Ecole Normale Supérieure
FIMFA – Statistique 2012-2013

Feuille d'exercices n°2.

Exercice 1. On rappelle que la loi Gamma $\mathcal{G}(\theta,\lambda)$, $\theta,\lambda>0$, est une v.a. réelle de densité

$$x \mapsto \frac{\lambda^\theta}{\Gamma(\theta)} x^{\theta-1}\exp(-\lambda x), \quad x>0.$$

Soit $(\mathbb{R}^n, (\mathcal{G}(\theta,\lambda))^{\otimes n}_{\theta>0,\lambda>0})$ un modèle statistique.

1. On suppose le paramètre λ connu. Proposer un estimateur de θ par la méthode des moments.
2. On suppose à présent que les deux paramètres θ,λ sont inconnus. Trouver un estimateur de (θ,λ) par la méthode des moments.

Exercice 2. Soit $(\mathbb{R}^n, (\mathbb{P}_\theta^{\otimes n})_{\theta\in\Theta})$ un modèle statistique. Dans chacun des cas suivants, calculer l'estimateur du maximum de vraisemblance. Est-il sans biais ? Consistant ? Dans l'affirmative, préciser sa vitesse de convergence.

1. Pour $\theta\in\Theta=]-1,1[$, $\mathbb{P}_\theta$ a une densité sur $\mathbb{R}$ donnée par

$$\begin{cases} 1-\theta \text{ si } x\in]-1/2,0] \\ 1+\theta \text{ si } x\in]0,1/2] \\ 0 \text{ sinon.}\end{cases}$$

2. Pour $\theta\in\Theta=]0,1[$, $\mathbb{P}_\theta$ prend les valeurs $0,1,-1$ avec les probabilités respectives $1-\theta$, $\theta/2$, $\theta/2$.
3. Pour $\theta=(a,b)\in\Theta=\{(x,y):y>x\}$, $\mathbb{P}_\theta=\mathcal{U}[a,b]$.

Exercice 3. Soit $(\mathbb{R}_+, (\mathcal{E}(\theta))_{\theta>0})$ un modèle statistique ($\mathcal{E}(\theta)$ désigne la loi exponentielle de paramètre θ).

1. Montrer qu'il n'existe pas d'estimateur sans biais de θ.

   On considère dans ce qui suit un échantillon $(X_1,...,X_n)$ de loi $\mathcal{E}(\theta)$, $\theta>0$.

2. Un statisticien propose d'estimer θ par

$$\hat\theta = \frac{1}{\overline{X}_n}.$$

   Expliquer ce choix. Cet estimateur est-il (fortement) consistant ? Asymptotiquement normal ?

3. Rappeler la loi de $\sum_{i=1}^n X_i$ et en déduire $\mathbb{E}_\theta[\hat{\theta}]$, $\mathrm{Var}_\theta(\hat{\theta})$ et $\mathbb{E}_\theta[(\hat{\theta}-\theta)^2]$.
4. Un deuxième statisticien propose l'estimateur

$$\theta^* = \frac{n-1}{n}\hat{\theta}.$$

Calculer $\mathbb{E}_\theta[\theta^*]$, $\mathrm{Var}_\theta(\theta^*)$ et $\mathbb{E}_\theta[(\theta^*-\theta)^2]$. Lequel de ces deux estimateurs préférez-vous ?

Exercice 4. On considère $(Z_1, ..., Z_n)$ un échantillon de loi exponentielle $\mathcal{E}(\lambda)$ et $(C_1, ..., C_n)$ un échantillon de loi exponentielle $\mathcal{E}(\mu)$, avec $\lambda, \mu > 0$. Ces deux échantillons sont indépendants.

1. Observation parfaite.
   (a) Déterminer l'estimateur du maximum de vraisemblance du paramètre (λ, μ).
   (b) Quelle est sa loi asymptotique ?
   (c) Déterminer sa loi. Est-il biaisé ? Calculer sa matrice de variance-covariance.
2. Observation imparfaite. On suppose dans cette partie que les seules observations disponibles portent sur $X_i = \min(Z_i, C_i)$, $i \in [[1, n]]$.
   (a) Rappeler la loi de X_1. En déduire le modèle statistique identifiable correspondant.
   (b) Trouver les estimateurs du maximum de vraisemblance du paramètre $\gamma = \lambda+\mu$ fondés :
      i. sur les X_i, $i \in [[1, n]]$;
      ii. sur les (Z_i, C_i), $i \in [[1, n]]$.

      Est-il naturel qu'ils soient différents ?
   (c) Comparer les propriétés asymptotiques de ces deux estimateurs.

Exercice 5. Pour un échantillon $(X_1, ..., X_n)$ de la loi $\mathcal{N}(0, \theta^2)$, $\theta > 0$, on considère les estimateurs suivants de θ :

$$\hat{S} = \left(\frac{\pi}{2}\right)^{1/2} \frac{1}{n}\sum_{i=1}^n |X_i| \quad \text{et} \quad \hat{T} = \sqrt{\frac{1}{n}\sum_{i=1}^n X_i^2}.$$

On rappelle que les moments d'ordre 2 et 4 de la loi $\mathcal{N}(0, \theta^2)$ valent respectivement θ^2 et $3\theta^4$.

