Lista de exercícios 4

Estatística Bayesiana

3.15 Show that every distribution from an exponential family can be generalized into a pseudo-exponential family by adding parametrized constraints on the support of x. Elaborate on the modification in the sufficient statistics.

Família exponencial: $f(x|\theta) = C(\theta)\, h(x)\, e^{R(\theta)\cdot T(x)}$, com $\theta \in \Theta$ e $x \in \mathcal{X}$.

Defina $D_\theta = \{x \in \mathcal{X} \mid \|T(x)\| \leq \|R(\theta)\|\}$. Note que

$$\int_{D_\theta} f(x|\theta)\, d\mu(x) \leq \int_{\mathcal{X}} f(x|\theta)\, d\mu(x) < +\infty.$$

Defina

$$\tilde{f}(x|\theta) = \begin{cases} f(x|\theta)/K_\theta, & x \in D_\theta \\ 0, & x \notin D_\theta, \end{cases}$$

em que $K_\theta = \int_{D_\theta} f(x|\theta)\, d\mu(x)$. Assim $\tilde{f}$ define uma família pseudo-exponencial. Em particular, definindo

$$g_\theta(t) = \begin{cases} \dfrac{C(\theta)}{K_\theta} \exp\{R(\theta)\cdot t\}; & \|t\| \leq \|R(\theta)\| \\ 0; & \text{c.c.}, \end{cases}$$

temos que $f_\theta(x) = h(x)\, g_\theta(T(x))$ e, pela fatorização de Fisher-Neyman, T é suficiente para θ.

3.21 *(Lauritzen (1996)) Consider $X = (x_{ij})$ and $\Sigma = (\sigma_{ij})$ symmetric positive-definite $m \times m$ matrices. The *Wishart distribution*, $\mathcal{W}_m(\alpha, \Sigma)$, is defined by the density

$$p_{\alpha,\Sigma}(X) = \frac{|X|^{\frac{\alpha-(m+1)}{2}} \exp(-\mathrm{tr}(\Sigma^{-1}X)/2)}{\Gamma_m(\alpha)|\Sigma|^{\alpha/2}},$$

with $\mathrm{tr}(A)$ the trace of A and

$$\Gamma_m(\alpha) = 2^{\alpha m/2}\pi^{m(m-1)/4}\prod_{i=1}^{m}\Gamma\left(\frac{\alpha-i+1}{2}\right).$$

a. Show that this distribution belongs to an exponential family. Give its natural representation and derive the mean of $\mathcal{W}_m(\alpha, \Sigma)$.

b. Show that, if $z_1, \ldots, z_n \sim \mathcal{N}_m(0, \Sigma)$,

$$\sum_{i=1}^{n} z_i z_i' \sim \mathcal{W}_m(n, \Sigma).$$

a) Defina $C(\theta) = [\Gamma_m(\alpha)|\Sigma|^{\alpha/2}]^{-1}$

$h(x) = |X|^{-\frac{(m+1)}{2}}$

$R(\theta).T(x) = \alpha.\log|X| - tr(\Sigma^{-1}X)/2$

$= \alpha.\log|x| - \sum_{i=1}^{m}\sum_{k=1}^{m}\sigma^{ik}x_{ki}/2,$

com $\sigma^{ik} = (\Sigma^{-1})_{ik}$. Logo

$R(\theta) = (2\alpha, -\sigma^{11}, -\sigma^{12}, \ldots, -\sigma^{mm})/2 \in \mathbb{R}^{m^2+1}$

$T(x) = (\log|X|, x_{11}, x_{12}, \ldots, x_{1m}) \in \mathbb{R}^{m^2+1}$

mostra que $W_m(\alpha, \Sigma)$ é da família exponencial.

Defina $M = -\Sigma^{-1}/2$

$\theta = (\alpha, M_{11}, M_{12}, \ldots, M_{mm})$.

