Dans tout l'énoncé,

- Les vecteurs sont des vecteurs colonnes.
- Leb est la mesure de Lebesque sur $\mathbb{R}$ et $\mathrm{Leb}^{\otimes n}$ est la mesure de Lebesgue sur $\mathbb{R}^n$.
- Pour $p > 0$ et $\lambda > 0$, on appelle loi $\mathrm{Gamma}(p, \lambda)$, la loi de densité donnée par [**attention : convention différente du polycopié**] :

$$f_{p,\lambda}(x) := \frac{\lambda^{-p}}{\Gamma(p)} \exp(-x/\lambda)\; x^{p-1} \;, \quad x > 0 \;.$$

On note Gcdf_p et Gqf_p la fonction de répartition et la fonction quantile de la $\mathrm{Gamma}(p, 1)$.

Nous rappelons

- Si la variable aléatoire Y est distribuée suivant une loi $\mathrm{Gamma}(p, \lambda)$ avec $p > 0$, $\lambda > 0$, alors Y/λ est distribuée suivant une loi $\mathrm{Gamma}(p, 1)$.
- Si la variable aléatoire Y est distribuée suivant une loi $\mathrm{Gamma}(p, \lambda)$ avec $p > 0$, $\lambda > 0$ alors $\mathbb{E}[Y] = p\lambda$ et $\mathrm{Var}(Y) = p\lambda^2$.
- La loi du χ^2 centré à n degrés de liberté est une loi $\mathrm{Gamma}(n/2, 2)$.
- Soit $(X_1, \ldots, X_n)$ n variables aléatoires indépendantes, et pour tout $i \in \{1, \ldots, n\}$, la loi de X_i est la loi $\mathrm{Gamma}(p_i, \lambda)$ pour $p_i > 0$, $\lambda > 0$. Alors, $\sum_{i=1}^n X_i$ est distribuée suivant une loi $\mathrm{Gamma}(\sum_{i=1}^n p_i, \lambda)$.

Exercice 1. Soit $Z = (X_1, \ldots, X_n)$ un n-échantillon du modèle $(\mathbb{R}_+, \mathcal{B}(\mathbb{R}_+), \{p_\theta.\mathrm{dLeb}, \theta \in \Theta := \mathbb{R}_+^*\})$ où p_θ est la densité de Weibull définie par

$$x \mapsto p_\theta(x) = \frac{c}{\theta} x^{c-1} \mathrm{e}^{-x^c/\theta} \mathbb{1}_{\mathbb{R}_+}(x) \;, \quad c \text{ est une constante positive connue}$$

Soit $\theta_0 > 0$. On considère le test

$$H_0 : \; \theta \leq \theta_0, \quad \text{contre} \quad H_1 : \; \theta > \theta_0 \;.$$

1. Montrer que, pour tout $\theta \in \Theta$, $(X_i)^c$ est, sous $\mathbb{P}_\theta$, distribuée suivant une loi exponentielle de paramètre θ, de densité $q_\theta(y) := \theta^{-1} \mathrm{e}^{-y/\theta} \mathbb{1}_{\mathbb{R}_+}(y)$ [qui est aussi $\mathrm{Gamma}(1, \theta)$].
2. Montrer que la famille $\{p_\theta(\cdot), \theta \in \Theta\}$ est à rapport de vraisemblance monotone.
3. Soit $\alpha \in \,]0, 1[$. Déterminer un test $\mathrm{U.P.P.}(\alpha)$ pour ces hypothèses. On explicitera le calcul des constantes dans la définition de la fonction critique du test en fonction de la fonction quantile $\mathrm{Gqf}_n(\cdot)$.
4. Montrer que la fonction puissance du test est donnée par

$$\theta \mapsto 1 - \mathrm{Gcdf}_n\left(\theta_0 \, \mathrm{Gqf}_n(1 - \alpha)/\theta\right)$$

Exercice 2. Soit $(X_1, \ldots, X_n)$ un n-échantillon du modèle $(\mathbb{R}, \mathcal{B}(\mathbb{R}), \{p_\theta.\mathrm{dLeb}, \theta \in \Theta\})$, où $\theta := (\mu, \sigma^2) \in \Theta := \mathbb{R} \times \mathbb{R}_+^*$, et p_θ est la densité de probabilité d'une loi Gaussienne de moyenne μ et variance σ^2. Nous considérons les estimateurs de μ^2 donnés par

$$T_{1,n} := \left(\bar{X}_n\right)^2 \quad \text{où} \quad \bar{X}_n := \frac{1}{n} \sum_{i=1}^n X_i$$

$$T_{2,n} := \left(\bar{X}_n\right)^2 - \frac{1}{n} S_n^2 \quad \text{où} \quad S_n^2 := \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X}_n)^2 \;.$$

On rappelle que Y est une variable aléatoire gaussienne centrée réduite, alors $\mathbb{E}[Y^4] = 3$.

1. Montrer que, pour tout $\theta \in \Theta$, sous $\mathbb{P}_{n,\theta}$
   (a) la variable aléatoire $\bar{X}_n$ suit une loi gaussienne dont on déterminera la moyenne et la variance.
   (b) la variable S_n^2 est indépendante de $\bar{X}_n$ et $(n-1)S_n^2/\sigma^2$ est distribuée suivant une loi Gamma dont on précisera les paramètres.
   (c) Montrer que $\mathbb{E}_\theta[S_n^2] = \sigma^2$ and $\mathbb{E}_\theta[S_n^4] = \sigma^4(1 + 2/(n-1))$.
2. Calculer le biais de l'estimateur et le risque quadratique de l'estimateur $T_{1,n}$ de μ^2.
3. Calculer le biais et le risque quadratique de l'estimateur $T_{2,n}$ de μ^2.
4. L'estimateur $T_{1,n}$ de μ^2 est-il admissible pour le risque quadratique ?
5. Montrer que les suites d'estimateurs $\{T_{1,n}, n \geq 2\}$ et $\{T_{2,n}, n \geq 2\}$ sont consistantes.
6. Montrer que pour tout $\theta = (\mu, \sigma^2) \in \mathbb{R}^* \times \mathbb{R}_+^*$ (i.e. $\mu \neq 0$), $\sqrt{n}(T_{1,n} - \mu^2)$ converge sous $\mathbb{P}_{n,\theta}$ vers une loi gaussienne dont on déterminera la variance. Même question pour $\sqrt{n}(T_{2,n} - \mu^2)$.
7. Montrer que pour tout θ tel que $\mu = 0$, $nT_{1,n}$ converge en loi sous $\mathbb{P}_{n,\theta}$ vers une loi que l'on caractérisera. Même question pour $nT_{2,n}$. Commenter.

