X}) = 1) = \alpha$, autrement dit $T_\alpha(\mathbb{X})$ est un test de taille α, donc aussi de niveau α.

Remarque : C'est très certainement le mieux que l'on puisse faire.

Le seuil c_α ainsi choisi, la fonction puissance du test $T_\alpha(\mathbb{X})$ se calcule alors facilement :

$$\begin{aligned}
\forall \theta > 0 : \quad \pi(\theta) &= \mathbb{P}_\theta(T_\alpha(\mathbb{X}) = 1) \\
&= \mathbb{P}_\theta\left(\widehat{\theta}_n > \frac{3}{n}F^{-1}_{\chi^2_n}(1-\alpha)\right) \\
&= \mathbb{P}_\theta\left(\frac{n\widehat{\theta}_n}{\theta} > \frac{3}{\theta}F^{-1}_{\chi^2_n}(1-\alpha)\right) \\
&= 1 - F_{\chi^2_n}\left(\frac{3}{\theta}F^{-1}_{\chi^2_n}(1-\alpha)\right).
\end{aligned}$$

Le graphe de π (pour $\alpha = 5\%$ et $n = 10$) est représenté sur la figure ci-dessous. On remarque en particulier que la valeur maximale de π sous H_0 est bien α (indiquant que le test est de taille α) et qu'il est atteint en 3, « frontière » entre H_0 et H_1.

(b) Par le cours, on sait qu'en notant $\Theta_0 = [0, 3]$, alors $\mathbb{1}_{\Theta_0 \cap I^{(1)}_{1-\alpha} = \emptyset}$, $\mathbb{1}_{\Theta_0 \cap I^{(2)}_{1-\alpha} = \emptyset}$ et $\mathbb{1}_{\Theta_0 \cap I^{(3)}_{1-\alpha} = \emptyset}$ sont des tests de niveau α pour les hypothèses considérées. Ces tests sont évidemment différents (certains seront plus conservateurs que d'autres) et donc de fonctions puissances différentes.

En particulier, il vient très rapidement $\mathbb{1}_{\Theta_0 \cap I^{(2)}_{1-\alpha} = \emptyset} = T_\alpha(\mathbb{X})$, qui est le meilleur des trois choix.

(c) Pour $n = 10$, et $\alpha = 5\%$, nous avons $c_\alpha = 3n^{-1}F^{-1}_{\chi^2_n}(1-\alpha) \approx 5.49$. Ainsi H_0 est acceptée lorsque $\widehat{\theta}_n = 4$.

De plus, la p-valeur est donnée par

$$\alpha_0 = \inf\{\alpha \in [0, 1], \text{ tel que } H_0 \text{ est rejetée au niveau } \alpha\}.$$

Or, nous avons

$$\begin{aligned}
H_0 \text{ est rejetée au niveau } \alpha &\iff \widehat{\theta}_n > 3n^{-1}F^{-1}_{\chi^2_n}(1-\alpha) \\
&\iff n\widehat{\theta}_n/3 > F^{-1}_{\chi^2_n}(1-\alpha) \\
&\iff F_{\chi^2_n}(n\widehat{\theta}_n/3) > 1-\alpha \\
&\iff \alpha > 1 - F_{\chi^2_n}(n\widehat{\theta}_n/3),
\end{aligned}$$

donc $\alpha_0 = 1 - F_{\chi^2_n}(n\widehat{\theta}_n/3)$. Pour les valeurs numériques plus haut, cela donne $\alpha_0 = 1 - F_{\chi^2_{10}}(40/3) \approx 0.2$. En particulier, on retrouve bien que le test accepte H_0 pour $\alpha = 0.05$.

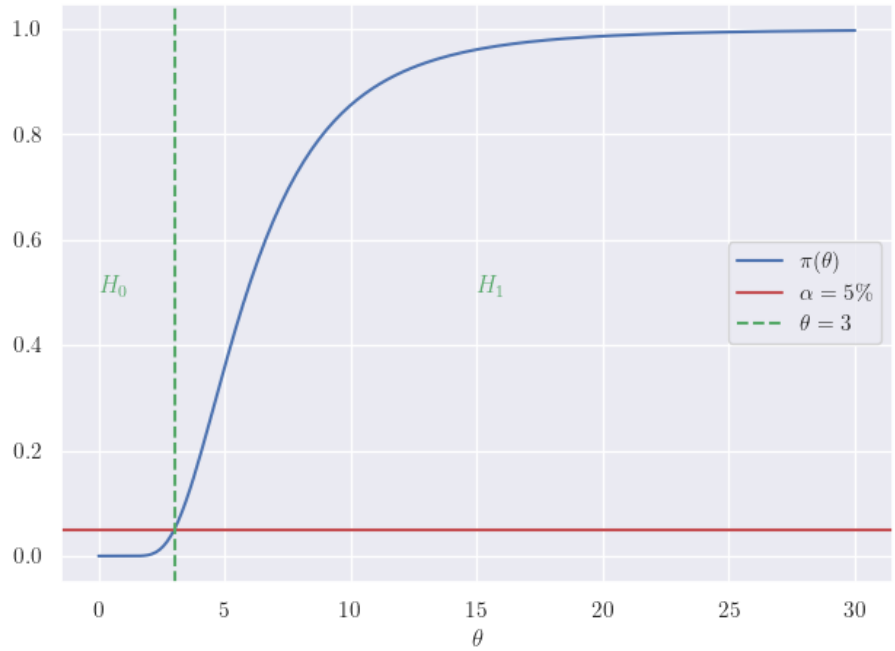

FIGURE 1 – Fonction puissance du test $T_\alpha(\mathbb{X})$ pour $\alpha = 5\%$ et $n = 10$.

