Feuille 2 : correction

Exercice 1 (Loi Gamma)

1. On suppose $q > 0$ connu et $\theta > 0$ inconnu.
   (a) Comme $\mathbb{E}[X] = \frac{q}{\theta}$, on a $\theta = q/\mathbb{E}[X]$ et on peut proposer $\widehat{\theta} = q/\overline{X}_n$ comme estimateur de θ (méthode des moments).
   (b) On applique le TCL aux variables X_1, ..., X_n qui sont i.i.d. de variance q/θ^2. On a
   $$\sqrt{n}(\overline{X}_n - q/\theta) \rightsquigarrow \mathcal{N}(0, q/\theta^2)$$
   et
   $$\sqrt{n}(\overline{X}_n\theta - q) \rightsquigarrow \mathcal{N}(0, q).$$
   Ainsi, en notant $q_\alpha = \Phi^{-1}(1-\alpha/2)$ le quantile d'ordre $1-\alpha/2$ d'une $\mathcal{N}(0,1)$, on a
   $$\begin{aligned}\mathbb{P}\left(-\frac{q_\alpha}{\sqrt{n}}\sqrt{q} \le \overline{X}_n\theta - q \le \frac{q_\alpha}{\sqrt{n}}\sqrt{q}\right) &= \mathbb{P}\left(q - \frac{q_\alpha}{\sqrt{n}}\sqrt{q} \le \overline{X}_n\theta \le q + \frac{q_\alpha}{\sqrt{n}}\sqrt{q}\right)\\ &= \mathbb{P}\left(\left(q - \frac{q_\alpha}{\sqrt{n}}\sqrt{q}\right)(\overline{X}_n)^{-1} \le \theta \le \left(q + \frac{q_\alpha}{\sqrt{n}}\sqrt{q}\right)(\overline{X}_n)^{-1}\right)\\ &\underset{n\to\infty}{\longrightarrow} 1-\alpha\end{aligned}$$
   D'où l'intervalle de confiance pour θ au niveau asymptotique $1-\alpha$:
   $$I = \left[\left(q - \frac{q_\alpha}{\sqrt{n}}\sqrt{q}\right)(\overline{X}_n)^{-1},\ \left(q + \frac{q_\alpha}{\sqrt{n}}\sqrt{q}\right)(\overline{X}_n)^{-1}\right].$$
2. On suppose $q > 0$ inconnu et $\theta > 0$ connu.
   (a) Comme $\mathbb{E}[X] = \frac{q}{\theta}$, on a $q = \theta\mathbb{E}[X]$ et on propose $\widehat{q} = \theta\overline{X}_n$ comme estimateur de q (méthode des moments).
   (b) Comme vu en 1.(b), on a
   $$\sqrt{n}(\widehat{q} - q) \rightsquigarrow \mathcal{N}(0, q),$$
   donc via Slutsky
   $$\sqrt{n}\frac{\widehat{q} - q}{\sqrt{\widehat{q}}} \rightsquigarrow \mathcal{N}(0, 1).$$
   En notant $q_\alpha = \Phi^{-1}(1-\alpha/2)$ le quantile d'ordre $1-\alpha/2$ d'une $\mathcal{N}(0,1)$, on en déduit un intervalle de confiance pour q au niveau asymptotique $1-\alpha$:
   $$I = \left[\widehat{q} - \frac{q_\alpha\sqrt{\widehat{q}}}{\sqrt{n}}\ ,\ \widehat{q} + \frac{q_\alpha\sqrt{\widehat{q}}}{\sqrt{n}}\right].$$
3. On suppose $q > 0$ et $\theta > 0$ inconnus. On cherche deux équations de moments : on a $\mathbb{E}[X] = \frac{q}{\theta}$ et $\mathrm{Var}(X) = \frac{q}{\theta^2}$. On a donc
   $$\theta = \frac{\mathbb{E}[X]}{\mathrm{Var}(X)} \quad \text{et} \quad q = \frac{(\mathbb{E}[X])^2}{\mathrm{Var}(X)},$$
   on propose donc les estimateurs empiriques :
   $$\widehat{\theta} = \frac{\overline{X}_n}{\overline{X_n^2} - (\overline{X}_n)^2} \quad \text{et} \quad \widehat{q} = \frac{(\overline{X}_n)^2}{\overline{X_n^2} - (\overline{X}_n)^2}.$$
   Comme d'habitude, la LFGN donne $\overline{X}_n \overset{p.s.}{\to} q/\theta$ et $\overline{X_n^2} \overset{p.s.}{\to} q/\theta^2 + (q/\theta)^2$ ce qui montre que les estimateurs sont consistants.

