Sorbonne Université, Master 1 **MU4MA015, Statistique, 2023-2024**
Cours : A. Guyader TD : A. Ben-Hamou, A. Godichon et M. Sangnier

Correction du TD 5

Exercice 1

1. D'après le cours, un modèle de translation $f_\theta(x) = f(x-\theta)$ est régulier si la densité f est continue sur $\mathbb{R}$ et C^1 par morceaux, avec
$$I := \int_{\mathbb{R}} \frac{f'(y)^2}{f(y)} \mathbb{1}_{f(y)>0}\, dy < +\infty,$$
auquel cas l'information de Fisher pour une observation est constante égale à $I(\theta) = I$ pour tout θ. Rappelons qu'une fonction définie sur un segment $[a,b]$ est dite continue et C^1 par morceaux si elle est continue sur $[a,b]$ et s'il existe une subdivision $a_0 = a < a_1 < \cdots < a_n = b$ telle que chaque restriction de f à $]a_i, a_{i+1}[$ se prolonge en une fonction de classe C^1 sur $[a_i, a_{i+1}]$. Une fonction définie sur $\mathbb{R}$ est dite continue et C^1 par morceaux si elle l'est sur tout segment contenu dans cet intervalle.
2. Etudions les différents cas.
   (a) Il s'agit du modèle $\mathcal{N}(\theta,1)$, qui est un modèle de translation de la loi $\mathcal{N}(0,1)$ de densité $f(y) = \exp(-y^2/2)/(2\pi)^{1/2}$. Les conditions de régularité sont vérifiées puisque f est C^1 sur $\mathbb{R}$. De plus, en notant Y une variable gaussienne centrée réduite, on obtient
$$\int_{\mathbb{R}} \frac{f'(y)^2}{f(y)} \mathbb{1}_{f(y)>0}\, dy = \int_{\mathbb{R}} y^2 \exp(-y^2/2)/(2\pi)^{1/2}\, dy = \mathbb{E}Y^2 = 1 < +\infty.$$
   L'information de Fisher est donc $I(\theta) = 1$.
   (b) Quel que soit $x > 0$ fixé, la fonction $\theta \mapsto f_\theta(x) = \theta^{-1}\mathbb{1}_{[x,+\infty[}(\theta)$ est discontinue en $\theta = x$, donc le modèle n'est pas régulier.
   (c) Il s'agit du modèle de translation de densité $f(y) = c_1(1-y^2)\mathbf{1}_{[-1,1]}(y)$. On a
$$\int_{-1}^{1} (1-y^2)dy = 2 - 2/3 = 4/3 \implies c_1 = 3/4.$$
   Clairement,
$$f(y) = (3/4)(1-y^2)\mathbb{1}_{y\in[-1,1]}$$
   est continue sur $\mathbb{R}$ et C^1 sur $\mathbb{R}\backslash\{-1,1\}$, points où f et f' admettent des limites à droite et à gauche, donc les restrictions $f\Big|_{]-\infty,1[}$, $f\Big|_{]-1,1[}$ et $f\Big|_{]1,\infty[}$ peuvent être prolongées en des fonctions C^1 respectivement sur $]-\infty,1]$, $[-1,1]$ et $[1,\infty[$. Ceci assure que f est C^1 par morceaux sur $\mathbb{R}$. La dérivée de f est presque partout donnée par $f'(y) = -(3/2)y\mathbb{1}_{y\in[-1,1]}$ donc $(f'(y))^2/f(y)$ vaut, à constante multiplicative près, $y^2/(1-y^2)\mathbb{1}_{y\in[-1,1]}$. Or, par parité, on a :
$$\int_{-1}^{1} \frac{y^2}{1-y^2}\,\mathrm{d}y = -2 + 2\int_0^1 \frac{1}{(1-y)(1+y)}\,\mathrm{d}y = -2 + 2\int_0^1 \frac{1}{u(2-u)}\,\mathrm{d}u \geq -2 + \int_0^1 \frac{\mathrm{d}u}{u}.$$
   Puisque $u \mapsto \frac{1}{u}$ n'est pas intégrable au voisinage de 0, la dernière intégrale n'est pas finie et le modèle n'est donc par régulier.

(d) On a encore un modèle de translation et cette fois

$$\int_{-1}^{1}(1-y^2)^2dy = 2\int_0^1(1-2y^2+y^4)dy = 16/15 \implies c_2 = 15/16.$$

Clairement, $f(y) = c_2(1-y^2)^2\mathbb{1}_{y\in[-1,1]}$ est C^1 sur $\mathbb{R}$: noter qu'il n'y a pas de problème en ±1 puisque $f'(y) = 2c_2(-2y)(1-y^2)\mathbb{1}_{y\in[-1,1]}$. Donc $(f'(y))^2/f(y)$ vaut, à constante multiplicative près, $y^2(1-y^2)^2/(1-y^2)^2\mathbb{1}_{y\in[-1,1]} = y^2\mathbb{1}_{y\in[-1,1]}$ qui est intégrable sur $\mathbb{R}$. Le modèle est régulier et l'information de Fisher vaut

$$I(\theta) = 16c_2\int_{-1}^{1}y^2dy = 16c_2 2/3 = 10.$$

Exercice 2

1. Le modèle considéré est le modèle de translation de la loi de Laplace de densité

$$f(x) = \frac{1}{2}\exp(-|x|),$$

sur $\mathbb{R}$. La fonction $y \mapsto f(y)$ est continue sur $\mathbb{R}$ et C^1 sur $\mathbb{R}$ sauf en 0, point où elle admet des dérivées à droite et à gauche qui sont les limites de $f'(y)$, donc elle est continue et C^1 par morceaux. La fonction $y \mapsto ((f'(y))^2/f(y)\mathbb{1}_{f(y)>0}(y)$ est intégrable, en effet

$$f'(y) = \begin{cases} -1/2e^{-y} & \text{si} \quad y>0 \\ 1/2e^{y} & \text{si} \quad y<0 \end{cases} = -\text{sign}(y)\frac{1}{2}e^{-|y|}.$$