Exercice 3. Soit $Z := (X_1, \ldots, X_n)$ un n-échantillon du modèle statistique du modèle exponentiel translaté $\mathcal{E}(a, \lambda)$ de densité par rapport à la mesure de Lebesgue

$$\left(\mathbb{R}; \mathcal{B}(\mathbb{R}); \{f_\theta \cdot \mathrm{Leb}, \theta := (a, \lambda) \in \Theta := \mathbb{R} \times \mathbb{R}_+^*\}\right), \quad \text{où} \quad f_\theta(x) := \lambda^{-1} \mathrm{e}^{-(x-a)/\lambda} \mathbb{1}_{[a,\infty[}(x)$$

Nous notons, pour $\theta \in \Theta$,

$$p_\theta(x_1, \cdots, x_n) := \prod_{k=1}^n f_\theta(x_k),$$

et $\mathbb{P}_\theta := p_\theta \cdot \mathrm{Leb}^{\otimes n}$. Soit $X_{1:n} \leq X_{2:n} \leq \cdots \leq X_{n:n}$ les statistiques d'ordre de l'échantillon ; puisque les variables aléatoires $(X_k)_{k=1}^n$ ont des lois à densité par rapport à la mesure de Puisque les variables aléatoires $(X_k)_{k=1}^n$ ont des lois à densité par rapport à la mesure de Lebesgue, on a $\mathbb{P}_\theta(X_i = X_j) = 0$ pour tout $i \neq j$ et $\theta \in \Theta$. Par suite, il existe, $\mathbb{P}_\theta$-p.s. une unique permutation, que nous notons $\sigma(Z) = [\sigma_1(Z), \ldots, \sigma_n(Z)]$ de $\{1, \cdots, n\}$ telle que

$$X_{k:n} = X_{\sigma_k(Z)}, \qquad X_{\sigma_1(Z)} < X_{\sigma_2(Z)} < \cdots < X_{\sigma_n}(Z), \quad \mathbb{P}_\theta - \text{p.s.}$$

Nous notons $\mathcal{S}_n$ l'ensemble des permutations de l'ensemble $\{1, \ldots, n\}$.

1. Montrer que pour toute fonction h mesurable bornée

$$\mathbb{E}_\theta\left[h\left(X_{1:n}, \cdots, X_{n:n}\right)\right] = \sum_{\tau \in \mathcal{S}_n} \mathbb{E}_\theta\left[h\left(X_{\tau_1}, \cdots, X_{\tau_n}\right) \mathbb{1}_{\{X_{\tau_1} < X_{\tau_2} < \cdots < X_{\tau_n}\}}\right] \tag{1}$$

2. En déduire que sous $\mathbb{P}_\theta$, la loi jointe des statistiques d'ordre $X_{1:n}, X_{2:n}, \ldots, X_{n:n}$ admet une densité par rapport à $\mathrm{Leb}^{\otimes n}$, densité donnée par

$$(y_1, \cdots, y_n) \mapsto \varphi(\theta; y_1,\ y_2,\ \ldots,\ y_n) := n!\, p_\theta(y_1, \cdots, y_n)\ \mathbb{1}_{\{a \leq y_1 < y_2 < \cdots < y_n\}}\ .$$

On considère la transformation

$$U_1 := nX_{1:n}, \qquad U_k := (n-k+1)(X_{k:n} - X_{(k-1):n}) \;, \quad \text{pour } k \in \{2,\dots,n\} \;; \tag{2}$$

Remarquons que pour tout $(u_1,\dots,u_n) \in \mathbb{R}^n$,

$$\sum_{i=1}^{n} u_i = nu_1 + \sum_{i=2}^{n} (n-i+1)(u_i - u_{i-1}) \;. \tag{3}$$

3. Montrer que, sous $\mathbb{P}_\theta$, les variables aléatoires $(U_1,\dots,U_n)$ sont indépendantes. Déterminer également la loi de chacune de ces variables.
4. Démontrer que sous $\mathbb{P}_\theta$, la statistique $T_1 := \sum_{i=1}^{n}(X_{i:n} - X_{1:n})$ est indépendante de $X_{1:n}$ et suit une loi Gamma$(n-1,\lambda)$.
5. Construire, en utilisant T_1, un intervalle de confiance pour λ avec une probabilité de couverture de $1-\alpha$, où $\alpha \in]0,1[$.
6. Établir que sous $\mathbb{P}_\theta$, la statistique $T_2 := X_{1:n}$ suit une loi $\mathcal{E}(a, \lambda/n)$.
7. Justifier *sans calcul* que la fonction $a \mapsto F_n(a) := (T_2 - a)/\{(n-1)^{-1}T_1\}$ est pivotale pour le paramètre θ.
8. **Question bonus !** Prouver que, pour tout $\theta \in \Theta$, la loi de $F_n(a)$ par rapport à la mesure de Lebesgue est donnée par l'équation suivante :

$$t \mapsto n\left(1 + \frac{nt}{n-1}\right)^{-n} \mathbb{1}_{[0,\infty[}(t) \;. \tag{4}$$

9. En se servant de la fonction pivotale $F_n(a)$, construire un intervalle de confiance pour a avec une probabilité de couverture de $1-\alpha$.

