1 Exercices

Exercice 1. Soit $(X_1,\dots,X_n)$ un n-échantillon d'un modèle $(\mathbb{R}_+,\mathscr{B}(\mathbb{R}_+),\{\mathrm{Unif}([0,\theta])\ :\ \theta\in\mathbb{R}_+^*\})$. Soit $g:\mathbb{R}_+^*\to\mathbb{R}$ une fonction différentiable vérifiant $\lim_{\theta\to 0^+}\theta g(\theta)=0$.

1. Peut-on choisir $T:\mathbb{R}^+\to\mathbb{R}$ de telle sorte que $\eta_n^{(1)}=n^{-1}\sum_{i=1}^n T(X_i)$ soit un estimateur sans biais de $g:\theta\to g(\theta)$?
2. Montrer que l'on peut choisir $\tilde{T}:\mathbb{R}^+\to\mathbb{R}$ de telle sorte que $\eta_n^{(2)}=\tilde{T}(X_{n:n})$ où $X_{n:n}=\max(X_1,\dots,X_n)$ soit un estimateur sans biais de $g:\theta\mapsto g(\theta)$.

Dans la suite de l'exercice, on pose $g(\theta)=\theta$.

3. Comparer le risque quadratique des deux estimateurs $\eta_n^{(1)}$ et $\eta_n^{(2)}$. Montrer que l'estimateur $\eta_n^{(1)}$ est inadmissible pour $n>1$.
4. Pour $a>0$, calculer le risque quadratique de l'estimateur $\eta_{n,a}^{(3)}=aX_{n:n}$. Montrer que l'estimateur $\eta_n^{(2)}$ est inadmissible.

Exercice 2. Soit $(\mathsf{Z},\mathscr{Z},\{p_\theta\cdot\mu\ :\ \theta\in\Theta\})$ où $\Theta\subset\mathbb{R}$ un modèle statistique dominé. Soit T un estimateur sans biais de la fonction $g:\Theta\to\mathbb{R}$ telle que pour tout $\theta\in\Theta$, $\mathbb{E}_\theta[T^2(Z)]<\infty$.

1. Montrer que pour tout variable aléatoire Ψ telle que $\mathbb{E}_\theta[\Psi^2(Z)]<\infty$,

$$\mathrm{Var}_\theta(T(Z))\geq\frac{\{\mathrm{Cov}_\theta(T(Z),\Psi(Z))\}^2}{\mathrm{Var}_\theta(\Psi(Z))}\ .$$

2. Soient $\theta,\theta+\delta\in\Theta$. On pose

$$\Psi(z)=\begin{cases}\frac{p_{\theta+\delta}(z)}{p_\theta(z)}-1 & p_\theta(z)>0\\ 0 & \text{sinon}\end{cases}$$

On suppose de plus que $p_\theta(z)=0$ implique $p_{\theta+\delta}(z)=0$. Montrer que $\mathbb{E}_\theta[\Psi(Z)]=0$ et

$$\mathrm{Cov}_\theta(T(Z),\Psi(Z))=g(\theta+\delta)-g(\theta)\ .$$

3. En déduire une borne sur les estimateurs sans biais de $g:\theta\to g(\theta)$.
4. Sous quelles conditions peut-on en déduire la borne de Cramér-Rao ?

On suppose maintenant que $(X_1,\dots,X_n)$ est un n échantillon du modèle

$$(\mathbb{R}_+,\mathscr{B}(\mathbb{R}_+),\{\mathrm{Unif}([0,\theta])\ :\ \theta\in\Theta=\mathbb{R}_+^*\}).$$

Nous considérons la fonction $g(\theta)=\theta$.

5. Montrer que pour tout $\delta\in[-\theta,0[$, nous avons pour tout estimateur sans biais de θ

$$\mathrm{Var}_\theta(T)\geq\frac{\delta^2}{([\theta^n/(\theta+\delta)^n]-1)^2}$$

6. Soit $0 < c \leq n$. En posant $\delta = -c\theta/n$, montrer que la borne inférieure peut s'écrire $\theta^2 g_n(c)/n^2$. Déterminer $\lim_{n\to\infty} g_n(c)$.
7. Discuter le résultat en utilisant l'Exercice 1.

Exercice 3. Soient $t_1,\ldots,t_n$ des réels tels que $t_i \neq t_j$ pour deux indices $i \neq j$. On considère le modèle statistique

$$\left(\mathbb{R}^n, \mathscr{B}(\mathbb{R}^n), \left\{ \bigotimes_{i=1}^n \mathrm{N}(\beta_1 + \beta_2 t_i, \sigma^2) \,:\, (\beta_1,\beta_2,\sigma) \in \mathbb{R}^2 \times \mathbb{R}_+^* \right\}\right) .$$

On suppose que $\sigma^2 > 0$ est connu.

1. Déterminer la matrice d'information de Fisher $\mathbb{I}(\beta_1,\beta_2)$.
2. Déterminer une borne inférieure sur la variance d'un estimateur sans biais de β_1.
3. On suppose que β_2 est connu. Donner une borne inférieure sur la variance d'un estimateur sans biais de β_1.
4. Déterminer une borne inférieure d'un estimateur sans biais de $\beta_1\beta_2$.

