1 Location-Scale Family

Definition 1. *A location-scale family of distributions has densities (pdf's) of the form*

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

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

1.1 Properties

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

Prove 2 and 3 by the material in Section 2.1

1.2 Examples

1. **Gaussian:** Take $\psi(x) = \frac{1}{\sqrt{2\pi}}e^{-x^2/2}$, the standard normal density. Then

   $$g(x \mid \mu, \sigma) = \frac{1}{\sigma}\psi\left(\frac{x-\mu}{\sigma}\right) = \frac{1}{\sigma}\frac{1}{\sqrt{2\pi}}e^{-\left(\frac{x-\mu}{\sigma}\right)^2/2} = \frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \quad -\infty < x < \infty$$

   is the pdf of a $\mathrm{N}(\mu, \sigma^2)$ distribution.
2. **Cauchy:** Take $\psi(x) = \frac{1}{\pi}\frac{1}{1+x^2}$, Then

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

   defines the Cauchy L-S family.
   Note: for this family of distributions, μ is not the mean and σ is not the standard deviation. Same remark applies in the next example

3. **Uniform:** 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 Unif$(0,1)$ distribution.

$$g(x \mid \mu, \sigma) \quad = \quad \frac{1}{\sigma}\psi\left(\frac{x-\mu}{\sigma}\right) = \frac{1}{\sigma}I_{(0,1)}\left(\frac{x-\mu}{\sigma}\right) = \frac{1}{\sigma}I_{(\mu,\mu+\sigma)}(x)$$

where the last equality follows from the fact that $\frac{x-\mu}{\sigma} \in (0,1)$ iff $x \in (\mu, \mu+\sigma)$. This is the pdf of the of the Unif$(\mu, \mu+\sigma)$.

2 Location / Scale Family

Let ψ be a pdf.

Definition 2. *A* <u>*scale family*</u> *of distributions has densities of the form* $g(x \mid \sigma) = \frac{1}{\sigma}\psi\left(\frac{x}{\sigma}\right)$ *where* $\sigma > 0$.

σ is the scale parameter.

Definition 3. *A* <u>*location family*</u> *of distributions has densities of the form* $g(x \mid \mu) = \psi(x-\mu)$ *where* $-\infty < \mu < \infty$.

2.1 Examples

1. The $\text{N}(\mu, 1)$ distributions form a location family and $\text{N}(0, \sigma^2)$ distributions 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 $\text{N}(\mu, 1)$ and $\frac{1}{\sigma}\psi\left(\frac{x}{\sigma}\right) = \frac{1}{\sqrt{2\pi}\sigma}e^{-x^2/2\sigma^2}$, which is the pdf of $\text{N}(0, \sigma^2)$.

2. The family of Gamma(α_0, β) distributions (where α_0 is any fixed value of α forms a scale family. Take $\psi(x) = \frac{x^{\alpha_0 - 1}e^{-x}}{\Gamma(\alpha_0)}, x > 0$. 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) distribution.
<u>Note</u>: If we permit both α and β to vary, the family of Gamma(α, β) distributions does <u>not</u> form a location-scale family. (α is a "shape" parameter).

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

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

Then

$$\frac{X-\mu}{\sigma} \stackrel{d}{=} Z, \quad X \stackrel{d}{=} \sigma Z + \mu$$

so that

$$\begin{aligned} P(X > b) &= P\left(\frac{X-\mu}{\sigma} > \frac{b-\mu}{\sigma}\right) = P\left(Z > \frac{b-\mu}{\sigma}\right) \\ E(X) &= E(\sigma Z + \mu) = \sigma EZ + \mu, \text{ if } EZ \text{ is finite} \\ Var(X) &= Var(\sigma Z + \mu) = \sigma^2 Var(Z) \text{ if } Var(Z) \text{ is finite} \end{aligned}$$

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.

2.2 Examples

1. The $\mathrm{N}(\sigma, \sigma^2), \sigma > 0$, distributions form a scale family. The density of the $N(\sigma, \sigma^2)$ distribution is:

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

   where ψ is the $N(1,1)$ pdf.

2. The $\mathrm{N}(1, \lambda), \lambda > 0$, distributions do <u>not</u> form a scale family. One way to see this is to note that if $X \sim \mathrm{N}(1, \lambda)$ 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 $EZ = 0$).

3 Exponential families

Definition 4. *The family of pdf's or pmf's $\{f(x \mid \theta) : \theta \in \Theta\}$, where Θ is the parameter space and θ can represent a single parameter or a vector of parameters is an <u>exponential family</u> if we can write*

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

for real valued functions $c(\theta), h(x), w(\theta), t(x)$ with $c(\theta), h(x) \geq 0$ for all x and all $\theta \in \Theta$. <u>In addition $c(\theta) > 0$ for all $\theta \in \Theta$.</u>

This is the general <u>k-parameter exponential family</u>(kpef). For $k = 1$, the general <u>one parameter exponential family</u>(1pef) has the form

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

for all x and all $\theta \in \Theta$.
<u>Note</u>: We allow h to be degenerate (constant), but require <u>all</u> the other functions to be nondegenerate (nonconstant).