Assim

$$p_{\alpha,M}(x) = \frac{|M|^{\alpha/2}}{(-2)^{m\alpha/2}\Gamma_m(\alpha)}|X|^{-\frac{(m+1)}{2}}\exp\left\{\alpha\log|X| + \sum_{i,j}M_{ij}X_{ji}\right\},$$

com parâmetro natural $\theta = (\alpha, -\Sigma^{-1}/2)$

$$\psi(\theta) = \ln\left(\Gamma_m(\alpha)|\Sigma|^{\alpha/2}\right) = \ln\Gamma_m(\alpha) + \frac{\alpha}{2}\ln(|\Sigma|)$$

$$= \ln\Gamma_m(\alpha) - \frac{\alpha}{2}\ln|-2M|$$

$$\nabla \psi(\theta) = \left(\frac{\Gamma_m'(\alpha)}{\Gamma_m(\alpha)} - \ln|-2M| , \cancel{-} \frac{\alpha}{\cancel{2}} \cancel{\frac{|-2M|(-2M)^{-1}}{|-2M|}} \cdot (-\cancel{2}) \right),$$

logo $E_\theta[X] = \alpha\Sigma$.

b) ver Gupta e Nagar (1999), página 88. Aqui vai um rascunho. Defina $(Z)_{ij} = (z_j)_i$, isto é, $Z = [z_1, \dots, z_n]$. Assim $S = \sum_{i=1}^n z_i z_i^T = ZZ^T$ supõe $n \geq m$

- Escreva $Z = TH$, em que T é triangular inferior com diagonal positiva e $HH^T = I$. Essa escrita é única (Teo. 1.2.15)
- A transformação $X \mapsto (T, H)$ é bijetiva com Jacobiano
$$J = \prod_{i=1}^m t_{ii}^{n-i} g_{n,m}(H).$$
- $\displaystyle\int_{HH^T=I} g_{n,m}(H)\,dH = \frac{2^m \pi^{\frac{1}{2}nm}}{\Gamma_m(\frac{1}{2}n)}$, logo a distribuição de T pode ser obtida.
- Por fim $ZZ^T = THH^TT^T = TT^T$ e a transformação $T \mapsto TT^T$ tem Jacobiano
$$J = (2^m \prod_{i=1}^m t_{ii}^{m-i+1})^{-1}$$

que implica a transformação de $S \sim W(n, \Sigma)$.

c) Já vimos que $E[X|\alpha, \Sigma] = \alpha\Sigma$. Agora $Cov(X)$ é definida pela Hessiana de ψ. Como
$$\nabla_M \psi(\theta) = \alpha(-2M)^{-1},$$
temos que
$$\nabla_{M,M} \psi(\theta) = \alpha(-2M)^{-1} \otimes (-2M)^{-1} \cdot (-1)(-2)$$
$$= 2\alpha\,\Sigma \otimes \Sigma.$$

3.35 Proposition 3.3.13 exhibits a conjugate family for every exponential family of the form (3.3.4),

$$\pi(\theta|\lambda,\mu) = \exp\{\theta\cdot\mu - \lambda\psi(\theta)\}K(\mu,\lambda).$$

a. Show that the distribution (3.3.4) is actually well defined when $\lambda > 0$ and $(\mu/\lambda) \in \overset{\circ}{N}$.

$$f(x|\theta) = h(x)\, e^{\theta\cdot x - \psi(\theta)}, \text{ com respeito a } \mu.$$

$$N = \left\{ \theta \in \Theta \,\middle|\, \int_{\mathcal{X}} e^{\theta\cdot x} h(x)\, d\mu(x) < +\infty \right\}$$

Como $\int_{\mathcal{X}} f(x|\theta)\, d\mu(x) = 1$, temos que

$$\psi(\theta) = \ln \int_{\mathcal{X}} e^{\theta\cdot x} h(x)\, d\mu(x)$$

Em particular, $N = \{\theta \in \Theta : \psi(\theta) < +\infty\}$. Queremos provar que

$$\int_N e^{\theta\cdot\mu - \lambda\psi(\theta)}\, d\theta < +\infty;$$

assumindo que $\lambda > 0$ e $(\mu/\lambda) \in \overset{\circ}{N}$. Tome $\theta \in N$. Seja $A \subseteq \mathcal{X}$ compacto com $\mu(A) > 0$. Denote $\mu_A(\cdot) = \mu(A \cap \cdot)/\mu(A)$. Dado $\theta \in N$, seja $c = \min_A e^{z\cdot\theta} > 0$, pois A é compacto. Assim $\mu(A) = \int_A d\mu(x) \leq \frac{1}{c}\int_A e^{x\cdot\theta}\, d\mu(x) \leq \frac{1}{c}\int_{\mathcal{X}} e^{x\cdot\theta}\, d\mu(x) < +\infty.$

