Exponential Family

Single parameter exponential family

A single-parameter exponential family is a set of probability distributions whose probability density function (or probability mass function, for the case of a discrete distribution) can be expressed in the form

$$f_\theta(y) = h(y) \times \exp[\eta(\theta)T(y) - B(\theta)]$$

K-parameter exponential family

A k-parameter exponential family is a set of probability distributions whose probability density function (or probability mass function, for the case of a discrete distribution) can be expressed in the form

$$f_\theta(y) = h(y) \times \exp\left[\sum_{i=1}^{k} \eta(\theta)T(y) - B(\theta)\right]$$

Canonical Form

For an distribution in exponential family, if $\eta(\theta)$ and $T(y)$ are the identity functions, and $f_\theta(y)$ can be rewritten as

$$f_\theta(y) = \exp\left[\frac{y\theta - b(\theta)}{\phi} + c(y, \phi)\right]$$

where $\phi, b(.)$ and $c(.)$ are known

then θ is called the canonical parameter (or natural parameter).

Bernoulli(p)

The probability density function is

$f_p(y) = p^y(1-p)^{1-y}$

Rewrite $f_p(y)$ as

$$\begin{aligned}
f_p(y) &= p^y(1-p)^{1-y} \\
&= \exp(ln[p^y(1-p)^{1-y}]) \\
&= \exp[y \times ln(p) + (1-y) \times ln(1-p)] \\
&= \exp[y \times ln(p) + ln(1-p) - y \times ln(1-p)] \\
&= 1 \times \exp\left[y \times ln\left(\frac{p}{1-p}\right) + ln(1-p)\right] \\
&= 1 \times \exp[\theta \times y - ln(1 + e^\theta)]
\end{aligned}$$

where

$$\theta = ln\left(\frac{p}{1-p}\right),\ p = \frac{e^\theta}{1+e^\theta}$$

$h(y) = 1$

$\eta(\theta) = \theta$

$T(y) = y$

$B(\theta) = ln(1 + e^\theta)$

Canonical form

$$f_\theta(y) = \exp[\theta y - ln(1+e^\theta)]$$
$$= \exp\left[\frac{\theta y - ln(1+e^\theta)}{1} + 0\right]$$
$$= \exp\ [\frac{y\theta - b(\theta)}{\phi} + c(y,\phi)]$$

where

$b(\theta) = ln(1+e^\theta)$

$\phi = 1$

$c(y,\phi) = 0$

Binomial(n, p)

The probability density function is

$f_p(y) = \binom{n}{y} p^y (1-p)^{1-y}$

Rewrite $f_p(y)$ as

$$f_p(y) = \binom{n}{y} p^y (1-p)^{n-y}$$

$$= \binom{n}{y} \exp(ln[p^y (1-p)^{n-y}])$$

$$= \binom{n}{y} \exp[y \times ln(p) + (n-y) \times ln(1-p)]$$

$$= \binom{n}{y} \exp[y \times ln(p) + n \times ln(1-p) - y \times ln(1-p)]$$

$$= \binom{n}{y} \times \exp\left[y \times ln\left(\frac{p}{1-p}\right) + n \times ln(1-p)\right]$$

$$= \binom{n}{y} \times \exp[\theta \times y - n \times ln(1+e^\theta)]$$

where

$\theta = \ln\left(\frac{p}{1-p}\right),\ p = \frac{e^\theta}{1+e^\theta}$

$h(y) = \binom{n}{y}$

$\eta(\theta) = \theta$

$T(y) = y$

$B(\theta) = n \times ln(1+e^\theta)$

Canonical form

$$f_\theta(y) = \binom{n}{y} \times \exp[\theta \times y - n \times ln(1+e^\theta)]$$

$$= \exp\left[\frac{\theta y - n \times ln(1+e^{\theta})}{1} + ln\left(\binom{n}{y}\right)\right]$$

$$= \exp\left[\frac{y\theta - b(\theta)}{\phi} + c(y,\phi)\right]$$

where

$b(\theta) = n \times ln(1+e^{\theta})$

$\phi = 1$

$c(y,\phi) = ln\left(\binom{n}{y}\right)$

Exponential(λ)

The probability density function is

$f_\lambda(y) = \lambda e^{-\lambda y}$

Rewrite $f_\lambda(y)$ as

$$f_\lambda(y) = exp\ [ln(\lambda e^{-\lambda y})]$$

$$= 1 \times exp[-\lambda y + ln(\lambda)]$$

$$= 1 \times exp[\theta y + ln(-\theta)]$$

where

$\theta = -\lambda, \lambda = -\theta$

$h(y) = 1$

$\eta(\theta) = \theta$

$T(y) = y$

$B(\theta) = -\ln(-\theta)$

Canonical form

$$f_\theta(y) = exp[\theta y + ln(-\theta)]$$

$$= \exp\left[\frac{\theta y + ln(-\theta)}{1} + 0\right]$$

$$= \exp\left[\frac{y\theta - b(\theta)}{\phi} + c(y,\phi)\right]$$

where

$b(\theta) = -ln(-\theta)$

$\phi = 1$

$c(y,\phi) = 0$

Poisson(λ)

