Note on Exponential Family

January 30, 2024

- A distribution with a random variable Y with a parameter θ is said to belong to an *exponential family* if $f_Y(y)$ or $P(Y = y)$ can be written as

$$h(y)c(\theta)\exp(w(\theta)t(y))$$

  - For example, if $Y \sim \text{Exponential}(\theta)$, then Y belongs to an exponential family since

$$f_Y(y) = \frac{1}{\theta}e^{-y/\theta} = \frac{1}{\theta}\exp\left(-\frac{1}{\theta}y\right) = h(y)c(\theta)\exp(w(\theta)t(y)), \quad y > 0$$

    where

$$h(y) = 1, \quad c(\theta) = \frac{1}{\theta}, \quad w(\theta) = -\frac{1}{\theta}, \quad t(y) = y$$

  - For another example, if $Y \sim \text{Binomial}(n, \theta)$ (with n fixed), then

$$\begin{aligned} P(Y = y) &= \binom{n}{y}\theta^y(1-\theta)^{n-y} = \binom{n}{y}(1-\theta)^n\left(\frac{\theta}{1-\theta}\right)^y \\ &= \binom{n}{y}(1-\theta)^n\exp\left(y\ln\left(\frac{\theta}{1-\theta}\right)\right) = \binom{n}{y}(1-\theta)^n\exp\left(\ln\left(\frac{\theta}{1-\theta}\right)y\right) \\ &= h(y)c(\theta)\exp(w(\theta)t(y)), \quad y = 0, 1, \ldots, n \end{aligned}$$

    where

$$h(y) = \binom{n}{y}, \quad c(\theta) = (1-\theta)^n, \quad w(\theta) = \ln\left(\frac{\theta}{1-\theta}\right), \quad t(y) = y$$

- If a random variable Y has a distribution with *multiple* parameters $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_k)$, then Y belongs to an exponential family if $f_Y(y)$ or $P(Y = y)$ can be written as

$$h(y)c(\boldsymbol{\theta})\exp\left(\sum_{i=1}^{k} w_i(\boldsymbol{\theta})t_i(y)\right)$$

  - For example, if $Y \sim N(\mu, \sigma^2)$, then by letting $\boldsymbol{\theta} = (\mu, \sigma^2)$, we have

$$\begin{aligned} f_Y(y) &= \frac{1}{\sqrt{2\pi}\sigma}\exp\left(-\frac{(y-\mu)^2}{2\sigma^2}\right) = \frac{1}{\sqrt{2\pi}\sigma}\exp\left(-\frac{\mu^2}{2\sigma^2}\right)\exp\left(-\frac{y^2}{2\sigma^2} + \frac{\mu y}{\sigma^2}\right) \\ &= h(y)c(\boldsymbol{\theta})\exp(w_1(\boldsymbol{\theta})t_1(y) + w_2(\boldsymbol{\theta})t_2(y)), \quad -\infty < y < \infty \end{aligned}$$

    where

$$h(y) = 1,\ c(\boldsymbol{\theta}) = \frac{1}{\sqrt{2\pi}\sigma}\exp\left(-\frac{\mu^2}{2\sigma^2}\right),\ w_1(\boldsymbol{\theta}) = \frac{-1}{2\sigma^2},\ t_1(y) = y^2,\ w_2(\boldsymbol{\theta}) = \frac{\mu}{\sigma^2},\ t_2(y) = y$$

- Not all distributions belong to an exponential family. For example, if $Y \sim \text{Uniform}(0, \theta)$, then

$$f_Y(y) = \frac{1}{\theta}, \quad 0 < y < \theta$$

We would like to rewrite the pdf $f_Y(y)$ to incorporate the region $0 < y < \theta$. For this task, we introduce an *indicator function* of set A, defined as

$$1_A(y) = \begin{cases} 1 & \text{if } y \in A \\ 0 & \text{if } y \notin A \end{cases}$$

Using the definition, it is often convenient to rewrite in the form

$$1_A(y) = 1_{\{y \in A\}}$$

So, for the pdf of $Y \sim \text{Uniform}(0, \theta)$, we can write

$$f_Y(y) = \frac{1}{\theta} 1_{(0,\theta)}(y) = \frac{1}{\theta} 1_{\{0 < y < \theta\}}$$

However, $f_Y(y)$ above *cannot* be written in the form $h(y)c(\theta)\exp(w(\theta)t(y))$, and therefore $Y \sim \text{Uniform}(0, \theta)$ *does not* belong to an exponential family.