Exercice 2 (**Loi gaussienne de variance inconnue**))

1. Si $Y \sim \mathcal{N}(0,1)$, on a, par symétrie
$$\mathbb{E}|Y| = \frac{2}{\sqrt{2\pi}} \int_0^{+\infty} xe^{-x^2/2}dx = \sqrt{\frac{2}{\pi}}\left[-e^{-x^2/2}\right]_0^{+\infty} = \sqrt{\frac{2}{\pi}}\,.$$
Ainsi, si $X = \sigma Y \sim \mathcal{N}(0,\sigma^2)$, on a $\mathbb{E}|X| = \sigma\sqrt{\frac{2}{\pi}}$. L'estimateur S_n est donc sans biais. Pour T_n, on a par l'inégalité de Jensen,
$$\mathbb{E}T_n < \sqrt{\mathbb{E}X^2} = \sigma\,.$$
(L'inégalité est stricte car $x \mapsto \sqrt{x}$ est strictement concave et $\frac{1}{n}\sum_{i=1}^n X_i^2$ est non-constante.) L'estimateur T_n a donc un biais négatif. Mais on peut remarquer que si $Y_1,\dots,Y_n$ sont i.i.d. de loi $\mathcal{N}(0,1)$, alors $\sum_{i=1}^n Y_i^2 \sim \gamma(n/2,1/2)$ donc $n^{-1}\sum_{i=1}^n Y_i^2 \sim \gamma(n/2,n/2)$. Ainsi
$$\begin{aligned}\mathbb{E}\left[\sqrt{\frac{1}{n}\sum_{i=1}^n Y_i^2}\right] &= \frac{(n/2)^{n/2}}{\Gamma(n/2)}\int_0^{+\infty} z^{\frac{n}{2}-1+\frac{1}{2}}e^{-\frac{n}{2}z}dz \\ &= \frac{(n/2)^{n/2}}{\Gamma(n/2)}\frac{\Gamma((n+1)/2)}{(n/2)^{\frac{n+1}{2}}}\,.\end{aligned}$$
Par la formule de Stirling, $\Gamma(z) \sim \sqrt{2\pi}z^{z-\frac{1}{2}}e^{-z}$ quand $z \to +\infty$. Ainsi
$$\mathbb{E}T_n = \sigma\mathbb{E}\left[\sqrt{\frac{1}{n}\sum_{i=1}^n Y_i^2}\right] \sim \sigma\sqrt{\frac{2}{n}}\frac{\left(\frac{n+1}{2}\right)^{n/2}e^{-\frac{n+1}{2}}}{\left(\frac{n}{2}\right)^{n/2-1/2}e^{-n/2}} = \sigma\left(1+\frac{1}{n}\right)^{n/2}e^{-1/2} \sim \sigma\,.$$
L'estimateur T_n est donc asymptotiquement sans biais.

2. Par le TCL, et en remarquant que $\mathrm{Var}(|X|) = \mathbb{E}|X|^2 - (\mathbb{E}|X|)^2 = \frac{\pi-2}{\pi}\sigma^2$, on a
$$\sqrt{n}\left(\frac{1}{n}\sum_{i=1}^n |X_i| - \sqrt{\frac{2}{\pi}}\sigma\right) \rightsquigarrow \mathcal{N}\left(0,\frac{\pi-2}{\pi}\sigma^2\right)\,,$$
donc
$$\sqrt{n}\,(S_n-\sigma) \rightsquigarrow \mathcal{N}\left(0,\frac{\pi-2}{2}\sigma^2\right)\,.$$
De plus, comme $\mathrm{Var}(X^2) = \mathbb{E}X^4 - (\mathbb{E}X^2)^2 = 2\sigma^4$, on a
$$\sqrt{n}\left(\frac{1}{n}\sum_{i=1}^n X_i^2 - \sigma^2\right) \rightsquigarrow \mathcal{N}\left(0,2\sigma^4\right)\,.$$
En appliquant la méthode delta avec $g : x \mapsto \sqrt{x}$, on obtient
$$\sqrt{n}\,(T_n-\sigma) \rightsquigarrow \mathcal{N}\left(0,\frac{1}{2}\sigma^2\right)\,.$$
Comme $\pi - 2 > 1$, on choisit T_n.

