Université de Montpellier — Année 2015-2016

M1 BIOSTATISTIQUE & MIND - HMMA101

Examen partiel - 6 novembre 2015
Durée 1h - Documents interdits

Exercice 1 (6 pts) On veut estimer la proportion θ des étudiants qui dorment plus de 8 heures par nuit. Les observations sur un échantillon de 30 étudiants donnent : 10 étudiants dorment plus de 8 heures et 20 moins de 8 heures. Nous obtenons ainsi une réalisation d'une loi binomiale $\mathcal{B}(30,\theta)$ où $\theta \in [0,1]$. On se place dans le paradigme bayésien et on envisage deux lois a priori pour θ :

— la loi uniforme sur $[0,1]$,

— une loi Beta de paramètre $(2,2)$.

On rappelle que si la variable aléatoire θ suit une loi Beta de paramètre (a,b), elle admet comme densité

$$\pi(\theta) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}\theta^{a-1}(1-\theta)^{b-1}\mathbb{I}_{\theta\in[0,1]}$$

avec $a > 0$ et $b > 0$.

1) (3 pts) Donner les deux lois a posteriori correspondantes.

2) (3 pts) Calculer les deux estimateurs de Bayes correspondants aux maximums des densités a posteriori et les estimations associées.

Exercice 2 (8 pts) Soit $(x_1,\dots,x_n)$ un n-échantillon d'une loi normale de paramètre $(\mu, 1/\theta)$. On suppose que μ est connu et on cherche à estimer θ. On se place dans le paradigme bayésien et on considère comme loi a priori sur θ une loi Gamma de paramètres (α,β) :

$$\pi(\theta) = \frac{\beta^\alpha}{\Gamma(\alpha)}\theta^{\alpha-1}\exp\left(-\beta\theta\right)\mathbb{I}_{\theta>0}$$

avec $\alpha > 0$ et $\beta > 0$.

1) (2 pts) Donner la loi a posteriori de θ.

2) (4 pts) Donner l'estimateur de Bayes de θ pour la fonction de coût quadratique.

3) (2 pts) Donner l'estimateur du maximum de vraisemblance de θ et comparer à celui obtenu à la question précédente.

Exercice 3 (6 pts) Soit X une loi binomiale négative de paramètre (n,θ), $\forall x \in \mathbb{N}$,

$$\mathbb{P}(X = x|\theta) = C_{n+x-1}^{n-1}\theta^x(1-\theta)^n$$

avec $\theta \in]0,1[$. On suppose n fixé et l'on souhaite estimer θ. On se place dans le paradigme bayésien.

1) (3 pts) Donner la loi non informative de Jeffreys pour le paramètre θ.

2) (3 pts) Soit $(x_1,\dots,x_N)$ un N-échantillon de la loi binomiale négative de paramètre (n,θ). Donner la loi a posteriori de θ pour la loi a priori obtenue à la question précédente.

Examen partiel
HΠΠA 101
06/11/2015

Exercice 1

1] $y|\theta \sim B(n,\theta)$

a] $\pi_1(\theta) = \mathbb{1}_{[0,1]}(\theta)$

$\pi_1(\theta|y) \propto \theta^y(1-\theta)^{n-y}\mathbb{1}_{[0,1]}(\theta)$

$\Rightarrow \theta|y \sim Beta(y+1, n-y+1)$

Pour $n=30$ et $y=10$, nous obtenons

$\theta|y=10 \sim Beta(11,21)$

b] $\pi_2(\theta) \propto \theta(1-\theta)\mathbb{1}_{[0,1]}(\theta)$

$\pi_2(\theta|y) \propto \theta^{y+1}(1-\theta)^{n-y+1}$

$\Rightarrow \theta|y \sim Beta(y+2, n-y+2)$

Pour $n = 30$ et $y = 10$, nous obtenons

$\theta \mid y = 10 \sim \mathrm{Beta}(12, 22)$

2) a) $\hat{\theta}^1(y) \in \arg\max_{\theta \in [0,1]} \pi_1(\theta \mid y)$

Si $y = 0$, $\hat{\theta}^1(0) \in \arg\max_{\theta \in [0,1]} (1-\theta)^n$.

