Estatística MAEG 2011/12
Exercícios Capítulo 5

1. Considere o modelo exponencial de média $1/\lambda$, $\lambda > 0$, i.e.,

$$f(y \mid \lambda) = \lambda \exp(-\lambda y), \ y > 0 \ .$$

   (a) Determine a informação de Fisher sobre λ contida numa observação do modelo.

   (b) Considere agora a parametrização alternativa $\mu = 1/\lambda$. Determine a informação de Fisher sobre μ contida numa observação do modelo de duas formas: directamente, e utilizando o resultado que a com relaciona a informação calculada na alínea anterior.

   (c) Determine a distribuição *a priori* de Jeffreys para λ. Trata-se de uma distribuição imprópria?

   (d) Suponha que dispõe de uma amostra casual de tamanho n proveniente do modelo exponencial. Determine a distribuição *a posteriori* de λ associada à distribuição *a priori* de Jeffreys.

2. Uma variável aleatória Y segue a distribuição de pareto se a sua função densidade é da forma

$$f(y \mid a, b) = ab^a y^{-(a+1)} I_{(b,+\infty)}(y)$$

   com $a, b > 0$.

   (a) Mostre que a família pareto é a família conjugada natural do modelo estatístico $\{\mathrm{U}(0,\theta) : \ \theta > 0\}$.

   (b) Assuma que $x_1, \dots, x_n$ corresponde ao valor observado de uma amostra casual de tamanho n proveniente do modelo $\{\mathrm{U}(0,\theta) : \ \theta > 0\}$. Usando uma distribuição *a priori* conjugada natural, determine a distribuição *a posteriori* de θ.

   (c) A distribuição *a posteriori* acima depende dos dados unicamente através do máximo da amostra. Podia este facto ter sido antecipado? Porquê?

   (d) Suponha que a informação disponível sobre θ permite ao analista especificar que o valor médio e a variância *a priori* de θ são, respectivamente, m e v. Seleccione um membro da família conjugada natural que reflicta esta informação.

3. Diz-se que uma variável aleatória X segue uma distribuição gama invertida, simbolicamente, $X \sim \mathrm{IG}(\alpha, \beta)$ sse $X = 1/Y$ com $Y \sim \mathrm{G}(\alpha, \beta)$, ou seja,

$$f_Y(y) = \frac{\beta^\alpha}{\Gamma(\alpha)} \, y^{\alpha-1} \, \exp(-\beta y) \ , \qquad y > 0$$

   onde $\alpha, \beta > 0$.

   (a) Determine a função densidade de probabilidade de X se $X \sim \mathrm{IG}(\alpha, \beta)$.

   (b) Considere o modelo $\mathcal{F} = \{\mathrm{N}(0, \sigma^2) : \sigma > 0\}$ e seja $W_1, \dots, W_n$ uma amostra casual de dimensão n dele proveniente.

i. Determine a distribuição *a priori* de Jeffreys para σ^2 nestas condições.
ii. Mostre que utilizando a distribuição *a priori* acima se tem $\sigma^2 \mid w_1, \ldots, w_n \sim \mathrm{IG}(n/2, \sum w_i^2/2)$.

4. Diz-se que uma variável aleatória X segue uma distribuição geométrica se, para algum $\theta \in (0, 1)$, a sua função probabilidade é da forma

$$f(x \mid \theta) = \theta(1-\theta)^{x-1}, \quad x = 1, 2, 3, \ldots$$

Pode assumir sem demonstração que $E[X \mid \theta] = 1/\theta$. Seja $X_1, \ldots, X_n$ uma amostra casual proveniente de uma população geométrica de parâmetro θ.

(a) Mostre que a distribuição *a priori* de Jeffreys é $\pi^J(\theta) \propto \theta^{-1}(1-\theta)^{-1/2}$, $\theta \in (0, 1)$.

(b) Verifique que a distribuição *a posteriori* de θ associada a π^J é tal que

$$\theta \mid x_1, \ldots, x_n \sim \mathrm{Be}(n, \textstyle\sum x_i - n + 1/2) \ .$$

(c) Suponha que dispomos de uma moeda que, lançada ao ar, fica com o lado da coroa para cima com probabilidade θ. Dada a informação contida em $x_1, \ldots, x_n$, determine uma estimativa da probabilidade de, num lançamento futuro dessa moeda, independente das experiências anteriores, obtermos uma coroa.

5. Seja $X_1, \ldots, X_n$ uma amostra casual proveniente do modelo Poisson de valor médio λ, $\lambda > 0$.

(a) Mostre que se $\lambda \sim \mathrm{G}(a, b)$, $a, b > 0$, então $\lambda \mid x_1, \ldots, x_n \sim \mathrm{G}(a + t, b + n)$, onde $t = \sum x_i$.