Exercice 3 (**Loi Bêta dilatée**)

1. (a) On a pour tout $x \in \mathbb{R}$:
$$\mathbb{P}\big(X_{(n)} \le x\big) \quad = \quad \big(\mathbb{P}\big(X_1 \le x\big)\big)^n = \begin{cases} 0 & \text{si } x < 0 \\ \left(\frac{2}{\theta^2}\int_0^x s ds\right)^n = \left(\frac{x}{\theta}\right)^{2n} & \text{si } 0 \le x \le \theta \\ 1 & \text{sinon.}\end{cases}$$
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,
$$f_{X_{(n)}}(x) \quad = \quad \frac{2nx^{2n-1}}{\theta^{2n}}\mathbb{1}_{[0,\theta]}(x).$$

(b) On calcule :

$$\begin{aligned}
\mathbb{E}\big(X_{(n)}\big) &= \int_0^\theta s\frac{2ns^{2n-1}}{\theta^{2n}}ds = \frac{2n}{2n+1}\theta; \\
\mathbb{E}\big(X_{(n)}^2\big) &= \int_0^\theta s^2\frac{2ns^{2n-1}}{\theta^{2n}}ds = \frac{2n}{2n+2}\theta^2 = \frac{n}{n+1}\theta^2; \\
\mathrm{Var}\left(X_{(n)}\right) &= \left(\frac{n}{n+1} - \frac{4n^2}{(2n+1)^2}\right)\theta^2 = \frac{n(2n+1)^2 - 4n^2(n+1)}{(n+1)(2n+1)^2}\theta^2 \\
&= \frac{n}{(n+1)(2n+1)^2}\theta^2; \\
R(X_{(n)},\theta) &= \left(\mathbb{E}(X_{(n)}) - \theta\right)^2 + \mathrm{Var}(X_{(n)}) = \left(\frac{2n}{2n+1}\theta - \theta\right)^2 + \frac{n}{(n+1)(2n+1)^2}\theta^2 \\
&= \frac{\theta^2}{(n+1)(2n+1)} \underset{n\to\infty}{\sim} \frac{\theta^2}{2n^2}.
\end{aligned}$$

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

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

Dans tous les cas $\mathbb{P}\left(|X_{(n)} - \theta| > \epsilon\right) \to 0$ quel que soit $\epsilon > 0$, ce qui prouve que $X_{(n)} \xrightarrow{\mathbb{P}} \theta$. Montrons que cette convergence a lieu à vitesse $1/n$. La variable $n(\theta - X_{(n)})$ est à valeurs dans $[0, n\theta]$ et pour tout $x \in [0, n\theta]$, on a

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

Ainsi, pour tout $x \in \mathbb{R}$, on a

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

Comme $\mathbb{1}_{x>n\theta} \underset{n\to\infty}{\longrightarrow} 0$, $\mathbb{1}_{x\in[0,n\theta]} \underset{n\to\infty}{\longrightarrow} \mathbb{1}_{x\geq 0}$ et $\left(1 - \frac{x}{\theta n}\right)^{2n} \underset{n\to\infty}{\longrightarrow} e^{-\frac{2x}{\theta}}$, on a $n(\theta - X_{(n)}) \rightsquigarrow \mathcal{E}(2/\theta)$.

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

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

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

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

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

donc $R(3\overline{X}_n/2, \theta) = \frac{\theta^2}{8n}$. D'autre part, on a vu $R(X_{(n)},\theta) \sim \frac{\theta^2}{2n^2}$. L'estimateur $X_{(n)}$ est donc préférable à $3\overline{X}_n/2$ au sens du risque quadratique dès que n est assez grand.

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

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

Ainsi,

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

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 est

$$\alpha_0 = (X_{(n)})^{2n}.$$

Si on observe $x_{(20)} = 0.85$, cela donne $\alpha_0 = 0.85^{40} \simeq 1.5 10^{-3}$. On rejette donc H_0 au niveau 0.05.

