Corrigé : Feuille de travaux dirigés 2

Solution Exercice 1

1. Le modèle statistique est bien de la forme $X_i = \phi(\theta, a_i) + Z_i$ donc on peut appliquer la méthode des moindres carrés. Ici, les variables aléatoires Z_i sont interprétées comme du bruit et $\phi(\theta, a_i)$ est interprété comme le signal. Il s'agit de minimiser le carré de la norme de la différence entre les observations X_i et le signal $\phi(\theta, a_i)$. La variable d'optimisation est le paramètre d'intérêt θ_1. On obtient l'estimateur

$$\widehat{\theta}_1 = \arg\min_{t\in\mathbb{R}} \sum_{i=1}^{n} \Big(X_i - a_i t\Big)^2 .$$

Notez que c'est bien une fonction de X et de a uniquement.
Posons $f(t) = \sum_{i=1}^{n} \Big(X_i - a_i t\Big)^2$. Sa dérivée vaut

$$f'(t) = \sum_{i=1}^{n} a_i(a_i t - X_i).$$

Comme $\widehat{\theta}_1(X)$ annule f', $\widehat{\theta}_1$ est solution de l'équation

$$\sum_{i=1}^{n} a_i(a_i t - X_i) = 0$$

c'est à dire $\widehat{\theta}_1 = \frac{\sum_{i=1}^{n} a_i X_i}{\sum_{j=1}^{n} a_i^2} = \frac{a^T X}{\|a\|^2}$.
Comme $f''(\widehat{\theta}) = \|a\|^2 > 0$, $\widehat{\theta}$ est bien un minimum local de f. Comme c'est le seul minimum local, c'est le minimum global.

2. On remarque qu'on retrouve le même estimateur qu'avec la méthode du maximum de vraisemblance.

3. Dans le calcul de la question 1, on n'a utilisé nulle part la loi de Z_i. On obtiendra donc le même résultat pour l'estimateur des moindres carrés si la loi de Z_i change. Par contre, l'estimateur du maximum de vraisemblance sera différent.

Par exemple, si $Z_i \overset{\text{i.i.d.}}{\sim} U([-\sqrt{3}\sigma, +\sqrt{3}\sigma])$, on a

$$p_{\boldsymbol{\theta}}(x) = \prod_{i=1}^{n} \frac{1}{2\sqrt{3}\sigma} \mathbb{1}_{[a_i\theta - \frac{1}{\sqrt{3}\sigma}, a_i\theta + \frac{1}{\sqrt{3}\sigma}]}(x)$$

L'estimateur du maximum de vraisemblance devient donc

$$\widehat{\theta}_U \in \arg\min_{\theta} \min_{\sigma>0} \{\log(2\sqrt{3}\sigma) \text{ sous la contrainte } a_i\theta - \sqrt{3}\sigma \le x_i \le a_i\theta - \sqrt{3}\sigma, \forall i\}.$$

Pour chaque θ, le problème interne a pour solution $\sigma = \frac{1}{\sqrt{3}} \max_i |x_i - a_i\theta|$.

Avec des bruits qui suivent une loi uniforme, on obtient

$$\widehat{\theta}_U \in \arg\min_{\theta} \max_{i} |x_i - a_i\theta|.$$

Autrement dit, avec la méthode du maximum de vraisemblance $\widehat{\theta}_U$ minimise la norme infinie des écarts alors que l'estimateur des moindres carrés minimise la norme 2.

Solution Exercice 2 1. On veut estimer un paramètre de dimension 2 donc il semble raisonnable d'utiliser les 2 premiers moments.

$$\mathbb{E}_\theta(X_1) = \int_{\mathbb{R}_+} \frac{\lambda^\alpha}{\Gamma(\alpha)} x^\alpha \exp(-\lambda x)dx = \int_{\mathbb{R}_+} \frac{\lambda^\alpha}{\Gamma(\alpha)} \left(\frac{y}{\lambda}\right)^\alpha \exp(-y)\frac{1}{\lambda}dy = \frac{1}{\lambda}\frac{\Gamma(\alpha+1)}{\Gamma(\alpha)} = \frac{\alpha}{\lambda}$$

$$\mathbb{E}_\theta(X_1^2) = \int_{\mathbb{R}_+} \frac{\lambda^\alpha}{\Gamma(\alpha)} x^{\alpha+1} \exp(-\lambda x)dx = \int_{\mathbb{R}_+} \frac{\lambda^\alpha}{\Gamma(\alpha)} \left(\frac{y}{\lambda}\right)^{\alpha+1} \exp(-y)\frac{1}{\lambda}dy = \frac{1}{\lambda^2}\frac{\Gamma(\alpha+2)}{\Gamma(\alpha)}$$
$$= \frac{\alpha(\alpha+1)}{\lambda^2}$$

$$\mathrm{var}_\theta(X_1) = \mathbb{E}_\theta(X_1^2) - \mathbb{E}_\theta(X_1)^2 \frac{\alpha}{\lambda^2}$$

