The GAMLSS distributions

Robert A. Rigby , Mikis D. Stasinopoulos, Gillian Z. Heller, and Fernanda De Bastiani

September 6, 2025

Contents

Chapter 1

Continuous distributions on $(-\infty, \infty)$

This chapter gives summary tables and plots for the explicit **gamlss.dist** continuous distributions with range $(-\infty, \infty)$. These are discussed in Chapter **??**.

1.1 Location-scale family of distributions

A continuous random variable Y defined on $(-\infty, \infty)$ is said to have a location-scale family of distributions with location shift parameter θ_1 and scaling parameter θ_2 (for fixed values of all other parameters of the distribution) if

$$Z = \frac{Y - \theta_1}{\theta_2}$$

has a cdf which does not depend on θ_1 or θ_2. Hence

$$F_Y(y) = F_Z\left(\frac{y - \theta_1}{\theta_2}\right)$$

and

$$f_Y(y) = \frac{1}{\theta_2} f_Z\left(\frac{y - \theta_1}{\theta_2}\right) \ ,$$

so $F_Y(y)$ and $\theta_2 f_Y(y)$ only depend on y, θ_1 and θ_2 through the function $z = (y - \theta_1)/\theta_2$. Note $Y = \theta_1 + \theta_2 Z$.

Example: Let Y have a Gumbel distribution, $Y \sim$ `GU`(μ, σ), then Y has a location-scale family of distributions with location shift parameter μ and scaling parameter σ, since $F_Y(y) = 1 - \exp\left[-\exp\left(\frac{y-\mu}{\sigma}\right)\right]$. Hence $Z = (Y - \mu)/\sigma$ has cdf $F_Z(z) = 1 - \exp\left[-\exp(z)\right]$ which does not depend on either μ or σ. Note $Z \sim$ `GU`$(0, 1)$.

All distributions with range $(-\infty, \infty)$ in **gamlss.dist** are location-scale families of distributions with location shift parameter μ and scaling parameter σ, (for fixed values of all other parameters of the distribution), with the exception of `NO2`(μ, σ), `NOF`(μ, σ, ν), and `exGAUS`(μ, σ, ν). Hence for these location-scale family distributions, $Y = \mu + \sigma Z$ where the distribution of Z does not depend on μ or σ. Hence, for fixed parameters ν and τ, $\mathrm{Var}(Y)$ is proportional to σ^2 and $\mathrm{SD}(Y)$ is proportional to σ, provided $\mathrm{Var}(Y)$ is finite.

For the location-scale family distributions plotted in Sections 1.3 and 1.4 we fix $\mu = 0$ and $\sigma = 1$, since changing σ from 1 to say σ_1 merely scales the horizontal axis in the figures by the factor σ_1, and scales the vertical axis by the factor $1/\sigma_1$. After this scaling, changing μ from 0 to say μ_1 simply shifts the horizontal axis by μ_1. (The effect of changing μ and σ can be seen, for example, for the `NO`(μ, σ) distribution in Figure 1.3.)

1.2 Continuous two-parameter distributions on $(-\infty, \infty)$

1.2.1 Gumbel: `GU`

The pdf of the Gumbel distribution (or reverse extreme value distribution), denoted by `GU`(μ, σ), is given by

$$f_Y(y \mid \mu, \sigma) = \frac{1}{\sigma} \exp\left[\left(\frac{y-\mu}{\sigma}\right) - \exp\left(\frac{y-\mu}{\sigma}\right)\right] \tag{1.1}$$

for $-\infty < y < \infty$, where $-\infty < \mu < \infty$ and $\sigma > 0$.

Note if $Y \sim$ `GU`(μ, σ) and $W = -Y$ then $W \sim$ `RG`$(-\mu, \sigma)$, from which the results in Table 1.1 were obtained. The Gumbel distribution is appropriate for moderately negatively skewed data.

Table 1.1: Gumbel distribution.

$\texttt{GU}(\mu, \sigma)$	
Ranges	
Y	$-\infty < y < \infty$
μ	$-\infty < \mu < \infty$, mode, location shift par.
σ	$0 < \sigma < \infty$, scaling parameter
Distribution measures	
mean	$\mu - C\sigma \approx \mu - 0.57722\sigma$
median	$\mu - 0.36651\sigma$
mode	μ
variance	$\pi^2\sigma^2/6 \approx 1.64493\sigma^2$
skewness	-1.13955
excess kurtosis	2.4
MGF	$e^{\mu t}\Gamma(1+\sigma t)$, for $\sigma\vert t\vert < 1$
pdf	$\frac{1}{\sigma}\ \exp\left[\left(\frac{y-\mu}{\sigma}\right) - \exp\left(\frac{y-\mu}{\sigma}\right)\right]$
cdf	$1 - \exp\left[-\exp\left(\dfrac{y-\mu}{\sigma}\right)\right]$
inverse cdf	$\mu + \sigma\log[-\log(1-p)]$
Reference	Results derived from $Y = -W$ where $W \sim \texttt{RG}(-\mu, \sigma)$
Note	$C \approx 0.57722$ above is Euler's constant

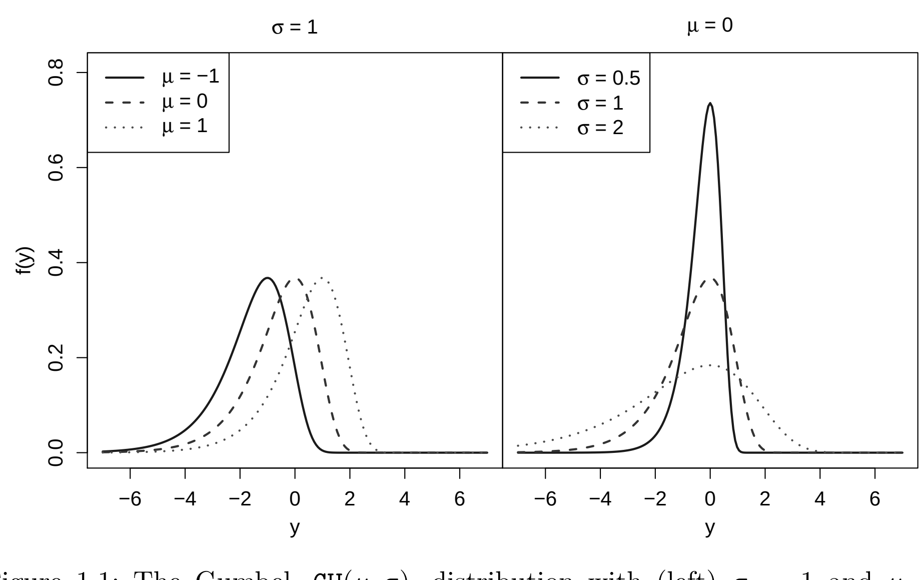

Figure 1.1: The Gumbel, $\texttt{GU}(\mu, \sigma)$, distribution with (left) $\sigma = 1$ and $\mu = -1, 0, 1$, and (right) $\mu = 0$ and $\sigma = 0.5, 1, 2$.

1.2.2 Logistic: LO

The pdf of the logistic distribution, denoted by $\texttt{LO}(\mu, \sigma)$, is given by

$$f_Y(y \mid \mu, \sigma) = \frac{1}{\sigma} \left\{ \exp\left[-\left(\frac{y-\mu}{\sigma} \right) \right] \right\} \left\{ 1 + \exp\left[-\left(\frac{y-\mu}{\sigma} \right) \right] \right\}^{-2} \tag{1.2}$$

for $-\infty < y < \infty$, where $-\infty < \mu < \infty$ and $\sigma > 0$. The logistic distribution is symmetric about μ, and is appropriate for a moderately leptokurtic response variable.

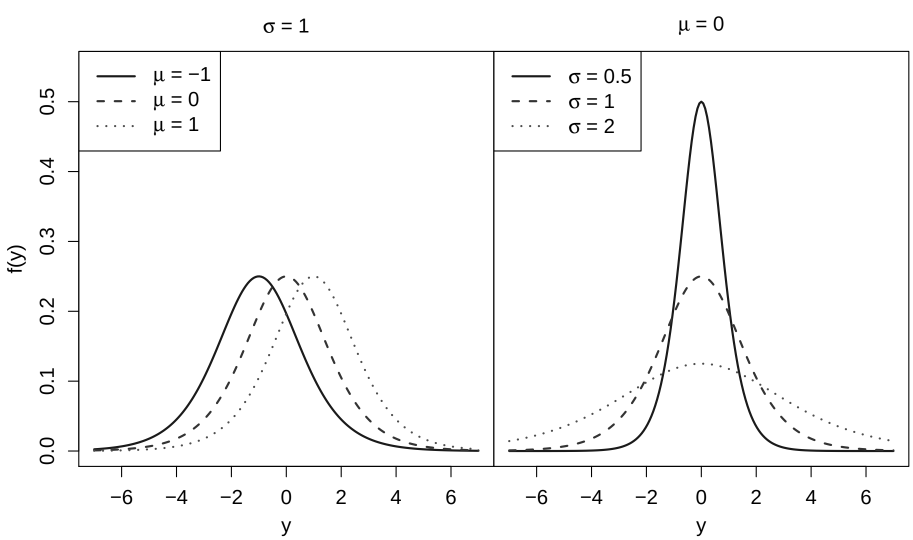

Figure 1.2: The logistic, $\texttt{LO}(\mu, \sigma)$, distribution, with (left) $\sigma = 1$ and $\mu = -1, 0, 1$, and (right) $\mu = 0$ and $\sigma = 0.5, 1, 2$.

Table 1.2: Logistic distribution.

$\texttt{LO}(\mu, \sigma)$	
Ranges	
Y	$-\infty < y < \infty$
μ	$-\infty < \mu < \infty$, mean, median, mode, location shift par.
σ	$0 < \sigma < \infty$, scaling parameter
Distribution measures	
mean [a]	μ
median	μ
mode	μ
variance [a]	$\pi^2\sigma^2/3$
skewness [a]	0
excess kurtosis [a]	1.2
MGF	$e^{\mu t}B(1-\sigma t, 1+\sigma t)$
pdf [a]	$\frac{1}{\sigma}\left\{\exp\left[-\left(\frac{y-\mu}{\sigma}\right)\right]\right\}\left\{1+\exp\left[-\left(\frac{y-\mu}{\sigma}\right)\right]\right\}^{-2}$
cdf [a]	$\left\{1+\exp\left[-\left(\dfrac{y-\mu}{\sigma}\right)\right]\right\}^{-1}$
inverse cdf	$\mu + \sigma \log\left(\dfrac{p}{1-p}\right)$
Reference	[a] Johnson et al. [1995, p115-117] with $\alpha = \mu$ and $\beta = \sigma$

1.2.3 Normal (or Gaussian): `NO`, `NO2`

First parameterization, `NO`

The normal distribution is the default distribution of the `family` argument of the `gamlss()` function. The parameterization used for the normal (or Gaussian) pdf, denoted by $\texttt{NO}(\mu, \sigma)$, is

$$f_Y(y \mid \mu, \sigma) = \frac{1}{\sqrt{2\pi}\sigma}\ \exp\left[-\frac{(y-\mu)^2}{2\sigma^2}\right] \tag{1.3}$$

for $-\infty < y < \infty$, where $-\infty < \mu < \infty$ and $\sigma > 0$. The $\texttt{NO}(\mu, \sigma)$ distribution is symmetric about μ. Note, in Table 1.3, $\Phi()$ and $\Phi^{-1}()$ are the cdf and inverse cdf of a standard normal, $\texttt{NO}(0, 1)$, distribution.

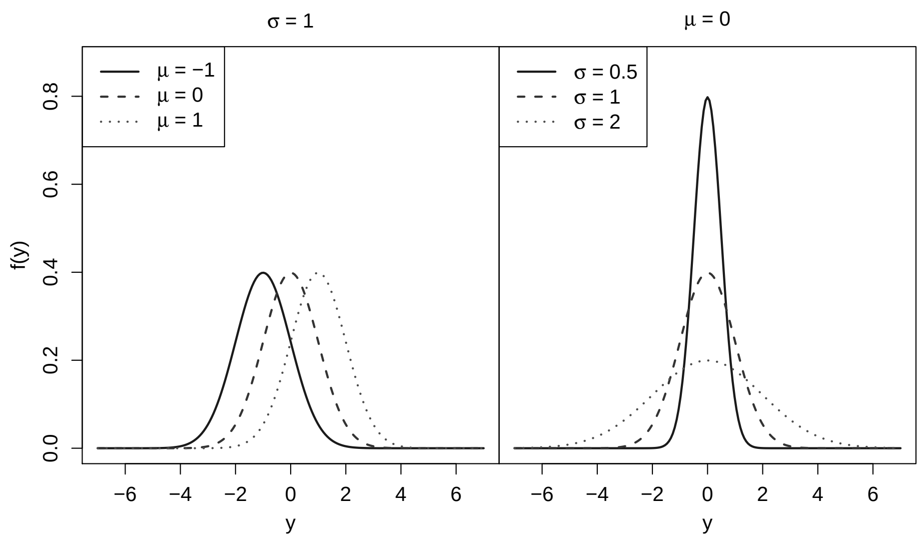

Figure 1.3: The normal, NO(μ, σ), distribution, with (left) $\sigma = 1$ and $\mu = -1, 0, 1$ and (right) $\mu = 0$ and $\sigma = 0.5, 1, 2$.

Table 1.3: Normal distribution.

	NO(μ, σ)
	Ranges
Y	$-\infty < y < \infty$
μ	$-\infty < \mu < \infty$, mean, median, mode, location shift par.
σ	$0 < \sigma < \infty$, standard deviation, scaling parameter
	Distribution measures
mean	μ
median	μ
mode	μ
variance	σ^2
skewness	0
excess kurtosis	0
MGF	$\exp\left(\mu t + \frac{1}{2}\sigma^2 t^2\right)$
pdf	$\frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{(y-\mu)^2}{2\sigma^2}\right]$
cdf	$\Phi[(y-\mu)/\sigma]$
inverse cdf	$\mu + \sigma z_p$ where $z_p = \Phi^{-1}(p)$
Reference	Johnson et al. [1994] Chapter 13, p80-89.

Second parameterization, NO2

The `NO2`(μ, σ) distribution is a parameterization of the normal distribution where μ is the mean and σ is the variance of Y, with pdf given by

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

The `NO2`(μ, σ) distribution is symmetric about μ. Note `NO2`$(\mu, \sigma) =$ `NO`$(\mu, \sigma^{1/2})$.

1.2.4 Reverse Gumbel: `RG`

The reverse Gumbel distribution, also known as the *type I extreme value distribution*, is a special case of the generalized extreme value distribution, see Johnson et al. [1995] p2, p11-13, and p75-76. The pdf of the reverse Gumbel distribution, denoted by `RG`(μ, σ), is given by

$$f_Y(y \,|\, \mu, \sigma) = \frac{1}{\sigma} \exp\left\{-\left(\frac{y-\mu}{\sigma}\right) - \exp\left[-\left(\frac{y-\mu}{\sigma}\right)\right]\right\} \tag{1.4}$$

for $-\infty < y < \infty$, where $-\infty < \mu < \infty$ and $\sigma > 0$.

Note that if $Y \sim$ `RG`(μ, σ) and $W = -Y$, then $W \sim$ `GU`$(-\mu, \sigma)$. The reverse Gumbel distribution is appropriate for moderately positively skewed data.

Since the reverse Gumbel distribution is the type I extreme value distribution, it is the reparameterized limiting distribution of the standardized maximum of a sequence of independent and identically distributed random variables from an 'exponential type distribution', which includes the exponential and gamma.

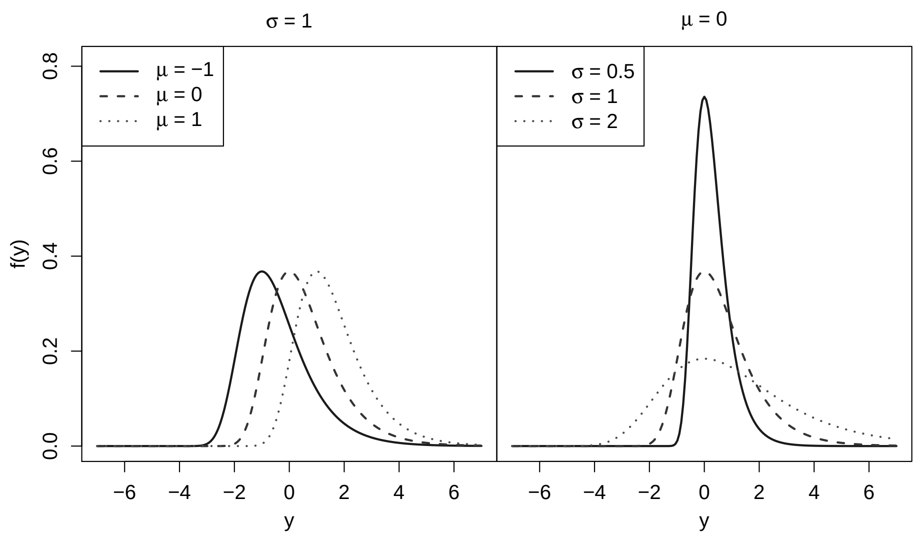

Figure 1.4: The reverse Gumbel, $\texttt{RG}(\mu, \sigma)$, distribution with (left) $\sigma = 1$ and $\mu = -1, 0, 1$ and (right) $\mu = 0$ and $\sigma = 0.5, 1, 2$.

Table 1.4: Reverse Gumbel distribution.

$\texttt{RG}(\mu, \sigma)$	
Ranges	
Y	$-\infty < y < \infty$
μ	$-\infty < \mu < \infty,$ mode, location shift parameter
σ	$0 < \sigma < \infty,$ scaling parameter
Distribution measures	
mean	$\mu + C\sigma \approx \mu + 0.57722\sigma$
median	$\mu + 0.36651\sigma$
mode	μ
variance	$\pi^2\sigma^2/6 \approx 1.64493\sigma^2$
skewness	1.13955
excess kurtosis	2.4
MGF	$e^{\mu t}\Gamma(1-\sigma t),$ for $\sigma\|t\| < 1$
pdf	$\frac{1}{\sigma}\ \exp\left\{-\left(\frac{y-\mu}{\sigma}\right) - \exp\left[-\left(\frac{y-\mu}{\sigma}\right)\right]\right\}$
cdf	$\exp\left\{-\exp\left[-\left(\frac{y-\mu}{\sigma}\right)\right]\right\}$
inverse cdf	$\mu - \sigma\log[-\log p]$
Reference	Johnson et al. [1995] Chapter 22, p2, p11-13, with $\xi = \mu$ and $\theta = \sigma$.
Note	$C \approx 0.57722$ above is Euler's constant.

1.3 Continuous three-parameter distributions on $(-\infty, \infty)$

1.3.1 Exponential Gaussian: exGAUS

The pdf of the exponential Gaussian distribution, denoted by $\texttt{exGAUS}(\mu, \sigma, \nu)$, is

$$f_Y(y \mid \mu, \sigma, \nu) = \frac{1}{\nu}\exp\left[\frac{\mu - y}{\nu} + \frac{\sigma^2}{2\nu^2}\right]\Phi\left(\frac{y-\mu}{\sigma} - \frac{\sigma}{\nu}\right) \tag{1.5}$$

for $-\infty < y < \infty$, where $-\infty < \mu < \infty$, $\sigma > 0$ and $\nu > 0$, and $\Phi(\cdot)$ is the cdf of the standard normal distribution. Note that Y is the sum of a normal and an exponential random variable, i.e. $Y = Y_1 + Y_2$ where $Y_1 \sim \texttt{NO}(\mu, \sigma)$ and $Y_2 \sim \texttt{EXP}(\nu)$ are independent. This distribution has also been called the lagged normal distribution, Johnson et al. [1994], p172. See also Davis and Kutner [1976] and Lovison and Schindler [2014]. The exGAUS distribution is popular in psychology where it has been used to model response times.

Note that if $Y \sim$ `exGAUS`(μ, σ, ν) then $Y_0 = a + bY \sim$ `exGAUS`$(a + b\mu, b\sigma, b\nu)$. So $Z = Y - \mu \sim$ `exGAUS`$(0, \sigma, \nu)$. Hence, for fixed σ and ν, the `exGAUS`(μ, σ, ν) distribution is a location family of distributions with location shift parameter μ. Also `exGAUS`(μ, σ, ν) is a location-scale family of distributions, but σ is not the scaling parameter. If `exGAUS`(μ, σ, ν) is reparameterized by setting $\alpha_1 = \sigma + \nu$ and $\alpha_2 = \sigma/\nu$ then the resulting `exGAUS2`$(\mu, \alpha_1, \alpha_2)$ distribution is a location-scale family of distributions, for fixed α_2, with location shift parameter μ and scaling parameter α_1, since if $Y \sim$ `exGAUS2`$(\mu, \alpha_1, \alpha_2)$ then $Z = (Y - \mu)/\alpha_1 \sim$ `exGAUS2`$(0, 1, \alpha_2)$.

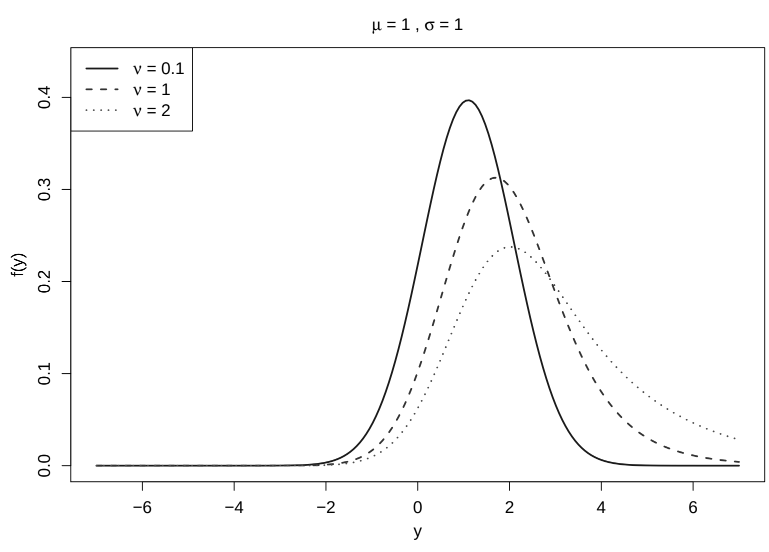

Figure 1.5: The exponential Gaussian, `exGAUS`(μ, σ, ν), distribution, with $\mu = 1$, $\sigma = 1$, and $\nu = 0.1, 1, 2$.

Table 1.5: Exponential Gaussian distribution.

exGAUS(μ, σ, ν)	
Ranges	
Y	$-\infty < y < \infty$
μ	$-\infty < \mu < \infty,$ mean of normal comp., location shift par.
σ	$0 < \sigma < \infty,$ standard deviation of normal component
ν	$0 < \nu < \infty,$ mean of exponential component
Distribution measures	
mean [a]	$\mu + \nu$
variance [a]	$\sigma^2 + \nu^2$
skewness [a]	$2\left(1 + \frac{\sigma^2}{\nu^2}\right)^{-1.5}$
excess kurtosis [a]	$6\left(1 + \frac{\sigma^2}{\nu^2}\right)^{-2}$
MGF	$(1-\nu t)^{-1}\exp(\mu t + \frac{1}{2}\sigma^2 t^2)$
pdf [a]	$\frac{1}{\nu}\exp\left(\frac{\mu - y}{\nu} + \frac{\sigma^2}{2\nu^2}\right)\Phi\left(\frac{y-\mu}{\sigma} - \frac{\sigma}{\nu}\right)$
Reference	[a]Lovison and Schindler [2014]

1.3.2 Normal family of variance-mean relationships: `NOF`

`NOF`(μ, σ, ν) defines a normal distribution family with three parameters. The third parameter ν allows the variance of the distribution to be proportional to a power of the mean. The mean of `NOF`(μ, σ, ν) is μ, while the variance is $\text{Var}(Y) = \sigma^2\mu^\nu$, so the standard deviation is $\sigma\mu^{\nu/2}$. The pdf of the `NOF`(μ, σ, ν) distribution is

$$f_Y(y \,|\, \mu, \sigma, \nu) = \frac{1}{\sqrt{2\pi}\sigma\mu^{\nu/2}} \exp\left[-\frac{(y-\mu)^2}{2\sigma^2\mu^\nu}\right] \tag{1.6}$$

for $-\infty < y < \infty$, where $\mu > 0$, $\sigma > 0$ and $-\infty < \nu < \infty$. The `NOF`(μ, σ, ν) distribution is symmetric about μ. Note that if $Y \sim$ `NOF`(μ, σ, ν), then $Y \sim$ `NO`$(\mu, \sigma\mu^{\nu/2})$.

`NOF`(μ, σ, ν) is appropriate for normally distributed regression type models where the variance of the response variable is proportional to a power of the mean. Models of this type are related to the "pseudo likelihood" models of Carroll and Ruppert [1988], but here a proper likelihood is maximized. The ν parameter is usually modeled as a constant, used as a device to model the variance-mean relationship. Note that, due to the high correlation between the σ and ν parameters, the `method=mixed()` and `c.crit=0.0001` method arguments are strongly recommended in the `gamlss()` fitting function to speed the convergence and avoid converging too early. Alternatively a constant ν can be estimated from its profile function, obtained using the **gamlss** function `prof.dev()`.

Table 1.6: Normal family (of variance-mean relationships) distribution.

$\texttt{NOF}(\mu,\sigma,\nu)$	
Ranges	
Y	$-\infty < y < \infty$
μ	$0 < \mu < \infty$, mean, median, mode
σ	$0 < \sigma < \infty$
ν	$-\infty < \nu < \infty$
Distribution measures	
mean	μ
median	μ
mode	μ
variance	$\sigma^2\mu^\nu$
skewness	0
excess kurtosis	0
MGF	$\exp\left(\mu t + \frac{1}{2}\sigma^2\mu^\nu t^2\right)$
pdf	$\frac{1}{\sqrt{2\pi}\sigma\mu^{\nu/2}} \exp\left[-\frac{(y-\mu)^2}{2\sigma^2\mu^\nu}\right]$
cdf	$\Phi[(y-\mu)/(\sigma\mu^{\nu/2})]$
inverse cdf	$\mu + \sigma\mu^{\nu/2} z_p$ where $z_p = \Phi^{-1}(p)$
Reference	Reparameterize σ to $\sigma\mu^{\nu/2}$ in $\texttt{NO}(\mu,\sigma)$

1.3.3 Power exponential: PE, PE2

First parameterization, $\texttt{PE}(\mu,\sigma,\nu)$

The pdf of the power exponential family distribution, denoted by $\texttt{PE}(\mu,\sigma,\nu)$, is given by

$$f_Y(y \mid \mu,\sigma,\nu) \quad = \quad \frac{\nu \exp\left[-\left|z\right|^\nu\right]}{2c\sigma\Gamma\left(\frac{1}{\nu}\right)} \tag{1.7}$$

for $-\infty < y < \infty$, where $-\infty < \mu < \infty$, $\sigma > 0$, $\nu > 0$ and where $z = (y-\mu)/(c\sigma)$ and $c^2 = \Gamma(1/\nu)[\Gamma(3/\nu)]^{-1}$. Note $S = |Z|^\nu \sim \texttt{GA}(\nu^{-1}, \nu^{1/2})$, where $Z = (Y-\mu)/(c\sigma)$. Note $\texttt{PE}(\mu,\sigma,\nu) = \texttt{PE2}(\mu, c\sigma, \nu)$.

In the parameterization (1.7), used by Nelson [1991], $\mathrm{E}(Y) = \mu$ and $\mathrm{Var}(Y) = \sigma^2$. The $\texttt{PE}(\mu,\sigma,\nu)$ distribution includes the Laplace (i.e. two-sided exponential) and normal, $\texttt{NO}(\mu,\sigma)$, distributions as special cases $\nu = 1$ and $\nu = 2$, respectively, while the uniform distribution is the limiting case as $\nu \to \infty$. The PE(μ, σ, ν) distribution is symmetric about μ, moment leptokurtic for $0 < \nu < 2$, and moment platykurtic for $\nu > 2$.

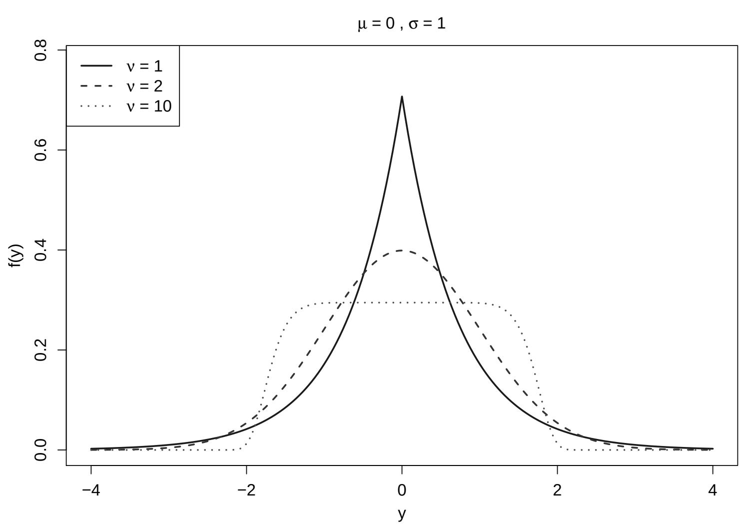

Figure 1.6: The power exponential, PE(μ, σ, ν), distribution, with $\mu = 0$, $\sigma = 1$, and $\nu = 1, 2, 10$.

Table 1.7: Power exponential distribution.

$\texttt{PE}(\mu,\sigma,\nu)$	
Ranges	
Y	$-\infty < y < \infty$
μ	$-\infty < \mu < \infty$, mean, median, mode, location shift par.
σ	$0 < \sigma < \infty$, standard deviation, scaling parameter
ν	$0 < \nu < \infty$, a primary true moment kurtosis parameter
Distribution measures	
mean	μ
median	μ
mode	μ
variance	σ^2
skewness	0
excess kurtosis	$\dfrac{\Gamma\left(5\nu^{-1}\right)\Gamma\left(\nu^{-1}\right)}{\left[\Gamma\left(3\nu^{-1}\right)\right]^2}-3$
pdf	$\dfrac{\nu\exp\left[-\lvert z\rvert^{\nu}\right]}{2c\sigma\Gamma\left(\nu^{-1}\right)}$ where $c^2=\Gamma(\nu^{-1})[\Gamma(3\nu^{-1})]^{-1}$ and $z=(y-\mu)/(c\sigma)$
cdf	$\dfrac{1}{2}\left[1+\dfrac{\gamma\left(\nu^{-1},\lvert z\rvert^{\nu}\right)}{\Gamma\left(\nu^{-1}\right)}\text{sign}(y-\mu)\right]$
inverse cdf	$\begin{cases}\mu-c\sigma\left[F_S^{-1}(1-2p)\right]^{1/\nu} & \text{if } p\le 0.5\\ \mu+c\sigma\left[F_S^{-1}(2p-1)\right]^{1/\nu} & \text{if } p>0.5\\ \text{where } S\sim\texttt{GA}(\nu^{-1},\nu^{1/2}) & \end{cases}$
Reference	Reparameterize σ to $c\sigma$ in $\texttt{PE2}(\mu,\sigma,\nu)$
Note	$\gamma(a,x)=\int_0^x t^{a-1}e^{-t}dt$ is the incomplete gamma function

Second parameterization, PE2

An alternative parameterization, the power exponential type 2 distribution, denoted by $\texttt{PE2}(\mu,\sigma,\nu)$, has pdf

$$f_Y(y\mid\mu,\sigma,\nu) \quad = \quad \frac{\nu\exp[-\lvert z\rvert^{\nu}]}{2\sigma\Gamma\left(\nu^{-1}\right)} \tag{1.8}$$

for $-\infty < y < \infty$, where $-\infty < \mu < \infty$, $\sigma > 0$ and $\nu > 0$ and where $z=(y-\mu)/\sigma$. Note $\texttt{PE2}(\mu,\sigma,\nu)=\texttt{PE}(\mu,\sigma/c,\nu)$.

This is a reparameterization of a version by Subbotin [1923] given in Johnson et al. [1995] Section 24.6, p195-196, equation (24.83).The cdf is given by

$$F_Y(y) = \frac{1}{2}\left[1 + F_S(s)\text{sign}(z)\right]$$

where $S = |Z|^\nu$ has a gamma, $\texttt{GA}(\nu^{-1}, \nu^{1/2})$, distribution with pdf $f_S(s) = s^{(1/\nu)-1}\exp(-s)/\Gamma\left(\nu^{-1}\right)$ and $Z = (Y - \mu)/\sigma$, from which the cdf and inverse cdf results in Table 1.8 was obtained.

The $\texttt{PE2}(\mu, \sigma, \nu)$ distribution is symmetric about μ. It includes the (reparameterized) normal and Laplace (i.e. two-sided exponential) distributions as special cases when $\nu = 2$ and $\nu = 1$, respectively, and the uniform distribution as a limiting case as $\nu \to \infty$. The $\texttt{PE2}(\mu, \sigma, \nu)$ distribution is moment leptokurtic for $0 < \nu < 2$ and moment platykurtic for $\nu > 2$.

Table 1.8: Second parameterization of power exponential distribution.

	$\texttt{PE2}(\mu, \sigma, \nu)$
	Ranges
Y	$-\infty < y < \infty$
μ	$-\infty < \mu < \infty,$ mean, median, mode, location shift par.
σ	$0 < \sigma < \infty,$ scaling parameter
ν	$0 < \nu < \infty,$ a primary true moment kurtosis parameter
	Distribution measures
mean	μ
median	μ
mode	μ
variance [a]	$\dfrac{\sigma^2}{c^2}$ where $c^2 = \Gamma(\nu^{-1})[\Gamma(3\nu^{-1})]^{-1}$
skewness	0
excess kurtosis [a]	$\dfrac{\Gamma\left(5\nu^{-1}\right)\Gamma\left(\nu^{-1}\right)}{\left[\Gamma\left(3\nu^{-1}\right)\right]^2} - 3$
pdf [a]	$\dfrac{\nu \exp[-\lvert z\rvert^{\nu}]}{2\sigma\Gamma\left(\nu^{-1}\right)}$
cdf	$\dfrac{1}{2}\left[1 + \dfrac{\gamma\left(\nu^{-1}, \lvert z\rvert^{\nu}\right)}{\Gamma\left(\nu^{-1}\right)}\text{sign}(y-\mu)\right]$ where $z = \dfrac{y-\mu}{\sigma}$
inverse cdf	$\begin{cases} \mu - \sigma\left[F_S^{-1}(1-2p)\right]^{1/\nu} & \text{if } p \le 0.5 \\ \mu + \sigma\left[F_S^{-1}(2p-1)\right]^{1/\nu} & \text{if } p > 0.5 \\ \text{where } S \sim \texttt{GA}(\nu^{-1}, \nu^{1/2}) & \end{cases}$
Reference	[a] Johnson et al. [1995] Section 24.6, p195-196, equation (24.83), reparameterized by $\theta = \mu$, $\phi = 2^{-1/\nu}\sigma$ and $\delta = 2/\nu$ and hence $\mu = \theta$, $\sigma = \phi 2^{\delta/2}$ and $\nu = 2/\delta$.

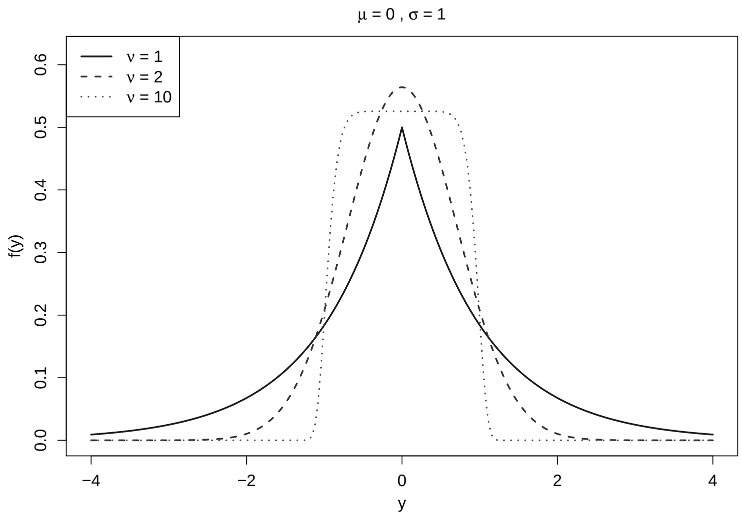

Figure 1.7: The power exponential type 2, $\texttt{PE2}(\mu, \sigma, \nu)$, distribution, with $\mu = 0$, $\sigma = 1$, and $\nu = 1, 2, 10$.

1.3.4 Skew normal type 1: SN1

The pdf of a skew normal type 1 distribution, denoted by $\texttt{SN1}(\mu, \sigma, \nu)$, is given by

$$f_Y(y \,|\, \mu, \sigma, \nu) = \frac{2}{\sigma}\, \phi(z)\, \Phi(\nu z) \tag{1.9}$$

for $-\infty < y < \infty$, where $-\infty < \mu < \infty$, $\sigma > 0$ and $-\infty < \nu < \infty$, and where $z = (y - \mu)/\sigma$ and $\phi(\cdot)$ and $\Phi(\cdot)$ are the pdf and cdf of a standard normal $\texttt{NO}(0, 1)$ variable, respectively [Azzalini, 1985]. See also Section **??**. The skew normal type 1 distribution is a special case of the skew exponential power type 1 distribution where $\tau = 2$, i.e. $\texttt{SN1}(\mu, \sigma, \nu) = \texttt{SEP1}(\mu, \sigma, \nu, 2)$.

The $\texttt{SN1}(\mu, \sigma, \nu)$ distribution includes the normal $\texttt{NO}(\mu, \sigma)$ as a special case when $\nu = 0$, the half normal as a limiting case as $\nu \to \infty$ and the reflected half normal (i.e. $Y < \mu$) as a limiting case as $\nu \to -\infty$. The $\texttt{SN1}(\mu, \sigma, \nu)$ distribution is positively moment skewed if $\nu > 0$, and negatively moment skewed if $\nu < 0$.

Note if $Y \sim \texttt{SN1}(\mu, \sigma, \nu)$ then $-Y \sim \texttt{SN1}(-\mu, \sigma, -\nu)$ is a reflection of the distribution of Y about zero. Also $W = (2\mu - Y) \sim \texttt{SN1}(\mu, \sigma, -\nu)$ is a reflection of the distribution of Y about μ.

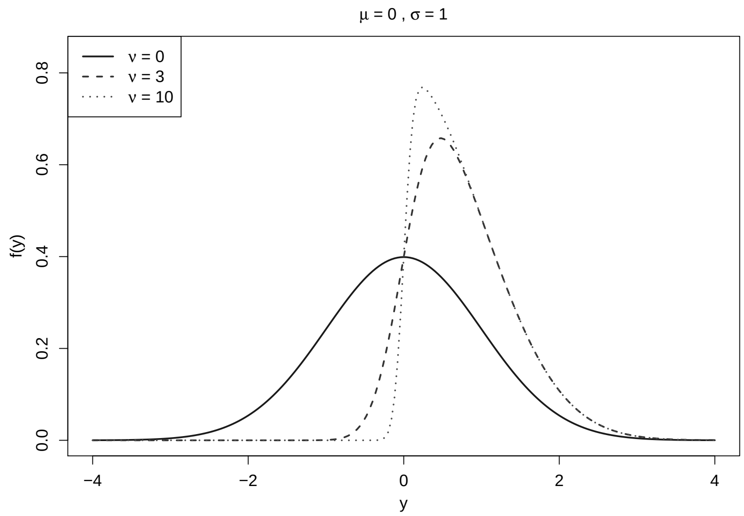

Figure 1.8: The skew normal type 1, $\texttt{SN1}(\mu, \sigma, \nu)$, distribution with $\mu = 0$, $\sigma = 1$, and $\nu = 0, 3, 10$. It is symmetric if $\nu = 0$. Changing ν to $-\nu$ reflects the distribution about $y = \mu$, i.e. $y = 0$ above.

Table 1.9: Skew normal type 1 distribution.

	$\texttt{SN1}(\mu, \sigma, \nu)$
	Ranges
Y	$-\infty < y < \infty$
μ	$-\infty < \mu < \infty$, location shift parameter
σ	$0 < \sigma < \infty$, scaling parameter
ν	$-\infty < \nu < \infty$, a primary true moment skewness parameter
	Distribution measures
mean	$\mu + \sigma\nu \left[2(1+\nu^2)^{-1}\pi^{-1}\right]^{1/2}$
variance	$\sigma^2 \left[1 - 2\nu^2(1+\nu^2)^{-1}\pi^{-1}\right]$
skewness	$\frac{1}{2}(4-\pi)\left[\frac{\pi}{2}(1+\nu^{-2}) - 1\right]^{-3/2} \text{sign}(\nu)$
excess kurtosis	$2(\pi - 3)\left[\frac{\pi}{2}(1+\nu^{-2}) - 1\right]^{-2}$
MGF	$2\exp\left(\mu t + \frac{1}{2}\sigma^2 t^2\right)\Phi\left(\frac{\sigma\nu t}{\sqrt{1+\nu^2}}\right)$
pdf	$\frac{2}{\sigma}\phi(z)\Phi(\nu z)$ where $z = (y-\mu)/\sigma$
Reference	From Azzalini [1985], p172 and p174, where $\lambda = \nu$ and here $Y = \mu + \sigma Z$

1.3.5 Skew normal type 2: SN2

The pdf of the skew normal type 2 distribution, denoted by SN2(μ, σ, ν), is given by

$$f_Y(y \mid \mu, \sigma\, \nu, \tau) = \begin{cases} \dfrac{c}{\sigma} \exp\left[-\dfrac{1}{2}(\nu z)^2\right] & \text{if } y < \mu \\[2ex] \dfrac{c}{\sigma} \exp\left[-\dfrac{1}{2}\left(\dfrac{z}{\nu}\right)^2\right] & \text{if } y \geq \mu \end{cases} \tag{1.10}$$

for $-\infty < y < \infty$, where $-\infty < \mu < \infty$, $\sigma > 0$ and $\nu > 0$, and where $z = (y - \mu)/\sigma$ and $c = \sqrt{2}\nu/[\sqrt{\pi}(1 + \nu^2)]$. This distribution is also called the two-piece normal distribution [Johnson et al., 1994, Section 10.3, p173]. SN2 is a 'scale-spliced' distribution, see Section **??**.

The skew normal type 2 distribution is a special case of the skew exponential power type 3 distribution where $\tau = 2$, i.e. SN2$(\mu, \sigma, \nu) =$ SEP3$(\mu, \sigma, \nu, 2)$. Also SN2(μ, σ, ν) is the limiting case of ST3(μ, σ, ν, τ) as $\tau \to \infty$. The SN2(μ, σ, ν) distribution includes the normal NO(μ, σ) as a special case when $\nu = 1$, and is positively moment skewed if $\nu > 1$ and negatively moment skewed if $0 < \nu < 1$.

Note if $Y \sim$ SN2(μ, σ, ν) then $-Y \sim$ SN2$(-\mu, \sigma, \nu^{-1})$ is a reflection of the distribution of Y about zero. Also $W = (2\mu - Y) \sim$ SN2(μ, σ, ν^{-1}) is a reflection of the distribution of Y about μ.

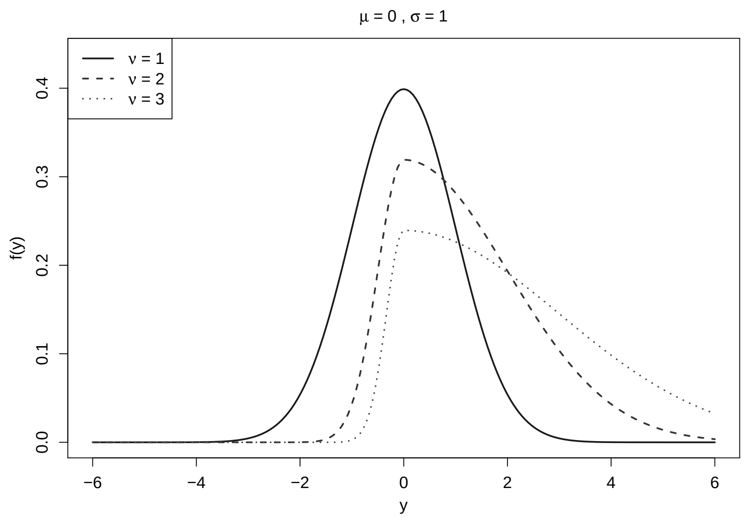

Figure 1.9: The skew normal type 2, $\texttt{SN2}(\mu, \sigma, \nu)$, distribution with $\mu = 0$, $\sigma = 1$, and $\nu = 1, 2, 3$. It is symmetric if $\nu = 1$. Changing ν to $1/\nu$ reflects the distribution about $y = \mu$, i.e. $y = 0$ above.

Table 1.10: Skew normal type 2 distribution.

	$\texttt{SN2}(\mu, \sigma, \nu)$
	Ranges
Y	$-\infty < y < \infty$
μ	$-\infty < \mu < \infty$, mode, location shift parameter
σ	$0 < \sigma < \infty$, scaling parameter
ν	$0 < \nu < \infty$, skewness parameter
	Distribution measures
mean [a]	$\mu + \sigma \mathrm{E}(Z) = \mu + \sigma \frac{\sqrt{2}}{\sqrt{\pi}}(\nu - \nu^{-1})$ where $Z = (Y - \mu)/\sigma$
median [a2]	$\begin{cases} \mu + \frac{\sigma}{\nu}\Phi^{-1}\left(\frac{1+\nu^2}{2}\right) & \text{if } \nu \le 1 \\ \mu + \sigma\nu\Phi^{-1}\left(\frac{3\nu^2 - 1}{4\nu^2}\right) & \text{if } \nu > 1 \end{cases}$
mode	μ
variance [a]	$\sigma^2 \mathrm{Var}(Z) = \sigma^2\{(\nu^2 + \nu^{-2} - 1) - [\mathrm{E}(Z)]^2\}$
skewness [a]	$\begin{cases} \mu_{3Y}/[\mathrm{Var}(Y)]^{1.5} \text{ where} \\ \mu_{3Y} = \sigma^3 \mu_{3Z} = \sigma^3\{\mu'_{3Z} - 3\mathrm{Var}(Z)\mathrm{E}(Z) - [\mathrm{E}(Z)]^3\} \\ \mu'_{3Z} = \mathrm{E}(Z^3) = \frac{2\sqrt{2}(\nu^4 - \nu^{-4})}{\sqrt{\pi}(\nu + \nu^{-1})} \end{cases}$
excess kurtosis [a]	$\begin{cases} \mu_{4Y}/[\mathrm{Var}(Y)]^2 - 3 \text{ where} \\ \mu_{4Y} = \sigma^4 \mu_{4Z} = \sigma^4\{\mu'_{4Z} - 4\mu'_{3Z}\mathrm{E}(Z) + 6\mathrm{Var}(Z)[\mathrm{E}(Z)]^2 \\ \quad +3[E(Z)]^4\} \\ \mu'_{4Z} = \mathrm{E}(Z^4) = 3(\nu^5 + \nu^{-5})/(\nu + \nu^{-1}) \end{cases}$
pdf [a]	$\begin{cases} \frac{c}{\sigma}\exp\left[-\frac{1}{2}(\nu z)^2\right] & \text{if } y < \mu \\ \frac{c}{\sigma}\exp\left[-\frac{1}{2}\left(\frac{z}{\nu}\right)^2\right] & \text{if } y \ge \mu \\ \text{where } z = (y-\mu)/\sigma \text{ and } \quad c = \frac{\sqrt{2}\nu}{\sqrt{\pi}(1+\nu^2)} \end{cases}$
cdf [a2]	$\begin{cases} \frac{2\Phi\left[\nu(y-\mu)/\sigma\right]}{(1+\nu^2)} & \text{if } y < \mu \\ \frac{1}{(1+\nu^2)}\left\{1 - \nu^2 + 2\nu^2\Phi\left[(y-\mu)/(\sigma\nu)\right]\right\} & \text{if } y \ge \mu \end{cases}$
inverse cdf [a2]	$\begin{cases} \mu + \frac{\sigma}{\nu}\Phi^{-1}\left[\frac{p(1+\nu^2)}{2}\right] & \text{if } p \le (1+\nu^2)^{-1} \\ \mu + \sigma\nu\Phi^{-1}\left[\frac{p(1+\nu^2) - 1 + \nu^2}{2\nu^2}\right] & \text{if } p > (1+\nu^2)^{-1} \end{cases}$
Note	[a] Set $\tau = 2$ in $\texttt{SEP3}(\mu, \sigma, \nu, \tau)$ [a2] Set $\tau \to \infty$ in $\texttt{ST3}(\mu, \sigma, \nu, \tau)$

1.3.6 t family: TF

The pdf of the t family distribution, denoted by $\texttt{TF}(\mu, \sigma, \nu)$, is given by

$$f_Y(y \mid \mu, \sigma, \nu) = \frac{1}{\sigma B\left(1/2, \nu/2\right) \nu^{1/2}} \left[1 + \frac{(y-\mu)^2}{\sigma^2 \nu}\right]^{-(\nu+1)/2} \tag{1.11}$$

for $-\infty < y < \infty$, where $-\infty < \mu < \infty$, $\sigma > 0$ and $\nu > 0$, where $B(a, b) = \Gamma(a)\Gamma(b)/\Gamma(a+b)$ is the beta function. Note that $T = (Y - \mu)/\sigma$ has a standard t distribution with ν degrees of freedom, i.e. $T \sim \texttt{TF}(0, 1, \nu) = t_\nu$, with pdf given by Johnson et al. [1995], p363, equation (28.2). Note if $T \sim \texttt{TF}(0, 1, \nu) = t_\nu$ then $W = \nu/(\nu + T^2) \sim \texttt{BEo}(\nu/2, 1/2)$ and $1 - W = T^2/(\nu + T^2) \sim \texttt{BEo}(1/2, \nu/2)$.

The $\texttt{TF}(\mu, \sigma, \nu)$ distribution is symmetric about μ. It is suitable for modeling leptokurtic data, since it has higher kurtosis than the normal distribution. See Lange et al. [1989] for modeling a response variable using the $\texttt{TF}(\mu, \sigma, \nu)$ distribution, including parameter estimation.

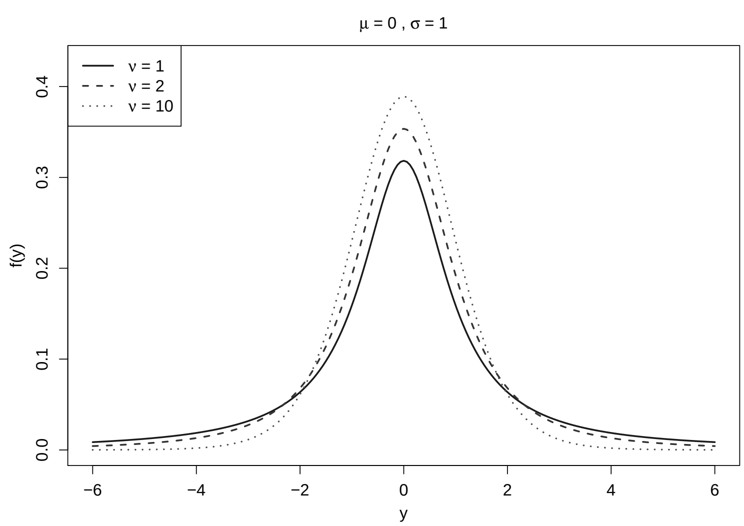

Figure 1.10: The t family, $\texttt{TF}(\mu, \sigma, \nu)$, distribution, with $\mu = 0$, $\sigma = 1$, and $\nu = 1, 2, 10$.

Table 1.11: t family distribution.

	$\texttt{TF}(\mu, \sigma, \nu)$
	Ranges
Y	$-\infty < y < \infty$
μ	$-\infty < \mu < \infty,$ mean, median, mode, location shift parameter
σ	$0 < \sigma < \infty,$ scaling parameter
ν	$\begin{cases} 0 < \nu < \infty, \ \nu \text{ is a primary true moment kurtosis parameter} \\ (\text{for } \nu > 4), \text{ and the degrees of freedom parameter} \end{cases}$
	Distribution measures
mean	$\begin{cases} \mu & \text{if } \nu > 1, \\ \text{undefined} & \text{if } \nu \leq 1 \end{cases}$
median	μ
mode	μ
variance [a]	$\begin{cases} \dfrac{\sigma^2\nu}{v-2} & \text{if } \nu > 2 \\ \infty & \text{if } 1 < \nu \leq 2 \\ \text{undefined} & \text{if } \nu \leq 1 \end{cases}$
skewness	$\begin{cases} 0 & \text{if } \nu > 3 \\ \text{undefined} & \text{if } \nu \leq 3 \end{cases}$
excess kurtosis [a]	$\begin{cases} \dfrac{6}{v-4} & \text{if } \nu > 4 \\ \infty & \text{if } 2 < \nu \leq 4 \\ \text{undefined} & \text{if } \nu \leq 2 \end{cases}$
pdf [a]	$\dfrac{1}{\sigma B\left(1/2, \nu/2\right) \nu^{1/2}} \left[1 + \dfrac{(y-\mu)^2}{\sigma^2 \nu}\right]^{-(\nu+1)/2}$
cdf [a]	$0.5 + \dfrac{B\left(1/2, \nu/2, z^2/[\nu + z^2]\right)}{2B(1/2, \nu/2)} \text{sign}(z) \text{ where } z = \dfrac{(y-\mu)}{\sigma}$
inverse cdf	$\begin{cases} \mu + \sigma t_{p,\nu} \text{ where } t_{p,\nu} = F_T^{-1}(p) \text{ is the } p \text{ quantile or } 100p \text{ centile value} \\ \text{of } T \sim t_\nu, \text{ i.e. } \mathrm{P}(T < t_{p,\nu}) = p \end{cases}$
Reference	[a] See Johnson et al. [1995], Section 28.2, p363-365.
Note	$B(a, b, x) = \int_o^x t^{a-1}(1-t)^{b-1} dt$ is the incomplete beta function.

1.3.7 t family type 2: TF2

The pdf of the t family type 2 distribution, denoted by $\texttt{TF2}(\mu, \sigma, \nu)$, is

$$f_Y(y \mid \mu, \sigma, \nu) = \frac{1}{\sigma B\left(1/2, \nu/2\right) (\nu-2)^{1/2}} \left[1 + \frac{(y-\mu)^2}{\sigma^2(\nu-2)}\right]^{-(\nu+1)/2} \qquad (1.12)$$

for $-\infty < y < \infty$, where $-\infty < \mu < \infty$, $\sigma > 0$ and $\nu > 2$, where $B(a, b) = \Gamma(a)\Gamma(b)/\Gamma(a+b)$ is the beta function. Note that since $\nu > 2$, the mean $\mathrm{E}(Y) = \mu$ and variance $\mathrm{Var}(Y) = \sigma^2$ of $Y \sim \mathtt{TF2}(\mu, \sigma, \nu)$ are always defined and finite.

The $\mathtt{TF2}(\mu, \sigma, \nu)$ distribution is symmetric about μ. It is suitable for modeling leptokurtic data, since it has higher kurtosis than the normal distribution.

Table 1.12: t family type 2 distribution.

	$\mathtt{TF2}(\mu, \sigma, \nu)$
	Ranges
Y	$-\infty < y < \infty$
μ	$-\infty < \mu < \infty$, mean, median, mode, location shift parameter
σ	$0 < \sigma < \infty$, standard deviation, scaling parameter
ν	$\begin{cases} 2 < \nu < \infty, \ \nu \text{ is a primary true moment kurtosis parameter} \\ (\text{for } \nu > 4), \ \text{and the degrees of freedom parameter} \end{cases}$
	Distribution measures
mean	μ
median	μ
mode	μ
variance	σ^2
skewness	$\begin{cases} 0 & \text{if } \nu > 3 \\ \text{undefined} & \text{if } 2 < \nu \leq 3 \end{cases}$
excess kurtosis	$\begin{cases} \dfrac{6}{v-4} & \text{if } \nu > 4 \\ \infty & \text{if } 2 < \nu \leq 4 \end{cases}$
pdf	$\dfrac{1}{\sigma B\left(1/2, \nu/2\right) (\nu-2)^{1/2}} \left[1 + \dfrac{(y-\mu)^2}{\sigma^2(\nu-2)}\right]^{-(\nu+1)/2}$
cdf	$0.5 + \dfrac{B\left(1/2, \nu/2, z^2/[\nu+z^2]\right)}{2B(1/2, \nu/2)} \mathrm{sign}(z)$ where $z = \dfrac{(y-\mu)\nu^{1/2}}{\sigma(\nu-2)^{1/2}}$
inverse cdf	$\mu + \sigma(\nu-2)^{1/2}\nu^{-1/2} t_{p,\nu}$
Note	Reparameterize σ to $\sigma(\nu-2)^{1/2}\nu^{-1/2}$ in $\mathtt{TF}(\mu, \sigma, \nu)$.

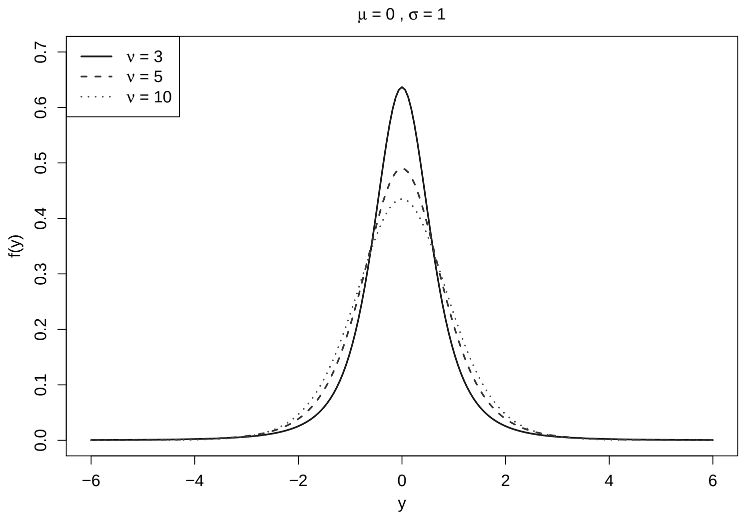

Figure 1.11: The t family type 2, `TF2`(μ,σ,ν), distribution, with $\mu=0$, $\sigma=1$, and $\nu=3,5,10$.

1.4 Continuous four-parameter distributions on $(-\infty,\infty)$

1.4.1 Exponential generalized beta type 2: `EGB2`

The pdf of the exponential generalized beta type 2 distribution, denoted by `EGB2`(μ,σ,ν,τ), is given by

$$f_Y(y\,|\,\mu,\sigma,\nu,\tau)=e^{\nu z}\{|\sigma|B(\nu,\tau)[1+e^{z}]^{\nu+\tau}\}^{-1} \tag{1.13}$$

for $-\infty<y<\infty$, where $-\infty<\mu<\infty$, $\sigma>0$, $\nu>0$ and $\tau>0$, and where $z=(y-\mu)/\sigma$, [McDonald and Xu, 1995, p141]. Note that McDonald and Xu [1995] appear to allow $\sigma<0$, however this is unnecessary since `EGB2`$(\mu,-\sigma,\nu,\tau)=$ `EGB2`(μ,σ,τ,ν). So we assume $\sigma>0$ and $|\sigma|$ can be replaced by σ in (1.13).

If $Y\sim$ `EGB2`(μ,σ,ν,τ), then $-Y\sim$ `EGB2`$(-\mu,\sigma,\tau,\nu)$ is a reflection of the distribution of Y about zero. Also $W=(2\mu-Y)\sim$ `EGB2`(μ,σ,τ,ν) is a reflection the distribution of Y about μ, changing the skewness from positive to negative, or vice-versa. The `EGB2` distribution is also called the type IV generalized logistic distribution [Johnson et al., 1995, Section 23.10, p142].

If $Y\sim$ `EGB2`(μ,σ,ν,τ) then $R=\{1+\exp[-(Y-\mu)/\sigma]\}^{-1}\sim$ `BEo`(ν,τ) from which the cdf in Table 1.13 is obtained. Also $Z=(\tau/\nu)\exp\left[(Y-\mu)/\sigma\right]\sim F_{2\nu,2\tau}$, an F distribution, defined in equation (**??**). The `EGB2`(μ,σ,ν,τ) distribution is symmetric if $\nu=\tau$, and is positively moment skewed if $\nu>\tau$ and negatively moment skewed if $\nu < \tau$. EGB2(μ, σ, ν, τ) is always moment leptokurtic with the normal distribution a limiting case [McDonald, 1996, p 437]. Johnson et al. [1995, p141] indicates that $\sqrt{2/\nu}(Y-\mu)/\sigma$ has a standard normal limiting distribution as $\nu = \tau \to \infty$. Figures **??**(e), **??**(d) and **??**(d) show moment and centile kurtosis-skewness plots for EGB2(μ, σ, ν, τ). The figures indicate that the EGB2 distribution is restricted in kurtosis to be higher (but not extremely higher) than of the normal distribution.

Table 1.13: Exponential generalized beta type 2 distribution.

EGB2(μ, σ, ν, τ)	
Ranges	
Y	$-\infty < y < \infty$
μ	$-\infty < \mu < \infty$, location shift parameter
σ	$0 < \sigma < \infty$, scaling parameter
ν	$0 < \nu < \infty$
τ	$0 < \tau < \infty$
Distribution measures	
mean [a2]	$\mu + \sigma[\Psi(\nu) - \Psi(\tau)]$
mode	$\mu + \sigma \log(\nu/\tau)$
variance [a2]	$\sigma^2\left[\Psi^{(1)}(\nu) + \Psi^{(1)}(\tau)\right]$
skewness [a2]	$\dfrac{\Psi^{(2)}(\nu) - \Psi^{(2)}(\tau)}{[\Psi^{(1)}(\nu) + \Psi^{(1)}(\tau)]^{1.5}}$
excess kurtosis[a2]	$\dfrac{\Psi^{(3)}(\nu) + \Psi^{(3)}(\tau)}{[\Psi^{(1)}(\nu) + \Psi^{(1)}(\tau)]^{2}}$
MGF [a]	$\dfrac{e^{\mu t} B(\nu + \sigma t, \tau - \sigma t)}{B(\nu, \tau)}$
pdf [a]	$e^{\nu z}\{\lvert\sigma\rvert B(\nu, \tau)[1 + e^z]^{\nu+\tau}\}^{-1}$, where $z = (y-\mu)/\sigma$
cdf	$\dfrac{B(\nu, \tau, c)}{B(\nu, \tau)}$, where $c = \{1 + \exp[-(y-\mu)/\sigma]\}^{-1}$
Reference	[a] McDonald and Xu [1995], p140-141, p150-151, where $\delta = \mu, \sigma = \sigma, p = \nu$ and $q = \tau$. [a2] McDonald [1996], p436-437, where $\delta = \mu, \sigma = \sigma$, $p = \nu$ and $q = \tau$
Note	$\Psi^{(r)}(x) = d^{(r)}\Psi(x)/dx^{(r)}$ is the rth derivative of the psi (or digamma) function $\Psi(x)$

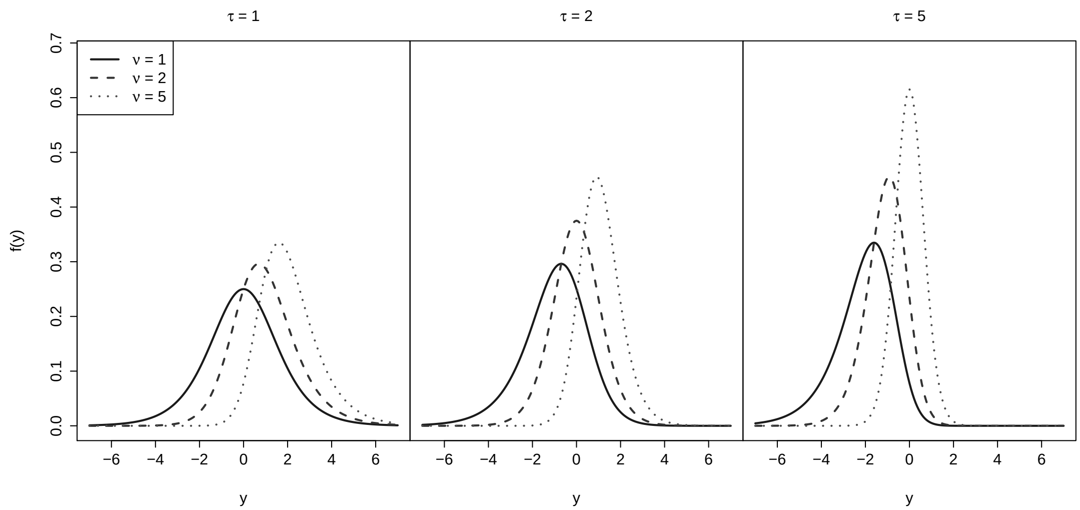

Figure 1.12: The exponential generalized beta type 2, $\texttt{EGB2}(\mu, \sigma, \nu, \tau)$, distribution, with $\mu = 0$, $\sigma = 1$, $\nu = 1, 2, 5$, and $\tau = 1, 2, 5$. It is symmetric if $\nu = \tau$. Interchanging ν and τ reflects the distribution about $y = \mu$, i.e. $y = 0$ above.