EXERCICE 4

1. (a) La v.a. Y_n est bien définie sur l'événement $\{\sum_{i=1}^n X_i > 0\}$ qui est de probabilité 1 car les X_i sont strictement positifs p.s. (la loi de X_i ne charge pas 0 car elle est absolument continue par rapport à la mesure de Lebesgue). Ainsi, Y_n est bien définie p.s.

   (b) Les v.a. X_i suivent une loi exponentielle $\mathcal{E}(\theta)$. On a ainsi $\mathbb{E}(X_1) = 1/\theta$. La loi des grands nombres appliquée aux X_i (i.i.d. et intégrables) montre que $\overline{X}_n$ converge p.s. vers $1/\theta > 0$. Par le théorème de continuité, $Y_n = 1/\overline{X}_n$ converge p.s. vers θ. Cela indique que $Y_n = 1/\overline{X}_n$ est un estimateur consistant et donc raisonnable de θ. Comme nous le verrons plus tard, ce dernier correspond à l'estimateur par la méthode des moments.

   (c) Comme $\mathrm{Var}X_1 = 1/\theta^2 < \infty$, le TCL donne que $\sqrt{n}(\overline{X}_n - \mathbb{E}X_1) \xrightarrow[n\to\infty]{d} \mathcal{N}(0, \mathrm{Var}X_1)$ c'est-à-dire

$$\sqrt{n}(\overline{X}_n - 1/\theta) \xrightarrow[n\to\infty]{d} \mathcal{N}(0, 1/\theta^2).$$

   A présent, on utilise la méthode delta avec la fonction $x \mapsto 1/x$, qui est bien dérivable sur $\mathbb{R}_+^*$ de dérivée en $1/\theta$ égale à $-1/(1/\theta)^2 = -\theta^2$:

$$\sqrt{n}(Y_n - \theta) \xrightarrow[n\to\infty]{d} \mathcal{N}(0, (-\theta^2)^2/\theta^2)$$

   donc

$$\sqrt{n}(Y_n - \theta) \xrightarrow[n\to\infty]{d} \mathcal{N}(0, \theta^2).$$

   (d) D'après les propriétés de la loi Gamma (cf. TD 1), comme les X_i sont indépendantes et puisque chaque X_i est de loi $\mathcal{E}(\theta) = \gamma(1, \theta)$, nous avons $\sum_{i=1}^n X_i \sim \gamma(n, \theta)$. On en déduit $\mathbb{E}(Y_n)$, $\mathrm{Var}(Y_n)$ et $\mathbb{E}[(Y_n - \theta)^2]$:

$$\begin{aligned}
\mathbb{E}(Y_n) &= \mathbb{E}\left(n/\sum_{i=1}^n X_i\right) = n\mathbb{E}\left(1/\sum_{i=1}^n X_i\right) = n\int_{\mathbb{R}_+} \frac{1}{x}\frac{\theta^n}{\Gamma(n)} x^{n-1}\exp(-\theta x)dx \\
&= \frac{n\theta}{\Gamma(n)}\int_{\mathbb{R}_+} u^{n-2}\exp(-u)du = \frac{n\theta}{\Gamma(n)}\Gamma(n-1) = \frac{n\,\theta}{n-1},
\end{aligned}$$

en posant $u = \theta x$ et en reconnaissant sous la dernière intégrale la densité de la $\gamma(n-1,1)$, au facteur $\Gamma(n-1)$ près. On obtient de la même manière :

$$\mathbb{E}(Y_n^2) = \frac{n^2\theta^2}{\Gamma(n)}\Gamma(n-2) = \frac{n^2\,\theta^2}{(n-1)(n-2)},$$

donc, sans même passer par la décomposition biais-variance,

$$R(Y_n,\theta) = \mathbb{E}[(Y_n-\theta)^2] = \mathbb{E}[Y_n^2] - 2\theta\mathbb{E}[Y_n] + \theta^2 = \frac{n+2}{(n-1)(n-2)}\theta^2.$$

2. (a) Les propriétés asymptotiques de Z_n sont similaires à celles de Y_n : comme $\frac{n-1}{n} \xrightarrow[n\to\infty]{} 1$ et $Y_n \xrightarrow[n\to\infty]{p.s.} \theta$, on a bien sûr $Z_n \xrightarrow[n\to\infty]{p.s.} \theta$. De plus, comme $\sqrt{n}(Y_n-\theta) \xrightarrow[n\to\infty]{d} \mathcal{N}(0,\theta^2)$, nous avons

$$\sqrt{n}(Z_n-\theta) = \sqrt{n}(Y_n-\theta) - \sqrt{n}(Y_n-Z_n) = \sqrt{n}(Y_n-\theta) - Y_n/\sqrt{n} \xrightarrow[n\to\infty]{d} \mathcal{N}(0,\theta^2),$$

où nous avons utilisé le lemme de Slutsky car $Y_n/\sqrt{n}$ converge en probabilité vers 0 (constante).

