3.4 Exponential Families

A family of pdfs or pmfs is called an exponential family if it can be expressed as

$$f(x|\boldsymbol{\theta}) = h(x)c(\boldsymbol{\theta}) \exp\left(\sum_{i=1}^{k} w_i(\boldsymbol{\theta})t_i(x)\right). \tag{1}$$

Here $h(x) \geq 0$ and $t_1(x), \ldots, t_k(x)$ are real-valued functions of the observation x (they cannot depend on $\boldsymbol{\theta}$), and $c(\boldsymbol{\theta}) \geq 0$ and $w_1(\boldsymbol{\theta}), \ldots, w_k(\boldsymbol{\theta})$ are real-valued functions of the possibly vector-valued parameter $\boldsymbol{\theta}$ (they cannot depend on x). Many common families introduced in the previous section are exponential families. They include the continuous families—normal, gamma, and beta, and the discrete families—binomial, Poisson, and negative binomial.

<u>Example 3.4.1</u> (Binomial exponential family)

Let n be a positive integer and consider the binomial(n, p) family with $0 < p < 1$. Then the pmf for this family, for $x = 0, 1, \ldots, n$ and $0 < p < 1$, is

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

Define

$$h(x) = \begin{cases} \binom{n}{x} & x = 0, 1, \ldots, n \\ 0 & \text{otherwise,} \end{cases}$$

$$c(p) = (1-p)^n, \quad 0 < p < 1, \quad w_1(p) = \log\left(\frac{p}{1-p}\right), \quad 0 < p < 1,$$

and

$$t_1(x) = x.$$

Then we have

$$f(x|p) = h(x)c(p)\exp\{w_1(p)t_1(x)\}.$$

1

<u>Example 3.4.4</u> (Normal exponential family)

Let $f(x|\mu, \sigma^2)$ be the $N(\mu, \sigma^2)$ family of pdfs, where $\boldsymbol{\theta} = (\mu, \sigma^2)$, $-\infty < \mu < \infty$, $\sigma > 0$.
Then

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

$$= \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{\mu^2}{2\sigma^2}\right) \exp\left(-\frac{x^2}{2\sigma^2} + \frac{\mu x}{\sigma^2}\right).$$

Define

$$h(x) = 1 \text{ for all } x;$$

$$c(\boldsymbol{\theta}) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{\mu^2}{2\sigma^2}\right), -\infty < \mu < \infty, \sigma > 0;$$

$$w_1(\boldsymbol{\theta}) = \frac{1}{\sigma^2}, \quad \sigma > 0; \qquad w_2(\boldsymbol{\theta}) = \frac{\mu}{\sigma^2}, \quad \sigma > 0;$$

$$t_1(x) = -x^2/2; \quad \text{and} \quad t_2(x) = x.$$

Then

$$f(x|\mu, \sigma^2) = h(x)c(\mu, \sigma) \exp[w_1(\mu, \sigma)t_1(x) + w_2(\mu, \sigma)t_2(x)].$$

In general, the set of x values for which $f(x|\boldsymbol{\theta} > 0$ cannot depend on $\boldsymbol{\theta}$ in an exponential family. For example, the set of pdfs given by

$$f(x|\theta) = \theta^{-1} \exp(1 - x/\theta), \quad 0 < \theta < x < \infty,$$

is not an exponential family. We have

$$f(x|\theta) = \theta^{-1} \exp(1 - x/\theta) I_{[\theta, \infty)}(x).$$

The indicator function can not be incorporated into any of the functions of (1) since it is not a function of x alone, not a function of θ alone, and cannot be expressed as an exponential.

3.5 Location and Scale Families

<u>Theorem 3.5.1</u>

Let $f(x)$ be any pdf and let μ and $\sigma > 0$ be any given constants. Then the function

$$g(x|\mu, \sigma) = \frac{1}{\sigma} f\left(\frac{x - \mu}{\sigma}\right)$$

is a pdf.

PROOF: Since $f(x) \geq 0$ for all values of x. So, $\frac{1}{\sigma}f(\frac{x-\mu}{\sigma}) \geq 0$ for all values of x, μ and σ.
Next,

$$\int_{-\infty}^{\infty} \frac{1}{\sigma}f(\frac{x-\mu}{\sigma})dx = \int_{-\infty}^{\infty} f(y)dy = 1.$$

$\square$

Definition 3.5.2 Let $f(x)$ be any pdf. Then the family of pdfs $f(x-\mu)$, indexed by the parameter μ, $-\infty < \mu < \infty$, is called the location family with standard pdf $f(x)$ and μ is called the location parameter for the family.

The location parameter μ simply shifts the pdf $f(x)$ so that the shape of the graph is unchanged but the point on the graph that was above $x = 0$ for $f(x)$ is above $x = \mu$ for $f(x-\mu)$.

Example 3.5.3 (Exponential location family)
Let $f(x) = e^{-x}$, $x \geq 0$, and $f(x) = 0$, $x < 0$. To form a location family we replace x with $x - \mu$ to obtain

$$f(x|\mu) = \begin{cases} e^{-(x-\mu)} & x - \mu \geq 0 \\ 0 & x - \mu < 0. \end{cases}$$

That is,

$$f(x|\mu) = \begin{cases} e^{-(x-\mu)} & x \geq \mu \\ 0 & x < \mu. \end{cases}$$

In this type of model, where μ denotes a bound on the range of X, μ is sometimes called a threshold parameter.

Definition 3.5.4
Let $f(x)$ be any pdf. Then for any $\sigma > 0$, the family of pdfs $(1/\sigma)f(x/\sigma)$, indexed by the parameter σ, is called the scale family with standard pdf $f(x)$ and σ is called the scale parameter of the family.

The effect of introducing the scale parameter σ is either to stretch $(\sigma > 1)$ or to contract $(\sigma < 1)$ the graph of $f(x)$ while still maintaining the same basic shape of the graph.

3

<u>Definition 3.5.5</u>

Let $f(x)$ be any pdf. Then for any μ, $-\infty < \mu < \infty$, and any $\sigma > 0$, the family of pdfs $(1/\sigma)f((x-\mu)/\sigma)$, indexed by the parameter (μ, σ), is called the location-scale family with standard pdf $f(x)$; μ is called the location parameter and σ is called the scale parameter.

The effect of introducing both the location and scale parameters is to stretch $(\sigma > 1)$ or contract $(\sigma < 1)$ the graph with the scale parameter and then shift the graph so that the point that was above 0 is now above μ. The normal and double exponential families are examples of location-scale families.

<u>Theorem 3.5.6</u>

Let $f(\cdot)$ be any pdf. Let μ be any real number, and let σ be any positive real number. Then X is a random variable with pdf $(1/\sigma)f((x-\mu)/\sigma)$ if and only if there exists a random variable Z with pdf $f(z)$ and $X = \sigma Z + \mu$.

<u>Theorem 3.5.7</u>

Let Z be a random variable with pdf $f(z)$. Suppose EZ and $\text{Var}Z$ exist. If X is a random variable with pdf $(1/\sigma)f((x-\mu)/\sigma)$, then

$$EX = \sigma EZ + \mu \quad \text{and} \quad \text{Var}X = \sigma^2 \text{Var}Z.$$

In particular, if $EZ = 0$ and $\text{Var}Z = 1$, then $EX = \mu$ and $\text{Var}X = \sigma^2$.