1.4.2 Generalized t: GT

This pdf of the generalized t distribution, denoted by $\texttt{GT}(\mu, \sigma, \nu, \tau)$, is given by

$$f_Y(y \,|\, \mu, \sigma\, \nu, \tau) = \tau \left\{ 2\sigma \nu^{1/\tau} B\left(1/\tau, \nu\right) \left[1 + |z|^{\tau}/\nu\right]^{\nu + (1/\tau)} \right\}^{-1} \tag{1.14}$$

for $-\infty < y < \infty$, where $-\infty < \mu < \infty$, $\sigma > 0$, $\nu > 0$ and $\tau > 0$, and where $z = (y - \mu)/\sigma$, [McDonald and Newey [1988] p430, equation (2.1), where $q = \nu$ and $p = \tau$]. See also Butler et al. [1990] and McDonald [1991]. Note that the t family distribution, $\texttt{TF}(\mu, \sigma, \nu)$, is a special case of the generalized t distribution given by $\texttt{GT}(\mu, \sqrt{2}\sigma, \nu/2, 2)$. The power exponential distribution $\texttt{PE2}(\mu, \sigma, \tau)$ is a limiting distribution of $\texttt{GT}(\mu, \sigma, \nu, \tau)$ as $\nu \to \infty$.

The $\texttt{GT}(\mu, \sigma, \nu, \tau)$ distribution is symmetric about μ and can be moment leptokurtic or platykurtic. It has two kurtosis parameters ν and τ. Parameter τ affects the peakedness around $y = \mu$, with $\tau \leq 1$ resulting in a spike in the pdf at $y = \mu$, and $\tau > 1$ resulting in a flat (i.e. zero derivative) pdf at $y = \mu$. For fixed τ, decreasing ν increases the tail heaviness (and vice-versa), since the pdf of $\texttt{GT}(\mu, \sigma, \nu, \tau)$ has order $O(|y|^{-\nu\tau-1})$ as $|y| \to \infty$, the same order as a t distribution with $\nu\tau$ degrees of freedom, and so the tails are heavier as $\nu\tau$ decreases and are always heavier than the normal distribution.

Table 1.14: Generalized t distribution.

	$\mathtt{GT}(\mu, \sigma, \nu, \tau)$
	Ranges
Y	$-\infty < y < \infty$
μ	$-\infty < \mu < \infty$, mean, median, mode, location shift parameter
σ	$0 < \sigma < \infty$, scaling parameter
ν	$0 < \nu < \infty$, first kurtosis parameter, ν is a true moment kurtosis parameter for $\nu\tau > 4$
τ	$0 < \tau < \infty$, second kurtosis parameter, τ is a true moment kurtosis parameter for $\nu\tau > 4$
	Distribution measures
mean	$\begin{cases} \mu & \text{if } \nu\tau > 1 \\ \text{undefined} & \text{if } \nu\tau \le 1 \end{cases}$
median	μ
mode	μ
variance [a2]	$\begin{cases} \dfrac{\sigma^2 \nu^{2/\tau} B(3\tau^{-1}, \nu - 2\tau^{-1})}{B(\tau^{-1}, \nu)} & \text{if } \nu\tau > 2 \\ \infty & \text{if } 1 < \nu\tau \le 2 \\ \text{undefined} & \text{if } \nu\tau \le 1 \end{cases}$
skewness	$\begin{cases} 0 & \text{if } \nu\tau > 3 \\ \text{undefined} & \text{if } \nu\tau \le 3 \end{cases}$
excess kurtosis [a2]	$\begin{cases} \dfrac{B(5\tau^{-1}, \nu - 4\tau^{-1}) B(\tau^{-1}, \nu)}{[B(3\tau^{-1}, \nu - 2\tau^{-1})]^2} - 3 & \text{if } \nu\tau > 4 \\ \infty & \text{if } 2 < \nu\tau \le 4 \\ \text{undefined} & \text{if } \nu\tau \le 2 \end{cases}$
pdf [a]	$\tau \left\{ 2\sigma\nu^{1/\tau} B\left(1/\tau, \nu\right) \left[1 + \|z\|^{\tau}/\nu\right]^{\nu + (1/\tau)} \right\}^{-1}$, where $z = (y - \mu)/\sigma$
Reference	[a] McDonald and Newey [1988], p430, equation (2.1), where $q = \nu$ and $p = \tau$. [a2] McDonald [1991], p 274, where $q = \nu$ and $p = \tau$.

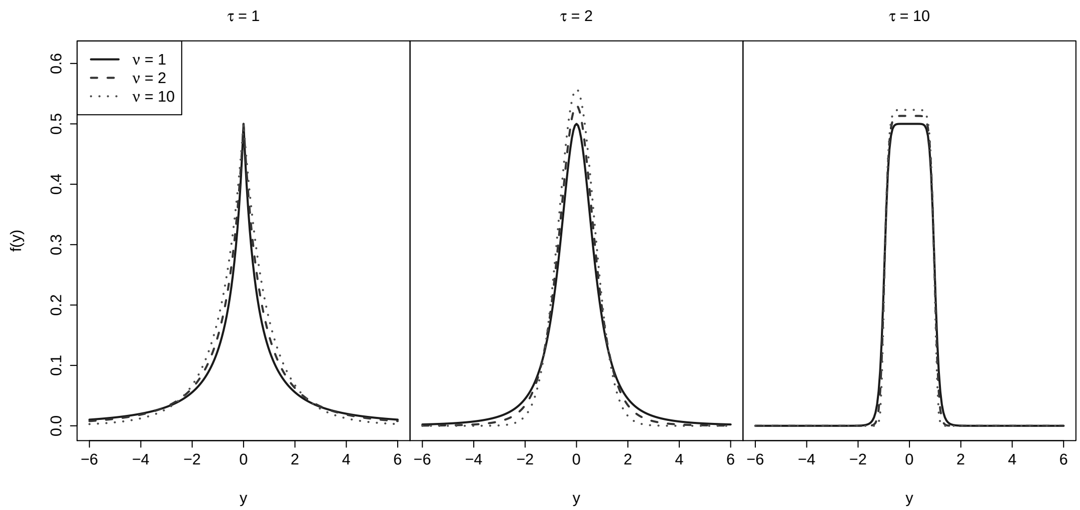

Figure 1.13: The generalized t, $\texttt{GT}(\mu, \sigma, \nu, \tau)$, distribution, with $\mu = 0$, $\sigma = 1$, $\nu = 1, 2, 10$, and $\tau = 1, 2, 10$. It is always symmetric.

1.4.3 Johnson's SU: `JSUo`, `JSU`

First parameterization, `JSUo`

This is the original parameterization of the Johnson's S_u distribution [Johnson, 1949] and is denoted by $\texttt{JSUo}(\mu, \sigma, \nu, \tau)$. Its pdf is given by

$$f_Y(y \mid \mu, \sigma\, \nu, \tau) = \frac{\tau}{\sigma (s^2+1)^{1/2} \sqrt{2\pi}} \; \exp\left[-\frac{1}{2} z^2 \right] \tag{1.15}$$

for $-\infty < y < \infty$, where $-\infty < \mu < \infty$, $\sigma > 0$, $-\infty < \nu < \infty$ and $\tau > 0$, and where

$$z = \nu + \tau \sinh^{-1}(s) = \nu + \tau \log\left[s + (s^2+1)^{1/2} \right], \tag{1.16}$$

where $s = (y - \mu)/\sigma$, [Johnson [1949] p162, equation (33) and p152, where $\xi = \mu, \lambda = \sigma, \gamma = \nu, \delta = \tau$ and $x = y, y = s, z = z$]. Hence $s = \sinh\left[(z-\nu)/\tau\right] = \frac{1}{2}\left\{\exp\left[(z-\nu)/\tau\right] - \exp\left[-(z-\nu)/\tau\right]\right\}$ and $y = \mu + \sigma s$. Note that $Z \sim \texttt{NO}(0, 1)$, where $Z = \nu + \tau \sinh^{-1}[(Y - \mu)/\sigma]$, from which the results for the cdf, inverse cdf, and median in Table 1.15 are obtained. Also

$$\mathrm{E}\left[\frac{(Y-\mu)^r}{\sigma^r} \right] = \frac{1}{2^r} \sum_{j=0}^{r} (-1)^{r-j} C_j^r \exp\left[\frac{1}{2\tau^2}(r-2j)^2 + \frac{\nu}{\tau}(r-2j) \right] \tag{1.17}$$

for $r = 1, 2, 3, ..,$ where $C_j^r = r!/[j!(r-j)!]$.

The parameter ν affects the skewness of the distribution, which is symmetric if $\nu = 0$, positively (moment, centile, and van Zwet) skew if $\nu < 0$, and negatively skew if $\nu > 0$. Parameter ν is a true van Zwet skewness parameter (see Section **??**) and hence a true centile skewness parameter and a true moment skewness parameter. Decreasing ν (for fixed τ) increases the moment and centile skewness. Chapter **??** shows moment and centile kurtosis-skewness plots for `JSU`(μ, σ, ν, τ). The plots for `JSUo`$(\mu, \sigma, -\nu, \tau)$ are the same.

The distribution is always (centile and Balanda-MacGillivray) leptokurtic. The parameter τ affects the kurtosis of the distribution and as $\tau \to \infty$ the distribution tends to the normal distribution. Parameter τ is a primary true Balanda-MacGillivray kurtosis parameter (see Section **??**) and hence it is a primary true centile kurtosis parameter. Increasing τ decreases the centile (and Balanda-MacGillivray) kurtosis.

If $Y \sim$ `JSUo`(μ, σ, ν, τ) then $-Y \sim$ `JSUo`$(-\mu, \sigma, -\nu, \tau)$ is a reflection of the distribution of Y about zero. Also $W = (2\mu - Y) \sim$ `JSUo`$(\mu, \sigma, -\nu, \tau)$ is a reflection of the distribution of Y about μ, changing the skewness from positive to negative (or vice-versa).

Table 1.15: Original parameterization Johnson's S_u distribution.

$\texttt{JSUo}(\mu, \sigma, \nu, \tau)$	
Ranges	
Y	$-\infty < y < \infty$
μ	$-\infty < \mu < \infty,$ location shift parameter
σ	$0 < \sigma < \infty,$ scaling parameter
ν	$-\infty < \nu < \infty,$ true van Zwet skewness parameter
τ	$0 < \tau < \infty,$ primary true Balanda-MacGillivray kurtosis parameter
Distribution measures	
mean [a]	$\mu - \sigma\omega^{1/2}\sinh(\nu/\tau),$ where $w = \exp(1/\tau^2)$
median	$\mu - \sigma\sinh(\nu/\tau)$
variance [a]	$\frac{1}{2}\sigma^2(\omega-1)[\omega\cosh(2\nu/\tau)+1]$
skewness [a]	$\begin{cases} \mu_3/[\mathrm{Var}(Y)]^{1.5} \text{ where} \\ \mu_3 = -\frac{1}{4}\sigma^3\omega^{1/2}(\omega-1)^2[\omega(\omega+2)\sinh(3\nu/\tau)+3\sinh(\nu/\tau)] \end{cases}$
excess kurtosis [a]	$\begin{cases} \mu_4/[\mathrm{Var}(Y)]^2 - 3 \text{ where} \\ \mu_4 = \frac{1}{8}\sigma^4(\omega-1)^2[\omega^2(\omega^4+2\omega^3+3\omega^2-3)\cosh(4\nu/\tau) \\ +4\omega^2(\omega+2)\cosh(2\nu/\tau)+3(2\omega+1)] \end{cases}$
cdf	$\Phi(\nu + \tau\sinh^{-1}[(y-\mu)/\sigma])$
inverse cdf	$\mu + \sigma\sinh[(z_p - \nu)/\tau]$ where $z_p = \Phi^{-1}(p)$
Reference	[a] Johnson [1949], p163, equation (37), p152 and p162 where $\xi = \mu, \lambda = \sigma$, $\gamma = \nu, \delta = \tau$ and $x = y, y = s, z = z.$

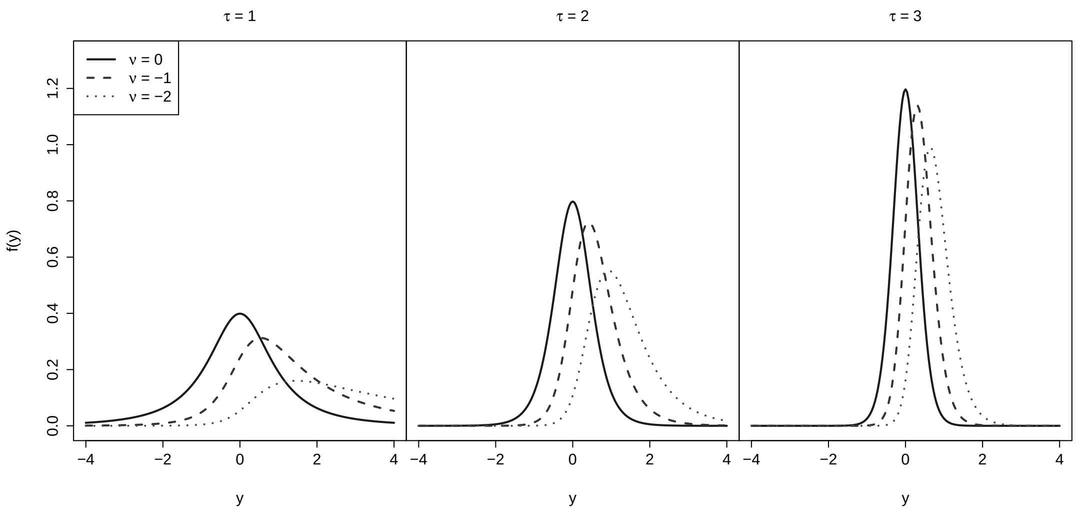

Figure 1.14: The original Johnson's S_u, `JSUo`(μ, σ, ν, τ), distribution, with $\mu = 0$, $\sigma = 1$, $\nu = 0, -1, -2$, and $\tau = 1, 2, 3$. It is symmetric if $\nu = 0$. Changing ν to $-\nu$ reflects the distribution about $y = \mu$, i.e. $y = 0$ above.

Second parameterization, `JSU`

This is a reparameterization of the original Johnson's S_u distribution [Johnson, 1949], so that parameters μ and σ are the mean and the standard deviation of the distribution. The `JSU`(μ, σ, ν, τ) is obtained by reparameterizing `JSUo`$(\mu_1, \sigma_1, \nu_1, \tau_1)$ to $\mu = \mu_1 - \sigma_1 \omega^{1/2} \sinh(\nu_1/\tau_1)$, $\sigma = \sigma_1/c$, $\nu = -\nu_1$ and $\tau = \tau_1$, where c and ω are defined in equations (1.20) and (1.21) below. Hence $\mu_1 = \mu - c\sigma\omega^{1/2} \sinh(\nu/\tau)$, $\sigma_1 = c\sigma$, $\nu_1 = -\nu$ and $\tau_1 = \tau$.

The pdf of the reparameterized Johnson's S_u, denoted by `JSU`(μ, σ, ν, τ), is given by

$$f_Y(y \mid \mu, \sigma\, \nu, \tau) = \frac{\tau}{c\sigma(s^2+1)^{1/2}\sqrt{2\pi}} \, \exp\left[-\frac{1}{2}z^2\right] \tag{1.18}$$

for $-\infty < y < \infty$, where $-\infty < \mu < \infty$, $\sigma > 0$, $-\infty < \nu < \infty$, $\tau > 0$, and where

$$
\begin{aligned}
z &= -\nu + \tau \sinh^{-1}(s) = -\nu + \tau \log\left[s + (s^2+1)^{1/2}\right] && (1.19)\\
s &= \frac{y - \mu + c\sigma w^{1/2}\sinh(\nu/\tau)}{c\sigma} \\
c &= \left\{\frac{1}{2}(w-1)\left[w\cosh(2\nu/\tau)+1\right]\right\}^{-1/2} && (1.20)\\
w &= \exp(1/\tau^2) \ . && (1.21)
\end{aligned}
$$

Note that $Z \sim$ `NO`$(0,1)$, where Z is obtained from (1.19) with y replaced by Y. For $Y \sim$ `JSU`(μ, σ, μ, τ) then $\mathrm{E}(Y) = \mu$ and $\mathrm{Var}(Y) = \sigma^2$.

Parameter ν is a true van Zwet skewness parameter and hence a true centile skewness parameter and a true moment skewness parameter. Increasing ν (for fixed τ) increases the moment and centile skewness. The distribution is symmetric if $\nu = 0$, and positively (moment, centile, and van Zwet) skew if $\nu > 0$, and negatively skew if $\nu < 0$. As $\nu \to \infty$, the `JSU` distribution tends to the log normal. Chapter **??** shows moment and centile kurtosis-skewness plots for `JSU`(μ, σ, ν, τ), see Figures **??**(d), **??**(c) and **??**(c). The distribution is always (centile and Balanda-MacGillivray) leptokurtic. The parameter τ affects the kurtosis of the distribution, and as $\tau \to \infty$ the distribution tends to the normal. Parameter τ is a primary true Balanda-MacGillivray kurtosis parameter, see Section **??**, and hence a primary true centile kurtosis parameter. Increasing τ decreases the centile (and Balanda-MacGillivray) kurtosis.

If $Y \sim$ `JSU`(μ, σ, ν, τ) then $-Y \sim$ `JSU`$(-\mu, \sigma, -\nu, \tau)$ is a reflection of the distribution of Y about zero. Also $W = (2\mu - Y) \sim$ `JSU`$(\mu, \sigma, -\nu, \tau)$ is a reflection of the distribution of Y about μ.

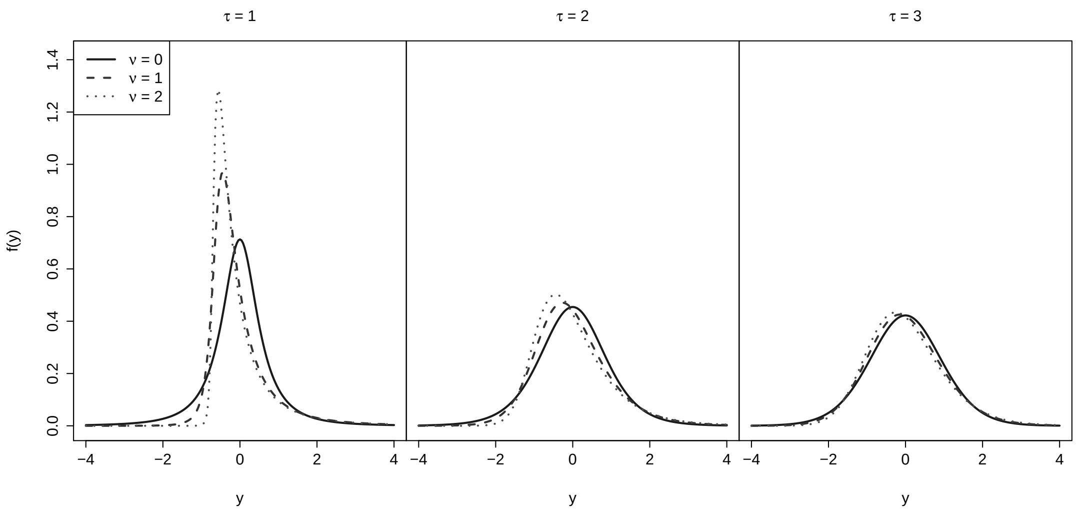

Figure 1.15: The reparameterized Johnson S_u, JSU(μ, σ, ν, τ), distribution, with $\mu = 0$, $\sigma = 1$, $\nu = 0, 1, 2$, and $\tau = 1, 2, 3$. It is symmetric if $\nu = 0$. Changing ν to $-\nu$ reflects the distribution about $y = \mu$, i.e. $y = 0$ above.

Table 1.16: Second parameterization Johnson's S_u distribution.

	`JSU`(μ,σ,ν,τ)
	Ranges
Y	$-\infty < y < \infty$
μ	$-\infty < \mu < \infty,$ mean, location shift parameter
σ	$0 < \sigma < \infty,$ standard deviation, scaling parameter
ν	$-\infty < \nu < \infty,$ true van Zwet skewness parameter
τ	$0 < \tau < \infty,$ primary true Balanda-MacGillivray kurtosis parameter
	Distribution measures
mean	μ
median	$\begin{cases} \mu_1 + c\sigma\sinh(\nu/\tau) \text{ where} \\ \mu_1 = \mu - c\sigma\omega^{1/2}\sinh(\nu/\tau)\ ,\ \omega = \exp(1/\tau^2) \text{ and } c \text{ is given by (1.20)} \end{cases}$
variance	σ^2
skewness	$\begin{cases} \mu_3/[\text{Var}(Y)]^{1.5} \text{ where} \\ \mu_3 = \dfrac{1}{4}c^3\sigma^3\omega^{1/2}(\omega-1)^2[\omega(\omega+2)\sinh(3\nu/\tau) + 3\sinh(\nu/\tau)] \end{cases}$
excess kurtosis	$\begin{cases} \mu_4/[\text{Var}(Y)]^2 - 3 \text{ where} \\ \mu_4 = \dfrac{1}{8}c^4\sigma^4(\omega-1)^2[\omega^2(\omega^4+2\omega^3+3\omega^2-3)\cosh(4\nu/\tau) \\ +4\omega^2(\omega+2)\cosh(2\nu/\tau) + 3(2\omega+1)] \end{cases}$
cdf	$\Phi(-\nu + \tau\sinh^{-1}[(y-\mu_1)/(c\sigma)])$
inverse cdf	$\mu_1 + c\sigma\sinh[(z_p+\nu)/\tau]$ where $z_p = \Phi^{-1}(p)$
Note	Reparameterize $\mu_1,\sigma_1,\nu_1,\tau_1$ in `JSUo`$(\mu_1,\sigma_1,\nu_1,\tau_1)$ by letting $\mu_1 = \mu - c\sigma\omega^{1/2}\sinh(\nu/\tau), \sigma_1 = c\sigma, \nu_1 = -\nu$ and $\tau_1 = \tau$ to give `JSU`(μ,σ,ν,τ).

1.4.4 Normal-exponential-t: `NET`

The `NET` is a four-parameter continuous distribution, although in **gamlss** it is used as a two-parameter distribution (μ and σ) with the other two parameters (ν and τ) fixed as constants, by default $\nu = 1.5$ and $\tau = 2$. (These values can be changed by the user.) The `NET` distribution is symmetric about its mean, median, and mode μ. If $Y \sim$ `NET`(μ,σ,ν,τ), then $Z = (Y-\mu)/\sigma$ has a standardized `NET`$(0,1,\nu,\tau)$ distribution, with pdf $f_Z(z)$, which is standard normal for $|z| \le \nu$, exponential with mean $1/\nu$ for $\nu < |z| \le \tau$, and has t distribution with $(\nu\tau - 1)$ degrees of freedom type tails for $|z| > \tau$, where $z = (y-\mu)/\sigma$. In **gamlss**, μ and σ can be modeled. Parameters ν and τ may be chosen as fixed constants by the user; alternatively estimates of ν and τ can be obtained using the `prof.dev()` function.

The normal-exponential-t distribution, `NET`, was introduced by Rigby and Stasinopoulos [1994] as a robust method of fitting the location and scale parameters of a symmetric distribution as functions of explanatory variables. That is, the distribution is appropriate if the response variable is contaminated and the user wants to robustify the fitting of the mean and variance models. Note that the Huber 'proposal 2' robust estimators of μ and σ [Huber, 1964, 1967], are equivalent to using a normal-exponential distribution for the response variable when fitting μ and a normal-Student-t distribution when fitting σ, and alternating between the fits. (Huber also includes a bias correction for the estimator of σ, which `NET` does not do.) `NET` combines these into one distribution. Figure 1.16 (left plot) shows a typical standardized $\texttt{NET}(0, 1, \nu, \tau)$ distribution with $\nu = 1.5$ and $\tau = 2$. The distribution in the interval $(-1.5, 1.5)$ behaves like a normal distribution, in the intervals $(-2, -1.5)$ and $(1.5, 2)$ it behaves like an exponential distribution, while in the intervals $(-\infty, -2)$ and $(2, \infty)$ it behaves like a t distribution. The parameters ν and τ are the breakpoints of the `NET` distribution and they are *not* modeled in **gamlss**.

The pdf of the $\texttt{NET}(\mu, \sigma, \nu, \tau)$ distribution is given by Rigby and Stasinopoulos [1994] as

$$f_Y(y \mid \mu, \sigma, \nu, \tau) = \frac{c}{\sigma} \begin{cases} \exp\left(-\dfrac{z^2}{2}\right) & \text{if} \quad |z| \le \nu \\ \exp\left(-\nu|z| + \dfrac{\nu^2}{2}\right) & \text{if} \quad \nu < |z| \le \tau \\ \exp\left(-\nu\tau \log\left(|z|/\tau\right) - \nu\tau + \dfrac{\nu^2}{2}\right) & \text{if} \quad |z| > \tau \end{cases} \tag{1.22}$$

for $-\infty < y < \infty$, where $-\infty < \mu < \infty$, $\sigma > 0$, $\nu > 0$, $\tau > \max(\nu, \nu^{-1})$, $z = (y - \mu)/\sigma$ and $c = (c_1 + c_2 + c_3)^{-1}$, where

$$\begin{aligned} c_1 &= \sqrt{2\pi}[2\Phi(\nu) - 1] \\ c_2 &= \frac{2}{\nu} \exp\left\{-\frac{\nu^2}{2}\right\} \\ c_3 &= \frac{2}{\nu(\nu\tau - 1)} \exp\left\{-\nu\tau + \frac{\nu^2}{2}\right\}. \end{aligned} \tag{1.23}$$

The $\texttt{NET}(\mu, \sigma, \nu, \tau)$ is symmetric about μ and is leptokurtic. The pdf of $\texttt{NET}(\mu, \sigma, \nu, \tau)$ has order $O(|y|^{-\nu\tau})$ as $|y| \to \infty$, and so the tails are heavier as $\nu\tau$ decreases. The excess kurtosis is $\gamma_2 = \mu_4/[\mathrm{Var}(Y)]^2 - 3$ where

$$\mu_4 = 2\sigma^4 c \left[\begin{array}{l} 3\sqrt{2\pi}[\Phi(\nu) - 0.5] + \left(\nu + \dfrac{12}{\nu} + \dfrac{24}{\nu^3} + \dfrac{24}{\nu^5}\right) e^{-\nu^2/2} \\ + \left(\dfrac{\tau^5}{\nu\tau - 5} - \dfrac{\tau^4}{\nu} - \dfrac{4\tau^3}{\nu^2} - \dfrac{12\tau^2}{\nu^3} - \dfrac{24\tau}{\nu^4} - \dfrac{24}{\nu^5}\right) e^{-\nu\tau + \nu^2/2} \end{array} \right], \tag{1.24}$$

for $\nu\tau > 5$.

The cdf of $Y \sim \texttt{NET}(\mu, \sigma, \nu, \tau)$ is given by $F_Y(y) = F_Z(z)$ where $Z = (Y - \mu)/\sigma$, $z = (y - \mu)/\sigma$ and

$$F_Z(z) = \begin{cases} \dfrac{c\tau^{\nu\tau}|z|^{-\nu\tau+1}}{(\nu\tau - 1)} \exp(-\nu\tau + \nu^2/2) & \text{if } z < -\tau \\ \dfrac{c}{\nu(\nu\tau - 1)} \exp(-\nu\tau + \nu^2/2) + \dfrac{c}{\nu} \exp(-\nu|z| + \nu^2/2) & \text{if } -\tau \le z < -\nu \\ \dfrac{c}{\nu(\nu\tau - 1)} \exp(-\nu\tau + \nu^2/2) + \dfrac{c}{\nu} \exp(-\nu^2/2) & \\ + c\sqrt{2\pi}[\Phi(z) - \Phi(-\nu)] & \text{if } -\nu \le z \le 0 \\ 1 - F_Z(-z) & \text{if } z > 0. \end{cases} \tag{1.25}$$

The inverse cdf of $Y \sim \texttt{NET}(\mu, \sigma, \nu, \tau)$ is given by

$$y_p = \mu + \sigma z_p \text{ where}$$

$$z_p = \begin{cases} -\left(\dfrac{b\nu\tau^{\nu\tau}}{p}\right)^{1/(\nu\tau-1)} & \text{if } p \le \nu\tau b \\ -\dfrac{\nu}{2} + \dfrac{1}{\nu} \log\left[\dfrac{\nu}{c}(p - b)\right] & \text{if } \nu\tau b < p \le b + \dfrac{c}{\nu}\exp\left(-\nu^2/2\right) \\ \Phi^{-1}\left\{\Phi(-\nu) + \dfrac{1}{\sqrt{2\pi}}\left[\dfrac{(p-b)}{c} - \dfrac{1}{\nu}\exp\left(-\nu^2/2\right)\right]\right\} & \\ & \text{if } b + \dfrac{c}{\nu}\exp\left(-\nu^2/2\right) < p \le \dfrac{1}{2} \\ -z_{1-p} & \text{if } p > \dfrac{1}{2}, \end{cases} \tag{1.26}$$

where $b = \dfrac{c}{\nu(\nu\tau - 1)} \exp\left(-\nu\tau + \nu^2/2\right)$.

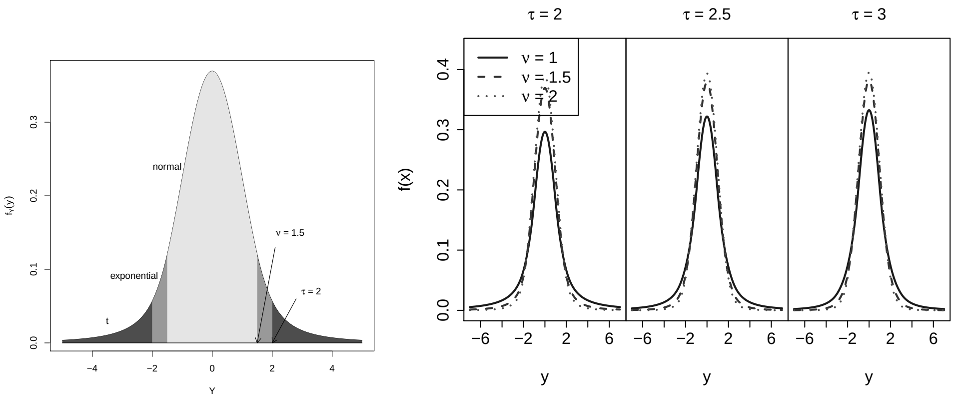

Figure 1.16: Left: schematic diagram of the NET$(0, 1, 1.5, 2)$ distribution. Right: the NET(μ, σ, ν, τ) distribution with $\mu = 0$, $\sigma = 1$, $\nu = 1, 1.5, 2$, and $\tau = 2, 2.5, 3$. Parameters ν and τ exist to make the distribution fitting robust and cannot be modeled as functions of explanatory variables. The distribution is always symmetric.

Table 1.17: Normal-exponential-Student-t distribution.

	$\texttt{NET}(\mu, \sigma, \nu, \tau)$
	Ranges
Y	$-\infty < y < \infty$
μ	$-\infty < \mu < \infty$, mean, median, mode, location shift parameter
σ	$0 < \sigma < \infty$, scaling parameter
ν	$0 < \nu < \infty$, first kurtosis parameter (fixed constant)
τ	$\max(\nu, \nu^{-1}) < \tau < \infty$, second kurtosis parameter (fixed constant)
	Distribution measures
mean	$\begin{cases} \mu & \text{if } \nu\tau > 2 \\ \text{undefined} & \text{if } \nu\tau \le 2 \end{cases}$
median	μ
mode	μ
variance	$\begin{cases} 2\sigma^2 c \Big\{ \sqrt{2\pi}\,[\Phi(\nu) - 0.5] & \\ + \left(2/\nu + 2/\nu^3\right) \exp(-\nu^2/2) & \\ + \dfrac{(\nu^2\tau^2 + 4\nu\tau + 6)}{\nu^3(\nu\tau - 3)} \exp(-\nu\tau + \nu^2/2) \Big\} & \text{if } \nu\tau > 3 \\ \infty & \text{if } 2 < \nu\tau \le 3 \\ \text{undefined} & \text{if } \nu\tau \le 2 \\ \text{where } c \text{ is given by (1.23)} & \end{cases}$
skewness	$\begin{cases} 0 & \text{if } \nu\tau > 4 \\ \text{undefined,} & \text{if } \nu\tau \le 4 \end{cases}$
excess kurtosis	see equation (1.24)
pdf	$\dfrac{c}{\sigma} \begin{cases} \exp\left(-z^2/2\right) & \text{if} \quad \lvert z\rvert \le \nu \\ \exp\left(-\nu\lvert z\rvert + \nu^2/2\right) & \text{if} \quad \nu < \lvert z\rvert \le \tau \\ \exp\left[-\nu\tau \log\left(\lvert z\rvert/\tau\right) - \nu\tau + \nu^2/2\right] & \text{if} \quad \lvert z\rvert > \tau \end{cases}$
cdf	see equation (1.25)
inverse cdf	see equation (1.26)

1.4.5 Sinh-arcsinh: SHASH

The pdf of the sinh-arcsinh distribution [Jones, 2005], denoted by $\texttt{SHASH}(\mu, \sigma, \nu, \tau)$, is given by

$$f_Y(y \mid \mu, \sigma\, \nu, \tau) = \frac{c}{\sqrt{2\pi}\sigma(1 + z^2)^{1/2}} \exp\left(-r^2/2\right) \tag{1.27}$$

where

$$\begin{aligned} r &= \tfrac{1}{2}\left\{\exp\left[\tau \sinh^{-1}(z)\right] - \exp\left[-\nu \sinh^{-1}(z)\right]\right\} \qquad (1.28) \\ c &= \tfrac{1}{2}\left\{\tau \exp\left[\tau \sinh^{-1}(z)\right] + \nu \exp\left[-\nu \sinh^{-1}(z)\right]\right\} \end{aligned}$$

and $z = (y-\mu)/\sigma$ for $-\infty < y < \infty$, where $-\infty < \mu < \infty$, $\sigma > 0$, $\nu > 0$ and $\tau > 0$. [Note $\sinh^{-1}(z) = \log(u)$ where $u = z + \left(z^2+1\right)^{1/2}$. Hence $r = \left(u^{\tau} - u^{-\nu}\right)/2$.] Note that $R \sim \texttt{NO}(0,1)$ where R is obtained from (1.28) and $Z = (Y-\mu)/\sigma$. Hence μ is the median of Y (since $y = \mu$ gives $z = 0$, $u = 1$, $\sinh^{-1}(z) = 0$ and hence $r = 0$). Note also R is the normalized quantile residual (or Z score).

The parameter ν controls the left tail heaviness, with the left tail being heavier than the normal distribution if $\nu < 1$ and lighter if $\nu > 1$. Similarly τ controls the right tail. The distribution is symmetric if $\nu = \tau$.

If $Y \sim \texttt{SHASH}(\mu, \sigma, \nu, \tau)$ then $-Y \sim \texttt{SHASH}(-\mu, \sigma, \tau, \nu)$ is a reflection of the distribution of Y about zero. Also $W = (2\mu - Y) \sim \texttt{SHASH}(\mu, \sigma, \tau, \nu)$ is a reflection of the distribution of Y about μ.

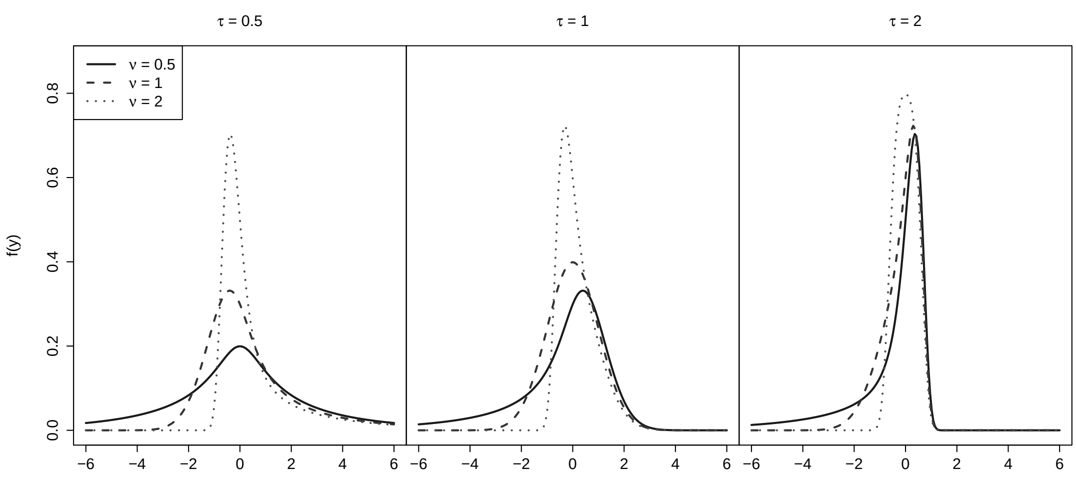

Figure 1.17: The sinh-arcsinh, $\texttt{SHASH}(\mu, \sigma, \nu, \tau)$, distribution, with $\mu = 0$, $\sigma = 1$, $\nu = 0.5, 1, 2$ and $\tau = 0.5, 1, 2$. It is symmetric if $\nu = \tau$. Interchanging ν and τ reflects the distribution about $y = \mu$, i.e. $y = 0$ above.

Table 1.18: Sinh-arcsinh distribution.

$\texttt{SHASH}(\mu, \sigma, \nu, \tau)$	
Ranges	
Y	$-\infty < y < \infty$
μ	$-\infty < \mu < \infty$, median, location shift parameter
σ	$0 < \sigma < \infty$, scaling parameter
ν	$0 < \nu < \infty$, left tail heaviness parameter
τ	$0 < \tau < \infty$, right tail heaviness parameter
Distribution measures	
median	μ
pdf	$\begin{cases} \dfrac{c}{\sigma\sqrt{2\pi}(1+z^2)^{1/2}} \exp\left(-r^2/2\right) \\ \text{where } r = \dfrac{1}{2}\left\{\exp[\tau \sinh^{-1}(z)] - \exp[-\nu \sinh^{-1}(z)]\right\} \text{ and} \\ z = \dfrac{y-\mu}{\sigma}, \; c = \dfrac{1}{2}\left\{\tau \exp[\tau \sinh^{-1}(z)] + \nu \exp[-\nu \sinh^{-1}(z)]\right\} \end{cases}$
cdf	$\Phi(r)$
Reference	Jones and Pewsey [2009], page 777 with $(\xi, \eta, \gamma, \delta, x, z)$ replaced by $(\mu, \sigma, \nu, \tau, z, r)$.

1.4.6 Sinh-arcsinh original: SHASHo, SHASHo2

First parameterization, SHASHo

The original sinh-arcsinh distribution, developed by Jones and Pewsey [2009] is denoted by $\texttt{SHASHo}(\mu, \sigma, \nu, \tau)$, with pdf given by

$$f_Y(y \mid \mu, \sigma, \nu, \tau) = \frac{\tau c}{\sigma\sqrt{2\pi}(1+z^2)^{1/2}} \exp\left(-\frac{1}{2}r^2\right) \tag{1.29}$$

where $r = \sinh[\tau \sinh^{-1}(z) - \nu]$, $c = \cosh[\tau \sinh^{-1}(z) - \nu]$ and $z = (y-\mu)/\sigma$ for $-\infty < y < \infty$, $-\infty < \mu < \infty$, $\sigma > 0$, $-\infty < \nu < \infty$ and $\tau > 0$. Note that $c^2 - r^2 = 1$. Note also that $\sinh^{-1}(z) = \log(\mu)$ where $u = z + (z^2+1)^{1/2}$. Hence $z = \frac{1}{2}(u - u^{-1})$. Note also that $R = \sinh[\tau \sinh^{-1}(Z) - \nu)] \sim \texttt{NO}(0,1)$ where $Z = (Y-\mu)/\sigma$. Hence R is the normalized quantile residual (or z-score).

Parameter ν is a true van Zwet skewness parameter, [see Jones and Pewsey, 2009, p763 and Section 14.5], with $\nu > 0$ and $\nu < 0$ corresponding to positive and negative (moment, centile, and van Zwet) skewness, respectively, and $\nu = 0$ corresponding to a symmetrical distribution. Hence it is a true centile skewness parameter and a true moment skewness parameter, see Section **??**. Increasing ν (for fixed τ) increases the (moment, centile, and van Zwet) skewness.

Parameter τ is a primary true Balanda-MacGillivray kurtosis parameter, see Section **??**, and hence a primary true centile kurtosis parameter, with $\tau < 1$ and $\tau > 1$ corresponding to heavier and lighter tails than the normal distribution, respectively [Jones and Pewsey, 2009, p762]. Increasing τ decreases the centile (and Balanda-MacGillivray) kurtosis. Figures **??**(a), **??**(a), and

??(a) show moment and centile kurtosis-skewness plots, for `SHASHo`(μ, σ, ν, τ). If $Y \sim$ `SHASHo`(μ, σ, ν, τ) then $-Y \sim$ `SHASHo`$(-\mu, \sigma, -\nu, \tau)$ is a reflection of the distribution of Y about zero. Also $W = (2\mu - Y) \sim$ `SHASHo`$(\mu, \sigma, -\nu, \tau)$ is a reflection of the distribution of Y about μ.

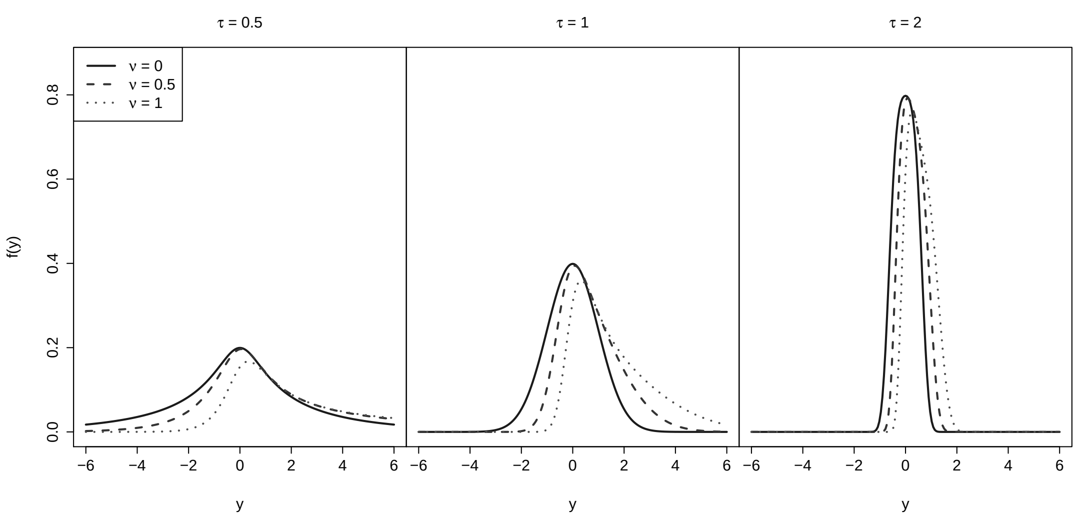

Figure 1.18: The original sinh-arcsinh, `SHASHo`(μ, σ, ν, τ), distribution, with $\mu = 0$, $\sigma = 1$, $\nu = 0, 0.5, 1$, and $\tau = 0.5, 1, 2$. It is symmetric if $\nu = 0$. Changing ν to $-\nu$ reflects the distribution about $y = \mu$, i.e. $y = 0$ above.

Table 1.19: Sinh-arcsinh original distribution.

	SHASHo(μ, σ, ν, τ)
	Ranges
Y	$-\infty < y < \infty$
μ	$-\infty < \mu < \infty,$ location shift parameter
σ	$0 < \sigma < \infty,$ scaling parameter
ν	$-\infty < \nu < \infty,$ true van Zwet skewness parameter
τ	$0 < \tau < \infty,$ primary true Balanda-MacGillivray kurtosis parameter
	Distribution measures
mean	$\begin{cases} \mu + \sigma \mathrm{E}(Z) = \mu + \sigma \sinh(\nu/\tau) P_{1/\tau} \text{ where } Z = (Y-\mu)/\sigma \text{ and} \\ P_q = \dfrac{\exp(0.25)}{(8\pi)^{0.5}} \left[K_{(q+1)/2}(0.25) + K_{(q-1)/2}(0.25)\right] \end{cases}$
median	$\mu + \sigma \sinh(\nu/\tau) = \mu + \dfrac{\sigma}{2}[\exp(\nu/\tau) - \exp(-\nu/\tau)]$
mode	μ only when $\nu = 0$
variance	$\sigma^2 \mathrm{Var}(Z) = \dfrac{\sigma^2}{2}[\cosh(2\nu/\tau) P_{2/\tau} - 1] - \sigma^2 [\sinh(\nu/\tau) P_{1/\tau}]^2$
skewness	$\begin{cases} \mu_{3Y}/[\mathrm{Var}(Y)]^{1.5} \text{ where} \\ \mu_{3Y} = \sigma^3 \mu_{3Z} = \sigma^3 \{\mu'_{3Z} - 3\mathrm{Var}(Z)\mathrm{E}(Z) - [\mathrm{E}(Z)]^3\} \\ \mu'_{3Z} = \mathrm{E}(Z^3) = \dfrac{1}{4}[\sinh(3\nu/\tau) P_{3/\tau} - 3\sinh(\nu/\tau) P_{1/\tau}] \end{cases}$
excess kurtosis	$\begin{cases} \mu_{4Y}/[\mathrm{Var}(Y)]^2 - 3 \text{ where} \\ \mu_{4Y} = \sigma^4 \mu_{4Z} = \sigma^4 \{\mu'_{4Z} - 4\mu'_{3Z}\mathrm{E}(Z) + 6\mathrm{Var}(Z)[\mathrm{E}(Z)]^2 + 3[\mathrm{E}(Z)]^4\} \\ \mu'_{4Z} = \mathrm{E}(Z^4) = \dfrac{1}{8}[\cosh(4\nu/\tau) P_{4/\tau} - 4\cosh(2\nu/\tau) P_{2/\tau} + 3] \end{cases}$
pdf	$\begin{cases} \dfrac{\tau c}{\sigma\sqrt{2\pi}(1+z^2)^{1/2}} \exp(-r^2/2) \\ \text{where } r = \sinh[\tau \sinh^{-1}(z) - \nu] \\ \text{and } z = (y-\mu)/\sigma \text{ and } c = \cosh[\tau \sinh^{-1}(z) - \nu] \end{cases}$
cdf	$\Phi(r)$
inverse cdf	$\mu + \sigma \sinh\left\{\dfrac{\nu}{\tau} + \dfrac{1}{\tau} \sinh^{-1}\left[\Phi^{-1}(p)\right]\right\}$
Reference	Jones and Pewsey [2009], page 762-764, with $(\xi, \eta, \epsilon, \delta, x, z)$ replaced by $(\mu, \sigma, \nu, \tau, z, r)$.
Note	$K_\lambda(t)$ is a modified Bessel function of the second kind

Second parameterization, SHASHo2

Jones and Pewsey [2009, p768] suggest reparameterizing SHASHo(μ, σ, ν, τ) in order to provide a more orthogonal parameterization, SHASHo2(μ, σ, ν, τ), with pdf given by (1.29) with σ replaced by $\sigma\tau$. This solves numerical problems encountered in their original parameterization, i.e. SHASHo(μ, σ, ν, τ), when $\tau > 1$. The summary table for SHASHo2(μ, σ, ν, τ) is given by replacing σ by $\sigma\tau$ in Table 1.19.

1.4.7 Skew exponential power type 1: SEP1

The pdf of the skew exponential power type 1 distribution, denoted by SEP1(μ, σ, ν, τ), is given by

$$f_Y(y \mid \mu, \sigma, \nu, \tau) = \frac{2}{\sigma} \, f_{Z_1}(z) \; F_{Z_1}(\nu z) \tag{1.30}$$

for $-\infty < y < \infty$, where $-\infty < \mu < \infty$, $\sigma > 0$, $-\infty < \nu < \infty$ and $\tau > 0$, and where $z = (y - \mu)/\sigma$ and $f_{Z_1}(\cdot)$ and $F_{Z_1}(\cdot)$ are the pdf and cdf of $Z_1 \sim$ PE2$(0, \tau^{1/\tau}, \tau)$, given in Table 1.8. The SEP1(μ, σ, ν, τ) distribution was introduced by Azzalini [1986] as his type I distribution. See Section **??**. The skew normal type 1 distribution, SN1(μ, σ, ν), is a special case of SEP1(μ, σ, ν, τ) given by $\tau = 2$.

The SEP1(μ, σ, ν, τ) distribution is symmetric if $\nu = 0$. However from Figures 1 and 2 of Azzalini [1986], for fixed $\tau > 2$, the moment skewness of SEP1(μ, σ, ν, τ) can be positive or negative, but is not always positive if $\nu > 0$ (and not always negative if $\nu < 0$) and is not always monotonically increasing with ν. The parameter τ affects the moment kurtosis of the distribution, with increasing τ (for fixed ν) appearing to decrease the moment kurtosis, which can be moment leptokurtosis or platykurtosis. However the shape of the distribution is very flexible.

If $Y \sim$ SEP1(μ, σ, ν, τ) then $-Y \sim$ SEP1$(-\mu, \sigma, -\nu, \tau)$ is a reflection of the distribution of Y about zero. Also $W = (2\mu - Y) \sim$ SEP1$(\mu, \sigma, -\nu, \tau)$ is a reflection of the distribution of Y about μ. As $\nu \to \infty$ the SEP1(μ, σ, ν, τ) distribution tends to a half power exponential type 2 distribution, i.e. a half PE2$(\mu, \sigma\tau^{1/\tau}, \tau)$, (see Table 1.8). Also SEP1$(\mu, \sigma, 0, \tau) \equiv$ PE2$(\mu, \sigma\tau^{1/\tau}, \tau)$.

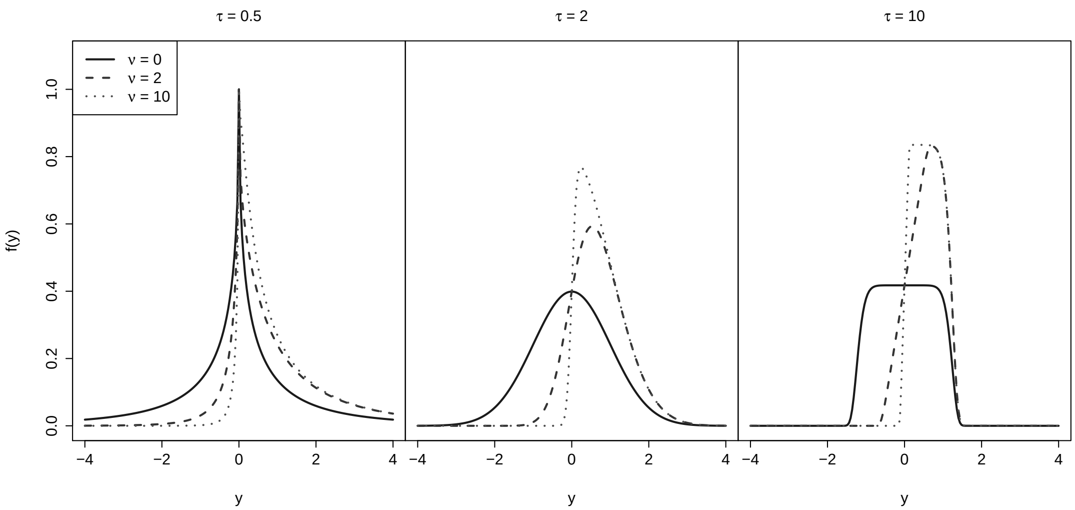

Figure 1.19: The skew exponential power type 1, `SEP1`(μ, σ, ν, τ), distribution, with $\mu = 0$, $\sigma = 1$, $\nu = 0, 2, 10$, and $\tau = 0.5, 2, 10$. It is symmetric if $\nu = 0$. Changing ν to $-\nu$ reflects the distribution about $y = \mu$, i.e. $y = 0$ above.

Table 1.20: Skew exponential power type 1 distribution.

	$\texttt{SEP1}(\mu, \sigma, \nu, \tau)$
	Ranges
Y	$-\infty < y < \infty$
μ	$-\infty < \mu < \infty$, location shift parameter
σ	$0 < \sigma < \infty$, scaling parameter
ν	$-\infty < \nu < \infty$, skewness parameter
τ	$0 < \tau < \infty$, kurtosis parameter
	Distribution measures
mean	$\begin{cases} \mu + \sigma \mathrm{E}(Z) = \mu + \dfrac{\sigma \mathrm{sign}(\nu) \tau^{1/\tau} \Gamma(2\tau^{-1}) B(\tau^{-1}, 2\tau^{-1}, \nu^\tau / [1+\nu^\tau])}{\Gamma(\tau^{-1}) B(\tau^{-1}, 2\tau^{-1})} \\ \text{where } Z = (Y-\mu)/\sigma \end{cases}$
variance	$\sigma^2 \mathrm{Var}(Z) = \sigma^2 \left\{ \dfrac{\tau^{2/\tau} \Gamma(3\tau^{-1})}{\Gamma(\tau^{-1})} - [\mathrm{E}(Z)]^2 \right\}$
skewness	$\begin{cases} \mu_{3Y} / [\mathrm{Var}(Y)]^{1.5} \text{ where} \\ \mu_{3Y} = \sigma^3 \mu_{3Z} = \sigma^3 \{ \mu'_{3Z} - 3\mathrm{Var}(Z)\mathrm{E}(Z) - [\mathrm{E}(Z)]^3 \} \\ \mu'_{3Z} = \mathrm{E}(Z^3) = \dfrac{\mathrm{sign}(\nu) \tau^{3/\tau} \Gamma(4\tau^{-1}) B(\tau^{-1}, 4\tau^{-1}, \nu^\tau/[1+\nu^\tau])}{\Gamma(\tau^{-1}) B(\tau^{-1}, 4\tau^{-1})} \end{cases}$
excess kurtosis	$\begin{cases} \mu_{4Y} / [\mathrm{Var}(Y)]^2 - 3 \text{ where} \\ \mu_{4Y} = \sigma^4 \mu_{4Z} = \sigma^4 \{ \mu'_{4Z} - 4\mu'_{3Z} \mathrm{E}(Z) + 6\mathrm{Var}(Z)[\mathrm{E}(Z)]^2 + 3[\mathrm{E}(Z)]^4 \} \\ \mu'_{4Z} = \mathrm{E}(Z^4) = \dfrac{\tau^{4/\tau} \Gamma(5\tau^{-1})}{\Gamma(\tau^{-1})} \end{cases}$
pdf	$\begin{cases} \dfrac{2}{\sigma} f_{Z_1}(z) F_{Z_1}(\nu z) \\ \text{where } z = (y-\mu)/\sigma \text{ and } Z_1 \sim \texttt{PE2}(0, \tau^{1/\tau}, \tau) \end{cases}$
Reference	Azzalini [1986], page 202-203, with (λ, ω) replaced by (ν, τ) giving the pdf and moments of Z.

1.4.8 Skew exponential power type 2: SEP2

The pdf of the skew exponential power type 2 distribution, denoted by $\texttt{SEP2}(\mu, \sigma, \nu, \tau)$, is given by

$$f_Y(y \mid \mu, \sigma \, \nu, \tau) = \frac{2}{\sigma} f_{Z_1}(z) \; \Phi(\omega) \tag{1.31}$$

for $-\infty < y < \infty$, where $-\infty < \mu < \infty$, $\sigma > 0$, $-\infty < \nu < \infty$, and $\tau > 0$, and where $z = (y-\mu)/\sigma$; $\omega = \mathrm{sign}(z)|z|^{\tau/2} \nu \sqrt{2/\tau}$; $f_{Z_1}(\cdot)$ is the pdf of $Z_1 \sim \texttt{PE2}(0, \tau^{1/\tau}, \tau)$ and $\Phi(\cdot)$ is the standard normal cdf.

This distribution was introduced by Azzalini [1986] as his type II distribution and developed by DiCiccio and Monti [2004]. See Section **??**. The $\texttt{SEP2}(\mu, \sigma, \nu, \tau)$ distribution is symmetric if $\nu = 0$. However from Figures 3 and 4 of Azzalini [1986], for fixed $\tau > 2$, the moment skewness of $\texttt{SEP2}(\mu, \sigma, \nu, \tau)$ is not always positive if $\nu > 0$ (and not always negative if $\nu < 0$) and is not always monotonically increasing with ν. The parameter τ affects the moment kurtosis of the distribution, with increasing τ (for fixed ν) appearing to decrease the moment kurtosis, which can be moment leptokurtosis or platykurtosis. However the shape of the distribution is very flexible.

If $Y \sim \texttt{SEP2}(\mu, \sigma, \nu, \tau)$ then $-Y \sim \texttt{SEP2}(-\mu, \sigma, -\nu, \tau)$ is a reflection of the distribution of Y about zero. Also $W = (2\mu - Y) \sim \texttt{SEP2}(\mu, \sigma, -\nu, \tau)$ is a reflection of the distribution of Y about μ. We have $\mathrm{E}(Y) = \mu + \sigma \mathrm{E}(Z)$ where $Z = (Y - \mu)/\sigma$ and

$$\mathrm{E}(Z) = \frac{2\tau^{1/\tau}\nu}{\sqrt{\pi}\Gamma\left(\tau^{-1}\right)(1+\nu^2)^{(2/\tau)+0.5}} \sum_{n=0}^{\infty} \frac{\Gamma\left(2\tau^{-1}+n+0.5\right)}{(2n+1)!!}\left(\frac{2\nu^2}{1+\nu^2}\right)^n \tag{1.32}$$

where $(2n+1)!! = 1.3.5\ldots(2n+1)$, and $\mu_{3Z}^{'} = \mathrm{E}(Z^3)$ is given by

$$\mathrm{E}(Z^3) = \frac{2\tau^{3/\tau}\nu}{\sqrt{\pi}\Gamma\left(\tau^{-1}\right)(1+\nu^2)^{(4/\tau)+0.5}} \sum_{n=0}^{\infty} \frac{\Gamma\left(4\tau^{-1}+n+0.5\right)}{(2n+1)!!}\left(\frac{2\nu^2}{1+\nu^2}\right)^n, \tag{1.33}$$

obtained from DiCiccio and Monti [2004, p440].

The $\texttt{SEP2}(\mu, \sigma, \nu, 2)$ is the skew normal type 1, $\texttt{SN1}(\mu, \sigma, \nu)$, distribution, [Azzalini, 1985], while the $\texttt{SEP2}(\mu, \sigma, 0, 2)$ is the normal distribution, $\texttt{NO}(\mu, \sigma)$. As $\nu \to \infty$ the $\texttt{SEP2}(\mu, \sigma, \nu, \tau)$ distribution tends to a half power exponential type 2 distribution, i.e. a half $\texttt{PE2}(\mu, \sigma\tau^{1/\tau}, \tau)$, (see Table 13.2). Also $\texttt{SEP2}(\mu, \sigma, 0, \tau) = \texttt{PE2}(\mu, \sigma\tau^{1/\tau}, \tau)$.

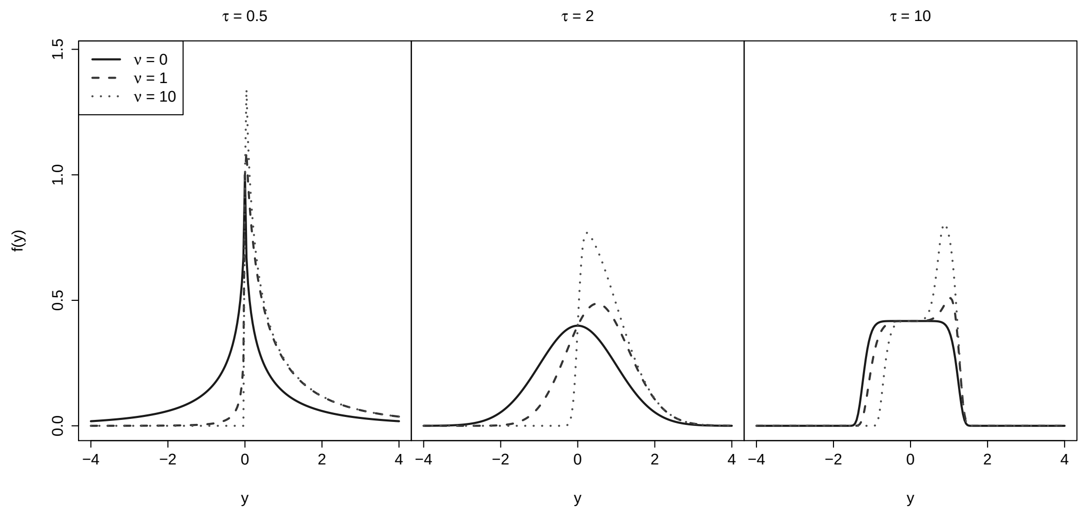

Figure 1.20: The skew exponential power type 2, SEP2(μ, σ, ν, τ), distribution with $\mu = 0$, $\sigma = 1$, $\nu = 0, 1, 10$, and $\tau = 0.5, 2, 10$. It is symmetric if $\nu = 0$. Changing ν to $-\nu$ reflects the distribution about $y = \mu$, i.e. $y = 0$ above.

Table 1.21: Skew exponential power type 2 distribution.

