Université Paris-Dauphine Année 2014-2015
L3 - Statistique (1) : Estimation

Examen Final 06/01/2015 - Durée 2h00 - Documents non autorisés

On rappelle que

— la loi normale $\mathcal{N}(\mu, \sigma^2)$ a pour densité $f_{\mu,\sigma^2}(x) = e^{-(x-\mu)^2/(2\sigma^2)}/\sqrt{2\pi}\sigma$

— la loi Exponentielle $\mathcal{E}(\lambda)$ a pour densité $f_\lambda(x) = \lambda e^{-\lambda x}\mathbf{1}_{x>0}$

— la loi $\mathcal{B}eta(\alpha, \beta)$ a pour densité $f_{a,b}(x) = x^{a-1}(1-x)^{b-1}\mathbf{1}_{x\in(0,1)}\Gamma(a+b)/\Gamma(a)\Gamma(b)$

— la loi $\Gamma(a, b)$ a pour densité $f_{a,b}(x) = \mathbf{1}_{x>0}x^{a-1}e^{-bx}b^a/\Gamma(a)$

Exercice 1 (5 pts)

Dans cet exercice il vous est demandé de donner la bonne réponse, seules les réponses justifiées seront validées. Il n'y a pas de points négatifs.

1 Un modèle statistique est associé au vecteur $(X_1, \ldots, X_n)$ en supposant que $X_i \stackrel{\text{ind}}{\sim} \mathcal{N}(\mu, \sigma_i^2)$, avec μ et les σ_i supposés inconnus. Pour une observation $(x_1, \ldots, x_n)$, un estimateur du maximum de vraisemblance de μ est donné par
(a) $\hat{\mu} = x_2$ (b) $\hat{\mu} = (x_1 + \ldots + x_n)/n$ (c) $\sum_{i=1}^n (x_i - \hat{\mu})^{-1} = 0$ (d) $|\hat{\mu}| = \min |x_i|$.

2 Si $X_1, \ldots, X_n$ et $Y_1, \ldots, Y_n$ sont des échantillons iid aléatoires de lois $\mathcal{E}(\lambda)$ et $\mathcal{E}(1/\lambda)$, respectivement, l'estimateur du maximum de vraisemblance de λ associé à l'ensemble des observations $(x_1, \ldots, x_n, y_1, \ldots, y_n)$ est donné par
(a) $\hat{\lambda} = \bar{y}/\bar{x}$ (b) $\hat{\lambda} = (\bar{x} + \bar{y})/2$ (c) $\hat{\lambda} = \sqrt{\bar{x}\bar{y}}$ (d) $\hat{\lambda} = \sqrt{\bar{y}/\bar{x}}$ (e) $\hat{\lambda} = \sqrt{\bar{x}/\bar{y}}$.

3 Si $X_1, \ldots, X_n$ est un échantillon iid de loi $\mathcal{B}eta(\lambda, 1)$, l'estimateur du maximum de vraisemblance de λ associé à l'échantillon $(x_1, \ldots, x_n)$ est donné par
(a) $\hat{\lambda} = -n/\sum_{i=1}^n \log\{x_i\}$ (b) $\hat{\lambda} = 1/\bar{x}$ (c) $\hat{\lambda} = -\sum_{i=1}^n \log\{x_i\}/n$ (d) $\hat{\lambda} = -\log\left\{\prod_{i=1}^n (1 - x_i)\right\}$.

4 Étant donné $X_1, \ldots, X_n$ un échantillon iid de loi $\mathcal{E}(\lambda)$, on observe $(Z_1, \ldots, Z_n) = (\mathbb{I}(X_1 \leq \delta), \ldots, \mathbb{I}(X_1 \leq \delta))$, avec δ connu. Si $(z_1, \ldots, z_n) = (0, \ldots, 0)$, l'estimateur du maximum de vraisemblance de λ est donné par :
(a) $\hat{\lambda} = 1/\bar{x}$ (b) $\hat{\lambda} = \delta$ (c) $\hat{\lambda} = \bar{x}$ (d) $\hat{\lambda} = 1/\delta$ (e) $\hat{\lambda} = +\infty$ (f) $\hat{\lambda} = 0$.