Exercice 4. On cherche à expliquer une variable quantitative Y_i (réponse) par une régression linéaire à un facteur $[1, x_i]^\top$, $i \in \{1,\dots,n\}$, $x_i \in \mathbb{R}$. On pose $\theta := (\boldsymbol{\beta}, \sigma^2) \in \Theta := \mathbb{R} \times \mathbb{R}^* \times \mathbb{R}_+^*$, avec $\boldsymbol{\beta} := [\beta_0, \beta_1]^\top$. Nous posons $\mathbf{Y} := [Y_1,\dots,Y_n]^\top$, $\mathbf{y} := [y_1,\dots,y_n]^\top$. Nous supposons que, pour tout $\theta \in \Theta$, $\mathbf{Y}$ est un vecteur Gaussien, de moyenne $[\beta_0 + \beta_1 x_1, \dots, \beta_0 + \beta_1 x_n]^\top$ et de covariance $\sigma^2 \mathrm{I}_n$, i.e. la densité de $\mathbf{Y}$ par rapport à la mesure de Lebesgue sur $\mathbb{R}^n$ est

$$p_\theta(\mathbf{y}) := \frac{1}{\sqrt{2\pi\sigma^2}^n} \exp\left(-\frac{1}{2}\sum_{i=1}^{n}(Y_i - \beta_0 - \beta_1 x_i)^2\right)$$

Posons $\mathbf{x} := [x_1,\dots,x_n]^\top$, $\mathbf{1} := [1,\dots,1]^\top \in \mathbb{R}^n$, et $\mathbf{X} := [\mathbf{1}, \mathbf{x}]$. Avec ces notations, nous pouvons écrire de façon équivalente

$$p_\theta(\mathbf{y}) = \frac{1}{\sqrt{2\pi\sigma^2}^n} \exp\left(-\frac{1}{2}\|\mathbf{Y} - \mathbf{X}\boldsymbol{\beta}\|^2\right)$$

Nous supposons que la matrice $\mathbf{X}$, de taille $n \times 2$, est de rang 2 et nous posons $H := \mathbf{X}(\mathbf{X}^\top\mathbf{X})^{-1}\mathbf{X}^\top$ le projecteur orthogonal sur Vect$(\mathbf{X})$.

Nous notons

$$\bar{Y} := \frac{1}{n}\sum_{i=1}^{n} Y_i = \frac{\mathbf{1}^\top \mathbf{Y}}{\|\mathbf{1}\|^2} \quad \text{et} \quad \bar{x} := \frac{1}{n}\sum_{i=1}^{n} x_i = \frac{\mathbf{1}^\top \mathbf{x}}{\|\mathbf{1}\|^2} \;.$$

1. Montrer que les estimateurs du maximum de vraisemblance $\hat{\boldsymbol{\beta}} := [\hat\beta_0, \hat\beta_1]^\top$ et $\hat\sigma^2$ de (β_0, β_1) et σ^2 sont donnés par

$$\hat{\boldsymbol{\beta}} = \begin{bmatrix} \hat\beta_0 \\ \hat\beta_1 \end{bmatrix} := (\mathbf{X}^\top\mathbf{X})^{-1}\mathbf{X}^\top\mathbf{Y} = \begin{bmatrix} \bar{Y} - \hat\beta_1 \bar{x} \\ \frac{\sum_{i=1}^n Y_i(x_i - \bar{x})}{\sum_{i=1}^n (x_i-\bar{x})^2} \end{bmatrix} \tag{5}$$

$$\hat\sigma^2 := \frac{1}{n}\sum_{i=1}^n (Y_i - \hat\beta_0 - \hat\beta_1 x_i)^2 = \frac{1}{n}\|(\mathrm{I}_n - H)\mathbf{Y}\|^2 \tag{6}$$

Nous remarquons que $\mathbb{P}_\theta$-p.s., $\hat\beta_1 \neq 0$.

2. Démontrer que, sous $\mathbb{P}_{\boldsymbol{\beta},\sigma^2}$, $\hat{\boldsymbol{\beta}}$ est un vecteur gaussien dont on précisera la moyenne et la covariance.
3. Montrer que les statistiques $\hat\sigma^2$ et $\hat{\boldsymbol{\beta}}$ sont indépendantes. Déterminer la distribution de la statistique $\hat\sigma^2$.

Nous posons, $S^2 := (n/(n-2))\hat\sigma^2$ et pour $\boldsymbol{\beta} := [\beta_0, \beta_1]^\top \in \mathbb{R} \times \mathbb{R}^*$: $\phi := -\beta_0/\beta_1$, et $\delta := -\phi + \bar{x}$.

4. Montrer que, pour tout $\theta \in \Theta$,
$$\mathbb{E}_\theta\left[\bar{Y} - \delta\hat\beta_1\right] = 0\ .$$
5. Montrer que, pour tout $\theta \in \Theta$,
$$\mathrm{Cov}_\theta(\bar{Y}, \hat\beta_1) = 0\,.$$
6. Montrer que, pour tout $\theta \in \Theta$,
$$\mathrm{Var}_\theta\left(\bar{Y} - \delta\hat\beta_1\right) = \sigma^2 w(\delta) \quad \text{où} \quad w(\delta) := \frac{1}{n} + \frac{\delta^2}{\sum_{i=1}^n (x_i - \bar{x})^2}\ .$$
7. Montrer que pour tout $\theta \in \Theta$, sous $\mathbb{P}_\theta$, $\bar{Y} - \delta\hat\beta_1$ est distribué suivant une loi normale centrée de variance $\sigma^2 w(\delta)$ et est indépendante de S^2.

Nous posons
$$T(\delta) := \frac{\bar{Y} - \delta\hat\beta_1}{S\sqrt{w(\delta)}}$$

8. Montrer que pour tout $\theta \in \Theta$, sous $\mathbb{P}_\theta$, $T^2(\delta)$ est distribuée suivant une loi de Fisher à $(1, n-2)$ degrés de liberté.
9. En déduire une région de confiance pour δ de probabilité de couverture $1-\alpha$, $\alpha \in]0,1[$, de la forme $A(\alpha)\delta^2 + B(\alpha)\delta + C(\alpha) \leq 0$ [on déterminera $A(\alpha)$, $B(\alpha)$, et $C(\alpha)$]. Sous quelle condition cette région de confiance est-elle un intervalle ?