	`SEP2`(μ, σ, ν, τ)
	Ranges
Y	$-\infty < y < \infty$
μ	$-\infty < \mu < \infty,$ location shift parameter
σ	$0 < \sigma < \infty,$ scaling parameter
ν	$-\infty < \nu < \infty,$ skewness parameter
τ	$0 < \tau < \infty,$ kurtosis parameter
	Distribution measures
mean	$\begin{cases} \mu + \sigma \mathrm{E}(Z) \\ \text{where } Z = (Y-\mu)/\sigma \text{ and } \mathrm{E}(Z) \text{ is given by (1.32)} \end{cases}$
variance	$\sigma^2 \mathrm{Var}(Z) = \sigma^2 \left\{ \dfrac{\tau^{2/\tau}\Gamma(3\tau^{-1})}{\Gamma(\tau^{-1})} - [E(Z)]^2 \right\}$
skewness	$\begin{cases} \mu_{3Y}/[\mathrm{Var}(Y)]^{1.5} \text{ where} \\ \mu_{3Y} = \sigma^3 \mu_{3Z} = \sigma^3 \{\mu_{3Z}^{'} - 3\mathrm{Var}(Z)\mathrm{E}(Z) - [\mathrm{E}(Z)]^3\} \\ \mu_{3Z}^{'} \text{ is given by (1.33)} \end{cases}$
excess kurtosis	$\begin{cases} \mu_{4Y}/[\mathrm{Var}(Y)]^2 - 3 \text{ where} \\ \mu_{4Y} = \sigma^4 \mu_{4Z} = \sigma^4 \{\mu_{4Z}^{'} - 4\mu_{3Z}^{'}\mathrm{E}(Z) + 6\mathrm{Var}(Z)[\mathrm{E}(Z)]^2 + 3[E(Z)]^4\} \\ \mu_{4Z}^{'} = \mathrm{E}(Z^4) = \tau^{4/\tau}\Gamma(5\tau^{-1})/\Gamma(\tau^{-1}) \end{cases}$
pdf	$\begin{cases} \dfrac{2}{\sigma} f_{Z_1}(z)\Phi(\omega) \quad \text{where} \\ z = (y-\mu)/\sigma\,, Z_1 \sim \mathtt{PE2}(0, \tau^{1/\tau}, \tau) \text{ and } \omega = \mathrm{sign}(z)\lvert z\rvert^{\tau/2}\nu\sqrt{2/\tau} \end{cases}$
Reference	DiCiccio and Monti [2004], page 439-440, with (λ, α) replaced by (ν, τ), giving the pdf and moments of Z.

1.4.9 Skew exponential power type 3: `SEP3`

This is a ‘scale-spliced’ distribution (see Section **??**), denoted by `SEP3`(μ, σ, ν, τ), with pdf

$$f_Y(y \mid \mu, \sigma\, \nu, \tau) = \begin{cases} \dfrac{c}{\sigma} \exp\left[-\dfrac{1}{2}\lvert \nu z\rvert^{\tau}\right] & \text{if } y < \mu \\ \dfrac{c}{\sigma} \exp\left[-\dfrac{1}{2}\left\lvert \dfrac{z}{\nu}\right\rvert^{\tau}\right] & \text{if } y \geq \mu, \end{cases} \tag{1.34}$$

for $-\infty < y < \infty$, where $-\infty < \mu < \infty$, $\sigma > 0$, $\nu > 0$, and $\tau > 0$, and where $z = (y-\mu)/\sigma$ and $c = \nu\tau\left[(1+\nu^2)2^{1/\tau}\Gamma\left(1/\tau\right)\right]^{-1}$, [Fernandez et al., 1995, p 1333]. Note that μ is the mode of Y.

The `SEP3`(μ, σ, ν, τ) distribution is symmetric if $\nu = 0$. It is moment positively skew if $\nu > 1$ and moment negatively skew if $0 < \nu < 1$, [Fernandez et al., 1995, p1334]. The distribution appears to be moment leptokurtic if $\tau < 2$, but can be moment platykurtic or leptokurtic if $\tau > 2$. Figures ??(b), ??(a) and ??(a) show moment and centile kurtosis-skewness plots for the SEP3 distribution.

If $Y \sim$ SEP3(μ, σ, ν, τ) then $-Y \sim$ SEP3$(-\mu, \sigma, 1/\nu, \tau)$ is a reflection of the distribution of Y about zero. Also $W = (2\mu - Y) \sim$ SEP3$(\mu, \sigma, 1/\nu, \tau)$ is a reflection of the distribution of Y about μ. The skew normal type 2 (or two-piece normal) distribution, SN2(μ, σ, ν), is the special case SEP3$(\mu, \sigma, \nu, 2)$. The PE2 distribution is a reparameterized special case of SEP3 given by PE2$(\mu, \sigma, \nu) =$ SEP3$(\mu, \sigma 2^{-1/\nu}, 1, \nu)$. As $\nu \to \infty$, the SEP3 distribution tends to a half power exponential type 2 distribution, see Table ??. Also SEP3$(\mu, \sigma, 1, \tau) \equiv$ PE2$(\mu, \sigma 2^{1/\tau}, \tau)$. The shape of the distribution can be very flexible.

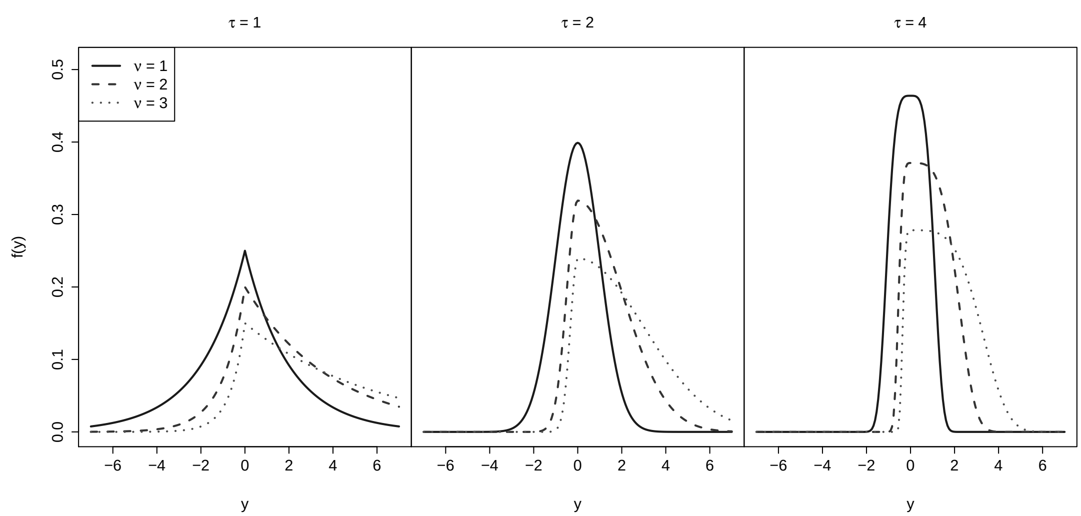

Figure 1.21: The skew exponential power type 3, SEP3(μ, σ, ν, τ), distribution, with $\mu = 0$, $\sigma = 1$, $\nu = 1, 2, 3$, and $\tau = 1, 2, 4$. It is symmetric if $\nu = 1$. Changing ν to $1/\nu$ reflects the distribution about $y = \mu$, i.e. $y = 0$ above.

Table 1.22: Skew exponential power type 3 distribution.

$\texttt{SEP3}(\mu, \sigma, \nu, \tau)$	
Ranges	
Y	$-\infty < y < \infty$
μ	$-\infty < \mu < \infty$, mode, location shift parameter
σ	$0 < \sigma < \infty$, scaling parameter
ν	$0 < \nu < \infty$, skewness parameter
τ	$0 < \tau < \infty$, kurtosis parameter
Distribution measures	
mean	$\begin{cases} \mu + \sigma E(Z) = \mu + \dfrac{\sigma 2^{1/\tau}\Gamma(2\tau^{-1})(\nu - \nu^{-1})}{\Gamma(\tau^{-1})} \\ \text{where } Z = (Y-\mu)/\sigma \end{cases}$
mode	μ
variance	$\sigma^2 \mathrm{Var}(Z) = \sigma^2 \left\{ \dfrac{2^{2/\tau}\Gamma(3\tau^{-1})(\nu^2 + \nu^{-2} - 1)}{\Gamma(\tau^{-1})} - [\mathrm{E}(Z)]^2 \right\}$
skewness	$\begin{cases} \mu_{3Y}/[\mathrm{Var}(Y)]^{1.5} \text{ where} \\ \mu_{3Y} = \sigma^3 \mu_{3Z} = \sigma^3 \{\mu'_{3Z} - 3\mathrm{Var}(Z)\mathrm{E}(Z) - [\mathrm{E}(Z)]^3\} \\ \mu'_{3Z} = \mathrm{E}(Z^3) = \dfrac{2^{3/\tau}\Gamma(4\tau^{-1})(\nu^4 - \nu^{-4})}{\Gamma(\tau^{-1})(\nu + \nu^{-1})} \end{cases}$
excess kurtosis	$\begin{cases} \mu_{4Y}/[\mathrm{Var}(Y)]^2 - 3 \text{ where} \\ \mu_{4Y} = \sigma^4 \mu_{4Z} = \sigma^4 \left\{ \mu'_{4Z} - 4\mu'_{3Z}\mathrm{E}(Z) + 6\mathrm{Var}(Z)[\mathrm{E}(Z)]^2 \right. \\ \quad \left. +3[\mathrm{E}(Z)]^4 \right\} \\ \mu'_{4Z} = \mathrm{E}(Z^4) = \dfrac{2^{4/\tau}\Gamma(5\tau^{-1})(\nu^5 + \nu^{-5})}{\Gamma(\tau^{-1})(\nu + \nu^{-1})} \end{cases}$
pdf	$\begin{cases} \dfrac{c}{\sigma} \exp\left(-\frac{1}{2}\lvert\nu z\rvert^\tau\right) & \text{if } y < \mu \\ \dfrac{c}{\sigma} \exp\left(-\frac{1}{2}\left\lvert\dfrac{z}{\nu}\right\rvert^\tau\right) & \text{if } y \geq \mu, \\ \text{where } z = (y-\mu)/\sigma \text{ and} & c = \nu\tau\left[(1+\nu^2)2^{1/\tau}\Gamma(1/\tau)\right]^{-1} \end{cases}$
cdf	$\begin{cases} \dfrac{1}{1+\nu^2}\left[\dfrac{\Gamma(\tau^{-1}, \alpha_1)}{\Gamma(\tau^{-1})}\right] & \text{if } y < \mu \\ 1 - \dfrac{\nu^2\Gamma(\tau^{-1}, \alpha_2)}{(1+\nu^2)\Gamma(\tau^{-1})} & \text{if } y \geq \mu \\ \text{where } \alpha_1 = \dfrac{\nu^2(\mu - y)^\tau}{2\sigma^\tau} \text{ and} & \alpha_2 = \dfrac{(y-\mu)^\tau}{2\sigma^\tau\nu^\tau} \end{cases}$
Reference	Fernandez et al. [1995], p1333, equations (8) and (12), with (γ, q) replaced by (ν, τ) giving the pdf and moments of Z.

1.4.10 Skew exponential power type 4: SEP4

This is a 'shape-spliced' distribution (see Section **??**), denoted by $\texttt{SEP4}(\mu, \sigma, \nu, \tau)$, with pdf

$$f_Y(y \mid \mu, \sigma, \nu, \tau) = \begin{cases} \dfrac{c}{\sigma} \exp\left[-|z|^{\nu}\right] & \text{if } y < \mu \\ \dfrac{c}{\sigma} \exp\left[-|z|^{\tau}\right] & \text{if } y \geq \mu \end{cases} \tag{1.35}$$

for $-\infty < y < \infty$, where $-\infty < \mu < \infty$, $\sigma > 0$, $\nu > 0$, and $\tau > 0$, and where $z = (y - \mu)/\sigma$ and $c = \left[\Gamma\left(1 + \tau^{-1}\right) + \Gamma\left(1 + \nu^{-1}\right)\right]^{-1}$, [Jones, 2005]. Note that μ is the mode of Y.

Parameters ν and τ affect the left- and right-tail heaviness, respectively, with $0 < \nu < 2$ or $0 < \tau < 2$ a heavier tail than the normal distribution, and $\nu > 2$ or $\tau > 2$ a lighter tail. The $\texttt{SEP4}(\mu, \sigma, \nu, \tau)$ distribution is symmetric if $\nu = \tau$ and $\texttt{SEP4}(\mu, \sigma, \nu, \nu) = \texttt{PE2}(\mu, \sigma, \nu)$. If $Y \sim \texttt{SEP4}(\mu, \sigma, \nu, \tau)$ then $-Y \sim \texttt{SEP4}(-\mu, \sigma, \tau, \nu)$ is a reflection of the distribution of Y about zero. Also $W = (2\mu - Y) \sim \texttt{SEP4}(\mu, \sigma, \tau, \nu)$ is a reflection of the distribution of Y about μ.

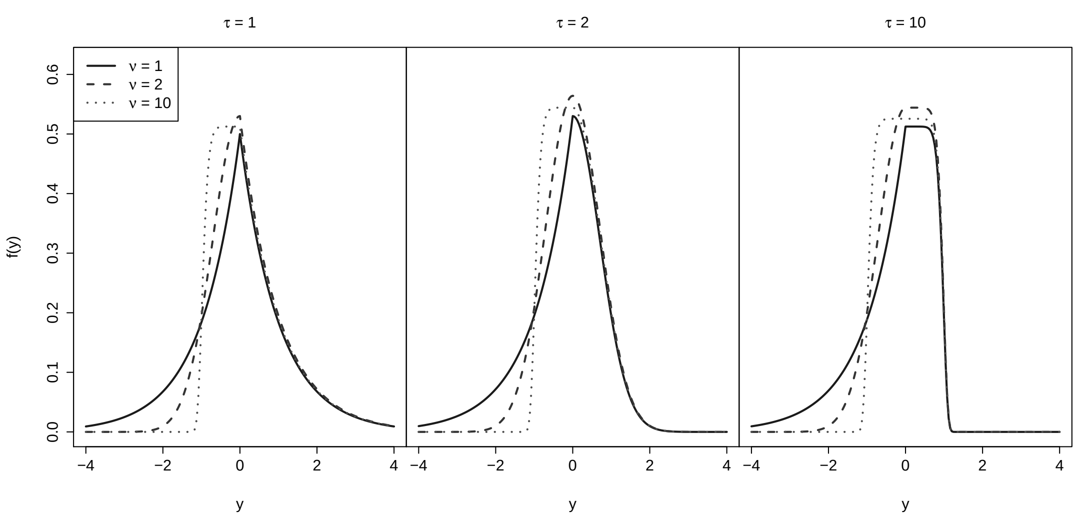

Figure 1.22: The skew exponential power type 4, $\texttt{SEP4}(\mu, \sigma, \nu, \tau)$, distribution, with $\mu = 0$, $\sigma = 1$, $\nu = 1, 2, 10$, and $\tau = 1, 2, 10$. It is symmetric if $\nu = \tau$. Interchanging ν and τ reflects the distribution about $y = \mu$, i.e. $y = 0$ above.

Table 1.23: Skew exponential power type 4 distribution.

SEP4(μ, σ, ν, τ)	
Ranges	
Y	$-\infty < y < \infty$
μ	$-\infty < \mu < \infty$, mode, location shift parameter
σ	$0 < \sigma < \infty$, scaling parameter
ν	$0 < \nu < \infty$, left tail heaviness parameter
τ	$0 < \tau < \infty$, right tail heaviness parameter
Distribution measures	
mean	$\begin{cases} \mu + \sigma \mathrm{E}(Z) = \mu + \sigma c \left[\tau^{-1}\Gamma\left(2\tau^{-1}\right) - \nu^{-1}\Gamma\left(2\nu^{-1}\right)\right] \\ \text{where } Z = (Y-\mu)/\sigma \text{ and } c = \left[\Gamma\left(1+\tau^{-1}\right) + \Gamma\left(1+\nu^{-1}\right)\right]^{-1} \end{cases}$
mode	μ
variance	$\sigma^2 \mathrm{Var}(Z) = \sigma^2 \left\{c\left[\tau^{-1}\Gamma\left(3\tau^{-1}\right) + \nu^{-1}\Gamma\left(3\nu^{-1}\right)\right] - [\mathrm{E}(Z)]^2\right\}$
skewness	$\begin{cases} \mu_{3Y}/[\mathrm{Var}(Y)]^{1.5} \text{ where} \\ \mu_{3Y} = \sigma^3 \mu_{3Z} = \sigma^3\{\mu'_{3Z} - 3\mathrm{Var}(Z)\mathrm{E}(Z) - [\mathrm{E}(Z)]^3\} \\ \mu'_{3Z} = \mathrm{E}(Z^3) = c\left[\tau^{-1}\Gamma\left(4\tau^{-1}\right) - \nu^{-1}\Gamma\left(4\nu^{-1}\right)\right] \end{cases}$
excess kurtosis	$\begin{cases} \mu_{4Y}/[\mathrm{Var}(Y)]^2 - 3 \text{ where} \\ \mu_{4Y} = \sigma^4 \mu_{4Z} = \sigma^4\{\mu'_{4Z} - 4\mu'_{3Z}\mathrm{E}(Z) + 6\mathrm{Var}(Z)[\mathrm{E}(Z)]^2 + \\ \qquad 3[\mathrm{E}(Z)]^4\} \\ \mu'_{4Z} = \mathrm{E}(Z^4) = c\left[\tau^{-1}\Gamma\left(5\tau^{-1}\right) + \nu^{-1}\Gamma\left(5\nu^{-1}\right)\right] \end{cases}$
pdf [a]	$\begin{cases} \dfrac{c}{\sigma}\exp\left(-\lvert z\rvert^{\nu}\right) & \text{if } y < \mu \\ \dfrac{c}{\sigma}\exp\left(-\lvert z\rvert^{\tau}\right) & \text{if } y \ge \mu, \\ \text{where } z = (y-\mu)/\sigma & \end{cases}$
cdf	$\begin{cases} \dfrac{c}{\nu}\Gamma(\nu^{-1}, \lvert z\rvert^{\nu}) & \text{if } y < \mu \\ 1 - \dfrac{c}{\tau}\Gamma(\tau^{-1}, \lvert z\rvert^{\tau}) & \text{if } y \ge \mu \end{cases}$
Reference	[a] Jones [2005]

1.4.11 Skew Student t: SST

Würtz et al. [2006] reparameterized the ST3 distribution, [Fernandez and Steel, 1998], so that in the new parameterization μ is the mean and σ is the standard deviation. They called this the skew Student t distribution, which we denote as SST.

Let $Z_0 \sim$ `ST3`$(0, 1, \nu, \tau)$ and $Y = \mu + \sigma \left(\frac{Z_0 - m}{s} \right)$, where

$$m = \mathrm{E}(Z_0) = \frac{2\tau^{1/2}(\nu - \nu^{-1})}{(\tau - 1)B\,(1/2, \tau/2)} \qquad \text{for } \tau > 1 \qquad (1.36)$$

$$s^2 = \mathrm{Var}(Z_0) = \frac{\tau}{(\tau - 2)}(\nu^2 + \nu^{-2} - 1) - m^2 \qquad \text{for } \tau > 2\ . \qquad (1.37)$$

Hence $Y = \mu_0 + \sigma_0 Z_0$, where $\mu_0 = \mu - \sigma m/s$ and $\sigma_0 = \sigma/s$, and so $Y \sim$ `ST3`$(\mu_0, \sigma_0, \nu, \tau)$ with $\mathrm{E}(Y) = \mu$ and $\mathrm{Var}(Y) = \sigma^2$ for $\tau > 2$. Let $Y \sim$ `SST`$(\mu, \sigma, \nu, \tau) =$ `ST3`$(\mu_0, \sigma_0, \nu, \tau)$, for $\tau > 2$.

Hence the pdf of the skew Student t distribution, denoted by $Y \sim$ `SST`(μ, σ, ν, τ), is given by

$$f_Y(y \,|\, \mu, \sigma, \nu, \tau) = \begin{cases} \dfrac{c}{\sigma_0} \left(1 + \dfrac{\nu^2 z^2}{\tau} \right)^{-(\tau+1)/2} & \text{if } y < \mu_0 \\ \dfrac{c}{\sigma_0} \left(1 + \dfrac{z^2}{\nu^2 \tau} \right)^{-(\tau+1)/2} & \text{if } y \geq \mu_0, \end{cases}$$

for $-\infty < y < \infty$, where $-\infty < \mu < \infty$, $\sigma > 0, \nu > 0$ and $\tau > 2$ and where $z = (y - \mu_0)/\sigma_0$, $\mu_0 = \mu - \sigma m/s$, $\sigma_0 = \sigma/s$ and $c = 2\nu \left[(1 + \nu^2) B\,(1/2, \tau/2)\, \tau^{1/2} \right]^{-1}$.

Note that $\mathrm{E}(Y) = \mu$ and $\mathrm{Var}(Y) = \sigma^2$ and the moment based skewness and excess kurtosis of `SST`(μ, σ, ν, τ) are the same as for `ST3`$(\mu_0, \sigma_0, \nu, \tau)$, and hence the same as for `ST3`$(0, 1, \nu, \tau)$, depending only on ν and τ. In **gamlss.dist** the default link function for τ is a shifted log link function, $\log(\tau - 2)$, which ensures that τ is always in its valid range, i.e. $\tau > 2$.

Parameter ν mainly affects the skewness, with symmetry when $\nu = 1$, while parameter τ mainly affects the heaviness of the tails, with increasing τ decreasing the heaviness of the tails.

If $Y \sim$ `SST`(μ, σ, ν, τ) then $-Y \sim$ `SST`$(-\mu, \sigma, 1/\nu, \tau)$ is a reflection of the distribution of Y about zero. Also $W = (2\mu - Y) \sim$ `SST`$(\mu, \sigma, 1/\nu, \tau)$ is a reflection of the distribution of Y about μ. Also `SST`$(\mu, \sigma, 1, \tau) =$ `TF2`(μ, σ, τ).

Table 1.24: Skew Student t distribution.

$\texttt{SST}(\mu, \sigma, \nu, \tau)$	
Ranges	
Y	$-\infty < y < \infty$
μ	$-\infty < \mu < \infty$, mean, location shift parameter
σ	$0 < \sigma < \infty$, standard deviation, scaling parameter
ν	$0 < \nu < \infty$, skewness parameter
τ	$2 < \tau < \infty$, kurtosis parameter
Distribution measures	
mean	μ
median	$\begin{cases} \mu_0 + \dfrac{\sigma_0}{\nu} t_{\alpha_1,\tau} & \text{if } \nu \le 1 \\ \mu_0 + \sigma_0 \nu t_{\alpha_2,\tau} & \text{if } \nu > 1 \\ \text{where } \alpha_1 = \frac{(1+\nu^2)}{4}, \ \alpha_2 = \frac{(3\nu^2-1)}{4\nu^2} & \\ t_{\alpha,\tau} = F_T^{-1}(\alpha) \text{ where } T \sim t_\tau & \end{cases}$
mode	μ_0
variance	σ^2
skewness	equal to skewness of $\texttt{ST3}(0, 1, \nu, \tau)$
excess kurtosis	equal to excess kurtosis of $\texttt{ST3}(0, 1, \nu, \tau)$
pdf	$\begin{cases} \dfrac{c}{\sigma_0}\left[1 + \dfrac{\nu^2 z^2}{\tau}\right]^{-(\tau+1)/2} & \text{if } y < \mu_0 \\ \dfrac{c}{\sigma_0}\left[1 + \dfrac{z^2}{\nu^2 \tau}\right]^{-(\tau+1)/2} & \text{if } y \ge \mu_0, \\ \text{where } z = (y-\mu_0)/\sigma_0 \text{ and} & \\ c = 2\nu\left[(1+\nu^2) B\left(1/2, \tau/2\right) \tau^{1/2}\right]^{-1} & \end{cases}$
cdf	$\begin{cases} \dfrac{2}{(1+\nu^2)} F_T[\,\nu(y-\mu_0)/(\sigma_0)\,] & \text{if } y < \mu_0 \\ \dfrac{1}{(1+\nu^2)}\left\{1 + 2\nu^2\left[F_T\left(\dfrac{y-\mu_0}{\sigma_0 \nu}\right) - 0.5\right]\right\} & \text{if } y \ge \mu_0 \\ \text{where } T \sim t_\tau & \end{cases}$
inverse cdf	$\begin{cases} \mu_0 + \dfrac{\sigma_0}{\nu} t_{\alpha_3,\tau} & \text{if } p \le (1+\nu^2)^{-1} \\ \mu_0 + \sigma_0 \nu t_{\alpha_4,\tau} & \text{if } p > (1+\nu^2)^{-1} \end{cases}$ where $\alpha_3 = \frac{p(1+\nu^2)}{2}$ and $\alpha_4 = \frac{p(1+\nu^2)-1+\nu^2}{2\nu^2}$
Note	μ_0 and σ_0 are defined below equation (1.37)

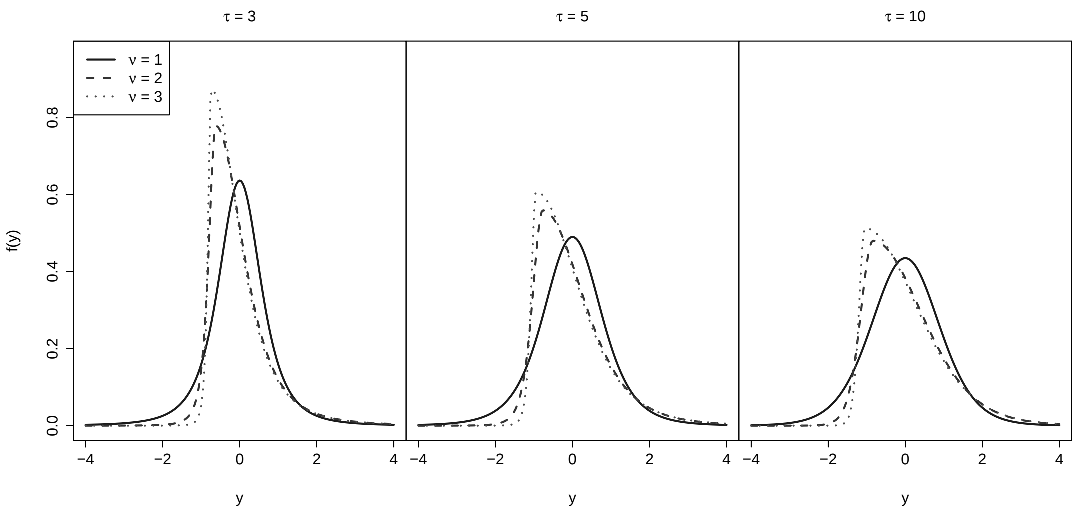

Figure 1.23: The skew Student t, $\texttt{SST}(\mu, \sigma, \nu, \tau)$, distribution, with $\mu = 0$, $\sigma = 1$, $\nu = 1, 2, 3$, and $\tau = 3, 5, 10$. It is symmetric if $\nu = 1$. Changing ν and $1/\nu$ reflects the distribution about $y = \mu$, i.e. $y = 0$ above.

1.4.12 Skew t type 1: ST1

The pdf of the skew t type 1 distribution, denoted by $\texttt{ST1}(\mu, \sigma, \nu, \tau)$, is given by

$$f_Y(y \mid \mu, \sigma, \nu, \tau) = \frac{2}{\sigma} f_{Z_1}(z) F_{Z_1}(\nu z) \tag{1.38}$$

for $-\infty < y < \infty$, where $-\infty < \mu < \infty$, $\sigma > 0$, $-\infty < \nu < \infty$ and $\tau > 0$, and where $z = (y - \mu)/\sigma$ and $f_{Z_1}(\cdot)$ and $F_{Z_1}(\cdot)$ are the pdf and cdf of $Z_1 \sim \texttt{TF}(0, 1, \tau) = t_\tau$, the t distribution with $\tau > 0$ degrees of freedom, with τ treated as a continuous parameter. This distribution is in the form of a type I distribution of Azzalini [1986], see Section **??**. (No summary table is given for ST1.)

If $Y \sim \texttt{ST1}(\mu, \sigma, \nu, \tau)$ then $-Y \sim \texttt{ST1}(-\mu, \sigma, -\nu, \tau)$ is a reflection of the distribution of Y about zero. Also $W = (2\mu - Y) \sim \texttt{ST1}(\mu, \sigma, -\nu, \tau)$ is a reflection of the distribution of Y about μ.

As $\nu \to \infty$, $\texttt{ST1}(\mu, \sigma, \nu, \tau)$ tends to a half t family distribution, i.e. a half $\texttt{TF}(\mu, \sigma, \tau)$, see Table **??**. Also $\texttt{ST1}(\mu, \sigma, 0, \tau) \equiv \texttt{TF}(\mu, \sigma, \tau)$.

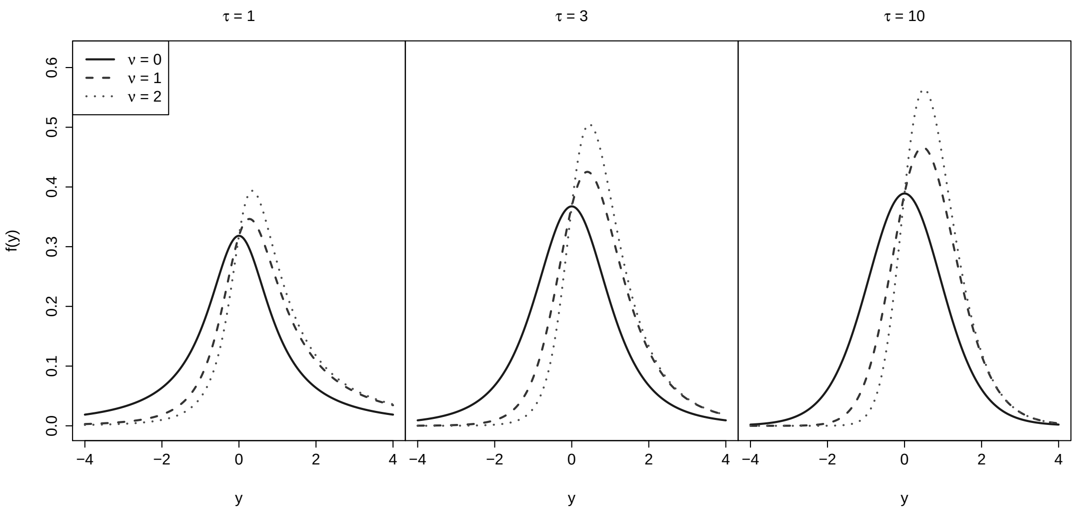

Figure 1.24: The skew t type 1, $\texttt{ST1}(\mu, \sigma, \nu, \tau)$, distribution, with $\mu = 0$, $\sigma = 1$, $\nu = 0, 1, 2$, and $\tau = 1, 3, 10$. It is symmetric if $\nu = 0$. Changing ν and $-\nu$ reflects the distribution about $y = \mu$, i.e. $y = 0$ above.

1.4.13 Skew t type 2: ST2

The pdf of the skew t type 2 distribution, denoted by $\texttt{ST2}(\mu, \sigma, \nu, \tau)$, is given by

$$f_Y(y \,|\, \mu, \sigma, \nu, \tau) = \frac{2}{\sigma} f_{Z_1}(z) F_{Z_2}(\omega) \tag{1.39}$$

for $-\infty < y < \infty$, where $-\infty < \mu < \infty$, $\sigma > 0$, $-\infty < \nu < \infty$, and $\tau > 0$, and where $z = (y - \mu)/\sigma$, $\omega = \nu \lambda^{1/2} z$, $\lambda = (\tau + 1)/(\tau + z^2)$, $f_{Z_1}(\cdot)$ is the pdf of $Z_1 \sim \texttt{TF}(0, 1, \tau) = t_\tau$ and $F_{Z_2}(\cdot)$ is the cdf of $Z_2 \sim \texttt{TF}(0, 1, \tau + 1) = t_{\tau+1}$. This distribution is the univariate case of the multivariate skew t distribution introduced by Azzalini and Capitanio [2003, p380, equation (26)]. See Section **??**.

If $Y \sim \texttt{ST2}(\mu, \sigma, \nu, \tau)$ then $-Y \sim \texttt{ST2}(-\mu, \sigma, -\nu, \tau)$ is a reflection of the distribution of Y about zero. Also $W = (2\mu - Y) \sim \texttt{ST2}(\mu, \sigma, -\nu, \tau)$ is a reflection of the distribution of Y about μ.

As $\nu \to \infty$, $\texttt{ST2}(\mu, \sigma, \nu, \tau)$ tends to a half t family distribution, i.e. a half $\texttt{TF}(\mu, \sigma, \tau)$, see Table 13.2. Also $\texttt{ST2}(\mu, \sigma, 0, \tau) = \texttt{TF}(\mu, \sigma, \tau)$.

Table 1.25: Skew t type 2 distribution.

	$\texttt{ST2}(\mu, \sigma, \nu, \tau)$
	Ranges
Y	$-\infty < y < \infty$
μ	$-\infty < \mu < \infty,$ location shift parameter
σ	$0 < \sigma < \infty,$ scaling parameter
ν	$-\infty < \nu < \infty,$ skewness parameter
τ	$0 < \tau < \infty,$ kurtosis parameter
	Distribution measures
mean	$\begin{cases} \mu + \sigma \mathrm{E}(Z) \\ \text{where } Z = (Y-\mu)/\sigma \text{ and} \\ \mathrm{E}(Z) = \dfrac{\nu \tau^{1/2} \Gamma([\tau-1]/2)}{(1+\nu^2)^{1/2} \pi^{1/2} \Gamma(\tau/2)}, \text{ for } \tau > 1 \end{cases}$
variance	$\sigma^2 \mathrm{Var}(Z) = \sigma^2 \left\{ \left(\dfrac{\tau}{\tau-2} \right) - [\mathrm{E}(Z)]^2 \right\}, \text{ for } \tau > 2$
skewness	$\begin{cases} \mu_{3Y}/[\mathrm{Var}(Y)]^{1.5} \text{ where} \\ \mu_{3Y} = \sigma^3 \mu_{3Z} = \sigma^3 \{ \mu_{3Z}^{'} - 3\mathrm{Var}(Z)\mathrm{E}(Z) - [\mathrm{E}(Z)]^3 \} \\ \mu_{3Z}^{'} = \mathrm{E}(Z^3) = \dfrac{\tau(3-\delta^2)}{(\tau-3)} \mathrm{E}(Z) \text{ for } \tau > 3 \\ \delta = \nu(1+\nu^2)^{-1/2} \end{cases}$
excess kurtosis	$\begin{cases} \mu_{4Y}/[\mathrm{Var}(Y)]^2 - 3 \text{ where} \\ \mu_{4Y} = \sigma^4 \mu_{4Z} = \sigma^4 \{ \mu_{4Z}^{'} - 4\mu_{3Z}^{'} \mathrm{E}(Z) + 6\mathrm{Var}(Z)[\mathrm{E}(Z)]^2 \\ \quad +3[\mathrm{E}(Z)]^4 \} \\ \mu_{4Z}^{'} = \mathrm{E}(Z^4) = \dfrac{3\tau^2}{(\tau-2)(\tau-4)} \text{ for } \tau > 4 \end{cases}$
pdf	$\begin{cases} \dfrac{2}{\sigma} f_{Z_1}(z) F_{Z_2}(\omega) \\ \text{where } z = (y-\mu)/\sigma,\ \omega = \nu \lambda^{1/2} z,\ \lambda = (\tau+1)/(\tau+z^2) \\ \text{and } Z_1 \sim t_\tau \text{ and } Z_2 \sim t_{\tau+1} \end{cases}$
Reference	Azzalini and Capitanio [2003], p380, equation (26) and p382 with dimension $d = 1$ and $(\xi, \omega, \alpha, \nu)$ and Ω replaced by (μ, σ, ν, τ) and σ^2, respectively, giving the pdf of Y and moments of Z.

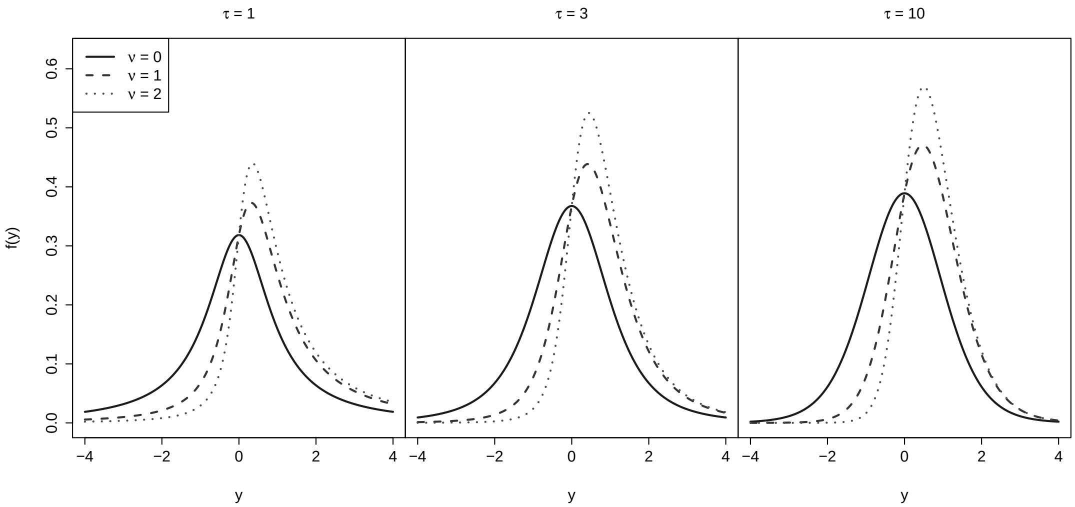

Figure 1.25: The skew t type 2, ST2(μ, σ, ν, τ), distribution, with $\mu = 0$, $\sigma = 1$, $\nu = 0, 1, 2$, and $\tau = 1, 3, 10$. It is symmetric if $\nu = 0$. Changing ν to $-\nu$ reflects the distribution about $y = \mu$, i.e. $y = 0$ above.

1.4.14 Skew t type 3: ST3

This is a 'scale-spliced' distribution (see Section **??**), denoted by ST3(μ, σ, ν, τ), with pdf

$$f_Y(y \mid \mu, \sigma, \nu, \tau) = \begin{cases} \dfrac{c}{\sigma}\left(1 + \dfrac{\nu^2 z^2}{\tau}\right)^{-(\tau+1)/2} & \text{if } y < \mu \\ \dfrac{c}{\sigma}\left(1 + \dfrac{z^2}{\nu^2 \tau}\right)^{-(\tau+1)/2} & \text{if } y \geq \mu \end{cases} \tag{1.40}$$

for $-\infty < y < \infty$, where $-\infty < \mu < \infty$, $\sigma > 0$, $\nu > 0$, and $\tau > 0$, and where $z = (y - \mu)/\sigma$ and $c = 2\nu[(1 + \nu^2)B(1/2, \tau/2)\tau^{1/2}]^{-1}$, [Fernandez and Steel [1998], p362, equation (13), with (γ, ν) replaced by (ν, τ)]. Note that μ is the mode of Y. The moments of Y in Table 1.26 are obtained using Fernandez and Steel [1998], p360, equation (5).

Parameter ν mainly affects the skewness, with symmetry when $\nu = 1$, while parameter τ mainly affects the heaviness of the tails with increasing τ decreasing the heaviness of the tails. The ST3 distribution appears to be positively skew if $\nu > 1$, negatively skew if $0 < \nu < 1$, and always leptokurtic. Figures **??**(c), **??**(b), and **??**(b) show moment and centile kurtosis-skewness plots for ST3(μ, σ, ν, τ).

If $Y \sim$ `ST3`(μ, σ, ν, τ) then $-Y \sim$ `ST3`$(-\mu, \sigma, 1/\nu, \tau)$ is a reflection of the distribution of Y about zero. Also $W = (2\mu - Y) \sim$ `ST3`$(\mu, \sigma, 1/\nu, \tau)$ is a reflection of the distribution of Y about μ. As $\nu \to \infty$, `ST3`(μ, σ, ν, τ) tends to a half t family distribution, i.e. a half `TF`(μ, σ, τ), see Table **??**. As $\tau \to \infty$, `ST3`(μ, σ, ν, τ) tends to a skew normal type 2, `SN2`(μ, σ, ν), distribution, see Table **??**. As $\nu \to \infty$ and $\tau \to \infty$, `ST3`(μ, σ, ν, τ) tends to a half normal distribution, i.e. a half `NO`(μ, σ). Also `ST3`$(\mu, \sigma, 1, \tau) =$ `TF`(μ, σ, τ).

Table 1.26: Skew t type 3 distribution.

$\texttt{ST3}(\mu, \sigma, \nu, \tau)$	
Ranges	
Y	$-\infty < y < \infty$
μ	$-\infty < \mu < \infty$, mode, location shift parameter
σ	$0 < \sigma < \infty$, scaling parameter
ν	$0 < \nu < \infty$, skewness parameter
τ	$0 < \tau < \infty$, kurtosis parameter
Distribution measures	
mean	$\begin{cases} \mu + \sigma \mathrm{E}(Z) = \mu + \dfrac{2\sigma\tau^{1/2}(\nu - \nu^{-1})}{(\tau-1)B(1/2, \tau/2)}, \text{ for } \tau > 1 \\ \text{where } Z = (Y-\mu)/\sigma \end{cases}$
median	$\begin{cases} \mu + \frac{\sigma}{\nu} t_{\alpha_1,\tau} & \text{if } \nu \le 1 \\ \mu + \sigma\nu t_{\alpha_2,\tau} & \text{if } \nu > 1 \\ \text{where } \alpha_1 = \frac{(1+\nu^2)}{4}, \ \alpha_2 = \frac{(3\nu^2-1)}{4\nu^2}, & t_{\alpha,\tau} = F_T^{-1}(\alpha) \text{ where } T \sim t_\tau \end{cases}$
mode	μ
variance	$\sigma^2 \mathrm{Var}(Z) = \sigma^2 \left\{ \left(\dfrac{\tau}{\tau-2} \right) (\nu^2 + \nu^{-2} - 1) - [\mathrm{E}(Z)]^2 \right\}$, for $\tau > 2$
skewness	$\begin{cases} \mu_{3Y}/[\mathrm{Var}(Y)]^{1.5} \text{ where} \\ \mu_{3Y} = \sigma^3 \mu_{3Z} = \sigma^3 \{ \mu'_{3Z} - 3\mathrm{Var}(Z)\mathrm{E}(Z) - [\mathrm{E}(Z)]^3 \} \\ \mu'_{3Z} = \mathrm{E}(Z^3) = \dfrac{4\tau^{3/2}\left(\nu^4 - \nu^{-4}\right)}{(\tau-1)(\tau-3)B(1/2, \tau/2)\left(\nu + \nu^{-1}\right)}, \text{ for } \tau > 3 \end{cases}$
excess kurtosis	$\begin{cases} \mu_{4Y}/[\mathrm{Var}(Y)]^2 - 3 \text{ where} \\ \mu_{4Y} = \sigma^4 \mu_{4Z} = \sigma^4 \{ \mu'_{4Z} - 4\mu'_{3Z}\mathrm{E}(Z) + 6\mathrm{Var}(Z)[\mathrm{E}(Z)]^2 + 3[\mathrm{E}(Z)]^4 \} \\ \mu'_{4Z} = \mathrm{E}(Z^4) = \dfrac{3\tau^2\left(\nu^5 + \nu^{-5}\right)}{(\tau-2)(\tau-4)\left(\nu + \nu^{-1}\right)}, \text{ for } \tau > 4 \end{cases}$
pdf [a]	$\begin{cases} \dfrac{c}{\sigma}\left(1 + \dfrac{\nu^2 z^2}{\tau}\right)^{-(\tau+1)/2} & \text{if } y < \mu \\ \dfrac{c}{\sigma}\left(1 + \dfrac{z^2}{\nu^2 \tau}\right)^{-(\tau+1)/2} & \text{if } y \ge \mu \\ \text{where } z = (y-\mu)/\sigma \text{ and} \\ c = 2\nu[(1+\nu^2)B(1/2, \tau/2)\tau^{1/2}]^{-1} \end{cases}$
cdf	$\begin{cases} \dfrac{2}{(1+\nu^2)} F_T[\nu(y-\mu)/\sigma] & \text{if } y < \mu \\ \dfrac{1}{(1+\nu^2)} \left[1 + 2\nu^2 \left\{ F_T[(y-\mu)/(\sigma\nu)] - \frac{1}{2} \right\} \right] & \text{if } y \ge \mu \end{cases}$
inverse cdf	$\begin{cases} \mu + \dfrac{\sigma}{\nu} t_{\alpha_3,\tau} & \text{if } p \le (1+\nu^2)^{-1} \\ \mu + \sigma\nu t_{\alpha_4,\tau} & \text{if } p > (1+\nu^2)^{-1} \\ \text{where } \alpha_3 = \frac{p(1+\nu^2)}{2}, \ \alpha_4 = \frac{p(1+\nu^2) - 1 + \nu^2}{2\nu^2} \end{cases}$
Reference	[a] From Fernandez and Steel [1998], p362, equation (13), with (γ, ν) replaced by (ν, τ).

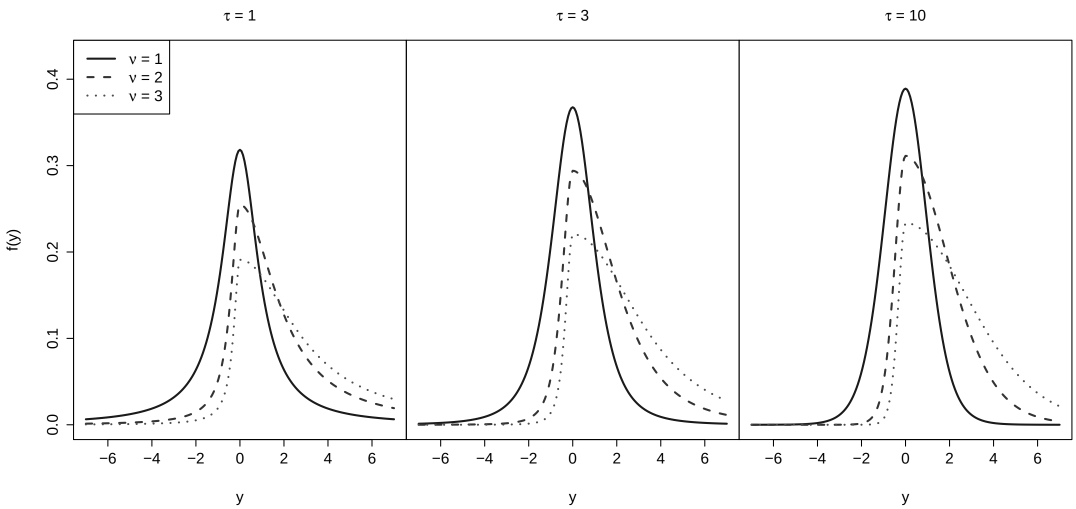

Figure 1.26: The skew t type 3, ST3(μ, σ, ν, τ), distribution, with $\mu = 0$, $\sigma = 1$, $\nu = 1, 2, 3$, and $\tau = 1, 3, 10$. It is symmetric if $\nu = 1$. Changing ν to $1/\nu$ reflects the distribution about $y = \mu$, i.e. $y = 0$ above.

1.4.15 Skew t type 4: ST4

This is a 'shape-spliced', distribution, (see Section **??**), denoted by ST4(μ, σ, ν, τ), with pdf

$$f_Y(y \mid \mu, \sigma, \nu, \tau) = \begin{cases} \dfrac{c}{\sigma}\left(1 + \dfrac{z^2}{\nu}\right)^{-(\nu+1)/2} & \text{if } y < \mu \\ \dfrac{c}{\sigma}\left(1 + \dfrac{z^2}{\tau}\right)^{-(\tau+1)/2} & \text{if } y \geq \mu \end{cases} \tag{1.41}$$

for $-\infty < y < \infty$, where $-\infty < \mu < \infty$, $\sigma > 0$, $\nu > 0$ and $\tau > 0$, and where $z = (y - \mu)/\sigma$ and $c = 2\left[\nu^{1/2} B\left(1/2, \nu/2\right) + \tau^{1/2} B\left(1/2, \tau/2\right)\right]^{-1}$.

Parameters ν and τ affect the left and right tail heaviness respectively, with increasing ν (τ) decreasing the left (right) tail heaviness, both always heavier than the normal distribution tails. The ST4(μ, σ, ν, τ) distribution is symmetric if $\nu = \tau$ and ST4$(\mu, \sigma, \nu, \nu) \equiv$ TF(μ, σ, ν). If $Y \sim$ ST4(μ, σ, ν, τ) then $-Y \sim$ ST4$(-\mu, \sigma, \tau, \nu)$ is a reflection of the distribution of Y about zero. Also $W = (2\mu - Y) \sim$ ST4(μ, σ, τ, ν) is a reflection of the distribution of Y about μ.

Table 1.27: Skew t type 4 distribution.

	$\texttt{ST4}(\mu, \sigma, \nu, \tau)$
	Ranges
Y	$-\infty < y < \infty$
μ	$-\infty < \mu < \infty,$ mode, location shift parameter
σ	$0 < \sigma < \infty,$ scaling parameter
ν	$0 < \nu < \infty,$ left tail heaviness parameter
τ	$0 < \tau < \infty,$ right tail heaviness parameter
	Distribution measures
mean	$\begin{cases} \mu + \sigma E(Z) = \mu + \sigma c\left[\tau/(\tau-1) - \nu/(\nu-1)\right] \\ \text{for } \nu > 1 \text{ and } \tau > 1 \text{ where } Z = (Y-\mu)/\sigma \text{ , } c = 2[b_1 + b_2]^{-1} \\ b_1 = \nu^{1/2} B(1/2, \nu/2) \text{ , } b_2 = \tau^{1/2} B(1/2, \tau/2) \end{cases}$
median	$\begin{cases} \mu + \sigma t_{\alpha_1, \nu} & \text{if } k \le 1 \\ \mu + \sigma t_{\alpha_2, \tau} & \text{if } k > 1 \\ \text{where } \alpha_1 = \frac{(1+k)}{4} \text{ , } \alpha_2 = \frac{(3k-1)}{4k} \text{ , } k = b_2/b_1, & \\ t_{\alpha,\tau} = F_T^{-1}(\alpha) \text{ where } T \sim t_\tau & \end{cases}$
mode	μ
Variance	$\begin{cases} \sigma^2 \text{Var}(Z) = \sigma^2\{\text{E}(Z^2) - [\text{E}(Z)]^2\} \text{ where} \\ \text{E}(Z^2) = \dfrac{c\tau b_2}{2(\tau-2)} + \dfrac{c\nu b_1}{2(\nu-2)} \text{ for } \nu > 2 \text{ and } \tau > 2 \end{cases}$
skewness	$\begin{cases} \mu_{3Y}/[\text{Var}(Y)]^{1.5} \text{ where} \\ \mu_{3Y} = \sigma^3 \mu_{3Z} = \sigma^3\{\mu_{3Z}^{'} - 3\text{Var}(Z)\text{E}(Z) - [\text{E}(Z)]^3\} \\ \mu_{3Z}^{'} = \text{E}(Z^3) = 2c\left[\dfrac{\tau^2}{(\tau-1)(\tau-3)} - \dfrac{\nu^2}{(\nu-1)(\nu-3)}\right] \\ \text{for } \nu > 3 \text{ and } \tau > 3 \end{cases}$
excess kurtosis	$\begin{cases} \mu_{4Y}/[\text{Var}(Y)]^2 - 3 \text{ where} \\ \mu_{4Y} = \sigma^4 \mu_{4Z} = \\ \qquad \sigma^4\{\mu_{4Z}^{'} - 4\mu_{3Z}^{'}\text{E}(Z) + 6\text{Var}(Z)[\text{E}(Z)]^2 + 3[\text{E}(Z)]^4\} \\ \mu_{4Z}^{'} = E(Z^4) = 3 + 3c\left[\dfrac{b_2}{(\tau-4)} + \dfrac{b_1}{(\nu-4)}\right] \text{ for } \nu > 4 \text{ and } \tau > 4 \end{cases}$
pdf	$\begin{cases} \dfrac{c}{\sigma}\left(1 + \dfrac{z^2}{\nu}\right)^{-(\nu+1)/2} & \text{if } y < \mu \\ \dfrac{c}{\sigma}\left(1 + \dfrac{z^2}{\tau}\right)^{-(\tau+1)/2} & \text{if } y \ge \mu \\ \text{where } z = (y-\mu)/\sigma & \end{cases}$
cdf	$\begin{cases} \frac{2}{(1+k)} F_{T_1}(z) & \text{if } y < \mu \\ \{1 + 2k\left[F_{T_2}(z) - 0.5\right]\}/(1+k) & \text{if } y \ge \mu \\ \text{where } T_1 \sim t_\nu,\ T_2 \sim t_\tau,\ z = (y-\mu)/\sigma & \end{cases}$
inverse cdf	$\begin{cases} \mu + \sigma t_{\alpha_3, \nu} & \text{if } p \le (1+k)^{-1} \\ \mu + \sigma t_{\alpha_4, \tau} & \text{if } p > (1+k)^{-1} \\ \text{where } \alpha_3 = \frac{p(1+k)}{2} \text{ , } \alpha_4 = \frac{p(1+k)-1}{2k} + 0.5 & \end{cases}$

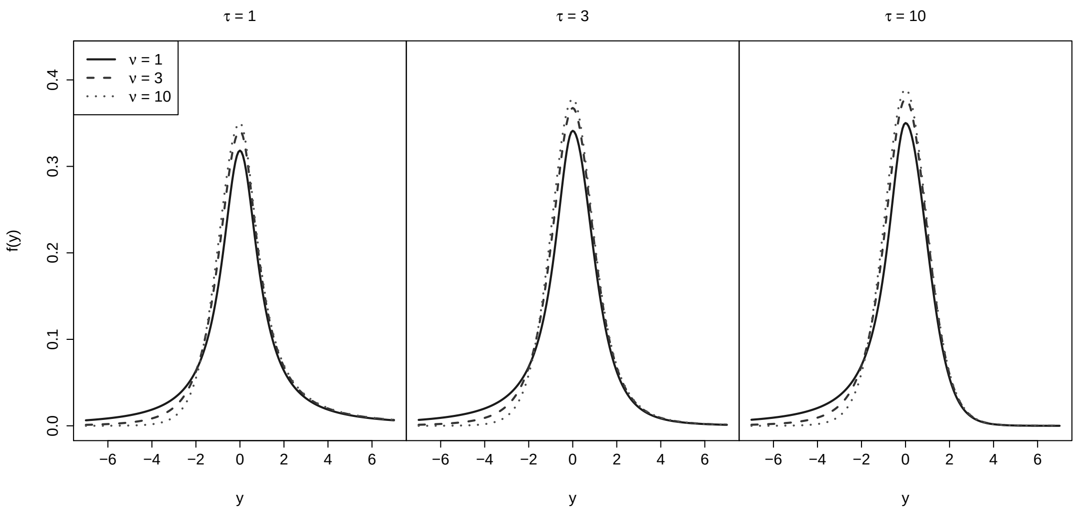

Figure 1.27: The skew t type 4, ST4(μ, σ, ν, τ), distribution, with $\mu = 0$, $\sigma = 1$, $\nu = 1, 3, 10$, and $\tau = 1, 3, 10$. It is symmetric if $\nu = \tau$. Interchanging ν and τ reflects the distribution about $y = \mu$, i.e. $y = 0$ above.

1.4.16 Skew t type 5: ST5

The pdf of the skew t distribution type 5, denoted by ST5(μ, σ, ν, τ), [Jones and Faddy, 2003] is given by

$$f_Y(y \mid \mu, \sigma, \nu, \tau) = \frac{c}{\sigma} \left[1 + \frac{z}{(a+b+z^2)^{1/2}}\right]^{a+1/2} \left[1 - \frac{z}{(a+b+z^2)^{1/2}}\right]^{b+1/2}$$

for $-\infty < y < \infty$, where $-\infty < \mu < \infty$, $\sigma > 0$, $-\infty < \nu < \infty$ and $\tau > 0$, and where $z = (y - \mu)/\sigma$ and

$$\begin{aligned} c &= \left[2^{a+b-1}(a+b)^{1/2} B(a,b)\right]^{-1} \\ \nu &= (a-b)/\left[ab(a+b)\right]^{1/2} \\ \tau &= 2/(a+b) \ . \end{aligned}$$

Hence

$$\begin{aligned} a &= \tau^{-1}[1 + \nu(2\tau + \nu^2)^{-1/2}] \\ b &= \tau^{-1}[1 - \nu(2\tau + \nu^2)^{-1/2}] \ . \end{aligned} \quad (1.42)$$

From Jones and Faddy [2003] p160, equation (2), if $B \sim$ BEo(a, b) then

$$Z = \frac{(a+b)^{1/2}(2B-1)}{2[B(1-B)]^{1/2}} \sim \texttt{ST5}(0, 1, \nu, \tau) \ .$$

Hence $B = \frac{1}{2}\left[1 + Z(a+b+Z^2)^{-1/2}\right] \sim \texttt{BEo}(a,b)$, from which the cdf of Y is obtained. Parameter ν is a skewess parameter with symmetry when $\nu = 0$. The distribution is positively moment skewed if $\nu > 0$ and negatively moment skewed if $\nu < 0$, [Jones and Faddy, 2003, p159]. If $Y \sim \texttt{ST5}(\mu, \sigma, \nu, \tau)$ then $-Y \sim \texttt{ST5}(-\mu, \sigma, -\nu, \tau)$ is a reflection of the distribution of Y about zero. Also $W = (2\mu - Y) \sim \texttt{ST5}(\mu, \sigma, -\nu, \tau)$ is a reflection of the distribution of Y about μ.

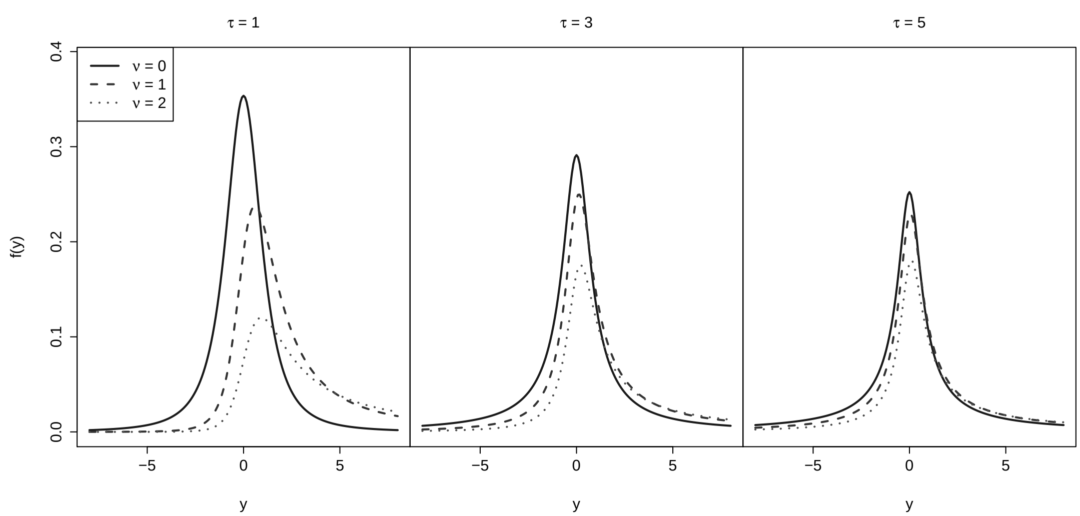

Figure 1.28: The skew t type 5, $\texttt{ST5}(\mu, \sigma, \nu, \tau)$, distribution, with $\mu = 0$, $\sigma = 1$, $\nu = 0, 1, 2$, and $\tau = 1, 3, 5$. It is symmetric if $\nu = 0$. Changing ν to $-\nu$ reflects the distribution about $y = \mu$, i.e. $y = 0$ above.

Table 1.28: Skew t type 5 distribution.

	`ST5`(μ,σ,ν,τ)
	Ranges
Y	$-\infty < y < \infty$
μ	$-\infty < \mu < \infty$, location shift parameter
σ	$0 < \sigma < \infty$, scaling parameter
ν	$-\infty < \nu < \infty$, skewness parameter
τ	$0 < \tau < \infty$, kurtosis parameter
	Distribution measures
mean	$\begin{cases} \mu + \sigma \mathrm{E}(Z) = \mu + \dfrac{\sigma(a+b)^{1/2}(a-b)\Gamma(a-1/2)\Gamma(b-1/2)}{2\Gamma(a)\Gamma(b)} \\ \text{for } a > 1/2 \text{ and } b > 1/2 \text{ where } Z = (Y-\mu)/\sigma \end{cases}$
mode	$\mu + \dfrac{\sigma(a+b)^{1/2}(a-b)}{(2a+1)^{1/2}(2b+1)^{1/2}}$
variance	$\begin{cases} \sigma^2 \mathrm{Var}(Z) = \sigma^2 \left\{ \dfrac{(a+b)[(a-b)^2+a+b-2]}{4(a-1)(b-1)} - [\mathrm{E}(Z)]^2 \right\} \\ \text{for } a > 1 \text{ and } b > 1 \end{cases}$
skewness	$\begin{cases} \mu_{3Y}/[\mathrm{Var}(Y)]^{1.5} \text{ where} \\ \mu_{3Y} = \sigma^3 \mu_{3Z} = \sigma^3 \{\mu'_{3Z} - 3\mathrm{Var}(Z)\mathrm{E}(Z) - [\mathrm{E}(Z)]^3\} \\ \mu'_{3Z} = \mathrm{E}(Z^3) = \dfrac{(a+b)^{3/2}\Gamma(a-3/2)\Gamma(b-3/2)}{8\Gamma(a)\Gamma(b)} \times \\ \quad \left[a^3+3a^2-7a-b^3-3b^2+7b+3ab^2-3a^2b\right] \\ \text{for } a > 3/2 \text{ and } b > 3/2 \end{cases}$
excess kurtosis	$\begin{cases} \mu_{4Y}/[\mathrm{Var}(Y)]^2 - 3 \text{ where} \\ \mu_{4Y} = \sigma^4 \mu_{4Z} = \sigma^4 \{\mu'_{4Z} - 4\mu'_{3Z}\mathrm{E}(Z) + 6\mathrm{Var}(Z)[\mathrm{E}(Z)]^2 + 3[\mathrm{E}(Z)]^4\} \\ \text{where } \mu'_{4Z} = \mathrm{E}(Z^4) = \dfrac{(a+b)^2}{16(a-1)(a-2)(b-1)(b-2)} \times \\ \quad \left[a^4-2a^3-a^2+2a+b^4-2b^3-b^2+2b+ \right. \\ \quad \left. 2(a-2)(b-2)(3ab-2a^2-2b^2-a-b+3)\right] \\ \text{for } a > 2 \text{ and } b > 2 \end{cases}$
pdf	$\begin{cases} f_Y(y \mid \mu,\sigma,\nu,\tau) = \dfrac{c}{\sigma}\left[1+\frac{z}{(a+b+z^2)^{1/2}}\right]^{a+1/2}\left[1-\frac{z}{(a+b+z^2)^{1/2}}\right]^{b+1/2} \\ \text{where } z = (y-\mu)/\sigma, c = \left[2^{a+b-1}(a+b)^{1/2}B(a,b)\right]^{-1} \\ a = \tau^{-1}[1+d] \text{ and } b = \tau^{-1}[1-d] \text{ where } d = \nu(2\tau+\nu^2)^{-1/2} \end{cases}$
cdf	$\begin{cases} \dfrac{B(a,b,r)}{B(a,b)} \\ \text{where } r = \frac{1}{2}[1+z(a+b+z^2)^{-1/2}] \text{ and } z = (y-\mu)/\sigma \end{cases}$
Reference	Jones and Faddy [2003], p159, equation (1) reparameterized as on their p164 with (q,p) replaced by (ν,τ), and p162 equations (4b) and (5) giving the pdf, moments and mode of Z, respectively.