■

Exercice 4 (Modèle gaussien ordinaire)

1. Le modèle de régression associé est $Y = X\beta + \sigma\varepsilon$, avec $X = (1, \cdots, 1)^t \in \mathbb{R}^n$, $\beta = m \in \mathbb{R}$ et $\varepsilon = (\varepsilon_1, \cdots, \varepsilon_n)^t \in \mathbb{R}^n$, $\varepsilon \sim \mathcal{N}_n(0, I_n)$. L'estimateur des moindres carrés est ainsi $\widehat{\beta} = (X^tX)^{-1}X^tY = \overline{Y}_n$.
2. En notant $u = (1 \cdots 1)^t \in \mathbb{R}^n$, on a $Y \sim \mathcal{N}(mu, \sigma^2 I_n)$ et la matrice de projection associée est :

$$P = X(X^tX)^{-1}X^t = \frac{1}{n}\begin{pmatrix} 1 & \dots & 1 \\ \vdots & \ddots & \vdots \\ 1 & \dots & 1 \end{pmatrix}.$$

De plus, $PY = \overline{Y}_n u$, $P_\perp Y = (I_n - P)Y = Y - \overline{Y}_n u$, $Pmu = mu$ et $P_\perp mu = 0u$. D'après le théorème de Cochran,
— $PY \sim \mathcal{N}(mu, \sigma^2 P)$, $P_\perp Y \sim \mathcal{N}(0u, \sigma^2 P_\perp)$;
— PY et $P_\perp Y$ sont indépendants ;
— $\frac{\|P(Y-mu)\|^2}{\sigma^2} = \frac{n(\overline{Y}_n - m)^2}{\sigma^2} \sim \chi^2(1)$, $\frac{\|Y - \overline{Y}_n u\|^2}{\sigma^2} = \frac{\sum_{i=1}^n (Y_i - \overline{Y}_n)^2}{\sigma^2} \sim \chi^2(n-1)$.

3. On vient de voir que

$$(n-1)\widehat{\sigma}_n^2/\sigma^2 = \sum_{i=1}^n (Y_i - \overline{Y}_n)^2/\sigma^2 \sim \chi^2(n-1).$$

Ainsi, $\mathbb{E}[(n-1)\widehat{\sigma}_n^2/\sigma^2] = n - 1$, c'est-à-dire $\mathbb{E}[\widehat{\sigma}_n^2] = \sigma^2$ donc $\widehat{\sigma}_n^2$ est sans biais. De plus, la loi des grands nombres et le théorème de continuité impliquent que

$$s_n^2 = \frac{1}{n}\sum_{i=1}^n (Y_i - \overline{Y}_n)^2 = \frac{1}{n}\sum_{i=1}^n Y_i^2 - \overline{Y}_n^2 \xrightarrow[n\to\infty]{p.s.} \mathbb{E}[Y_1^2] - m^2 = \sigma^2,$$

i.e. s_n^2 est un estimateur convergent de σ^2. Ainsi, $\widehat{\sigma}_n^2 = \frac{n}{n-1}s_n^2 \overset{p.s.}{\to} \sigma^2$.