Exercice 5. Dans cet exercice, nous considérons des observations i.i.d. $(X_i, Y_i)_{1\leq i\leq n}$ où Y_i prend des valeurs dans $\{-1, 1\}$ et X_i prend des valeurs dans $\mathbb{R}^d$. Pour un x donné dans $\mathbb{R}^d$, nous considérerons le classificateur suivant : $\hat{h}_n(x) = Y_{\phi_n(x)}$ où

$$\phi_n(x) = \arg\min_{i\in[1:n]} \|x - X_i\|$$

En d'autres termes, $\phi_n(x)$ est l'indice du voisin le plus proche de x parmi l'ensemble de données $\{X_i; i \in [1:n]\}$. Dans tout l'exercice, nous considérons une paire de variables aléatoires (X, Y) ayant la même loi que (X_i, Y_i) pour tout $i \geq 1$. Et pour éviter toute ambiguïté dans la définition de l'indice $\phi_n(X)$, nous supposons qu'avec probabilité 1, tous les $\|X - X_i\|$ pour tout $i \in \mathbb{N}$ sont strictement différents.

Rappelons que le classificateur optimal de Bayes est défini par

$$h^\star(X) = \begin{cases} 1 & \text{si } \mathbb{P}(Y = 1|X) > \mathbb{P}(Y = -1|X) \\ -1 & \text{sinon} \end{cases}$$

et rappelons qu'en définissant $\eta(X) = \mathbb{P}(Y = 1|X)$ et $r^\star(X) = \min(\eta(X), 1 - \eta(X))$, nous avons

$$R^\star = \mathbb{P}(Y \neq h^\star(X)) = \mathbb{E}[r^\star(X)] \leq \mathbb{P}(Y \neq h(X))$$

pour tout autre classificateur $h : \mathbb{R}^d \to \{-1, 1\}$. Le but de cet exercice est de montrer la borne

$$R^\star \leq \lim_{n\to\infty} \mathbb{P}(Y \neq \hat{h}_n(X)) \leq 2R^\star(1 - R^\star)$$

Définissons $X_{(n)} = X_{\phi_n(X)}$ le voisin le plus proche de X parmi l'ensemble $\{(X_i); i \in [1:n]\}$. Nous admettons que,

— la suite $\{X_{(n)},\ (n) \in \mathbb{N}\}$ *converge en probabilité* vers X :

$$X_{(n)} \overset{\mathbb{P}-\text{prob}}{\longrightarrow} X\ ,$$

— la fonction η est continue,

1. Montrez que, pour tout $x \in \mathbb{R}^d$, $r^\star(x)(1 - r^\star(x)) = \eta(x)(1 - \eta(x))$.
2. Montrez que pour tout $i \in [1:n]$, nous avons presque sûrement,

$$\mathbb{P}(Y = -1, Y_i = 1, \phi_n(X) = i|X, X_{1:n}) = (1 - \eta(X))\eta(X_i)\mathbb{1}_{\{i\}}(\phi_n(X))$$

3. En notant que, $\eta(X_{(n)}) = \sum_{i=1}^n \eta(X_i)\mathbb{1}_{\{i\}}(\phi_n(X))$, déduisez que, presque sûrement,

$$\mathbb{P}(Y = -1, \hat{h}_n(X) = 1|X, X_{1:n}) = (1 - \eta(X))\eta(X_{(n)}),$$

4. De la même manière, montrez que

$$\mathbb{P}(Y = 1, \hat{h}_n(X) = -1|X, X_{1:n}) = \eta(X)(1 - \eta(X_{(n)})),$$

5. Montrez que $\lim_{n\to\infty} \mathbb{E}[\eta(X_{(n)})] = \mathbb{E}[\eta(X)]$.
6. Montrez que :

$$\lim_{n\to\infty} \mathbb{P}(Y \neq \hat{h}_n(X)) = 2\mathbb{E}[r^\star(X)(1 - r^\star(X))]$$

7. Déduisez que

$$\lim_{n\to\infty} \mathbb{P}(Y \neq \hat{h}_n(X)) \leq 2R^\star(1 - R^\star)$$

et conclure.

Solution . [Exercice 1]

1. Soit h une fonction borélienne bornée. Pour tout $\theta \in \Theta$, nous avons, en utilisant le changement de variable $y = x^c$

$$\begin{aligned}\mathbb{E}_\theta[h(X_1^c)] &= \int_0^{+\infty} h(x^c)\theta^{-1} c x^{c-1} \exp(-x^c/\theta) \mathrm{d}x \\ &= \int_0^{+\infty} h(y)\theta^{-1} \exp(-y/\theta) \mathrm{d}y \ .\end{aligned}$$

2. Pour tout $(x_1, \dots, x_n) \in \mathbb{R}^+$, nous avons

$$p_\theta(x_1, \dots, x_n) = \theta^{-n} c^n \prod_{i=1}^n x_i^c \mathrm{e}^{-\sum_{i=1}^n x_i^c/\theta}.$$

qui est de la forme $p_\theta(z) = h(z)\psi(T(z);\theta)$ avec $T(z) = \sum_{i=1}^n x_i^c$, $h(z) = c^n \prod_{i=1}^n x_i^c$ et $\psi(t,\theta) = \exp(-t/\theta - n\ln(\theta))$.

3. Pour $\theta > \theta'$, le rapport

$$\frac{\psi(t;\theta)}{\psi(t;\theta')} = \exp\left(t\frac{\theta - \theta'}{\theta\theta'} - n\ln(\theta/\theta')\right)$$

est une fonction monotone croissante de t. La fonction critique du test U. P. P.(α) est donné par $\phi(z) = \mathbb{1}_{\{T(z)>c\}}$ où c est donné par

$$\mathbb{P}_{\theta_0}(T(Z) > c) = \alpha \ .$$

Sous $\mathbb{P}_{\theta_0}$, $T(Z)$ est distribué suivant (voir (1)) suivant une loi Gamma(n, θ_0) ; par conséquent, $T(Z)/\theta_0$ est distribué suivant une loi Gamma$(n, 1)$. Le seuil critique du test est donc donné par $c = \theta_0 \,\mathrm{Gqf}_n(1 - \alpha)$.

