خانواده نمایی توزیع ها

Exponential Family of Distributions

گروه دانش آماری

http://MiladZaman005.blogfa.com

نویسنده : سید جمال میرکمالی

اسفند ۸۵

معرفی خانواده نمایی

بعضی از توابع چگالی (احتمال) دارای یک شکل مشترک هستند که می توان آنها را در یک خانواده کلی معرفی کرد و خواص آنها را از خواص این خانواده کلی نتیجه گرفت. یکی از این خانواده ها ، خانواده نمایی می باشد.

تعریف: خانواده توابع چگالی (یا توابع احتمال) $\{ f_{\theta}(x) : \boldsymbol{\theta} = (\theta_1, \theta_2, \ldots, \theta_k) \in \boldsymbol{\Theta} \subseteq R^k \}$ را عضو خانواده نمایی توزیع ها (خانواده توزیع های نمایی k پارامتری) گوییم هرگاه بتوان $f_{\theta}(x)$ را به ازای هر $\boldsymbol{\theta} \in \boldsymbol{\Theta}$ و $x \in S_X^*$ (تکیه گاه X) به صورت زیر نوشت:

$$f_{\theta}(x) = \exp\{\sum_{j=1}^{k} c_j(\boldsymbol{\theta}) d_j(x) + Q(\boldsymbol{\theta})\} h(x) \qquad (1)$$

که

۱- تکیه گاه X به هیچ یک از مقادیر پارامتر $\boldsymbol{\theta} = (\theta_1, \theta_2, \ldots, \theta_k)$ بستگی ندارد.

۲- $j = 1, 2, \ldots, k \quad c_j(\boldsymbol{\theta})$ ها توابعی غیر صفر و پیوسته از $\boldsymbol{\theta} = (\theta_1, \theta_2, \ldots, \theta_k)$ هستند.

۳- در حالت پیوسته $j = 1, 2, \ldots, k \quad d_j(x)$ توابع پیوسته روی S_X^* هستند و در حالت گسسته توابعی غیر صفر روی S_X^* می باشند.

٤- در حالت پیوسته $h(x)$ تابعی پیوسته روی S_X^* است.

خانواده $f_{\theta}(x)$ در (۱) را خانواده نمایی پر رتبه با بعد k گوییم هرگاه علاوه بر ۴ شرط فوق :

الف - توابع $j = 1, 2, \ldots, k \quad c_j(\boldsymbol{\theta})$ به طور تابعی از یکدیگر مستقل باشند.

ب- هیچ یک از $j = 1, 2, \ldots, k \quad d_j(x)$ تابعی خطی از بقیه نباشد(مستقل خطی باشند).

مثال : اگر $c_1(\theta)=\theta$ و $c_2(\theta)=\theta^2/2$ در این صورت این دو ، به طور تابعی وابسته هستند به این معنی که با داشتن یکی ، مقدار دیگری معلوم خواهد بود. اما اگر $d_1(x)=x$ و $d_2(x)=x^2$ این دو ، مستقل خطی هستند به این معنی که یکی را نمی توان به طور تابعی خطی از دیگری نوشت.

مثال: خانواده توزیع های زیر به یک خانواده نمایی یک پارامتری متعلق اند:

a) $Ber(\theta) \quad \theta\in(0,1)$
b) $Bin(n,\theta) \quad \theta\in(0,1)$, n معلوم
c) $NB(n,\theta) \quad \theta\in(0,1)$, n معلوم
d) $Ge(\theta) \quad \theta\in(0,1)$
e) $P(\theta) \quad \theta\in(0,\infty)$
f) $N(\mu,1) \quad \mu\in R$
g) $N(1,\sigma^2) \quad \sigma>0$
h) $N(\theta,\theta) \quad \theta>0$
i) $G(\alpha,\beta) \quad \alpha>0$, β معلوم
j) $G(\alpha,\beta) \quad \beta>0$, α معلوم
k) $Beta(\alpha,\beta) \quad \alpha>0$, β معلوم
l) $Beta(\alpha,\beta) \quad \beta>0$, α معلوم

مثلا برای گزینه های k و h داریم:

k) $X\sim Beta(\alpha,\beta) \quad \alpha>0$, β معلوم

$$f_\alpha(x)=\frac{1}{B(\alpha,\beta)}x^{\alpha-1}(1-x)^{\beta-1}=e^{(\alpha-1)\ln x-\ln B(\alpha,\beta)}(1-x)^{\beta-1}$$

$$c(\alpha)=\alpha-1 \qquad d(x)=\ln x \qquad Q(\alpha)=-\ln B(\alpha,\beta) \qquad h(x)=(1-x)^{\beta-1}$$

h) $X\sim N(\theta,\theta) \quad \theta>0$

$$f_\theta(x)=\frac{1}{\sqrt{2\pi\theta}}e^{-\frac{1}{2\theta}(x-\theta)^2}=e^{-\frac{1}{2\theta}x^2-\frac{1}{2}(\theta+\ln(2\pi\theta))}.e^x$$

$$c(\theta)=-\frac{1}{2\theta} \qquad d(x)=x^2 \qquad Q(\theta)=-\frac{1}{2}(\theta+\ln(2\pi\theta)) \qquad h(x)=e^x$$

