Lista de exercícios 6

Estatística Bayesiana

4.5 Show that a setting opposite to Example 4.1.2 may happen, namely, a case when the prior information is negligible. (*Hint:* Consider $\pi(\theta)$ to be $\mathcal{C}(\mu, 1)$ and $f(x|\theta) \propto \exp -|x-\theta|$, and show that the MAP estimator does not depend on μ.)

No exemplo 4.1.2, verificamos que o MAP de θ é sempre 0 quando $x \sim \mathcal{C}(\theta, 1)$ e $\pi(\theta) \propto e^{-|\theta|}$, isto é, x é não importante, nesse caso.

Agora considere a dica. Queremos maximizar em θ

$$f(x|\theta)\pi(\theta) \propto \frac{e^{-|x-\theta|}}{1+(\theta-\mu)^2},$$

Quando $\theta < x$, o ponto crítico é $\theta = \mu + 1$. Já quando $\theta > x$, o ponto crítico é $\theta = \mu - 1$. Além do mais,

$$\lim_{\theta \to +\infty} f(x|\theta)\pi(\theta) = \lim_{\theta \to -\infty} f(x|\theta)\pi(\theta) = 0,$$

portanto existe um compacto $[-M, M]$ de forma que se $|\theta| > M$, então $f(x|\theta)\pi(\theta) < e^{-|x|}/(1+\mu^2)$. Pelo Teorema de Weierstrass, sabemos que existe máximo global em $[-M, M]$. Logo, ele pode ser $\theta = x$, $\theta = \mu + 1$ ou $\theta = \mu - 1$, bastando verificar

$$\frac{1}{1+(x-\mu)^2}, \quad \frac{e^{-|x-\mu-1|}}{2}, \quad \frac{e^{-|x-\mu+1|}}{2}.$$

Defina $z = x - \mu$. Note que

$$2e^{|z \pm 1|} \geq 1 + z^2.$$

é resultado da expansão de Taylor. Isso mostra que o MAP é $\hat{\theta}_{MAP} = x$.

4.17 Consider $x \sim \mathcal{B}(n,p)$ and $p \sim \mathcal{B}e(\alpha,\beta)$.

a. Derive the posterior and marginal distributions. Deduce the Bayes estimator under quadratic loss.

b. If the prior distribution is $\pi(p) = [p(1-p)]^{-1}\mathbb{I}_{(0,1)}(p)$, give the generalized Bayes estimator of p (when it is defined).

c. Under what condition on (α,β) is δ^π unbiased? Is there a contradiction with Exercise 4.16?

d. Give the Bayes estimator of p under the loss

$$\mathrm{L}(p,\delta) = \frac{(\delta-p)^2}{p(1-p)}.$$

a) Como Beta forma uma família conjugada para a binomial, $p \mid x \sim \text{Beta}(\alpha + x, \beta + n - x)$. A distribuição marginal de x é

$$m(x) = \frac{\binom{n}{x} \cancel{p^{x}}\, \cancel{(1-p)}^{n-x} \frac{1}{B(\alpha,\beta)} \cancel{p}^{\alpha-1} \cancel{(1-p)}^{\beta-1}}{\frac{1}{B(\alpha+x,\beta+n-x)} \cancel{p}^{\alpha+x-1} \cancel{(1-p)}^{\beta+n-x-1}}$$

$$= \binom{n}{x} \frac{B(\alpha+x, \beta+n-x)}{B(\alpha,\beta)}$$

O estimador de Bayes para θ é $\delta^\pi(x) = \dfrac{\alpha+x}{\alpha+\beta+n}$.

b) Nesse caso, teremos $p \sim \text{Beta}(0,0)$ e, portanto, o estimador de Bayes é $\delta^\pi(x) = x/n$, mas ele só é definido quando $x \neq 0, n$. Caso contrário, a posteriori é imprópria.

c) $E_p\left[\dfrac{\alpha+x}{\alpha+\beta+n}\right] = \dfrac{\alpha+np}{\alpha+\beta+n}$. Temos que

$$\frac{\alpha+np}{\alpha+\beta+n} = p \iff \alpha = p(\alpha+\beta).$$

Isso só vale se $\alpha = \beta = 0$. Nesse caso, a priori é imprópria, o que não contradiz 4.16.

d) Queremos $\min_{\delta} \rho(\pi, \delta|x)$, em que

$$\rho(\pi, \delta|x) = \int_0^1 \frac{(\delta-p)^2}{p(1-p)} \cdot \frac{p^{\alpha+x-1}(1-p)^{\beta+n-x-1}}{B(\alpha+x, \beta+n-x)} dp$$

$$= \frac{B(\alpha+x-1, \beta+n-x-1)}{B(\alpha+x, \beta+n-x)} \int_0^1 (\delta-p)^2 \frac{p^{\alpha+x-2}(1-p)^{\beta+n-x-2}}{B(\alpha+x-1, \beta+n-x-1)} dp$$

