3.5 Location and scale

In this section we consider continuous distributions only.

Definition. Let $f(x)$ denote a pdf.

(i) The **location family** associated with f consists of all pdf's of the form

$$g(x) = f(x - c), \quad c \in \mathbb{R}.$$

(ii) The **scale family** associated with f consists of all pdf's of the form

$$g(x) = \frac{1}{s} f(\frac{x}{s}), \quad s > 0.$$

(iii) The **location-scale family** associated with f consists of all pdf's of the form

$$g(x) = \frac{1}{s} f(\frac{x - c}{s}), \quad c \in \mathbb{R}, s > 0.$$

Call c location parameter and s scale parameter.

Example. Consider $Z \sim N(0,1)$ and $X = \mu + \sigma Z \sim N(\mu, \sigma^2)$. Then, using $Z = (X - \mu)/\sigma$,

$$f_X(x) = \left|\frac{dz}{dx}\right| f_Z(z) = \frac{1}{\sigma} f_Z(\frac{x - \mu}{\sigma}) \quad \text{(location-scale family).}$$

Since $f_Z(z) = \frac{1}{\sqrt{2\pi}} e^{-z^2/2}$, this is $f_X(x) = \frac{1}{\sigma}\frac{1}{\sqrt{2\pi}} e^{-\{(x-\mu)/\sigma\}^2/2}$.

Note. The family can be re-parameterized, i.e. choose a different *standard* density, e.g. $f(z) = e^{-\pi(x-1)^2}$ (instead of the N(0,1) density f_Z). Any normal density can be expressed with f: choose $s = \sqrt{2\pi}\sigma$ and $c = \mu - \sqrt{2\pi}\sigma$ to obtain

$$\frac{1}{s} f(\frac{x - c}{s}) = \frac{1}{\sqrt{2\pi}\sigma} \exp\Big\{-\pi\Big(\frac{x - \mu + \sqrt{2\pi}\sigma}{\sqrt{2\pi}\sigma} - 1\Big)^2\Big\} = \frac{1}{\sqrt{2\pi}\sigma} \exp\Big\{-\frac{(x-\mu)^2}{2\sigma^2}\Big\}.$$

Example. Consider $X \sim \text{gamma}(\alpha, \beta)$ with pdf $f_X(x) = \{\Gamma(\alpha)\beta^\alpha\}^{-1} x^{\alpha-1} e^{-x/\beta} \times 1_{(0,\infty)}(x)$. Assume α fixed. This gives a scale family with scale parameter β,

$$f_X(x) = \frac{1}{\beta}\frac{1}{\Gamma(\alpha)}\Big(\frac{x}{\beta}\Big)^{\alpha-1} e^{-\frac{x}{\beta}} 1_{(0,\infty)}\Big(\frac{x}{\beta}\Big) = \frac{1}{\beta} f_Y\Big(\frac{x}{\beta}\Big)$$

with standard density $f_Y(y) = \Gamma(\alpha)^{-1} y^{\alpha-1} e^{-y}$ for $y > 0$.

Example (standard logistic distribution function):

$$F(x) = \frac{1}{1 + e^{-x}}, \quad f(x) = \frac{e^{-x}}{(1 + e^{-x})^2}.$$

We can create a location family (L), or a location scale family (LS):

$$\text{(L)} \quad g(x) = f(x - \mu) = \frac{e^{-(x-\mu)}}{(1 + e^{-(x-\mu)})^2} \qquad \text{(LS)} \quad g(x) = \frac{1}{\beta}\frac{e^{-(x-\mu)/\beta}}{(1 + e^{-(x-\mu)/\beta})^2}.$$

Note. We can make this for any pdf, i.e. for any distribution shape we can compute a pdf if we have values for the mean and the variance.

3.6 Exponential families

A family of pmf's or pdf's is a **one-parameter exponential family** (with parameter θ) if
(i) the support of f ($=\{x : f(x) > 0\}$) is the same for all f in the family;
(ii) there are functions c, h, w and t so that the pdf/pmf can be written in the form

$$f(x) \;=\; c(\theta)\,h(x)\,e^{w(\theta)\,t(x)}.$$

Alternatively consider the log transform: $\log f(x) = a(\theta) + g(x) + w(\theta)t(x)$.

