Corrigé : Feuille de travaux dirigés 2

Solution Exercice 2

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

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

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 3

1. Soit X la durée de vie d'une cellule choisie au hasard (uniforme) dans la population à l'étude et soit Y la variable binaire décrivant son état ($Y = 0$ si la cellule est pathogène, $Y = 1$ sinon). Par hypothèse, Y suit une loi de Bernoulli de paramètre θ. La durée de vie Z_s d'une celleule saine suit, par hypothèse, une loi exponentielle $\mathcal{E}(1)$ et une cellule pathogène Z_p vérifie $\mathbb{P}(Z_p > x) = \mathbb{P}(Z_s/2 > x) = e^{-2x}$, de sorte que $Z_p \sim \exp(2)$. Finalement, d'après les hypothèses de l'énoncé, on peut écrire :

$$X_1 \overset{loi}{=} YZ_s + (1-Y)Z_p.$$

Ainsi la fonction de répartition de X_1 s'écrit

$$\begin{aligned}\forall x \in \mathbb{R}, \quad F_\theta(x) &= \mathbb{P}(X_1 \le x|U=1)\mathbb{P}(U=1) + \mathbb{P}(X_1 \le x|U=0)\mathbb{P}(U=0)\\ &= \theta\mathbb{P}(Z_s \le x) + (1-\theta)\mathbb{P}(Z_p \le x)\\ &= \mathbb{1}_{x>0}\left[\theta(1-e^{-x}) + (1-\theta)e^{-2x}\right]\end{aligned}$$

Cette fonction est continue sur $\mathbb{R}$ et de classe $\mathcal{C}^1$ par morceaux (partout sauf en 0). Ainsi, X_1 admet une densité par rapport à Lebesgue donnée par

$$\forall x \neq 0, f_\theta(x) = F'_\theta(x) = \mathbb{1}_{x>0}\theta\, e^{-x} + 2(1-\theta)e^{-2x}$$

La valeur de f_θ en 0 peut être choisie arbitrairement puisque $\{0\}$ est de mesure de Lebesgue nulle.

2. La densité permet de calculer $\mathbb{E}(X^p)$:

$$\begin{aligned}\mathbb{E}_\theta(X^p) &= \int_{\mathbb{R}} x^p f_\theta(x)\mathrm{d}x\\ &= \theta\int_{\mathbb{R}^+} x^p e^{-x}\mathrm{d}x + 2(1-\theta)\int_{\mathbb{R}^+} x^p e^{-2x}.\end{aligned}$$

On reconnait dans la première intégrale la quantité $\Gamma(p+1) = p!$. Par changement de variable $y = 2x$, la deuxième intégrale vaut $2^{-p-1}p!$, d'où

$$\mathbb{E}_\theta(X^p) = p!\left(\theta + 2^{-p}(1-\theta)\right).$$

Posons $\Phi(\theta) = \mathbb{E}_\theta(X^p)$, dont l'expression est donnée ci-dessus et soit $\widehat{\Phi}_n$ la version empirique du moment d'ordre p, $\widehat{\Phi}_n = \frac{1}{n}\sum_1^n X_i^p$.
La méthode des moments consiste alors à définir $\widehat{\theta}(X)$ comme le M-estimateur

$$\widehat{\theta}_n(X) = \underset{t\in\Theta}{\operatorname{argmin}} \|\Phi(t) - \widehat{\Phi}_n\|^2.$$

La fonction Φ est inversible de $[0,1]$ dans son image $Im(\Phi) = \left[\frac{p!}{2^p}, p!\right]$ d'inverse pour $p \geq 1$:

$$\forall y \in \ Im(\Phi), \quad \Phi^{-1}(y) = \frac{\frac{2^p}{p!}y - 1}{2^p - 1}.$$

Ainsi, l'estimateur par la méthode des moments est

$$\widehat{\theta}(X) = \begin{cases} 1 & \text{si } \widehat{\Phi}_n \geq p! \\ 0 & \text{si } \widehat{\Phi}_n \leq \frac{p!}{2^p} \\ \Phi^{-1}(\widehat{\Phi}_n) = \frac{\frac{2^p}{p!}\widehat{\Phi}_n - 1}{2^p - 1} & \text{sinon .} \end{cases}$$

Solution Exercice 4 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.