Lista de exercícios 5

Estatística Bayesiana

3.44 *(Dawid et al. (1973)) Consider n random variables $x_1, \dots, x_n$, such that the first ξ of these variables has an $\mathcal{E}xp(\eta)$ distribution and the $n-\xi$ other have a $\mathcal{E}xp(c\eta)$ distribution, where c is known and ξ takes its values in $\{1, 2, \dots, n-1\}$.

a. Give the shape of the posterior distribution of ξ when $\pi(\xi, \eta) = \pi(\xi)$ and show that it only depends on $z = (z_2, \dots, z_n)$, with $z_i = x_i/x_1$.

b. Show that the distribution of z, $f(z|\xi)$, only depends on ξ.

c. Show that the posterior distribution $\pi(\xi|x)$ cannot be written as a posterior distribution for $z \sim f(z|\xi)$, whatever $\pi(\xi)$, although it only depends on z. How do you explain this?

d. Show that the paradox does not occur when $\pi(\xi, \eta) = \pi(\xi)\eta^{-1}$.

a) $X_i | \xi, \eta \sim \text{Exp}(c_i \eta)$ com $c_i = \begin{cases} 1, & i \leq \xi \\ c, & i > \xi \end{cases}$.

$$f(x_1, \dots, x_n | \xi, \eta) = \prod_{i=1}^{\xi} \eta e^{-x_i \eta} \prod_{i=\xi+1}^{n} c\eta e^{-x_i c\eta}$$
$$= c^{n-\xi} \eta^n e^{-\eta \sum_{i=1}^n c_i x_i}.$$

$$\pi(\xi, \eta | x_1, \dots, x_n) \propto c^{n-\xi} \eta^n e^{-\eta \sum_{i=1}^n c_i x_i} \pi(\xi).$$

Assim

$$\pi(\xi | x_1, \dots, x_n) \propto \int_0^{+\infty} c^{n-\xi} \eta^n e^{-\eta \sum_{i=1}^n c_i x_i} \pi(\xi) \, d\eta$$
$$= c^{n-\xi} \pi(\xi) \int_0^{+\infty} \underbrace{\eta^n e^{-\eta \sum_{i=1}^n c_i x_i}}_{\text{Gamma}(n+1, \sum c_i x_i)} d\eta$$
$$\propto c^{-\xi} \left(\sum_{i=1}^{\xi} x_i + c \sum_{i=\xi+1}^{n} x_i \right)^{-(n+1)} \pi(\xi)$$
$$\propto c^{-\xi} \left(\sum_{i=1}^{\xi} z_i + c \sum_{i=\xi+1}^{n} z_i \right)^{-(n+1)} \pi(\xi),$$

sendo que $z_1 = 1$.

b) $x \mapsto (x_1, z)$ é bijetiva, com

$$J = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ -\frac{x_2}{x_1^2} & \frac{1}{x_1} & 0 \cdots & 0 \\ \vdots & 0 & \ddots & 0 \\ -\frac{x_n}{x_1^2} & 0 & \cdots 0 & 1/x_1 \end{pmatrix} \Rightarrow \det J = x_1^{-n+1}.$$

Assim $f(x_1, z \mid \xi, \eta) = c^{n-\xi} \eta^n e^{-\eta \sum_{i=1}^n c_i x_1 z_i} \cdot x_1^{n-1}$

e $f(z \mid \xi, \eta) = c^{n-\xi} \eta^n \int_0^{+\infty} e^{-\eta x_1 \sum_{i=1}^n c_i z_i} x_1^{n-1} \, dx_1$ Gamma$(n, \eta \sum c_i z_i)$

$$= c^{n-\xi} \eta^n \cdot \frac{(n-1)!}{\eta^n \left(\sum_{i=1}^n c_i z_i\right)^n}$$

$$= (n-1)! \, c^{n-\xi} \left(\sum_{i=1}^n c_i z_i\right)^{-n},$$

que independe de η.

c) Por um fator $(\sum c_i z_i)^{-1}$, temos que independente da escolha de $\pi(\xi)$, teremos que $f(z \mid \xi)\pi(\xi)$ não é proporcional a $\pi(\xi \mid x)$, embora esse só dependa de x através de z.

O problema parece advir da priori imprópria em η.

d) Quando $\pi(\xi, \eta) = \pi(\xi)/\eta$,

$\pi(\xi \mid x) \propto c^{-\xi} \left(\sum_{i=1}^{\xi} z_i + c \sum_{i=\xi+1}^n z_i\right)^{-n} \pi(\xi)$ e o problema é resolvido.

3.51 In relation to Example 3.5.7, if $x \sim \mathcal{B}(n,p)$, find a prior distribution on n such that $\pi(n|x)$ is $\mathcal{N}eg(x,p)$.