Chapter 2

Continuous distributions on $(0, \infty)$

This chapter gives summary tables and plots for the explicit **gamlss.dist** continuous distributions with range $(0, \infty)$. These are discussed in Chapter **??**. Section **??** discuesses creating distributions on $(0, \infty)$ in **gamlss**, either by an inverse log transform or by truncation below zero, from any **gamlss.dist** distribution on $(-\infty, \infty)$.

2.1 Scale family of distributions

A continuous random variable Y defined on $(0, \infty)$, is said to have a scale family of distributions with scaling parameter θ (for fixed values of all other parameters of the distribution) if

$$Z = \frac{Y}{\theta}$$

has a cdf which does not depend on θ. Hence

$$F_Y(y) = F_Z\left(\frac{y}{\theta}\right)$$

and

$$f_Y(y) = \frac{1}{\theta} f_Z\left(\frac{y}{\theta}\right),$$

so $F_Y(y)$ and $\theta f_Y(y)$ only depend on y and θ through the function $z = y/\theta$.

Example: Let Y have a Weibull distribution, $Y \sim \texttt{WEI}(\mu, \sigma)$, then Y has a scale family of distributions with scaling parameter μ (for a fixed value of σ), since

$F_Y(y) = 1 - \exp\left[-\left(y/\mu\right)^{\sigma}\right]$ and hence $Z = Y/\mu$ has cdf $F_Z(z) = 1 - \exp\left[-\left(z\right)^{\sigma}\right]$ which does not depend on μ. Note $Z \sim$ `WEI`$(1, \sigma)$.

All distributions in **gamlss.dist** with range $(0, \infty)$ are scale families of distributions with scaling parameter μ, except for `GAF`, `IG`, `LNO`, `LOGNO`, and `WEI2`. Hence for these scale family distributions, $Y = \mu Z$ where the distribution of Z does not depend on μ. Hence for fixed parameers other than μ, $\mathrm{E}(Y) = \mu \mathrm{E}(Z)$ and $\mathrm{Var}(Y) = \mu^2 \mathrm{Var}(Z)$. Hence, $\mathrm{E}(Y)$ is proportional to μ provided $\mathrm{E}(Z)$ is finite, $\mathrm{Var}(Y)$ is proportional to μ^2, and the standard deviation of Y is proportional to μ, provided $\mathrm{Var}(Z)$ is finite. Note also that $\mathrm{Var}(Y) = [\mathrm{E}(Y)]^2 \mathrm{Var}(Z)/[\mathrm{E}(Z)]^2$, so for these scale family distributions the variance-mean relationship is squared, for fixed parameters other than μ. The gamma family, `GAF`(μ, σ, ν), distribution allows for a power variance-mean relationship, since it has mean μ and variance $\sigma^2 \mu^{\nu}$.

For scale family distributions plotted in this chapter we fix $\mu = 1$, since changing μ from 1 to say μ_1 just scales the horizontal axis in the figures by the factor μ_1 and scales the vertical axis by the factor $1/\mu_1$.

2.2 Continuous one-parameter distribution on $(0, \infty)$

2.2.1 Exponential: `EXP`

This is the only one-parameter continuous distribution in the **gamlss.dist** package. The pdf of the exponential distribution, denoted by `EXP`(μ), is given by

$$f_Y(y \mid \mu) = \frac{1}{\mu} \exp\left(-\frac{y}{\mu}\right) \tag{2.1}$$

for $y > 0$, where $\mu > 0$. Hence $\mathrm{E}(Y) = \mu$ and $\mathrm{Var}(Y) = \mu^2$. The exponential distribution is appropriate for moderately positively skewed data see Figure **??**.

Table 2.1: Exponential distribution.

EXP(μ)	
Ranges	
Y	$0 < y < \infty$
μ	$0 < \mu < \infty$, mean, scaling parameter
Distribution measures	
mean	μ
median	$\mu \log 2 \approx 0.69315\mu$
mode	$\to 0$
variance	μ^2
skewness	2
excess kurtosis	6
MGF	$(1 - \mu t)^{-1}$ for $t < 1/\mu$
pdf	$(1/\mu) \exp(-y/\mu)$
cdf	$1 - \exp(-y/\mu)$
inverse cdf	$-\mu \log(1 - p)$
Reference	Johnson et al. [1994] Chapter 19, p494-499, or set $\sigma = 1$ in GA(μ, σ)

2.3 Continuous two-parameter distributions on $(0, \infty)$

2.3.1 Gamma: GA

The pdf of the gamma distribution, denoted by GA(μ, σ), is given by

$$f_Y(y \mid \mu, \sigma) = \frac{y^{1/\sigma^2 - 1} e^{-y/(\sigma^2 \mu)}}{(\sigma^2 \mu)^{1/\sigma^2} \Gamma(1/\sigma^2)} \tag{2.2}$$

for $y > 0$, where $\mu > 0$ and $\sigma > 0$. Here $\mathrm{E}(Y) = \mu$ and $\mathrm{Var}(Y) = \sigma^2 \mu^2$ and $\mathrm{E}(Y^r) = \mu^r \sigma^{2r} \Gamma\left(1/\sigma^2 + r\right) / \Gamma\left(1/\sigma^2\right)$ for $r > -1/\sigma^2$.

This a reparameterization of Johnson et al. [1994] p343, equation (17.23), obtained by setting $\alpha = 1/\sigma^2$ and $\beta = \mu\sigma^2$. Hence $\mu = \alpha\beta$ and $\sigma^2 = 1/\alpha$.

The gamma distribution is appropriate for positively skewed data.

Parameter σ is a primary true van Zwet skewness parameter, (see van Zwet [1964a,b]). Note also that σ is a true variance parameter (since the variance inscreases monotonically with σ, for fixed μ), and σ is also a primary true moment skewness parameter and a primary true moment kurtosis parameter (since the moment skewness and moment kurtosis increase monotonically with σ, irrespective of the value of μ), see Sections **??** and **??**.

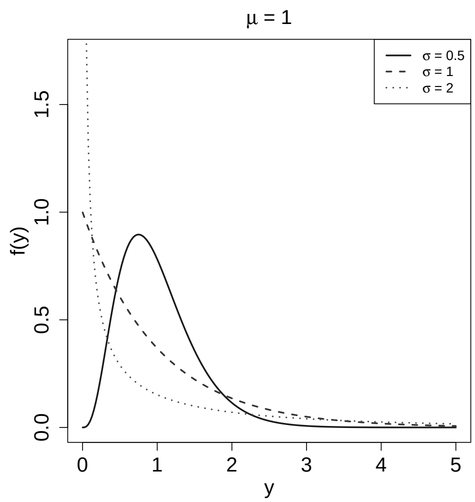

Figure 2.1: The gamma, `GA`(μ, σ), distribution, with $\mu = 1$ and $\sigma = 0.5, 1, 2$.

Table 2.2: Gamma distribution.

	$\texttt{GA}(\mu, \sigma)$
	Ranges
Y	$0 < y < \infty$
μ	$0 < \mu < \infty$, mean, scaling parameter
σ	$0 < \sigma < \infty$, coefficient of variation
	Distribution measures
mean [a]	μ
mode	$\begin{cases} \mu(1-\sigma^2) & \text{if } \sigma < 1 \\ \to 0 & \text{if } \sigma \geq 1 \end{cases}$
variance [a]	$\sigma^2\mu^2$
skewness [a]	2σ
excess kurtosis [a]	$6\sigma^2$
MGF	$(1-\mu\sigma^2 t)^{-1/\sigma^2}$ for $t < (\mu\sigma^2)^{-1}$
pdf [a,a2]	$\dfrac{y^{1/\sigma^2-1}e^{-y/(\sigma^2\mu)}}{(\sigma^2\mu)^{1/\sigma^2}\Gamma(\sigma^{-2})}$
cdf	$\dfrac{\gamma(\sigma^{-2}, y\mu^{-1}\sigma^{-2})}{\Gamma(\sigma^{-2})}$
Reference	[a] Derived from McCullagh and Nelder [1989] p287 reparameterized by $\mu = \mu$ and $\nu = 1/\sigma^2$ [a2] Johnson et al. [1994] p343, equation (17.23) reparameterized by $\alpha = 1/\sigma^2$ and $\beta = \mu\sigma^2$.
Note	$\gamma(a, x) = \int_0^x t^{a-1}e^{-t}dt$, the incomplete gamma function.

2.3.2 Inverse gamma: IGAMMA

The pdf of the inverse gamma distribution, denoted by IGAMMA(μ, σ), is given by

$$f_Y(y \,|\, \mu, \sigma) = \frac{\mu^\alpha(\alpha+1)^\alpha y^{-(\alpha+1)}}{\Gamma(\alpha)} \exp\left[-\frac{\mu(\alpha+1)}{y}\right]$$

for $y > 0$, where $\mu > 0$ and $\sigma > 0$ and where $\alpha = 1/\sigma^2$. The inverse gamma distribution is a reparameterized special case of the generalized gamma (GG) distribution given by IGAMMA$(\mu, \sigma) =$ GG$((1 + \sigma^2)\mu,\ \sigma, -1)$.

The inverse gamma distribution is appropriate for highly positively skewed data. Note that σ is a true variance parameter (for $\sigma^2 < 1/2$), and a primary true moment skewness parameter (for $\sigma^2 < 1/3$), and a primary true moment kurtosis parameter (for $\sigma^2 < 1/4$), see Sections **??** and **??**.

Table 2.3: Inverse gamma distribution.

	`IGAMMA`(μ,σ)
	Ranges
Y	$0<y<\infty$
μ	$0<\mu<\infty$, mode, scaling parameter
σ	$0<\sigma<\infty$
	Distribution measures
mean [a]	$\begin{cases} \dfrac{(1+\sigma^2)\mu}{(1-\sigma^2)} & \text{if } \sigma^2<1 \\ \infty & \text{if } \sigma^2\ge 1 \end{cases}$
mode [a]	μ
variance [a]	$\begin{cases} \dfrac{(1+\sigma^2)^2\mu^2\sigma^2}{(1-\sigma^2)^2(1-2\sigma^2)} & \text{if } \sigma^2<1/2 \\ \infty & \text{if } \sigma^2\ge 1/2 \end{cases}$
skewness [a]	$\begin{cases} \dfrac{4\sigma(1-2\sigma^2)^{1/2}}{(1-3\sigma^2)} & \text{if } \sigma^2<1/3 \\ \infty & \text{if } \sigma^2\ge 1/3 \end{cases}$
excess kurtosis [a]	$\begin{cases} \dfrac{3\sigma^2(10-22\sigma^2)}{(1-3\sigma^2)(1-4\sigma^2)} & \text{if } \sigma^2<1/4 \\ \infty & \text{if } \sigma^2\ge 1/4 \end{cases}$
pdf [a]	$\dfrac{\mu^\alpha(\alpha+1)^\alpha y^{-(\alpha+1)}}{\Gamma(\alpha)}\exp\left[-\dfrac{\mu(\alpha+1)}{y}\right]$, where $\alpha=1/\sigma^2$
cdf [b]	$\dfrac{\Gamma\left[\,\alpha,\mu(\alpha+1)/y\,\right]}{\Gamma(\alpha)}$
Note	[a] Set $\mu_1=(1+\sigma^2)\mu$, $\sigma_1=\sigma$ and $\nu_1=-1$ (so $\theta=1/\sigma_1^2$) in `GG`(μ_1,σ_1,ν_1). [b] $\Gamma(a,x)=\int_x^\infty t^{a-1}e^{-t}dt$, the complement of the incomplete gamma function.

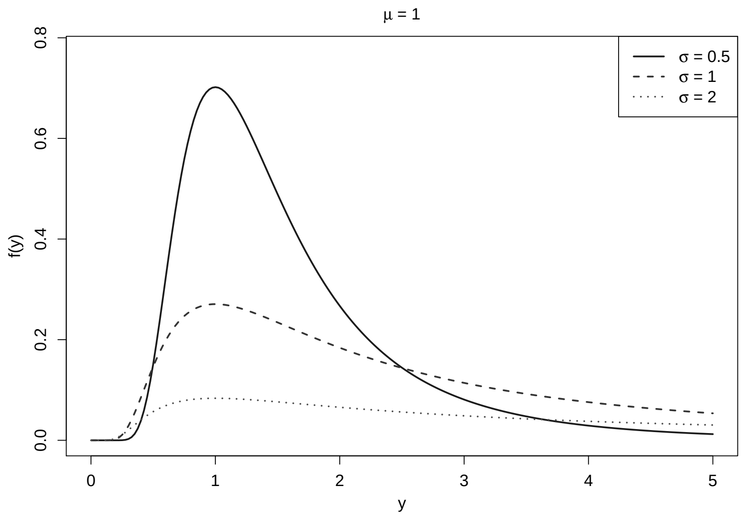

Figure 2.2: The inverse gamma, IGAMMA(μ, σ), distribution, with $\mu = 1$ and $\sigma = 0.5, 1, 2$.

2.3.3 Inverse Gaussian: IG

The pdf of the inverse Gaussian distribution, denoted by IG(μ, σ) is given by

$$f_Y(y \mid \mu, \sigma) = \frac{1}{\sqrt{2\pi\sigma^2 y^3}} \exp\left[-\frac{1}{2\mu^2\sigma^2 y} (y-\mu)^2\right] \tag{2.3}$$

for $y > 0$, where $\mu > 0$ and $\sigma > 0$. Hence $\mathrm{E}(Y) = \mu$ and $\mathrm{Var}(Y) = \sigma^2\mu^3$. This is a reparameterization of Johnson et al. [1994, p261 equation (15.4a)], obtained by setting $\sigma^2 = 1/\lambda$. Note that the inverse Gaussian distribution is a reparameterized special case of the generalized inverse Gaussian distribution, given by IG$(\mu, \sigma) =$ GIG$(\mu, \sigma\mu^{1/2}, -0.5)$.

The inverse Gaussian distribution is appropriate for highly positively skewed data.

Note also that if $Y \sim$ IG(μ, σ) then $Y_1 = aY \sim$ IG$(a\mu, a^{-1/2}\sigma)$. Hence IG(μ, σ) is a scale family of distributions, but neither μ nor σ is a scaling parameter. The shape of the IG(μ, σ) distribution depends only on the value of $\sigma^2\mu$, so if $\sigma^2\mu$ is fixed, then changing μ changes the scale but not the shape of the IG(μ, σ) distribution. If σ in the distribution IG(μ, σ) is reparameterized by setting $\alpha = \sigma\mu^{1/2}$, then for fixed α the resulting IG2(μ, α) distribution is a scale family of distributions with scaling parameter μ, since $Z = Y/\mu \sim$ IG2$(1, \alpha)$. Note IG2 is not currently available in **gamlss.dist**.

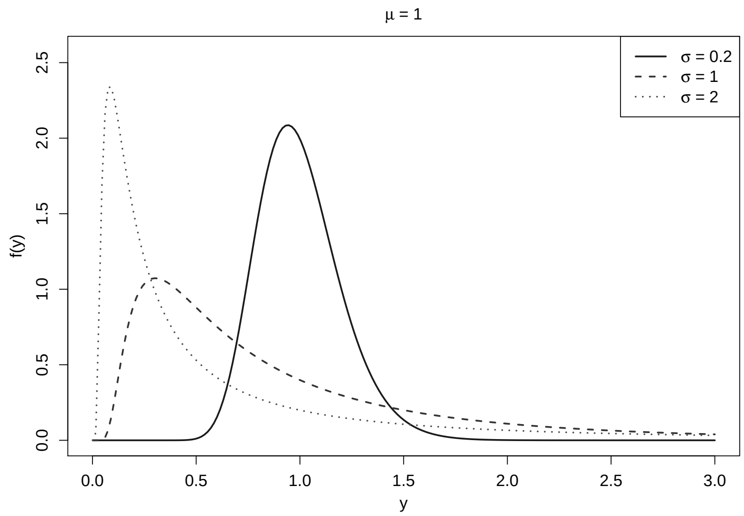

Figure 2.3: The inverse Gaussian, $\texttt{IG}(\mu, \sigma)$, distribution, with $\mu = 1$ and $\sigma = 0.2, 1, 2$.

Table 2.4: Inverse Gaussian distribution.

	$\texttt{IG}(\mu,\sigma)$
	Ranges
Y	$0 < y < \infty$
μ	$0 < \mu < \infty$, mean, *not* a scaling parameter
σ	$0 < \sigma < \infty$
	Distribution measures
mean [a]	μ
mode [a]	$\dfrac{-3\mu^2\sigma^2 + \mu(9\mu^2\sigma^4+4)^{1/2}}{2}$
variance [a]	$\sigma^2\mu^3$
skewness [a]	$3\mu^{1/2}\sigma$
excess kurtosis[a]	$15\mu\sigma^2$
MGF [a]	$\exp\left\{\dfrac{1}{\mu\sigma^2}\left[1-\left(1-2\mu^2\sigma^2 t\right)^{1/2}\right]\right\}$ for $t < (2\mu^2\sigma^2)^{-1}$
pdf [a]	$(2\pi\sigma^2 y^3)^{-1/2}\exp\left[-\dfrac{1}{2\mu^2\sigma^2 y}(y-\mu)^2\right]$
cdf [a]	$\Phi\left[(\sigma^2 y)^{-1/2}\left(\dfrac{y}{\mu}-1\right)\right] + e^{2(\mu\sigma^2)^{-1}}\Phi\left[-(\sigma^2 y)^{-1/2}\left(\dfrac{y}{\mu}+1\right)\right]$
Reference	[a] Johnson et al. [1994] Chapter 15, p261-263 and p268 with equation (15.4a) reparameterized by $\mu=\mu$ and $\lambda = 1/\sigma^2$, or set $\mu_1=\mu, \sigma_1 = \sigma\mu^{1/2}$ and $\nu_1 = -1/2$, (and so $b=1$), in $\texttt{GIG}(\mu_1,\sigma_1,\nu_1)$.

2.3.4 Log normal: LOGNO, LOGNO2

First parameterization, LOGNO

The pdf of the log normal distribution, denoted by $\texttt{LOGNO}(\mu,\sigma)$, is given by

$$f_Y(y\,|\,\mu,\sigma) = \frac{1}{\sqrt{2\pi\sigma^2}}\,\frac{1}{y}\exp\left[-\frac{(\log y-\mu)^2}{2\sigma^2}\right] \tag{2.4}$$

for $y>0$, where $-\infty<\mu<\infty$ and $\sigma>0$. Note that $\log Y \sim \texttt{NO}(\mu,\sigma)$. Note also that if $Y\sim\texttt{LOGNO}(\mu,\sigma)$ then $Y_1 = aY \sim \texttt{LOGNO}(\mu+\log a,\sigma)$ and so $Z = Y/e^{\mu} \sim \texttt{LOGNO}(0,\sigma)$. So, for fixed σ, $\texttt{LOGNO}(\mu,\sigma)$ is a scale family of distributions with scaling parameter e^{μ}. Hence μ itself is not a scaling parameter.

The log normal distribution is appropriate for positively skewed data. Note that, for both $\texttt{LOGNO}(\mu,\sigma)$ and $\texttt{LOGNO2}(\mu,\sigma)$, σ is a true variance parameter and a primary true moment skewness parameter, and also a primary true, both Balanda-MacGillivray and moment, kurtosis parameter, see Sections **??** and **??**.

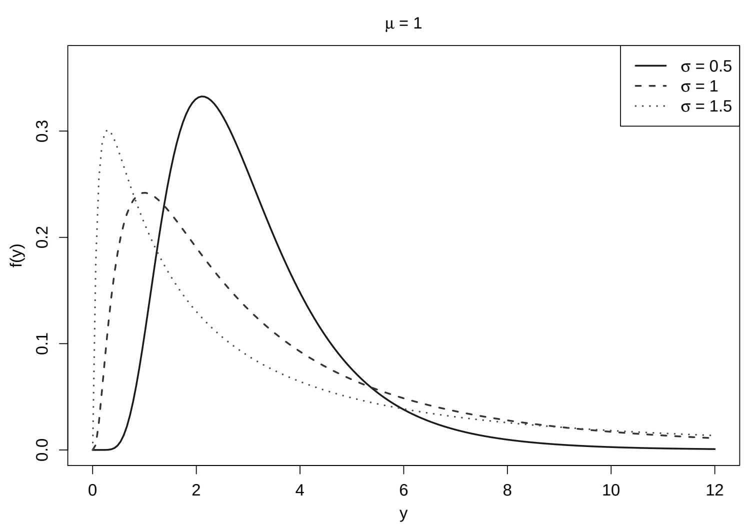

Figure 2.4: The log normal, LOGNO(μ, σ), distribution, with $\mu = 1$ and $\sigma = 0.5, 1, 1.5$.

Table 2.5: Log normal distribution.

LOGNO(μ,σ)	
Ranges	
Y	$0 < y < \infty$
μ	$-\infty < \mu < \infty$, *not* a scaling parameter
σ	$0 < \sigma < \infty$
Distribution measures	
mean [a]	$e^{\mu+\sigma^2/2}$
median [a]	e^{μ}
mode [a]	$e^{\mu-\sigma^2}$
variance [a]	$e^{2\mu+\sigma^2}(e^{\sigma^2}-1)$
skewness [a]	$(e^{\sigma^2}-1)^{0.5}(e^{\sigma^2}+2)$
excess kurtosis [a]	$e^{4\sigma^2}+2e^{3\sigma^2}+3e^{2\sigma^2}-6$
pdf [a2]	$\frac{1}{\sqrt{2\pi\sigma^2}}\,\frac{1}{y}\exp\left[-\frac{(\log\, y-\mu)^2}{2\sigma^2}\right]$
cdf	$\Phi\left(\frac{\log\, y-\mu}{\sigma}\right)$
inverse cdf	$e^{\mu+\sigma z_p}$ where $z_p = \Phi^{-1}(p)$
Reference	[a] Johnson et al. [1994] Chapter 14, p208-213 [a2] Johnson et al. [1994] Chapter 14, p208, equation (14.2) where $\xi=\mu$, $\sigma=\sigma$, $\theta=0$.

Second parameterization, LOGNO2

The pdf of the second parameterization of the log normal distribution, denoted by LOGNO2(μ,σ), is given by

$$f_Y(y\,|\,\mu,\sigma) = \frac{1}{\sqrt{2\pi\sigma^2}}\,\frac{1}{y}\exp\left[-\frac{(\log\, y-\log\,\mu)^2}{2\sigma^2}\right] \tag{2.5}$$

for $y>0$, where $\mu>0$ and $\sigma>0$. Note that $\log Y \sim$ NO$(\log\mu,\sigma)$. In this parameterization μ is the median of Y and μ is a scaling parameter, so if $Y \sim$ LOGNO2(μ,σ) then $Z = Y/\mu \sim$ LOGNO2$(0,\sigma)$ which does not depend on μ.

Table 2.6: Second parameterization of the log normal distribution.

$\texttt{LOGNO2}(\mu, \sigma)$	
Ranges	
Y	$0 < y < \infty$
μ	$0 < \mu < \infty$, median, scaling parameter
σ	$0 < \sigma < \infty$
Distribution measures	
mean	$\mu e^{\sigma^2/2}$
median	μ
mode	$\mu e^{-\sigma^2}$
variance	$\mu^2 e^{\sigma^2}(e^{\sigma^2} - 1)$
skewness	$(e^{\sigma^2} - 1)^{0.5}(e^{\sigma^2} + 2)$
excess kurtosis [a]	$e^{4\sigma^2} + 2e^{3\sigma^2} + 3e^{2\sigma^2} - 6$
pdf	$\frac{1}{\sqrt{2\pi\sigma^2}} \frac{1}{y} \exp\left[-\frac{(\log\, y - \log\, \mu)^2}{2\sigma^2}\right]$
cdf	$\Phi\left(\frac{\log\, y - \log\, \mu}{\sigma}\right)$
inverse cdf	$\mu e^{\sigma z_p}$ where $z_p = \Phi^{-1}(p)$
Reference	Set $\mu_1 = \log \mu$ in $\texttt{LOGNO}(\mu_1, \sigma)$

2.3.5 Pareto type 1: `PARETO1o`

The pdf of the Pareto type 1 original distribution, denoted by $\texttt{PARETO1o}(\mu, \sigma)$, is given by

$$f_Y(y \,|\, \mu, \sigma) = \frac{\sigma \mu^\sigma}{y^{\sigma+1}} \tag{2.6}$$

for $y > \mu$, where $\mu > 0$ and $\sigma > 0$. This is also called the Pareto distribution of the first kind [Johnson et al., 1994, p574]. It is especially important in the discussion of heavy tails, because its survivor function is $S_Y(y) = \mu^\sigma y^{-\sigma}$, for $y > \mu$, i.e. exactly proportional to $y^{-\sigma}$ (see Chapter **??**). Note that σ is a true variance parameter (for $\sigma > 2$), a primary true van Zwet skewness parameter and also (for $\sigma > 4$) a primary true moment kurtosis parameter, see Sections **??** and **??**.

Table 2.7: Pareto type 1 original distribution.

PARETO1o(μ,σ)	
Ranges	
Y	$\mu < y < \infty$
μ	$0 < \mu < \infty,$ scaling parameter (fixed constant)
σ	$0 < \sigma < \infty$
Distribution measures	
mean	$\begin{cases} \dfrac{\mu\sigma}{(\sigma-1)} & \text{if } \sigma > 1 \\ \infty & \text{if } \sigma \le 1 \end{cases}$
median	$\mu 2^{1/\sigma}$
mode	$\to 0$
variance [a]	$\begin{cases} \dfrac{\sigma\mu^2}{(\sigma-1)^2(\sigma-2)} & \text{if } \sigma > 2 \\ \infty & \text{if } \sigma \le 2 \end{cases}$
skewness [a]	$\begin{cases} \dfrac{2(\sigma+1)(\sigma-2)^{1/2}}{(\sigma-3)\sigma^{1/2}} & \text{if } \sigma > 3 \\ \infty & \text{if } \sigma \le 3 \end{cases}$
excess kurtosis[a]	$\begin{cases} \dfrac{3(\sigma-2)(3\sigma^2+\sigma+2)}{\sigma(\sigma-3)(\sigma-4)} - 3 & \text{if } \sigma > 4 \\ \infty & \text{if } \sigma \le 4 \end{cases}$
pdf [a2]	$\dfrac{\sigma\mu^\sigma}{y^{\sigma+1}}$
cdf [a]	$1 - \left(\dfrac{\mu}{y}\right)^\sigma$
inverse cdf	$\mu(1-p)^{-1/\sigma}$
Reference	[a]Johnson et al. [1994] Sections 20.3, 20.4 p574-579. [a2] Johnson et al. [1994], equation (20.3), p574 with $k = \mu$, $a = \sigma$.

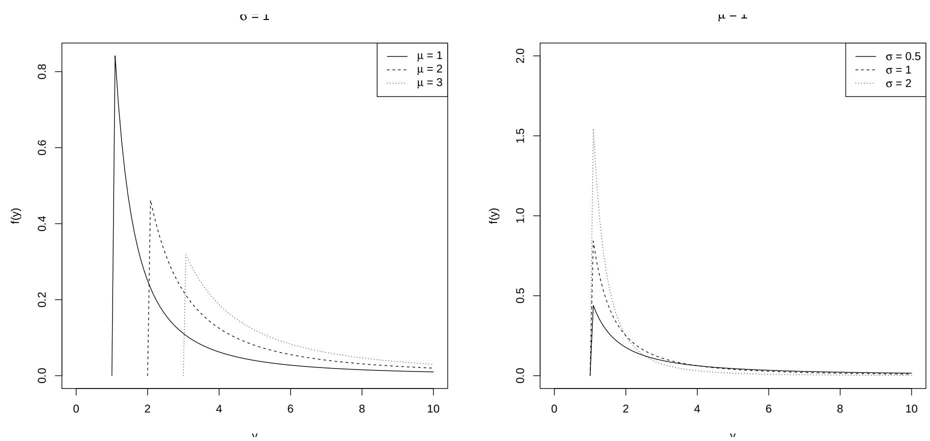

Figure 2.5: The Pareto, PARETO1o(μ,σ), distribution, (left) $\sigma = 1$ and $\mu = 1,2,3$, and (right) $\mu = 1$ and $\sigma = 0.5, 1, 2$.

2.3.6 Pareto type 2: PARETO2o, PARETO2

First parameterization, PARETO2o(μ,σ)

The pdf of the Pareto type 2 original distribution, denoted by PARETO2o(μ,σ), is

$$f_Y(y\,|\,\mu,\sigma) = \frac{\sigma\mu^{\sigma}}{(y+\mu)^{\sigma+1}} \tag{2.7}$$

for $y > 0$, where $\mu > 0$ and $\sigma > 0$. This was called the Pareto distribution of the second kind or Lomax distribution by Johnson et al. [1994, p575], where $C = \mu$ and $a = \sigma$. Note that σ is a true variance parameter (for $\sigma < 2$), a primary true van Zwet skewness parameter, and also (for $\sigma > 4$) a primary true moment kurtosis parameter, see Sections **??** and **??**. Note that if $Y \sim$ PARETO2o(μ,σ) then $Z = Y + \mu \sim$ PARETO1o(μ,σ).

Table 2.8: Pareto type 2 original distribution.

PARETO2o(μ, σ)	
Ranges	
Y	$0 < y < \infty$
μ	$0 < \mu < \infty$, scaling parameter
σ	$0 < \sigma < \infty$
Distribution measures	
mean	$\begin{cases} \dfrac{\mu}{(\sigma-1)} & \text{if } \sigma > 1 \\ \infty & \text{if } \sigma \le 1 \end{cases}$
median	$\mu(2^{1/\sigma} - 1)$
mode	$\to 0$
variance [a]	$\begin{cases} \dfrac{\sigma\mu^2}{(\sigma-1)^2(\sigma-2)} & \text{if } \sigma > 2 \\ \infty & \text{if } \sigma \le 2 \end{cases}$
skewness [a]	$\begin{cases} \dfrac{2(\sigma+1)(\sigma-2)^{1/2}}{(\sigma-3)\sigma^{1/2}} & \text{if } \sigma > 3 \\ \infty & \text{if } \sigma \le 3 \end{cases}$
excess kurtosis[a]	$\begin{cases} \dfrac{3(\sigma-2)(3\sigma^2+\sigma+2)}{\sigma(\sigma-3)(\sigma-4)} - 3 & \text{if } \sigma > 4 \\ \infty & \text{if } \sigma \le 4 \end{cases}$
pdf [a2]	$\dfrac{\sigma\mu^\sigma}{(y+\mu)^{\sigma+1}}$
cdf [a]	$1 - \dfrac{\mu^\sigma}{(y+\mu)^\sigma}$
inverse cdf	$\mu[(1-p)^{-1/\sigma} - 1]$
Reference	[a]Johnson et al. [1994] Sections 20.3, 20.4 p574-579. [a2] Johnson et al. [1994, p574] equation (20.3), p574 with $k = \mu$, $a = \sigma$ and $x = y + \mu$, so $X = Y + \mu$.

Second parameterization, PARETO2(μ, σ)

The pdf of the Pareto type 2 distribution, denoted by PARETO2(μ, σ), is given by

$$f_Y(y \mid \mu, \sigma) = \frac{\sigma^{-1}\mu^{1/\sigma}}{(y+\mu)^{(1/\sigma)+1}} \tag{2.8}$$

for $y > 0$, where $\mu > 0$ and $\sigma > 0$. PARETO2 is given by reparameterizing σ to $1/\sigma$ in PARETO2o, i.e. PARETO2(μ, σ) = PARETO2o$(\mu, 1/\sigma)$. Note that σ is a true variance parameter (for $\sigma < 1/2$), a primary true van Zwet skewness parameter, and (for $\sigma < 1/4$) a primary true moment kurtosis parameter, see Sections **??** and **??**.

Table 2.9: Pareto type 2 distribution.

PARETO2(μ, σ)	
Ranges	
Y	$0 < y < \infty$
μ	$0 < \mu < \infty$ scaling parameter
σ	$0 < \sigma < \infty$
Distribution measures	
mean	$\begin{cases} \dfrac{\mu\sigma}{(1-\sigma)} & \text{if } \sigma < 1 \\ \infty & \text{if } \sigma \geq 1 \end{cases}$
median	$\mu(2^{\sigma} - 1)$
mode	$\to 0$
variance	$\begin{cases} \dfrac{\sigma^2\mu^2}{(1-\sigma)^2(1-2\sigma)} & \text{if } \sigma < 1/2 \\ \infty & \text{if } \sigma \geq 1/2 \end{cases}$
skewness	$\begin{cases} \dfrac{2(1+\sigma)(1-2\sigma)^{1/2}}{(1-3\sigma)} & \text{if } \sigma < 1/3 \\ \infty & \text{if } \sigma \geq 1/3 \end{cases}$
excess kurtosis	$\begin{cases} \dfrac{3(1-2\sigma)(2\sigma^2+\sigma+3)}{(1-3\sigma)(1-4\sigma)} - 3 & \text{if } \sigma < 1/4 \\ \infty & \text{if } \sigma \geq 1/4 \end{cases}$
pdf	$\dfrac{\sigma^{-1}\mu^{1/\sigma}}{(y+\mu)^{(1/\sigma)+1}}$
cdf	$1 - \dfrac{\mu^{1/\sigma}}{(y+\mu)^{1/\sigma}}$
inverse cdf	$\mu[(1-p)^{-\sigma} - 1]$
Reference	Set σ to $1/\sigma$ in PARETO2o(μ, σ).

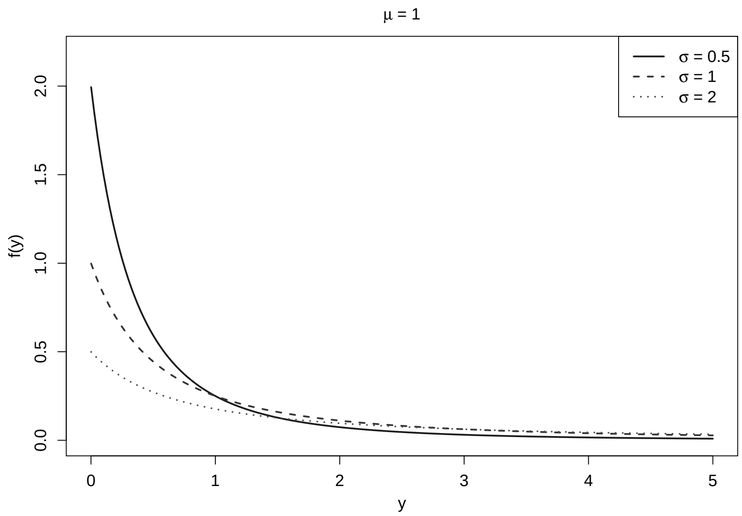

Figure 2.6: The Pareto, PARETO2(μ, σ), distribution, with $\mu = 1$ and $\sigma = 0.5, 1, 2$.

2.3.7 Weibull: WEI, WEI2, WEI3

First parameterization, WEI

There are three versions of the two-parameter Weibull distribution implemented in **gamlss**. The first, denoted by WEI(μ, σ), has the following parameterization

$$f_Y(y \,|\, \mu, \sigma) = \frac{\sigma y^{\sigma-1}}{\mu^\sigma} \, \exp\left[-\left(\frac{y}{\mu}\right)^\sigma\right] \tag{2.9}$$

for $y > 0$, where $\mu > 0$ and $\sigma > 0$, see Johnson et al. [1994, p629]. Note that σ is a primary true van Zwet skewness parameter, see Section **??**. The moment skewness is positive for $\sigma \leq 3.60$ and negative for $\sigma > 3.60$. Note that the exponential distribution, EXP(μ), is a special case of WEI(μ, σ) when $\sigma = 1$.

Table 2.10: Weibull distribution.

	$\texttt{WEI}(\mu,\sigma)$
	Ranges
Y	$0 < y < \infty$
μ	$0 < \mu < \infty$, scaling parameter
σ	$0 < \sigma < \infty$
	Distribution measures
mean	$\mu\Gamma(\sigma^{-1}+1)$
median	$\mu(\log 2)^{1/\sigma}$
mode	$\begin{cases} \mu(1-\sigma^{-1})^{1/\sigma} & \text{if } \sigma > 1 \\ \to 0 & \text{if } \sigma \le 1 \end{cases}$
variance	$\mu^2\{\Gamma(2\sigma^{-1}+1) - [\Gamma(\sigma^{-1}+1)]^2\}$
skewness	$\begin{cases} \mu_3/[\text{Var}(Y)]^{1.5} \quad \text{where} \\ \quad \mu_3 = \mu^3\{\Gamma(3\sigma^{-1}+1) - 3\Gamma(2\sigma^{-1}+1)\Gamma(\sigma^{-1}+1) \\ \qquad +2[\Gamma(\sigma^{-1}+1)]^3\} \end{cases}$
excess kurtosis	$\begin{cases} \mu_4/[\text{Var}(Y)]^2 - 3 \quad \text{where} \\ \quad \mu_4 = \mu^4\{\Gamma(4\sigma^{-1}+1) - 4\Gamma(3\sigma^{-1}+1)\Gamma(\sigma^{-1}+1) \\ \qquad +6\Gamma(2\sigma^{-1}+1)[\Gamma(\sigma^{-1}+1)]^2 - 3[\Gamma(\sigma^{-1}+1)]^4\} \end{cases}$
pdf [a]	$\dfrac{\sigma y^{\sigma-1}}{\mu^\sigma}\exp\left[-(y/\mu)^\sigma\right]$
cdf	$1-\exp\left[-(y/\mu)^\sigma\right]$
inverse cdf	$\mu[-\log(1-p)]^{1/\sigma}$
Reference	Johnson et al. [1994] Chapter 21, p628-632. [a] Johnson et al. [1994] equation (21.3), p629, with $\alpha=\mu, c=\sigma$ and $\xi_0=0$.

Second parameterization, WEI2

The second parameterization of the Weibull distribution, denoted by $\texttt{WEI2}(\mu,\sigma)$, is defined as

$$f_Y(y \mid \mu,\sigma) = \sigma\mu y^{\sigma-1}\exp\left(-\mu y^\sigma\right) \tag{2.10}$$

for $y > 0$, where $\mu > 0$ and $\sigma > 0$. The parameterization (2.10) gives the usual proportional hazards Weibull model. Note $\texttt{WEI2}(\mu,\sigma) = \texttt{WEI}(\mu^{-1/\sigma},\sigma)$, and so $\texttt{WEI}(\mu,\sigma) = \texttt{WEI2}(\mu^{-\sigma},\sigma)$. In this second parameterization, $\texttt{WEI2}(\mu,\sigma)$, of the Weibull distribution, the two parameters μ and σ are highly correlated. As a result the `RS` method of fitting is very slow and therefore the `CG()` or `mixed()` method of fitting should be used.

Table 2.11: Second parameterization of Weibull distribution.

	`WEI2`(μ, σ)
	Ranges
Y	$0 < y < \infty$
μ	$0 < \mu < \infty$, *not* a scaling parameter
σ	$0 < \sigma < \infty$
	Distribution measures
mean	$\mu^{-1/\sigma}\Gamma(\sigma^{-1}+1)$
median	$\mu^{-1/\sigma}(\log 2)^{1/\sigma}$
mode	$\begin{cases} \mu^{-1/\sigma}(1-\sigma^{-1})^{1/\sigma} & \text{if } \sigma > 1 \\ \to 0 & \text{if } \sigma \le 1 \end{cases}$
variance	$\mu^{-2/\sigma}\{\Gamma(2\sigma^{-1}+1) - [\Gamma(\sigma^{-1}+1)]^2\}$
skewness	$\begin{cases} \mu_3/[\mathrm{Var}(Y)]^{1.5} \quad \text{where} \\ \mu_3 = \mu^{-3/\sigma}\{\Gamma(3\sigma^{-1}+1) - 3\Gamma(2\sigma^{-1}+1)\Gamma(\sigma^{-1}+1) \\ \quad +2[\Gamma(\sigma^{-1}+1)]^3\} \end{cases}$
excess kurtosis	$\begin{cases} \mu_4/[\mathrm{Var}(Y)]^2 - 3 \quad \text{where} \\ \mu_4 = \mu^{-4/\sigma}\{\Gamma(4\sigma^{-1}+1) - 4\Gamma(3\sigma^{-1}+1)\Gamma(\sigma^{-1}+1) \\ \quad +6\Gamma(2\sigma^{-1}+1)[\Gamma(\sigma^{-1}+1)]^2 - 3[\Gamma(\sigma^{-1}+1)]^4\} \end{cases}$
pdf	$\mu\sigma y^{\sigma-1}\exp(-\mu y^{\sigma})$
cdf	$1 - \exp(-\mu y^{\sigma})$
inverse cdf	$\mu^{-1/\sigma}[-\log(1-p)]^{1/\sigma}$
Note	Set $\mu_1 = \mu^{-1/\sigma}$ in `WEI`(μ_1, σ).

Third parameterization, `WEI3`

This is a parameterization of the Weibull distribution where μ is the mean of the distribution. This parameterization of the Weibull distribution, denoted by `WEI3`(μ, σ), is defined as

$$f_Y(y \mid \mu, \sigma) = \frac{\sigma y^{\sigma-1}}{\beta^{\sigma}} \exp\left[-\left(\frac{y}{\beta}\right)^{\sigma}\right] \tag{2.11}$$

for $y > 0$, where $\mu > 0$ and $\sigma > 0$ and $\beta = \mu[\,\Gamma(\sigma^{-1}+1)\,]^{-1}$. The parameterization (2.11) gives the usual accelerated lifetime Weibull model. Note that

$\texttt{WEI3}(\mu, \sigma) = \texttt{WEI}(\beta, \sigma)$. Note that the exponential distribution, $\texttt{EXP}(\mu)$, is a special case of $\texttt{WEI}(\mu, \sigma)$ when $\sigma = 1$.

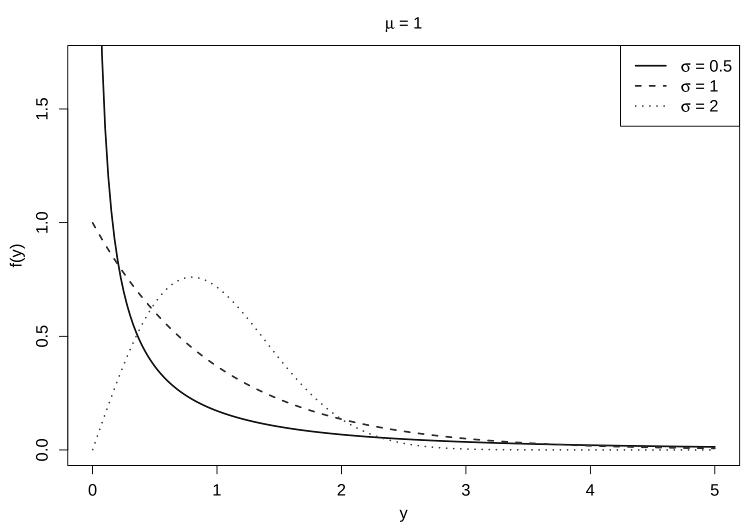

Figure 2.7: The third parameterization of Weibull, $\texttt{WEI3}(\mu, \sigma)$, distribution, with $\mu = 1$ and $\sigma = 0.5, 1, 2$.

Table 2.12: Third parameterization of Weibull distribution.

	`WEI3`(μ,σ)
	Ranges
Y	$0 < y < \infty$
μ	$0 < \mu < \infty$, mean, scaling parameter
σ	$0 < \sigma < \infty$
	Distribution measures
mean	μ
median	$\beta(\log 2)^{1/\sigma}$ where $\beta = \mu[\,\Gamma(\sigma^{-1}+1)\,]^{-1}$
mode	$\begin{cases} \beta(1-\sigma^{-1})^{1/\sigma} & \text{if } \sigma > 1 \\ \to 0, & \text{if } \sigma \le 1 \end{cases}$
variance	$\beta^2\{\Gamma(2\sigma^{-1}+1) - [\Gamma(\sigma^{-1}+1)]^2\}$
skewness	$\begin{cases} \mu_3/[\mathrm{Var}(Y)]^{1.5} \quad \text{where} \\ \mu_3 = \beta^3\{\Gamma(3\sigma^{-1}+1) - 3\Gamma(2\sigma^{-1}+1)\Gamma(\sigma^{-1}+1) \\ \quad +2[\Gamma(\sigma^{-1}+1)]^3\} \end{cases}$
excess kurtosis	$\begin{cases} \mu_4/[\mathrm{Var}(Y)]^2 - 3 \quad \text{where} \\ \mu_4 = \beta^4\{\Gamma(4\sigma^{-1}+1) - 4\Gamma(3\sigma^{-1}+1)\Gamma(\sigma^{-1}+1) \\ \quad +6\Gamma(2\sigma^{-1}+1)[\Gamma(\sigma^{-1}+1)]^2 - 3[\Gamma(\sigma^{-1}+1)]^4\} \end{cases}$
pdf	$\dfrac{\sigma y^{\sigma-1}}{\beta^\sigma}\exp\left[-\left(y/\beta\right)^\sigma\right]$
cdf	$1-\exp\left[-\left(y/\beta\right)^\sigma\right]$
inverse cdf	$\beta[\,-\log(1-p)\,]^{1/\sigma}$
Reference	Set μ_1 to β in `WEI`(μ_1,σ), where $\beta = \mu[\Gamma(\sigma^{-1}+1)]^{-1}$.

2.4 Continuous three-parameter distributions on $(0,\infty)$

2.4.1 Box-Cox Cole and Green: `BCCG`, `BCCGo`

The Box-Cox Cole and Green distribution is suitable for positively or negatively skewed data. Let $Y > 0$ be a positive random variable having a Box-Cox Cole and Green distribution, denoted by `BCCG`(μ,σ,ν), defined through the transformed random variable Z given by

$$Z = \begin{cases} \dfrac{1}{\sigma\nu}\left[\left(\dfrac{Y}{\mu}\right)^{\nu} - 1\right] & \text{if} \quad \nu \neq 0 \\[2ex] \dfrac{1}{\sigma}\log\left(\dfrac{Y}{\mu}\right) & \text{if} \quad \nu = 0 \end{cases} \tag{2.12}$$

for $Y > 0$, where $\mu > 0$, $\sigma > 0$ and $-\infty < \nu < \infty$, and where the random variable Z is assumed to follow a truncated standard normal distribution. The condition $Y > 0$ (required for Y^ν to be real for all ν) leads to the conditions

$$\begin{aligned} -\frac{1}{\sigma\nu} < Z < \infty & \qquad \text{if } \nu > 0 \\ -\infty < Z < -\frac{1}{\sigma\nu} & \qquad \text{if } \nu < 0 \ , \end{aligned}$$

which necessitates the truncated standard normal distribution for Z. Note that

$$Y = \begin{cases} \mu(1+\sigma\nu Z)^{1/\nu} & \text{if } \nu \neq 0 \\ \mu\exp(\sigma Z) & \text{if } \nu = 0 \ . \end{cases} \tag{2.13}$$

[See Figure 2.8 for a plot of the relationship between Y and Z for $\mu = 1$, $\sigma = 0.2$ and $\nu = -2$. For this case $-\infty < Z < 2.5$.]

The pdf of $Y \sim$ `BCCG`(μ,σ,ν) is given by

$$f_Y(y) = \frac{y^{\nu-1}\exp\left(-\frac{1}{2}z^2\right)}{\mu^\nu\sigma\sqrt{2\pi}\,\Phi\left[(\sigma|\nu|)^{-1}\right]} \tag{2.14}$$

for $y > 0$, where $\mu > 0$, $\sigma > 0$ and $-\infty < \nu < \infty$, and where z is given by (2.12). [The distribution `BCCGo`(μ,σ,ν) also has pdf given by (2.14). It differs from `BCCG`(μ,σ,ν) only in having a default log link function for μ, instead of the default identity link function for μ in `BCCG`(μ,σ,ν), in **gamlss.dist**]. Note that `LOGNO2`(μ,σ) is the special case `BCCG`$(\mu,\sigma,0)$.

The exact cdf of $Y \sim$ `BCCG`(μ,σ,ν) is given by

$$F_Y(y) = \begin{cases} \dfrac{\Phi(z)}{\Phi\left[(\sigma|\nu|)^{-1}\right]} & \text{if } \nu \leq 0 \\[2ex] \dfrac{\Phi(z) - \Phi\left[-(\sigma|\nu|)^{-1}\right]}{\Phi\left[(\sigma|\nu|)^{-1}\right]} & \text{if } \nu > 0 \end{cases}$$

where z is given by (2.12). The term $\Phi\left(-(\sigma|\nu|)^{-1}\right)$ is the truncation probability. If this is negligible, and hence $\Phi\left((\sigma|\nu|)^{-1}\right) \approx 1$, then $F_Y(y) \approx \Phi(z)$. Also if the truncation probability is negligible, Y has median μ. Note if $\nu = 0$ there is no truncation and the truncation probability is zero.

The exact inverse cdf (or quantile) y_p of $Y \sim$ `BCCG`(μ,σ,ν), defined by $P(Y \leq y_p) = p$, is given by

$$y_p = \begin{cases} \mu\left(1+\sigma\nu z_T\right)^{1/\nu}, & \text{if } \nu \neq 0 \\ \mu\exp(\sigma z_T), & \text{if } \nu = 0 \end{cases}$$

where

$$z_T = \begin{cases} \Phi^{-1}\left\{p\Phi\left[(\sigma|\nu|)^{-1}\right]\right\} & \text{if } \nu \leq 0 \\ \Phi^{-1}\left\{p\Phi\left[(\sigma|\nu|)^{-1}\right] + \Phi\left[-(\sigma|\nu|)^{-1}\right]\right\} & \text{if } \nu > 0 \end{cases}$$

Hence for $\nu \neq 0$, $z_T \approx \Phi^{-1}(p) = z_p$ provided the truncation probability is negligible. The parameterization in (2.12) was originally used by Cole and Green [1992], who assumed a standard normal distribution for Z but also assumed a negligible truncation probability. The BCCG(μ, σ, ν) is a special case of the BCPE(μ, σ, ν, τ), given by setting $\tau = 2$, and hence the results in this section are obtained from Rigby and Stasinopoulos [2004], giving a proper distribution (i.e. not requiring the assumption of a negligible truncation probability).

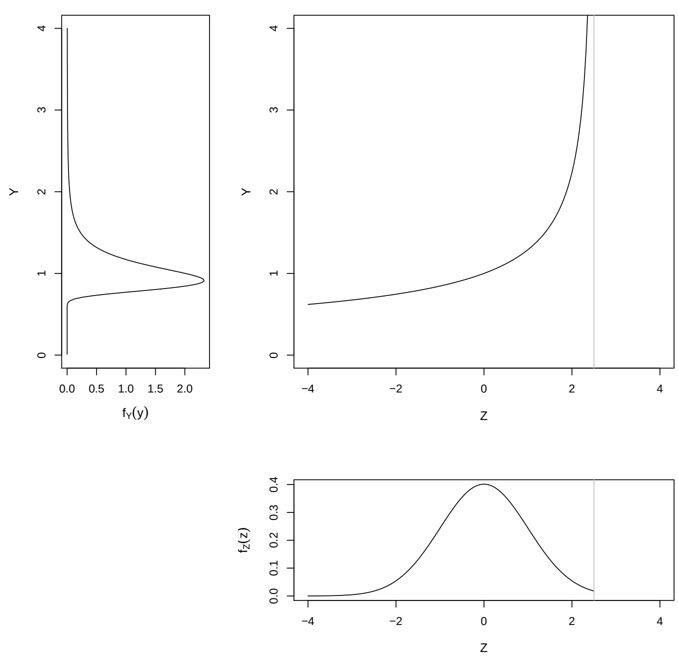

Figure 2.8: Relationship between Y and Z for the BCCG$(1, 0.2, -2)$ distribution (top right), $f_Z(z)$ (bottom), and $f_Y(y)$ (top left).

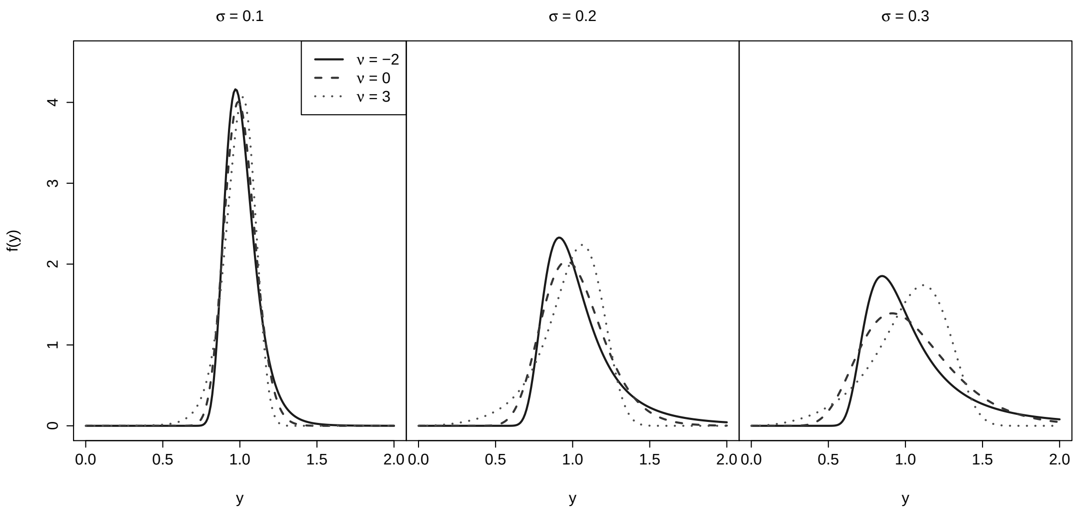

Figure 2.9: The Box-Cox Cole and Green, BCCG(μ, σ, ν), distribution, with $\mu = 1$, $\sigma = 0.1, 0.2, 0.3$ and $\nu = -2, 0, 3$.

Table 2.13: Box-Cox Cole and Green distribution.

	$\texttt{BCCG}(\mu,\sigma,\nu)$
	Ranges
Y	$0<y<\infty$
μ	$0<\mu<\infty,$ median[a], scaling parameter
σ	$0<\sigma<\infty,$ approximate coefficient of variation[a,a2]
ν	$-\infty<\nu<\infty,$ skewness parameter
	Distribution measures
median [a]	μ
mode [a3]	$\begin{cases} \mu\omega^{1/\nu} & \text{if } \nu\neq 0 \\ \mu e^{-\sigma^2} & \text{if } \nu=0 \\ \text{where } \omega=\{1+[1+4\sigma^2\nu(\nu-1)]^{1/2}\}/2 \end{cases}$
pdf	$\dfrac{y^{\nu-1}\exp\left(-\frac{1}{2}z^2\right)}{\mu^\nu\sigma\sqrt{2\pi}\,\Phi\left((\sigma\lvert\nu\rvert)^{-1}\right)}$
cdf [a]	$\Phi(z)$ where z is given by (2.12)
inverse cdf[a]	$\begin{cases} \mu(1+\sigma\nu z_p)^{1/\nu} & \text{if } \nu\neq 0 \\ \mu\exp(\sigma z_p) & \text{if } \nu=0 \\ \text{where } z_p=\Phi^{-1}(p) \end{cases}$
Reference	Set $\tau=2$ in $\texttt{BCPE}(\mu,\sigma,\nu,\tau)$.
Note	[a] Provided $\Phi\left(-(\sigma\lvert\nu\rvert)^{-1}\right)$ is negligible, e.g. <0.0001 for $\sigma\lvert\nu\rvert<0.27$. [a2] Provided ν is reasonably close to 1 ($\nu\geq 0$ and not too large) and σ is small, say $\sigma\leq 0.3$. [a3] If $0<\nu<1$ there is a second mode $\to 0$.

2.4.2 Gamma family: GAF

$\texttt{GAF}(\mu,\sigma,\nu)$ defines a gamma distribution family with three parameters. The third parameter ν allows the variance of the distribution to be proportional to a power of the mean. The mean of $\texttt{GAF}(\mu,\sigma,\nu)$ is equal to μ, while the variance is $\text{Var}(Y)=\sigma^2\mu^\nu$.

The pdf of the gamma family distribution, $\texttt{GAF}(\mu,\sigma,\nu)$, is given by

$$f_Y(y\,|\,\mu,\sigma,\nu)=\frac{y^{\sigma_1^{-2}-1}\exp[-y/(\sigma_1^2\mu)]}{(\sigma_1^2\mu)^{\sigma_1^{-2}}\Gamma(\sigma_1^{-2})} \tag{2.15}$$

where $y>0$, $\sigma_1=\sigma\mu^{(\nu/2)-1}$, $\mu>0$, $\sigma>0$ and $-\infty<\nu<\infty$.

Note that if $Y\sim\texttt{GAF}(\mu,\sigma,\nu)$, then $Y\sim\texttt{GA}(\mu,\sigma_1)=\texttt{GA}(\mu,\sigma\mu^{(\nu/2)-1})$. Hence $\texttt{GAF}(\mu,\sigma,\nu)$ is appropriate for a gamma distributed response variable where the variance of the response variable is proportional to a power of the mean. The parameter ν is usually modeled as a constant, used as a device to model the variance-mean relationship. Note that, due to the high correlation between the σ and ν parameters, the `method=mixed()` and `c.crit=0.0001` arguments are strongly recommended to speed up the convergence, and avoid converging too early.

Table 2.14: Gamma family (of variance-mean relationships) distribution.

	$\texttt{GAF}(\mu, \sigma, \nu)$
	Ranges
Y	$0 < y < \infty$
μ	$0 < \mu < \infty$, mean
σ	$0 < \sigma < \infty$
ν	$-\infty < \nu < \infty$
	Distribution measures
mean	μ
mode	$\begin{cases} \mu(1-\sigma^2\mu^{\nu-2}) & \text{if } \sigma^2\mu^{\nu-2} < 1 \\ \to 0 & \text{if } \sigma^2\mu^{\nu-2} \geq 1 \end{cases}$
variance	$\sigma^2\mu^\nu$
skewness	$2\sigma\mu^{(\nu/2)-1}$
excess kurtosis	$6\sigma^2\mu^{\nu-2}$
MGF	$(1-\sigma^2\mu^{\nu-1}t)^{-1/(\sigma^2\mu^{\nu-2})}$, for $t < (\sigma^2\mu^{\nu-1})^{-1}$
pdf	$\dfrac{y^{\sigma_1^{-2}-1}\exp[-y/(\sigma_1^2\mu)]}{(\sigma_1^2\mu)^{\sigma_1^{-2}}\Gamma(\sigma_1^{-2})}$, where $\sigma_1 = \sigma\mu^{(\nu/2)-1}$
cdf	$\dfrac{\gamma(\sigma_1^2,\, y\mu^{-1}\sigma_1^{-2})}{\Gamma(\sigma_1^{-2})}$
Reference	Set σ_1 to $\sigma\mu^{(\nu/2)-1}$ in $\texttt{GA}(\mu, \sigma_1)$

2.4.3 Generalized gamma: GG

First parameterization, $\texttt{GG}(\mu, \sigma, \nu)$

The parameterization of the generalized gamma distribution used here and denoted by $\texttt{GG}(\mu,\sigma,\nu)$ was used by Lopatatzidis and Green [2000], with pdf

$$f_Y(y \mid \mu, \sigma, \nu) = \frac{|\nu|\, \theta^\theta z^\theta \exp(-\theta z)}{\Gamma(\theta) y} \tag{2.16}$$

for $y > 0$, where $\mu > 0$, $\sigma > 0$ and $-\infty < \nu < \infty$, $\nu \neq 0$, and where $z = (y/\mu)^\nu$ and $\theta = 1/(\sigma^2\nu^2)$. Note that $Z = (Y/\mu)^\nu \sim \texttt{GA}(1, \sigma\nu)$, from which the results in Table 2.15 are obtained. Note also that the gamma, $\texttt{GA}(\mu,\sigma)$, and Weibull, $\texttt{WEI}(\mu,\sigma)$, distributions are special cases of $\texttt{GG}(\mu,\sigma,\nu)$ given by $\texttt{GG}(\mu,\sigma,1) = \texttt{GA}(\mu,\sigma)$ and $\texttt{GG}(\mu,\sigma^{-1},\sigma) = \texttt{WEI}(\mu,\sigma)$, for $\mu > 0$ and $\sigma > 0$. Note also that

$$\mathrm{E}(Y^r) = \frac{\mu^r \Gamma(\theta + r/\nu)}{\theta^{r/\nu}\Gamma(\theta)} \qquad \text{if } \theta > -r/\nu \ .$$

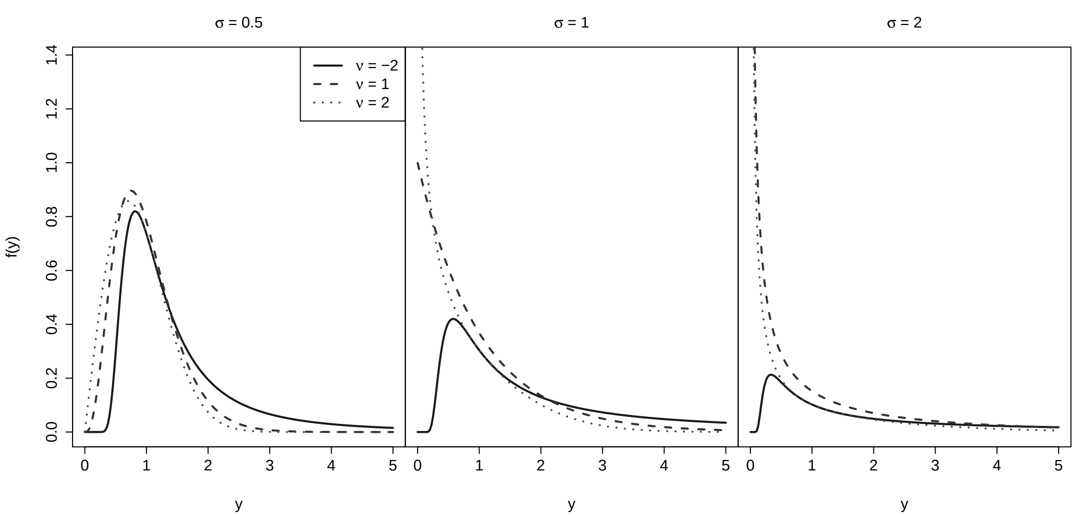

Figure 2.10: The generalized gamma, $\texttt{GG}(\mu,\sigma,\nu)$, distribution, with $\mu = 1$, $\sigma = 0.5, 1, 2$, and $\nu = -2, 1, 2$.

Table 2.15: Generalized gamma distribution.