مثال : خانواده توزیع های زیر به یک خانواده نمایی دو پارامتری پر رتبه متعل اند:

a) $N(\mu,\sigma^2) \quad \mu\in R,\sigma>0$
b) $G(\alpha,\beta) \quad \alpha,\beta>0$

c) $Beta(\alpha,\beta) \quad \alpha,\beta>0$

برای مثال در قسمت a داریم:

$$a)\ X\sim N(\mu,\sigma^2) \quad \mu\in R,\sigma>0$$

$$f_{\mu,\sigma^2}(x)=\frac{1}{\sqrt{2\pi\sigma^2}}e^{-\frac{1}{2\sigma^2}(x-\mu)^2}=e^{-\frac{1}{2\sigma^2}x^2+\frac{\mu}{\sigma^2}x-\frac{1}{2}(\frac{\mu^2}{\sigma^2}+\ln(2\pi\sigma^2))}$$

$$c_1(\mu,\sigma^2)=-\frac{1}{2\sigma^2} \qquad c_2(\mu,\sigma^2)=\frac{\mu}{\sigma^2} \qquad d_1(x)=x^2 \qquad d_2(x)=x$$

مثال : خانواده $\theta>0 \quad X\sim N(\theta,\theta^2)$ به یک خانواده نمایی متعلق است اما پر رتبه نیست زیرا:

$$f_\theta(x)=\frac{1}{\sqrt{2\pi\theta^2}}e^{-\frac{1}{2\theta^2}(x-\theta)^2}=e^{-\frac{1}{2\theta^2}x^2+\frac{x}{\theta}-\frac{1}{2}\ln(2\pi\theta^2)}.e^{-\frac{1}{2}}$$

$$c_1(\theta)=-\frac{1}{2\theta^2} \qquad c_2(\theta)=\frac{1}{\theta}$$

شرط الف برای پر رتبه بودن برقرار نیست.

مثال : خانواده توزیع های زیر به یک خانواده نمایی متعلق نیستند:

a) $U(0,\theta) \quad \theta>0$ شرط ۱ برقرار نیست

b) $Exp(\alpha,\beta) \quad \alpha\in R,\beta>0$ شرط ۱ برقرار نیست

خواص خانواده نمایی

۱- اگر $X_1,X_2,\dots,X_n\overset{iid}{\sim} f_\theta$ متعلق به خانواده نمایی باشند آنگاه توزیع توام این نمونه تصادفی به خانواده نمایی تعلق دارد

$$f_\theta(\boldsymbol{x})=\prod_{i=1}^{n}\exp\{\sum_{j=1}^{k}c_j(\boldsymbol{\theta})d_j(x_i)+Q(\boldsymbol{\theta})\}h(x_i)$$

$$=\exp\{\sum_{j=1}^{k}\left[c_j(\boldsymbol{\theta})\left(\sum_{i=1}^{n}d_j(x_i)\right)\right]+nQ(\boldsymbol{\theta})\}\prod_{i=1}^{n}h(x_i)$$

$$=\exp\{\sum_{j=1}^{k}\left[c_j(\boldsymbol{\theta})d'_j(\boldsymbol{x})\right]+Q'(\boldsymbol{\theta})\}h'(\boldsymbol{x})$$

که در آن $d'_j(x)=\sum_{i=1}^{n}d_j(x_i)$ و $Q'(\boldsymbol{\theta})=nQ(\boldsymbol{\theta})$ و $h'(\boldsymbol{x})=\prod_{i=1}^{n}h(x_i)$.

فرم کانونی خانواده نمایی: فرم کانونی یک تابع چگالی متعلق به خانواده نمایی به صورت زیر ساخته می شود:

$$f_\theta(\boldsymbol{x})=\exp\{\sum_{j=1}^{k}c_j(\boldsymbol{\theta})d_j(x)+Q(\boldsymbol{\theta})\}h(x)$$

$$\eta_j=c_j(\boldsymbol{\theta})$$

$$f_{\boldsymbol{\eta}}(\boldsymbol{x})=\exp\{\sum_{j=1}^{k}\eta_j d_j(x)+A(\boldsymbol{\eta})\}h(x)$$

در این حالت مجموعه $\left\{\boldsymbol{\eta}=(\eta_1,\eta_1,\ldots,\eta_k)\,|\int_{-\infty}^{+\infty}h(x)e^{\sum_{j=1}^{k}\eta_j d_j(x)}dx<\infty\right\}$ را فضای پارامتر طبیعی می گوییم و $\boldsymbol{\eta}$ را پارامتر طبیعی گوییم.

