TD5 - Estimation efficace, Borne de Cramér-Rao

Exercice 1.

Soit $N \in \mathbb{N}^*$ supposé connu. On considère le modèle $\mathbb{P}_\theta = \mathcal{B}(N, \theta)$ avec $\theta \in]0,1[$.
On suppose que l'on dispose d'un n-échantillon $(X_1, \dots, X_n)$ de loi $\mathbb{P}_\theta$.
On admettra que ce modèle est régulier.

1. Calculer l'information de Fisher $I_n(\theta)$ associée à ce modèle.

2. On rappelle que $\widehat{\theta}_n^{\mathrm{MV}} = \dfrac{1}{N}\overline{X}_n$. Montrer que $\widehat{\theta}_n^{\mathrm{MV}}$ est un estimateur efficace de θ.

3. Dans cette question on prend $N = 1$, c'est-à-dire qu'on considère le modèle Bernoulli $\mathcal{B}(\theta)$.
Considérons l'estimateur de θ donné par

$$T_n = \frac{X_1 + \ldots + X_n + 1}{n+2} \; .$$

a. Calculer le biais de T_n.
b. Comparer T_n et $\widehat{\theta}_n^{\mathrm{MV}}$.

Exercice 2.

On cherche à estimer le paramètre d'intensité $\lambda > 0$ de la loi exponentielle.
On considère donc le modèle exponentiel $\mathbb{P}_\theta = \mathcal{E}(\theta)$ avec $\theta > 0$.
On suppose que l'on dispose d'un n-échantillon $(X_1, \dots, X_n)$ de loi $\mathbb{P}_\theta$.
On admettra que ce modèle est régulier.

1. Calculer l'information de Fisher $I_n(\theta)$ associée à ce modèle.

2. Montrer que $\overline{X}_n$ est un estimateur efficace de $\frac{1}{\theta}$.

3. On admettra que $T_n = \dfrac{n-1}{X_1 + \ldots + X_n}$ est un ESBVM pour θ, et que sa variance est donnée par $\frac{\theta^2}{n-2}$.
Est-ce que T_n est un estimateur efficace de θ ?

Exercice 3.

Soit un modèle statistique régulier $\mathcal{P} = \{\mathbb{P}_\theta, \theta \in \Theta\}$ avec $\Theta \subset \mathbb{R}$ ouvert.
On suppose que $\theta \mapsto p(x;\theta)$ est deux fois dérivable avec

$$\int \left| \frac{\partial^2 p}{\partial \theta^2}(x;\theta) \right| \mu(dx) < \infty \quad \text{et} \quad \frac{\partial^2}{\partial \theta^2}\left(\int p(x;\theta)\mu(dx) \right) = \int \frac{\partial^2 p}{\partial \theta^2}(x;\theta)\mu(dx) \; .$$

Montrer que

$$I(\theta) = -\mathbb{E}_\theta\left[\frac{\partial^2 \log p}{\partial \theta^2}(X;\theta) \right] \; .$$

Exercice 4. Un exemple de modèle non régulier

On considère donc le modèle uniforme $\mathbb{P}_\theta = \mathcal{U}(0,\theta)$ avec $\theta > 0$.
On suppose que l'on dispose d'un n-échantillon $(X_1, \dots, X_n)$ de loi $\mathbb{P}_\theta$.

1. En écartant les problèmes de dérivabilité, que vaudrait l'information de Fisher ?
Et que vaudrait la borne de Cramér-Rao pour l'estimation de θ ?

2. Comparer cette borne avec le risque quadratique de l'estimateur $2\overline{X}_n$. Commenter.

Exercice 5.

On cherche à estimer les paramètres $m \in \mathbb{R}$ et $v = \sigma^2 > 0$ de la loi gaussienne.
On considère donc le modèle gaussien $\mathbb{P}_\theta = \mathcal{N}(m, v)$ avec $\theta = (m, v) \in \mathbb{R}\times]0, \infty[$.
On suppose que l'on dispose d'un n-échantillon $(X_1, \dots, X_n)$ de loi $\mathbb{P}_\theta$.
On admettra que ce modèle est régulier.

1. Rappeler l'expression de la densité gaussienne $p(x; \theta)$ en fonction de $\theta = (m, v)$.

2. Calculer la matrice d'information de Fisher $I_n(\theta)$ associée à ce modèle.

3. Expliciter les bornes de Cramér-Rao pour l'estimation de m et v.
(On prendra d'abord $g(\theta) = \theta_1$ puis $g(\theta) = \theta_2$.)

4. Est-ce que

$$\widehat{m}_n = \overline{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i \quad \text{et} \quad \widehat{v}_n = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X}_n)^2$$

sont des estimateurs efficaces de m et v ?

Exercice 6.

Démontrer que le modèle $\mathbb{P}_\theta = \mathcal{E}(\theta)$ avec $\theta > 0$ est régulier.