Ainsi

$$(f'(y))^2/f(y)\mathbb{1}_{f(y)>0}(y) = \frac{1}{2}e^{-|y|}.$$

Donc le modèle est régulier, d'information de Fisher constante égale à

$$I = \int_{\mathbb{R}}\frac{1}{2}e^{-|y|}dy = 1.$$

2. L'estimateur au maximum de vraisemblance est défini par :

$$\widehat{\theta}_{MV} = \arg\max_{\theta\in\mathbb{R}}\Big(\frac{1}{2^n}\exp\big(-\sum_{i=1}^{n}|X_i-\theta|\big)\Big) = \arg\min_{\theta\in\mathbb{R}}\Big(\sum_{i=1}^{n}|X_i-\theta|\Big) \ni x_{1/2}(n),$$

où $x_{1/2}(n)$ est la médiane empirique de $(X_1,\ldots,X_n)$ (cf. TD 3 Exercice 1). De plus, comme f_θ est symétrique par rapport à θ et de fonction de répartition strictement croissante dans un voisinage de θ ($f_\theta(\theta) = 1/2 > 0$), sa médiane vaut $x_{1/2} = \theta$. Ainsi, puisque $f_\theta(x_{1/2}) = 1/2 > 0$, le théorème de normalité asymptotique des quantiles empiriques s'applique et on obtient :

$$\sqrt{n}\big(x_{1/2}(n) - x_{1/2}\big) = \sqrt{n}(x_{1/2}(n) - \theta) \rightsquigarrow \mathcal{N}(0,1) = \mathcal{N}(0,1/I(\theta)).$$

3. On en déduit que cet estimateur est asymptotiquement efficace.

Exercice 3

1. Nous avons $f_\theta(x) = f(x/\theta)/\theta$ par la méthode de la fonction muette et le changement de variable $x = \theta y$. Ainsi, nous obtenons
$$f_\theta(x) = \frac{1}{2\theta(1 - x/\theta)^{1/2}} \mathbb{1}_{x \in [0,\theta[}.$$

2. Pour tout $x > 0$ fixé, la fonction
$$\theta \mapsto f_\theta(x) = \frac{1}{2\theta(1 - x/\theta)^{1/2}} \mathbb{1}_{]x,+\infty[}(\theta)$$
n'est pas continue en $\theta = x$ donc le modèle n'est pas régulier.

3. La vraisemblance du modèle s'écrit
$$\prod_{i=1}^{n} f_\theta(X_i) = \left(\frac{1}{2\theta}\right)^n \prod_{i=1}^{n} (1 - X_i/\theta)^{-1/2} \mathbb{1}_{]X_{(n)}+\infty[}(\theta),$$
qui est une fonction nulle sur $]0, X_{(n)}]$ et non nulle strictement décroissante sur $]X_{(n)}, +\infty[$. Il est clair que la vraisemblance n'atteint pas son maximum et que l'EMV n'est pas défini. Toutefois, la vraisemblance tend vers l'infini lorsque $\theta \to X_{(n)}$ par la droite. On propose donc de prendre comme estimateur :
$$\widehat{\theta}_n = X_{(n)}.$$
On obtient la loi limite en calculant pour tout $x \geq 0$,
$$\begin{aligned} \mathbb{P}_\theta(n(\theta - \widehat{\theta}_n)^{1/2} \leq x) &= \mathbb{P}_\theta(\theta - (x/n)^2 \leq \widehat{\theta}_n) \\ &= 1 - F(1 - x^2/(n^2\theta))^n, \end{aligned}$$
où F est la fonction de répartition associée à la densité f. On obtient facilement pour tout $u \in \mathbb{R}$,
$$F(u) = \begin{cases} 1 & \text{si } u \geq 1 \\ 1 - (1-u)^{1/2} & \text{si } 0 \leq u \leq 1 \\ 0 & \text{sinon,} \end{cases}$$
de sorte que pour tout $x \geq 0$
$$\begin{aligned} \mathbb{P}_\theta(n(\theta - \widehat{\theta}_n)^{1/2} \leq x) &= 1 - \left(1 - x/(n\theta^{1/2})\right)^n \mathbb{1}_{[0, n\theta^{1/2}]}(x) \\ &\to 1 - e^{-x/\theta^{1/2}}. \end{aligned}$$
Ainsi, on déduit que
$$n(\theta - \widehat{\theta}_n)^{1/2} \rightsquigarrow \mathcal{E}(1/\sqrt{\theta}).$$

4. En observant que la médiane de la loi de densité f_θ est égale à $x_{1/2} = 3\theta/4$, un estimateur obtenu par la médiane empirique est $\widetilde{\theta}_n = 4x_{1/2}(n)/3$ où $x_{1/2}(n)$ est la médiane empirique de l'échantillon. Puisque $f_\theta(x_{1/2}) > 0$, on en déduit que $\widetilde{\theta}_n$ est un estimateur asymptotiquement normal de θ, donc convergeant à une vitesse en $\sqrt{n}$. Avec $\widehat{\theta}_n$ on obtient une vitesse en n^2, donc bien meilleure : en effet, en notant $T \sim \mathcal{E}(1/\sqrt{\theta})$, la convergence en loi ci-dessus et le théorème de continuité donnent
$$n^2(\theta - \widehat{\theta}_n) \rightsquigarrow T^2.$$
Bref, on choisira clairement $\widehat{\theta}_n$ plutôt que $\widetilde{\theta}_n$.