4. On sait que :
$$\frac{\widehat{\beta}-\beta}{\widehat{\sigma}_n\sqrt{1/n}} = \sqrt{n}\frac{\overline{Y}_n - m}{\widehat{\sigma}_n} \sim \mathcal{T}(n-1).$$
5. D'après ce qui vient d'être dit, un intervalle de confiance de niveau $(1-\alpha)$ pour σ^2 est donné par
$$\left[\frac{(n-1)\widehat{\sigma}_n^2}{c_{n-1}(1-\alpha/2)}, \frac{(n-1)\widehat{\sigma}_n^2}{c_{n-1}(\alpha/2)}\right]$$
où $c_{n-1}(\alpha/2)$ et $c_{n-1}(1-\alpha/2)$ sont les quantiles d'ordres $\alpha/2$ et $1-\alpha/2$ d'une loi χ^2_{n-1} (ici $p=1$ si p est le nombre de variables explicatives). Ainsi, le test consistant à rejeter H_0 si
$$3 \notin \left[\frac{(n-1)\widehat{\sigma}_n^2}{c_{n-1}(1-\alpha/2)}, \frac{(n-1)\widehat{\sigma}_n^2}{c_{n-1}(\alpha/2)}\right]$$
i.e., si
$$\widehat{\sigma}_n^2 > 3\frac{c_{n-1}(1-\alpha/2)}{n-1} \quad \text{ou} \quad \widehat{\sigma}_n^2 < 3\frac{c_{n-1}(\alpha/2)}{n-1}$$
est de niveau α.
6. Un intervalle de confiance de niveau $1-\alpha$ pour m est donné par
$$J(m) = \left[\overline{Y}_n - \widehat{\sigma}_n\frac{t_{n-1}(1-\alpha/2)}{\sqrt{n}}, \overline{Y}_n + \widehat{\sigma}_n\frac{t_{n-1}(1-\alpha/2)}{\sqrt{n}}\right].$$
où $t_{n-1}(1-\alpha/2)$ est le quantile d'ordre $1-\alpha/2$ d'une loi de Student $\mathcal{T}(n-1)$.
On peut construire un test de niveau α en rejetant H_0 si $m_0 \notin J(m)$, c'est-à-dire si
$$\sqrt{n}\left|\frac{\overline{Y}_n - m_0}{\widehat{\sigma}_n}\right| > t_{n-1}(1-\alpha/2).$$
7. Pour ce test unilatéral, on cherche une région de rejet de la forme $\mathcal{R} =]-\infty, c_\alpha[$ pour $\overline{Y}_n$ telle que :
$$\begin{aligned}\alpha &= \sup_{m\geq m_0} \mathbb{P}(\overline{Y}_n < c_\alpha)\\ &= \sup_{m\geq m_0} \mathbb{P}\left(\sqrt{n}\frac{\overline{Y}_n - m}{\widehat{\sigma}_n} < \sqrt{n}\frac{c_\alpha - m}{\widehat{\sigma}_n}\right)\\ &= \sup_{m\geq m_0} F_{\mathcal{T}(n-1)}\left(\sqrt{n}\frac{c_\alpha - m}{\widehat{\sigma}_n}\right)\\ &= F_{\mathcal{T}(n-1)}\left(\sqrt{n}\frac{c_\alpha - m_0}{\widehat{\sigma}_n}\right),\end{aligned}$$
i.e.
$$c_\alpha = m_0 + \frac{\widehat{\sigma}_n}{\sqrt{n}}t_{n-1}(\alpha).$$
Pour ce test,
$$\begin{aligned}T(Y) = 1 &\Leftrightarrow \overline{Y}_n < c_\alpha\\ &\Leftrightarrow \sqrt{n}\frac{\overline{Y}_n - m_0}{\widehat{\sigma}_n} < t_{n-1}(\alpha)\\ &\Leftrightarrow F_{\mathcal{T}(n-1)}\left(\sqrt{n}\frac{\overline{Y}_n - m_0}{\widehat{\sigma}_n}\right) < \alpha.\end{aligned}$$
Ainsi, la p-valeur du test $T(Y) = 1 \Leftrightarrow \overline{Y}_n < c_\alpha$ est :
$$\alpha_0 = F_{\mathcal{T}(n-1)}\left(\sqrt{n}\frac{\overline{Y}_n - m_0}{\widehat{\sigma}_n}\right).$$
Pour $m_0 = 12,5$, $n = 25$, $\overline{y}_{25} = 12$ et $\widehat{\sigma}_n^2 = 1,69$, la p-valeur est :
$$\alpha_0 = F_{\mathcal{T}(24)}\left(\sqrt{25}\left(\frac{12-12.5}{\sqrt{1.69}}\right)\right) \simeq F_{\mathcal{T}(24)}(-1,92) \simeq 0.03.$$
On rejette donc H_0 au niveau 5%.

$\blacksquare$

Exercice 5 (Réciproque du théorème de Cochran)

1. Notons $u = (1, \cdots, 1)^t$. La projection orthogonale de X sur $\mathrm{Vect}(u)$ est $\overline{X}_n u$. Ainsi, le théorème de Cochran assure que $\overline{X}_n u$ est indépendant de $X - \overline{X}_n u$. Ceci implique que $\overline{X}_n$ est indépendant de $s_n^2 = n^{-1}\|X - \overline{X}_n u\|^2$. De plus, nous avons $\sqrt{n}(\overline{X}_n - m) \sim \mathcal{N}(0, \sigma^2)$ et le théorème de Cochran nous dit que

$$\sigma^{-2}\|X - \overline{X}_n u\|^2 = \sigma^{-2} n s_n^2 \sim \chi^2(n-1).$$

En particulier, $\frac{\sqrt{n}(\overline{X}_n - m)}{\sqrt{\frac{n s_n^2}{n-1}}} \sim \mathcal{T}(n-1)$.