Seja p fixo. Então

$$f(x|n) = \binom{n}{x} p^x (1-p)^{n-x} \mathbb{1}_{\{0,\dots,n\}}(x)$$

Queremos encontrar $\pi(n)$ tal que

$$\pi(n|x) = \binom{n-1}{x-1} p^x (1-p)^{n-x}$$

Aqui $n \sim Neg(x,p)$ significa n tentativas com x sucessos fixados a priori. (Errata do livro). Sejam $\pi(n)$ a priori desejada e $m(x)$ a marginal de x. Assim

$$\frac{\pi(n)}{m(x)} = \frac{\pi(n|x)}{f(x|n)} = \frac{(n-1)!\,x!\,(n-x)!}{n!\,(x-1)!\,(n-x)!} = \frac{1/n}{1/x}.$$

Considere a priori, imprópria, $\pi(n) = 1/n$. Assim

$$\begin{aligned}
\pi(n|x) &\propto \frac{n!}{x!(n-x)!} p^x (1-p)^{n-x} \cdot \frac{1}{n} \\
&= \frac{(n-1)!}{(x-1)!(n-x)!} p^x (1-p)^{n-x} \cdot \frac{1}{x} \\
&\propto \binom{n-1}{x-1} p^x (1-p)^{n-x},
\end{aligned}$$

como desejávamos.

3.57 Show that a second-order approximation of both the entropy and Hellinger losses introduced in Section 2.5.4 is $(\theta-\delta)^2 I(\theta)$. Does this result give additional support to the use of the Jeffreys prior?

Ver exercício 2.48 resolvido na lista 2.

$$L_E(\theta,\delta) = C_E(\theta-\delta)^T I(\theta)(\theta-\delta) + O(\|\theta-\delta\|^2)$$
$$L_H(\theta,\delta) = C_H(\theta-\delta)^T I(\theta)(\theta-\delta) + O(\|\theta-\delta\|^2)$$

Sugiro o artigo "Jeffreys' prior is asymtotically least favorable under entropy risk" de Clarke e Barron.

Note que tomando $\tilde{\pi}(\theta) \propto \pi(\theta) I(\theta)$, temos que sob perda quadrática, $\delta^{\tilde{\pi}}$ é equivalente a δ^{π} sob
$$L(\theta,\delta) = I(\theta)(\theta-\delta)^2.$$
Com isso, o uso da priori de Jeffreys está associado a adotar L em um contexto frequentista, por exemplo.

3.58 Consider $x \sim \mathcal{P}(\theta)$.

a. Determine the Jeffreys prior π^J and discuss whether the scale-invariant prior $\pi_0(\theta) = 1/\theta$ is preferable.

b. Find the maximum entropy prior for the reference measure π_0^J and the constraints $\mathbb{E}^\pi[\theta] = 1$, $\text{var}^\pi(\theta) = 1$. What about using π_0 instead?

c. Actually, x is the number of cars crossing a railroad in a period T. Show that x is distributed according to a Poisson distribution $\mathcal{P}(\theta)$ if the interval between two arrivals is distributed according to $\mathcal{E}xp(\lambda)$. Note that $\theta = \lambda T$.

d. Use the above derivation of the Poisson distribution to justify the use of π_0.

$$f(x|\theta) = \frac{\theta^x e^{-\theta}}{x!}$$

a)
$$\frac{\partial^2}{\partial\theta^2} \log f(x|\theta) = \frac{\partial^2}{\partial\theta^2}\left(x\log\theta - \theta - \log x!\right)$$
$$= \frac{\partial}{\partial\theta}\left(\frac{x}{\theta} - 1\right)$$
$$= -\frac{x}{\theta^2}$$

$I(\theta) = -E[-x/\theta^2] = 1/\theta$. Assim $\pi^J(\theta) \propto 1/\sqrt{\theta}$

Nesse caso, $\theta|x \sim \text{Gamma}(x+1/2, 1)$. Por outro lado, se $\pi_0(\theta) = 1/\theta$, $\theta|x \sim \text{Gamma}(x, 1)$. Um primeiro problema com π_0 é que se $x=0$, a posteriori é indefinida.

$$\mathbb{P}(x=0|\theta) = e^{-\theta}.$$

Além disso, θ não é somente um parâmetro de escala, o que reduz o apelo por π_0. Por fim, se crenças sobre θ_1 e θ_2 só dependam de θ_1/θ_2, π_0 é vantajosa.

b) Suponha $E_\pi[\theta] = 1$ e $Var_\pi(\theta) = 1 \Rightarrow E[\theta^2] = 2$

$$\mathcal{E}(\pi) = \int \log\left(\frac{\pi(\theta)}{\theta^{-1/2}}\right)\theta^{-1/2}\, d\theta$$