Ainsi, $\Phi(\theta) = \Phi(\alpha, \lambda) = (\frac{\alpha}{\lambda}, \frac{\alpha(\alpha+1)}{\lambda^2})$.
Φ est inversible : on peut retrouver λ et α par $\lambda = \frac{\mathbb{E}_\theta(X_1)}{\mathrm{var}_\theta(X_1)}$ et $\alpha = \lambda\mathbb{E}_\theta(X_1) = \frac{\mathbb{E}_\theta(X_1)^2}{\mathrm{var}_\theta(X_1)}$.
Par ailleurs, la version empirique de Φ est la statistique

$$\widehat{\Phi}(X) = (\frac{1}{n}\sum_{1}^{n} X_i, \frac{1}{n}\sum_{1}^{n} X_i^2).$$

On a donc,

$$\widehat{\theta}_M = \Phi^{-1}\left(\widehat{\Phi}(X)\right) = \left(\frac{(\frac{1}{n}\sum_{i=1}^n X_i)^2}{\frac{1}{n}\sum_{i=1}^n X_i^2 - (\frac{1}{n}\sum_{i=1}^n X_i)^2}, \frac{\frac{1}{n}\sum_{i=1}^n X_i}{\frac{1}{n}\sum_{i=1}^n X_i^2 - (\frac{1}{n}\sum_{i=1}^n X_i)^2}\right)$$

2. Comme les observations sont i.i.d., la vraisemblance est

$$p_\theta(x) = \prod_{i=1}^n \frac{\lambda^\alpha}{\Gamma(\alpha)} x_i^{\alpha-1} \exp(-\lambda x_i)\mathbb{1}_{\mathbb{R}_+}(x_i).$$

On doit donc résoudre le problème

$$\widehat{\theta}_V = \arg\min_{\theta=(\alpha,\lambda)} \sum_{i=1}^n -\alpha\log(\lambda) + \log(\Gamma(\alpha)) - (\alpha-1)\log(x_i) + \lambda x_i.$$

Notons ψ_0 la fonction digamma, qui vérifie $\psi_0(\alpha) = \frac{\Gamma'(\alpha)}{\Gamma(\alpha)}$. $\widehat{\alpha}_V$ et $\widehat{\lambda}_V$ sont donc solutions du système

$$\begin{cases} -n\log(\lambda) + n\psi_0(\alpha) - \sum_{i=1}^n \log(x_i) = 0 \\ -n\frac{\alpha}{\lambda} + \sum_{i=1}^n x_i = 0 \end{cases}$$

On peut éliminer une des deux variables grâce à la deuxième équation, λ par exemple :

$$\begin{cases} \lambda = \frac{n\alpha}{\sum_{i=1}^n x_i} \\ n\log\left(\frac{1}{n}\sum_{i=1}^n x_i\right) - n\log(\alpha) + n\psi_0(\alpha) - \sum_{i=1}^n \log(x_i) = 0 \end{cases}$$

L'équation qui reste doit être résolue numériquement, par exemple grâce à la méthode de Newton.

Solution Exercice 3 1. La fonction de répartition est continue et de classe C_1 par morceaux sur $\mathbb{R}$ donc la densité est obtenue en la dérivant, d'où le résultat :

$$p_\theta(x) = \mathbb{1}_{\{x>u\}}\, \theta \frac{u^\theta}{x^{\theta+1}}$$

2. Pour une observation, on a $\log p_\theta(x) = \log\theta + \theta\log u - (\theta+1)\log x$, d'où

$$\log p_\theta^{\otimes n}(\mathbf{x}) = n(\log\theta + \theta\log u) - (\theta+1)\sum_{i=1}^n \log x_i$$

Cette quantité est finie si et seulement si $x_i > 0$ pour $i \in \{1, \dots n\}$.

3. On obtient le MV est annulant la dérivée partielle en θ de la log vraisemblance (qui est bien différentiable et strictement concave sur $]0, +\infty[$, en tant que fonction de θ) :

$$\frac{\partial}{\partial\theta}(\log p_\theta^{\otimes n}(\mathbf{x})) = n\left(\frac{1}{\theta} + \log u\right) - \sum \log X_i = \frac{n}{\theta} - \sum \log(X_i/u),$$

d'où :

$$\widehat{\theta}_{MV}(X) = \frac{n}{\sum_{i=1}^n \log\left(\frac{X_i}{u}\right)}.$$