Remarque : Le critère de la vitesse de convergence domine les autres critères, au sens où, entre deux estimateurs, on préférera toujours celui dont la vitesse de convergence est la plus grande.

Exercice 4

1. De manière triviale, l'EMV est $\widehat{\theta}_n = \overline{X}_n = \frac{1}{n}\sum_{i=1}^n X_i$. De plus, par le TCL $\sqrt{n}(\widehat{\theta}_n - \theta) \xrightarrow{\mathcal{L}} \mathcal{N}(0,1)$. En outre, on sait de l'exercice 1 que le modèle $\mathcal{P} = \{\mathcal{N}(\theta,1)\}_{\theta\in\mathbb{R}}$ est régulier et pour tout $\theta \in \mathbb{R}$, l'information de Fisher est $I(\theta) = 1$. L'estimateur $\widehat{\theta}_n$ est donc asymptotiquement efficace.
2. Il est aisé de voir que l'EMV de η est $\widehat{\eta}_n = (\overline{X}_n)^{1/3}$. Soit alors $\varphi : x \mapsto x^{1/3}$. Nous avons $\eta = \varphi(\theta)$. Par continuité de φ sur $\mathbb{R}$ et par la LGN, $\widehat{\eta}_n = \varphi(\overline{X}_n) \xrightarrow{\mathbb{P}} \varphi(\mathbb{E}X_1) = \varphi(\eta^3) = \eta$ (ce qui montre la consistance de $\widehat{\eta}_n$).
   En outre, le modèle $\mathcal{P}$ étant régulier, le reparamétrage $\eta = \varphi(\theta)$, pour lequel $\psi = \varphi^{-1}$ est une fonction de classe C^1, conduit à l'information de Fisher :
   $$J(\eta) = \psi'(\eta)^2 I(\psi(\eta)) = 9\eta^4.$$
   On remarque qu'avec ce reparamétrage, l'information de Fisher n'est plus constante et s'annule pour $\eta = 0$.
3. Par le TCL, $\sqrt{n}(\overline{X}_n - \eta^3) \xrightarrow{\mathcal{L}} \mathcal{N}(0,1)$ (en réalité, on a même $\sqrt{n}(\overline{X}_n - \eta^3) \sim \mathcal{N}(0,1)$). De plus, φ est dérivable sur $\mathbb{R}^*$, avec pour tout $x \neq 0$: $\varphi'(x) = \frac{1}{3}x^{-2/3}$. Ainsi, si $\eta \neq 0$, $\varphi'(\eta^3) = \frac{1}{3\eta^2} \neq 0$. Par la méthode Delta, on obtient donc $\sqrt{n}(\widehat{\eta}_n - \eta) \xrightarrow{\mathcal{L}} \mathcal{N}\left(0, \frac{1}{9\eta^4}\right)$, avec $9\eta^4 = J(\eta)$.
4. Le TCL, appliqué à la question précédente, nous donne : $\sqrt{n}(\overline{X}_n - \eta^3) = \sqrt{n}\overline{X}_n \xrightarrow{\mathcal{L}} Z$, où $Z \sim \mathcal{N}(0,1)$. Ainsi, par le théorème de continuité, $\varphi(\sqrt{n}\overline{X}_n) \xrightarrow{\mathcal{L}} \varphi(Z)$, c'est à dire $n^{1/6}\widehat{\eta}_n \xrightarrow{\mathcal{L}} Z^{1/3}$.
5. Soient ϕ et Φ respectivement la densité et la fonction de répartition de Z.
   $$\forall x \in \mathbb{R} : \mathbb{P}(Z^{1/3} \leq x) = \mathbb{P}(Z \leq x^3) = \Phi(x^3).$$
   Ainsi, la densité f de $Z^{1/3}$ vaut, pour tout $x \in \mathbb{R}$, $f(x) = 3x^2\phi(x^3) = \frac{3x^2}{\sqrt{2\pi}}e^{-\frac{x^6}{2}}$. Il apparaît clairement que la variable aléatoire $Z^{1/3}$ n'est pas gaussienne (elle est même bimodale, comme l'illustre la figure 1).

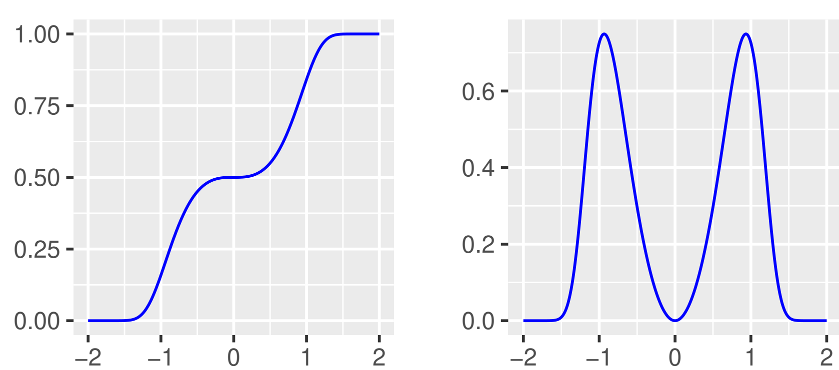

Figure 1 – Fonction de répartition (à gauche) et densité (à droite) de $Z^{1/3}$.