۲- $E_{\boldsymbol{\eta}}(d_i(X))=-\frac{\partial}{\partial\eta_i}A(\boldsymbol{\eta})$ چون:

$$1=\int e^{\sum\eta_j d_j(x)+A(\boldsymbol{\eta})}h(x)\Rightarrow\frac{\partial}{\partial\eta_i}1=\frac{\partial}{\partial\eta_i}\int e^{\sum\eta_j d_j(x)+A(\boldsymbol{\eta})}h(x)$$

$$\Rightarrow 0=\int\frac{\partial}{\partial\eta_i}e^{\sum\eta_j d_j(x)+A(\boldsymbol{\eta})}h(x)=\int\left[d_i(x)+\frac{\partial}{\partial\eta_i}A(\boldsymbol{\eta})\right]e^{\sum\eta_j d_j(x)+A(\boldsymbol{\eta})}h(x)=$$

$$=\int d_i(x)e^{\sum\eta_j d_j(x)+A(\boldsymbol{\eta})}h(x)+\int\frac{\partial}{\partial\eta_i}A(\boldsymbol{\eta})e^{\sum\eta_j d_j(x)+A(\boldsymbol{\eta})}h(x)$$

$$\Rightarrow 0=E(d_i(X))+\frac{\partial}{\partial\eta_i}A(\boldsymbol{\eta})\Rightarrow E(d_i(X))=-\frac{\partial}{\partial\eta_i}A(\boldsymbol{\eta})$$

اگر $f_\theta(x)=\exp\{c(\theta)d(x)+Q(\theta)\}h(x)$ باشد ، آنگاه $E_{\boldsymbol{\eta}}(d(X))=-\frac{Q'(\theta)}{c'(\theta)}$.

۳- $\operatorname{cov}_{\boldsymbol{\eta}}(d_i(X),d_j(X))=-\frac{\partial^2}{\partial\eta_i\partial\eta_j}A(\boldsymbol{\eta})$ چون:

$$1=\int e^{\sum\eta_t d_t(x)+A(\boldsymbol{\eta})}h(x)\Rightarrow\frac{\partial^2}{\partial\eta_i\partial\eta_j}1=\frac{\partial^2}{\partial\eta_i\partial\eta_j}\int e^{\sum\eta_t d_t(x)+A(\boldsymbol{\eta})}h(x)$$

$$\Rightarrow 0=\int\frac{\partial^2}{\partial\eta_i\partial\eta_j}e^{\sum\eta_t d_t(x)+A(\boldsymbol{\eta})}h(x)=\int\frac{\partial}{\partial\eta_j}\left[d_i(x)+\frac{\partial}{\partial\eta_i}A(\boldsymbol{\eta})\right]e^{\sum\eta_t d_t(x)+A(\boldsymbol{\eta})}h(x)=$$

$$=\int\left\{\frac{\partial^2}{\partial\eta_i\partial\eta_j}A(\boldsymbol{\eta})+\left[d_i(x)+\frac{\partial}{\partial\eta_i}A(\boldsymbol{\eta})\right]\left[d_j(x)+\frac{\partial}{\partial\eta_j}A(\boldsymbol{\eta})\right]\right\}e^{\sum\eta_t d_t(x)+A(\boldsymbol{\eta})}h(x)=$$

$$= \frac{\partial^{۲}}{\partial\eta_i\partial\eta_j}A(\boldsymbol{\eta}) + E(d_i(X),d_j(X)) - ۲E(d_i(X))E(d_j(X)) + E(d_i(X))E(d_j(X))$$

$$= \frac{\partial^{۲}}{\partial\eta_i\partial\eta_j}A(\boldsymbol{\eta}) + \operatorname{cov}(d_i(X),d_j(X)) \Rightarrow \operatorname{cov}(d_i(X),d_j(X)) = -\frac{\partial^{۲}}{\partial\eta_i\partial\eta_j}A(\boldsymbol{\eta})$$

٤- تابع مولد گشتاور X عبارتست از $M_X(t) = \exp\{A(\boldsymbol{\eta}) - A(\boldsymbol{\eta}+t)\}$ و بعبارتی تابع مولد انباشتک X برابر است با $K_X(t) = \ln M_X(t) = A(\boldsymbol{\eta}) - A(\boldsymbol{\eta}+t)$.

