
2. Information de Fisher. Estimation par maximum de vraisemblance

Objectifs : Calculer une information de Fisher, l'interpréter. Savoir donner la borne de Cramer-Rao. Pratiquer l'estimation par maximum de vraisemblance. Les exercices 2.1 et 2.2 sont à faire pendant le TD, les autres sont à chercher de votre côté.

Exercice 2.1 (Information de Fisher dans un modèle à deux paramètres). On considère le modèle statistique où X_1, \dots, X_n sont des variables réelles indépendantes, avec pour tout $1 \leq i \leq n$, $X_i \sim \mathcal{N}(\alpha + \beta t_i, 1)$, où les constantes $(t_i)_{1 \leq i \leq n}$ sont connues et $\alpha, \beta \in \mathbb{R}$ sont des paramètres inconnus. On admettra que ce modèle est régulier.

1. Montrer que la matrice d'information de Fisher $I(\alpha, \beta)$ s'écrit

$$I(\alpha, \beta) = \begin{pmatrix} n & \sum_{i=1}^n t_i \\ \sum_{i=1}^n t_i & \sum_{i=1}^n t_i^2 \end{pmatrix}.$$

2. Déterminer une condition nécessaire et suffisante sur les $(t_i)_{1 \leq i \leq n}$ pour que $I(\alpha, \beta)$ soit inversible. Interpréter cette condition.
3. Sous la condition de la question 2, donner une borne inférieure sur la variance de tout estimateur non biaisé de α . (On prendra donc $h(\alpha, \beta) = \alpha$).
4. Supposons que l'on connaît β . Reprendre les calculs pour trouver $\mathcal{I}(\alpha)$ et donner dans ce cas une borne inférieure sur la variance de tout estimateur non biaisé de α .
5. Comment les bornes inférieures des questions 3 et 4 se comparent-elles ? Interpréter.

Exercice 2.2 (Deux estimateurs dans le modèle uniforme). On considère le modèle dans lequel X_1, \dots, X_n sont des variables i.i.d. suivant une loi uniforme sur $[0, \theta]$, où $\theta > 0$ est un paramètre inconnu.

1. Trouver un estimateur $\hat{\theta}_1$ de θ en utilisant la méthode des moments appliquée au premier moment.
2. Déterminer la vraisemblance dans ce modèle, et montrer qu'il existe un unique estimateur du maximum de vraisemblance $\hat{\theta}_2$. Donner son expression.
3. Calculer le biais, la variance et le risque quadratique de $\hat{\theta}_1$ et de $\hat{\theta}_2$. On calculera la densité de $\hat{\theta}_2$.
4. Quel est le meilleur estimateur ?

Exercice 2.3 (Estimation par maximum de vraisemblance dans le cas régulier). On considère un échantillon $X = (X_1, \dots, X_n)$ i.i.d. de loi η . Pour les modèles suivants, dont on admettra qu'ils sont réguliers, déterminer l'estimateur du maximum de vraisemblance, puis énoncer le théorème central limite associé.

1. η est la loi exponentielle de paramètre¹ $\lambda > 0$, et l'on veut estimer λ .
2. η est la loi de puissance à densité sur \mathbb{R} donnée par $x \mapsto \mathbf{1}_{x \geq 1} \frac{a-1}{x^a}$ avec $a > 1$ paramètre et l'on veut estimer a .

Exercice 2.4 (Information de Fisher via reparamétrisation). Soit $\{\mathbb{P}_\theta\}_{\theta \in \Theta}$ un modèle régulier unidimensionnel. Considérons une fonction différentiable bijective $h : \Lambda \rightarrow \Theta$. La famille $\{\mathbb{P}_\theta\}_{\theta \in \Theta}$ peut ainsi être reparamétrisée par $\{\mathbb{Q}_\lambda\}_{\lambda \in \Lambda}$. On note $\tilde{I}(\lambda)$ l'information de Fisher dans ce modèle reparamétrisé. Pour λ tel que $h(\lambda) = \theta$, donner $\tilde{I}(\lambda)$ en fonction de $I(\theta)$. Donner une interprétation physique de ce résultat.

¹sa densité sur \mathbb{R}_+ est donnée par $x \mapsto \lambda e^{-\lambda x}$.

Exercice 2.5 (Un modèle autorégressif). Nous considérons un modèle avec (X_1, \dots, X_n) tels que $X_1 = \theta + Z_1$, et récursivement, $X_{j+1} = \rho(X_j - \theta) + \theta + Z_{j+1}$ avec les Z_j i.i.d. de loi $\mathcal{N}(0, \sigma^2)$. Nous supposons que $\rho \in [0, 1]$ et qu'il est connu. On admettra la régularité de ce modèle.

1. Calculer la matrice d'information de Fisher $I(\theta, \sigma^2)$ en fonction de θ, σ^2, ρ et n .
2. Donner une borne inférieure pour la variance d'un estimateur sans biais de θ . Commenter et interpréter les valeurs de cette borne lorsque $\rho = 0$, et lorsque $\rho = 1$.