   De plus, $\widehat{\eta}_n$ ne converge pas à la vitesse $n^{-1/2}$ mais $n^{-1/6}$, ce qui est bien plus lent.
   L'estimateur $\widehat{\eta}_n$ est asymptotiquement efficace puisque, pour tout η tel que $J(\eta) \neq 0$, il est asymptotiquement normal de variance limite $1/J(\eta)$. Cet exemple permet simplement de constater que, en

un point où l'information de Fisher s'annule, le comportement d'un estimateur asymptotiquement efficace peut être complètement différent de ce qui se passe partout ailleurs : ici, lorsque $J(\eta) = 0$, i.e. lorsque $\eta = 0$, la vitesse n'est plus en $1/\sqrt{n}$ et la loi limite n'est plus gaussienne.

Exercice 5

1. La vraisemblance s'écrit
$$L_n(\theta) = \prod_{i=1}^{n} \theta(1-\theta)^{X_i - 1} = \theta^n (1-\theta)^{n(\overline{X}_n - 1)},$$
d'où la log-vraisemblance
$$\ell_n(\theta) = n\left(\ln\theta + (\overline{X}_n - 1)\ln(1-\theta)\right).$$
On voit que $\lim_{\theta\to 0^+} \ell_n(\theta) = -\infty$.

   Sur l'événement $\{\overline{X}_n > 1\}$ On a $\lim_{\theta\to 1^-} \ell_n(\theta) = -\infty$ et la fonction $\theta \mapsto \ell_n(\theta)$ admet un unique maximum au point où sa dérivée s'annule :
$$\ell_n'(\theta) = n\left(\frac{1}{\theta} - \frac{\overline{X}_n - 1}{1-\theta}\right) = 0 \iff \theta = \frac{1}{\overline{X}_n}.$$

   Sur l'événement $\{\overline{X}_n = 1\}$ ℓ_n est strictement croissante sur $]0,1[$ et n'atteint jamais son maximum. L'EMV n'est donc pas défini (cela arrive avec probabilité θ^n, qui est exponentiellement petite) mais on convient de prendre comme estimateur $\widehat{\theta}_n = \frac{1}{\overline{X}_n}$.
Par la loi des grands nombres, on sait que $\overline{X}_n$ tend presque sûrement vers $\mathbb{E}[X_1] = 1/\theta > 0$, donc par continuité
$$\widehat{\theta}_n = \frac{1}{\overline{X}_n} \xrightarrow[n\to\infty]{p.s.} \frac{1}{\mathbb{E}[X_1]} = \theta,$$
et $\widehat{\theta}_n$ est consistant. Par le TCL, on sait aussi que
$$\sqrt{n}\left(\overline{X}_n - \frac{1}{\theta}\right) \xrightarrow[n\to\infty]{\mathcal{L}} \mathcal{N}(0, (1-\theta)/\theta^2),$$
donc la delta-méthode assure que
$$\sqrt{n}\left(\widehat{\theta}_n - \theta\right) \xrightarrow[n\to\infty]{\mathcal{L}} \mathcal{N}(0, \theta^2(1-\theta)).$$

2. Notons X une v.a. de densité f_θ avec $0 < \theta < 1$. Puisque
$$f_\theta(x) = \theta(1-\theta)^{x-1} = \theta e^{(x-1)\ln(1-\theta)},$$
il est clair que, pour tout $x \in \mathbb{N}^\star$, la fonction $\theta \mapsto f_\theta(x)$ est C^1 sur $]0,1[$ donc les deux premiers points de la définition d'un modèle régulier sont satisfaits. Pour le dernier point, le score vaut, d'après ce qui précède et le rappel de l'énoncé sur la loi géométrique,
$$\ell_\theta'(X) = \frac{1}{\theta} - \frac{X-1}{1-\theta} = \frac{X - \mathbb{E}_\theta[X]}{\theta - 1},$$
donc son moment d'ordre 2 est
$$\mathbb{E}_\theta\left[\left(\ell_\theta'(X)\right)^2\right] = \frac{\mathrm{Var}_\theta(X)}{(\theta-1)^2} = \frac{1}{\theta^2(1-\theta)},$$
qui est bien une fonction continue sur $]0,1[$. Par conséquent, le modèle $(f_\theta)_{0<\theta<1}$ est régulier, d'information de Fisher $I(\theta) = I_1(\theta) = (\theta^2(1-\theta))^{-1}$. En particulier, d'après la question précédente, l'estimateur $\widehat{\theta}_n$ est asymptotiquement efficace.

3. Soit $0 < \alpha < 1$. On veut tester $H_0 : \theta = 1/2$ contre $H_1 : \theta \neq 1/2$. Sous H_0, d'après ce qui précède,
$$\sqrt{n}\left(\widehat{\theta}_n - \frac{1}{2}\right) \xrightarrow[n\to\infty]{\mathcal{L}} \mathcal{N}(0,1/8) \iff 2\sqrt{2n}\left(\widehat{\theta}_n - \frac{1}{2}\right) \xrightarrow[n\to\infty]{\mathcal{L}} \mathcal{N}(0,1).$$
En notant $\Phi^{-1}(1-\alpha/2)$ le quantile d'ordre $1-\alpha/2$ d'une loi normale centrée réduite, on en déduit que le test consistant à accepter H_0 si et seulement si
$$2\sqrt{2n}\left|\widehat{\theta}_n - \frac{1}{2}\right| \leq \Phi^{-1}(1-\alpha/2)$$
est de niveau asymptotique α. Réciproquement, pour une réalisation $\widehat{\theta}_n(\omega)$ de l'estimateur, la p-valeur est donc
$$\alpha_0 = 2\left(1 - \Phi(2\sqrt{2n}|\widehat{\theta}_n(\omega) - 1/2|)\right).$$
Par exemple, si $n = 50$ et $\overline{x}_n = \overline{x}_{50} = 5/3$, donc $\widehat{\theta}_n(\omega) = 3/5$, la p-valeur associée est
$$\alpha_0 = 2(1-\Phi(2)) \approx 0.05.$$