4. La fonction puissance du test est donnée par

$$\beta_\phi(\theta) = \mathbb{P}_\theta(T(Z) > \theta_0 \,\mathrm{Gqf}_n(1 - \alpha))$$

Comme sous $\mathbb{P}_\theta$, pour tout $\theta \in \Theta$, $T(Z)/\theta$ est distribuée suivant une loi Gamma$(n, 1)$, nous avons

$$\beta_\phi(\theta) = \mathbb{P}_\theta(T(Z)/\theta > \theta_0 \,\mathrm{Gqf}_n(1 - \alpha)/\theta) = 1 - \mathrm{Gcdf}_n(\theta_0 \,\mathrm{Gqf}_n(1 - \alpha)/\theta)$$

Solution . [Exercice 2]

1. Il s'agit d'une application directe du théorème IV.3.24
   (a) $\bar{X}_n$ est distribuée suivant une loi gaussienne de moyenne μ et de variance σ^2/n.
   (b) $(n-1)S_n^2/2$ est distribuée suivant une loi du χ^2 centré à $(n-1)$ degrés de liberté, donc une loi Gamma$((n-1)/2, 2)$ en utilisant le rappel.
   (c) Nous avons

$$\mathbb{E}_\theta[S_n^4] = \frac{\sigma^4}{(n-1)^2}\left[\left(\frac{n-1}{\sigma^2}S_n^2\right)^2\right] = \sigma^4\left(1 + \frac{2}{n-1}\right)$$

2. En utilisant la décomposition biais-variance de l'erreur quadratique et les résultats de la question précédente, il vient

$$\mathbb{E}_{n,\theta}[\bar{X}_n^2] = \mu^2 + \sigma^2/n \ ,$$

dont on déduit que le biais de l'estimateur est σ^2/n.
Sous $\mathbb{P}_\theta$, $\bar{X}_n - \mu$ est une variable aléatoire gaussienne de moyenne nulle et de variance σ^2/n. centrée ; remarquons que son moment d'ordre 3 sont nuls ; son moment d'ordre 2 est égal à sa variance. Le risque quadratique vaut, par définition : $\mathbb{E}_{n,\theta}[(\bar{X}_n^2 - \mu^2)^2]$.
Nous avons

$$\begin{aligned}\mathbb{E}_\theta[(\bar{X}_n^2 - \mu^2)^2] &= \mathbb{E}_\theta[(\{\bar{X}_n - \mu + \mu)^2 - \mu^2)^2] \\ &= \mathbb{E}_\theta[(\{\bar{X}_n - \mu\}^2 + 2\mu\{\bar{X}_n - \mu\})^2] = \mathbb{E}_\theta[(\bar{X}_n - \mu)^4] + 4\mu^2\sigma^2/n\end{aligned}$$

en utilisant la question précédente dans la dernière égalité. En utilisant le rappel sur le moment d'ordre 4 d'une loi gaussienne centrée réduite, nous en déduisons

$$\mathbb{E}_\theta[(\bar{X}_n^2 - \mu^2)^2] = 3\frac{\sigma^4}{n^2} + 4\mu^2\frac{\sigma^2}{n}.$$

3. Nous avons

$$\mathbb{E}_{n,\theta}\left[T_{2,n}\right] = \mathbb{E}_{n,\theta}\left[T_{1,n}\right] - \frac{1}{n}\mathbb{E}_{n,\theta}\left[S_n^2\right] = \mu^2 + \frac{\sigma^2}{n} - \frac{1}{n}\mathbb{E}_{n,\theta}\left[S_n^2\right].$$

En utilisant (1), S_n^2 est une estimateur sans biais de la variance des variables aléatoires $\{X_i\}_{i=1}^n$; par suite : $\mathbb{E}_{n,\theta}\left[S_n^2\right] = \sigma^2$. On en déduit que

$$\mathbb{E}_{n,\theta}\left[T_{2,n}\right] = \mu^2$$

et l'estimateur est sans biais.

4. Soit $\varepsilon > 0$. D'après l'inégalité de Markov

$$\mathbb{P}_{n,\theta}\left[|T_{1,n} - \mu^2| \geq \varepsilon\right] \leq \frac{1}{\varepsilon^2}\mathbb{E}_{n,\theta}\left[|T_{1,n} - \mu^2|^2\right].$$

En utilisant la question 2, $\lim_{n\to\infty}\mathbb{E}_{n,\theta}\left[|T_{1,n} - \mu^2|^2\right] = 0$ dont on déduit la convergence en probabilité de $T_{1,n}$ vers μ^2. Ainsi, la suite $\{T_{1,n}, n \geq 2\}$ est consistante.
Nous avons

$$T_{2,n} = T_{1,n} - \frac{1}{n}S_n^2.$$

Nous venons d'établir que $T_{1,n} \overset{\mathbb{P}_{n,\theta}-\text{prob}}{\longrightarrow} \mu^2$. Par ailleurs, la loi faible des grands nombres entraine que $(n-1)S_n^2/n \overset{\mathbb{P}_{n,\theta}-\text{prob}}{\longrightarrow} \sigma^2$. Nous en déduisons, par le lemme de Slutsky, que $S_n^2/n \overset{\mathbb{P}_{n,\theta}}{\Longrightarrow} 0$ et donc $S_n^2/n \overset{\mathbb{P}_{n,\theta}-\text{prob}}{\longrightarrow} 0$. Par suite $T_{2,n} \overset{\mathbb{P}_{n,\theta}-\text{prob}}{\longrightarrow} \mu^2$.