مثال: فرض کنیم $X \sim P(\theta)$ در این صورت:

$$f_\theta(x) = \frac{e^{-\theta}\theta^x}{x!} = e^{x\ln\theta-\theta}\frac{۱}{x!} = e^{x\eta+A(\eta)}\frac{۱}{x!}$$

$$\eta = \ln\theta \Leftrightarrow \theta = e^\eta \qquad A(\eta) = -\theta = -e^\eta$$

بنابر خواص ۲، ۳ و ۴ خانواده نمایی داریم:

$$E(d(X)) = E(X) = -\frac{\partial}{\partial\eta}A(\eta) = -(-e^\eta) = \theta$$

$$V(d(X)) = -\frac{\partial^{۲}}{\partial\eta^{۲}}A(\eta) = e^\eta = \theta$$

$$M_X(t) = e^{A(\eta)-A(\eta+t)} = \exp\{-e^\eta + e^{\eta+t}\} = \exp\{e^\eta(e^t-۱)\} = \exp\{\theta(e^t-۱)\}$$

مثال : در خانواده توزیع های زیر $E(X_i)$ و $\operatorname{cov}(X_i,X_j)$ را به دست آورید:

$$\boldsymbol{X} = (X_۱,X_۲,\ldots,X_n) \sim M(n,p_۱,p_۲,\ldots,p_n)$$

$$f_{\boldsymbol{p}}(\boldsymbol{x}) = \frac{n!}{\prod_{i=۱}^n x_i!}\prod_{i=۱}^n p_i^{x_i} \qquad x_i = ۰,۱,\ldots,n \quad , \quad \Sigma x_i = n \ , \ \Sigma p_i = ۱$$

$$f_{\boldsymbol{p}}(\boldsymbol{x}) = \frac{n!}{\prod_{i=۱}^n x_i!}\exp\{x_۱\ln\frac{p_۱}{p_n} + x_۲\ln\frac{p_۲}{p_n} + \ldots + x_{n-۱}\ln\frac{p_{n-۱}}{p_n} + n\ln p_n\}$$

$$\eta_j = \ln\frac{p_j}{p_n} \qquad j = ۱,\ldots,n-۱ \Rightarrow p_j = p_n e^{\eta_j} \quad j = ۱,\ldots,n-۱$$

$$\sum_{j=۱}^n p_j = ۱ \Rightarrow p_n e^{\eta_۱} + p_n e^{\eta_۲} + \ldots + p_n e^{\eta_{n-۱}} + p_n = ۱ \Rightarrow p_n = \frac{۱}{۱+e^{\eta_۱}+\ldots+e^{\eta_{n-۱}}}$$

لذا فرم کانونی تابع احتمال چندجمله ای به شکل زیر است:

$$f_{\boldsymbol{p}}(\boldsymbol{x})=\frac{n!}{\prod_{i=1}^{n}x_i!}\exp\{\sum_{j=1}^{n-1}x_j\eta_j-n\ln(1+e^{\eta_1}+\ldots+e^{\eta_{n-1}})\}=h(\boldsymbol{x})e^{\sum_{j=1}^{n-1}x_j\eta_j+A(\boldsymbol{\eta})}$$

$$E(d_i(\boldsymbol{X}))=E(X_i)=-\frac{\partial}{\partial\eta_i}A(\boldsymbol{\eta})=n\frac{e^{\eta_i}}{1+e^{\eta_1}+\ldots+e^{\eta_{n-1}}}=np_n\frac{p_i}{p_n}=np_i$$

$$V(X_i)=-\frac{\partial^2}{\partial\eta_i^2}A(\boldsymbol{\eta})=-n\frac{e^{\eta_i}(1+e^{\eta_1}+\ldots+e^{\eta_{n-1}})-e^{2\eta_i}}{(1+e^{\eta_1}+\ldots+e^{\eta_{n-1}})^2}=np_i(1-p_i)$$

$$\operatorname{cov}(X_i,X_j)=-\frac{\partial^2}{\partial\eta_i\partial\eta_j}A(\boldsymbol{\eta})=-n\frac{e^{\eta_i}e^{\eta_j}}{(1+e^{\eta_1}+\ldots+e^{\eta_{n-1}})^2}=-np_ip_j \qquad i\neq j$$

اما برای X_n شاخص ها را باید به صورت زیر محاسبه نمود:

$$X_n=n-X_1-\ldots-X_{n-1}\Rightarrow E(X_n)=n-E(X_1)-\ldots-E(X_{n-1})\Rightarrow$$
$$E(X_n)=n-np_1-\ldots-np_{n-1}=np_n$$

منبع: *مبانی آمار ریاضی* ، دکتر احمد پارسیان ، دانشگاه صنعتی اصفهان، ۱۳۷۸