Com isso,

$$e^{-\psi(\theta)} = \left[\int_{\mathcal{X}} e^{\theta\cdot x} h(x)\, d\mu(x)\right]^{-1} \leq \left[\int_A e^{\theta\cdot x + \ln(h(x))}\, d\mu(x)\right]^{-1}$$

$$= \mu(A)^{-1}\left[\int_{\mathcal{X}} e^{\theta\cdot x + \ln(h(x))}\, d\mu_A(x)\right]^{-1}$$

$$\leq \mu(A)^{-1} \exp\left\{-\theta \underbrace{\int_{\mathcal{X}} x\, d\mu_A(x)}_{y_A} - \overbrace{\int_{\mathcal{X}} \ln(h(x))\, d\mu_A(x)}^{c_A}\right\}$$

$$= \mu(A)^{-1} \exp\{-\theta\, y_A - c_A\}$$

Suponha Θ convexo. É fácil ver que $\mathcal{N}$ também será. Logo $\mu/\lambda = \sum_{j=1}^{K+1} t_j x_j$ para $\sum_{j=1}^{K+1} t_j = 1$ e $t_j \geq 0$. Como está no interior, podemos assumir que $t_j > 0$. Em particular, existe A_j compacto tal que

$$x_j = \int x \, d\mu_{A_j}(x) \rightsquigarrow \text{pode-se provar}$$

Seja $\mathcal{N}_K = \mathcal{N} \cap \{\theta \mid \theta \cdot x_K = \max_j \theta \cdot x_j\}$. Assim:

$$\int_{\mathcal{N}} e^{\theta\cdot\mu - \lambda\psi(\theta)} d\theta \leq \sum_{K=1}^{K+1} \int_{\mathcal{N}_K} e^{\theta\cdot\mu - \lambda\psi(\theta)} d\theta$$

$$\leq \sum_{K=1}^{K+1} \frac{e^{-\lambda C_{A_K}}}{\mu(A_K)} \int_{\Theta_K} e^{\lambda\theta\cdot\left(\frac{\mu}{\lambda} - \overset{x_K}{\overset{\shortparallel}{y_{A_K}}}\right)} d\theta$$

$$\leq C \sum_{K=1}^{K+1} \int_{\Theta_K} \exp\Big\{\lambda \sum_{j=1}^{K+1} t_j \overbrace{\theta\cdot(x_j - x_K)}^{\leq 0 \text{ em } \Theta_K}\Big\} d\theta$$

$$< +\infty$$

b. Give the constant K for normal, gamma, and negative binomial distributions.

b) Contas. Primeiro encontramos a representação natural da distribuição. Então calculamos

$$\left[\int_{\Theta} \exp\{\theta\cdot\mu - \lambda\psi(\theta)\} d\theta\right]^{-1}$$

c. Deduce that the likelihood function $\ell(\theta|x)$ is a particular prior distribution for exponential families (by mean of a reparameterization) and give the corresponding prior for the above families.

$$\ell(\theta|x) \propto \exp\{\theta\cdot x - \psi(\theta)\}$$

Quando $\lambda = 1$ e $\mu = x$, se $x \in \mathring{\mathcal{N}}$, vale (a)

d. Is this property characterizing exponential families? Give a counter-example.

Seja $f(x|\theta) = \frac{1}{2\theta} \mathbb{1}[x \in [-\theta, \theta]]$.

Uma conjugada nesse caso é Pareto (α, λ),

$$\pi(\theta|\alpha,\lambda) = \begin{cases} \frac{\alpha\lambda^2}{\theta^{\alpha+1}}, & \theta \geqslant \lambda \\ 0, & \theta < \lambda \end{cases}$$

Logo $\pi(\theta|x,\alpha,x_m) \propto \begin{cases} \frac{\alpha\lambda^2}{2\theta^{\alpha+2}}, & \theta \geqslant \max\{|x|, \lambda\} \\ 0, & \text{c.c.} \end{cases}$.

Logo $\theta|x \sim \text{Pareto}(\alpha+1, \max\{|x|, \lambda\})$

Note que a verossimilhança é Pareto$(0, |x|)$ que é distribuição imprópria.