5. Etudions la loi asymptotique de $\sqrt{n}(T_{1,n} - \mu^2)$. Nous écrivons $T_{1,n} = g(\bar{X}_n)$ où g est la fonction de classe C^1 sur $\mathbb{R}$ définie par $g(x) := x^2$. Le TLC pour des v.a. i.i.d. entraine

que $\sqrt{n}(\bar{X}_n - \mu) \overset{\mathbb{P}_{n,\theta}}{\Longrightarrow} \mathrm{N}(0, \sigma^2)$. En appliquant la méthode delta avec la fonction g, nous avons

$$\sqrt{n}(T_{1,n} - g(\mu)) \overset{\mathbb{P}_{n,\theta}}{\Longrightarrow} g'(\mu)\, \mathrm{N}(0, \sigma^2) \equiv \mathrm{N}(0, 4\mu^2\sigma^2).$$

Etudions la loi asymptotique de $\sqrt{n}(T_{2,n} - \mu^2)$. Nous écrivons

$$\sqrt{n}(T_{2,n} - \mu^2) = \sqrt{n}(T_{1,n} - \mu^2) - \frac{1}{\sqrt{n}} S_n^2.$$

Nous avons $S_n^2 \overset{\mathbb{P}_{n,\theta}-\text{prob}}{\longrightarrow} \sigma^2$, dont nous déduisons que $S_n^2/\sqrt{n} \overset{\mathbb{P}_{n,\theta}-\text{prob}}{\longrightarrow} 0$. De plus, nous venons de prouver la normalité asymptotique du premier terme. Par le lemme de Slutsky, nous en déduisons que $\sqrt{n}(T_{2,n} - \mu^2)$ et $\sqrt{n}(T_{1,n} - \mu^2)$ ont même limite en loi :

$$\sqrt{n}(T_{2,n} - \mu^2) \overset{\mathbb{P}_{n,\theta}}{\Longrightarrow} \mathrm{N}(0, 4\mu^2\sigma^2).$$

6. Soit θ tel que $\mu = 0$. Nous avons

$$nT_{1,n} = \frac{1}{n}\left(\sum_{i=1}^n X_i\right)^2 = \left(\frac{1}{\sqrt{n}}\sum_{i=1}^n X_i\right)^2.$$

Or, sous $\mathbb{P}_{n,\theta}$, $X_i \sim \mathrm{N}(0, \sigma^2)$ et les v.a. X_i sont i.i.d. Par suite, sous $\mathbb{P}_{n,\theta}$, $\sum_{i=1}^n X_i \sim \mathrm{N}(0, n\sigma^2)$ dont on déduit que

$$\frac{1}{\sqrt{n}\sigma}\sum_{i=1}^n X_i \sim \mathrm{N}(0,1).$$

Nous avons donc établi que sous $\mathbb{P}_{n,\theta}$

$$nT_{1,n} = \left(\frac{1}{\sqrt{n}}\sum_{i=1}^n X_i\right)^2 \sim \sigma^2\chi^2(1).$$

Nous avons

$$nT_{2,n} = nT_{1,n} - S_n^2.$$

Puisque $S_n^2 \overset{\mathbb{P}_{n,\theta}-\text{prob}}{\longrightarrow} \sigma^2$, nous déduisons du lemme de Slutsky que

$$nT_{2,n} \overset{\mathbb{P}_{n,\theta}}{\Longrightarrow} \sigma^2\left(\chi^2(1) - 1\right).$$

Solution . 1. Par construction nous avons

$$\begin{aligned}&\mathbb{E}_\theta[h(X_{1:n}, \dots, X_{n:n})] = \mathbb{E}_\theta[h(X_{\sigma_1(Z)}), \dots, h(X_{\sigma_n(Z)})] \\ &= \sum_{\tau \in \mathcal{S}_n} \mathbb{E}_\theta[h(X_{\tau_1}, \dots, X_{tau_n})\mathbb{1}_{\{\tau = \sigma(Z)\}}] == \sum_{\tau \in \mathcal{S}_n} \mathbb{E}_\theta[h(X_{\tau_1}, \dots, X_{tau_n})\mathbb{1}_{\{X_{\tau_1} < \dots X_{\tau_n}\}}]\end{aligned}$$

2. Comme sous $\mathbb{P}_\theta$, $(X_{\tau_1}, \dots, X_{\tau_n})$ a la même loi que $X_1, \dots, X_n$, nous en déduisons

$$\mathbb{E}_\theta[h(X_{1:n}, \dots, X_{n:n})] = n! \iint h(y_1, \dots, y_n) p_\theta(y_1, \dots, y_n) \mathbb{1}_{\{y_1 < \dots < y_n\}} \mathrm{dLeb}^{\otimes n}(y_1, \dots, y_n)$$

3. Soit h une fonction mesurable bornée. Notons que

$$
\begin{aligned}
&\mathbb{E}_\theta[h(U_1,\dots,U_n)] = \mathbb{E}_\theta[h(nX_{1:n},(n-1)(X_{2:n}-X_{1:n}),\dots,X_{n:n})] \\
&= n! \iint h(ny_1,(n-2)(y_2-y_1),\dots,y_n)\lambda^{-n}\exp\left(-\lambda^{-1}\sum_{i=1}^n (y_i-a)\right)\mathbb{1}_{\{a<y_1<\dots<y_n\}}\mathrm{dLeb}^{\otimes}(y_1,\dots,y_n) \\
&= n! \iint h(ny_1,(n-2)(y_2-y_1),\dots,y_n)\lambda^{-n}\exp\left(-\lambda^{-1}n(y_1-a)\right) \\
&\qquad \times \prod_{i=2}^n \exp\left(\lambda^{-1}(n-i+2)(y_i-y_{i-1})\right)\mathbb{1}_{\{a<y_1<\dots<y_n\}}\mathrm{dLeb}^{\otimes}(y_1,\dots,y_n)
\end{aligned}
$$

où nous avons utilisé (3). Nous en déduisons

$$
\begin{aligned}
&\mathbb{E}_\theta[h(U_1,\dots,U_n)] \\
&= \iint h(u_1,\dots,u_n)\lambda^{-n}\exp(-\lambda^{-1}(u_1-na))\prod_{i=2}^n \exp(-\lambda^{-1}u_i)\mathbb{1}_{\mathbb{R}^+}(u_i)\mathbb{1}_{\{a<u_1\}}\mathrm{dLeb}^{\otimes n}(u_1,\dots,u_n)
\end{aligned}
$$

ce qui démontre que les variables $U_1, U_2, \dots, U_n$ sont indépendantes sous $\mathbb{P}_\theta$. La variable U_1 est distribuée suivant une loi $\mathcal{E}(na,\lambda)$. Les variables $U_2,\dots,U_n$ sont distribuées suivant une loi $\mathcal{E}(0,\lambda)$.