Assim basta minimizar δ sob perda quadrática e a posteriori $p|x \sim Beta(\alpha+x-1, \beta+n-x-1)$ e teremos

$$\delta^{\pi}(x) = \frac{\alpha+x-1}{\alpha+\beta+n-2}$$

4.40 (Jeffreys (1961)) Consider $x_1, \ldots, x_{n_1}$ i.i.d. $\mathcal{N}(\theta, \sigma^2)$. Let $\bar{x}_1$, s_1^2 be the associated statistics. For a second sample of observations, give the predictive distribution of $(\bar{x}_2, s_2^2)$ under the noninformative distribution $\pi(\theta, \sigma) = \frac{1}{\sigma}$. If $s_2^2 = s_1^2/y$ and $y = e^z$, deduce that z follows a Fisher's F distribution.

Sabemos que $\bar{x}_1$ e $\bar{s}_1$ são independentes, condicionado em θ e σ^2, e $\bar{x}_1 \sim N(\theta, \sigma^2/n_1)$, $(n_1-1)s_1^2/\sigma^2 \sim \text{Gamma}(\frac{n-1}{2}, \frac{1}{2})$. Em particular, $s_1^2 \sim \text{Gamma}(\frac{n-1}{2}, \frac{n-1}{2\sigma^2})$. Agora vamos calcular a posteriori:

$$p(\theta, \sigma^2 | x) \propto \frac{1}{\sigma^n} \exp\left\{-\frac{1}{2\sigma^2} \sum_{i=1}^{n} (x_i - \theta)^2\right\} \cdot \frac{1}{\sigma}$$

$$= \frac{1}{\sigma}\left(\frac{1}{\sigma^2}\right)^{n/2} \exp\left\{-\left(\frac{(n-1)s_1^2 + n(\bar{x}-\theta)^2}{2\sigma^2}\right)\right\}$$

Em particular

$$\theta, \sigma^2 | x \sim N\text{-}\Gamma^{-1}\left(\bar{x}, n, \frac{n}{2} - 1, \frac{n-1}{2} s_1^2\right)$$

Para calcular $(\bar{x}_2, s_2^2) | (\bar{x}_1, s_1^2)$, precisamos fazer

$$g^{\pi}(\bar{x}_2, s_2^2 | \bar{x}_1, s_1^2) = \int_{-\infty}^{\infty}\int_0^{\infty} N\left(\theta, \frac{\sigma^2}{n_2}\right) \cdot G\left(\frac{n-1}{2}, \frac{n-1}{2\sigma^2}\right) \cdot N\text{-}\Gamma^{-1}\left(\bar{x}, n, \frac{n}{2}-1, \frac{n-1}{2} s_1^2\right) d\sigma^2 d\theta$$

Como $(n_1-1) s_1^2/\sigma^2 \sim \chi^2_{n-1}$ e $(n_2-1) s_2^2/\sigma^2 \sim \chi^2_{n-1}$,

$$\frac{s_1^2}{s_2^2} = \frac{n_1-1}{n_2-1} \frac{s_1^2/(n_1-1)}{s_2^2/(n_2-1)} \sim F(n_1-1, n_2-1),$$

isto é, $y \sim F(n_1-1, n_2-1)$.

4.44 *For a normal model $\mathcal{N}_k(X\beta, \Sigma)$ where the covariance matrix Σ is totally unknown, give the noninformative Jeffreys prior.

a. Show that the posterior distribution of Σ conditional upon β is a Wishart distribution and deduce that there is no proper marginal posterior distribution on β when the number of observations is smaller than k.

b. Explain why it is not possible to derive a conjugate distribution in this setting. Consider the particular case when Σ has a Wishart distribution.

c. What is the fundamental difference in this model which prevents what was possible in Section 4.4.2?

$$f(y \mid \beta, \Sigma) = \frac{1}{\sqrt{(2\pi)^k |\Sigma|}} \exp\left\{-\frac{1}{2}(y - X\beta)^T \Sigma^{-1} (y - X\beta)\right\}$$

Seja $\ell(\beta, \Sigma \mid y, X) = \log f(y \mid \beta, \Sigma, X)$.

$$= -\frac{1}{2}\underbrace{(y - X\beta)^T \Sigma^{-1} (y - X\beta)}_{tr((y-X\beta)(y-X\beta)^T \Sigma^{-1})} - \frac{1}{2}\log|\Sigma|$$

$$\frac{\partial \ell}{\partial \beta} = X^T \Sigma^{-1} (y - X\beta)$$

$$\frac{\partial \ell}{\partial \Sigma} = -(y - X\beta)(y - X\beta)^T \Sigma^{-2} - \Sigma^{-1} + \frac{1}{2}(\Sigma^{-1} \circ I)$$

$$\frac{\partial^2 \ell}{\partial \beta^2} = -X^T \Sigma^{-1} X \Rightarrow I(\beta) = X^T \Sigma^{-1} X$$

$$\frac{\partial^2 \ell}{\partial \Sigma^2} =$$