(b) On calcule le risque de Z_n :

$$\begin{aligned}
\mathbb{E}(Z_n) &= \frac{n-1}{n}\mathbb{E}(Y_n) = \theta \text{ (estimateur sans biais)} \\
\mathbb{E}(Z_n^2) &= \left(\frac{n-1}{n}\right)^2 \mathbb{E}(Y_n^2) = \frac{n-1}{n-2}\theta^2 \\
R(Z_n,\theta) &= \mathbb{E}[(Z_n-\theta)^2] = \mathbb{E}[Z_n^2] - 2\theta\mathbb{E}[Z_n] + \theta^2 = \frac{\theta^2}{n-2}.
\end{aligned}$$

On voit facilement que le risque quadratique de Z_n est inférieur à celui de Y_n (pour toute valeur de θ, puisque $\frac{n+2}{n-1} > 1$), l'estimateur Z_n est donc un peu meilleur au sens du risque quadratique. La différence devient cependant négligeable lorsque n grandit.

3. Puisque $\frac{\sqrt{n}}{\sqrt{\theta^2}}(Z_n-\theta) \xrightarrow[n\to\infty]{d} \mathcal{N}(0,1)$, nous avons :

$$\mathbb{P}\left(-\Phi^{-1}(1-\alpha/2) \le \sqrt{n}(Z_n/\theta - 1) \le \Phi^{-1}(1-\alpha/2)\right) \xrightarrow[n\to\infty]{} 1-\alpha.$$

On obtient un intervalle de confiance asymptotique en réorganisant les termes des inégalités :

$$\mathbb{P}\left(\frac{Z_n}{1+\Phi^{-1}(1-\alpha/2)/\sqrt{n}} \le \theta \le \frac{Z_n}{1-\Phi^{-1}(1-\alpha/2)/\sqrt{n}}\right) \xrightarrow[n\to\infty]{} 1-\alpha,$$

ce qui correspond à l'intervalle de confiance asymptotique

$$\left[\frac{Z_n}{1+\Phi^{-1}(1-\alpha/2)/\sqrt{n}}\ ;\ \frac{Z_n}{1-\Phi^{-1}(1-\alpha/2)/\sqrt{n}}\right].$$

4. Pour tout $\theta > 0$, par le même raisonnement que ci-dessus, un intervalle de confiance unilatère de niveau asymptotique $1-\alpha$ pour θ est (en prenant comme convention $]a,b] = \emptyset$ pour $b \le a$)

$$\left]0, \frac{Z_n}{1+\frac{\Phi^{-1}(\alpha)}{\sqrt{n}}}\right].$$

Le lien entre intervalles de confiance (asymptotiques) et tests (asymptotiques) assure donc que le test consistant à rejeter H_0 ssi

$$[1,+\infty[\cap\left]0, \frac{Z_n}{1+\frac{\Phi^{-1}(\alpha)}{\sqrt{n}}}\right] = \emptyset$$

est de niveau asymptotique α. Dit autrement, on rejette H_0 ssi

$$Z_n < 1 + \frac{\Phi^{-1}(\alpha)}{\sqrt{n}}.$$

Remarque 1 : en mettant le test sous cette dernière forme, on a supposé que $1+\Phi^{-1}(\alpha)/\sqrt{n} > 0$, or à n fixé petit ceci n'est pas toujours vrai puisque pour les valeurs de α courantes (i.e. inférieures à $1/2$, par exemple $\alpha = 5\%$), on a $\Phi^{-1}(\alpha) < 0$ (par exemple $\Phi^{-1}(0.05) \approx -1,64$), donc pour $n < \Phi^{-1}(\alpha)^2$ (par exemple $n < \Phi^{-1}(0.05)^2 \approx 2.7$), il est clair que $1 + \Phi^{-1}(\alpha)/\sqrt{n} < 0$ et le passage à la dernière forme n'est pas licite. Néanmoins, il faut bien garder à l'esprit que tout ce qu'on dit n'est valable que pour $n \to \infty$, or quel que soit $\alpha \in]0,1[$ on a bien asymptotiquement $1 + \Phi^{-1}(\alpha)/\sqrt{n} > 0$, donc le passage à la dernière expression est correct. Quoi qu'il en soit, si on veut éviter toute discussion, il suffit de garder le test sous la forme : rejet de H_0 ssi

$$\frac{Z_n}{1 + \frac{\Phi^{-1}(\alpha)}{\sqrt{n}}} < 1.$$

Remarque 2 : pour construire un test, il suffit que l'intervalle de confiance soit vrai sous H_0, c'est-à-dire pour tout $\theta \in \Theta_0 = [1, \infty)$, et pas nécessairement sur tout $\mathbb{R}$. On aurait donc très bien pu écrire que pour tout $\theta \geq 1$, $\left]0, \frac{Z_n}{1+\frac{\Phi^{-1}(\alpha)}{\sqrt{n}}}\right]$ est un intervalle de confiance de niveau asymptotique $1-\alpha$ pour θ sous H_0, et construire le test par intersection avec $[1, +\infty[$, ce qui sur cet exemple serait revenu au même. La situation typique où il est plus judicieux de construire l'intervalle de confiance en tenant compte du fait que $\theta \in \Theta_0$ est celle où $\Theta_0 = \{\theta_0\}$, c'est-à-dire que H_0 est une hypothèse simple (cf. question suivante).