(b) Escreva a média *a posteriori* de λ como uma combinação linear da média amostral e da média *a priori* de λ.

(c) O que acontece à média *a posteriori* de λ quando $n \to +\infty$ de tal forma que $\bar{x}$ se mantém constante? Interprete este fenómeno.

(d) Suponha que se pretende predizer uma variável aleatória Y, independente de $X_1, \ldots, X_n$ dado λ, e tal que $Y \mid \lambda \sim \mathrm{Ex}(\lambda)$. Determine a distribuição preditiva *a posteriori* de Y.

6. Considere o seguinte modelo estatístico para o par (X, Y):

$$Y \mid X = x, \lambda \sim \mathrm{Po}(\lambda x) \qquad \text{e} \qquad X \mid \lambda \sim \mathrm{Ex}(\lambda)$$

Denote-se por $(x_1, y_1), \ldots, (x_n, y_n)$ a concretização de uma amostra casual de dimensão n proveniente de tal modelo. Pretende-se implementar uma análise bayesiana do problema baseada na distribuição *a priori* $\pi(\lambda) \propto \lambda^{-1}$.

(a) Mostre que $\pi(\lambda)$ corresponde à distribuição *a priori* de Jeffreys.

(b) Verifique que a estimativa de máxima verosimilhança de λ é dada por $\hat{\lambda} = (\bar{y} + 1)/(2\bar{x})$.

(c) Verifique que a média *a posteriori* e a estimativa de máxima verosimilhança de λ coincidem.

(d) Construa um intervalo de credibilidade central $(1-\alpha)$ para λ.

7. Seja $X_1, \ldots X_n$ uma amostra casual de tamanho n proveniente de uma população X com função densidade de probabilidade dada por

$$f(x \mid \theta) = \frac{2}{\theta^2}\, x, \quad 0 < x < \theta$$

onde $\theta > 0$. Designe-se por $x_1, \ldots, x_n$ a correspondente amostra observada. Pretende-se implementar uma análise bayesiana destes dados.

(a) Considere que, *a priori*,

$$\pi(\theta) \propto \frac{1}{\theta^2}\, I_{(1,+\infty)}(\theta)\ .$$

Esta distribuição *a priori* corresponde a uma distribuição própria? Em caso afirmativo, determine a correspondente constante de normalização.

(b) Assumindo a distribuição *a priori* acima, mostre que

$$\pi(\theta \mid x_1, \ldots, x_n) = (2n+1)\ (x_{(n)} \vee 1)^{2n+1}\ \theta^{-(2n+2)}\ I_{(x_{(n)} \vee 1, +\infty)}(\theta)$$

onde $x_{(n)}$ designa o máximo da amostra e $a \vee b \equiv \max\{a, b\}$.

(c) Mostre que

$$\left(\frac{x_{(n)} \vee 1}{(1-\alpha/2)^{1/(2n+1)}}\ ,\ \frac{x_{(n)} \vee 1}{(\alpha/2)^{1/(2n+1)}} \right)$$

corresponde a um intervalo de credibilidade central $(1-\alpha)$ para θ.

Soluções:

1. a) $I_X(\lambda) = \lambda^{-2}$ b) $I^*_X(\mu) = \mu^{-2}$ c) $\pi^J(\lambda) \propto \lambda^{-1},\ \lambda > 0$; imprópria d) $\lambda \mid \boldsymbol{x} \sim \mathrm{G}(n, \sum x_i)$

2. b) $\theta \mid \boldsymbol{x} \sim \mathrm{Pa}(n+a, b \vee x_{(n)})$, onde $a \vee b \equiv \max\{a, b\}$ c) Sim, $X_{(n)}$ é suficiente para θ

3. a) $f_X(x) = \beta^\alpha/\Gamma(\alpha)\ x^{-(\alpha+1)} \exp(-x/\alpha)$ b)i. $\pi^J(\sigma^2) \propto 1/\sigma^2$

4. c) $P(Y = 1 \mid \boldsymbol{x}) = n/(\sum x_i + 1/2)$

5. b) $E[\lambda \mid \boldsymbol{x}] = n/(n+b)\ \bar{x} + b/(n+b) a/b$ onde $\bar{x}$ é a média amostral e a/b corresponde à média *a priori* de λ c) $E[\lambda \mid \boldsymbol{x}] \to \bar{x}$ d) $f(y \mid \boldsymbol{x}) = (b+n)^{a+t}/(a+t-1)\ (b+n+y)^{-(a+t-1)}$, $y > 0$.

6. d) $(F^{-1}_{\chi^2(m)}(\alpha/2)/(4\sum x_i), F^{-1}_{\chi^2(m)}(1-\alpha/2)/(4\sum x_i))$, onde $m = 2n(\bar{y}+1)$

7. a) própria, constante de normalização igual a 1.