2. (a) En développant on obtient

$$n s_n^2 = \sum_{i=1}^n (X_i - \overline{X}_n)^2 = \sum_{i=1}^n X_i^2 - n\overline{X}_n^2 .$$

On a $\mathbb{E}[\overline{X}_n^2] = \mathrm{Var}(\overline{X}_n) + (\mathbb{E}\overline{X}_n)^2 = \frac{\sigma^2}{n} + m^2$, et ainsi

$$\mathbb{E}[n s_n^2] = (n-1)\sigma^2 .$$

Par suite, comme $\overline{X}_n$ et $n s_n^2$ sont indépendants et que les X_i sont i.i.d., on a pour tout $t \in \mathbb{R}$,

$$\mathbb{E}\left(s_n^2 \exp\left(itn\overline{X}_n\right)\right) = \mathbb{E}(s_n^2)\mathbb{E}\left(\exp\left(it\sum_{i=1}^n X_i\right)\right) = \mathbb{E}(s_n^2)\left(\mathbb{E}\big(\exp\left(itX_1\right)\big)\right)^n = \phi^n(t)\mathbb{E}(s_n^2).$$

(b) Remarquons tout d'abord que $\phi'(t) = i\,\mathbb{E}\big(X_1 \exp\left(itX_1\right)\big)$ et $\phi''(t) = -\mathbb{E}\big(X_1^2 \exp\left(itX_1\right)\big)$. En particulier, $\phi'(0) = im$. De plus, comme

$$\begin{aligned} n s_n^2 &= \sum_{j=1}^n X_j^2 - n(\overline{X}_n)^2 \\ &= \sum_{j=1}^n X_j^2 - n^{-1}\sum_{k,\ell=1}^n X_k X_\ell \\ &= (1 - n^{-1})\sum_{j=1}^n X_j^2 - n^{-1}\sum_{k\neq\ell} X_k X_\ell, \end{aligned}$$

on obtient, en utilisant le fait que les X_i sont i.i.d.,

$$\begin{aligned} \mathbb{E}\left(n s_n^2 e^{itn\overline{X}_n}\right) &= (1 - n^{-1})\sum_{j=1}^n \mathbb{E}\left(X_j^2 \prod_{i=1}^n e^{itX_i}\right) - n^{-1}\sum_{k\neq\ell}\mathbb{E}\left(X_k X_\ell \prod_{i=1}^n e^{itX_i}\right) \\ &= (n-1)\mathbb{E}\left[X_1^2 e^{itX_1}\right]\prod_{i=2}^n \mathbb{E}\left[e^{itX_i}\right] - (n-1)\mathbb{E}\left[X_1 e^{itX_1}\right]\mathbb{E}\left[X_2 e^{itX_2}\right]\prod_{i=3}^n \mathbb{E}\left[e^{itX_i}\right] \\ &= -(n-1)\phi''(t)\phi(t)^{n-1} + (n-1)\phi'(t)^2\phi(t)^{n-2} . \end{aligned}$$

En utilisant la question (a), on obtient donc

$$\phi^n(t)(n-1)\sigma^2 = -(n-1)\phi''(t)\phi^{n-1}(t) + (n-1)(\phi'(t))^2\phi^{n-2}(t),$$

ce qui donne la relation voulue

$$\frac{\phi''}{\phi} - \left(\frac{\phi'}{\phi}\right)^2 = -\sigma^2, \quad \phi(0) = 1, \phi'(0) = im.$$

(c) On résout l'équation différentielle. Comme

$$(\log\phi)''(t) = \frac{\phi''}{\phi} - \left(\frac{\phi'}{\phi}\right)^2,$$

elle s'écrit

$$(\log\phi)''(t) = -\sigma^2.$$

D'où $\log\phi(t) = -\sigma^2 t^2/2 + at + b$, pour $a, b \in \mathbb{R}$, et les conditions initiales $\phi(0) = 1$ et $\phi'(0) = im$ donnent

$$\phi(t) = \exp\left(\frac{-\sigma^2 t^2}{2} + imt\right) \text{ pour tout } t \in \mathbb{R}.$$

On reconnaît la fonction caractéristique de la loi $\mathcal{N}(m, \sigma^2)$ et on en déduit que les v.a. X_i suivent la loi $\mathcal{N}(m, \sigma^2)$, pour tout $i = 1, \ldots, n$.

Exercice 6 (Théorème de Gauss–Markov)

1. L'estimateur des moindres carrés est donné par $\widehat{\beta} = (X^T X)^{-1} X^T Y$.
2. Pour tout $\beta \in \mathbb{R}^p$, on a $\mathbb{E}_\beta[CY] = C\mathbb{E}_\beta[Y] = CX\beta = \beta$ car $\tilde{\beta}$ est supposé sans biais. Ainsi $CX = I_n$.
3. Posons $D = C - (X^T X)^{-1} X^T$. Par la question précédente, on a $DX = 0$, et

$$\begin{aligned}
\mathrm{Cov}(\tilde{\beta}) &= C\,\mathrm{Cov}(Y) C^T \\
&= \sigma^2 \left(D + (X^T X)^{-1} X^T\right)\left(X(X^T X)^{-1} + D^T\right) \\
&= \sigma^2 \left((X^T X)^{-1} + DD^T\right) \\
&= \mathrm{Cov}(\widehat{\beta}) + \sigma^2 DD^T,
\end{aligned}$$

et la matrice DD^T est bien semi-définie positive.