5. Sous $H_0 : \theta = 1$, on sait d'après ci-dessus que

$$\sqrt{n}(Z_n - 1) \xrightarrow[n\to\infty]{d} \mathcal{N}(0,1),$$

donc un intervalle de confiance bilatère de niveau asymptotique $1-\alpha$ pour la valeur 1 est

$$\left[Z_n - \frac{\Phi^{-1}(1-\alpha/2)}{\sqrt{n}}, Z_n + \frac{\Phi^{-1}(1-\alpha/2)}{\sqrt{n}}\right]$$

et le test consistant à rejeter H_0 ssi 1 n'est pas dans cet intervalle est de niveau asymptotique α, ce qui revient à dire que l'on rejette H_0 ssi

$$|Z_n - 1| > \frac{\Phi^{-1}(1-\alpha/2)}{\sqrt{n}}.$$

Remarque : comme dans l'exemple du Pile ou Face vu en cours, on peut noter que l'intervalle de confiance n'a en soi aucun intérêt (un intervalle de confiance de longueur minimale, i.e. 0, et de niveau de confiance maximale, i.e. 100%, est tout simplement $\{1\}$), mais que le test qui en découle est, lui, pertinent.

Exercice 5

1. (a) On a pour tout $x \in \mathbb{R}$:

$$F_{X_{(n)}}(x) = \mathbb{P}\big(X_{(n)} \leq x\big) \quad = \quad \big(\mathbb{P}\big(X_1 \leq x\big)\big)^n = \begin{cases} 0 & \text{si } x < 0 \\ \left(\frac{2}{\theta^2}\int_0^x s ds\right)^n = \left(\frac{x}{\theta}\right)^{2n} & \text{si } 0 \leq x \leq \theta \\ 1 & \text{sinon.} \end{cases}$$

Ainsi, la fonction de répartition de $X_{(n)}$ est continue et C^1 par morceaux sur $\mathbb{R}$. La loi de $X_{(n)}$ est donc absolument continue par rapport à la mesure de Lebesgue sur $\mathbb{R}$. Une densité est donnée par la dérivée (p.p.) de cette fonction de répartition, c'est-à-dire,

$$f_{X_{(n)}}(x) = \frac{2nx^{2n-1}}{\theta^{2n}}\mathbb{1}_{[0,\theta]}(x).$$

(b) On calcule :

$$\begin{aligned}
\mathbb{E}\big(X_{(n)}\big) &= \int_0^\theta s\frac{2ns^{2n-1}}{\theta^{2n}}ds = \frac{2n}{2n+1}\theta;\\
\mathbb{E}\big(X_{(n)}^2\big) &= \int_0^\theta s^2\frac{2ns^{2n-1}}{\theta^{2n}}ds = \frac{2n}{2n+2}\theta^2;\\
\mathrm{Var}\big(X_{(n)}\big) &= \frac{2n}{2n+2}\theta^2 - \left(\frac{2n}{2n+1}\theta\right)^2 = \frac{2n\big((2n+1)^2-2n(2n+2)\big)}{(2n+2)(2n+1)^2}\theta^2\\
&= \frac{n}{(n+1)(2n+1)^2}\theta^2.
\end{aligned}$$

(c) Comme $X_{(n)}$ est à valeurs dans $[0,\theta]$, on a $X_{(n)} \leq \theta$ $p.s.$ Ainsi, Pour tout $\varepsilon > 0$:

$$\mathbb{P}\left(|X_{(n)}-\theta| > \varepsilon\right) = \mathbb{P}\left(\theta - X_{(n)} > \varepsilon\right) = \mathbb{P}\left(X_{(n)} < \theta-\varepsilon\right) = \begin{cases} 0 & \text{si } \varepsilon > \theta \\ \frac{\left(\theta-\varepsilon\right)^{2n}}{\theta^{2n}} & \text{sinon.}\end{cases}$$

Dans tous les cas $\mathbb{P}\left(|X_{(n)}-\theta| > \varepsilon\right) \xrightarrow[n\to\infty]{} 0$ quel que soit $\varepsilon > 0$, ce qui prouve que $X_{(n)} \xrightarrow[n\to\infty]{\mathbb{P}} \theta$.

(d) Pour tout $\varepsilon > 0$, $\mathbb{P}\left(|X_{(n)}-\theta| > \varepsilon\right) = r^n$ avec $r \in [0,1[$. Ainsi $\sum_{n\geq 1}\mathbb{P}\left(|X_{(n)}-\theta| > \varepsilon\right) < \infty$ et par le Lemme de Borel-Cantelli, $X_{(n)} \xrightarrow{\text{p.s.}} \theta$.

2. Les X_i sont des v.a. bornées, elles sont donc intégrables. Leur moment d'ordre 1 est

$$\mathbb{E}(X_1) = \frac{2}{\theta^2}\int_0^\theta x^2dx = \frac{2\theta^3}{3\theta^2} = \frac{2\theta}{3}.$$

On peut donc appliquer la LFGN (X_i i.i.d.) et on obtient $\overline{X}_n \xrightarrow[n\to\infty]{p.s.} 2\theta/3$. Comme $3\overline{X}_n/2 \xrightarrow[n\to\infty]{p.s.} \theta$, alors $3\overline{X}_n/2 \xrightarrow[n\to\infty]{\mathbb{P}} \theta$ et la suite de v.a. $3\overline{X}_n/2$ est un estimateur consistant de θ, et même fortement consistant.