Exercice 6

Partie 1

1. (a) On vérifie que $\mathbb{P}_\theta$-p.s., $\theta \in]X_{(n)} - 1, X_{(1)} + 1[$ donc $]X_{(n)} - 1, X_{(1)} + 1[\neq \emptyset$. Par ailleurs, la vraisemblance est égale à
$$L_n(\theta') = c_k^n \prod_{i=1}^{n} (1 - (X_i - \theta')^2)^{k-1} \mathbb{1}_{\theta' \in [X_{(n)}-1, X_{(1)}+1]}.$$
$L_n(\theta')$ est nulle en dehors de $]X_{(n)} - 1, X_{(1)} + 1[$ et est strictement positive sur $]X_{(n)} - 1, X_{(1)} + 1[$. Donc, s'il existe $\widehat{\theta}_n$ qui la maximise, on a $\mathbb{P}_\theta$-p.s., $\widehat{\theta}_n \in]X_{(n)} - 1, X_{(1)} + 1[$

   (b) La log-vraisemblance s'écrit, pour $\theta' \in]X_{(n)} - 1, X_{(1)} + 1[$:
$$\ell_n(\theta') = n \ln c_k + (k-1) \sum_{i=1}^{n} \ln(1 - (X_i - \theta')^2),$$
et sa dérivée est égale à
$$\begin{aligned} \ell_n'(\theta') &= 2(k-1) \sum_{i=1}^{n} \frac{X_i - \theta'}{1 - (X_i - \theta')^2} \\ &= (k-1)\left(\sum_{i=1}^{n} \frac{1}{1 - (X_i - \theta')} - \sum_{i=1}^{n} \frac{1}{1 + (X_i - \theta')}\right). \end{aligned}$$
Lorsque $\theta' \to X_{(n)} - 1$, alors $1 - (X_i - \theta') \to X_{(n)} - X_i$, qui vaut 0 si i est tel que $X_i = X_{(n)}$, et qui est strictement positif sinon. De plus $1 + (X_i - \theta') \to 2 + X_i - X_{(n)} > 0$ $\mathbb{P}_\theta$-p.s., pour tout i. Donc $\ell_n'(\theta') \to +\infty$ quand $\theta' \to X_{(n)} - 1$.
Le raisonnement est le même pour montrer que $\ell_n'(\theta') \to -\infty$ quand $\theta' \to X_{(1)} + 1$. Par ailleurs, l'application $\theta' \mapsto \ell_n'(\theta')$ est strictement décroissante sur $]X_{(n)} - 1, X_{(1)} + 1[$ puisque pour tout $x \in \mathbb{R}$, les fonctions $\theta' \mapsto \frac{1}{1-(x-\theta')}$ et $\theta' \mapsto -\frac{1}{1+(x-\theta')}$ sont décroissantes sur $]x-1, +\infty[$ et $]-\infty, 1+x[$ respectivement.

(c) L_n est continue et strictement positive sur $]X_{(n)} - 1, X_{(1)} + 1[$ donc elle atteint son maximum dans $]X_{(n)} - 1, X_{(1)} + 1[$, segment ouvert. Donc les lieux de maximum sont des points critiques de L_n, et aussi de ℓ_n. Or, par les questions précédentes, ℓ'_n est continue sur $]X_{(n)} - 1, X_{(1)} + 1[$, strictement décroissante et telle que $\ell'_n(X_{(1)}+1) = -\infty$ et $\ell'_n(X_{(n)}-1) = +\infty$. Elle ne s'annule donc qu'en un unique point, qui est l'unique lieu de maximum de la vraisemblance. Ce point, $\widehat{\theta}_n$, vérifie :

$$\sum_{i=1}^{n}(X_i - \widehat{\theta}_n)\prod_{j\neq i}(1-(X_j-\widehat{\theta}_n)^2) = 0,$$

donc il est racine d'un polynôme de degré $1 + 2(n-1) = 2n-1$.

2. (a) On a, pour $0 \leq \epsilon \leq 2$:

$$\begin{aligned}\mathbb{P}_\theta[X_1 \leq \theta - 1 + \epsilon] &= \int_{\theta-1}^{\theta-1+\epsilon} c_k(1-(t-\theta)^2)^{k-1}dt \\ &= c_k\int_{-1}^{-1+\epsilon}(1-y^2)^{k-1}dy \\ &= c_k\int_0^{\epsilon} u^{k-1}(2-u)^{k-1}du.\end{aligned}$$

Si $0 \leq \epsilon \leq 1$, alors $1 \leq 2-u \leq 2$ donc

$$\mathbb{P}_\theta[X_1 \leq \theta - 1 + \epsilon] \geq c_k\int_0^{\epsilon} u^{k-1}du = \frac{c_k}{k}\epsilon^k.$$

On a alors

$$\begin{aligned}\mathbb{P}_\theta[X_{(1)} + 1 - \theta \geq \epsilon] &= (\mathbb{P}_\theta[X_1 \geq \epsilon + \theta - 1])^n \\ &= (1 - \mathbb{P}_\theta[X_1 < \epsilon + \theta - 1])^n \\ &\leq \left(1 - \frac{c_k\epsilon^k}{k}\right)^n.\end{aligned}$$

D'où la deuxième inégalité.