Example: bin(n, p) with n fixed, support $A = \{0, 1, \ldots, n\}$, $\theta = p$.

$$\begin{aligned} f(y) &= \binom{n}{y} p^y (1-p)^{n-y} 1_A(y) = \binom{n}{y} 1_A(y)(1-p)^n \Big(\frac{p}{1-p}\Big)^y \\ &= \binom{n}{y} 1_A(y)(1-p)^n \exp\Big(y \log \frac{p}{1-p}\Big) = c(p)\,h(y)\,e^{w(p)t(y)} \end{aligned}$$

with $c(p) = (1-p)^n$, $h(y) = \binom{n}{y} 1_A(y)$, $t(y) = y$, $w(p) = \log\{p/(1-p)\}$, $p \in (0, 1)$.
[$p = 0$: $A = \{0\}$, $p = 1$: $A = \{n\}$; definition (i) is violated.]

Example: $\mathrm{N}(0, \sigma^2)$, support $A = \mathbb{R}$, $f(x) = \underbrace{\frac{1}{\sqrt{2\pi}\sigma}}_{c(\sigma)} \exp(\underbrace{-\frac{1}{2\sigma^2}}_{w(\sigma)} \underbrace{x^2}_{t(x)}) \cdot \underbrace{1}_{h(x)}$.

Definition. Suppose $\theta \in \mathbb{R}^d$, $d \in \mathbb{N}$. A family of pdf's or pmf's forms an **exponential family** if
(i) the support of f is the same for all f in the family;
(ii) $f(x) = c(\theta)\,h(x) \exp\{w_1(\theta)t_1(x) + \ldots + w_k(\theta)t_k(x)\}$, $k \ge d$.

Note: the condition on the dimension k is necessary to guarantee uniqueness: if $k < d$, for example $k = 1 < d = 2$, then $w_1(\theta_1, \theta_2)$ can have the same value for different parameters θ_1, θ_2 and therefore yield the same distribution.

Example: $\mathrm{N}(\mu, \sigma^2)$, $\theta = (\mu, \sigma^2)$.

$$\begin{aligned} f(x) &= \frac{1}{\sqrt{2\pi}\sigma} e^{-(\frac{x-\mu}{\sigma})^2/2} = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2\sigma^2}(x^2+\mu^2-2x\mu)} \\ &= \underbrace{\frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{\mu^2}{2\sigma^2}}}_{c(\mu,\sigma^2)} \exp\{\underbrace{-\frac{1}{2\sigma^2}}_{w_1} \underbrace{x^2}_{t_1} + \underbrace{\frac{\mu}{\sigma^2}}_{w_2} \underbrace{x}_{t_2}\}. \end{aligned}$$

Example: gamma(α, β), $A = \mathbb{R}^+$; two-dimensional exponential family:
$f(x) = \frac{1}{\Gamma(\alpha)\beta^\alpha} x^{\alpha-1} e^{-x/\beta} 1_A(x) = \frac{1}{\Gamma(\alpha)\beta^\alpha} \exp\{\underline{(\alpha-1)}\,\underline{\log x} + \underline{(-\beta^{-1})}\underline{x}\} 1_A(x)$.
Fixing α (or β) gives a one-parameter exponential family:

$$f(x) \;=\; \underbrace{\frac{1}{\Gamma(\alpha)\beta^\alpha}} \underbrace{1_A(x) x^{\alpha-1}} e^{-\beta^{-1}x}.$$

Example: Pareto pdf $f(x) = \alpha x^{-(\alpha+1)} 1_{[1,\infty)}(x) = \alpha 1_{[1,\infty)}(x) \exp\{-(\alpha+1)\log x\}$ (one-parameter exponential family). Adding a scale parameter gives

$$g(x) = \frac{1}{\beta} f\Big(\frac{x}{\beta}\Big) = \frac{1}{\beta} \alpha \Big(\frac{x}{\beta}\Big)^{-(\alpha+1)} 1_{[1,\infty)}\Big(\frac{x}{\beta}\Big).$$

This is not an exponential family anymore: the support is $[\beta, \infty)$.