3. Pour comparer plus précisément les estimateurs $X_{(n)}$ et $3\overline{X}_n/2$, on va calculer leurs risques quadratiques respectifs. D'une part, nous avons

$$\begin{aligned}
\mathbb{E}\big(3\overline{X}_n/2\big) &= \frac{3}{2}\mathbb{E}\big(X_1\big) = \theta.\\
\mathrm{Var}\big(3\overline{X}_n/2\big) &= \frac{9}{4}\mathrm{Var}(\overline{X}_n) = \frac{9}{4}\frac{1}{n}\mathrm{Var}(X_1) = \frac{9}{4n}\left(\frac{2}{\theta^2}\int_0^\theta x^3dx - \left(\frac{2\theta}{3}\right)^2\right)\\
&= \frac{9}{4n}\left(\frac{2}{\theta^2}\int_0^\theta x^3dx - \left(\frac{2\theta}{3}\right)^2\right) = \frac{9}{4n}\left(\frac{2\theta^4}{4\theta^2} - \frac{4\theta^2}{9}\right) = \frac{9}{4n}\left(\theta^2\left(\frac{1}{2}-\frac{4}{9}\right)\right)\\
&= \frac{\theta^2}{8n},
\end{aligned}$$

donc $R(3\overline{X}_n/2,\theta) = \frac{\theta^2}{8n}$. D'autre part,

$$\begin{aligned}
R(X_{(n)},\theta) &= \big(\mathbb{E}(X_{(n)})-\theta\big)^2 + \mathrm{Var}(X_{(n)}) = \left(\frac{2n}{2n+1}\theta-\theta\right)^2 + \frac{n}{(n+1)(2n+1)^2}\theta^2\\
&= \frac{\theta^2}{(n+1)(2n+1)} \underset{n\to\infty}{\sim} \frac{\theta^2}{2n^2}.
\end{aligned}$$

Pour tout $\theta > 0$, on a $R(X_{(n)}, \theta)/R(3\overline{X}_n/2, \theta) = \frac{8n}{(n+1)(2n+1)}$, l'estimateur $X_{(n)}$ est donc préférable à $3\overline{X}_n/2$ au sens du risque quadratique dès que $n \geq 3$.

4. On construit un IC *non-asymptotique* pour θ à partir de l'estimateur $X_{(n)}$ (le meilleur au sens du risque quadratique). Rappelons ce que nous savons sur $X_{(n)}$:

$$\mathbb{P}\left(0 \leq X_{(n)} \leq \theta\right) = 1,$$
$$\forall x \in [0, \theta] : \mathbb{P}\left(X_{(n)} \leq x\right) = \frac{x^{2n}}{\theta^{2n}}.$$

Première option On cherche $c_\alpha \geq 0$ et $d_\alpha \geq 0$, ainsi qu'un intervalle de la forme

$$\left[X_{(n)} - c_\alpha, X_{(n)} + d_\alpha\right],$$

tels que

$$\mathbb{P}\left(\theta \in \left[X_{(n)} - c_\alpha, X_{(n)} + d_\alpha\right]\right) = 1 - \alpha. \tag{2}$$

Or,

$$\begin{aligned}\mathbb{P}\left(\theta \in \left[X_{(n)} - c_\alpha, X_{(n)} + d_\alpha\right]\right) &= \mathbb{P}\left(\theta - d_\alpha \leq X_{(n)} \leq \theta + c_\alpha\right) \\ &= F_{X_{(n)}}(\theta + c_\alpha) - F_{X_{(n)}}(\theta - d_\alpha) \\ &= 1 - \left(\frac{\theta - d_\alpha}{\theta}\right)^{2n}.\end{aligned}$$

Il suffit donc que $d_\alpha = \theta\left(1 - \alpha^{1/2n}\right)$ pour avoir l'égalité (2). Par ailleurs, la valeur de c_α n'importe pas tant que $c_\alpha \geq 0$. Puisque l'on désire un intervalle de confiance le plus court possible, on peut prendre $c_\alpha = 0$. Ainsi, l'intervalle recherché devient :

$$\left[X_{(n)} - c_\alpha, X_{(n)} + d_\alpha\right], = \left[X_{(n)}, X_{(n)} + \theta\left(1 - \alpha^{1/2n}\right)\right],$$

qui **n'est pas un intervalle de confiance** puisque la borne de droite dépend de θ ! Toutefois,

$$\begin{aligned}\mathbb{P}\left(\theta \in \left[X_{(n)}, X_{(n)} + \theta\left(1 - \alpha^{1/2n}\right)\right]\right) &= \mathbb{P}\left(X_{(n)} \leq \theta \leq X_{(n)} + \theta\left(1 - \alpha^{1/2n}\right)\right) \\ &= \mathbb{P}\left(X_{(n)} \leq \theta \leq X_{(n)}\alpha^{-1/2n}\right).\end{aligned}$$

On obtient donc l'intervalle de confiance de niveau $1 - \alpha$ pour θ :

$$\left[X_{(n)}, X_{(n)}\alpha^{-1/2n}\right].$$

Remarque : On aurait obtenu directement ce résultat en cherchant un intervalle de la forme $\left[X_{(n)}c_\alpha, X_{(n)}d_\alpha\right]$, avec $c_\alpha \in [0, 1]$ et $d_\alpha \geq 1$. Il faut laisser parler son expérience !

Seconde option

$$\forall x \in [0, \theta] : \mathbb{P}\left(x \leq X_{(n)} \leq \theta\right) = 1 - \frac{x^{2n}}{\theta^{2n}}.$$

Ainsi, pour $\alpha \in]0, 1[$, en choisissant $x = \theta\alpha^{1/(2n)}$, on obtient $x \in]0, \theta[$, puis :

$$\mathbb{P}\left(\theta\alpha^{1/(2n)} \leq X_{(n)} \leq \theta\right) = 1 - \alpha.$$

En réordonnant les inégalités, on en déduit :

$$\mathbb{P}\left(X_{(n)} \leq \theta \leq X_{(n)}\alpha^{-1/(2n)}\right) = 1 - \alpha,$$

d'où un intervalle de confiance pour θ de niveau $1 - \alpha$ donné par $\left[X_{(n)}, X_{(n)}\alpha^{-1/(2n)}\right]$.

