A Location-Scale Family of distributions has densities (pdf's) of the form

$$g(x|\mu,\sigma) = \frac{1}{\sigma}\psi\left(\frac{x-\mu}{\sigma}\right) \quad \text{where}$$

ψ is a pdf, $\sigma > 0$, $-\infty < \mu < \infty$.

Properties

1. $g(x|0,1) = \psi(x)$
2. If $X \sim g(\cdot|\mu,\sigma)$, then $\frac{X-\mu}{\sigma} \sim g(\cdot|0,1)$.
3. If $X \sim g(\cdot|0,1)$, then $\sigma X+\mu \sim g(\cdot|\mu,\sigma)$.

(Prove 2 and 3 by the material in Section 2.1.)

Example: The normal distributions are a Location-scale family.

Take $\psi(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$. (The standard normal distn.)

Then
$$g(x|\mu,\sigma) = \frac{1}{\sigma}\psi\left(\frac{x-\mu}{\sigma}\right)$$
$$= \frac{1}{\sigma}\frac{1}{\sqrt{2\pi}} e^{-\left(\frac{x-\mu}{\sigma}\right)^2/2}$$
$$= \frac{1}{\sqrt{2\pi}\,\sigma} e^{-(x-\mu)^2/2\sigma^2}$$
is the pdf of a $N(\mu,\sigma^2)$ distn.

Example: The Cauchy Location-Scale family.

Take $\psi(x) = \frac{1}{\pi} \frac{1}{1+x^2}, \ -\infty < x < \infty.$

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

$$= \frac{1}{\sigma} \frac{1}{1+\left(\frac{x-\mu}{\sigma}\right)^2}, \ -\infty < x < \infty$$

defines the Cauchy L-S family.

Note: For this family of distns, μ is not the mean and σ is not the standard deviation.
(Same remark applies in next example.)

Example: The Uniform distn's form a L-S family.

Take $\psi(x) = I_{(0,1)}(x) = \begin{cases} 1 & \text{if } 0<x<1 \\ 0 & \text{otherwise}. \end{cases}$

ψ is the pdf of the Uniform(0,1) distn.

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

$$= \frac{1}{\sigma} I_{(0,1)}\left(\frac{x-\mu}{\sigma}\right)$$

(Note: $\frac{x-\mu}{\sigma} \in (0,1)$ iff $x-\mu \in (0,\sigma)$
iff $x \in (\mu, \mu+\sigma)$)

$$= \frac{1}{\sigma} I_{(\mu,\mu+\sigma)}(x)$$

which is the pdf of the Uniform$(\mu,\mu+\sigma)$ distn.

Let ψ be a pdf.

A scale family of distns. has densities of the form $g(x|\sigma) = \frac{1}{\sigma}\psi\left(\frac{x}{\sigma}\right)$ where $\sigma > 0$.

(σ is the scale parameter.)

A location family of distns. has densities of the form $g(x|\mu) = \psi(x-\mu)$ where $-\infty < \mu < \infty$.

Example: The $N(\mu,1)$ distns. form a location family.

The $N(0,\sigma^2)$ distns. form a scale family.

Take $\psi(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$. Then

$$\psi(x-\mu) = \frac{1}{\sqrt{2\pi}} e^{-(x-\mu)^2/2}$$ which is the pdf of $N(\mu,1)$,

$$\frac{1}{\sigma}\psi\left(\frac{x}{\sigma}\right) = \frac{1}{\sqrt{2\pi}\,\sigma} e^{-x^2/2\sigma^2}$$ which is the pdf of $N(0,\sigma^2)$.

Example: The family of Gamma(α_0, β) distns (where α_0 is any fixed value of α) forms a scale family.

Take $\psi(x) = \dfrac{x^{\alpha_0 - 1} e^{-x}}{\Gamma(\alpha_0)}, \quad x > 0.$

Then
$$\frac{1}{\sigma}\psi\left(\frac{x}{\sigma}\right) = \frac{\frac{1}{\sigma}\left(\frac{x}{\sigma}\right)^{\alpha_0 - 1} e^{-x/\sigma}}{\Gamma(\alpha_0)}$$

$$= \frac{x^{\alpha_0 - 1} e^{-x/\sigma}}{\sigma^{\alpha_0}\,\Gamma(\alpha_0)}$$

which is the pdf of the Gamma(α_0, σ) distn.

Note: If we permit both α and β to vary, the family of Gamma(α, β) distns. does not form a Location-Scale family.

