Família Exponencial

Tópicos de Estatística

José Passos

ISEG-UTL

29 de Outubro de 2011

- Alguns dos modelos paramétricos estudados até aqui são casos particulares de uma classe mais geral de modelos designada família exponencial.
- Esta família tem um conjunto de propriedades que são extremamente úteis.

Definição: família exponencial com um parâmetro

Dizemos que $\{f(x \mid \theta) : \theta \in \Theta\}$ é uma família exponencial se existem funções $\eta(\theta)$, $c(\theta)$, $t(x)$ e $h(x)$ tal que,

$$f(x \mid \theta) = h(x)c(\theta)\exp\{\eta(\theta)t(x)\}$$

onde $\theta \in \Theta$, $\{x : f(x|\theta) > 0\}$ não depende de θ, $h(x) \geq 0$, $t(x)$ não pode depender de θ, $c(\theta) \geq 0$ e $\eta(\theta)$ não pode depender de x.

Definição: família exponencial com um ou mais parâmetros

Dizemos que $\{f(x \mid \theta) : \theta \in \Theta\}$ é uma família exponencial se existem funções $\eta(\theta)$, $c(\theta)$, $t(x)$ e $h(x)$ tal que,

$$f(x \mid \theta) = h(x)c(\theta)\exp\left\{\sum_{j=1}^{k} \eta_j(\theta)t_j(x)\right\}$$

onde $x \in \chi$ (que pode ser discreto ou contínuo), $\theta \in \Theta$, o suporte não depende de θ, $h(x) \geq 0$, $t_j(x)$ são funções reais dependentes apenas de x (não podem depender de θ), $c(\theta) \geq 0$ e $\eta_j(\theta)$ são funções reais dependentes apenas de θ (não podem depender de x).

Notas:

- Se $f(x \mid \theta)$ é da família exponencial o suporte, $\{x : f(x \mid \theta) > 0\}$, não depende de θ;
- Esta família tem como casos particulares a normal, exponencial, gama, qui-quadrado, beta, binomial, Poisson, geométrica, etc.;
- Para verificar se uma família de funções densidade ou funções probabilidade está na família exponencial, temos que identificar as funções $h(x)$, $c(\theta)$, $\eta_j(\theta)$ e $t_j(x)$ e mostrar que a família se pode expressar na forma anterior;
- As k estatísticas $t_1, \ldots, t_k$ são suficientes.

- Se $\dim(\theta) = k$ diz-se que a família exponencial é completa
- Se $\dim(\theta) > k$ os parâmetros não são identificáveis
- Se $\dim(\theta) < k$ diz-se que a família exponencial é curvada
- Numa família exponencial curvada o estimador da máxima verosimilhança não é estatística suficiente pelo que a inferência deve ser condicionada numa estatística ancilária.
- Qualquer família paramétrica que seja suficientemente lisa (smooth) pode ser aproximada localmente, em torno do verdadeiro valor do parâmetro por uma família exponencial curvada.

Exemplos:

- Distribuição binomial:

$$
\begin{aligned}
f(x \mid p) &= \binom{n}{x} p^x (1-p)^{n-x} \\
&= \binom{n}{x} (1-p)^n \left(\frac{p}{1-p}\right)^x \\
&= \binom{n}{x} (1-p)^n \exp\left\{\log\left(\frac{p}{1-p}\right) x\right\}.
\end{aligned}
$$

com $k = 1$, $h(x) = \binom{n}{x}$, $c(p) = (1-p)^n$, $\eta_1(p) = \log\left(\frac{p}{1-p}\right)$ e $t_1(x) = x$.

- Distribuição Poisson:

$$
\begin{aligned}
f(x \mid \theta) &= \frac{e^{-\theta}\theta^x}{x!} \\
&= \frac{1}{x!} e^{-\theta} \exp\{x \log \theta\}.
\end{aligned}
$$

com $k = 1$, $h(x) = 1/x!$, $c(\theta) = e^{-\theta}$, $\eta_1(\theta) = \log\theta$ e $t_1(x) = x$.

- Distribuição Normal:

$$f(x \mid \mu, \sigma^2) = (2\pi\sigma^2)^{-1/2} \exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\}$$
$$= (2\pi\sigma^2)^{-1/2} \exp\left\{-\frac{\mu^2}{2\sigma^2}\right\} \exp\left\{\frac{\mu}{\sigma^2}x - \frac{1}{2\sigma^2}x^2\right\}.$$

com $k = 2,\ h(x) = 1,\ c(\mu, \sigma^2) = (2\pi\sigma^2)^{-1/2} \exp\left\{-\frac{\mu^2}{2\sigma^2}\right\}$, $\eta_1(\mu, \sigma^2) = \mu/\sigma^2$, $\eta_2(\mu, \sigma^2) = -1/(2\sigma^2)$, $t_1(x) = x$ e $t_2(x) = x^2$.