5 Étant donné $X_1, \ldots, X_n$ un échantillon iid de loi $\mathcal{E}(\lambda)$, on observe $(Z_1, \ldots, Z_n) = (\mathbb{I}(X_1 \leq \delta), \ldots, \mathbb{I}(X_1 \leq \delta))$, avec δ connu. Si $z_1 = 1$ et $(z_2, \ldots, z_n) = (0, \ldots, 0)$, l'estimateur du maximum de vraisemblance de λ est donné par :
(a) $\hat{\lambda} = -\log\{\frac{n-1}{n}\}/\delta$ (b) $\hat{\lambda} = \frac{(n-1)\delta}{n}$ (c) $\hat{\lambda} = \frac{\log\{\delta\}}{n-1}$ (d) $\hat{\lambda} = \frac{n+1}{\delta}$ (e) $\hat{\lambda} = 0$.

Exercice 2 (7 pts)

Soit un échantillon aléatoire

$$X_1, X_2, \ldots, X_n \stackrel{\text{iid}}{\sim} \mathcal{N}(0, \theta^2), \quad \theta > 0$$

1 Représenter ce modèle comme une famille exponentielle : donner la paramétrisation naturelle, la représentation minimale, et indiquer si la famille est régulière.

2 Donner une statistique exhaustive minimale T associée à l'échantillon et sa densité.

3 Donner un estimateur sans biais de $\Phi(-1/\theta)$ fondé sur X_1, noté $\delta(X_1)$. *Indication* : Remarquer que $\Phi(-1/\theta)$ est la probabilité que X_1 soit inférerieur à une valeur μ à déterminer.

4 Donner la loi conditionnelle de X_1 sachant $T(X_1, \ldots, X_n)$. *Indication* : Elle doit être constante en θ.

5 En déduire $\delta^\star(X_1, \ldots, X_n) = \mathbb{E}[\delta(X_1)|T(X_1, \ldots, X_n)]$. Expliquer pourquoi la variance de $\delta^\star(X_1, \ldots, X_n)$ est inférieure à celle de $\delta(X_1)$.

Exercice 3 (8 pts)

Soit un échantillon aléatoire $X_1, \cdots, X_n$ de variables indépendantes et identiquement distribuées selon une loi $\mathcal{E}(\lambda)$, $\lambda > 0$. On observe une réalisation $(x_1, \ldots, x_n)$.

1 Ecrire la fonction de vraisemblance de l'échantillon.

2 Calculer le maximum de vraisemblance de λ et déterminer sa loi pour tout $n \geq 1$. *Indication* : On commencera par montrer que $\sum_{i=1}^n X_i$ suit une loi Gamma dont on donnera les paramètres.

En déduire l'estimateur du maximum de vraisemblance de $\exp\{-\lambda\}$.

3 On considère, dans une approche bayésienne, une loi de probabilité a priori Gamma sur λ, $\lambda \sim \Gamma(a, b)$.
Pour tout $(a, b) \in (\mathbb{R}^{+*})^2$, déterminer la loi a posteriori de λ et en déduire que la famille des lois Gamma est une famille conjuguée pour le modèle $\mathcal{E}(\lambda)$.

4 Déterminer l'espérance a posteriori $\mathbb{E}^\pi[\lambda|X_1, \cdots, X_n]$, issue de la loi a posteriori de la question précédente. On notera $\hat{\lambda}_n^\pi$ cette quantité.

5 Montrer que $\sqrt{n}(\hat{\lambda}_n^\pi - \lambda)$ converge en loi vers une loi normale dont on déterminera l'espérance et la variance.

Exercice 4 (4 pts)

1 Donner la définition d'une statistique exhaustive.

2 Dans une famille exponentielle

$$f_\theta(x) = h(x)e^{\theta^t T(x) - \psi(\theta)} \tag{1}$$

donner l'expression de l'espérance et de la variance de $T(X)$ en fonction des dérivées de ψ.

3 Calculer l'information de Fisher en θ dans le modèle (1).

4 Quelle est l'expression analytique de la variance associée à la distribution bootstrap de l'échantillon $(x_1, \ldots, x_n)$?

French – English Lexicon

bayésien(ne) – Bayesian
échantillon – sample
famille exponentielle – exponential family
loi *a posteriori* – posterior distribution
loi *a priori* – prior distribution
statistique exhaustive – sufficient statistic
vraisemblance – likelihood