$\texttt{GG}(\mu,\sigma,\nu)$	
Ranges	
Y	$0<y<\infty$
μ	$0<\mu<\infty$, scaling parameter
σ	$0<\sigma<\infty$
ν	$-\infty<\nu<\infty, \quad \nu\neq 0$
Distribution measures	
mean	$\begin{cases} \dfrac{\mu\Gamma(\theta+1/\nu)}{\theta^{1/\nu}\Gamma(\theta)} & \text{if } (\nu>0) \text{ or } (\nu<0 \text{ and } \sigma^2\|\nu\|<1) \\ \infty & \text{if } \nu<0 \text{ and } \sigma^2\|\nu\|\geq 1 \\ \text{where } \theta=1/(\sigma^2\nu^2) & \end{cases}$
mode	$\begin{cases} \mu(1-\sigma^2\nu)^{1/\nu} & \text{if } \sigma^2\nu<1 \\ \to 0 & \text{if } \sigma^2\nu\geq 1 \end{cases}$
variance	$\begin{cases} \dfrac{\mu^2}{\theta^{2/\nu}\left[\Gamma(\theta)\right]^2}\left\{\Gamma\left(\theta+2/\nu\right)\Gamma(\theta)-\left[\Gamma\left(\theta+1/\nu\right)\right]^2\right\} & \\ \qquad \text{if } (\nu>0) \text{ or } (\nu<0 \text{ and } \sigma^2\|\nu\|<1/2) & \\ \infty \qquad \text{if } \nu<0 \text{ and } \sigma^2\|\nu\|\geq 1/2 & \end{cases}$
skewness	$\begin{cases} \mu_3/[\mathrm{Var}(Y)]^{1.5} & \text{if } (\nu>0) \text{ or } (\nu<0 \text{ and } \sigma^2\|\nu\|<1/3) \\ \infty & \text{if } \nu<0 \text{ and } \sigma^2\|\nu\|\geq 1/3 \\ \text{where } \mu_3= & \dfrac{\mu^3}{\theta^{3/\nu}[\Gamma(\theta)]^3}\left\{\Gamma\left(\theta+3/\nu\right)[\Gamma(\theta)]^2-3\Gamma\left(\theta+2/\nu\right)\times\right. \\ & \left.\Gamma\left(\theta+1/\nu\right)\Gamma(\theta)+2\left[\Gamma\left(\theta+1/\nu\right)\right]^3\right\} \end{cases}$
excess kurtosis	$\begin{cases} \mu_4/[\mathrm{Var}(Y)]^2-3 & \text{if } (\nu>0) \text{ or } (\nu<0 \text{ and } \sigma^2\|\nu\|<1/4) \\ \infty & \text{if } \nu<0 \text{ and } \sigma^2\|\nu\|\geq 1/4 \\ \text{where } \mu_4=\dfrac{\mu^4}{\theta^{4/\nu}[\Gamma(\theta)]^4}\left\{\Gamma\left(\theta+4/\nu\right)[\Gamma(\theta)]^3-\right. & \\ \quad 4\Gamma\left(\theta+3/\nu\right)\Gamma\left(\theta+1/\nu\right)\left[\Gamma(\theta)\right]^2+ & \\ \quad \left.6\Gamma\left(\theta+2/\nu\right)\left[\Gamma\left(\theta+1/\nu\right)\right]^2\Gamma(\theta)-3\left[\Gamma\left(\theta+1/\nu\right)\right]^4\right\} & \end{cases}$
pdf	$\dfrac{\|\nu\|\,\theta^\theta z^\theta\exp\left(-\theta z\right)}{\Gamma(\theta)y}$, where $z=(y/\mu)^\nu$ and $\theta=1/(\sigma^2\nu^2)$
cdf	$\begin{cases} \gamma\left[\theta,\theta\left(y/\mu\right)^\nu\right]/\Gamma(\theta) & \text{if } \nu>0 \\ \Gamma\left[\theta,\theta\left(y/\mu\right)^\nu\right]/\Gamma(\theta) & \text{if } \nu<0 \end{cases}$

Second parameterization, `GG2`(μ, σ, ν)

A second parameterization, given by Johnson et al. [1994] p393, equation (17.128b), denoted here by `GG2`(μ, σ, ν), is defined as

$$f_Y(y \mid \mu, \sigma, \nu) = \frac{|\mu| \, y^{\mu\nu - 1}}{\Gamma(\nu) \, \sigma^{\mu\nu}} \exp\left[- \left(\frac{y}{\sigma} \right)^{\mu} \right] \tag{2.17}$$

for $y > 0$, where $-\infty < \mu < \infty$, $\mu \neq 0$, $\sigma > 0$ and $\nu > 0$. The parameterization (2.17) was suggested by Stacy and Mihram [1965], allowing μ to be negative. The moments of $Y \sim$ `GG2`(μ, σ, ν) can be obtained from those of `GG`(μ, σ, ν), since `GG`$(\mu, \sigma, \nu) =$ `GG2`$(\nu, \mu\theta^{-1/\nu}, \theta)$ and `GG2`$(\mu, \sigma, \nu) =$ `GG`$(\sigma\nu^{1/\mu}, \left[\mu^2\nu\right]^{-1/2}, \mu)$.

Note that `GG2`(μ, σ, ν) is not currently implemented in **gamlss.dist**.

2.4.4 Generalized inverse Gaussian: `GIG`

The parameterization of the generalized inverse Gaussian distribution, denoted by `GIG`(μ, σ, ν), is given by

$$f_Y(y \mid \mu, \sigma, \nu) = \left(\frac{b}{\mu} \right)^{\nu} \left[\frac{y^{\nu - 1}}{2K_{\nu}\left(\sigma^{-2}\right)} \right] \exp\left[-\frac{1}{2\sigma^2} \left(\frac{b\,y}{\mu} + \frac{\mu}{b\,y} \right) \right] \tag{2.18}$$

for $y > 0$, where $\mu > 0$, $\sigma > 0$ and $-\infty < \nu < \infty$, $b = \left[K_{\nu+1}\left(\sigma^{-2}\right)\right]\left[K_{\nu}\left(\sigma^{-2}\right)\right]^{-1}$ and $K_{\lambda}(t)$ is a modified Bessel function of the second kind [Abramowitz and Stegun, 1965], $K_{\lambda}(t) = \frac{1}{2}\int_0^{\infty} x^{\lambda - 1} \exp\{-\frac{1}{2}t(x + x^{-1})\}\, dx$.

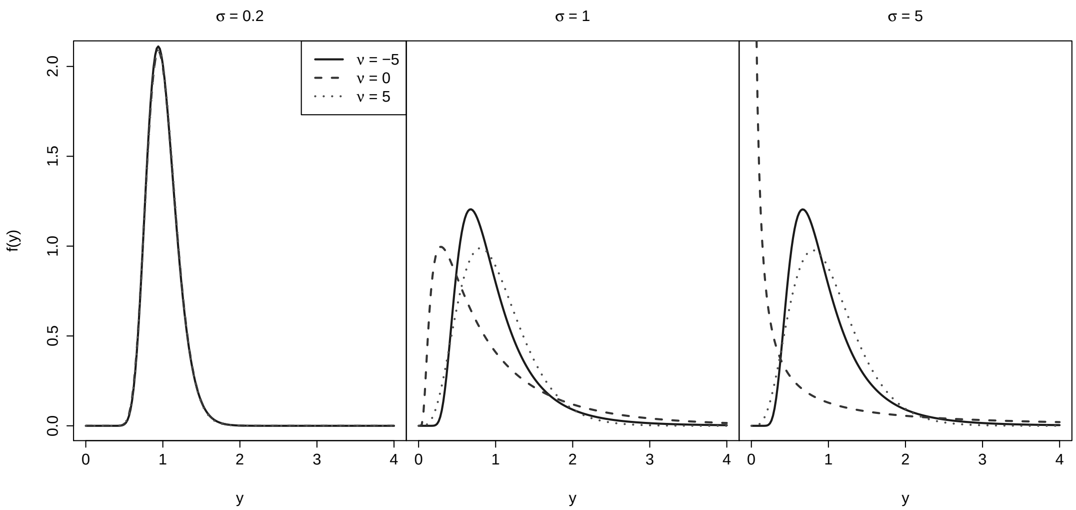

Figure 2.11: The generalized inverse Gaussian, `GIG`(μ, σ, ν), distribution, with $\mu = 1$, $\sigma = 0.2, 1, 5$ and $\nu = -5, 0, 5$.

`GIG`(μ, σ, ν) is a reparameterization of the generalized inverse Gaussian distribution of Jørgensen [1982]. Note also that `GIG`$(\mu, \sigma, -0.5) =$ `IG`$(\mu, \sigma\mu^{-1/2})$, a reparameterization of the inverse Gaussian distribution. Note that for $\nu > 0$, `GIG`(μ, σ, ν) has limiting distribution `GA`$(\mu, \nu^{-1/2})$ as $\sigma \to \infty$. (Derived using Jørgensen [1982, p171], equation (17.7): $K_\lambda(t) \sim \Gamma(\lambda) 2^{\lambda-1} t^{-\lambda}$ as $t \to 0$, for $\lambda > 0$.)

Table 2.16: Generalized inverse Gaussian distribution.

	$\texttt{GIG}(\mu,\sigma,\nu)$
	Ranges
Y	$0 < y < \infty$
μ	$0 < \mu < \infty$, mean, scaling parameter
σ	$0 < \sigma < \infty$
ν	$-\infty < \nu < \infty$
	Distribution measures
mean [a]	μ
mode [a2]	$\frac{\mu}{b}\{(\nu-1)\sigma^2 + [(\nu-1)^2\sigma^4+1]^{1/2}\}$
variance [a]	$\mu^2\left[\frac{2\sigma^2}{b}(\nu+1)+\frac{1}{b^2}-1\right]$
skewness [a]	$\begin{cases} \mu_3/[\mathrm{Var}(Y)]^{1.5} \text{ where} \\ \mu_3 = \mu^3\left[2-\frac{6\sigma^2}{b}(\nu+1)+\frac{4}{b^2}(\nu+1)(\nu+2)\sigma^4-\frac{2}{b^2}+\frac{2\sigma^2}{b^3}(\nu+2)\right] \end{cases}$
excess kurtosis [a]	$\begin{cases} k_4/[\mathrm{Var}(Y)]^2 \qquad \text{where} \\ k_4 = \mu^4\{-6+24\frac{\sigma^2}{b}(\nu+1)+\frac{4}{b^2}[2-\sigma^4(\nu+1)(7\nu+11)] \\ \qquad +4\frac{\sigma^2}{b^3}[2\sigma^4(\nu+1)(\nu+2)(\nu+3)-4\nu-5] \\ \qquad +\frac{1}{b^4}[4\sigma^4(\nu+2)(\nu+3)-2]\} \end{cases}$
MGF [a2]	$\left(1-\frac{2\mu\sigma^2 t}{b}\right)^{-\nu/2}\left[K_\nu\left(\sigma^{-2}\right)\right]^{-1}K_\nu\left[\sigma^{-2}\left(1-\frac{2\mu\sigma^2 t}{b}\right)^{1/2}\right]$
pdf [a]	$\left(\frac{b}{\mu}\right)^\nu\left[\frac{y^{\nu-1}}{2K_\nu\left(\sigma^{-2}\right)}\right]\exp\left[-\frac{1}{2\sigma^2}\left(\frac{by}{\mu}+\frac{\mu}{by}\right)\right]$
Reference	[a] Jørgensen [1982] page 6, equation (2.2), reparameterized by $\eta=\mu/b$, $\omega=1/\sigma^2$ and $\lambda=\nu$, and pages 15-17. [a2] Jørgensen [1982] p1, equation (1.1) reparameterized by $\chi=\mu/(\sigma^2 b)$ and $\psi=b/(\sigma^2\mu)$, with p7, equation (2.6), and p12, equation (2.9) where t is replaced by $-t$.
Note	$b=\left[K_{\nu+1}\left(\sigma^{-2}\right)\right]\left[K_\nu\left(\sigma^{-2}\right)\right]^{-1}$

2.4.5 Log normal (Box-Cox) family: `LNO`

The `gamlss.family` distribution $\texttt{LNO}(\mu,\sigma,\nu)$ allows the use of the Box-Cox power transformation approach [Box and Cox, 1964], where a transformation is applied to Y in order to remove skewness, given by

$$Z = \begin{cases} (Y^\nu - 1)/\nu & \text{if } \nu \neq 0 \\ \log(Y) & \text{if } \nu = 0 \ . \end{cases} \tag{2.19}$$

The transformed variable Z is then assumed to have a normal, `NO`(μ, σ), distribution. The resulting distribution of Y is denoted by `LNO`(μ, σ, ν) with pdf

$$f_Y(y \mid \mu, \sigma, \nu) = \frac{y^{\nu-1}}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{(z-\mu)^2}{2\sigma^2}\right] \tag{2.20}$$

for $y > 0$, where $-\infty < \mu < \infty$, $\sigma > 0$ and $-\infty < \nu < \infty$, and where z is defined as in (2.19). Strictly (2.20) is not exactly a proper distribution, since Y is positive and hence Z should have a truncated normal distribution. Hence the integral of $f_Y(y \mid \mu, \sigma, \nu)$ for y from 0 to ∞ is not exactly equal to 1 (and can be far from 1 especially if σ is large relative to μ). This is the *only* distribution in **gamlss.dist** which is not a proper distribution. When $\nu = 0$, this results in the `LOGNO2` distribution (equation (2.5)). The distribution in (2.20) can be fitted for fixed ν only, e.g. $\nu = 0.5$, using the following arguments of `gamlss()`:

```
family = LNO, nu.fix = TRUE, nu.start = 0.5
```

If ν is unknown, it can be estimated from its profile likelihood. Alternatively instead of (2.20), the more orthogonal parameterization given by the `BCCG` distribution in Section 2.4.1 is recommented, and `BCCG` is a proper distribution.

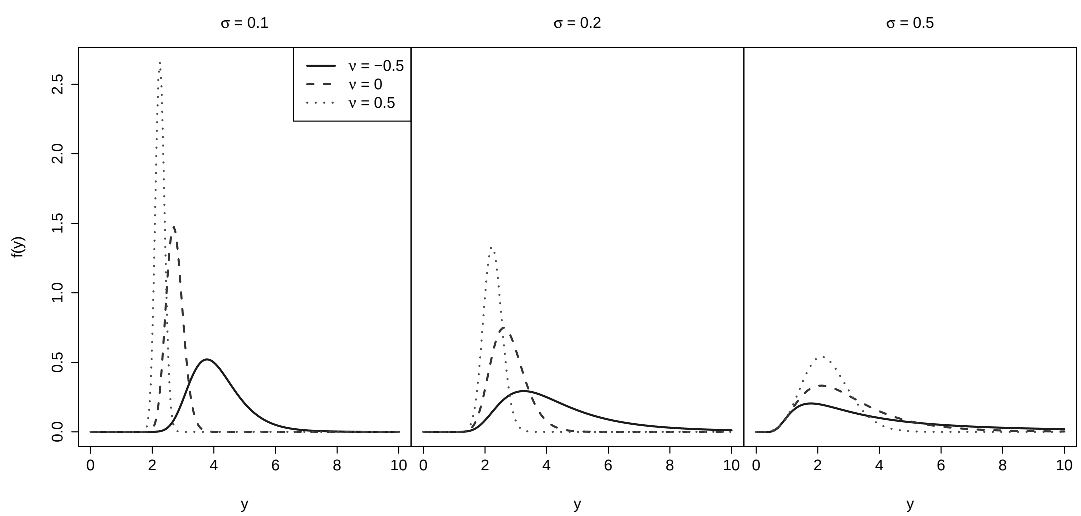

Figure 2.12: The log normal family, `LNO`(μ, σ, ν), distribution, with $\mu = 1$, $\sigma = 0.1, 0.2, 0.5$, and $\nu = -0.5, 0, 0.5$.

2.5 Continuous four-parameter distributions on $(0, \infty)$

2.5.1 Box-Cox t: `BCT`, `BCTo`

The Box-Cox t distribution, $\texttt{BCT}(\mu, \sigma, \nu, \tau)$, has been found to provide a good model for many response variables Y on $(0, \infty)$ when Y is not very close to zero.

Table 2.17: Box-Cox t distribution.

$\texttt{BCT}(\mu, \sigma, \nu, \tau)$	
Ranges	
Y	$0 < y < \infty$
μ	$0 < \mu < \infty$, median[a], scaling parameter
σ	$0 < \sigma < \infty$, approximate coefficient of variation[a,a2]
ν	$-\infty < \nu < \infty$, skewness parameter
τ	$0 < \tau < \infty$, kurtosis parameter
Distribution measures	
median[a]	μ
pdf	$\dfrac{y^{\nu-1} f_T(z)}{\mu^\nu \sigma F_T\left[(\sigma\lvert\nu\rvert)^{-1}\right]}$ where $T \sim t_\tau = \texttt{TF}(0, 1, \tau)$
cdf[a]	$F_T(z)$ where $T \sim t_\tau$ and z is given by (2.12)
inverse cdf[a]	$\begin{cases} \mu(1 + \sigma\nu t_{p,\tau})^{1/\nu} & \text{if } \nu \neq 0 \\ \mu \exp(\sigma t_{p,\tau}) & \text{if } \nu = 0 \\ \text{where } t_{p,\tau} = F_T^{-1}(p) \text{ and } T \sim t_\tau & \end{cases}$
Reference	Rigby and Stasinopoulos [2006]
Notes	[a] Provided $F_T\left(-(\sigma\lvert\nu\rvert)^{-1}\right)$ is negligible (say < 0.0001). [a2] Provided ν is reasonably close to 1 ($\nu \geq 0$ and not too large), and σ is small, say $\sigma < 0.3$.

Let Y be a positive random variable having a Box-Cox t distribution, [Rigby and Stasinopoulos, 2006], denoted by $\texttt{BCT}(\mu, \sigma, \nu, \tau)$, defined through the transformed random variable Z given by (2.12). The random variable Z is assumed to follow a truncated t_τ, i.e. truncated $\texttt{TF}(0, 1, \tau)$, distribution, with degrees of freedom τ treated as a continuous parameter, where $-1/(\sigma\nu) < Z < \infty$ if $\nu > 0$ and $-\infty < Z < -1/(\sigma\nu)$ if $\nu < 0$. The pdf of the $\texttt{BCT}(\mu, \sigma, \nu, \tau)$ distribution is given by

$$f_Y(y \mid \mu, \sigma, \nu, \tau) = \frac{y^{\nu-1} f_T(z)}{\mu^\nu \sigma F_T\left[(\sigma|\nu|)^{-1}\right]} \tag{2.21}$$

for $y > 0$, where $\mu > 0$, $\sigma > 0$, $-\infty < \nu < \infty$, $\tau > 0$, and z is given by (2.12) and $f_T(\cdot)$ and $F_T(\cdot)$ are the pdf and cdf, respectively, of $T \sim t_\tau = \texttt{TF}(0, 1, \tau)$, see Section 1.3.6, Hence

$$f_T(z) = \frac{1}{B\left(\frac{1}{2}, \frac{\tau}{2}\right)\tau^{1/2}}\left(1 + z^2/\tau\right)^{-(\tau+1)/2}.$$

[The distribution $\texttt{BCTo}(\mu, \sigma, \nu, \tau)$ also has pdf given by (2.21). It differs from $\texttt{BCT}(\mu, \sigma, \nu, \tau)$ only in having default log link function for μ, instead of the default identity link function for μ in $\texttt{BCT}(\mu, \sigma, \nu, \tau)$ in **gamlss.dist**.] Note that the special case $\texttt{BCT}(\mu, \sigma, 0, \tau)$ gives a log t family distribution. Note that $\texttt{BCCG}(\mu, \sigma, \nu)$ is the limiting distribution of $\texttt{BCT}(\mu, \sigma, \nu, \tau)$ as $\tau \to \infty$.

The exact cdf of Y is given by

$$F_Y(y) = \begin{cases} \dfrac{F_T(z)}{F_T\left[(\sigma|\nu|)^{-1}\right]} & \text{if } \nu \le 0 \\ \dfrac{F_T(z) - F_T\left[-(\sigma|\nu|)^{-1}\right]}{F_T\left[(\sigma|\nu|)^{-1}\right]} & \text{if } \nu > 0 \end{cases} \tag{2.22}$$

where z is given by (2.12).

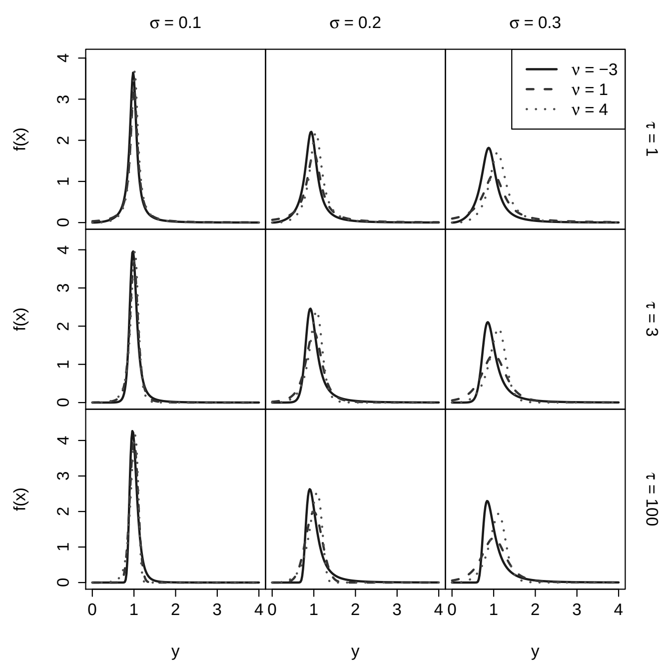

Figure 2.13: The Box-Cox t, $\texttt{BCT}(\mu, \sigma, \nu, \tau)$, distribution, with $\mu = 1$, $\sigma = 0.1, 0.2, 0.3$, $\nu = -3, 1, 4$, and $\tau = 1, 3, 100$.

The exact inverse cdf y_p of Y, defined by $P(Y \le y_p) = p$, is given by

$$y_p = \begin{cases} \mu(1+\sigma\nu z_T)^{1/\nu}, & \text{if } \nu \neq 0 \\ \mu\exp(\sigma z_T), & \text{if } \nu = 0 \ , \end{cases} \tag{2.23}$$

where, [Rigby and Stasinopoulos, 2006],

$$z_T = \begin{cases} F_T^{-1}\left[pF_T\left((\sigma|\nu|)^{-1}\right)\right] & \text{if } \nu \le 0 \\ F_T^{-1}\left[pF_T\left((\sigma|\nu|)^{-1}\right) + F_T\left(-(\sigma|\nu|)^{-1}\right)\right] & \text{if } \nu > 0 \ . \end{cases} \tag{2.24}$$

If the truncation probability $F_T\left(-(\sigma|\nu|)^{-1}\right)$ is negligible, then $F_Y(y) = F_T(z)$ in (2.22) and $z_T = F_T^{-1}(p) = t_{p,\tau}$ in (2.23) and (2.24), so Y has median μ. Note if $\nu = 0$ there is no truncation and the truncation probability is zero. The mean of Y is finite if $\nu < -1$ and also for $\tau > 1/\nu$ if $\nu > 0$. The variance of Y is finite if $\nu < -2$ and also for $\tau > 2/\nu$ if $\nu > 0$.

2.5.2 Box-Cox power exponential: `BCPE`, `BCPEo`

The Box-Cox exponential distribution, BCPE(μ,σ,ν,τ), is a very flexible distribution for a response variable Y on $(0,\infty)$, when Y is not very close to 0.

Let Y be a positive random variable having a Box-Cox power exponential distribution [Rigby and Stasinopoulos, 2004], denoted by BCPE(μ,σ,ν,τ), defined through the transformed random variable Z given by (2.12). Z is assumed to follow a truncated standard power exponential, i.e. truncated PE$(0,1,\tau)$ distribution, with power parameter $\tau > 0$ treated as a continuous parameter, where $-1/(\sigma\nu) < Z < \infty$ if $\nu > 0$ and $-\infty < Z < -1/(\sigma\nu)$ if $\nu < 0$. The pdf of Y is given by (2.21), where Z is given by (2.12), and where $f_T(\cdot)$ and $F_T(\cdot)$ are the pdf and cdf, respectively, of T having a standard power exponential distribution, i.e. $T \sim$ PE$(0,1,\tau)$, see Section 1.3.3. Hence

$$f_T(z) = \frac{\tau\exp\left[-\left|z/c\right|^\tau\right]}{2c\Gamma\left(\tau^{-1}\right)}$$

where $c^2 = \Gamma(\tau^{-1})[\Gamma(3\tau^{-1})]^{-1}$. Note that BCCG$(\mu,\sigma,\nu)$ = BCPE$(\mu,\sigma,\nu,2)$ is a special case of BCPE(μ,σ,ν,τ) when $\tau = 2$. Note also that the special case BCPE$(\mu,\sigma,0,\tau)$ given a log power exponential distribution. [The BCPEo(μ,σ,ν,τ) distribution also has the same pdf as BCPE(μ,σ,ν,τ) distribution. It differs from BCPE(μ,σ,ν,τ) only in having default log link function for μ, instead of the default identity link function for μ in BCPE(μ,σ,ν,τ), in **gamlss.dist**.]

The exact cdf of Y is given by (2.22) where $T \sim$ PE$(0,1,\tau)$. The exact inverse cdf y_p of Y is given by (2.23) where z_T is given by (2.24) and $T \sim$ PE$(0,1,\tau)$, [Rigby and Stasinopoulos, 2004]. If the truncation probability $F_T\left[-(\sigma|\nu|)^{-1}\right]$ is negligible, then $F_Y(y) = F_T(z)$ in (2.22) and $z_T = F_T^{-1}(p)$ in (2.23) and (2.24) where $T \sim$ PE$(0, 1, \tau)$, so Y has median μ. Note if $\nu = 0$ there is no truncation and the truncation probability is zero.

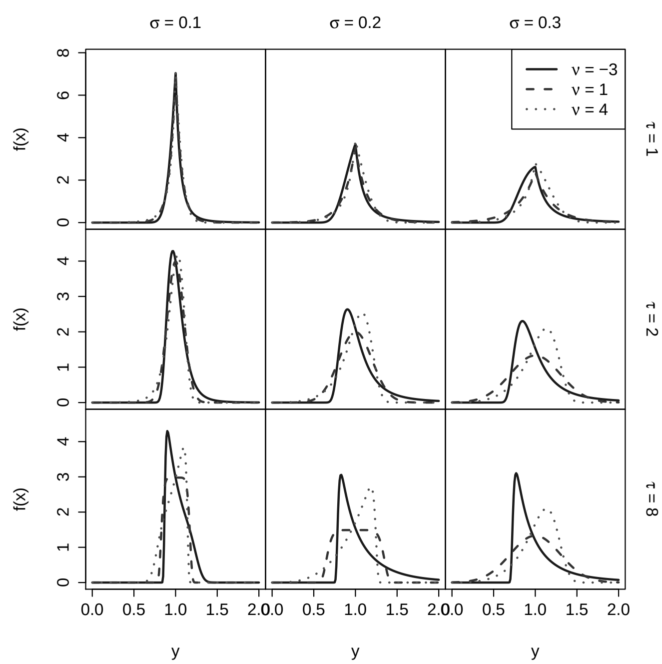

Figure 2.14: The Box-Cox power exponential, BCPE(μ, σ, ν, τ), distribution, with $\mu = 1$, $\sigma = 0.1, 0.2, 0.3$, $\nu = -3, 1, 4$, and $\tau = 1, 2, 8$.

Table 2.18: Box-Cox power exponential distribution.

BCPE(μ,σ,ν,τ)	
Ranges	
Y	$0<y<\infty$
μ	$0<\mu<\infty$, median[a], scaling parameter
σ	$0<\sigma<\infty$, approximate coefficient of variation[a,a2]
ν	$-\infty<\nu<\infty$, skewness parameter
τ	$0<\tau<\infty$ kurtosis parameter
Distribution measures	
median [a]	μ
pdf	$\dfrac{y^{\nu-1}f_T(z)}{\mu^\nu \sigma F_T\left[(\sigma\|\nu\|)^{-1}\right]}$ where $T\sim$ PE$(0,1,\tau)$
cdf [a]	$F_T(z)$ where $T\sim$ PE$(0,1,\tau)$ and z is given by (2.12)
inverse cdf[a]	$\begin{cases}\mu\left[1+\sigma\nu F_T^{-1}(p)\right]^{1/\nu} & \text{if } \nu\neq 0\\ \mu\exp\left[\sigma F_T^{-1}(p)\right] & \text{if } \nu=0\\ \text{where } T\sim \texttt{PE}(0,1,\tau) & \end{cases}$
Reference	Rigby and Stasinopoulos [2004]
Notes	[a] Provided $F_T\left(-(\sigma\|\nu\|)^{-1}\right)$ is neglibible (e.g. say <0.0001). [a2] Provided ν is reasonably close to 1 ($\nu\geq 0$ and not too large), and σ is small, say $\sigma<0.3$.

2.5.3 Generalized beta type 2: GB2

This pdf of the generalized beta type 2 distribution, denoted by GB2(μ,σ,ν,τ), is given by

$$\begin{aligned} f_Y(y\,|\,\mu,\sigma,\nu,\tau) &= |\sigma|\,y^{\sigma\nu-1}\left\{\mu^{\sigma\nu}\,B(\nu,\tau)\,\left[1+(y/\mu)^\sigma\right]^{\nu+\tau}\right\}^{-1}\\ &= \frac{\Gamma(\nu+\tau)}{\Gamma(\nu)\Gamma(\tau)}\frac{|\sigma|\,(y/\mu)^{\sigma\nu}}{y\left[1+(y/\mu)^\sigma\right]^{\nu+\tau}} \end{aligned} \tag{2.25}$$

for $y>0$, where $\mu>0$, $\sigma>0$, $\nu>0$ and $\tau>0$, [McDonald [1984] p648, equation (3), who called it the generalized beta of the second kind.] Note that McDonald and Xu [1995] appear to allow for $\sigma<0$, however this is unnecessary since GB2$(\mu,-\sigma,\nu,\tau)=$ GB2(μ,σ,τ,ν). So we assume $\sigma>0$, and $|\sigma|$ can be replaced by σ in (2.25). The GB2 distribution is also considered by McDonald and Xu [1995, p136-140] and McDonald [1996, p433-435].

If $Y\sim$ GB2(μ,σ,ν,τ) then $Y_1=[1+(Y/\mu)^{-\sigma}]^{-1}\sim$ BEo(ν,τ), from which the cdf of Y in Table 2.19 is obtained. Also $Z=(\tau/\nu)(Y/\mu)^\sigma\sim F_{2\nu,2\tau}$, an F distribution defined in equation (**??**).

Note

$$\mathrm{E}(Y^r) = \mu^r \frac{B(\nu + r\sigma^{-1}, \tau - r\sigma^{-1})}{B(\nu, \tau)} \quad \text{for } \tau > r\sigma^{-1}$$

[McDonald and Xu, 1995, p136, equation (2.8)], from which the mean, variance, skewness, and excess kurtosis in Table 2.19 are obtained.

Setting $\sigma = 1$ in (2.25) gives a form of the Pearson type VI distribution:

$$f_Y(y \,|\, \mu, \nu, \tau) = \frac{\Gamma(\nu + \tau)}{\Gamma(\nu)\Gamma(\tau)} \frac{\mu^\tau y^{\nu - 1}}{(y + \mu)^{\nu + \tau}} \ . \tag{2.26}$$

Setting $\nu = 1$ in (2.25) gives the Burr XII (or Singh-Maddala) distribution:

$$f_Y(y \,|\, \mu, \sigma, \tau) = \frac{\tau\sigma(y/\mu)^\sigma}{y\left[1 + (y/\mu)^\sigma\right]^{\tau + 1}} \ . \tag{2.27}$$

Setting $\tau = 1$ in (2.25) gives the Burr III (or Dagum) distribution

$$f_Y(y \,|\, \mu, \sigma, \nu) = \frac{\nu\sigma(y|\mu)^{\sigma\nu}}{y[1 + (y/\mu)^\sigma]^{\nu + 1}} \ .$$

Setting $\sigma = 1$ and $\nu = 1$ in (2.25) gives the Pareto type 2 original distribution, `PARETO2o`(μ, τ). Setting $\nu = 1$ and $\tau = 1$ in (2.25) gives the log logistic distribution. Other special cases and limiting cases are given in McDonald and Xu [1995, pp136, 139].

Table 2.19: Generalized beta type 2 distribution.

$\texttt{GB2}(\mu,\sigma,\nu,\tau)$	
Ranges	
Y	$0<y<\infty$
μ	$0<\mu<\infty$, scaling parameter
σ	$0<\sigma<\infty$
ν	$0<\nu<\infty$
τ	$0<\tau<\infty$
Distribution measures	
mean	$\begin{cases} \mu\dfrac{B(\nu+\sigma^{-1},\tau-\sigma^{-1})}{B(\nu,\tau)} & \text{if } \tau>\sigma^{-1} \\ \infty & \text{if } \tau\le\sigma^{-1} \end{cases}$
mode	$\begin{cases} \mu\left(\dfrac{\sigma\nu-1}{\sigma\tau+1}\right)^{1/\sigma} & \text{if } \nu>\sigma^{-1} \\ \to 0 & \text{if } \nu\le\sigma^{-1} \end{cases}$
variance	$\begin{cases} \dfrac{\mu^2\{B(\nu+2\sigma^{-1},\tau-2\sigma^{-1})B(\nu,\tau)-[B(\nu+\sigma^{-1},\tau-\sigma^{-1})]^2\}}{[B(\nu,\tau)]^2} & \text{if } \tau>2\sigma^{-1} \\ \infty & \text{if } \tau\le 2\sigma^{-1} \end{cases}$
skewness	$\begin{cases} \mu_3/[\mathrm{Var}(Y)]^{1.5} & \text{if } \tau>3\sigma^{-1} \\ \infty & \text{if } \tau\le 3\sigma^{-1} \quad \text{where} \\ \mu_3=\mu^3\dfrac{1}{[B(\nu,\tau)]^3}\{B(\nu+3\sigma^{-1},\tau-3\sigma^{-1})[B(\nu,\tau)]^2 & \\ \quad -3B(\nu+2\sigma^{-1},\tau-2\sigma^{-1})B(\nu+\sigma^{-1},\tau-\sigma^{-1})B(\nu,\tau) & \\ \quad +2[B(\nu+\sigma^{-1},\tau-\sigma^{-1})]^3\} & \end{cases}$
excess kurtosis	$\begin{cases} \mu_4/[\mathrm{Var}(Y)]^2-3 & \text{if } \tau>4\sigma^{-1} \\ \infty & \text{if } \tau\le 4\sigma^{-1} \quad \text{where} \\ \mu_4=\mu^4\dfrac{1}{[B(\nu,\tau)]^4}\{B(\nu+4\sigma^{-1},\tau-4\sigma^{-1})[B(\nu,\tau)]^3 & \\ \quad -4B(\nu+3\sigma^{-1},\tau-3\sigma^{-1})B(\nu+\sigma^{-1},\tau-\sigma^{-1})[B(\nu,\tau)]^2 & \\ \quad +6B(\nu+2\sigma^{-1},\tau-2\sigma^{-1})[B(\nu+\sigma^{-1},\tau-\sigma^{-1})]^2B(\nu,\tau) & \\ \quad -3[B(\nu+\sigma^{-1},\tau-\sigma^{-1})]^4\} & \end{cases}$
pdf[a]	$\lvert\sigma\rvert y^{\sigma\nu-1}\left\{\mu^{\sigma\nu}B(\nu,\tau)\left[1+(y/\mu)^{\sigma}\right]^{\nu+\tau}\right\}^{-1}$
cdf	$\dfrac{B(\nu,\tau,c)}{B(\nu,\tau)}$ where $c=1/[1+(y/\mu)^{-\sigma}]$
Reference	[a] McDonald [1984] p648, equation (3) where $b=\mu$, $a=\sigma$, $p=\nu$ and $q=\tau$.

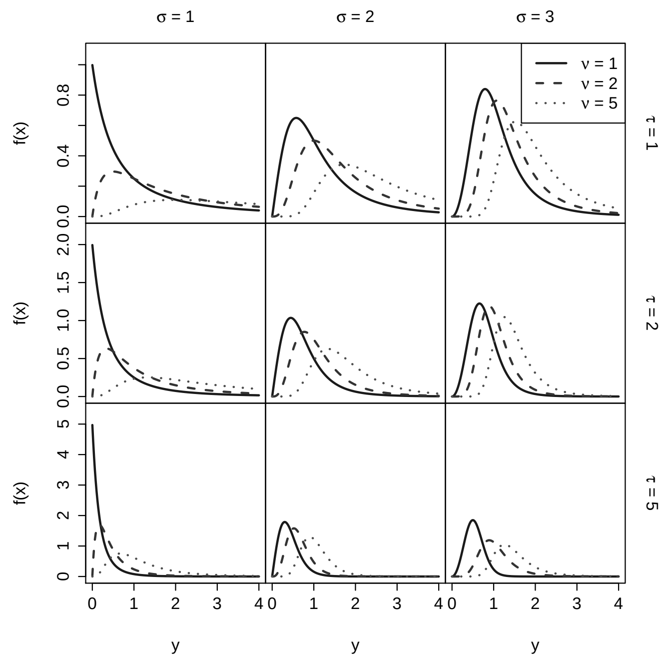

Figure 2.15: The generalized beta type 2, GB2(μ, σ, ν, τ), distribution, with $\mu = 1$, $\sigma = 1, 2, 3$, $\nu = 1, 2, 5$, and $\tau = 1, 2, 5$.

Chapter 3

Mixed distributions on $[0, \infty)$

This chapter gives summary tables and plots for the explicit **gamlss.dist** mixed distributions with range $[0, \infty)$, which are continuous on $(0, \infty)$ and include a point probability at zero. These are discussed in Section **??**, which also discusses creating distributions on $[0, \infty)$ in the **gamlss** packages either by generating a zero-adjusted distribution [i.e. adding a point probability at zero to any `gamlss.family` distribution on $(0, \infty)$] or by generalized Tobit model [i.e. censoring below zero any **gamlss.dist** distribution on $(-\infty, \infty)$ to give a point probability at zero].

3.1 Zero-adjusted gamma: `ZAGA`

The zero-adjusted gamma distribution, `ZAGA`, is appropriate when the response variable Y takes values from zero to infinity, including exact value zero, i.e. $[0,\infty)$. Here $Y = 0$ with nonzero probability ν, and $Y \sim$ `GA`(μ, σ) with probability $(1 - \nu)$. The mixed probability function of the zero-adjusted gamma distribution, denoted by `ZAGA`(μ, σ, ν), is given (informally) by

$$f_Y(y|\mu, \sigma, \nu) = \begin{cases} \nu & \text{if } y = 0 \\ (1-\nu)\left[\dfrac{y^{1/\sigma^2-1}e^{-y/(\sigma^2\mu)}}{(\sigma^2\mu)^{1/\sigma^2}\Gamma(1/\sigma^2)}\right] & \text{if } y > 0 \end{cases} \tag{3.1}$$

for $y \geq 0$, where $\mu > 0$, $\sigma > 0$ and $0 < \nu < 1$.

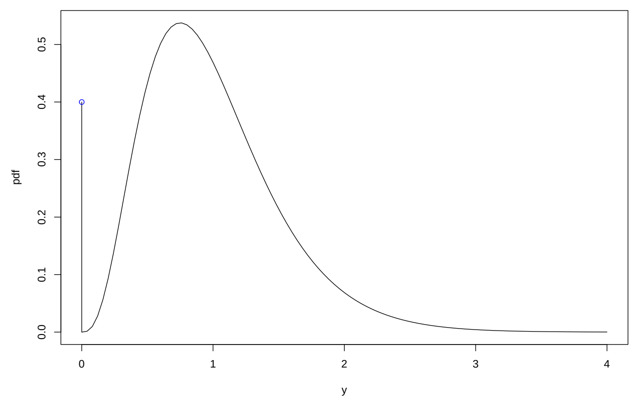

Figure 3.1: Zero-adjusted gamma, ZAGA(μ, σ, ν), distribution, with $\mu = 1$, $\sigma = 0.5$, and $\nu = 0.4$.

Table 3.1: Zero-adjusted gamma distribution.

	$\texttt{ZAGA}(\mu, \sigma, \nu)$
	Ranges
Y	$0 \le y < \infty$,
μ	$0 < \mu < \infty$, mean of gamma component
σ	$0 < \sigma < \infty$, coefficient of variation of gamma component
ν	$0 < \nu < 1$, $\mathrm{P}(Y=0)$
	Distribution measures
mean	$(1-\nu)\mu$
mode	$\begin{cases} 0 & \text{if } \sigma \ge 1 \\ 0 \text{ and } \mu(1-\sigma^2) & \text{if } \sigma < 1 \end{cases}$
variance	$(1-\nu)\mu^2(\sigma^2+\nu)$
skewness	$\begin{cases} \mu_3/[\mathrm{Var}(Y)]^{1.5} \text{ where} \\ \mu_3 = \mu^3(1-\nu)\left(2\sigma^4+3\nu\sigma^2+2\nu^2-\nu\right) \end{cases}$
excess kurtosis	$\begin{cases} \mu_4/[\mathrm{Var}(Y)]^2 - 3 \text{ where} \\ \mu_4 = \mu^4(1-\nu)\left[6\sigma^6+3\sigma^4+8\nu\sigma^4+6\nu^2\sigma^2+\nu(1-3\nu+3\nu^2)\right] \end{cases}$
MGF	$\nu + (1-\nu)(1-\mu\sigma^2 t)^{-1/\sigma^2}$ for $t < (\mu\sigma^2)^{-1}$
pdf	$\begin{cases} \nu & \text{if } y = 0 \\ (1-\nu)\left[\dfrac{y^{1/\sigma^2-1}e^{-y/(\sigma^2\mu)}}{(\sigma^2\mu)^{1/\sigma^2}\Gamma(1/\sigma^2)}\right] & \text{if } y > 0 \end{cases}$
cdf	$\begin{cases} \nu & \text{if } y = 0 \\ \nu + (1-\nu)\dfrac{\gamma(\sigma^{-2}, y\mu^{-1}\sigma^{-2})}{\Gamma(\sigma^{-2})} & \text{if } y > 0 \end{cases}$
Reference	Obtained from equations (**??**), (**??**), (**??**), and (**??**), where $Y_1 \sim \texttt{GA}(\mu, \sigma)$.

3.2 Zero-adjusted inverse Gaussian: ZAIG

The zero-adjusted inverse Gaussian distribution, ZAIG, is appropriate when the response variable Y takes values from zero to infinity, including exact value zero, i.e. $[0, \infty)$. Here $Y = 0$ with nonzero probability ν, and $Y \sim \texttt{IG}(\mu, \sigma)$ with probability $(1-\nu)$. The mixed probability function of the zero-adjusted inverse Gaussian distribution, denoted by $\texttt{ZAIG}(\mu, \sigma, \nu)$, is given (informally) by

$$f_Y(y|\mu,\sigma,\nu) = \begin{cases} \nu & \text{if } y = 0 \\ (1-\nu)\dfrac{1}{\sqrt{2\pi\sigma^2 y^3}} \exp\left[-\dfrac{1}{2\mu^2\sigma^2 y}(y-\mu)^2\right] & \text{if } y > 0 \end{cases} \tag{3.2}$$

for $y \ge 0$, where $\mu > 0$, $\sigma > 0$ and $0 < \nu < 1$.

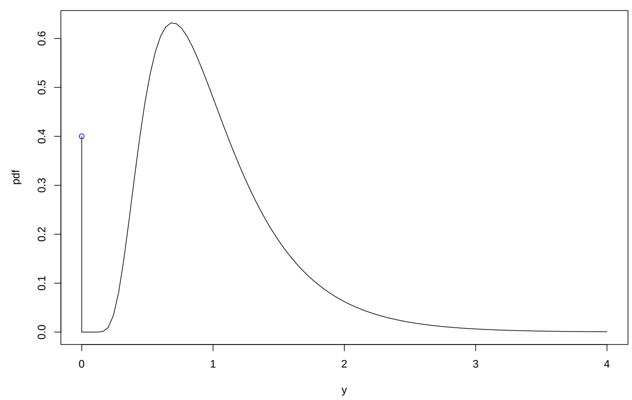

Figure 3.2: Zero-adjusted inverse Gaussian, `ZAIG`(μ, σ, ν), distribution, with $\mu = 1$, $\sigma = 0.5$, and $\nu = 0.4$.

Table 3.2: Zero-adjusted inverse Gaussian distribution.

	ZAIG(μ, σ, ν)
	Ranges
Y	$0 \leq y < \infty$
μ	$0 < \mu < \infty$, mean of inverse Gaussian component
σ	$0 < \sigma < \infty$
ν	$0 < \nu < 1$, $\quad \mathrm{P}(Y=0)$
	Distribution measures
mean	$(1-\nu)\mu$
mode	0 and $\left[-3\mu^2\sigma^2 + \mu(9\mu^2\sigma^4+4)^{1/2}\right]/2$
variance	$(1-\nu)\mu^2(\nu+\mu\sigma^2)$
skewness	$\begin{cases} \mu_3/[\mathrm{Var}(Y)]^{1.5} \text{ where} \\ \mu_3 = \mu^3(1-\nu)\left[3\mu^2\sigma^4 + 3\mu\sigma^2\nu + 2\nu^2 - \nu\right] \end{cases}$
excess kurtosis	$\begin{cases} \mu_4/[\mathrm{Var}(Y)]^2 - 3 \text{ where} \\ \mu_4 = \mu^4(1-\nu)\left[3\mu^2\sigma^4 + 15\mu^3\sigma^6 + \nu(1+12\mu^2\sigma^4) + \right. \\ \left. \nu^2(6\mu\sigma^2-3) + 3\nu^3\right] \end{cases}$
MGF	$\nu + (1-\nu)\exp\left\{\dfrac{1}{\mu\sigma^2}\left[1-\left(1-2\mu^2\sigma^2 t\right)^{1/2}\right]\right\}$ for $t < (2\mu^2\sigma^2)^{-1}$
pdf	$\begin{cases} \nu & \text{if } y=0 \\ (1-\nu)\dfrac{1}{\sqrt{2\pi\sigma^2 y^3}} \exp\left[-\dfrac{1}{2\mu^2\sigma^2 y}(y-\mu)^2\right] & \text{if } y>0 \end{cases}$
cdf	$\begin{cases} \nu & \text{if } y=0 \\ \nu + (1-\nu)\Phi\left[(\sigma^2 y)^{-1/2}\left(\dfrac{y}{\mu}-1\right)\right] & \\ +e^{2(\mu\sigma^2)^{-1}}\Phi\left[-(\sigma^2 y)^{-1/2}\left(\dfrac{y}{\mu}-1\right)\right] & \text{if } y>0 \end{cases}$
Reference	Obtained from equations (**??**), (**??**), (**??**), and (**??**), where $Y_1 \sim$ IG(μ, σ).

Chapter 4

Continuous and mixed distributions on $[0, 1]$

This chapter gives summary tables and plots for the explicit **gamlss.dist** continuous distributions with range $(0, 1)$. These are discussed in Chapter **??**, which also discusses creating distributions on $(0, 1)$ in the **gamlss** packages, either by an inverse logit transform, or by truncation below zero and above one, from any `gamlss.family` distribution on $(-\infty, \infty)$, or by truncation above one from any `gamlss.family` distribution on $(0, \infty)$.

This chapter also gives summary tables and plots for the explicit **gamlss.dist** mixed distributions with ranges $[0, 1)$, $(0, 1]$, and $[0, 1]$, which are continuous on $(0, 1)$ and include point probabilities at zero or one, or both, respectively. These are discussed in Section **??**, which also discusses creating distributions on $[0, 1)$, $(0, 1]$, or $[0, 1]$ in **gamlss**, firstly by generating an 'inflated distribution' [i.e. by adding point probabilities at zero, or one, or both, to any `gamlss.family` continuous distribution on $(0, 1)$], and secondly by a generalized Tobit model [i.e. obtaining the point probabilities at zero, or one, or both, by censoring].

Note that if a random variable Y has any range from a to b, where a and b are known and finite, then $Z = (Y - a)/(b - a)$ has range from 0 to 1 and can be modeled using distributions in this chapter.

4.1 Continuous two-parameter distributions on $(0, 1)$

4.1.1 Beta: BE, BEo

The beta distribution is appropriate when the response variable takes values in a known restricted range, excluding the endpoints of the range. Appropriate standardization can be applied to make the range of the response variable (0,1). Note that the exact values $Y = 0$ and $Y = 1$ are not included in the range and so have zero density under the model.

First parameterization, BEo

The original parameterization of the beta distribution, denoted by $\texttt{BEo}(\mu, \sigma)$, has pdf

$$f_Y(y \mid \mu, \sigma) = \frac{1}{B(\mu, \sigma)}\, y^{\mu-1}(1-y)^{\sigma-1}$$

for $0 < y < 1$, with parameters $\mu > 0$ and $\sigma > 0$. Here $\mathrm{E}(Y) = \mu/(\mu + \sigma)$ and $\mathrm{Var}(y) = \mu\sigma(\mu + \sigma)^{-2}(\mu + \sigma + 1)^{-1}$.

Table 4.1: The beta distribution (original parameterization).

	`BEo`(μ, σ)
	Ranges
Y	$0 < y < 1$
μ	$0 < \mu < \infty$
σ	$0 < \sigma < \infty$
	Distribution measures
mean [a]	$\dfrac{\mu}{\mu+\sigma}$
mode	$\begin{cases} \dfrac{\mu-1}{\mu+\sigma-2} & \text{if } \mu > 1 \text{ and } \sigma > 1 \\ \to 0 & \text{if } 0 < \mu < 1 \text{ and } \sigma > 1 \\ \to 1 & \text{if } \mu > 1 \text{ and } 0 < \sigma < 1 \\ \to 0 \text{ and } 1 & \text{if } 0 < \mu < 1 \text{ and } 0 < \sigma < 1 \end{cases}$
variance [a]	$\dfrac{\mu\sigma}{(\mu+\sigma)^2(\mu+\sigma+1)}$
skewness [a]	$\dfrac{2(\sigma-\mu)(\mu+\sigma+1)^{0.5}}{\mu^{0.5}\sigma^{0.5}(\mu+\sigma+2)}$
excess kurtosis[a]	$\begin{cases} \beta_2 - 3 \text{ where} \\ \beta_2 = \dfrac{3(\mu+\sigma+1)[2(\mu+\sigma)^2+\mu\sigma(\mu+\sigma-6)]}{\mu\sigma(\mu+\sigma+2)(\mu+\sigma+3)} \end{cases}$
pdf [a]	$\dfrac{1}{B(\mu,\sigma)} y^{\mu-1}(1-y)^{\sigma-1}$
cdf	$\dfrac{B(\mu,\sigma,y)}{B(\mu,\sigma)}$
Reference	[a] Johnson et al. [1995] chapter 25, p210, equation (25.2) and p217, with (p, q) set to (μ, σ)

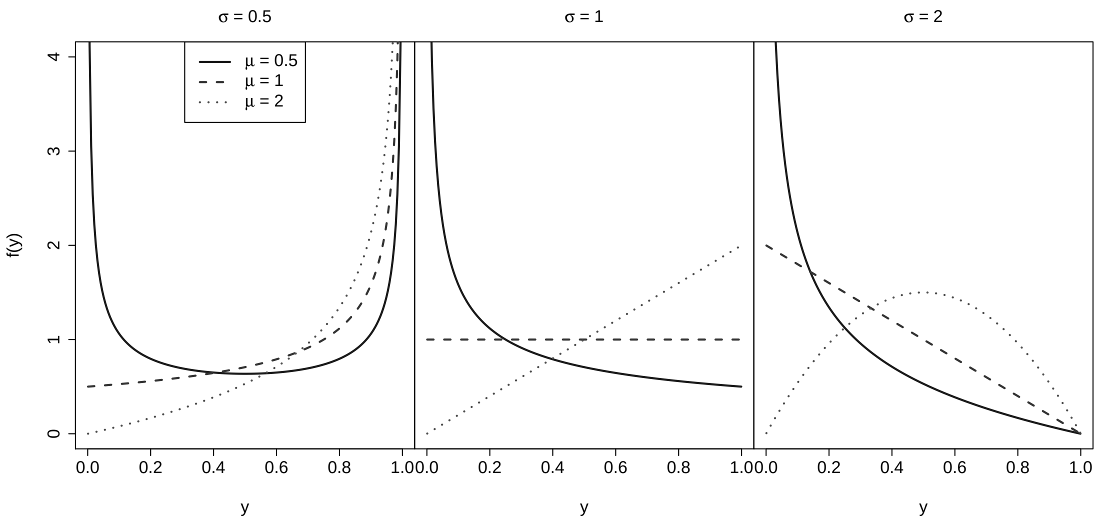

Figure 4.1: Beta, `BEo`(μ, σ), distribution, with $\mu = 0.5, 1, 2$ and $\sigma = 0.5, 1, 2$.

Second parameterization, `BE`

The pdf of the beta distribution, denoted by `BE`(μ, σ), is given by

$$f_Y(y \,|\, \mu, \sigma) = \frac{1}{B(\alpha, \beta)}\; y^{\alpha-1}(1-y)^{\beta-1} \tag{4.1}$$

for $0 < y < 1$, where

$$\begin{aligned} \alpha &= \mu(1-\sigma^2)/\sigma^2 \\ \beta &= (1-\mu)(1-\sigma^2)/\sigma^2 \ , \end{aligned} \tag{4.2}$$

$\alpha > 0$, and $\beta > 0$ and hence $0 < \mu < 1$ and $0 < \sigma < 1$. Note the relationship between parameters (μ,σ) and (α,β) is given by

$$\begin{aligned} \mu &= \alpha/(\alpha+\beta) \\ \sigma &= (\alpha+\beta+1)^{-1/2} \ . \end{aligned}$$

Hence `BE`(μ, σ) = `BEo`(α, β). In the parameterization $Y \sim$ `BE`(μ, σ), the mean and variance of Y are $\mathrm{E}(Y) = \mu$ and $\mathrm{Var}(Y) = \sigma^2\mu(1-\mu)$. Hence $\mathrm{SD}(Y) = \sigma\mu^{0.5}(1-\mu)^{0.5}$ and σ scales the standard deviation of Y.

Table 4.2: The beta distribution.

	$\texttt{BE}(\mu,\sigma)$
	Ranges
Y	$0<y<1$
μ	$0<\mu<1$, mean
σ	$0<\sigma<1$
	Distribution measures
mean	μ
mode	$\begin{cases} \mu+\frac{(2\mu-1)\sigma^2}{1-3\sigma^2} & \text{if } \sigma^2<\left[1+\max\left(\frac{1}{\mu},\frac{1}{1-\mu}\right)\right]^{-1} \\ \to 0 & \text{if } \left(1+\frac{1}{\mu}\right)^{-1}<\sigma^2<\left(1+\frac{1}{1-\mu}\right)^{-1} \\ \to 1 & \text{if } \left(1+\frac{1}{1-\mu}\right)^{-1}<\sigma^2<\left(1+\frac{1}{\mu}\right)^{-1} \\ \to 0 \text{ and } 1 & \text{if } \sigma^2>\left[1+\min\left(\frac{1}{\mu},\frac{1}{1-\mu}\right)\right]^{-1} \end{cases}$
variance	$\sigma^2\mu(1-\mu)$
skewness	$\dfrac{2(1-2\mu)\sigma}{\mu^{0.5}(1-\mu)^{0.5}(1+\sigma^2)}$
excess kurtosis	$\begin{cases} \beta_2-3 \text{ where} \\ \beta_2=\dfrac{6\sigma^2+3\mu(1-\mu)(1-7\sigma^2)}{\mu(1-\mu)(1+\sigma^2)(1+2\sigma^2)} \end{cases}$
pdf	$\frac{1}{B(\alpha,\beta)}y^{\alpha-1}(1-y)^{\beta-1}$ where $\alpha=\mu(1-\sigma^2)/\sigma^2$ and $\beta=(1-\mu)(1-\sigma^2)/\sigma^2$
cdf	$B(\alpha,\beta,y)/B(\alpha,\beta)$
Reference	Reparameterize $\texttt{BEo}(\alpha,\beta)$ by setting $\alpha=\mu(1-\sigma^2)/\sigma^2$ and $\beta=(1-\mu)(1-\sigma^2)/\sigma^2$

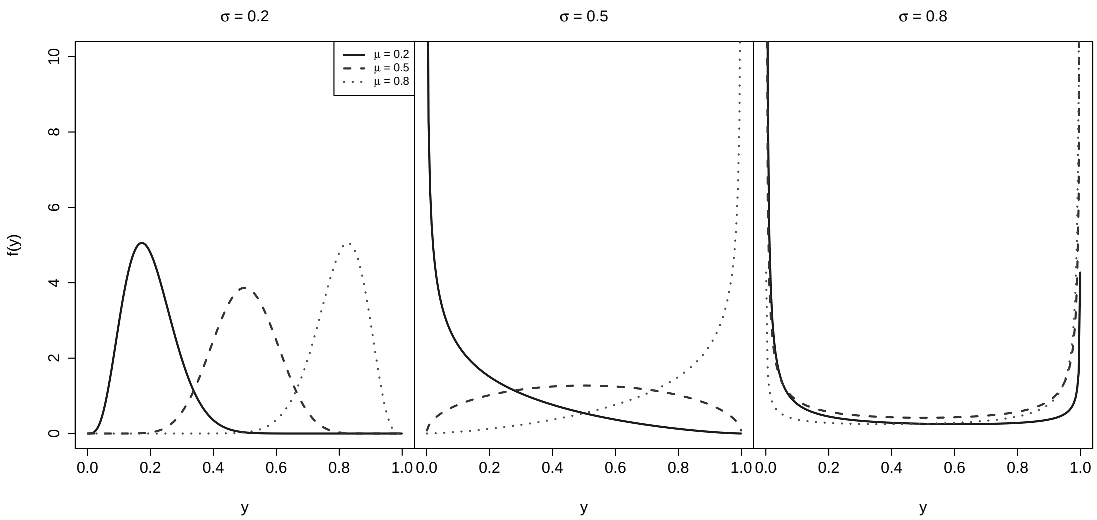

Figure 4.2: Beta, $\texttt{BE}(\mu, \sigma)$, distribution, with $\mu = 0.2, 0.5, 0.8$ and $\sigma = 0.2, 0.5, 0.8$

4.1.2 Logit normal: `LOGITNO`

The pdf of the logit normal distribution, denoted by $\texttt{LOGITNO}(\mu, \sigma)$, is

$$f_Y(y \mid \mu, \sigma) = \frac{1}{\sqrt{2\pi\sigma^2}} \frac{1}{y(1-y)} \exp\left(-\frac{\{\log[y/(1-y)] - \log[\mu/(1-\mu)]\}^2}{2\sigma^2} \right) \tag{4.3}$$

for $0 < y < 1$, where $0 < \mu < 1$ and $\sigma > 0$.

Note that $\log[Y/(1-Y)] \sim \texttt{NO}(\log[\mu/(1-\mu)], \sigma)$ and μ is the median of Y.

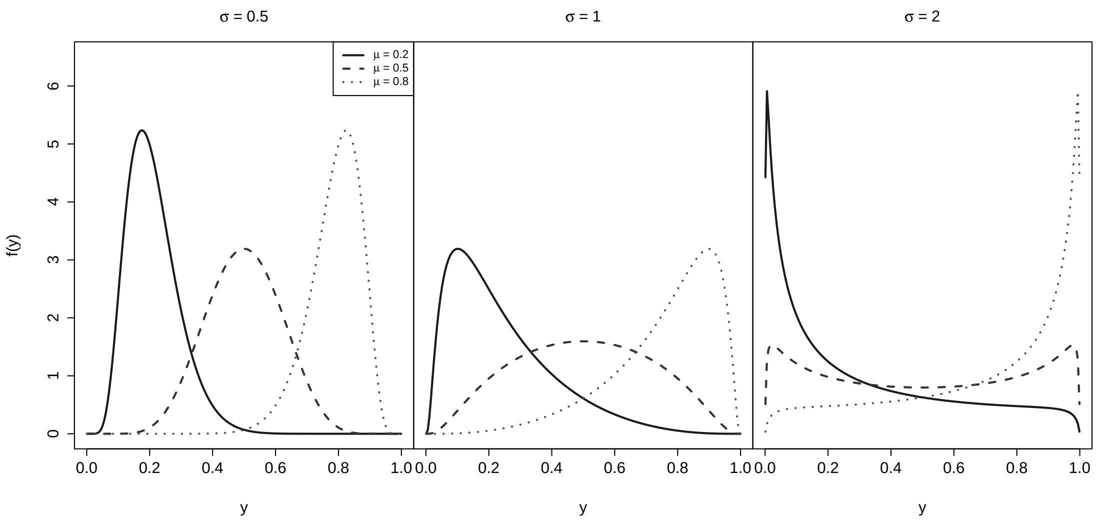

Figure 4.3: The logit normal, `LOGITNO`(μ, σ), distribution, for $\mu = 0.2, 0.5, 0.8$ and $\sigma = 0.5, 1, 2$.

4.1.3 Simplex: `SIMPLEX`

The pdf of the simplex distribution, denoted by `SIMPLEX`(μ, σ), is given by

$$f_Y(y \mid \mu, \sigma) = \frac{1}{[2\pi\sigma^2 y^3 (1-y)^3]^{1/2}} \exp\left[-\frac{(y-\mu)^2}{2\sigma^2 y (1-y) \mu^2 (1-\mu)^2}\right], \qquad (4.4)$$

for $0 < y < 1$, where $0 < \mu < 1$ and $\sigma > 0$. The mean of the distribution is μ. This distribution was called the 'standard simplex distribution' by Jørgensen [1997, p199].

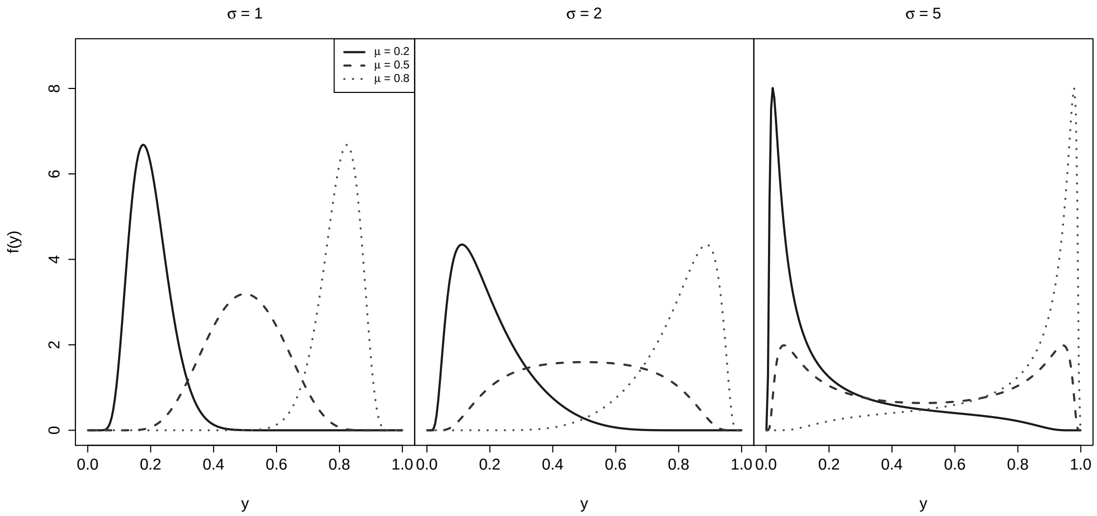

Figure 4.4: The simplex, SIMPLEX(μ,σ), distribution, with $\mu = 0.2, 0.5, 0.8$ and $\sigma = 1, 2, 5$

4.2 Continuous four-parameter distribution on $(0,1)$: Generalized beta type 1, GB1

The generalized beta type 1 distribution is defined by assuming

$$Z = \frac{Y^\tau}{\nu + (1-\nu)Y^\tau} \sim \texttt{BE}(\mu,\sigma)\ .$$

Hence, the pdf of generalized beta type 1 distribution, denoted by GB1(μ,σ,ν,τ), is given by

$$f_Y(y\,|\,\mu,\sigma,\nu,\tau) = \frac{\tau\nu^\beta y^{\tau\alpha-1}(1-y^\tau)^{\beta-1}}{B(\alpha,\beta)[\nu+(1-\nu)y^\tau]^{\alpha+\beta}} \tag{4.5}$$

for $0 < y < 1$, where $0 < \mu < 1$, $0 < \sigma < 1$, $\nu > 0$ and $\tau > 0$, and where α and β are defined as in (4.2), for $\alpha > 0$ and $\beta > 0$. Hence GB1(μ,σ,ν,τ) has adopted parameters $\mu = \alpha/(\alpha+\beta)$, $\sigma = (\alpha+\beta+1)^{-1/2}$, ν and τ.