(b) Les autres inégalités se montrent par symétrie car $X - \theta \sim \theta - X$.
Remarque : les calculs ci-dessus montrent plus précisément que, lorsque $\epsilon \to 0$,

$$\mathbb{P}_\theta[X_1 \leq \theta - 1 + \epsilon] = \alpha_k\epsilon^k + o(\epsilon^k) \sim \alpha_k\epsilon^k,$$

ce qui implique que

$$\mathbb{P}_\theta[n^{1/k}(X_{(1)} + 1 - \theta) \geq \epsilon] \xrightarrow[n\to\infty]{} e^{-\alpha_k\epsilon^k}.$$

Ceci signifie que $n^{1/k}(X_{(1)} + 1 - \theta)$ converge en loi vers une loi de Weibull. Par symétrie, il en va de même pour $n^{1/k}(\theta - (X_{(n)} - 1))$. Au total, l'intervalle $[X_{(n)} - 1; X_{(1)} + 1]$, qui contient θ, a une longueur de l'ordre de $n^{-1/k}$.

(c) Pour $0 < \epsilon \leq 1$, puisque $X_{(n)} - 1 \leq \widehat{\theta}_n \leq X_{(1)} + 1$ p.s.,

$$\begin{aligned}\mathbb{P}\left(|\widehat{\theta}_n - \theta| \geq \epsilon\right) &\leq \mathbb{P}\left(\widehat{\theta}_n - \theta \geq \epsilon\right) + \mathbb{P}\left(\theta - \widehat{\theta}_n \geq \epsilon\right) \\ &\leq \mathbb{P}\left((X_{(1)} + 1) - \theta \geq \epsilon\right) + \mathbb{P}\left(\theta - (X_{(n)} - 1) \geq \epsilon\right) \\ &\leq 2(1 - C\epsilon^k)^n \\ &\xrightarrow[n\to\infty]{} 0,\end{aligned}$$

car $(1 - C\epsilon^k) \in [0, 1[$. Ainsi, $\widehat{\theta}_n \xrightarrow{\mathbb{P}} \theta$ et $\widehat{\theta}_n$ est consistant.
En outre, puisque pour $0 < \epsilon \leq 1$,

$$\mathbb{P}\left(|\widehat{\theta}_n - \theta| \geq \epsilon\right) \leq 2(1 - C\epsilon^k)^n,$$

et $(1 - C\epsilon^k) \in [0, 1[$, le Lemme de Borel-Cantelli assure que $\widehat{\theta}_n \xrightarrow{\text{p.s.}} \theta$, d'où la consistance forte.
Remarque : Nous avons aussi que, pour $0 < \epsilon \leq 1$,

$$\mathbb{P}_\theta\left(|X_{(1)} - (\theta - 1)| \geq \epsilon\right) = \mathbb{P}_\theta\left(X_{(1)} - (\theta - 1) \geq \epsilon\right) \leq (1 - C\epsilon^k)^n.$$

Ainsi, par le Lemme de Borel-Cantelli, $X_{(1)} \xrightarrow{\text{p.s.}} \theta - 1$ et $X_{(1)} + 1 \xrightarrow{\text{p.s.}} \theta$. De même, $X_{(n)} - 1 \xrightarrow{\text{p.s.}} \theta$. La consistance forte de $\widehat{\theta}_n$ vient alors de l'encadrement $X_{(n)} - 1 \leq \widehat{\theta}_n \leq X_{(1)} + 1$ p.s.

3. Le modèle considéré est un modèle de translation construit à partir de la densité $f(x) = c_k(1 - x^2)^{k-1}\mathbb{1}_{x\in[-1,1]}$. Cette densité est continue et C^1 par morceaux. De plus, $f'(x) = c_k(k-1)(-2x)(1 - x^2)^{k-2}\mathbb{1}_{x\in[-1,1]}$ donc pour tout $x \in [-1, 1]$,

$$(f'(x))^2/f(x) = 4c_k(k-1)^2x^2(1-x^2)^{2k-4-k+1} = 4c_k(k-1)^2x^2(1-x^2)^{k-3}.$$

Par conséquent, en utilisant l'indication,

$$\begin{aligned}
\int_{\mathbb{R}} (f'(x))^2/f(x)dx &= 4c_k(k-1)^2 \int_{-1}^{1} x^2(1-x^2)^{k-3}dx = 8c_k(k-1)^2 \int_0^1 x^2(1-x^2)^{k-3}dx \\
&= 4c_k(k-1)^2 \int_0^1 y^{1/2}(1-y)^{k-3}dy \\
&= 4c_k(k-1)^2 \frac{\Gamma(3/2)\Gamma(k-2)}{\Gamma(k-1/2)} < \infty.
\end{aligned}$$

Le modèle est donc régulier et l'information de Fisher vaut

$$I(\theta) = 4c_k(k-1)^2 \frac{\Gamma(3/2)\Gamma(k-2)}{\Gamma(k-1/2)}.$$

Partie 2

1. Lorsque $k \leq 2$, les calculs des questions 1 et 2 restent valident, donc $\widehat{\theta}_n$ existe (il n'est pas unique pour $k = 1$) et est fortement consistant. En revanche, pour $k = 1$, la densité f n'est pas continue, donc le modèle n'est pas régulier. Pour $k \in]1, 2]$ le calcul de la régularité débouche sur

$$\int_{\mathbb{R}} (f'(x))^2/f(x)dx = \infty,$$

donc, de nouveau, le modèle n'est pas régulier. Ainsi, on peut légitimement supposer que $\widehat{\theta}_n$ ne sera pas asymptotiquement normal lorsque $k \in [1, 2]$.

