TD 5 : Comparaison et optimalité des estimateurs

Exercice 1 (Régularité et information de Fisher)

1. Soit $X_1, \ldots, X_n$ i.i.d. de densité $f_\theta(x) = f(x-\theta)$ par rapport à la mesure de Lebesgue, pour une certaine densité f sur $\mathbb{R}$ connue et un paramètre inconnu $\theta \in \mathbb{R}$. Donner des conditions suffisantes sur f pour que le modèle $\mathcal{P} = \{f_t \cdot \lambda_\mathbb{R}, t \in \mathbb{R}\}$, qui est dominé par la mesure de Lebesgue sur $\mathbb{R}$, notée $\lambda_\mathbb{R}$, soit régulier et préciser la valeur de l'information de Fisher.
2. Pour chacun des modèles suivants, dire si le modèle est régulier et si oui calculer l'information de Fisher.
   (a) $\mathcal{P} = \{f_t \cdot \lambda_\mathbb{R}, t \in \mathbb{R}\}$, avec $f_t(x) = \exp(-(x-t)^2/2)/(2\pi)^{1/2}$;
   (b) $\mathcal{P} = \{f_t \cdot \lambda_\mathbb{R}, t > 0\}$, avec $f_t(x) = t^{-1}\mathbb{1}_{[0,t]}(x)$;
   (c) $\mathcal{P} = \{f_t \cdot \lambda_\mathbb{R}, t \in \mathbb{R}\}$, avec $f_t(x) = c_1(1-(x-t)^2)\mathbb{1}_{[t-1,t+1]}(x)$ et pour une constante $c_1 > 0$ à déterminer ;
   (d) $\mathcal{P} = \{f_t \cdot \lambda_\mathbb{R}, t \in \mathbb{R}\}$, avec $f_t(x) = c_2(1-(x-t)^2)^2\mathbb{1}_{[t-1,t+1]}(x)$ et pour une constante $c_2 > 0$ à déterminer.

Exercice 2 (Loi de Laplace) Soient $X_1, \ldots, X_n$ des v.a. i.i.d. de densité

$$f_\theta(x) = \frac{1}{2}e^{-|x-\theta|}, \ \ x \in \mathbb{R},$$

pour un paramètre $\theta \in \mathbb{R}$.

1. Le modèle $\mathcal{P} = \{f_t \cdot \lambda_\mathbb{R}, t \in \mathbb{R}\}$, qui est dominé par la mesure de Lebesgue sur $\mathbb{R}$, notée $\lambda_\mathbb{R}$, est-il régulier ? Si oui, calculer l'information de Fisher $I_1(\theta)$ du modèle à une observation.
2. Donner un estimateur du maximum de vraisemblance ainsi que sa loi limite.
3. Est-il asymptotiquement efficace ?

Exercice 3 (Loi Bêta) On considère la densité

$$f(y) = \frac{1}{2(1-y)^{1/2}}\mathbb{1}_{[0,1[}(y).$$

1. Si Y a pour densité f, quelle est la densité f_θ de $X = \theta Y$, où $\theta > 0$?
2. Le modèle $\mathcal{P} = \{f_t \cdot \lambda_\mathbb{R}, t > 0\}$, qui est dominé par la mesure de Lebesgue sur $\mathbb{R}$, notée $\lambda_\mathbb{R}$, est-il régulier ?
3. On observe un échantillon $X_1, \ldots, X_n$ de même loi que X. L'estimateur du maximum de vraisemblance est-il bien défini ? Proposer un estimateur $\widehat{\theta}_n$ de θ et trouver la loi limite de $n(\theta - \widehat{\theta}_n)^{1/2}$.
4. Déterminer la médiane de la loi de X. En déduire un nouvel estimateur $\widetilde{\theta}_n$. Entre $\widehat{\theta}_n$ et $\widetilde{\theta}_n$, lequel est préférable ?

Exercice 4 ($\star$) (Information de Fisher nulle) Soient $\theta \in \mathbb{R}$ un paramètre inconnu et $X_1, \ldots, X_n$ un échantillon i.i.d. de loi $\mathcal{N}(\theta, 1)$.

1. Déterminer l'EMV $\widehat{\theta}_n$ de θ et montrer qu'il est asymptotiquement efficace.
2. On définit le paramètre inconnu $\eta \in \mathbb{R}$ de sorte que $\theta = \eta^3$. Proposer un estimateur $\widehat{\eta}_n$ consistant de η. Quelle est l'information de Fisher pour ce nouveau modèle ? Que remarque-t-on ?
3. Supposons $\eta \neq 0$. Quelle est la loi limite correctement normalisée de $\widehat{\eta}_n$?
4. Plaçons-nous maintenant dans le cas où $\eta = 0$. Déterminer la loi limite de $\sqrt{n}\overline{X}_n = \frac{1}{\sqrt{n}} \sum_{i=1}^n X_i$ et en déduire la loi limite convenablement normalisée de $\widehat{\eta}_n$.
5. Calculer la densité de la loi limite obtenue à la question précédente. Conclure sur l'efficacité asymptotique de $\widehat{\eta}_n$.