5. A partir de l'IC précédent, on peut construire un test de niveau α pour l'hypothèse $H_0 : \theta = 1$. La règle de décision est la suivante : on rejette H_0 si $1 \notin [X_{(n)}, X_{(n)}\alpha^{-1/(2n)}]$, on accepte H_0 sinon. Ainsi, pour une valeur observée de $X_{(n)} < 1$, on rejette au niveau α si et seulement si $X_{(n)}\alpha^{-1/(2n)} < 1$, c'est-à-dire $(X_{(n)})^{2n} < \alpha$. Donc la p-valeur associé à une réalisation $x_{(n)} < 1$ de la variable aléatoire $X_{(n)}$ est
$$\alpha_0(x_{(n)}) = (x_{(n)})^{2n}.$$
Si on observe $x_{(20)} = 0.85 < 1$, cela donne $\alpha_0(x_{(20)}) = 0.85^{40} \approx 1.5 \times 10^{-3}$. On rejette donc H_0 au niveau 0.05.

Exercice 6

1. On remarque que la densité est obtenue par translation : $f_\theta(x) = g_\theta(x - \theta)$, où la densité
$$g_\theta(u) = \theta^{-1} e^{-\theta^{-1}u} \mathbb{1}_{u \geq 0}$$
est la densité d'une loi exponentielle de paramètre θ^{-1}. Ainsi $X_i - \theta$ est de loi $\mathcal{E}(1/\theta) = \gamma(1, 1/\theta)$ donc $X_i/\theta - 1 \sim \mathcal{E}(1)$. Ceci peut aussi se vérifier par la formule de transfert : pour toute fonction $\phi\colon \mathbb{R} \to \mathbb{R}$ borélienne bornée,
$$\begin{aligned}\mathbb{E}(\phi(X_i/\theta - 1)) &= \int_\theta^{+\infty} \phi(x/\theta - 1)\theta^{-1} e^{-(x/\theta - 1)}\, dx \\ &= \int_0^{+\infty} \phi(u) e^{-u}\, du,\end{aligned}$$
via le changement de variable $u = x/\theta - 1$. On reconnaît bien la densité de la loi $\mathcal{E}(1)$.
2. Comme les X_i sont i.i.d., nous avons pour tout x
$$\mathbb{P}\big(X_{(1)} \leq x\big) = 1 - \mathbb{P}\big(X_{(1)} > x\big) = 1 - \big(\mathbb{P}\big(X_1 > x\big)\big)^n,$$
Pour calculer $\mathbb{P}\big(X_1 > x\big)$, on peut faire un calcul direct avec un changement de variable, ou on peut utiliser $X_1/\theta - 1 \sim \mathcal{E}(1)$ comme expliqué à la question 1 :
$$\mathbb{P}\big(X_1 > x\big) = \mathbb{P}\big(X_1/\theta - 1 > x/\theta - 1\big) = \begin{cases} e^{1-x/\theta} & \text{pour } x/\theta - 1 \geq 0 \\ 1 & \text{pour } x/\theta - 1 < 0 \end{cases}$$
Finalement, nous obtenons
$$\mathbb{P}\big(X_{(1)} \leq x\big) = \begin{cases} 1 - e^{n(1-x/\theta)} & \text{pour } x \geq \theta \\ 0 & \text{pour } x < \theta \end{cases}. \tag{3}$$
3. Calculons $\mathbb{E}[(X_{(1)} - \theta)^2]$. Comme la fonction de répartition de $X_{(1)}$ est continue et C^1 par morceaux sur $\mathbb{R}$, la loi de $X_{(1)}$ est absolument continue par rapport à la mesure de Lebesgue sur $\mathbb{R}$. Une densité est donnée par la dérivée de cette fonction de répartition, puis on peut calculer $\mathbb{E}[(X_{(1)} - \theta)^2]$ par intégration. On propose ici une autre solution, plus rapide et utilisant directement (3) avec une petite astuce classique (Fubini) et la formulation suivante de l'espérance d'une v.a. positive : $\forall Y \geq 0$, $\mathbb{E}[Y] = \int_0^{+\infty} \mathbb{P}(Y > y)dy$. Ainsi,
$$\begin{aligned}\mathbb{E}[(X_{(1)} - \theta)^2] &= \int_0^\infty \mathbb{P}((X_{(1)} - \theta)^2 > s)ds = \int_0^\infty \mathbb{P}(X_{(1)} > \sqrt{s} + \theta)ds \quad (\text{car } X_{(1)} - \theta \geq 0 \text{ p.s.}) \\ &= \int_0^\infty e^{-n\sqrt{s}/\theta} ds = 2(\theta/n)^2 \int_0^\infty u e^{-u} du = 2(\theta/n)^2,\end{aligned}$$

où on a utilisé le changement de variable $u = n\sqrt{s}/\theta$ et l'espérance de la loi exponentielle de paramètre 1. Ainsi, on a finalement

$$R(X_{(1)}, \theta) = 2(\theta/n)^2 \xrightarrow[n\to\infty]{} 0$$

et $X_{(1)}$ converge vers θ en moyenne quadratique.