Pela equação 3.2.2., a priori que maximiza $\mathcal{E}$ é

$$\pi^*(\theta) \propto \exp\left\{\lambda_1\theta + \lambda_2\theta^2\right\}\theta^{-1/2}$$

$$\propto \exp\left\{\lambda_2\left(\theta + \frac{\lambda_1}{2\lambda_2}\right)^2\right\}\theta^{-1/2}$$

Se π_0 é a priori de referência,

$$\pi^*(\theta) \propto \exp\left\{\lambda_2\left(\theta + \frac{\lambda_1}{2\lambda_2}\right)^2\right\}\theta^{-1}$$

c) Defina T_i = passagem do i-th carro. Suponha que $\{T_{i+1} - T_i\}_{i=0}^{\infty}$ seja iid com $T_2 - T_1 \sim Exp(\lambda)$ e $T_0 = 0$. Também defina

$$X = \max\{n : T_n \leq T\}$$

Note que

$$P(X \geq K) = P(T_1 \leq T, \ldots, T_K \leq T) = P(T_K \leq T).$$

Logo
$$\begin{aligned} P(X = K) &= P(T_K \leq T) - P(T_{K+1} \leq T). \\ &= P(T_{K+1} > T) - P(T_K > t). \\ &= \sum_{i=0}^{K} (\lambda T)^i / i!\, e^{-\lambda T} - \sum_{i=0}^{K-1} (\lambda T)^i / i!\, e^{-\lambda T} \\ &= \frac{(\lambda T)^K}{K!} e^{-\lambda T}, \end{aligned}$$

em que a 3º igualdade se deve ao fato que

$$T_K = \sum_{i=0}^{K-1} T_{i+1} - T_i \sim Gamma(K, \lambda).$$

Note que $\theta = \lambda T$.

d) Vamos encontrar a priori de Jeffreys para λ:

$$-\frac{\partial^2}{\partial\lambda} \log f(x|\lambda) = \frac{1}{\lambda^2} \Rightarrow I(\lambda) = 1/\lambda^2.$$

Com isso $\pi(\lambda) = \lambda^{-1}$. Usando a mudança de variáveis,

$$\pi(\theta) \propto 1/\theta = \pi_0(\theta).$$

Isso justifica usar π_0.

3.63 Consider the class of prior distributions

$$\Gamma = \{\mathcal{N}(\mu, \tau^2),\ 0 \le \mu \le 2,\ 2 \le \tau^2 \le 4\}$$

when $x \sim \mathcal{N}(\theta, 1)$.

a. Study the variations of $\mathbb{E}^\pi[\theta|x]$ and $\mathrm{var}^\pi(\theta|\mathrm{x})$ when $\pi \in \Gamma$.

b. Study $\varrho(\pi, \delta^{\pi'})$ when $\pi, \pi' \in \Gamma$ and $\delta^\pi(x) = \mathbb{E}^\pi[\theta|x]$, $\mathrm{L}(\theta, \delta) = (\theta - \delta)^2$ in order to determine the Γ-minimax estimator.

a) Se $x \sim N(\theta, 1)$ e $\theta \sim N(\mu, \tau^2)$, temos que

$$\theta | x \sim N\left(\frac{\mu/\tau^2 + x}{1/\tau^2 + 1}, \frac{\tau^2}{1+\tau^2}\right),$$

pois $N(\mu, \tau^2)$ é família conjugada para $N(\theta, 1)$.

$$E^\pi[\theta|x] = -\frac{\mu + \tau^2 \cdot x}{1+\tau^2}$$

$$\mathrm{var}^\pi[\theta|x] = \left[\tau^{-2} + 1\right]^{-1}.$$

Se $\mu \in [0,2]$ e $\tau^2 \in [2,4]$,

- $E^\pi[\theta|x] \in \left[\frac{2x}{3}, \frac{2+4x}{5}\right]$
- $\mathrm{Var}^\pi[\theta|x] \in \left[\frac{2}{3}, \frac{4}{5}\right]$

b) $\varrho(\pi, \delta^{\pi'}) = E^\pi[(\theta - E^{\pi'}[\theta|x])^2|x]$

$$= E^\pi[\theta^2|x] - 2E^{\pi'}[\theta|x]E^\pi[\theta|x] + (E^{\pi'}[\theta|x])^2$$

$$= \frac{\tau^2}{1+\tau^2} + \left(\frac{\mu+\tau^2 x}{1+\tau^2}\right)^2 - 2\left(\frac{\mu+\tau^2 x}{1+\tau^2}\right)\left(\frac{\mu'+\tau'^2 x}{1+\tau'^2}\right) + \left(\frac{\mu'+\tau'^2 x}{1+\tau'^2}\right)^2$$

$$= \frac{\tau^2}{1+\tau^2} + \left(\frac{\mu+\tau^2 x}{1+\tau^2} - \frac{\mu'+\tau'^2 x}{1+\tau'^2}\right)^2$$

$$\in \left[\frac{2}{3}, \frac{2x+18}{5}\right]$$