Exercice 5 (Loi géométrique, décembre 2016) Soit $0 < \theta < 1$ un paramètre inconnu. On dit que la variable aléatoire X suit une loi géométrique de paramètre θ, notée $\mathcal{G}(\theta)$, si elle est à valeurs dans $\mathbb{N}^\star$ avec

$$\forall x \in \mathbb{N}^\star \qquad \mathbb{P}(X = x) = f_\theta(x) = \theta(1-\theta)^{x-1},$$

auquel cas $\mathbb{E}[X] = 1/\theta$ et $\mathrm{Var}(X) = (1-\theta)/\theta^2$. Dans toute la suite, on suppose disposer de n observations i.i.d. $X_1, \ldots, X_n$ de densité f_θ par rapport à la mesure de comptage sur $\mathbb{N}^\star$.

1. Déterminer un estimateur $\widehat{\theta}_n$ par maximum de vraisemblance. Est-il consistant ? Asymptotiquement normal ?
2. Montrer que le modèle $\mathcal{P} = \{\mathcal{G}(t), 0 < t < 1\}$ est régulier. L'estimateur $\widehat{\theta}_n$ est-il asymptotiquement efficace ?
3. Soit $0 < \alpha < 1$. On veut tester $H_0 : \theta = 1/2$ contre $H_1 : \theta \neq 1/2$. Proposer un test de niveau asymptotique α. Pour $n = 50$ et une moyenne empirique $\overline{x}_n = \overline{x}_{50} = 5/3$, préciser la p-valeur associée.

Exercice 6 (Loi Bêta dilatée et translatée) Soient $\theta \in \mathbb{R}$ et $X_1, \ldots, X_n$ i.i.d. de densité :

$$x \in \mathbb{R} \mapsto f_k(x, \theta) = c_k \left[1 - (x-\theta)^2\right]^{k-1} \mathbb{1}_{[\theta-1, \theta+1]}(x),$$

où $k \geq 1$ est connu et $c_k = \frac{\Gamma(2k)}{\Gamma(k)^2 2^{2k-1}}$. Cette densité est celle d'une variable aléatoire qui s'obtient par homothétie de rapport 2 et translation de $\theta - 1$ d'une variable de loi Bêta(k, k).
Dans toute la suite, on appellera $\widehat{\theta}_n$ un estimateur du maximum de vraisemblance et on étudiera ses propriétés.

Partie 1 On se place tout d'abord dans la situation où $k > 2$.

1. Existence de $\widehat{\theta}_n$.
   (a) Pour tout $\theta \in \mathbb{R}$, montrer que l'intervalle $[X_{(n)} - 1, X_{(1)} + 1]$ est d'intérieur non-vide $\mathbb{P}_\theta$-p.s. et que si $\widehat{\theta}_n$ existe, alors $\widehat{\theta}_n \in]X_{(n)} - 1, X_{(1)} + 1[$.
   (b) Montrer que pour $t \in]X_{(n)} - 1, X_{(1)} + 1[$, la dérivée $\ell_n'(t)$ de la log-vraisemblance s'écrit :

$$\ell_n'(t) = (k-1)\left(\sum_{i=1}^n \frac{1}{1-(X_i - t)} - \sum_{i=1}^n \frac{1}{1+(X_i - t)}\right).$$

   Prouver que

i. $\ell_n'(t) \xrightarrow[t \to X_{(n)}-1]{} \infty$;

ii. $\ell_n'(t) \xrightarrow[t \to X_{(1)}+1]{} -\infty$;

iii. $t \in]X_{(n)} - 1, X_{(1)} + 1[\mapsto \ell_n'(t)$ est strictement décroissante.

(c) En déduire que l'estimateur du maximum de vraisemblance est l'unique solution de $\ell_n'(\theta) = 0$. Montrer que la résolution de $\ell_n'(\theta) = 0$ nécessite de rechercher les racines d'un polynôme de degré élevé.