The probability density function is

$$f_\lambda(y) = \frac{\lambda^y e^{-\lambda}}{y!}$$

Rewrite $f_\lambda(y)$ as

$$f_\lambda(y) = \frac{1}{y!}\exp\left[ln\left(\lambda^y e^{-\lambda}\right)\right]$$

$$= \frac{1}{y!}\exp[yln(\lambda) - \lambda)]$$

$$= \frac{1}{y!} \times \exp\left(\theta y - e^\theta\right)$$

where

$$\theta = ln(\lambda), \lambda = e^\theta$$

$$h(y) = \frac{1}{y!}$$

$$\eta(\theta) = \theta$$

$$T(y) = y$$

$$B(\theta) = e^\theta$$

Canonical form

$$f_\theta(y) = \frac{1}{y!} \times \exp\left(\theta y - e^\theta\right)$$

$$= \exp\left[\frac{\theta y - e^\theta}{1} - \ln(y!)\right]$$

$$= \exp\left[\frac{y\theta - b(\theta)}{\phi} + c(y, \phi)\right]$$

where

$$b(\theta) = e^\theta$$

$$\phi = 1$$

$$c(y, \phi) = -\ln(y!)$$

Normal(μ, σ^2)

The probability density function is

$$f_{\mu,\sigma^2}(y) = \frac{1}{\sigma\sqrt{2\pi}}\exp\left[-\frac{1}{2\sigma^2}(y - \mu)^2\right]$$

Rewrite $f_{\mu,\sigma^2}(y)$ as

$$f_\lambda(y) = \frac{1}{\sigma\sqrt{2\pi}}\exp\left[-\frac{1}{2\sigma^2}(y - \mu)^2\right]$$

$$= \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{1}{2\sigma^2}(y^2 + \mu^2 - 2\mu y) - \ln(\sigma)\right]$$
$$= \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{y^2}{2\sigma^2} - \frac{\mu^2}{2\sigma^2} + \frac{\mu y}{\sigma^2} - \ln(\sigma)\right]$$
$$= \frac{1}{\sqrt{2\pi}} \exp\left[\left(\frac{\mu y}{\sigma^2} - \frac{y^2}{2\sigma^2}\right) - \left(\ln(\sigma) + \frac{\mu^2}{2\sigma^2}\right)\right]$$
$$= \frac{1}{\sqrt{2\pi}} \exp\left[(\theta_1 y + \theta_2 y^2) - \left(-\frac{1}{2}\ln(-2\theta_2) - \frac{\theta_1^2}{4\theta_2}\right)\right]$$

where

$$\theta_1 = \frac{\mu}{\sigma^2}, \theta_2 = -\frac{1}{2\sigma^2}$$

$$\mu = -\frac{\theta_1}{2\theta_2}, \sigma^2 = -\frac{1}{2\theta_2}$$

$$h(y) = \frac{1}{\sqrt{2\pi}}$$

$$\eta_1(\theta_1, \theta_2) = \theta_1$$

$$T_1(y) = y$$

$$\eta_1(\theta_1, \theta_2) = \theta_2$$

$$T_2(y) = y^2$$

$$B(\mu, \sigma^2) = -\frac{1}{2}\ln(-2\theta_2) - \frac{\theta_1^2}{4\theta_2}$$

Normal(μ, σ^2), σ^2 is known

The probability density function is

$$f_\mu(y) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{1}{2\sigma^2}(y-\mu)^2\right]$$

Rewrite $f_\mu(y)$ as

$$f_\lambda(y) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{1}{2\sigma^2}(y-\mu)^2\right]$$
$$= \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{1}{2\sigma^2}(y^2 + \mu^2 - 2\mu y) - \ln(\sigma)\right]$$
$$= \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{\mu^2}{2\sigma^2} + \frac{\mu y}{\sigma^2}\right) \times \exp\left(-\frac{y^2}{2\sigma^2} - \ln(\sigma)\right)$$
$$= \frac{\exp\left(-\frac{y^2}{2\sigma^2}\right)}{\sigma\sqrt{2\pi}} \exp\left(\frac{\mu y}{\sigma^2} - \frac{\mu^2}{2\sigma^2}\right)$$
$$= \frac{\exp\left(-\frac{y^2}{2\sigma^2}\right)}{\sigma\sqrt{2\pi}} \exp\left[\left(\theta y - \frac{(\theta\sigma^2)^2}{2\sigma^2}\right)\right]$$
$$= \frac{\exp\left(-\frac{y^2}{2\sigma^2}\right)}{\sigma\sqrt{2\pi}} \exp\left[\left(\theta y - \frac{\sigma^2\theta^2}{2}\right)\right]$$

where

$$\theta = \frac{\mu}{\sigma^2}, \mu = \theta\sigma^2$$

$$h(y) = \frac{\exp\left(-\frac{y^2}{2\sigma^2}\right)}{\sigma\sqrt{2\pi}}$$

$$\eta(\theta) = \theta$$

$$T(y) = y$$

$$B(\theta) = \frac{\sigma^2\theta^2}{2}$$

Canonical form

$$f_\theta(y) = \frac{\exp\left(-\frac{y^2}{2\sigma^2}\right)}{\sigma\sqrt{2\pi}} \exp\left(\frac{\mu y}{\sigma^2} - \frac{\mu^2}{2\sigma^2}\right)$$

$$= \exp\left[\frac{\mu y - \frac{1}{2}\mu^2}{\sigma^2} + \left(-\frac{y^2}{2\sigma^2}\right) - \ln\left(\sigma\sqrt{2\pi}\right)\right] = \exp\left[\frac{y\theta - b(\theta)}{\phi} + c(y,\phi)\right]$$

where

$$\theta = \mu$$

$$b(\theta) = \frac{1}{2}\theta^2$$

$$\phi = \sigma^2$$

$$c(y,\phi) = \left(-\frac{y^2}{2\sigma^2}\right) - \ln\left(\sigma\sqrt{2\pi}\right) = -\frac{1}{2}\left(\frac{y^2}{\phi}\right) - ln(2\pi\phi)^{\frac{1}{2}} = -\frac{1}{2}\left[\frac{y^2}{\phi} + ln(2\pi\phi)\right]$$