4. En utilisant (3), nous avons

$$
\sum_{i=1}^n (X_{i:n}-X_{1:n}) = \sum_{i=2}^n (n-i+1)(X_{i:n}-X_{i-1:n}) = \sum_{i=2}^n U_i.
$$

Comme sous $\mathbb{P}_\theta$, les variables aléatoires $U_2,\dots,U_n$ sont indépendantes de loi Gamma$(1,\lambda)$, $\sum_{i=2}^n U_i$ est distribuée suivant une loi Gamma$(n-1,\lambda)$.

5. Pour tout $\theta \in \Theta$, la fonction T_1/λ est distribuée suivant la loi Gamma$(n-1,1)$. Elle est donc pivotale. En notant $a_\alpha = \mathrm{Gqf}(\alpha/2)$ et $b_\alpha = \mathrm{Gqf}(1-\alpha/2)$ on a donc pour tout $\theta \in \Theta$,

$$
\mathbb{P}_\theta(T_1/\lambda \in [a_\alpha, b_\alpha]) = 1-\alpha\ .
$$

Par conséquent, $[T_1/b_\alpha, T_1/a_\alpha]$ est un intervalle de confiance de couverture $1-\alpha$.

6. Comme $nX_{1:n}$ est distribué suivant une loi $\mathcal{E}(na,\lambda)$, nous avons pour toute fonction h mesurable bornée

$$
\begin{aligned}
\mathbb{E}_\theta[h(X_{1:n})] = \mathbb{E}_\theta[h(U_1/n)] &= \int h(u_1/n)\lambda^{-1}\exp(\lambda^{-1}(u_1-na))\mathbb{1}_{\{u_1\geq na\}}\mathrm{dLeb}(u_1) \\
&\int (n/\lambda)h(v_1)\exp((n/\lambda)(v_1-a))\mathbb{1}_{\{v_1\geq a\}}\mathrm{dLeb}(v_1)
\end{aligned}
$$

et donc $X_{1:n}$ est $\mathcal{E}(a,\lambda/n)$.

7. Sous $\mathbb{P}_\theta$, $(T_2-a)/\lambda$ est $\mathcal{E}(0,1)$ et est indépendant de T_1/λ qui est distribuée suivant une loi Gamma$(n-1,1)$. La fonction $F_n(a)$ est donc pivotale.

8. **[Question bonus]** : Soit h une fonction mesurable bornée. Nous avons

$$\begin{aligned}\mathbb{E}_\theta[h(F_n(a))] &= \int_a^{+\infty}\int_0^{+\infty} h\left(\frac{(n-1)(t_2-a)}{t_1}\right)\frac{n}{\lambda}\exp\left(-\frac{n}{\lambda}(t_2-a)\right)\frac{t_1^{n-2}}{(n-2)!\lambda^{n-1}}\exp\left(-\frac{t_1}{\lambda}\right)\mathrm{d}t_1\mathrm{d}t_2 \\ &= \int_0^{\infty} h(t)\left\{\int \frac{nt_1^{n-1}}{(n-1)!\lambda^n}\exp\left(-\frac{t_1}{\lambda}\left(1+\frac{nt}{n-1}\right)\right)\mathrm{d}t_1\right\}\mathrm{d}t\end{aligned}$$

Nous concluons en remarquant que pour tout $s > 0$,

$$\int_0^{+\infty}\frac{v^{n-1}}{\Gamma(n)}\exp(-v/s)\mathrm{dLeb}(v) = s^n\ .$$

9. Choissons deux constantes $0 < c_1 < c_2$ telles que

$$\int_{c_1}^{c_2} n\left(1+\frac{nt}{n-1}\right)^{-n} dt = 1- \mid \alpha.$$

Alors un intervalle de confiance pour a de niveau de couverture $1-\alpha$ est

$$(T_2 - c_2T_1, T_2 - c_1T_1)\,.$$

Solution . 1.

2. Remarquons que $\mathbb{E}_\theta[\hat\beta_1] = \beta_1$ et $\mathbb{E}_\theta[\bar Y] = \beta_0 + \beta_1\bar x$. Par conséquent :

$$\mathbb{E}_\theta[\bar Y - \delta\hat\beta_1] = \beta_0 + \beta_1\bar x - \delta\beta_1 = 0$$

3. En posant $\mathbf{u} = \mathbf{x} - \bar x\mathbf{1}$, nous avons

$$\mathrm{Cov}_\theta(\bar Y, \hat\beta_1) = \frac{1}{n\|\mathbf{u}\|^2}\mathbf{1}^\top\,\mathrm{Cov}_\theta(Y)\mathbf{u} = \sigma^2\frac{\mathbf{1}^\top\mathbf{u}}{n\|\mathbf{u}\|^2} = 0$$

4. Nous avons

$$\begin{aligned}\mathrm{Var}_\theta(\bar Y - \delta\hat\beta_1) &= \mathrm{Var}_\theta(\bar Y) + \delta^2\,\mathrm{Var}_\theta(\hat\beta_1) \\ &= \frac{\sigma^2}{n} + \delta^2\frac{\sigma^2}{\|\mathbf{u}\|^2}\ .\end{aligned}$$

5. Remarquons que

$$\mathbf{Y} - \delta\hat\beta_1 = \mathbf{v}^\top\mathbf{Y} \quad \text{où}\mathbf{v} = \left(\frac{1}{n}\mathbf{1} - \frac{\delta}{\|\mathbf{u}\|^2}\mathbf{u}\right)^\top\mathbf{Y},$$

Par conséquent $\mathbf{Y} - \delta\hat\beta_1$ est un vecteur Gaussien ; la moyenne et la covariance ont été déterminées dans les deux questions précédentes.