Remarque : Une méthode encore plus expéditive consiste à se souvenir que le minimum de deux variables exponentielles indépendantes est encore une variable exponentielle, de paramètre la somme des deux paramètres. Ainsi, les $X_i/\theta - 1$ étant i.i.d. selon une loi exponentielle de paramètre 1, on en déduit que $X_{(1)}/\theta - 1 \sim \mathcal{E}(n)$ donc [2]

$$\mathbb{E}\left[(X_{(1)}/\theta - 1)^2\right] = \frac{2}{n^2},$$

d'où

$$R(X_{(1)}, \theta) = \theta^2 \mathbb{E}\left[(X_{(1)}/\theta - 1)^2\right] = \frac{2\theta^2}{n^2} \xrightarrow[n\to\infty]{} 0.$$

4. Pour construire un intervalle de confiance, résumons ce que nous savons sur la variable aléatoire $X_{(1)}$:

$$\forall x \geq \theta : \mathbb{P}\big(X_{(1)} \leq x\big) = 1 - e^{n(1-x/\theta)},$$
$$\mathbb{P}(\theta \leq X_{(1)}) = 1.$$

On a ainsi pour tout $x \geq \theta$:

$$\mathbb{P}\big(\theta \leq X_{(1)} \leq x\big) = \mathbb{P}\big(X_{(1)} \leq x\big) - \mathbb{P}\big(X_{(1)} \leq \theta\big) = 1 - e^{n(1-x/\theta)}.$$

Pour $\alpha \in]0, 1]$, en choisissant $x = \theta(1 - (\ln\alpha)/n)$, on obtient $x \geq \theta$, puis :

$$\mathbb{P}\big(\theta \leq X_{(1)} \leq \theta(1 - (\ln\alpha)/n)\big) = 1 - \alpha.$$

En réordonnant les inégalités, il vient :

$$\mathbb{P}\left(\frac{X_{(1)}}{1 - (\ln\alpha)/n} \leq \theta \leq X_{(1)}\right) = 1 - \alpha.$$

Ainsi, un intervalle de confiance de niveau $1 - \alpha$ pour θ est donné par $\left[\frac{X_{(1)}}{1-(\ln\alpha)/n}, X_{(1)}\right]$.

Exercice 7

1. Le nombre N_n s'écrit $N_n = \sum_{i=1}^n \mathbb{1}(X_i = 0)$. Pour tout $i = 1, \ldots, n$, la v.a. $\mathbb{1}(X_i = 0)$ suit une loi de Bernoulli de paramètre $\mathbb{P}(X_i = 0)$.
   (a) En supposant les X_i i.i.d., on sait alors que N_n suit une loi $\mathcal{B}(n, \mathbb{P}(X = 0))$. On en déduit que $\mathbb{E}(N_n/n) = \mathbb{P}(X = 0)$ donc que $\widehat{p}_1 = N_n/n$ est un estimateur sans biais de $\mathbb{P}(X = 0)$. Son risque quadratique est alors égal à sa variance :

$$R(\widehat{p}_1, \mathbb{P}(X = 0)) = \mathrm{Var}(\widehat{p}_1) = (1/n)\mathbb{P}(X = 0)(1 - \mathbb{P}(X = 0)).$$

   Par ailleurs, la suite N_n/n (X_i i.i.d. et de carré intégrable) satisfait le TCL :

$$\sqrt{n}\frac{N_n/n - \mathbb{P}(X = 0)}{\sqrt{\mathbb{P}(X = 0)(1 - \mathbb{P}(X = 0))}} \xrightarrow[n\to\infty]{d} \mathcal{N}(0, 1).$$

2. Si $T \sim \mathcal{E}(\lambda)$, alors $\mathbb{E}[X^2] = 2/\lambda^2$.

(b) Une première solution consiste à adopter une démarche asymptotique : on utilise la convergence précédente, la LFGN $N_n/n \xrightarrow[n\to\infty]{p.s.} \mathbb{P}(X=0)$, et le théorème de Slutsky pour obtenir

$$\sqrt{n}\frac{N_n/n - \mathbb{P}(X=0)}{\sqrt{N_n/n(1-N_n/n)}} \xrightarrow[n\to\infty]{d} \mathcal{N}(0,1).$$

On calcule alors facilement l'IC asymptotique de niveau $1-\alpha$:

$$\mathbb{P}\Big(N_n/n - \varepsilon_n \leq \mathbb{P}(X=0) \leq N_n/n + \varepsilon_n\Big) \xrightarrow[n\to\infty]{} 1-\alpha$$

pour

$$\varepsilon_n = \frac{\Phi^{-1}(1-\alpha/2)\sqrt{N_n/n(1-N_n/n)}}{\sqrt{n}}$$

Une seconde solution est de construire un IC en utilisant l'inégalité de Hoeffding :

$$\mathbb{P}(|N_n/n - \mathbb{P}(X=0)| \geq x) \leq 2e^{-2nx^2},$$

ce qui donne l'intervalle de confiance de niveau $1-\alpha$

$$\left[N_n/n - \sqrt{\frac{1}{2n}\ln(2/\alpha)}, N_n/n + \sqrt{\frac{1}{2n}\ln(2/\alpha)}\right].$$

2. (a) D'après le cours de probabilités, on sait que Y_1+Y_2 suit une loi de Poisson de paramètre $\lambda_1+\lambda_2$. Pour le voir, on peut passer par la fonction caractéristique d'une loi de Poisson, à savoir

$$\mathbb{E}\left[e^{itY_1}\right] = \exp(\lambda_1(e^{it}-1)) \qquad \text{et} \qquad \mathbb{E}\left[e^{itY_2}\right] = \exp(\lambda_2(e^{it}-1)),$$
$$\Longrightarrow \mathbb{E}\left[e^{it(Y_1+Y_2)}\right] = \mathbb{E}\left[e^{itY_1}e^{itY_2}\right] \overset{(\perp\!\!\!\perp)}{=} \mathbb{E}\left[e^{itY_1}\right]\mathbb{E}\left[e^{itY_2}\right] = \exp((\lambda_1+\lambda_2)(e^{it}-1)),$$

où l'on reconnaît la fonction caractéristique d'une loi de Poisson de paramètre $\lambda_1+\lambda_2$.