Exercice 7 (Validation croisée et levier)

1. Commençons par étudier l'erreur de prédiction, $Y_{n+1} - X_{n+1}^T \widehat{\theta}_n$, conditionnellement à $X_1, \ldots, X_{n+1}$. Notons $\mathbf{X}_{n+1} = (X_1, \ldots, X_{n+1})$. On a

$$\mathbb{E}\left[Y_{n+1} - X_{n+1}^T \widehat{\theta}_n \mid \mathbf{X}_{n+1}\right] = X_{n+1}^T \theta - X_{n+1}^T \mathbb{E}\left[\widehat{\theta}_n \mid \mathbf{X}_{n+1}\right] = 0,$$

et, comme ε_{n+1} est indépendant de $\mathbf{X}_{n+1}$ et de $\widehat{\theta}_n$,

$$\begin{aligned}
\mathbb{E}\left[\left(Y_{n+1} - X_{n+1}^T \widehat{\theta}_n\right)^2 \mid \mathbf{X}_{n+1}\right] &= \mathbb{E}\left[\left(\varepsilon_{n+1} + X_{n+1}^T(\theta - \widehat{\theta}_n\right)^2 \mid \mathbf{X}_{n+1}\right] \\
&= \sigma^2 + X_{n+1}^T \,\mathrm{Cov}\left(\widehat{\theta}_n \mid \mathbf{X}_{n+1}\right) X_{n+1} \\
&= \sigma^2 \left(1 + X_{n+1}^T (\mathbf{X}^T \mathbf{X})^{-1} X_{n+1}\right).
\end{aligned}$$

De même, pour tout $i \in [\![1, n]\!]$, on a

$$\mathbb{E}\left[\left(Y_i - X_i^T \widehat{\theta}^{(i)}\right)^2 \mid \mathbf{X}_{n+1}\right] = \sigma^2 \left(1 + X_i^T ((\mathbf{X}^{(i)})^T \mathbf{X}^{(i)})^{-1} X_i\right).$$

Par échangeabilité, on a

$$\mathbb{E}\left[X_i^T ((\mathbf{X}^{(i)})^T \mathbf{X}^{(i)})^{-1} X_i\right] = \mathbb{E}\left[X_{n+1}^T ((\mathbf{X}^{(n)})^T \mathbf{X}^{(n)})^{-1} X_{n+1}\right].$$

Il s'agit donc de montrer que

$$\mathbb{E}\left[X_{n+1}^T \left(((\mathbf{X}^{(n)})^T \mathbf{X}^{(n)})^{-1} - (\mathbf{X}^T \mathbf{X})^{-1}\right) X_{n+1})\right] \geq 0.$$

Or $\mathbf{X}^T \mathbf{X} = \sum_{i=1} X_i X_i^T = (\mathbf{X}^{(n)})^T \mathbf{X}^{(n)} + X_n X_n^T$. Donc la matrice $\mathbf{X}^T \mathbf{X} - (\mathbf{X}^{(n)})^T \mathbf{X}^{(n)}$ est semi-définie positive. Or, si A et B sont des matrices symétriques définies positives et si $A \succ B$ (i.e. $A - B$ semi-définie positive), alors $B^{-1} \succ A^{-1}$. En effet, si $A \succ B$, alors pour toute matrice C, on a $CAC^T \succ CBC^T$. Pour $C = A^{-1/2}$, on obtient $I \succ A^{-1/2} B A^{-1/2}$. Donc (car tout se passe bien avec la matrice I) $A^{1/2} B^{-1} A^{1/2} \succ I$, et donc $B^{-1} \succ A^{-1}$. On en déduit que $((\mathbf{X}^{(n)})^T \mathbf{X}^{(n)})^{-1} - (\mathbf{X}^T \mathbf{X})^{-1}$ est définie positive p.s. Comme ces quantités sont intégrables, l'inégalité souhaitée sur l'espérance s'ensuit.