2. Consistance de $\widehat{\theta}_n$.

(a) Montrer que l'on peut trouver une constante C (ne dépendant pas de θ) telle que, pour $\varepsilon \in [0, 1]$:

$$\mathbb{P}_\theta(X_1 \leq \theta - 1 + \varepsilon) \geq C\varepsilon^k.$$

En déduire que :

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

(b) Par symétrie, montrer que pour tout $\varepsilon \in [0, 1]$:

$$\mathbb{P}_\theta(\theta + 1 - \varepsilon \leq X_1) \geq C\varepsilon^k$$

et

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

(c) En déduire que $\widehat{\theta}_n$ est consistant. Est-il fortement consistant ?

3. Le modèle statistique $\mathcal{P} = \{f_k(\cdot, t) \cdot \mu, t \in \mathbb{R}\}$, où $\lambda_\mathbb{R}$ est la mesure de Lebesgue sur $\mathbb{R}$, est-il régulier ? <u>Indication</u> : Si $\alpha > 0$ et $\beta > 0$, alors $\int_0^1 x^{\alpha-1}(1-x)^{\beta-1}dx = \Gamma(\alpha)\Gamma(\beta)/\Gamma(\alpha+\beta)$.

Partie 2 On suppose maintenant que $k \leq 2$.

1. Que dire des propriétés de $\widehat{\theta}_n$? Que dire de la régularité du modèle ?
2. Soit $\alpha > 0$. À l'aide de la question 2, montrer que pour tout $\varepsilon > 0$:

$$\mathbb{P}_\theta\left(n^\alpha|\widehat{\theta}_n - \theta| > \varepsilon\right) \leq 2\left(1 - C\frac{\varepsilon^k}{n^{\alpha k}}\right)^n.$$

3. Conclure sur la vitesse de convergence de $\widehat{\theta}_n$. Quand $k < 2$, cette convergence est-elle plus rapide que lorsque le modèle est régulier (i.e. $k > 2$) ? Qu'en est-il lorsque $k = 2$?

Exercice 7 ($\star$) **(Modèle pathologique)** Soit $0 < \theta < 1$, U une variable de loi uniforme sur $[0, 1]$, et X définie par : $X = \mathbb{1}_{U \leq \theta/2}$ si $0 < \theta \leq 1/2$ et $X = \mathbb{1}_{U \leq \theta - 1/4}$ si $1/2 < \theta < 1$.

1. Quelles valeurs peut prendre la variable X ? En déduire une mesure dominante naturelle. Donner la densité $g_\theta(x)$ par rapport à cette mesure dominante.
2. Représenter les fonctions $\theta \mapsto g_\theta(x)$ pour $x = 0$ et $x = 1$. Quelle est leur régularité (continuité, etc.) ?
3. Le modèle $\mathcal{P} = \{g_t \cdot \mu, t \in]0, 1[\}$, où $\mu = \delta_0 + \delta_1$ est la mesure de comptage sur $\{0, 1\}$, est-il régulier ?

Exercice 8 ($\star$) **(Exponentielle translatée)** Soit $X_1, \ldots, X_n$ un échantillon de la loi de densité

$$f_\theta(x) = \exp(-(x - \theta))\mathbb{1}_{[\theta, +\infty[}(x),$$

où $\theta > 0$.

1. Déterminer l'estimateur du maximum de vraisemblance $\widetilde{\theta}_n$ de θ et donner la loi de $n(\widetilde{\theta}_n - \theta)$. Construire, à partir de cet estimateur, un estimateur sans biais que l'on notera $\widehat{\theta}_n$ et dont on calculera le risque quadratique.
2. Que peut-on dire de l'information de Fisher du modèle ?
3. Soit $\ell_1 \colon \theta > 0 \mapsto \ln(f_\theta(X_1))$ la log-vraisemblance de X_1. Justifier que ℓ_1 et ℓ_1' sont bien définies $\mathbb{P}_\theta$-p.s. et calculer
$$J_1(\theta) = \mathrm{Var}_\theta\left(\ell_1'(\theta)\right).$$
Que constate-t-on si l'on cherche à appliquer l'inégalité de Cramér-Rao à $\widehat{\theta}_n$ à tout prix ?

Exercice 9 (Loi normale et minimisation de l'information de Fisher) On considère une densité centrée réduite f de classe C^1 et strictement positive sur $\mathbb{R}$. Pour le modèle de translation associé, sous réserve de convergence de l'intégrale, l'information de Fisher est donc
$$I = \int_{\mathbb{R}} \frac{f'(x)^2}{f(x)} dx.$$

1. Si on note $f_1(x) = x\sqrt{f(x)}$ et $f_2(x) = f'(x)/\sqrt{f(x)}$, montrer que
$$\int_{\mathbb{R}} (f_1(x) + f_2(x))^2 dx = I - 1.$$
2. En déduire que $I \geq 1$ avec égalité si et seulement si f est la densité d'une gaussienne centrée réduite.