For $0 < \nu < 1$, GB1(μ,σ,ν,τ) is a reparameterized submodel, with range $0 < Y < 1$, of the five parameter generalized beta (GB) distribution of McDonald and Xu [1995, equation (2.8)] where GB1$(\mu,\sigma,\nu,\tau) =$ GB$(\tau,\nu^{1/\tau},1-\nu,\alpha,\beta)$. (Note that GB1 is different from the generalized beta of the first kind of McDonald and Xu [1995].) The generalized three-parameter beta (GB3) distribution of Pham-Gia and Duong (1989) and Johnson et al. [1995, p251] is a reparameterized submodel of GB1 where $\tau = 1$, given by G3B$(\alpha_1,\alpha_2,\lambda) =$

GB1$(\alpha_1(\alpha_1+\alpha_2)^{-1}, (\alpha_1+\alpha_2+1)^{-1}, \lambda^{-1}, 1)$. The BE$(\mu,\sigma)$ distribution is a submodel of GB1(μ,σ,ν,τ) where $\nu = 1$ and $\tau = 1$, i.e. BE(μ,σ) = GB1$(\mu,\sigma,1,1)$.

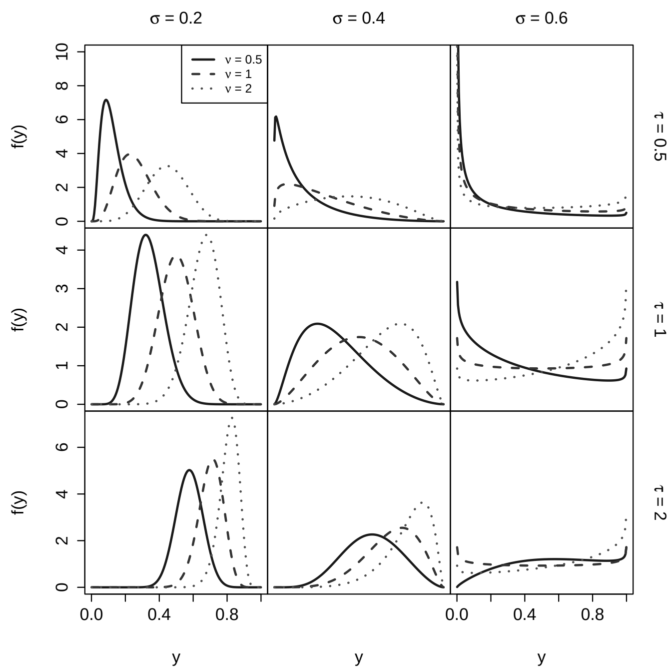

Figure 4.5: The generalized beta type 1 distribution, GB1(μ,σ,ν,τ), with $\mu = .5$, $\sigma = 0.2, 0.4, 0.6$, $\nu = 0.5, 1, 2$, and $\tau = 0.5, 1, 2$

4.3 Inflated distributions on [0,1), (0,1], and [0,1]

The three types of inflated distributions are shown in Figure 4.6, which is identical to Figure **??**.

```
## Warning in par(cex.axis = 2, lwd = 2, cex.lab = 2, mgp = c(2.5,
1, 0), par(mar = c(5, : argument 5 does not name a graphical parameter
```

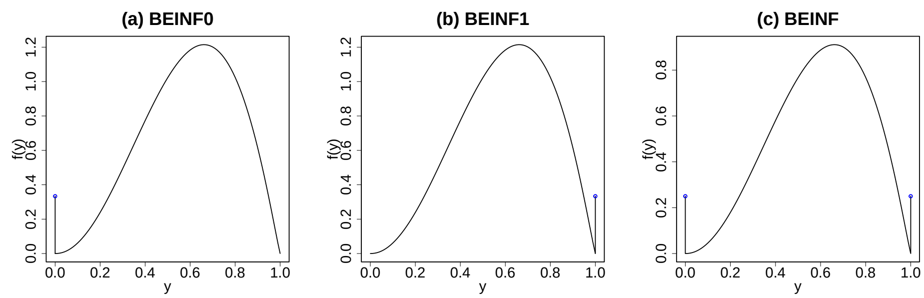

Figure 4.6: Three types of inflated distributions: (a) zero-inflated beta, BEINF0(μ, σ, ν), with $\mu = 0.6$, $\sigma = 0.4$, $\nu = 0.5$, (b) one-inflated beta, BEINF1(μ, σ, ν), with $\mu = 0.6$, $\sigma = 0.4$, $\nu = 0.5$, (c) zero- and one-inflated beta, BEINF(μ, σ, ν, τ), with $\mu = 0.6$, $\sigma = 0.4$, $\nu = 0.5$, $\tau = 0.5$.

4.3.1 Zero- and one-inflated beta: BEINF

The zero- and one-inflated beta distribution is appropriate when the response variable takes values in a known restricted range including both the endpoints of the range. Appropriate standardization can be applied to make the range of the response variable [0,1], i.e. from zero to one, including both the endpoints, i.e. exact values 0 and 1. Values zero and one have nonzero probabilities p_0 and p_1 respectively. Informally, the mixed probability function of the zero- and one-inflated beta distribution, denoted by BEINF(μ, σ, ν, τ), is given by

$$f_Y(y \mid \mu, \sigma, \nu, \tau) = \begin{cases} p_0 & \text{if } y = 0 \\ (1 - p_0 - p_1) \frac{1}{B(\alpha,\beta)}\, y^{\alpha-1}(1-y)^{\beta-1} & \text{if } 0 < y < 1 \\ p_1 & \text{if } y = 1 \end{cases} \tag{4.6}$$

for $0 \le y \le 1$, where α and β are given by (4.2), $p_0 = \nu(1 + \nu + \tau)$ and $p_1 = \tau/(1 + \nu + \tau)$ where $\alpha > 0$, $\beta > 0$, $0 < p_0 < 1$, $0 < p_1 < 1$ and $0 < p_0 + p_1 < 1$.

Hence BEINF(μ, σ, ν, τ) has parameters

$$\begin{aligned} \mu &= \alpha/(\alpha + \beta) \\ \sigma &= (\alpha + \beta + 1)^{-1/2} \\ \nu &= \frac{p_0}{1 - p_0 - p_1} \\ \tau &= \frac{p_1}{1 - p_0 - p_1}\ , \end{aligned}$$

and $0 < \mu < 1$, $0 < \sigma < 1$, $\nu > 0$ and $\tau > 0$. Note that $\mathrm{E}(Y) = (\mu+\tau)/(1+\nu+\tau)$.

Table 4.3: The zero- and one-inflated beta distribution.

$\texttt{BEINF}(\mu,\sigma,\nu,\tau)$	
Ranges	
Y	$0 \le y \le 1$
μ	$0 < \mu < 1$, mean of $\texttt{BE}(\mu,\sigma)$ component
σ	$0 < \sigma < 1$
ν	$0 < \nu < \infty$
τ	$0 < \tau < \infty$
Distribution measures	
mean	$(\mu+\tau)(1+\nu+\tau)^{-1}$
variance	$\dfrac{\sigma^2\mu(1-\mu)+\mu^2+\tau+(\mu+\tau)^2(1+\nu+\tau)^{-1}}{1+\nu+\tau}$
pdf	$\begin{cases} p_0 & \text{if } y=0 \\ (1-p_0-p_1)\dfrac{1}{B(\alpha,\beta)}y^{\alpha-1}(1-y)^{\beta-1} & \text{if } 0<y<1 \\ p_1 & \text{if } y=1 \\ \text{where} & \\ \alpha=\mu(1-\sigma^2)/\sigma^2,\ \beta=(1-\mu)(1-\sigma^2)/\sigma^2, & \\ p_0=\nu(1+\nu+\tau)^{-1}, & \\ p_1=\tau(1+\nu+\tau)^{-1} & \end{cases}$
cdf	$\begin{cases} p_0 & \text{if } y=0 \\ p_0+\dfrac{(1-p_0-p_1)B(\alpha,\beta,y)}{B(\alpha,\beta)} & \text{if } 0<y<1 \\ 1 & \text{if } y=1 \end{cases}$

4.3.2 Zero-inflated beta: BEINF0

Informally, the mixed probability function of the zero-inflated beta distribution, denoted by $\texttt{BEINF0}(\mu,\sigma,\nu)$, is given by

$$f_Y(y\,|\,\mu,\sigma,\nu)=\begin{cases} p_0 & \text{if } y=0 \\ (1-p_0)\frac{1}{B(\alpha,\beta)}\,y^{\alpha-1}(1-y)^{\beta-1} & \text{if } 0<y<1 \end{cases} \tag{4.7}$$

for $0 \le y < 1$, where α and β are given by (4.2) and $p_0 = \nu/(1+\nu)$, where $\alpha > 0$, $\beta > 0$, $0 < p_0 < 1$. Hence $\texttt{BEINF0}(\mu,\sigma,\nu)$ has parameters

$$\mu = \alpha/(\alpha+\beta)$$
$$\sigma = (\alpha+\beta+1)^{-1/2}$$
$$\nu = \frac{p_0}{1-p_0}\ ,$$

and $0 < \mu < 1$, $0 < \sigma < 1$ and $\nu > 0$. Note that for `BEINF0`(μ, σ, ν), $\mathrm{E}(Y) = \mu/(1+\nu)$.

Table 4.4: The zero-inflated beta distribution.

`BEINF0`(μ,σ,ν)	
Ranges	
Y	$0 \le y < 1$
μ	$0 < \mu < 1$, mean of `BE`(μ,σ) component
σ	$0 < \sigma < 1$
ν	$0 < \nu < \infty$
Distribution measures	
mean	$\dfrac{\mu}{1+\nu}$
variance	$\dfrac{\sigma^2\mu(1-\mu)+\mu^2+\mu^2(1+\nu)^{-1}}{1+\nu}$
pdf	$\begin{cases} p_0 & \text{if } y=0 \\ (1-p_0)\dfrac{1}{B(\alpha,\beta)}y^{\alpha-1}(1-y)^{\beta-1} & \text{if } 0<y<1 \\ \text{where} & \\ \alpha = \dfrac{\mu(1-\sigma^2)}{\sigma^2},\ \beta = \dfrac{(1-\mu)(1-\sigma^2)}{\sigma^2}, & \\ p_0 = \nu(1+\nu)^{-1} & \end{cases}$
cdf	$\begin{cases} p_0 & \text{if } y=0 \\ p_0 + \dfrac{(1-p_0)B(\alpha,\beta,y)}{B(\alpha,\beta)} & \text{if } 0<y<1 \\ 1 & \text{if } y=1 \end{cases}$
Reference	Set $\tau = 0$ in `BEINF`(μ,σ,ν,τ)

4.3.3 One-inflated beta: `BEINF1`

Informally, the mixed probability function of the one-inflated beta distribution, denoted by `BEINF1`(μ, σ, ν), is given by

$$f_Y(y \mid \mu, \sigma, \nu) = \begin{cases} (1-p_1)\frac{1}{B(\alpha,\beta)}\, y^{\alpha-1}(1-y)^{\beta-1} & \text{if } 0<y<1 \\ p_1 & \text{if } y=1 \end{cases} \tag{4.8}$$

for $0 < y \leq 1$, where α and β are given by (4.2), $p_1 = \nu/(1+\nu)$, where $\alpha > 0$, $\beta > 0$ and $0 < p_1 < 1$. Hence `BEINF1`(μ, σ, ν) has parameters

$$\mu = \alpha/(\alpha+\beta)$$
$$\sigma = (\alpha+\beta+1)^{-1/2}$$
$$\nu = \frac{p_1}{1-p_1}\ ,$$

and $0 < \mu < 1$, $0 < \sigma < 1$ and $\nu > 0$. Note that $\mathrm{E}(Y) = (\mu+\nu)/(1+\nu)$.

Table 4.5: The one-inflated beta distribution.

`BEINF1`(μ, σ, ν)	
Ranges	
Y	$0 < y \leq 1$
μ	$0 < \mu < 1$, mean of `BE`(μ, σ) component
σ	$0 < \sigma < 1$
ν	$0 < \nu < \infty$
Distribution measures	
mean	$\frac{\mu+\nu}{1+\nu}$
variance	$\frac{\sigma^2\mu(1-\mu)+\mu^2+\nu+(\mu+\nu)^2(1+\nu)^{-1}}{1+\nu}$
pdf	$\begin{cases} (1-p_1)\frac{1}{B(\alpha,\beta)}y^{\alpha-1}(1-y)^{\beta-1} & \text{if } 0<y<1 \\ p_1 & \text{if } y=1 \\ \text{where} & \\ \alpha = \frac{\mu(1-\sigma^2)}{\sigma^2},\ \beta = \frac{(1-\mu)(1-\sigma^2)}{\sigma^2} & \\ p_1 = \nu(1+\nu)^{-1} & \end{cases}$
cdf	$\begin{cases} \frac{(1-p_1)B(\alpha,\beta,y)}{B(\alpha,\beta)} & \text{if } 0<y<1 \\ 1 & \text{if } y=1 \end{cases}$
Reference	Set $\nu = 0$ and then $\tau = \nu$ in `BEINF`(μ, σ, ν, τ)

`BEZI`(μ, σ, ν) and `BEOI`(μ, σ, ν) are different parameterizations of the `BEINF0`(μ, σ, ν) and `BEINF1`(μ, σ, ν) distributions contributed to **gamlss.dist** by Raydonal Ospina, see Ospina and Ferrari [2010] and Section **??**.

Chapter 5

Discrete count distributions

This chapter gives summary tables and plots for explicit **gamlss.dist** discrete count distributions. All the distributions in this chapter have range $\{0, 1, 2, 3, \ldots\}$, with the exception of the logarithmic (`LG`(μ)) and the Zipf (`ZIPF`(μ)) distributions, which have range $\{1, 2, 3, \ldots\}$. Discrete count distributions are discussed in Chapter **??**. Section **??** discusses creating discrete count distributions by discretizing any continuous `gamlss.family` distribution defined on $(0, \infty)$.

5.1 One-parameter count distributions

5.1.1 Geometric: GEOM, GEOMo

There are two parameterizations of the geometric distribution in the **gamlss.dist** package: `GEOM`(μ) and `GEOMo`(μ).

First parameterization, GEOM

The probability function (pf) of the geometric distribution, `GEOM`(μ), is given by

$$P(Y = y \mid \mu) = \frac{\mu^y}{(\mu + 1)^{y+1}} \tag{5.1}$$

for $y = 0, 1, 2, 3, \ldots$, where $\mu > 0$. Figure 5.1 shows the shapes of the `GEOM`(μ) distribution for $\mu = 1, 3, 5$. The mode of Y is always at zero, and the probabilities always decrease as Y increases. For large values of y (and all y), $P(Y = y \mid \mu) = q \exp\left\{-y \log\left(1 + \mu^{-1}\right)\right\}$ where $q = (\mu + 1)^{-1}$, i.e. an exponential right tail.

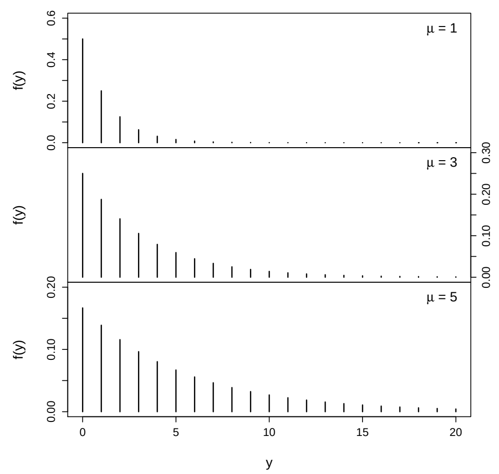

Figure 5.1: The geometric, GEOM(μ), distribution, with $\mu = 1, 3, 5$.

Table 5.1: Geometric distribution.

	GEOM(μ)
	Ranges
Y	$0, 1, 2, 3 \ldots$
μ	$0 < \mu < \infty$, mean
	Distribution measures
mean	μ
median	$\begin{cases} \lfloor \alpha \rfloor & \text{if } \alpha \text{ is not an integer} \\ \alpha - 1 & \text{if } \alpha \text{ is an integer} \end{cases}$ where $\alpha = \dfrac{-\log 2}{\log[\,\mu\,(\mu+1)^{-1}]}$
mode	0
variance	$\mu + \mu^2$
skewness	$(1 + 2\mu)\left(\mu + \mu^2\right)^{-0.5}$
excess kurtosis	$6 + (\mu + \mu^2)^{-1}$
PGF	$[1 + \mu(1 - t)]^{-1}$ for $0 < t < 1 + \mu^{-1}$
pf	$\dfrac{\mu^y}{(\mu + 1)^{y+1}}$
cdf	$1 - \left(\frac{\mu}{\mu+1}\right)^{y+1}$
inverse cdf	$\begin{cases} \lfloor \alpha_p \rfloor & \text{if } \alpha_p \text{ is not an integer} \\ \alpha_p - 1 & \text{if } \alpha_p \text{ is an integer} \end{cases}$ where $\alpha_p = \dfrac{\log(1-p)}{\log[\,\mu(\mu+1)^{-1}]}$
Reference	Set $\sigma = 1$ in NBI(μ, σ)
Notes	$\lfloor \alpha \rfloor$ is the largest integer less than or equal to α, i.e. the floor function.

Second parameterization, GEOMo

For the original parameterization of the geometric distribution, GEOMo(μ), replace μ in (5.1) by $(1 - \mu)/\mu$, giving

$$P(Y = y \,|\, \mu) = (1 - \mu)^y \mu$$

for $y = 0, 1, 2, 3, \ldots$, and $0 < \mu < 1$. Hence μ in this case is $P(Y = 0)$. Other characteristics of the GEOMo(μ) distribution are given in Table 5.2.

Table 5.2: Geometric distribution (original).

GEOMo(μ)	
Ranges	
Y	$0, 1, 2, 3 \dots$
μ	$0 < \mu < 1,$
Distribution measures	
mean	$(1-\mu)\mu^{-1}$
median	$\begin{cases} \lfloor \alpha \rfloor, & \text{if } \alpha \text{ is not an integer} \\ \alpha - 1, & \text{if } \alpha \text{ is an integer} \end{cases}$ where $\alpha = \dfrac{-\log 2}{\log(1-\mu)}$
mode	0
variance	$(1-\mu)\,\mu^{-2}$
skewness	$(2-\mu)\,(1-\mu)^{-0.5}$
excess kurtosis	$6 + \mu^2\,(1-\mu)^{-1}$
PGF	$\mu\left[1-(1-\mu)t\right]^{-1}$ for $0 < t < (1-\mu)^{-1}$
pf	$\mu(1-\mu)^y$
cdf	$1-(1-\mu)^{y+1}$
inverse cdf	$\begin{cases} \lfloor \alpha_p \rfloor, & \text{if } \alpha_p \text{ is not an integer} \\ \alpha_p - 1, & \text{if } \alpha_p \text{ is an integer} \end{cases}$ where $\alpha_p = \dfrac{\log(1-p)}{\log(1-\mu)}$
Reference	Set μ_1 to $(1-\mu)\mu^{-1}$ in GEOM(μ_1)

5.1.2 Logarithmic: LG

The probability function of the logarithmic distribution, denoted by LG(μ), is given by

$$P(Y = y \mid \mu) = \frac{\alpha \mu^y}{y} \tag{5.2}$$

for $y = 1, 2, 3, \ldots$, where $\alpha = -\left[\log(1-\mu)\right]^{-1}$ for $0 < \mu < 1$. Note that the range of Y starts from 1. As $y \to \infty$ (and all y), $P(Y = y \mid \mu)$ is proportional to $\exp\{y \log \mu - \log y\}$. Hence $\log P(Y = y \mid \mu) \sim y \log \mu$ as $y \to \infty$. The mode of Y is always at 1 and the probabilities always decrease as Y increases.

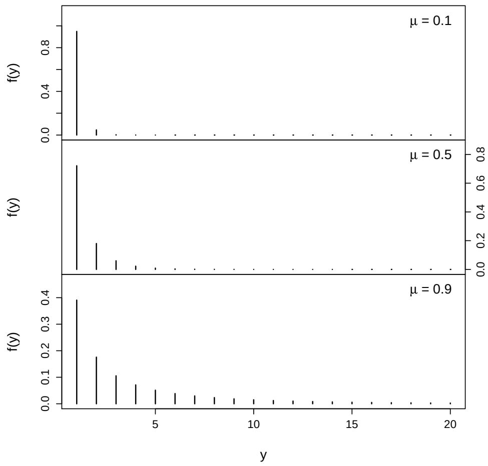

Figure 5.2: The logarithmic, $\texttt{LG}(\mu)$, distribution, with $\mu = 0.1, 0.5, 0.9$.

Table 5.3: Logarithmic distribution.

	$\texttt{LG}(\mu)$
	Ranges
Y	$1, 2, 3 \ldots$
μ	$0 < \mu < 1$
	Distribution measures
mean[a]	$\alpha\mu(1-\mu)^{-1}$ where $\alpha = -[\log(1-\mu)]^{-1}$
mode[a]	1
variance[a]	$\alpha\mu(1-\alpha\mu)(1-\mu)^{-2}$
skewness[a]	$\begin{cases} \mu_3/[\text{Var}(Y)]^{1.5} \text{ where} \\ \mu_3 = \alpha\mu(1+\mu-3\alpha\mu+2\alpha^2\mu^2)(1-\mu)^{-3} \end{cases}$
excess kurtosis	$\begin{cases} k_4/[\text{Var}(Y)]^2 \text{ where} \\ k_4 = \alpha\mu[1+4\mu+\mu^2-\alpha\mu(7+4\mu)+12\alpha^2\mu^2-6\alpha^3\mu^3](1-\mu)^{-4} \end{cases}$
PGF[a]	$\dfrac{\log(1-\mu t)}{\log(1-\mu)}$ for $0 < t < \mu^{-1}$
pf [a]	$\dfrac{\alpha\mu^y}{y}$
Reference	[a] Johnson et al. [2005], Section 7.1, p302-307, parameterized by $\theta = \mu$

5.1.3 Poisson: PO

The probability function of the Poisson distribution, denoted by $\texttt{PO}(\mu)$, is given by

$$P(Y = y \mid \mu) \quad = \quad \frac{e^{-\mu}\mu^y}{y!} \tag{5.3}$$

for $y = 0, 1, 2, 3, \ldots$, where $\mu > 0$. Properties of the distribution are given in Table 5.4. Note in particular that $E(Y) = \text{Var}(Y) = \mu$, and hence the index of dispersion $\text{Var}(Y)/E(Y)$ is equal to one. For distributions with $\text{Var}(Y) > E(Y)$ we have overdispersion relative to the Poisson distribution, and for those with $\text{Var}(Y) < E(Y)$ we have underdispersion. The distribution is positively skew for small values of μ, but almost symmetric for large μ values. As $y \to \infty$, $P(Y = y) \sim q \exp[y(\log \mu + 1) - (y + 0.5)\log y]$ where $q = e^{-\mu}\sqrt{2\pi}$, (using Stirling's approximation to $y!$), which decreases faster than an exponential $\exp(-y)$. Note that $\log P(Y = y) \sim -y \log y$ as $y \to \infty$.

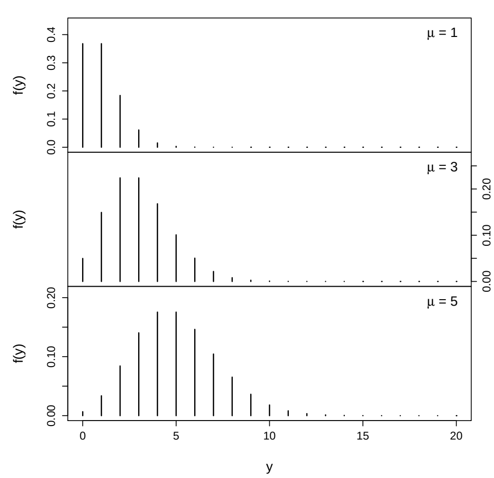

Figure 5.3: The Poisson, $\texttt{PO}(\mu)$, distribution, with $\mu = 1, 3, 5$.

Table 5.4: Poisson distribution.

PO(μ)	
Ranges	
Y	$0, 1, 2, 3 \ldots$
μ	$0 < \mu < \infty,$ mean
Distribution measures	
mean	μ
mode	$\begin{cases} \lfloor \mu \rfloor & \text{if } \mu \text{ is not an integer} \\ \mu - 1 \text{ and } \mu & \text{if } \mu \text{ is an integer} \end{cases}$
variance	μ
skewness	$\mu^{-0.5}$
excess kurtosis	μ^{-1}
PGF	$e^{\mu(t-1)}$
pf	$\dfrac{e^{-\mu}\mu^y}{y!}$
cdf	$\dfrac{\Gamma(y+1,\mu)}{\Gamma(y)}$
Reference	Johnson et al. [2005] Sections 4.1, 4.3, 4.4, p156, p161-165, p307
Note	$\lfloor \alpha \rfloor$ is the largest integer less than or equal to α $\Gamma(\alpha, x) = \int_x^\infty t^{\alpha-1}e^{-t}dt$ is the complement of the incomplete gamma function

5.1.4 Yule: YULE

The probability function of the Yule distribution, denoted by YULE(μ), is given by

$$P(Y = y \mid \mu) = (\mu^{-1} + 1)B(y + 1, \mu^{-1} + 2) \tag{5.4}$$

for $y = 0, 1, 2, 3, \ldots$, where $\mu > 0$. Note that this parameterization only includes Yule distributions with a finite mean μ. The mode is always at zero, and the probabilities always decrease with increasing Y. The YULE(μ) distribution is a special case of the WARING(μ, σ) where $\mu = \sigma$. It can be shown (using equation (1.32) of Johnson et al. [2005, p8]) that as $y \to \infty$, $P(Y = y \mid \mu) \sim qy^{-(\mu^{-1}+2)}$ where $q = (\mu^{-1} + 1)\Gamma(\mu^{-1} + 2)$. Hence the YULE($\mu$) distribution has a heavy tail, especially for large μ .

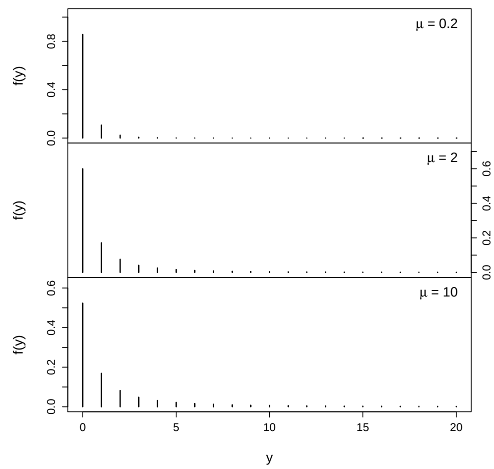

Figure 5.4: The Yule, YULE(μ), distribution, with $\mu = 0.2, 2, 10$.

Table 5.5: Yule distribution.

	YULE(μ)
	Ranges
Y	$0, 1, 2, 3 \ldots$
μ	$0 < \mu < \infty$, mean
	Distribution measures
mean	μ
mode	0
variance	$\begin{cases} \mu(\mu+1)^2(1-\mu)^{-1} & \text{if } \mu < 1 \\ \infty & \text{if } \mu \geq 1 \end{cases}$
skewness	$\begin{cases} (2\mu+1)^2(1-\mu)^{0.5}(\mu+1)^{-1}(1-2\mu)^{-1}\mu^{-0.5} & \text{if } \mu < 1/2 \\ \infty & \text{if } \mu \geq 1/2 \end{cases}$
excess kurtosis	$\begin{cases} \dfrac{\left(1+11\mu+18\mu^2-6\mu^3-36\mu^4\right)}{\mu(\mu+1)(1-2\mu)(1-3\mu)} & \text{if } \mu < 1/3 \\ \infty & \text{if } \mu \geq 1/3 \end{cases}$
PGF	$(\mu+1)(2\mu+1)^{-1}{}_2F_1(1,1;3+\mu^{-1};t)$
pf	$(\mu^{-1}+1)B(y+1, \mu^{-1}+2)$
cdf	$1-(y+1)B(y+1, 2+\mu^{-1})$
Reference	Set $\sigma = \mu$ in WARING(μ, σ)

5.1.5 Zipf: ZIPF

The probability function of the Zipf distribution, denoted by ZIPF(μ), is given by

$$P(Y = y \mid \mu) = \frac{y^{-(\mu+1)}}{\zeta(\mu+1)} \tag{5.5}$$

for $y = 1, 2, 3, \ldots$, where $\mu > 0$ and $\zeta(b) = \Sigma_{i=1}^{\infty} i^{-b}$ is the Riemann zeta function. Note that the range of Y starts from 1. The mode of Y is always at 1, and the probabilities always decrease with increasing Y.

This distribution is also known as the Riemann zeta distribution or the discrete Pareto distribution. as $y \to \infty$ (and all y), $P(Y = y \mid \mu) = qy^{-(\mu+1)}$ where $q = [\zeta(\mu+1)]^{-1}$, so the ZIPF(μ) distribution has a very heavy tail especially for μ close to zero. It is suitable for very heavy-tailed count data, as can be seen from Figure 5.5 where it is plotted for $\mu = 0.1, 0.5, 1$.

The rth raw moment is $E(Y^r) = \zeta(\mu - r + 1)/\zeta(\mu + 1)$, provided that $\mu > r$, [Johnson et al., 2005, p528]. Hence this gives the mean, variance, skewness, and excess kurtosis of Y. The mean increases as μ decreases and is infinite if $\mu \leq 1$.

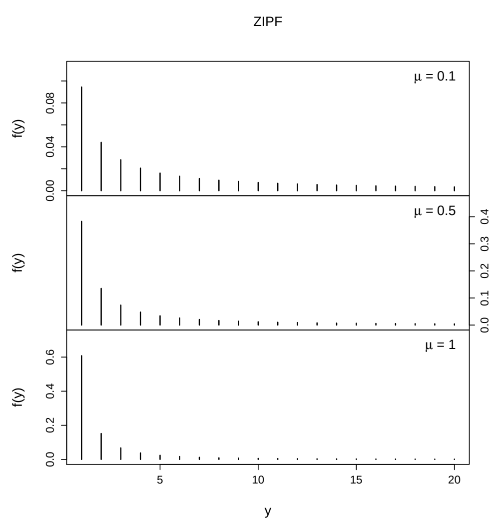

Figure 5.5: The ZIPF(μ) distribution, with $\mu = 0.1, 0.5, 1$.

Table 5.6: ZIPF distribution.

	$\texttt{ZIPF}(\mu)$
	Ranges
Y	$1, 2, 3 \ldots$
μ	$0 < \mu < \infty$
	Distribution measures
mean	$\begin{cases} \zeta(\mu)/\zeta(\mu+1) & \text{if } \mu > 1 \\ \infty & \text{if } \mu \le 1 \end{cases}$
mode	1
variance	$\begin{cases} \left\{\zeta(\mu+1)\zeta(\mu-1) - [\zeta(\mu)]^2\right\}/[\zeta(\mu+1)]^2 & \text{if } \mu > 2 \\ \infty & \text{if } \mu \le 2 \end{cases}$
skewness	$\begin{cases} \mu_3/[\text{Var}(Y)]^{1.5} \quad \text{where} & \\ \mu_3 = \{[\zeta(\mu+1)]^2\zeta(\mu-2) - 3\zeta(\mu+1)\zeta(\mu)\zeta(\mu-1) + & \\ \qquad 2[\zeta(\mu)]^3\}/[\zeta(\mu+1)]^3 & \text{if } \mu > 3 \\ \infty & \text{if } \mu \le 3 \end{cases}$
excess kurtosis	$\begin{cases} \{\mu_4/[\text{Var}(Y)]^2\} - 3 \quad \text{where} & \\ \mu_4 = \{[\zeta(\mu+1)]^3\zeta(\mu-3) - 4[\zeta(\mu+1)]^2\zeta(\mu)\zeta(\mu-2) + & \\ \qquad 6\zeta(\mu+1)[\zeta(\mu)]^2\zeta(\mu-1) - 3[\zeta(\mu)]^4\}/[\zeta(\mu+1)]^4 & \text{if } \mu > 4 \\ \infty & \text{if } \mu \le 4 \end{cases}$
PGF[a]	$t\phi(t, \mu+1, 1)/\phi(1, \mu+1, 1)$
pf [a]	$[y^{(\mu+1)}\zeta(\mu+1)]^{-1}$
Reference	[a]Johnson et al. [2005] Section 11.2.20, p527-528, where $\rho = \mu$
Notes	$\zeta(b) = \sum_{i=1}^{\infty} i^{-b}$ is the Riemann zeta function
	$\phi(a, b, c) = \sum_{i=0}^{\infty} \frac{a^i}{(i+c)^b}$, for $c \ne 0, -1, -2$, is the Lerch function

5.2 Two-parameter count distributions

5.2.1 Double Poisson: DPO

The double Poisson distribution, denoted by $\texttt{DPO}(\mu, \sigma)$, has probability function

$$P\left(Y = y \,|\, \mu, \sigma\right) = c(\mu, \sigma)\sigma^{-1/2} e^{-\mu/\sigma} \left(\frac{\mu}{y}\right)^{y/\sigma} \frac{e^{y/\sigma - y} y^y}{y!} \tag{5.6}$$

for $y = 0, 1, 2, 3, \ldots$, where $\mu > 0$, $\sigma > 0$ and $c(\mu, \sigma)$ is a normalizing constant (ensuring that the distribution probabilities sum to one) given by

$$c(\mu, \sigma) = \left[\sum_{y=0}^{\infty} \sigma^{-1/2} e^{-\mu/\sigma} \left(\frac{\mu}{y} \right)^{y/\sigma} \frac{e^{y/\sigma - y} y^y}{y!} \right]^{-1} \tag{5.7}$$

(Lindsey [1995, p131], reparameterized by $\nu = \mu$ and $\psi = 1/\sigma$).

The double Poisson distribution is a special case of the double exponential family of Efron [1986]. It has approximate mean μ and approximate variance $\sigma\mu$. The `DPO`(μ, σ) distribution is a `PO`(μ) distribution if $\sigma = 1$. It is (approximately) an overdispersed Poisson distribution if $\sigma > 1$ and underdispersed Poisson if $\sigma < 1$. Unlike some other implementations, **gamlss.dist** approximates $c(\mu, \sigma)$ using a finite sum with a very large number of terms ($3 \times \max(y)$) in equation (5.7), rather than a potentially less accurate functional approximation of $c(\mu, \sigma)$. As $y \to \infty$, $\log \mathrm{P}(Y = y \mid \mu, \sigma) \sim -(y \log y)/\sigma$.

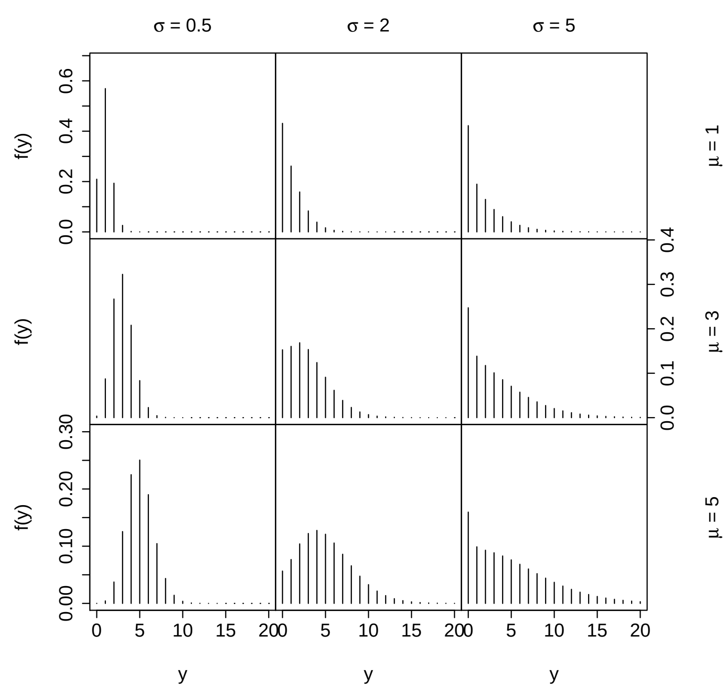

Figure 5.6: The double Poisson, `DPO`(μ, σ), distribution with $\mu = 1, 3, 5$ and $\sigma = .5, 2, 5$.

5.2.2 Generalized Poisson: `GPO`

The probability function of the generalized Poisson distribution [Consul and Jain, 1973], [Consul, 1989], [Poortema, 1999], [Johnson et al., 2005, p336], is given by

$$\mathrm{P}(Y = y \mid \theta, \lambda) = \frac{\theta(\theta + \lambda y)^{y-1} \, e^{-(\theta+\lambda y)}}{y!} \tag{5.8}$$

for $y = 0, 1, 2, ...$ where $\theta > 0$ and $0 \leq \lambda \leq 1$. The generalized Poisson distribution was derived by Consul and Jain [1973] as an approximation of a generalized negative binomial distribution. In recognition of P. C. Consul's contribution to this distribution, it is sometimes called *Consul's generalized Poisson distribution.* The mean and the variance of the generalized Poisson distribution are given by $\mathrm{E}(Y) = \theta/(1-\lambda)$ and $\mathrm{Var}(Y) = \theta/(1-\lambda)^3$, [Johnson et al., 2005, p337-338].

In **gamlss.dist**, a parameterization of the generalized Poisson distribution which is more suitable for regression modeling is used. This is given by reparameterizing (5.8) by setting

$$\mu = \frac{\theta}{1-\lambda} \qquad \text{and} \qquad \sigma = \frac{\lambda}{\theta}$$

and hence

$$\theta = \frac{\mu}{1+\sigma\mu} \qquad \text{and} \qquad \lambda = \frac{\sigma\mu}{1+\sigma\mu}\ .$$

This gives the probability function, denoted by `GPO`(μ, σ):

$$\mathrm{P}(Y = y \mid \mu, \sigma) = \left(\frac{\mu}{1+\sigma\mu}\right)^{y} \frac{(1+\sigma y)^{y-1}}{y!} \exp\left[\frac{-\mu\,(1+\sigma y)}{1+\sigma\mu}\right] \tag{5.9}$$

for $y = 0, 1, 2, ...,$ where $\mu > 0$ and $\sigma > 0$. The mean and variance of this version of the distribution are $\mathrm{E}(Y) = \mu$ and $\mathrm{Var}(Y) = \mu(1+\sigma\mu)^2$, which is overdispersed Poisson since $\sigma > 0$. As $y \to \infty$, $\log \mathrm{P}(Y = y \mid \mu, \sigma) \sim -y\left\{\log\left[1+(\mu\sigma)^{-1}\right] - (1+\mu\sigma)^{-1}\right\}$.

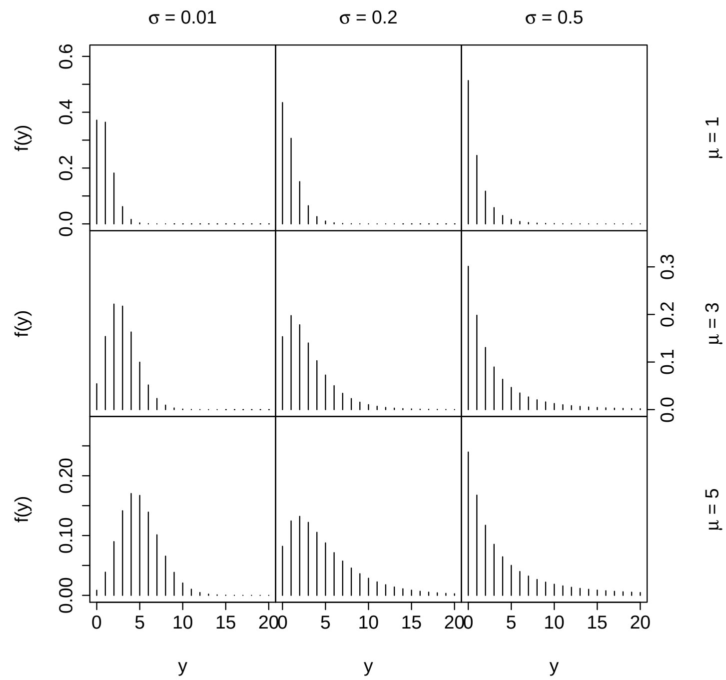

Figure 5.7: The generalized Poisson, `GPO`(μ, σ), distribution, with $\mu = 1, 3, 5$ and $\sigma = 0.01, 0.2, 0.5$.

Table 5.7: Generalized Poisson distribution.

`GPO`(μ,σ)	
Ranges	
Y	$0, 1, 2, 3 \ldots$
μ	$0 < \mu < \infty$, mean
σ	$0 < \sigma < \infty$, dispersion parameter
Distribution measures	
mean	μ
variance	$\mu\,(1+\sigma\mu)^2$
skewness	$(1+3\sigma\mu)\,/\mu^{0.5}$
excess kurtosis	$\left(1+10\sigma\mu+15\sigma^2\mu^2\right)/\mu$
pf	$\left(\dfrac{\mu}{1+\sigma\mu}\right)^y \dfrac{(1+\sigma y)^{y-1}}{y!} \exp\left[\dfrac{-\mu\,(1+\sigma y)}{1+\sigma\mu}\right]$
Reference	Johnson et al. [2005] p336-339 reparameterized by $\theta = \mu/(1+\sigma\mu)$ and $\lambda = \sigma\mu/(1+\sigma\mu)$ and hence $\mu = \theta/(1-\lambda)$ and $\sigma = \lambda/\theta$

5.2.3 Negative binomial: `NBI`, `NBII`

Negative binomial type I, `NBI` The probability function of the negative binomial distribution type I, denoted by `NBI`(μ,σ), is given by

$$P(Y = y \,|\, \mu, \sigma) = \frac{\Gamma(y+\frac{1}{\sigma})}{\Gamma(\frac{1}{\sigma})\Gamma(y+1)} \left(\frac{\sigma\mu}{1+\sigma\mu}\right)^y \left(\frac{1}{1+\sigma\mu}\right)^{1/\sigma}$$

for $y = 0, 1, 2, 3 \ldots$, where $\mu > 0$, $\sigma > 0$. The mean and variance are $E(Y) = \mu$ and $\mathrm{Var}(Y) = \mu + \sigma\mu^2$. The above parameterization is equivalent to that used by Anscombe [1950], with the exception that he used $\alpha = 1/\sigma$, [Johnson et al., 2005, p209]. The Poisson, `PO`(μ), distribution is the limiting case of `NBI`(μ,σ) as $\sigma \to 0$.

[Note that the original parameterization of the negative binomial distribution, `NBo`(k,π), is given by setting $\mu = k(1-\pi)/\pi$ and $\sigma = 1/k$, giving

$$P(Y = y \,|\, \pi) = \frac{\Gamma(y+k)}{\Gamma(y+1)\Gamma(k)} (1-\pi)^y \pi^k$$

for $y = 0, 1, 2, 3 \ldots$, where $k > 0$ and $0 < \pi < 1$, [Johnson et al., 2005, p209]. Hence $\pi = 1/(1+\mu\sigma)$ and $k = 1/\sigma$. When k is an integer, this version has the interpretation of being the distribution of the number of failures until the

kth success, in independent Bernoulli trials, where the probability of success in each trial is π. The $\texttt{NBo}(k, \pi)$ distribution is not available in **gamlss.dist**.]

As $y \to \infty$, $\mathrm{P}(Y = y \mid \mu, \sigma) \sim q \exp\left[-y \log\left(1 + \frac{1}{\mu\sigma}\right) + \left(\frac{1}{\sigma} - 1\right) \log y\right]$ where $q = \left[\Gamma(1/\sigma)(1 + \sigma\mu)^{1/\sigma}\right]^{-1}$, essentially an exponential tail. Note $\log \mathrm{P}(Y = y \mid \mu, \sigma) \sim -y \log\left(1 + \frac{1}{\mu\sigma}\right)$ as $y \to \infty$.

See $\texttt{NBII}(\mu, \sigma)$ below with variance $\mathrm{Var}(Y) = \mu + \sigma\mu$ for an alternative parameterization. Also see $\texttt{NBF}(\mu, \sigma, \nu)$ in Section 5.3.4 with variance $\mathrm{Var}(Y) = \mu + \sigma\mu^\nu$ for a family of reparameterizations of the $\texttt{NBI}(\mu, \sigma)$.

Table 5.8: Negative binomial type I distribution.

	$\texttt{NBI}(\mu, \sigma)$
	Ranges
Y	$0, 1, 2, 3 \ldots$
μ	$0 < \mu < \infty$, mean
σ	$0 < \sigma < \infty$, dispersion
	Distribution measures
mean	μ
mode	$\begin{cases} \lfloor(1-\sigma)\mu\rfloor & \text{if } (1-\sigma)\mu \text{ is not an integer and } \sigma < 1 \\ (1-\sigma)\mu - 1 \text{ and } (1-\sigma)\mu & \text{if } (1-\sigma)\mu \text{ is an integer and } \sigma < 1 \\ 0 & \text{if } \sigma \geq 1 \end{cases}$
variance	$\mu + \sigma\mu^2$
skewness	$(1 + 2\mu\sigma)(\mu + \sigma\mu^2)^{-0.5}$
excess kurtosis	$6\sigma + (\mu + \sigma\mu^2)^{-1}$
PGF	$[1 + \mu\sigma(1 - t)]^{-1/\sigma}$ for $0 < t < 1 + (\mu\sigma)^{-1}$
pf [a]	$\dfrac{\Gamma(y + \sigma^{-1})}{\Gamma(\sigma^{-1})\Gamma(y + 1)} \left(\dfrac{\sigma\mu}{1 + \sigma\mu}\right)^y \left(\dfrac{1}{1 + \sigma\mu}\right)^{1/\sigma}$
cdf	$1 - \dfrac{B(y + 1, \sigma^{-1}, \mu\sigma(1 + \mu\sigma)^{-1})}{B(y + 1, \sigma^{-1})}$
Reference	Johnson et al. [2005], Sections 5.1 to 5.5, p209-217, reparameterized by $p = 1/(1 + \mu\sigma)$ and $k = 1/\sigma$ and hence $\mu = k(1 - p)/p$ and $\sigma = 1/k$ [a] McCullagh and Nelder [1989] p 373 reparameterized by $\alpha = \mu\sigma$ and $k = 1/\sigma$.
Note	$\lfloor\alpha\rfloor$ is the largest integer less than or equal to α $B(\alpha, \beta, x) = \int_0^x t^{\alpha-1} (1 - t)^{\beta-1} \, dt$ is the incomplete beta function

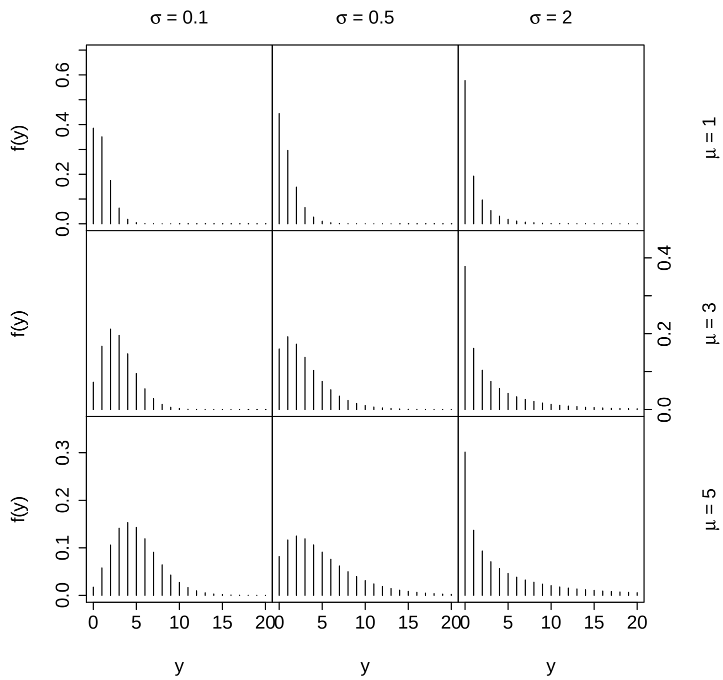

Figure 5.8: The negative binomial type I, `NBI`(μ, σ), distribution, with $\mu = 1, 3, 5$ and $\sigma = 0.1, 0.5, 2$.

Negative binomial type II, `NBII`

The probability function of the negative binomial type II distribution, denoted by `NBII`(μ, σ), is given by

$$P(Y = y \mid \mu, \sigma) = \frac{\Gamma(y + \frac{\mu}{\sigma})\sigma^y}{\Gamma(\frac{\mu}{\sigma})\Gamma(y+1)(1+\sigma)^{y+\mu/\sigma}}$$

for $y = 0, 1, 2, 3, \ldots$, where $\mu > 0$ and $\sigma > 0$.

This parameterization was used by Evans [1953], and is obtained by reparameterizing σ to σ/μ in `NBI`(μ, σ). The important difference between the `NBI` and `NBII` parameterizations is the variance-mean relationship: both distributions have $E(Y) = \mu$; in `NBI` the variance is $\text{Var}(Y) = \mu(1 + \sigma\mu)$, (i.e. the variance is quadratic in μ), and in `NBII` it is $\text{Var}(Y) = \mu(1 + \sigma)$, (i.e. the variance is linear in μ).

Table 5.9: Negative binomial Type II distribution.

	`NBII`(μ, σ)
	Ranges
Y	$0, 1, 2, 3 \ldots$
μ	$0 < \mu < \infty$, mean
σ	$0 < \sigma < \infty$, dispersion
	Distribution measures
mean	μ
mode	$\begin{cases} \lfloor \mu - \sigma \rfloor & \text{if } (\mu - \sigma) \text{ is not an integer and } \sigma < \mu \\ (\mu - \sigma - 1) \text{ and } (\mu - \sigma) & \text{if } (\mu - \sigma) \text{ is an integer and } \sigma < \mu \\ 0 & \text{if } \sigma \geq \mu \end{cases}$
variance	$\mu + \sigma\mu$
skewness	$(1 + 2\sigma)(\mu + \sigma\mu)^{-0.5}$
excess kurtosis	$6\sigma\mu^{-1} + (\mu + \sigma\mu)^{-1}$
PGF	$[1 + \sigma(1 - t)]^{-\mu/\sigma}$ for $0 < t < 1 + \sigma^{-1}$
pf	$\dfrac{\Gamma(y + \mu\sigma^{-1})}{\Gamma(\mu\sigma^{-1})\Gamma(y + 1)} \left(\dfrac{\sigma}{1 + \sigma}\right)^y \left(\dfrac{1}{1 + \sigma}\right)^{\mu/\sigma}$
cdf	$1 - \dfrac{B(y + 1, \mu\sigma^{-1}, \sigma(1 + \sigma)^{-1})}{B(y + 1, \mu\sigma^{-1})}$
Reference	Reparameterize σ_1 to σ/μ in `NBI`(μ, σ_1)

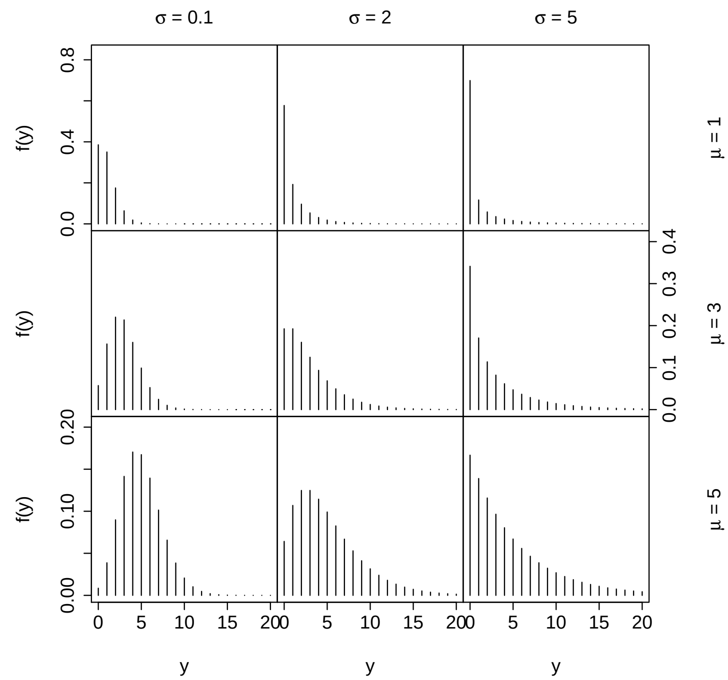

Figure 5.9: The negative binomial type II, NBII(μ, σ), distribution, with $\mu = 1, 3, 5$ and $\sigma = 0.1, 2, 5$.

5.2.4 Poisson-inverse Gaussian: PIG, PIG2

First parameterization, PIG

The probability function of the Poisson-inverse Gaussian distribution, denoted by PIG(μ, σ), is given by

$$P(Y = y \mid \mu, \sigma) = \left(\frac{2\alpha}{\pi}\right)^{1/2} \frac{\mu^y e^{1/\sigma} K_{y-\frac{1}{2}}(\alpha)}{y!(\alpha\sigma)^y} \tag{5.10}$$

for $y = 0, 1, 2, 3, \ldots$, where $\mu > 0$, $\sigma > 0$, $\alpha^2 = \sigma^{-2} + 2\mu\sigma^{-1}$ and $\alpha > 0$ and $K_\lambda(t)$ is the modified Bessel function of the second kind given by $K_\lambda(t) = \frac{1}{2}\int_0^\infty x^{\lambda-1} \exp\left[-\frac{1}{2}t(x + x^{-1})\right] dx$, [Abramowitz and Stegun, 1965, Section 9.6, p 374]. Note that $\sigma = \left[\left(\mu^2 + \alpha^2\right)^{0.5} - \mu\right]^{-1}$. This parameterization was used by Dean et al. [1989], and it is the special case of the **gamlss.dist** distributions SI(μ, σ, ν) and SICHEL (μ, σ, ν) when $\nu = -\frac{1}{2}$. The Poisson distribution is the limiting case of PIG(μ, σ) as $\sigma \to 0$.

The Bessel function $K_\lambda(t)$ in general presents challenges in computation; however, when the order λ of the Bessel function is a half-integer, as it is in (5.10), computations are simplified considerably as

$$K_{-1/2}(t) = K_{1/2}(t) = \sqrt{\frac{\pi}{2t}}\, e^{-t}$$

and use is made of the recurrence relation

$$K_{\lambda+1}(t) = \frac{2\lambda}{t} K_\lambda(t) + K_{\lambda-1}(t)\ .$$

As $y \to \infty$, $\mathrm{P}(Y = y \,|\, \mu, \sigma) \sim r \exp\left[-y \log\left(1 + \frac{1}{2\mu\sigma}\right) - \frac{3}{2}\log y\right]$, where r does not depend on y, i.e. essentially an exponential tail. Note that $\log \mathrm{P}(Y = y \,|\, \mu, \sigma) \sim -y \log\left(1 + \frac{1}{2\mu\sigma}\right)$ as $y \to \infty$.

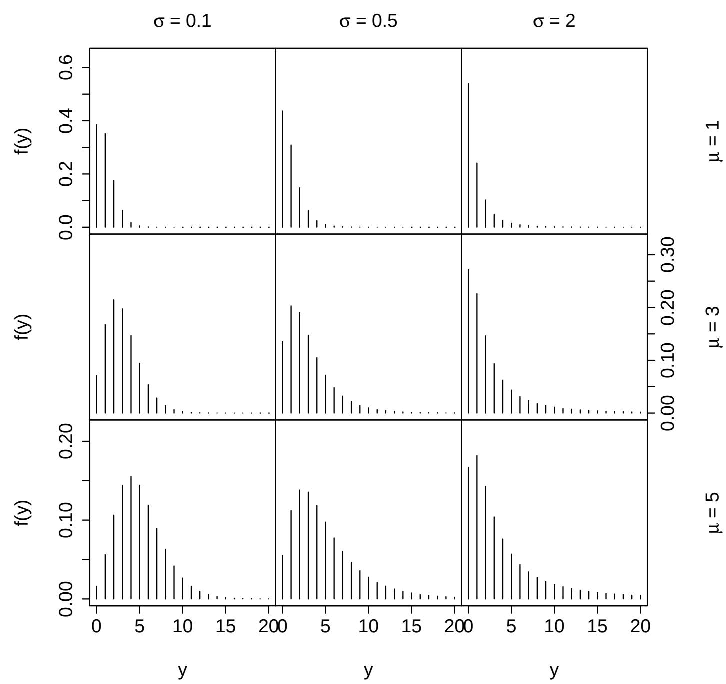

Figure 5.10: The Poisson-inverse Gaussian, $\texttt{PIG}(\mu, \sigma)$, distribution, with $\mu = 1, 3, 5$ and $\sigma = 0.1, 0.5, 2$.

,

Table 5.10: Poisson-inverse Gaussian distribution.

$\texttt{PIG}(\mu,\sigma)$	
Ranges	
Y	$0, 1, 2, 3 \ldots$
μ	$0 < \mu < \infty$, mean
σ	$0 < \sigma < \infty$, dispersion
Distribution measures	
mean	μ
variance	$\mu + \sigma\mu^2$
skewness	$(1 + 3\mu\sigma + 3\mu^2\sigma^2)(1 + \mu\sigma)^{-1.5}\mu^{-0.5}$
excess kurtosis	$(1 + 7\mu\sigma + 18\mu^2\sigma^2 + 15\mu^3\sigma^3)(1 + \mu\sigma)^{-2}\mu^{-1}$
PGF[a]	$e^{1/\sigma - q}$ where $q^2 = \sigma^{-2} + 2\mu(1 - t)\sigma^{-1}$
pf [a]	$\left(\frac{2\alpha}{\pi}\right)^{1/2} \frac{\mu^y e^{1/\sigma} K_{y-\frac{1}{2}}(\alpha)}{y!(\alpha\sigma)^y}$ where $\alpha^2 = \sigma^{-2} + 2\mu\sigma^{-1}$
Reference	Set $\nu = -\frac{1}{2}$ (and hence $c = 1$) in $\texttt{SICHEL}(\mu, \sigma, \nu)$, see also [a] Dean et al. [1989]

Second parameterization, `PIG2`

In the second parameterization of the Poisson-inverse Gaussian distribution, denoted $\texttt{PIG2}(\mu,\sigma)$, the probability function is given by

$$\mathrm{P}(Y = y \mid \mu, \sigma) = \left(\frac{2\sigma}{\pi}\right)^{1/2} \frac{\mu^y e^{1/\alpha} K_{y-\frac{1}{2}}(\sigma)}{y!(\sigma\alpha)^y} \tag{5.11}$$

for $y = 0, 1, 2, 3, \ldots$, where $\mu > 0$, $\sigma > 0$ and $\alpha = \left[(\mu^2 + \sigma^2)^{0.5} - \mu\right]^{-1}$. Note that $\sigma^2 = \alpha^{-2} + 2\mu\alpha^{-1}$. Currently under development in **gamlss.dist**.

This parameterization was given by Stein et al. [1987] and has the advantage that μ and σ are informationally orthogonal (and hence their maximum likelihood estimators are asymptotically uncorrelated). Heller et al. [2018] demonstrate that this results in (5.11) having more robust estimation of the model for the mean parameter μ, under misspecification of the model for σ. Note however that (5.11) has the disadvantage of a more complicated variance-mean relationship.

The `PIG2` parameterization (5.11) is given by interchanging σ and α in (5.10), i.e. in (5.10) σ is reparameterized to $[(\mu^2 + \sigma^2)^{0.5} - \mu]^{-1}$ giving

$$\texttt{PIG2}(\mu, \sigma) = \texttt{PIG}(\mu, [(\mu^2 + \sigma^2)^{0.5} - \mu]^{-1})$$

and

$$\texttt{PIG}(\mu,\sigma) = \texttt{PIG2}(\mu, \sqrt{\sigma^{-2} + 2\mu\sigma^{-1}}) \ .$$

The Poisson is the limiting distribution of $\texttt{PIG2}(\mu,\sigma)$ as $\sigma \to \infty$ and the variance of $\texttt{PIG2}(\mu,\sigma)$ increases with decreasing σ. As $y \to \infty$, $\mathrm{P}(Y = y \,|\, \mu,\sigma) \sim r \exp\left[-y \log\left(1 + \frac{1}{2\mu\alpha}\right) - \frac{3}{2}\log y\right]$, where r does not depend on y, i.e. essentially an exponential tail. Note that $\log P(Y = y|\mu,\sigma) \sim -y \log\left(1 + \frac{1}{2\mu\alpha}\right)$ as $y \to \infty$.

Table 5.11: Second parameterization of Poisson-inverse Gaussian distribution.

$\texttt{PIG2}(\mu,\sigma)$	
Ranges	
Y	$0, 1, 2, 3 \ldots$
μ	$0 < \mu < \infty$, mean
σ	$0 < \sigma < \infty$, dispersion
Distribution measures	
mean	μ
variance	$\mu + \alpha\mu^2 = \mu + \dfrac{\mu^2}{(\mu^2+\sigma^2)^{0.5} - \mu}$ where $\alpha = [(\mu^2+\sigma^2)^{0.5} - \mu]^{-1}$
skewness	$(1 + 3\mu\alpha + 3\mu^2\alpha^2)(1+\mu\alpha)^{-1.5}\mu^{-0.5}$
excess kurtosis	$(1 + 7\mu\alpha + 18\mu^2\alpha^2 + 15\mu^3\alpha^3)(1+\mu\alpha)^{-2}\mu^{-1}$
PGF[a]	$e^{1/\alpha - q}$ where $q^2 = \alpha^{-2} + 2\mu(1-t)\alpha^{-1}$
pf [a]	$\left(\dfrac{2\sigma}{\pi}\right)^{1/2} \dfrac{\mu^y e^{1/\alpha} K_{y-\frac{1}{2}}(\sigma)}{y!(\sigma\alpha)^y}$
Reference	[a] Stein et al. [1987], equation (2.1) with (ξ,α) replaced by (μ,σ)

5.2.5 Waring: `WARING`

The probability function of the Waring distribution (also called the beta geometric), denoted by $\texttt{WARING}(\mu,\sigma)$, is given by

$$\mathrm{P}(Y = y \,|\, \mu,\sigma) = \frac{B(y + \mu\sigma^{-1}, \sigma^{-1} + 2)}{B(\mu\sigma^{-1}, \sigma^{-1} + 1)} \tag{5.12}$$

for $y = 0, 1, 2, 3, \ldots$, where $\mu > 0$ and $\sigma > 0$. Note that this parameterization of $\texttt{WARING}(\mu,\sigma)$ only includes Waring distributions with finite mean μ. The

$\texttt{WARING}(\mu, \sigma)$ distribution is a reparameterization of the distribution given in Wimmer and Altmann [1999, p643], where $b = \sigma^{-1} + 1$ and $n = \mu\sigma^{-1}$. Hence $\mu = n(b-1)^{-1}$ and $\sigma = (b-1)^{-1}$. It is also a reparameterization of the distribution given in equation 16.131 in Johnson et al. [2005, p290], where $\alpha = \mu\sigma^{-1}$ and $c = (\mu+1)\,\sigma^{-1} + 1$.