(α is a "shape" parameter.)

Suppose $g(x \mid \mu, \sigma)$ is a location-scale family of densities, and

$$X \sim g(\cdot \mid \mu, \sigma)\,, \qquad Z \sim g(\cdot \mid 0, 1)\,.$$

Then $\quad \dfrac{X-\mu}{\sigma} \stackrel{d}{=} Z \quad$ and $\quad X \stackrel{d}{=} \sigma Z + \mu \quad$ so that

$$P(X > b) = P\left(\frac{X-\mu}{\sigma} > \frac{b-\mu}{\sigma}\right) = P\left(Z > \frac{b-\mu}{\sigma}\right), \text{ etc.}$$

$$E(X) = E(\sigma Z + \mu) = \sigma \cdot EZ + \mu \quad \text{(if } EZ \text{ is finite)}$$

$$\mathrm{Var}(X) = \mathrm{Var}(\sigma Z + \mu) = \sigma^2 \mathrm{Var}(Z) \quad \text{(if } \mathrm{Var}(Z) \text{ is finite)}$$

Similar facts hold for location families and scale families.

Erase μ (set $\mu = 0$) for facts for scale families. Erase σ (set $\sigma = 1$) for facts about location families.

Example: The $N(\sigma, \sigma^2)$, $\sigma > 0$, distributions form a scale family.

The density of the $N(\sigma, \sigma^2)$ distribution is:

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

Example: The $N(1, \lambda)$, $\lambda > 0$, distributions do NOT form a scale family.

One way to see this is to note that if $X \sim N(1, \lambda)$ then $EX = 1$ for all λ (it is constant). But a scale family with scale parameter σ satisfies $EX = \sigma EZ$ which cannot be constant (unless $EZ = 0$).

Exponential Families

The family of pdf's or pmf's

$$\{ f(x|\theta) : \theta \in \Theta \}$$

The parameter space.

(θ might represent a single parameter or a vector of parameters.)

is an exponential family if we can write

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

valid for all x and all $\theta \in \Theta$.

This is the general k parameter exponential family (kpef).

For $k=1$, the general one parameter exponential family (1pef) has the form

$$f(x|\theta) = h(x)\,c(\theta)\exp\{w(\theta)t(x)\}$$

valid for all x and all $\theta \in \Theta$.

Note: We allow h to be degenerate (constant), but require all the other functions to be nondegenerate (nonconstant).

Examples of 1pef's

Exponential distributions

$$f(x|\beta) = \frac{1}{\beta} e^{-x/\beta}, \quad x > 0, \beta > 0. \quad \text{(pdf)}$$

In this example

$$\theta = \beta, \quad \Theta = (0,\infty).$$

$$f(x|\beta) = \underbrace{I_{(0,\infty)}(x)}_{h(x)} \cdot \underbrace{\frac{1}{\beta}}_{c(\theta)} \cdot \exp\left\{\underbrace{-\frac{1}{\beta}}_{w(\theta)} \cdot \underbrace{x}_{t(x)}\right\}$$

Thus $f(x|\beta)$ forms a 1pef with the parts as identified above.

Binomial distributions

The family of Binomial(n,p) distributions with n known (fixed) is a 1pef. The pmf is

$$f(x|p) = \binom{n}{x} \underbrace{p^x (1-p)^{n-x}}_{(1-p)^n \left(\frac{p}{1-p}\right)^x}, \quad x = 0, 1, \ldots, n, \quad 0 < p < 1$$

In this example $\theta = p$, $\Theta = (0,1)$, and

$f(x|p) =$

$$\underbrace{\binom{n}{x} I_{\{0,1,\ldots,n\}}(x)}_{h(x)} \underbrace{(1-p)^n}_{c(\theta)} \exp\Big\{ \underbrace{x}_{t(x)} \cdot \underbrace{\log\left(\tfrac{p}{1-p}\right)}_{w(\theta)} \Big\}.$$

Examples of 2pef's

The family of $N(\mu, \sigma^2)$ distributions

The $N(\mu, \sigma^2)$ pdf is

$$f(x|\mu,\sigma^2) = \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \quad -\infty < x < \infty$$

valid for $\sigma^2 > 0$ and $-\infty < \mu < \infty$.