$(1-\theta)^n$ fonction décroissante en θ

et donc $\hat{\theta}^1(0) = 0$

Si $y = n$, $\hat{\theta}^1(n) \in \arg\max_{\theta \in [0,1]} \theta^n$

θ^n fonction croissante en θ

et donc $\hat{\theta}^1(n) = 1$

Si $y \in \{1, \dots, n-1\}$,

$\hat{\theta}^1(y) \in \arg\max_{\theta \in [0,1]} \theta^y (1-\theta)^{n-y}$

$\hat{\theta}^1(y) \in \arg\max_{\theta \in]0,1[} y \log(\theta) + (n-y) \log(1-\theta)$

③

Soit $q(\theta) = y \log(\theta) + (n-y)\log(1-\theta)$

$$\frac{dq}{d\theta}(\theta^*) = 0 \iff \frac{y}{\theta^*} - \frac{n-y}{1-\theta^*} = 0$$

$$\iff \theta^* = \boxed{\frac{y}{n}}$$

Par ailleurs, $\frac{d^2q}{(d\theta)^2}(\theta) = -\left(\frac{y}{\theta^2}\right) - \frac{n-y}{(1-\theta)^2}$

< 0

Ainsi, $\widehat{\theta}^1(y) = \boxed{\frac{y}{n}}$

Au final, $\widehat{\theta}^1(y) = \frac{y}{n}$

Pour $n = 30$ et $y = 10$, on obtient

$\widehat{\theta}^1(10) = \boxed{\frac{1}{3}}$.

b) $\widehat{\theta}^2(y) = \frac{y+1}{n+2}$

Exercice 2

1) $\pi(\theta) \propto \theta^{\alpha-1} e^{-\beta\theta} \mathbb{1}\{\theta > 0\}$

$\underline{x} = (x_1, \dots, x_n)$

$$f(\underline{x} \mid \theta) = \prod_{i=1}^{n} \left\{ \frac{\theta^{1/2}}{\sqrt{2\pi}} e^{-\frac{\theta}{2}(x_i-\mu)^2} \right\}$$

4

$$f(\underline{x}|\theta) = \left(\frac{\theta}{2\pi}\right)^{n/2} e^{-\frac{\theta}{2}\sum_{i=1}^{n}(x_i-\mu)^2}$$

$$\pi(\theta|\underline{x}) \propto \theta^{n/2+\alpha-1}$$

$$\exp\left\{-\theta\left[\frac{\sum_{i=1}^{n}(x_i-\mu)^2}{2}+\beta\right]\right\}$$

$$\mathbb{1}\{\theta>0\}$$

$$\Rightarrow \theta|\underline{x} \sim Ga\left(\frac{n}{2}+\alpha, \frac{\sum_{i=1}^{n}(x_i-\mu)^2}{2}+\beta\right)$$

2) L'estimateur de Bayes associé au coût quadratique est la moyenne de la loi a posteriori.

$$\mathbb{E}(\theta|\underline{x}) = \int_0^{+\infty} \frac{\gamma^{n/2+\alpha}}{\Gamma\left(\frac{n}{2}+\alpha\right)} \theta^{n/2+\alpha} e^{-\gamma\theta} d\theta$$

où $\gamma = \frac{\sum_{i=1}^{n}(x_i-\mu)^2}{2}+\beta$

$$\Longleftrightarrow \mathbb{E}(\theta|\underline{x}) = \frac{\gamma^{M/2+\alpha}}{\Gamma(M/2+\alpha)} \frac{\Gamma(M/2+\alpha+1)}{\gamma^{M/2+\alpha+1}}$$

$$\Longleftrightarrow \mathbb{E}(\theta|\underline{x}) = \left[\frac{M/2+\alpha}{\gamma}\right] = \left[\frac{M/2+\alpha}{\frac{\sum_{i=1}^{M}(x_i-M)^2}{2}+\beta}\right]$$

$(\Gamma(x+1) = x\,\Gamma(x),\ \forall x > 0)$

3] $V(\theta;\underline{x}) = \left(\frac{\theta}{2\pi}\right)^{M/2} e^{-\frac{\theta}{2}\sum_{i=1}^{M}(x_i-M)^2}$

Soit $L(\theta;\underline{x})$ la fonction de log-vraisemblance.