2. Par un raisonnement similaire à celui de la question 2c,

$$\begin{aligned}
\mathbb{P}_\theta\left(n^\alpha|\widehat{\theta}_n - \theta| > \epsilon\right) &\leq \mathbb{P}_\theta\left(\widehat{\theta}_n - \theta > \frac{\epsilon}{n^\alpha}\right) + \mathbb{P}_\theta\left(\theta - \widehat{\theta}_n > \frac{\epsilon}{n^\alpha}\right) \\
&\leq \mathbb{P}\left((X_{(1)} + 1) - \theta \geq \frac{\epsilon}{n^\alpha}\right) + \mathbb{P}\left(\theta - (X_{(n)} - 1) \geq \frac{\epsilon}{n^\alpha}\right) \\
&\leq 2\left(1 - C\frac{\epsilon^k}{n^{\alpha k}}\right)^n.
\end{aligned}$$

3. De la question précédente, il vient

$$\begin{aligned}
\mathbb{P}_\theta\left(n^\alpha|\widehat{\theta}_n-\theta|>\epsilon\right) &\leq 2\left(1-C\frac{\epsilon^k}{n^{\alpha k}}\right)^n \\
&= 2\,\mathrm{e}^{n\ln\left(1-C\frac{\epsilon^k}{n^{\alpha k}}\right)} \\
&= 2\,\mathrm{e}^{-C\frac{\epsilon^k}{n^{\alpha k-1}}+o\left(\frac{1}{n^{\alpha k-1}}\right)} && (\alpha k>0) \\
&\underset{n\to\infty}{\longrightarrow} \begin{cases} 2 & \text{si } \alpha k-1>0 \text{ i.e. si } \alpha>1/k \\ 0 & \text{si } \alpha k-1<0 \text{ i.e. si } \alpha<1/k \\ 2\,\mathrm{e}^{-C\epsilon^k} & \text{si } \alpha k=1 \text{ i.e. si } \alpha=1/k. \end{cases}
\end{aligned}$$

Ainsi, on a montré que le terme $|\widehat{\theta}_n-\theta|$ tend vers 0 au moins aussi rapidement que $1/n^\alpha$, tant que $\alpha<1/k$. Autrement dit, $\widehat{\theta}_n$ tend en probabilité vers θ avec une vitesse d'ordre au moins $1/n^\alpha$, pour $\alpha<1/k$.
En particulier, pour $k<2$, (en prenant $\alpha=1/2<1/k$) $\widehat{\theta}_n$ converge avec une vitesse d'ordre au moins $1/\sqrt{n}$ (donc plus rapidement que quand le modèle est régulier).
Lorsque $k=2$, la majoration précédente ne nous indique rien sur la convergence en probabilité de $\sqrt{n}|\widehat{\theta}_n-\theta|$. On ne peut donc pas affirmer que $\widehat{\theta}_n$ converge plus rapidement que la vitesse $1/\sqrt{n}$.
Pour rappel, lorsque $k>2$, nous avons vu que le modèle est régulier et que $\widehat{\theta}_n$ converge à la vitesse $1/\sqrt{n}$.

Exercice 7

1. La variable X ne pouvant prendre que les valeurs 0 ou 1, c'est une variable de Bernoulli et une mesure dominante naturelle est la mesure de comptage $\mu=\delta_0+\delta_1$. Pour $x=1$, la densité $g_\theta(x)$ est donc

$$g_\theta(1)=\mathbb{P}_\theta(X=1)=\mathbb{P}_\theta(U\leq\theta/2)=\theta/2$$

si $0<\theta\leq 1/2$, et

$$g_\theta(1)=\mathbb{P}_\theta(X=1)=\mathbb{P}_\theta(U\leq\theta-1/4)=\theta-1/4$$

si $1/2<\theta<1$. Par ailleurs, pour tout $0<\theta<1$,

$$g_\theta(0)=\mathbb{P}_\theta(X=0)=1-\mathbb{P}_\theta(X=1)=1-g_\theta(1).$$

2. Pour $x=1$, la fonction $\theta\mapsto g_\theta(1)$ est donc affine par morceaux et continue puisque

$$g_{1/2}(1)=1/4=\lim_{\theta\downarrow 1/2} g_\theta(1).$$

Il en va clairement de même pour $\theta\mapsto g_\theta(0)=1-g_\theta(1)$. Ainsi, pour tout $x\in\{0,1\}$, la fonction $\theta\mapsto g_\theta(x)$ est continue sur $]0,1[$ et $\mathcal{C}^\infty$ sauf au point $\theta=1/2$.

3. Pour tout x, la fonction $\theta\mapsto g_\theta(x)$ est absolument continue (car continue et $\mathcal{C}^1$ par morceaux) : on peut donc définir sa dérivée au sens de l'absolue continuité. En particulier, puisque quel que soit $x\in\{0,1\}$, la fonction $\theta\mapsto g_\theta(x)$ n'est pas dérivable au sens usuel en $\theta=1/2$, cette dérivée vaut par convention 0 en ce point : $g'_{1/2}(1)=g'_{1/2}(0)=0$. Or

$$\lim_{\theta\uparrow 1/2} g'_\theta(1)=1/2\neq 0\neq 1=\lim_{\theta\downarrow 1/2} g'_\theta(1),$$

et de même

$$\lim_{\theta\uparrow 1/2} g'_\theta(0)=-1/2\neq 0\neq -1=\lim_{\theta\downarrow 1/2} g'_\theta(0).$$

Ainsi, quel que soit $x \in \{0,1\}$, la fonction $\theta \mapsto g'_\theta(x)$ n'est pas continue en $\theta = 1/2$. Par conséquent, le modèle n'est pas régulier : il est impossible de définir une information de Fisher pour la valeur $\theta = 1/2$.