Here $\theta = (\mu, \sigma^2)$ and

$$\Theta = \{(\mu, \sigma^2) : \sigma^2 > 0 \text{ and } -\infty < \mu < \infty\}.$$

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

$$= \underbrace{\frac{1}{\sqrt{2\pi}}}_{h(x)} \underbrace{\frac{1}{\sigma}\exp\left(\frac{-\mu^2}{2\sigma^2}\right)}_{c(\theta)} \exp\Big\{\underbrace{\frac{-1}{2\sigma^2}\cdot x^2}_{w_1(\theta)t_1(x)} + \underbrace{\frac{\mu}{\sigma^2}\cdot x}_{w_2(\theta)t_2(x)}\Big\}$$

We have a 2pef with the parts as identified above.

Note: In writing an exponential family, h is allowed to be degenerate (constant), but not any of the other parts. Also, we require $h(x) \geq 0$ for all x.

Non-exponential families

There are many families of distributions which are <u>not</u> exponential families.

The Cauchy Location-Scale family

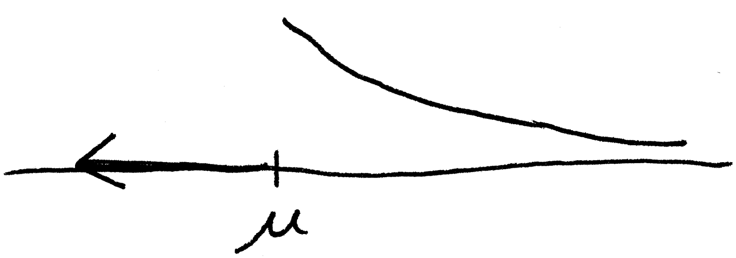
$$f(x|\mu,\sigma) = \frac{1}{\sigma} \cdot \frac{1}{\pi\left(1+\left(\frac{x-\mu}{\sigma}\right)^2\right)} \quad \text{for all } x$$

can<u>not</u> be written as an exponential family.
(Try it!)

A trickier example:

Consider this shifted exponential distribution with pdf

$$f(x|\mu,\beta) = \frac{1}{\beta} e^{-(x-\mu)/\beta} \quad \text{for } x > \mu$$



Plot of pdf

$$?? \quad f(x|\underbrace{\mu,\beta}_{\theta}) = \underbrace{1}_{h(x)} \cdot \underbrace{\frac{e^{\mu/\beta}}{\beta}}_{c(\theta)} \, \underbrace{e^{-\frac{1}{\beta}\cdot x}}_{w(\theta)t(x)}$$

This is <u>not</u> valid for <u>all</u> x, but only for $x > \mu$. To get an expression valid for all x, we need an indicator function.

$$f(x|\mu,\beta) = \underbrace{I_{(\mu,\infty)}(x)}_{\substack{\text{not a} \\ \text{function} \\ \text{of } x \text{ alone!}}} \underbrace{\frac{e^{\mu/\beta}}{\beta}}_{c(\theta)} e^{\overbrace{-\frac{1}{\beta}\cdot x}^{w(\theta)t(x)}}$$

This is <u>not</u> an exponential family.

<u>Definition</u>: The <u>support</u> of a pdf or pmf $f(x)$ is the set

$$\{x : f(x) > 0\}.$$

<u>Fact</u>: The support of an exponential family of pdf's (pmf's) $f(x|\theta)$ is the same for all θ.

<u>Proof</u> (for 1pef): The support of

$$f(x|\theta) = h(x)c(\theta)\exp\{w(\theta)t(x)\}$$

is $\{x : h(x) > 0\}$ which does not involve θ.

Return to Previous Example

The pdf

$$f(x|\mu,\beta) = \frac{1}{\beta} e^{-(x-\mu)/\beta} \quad \text{for } x>\mu$$

has support $\{x : x>\mu\}$ which depends on $\theta=(\mu,\beta)$ through the value μ.

Thus (without further work) we know this is <u>not</u> an exponential family.

<u>Example</u>: The family of Uniform(a,b) distributions with $-\infty < a < b < \infty$ is <u>not</u> an exponential family.

The Uniform(a,b) density

$$f(x|\underbrace{a,b}_{\theta}) = \begin{cases} \frac{1}{b-a} & \text{for } a<x<b \\ 0 & \text{otherwise} \end{cases}$$

has support $\{x : a<x<b\}$ which depends on $\theta=(a,b)$. Thus (without further work) we know this is <u>not</u> an exponential family.

<u>Example</u>: The Cauchy location-scale family is <u>not</u> an exponential family, but its support is the <u>same</u> for all $\theta=(\mu,\sigma)$.