$$L(\theta;\underline{x}) = \log(V(\theta;\underline{x}))$$

$$L(\theta;\underline{x}) = \frac{M}{2}\log\left(\frac{\theta}{2\pi}\right) - \frac{\theta}{2}\sum_{i=1}^{M}(x_i-M)^2$$

(6)

Soit $\hat{\theta}_n$ l'estimateur du maximum de vraisemblance de θ

$$\hat{\theta}_n \in \arg\max_{\theta > 0} L(\theta ; \underline{x})$$

L'espace de recherche de l'optimum n'étant pas compact,

si $\theta^* \in \arg\max_{\theta > 0} L(\theta ; \underline{x})$

alors $\frac{dL}{d\theta}(\theta^* ; \underline{x}) = 0 \quad (*)$

$$(*) \iff \frac{n}{2\theta^*} - \frac{1}{2} \sum_{i=1}^{n} (x_i - m)^2 = 0$$

$$(*) \iff \theta^* = \frac{n}{\sum_{i=1}^{n} (x_i - m)^2}$$

Par ailleurs, $\frac{d^2L}{d\theta^2}(\theta ; \underline{x}) = -\frac{n}{2\theta^2} < 0$

(7)

Ainsi, $\hat{\theta}_n = \left[\dfrac{n}{\sum_{i=1}^{n} (X_i - \mu)^2} \right]$

L'estimateur de Bayes est égal à l'estimateur du maximum de vraisemblance si $\alpha = \beta = 0$.

Dans ce cas, la loi a priori est impropre

$\pi(\theta) \propto \mathbb{1}\{\theta > 0\}$.

Exercice 3

$\boxed{1}$ $f(k|\theta) = C_{n+k-1}^{n-1} \theta^{k} (1-\theta)^{n}$

$\log(f(x|\theta)) = \log\left(C_{n+x-1}^{n-1}\right) + x \log(\theta)$
$+ n \log(1-\theta)$

$$\frac{d\log f}{d\theta}(x|\theta) = \frac{x}{\theta} - \frac{m}{1-\theta}$$

$$\frac{d^2\log f}{(d\theta)^2}(x|\theta) = -\frac{x}{\theta^2} - \frac{m}{(1-\theta)^2}$$

$$I_X(\theta) = \frac{\mathbb{E}(X|\theta)}{\theta^2} + \frac{m}{(1-\theta)^2}$$

$$\mathbb{E}(X|\theta) = \sum_{k\in\mathbb{N}} k\, C_{m+k-1}^{m-1}\, \theta^k (1-\theta)^m$$

$$\sum_{k\in\mathbb{N}} C_{m+k-1}^{m-1}\, \theta^k (1-\theta)^m = 1$$

$$\sum_{k\in\mathbb{N}} C_{m+k-1}^{m-1} \left[k\theta^{k-1}(1-\theta)^m - \theta^k m (1-\theta)^{m-1} \right] = 0$$

$$\sum_{k\in\mathbb{N}} k\, C_{m+k-1}^{m-1}\, \theta^k (1-\theta)^m =$$

$$\frac{m\theta}{1-\theta} \sum_{k\in\mathbb{N}} C_{m+k-1}^{m-1}\, \theta^k (1-\theta)^m = \left(\frac{m\theta}{1-\theta}\right)$$

9

$$I_X(\theta) = \frac{m}{\theta(1-\theta)} + \frac{m}{(1-\theta)^2}$$

$$I_X(\theta) = \left[\frac{m}{\theta(1-\theta)^2}\right]$$

Ainsi, $\pi^J(\theta) \propto \theta^{-1/2}(1-\theta)^{-1} \mathbb{1}_{]0,1[}(\theta)$

c'est une loi impropre

3) $f(\underline{x} \mid \theta) = (1-\theta)^{NM} \theta^{\sum_{i=1}^{N} x_i} \prod_{i=1}^{N} C_{M+x_i-1}^{M-1}$

$\underline{x} = (x_1, \ldots, x_N)$

$$\pi^J(\theta \mid \underline{x}) \propto (1-\theta)^{Nm-1} \theta^{\sum_{i=1}^{N} x_i - \frac{1}{2}} \mathbb{1}_{]0,1[}(\theta)$$

$$\Rightarrow \theta \mid \underline{x} \sim \mathrm{Beta}\left(\sum_{i=1}^{N} x_i + \frac{1}{2}, Nm\right)$$