Ecole Normale Supérieure
FIMFA – Statistique 2012-2013

Correction des exercices 4 ,5 ,9 Feuille 2.

Exercice 4. 1. (a) La vraisemblance du modèle est donnée pour tous les vecteurs $z = (z_1, ..., z_n), c = (c_1, ..., c_n)$, avec $z_i, c_i > 0, \forall i \in [[1, n]]$, et tous $\lambda > 0$, $\mu > 0$ par

$$V_{(z,c)}(\lambda, \mu) = V_z(\lambda) V_c(\mu), \text{ où } V_z(\lambda) = \lambda^n \exp\left(-\lambda \sum_{i=1}^n z_i\right), \; V_c(\mu) = \mu^n \exp\left(-\mu \sum_{i=1}^n c_i\right).$$

On a vu dans l'exercice 3 que $V_z(\lambda)$ (respectivement $V_c(\mu)$) est maximal pour $\lambda = 1/\bar{z}_n$ (resp. $\mu = 1/\bar{c}_n$). On en déduit que l'estimateur du maximum de vraisemblance du paramètre (λ, μ) est

$$\left(\hat{\lambda}_n, \hat{\mu}_n\right) = \left(\frac{1}{\bar{Z}_n}, \frac{1}{\bar{C}_n}\right).$$

(b) Comme on l'a vu dans l'exercice 3, en utilisant le Théorème Central Limite et la méthode delta on montre que

$$\sqrt{n}\left(\hat{\lambda}_n - \lambda\right) \overset{\text{loi}}{\to} \mathcal{N}(0, \lambda^2), \quad \sqrt{n}\left(\hat{\mu}_n - \mu\right) \overset{\text{loi}}{\to} \mathcal{N}(0, \mu^2).$$

Les variables $\hat{\lambda}_n$ et $\hat{\mu}_n$ étant indépendantes, on en déduit la convergence jointe,

$$\sqrt{n}\left(\left(\hat{\lambda}_n - \lambda\right), (\hat{\mu}_n - \mu)\right) \overset{\text{loi}}{\to} (N, M),$$

où N et M sont deux lois normales centrées indépendantes, de variances respectives λ^2 et μ^2.

(c) A n fixé, $\sum_{i=1}^n Z_i$ suit une loi Gamma(n, λ). On en déduit que $\hat{\lambda}_n$ admet pour densité la fonction

$$x \in \mathbb{R}_+^* \mapsto \frac{(n\lambda)^n}{(n-1)! x^{n+1}} \exp(-\lambda n / x).$$

De même pour $\hat{\mu}_n$ en remplaçant le paramètre λ par μ. La densité du couple $(\hat{\lambda}_n, \hat{\mu}_n)$ est alors le produit des densités puisque ces variables sont indépendantes. La matrice des variances-covariances est alors une matrice 2×2 diagonale, dont le terme en haut à gauche est $\mathrm{Var}(\hat{\lambda}_n) = \frac{n^2\lambda^2}{(n-1)^2(n-2)}$ et celui en bas à droite $\mathrm{Var}(\hat{\mu}_n) = \frac{n^2\mu^2}{(n-1)^2(n-2)}$.

2. (a) Pour $x \geq 0$, $\mathbb{P}(X_1 > x) = \mathbb{P}(Z_1 > x)\mathbb{P}(C_1 > x) = \exp(-(\lambda + \mu)x)$. Donc X_1 suit une loi exponentielle de paramètre $\lambda + \mu$. En observant les $X_i, 1 \leq i \leq n$, on ne va pouvoir avoir des informations que sur $\lambda + \mu$. Le modèle identifiable correspondant est $\left(\mathbb{R}_+^n, (\mathcal{E}(\gamma))_{\gamma > 0}^{\otimes n}\right)$, où le paramètre γ est relié à (λ, μ) par $\gamma = \lambda + \mu$.

(b) i) Comme précédemment, l'EMV $\hat{\gamma}_n = 1/\bar{X}_n$.

ii) L'estimateur de γ construit à partir des EMV de λ et μ est lui égal à $\frac{1}{Z_n} + \frac{1}{C_n}$. On le note $\tilde{\gamma}_n$. Ce n'est pas surprenant que ces deux estimateurs soient différents puisqu'on n'utilise pas les mêmes informations pour les construire.

(c) Plus précisément, on a moins d'informations pour construire $\hat{\gamma}_n$ que $\tilde{\gamma}_n$. On s'attend donc à ce que le premier estimateur soit moins bon que le deuxième. Ceci peut se formaliser ainsi : d'une part,

$$\sqrt{n}\,(\hat{\gamma}_n - (\lambda + \mu)) \overset{\text{loi}}{\to} \mathcal{N}(0, (\lambda + \mu)^2)$$

et d'autre part, comme conséquence de 1) (b),

$$\sqrt{n}\,(\tilde{\gamma}_n - (\lambda + \mu)) \overset{\text{loi}}{\to} \mathcal{N}(0, \lambda^2 + \mu^2).$$

La variance de la loi normale limite est plus petite pour l'estimateur $\tilde{\gamma}_n$. Celui-ci converge donc plus vite vers $\lambda + \mu$ que $\hat{\gamma}_n$.

