Partiel de Statistiques

Allan Merino

Master 1, UE Statistiques, Février 2016

Questions de cours

1. Soit le modèle statistique $\mathscr{M} = \{(\mathbb{R}^n, \mathscr{B}(\mathbb{R})^{\otimes n}, \mathscr{N}(\mu, \sigma^2)^{\otimes n}), (\mu, \sigma^2) \in \mathbb{R} \times \mathbb{R}_+^*\}$. Déterminer l'estimateur du maximum de vraisemblance du paramètre (μ, σ^2). Cet estimateur est-il sans biais ? Asymptotiquement sans biais ?
2. Soit $X_1, \ldots, X_n$ un n-échantillon de loi $\mathbb{P}_\theta$. Trouver un intervalle de confiance asymptotique de niveau $1 - \alpha$ pour le cas $\mathbb{P}_\theta = \mathscr{P}(\theta)$, $\theta \in \Theta = \mathbb{R}^+$ (loi de Poisson).

Exercice 1

La loi de Pareto a été introduite pour modéliser la distribution des revenus supérieurs à un seuil donné, puis s'est avérée utile pour d'autres phénomènes (par exemple la distribution de la taille de grains de sables passés au travers d'un tamis). Elle a deux paramètres strictement positifs : le paramètre de seuil a et un paramètre de forme θ. La fonction de répartition de cette loi est donnée par :

$$F(x) = \begin{cases} 1 - \left(\frac{a}{x}\right)^\theta & \text{si } x \geqslant a \\ 0 & \text{sinon} \end{cases}$$

1. On suppose a connu. Déterminer l'estimateur du maximum de vraisemblance pour le paramètre θ. Quel est son biais (Indication : on pourra déterminer la loi de $log(\frac{X}{a})$).
2. Calculer la borne de Cramer-Rao et constater que l'EMV est asymptotiquement efficace.

Exercice 2

Soit $(\Omega, \mathscr{A}, m)$ un espace mesuré. Soit $T : (\Omega, \mathscr{A}) \to \mathbb{R}^d$ une application mesurable et $f : \Theta \to \mathbb{R}^d$ telles que

$$(\forall \theta \in \Theta), \int_\Omega exp(< f(\theta), T(\omega) >)dm(\omega) \in \mathbb{R}_+^*.$$

On définit ϕ par $\phi(\theta) = ln\left(\int_\Omega exp(< f(\theta), T(\omega) >)dm(\omega)\right)$. La famille $(\mathbb{P}_\theta)_{\theta\in\Theta}$ de probabilités définies par $\mathbb{P}_\theta = p_\theta.m$, avec $p_\theta(\omega) = exp(< f(\theta), T(\omega) > -\phi(\theta))$ constitue sur $(\Omega, \mathscr{A})$ un modèle statistique appelé modèle exponentiel.

1. Soit un modèle exponentiel de la forme $\mathbb{P}_\theta(d\omega) = p_\theta(\omega)dm(\omega)$, avec Θ un ouvert de $\mathbb{R}^p$. Posons $f = Id$ (donc $d = p$). Montrer que T est un estimateur sans biais de $g(\theta) = grad\phi(\theta) = \left(\frac{\partial\phi(\theta)}{\partial\theta_1}, \ldots, \frac{\partial\phi(\theta)}{\partial\theta_p}\right)$.

2. Montrer que la matrice de covariance $K_T(\theta)$ a pour terme général $\frac{\partial^2 \phi(\theta)}{\partial\theta_i \partial\theta_j}$.
3. Montrer que T est de dispersion minimale.

Exercice 3

On considère X une variable aléatoire continue de densité :

$$f_\theta(x) = \frac{\theta}{1-\theta} x^{\frac{2\theta-1}{1-\theta}} \mathbb{1}_{\{0<x<1\}}, \quad (\theta \in]\frac{1}{2}, 1[).$$

Soit $(X_1, \dots, X_n)$ un n-échantillon pour cette variable aléatoire.

1. Donner une structure statistique adaptée.
2. Calculer l'espérance et la variance de X.
3. Trouver un estimateur de θ par la méthode des moments et un estimateur du maximum de vraisemblance.
4. Etudier leur biais, leur convergence (faible, forte, en moyenne quadratique).
5. Donner un intervalle de confiance asymptotique de niveau $1-\alpha$ pour le paramètre θ.

Exercice 4

La loi bêta de paramètres $a, b > 0$ a pour densité (par rapport à la mesure de Lebesgue) :

$$f(x) = B(a,b)x^{a-1}(1-x)^{b-1}\mathbb{1}_{[0,1]}(x).$$

où la constante de normalisation $B(a,b)$ vaut $\Gamma(a+b)/\Gamma(a)\Gamma(b)$ (Γ est définie comme suit : $\Gamma(a+1) = \int_{\mathbb{R}^+} x^a e^{-x} dx$).

1. (a) Exprimer le moment d'ordre k de cette loi à l'aide de la fonction B.
   (b) Exprimer $B(a+1, b)$ en fonction de $B(a,b)$.
   (c) Exprimer les deux premiers moments de cette loi en fonction de a et b.
2. On dispose d'un n-échantillon $X_1, \dots, X_n$ d'une loi bêta dont on ignore la valeur des paramètres $\theta = (a,b)$. Proposer un estimateur $\hat{\theta}_n$ de θ à l'aide de la méthode des moments.

Exercice 5

Soit le modèle statistique $\mathcal{M} = \{(\mathbb{R}^n, \mathcal{B}(\mathbb{R})^{\otimes n}, \mathbb{P}_\theta^{\otimes n}), \theta = (p, m, \sigma) \in \Theta =]0,1[\times\mathbb{R}\times\mathbb{R}_+^*\}$. $\mathbb{P}_\theta$ a pour densité $f_{(p,m,\sigma)}$ par rapport à la mesure de Lebesgue et pour tout $x \in \mathbb{R}$,

$$f_{(p,m,\sigma)}(x) = \frac{p}{\sigma\sqrt{2\pi}} exp\left(-\frac{(x-m)^2}{2\sigma^2}\right) + \frac{1-p}{\sqrt{2\pi}} exp\left(-\frac{(x-m)^2}{2}\right).$$

Montrer qu'il n'existe pas d'estimateur du maximum de vraisemblance de θ.

Bonne chance à vous !