Teorema: momentos

Se X é v.a. com f.d.p. ou f.p. na família exponencial, tem-se,

$$E\left(\sum_{j=1}^{k}\frac{\partial\eta_j(\theta)}{\partial\theta_l}t_j(X)\right)=-\frac{\partial}{\partial\theta_l}\log c(\theta)$$

$$Var\left(\sum_{j=1}^{k}\frac{\partial\eta_j(\theta)}{\partial\theta_l}t_j(X)\right)=-\frac{\partial^2}{\partial\theta_l^2}\log c(\theta)-E\left(\sum_{j=1}^{k}\frac{\partial^2\eta_j(\theta)}{\partial\theta_l^2}t_j(X)\right)$$

Nota: a vantagem deste resultado está no facto de os momentos poderem ser calculados com recurso a diferenciação, em vez de integração.

Exemplo: distribuição binomial,

- Para calcular $E(X)$, tem-se, do exemplo anterior,

$$\frac{d}{dp}\eta_1(p) = \frac{d}{dp}\log\frac{p}{1-p} = \frac{1}{p(1-p)}$$
$$\frac{d}{dp}\log c(p) = \frac{d}{dp}n\log(1-p) = -\frac{n}{1-p}$$

e portanto,

$$E\left(\frac{1}{p(1-p)}X\right) = \frac{n}{1-p}$$

donde se tira $E(X) = np$.

Exemplo (cont.):

- Para calcular $Var(X)$, tem-se,

$$\frac{d^2}{dp^2}\eta_1(p) = \frac{2p-1}{p^2(1-p)^2}$$
$$\frac{d^2}{dp^2}\log c(p) = -\frac{n}{(1-p)^2}$$

e portanto,

$$Var\left(\frac{1}{p(1-p)}X\right) = \frac{n}{(1-p)^2} - E\left(\frac{2p-1}{p^2(1-p)^2}X\right)$$

donde se tira $Var(X) = np(1-p)$.

Definição: família exponencial com parâmetrização natural

A família exponencial é por vezes parametrizada do seguinte modo,

$$f(x \mid \eta) = h(x) c^*(\eta) \exp\left\{ \sum_{j=1}^{k} \eta_j t_j(x) \right\}$$

O conjunto

$$\aleph = \left\{ \eta = (\eta_1, \ldots, \eta_k) : \int_{\chi} h(x) \exp(\sum_{j=1}^{k} \eta_j t_j(x)) dx < \infty \right\}$$

designa-se espaço parâmetro natural da família.

Exemplo: distribuição binomial:

- Espaço parâmetro: $\Theta = \{p : \quad 0 < p < 1\}$
- Espaço parâmetro natural: $\aleph = \{\eta_1 : \quad -\infty < \eta_1 < +\infty\}$ onde $\eta_1 = \eta_1(p) = \log \frac{p}{1-p}$. Na parametrização natural tem-se,

$$f(x \mid \eta_1) = \binom{n}{x} (1 + e^{\eta_1})^{-n} \exp\{\eta_1 x\}.$$

Exemplo: distribuição normal:

- Espaço parâmetro:
  $\Theta = \Big\{(\mu, \sigma^2) : \ -\infty < \mu < +\infty, \ 0 < \sigma^2 < +\infty\Big\}$
- Espaço parâmetro natural:
  $\aleph = \{(\eta_1, \eta_2) : \ -\infty < \eta_1 < +\infty, \ -\infty < \eta_2 < 0\}$ onde
  $\eta_1 = \mu/\sigma^2$ e $\eta_2 = -1/(2\sigma^2)$

Corolário:

Se X é v.a. com f.d.p. ou f.p. na família exponencial com parâmetro natural $\eta_1 = \eta_1(\theta)$, tem-se (para simplificar e sem perda de generalidade vamos fazer $k = 1$ e $\dim(\theta) = 1$),

$$E\left(t_1(X)\right) = -\frac{d}{d\eta_1}\log c(\eta_1)$$
$$Var\left(t_1(X)\right) = -\frac{d^2}{d\eta_1^2}\log c(\eta_1)$$

Nota: a parametrização natural simplifica o cálculo dos momentos da estatística $t_1(X)$.

- Dada uma amostra casual de dimensão n, a f.d.p. ou f.p. $f(x_1, \ldots, x_n)$, está na família exponencial se,

$$f(x_1, \ldots, x_n \mid \theta) = c(\theta)^n \left(\prod_{i=1}^{n} h(x_i) \right) \exp\left\{ \sum_{j=1}^{k} \eta_j(\theta) \sum_{i=1}^{n} T_j(x_i) \right\}$$

- Fazendo $T_j(x) = \sum_i^n T_j(x_i)$, as v.a.'s $T_j(X)$, $j = 1, \ldots, k$ designam-se estatísticas naturais,

$$f(x_1, \ldots, x_n \mid \theta) = c(\theta)^n \left(\prod_{i=1}^{n} h(x_i) \right) \exp\left\{ \sum_{j=1}^{k} \eta_j(\theta) T_j(x) \right\}$$

Teorema: distribuição conjunta de $T_j(X)$

A distribuição conjunta de $T_1(X), \ldots, T_k(X)$ é da família exponencial com parâmetro natural $\eta_1(\theta), \ldots, \eta_k(\theta)$

Bibliografia

- Casella and Berger (2002), *Statistical Inference*, 2nd Edition, Duxbury (pags. 111-116).