6. Notons que $(\mathrm{I}_n - H)\mathbf{v} = 0$. Par conséquent, en utilisant le corollaire IV.3.9, les vecteurs aléatoires $\mathbf{Y} - \delta\hat\beta_1$ et $(\mathrm{I}_n - H)\mathbf{Y}$ sont indépendants.

7. Considérons la décomposition

$$T(\delta) = \frac{\bar{Y} - \delta\hat{\beta}_1}{\sigma\sqrt{w(\delta)}} \frac{1}{\sqrt{S^2/\sigma^2}}$$

Le numérateur est une variable Gaussienne centrée réduite. $(n-2)S^2/\sigma^2$ suit une loi du χ^2 centré à $n-2$ degrés de liberté. Donc $T^2(\delta)$ est distribué suivant une loi de Fisher à $(1, n-2)$ degrés de liberté.

8. Nous remarquons que la fonction $T^2(\delta)$ est une fonction pivotale. En notant $\mathsf{F}^{\alpha}_{1,n-2}$ le quantile $1-\alpha$ de la loi de Fisher à $(1, n-2)$ d.l, nous avons pour tout $\theta \in \Theta$,

$$\mathbb{P}_\theta(T^2(\delta) \leq \mathsf{F}^{\alpha}_{1,n-2}) = 1 - \alpha \ .$$

Ceci définit la région de confiance

$$\delta^2 \left\{ \hat{\beta}_1^2 - \frac{S^2 F^{\alpha}_{1,n-2}}{\sum (x_i - \bar{x})^2} \right\} - 2\delta\bar{Y}\hat{\beta}_1 + \left(\bar{Y}^2 - \frac{1}{n} S^2 F^{\alpha}_{1,n-2} \right) \leq 0$$

Solution . — Si $\eta(X) \leq 1/2$, alors $r^\star(X) = \eta(X)$ et $r^\star(X)(1 - r^\star(X)) = \eta(X)(1-\eta(X))$. Si $\eta(X) > 1/2$, alors $r^\star(X) = 1 - \eta(X)$ et $r^\star(X)(1 - r^\star(X)) = \eta(X)(1-\eta(X))$.

— Nous avons

$$\begin{aligned}\mathbb{P}(Y = -1, Y_i = 1, \phi_n(X) = i | X, X_{1:n}) &= \mathbb{E}\left[\mathbb{1}_{\{Y=-1\}} \mathbb{1}_{\{Y_i=1\}} \mathbb{1}_{\{\phi_n(X)=i\}} | X, X_{1:n}\right] \\ &= \mathbb{E}(\mathbb{1}_{\{Y=-1\}} | X) \mathbb{E}(\mathbb{1}_{\{Y_i=1\}} | X_i) \mathbb{1}_{\{\phi_n(X)=i\}} \\ &= (1 - \eta(X))\eta(X_i)\mathbb{1}_{\{\phi_n(X)=i\}}\end{aligned}$$

—

$$\begin{aligned}\mathbb{P}(Y = -1, \hat{h}_n(X) = 1 | X, X_{1:n}) &= \sum_{i=1}^n \mathbb{P}(Y = -1, Y_i = 1, \phi_n(X) = i | X, X_{1:n}) \\ &= \sum_{i=1}^n (1 - \eta(X))\eta(X_i)\mathbb{1}_{\{\phi_n(X)=i\}} \\ &= (1 - \eta(X))\eta(X_{(n)})\end{aligned}$$

— Nous avons

$$\begin{aligned}\mathbb{P}(Y \neq \hat{h}_n(X)) &= \mathbb{E}[\mathbb{P}(Y = -1, \hat{h}_n(X) = 1 | X, X_{1:n})] + \mathbb{E}[\mathbb{P}(Y = 1, \hat{h}_n(X) = -1 | X, X_{1:n})] \\ &= \mathbb{E}[\eta(X_{(n)})(1 - \eta(X))] + \mathbb{E}[\eta(X)(1 - \eta(X_{(n)}))] \to 2\mathbb{E}[\eta(X)(1 - \eta(X))]\end{aligned}$$

où nous avons utilisé que $0 \leq \eta(X_{(n)})(1 - \eta(X)) \leq 1$ et $\lim_{n\to\infty} \eta(X_{(n)})(1 - \eta(X)) = \eta(X)(1 - \eta(X))$ presque sûrement. Et le même raisonnement vaut pour la deuxième espérance $\mathbb{E}[\eta(X)(1 - \eta(X_{(n)}))]$.

— C'est une simple adaptation des deux questions précédentes.

— Le résultat se tient immédiatement de la question précédente et de la première question.

— D'après la question précédente,

$$\begin{aligned}\lim_{n\to\infty} \mathbb{P}(Y \neq \hat{h}_n(X)) &= 2\mathbb{E}[r^\star(X)(1-r^\star(X))] \\ &= 2\left[\mathbb{E}[r^\star(X)] - \mathbb{E}[r^\star(X)^2]\right] \\ &\leq 2\left[\mathbb{E}[r^\star(X)] - \mathbb{E}^2[r^\star(X)]\right] = 2R^\star(1-R^\star)\end{aligned}$$

— Puisque $\hat{h}_n$ est un classificateur, nous avons aussi

$$R^\star \leq \mathbb{P}(Y \neq \hat{h}_n(X))$$

En combinant avec la question précédente, nous obtenons finalement

$$R^\star \leq \lim_{n\to\infty} \mathbb{P}(Y \neq \hat{h}_n(X)) \leq 2R^\star(1-R^\star)$$