3.1 Examples of 1pef's

1. **Exponential Distributions:** The pdf is given by

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

In this example, $\theta = \beta$ and $\Theta = (0, \infty)$ and

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

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

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

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

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

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

This $f(x \mid p)$ forms a 1pef with the parts as identified above.

3.2 Examples of 2pef's

1. **The family of $\mathbf{N}(\mu, \sigma^2)$ distributions:** The pdf is given by

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

valid for $\sigma > 0$ and $-\infty < \mu < \infty$. In this example, $\theta = (\mu, \sigma^2)$ and

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

$$\begin{aligned} f(x \mid \mu, \sigma^2) &= \frac{1}{\sqrt{2\pi}\sigma} \exp\left\{ -\frac{x^2}{2\sigma^2} + \frac{x\mu}{\sigma^2} - \frac{\mu^2}{2\sigma^2} \right\} \\ &= \underbrace{\frac{1}{\sqrt{2\pi}}}_{h(x)} \cdot \underbrace{\frac{1}{\sigma} \exp\left(-\frac{\mu^2}{2\sigma^2} \right)}_{c(\theta)} \cdot \exp\left\{ \underbrace{\frac{-1}{2\sigma^2}}_{w_1(\theta)} \cdot \underbrace{x^2}_{t_1(x)} + \underbrace{\frac{\mu}{\sigma^2}}_{w_2(\theta)} \cdot \underbrace{x}_{t_2(x)} \right\} \end{aligned}$$

This $f(x \mid \mu, \sigma^2)$ forms a 2pef with the parts as identified above.

3.3 Non-exponential families

There are many families of distributions which are <u>not</u> exponential families. The Cauchy L-S family family

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

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

A trickier Example: Consider this shifted exponential distribution with pdf

$$f(x \mid \mu, \beta) = \begin{cases} \frac{1}{\beta} e^{-(x-\mu)/\beta}, & x > \mu \\ 0, & \text{otherwise.} \end{cases}$$

Note that

$$f(x \mid \mu, \beta) = \underbrace{1}_{h(x)} \cdot \underbrace{\frac{e^{\mu/\beta}}{\beta}}_{c(\theta)} \exp\left\{ -\underbrace{\frac{1}{\beta}}_{w(\theta)} \cdot \underbrace{x}_{t(x)} \right\}, \qquad \text{Is it correct?}$$

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.

$$f(x \mid \mu, \beta) = \underbrace{I_{(\mu,\infty)}(x)}_{\underline{\text{not}} \text{ a function of } x \text{ alone}} \cdot \underbrace{\frac{e^{\mu/\beta}}{\beta}}_{c(\theta)} \exp\left\{ -\underbrace{\frac{1}{\beta}}_{w(\theta)} \cdot \underbrace{x}_{t(x)} \right\}.$$

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

Definition 5. *The* <u>*support*</u> *of an exponential family 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) $f(x \mid \theta)$ is the same for all θ.

Proof. (For 1pef) The support of $f(x \mid \theta) = h(x)c(\theta)\exp\{w(\theta)t(x)\}$ is $\{x : h(x) > 0\}$ which does not involve θ. □

<u>Fact:</u> If $f(x \mid \theta)$ is 1pef, then for any 4 distinct points x_1, x_2, x_3 and x_4 in the support, $g(x, \theta) = \log\{f(x \mid \theta)/h(x)\}$ satisfies

$$\frac{g(x_1, \theta) - g(x_2, \theta)}{g(x_3, \theta) - g(x_4, \theta)}$$

independent of θ.

Proof. Observe that if $f(x \mid \theta)$ is 1pef

$$\frac{g(x_1, \theta) - g(x_2, \theta)}{g(x_3, \theta) - g(x_4, \theta)} = \frac{t(x_1) - t(x_2)}{t(x_3) - t(x_4)}$$

and the proof follows immediately.

□

3.4 Examples

1. The pdf

$$f(x \mid \mu, \beta) = \begin{cases} \frac{1}{\beta} e^{-(x-\mu)/\beta}, & x > \mu \\ 0, & \text{otherwise.} \end{cases}$$

   has support $\{x : x > \mu\}$ which depends on $\theta = (\mu, \beta)$ through the value μ. Thus (without further work) we know that this is <u>not</u> an exponential family.

2. The family of Unif(a, b) distributions with $-\infty < a < b < \infty$ is <u>not</u> an exponential family. The Unif(a, b) density

$$f(x \mid a, b) = \begin{cases} \frac{1}{b-a}, \text{ for } a < x < b \\ 0, \text{ otherwise} \end{cases}$$

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

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

   Proof. We will show that for $\sigma = 1$, Cauchy Location family is not a 1pef. Observe that for Cauchy Location family

$$\frac{g(x_1, \theta) - g(x_2, \theta)}{g(x_3, \theta) - g(x_4, \theta)}$$

   is not a function of x alone.

   □