(b) Il estime λ par $\overline{X}_n$ puis $\exp(-\lambda)$ par $\widehat{p}_2 = \exp(-\overline{X}_n)$. La LFGN nous indique que $\overline{X}_n \xrightarrow[n\to\infty]{p.s.} \lambda$ et, en composant par la fonction $x \mapsto \exp(-x)$ (continue), on obtient

$$\widehat{p}_2 = \exp(-\overline{X}_n) \xrightarrow[n\to\infty]{p.s.} \exp(-\lambda) = \mathbb{P}(X=0).$$

Pour calculer le biais et la variance de $\widehat{p}_2$, on commence par noter que $\sum_{i=1}^n X_i \sim \mathcal{P}(n\lambda)$. On a alors, par le Théorème de Transfert,

$$\begin{aligned}
\mathbb{E}(\widehat{p}_2) &= \mathbb{E}(\exp(-\overline{X}_n)) = \sum_{k=0}^{\infty}\exp(-k/n)\frac{(n\lambda)^k\exp(-n\lambda)}{k!} \\
&= \exp(-n\lambda)\sum_{k=0}^{\infty}\frac{(e^{-1/n}n\lambda)^k}{k!} \\
&= \exp(-\lambda n(1-e^{-1/n})) \xrightarrow[n\to\infty]{} \exp(-\lambda).
\end{aligned}$$

Le biais de $\widehat{p}_2$ est donc

$$\mathbb{E}(\widehat{p}_2) - e^{-\lambda} = e^{-\lambda}\left(\exp\{-\lambda(n(1-e^{-1/n})-1)\}-1\right),$$

qui est non nul car $1-e^{-1/n} \neq 1/n$. Comme $n(1-e^{-1/n})-1 = -1/(2n)+o(1/n)$, l'équivalent asymptotique du biais est donné par

$$\mathbb{E}(\widehat{p}_2) - e^{-\lambda} \sim e^{-\lambda}\frac{\lambda}{2n}.$$

De même,

$$\begin{aligned}
\mathbb{E}(\widehat{p}_2^2) &= \mathbb{E}(\exp(-2\overline{X}_n)) = \sum_{k=0}^{\infty} \exp(-2k/n) \frac{(n\lambda)^k \exp(-n\lambda)}{k!} \\
&= \exp(-\lambda n(1-e^{-2/n})) \xrightarrow[n\to\infty]{} \exp(-2\lambda).
\end{aligned}$$

Sa variance est donnée par :

$$\begin{aligned}
\mathrm{Var}(\widehat{p}_2) &= \mathbb{E}(\widehat{p}_2^2) - \mathbb{E}(\widehat{p}_2)^2 = \exp(-\lambda n(1-e^{-2/n})) - \exp(-2\lambda n(1-e^{-1/n})) \\
&= \exp(-2\lambda)\left(\left[\exp(2\lambda/n + o(1/n)) - 1\right] - \left[\exp(\lambda/n + o(1/n)) - 1\right]\right) \\
&= \exp(-2\lambda)\left(\left[1 + \frac{2\lambda}{n} + o(1/n) - 1\right] - \left[1 + \frac{\lambda}{n} + o(1/n) - 1\right]\right) \\
&\underset{\infty}{\sim} e^{-2\lambda}\frac{\lambda}{n}.
\end{aligned}$$

(c) Le risque quadratique de $\widehat{p}_2$ est d'après ce qui précède :

$$R(\widehat{p}_2, \exp(-\lambda)) = \mathrm{Var}(\widehat{p}_2) + (\mathbb{E}(\widehat{p}_2) - e^{-\lambda})^2 \underset{\infty}{\sim} e^{-2\lambda}\frac{\lambda}{n}.$$

Le risque quadratique de $\widehat{p}_1$ est

$$R(\widehat{p}_1, e^{-\lambda}) = \frac{e^{-\lambda}(1-e^{-\lambda})}{n}.$$

Ainsi,

$$\frac{R(\widehat{p}_1, e^{-\lambda})}{R(\widehat{p}_2, e^{-\lambda})} \underset{\infty}{\sim} \frac{e^{-\lambda}(1-e^{-\lambda})}{\lambda e^{-2\lambda}} = \frac{1-e^{-\lambda}}{\lambda e^{-\lambda}}.$$

On peut montrer que pour tout $\lambda > 0$, $1 - e^{-\lambda} - \lambda e^{-\lambda} > 0$ par une étude de fonction, ce qui indique que $\widehat{p}_2$ est meilleur que $\widehat{p}_1$ au sens du risque quadratique. Il faut cependant noter que $\widehat{p}_1$ ne suppose aucune connaissance sur la loi des X_i, tandis que $\widehat{p}_2$ suppose et exploite le fait que les X_i sont des variables de Poisson.