1. Ces estimateurs sont-ils biaisés ? Consistants ?
2. Trouver une suite $(x_n)_{n\geq 1}$ et une probabilité non-dégénérée μ sur $\mathbb{R}$ telles que

$$x_n(\hat{S}-\theta) \underset{n\to\infty}{\overset{\text{loi}}{\to}} \mu.$$

Même question pour l'estimateur $\hat{T}$. Sur la base de ces résultats, lequel de ces estimateurs est le plus performant ?

Exercice 6. Pour $\theta \in \Theta =]0,1[$, soit $\mathbb{P}_\theta$ la probabilité définie sur $\mathbb{N} \cup \{-1\}$ par

$$\mathbb{P}_\theta(\{-1\}) = \theta, \quad \text{et} \quad \mathbb{P}_\theta(\{k\}) = (1-\theta)^2\theta^k, \quad \forall k \in \mathbb{N}.$$

On cherche un estimateur UMVU de $g(\theta) = (1-\theta)^2, \theta \in \Theta$.

1. Soit X une observation. Vérifier que $\hat{g} = \mathbf{1}_{\{X=0\}}$ est un estimateur sans biais de $g(\theta)$.
2. Soit g^* un estimateur sans biais de $g(\theta)$. Montrer qu'il existe $a \in \mathbb{R}$ tel que
$$g^* = \hat{g} - aX.$$
3. Conclure.

Exercice 7. Pour $\theta \in]0,1[$, on introduit la densité f_θ définie sur $\mathbb{R}$ par

$$f_\theta(x) = \theta^2 \exp(-\theta x)\mathbf{1}_{\{x\geq 0\}} + \theta(1-\theta)\exp(\theta x)\mathbf{1}_{\{x<0\}}$$

et on note $\mathbb{P}_\theta$ la loi de densité f_θ.

1. Calculer l'information de Fisher I du modèle $(\mathbb{R}, (\mathbb{P}_\theta)_{\theta\in]0,1[})$.
2. On note $\mathbb{P}'_\theta$ la loi suivie par la valeur absolue d'une variable de loi $\mathbb{P}_\theta$. Calculer l'information de Fisher I' du modèle $(\mathbb{R}, (\mathbb{P}'_\theta)_{\theta\in]0,1[})$.
3. Comparer I' et I et interpréter cette observation.

Exercice 8. Soit $(\{0,1\}^n, (\mathcal{B}(\theta)^{\otimes n})_{\theta\in]0,1[})$. Calculer l'estimateur du maximum de vraisemblance de θ et montrer que c'est un estimateur uniformément efficace.

Exercice 9. Soient $X_1, \ldots, X_n$ des variables aléatoires réelles i.i.d. de densité

$$f_\theta(x) = x\theta^{-x^2/2}\log\theta\,\mathbf{1}_{\{x\geq 0\}},$$

avec $\theta > 1$ (log désigne le logarithme népérien). On admettra que

$$\mathbb{E}_\theta[X_i^2] = \frac{2}{\log\theta} \quad \text{et} \quad \mathbb{E}_\theta[X_i^4] = \frac{8}{(\log\theta)^2}.$$

On note

$$T_n = \sum_{i=1}^n X_i^2.$$

1. Calculer l'estimateur du maximum de vraisemblance $\hat{\theta}_n$ de θ.
2. 2.a Montrer la convergence et la normalité asymptotique de la suite de variables aléatoires $T_n/(2n)$ (on précisera bien la limite et la loi limite).

   2.b En déduire la convergence et la normalité asymptotique de l'estimateur $\hat{\theta}_n$.
3. Calculer l'information de Fisher du modèle (à une observation) et vérifier ainsi que $\hat{\theta}_n$ est asymptotiquement efficace.
4. Soit X une variable aléatoire de densité $f_\theta(x)$. Calculer $\mathbb{E}_\theta[X]$ et en déduire un estimateur $\tilde{\theta}_n$ de θ fondé sur la méthode des moments.

Exercice 10. Soient $X_1, \dots, X_n$ des variables aléatoires i.i.d. de loi $\mathcal{N}(\mu, \sigma^2)$ où $\mu \in \mathbb{R}$ et $\sigma^2 > 0$ sont des paramètres inconnus.

1. Calculer l'information de Fisher associée à (μ, σ^2).
2. Calculer l'estimateur du maximum de vraisemblance $(\hat{\mu}, \hat{\sigma}^2)$ de (μ, σ^2). Montrer que $\hat{\sigma}^2$ est biaisé. Quelle est sa loi ? Vérifie-t-il l'inégalité de l'information ?
3. Proposer un estimateur sans biais de σ^2. Montrer qu'il a un plus grand risque quadratique que l'estimateur du maximum de vraisemblance. Comparer sa variance à la borne de l'information.