Exercice 5. On vérifie aisément que

$$\mathbb{E}_\theta[|X_1|] = \frac{\theta}{\sqrt{\pi/2}}, \quad \mathrm{Var}_\theta(|X_1|) = \theta^2\left(1 - \frac{2}{\pi}\right), \quad \mathbb{E}_\theta[X_1^2] = \theta^2 \text{ et } \mathrm{Var}_\theta(X_1^2) = 2\theta^4.$$

1. De ceci, on déduit d'une part que

   $$\mathbb{E}_\theta[\hat{S}] = \theta,$$

   $\hat{S}$ est donc un estimateur non biaisé. D'autre part, $\mathbb{E}_\theta[n^{-1}\sum_{i=1}^n X_i^2] = \theta^2$. Or, d'après l'inégalité de Cauchy-Schwarz (ou Jensen), si Y est une variable aléatoire réelle positive, telle qu'il n'existe pas de réel $a \geq 0$ t.q. $Y = a$ p.s., alors $\mathbb{E}[\sqrt{Y}] < \sqrt{\mathbb{E}[Y]}$. Donc

   $$\mathbb{E}_\theta[\hat{T}] < \theta,$$

   et $\hat{T}$ est un estimateur biaisé.
   Par ailleurs, les deux estimateurs $\hat{S}$ et $\hat{T}$ sont fortement consistants, c'est une conséquence immédiate de la Loi Forte des Grands Nombres appliquée à l'échantillon $(|X_1|, |X_2|, ..., |X_n|)$ pour $\hat{S}$ et à $(X_1^2, ..., X_n^2)$ pour $\hat{T}$.

2. En appliquant le Théorème Central Limite aux deux échantillons précédemment cités, on obtient :

   $$\sqrt{n}\left(\hat{S} - \theta\right) \tag{1}$$

   et

   $$\sqrt{n}\left(\frac{\sum_{i=1}^n X_i^2}{n} - \theta^2\right) \overset{\text{loi}}{\to} \mathcal{N}\left(0, 2\theta^4\right).$$

   Par la méthode delta, cette 2ème convergence en loi implique que

   $$\sqrt{n}\left(\hat{T} - \theta\right) \overset{\text{loi}}{\to} \frac{1}{2\theta}\mathcal{N}\left(0, 2\theta^4\right) \overset{\text{loi}}{=} \mathcal{N}\left(0, \frac{\theta^2}{2}\right). \tag{2}$$

Comme $\pi/2 - 1 > 0,5$, l'estimateur $\hat{T}$ converge plus vite vers θ que $\hat{S}$. En fait, en divisant par θ les membres de droite et gauche des convergences (1) et (2), on peut construire des intervalles de confiance pour θ. Celui donné par (2) sera plus étroit, donc plus précis, que celui obtenu à partir de (1).

Exercice 9 (correction succincte). 1. Le calcul est classique, on obtient

$$\hat{\theta}_n = \exp\left(\frac{2n}{\sum_{i=1}^n X_i^2}\right).$$

2. (a) Là aussi, c'est classique : $T_n/(2n)$ converge presque sûrement vers $1/\log(\theta)$ d'après la Loi Forte des Grands Nombres, et

$$\sqrt{n}\left(\frac{T_n}{2n} - \frac{1}{\log(\theta)}\right) \overset{\text{loi}}{\to} \mathcal{N}\left(0, \frac{1}{(\log\theta)^2}\right)$$

d'après le Théorème Central Limite.
(b) On en déduit que $\hat{\theta}_n$ converge presque sûrement vers θ, puis, en utilisant la méthode delta et la convergence en loi ci-dessus,

$$\sqrt{n}\left(\hat{\theta}_n - \theta\right) \overset{\text{loi}}{\to} \mathcal{N}\left(0, (\theta\log(\theta))^2\right).$$

3. On a (pour $n = 1$)

$$\partial_\theta \log V_X(\theta) = \frac{1}{\theta\log\theta} - \frac{X^2}{2\theta}.$$

On en déduit que

$$I_1(\theta) = \mathbb{E}_\theta\left[(\partial_\theta \log V_X(\theta))^2\right] = \operatorname{Var}\left(\frac{X^2}{2\theta}\right) = \frac{1}{4\theta^2}\left(\mathbb{E}[X_1^4] - \mathbb{E}[X_1^2]\right) = \frac{1}{(\theta\log\theta)^2}.$$

Donc on a bien

$$\sqrt{n}\left(\hat{\theta}_n - \theta\right) \overset{\text{loi}}{\to} \mathcal{N}\left(0, \frac{1}{I_1(\theta)}\right),$$

autrement dit, $\hat{\theta}_n$ est asymptotiquement efficace (remarque : on vérifie aisément que cet estimateur est biaisé, il n'y a donc pas de raison qu'il vérifie l'inégalité de l'information).
4. $\mathbb{E}_\theta[X] = \sqrt{\pi/(2\log\theta)}$, donc $\theta = \exp(\pi/(2\mathbb{E}_\theta[X])$ et par conséquent

$$\tilde{\theta}_n = \exp\left(\frac{\pi}{2\bar{X}_n}\right)$$

est un estimateur de θ fondé sur la méthode des moments.