Exercice 4. Soit q une densité sur $\mathbb{R}$ par rapport à la mesure de Lebesgue vérifiant

$$\int x q(x)\mathrm{d}x = 0 \,, \qquad \int x^2 q(x)\mathrm{d}x = 1 \,.$$

Soit $(X_1,\ldots,X_n)$, n variables aléatoires i.i.d. distribuées suivant un modèle de translation :

$$x \mapsto p(\theta,x) = \frac{1}{\sigma} q\left(\frac{x-\theta}{\sigma}\right) \,,$$

où $\theta \in \mathbb{R}$ est inconnu et $\sigma > 0$. Pout tout $\underline{a}_n = (a_{1,n},\cdots,a_{n,n}) \in \mathbb{R}^n$, on considère la classe des estimateurs linéaires, i.e. des estimateurs qui sont des combinaisons linéaires des échantillons.

$$\hat{\theta}_n(\underline{a}_n) := \sum_{i=1}^n a_{i,n} X_i \,.$$

1. Déterminer l'estimateur de θ de risque quadratique minimal dans la classe des estimateurs $(\hat{\theta}_n(\underline{a}_n), \underline{a}_n \in \mathbb{R}^n)$ et sans biais. Déterminer le risque quadratique minimal dans la classe des estimateurs linéaires sans biais.
2. Calculer une borne inférieure du risque quadratique dans la classe des estimateurs linéaires $(\hat{\theta}_n(\underline{a}_n), \underline{a}_n \in \mathbb{R}^n)$.

Exercice 5. Soient $(X_1,\ldots,X_n)$ n variables aléatoires réelles i.i.d. de densité exponentielle de paramètre $\theta > 0$:

$$x \mapsto f(\theta;x) = \theta^{-1}\mathrm{e}^{-x/\theta}\mathbb{1}_{\mathbb{R}_+}(x), \quad \theta \in \Theta := \mathbb{R}_*^+$$

par rapport à la mesure de Lebesgue sur $\mathbb{R}$. On note $X_{1:n} \leq X_{2:n} \leq \cdots \leq X_{n:n}$ les statistiques d'ordre de l'échantillon.

1. Montrer que la loi jointe des statistiques d'ordre $X_{1:n}, X_{2:n}, \ldots, X_{n:n}$ admet une densité par rapport à la mesure de Lebesgue sur $\mathbb{R}$, donnée par

$$(y_1, \cdots, y_n) \mapsto f_{1:n}(\theta; y_1,\ y_2,\ \ldots,\ y_n) = n!\, f(\theta; y_1) f(\theta; y_2) \cdots f(\theta; y_n) \mathbb{1}_{\{0 \le y_1 < y_2 < \cdots < y_n\}} .$$

On considère la transformation

$$Z_1 = nX_{1:n}, \qquad Z_k = (n-k+1)(X_{k:n} - X_{(k-1):n}) ,\ \ \forall k \geq 2 ; \tag{1}$$

ou de façon équivalente

$$X_{k:n} = \sum_{j=1}^{k} \frac{Z_j}{n-j+1} ,\ \ \forall k \geq 1 .$$

On remarquera que le jacobien de la transformation (1) est $n!$. On notera aussi que pour tout $x_i \in \mathbb{R}$,

$$\sum_{i=1}^{n} x_i = \sum_{i=1}^{n} (n-i+1)(x_i - x_{i-1}) , \qquad \text{où, par convention } \ x_0 = 0 .$$

2. Montrer que les v.a. $(Z_1, \ldots, Z_n)$ sont des variables aléatoires i.i.d. de loi exponentielle de paramètre $\theta > 0$.

Pour tout $\underline{a}_n = (a_{1,n}, \cdots, a_{n,n}) \in \mathbb{R}^d$, on considère l'estimateur donné par

$$\hat{\theta}_n(\underline{a}_n) := \sum_{i=1}^{n} a_{i,n} X_{i:n} .$$

3. Montrer que

$$\hat{\theta}_n(\underline{a}_n) = \sum_{j=1}^{n} Z_j m_{j,n} \qquad \text{où} \quad m_{j,n} := \frac{1}{n-j+1} \sum_{i=j}^{n} a_{i,n} .$$

4. Calculer le biais et le risque quadratique de l'estimateur.
5. Déterminer les poids $\underline{a}_n$ de l'estimateur sans biais de risque quadratique minimal, dans la classe des estimateurs de la forme $(\hat{\theta}_n(\underline{a}_n), \underline{a}_n \in \mathbb{R}^n)$ [indication : on déterminera les poids $m_{j,n}$ puis les poids $a_{j,n}$].
6. Déterminer la valeur des poids $\underline{a}_n$ pour lequel le risque quadratique est minimal dans la classe des estimateurs de la forme $(\hat{\theta}_n(\underline{a}_n), \underline{a}_n \in \mathbb{R}^n)$.