2. On a $\widehat{\theta}^{(i)} = ((\mathbf{X}^{(i)})^T\mathbf{X}^{(i)})^{-1}(\mathbf{X}^{(i)})^TY^{(i)}$.
3. On a

$$\mathbf{X}^T\mathbf{X} = \sum_{k=1}^n X_kX_k^T = (\mathbf{X}^{(i)})^T\mathbf{X}^{(i)} + X_iX_i^T\,.$$

   Ainsi, par la formule de Morrison–Sherman,

$$((\mathbf{X}^{(i)})^T\mathbf{X}^{(i)})^{-1} = (\mathbf{X}^T\mathbf{X})^{-1} + \frac{(\mathbf{X}^T\mathbf{X})^{-1}X_iX_i^T(\mathbf{X}^T\mathbf{X})^{-1}}{1-X_i^T(\mathbf{X}^T\mathbf{X})^{-1}X_i} = (\mathbf{X}^T\mathbf{X})^{-1} + \frac{(\mathbf{X}^T\mathbf{X})^{-1}X_iX_i^T(\mathbf{X}^T\mathbf{X})^{-1}}{1-\mathbf{H}[i,i]}\,,$$

   où $\mathbf{H} = \mathbf{X}(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T$. La matrice $\mathbf{H}$ est la matrice de projection sur l'espace de dimension p engendré par les colonnes de $\mathbf{X}$. Ainsi $\mathsf{tr}(\mathbf{H}) = p$. De plus, comme $\mathbf{H} = \mathbf{H}^T = \mathbf{H}^2$, on a

$$\mathbf{H}[i,i] = \sum_{j=1}^n \mathbf{H}[i,j]\mathbf{H}[j,i] = \mathbf{H}[i,i]^2 + \sum_{j\neq i}\mathbf{H}[i,j]^2\,.$$

   Ainsi $\mathbf{H}[i,i](1-\mathbf{H}[i,i]) \geq 0$ donc $0 \leq \mathbf{H}[i,i] \leq 1$.
4. En utilisant que $(\mathbf{X}^{(i)})^TY^{(i)} = \mathbf{X}^TY - X_iY_i$ puis que $X_i^T(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^TY = \widehat{Y}_i$, on obtient

$$\begin{aligned}\widehat{\theta}^{(i)} &= \left((\mathbf{X}^T\mathbf{X})^{-1} + \frac{(\mathbf{X}^T\mathbf{X})^{-1}X_iX_i^T(\mathbf{X}^T\mathbf{X})^{-1}}{1-\mathbf{H}[i,i]}\right)\left(\mathbf{X}^TY - X_iY_i\right)\\ &= \widehat{\theta}_n + \frac{(\mathbf{X}^T\mathbf{X})^{-1}X_i\widehat{Y}_i}{1-\mathbf{H}[i,i]} - \frac{(\mathbf{X}^T\mathbf{X})^{-1}X_i\mathbf{H}[i,i]Y_i}{1-\mathbf{H}[i,i]} - (\mathbf{X}^T\mathbf{X})^{-1}X_iY_i\\ &= \widehat{\theta}_n + \frac{(\mathbf{X}^T\mathbf{X})^{-1}X_i}{1-\mathbf{H}[i,i]}(\widehat{Y}_i - Y_i)\,.\end{aligned}$$

   Ainsi, plus $\mathbf{H}[i,i]$ est proche de 1, plus enlever l'observation i a un effet important sur l'estimation de θ, i.e. plus $\widehat{\theta}^{(i)}$ est différent de $\widehat{\theta}_n$.
5. Déjà vu en question 1.
6. Par la question 4., on a

$$\begin{aligned}\widehat{Y}_i^{(i)} = X_i^T\theta^{(i)} &= X_i^T\left(\widehat{\theta}_n + \frac{(\mathbf{X}^T\mathbf{X})^{-1}X_i(\widehat{Y}_i - Y_i)}{1-\mathbf{H}[i,i]}\right)\\ &= \widehat{Y}_i + \frac{\mathbf{H}[i,i](\widehat{Y}_i - Y_i)}{1-\mathbf{H}[i,i]}\,.\end{aligned}$$

   Ainsi $(1-\mathbf{H}[i,i])(\widehat{Y}_i^{(i)} - Y_i) = \widehat{Y}_i - Y_i$, et

$$\widehat{\mathrm{CV}}_n = \frac{1}{n}\sum_{i=1}^n\left(\frac{\widehat{Y}_i - Y_i}{1-\mathbf{H}[i,i]}\right)^2\,.$$

   En particulier,

$$\widehat{\mathrm{CV}}_n \geq \frac{1}{n}\sum_{i=1}^n\left(\widehat{Y}_i - Y_i\right)^2\,.$$