The $\texttt{WARING}(\mu, \sigma)$ distribution is a special case of the beta negative binomial, $\texttt{BNB}(\mu, \sigma, \nu)$, distribution where $\nu = 1$. It can be derived as a beta mixture of geometric distributions, by assuming $Y \,|\, \pi \sim \texttt{GEOMo}(\pi)$, where $\pi \sim \texttt{BEo}(b, n)$, $b = \sigma^{-1} + 1$ and $n = \mu\sigma^{-1}$. Hence the $\texttt{WARING}(\mu, \sigma)$ distribution can be considered as an overdispersed geometric distribution.

As $y \to \infty$,

$$\mathrm{P}(Y = y \,|\, \mu, \sigma) \sim q y^{-(\sigma^{-1}+2)}$$

where

$$q = \left(\sigma^{-1} + 1\right) \Gamma\left([\mu + 1]\,\sigma^{-1} + 1\right) / \Gamma\left(\mu\sigma^{-1}\right)$$

and hence the $\texttt{WARING}(\mu, \sigma)$ has a heavy right tail, especially for large σ.

Table 5.12: Waring distribution.

WARING(μ, σ)	
Ranges	
Y	$0, 1, 2, 3 \ldots$
μ	$0 < \mu < \infty$, mean
σ	$0 < \sigma < \infty$
Distribution measures	
mean	μ
mode	0
variance	$\begin{cases} \mu(\mu+1)(1+\sigma)(1-\sigma)^{-1} & \text{if } \sigma < 1 \\ \infty & \text{if } \sigma \geq 1 \end{cases}$
skewness	$\begin{cases} \dfrac{(2\mu+1)(1+2\sigma)(1-\sigma)^{1/2}}{\mu^{1/2}(\mu+1)^{1/2}(1+\sigma)^{1/2}(1-2\sigma)} & \text{if } \sigma < 1/2 \\ \infty & \text{if } \sigma \geq 1/2 \end{cases}$
excess kurtosis	$\begin{cases} \beta_2 - 3 & \text{if } \sigma < 1/3 \\ \infty & \text{if } \sigma \geq 1/3 \end{cases}$
	where $\beta_2 = \dfrac{(1-\sigma)\left\{1 + 7\sigma + 6\sigma^2 + \mu(\mu+1)[9 + 21\sigma + 18\sigma^2]\right\}}{\mu(\mu+1)(1+\sigma)(1-2\sigma)(1-3\sigma)}$
PGF [a]	$\dfrac{(\sigma+1)}{(\mu+\sigma+1)}\, {}_2F_1(\mu\sigma^{-1}, 1; [\mu + 2\sigma + 1]\sigma^{-1}; t)$
pf [a,a2]	$\dfrac{B(y + \mu\sigma^{-1}, \sigma^{-1} + 2)}{B(\mu\sigma^{-1}, \sigma^{-1} + 1)}$
cdf [a2]	$1 - \dfrac{\Gamma(y + \mu\sigma^{-1} + 1)\Gamma([\mu+1]\sigma^{-1} + 1)}{\Gamma(y + [\mu+1]\sigma^{-1} + 2)\Gamma(\mu\sigma^{-1})}$
Reference	[a] Wimmer and Altmann [1999], p643, reparameterized by $b = \sigma^{-1} + 1$ and $n = \mu\sigma^{-1}$ [a2] `http://reference.wolfram.com/language/ref/WaringYuleDistribution.html` reparameterized by $\alpha = \sigma^{-1} + 1$ and $\beta = \mu\sigma^{-1}$. Set $\nu = 1$ in BNB(μ, σ, ν)

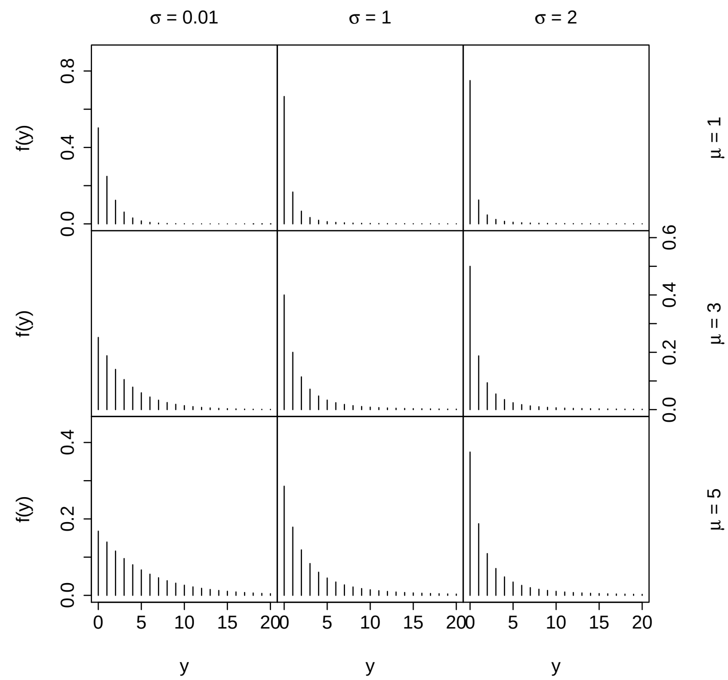

Figure 5.11: The `WARING`(μ, σ), distribution, with $\mu = 1, 3, 5$ and $\sigma = 0.01, 1, 2$.

5.2.6 Zero-adjusted logarithmic: `ZALG`

Let $Y = 0$ with probability σ, and $Y = Y_1$ where $Y_1 \sim$ `LG`(μ), (logarithmic distribution), with probability $(1-\sigma)$. Then Y has a zero-adjusted logarithmic distribution, denoted by `ZALG`(μ, σ), with probability function given by

$$P(Y = y \mid \mu, \sigma) = \begin{cases} \sigma & \text{if } y = 0 \\ (1-\sigma)(\alpha\mu^y)/y & \text{if } y = 1, 2, 3, \ldots \end{cases} \tag{5.13}$$

where $\alpha = -\left[\log(1-\mu)\right]^{-1}$ for $0 < \mu < 1$ and $0 < \sigma < 1$, see Johnson et al. [2005, p355]. For large (and all) y, $P(Y = y \mid \mu, \sigma) = (1-\sigma)\alpha \exp\left(y \log \mu - \log y\right)$, so $\log P(Y = y \mid \mu, \sigma) \sim y \log \mu$ as $y \to \infty$.

Table 5.13: Zero-adjusted logarithmic distribution.

ZALG(μ, σ)	
Ranges	
Y	$0, 1, 2, 3 \ldots$
μ	$0 < \mu < 1$
σ	$0 < \sigma < 1, \ \sigma = \mathrm{P}(Y = 0)$
Distribution measures	
mean	$(1-\sigma)\,\alpha\mu\,(1-\mu)^{-1}$ where $\alpha = -[\log(1-\mu)]^{-1}$
mode	$\begin{cases} 0 & \text{if } \sigma > \frac{\alpha\mu}{(1+\alpha\mu)} \\ 0 \ \text{ and } 1 & \text{if } \sigma = \frac{\alpha\mu}{(1+\alpha\mu)} \\ 1 & \text{if } \sigma < \frac{\alpha\mu}{(1+\alpha\mu)} \end{cases}$
variance	$(1-\sigma)\alpha\mu[1-(1-\sigma)\alpha\mu](1-\mu)^{-2}$
skewness	$\begin{cases} \mu_3/[\mathrm{Var}(Y)]^{1.5} \text{ where} \\ \mu_3 = (1-\sigma)\alpha\mu[1+\mu-3(1-\sigma)\alpha\mu+2(1-\sigma)^2\alpha^2\mu^2](1-\mu)^{-3} \end{cases}$
excess kurtosis	$\begin{cases} k_4/[\mathrm{Var}(Y)]^2 \quad \text{where} \\ k_4 = (1-\sigma)\alpha\mu[1+4\mu+\mu^2-(1-\sigma)\alpha\mu(7+4\mu)+ \\ 12(1-\sigma)^2\alpha^2\mu^2-6(1-\sigma)^3\alpha^3\mu^3](1-\mu)^{-4} \end{cases}$
PGF	$\sigma + (1-\sigma)\dfrac{\log(1-\mu t)}{\log(1-\mu)}$ for $0 < t < \mu^{-1}$
pf	$\begin{cases} \sigma & \text{if } y = 0 \\ (1-\sigma)\alpha\mu^y/y & \text{if } y = 1, 2, 3 \ldots \end{cases}$

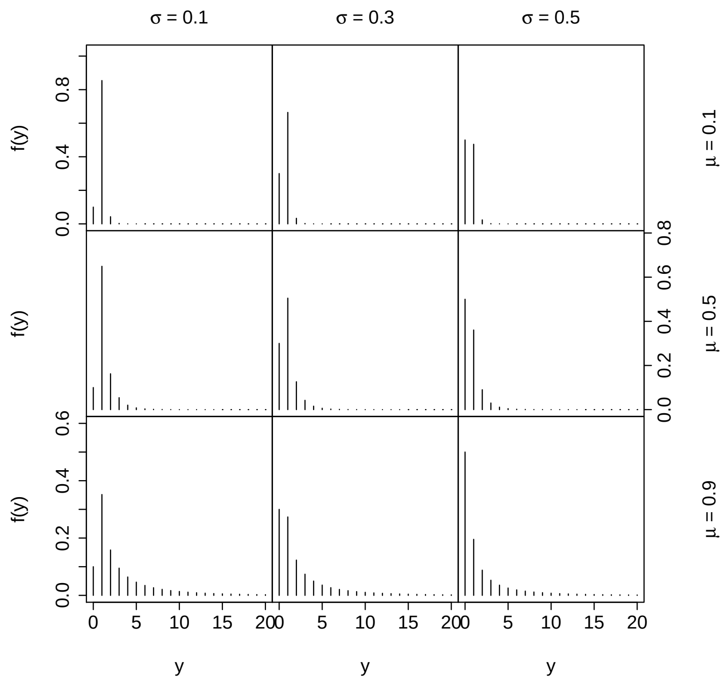

Figure 5.12: The zero-adjusted logarithmic, `ZALG`(μ, σ), distribution, with $\mu = 0.1, 0.5, 0.9$ and $\sigma = 0.1, 0.3, 0.5$.

5.2.7 Zero-adjusted Poisson: `ZAP`

Let $Y = 0$ with probability σ, and $Y = Y_1$ where $Y_1 \sim$ `POtr`(μ) with probability $(1-\sigma)$, where `POtr`(μ) is a Poisson distribution truncated at zero. Then Y has a zero-adjusted Poisson distribution, denoted by `ZAP`(μ, σ), with probability function

$$\mathrm{P}(Y = y \,|\, \mu, \sigma) = \begin{cases} \sigma & \text{if } y = 0 \\ (ce^{-\mu}\mu^y)/y! & \text{if } y = 1, 2, 3, \ldots \end{cases} \tag{5.14}$$

for $\mu > 0$ and $0 < \sigma < 1$, where $c = (1-\sigma)/(1-e^{-\mu})$.

As $y \to \infty$, $P(Y = y \,|\, \mu, \sigma) \sim q \exp[y\,(\log\mu + 1) - (y + 0.5)\log y]$. where $q = ce^{-\mu}\sqrt{2\pi}$, so $\log \mathrm{P}(Y = y \,|\, \mu, \sigma) \sim -y \log y$, as $y \to \infty$.

Table 5.14: Zero-adjusted Poisson distribution.

	ZAP(μ, σ)
	Ranges
Y	$0, 1, 2, 3 \dots$
μ	$0 < \mu < \infty$, mean of Poisson component before truncation at 0
σ	$0 < \sigma < 1$, $\sigma = \text{P}(Y = 0)$
	Distribution measures
mean [a,a2]	$c\mu$ where $c = (1-\sigma)/(1-e^{-\mu})$
variance [a,a2]	$c\mu + c\mu^2 - c^2\mu^2$
skewness [a,a2]	$\begin{cases} \mu_3/[\text{Var}(Y)]^{1.5} \\ \text{where } \mu_3 = c\mu[1 + 3\mu(1-c) + \mu^2(1 - 3c + 2c^2)] \end{cases}$
excess kurtosis [a,a2]	$\begin{cases} \mu_4/[\text{Var}(Y)]^2 - 3 \\ \text{where } \mu_4 = c\mu[1 + \mu(7-4c) + 6\mu^2(1 - 2c + c^2) + \\ \mu^3(1 - 4c + 6c^2 - 3c^3)] \end{cases}$
PGF [a,a2]	$(1-c) + ce^{\mu(t-1)}$
pf [a,a2]	$\begin{cases} \sigma & \text{if } y = 0 \\ (ce^{-\mu}\mu^y)/y! & \text{if } y = 1, 2, 3, \dots \end{cases}$
cdf [a2]	$\begin{cases} \sigma & \text{if } y = 0 \\ \sigma + c\left[\frac{\Gamma(y+1,\mu)}{\Gamma(y)} - e^{-\mu}\right] & \text{if } y = 1, 2, 3, \dots \end{cases}$
Reference	[a] let $\sigma \to 0$ in ZANBI(μ, σ, ν) and then set ν to σ [a2] obtained from equations (??), (??), (??), and (??), where $Y_2 \sim$ PO(μ)

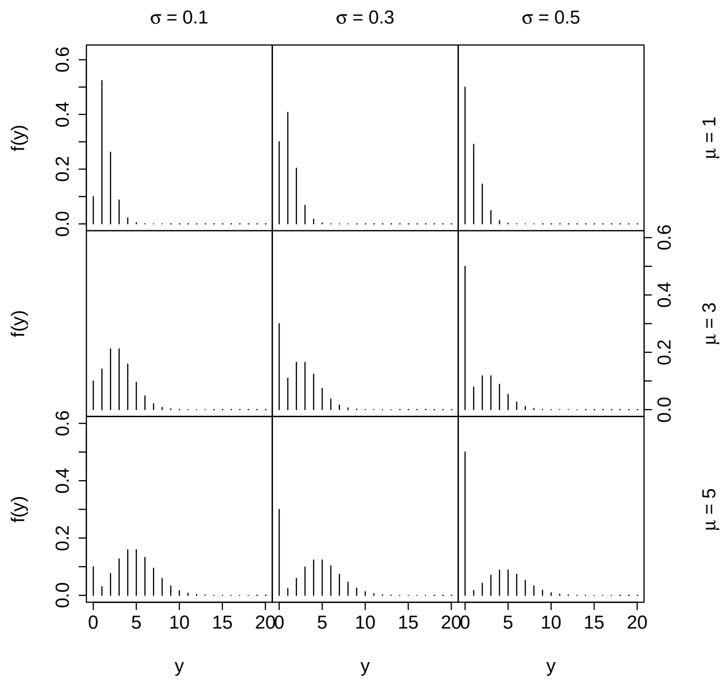

Figure 5.13: The zero-adjusted Poisson, ZAP(μ, σ), distribution, with $\mu = 1, 3, 5$ and $\sigma = 0.1, 0.3, 0.5$.

5.2.8 Zero-adjusted Zipf: ZAZIPF

Let $Y = 0$ with probability σ, and $Y = Y_1$ with probability $(1-\sigma)$, where $Y_1 \sim$ ZIPF(μ). Then Y has a zero-adjusted Zipf distribution, denoted by ZAZIPF(μ, σ), with probability function

$$P(Y = y \,|\, \mu, \sigma) = \begin{cases} \sigma & \text{if } y = 0 \\ (1-\sigma) y^{-(\mu+1)}/\zeta(\mu+1) & \text{if } y = 1, 2, 3, \ldots \end{cases} \tag{5.15}$$

for $\mu > 0$ and $0 < \sigma < 1$.

Table 5.15: Zero-adjusted Zipf distribution.

	ZAZIPF(μ, σ)
	Ranges
Y	$0, 1, 2, 3 \dots$
μ	$0 < \mu < \infty$
σ	$0 < \sigma < 1,\ \sigma = \mathrm{P}(Y = 0)$
	Distribution measures
mean[a]	$\begin{cases} b(1-\sigma) & \text{if } \mu > 1 \\ \infty & \text{if } \mu \le 1 \end{cases}$ where $b = \zeta(\mu)/\zeta(\mu+1)$
mode	$\begin{cases} 0 & \text{if} \quad \sigma > [1+\zeta(\mu+1)]^{-1} \\ 0 \quad \text{and} \quad 1 & \text{if} \quad \sigma = [1+\zeta(\mu+1)]^{-1} \\ 1 & \text{if} \quad \sigma < [1+\zeta(\mu+1)]^{-1} \end{cases}$
variance[a]	$\begin{cases} (1-\sigma)[\zeta(\mu-1)/\zeta(\mu+1)] - (1-\sigma)^2 b^2 & \text{if} \quad \mu > 2 \\ \infty & \text{if} \quad \mu \le 2 \end{cases}$
skewness[a]	$\begin{cases} \mu_3/[\mathrm{Var}(Y)]^{1.5} \quad \text{where} & \\ \mu_3 = \dfrac{(1-\sigma)\zeta(\mu-2) - 3b(1-\sigma)^2\zeta(\mu-1)}{\zeta(\mu+1)} + & \\ \qquad 2b^3(1-\sigma)^3 & \text{if } \mu > 3 \\ \infty & \text{if } \mu \le 3 \end{cases}$
excess kurtosis [a]	$\begin{cases} \mu_4/[\mathrm{Var}(Y)]^2 - 3 \quad \text{where} & \\ \mu_4 = \{\left[(1-\sigma)\zeta(\mu-3) - 4b(1-\sigma)^2\zeta(\mu-2) + \right. & \\ \qquad \left. 6b^2(1-\sigma)^3\zeta(\mu-1)\right]/\zeta(\mu+1)\} - 3b^4(1-\sigma)^4 & \text{if } \mu > 4 \\ \infty & \text{if } \mu \le 4 \end{cases}$
PGF	$\sigma + (1-\sigma)t\phi(t, \mu+1, 1)/\phi(1, \mu+1, 1)$
pf	$\begin{cases} \sigma & \text{if } y = 0 \\ (1-\sigma)y^{-(\mu+1)}/\zeta(\mu+1) & \text{if } y = 1, 2, 3, \dots \end{cases}$
Reference	[a] Obtained using equation (**??**)
Notes	$\zeta(b) = \sum_{i=1}^{\infty} i^{-b}$ is the Riemann zeta function
	$\phi(a, b, c) = \sum_{i=0}^{\infty} \frac{a^i}{(i+c)^b}$, for $c \ne 0, -1, -2$, is the Lerch function

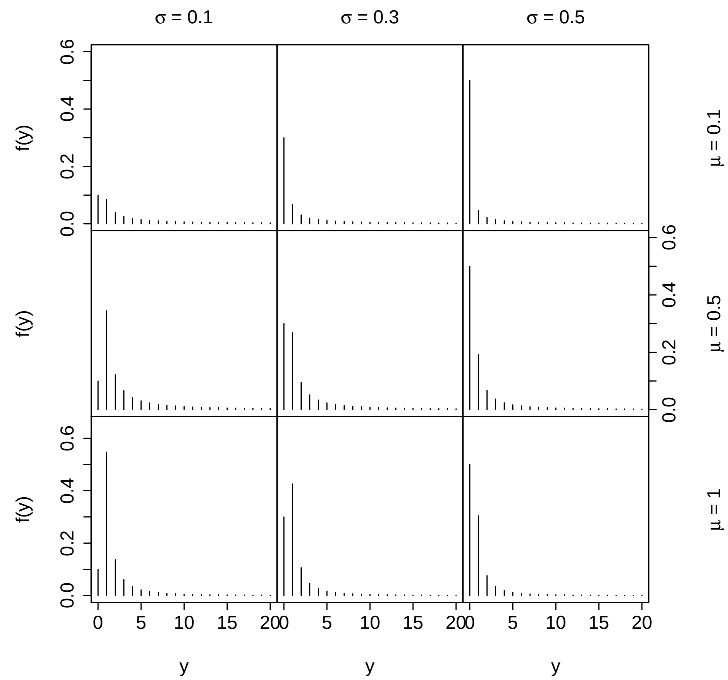

Figure 5.14: The zero-adjusted Zipf, ZAZIPF(μ, σ), distribution, with $\mu = 0.1, 0.5, 1$ and $\sigma = 0.1, 0.3, 0.5$.

5.2.9 Zero-inflated Poisson: ZIP, ZIP2

First parameterization, ZIP

Let $Y = 0$ with probability σ, and $Y = Y_1$ with probability $(1-\sigma)$, where $Y_1 \sim$ PO(μ). Then Y has a zero-inflated Poisson distribution, denoted by ZIP(μ, σ), given by

$$P(Y = y \mid \mu, \sigma) = \begin{cases} \sigma + (1-\sigma)e^{-\mu} & \text{if} \quad y = 0 \\ (1-\sigma)e^{-\mu}\mu^y/y! & \text{if} \quad y = 1, 2, 3, \ldots \end{cases} \tag{5.16}$$

for $\mu > 0$ and $0 < \sigma < 1$. See Johnson et al. [2005, p193] for this parameterization, which was also used by Lambert [1992]. As $y \to \infty$,

$$P(Y = y|\mu, \sigma) \sim q \exp\left[y\left(\log\mu + 1\right) - (y + 0.5)\log y\right]$$

where $q = (1-\sigma)\ e^{-\mu}\sqrt{2\pi}$, so, as $y \to \infty$,

$$\log P(Y = y \mid \mu, \sigma) \sim -y \log y \ .$$

Table 5.16: Zero-inflated Poisson distribution.

	$\texttt{ZIP}(\mu, \sigma)$
	Ranges
Y	$0, 1, 2, 3 \ldots$
μ	$0 < \mu < \infty$, mean of Poisson component
σ	$0 < \sigma < 1$, inflation probability at zero
	Distribution measures
mean[a,a2]	$(1-\sigma)\mu$
variance[a,a2]	$\mu(1-\sigma)(1+\mu\sigma)$
skewness[a,a2]	$\mu_3/[\text{Var}(Y)]^{1.5}$ where $\mu_3 = \mu(1-\sigma)[1+3\mu\sigma+\mu^2\sigma(2\sigma-1)]$
excess kurtosis [a,a2]	$k_4/[\text{Var}(Y)]^2$ where $k_4 = \mu(1-\sigma)[1+7\mu\sigma-6\mu^2\sigma+ 12\mu^2\sigma^2+\mu^3\sigma(1-6\sigma+6\sigma^2)]$
PGF[a,a2]	$\sigma + (1-\sigma)e^{\mu(t-1)}$
pf [a]	$\begin{cases} \sigma + (1-\sigma)e^{-\mu} & \text{if} \quad y = 0 \\ (1-\sigma)e^{-\mu}\mu^y/y! & \text{if} \quad y = 1, 2, 3, \ldots \end{cases}$
cdf[a,a2]	$\sigma + \dfrac{(1-\sigma)\Gamma(y+1,\mu)}{\Gamma(y)}$
Reference	[a] let $\sigma \to 0$ in $\texttt{ZINBI}(\mu, \sigma, \nu)$ and then set ν to σ [a2] obtained from equations (**??**), (**??**), (**??**), and (**??**), where $Y_1 \sim \texttt{PO}(\mu)$

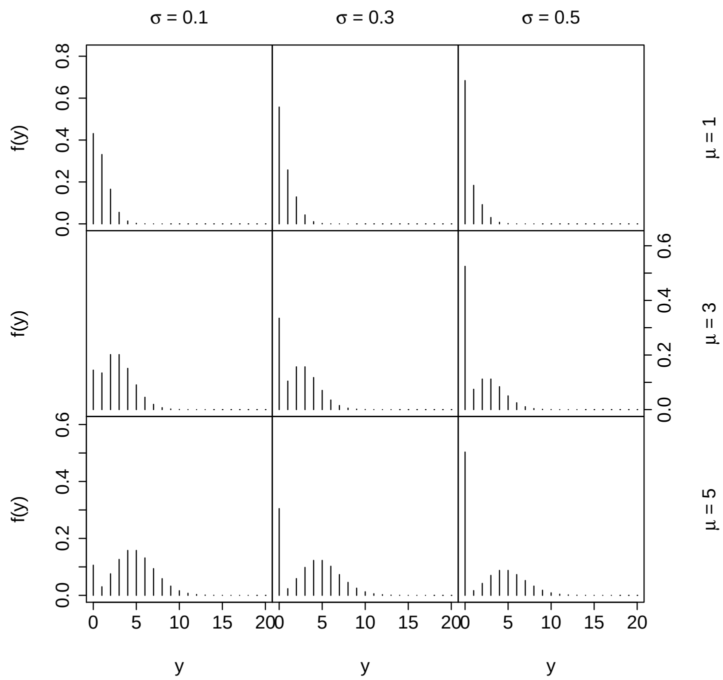

Figure 5.15: The zero-inflated Poisson, `ZIP`(μ, σ), distribution, with $\mu = 0.5, 1, 5$ and $\sigma = 0.1, 0.5$.

Second parameterization, `ZIP2`

A different parameterization of the zero-inflated Poisson distribution, denoted by `ZIP2`(μ, σ), has a probability function given by

$$\mathrm{P}(Y = y \mid \mu, \sigma) = \begin{cases} \sigma + (1-\sigma)e^{-\left(\frac{\mu}{1-\sigma}\right)} & \text{if} \quad y = 0 \\ \dfrac{\mu^y}{y!(1-\sigma)^{y-1}} e^{-\left(\frac{\mu}{1-\sigma}\right)} & \text{if} \quad y = 1, 2, 3, \ldots \end{cases} \tag{5.17}$$

for $\mu > 0$ and $0 < \sigma < 1$. The `ZIP2`(μ, σ) distribution is given by reparameterizing μ to $\mu/(1-\sigma)$ in the `ZIP`(μ, σ) distribution. This has the advantage that $\mathrm{E}(Y) = \mu$.

As $y \to \infty$,

$$\mathrm{P}(Y = y|\mu, \sigma) \sim q \exp\left[y\left(\log(\mu/(1-\sigma)) + 1\right) - (y + 0.5)\log y\right] \ ,$$

where $q = \sqrt{2\pi}(1-\sigma)\ e^{-\mu/(1-\sigma)}$, so $\log \mathrm{P}(Y = y \mid \mu, \sigma) \sim -y \log y \ \text{ as } y \to \infty$.

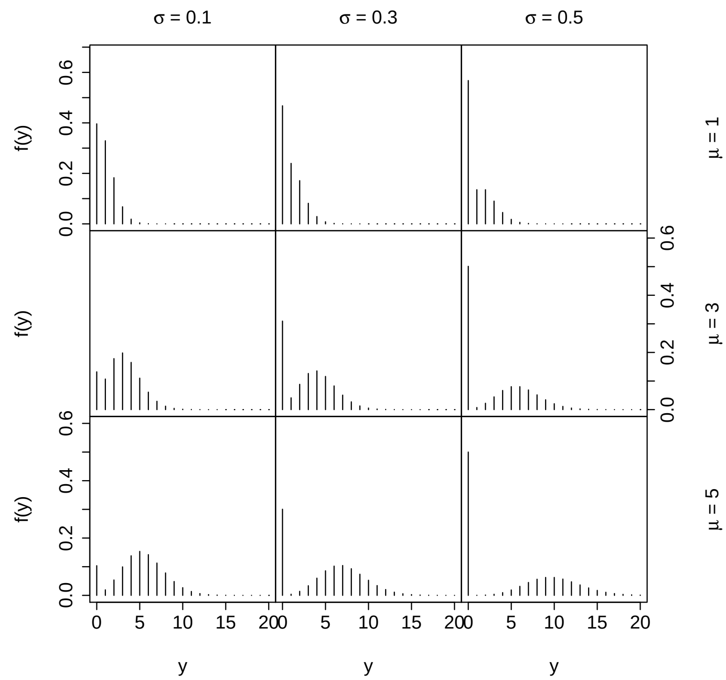

Figure 5.16: The zero-inflated Poisson, ZIP2(μ, σ), distribution type 2, with $\mu = 1, 3, 5$ and $\sigma = 0.1, 0.3, 0.5$.

Table 5.17: Zero-inflated Poisson distribution type 2.

$\texttt{ZIP2}(\mu, \sigma)$	
Ranges	
Y	$0, 1, 2, 3 \ldots$
μ	$0 < \mu < \infty$, mean
σ	$0 < \sigma < 1$, inflation probability at zero
Distribution measures	
mean	μ
variance	$\mu\,[1 + \mu\sigma/(1-\sigma)]$
skewness	$\mu_3/[\text{Var}(Y)]^{1.5}$ where $\mu_3 = \mu\left[1 + \frac{3\mu\sigma}{(1-\sigma)} + \frac{\mu^2\sigma(2\sigma-1)}{(1-\sigma)^2}\right]$
excess kurtosis	$k_4/[\text{Var}(Y)]^2$ where $k_4 = \mu[1 + \frac{7\mu\sigma}{(1-\sigma)} + \frac{6\mu^2\sigma(2\sigma-1)}{(1-\sigma)^2} + \frac{\mu^3\sigma}{(1-\sigma)^3}(1 - 6\sigma + 6\sigma^2)]$
PGF[a,b]	$\sigma + (1-\sigma)e^{\mu(t-1)/(1-\sigma)}$
pf[a]	$\begin{cases} \sigma + (1-\sigma)e^{-\left(\frac{\mu}{1-\sigma}\right)} & \text{if} \quad y = 0 \\ \dfrac{\mu^y}{y!(1-\sigma)^{y-1}} e^{-\left(\frac{\mu}{1-\sigma}\right)} & \text{if} \quad y = 1, 2, 3, \ldots \end{cases}$
cdf[b]	$\sigma + (1-\sigma)\Gamma(y+1, \frac{\mu}{1-\sigma})/\Gamma(y)$
Reference	Reparameterize μ_1 to $\mu/(1-\sigma)$ in $\texttt{ZIP}(\mu_1, \sigma)$

5.3 Three-parameter count distributions

5.3.1 Beta negative binomial: BNB

The probability function of the beta negative binomial distribution, denoted by $\texttt{BNB}(\mu, \sigma, \nu)$, is given by

$$P(Y = y \mid \mu, \sigma, \nu) = \frac{\Gamma(y + \nu^{-1})\;\; B(y + \mu\sigma^{-1}\nu, \sigma^{-1} + \nu^{-1} + 1)}{\Gamma(y+1)\;\; \Gamma(\nu^{-1})\;\; B(\mu\sigma^{-1}\nu, \sigma^{-1} + 1)} \tag{5.18}$$

for $y = 0, 1, 2, 3, \ldots$, where $\mu > 0$, $\sigma > 0$, and $\nu > 0$. Note that this parameterization only includes beta negative binomial distributions with a finite mean μ. The beta negative binomial distribution is also called the beta Pascal distribution or the generalized Waring distribution.

The $\texttt{BNB}(\mu, \sigma, \nu)$ distribution has mean μ and is an overdispersed negative binomial distribution. The Waring distribution, $\texttt{WARING}(\mu, \sigma)$, (which is an overdispersed geometric distribution), is a special case of $\texttt{BNB}(\mu, \sigma, \nu)$ where $\nu = 1$. The negative binomial distribution is a limiting case of the beta negative binomial, since $\texttt{BNB}(\mu, \sigma, \nu) \to \texttt{NBI}(\mu, \nu)$ as $\sigma \to 0$ (for fixed μ and ν).

The $\texttt{BNB}(\mu, \sigma, \nu)$ is a reparameterization of the distribution given in Wimmer and Altmann [1999, p19], where $m = \sigma^{-1} + 1$ and $n = \mu\sigma^{-1}\nu$ and $k = \nu^{-1}$.

Hence $\mu = kn(m-1)^{-1}$, $\sigma = (m-1)^{-1}$ and $\nu = 1/k$. It can be derived as a beta mixture of negative binomial distributions, by assuming $y \mid \pi \sim$ `NBo`(k, π) where $\pi \sim$ `BEo`(m, n). See Section 5.2.3 for `NBo`(k, π). Hence the `BNB`(μ, σ, ν) can be considered as an overdispersed negative binomial distribution.

As $y \to \infty$, $\mathrm{P}(Y = y \mid \mu, \sigma, \nu) \sim q y^{-(\sigma^{-1}+2)}$ where

$$q = \Gamma\left([\mu\nu + 1]\sigma^{-1} + 1\right) / [B(\sigma^{-1} + 1, \nu^{-1})\Gamma(\mu\sigma^{-1}\nu)]$$

and hence the `BNB`(μ, σ, ν) distribution has a heavy right tail, especially for large σ.

The probability function (5.18) and the mean, variance, skewness, and kurtosis in Table 5.18 can be obtained from Johnson et al. [2005, p259, 263] setting $\alpha = -\mu\nu\sigma^{-1}$, $b = (\mu\nu + 1)\sigma^{-1}$ and $n = -\nu^{-1}$. The equivalence of their probability function to (5.18) is shown using their equation (1.24).

To interpret the parameters of `BNB`(μ, σ, ν), μ is the mean, σ is a right tail heaviness parameter (and increasing σ increases the variance, for finite σ, i.e. $\sigma < 1$), and ν increases the variance (for $\nu^2 > \sigma/\mu$ and $\sigma < 1$), while the variance is infinite for $\sigma \geq 1$.

Table 5.18: Beta negative binomial distribution.

	$\texttt{BNB}(\mu, \sigma, \nu)$
	Ranges
Y	$0, 1, 2, 3 \ldots$
μ	$0 < \mu < \infty$, mean
σ	$0 < \sigma < \infty$
ν	$0 < \nu < \infty$
	Distribution measures
mean	μ
variance	$\begin{cases} \mu(1+\mu\nu)(1+\sigma\nu^{-1})(1-\sigma)^{-1} & \text{if } \sigma < 1 \\ \infty & \text{if } \sigma \geq 1 \end{cases}$
skewness	$\begin{cases} \dfrac{(2\mu\nu+1)(1+2\sigma\nu^{-1})(1-\sigma)^{1/2}}{\mu^{1/2}(\mu\nu+1)^{1/2}(1+\sigma\nu^{-1})^{1/2}(1-2\sigma)} & \text{if } \sigma < 1/2 \\ \infty & \text{if } \sigma \geq 1/2 \end{cases}$
excess kurtosis	$\beta_2 - 3$ where
	$\beta_2 = \begin{cases} (1-\sigma)\left[1+\sigma+6\sigma\nu^{-1}+6\sigma^2\nu^{-2}+ \right. \\ \left. 3\mu(\mu\nu+1)(1+2\nu+\sigma\nu^{-1}+6\sigma+6\sigma^2\nu^{-1})\right] \times \\ \left[\mu(\mu\nu+1)(1+\sigma\nu^{-1})(1-2\sigma)(1-3\sigma)\right]^{-1} & \text{if } \sigma < 1/3 \\ \infty & \text{if } \sigma \geq 1/3 \end{cases}$
PGF[a]	$\dfrac{{}_2F_1(\nu^{-1}, \mu\sigma^{-1}\nu; \mu\sigma^{-1}\nu + \sigma^{-1} + \nu^{-1} + 1; t)}{{}_2F_1(\nu^{-1}, \mu\sigma^{-1}\nu; \mu\sigma^{-1}\nu + \sigma^{-1} + \nu^{-1} + 1; 1)}$
pf[a]	$\dfrac{\Gamma(y+\nu^{-1})B(y+\mu\sigma^{-1}\nu, \sigma^{-1}+\nu^{-1}+1)}{\Gamma(y+1)\Gamma(\nu^{-1})B(\mu\sigma^{-1}\nu, \sigma^{-1}+1)}$
Reference	[a] Wimmer and Altmann [1999] p19, reparameterized by $m = \sigma^{-1}+1$, $n = \mu\sigma^{-1}\nu$ and $k = \nu^{-1}$ and hence $\mu = kn(m-1)^{-1}$, $\sigma = (m-1)^{-1}$ and $\nu = 1/k$

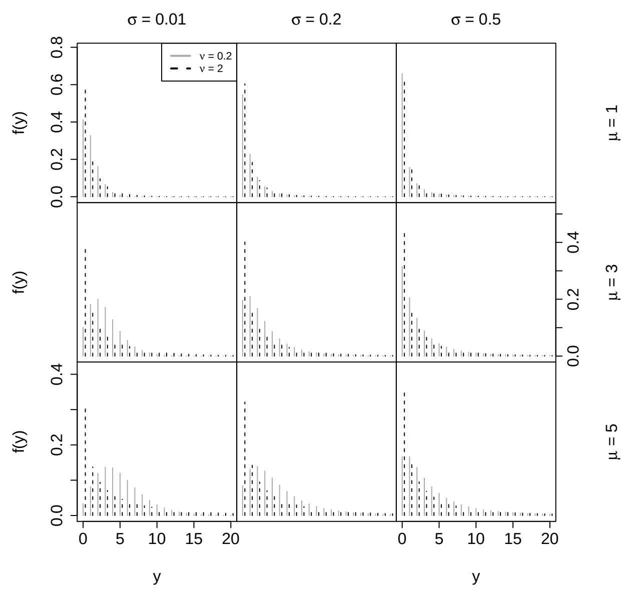

Figure 5.17: The beta negative binomial, BNB(μ, σ, ν), distribution, with $\mu = 1, 3, 5$, $\sigma = 0.01, 0.2, 0.5$, and $\nu = 0.2, 2$.

5.3.2 Discrete Burr XII: `DBURR12`

The discrete Burr XII distribution, denoted by DBURR12(μ, σ, ν), has probability function given by

$$\mathrm{P}(Y = y \mid \mu, \sigma, \nu) = \left[1 + \left(\frac{y}{\mu}\right)^{\sigma}\right]^{-\nu} - \left[1 + \left(\frac{y+1}{\mu}\right)^{\sigma}\right]^{-\nu} \tag{5.19}$$

for $y = 0, 1, 2, 3, \ldots$, where $\mu > 0$, $\sigma > 0$ and $\nu > 0$.

This distribution was investigated by Para and Jan [2016], who used parameterization $\gamma = \mu$, $c = \sigma$ and $\beta = \exp(-\nu)$, so $\nu = -\log(\beta)$ and hence (5.19) can be written in the form

$$\mathrm{P}(Y = y \mid \mu, \sigma, \beta) = \beta^{\log\left[1+\left(\frac{y}{\mu}\right)^{\sigma}\right]} - \beta^{\log\left[1+\left(\frac{y+1}{\mu}\right)^{\sigma}\right]} \tag{5.20}$$

since $\beta^{\log\alpha} = \exp\left[(\log\alpha)(\log\beta)\right] = \alpha^{\log\beta}$.

Note parameters μ and ν in DBURR12(μ, σ, ν) can be informationally highly correlated, so arguments `method=mixed(10,100)` and `c.crit=0.0001` are highly recommended in the `gamlss()` fitting function, in order to speed convergence and avoid converging too early. As $y \to \infty$, $\log \mathrm{P}(Y = y|\mu, \sigma, \nu) \sim -(1 + \sigma\nu)\log y$, so DBURR12$(\mu, \sigma, \nu)$ can have a very heavy tail if $\sigma\nu$ is close to 0.

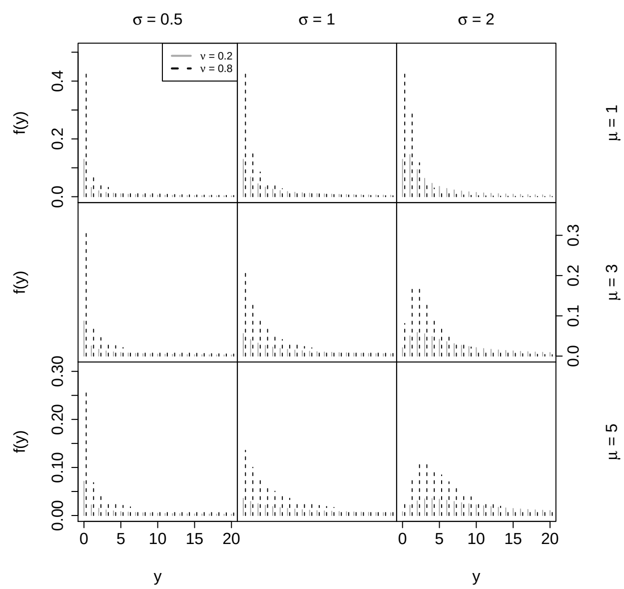

Figure 5.18: The discrete Burr XII, DBURR12(μ, σ, ν), distribution, with $\mu = 1, 3, 5$, $\sigma = 0.5, 1, 2$, and $\nu = 0.2, 0.8$.

Table 5.19: Discrete Burr XII distribution.

	`DBURR12`(μ, σ, ν)
	Ranges
Y	$0, 1, 2, 3 \ldots$
μ	$0 < \mu < \infty$
σ	$0 < \sigma < \infty$
ν	$0 < \nu < \infty$
	Distribution measures
mean	$\sum_{y=1}^{\infty} \left[1 + \left(\frac{y}{\mu}\right)^{\sigma}\right]^{-\nu}$, for $\sigma\nu > 1$
pf[a]	$\left[1 + \left(\frac{y}{\mu}\right)^{\sigma}\right]^{-\nu} - \left[1 + \left(\frac{y+1}{\mu}\right)^{\sigma}\right]^{-\nu}$
cdf	$1 - \left[1 + \left(\frac{y+1}{\mu}\right)^{\sigma}\right]^{-\nu}$
inverse cdf	$\left\lceil \mu \left\{ \exp\left[-\frac{\log(1-p)}{\nu}\right] - 1 \right\}^{1/\sigma} - 1 \right\rceil$
Reference	Para and Jan [2016] with parameters (γ, c, β) replaced by $(\mu, \sigma, \exp(-\nu))$.
Note	$\lceil x \rceil$ is the ceiling function, i.e. the smallest integer greater than or equal to x

5.3.3 Delaporte: `DEL`

The probability function of the Delaporte distribution, denoted by `DEL`(μ, σ, ν), is given by

$$P(Y = y \mid \mu, \sigma, \nu) = \frac{e^{-\mu\nu}}{\Gamma(1/\sigma)} \left[1 + \mu\sigma(1-\nu)\right]^{-1/\sigma} S \tag{5.21}$$

where

$$S = \sum_{j=0}^{y} \binom{y}{j} \frac{\mu^y \nu^{y-j}}{y!} \left[\mu + \frac{1}{\sigma(1-\nu)}\right]^{-j} \Gamma\left(1/\sigma + j\right)$$

for $y = 0, 1, 2, 3, \ldots$, where $\mu > 0$, $\sigma > 0$, and $0 < \nu < 1$. This distribution is a reparameterization of the distribution given by Wimmer and Altmann [1999, p515-516] where $\alpha = \mu\nu$, $k = 1/\sigma$ and $\rho = [1 + \mu\sigma(1-\nu)]^{-1}$. This parameterization is given by Rigby et al. [2008]. As $y \to \infty$,

$$\log P(Y = y \mid \mu, \sigma, \nu) \sim -y \log\left[1 + \frac{1}{\mu\sigma(1-\nu)}\right] .$$

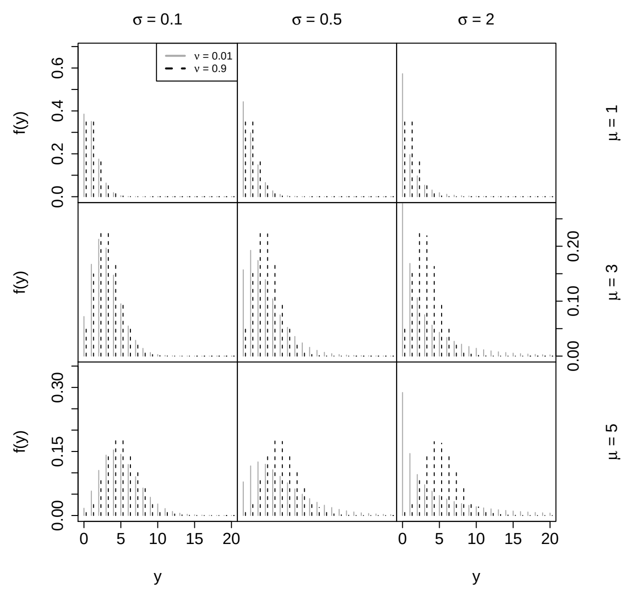

Figure 5.19: The Delaporte, $\texttt{DEL}(\mu, \sigma, \nu)$, distribution, with $\mu = 1, 3, 5$, $\sigma = 0.1, 0.5, 2$, and $\nu = 0.01, 0.9$.

Table 5.20: Delaporte distribution.

	DEL(μ,σ,ν)
	Ranges
Y	$0,1,2,3\ldots$
μ	$0<\mu<\infty$, mean
σ	$0<\sigma<\infty$
ν	$0<\nu<1$
	Distribution measures
mean	μ
variance	$\mu+\mu^2\sigma(1-\nu)^2$
skewness	$\mu_3/[\mathrm{Var}(Y)]^{1.5}$ where $\mu_3=\mu[1+3\mu\sigma(1-\nu)^2+2\mu^2\sigma^2(1-\nu)^3]$
excess kurtosis	$\begin{cases} k_4/[\mathrm{Var}(Y)]^2 \quad \text{where} \\ k_4=\mu[1+7\mu\sigma(1-\nu)^2+12\mu^2\sigma^2(1-\nu)^3+6\mu^3\sigma^3(1-\nu)^4] \end{cases}$
PGF	$e^{\mu\nu(t-1)}[1+\mu\sigma(1-\nu)(1-t)]^{-1/\sigma}$
pf	$\frac{\exp(-\mu\nu)}{\Gamma(1/\sigma)}\,[1+\mu\sigma(1-\nu)]^{-1/\sigma}\,S$ where
	$S=\sum_{j=0}^{y}\binom{y}{j}\frac{\mu^y\nu^{y-j}}{y!}\left[\mu+\frac{1}{\sigma(1-\nu)}\right]^{-j}\Gamma\left(1/\sigma+j\right)$
Reference	Rigby et al. [2008]

5.3.4 Negative binomial family: NBF

The probability function of the negative binomial family distribution, denoted NBF(μ,σ,ν), is given by

$$\mathrm{P}(Y=y\,|\,\mu,\sigma,\nu)=\frac{\Gamma\left(y+\sigma^{-1}\mu^{2-\nu}\right)\sigma^y\mu^{y(\nu-1)}}{\Gamma\left(\sigma^{-1}\mu^{2-\nu}\right)\Gamma\left(y+1\right)\left(1+\sigma\mu^{\nu-1}\right)^{\sigma^{-1}\mu^{2-\nu}+y}} \tag{5.22}$$

for $y=0,1,2,3,\ldots$, where $\mu>0$, $\sigma>0$ and $\nu>0$.

This family of reparameterizations of the negative binomial distribution is obtained by reparameterizing σ to $\sigma\mu^{\nu-2}$ in NBI(μ,σ). The variance of $Y\sim$ NBF(μ,σ,ν) is $\mathrm{Var}(Y)=\mu+\sigma\mu^\nu$. Hence ν is the power in the variance-mean relationship.

Table 5.21: Negative binomial family distribution.

	$\texttt{NBF}(\mu,\sigma,\nu)$
	Ranges
Y	$0,1,2,3\ldots$
μ	$0<\mu<\infty$ mean
σ	$0<\sigma<\infty$ dispersion
ν	$0<\nu<\infty$ power parameter in variance-mean relationship
	Distribution measures
mean	μ
variance	$\mu+\sigma\mu^{\nu}$
skewness	$(1+2\sigma\mu^{\nu-1})(\mu+\sigma\mu^{\nu})^{-0.5}$
excess kurtosis	$6\sigma\mu^{\nu-2}+(\mu+\sigma\mu^{\nu})^{-1}$
PGF	$\left[1+\sigma\mu^{\nu-1}(1-t)\right]^{-1/(\sigma\mu^{\nu-2})}$ for $0<t<1+(\sigma\mu^{\nu-1})^{-1}$
pf	$\dfrac{\Gamma\left(y+\sigma^{-1}\mu^{2-\nu}\right)\sigma^{y}\mu^{y(\nu-1)}}{\Gamma\left(\sigma^{-1}\mu^{2-\nu}\right)\Gamma\left(y+1\right)\left(1+\sigma\mu^{\nu-1}\right)^{\sigma^{-1}\mu^{2-\nu}+y}}$
cdf	$1-\dfrac{B\left(y+1,\sigma^{-1}\mu^{2-\nu},\sigma\mu^{\nu-1}(1+\sigma\mu^{\nu-1})^{-1}\right)}{B\left(y+1,\sigma^{-1}\mu^{2-\nu}\right)}$
Reference	Reparameterized σ_1 to $\sigma\mu^{\nu-2}$ in $\texttt{NBI}(\mu,\sigma_1)$

5.3.5 Sichel: SICHEL, SI

First parameterization, SICHEL

This parameterization of the Sichel distribution, [Rigby et al., 2008], denoted by $\texttt{SICHEL}(\mu,\sigma,\nu)$, has probability function

$$\mathrm{P}(Y=y\,|\,\mu,\sigma,\nu)=\frac{(\mu/b)^{y}K_{y+\nu}(\alpha)}{y!\,(\alpha\sigma)^{y+\nu}\,K_{\nu}\left(1/\sigma\right)} \tag{5.23}$$

for $y=0,1,2,3,\ldots$, where $\mu>0$, $\sigma>0$ and $-\infty<\nu<\infty$, and

$$\alpha^2=\sigma^{-2}+2\mu(b\sigma)^{-1}\,,\quad b=\frac{K_{\nu+1}(1/\sigma)}{K_{\nu}(1/\sigma)}$$

and $K_{\lambda}(t)$ is the modified Bessel function of the second kind [Abramowitz and Stegun, 1965]. Note

$$\sigma=\left[\left(\mu^2/b^2+\alpha^2\right)^{0.5}-\mu/b\right]^{-1}.$$

The $\texttt{SICHEL}(\mu,\sigma,\nu)$ distribution is a reparameterization of the $\texttt{SI}(\mu,\sigma,\nu)$ distribution given by setting μ to μ/b, so that the mean of $\texttt{SICHEL}(\mu,\sigma,\nu)$ is μ. As $\sigma \to 0$, $\texttt{SICHEL}(\mu,\sigma,\nu) \to \texttt{PO}(\mu)$. For $\nu > 0$, as $\sigma \to \infty$ $\texttt{SICHEL}(\mu,\sigma,\nu) \to \texttt{NBI}(\mu,\nu^{-1})$. For $\nu = 0$, as $\sigma \to \infty$, $\texttt{SICHEL}(\mu,\sigma,\nu) \to \texttt{ZALG}(\mu_1,\sigma_1)$ where $\mu_1 = (2\mu\log\sigma)/(1+2\mu\log\sigma)$ and $\sigma_1 = 1 - [\log(1+2\mu\log\sigma)]\,/\,(2\log\sigma)$. This tends to a degenerate point probability 1 at $y=0$ in the limit as $\sigma \to \infty$.

As $y \to \infty$, $\text{P}(Y = y\,|\,\mu,\sigma,\nu) \sim q\exp\left[-y\log\left(1+\frac{b}{2\mu\sigma}\right) - (1-\nu)\log y\right]$ where q does not depend on y, i.e. essentially an exponential tail. The tail becomes heavier as $2\mu\sigma/b$ increases.

Table 5.22: Sichel distribution, first parameterization.

$\texttt{SICHEL}(\mu,\sigma,\nu)$	
Ranges	
Y	$0,1,2,3\ldots$
μ	$0<\mu<\infty$, mean
σ	$0<\sigma<\infty$
ν	$-\infty<\nu<\infty$
Distribution measures	
mean	μ
variance	$\mu+\mu^2 g_1$ where $g_1 = \dfrac{2\sigma(\nu+1)}{b} + \dfrac{1}{b^2} - 1$
skewness	$\begin{cases} \mu_3/[\text{Var}(Y)]^{1.5} \text{ where } \mu_3 = \mu + 3\mu^2 g_1 + \mu^3(g_2 - 3g_1) \\ \text{where } g_2 = \dfrac{2\sigma(\nu+2)}{b^3} + \dfrac{[4\sigma^2(\nu+1)(\nu+2)+1]}{b^2} - 1 \end{cases}$
excess kurtosis	$\begin{cases} k_4/[\text{Var}(Y)]^2 \quad \text{where} \\ k_4 = \mu + 7\mu^2 g_1 + 6\mu^3(g_2 - 3g_1) + \mu^4(g_3 - 4g_2 + 6g_1 - 3g_1^2) \\ \text{and where} \\ g_3 = [1+4\sigma^2(\nu+2)(\nu+3)]/b^4 + \\ \quad [8\sigma^3(\nu+1)(\nu+2)(\nu+3) + 4\sigma(\nu+2)]/b^3 - 1 \end{cases}$
PGF	$\dfrac{K_\nu(q)}{(q\sigma)^\nu K_\nu(1/\sigma)}$ where $q^2 = \sigma^{-2} + 2\mu(1-t)(b\sigma)^{-1}$
pf	$\dfrac{(\mu/b)^y K_{y+\nu}(\alpha)}{y!\,(\alpha\sigma)^{y+\nu}\,K_\nu\,(1/\sigma)}$ where $\alpha^2 = \sigma^{-2} + 2\mu(b\sigma)^{-1}$
Reference	Rigby et al. [2008]
Notes	$b = K_{\nu+1}(1/\sigma)/K_\nu(1/\sigma)$
	$K_\lambda(t)$ is the modified Bessel function of the second kind

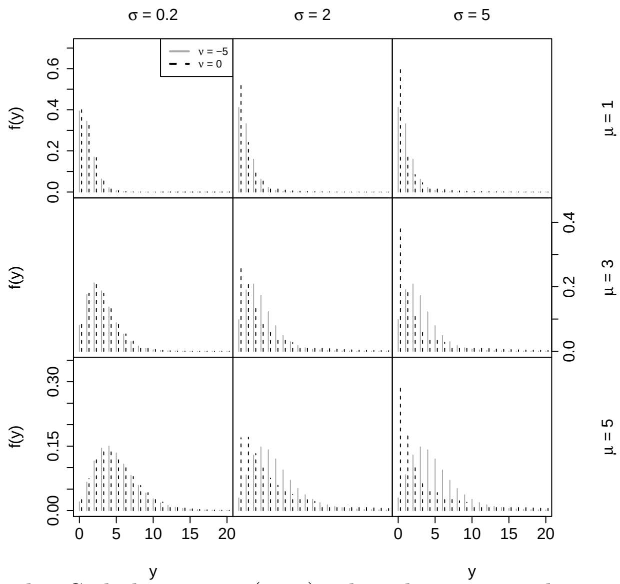

Figure 5.20: The Sichel, SICHEL(μ, σ), distribution, with $\mu = 1, 3, 5$, $\sigma = 0.2, 2, 5$, and $\nu = -5, 0$.

Second parameterization, SI

The probability function of the second parameterization of the Sichel distribution, denoted by SI(μ, σ, ν), is given by

$$\mathrm{P}(Y = y \,|\, \mu, \sigma, \nu) = \frac{\mu^y K_{y+\nu}(\alpha)}{y!(\alpha\sigma)^{y+\nu} K_\nu(1/\sigma)} \tag{5.24}$$

for $y = 0, 1, 2, 3, \ldots$, where $\mu > 0$, $\sigma > 0$, $-\infty < \nu < \infty$, and where $\alpha^2 = \sigma^{-2} + 2\mu\sigma^{-1}$, and $K_\lambda(t)$ is the modified Bessel function of the second kind [Abramowitz and Stegun, 1965]. Note that $\sigma = [(\mu^2 + \alpha^2)^{0.5} - \mu]^{-1}$. Note that the above parameterization (5.24) is different from that of Stein et al. [1987, Section 2.1], who use the above probability function but treat μ, α, and ν as the parameters.

As $\sigma \to 0$, SI$(\mu, \sigma, \nu) \to$ PO(μ). Following Stein et al. [1987], for $\nu = 0$, then as $\sigma \to \infty$, SI$(\mu, \sigma, \nu) \to$ ZALG(μ_1, σ_1) where $\mu_1 = 2\mu\sigma/(1 + 2\mu\sigma)$ and $\sigma_1 = 1 - [\log(1 + 2\mu\sigma)]/[2 \log \sigma]$. This tends to a degenerate ZALG$(1, 0.5)$ in the limit as $\sigma \to \infty$.

As $y \to \infty$, $\mathrm{P}(Y = y \,|\, \mu, \sigma, \nu) \sim q \exp\left[-y \log(1 + \frac{1}{2\mu\sigma}) - (1 - \nu) \log y\right]$, where q does not depend on y, i.e. essentially an exponential tail. The tail becomes heavier as $2\mu\sigma$ increases.

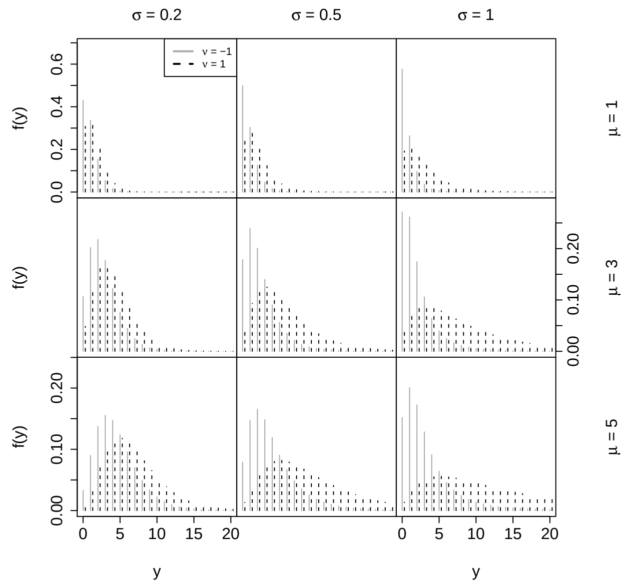

Figure 5.21: The Sichel, $\texttt{SI}(\mu, \sigma, \nu)$, distribution, with $\mu = 1, 3, 5$, $\sigma = 0.2, 0.5, 1$, and $\nu = -1, 1$.

Table 5.23: Sichel distribution, second parameterization.

	`SI`(μ, σ, ν)
	Ranges
Y	$0, 1, 2, 3 \ldots$
μ	$0 < \mu < \infty$
σ	$0 < \sigma < \infty$
ν	$-\infty < \nu < \infty$
	Distribution measures
mean	$b\mu$
variance	$b\mu + b^2\mu^2 g_1$
skewness	$\mu_3/[\mathrm{Var}(Y)]^{1.5}$ where $\mu_3 = b\mu + 3b^2\mu^2 g_1 + b^3\mu^3(g_2 - 3g_1)$
excess kurtosis	$\begin{cases} k_4/[\mathrm{Var}(Y)]^2 \quad \text{where} \\ k_4 = b\mu + 7b^2\mu^2 g_1 + 6b^3\mu^3(g_2 - 3g_1) \\ \qquad +b^4\mu^4(g_3 - 4g_2 + 6g_1 - 3g_1^2) \end{cases}$
PGF	$\dfrac{K_\nu(q)}{(q\sigma)^\nu K_\nu(1/\sigma)}$ where $q^2 = \sigma^{-2} + 2\mu(1-t)\sigma^{-1}$
pf	$\dfrac{\mu^y K_{y+\nu}(\alpha)}{y!(\alpha\sigma)^{y+\nu} K_\nu(1/\sigma)}$ where $\alpha^2 = \sigma^{-2} + 2\mu\sigma^{-1}$
Reference	Reparameterize μ_1 to $b\mu$ in `SICHEL`(μ_1, σ, ν)
Note	Formulae for b, g_1, g_2, and g_3 are given in Table 5.22

5.3.6 Zero-adjusted negative binomial: `ZANBI`

Let $Y = 0$ with probability ν, and $Y = Y_1$ with probability $(1-\nu)$, where $Y_1 \sim$ `NBItr`(μ, σ) and where `NBItr`(μ, σ) is a negative binomial (`NBI`(μ, σ)) truncated at zero. Then Y has a zero-adjusted negative binomial distribution, denoted by `ZANBI`(μ, σ, ν), with probability function

$$P(Y = y \mid \mu, \sigma, \nu) = \begin{cases} \nu & \text{if} \quad y = 0 \\ cP(Y_2 = y \mid \mu, \sigma) & \text{if} \quad y = 1, 2, 3, \ldots \end{cases} \tag{5.25}$$

where $\mu > 0$, $\sigma > 0$, $0 < \nu < 1$, $Y_2 \sim$ `NBI`(μ, σ), $c = (1-\nu)/(1-p_0)$ and $p_0 = P(Y_2 = 0 \mid \mu, \sigma) = (1 + \mu\sigma)^{-1/\sigma}$.

Table 5.24: Zero-adjusted negative binomial distribution.

	$\texttt{ZANBI}(\mu,\sigma,\nu)$
	Ranges
Y	$0,1,2,3\ldots$
μ	$0<\mu<\infty$, mean of NBI component before zero truncation
σ	$0<\sigma<\infty$, dispersion of NBI component before zero truncation
ν	$0<\nu<1,\ \nu=\mathrm{P}(Y=0)$
	Distribution measures
mean ^{a,a2}	$c\mu$ where $c=\dfrac{(1-\nu)}{[1-(1+\mu\sigma)^{-1/\sigma}]}$
variance ^{a,a_2}	$c\mu+c\mu^2(1+\sigma-c)$
skewness^{a,a_2}	$\begin{cases}\mu_3/[\mathrm{Var}(Y)]^{1.5} \ \text{ where}\\ \mu_3=c\mu\left[1+3\mu\,(1+\sigma-c)+\mu^2\left(1+3\sigma+2\sigma^2-3c-3c\sigma+2c^2\right)\right]\end{cases}$
excess kurtosis ^{a,a_2}	$\begin{cases}\mu_4/[\mathrm{Var}(Y)]^2-3 \ \text{ where}\\ \mu_4=c\mu\,[1+\mu\,(7+7\sigma-4c)+\\ \quad 6\mu^2\left(1+3\sigma+2\sigma^2-2c-2c\sigma+c^2\right)+\\ \quad \mu^3\left(1-4c+6c^2-3c^3+6\sigma+11\sigma^2+\right.\\ \quad \left.6\sigma^3-12c\sigma+6c^2\sigma-8c\sigma^2\right)\,]\end{cases}$
PGF ^{a}	$(1-c)+c[1+\mu\sigma(1-t)]^{-1/\sigma}$ for $0<t<1+(\mu\sigma)^{-1}$
pf ^{a}	$\begin{cases}\nu & \text{if} \quad y=0\\ c\mathrm{P}(Y_2=y\,\vert\,\mu,\sigma) & \text{if} \quad y=1,2,3,\ldots\end{cases}$ where $Y_2\sim\texttt{NBI}(\mu,\sigma)$
cdf ^{a}	$\nu+c\left[1-\dfrac{B(y+1,\sigma^{-1},\mu\sigma(1+\mu\sigma)^{-1})}{B(y+1,\sigma^{-1})}-(1+\mu\sigma)^{-1/\sigma}\right]$
Reference	^{a} obtained from equations (??), (??), (??), and (??), where $Y_2\sim\texttt{NBI}(\mu,\sigma)$. ^{a_2} set $\nu=(1-c)$ in moments of $\texttt{ZINBI}(\mu,\sigma,\nu)$

5.3.7 Zero-adjusted Poisson-inverse Gaussian: ZAPIG

Let $Y=0$ with probability ν, and $Y=Y_1$ with probability $(1-\nu)$, where $Y_1\sim$ $\texttt{PIGtr}(\mu,\sigma)$ and where $\texttt{PIGtr}(\mu,\sigma)$ is a Poisson-inverse Gaussian ($\texttt{PIG}(\mu,\sigma)$) truncated at zero. Then Y has a zero-adjusted Poisson-inverse Gaussian distribution, denoted by $\texttt{ZAPIG}(\mu,\sigma,\nu)$, with probability function

$$\mathrm{P}(Y=y\,\vert\,\mu,\sigma,\nu)=\begin{cases}\nu & \text{if} \quad y=0\\ c\mathrm{P}(Y_2=y\,\vert\,\mu,\sigma) & \text{if} \quad y=1,2,3,\ldots\end{cases} \tag{5.26}$$

for $y = 0, 1, 2, 3, \ldots$, where $\mu > 0$, $\sigma > 0$, $0 < \nu < 1$, $Y_2 \sim \texttt{PIG}(\mu, \sigma)$, $c = (1 - \nu)/(1 - p_0)$ and $p_0 = \mathrm{P}(Y_2 = 0 \,|\, \mu, \sigma) = \exp(1/\sigma - \alpha)$, where $\alpha^2 = \sigma^{-2} + 2\mu\sigma^{-1}$ and $\alpha > 0$.