Exercice 8

1. La vraisemblance du modèle s'écrit

$$\prod_{i=1}^{n} f_\theta(X_i) = e^{-\sum_{i=1}^n (X_i - \theta)} \mathbb{1}_{\theta \leq X_{(1)}}$$

qui est une fonction non nulle strictement croissante sur $]0, X_{(1)}]$ et nulle sur $]X_{(1)}, +\infty[$, donc le maximum de vraisemblance vaut

$$\widetilde{\theta}_n = X_{(1)}.$$

On obtient la loi de $n(\widetilde{\theta}_n - \theta)$ en calculant pour tout $x \geq 0$,

$$\begin{aligned}\mathbb{P}_\theta(n(\widetilde{\theta}_n - \theta) \leq x) &= \mathbb{P}_\theta(\widetilde{\theta}_n \leq \theta + x/n) \\ &= 1 - \mathbb{P}_\theta(X_1 > \theta + x/n)^n \\ &= 1 - e^{-x}.\end{aligned}$$

En particulier, on a

$$n(\widetilde{\theta}_n - \theta) \sim \mathcal{E}(1).$$

Comme $\mathbb{E}_\theta(n(\widetilde{\theta}_n - \theta)) = 1$, on peut définir le nouvel estimateur

$$\widehat{\theta} = \widetilde{\theta}_n - 1/n$$

qui est sans biais. Son risque vaut

$$R(\theta, \widehat{\theta}) = \mathrm{Var}(\widehat{\theta}) = \mathrm{Var}(\widetilde{\theta}_n) = n^{-2}\mathrm{Var}(n(\widetilde{\theta}_n - \theta)) = n^{-2}.$$

2. Pour tout $x \in \mathbb{R}$, nous avons $\theta \mapsto f_\theta(x)$ qui n'est pas continue en $\theta = x$ donc le modèle n'est pas régulier. Son information de Fisher n'est donc pas définie.

3. Presque sûrement, on a $X_1 > \theta$ donc

$$\ell_1(\theta) = \ln f_\theta(X_1) = \theta - X_1$$

et

$$\ell'_1(\theta) = 1 \implies \mathrm{Var}_\theta\left(\ell'_1(\theta)\right) = 0.$$

En oubliant les hypothèses de l'inégalité de Cramér-Rao, nous écririons

$$\mathrm{Var}(\widehat{\theta}) \geq \frac{1}{nI_1(\theta)},$$

où l'information de Fisher $I_1(\theta) = \mathrm{Var}_\theta\left(\ell'_1(\theta)\right) = J_1(\theta) = 0$. Ainsi, la borne de Cramér-Rao serait invalidée.

Exercice 9

1. Si X a pour densité f, alors par hypothèse X est centrée réduite donc

$$\int_{\mathbb{R}} f_1(x)^2 dx = \int_{\mathbb{R}} x^2 f(x) dx = \mathbb{E}[X^2] = 1.$$

   Qui plus est, par définition de l'information de Fisher, on a aussi

$$\int_{\mathbb{R}} f_2(x)^2 dx = \int_{\mathbb{R}} \frac{f'(x)^2}{f(x)} dx = I.$$

   Ainsi, $(L_2, \|\cdot\|_2)$ étant un espace vectoriel normé auquel appartiennent f_1 et f_2, c'est encore le cas pour la fonction $f_1 + f_2$ et la décomposition $f_1 f_2 = \frac{1}{2}((f_1 + f_2)^2 - f_1^2 - f_2^2)$ assure que $f_1 f_2$ est intégrable sur $\mathbb{R}$ et une intégration par parties donne

$$\int_{\mathbb{R}} f_1(x) f_2(x) dx = [x f(x)]_{-\infty}^{+\infty} - \int_{\mathbb{R}} f(x) dx = [x f(x)]_{-\infty}^{+\infty} - 1.$$

   Le membre de gauche étant fini, le terme entre crochets est bien défini, ce qui implique l'existence de deux constantes $L^- := \lim_{x \to -\infty} x f(x)$ et $L^+ := \lim_{x \to +\infty} x f(x)$. Supposons $L^+ > 0$, alors on aurait $f(x) \sim L^+/x$ en $+\infty$, ce qui est incompatible avec l'intégrabilité de f. Ainsi $L^+ = 0$ et par le même raisonnement $L^- = 0$, ce qui donne finalement le résultat souhaité :

$$\int_{\mathbb{R}} (f_1(x) + f_2(x))^2 dx = I - 1.$$

2. L'équation que l'on vient d'obtenir, dont le terme de gauche est positif, permet de conclure que $I \geq 1$, avec égalité si et seulement si $f_1 + f_2 = 0$ presque partout. Puisque f est de classe C^1, c'est en fait vrai partout et on a donc, pour tout $x \in \mathbb{R}$, $f'(x)/f(x) = -x$. L'intégration de cette équation différentielle donne $f(x) = ce^{-x^2/2}$, or f est une densité donc nécessairement $c = 1/\sqrt{2\pi}$. Ainsi, dans un modèle de translation régulier, parmi les lois centrées réduites, celle pour laquelle l'information de Fisher est minimale est la gaussienne standard. Dit autrement (et rapidement), c'est pour la gaussienne que le paramètre de centrage est le plus difficile à retrouver.