Feuille de travaux dirigés 2 : Estimation ponctuelle

Exercice 1 (Moindres carrés dans le modèle linéaire simple)**:**
On reprend l'exemple du dernier exercice de la feuille de TD précédente. Pour tout $i = 1, \ldots, n$, on considère

$$X_i = a_i \theta_1 + Z_i \tag{1}$$

où les Z_i sont des variables aléatoires indépendantes, de même loi $\mathcal{N}(0, \sigma^2)$. Les coefficients a_i sont des variables déterministes connues (et non toutes nulles). Les paramètres θ_1 et σ^2 sont inconnus. On observe $X = (X_1, \ldots, X_n)$. On paramètre le modèle pour X par $\theta = (\theta_1, \sigma^2) \in \mathbb{R} \times \mathbb{R}^+$.

1. Calculez l'estimateur de θ_1 donné par la méthode des moindres carrés.
2. Comparez à l'estimateur du maximum de vraisemblance (voir le TD précédent).
3. On change la loi des bruits Z_i par une autre loi connue ; comment sont modifiés ces deux estimateurs ?

Exercice 2 (Loi Gamma)**:**
Un opérateur mobile s'intéresse aux durées d'appel de ses clients. On observe les durées $X_1, \ldots, X_n$ de n appels, supposées indépendantes et modélisées chacune par une loi Gamma $\mathrm{P}_\theta = \gamma(\alpha, \lambda)$, de paramètre $\theta = (\alpha, \lambda)$ avec $\alpha > 0$ et $\lambda > 0$ dont la densité est

$$\frac{\lambda^\alpha}{\Gamma(\alpha)} x^{\alpha-1} \exp(-\lambda x) \mathbb{1}_{\mathbb{R}_+}(x) \ ,$$

où $\Gamma(\alpha)$ est la fonction Gamma. On s'intéresse au paramètre $\theta = (\alpha, \lambda) \in \Theta = (0, \infty)^2$.

1. Calculez l'estimateur de θ par la méthode des moments.
2. Donnez l'équation que doit vérifier l'estimateur de θ par la méthode du maximum de vraisemblance. Commentez.

Exercice 3 (Maximum de vraisemblance, modèle de Pareto)**:**
Pour la modélisation du trafic internet, on utilise couramment la loi de Pareto pour représenter la distribution de la taille des paquets envoyés, parmi les paquets dont la taille dépasse un certain seuil u suffisamment élevé. Par définition, une variable aléatoire X_1 suit une loi de Pareto $\mathcal{P}ar(u, \theta)$, avec $u > 0$ et $\theta > 0$, si la fonction de répartition de X_1 est

$$F_\theta(x) = \begin{cases} 0 & \text{si } x \leq u \\ 1 - \left(\frac{u}{x}\right)^\theta & \text{si } x > u. \end{cases}$$

Dans toute la suite de l'exercice on fixe $u > 0$, supposé connu et on considère le modèle statistique

$$\mathcal{P} = \{P_\theta, \theta \in \Theta\}, \qquad \text{avec } \Theta =]0, \infty[,$$

où P_θ désigne la loi de Pareto $\mathcal{P}ar(u, \theta)$. On considère alors n observations indépendantes $X = (X_1, \ldots, X_n)$, où $X_i \overset{i.i.d.}{\sim} \mathrm{P}_\theta$ $(1 \leq i \leq n)$ avec $n \geq 3$.

1. Montrer que la loi $\mathcal{P}ar(u, \theta)$ admet une densité par rapport à la mesure de Lebesgue que l'on calculera.
2. Exprimer la log-vraisemblance $\log p_\theta^{\otimes n}(\mathbf{x})$, pour $\mathbf{x} = (x_1, \ldots, x_n) \in U \subset \mathbb{R}^n$ où U est le domaine (que l'on précisera), où cette quantité est finie.
3. Calculez l'estimateur du maximum de vraisemblance pour le paramètre θ.