Table 5.25: Zero-adjusted Poisson-inverse Gaussian.

$\texttt{ZAPIG}(\mu, \sigma, \nu)$	
Ranges	
Y	$0, 1, 2, 3 \ldots$
μ	$0 < \mu < \infty$, mean of PIG component before zero truncation
σ	$0 < \sigma < \infty$, dispersion of PIG component before zero truncation
ν	$0 < \nu < 1$, $\nu = \mathrm{P}(Y = 0)$
Distribution measures	
mean [a,a2]	$c\mu$ where $c = (1 - \nu)/[1 - \exp(1/\sigma - \alpha)]$ and $\alpha^2 = \sigma^{-2} + 2\mu\sigma^{-1}$
variance [a,a2]	$c\mu + c\mu^2(1 + \sigma - c)$
skewness [a,a2]	$\begin{cases} \mu_3/[\mathrm{Var}(Y)]^{1.5} \quad \text{where} \\ \mu_3 = c\mu\left[1 + 3\mu(1 + \sigma - c) + \right. \\ \qquad \left. \mu^2\left(1 + 3\sigma + 3\sigma^2 - 3c - 3c\sigma + 2c^2\right)\right] \end{cases}$
excess kurtosis [a,a2]	$\begin{cases} k_4/[\mathrm{Var}(Y)]^2 \text{ where} \\ k_4 = \; c\mu\left[1 + 7\mu(1 + \sigma - c) + \right. \\ \qquad 6\mu^2\left(1 + 3\sigma + 3\sigma^2 - 3c - 3c\sigma + 2c^2\right) \\ \qquad +\mu^3\left(1 - 7c + 12c^2 - 6c^3 + 6\sigma + 15\sigma^2 + \right. \\ \qquad \left.\left. 15\sigma^3 - 18c\sigma + 12c^2\sigma - 15c\sigma^2\right)\right] \end{cases}$
PGF [a]	$(1 - c) + c\exp(1/\sigma - q) \quad \text{where } q^2 = \sigma^{-2} + 2\mu(1 - t)\sigma^{-1}$
pf [a]	$\begin{cases} \nu & \text{if} \quad y = 0 \\ c\mathrm{P}(Y_2 = y \,\|\, \mu, \sigma) & \text{if} \quad y = 1, 2, 3, \ldots \end{cases}$ where $Y_2 \sim \texttt{PIG}(\mu, \sigma)$.
Reference	[a] obtained from equations (??), (??), and (??), where $Y_2 \sim \texttt{PIG}(\mu, \sigma)$. [a2] set $\nu = (1 - c)$ in moments of $\texttt{ZIPIG}(\mu, \sigma, \nu)$

5.3.8 Zero-inflated negative binomial: ZINBI

Let $Y = 0$ with probability ν and $Y = Y_1$ with probability $(1 - \nu)$, where $Y_1 \sim \texttt{NBI}(\mu, \sigma)$. Then Y has a zero-inflated negative binomial distribution, denoted by ZINBI(μ, σ, ν), with probability function

$$P(Y = y \mid \mu, \sigma, \nu) = \begin{cases} \nu + (1-\nu)\,P(Y_1 = 0 \mid \mu, \sigma) & \text{if} \quad y = 0 \\ (1-\nu)P(Y_1 = y \mid \mu, \sigma) & \text{if} \quad y = 1, 2, 3, \ldots \end{cases} \tag{5.27}$$

for $\mu > 0$, $\sigma > 0$, $0 < \nu < 1$, where $Y_1 \sim$ NBI(μ, σ) and so $P(Y_1 = 0 \mid \mu, \sigma) = (1 + \mu\sigma)^{-1/\sigma}$. The mean is given by $\text{E}(Y) = (1 - \nu)\,\mu$ and the variance by $\text{Var}(Y) = \mu\,(1 - \nu) + \mu^2\,(1 - \nu)\,(\sigma + \nu)$. Hence $\text{Var}(Y) = \text{E}(Y) + [\text{E}(Y)]^2\,(\sigma + \nu)/(1 - \nu)$.

Table 5.26: Zero-inflated negative binomial distribution.

	ZINBI(μ, σ, ν)
	Ranges
Y	$0, 1, 2, 3 \ldots$
μ	$0 < \mu < \infty$, mean of the negative binomial component
σ	$0 < \sigma < \infty$, dispersion of the negative binomial component
ν	$0 < \nu < 1$, inflation probability at zero
	Distribution measures
mean [a]	$(1-\nu)\mu$
variance [a]	$\mu(1-\nu) + \mu^2(1-\nu)(\sigma+\nu)$
skewness [a]	$\begin{cases} \mu_3/[\text{Var}(Y)]^{1.5} \quad \text{where} \\ \mu_3 = \mu(1-\nu)\left[1 + 3\mu(\sigma+\nu) + \mu^2\left(2\sigma^2 + 3\sigma\nu + 2\nu^2 - \nu\right)\right] \end{cases}$
excess kurtosis [a]	$\begin{cases} \mu_4/[\text{Var}(Y)]^2 - 3 \quad \text{where} \\ \mu_4 = \mu(1-\nu)\left[1 + \mu(3 + 7\sigma + 4\nu) + 6\mu^2\left(\sigma + 2\sigma^2 + 2\sigma\nu + \nu^2\right) + \right. \\ \left. \mu^3\left(3\sigma^2 + 6\sigma^3 + 6\sigma\nu^2 + 8\sigma^2\nu + \nu - 3\nu^2 + 3\nu^3\right)\,\right] \end{cases}$
PGF [a]	$\nu + (1-\nu)[1 + \mu\sigma(1-t)]^{-1/\sigma}$ for $0 < t < 1 + (\mu\sigma)^{-1}$
pf [a]	$\begin{cases} \nu + (1-\nu)\,P(Y_1 = 0 \mid \mu, \sigma) & \text{if } y = 0 \\ (1-\nu)P(Y_1 = y \mid \mu, \sigma) & \text{if } y = 1, 2, 3, \ldots \end{cases}$ where $Y_1 \sim$ NBI(μ, σ)
cdf [a]	$1 - \dfrac{(1-\nu)B(y+1, \sigma^{-1}, \mu\sigma(1+\mu\sigma)^{-1})}{B(y+1, \sigma^{-1})}$
Reference	[a]Obtained from equations (??), (??), (??) and (??), where $Y_1 \sim$ NBI(μ, σ)

5.3.9 Zero-inflated Poisson-inverse Gaussian: `ZIPIG`

Let $Y = 0$ with probability ν, and $Y = Y_1$ with probability $(1-\nu)$, where $Y_1 \sim$ `PIG`(μ, σ). Then Y has a zero-inflated Poisson-inverse Gaussian distribution, denoted by `ZIPIG`(μ, σ, ν), with probability function

$$P(Y = y \mid \mu, \sigma, \nu) = \begin{cases} \nu + (1-\nu)\, P(Y_1 = 0 \mid \mu, \sigma) & \text{if} \quad y = 0 \\ (1-\nu)\, P(Y_1 = y \mid \mu, \sigma) & \text{if} \quad y = 1, 2, 3, \ldots \end{cases} \tag{5.28}$$

for $\mu > 0$, $\sigma > 0$ and $0 < \nu < 1$, where $Y_1 \sim$ `PIG`(μ, σ) and so $P(Y_1 = 0|\mu, \sigma) = \exp(1/\sigma - \alpha)$, where $\alpha^2 = \sigma^{-2} + 2\mu\sigma^{-1}$ and $\alpha > 0$. The mean of Y is given by $E(Y) = (1-\nu)\,\mu$ and the variance by $\mathrm{Var}(Y) = \mu(1-\nu) + \mu^2(1-\nu)(\sigma+\nu)$. Hence $\mathrm{Var}(Y) = E(Y) + [E(Y)]^2\,(\sigma+\nu)/(1-\nu)$.

Table 5.27: Zero-inflated Poisson-inverse Gaussian distribution.

	`ZIPIG`(μ, σ, ν)
	Ranges
Y	$0, 1, 2, 3 \ldots$
μ	$0 < \mu < \infty$, mean of the `PIG` component
σ	$0 < \sigma < \infty$, dispersion of the `PIG` component
ν	$0 < \nu < 1$, inflation probability at zero
	Distribution measures
mean[a]	$(1-\nu)\mu$
variance[a]	$\mu(1-\nu) + \mu^2(1-\nu)(\sigma+\nu)$
skewness[a]	$\begin{cases} \mu_3/[\mathrm{Var}(Y)]^{1.5} \quad \text{where} \\ \mu_3 = \mu(1-\nu)[1 + 3\mu(\sigma+\nu) + \mu^2(3\sigma^2 + 3\sigma\nu + 2\nu^2 - \nu)] \end{cases}$
excess kurtosis [a]	$\begin{cases} k_4/[\mathrm{Var}(Y)]^2 \quad \text{where} \\ k_4 = \mu(1-\nu)\left[1 + 7\mu(\sigma+\nu) + 6\mu^2\left(3\sigma^2 + 3\sigma\nu + 2\nu^2 - \nu\right) + \right. \\ \qquad \left. \mu^3\left(\nu - 6\nu^2 + 6\nu^3 + 15\sigma^3 - 6\sigma\nu + 12\sigma\nu^2 + 15\sigma^2\nu\right)\right] \end{cases}$
PGF[a]	$\nu + (1-\nu)e^{(1/\sigma)-q}$ where $q^2 = \sigma^{-2} + 2\mu(1-t)\sigma^{-1}$
pf[a]	$\begin{cases} \nu + (1-\nu)\, P(Y_1 = 0 \mid \mu, \sigma) & \text{if} \quad y = 0 \\ (1-\nu)\, P(Y_1 = y \mid \mu, \sigma) & \text{if} \quad y = 1, 2, 3, \ldots \end{cases}$ where $Y_1 \sim$ `PIG`(μ, σ)
Reference	[a]Obtained from equations (??), (??), and (??), $Y_1 \sim$ `PIG`(μ, σ)

5.4 Four-parameter count distributions

5.4.1 Poisson shifted generalized inverse Gaussian: PSGIG

The Poisson shifted generalized inverse Gaussian distribution, [Rigby et al., 2008], denoted by PSGIG(μ, σ, ν, τ), has probability function given by

$$\mathrm{P}(Y = y \,|\, \mu, \sigma, \nu, \tau) \quad = \quad \frac{e^{-\mu\tau} T}{K_\nu(1/\sigma)} \tag{5.29}$$

where $y = 0, 1, 2, 3, \ldots$, $\mu > 0$, $\sigma > 0$, $-\infty < \nu < \infty$, $0 < \tau < 1$,

$$T \quad = \quad \sum_{j=0}^{y} \binom{y}{j} \frac{\mu^y \tau^{y-j} K_{\nu+j}(\delta)}{y! d^j (\delta\sigma)^{\nu+j}} \ , \tag{5.30}$$

$d = b/(1-\tau)$, $b = K_{\nu+1}(1/\sigma)/K_\nu(1/\sigma)$ and $\delta^2 = \sigma^{-2} + 2\mu(d\sigma)^{-1}$. Note that

$$\sigma = \left[\left(\frac{\mu^2}{d^2} + \delta^2 \right)^{1/2} - \frac{\mu}{d} \right]^{-1} .$$

As $y \to \infty$, $\log \mathrm{P}(Y = y \,|\, \mu, \sigma, \nu, \tau) \sim -y \log \left[1 + \frac{b}{2\mu\sigma(1-\tau)} \right]$. Note that the distribution is currently under development in **gamlss.dist**.

Table 5.28: Poisson shifted generalized inverse Gaussian distribution.

	$\texttt{PSGIG}(\mu, \sigma, \nu, \tau)$
	Ranges
Y	$0, 1, 2, 3 \ldots$
μ	$0 < \mu < \infty$, mean
σ	$0 < \sigma < \infty$
ν	$-\infty < \nu < \infty$
τ	$0 < \tau < 1$
	Distribution measures
mean	μ
variance	$\mu + (1-\tau)^2 \mu^2 g_1$ where $g_1 = 2\sigma(\nu+1)b^{-1} + b^{-2} - 1$
skewness	$\begin{cases} \mu_3/[\mathrm{Var}(Y)]^{1.5} \text{ where} \\ \mu_3 = \mu + 3(1-\tau)^2\mu^2 g_1 + (1-\tau)^3\mu^3(g_2 - 3g_1) \\ g_2 = 2\sigma(\nu+2)b^{-3} + [4\sigma^2(\nu+1)(\nu+2)+1]b^{-2} - 1 \end{cases}$
excess kurtosis	$\begin{cases} k_4/[\mathrm{Var}(Y)]^2 \quad \text{where} \\ k_4 = \mu + 7(1-\tau)^2\mu^2 g_1 + 6(1-\tau)^3\mu^3(g_2 - 3g_1) + \\ \qquad (1-\tau)^4\mu^4(g_3 - 4g_2 + 6g_1 - 3g_1^2) \\ g_3 = [1 + 4\sigma^2(\nu+2)(\nu+3)]b^{-4} + \\ \quad [8\sigma^3(\nu+1)(\nu+2)(\nu+3) + 4\sigma(\nu+2)]b^{-3} - 1 \end{cases}$
PGF	$\dfrac{\exp[\mu\tau(t-1)]K_\nu(r)}{(r\sigma)^\nu K_\nu(1/\sigma)}$
	where $r^2 = \sigma^{-2} + 2\mu(1-t)(d\sigma)^{-1}$ and $d = b/(1-\tau)$
pf	$\dfrac{\exp(-\mu t)T}{K_\nu(1/\sigma)}$ where $T = \displaystyle\sum_{j=0}^{y} \binom{y}{j} \frac{\mu^y \tau^{y-j} K_{\nu+j}(\delta)}{y! d^j (\delta\sigma)^{\nu+j}}$
	where $\delta^2 = \sigma^{-2} + 2\mu(d\sigma)^{-1}$
Reference	Rigby et al. [2008]
Notes	$b = K_{\nu+1}(1/\sigma)/K_\nu(1/\sigma)$
	$K_\lambda(t)$ is the modified Bessel function of the second kind

5.4.2 Zero-adjusted beta negative binomial: ZABNB

Let $Y = 0$ with probability τ, and $Y = Y_1$ with probability $(1-\tau)$, where $Y_1 \sim \texttt{BNBtr}(\mu, \sigma, \nu)$, i.e. the beta negative binomial ($\texttt{BNB}(\mu, \sigma, \tau)$) distribution truncated at zero. Then Y has a zero-adjusted beta negative binomial distribution, denoted by ZABNB(μ, σ, ν, τ), with probability function

$$P(Y = y \mid \mu, \sigma, \nu, \tau) = \begin{cases} \tau & \text{if } \; y = 0 \\ cP(Y_2 = y \mid \mu, \sigma, \nu) & \text{if } \; y = 1, 2, 3, \ldots \end{cases} \tag{5.31}$$

where $\mu > 0$, $\sigma > 0$, $\nu > 0$, $0 < \tau < 1$, $Y_2 \sim$ BNB(μ, σ, ν) and $c = (1-\tau)/(1-p_0)$ where

$$p_0 = P(Y_2 = 0 \mid \mu, \sigma, \nu) = \frac{B(\mu\sigma^{-1}\nu, \sigma^{-1} + \nu^{-1} + 1)}{B(\mu\sigma^{-1}\nu, \sigma^{-1} + 1)} \; .$$

The moments of Y can be obtained from the moments of Y_2 using equation (**??**), from which the mean, variance, skewness, and excess kurtosis of Y can be found. In particular the mean of Y is $E(Y) = c\mu$.

Table 5.29: Zero-adjusted beta negative binomial.

ZABNB(μ, σ, ν, τ)	
Ranges	
Y	$0, 1, 2, 3 \ldots$
μ	$0 < \mu < \infty$, mean of BNB component before zero truncation
σ	$0 < \sigma < \infty$
ν	$0 < \nu < \infty$
τ	$0 < \tau < 1, \; \tau = P(Y = 0)$
Distribution measures	
mean [a,a2]	$c\mu$ where $c = (1-\tau)/(1-p_0)$ and $p_0 = P(Y_2 = 0 \mid \mu, \sigma, \nu)$ where $Y_2 \sim$ BNB(μ, σ, ν)
variance [a,a2]	$\begin{cases} c\mu(1+\mu\nu)(1+\sigma\nu^{-1})(1-\sigma)^{-1} + c(1-c)\mu^2 & \text{for } \; \sigma < 1 \\ \infty & \text{for } \; \sigma \geq 1 \end{cases}$
pf [a]	$\begin{cases} \tau & \text{if } \; y = 0 \\ cP(Y_2 = y \mid \mu, \sigma, \nu) & \text{if } \; y = 1, 2, 3, \ldots \end{cases}$ where $Y_2 \sim$ BNB(μ, σ, ν)
Reference	[a] Obtained from equations (**??**) and (**??**). where $Y_2 \sim$ BNB(μ, σ, ν). [a2] set $\nu = (1-c)$ in moments of ZIBNB(μ, σ, ν)

5.4.3 Zero-adjusted Sichel: ZASICHEL

Let $Y = 0$ with probability τ, and $Y = Y_1$ with probability $(1-\tau)$, where $Y_1 \sim$ SICHELtr(μ, σ, ν), i.e. SICHEL(μ, σ, ν) truncated at zero. Then Y has a zero-adjusted Sichel distribution, denoted by ZASICHEL(μ, σ, ν, τ), with probability function

$$\mathrm{P}(Y = y \mid \mu, \sigma, \nu) = \begin{cases} \tau & \text{if} \quad y = 0 \\ c\mathrm{P}(Y_2 = y \mid \mu, \sigma, \nu) & \text{if} \quad y = 1, 2, 3, \ldots \end{cases} \tag{5.32}$$

where $\mu > 0$, $\sigma > 0$, $-\infty < \nu < \infty$, $0 < \tau < 1$, $Y_2 \sim \texttt{SICHEL}(\mu, \sigma, \nu)$ and $c = (1-\tau)/(1-p_0)$ where

$$p_0 = \mathrm{P}(Y_2 = 0 \mid \mu, \sigma, \nu) = \frac{K_\nu(\alpha)}{(\alpha\sigma)^\nu K_\nu(1/\sigma)}$$

and $\alpha^2 = \sigma^{-2} + 2\mu(b\sigma)^{-1}$, $b = K_{\nu+1}(1/\sigma)/K_\nu(1/\sigma)$.

Table 5.30: Zero-adjusted Sichel distribution.

$\texttt{ZASICHEL}(\mu, \sigma, \nu, \tau)$	
Ranges	
Y	$0, 1, 2, 3 \ldots$
μ	$0 < \mu < \infty$, mean of SICHEL component before zero truncation
σ	$0 < \sigma < \infty$
ν	$-\infty < \nu < \infty$
τ	$0 < \tau < 1,\ \tau = \mathrm{P}(Y = 0)$
Distribution measures	
mean [a,a2]	$c\mu$ where $c = (1-\tau)/(1-p_0)$ and $p_0 = \mathrm{P}(Y_2 = 0)$ where $Y_2 \sim \texttt{SICHEL}(\mu, \sigma, \nu)$
variance [a,a2]	$c\mu + c^2\mu^2 h_1$ where $h_1 = c^{-1}\left[2\sigma(\nu+1)b^{-1} + b^{-2}\right] - 1$
skewness [a,a2]	$\begin{cases} \mu_3/[\mathrm{Var}(Y)]^{1.5} \quad \text{where} \\ \mu_3 = c\mu + 3c^2\mu^2 h_1 + c^3\mu^3(h_2 - 3h_1 - 1) \\ h_2 = c^{-2}\left\{2\sigma(\nu+2)b^{-3} + \left[4\sigma^2(\nu+1)(\nu+1)+1\right]b^{-2}\right\} \end{cases}$
excess kurtosis [a,a2]	$\begin{cases} k_4/[\mathrm{Var}(Y)]^2 \quad \text{where} \\ k_4 = c\mu + 7c^2\mu^2 h_1 + 6c^3\mu^3(h_2 - 3h_1 - 1) + \\ \qquad c^4\mu^4\left(h_3 - 4h_2 + 6h_1 - 3h_1^2 + 3\right) \\ h_3 = c^{-3}\left\{b^{-4}\left[1 + 4\sigma^2(\nu+2)(\nu+3)\right] + \right. \\ \qquad \left. b^{-3}\left[8\sigma^3(\nu+1)(\nu+2)(\nu+3) + 4\sigma(\nu+2)\right]\right\} \end{cases}$
PGF [a]	$(1-c) + \dfrac{cK_\nu(q)}{(q\sigma)^\nu K_\nu(1/\sigma)}$ where $q^2 = \sigma^{-2} + 2\mu(1-\tau)(b\sigma)^{-1}$
pf [a]	$\begin{cases} \tau & \text{if } y = 0 \\ c\mathrm{P}(Y_2 = y \mid \mu, \sigma, \nu) & \text{if } y = 1, 2, 3, \ldots \end{cases}$ where $Y_2 \sim \texttt{SICHEL}(\mu, \sigma, \nu)$
Reference	[a] Obtained from equations (**??**), (**??**), and (**??**) where $Y_2 \sim \texttt{SICHEL}(\mu, \sigma, \nu)$. [a2] set $\tau = (1-c)$ in moments of $\texttt{ZISICHEL}(\mu, \sigma, \nu, \tau)$.
Notes	b and $K_\lambda(t)$ are defined in Table 5.22.

5.4.4 Zero-inflated beta negative binomial: ZIBNB

Let $Y = 0$ with probability τ, and $Y = Y_1$ with probability $(1-\tau)$ where $Y_1 \sim \texttt{BNB}(\mu, \sigma, \nu)$. Then Y has a zero-inflated beta negative binomial distribution, denoted by ZIBNB(μ, σ, ν, τ), with probability function

$$P(Y = y \mid \mu, \sigma, \nu, \tau) = \begin{cases} \tau + (1-\tau)\, P(Y_1 = 0 \mid \mu, \sigma, \nu) & \text{if } y = 0 \\ (1-\tau) P(Y_1 = y \mid \mu, \sigma, \nu) & \text{if } y = 1, 2, 3, \ldots \end{cases} \tag{5.33}$$

for $\mu > 0$, $\sigma > 0$, $\nu > 0$, $0 < \tau < 1$ and $Y_1 \sim$ BNB(μ, σ, ν). The moments of Y can be obtained from the moments of Y_1, using equations (**??**), from which the mean, variance, skewness, and excess kurtosis of Y can be found. In particular the mean of Y is $E(Y) = (1-\tau)\mu$.

Table 5.31: Zero-inflated beta negative binomial distribution.

ZIBNB(μ, σ, ν, τ)	
Ranges	
Y	$0, 1, 2, 3 \ldots$
μ	$0 < \mu < \infty$, mean of the BNB component
σ	$0 < \sigma < \infty$
ν	$0 < \nu < \infty$
τ	$0 < \tau < 1$, inflation probability at zero
Distribution measures	
mean	$(1-\tau)\,\mu$
variance	$(1-\tau)\mu(1+\mu\nu)(1+\sigma\nu^{-1})(1-\sigma)^{-1} + \tau(1-\tau)\mu^2$
pf	$\begin{cases} \tau + (1-\tau)\, P(Y_1 = 0 \mid \mu, \sigma, \nu) & \text{if } y = 0 \\ (1-\tau) P(Y_1 = y \mid \mu, \sigma, \nu) & \text{if } y = 1, 2, 3, \ldots \end{cases}$ where $Y_1 \sim$ BNB(μ, σ, ν)
Reference	Obtained from equations (**??**) and (**??**), where $Y_1 \sim$ BNB(μ, σ, ν)

5.4.5 Zero-inflated Sichel: ZISICHEL

Let $Y = 0$ with probability τ, and $Y = Y_1$, with probability $(1-\tau)$ where $Y_1 \sim$ SICHEL(μ, σ, ν). Then Y has a zero-inflated Sichel distribution, denoted by ZISICHEL(μ, σ, ν, τ), with probability function

$$P(Y = y \mid \mu, \sigma, \nu, \tau) = \begin{cases} \tau + (1-\tau)\, P(Y_1 = 0 \mid \mu, \sigma, \nu) & \text{if } y = 0 \\ (1-\tau) P(Y_1 = y \mid \mu, \sigma, \nu) & \text{if } y = 1, 2, 3, \ldots \end{cases} \tag{5.34}$$

for $\mu > 0$, $\sigma > 0$, $-\infty < \nu < \infty$ and $0 < \tau < 1$, and $Y_1 \sim$ SICHEL(μ, σ, ν).

Table 5.32: Zero-inflated Sichel distribution.

ZISICHEL(μ, σ, ν, τ)	
Ranges	
Y	$0, 1, 2, 3 \ldots$
μ	$0 < \mu < \infty$, mean of the SICHEL component
σ	$0 < \sigma < \infty$
ν	$-\infty < \nu < \infty$
τ	$0 < \tau < 1,$ inflation probability at zero
Distribution measures	
mean[a]	$(1-\tau)\,\mu$
variance[a]	$\begin{cases} (1-\tau)\,\mu + (1-\tau)^2\,\mu^2 h_1 \\ \text{where } h_1 = (1-\tau)^{-1}\left[2\sigma\,(\nu+1)b^{-1} + b^{-2}\right] - 1 \end{cases}$
skewness[a]	$\begin{cases} \mu_3/\left[\mathrm{Var}(Y)\right]^{1.5} \text{ where} \\ \mu_3 = (1-\tau)\,\mu + 3\,(1-\tau)^2\,\mu^2 h_1 + (1-\tau)^3\,\mu^3\,(h_2 - 3h_1 - 1) \\ h_2 = (1-\tau)^{-2}\left\{2\sigma\,(\nu+2)/b^3 + \left[4\sigma^2\,(\nu+1)\,(\nu+2)+1\right]/b^2\right\} \end{cases}$
excess kurtosis [a]	$\begin{cases} k_4/\left[\mathrm{Var}(Y)\right]^2 \text{ where} \\ k_4 = (1-\tau)\,\mu + 7\,(1-\tau)^2\,\mu^2 h_1 + 6\,(1-\tau)^3\,\mu^3\,(h_2 - 3h_1 - 1) + \\ \qquad (1-\tau)^4\,\mu^4\left(h_3 - 4h_2 + 6h_1 - 3h_1^2 + 3\right) \\ h_3 = (1-\tau)^{-3}\left\{\left[1 + 4\sigma^2\,(\nu+2)\,(\nu+3)\right]/b^4 + \right. \\ \qquad \left.\left[8\sigma^3\,(\nu+1)\,(\nu+2)\,(\nu+3) + 4\sigma\,(\nu+2)\right]/b^3\right\} \end{cases}$
PGF[a]	$\tau + (1-\tau)\dfrac{K_\nu(q)}{(q\sigma)^\nu K_\nu(1/\sigma)}$ where $q^2 = \sigma^{-2} + 2\mu(1-t)(b\sigma)^{-1}$
pf[a]	$\begin{cases} \tau + (1-\tau)\,\mathrm{P}(Y_1 = 0 \mid \mu, \sigma, \nu) & \text{if } y = 0 \\ (1-\tau)\mathrm{P}(Y_1 = y \mid \mu, \sigma, \nu) & \text{if } y = 1, 2, 3, \ldots \end{cases}$ where $Y_1 \sim$ SICHEL(μ, σ, ν)
Reference	[a]Obtained from equations (**??**), (**??**) and (**??**), where $Y_1 \sim$ SICHEL(μ, σ, ν)
Note	b and $K_\nu(\cdot)$ are defined in Table 5.22.

Chapter 6

Binomial type distributions and multinomial distributions

This chapter gives summary tables and plots for the explicit **gamlss.dist** discrete binomial type distributions with range $\{0, 1, \ldots, n\}$, in which it is assumed throughout that n is a known positive integer. These are discussed in Chapter **??**. This chapter also covers the explicit **gamlss.dist** multinomial distributions.

6.1 One-parameter binomial type distribution

6.1.1 Binomial: `BI`

The probability function of the binomial distribution, denoted as `BI`(n, μ), is given by

$$\begin{aligned} \mathrm{P}(Y = y|n, \mu) &= \binom{n}{y} \mu^y (1-\mu)^{n-y} \\ &= \frac{n!}{y!(n-y)!}\, \mu^y (1-\mu)^{n-y} \end{aligned} \tag{6.1}$$

for $y = 0, 1, 2, \ldots, n$, where $0 < \mu < 1$ and n is a known positive integer.

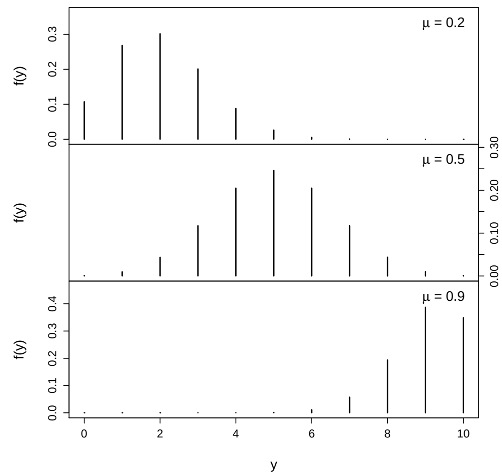

Figure 6.1: The binomial, BI(n, μ), distribution, with $\mu = 0.2, 0.5, 0.9$ and binomial denominator $n = 10$.

Table 6.1: Binomial distribution.

	BI(n, μ)	
	Ranges	
Y	$0, 1, 2, \ldots, n$ where n is a known positive integer	
μ	$0 < \mu < 1$	
	Distribution measures	
mean	$n\mu$	
mode	$\begin{cases} \lfloor (n+1)\mu \rfloor & \text{if } (n+1)\mu \text{ is not an integer} \\ 0 & \text{if } \mu < 1/(n+1) \\ n & \text{if } \mu > n/(n+1) \\ (n+1)\mu - 1 \text{ and } (n+1)\mu, & \text{if } (n+1)\mu \text{ is an integer} \end{cases}$	
variance	$n\mu(1-\mu)$	
skewness	$\dfrac{(1-2\mu)}{[n\mu(1-\mu)]^{0.5}}$	
excess kurtosis	$\dfrac{1-6\mu(1-\mu)}{n\mu(1-\mu)}$	
PGF	$(1-\mu+\mu t)^n$	
pf	$\binom{n}{y}\mu^y(1-\mu)^{n-y}$	
cdf	$1 - \dfrac{B(y+1, n-y, \mu)}{B(y+1, n-y)}$	
Reference	Johnson et al. [2005] p108, 110, 112, 113, with $p = \mu$	
Note	$\lfloor x \rfloor$ is the floor function or integer part of x	

6.2 Two-parameter binomial type distributions

6.2.1 Beta binomial: BB

The probability function of the beta binomial distribution, denoted as BB(n, μ, σ), is given by

$$
\begin{aligned}
P(Y = y|n, \mu, \sigma) &= \binom{n}{y} \frac{B\left(y + \mu/\sigma, n + (1-\mu)/\sigma - y\right)}{B(\mu/\sigma, (1-\mu)/\sigma)} \qquad (6.2) \\
&= \frac{\Gamma(n+1)}{\Gamma(y+1)\Gamma(n-y+1)} \frac{\Gamma(\frac{1}{\sigma})\Gamma(y+\frac{\mu}{\sigma})\Gamma(n+\frac{(1-\mu)}{\sigma}-y)}{\Gamma(n+\frac{1}{\sigma})\Gamma(\frac{\mu}{\sigma})\Gamma(\frac{1-\mu}{\sigma})}
\end{aligned}
$$

for $y = 0, 1, \ldots, n$, where $0 < \mu < 1$, $\sigma > 0$, and n is a known positive integer.

The binomial BI(n, μ) distribution is the limiting distribution of BB(n, μ, σ) as $\sigma \to 0$. For $\mu = 0.5$ and $\sigma = 0.5$, BB(n, μ, σ) is a discrete uniform distribution.

The probability function (6.2) and the mean, variance, skewness, and kurtosis in Table 6.2 can be obtained from Johnson et al. [2005], p259 equation (6.31) and p263, by setting $a = -\mu/\sigma$ and $b = -(1-\mu)/\sigma$. The equivalence of their probability function to (6.2) is shown using their equation (1.24).

The $\texttt{BB}(n, \mu, \sigma)$ distribution can be derived as a beta mixture of binomial distributions, by assuming $Y|\pi \sim \texttt{BI}(n, \pi)$ where $\pi \sim \texttt{BEo}(\frac{\mu}{\sigma}, \frac{1-\mu}{\sigma})$. Hence the $\texttt{BB}(n, \mu, \sigma)$ is an overdispersed binomial distribution.

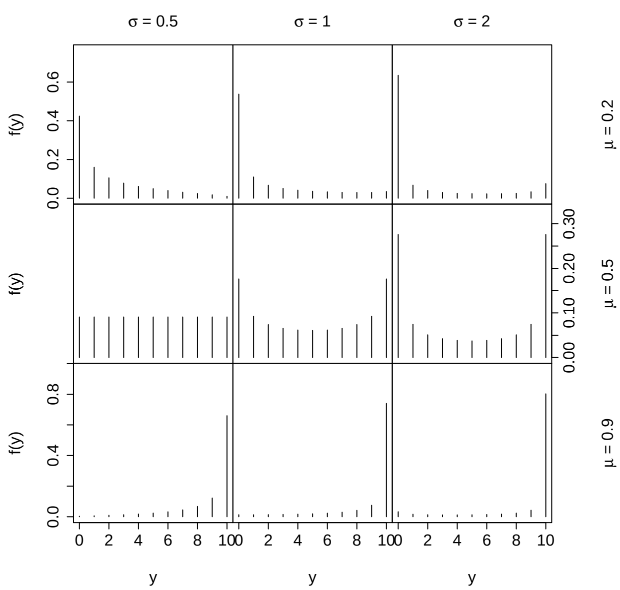

Figure 6.2: The beta binomial, $\texttt{BB}(n, \mu, \sigma)$, distribution, with $\mu = 0.2, 0.5, 0.9$, $\sigma = 0.5, 1, 2$ and binomial denominator $n = 10$.

Table 6.2: Beta binomial distribution.

	BB(n, μ, σ)
	Ranges
Y	$0, 1, 2, \ldots, n$ where n is a known positive integer
μ	$0 < \mu < 1$
σ	$0 < \sigma < \infty$
	Distribution measures
mean[a]	$n\mu$
mode	$\begin{cases} \lfloor \frac{(n+1)(\mu-\sigma)}{1-2\sigma} \rfloor & \text{if } \sigma < \min(\sigma_0, \sigma_1) \\ 0 & \text{if } \sigma_0 < \sigma < \sigma_1 \\ n & \text{if } \sigma_1 < \sigma < \sigma_0 \\ 0 \text{ and } n & \text{if } \sigma > \max(\sigma_0, \sigma_1) \\ \text{where} & \\ \sigma_0 = [(n+1)\mu - 1]/(n-1) & \\ \sigma_1 = [(n+1)(1-\mu) - 1]/(n-1) & \end{cases}$
variance[a]	$n\mu(1-\mu)\left[1 + \frac{(n-1)\sigma}{(1+\sigma)}\right]$
skewness[a]	$\left[\frac{(1+\sigma)}{n\mu(1-\mu)(1+n\sigma)}\right]^{0.5} \frac{(1-2\mu)(1+2n\sigma)}{(1+2\sigma)}$
excess kurtosis[a]	$\begin{cases} \beta_2 - 3 \text{ where} \\ \beta_2 = \frac{(1+\sigma)}{n\mu(1-\mu)(1+n\sigma)(1+2\sigma)(1+3\sigma)} [(1+6n\sigma - \\ \quad \sigma + 6n^2\sigma^2) - 3\mu(1-\mu)(6n^2\sigma^2 + 6n\sigma - n^2\sigma - n + 2)] \end{cases}$
pf[a]	$\binom{n}{y} \frac{B(y + \mu/\sigma, n + (1-\mu)/\sigma - y)}{B(\mu/\sigma, (1-\mu)/\sigma)}$
Reference	[a]Johnson et al. [2005] p259, 263 with $a = -\mu/\sigma$ and $b = -(1-\mu)/\sigma$

6.2.2 Double binomial: DBI

The probability function of the double binomial distribution, denoted by DBI(n, μ, σ), is given by

$$P(Y = y \mid n, \mu, \sigma) = \frac{c(\mu, \sigma) n! y^y (n-y)^{n-y} n^{n/\sigma} \mu^{y/\sigma} (1-\mu)^{(n-y)/\sigma}}{y!(n-y)! n^n y^{y/\sigma} (n-y)^{(n-y)/\sigma}} \tag{6.3}$$

for $y = 0, 1, \ldots, n$, where $0 < \mu < 1$, $\sigma > 0$, n is a known positive integer, and $c(\mu, \sigma)$ is a normalizing constant (ensuring that the distribution probabilities sum to one) given by

$$c(\mu,\sigma) = \left[\sum_{y=0}^{n} \frac{n! y^y (n-y)^{n-y} n^{n/\sigma} \mu^{y/\sigma} (1-\mu)^{(n-y)/\sigma}}{y!(n-y)! n^n y^{y/\sigma} (n-y)^{(n-y)/\sigma}} \right]^{-1} \tag{6.4}$$

obtained from Lindsey [1995, p131], reparameterized by $\pi = \mu$ and $\psi = 1/\sigma$.

The DBI(n, μ, σ) distribution is a special case of the double exponential family of Efron [1986]. It has approximate mean $n\mu$ and approximate variance $n\sigma\mu(1-\mu)$. The DBI(n, μ, σ) distribution is a binomial (BI(n, μ)) distribution if $\sigma = 1$. It is (approximately) an overdispersed binomial if $\sigma > 1$ and underdispersed if $\sigma < 1$. Unlike some other implementations, **gamlss.dist** calculates $c(\mu, \sigma)$ using a finite sum with a very large number of terms in equation (6.4), rather than a potentially less accurate functional approximation of $c(\mu, \sigma)$.

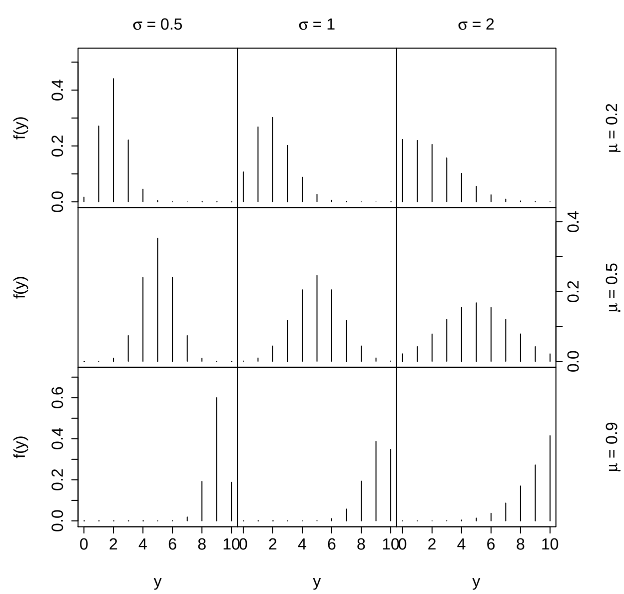

Figure 6.3: The double binomial, DBI(n, μ, σ), distribution, with $\mu = 0.2, 0.5, 0.9$, $\sigma = .5, 1, 2$, and binomial denominator $n = 10$.

Table 6.3: Double binomial distribution.

$\texttt{DBI}(n, \mu, \sigma)$	
Ranges	
Y	$0, 1, 2, \ldots, n$ where n is a known positive integer
μ	$0 < \mu < 1$
σ	$0 < \sigma < \infty$
Distribution measures	
mean[a]	$n\mu$
variance[a]	$n\sigma\mu(1-\mu)$
pf[a2]	$\dfrac{c(\mu,\sigma) n! y^y (n-y)^{n-y} n^{n/\sigma} \mu^{y/\sigma} (1-\mu)^{(n-y)/\sigma}}{y!(n-y)! n^n y^{y/\sigma} (n-y)^{(n-y)/\sigma}}$ where $c(\mu,\sigma)$ is given by equation (6.4)
Reference	[a2] Lindsey [1995] p131, where $\pi = \mu$ and $\psi = 1/\sigma$
Note	[a] approximate

6.2.3 Zero-adjusted binomial: ZABI

Let $Y = 0$ with probability σ, and $Y = Y_1$ with probability $(1-\sigma)$, where $Y_1 \sim \texttt{BItr}(n, \mu)$ a binomial distribution truncated at zero. Then Y has a zero-adjusted (or altered) binomial distribution, denoted by $\texttt{ZABI}(n, \mu, \sigma)$, given by

$$P(Y = y \mid n, \mu, \sigma) = \begin{cases} \sigma & \text{if } \ y = 0 \\ \dfrac{(1-\sigma)\, n! \mu^y\, (1-\mu)^{n-y}}{\left[1-(1-\mu)^n\right] y!\, (n-y)!} & \text{if } \ y = 1, 2, 3, \ldots \end{cases} \tag{6.5}$$

for $0 < \mu < 1$, $0 < \sigma < 1$ and n is a known positive integer. The mean and variance of Y are given by

$$\begin{aligned} \mathrm{E}(Y) &= \frac{(1-\sigma)\, n\mu}{\left[1-(1-\mu)^n\right]} \\ \mathrm{Var}(Y) &= \frac{n\mu\,(1-\sigma)\,(1-\mu+n\mu)}{\left[1-(1-\mu)^n\right]} - \left[\mathrm{E}\,(Y)\right]^2 \ . \end{aligned}$$

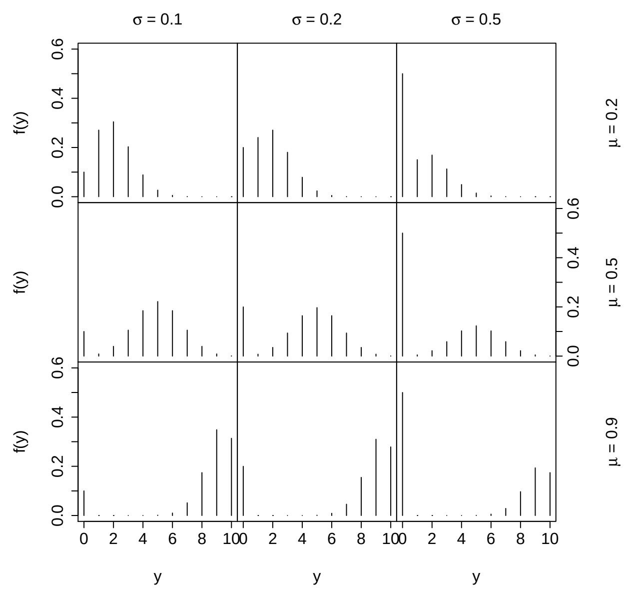

Figure 6.4: The zero-adjusted binomial, ZABI(n, μ, σ), distribution, with $\mu = 0.2, 0.5, 0.9$, $\sigma = 0.1, 0.2, 0.5$, and binomial denominator $n = 10$.

6.2.4 Zero-inflated binomial: ZIBI

Let $Y = 0$ with probability σ, and $Y = Y_1$ with probability $(1-\sigma)$, where $Y_1 \sim$ BI(n, μ), then Y has a zero-inflated binomial distribution, denoted by ZIBI(n, μ, σ), given by

$$P(Y = y \mid n, \mu, \sigma) \begin{cases} \sigma + (1-\sigma)\,(1-\mu)^n & \text{if } y = 0 \\ \dfrac{(1-\sigma)\, n!\mu^y\,(1-\mu)^{n-y}}{y!\,(n-y)!} & \text{if } y = 1, 2, 3, \ldots \end{cases} \tag{6.6}$$

for $0 < \mu < 1$, $0 < \sigma < 1$ and n is a known positive integer. The mean and variance of Y are given by

$$\begin{aligned} E(Y) &= (1-\sigma)\, n\mu \\ \text{Var}(Y) &= n\mu\,(1-\sigma)\,[1-\mu+n\mu\sigma] \ . \end{aligned}$$

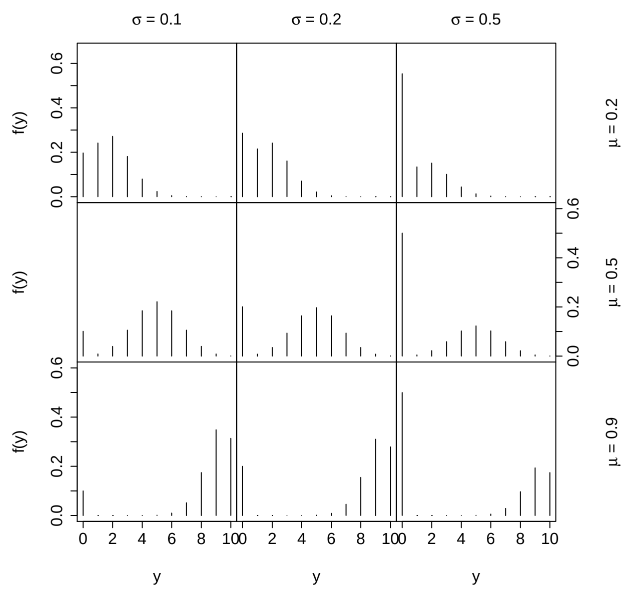

Figure 6.5: The zero-inflated binomial, $\texttt{ZIBI}(n, \mu, \sigma)$, distribution, with $\mu = 0.2, 0.5, 0.9$, $\sigma = 0.1, 0.2, 0.5$, and binomial denominator $n = 10$.

6.3 Three-parameter binomial type distributions

6.3.1 Zero-adjusted beta binomial: ZABB

Let $Y = 0$ with probability ν, and $Y = Y_1$ with probability $(1-\nu)$, where $Y_1 \sim \texttt{BBtr}(n, \mu, \sigma)$ a beta binomial distribution truncated at zero. Then Y has a zero-adjusted (or altered) beta binomial distribution, denoted by $\texttt{ZABB}(n, \mu, \sigma, \nu)$, given by

$$\mathrm{P}(Y = y \mid n, \mu, \sigma, \nu) = \begin{cases} \nu & \text{if } \; y = 0 \\ \dfrac{(1-\nu)\,\mathrm{P}(Y_2 = y \mid n, \mu, \sigma)}{1 - \mathrm{P}(Y_2 = 0 \mid n, \mu, \sigma)} & \text{if } \; y = 1, 2, 3, \ldots \end{cases} \tag{6.7}$$

for $0 < \mu < 1$, $\sigma > 0$ and $0 < \nu < 1$ and n is a known positive integer, where $Y_2 \sim \texttt{BB}(n, \mu, \sigma)$. The mean and variance of Y are given by

$$\mathrm{E}(Y) = \frac{(1-\nu)\, n\mu}{[1 - \mathrm{P}(Y_2 = 0 \mid n, \mu, \sigma)]}$$

$$\mathrm{Var}(Y) = \frac{(1-\nu)\left\{ n\mu\,(1-\mu)\left[1 + \frac{\sigma}{1+\sigma}\,(n-1)\right] - n^2\mu^2 \right\}}{[1 - \mathrm{P}(Y_2 = 0 \mid n, \mu, \sigma)]} - [E(Y)]^2 \ .$$

6.3.2 Zero-inflated beta binomial: `ZIBB`

Let $Y = 0$ with probability ν, and $Y = Y_1$ with probability $(1-\nu)$ where $Y_1 \sim \texttt{BB}(n, \mu, \sigma)$. Then Y has a zero-inflated beta binomial distribution, denoted by $\texttt{ZIBB}(n, \mu, \sigma, \nu)$, given by

$$P(Y = y \mid n, \mu, \sigma, \nu) = \begin{cases} \nu + (1-\nu)\, P(Y_1 = 0 \mid n, \mu, \sigma) & \text{if } y = 0 \\ (1-\nu)\, P(Y_1 = y \mid n, \mu, \sigma) & \text{if } y = 1, 2, 3, \ldots \end{cases} \tag{6.8}$$

for $0 < \mu < 1$, $\sigma > 0$ and $0 < \nu < 1$, and n is a known positive integer, where $Y_1 \sim \texttt{BB}(n, \mu, \sigma)$. The mean and variance of Y are given by

$$\begin{aligned} E(Y) &= (1-\nu)\, n\mu \\ \mathrm{Var}(Y) &= (1-\nu)\, n\mu\, (1-\mu) \left[1 + \tfrac{\sigma}{1+\sigma}\,(n-1)\right] + \nu\,(1-\nu)\, n^2\mu^2 \ . \end{aligned}$$

6.4 Multinomial distributions

6.4.1 Multinomial with five categories: `MN5`

Let Y be a categorical variable with five possible categories (or levels) labelled '0', '1', '2', '3' and '4'.

The probability function of the multinomial distribution with 5 categories, denoted by $\texttt{MN5}(\mu, \sigma, \nu, \tau)$, is given by

$$P(Y = y \mid n, \mu, \sigma, \nu, \tau) = \begin{cases} p_0 & \text{if } y = 0 \\ p_1 & \text{if } y = 1 \\ p_2 & \text{if } y = 2 \\ p_3 & \text{if } y = 3 \\ p_4 & \text{if } y = 4, \end{cases} \tag{6.9}$$

where $p_0 = \mu/s$, $p_1 = \sigma/s$, $p_2 = \nu/s$, $p_3 = \tau/s$, $p_4 = 1/s$, and $s = (1 + \mu + \sigma + \nu + \tau)$ and so $\sum_{j=0}^{4} p_j = 1$.

Hence $\mu = p_0/p_4$, $\sigma = p_1/p_4$, $\nu = p_2/p_4$, $\tau = p_3/p_4$, and $\mu > 0$, $\sigma > 0$, $\nu > 0$ and $\tau > 0$. The default link functions for μ, σ, ν, and τ in $\texttt{MN5}(\mu, \sigma, \nu, \tau)$ in **gamlss.dist** are all logarithmic.

6.4.2 Multinomial with four categories: `MN4`

The probability function of the multinomial distribution with four categories, denoted by $\texttt{MN4}(\mu, \sigma, \nu)$, is given by setting $p_3 = 0$ in (6.9) for $\texttt{MN5}(\mu, \sigma, \nu, \tau)$, hence $\tau = 0$, and then relabelling '4' to '3', and p_4 to p_3.

6.4.3 Multinomial with three categories: `MN3`

The probability function of the multinomial distribution with three categories, denoted by `MN3`(μ, σ), is given by setting $p_2 = p_3 = 0$ in (6.9) for `MN5`(μ, σ, ν, τ), hence $\nu = \tau = 0$, and then relabelling '4' to '2', and p_4 to p_2.

Bibliography

M. Abramowitz and I. A. Stegun. *Handbook of Mathematical Functions: with Formulas, Graphs, and Mathematical Tables.* Dover, New York, 1965.

F. J. Anscombe. Sampling theory of the negative binomial and logarithmic series approximations. *Biometrika*, 37:358–382, 1950.

A. Azzalini. A class of distributions which includes the normal ones. *Scandinavian Journal of Statistics*, 12:171–178, 1985.

A. Azzalini. Further results on a class of distributions which includes the normal ones. *Statistica*, 46:199:208, 1986.

A. Azzalini and A. Capitanio. Distributions generated by perturbation of symmetry with emphasis on a multivariate skew t-distribution. *Journal of the Royal Statistical Society: Series B*, 65:367–389, 2003.

G. E. P. Box and D. R. Cox. An analysis of transformations (with discussion). *Journal of the Royal Statistical Society: Series B*, 26:211–252, 1964.

R.J. Butler, J. B. McDonald, R.D. Nelson, and S.B. White. Robust and partially adaptive estimation of regression models. *The Review of Economics and Statistics*, 72(2):321–327, 1990.

R. J. Carroll and D. Ruppert. *Transformations and Weighting in Regression.* Chapman & Hall, London, 1988.

T. J. Cole and P. J. Green. Smoothing reference centile curves: The LMS method and penalized likelihood. *Statistics in Medicine.*, 11:1305–1319, 1992.

P. C. Consul. *Generalized Poisson Distributions.* Marcel Dekker, New York, 1989.

P. C. Consul and G. C. Jain. A generalization of the Poisson distribution. *Technometrics*, 15(4):791–799, 1973.

G. C. Jr. Davis and M. H. Kutner. The lagged normal family of probability density functions applied to indicator-dilution curves. *Biometrics*, 32:669–675, 1976.

C. Dean, J. F. Lawless, and G. E. Willmot. A mixed Poisson-inverse-Gaussian regression model. *Canadian Journal of Statistics*, 17(2):171–181, 1989.

M. L. Delignette-Muller and C. Dutang. fitdistrplus: An R package for fitting distributions. *Journal of Statistical Software*, 64(4):1–34, 2015. URL `http://www.jstatsoft.org/v64/i04/`.

T. J. DiCiccio and A. C. Monti. Inferential aspects of the skew exponential power distribution. *Journal of the American Statistical Association*, 99:439–450, 2004.

B. Efron. Double exponential families and their use in generalized linear regression. *Journal of the American Statistical Association*, 81:709–721, 1986.

D. A. Evans. Experimental evidence concerning contagious distributions in ecology. *Biometrika*, 40:186–211, 1953.

C. Fernandez and M. F. J. Steel. On Bayesian modelling of fat tails and skewness. *Journal of the American Statistical Association*, 93:359–371, 1998.

C. Fernandez, J. Osiewalski, and M. J. F. Steel. Modeling and inference with v-spherical distributions. *Journal of the American Statistical Association*, 90:1331–1340, 1995.

G. Z. Heller, D.-L. Couturier, and S. R. Heritier. Beyond mean modelling: Bias due to misspecification of dispersion in Poisson-inverse Gaussian regression. *Biometrical Journal*, 61(2):333–342, 2018. doi: 10.1002/bimj.201700218.

B. Hofner, A. Mayr, and M. Schmid. gamboostLSS: An R package for model building and variable selection in the GAMLSS framework. *Journal of Statistical Software*, 74(1):1–31, 2016.

P. J. Huber. Robust estimation of a location parameter. *Annals of Mathematical Statistics*, 35:73–101, 1964.

P. J. Huber. The behavior of maximum likelihood estimates under nonstandard conditions. In *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability*, volume 1, pages 221–233, 1967.

N. L. Johnson. Systems of frequency curves generated by methods of translation. *Biometrika*, 36:149–176, 1949.

N. L. Johnson, S. Kotz, and N. Balakrishnan. *Continuous Univariate Distributions*, volume 1. John Wiley & Sons, New York, 2nd edition, 1994.

N. L. Johnson, S. Kotz, and N. Balakrishnan. *Continuous Univariate Distributions*, volume 2. John Wiley & Sons, New York, 2nd edition, 1995.

N. L. Johnson, A. W. Kemp, and S. Kotz. *Univariate Discrete Distributions.* John Wiley & Sons, New York, 3rd edition, 2005.

M. C. Jones. In discussion of Rigby, R. A. and Stasinopoulos, D. M. (2005) Generalized additive models for location, scale and shape. *Applied Statistics*, 54:507–554, 2005.

M. C. Jones and M. J. Faddy. A skew extension of the t distribution, with applications. *Journal of the Royal Statistical Society: Series B*, 65:159–174, 2003.

M. C. Jones and A. Pewsey. Sinh-arcsinh distributions. *Biometrika*, 96:761–780, 2009.

B. Jørgensen. *Statistical properties of the generalized inverse Gaussian distribution.* Lecture Notes in Statistics No. 9. Springer-Verlag, New York, 1982.

B. Jørgensen. *The Theory of Dispersion Models.* Chapman & Hall, London, 1997.

R. Koenker. Quantile regression: 40 years on. *Annual Review of Economics*, 9: 155–176, 2017.

D. Lambert. Zero-inflated Poisson regression, with an application to defects in manufacturing. *Technometrics*, 34(1):1–14, 1992.

K. L. Lange, R. J. A. Little, and J. M. G. Taylor. Robust statistical modelling using the t distribution. *Journal of the American Statistical Association*, 84: 881–896, 1989.

J. K. Lindsey. *Modelling Frequency and Count Data.* Clarendon Press, Oxford, 1995.

A. Lopatatzidis and P. J. Green. Private Communication, 2000.

G. Lovison and C. Schindler. Separate regression modelling of the Gaussian and exponential components of an EMG response from respiratory physiology. In T. Kneib, F. Sobotka, J. Fahrenholz, and H. Irmer, editors, *Proceedings of the 29th International Workshop on Statistical Modelling*, pages 189–194, 2014.

G. Marra and R. Radice. *GJRM: Generalised Joint Regression Modelling*, 2019. URL `https://CRAN.R-project.org/package=GJRM`. R package version 0.2.0.

P. McCullagh and J. A. Nelder. *Generalized Linear Models*. Chapman & Hall, London, 2nd edition, 1989.

J. B. McDonald. Some generalized functions for the size distributions of income. *Econometrica*, 52:647–663, 1984.

J. B. McDonald. Parametric models for partially adaptive estimation with skewed and leptokurtic residuals. *Economic Letters*, 37:273–278, 1991.

J. B. McDonald. Probability distributions for financial models. In G. S. Maddala and C. R. Rao, editors, *Handbook of Statistics*, volume 14, pages 427–460. Elsevier Science, 1996.

J. B. McDonald and W. K. Newey. Partially adaptive estimation of regression models via the generalized t distribution. *Econometric Theory*, 4:428–457, 1988.

J. B. McDonald and Y. J. Xu. A generalisation of the beta distribution with applications. *Journal of Econometrics*, 66:133–152, 1995.

S. Nadarajah and R. Rocha. Newdistns: An R package for new families of distributions. *Journal of Statistical Software*, 69(10):1–32, 2016.

D. B. Nelson. Conditional heteroskedasticity in asset returns: a new approach. *Econometrica*, 59:347–370, 1991.

R. Ospina and S. L. P. Ferrari. Inflated beta distributions. *Statistical Papers*, 51:111–126, 2010.

B. A. Para and T. R. Jan. On discrete three parameter Burr type XII and discrete Lomax distributions and their applications to model count data from medical science. *Biometrics and Biostatistics International Journal*, 4(2):1–15, 2016.

K. Poortema. On modelling overdispersion of counts. *Statistica Neerlandica*, 53(1):5–20, 1999.

R. A. Rigby and D. M. Stasinopoulos. Robust fitting of an additive model for variance heterogeneity. In R. Dutter and W. Grossmann, editors, *COMPSTAT: Proceedings in Computational Statistics*, pages 263–268. Physica, Heidelberg, 1994.

R. A. Rigby and D. M. Stasinopoulos. Smooth centile curves for skew and kurtotic data modelled using the Box Cox power exponential distribution. *Statistics in Medicine*, 23:3053–3076, 2004.

R. A. Rigby and D. M. Stasinopoulos. Generalized additive models for location, scale and shape (with discussion). *Applied Statistics*, 54:507–554, 2005.

R. A. Rigby and D. M. Stasinopoulos. Using the Box-Cox t distribution in GAMLSS to model skewness and kurtosis. *Statistical Modelling*, 6(3):209, 2006.

R. A. Rigby and D. M. Stasinopoulos. Automatic smoothing parameter selection in GAMLSS with an application to centile estimation. *Statistical Methods in Medical Research*, 23(4):318–332, 2013.

R. A. Rigby, D. M. Stasinopoulos, and C. Akantziliotou. A framework for modelling overdispersed count data, including the Poisson-shifted generalized inverse Gaussian distribution. *Computational Statistics and Data Analysis*, 53(2):381–393, 2008. ISSN 0167-9473.

E. W. Stacy and G. A. Mihram. Parameter estimation for a generalized gamma distribution. *Technometrics*, 7(3):349–358, 1965.

S. Stadlmann. *distreg.vis: Framework for the Visualization of Distributional Regression Models*, 2019. URL `https://CRAN.R-project.org/package=distreg.vis`. R package version 1.1.0.

D. M. Stasinopoulos, R. A. Rigby, G. Z. Heller, V. Voudouris, and F. De Bastiani. *Flexible Regression and Smoothing: Using GAMLSS in R*. Chapman & Hall/CRC, Boca Raton, 2017.

G. Z. Stein, W. Zucchini, and J. M. Juritz. Parameter estimation of the Sichel distribution and its multivariate extension. *Journal of the American Statistical Association*, 82:938–944, 1987.

M. T. Subbotin. On the law of frequency of errors. *Mathematicheskii Sbornik*, 31:296–301, 1923.

N. Umlauf, N. Klein, and A. Zeileis. BAMLSS: Bayesian additive models for location, scale and shape (and beyond). *Journal of Computational and Graphical Statistics*, 27:612–627, 2017. URL `http://dx.doi.org/10.1080/10618600.2017.1407325`.

W. R. van Zwet. Convex transformations: a new approach to skewness and kurtosis. *Statistica Neerlandica*, 18:433–441, 1964a.

W. R. van Zwet. *Convex Transformations of Random Variables. Mathematica Centre Tracts 7*. Mathematisch Centrum, Amsterdam, 1964b.

G. Wimmer and G. Altmann. *Thesaurus of Univariate Discrete Probability Distributions*. Stamm Verlag, Essen, Germany, 1999.

S. N. Wood, N. Pya, and B. Säfken. Smoothing parameter and model selection for general smooth models. *Journal of the American Statistical Association*, 111(516):1548–1563, 2017.

D. Würtz, Y. Chalabi, and L. Luksan. Parameter estimation of ARMA models with GARCH/APARCH errors. An R and SPlus software implementation. *Journal of Statistical Software*, 55:28–33, 2006.

T. W. Yee. *VGAM: Vector Generalized Linear and Additive Models*, 2019. URL `https://CRAN.R-project.org/package=VGAM`. R package version 1.1-1.

A. Zeileis, C. Kleiber, and S. Jackman. Regression models for count data in R. *Journal of Statistical Software*, 27(8):1–25, 2008. URL `http://www.jstatsoft.org/v27/i08/`.

A. Ziegler. *Generalized Estimating Equations*. Springer Science and Business Media, New York, 2011.

Index