Location Scale Distributions

Alexios Galanos

2025-04-17

Introduction

The **tsdistributions** package implements a number of location-scale based statistical distributions parameterized in terms of their mean and standard deviation, in addition to the skew and shape parameters. The distributions functions follow R's distributions naming convention in terms of pdf (d), cdf (p), quantile (q) and random number generation (r).

Currently implemented distributions are the Normal ('norm'), Student ('std') and Generalized Error ('ged') distributions; the skewed variants of these based on the transformations in (Fernandez and Steel 1998), which are the Skew Normal ('snorm'), Skew Student ('sstd') and Skew Generalized Error ('sged') distributions; The re-parameterized version of (Johnson 1949) SU distribution ('jsu'); the Generalized Hyperbolic ('gh') distribution of (Barndorff-Nielsen 1977), the Normal Inverse Gaussian ('nig'), and the Generalized Hyperbolic Skew Student ('ghst') distribution of (Aas and Haff 2006).

Location Scale Distributions

For any random variable Z whose probability distribution function belongs to a location scale family, the distribution function of $X \overset{d}{=} \mu + \sigma Z$ also belongs to the family.

This implies that if Z has a continuous distribution with probability density function (pdf) g, then X also has a continuous distribution, with probability density function f given by:

$$f(x) = \frac{1}{\sigma} g\left(\frac{x - \mu}{\sigma}\right), \quad x \in \mathbb{R}$$

The parameter (moment) σ affects the pdf through a scaling operation which is why the standardized distribution g needs to be adjusted for this, whereas the centering by the mean (μ) is simply a translation operation which does not have any affect on the pdf.

The distributions implemented in this package all belong to this type of family and are parameterized in terms of the mean (μ) and standard deviation (σ), making for easier inference in addition to allowing one to concentrate out the first two moments if desired during estimation. Specific applications of such distributions and parameterizations can be found in time series analysis in particular, where the conditional mean (μ_t) and/or standard deviation (σ_t) have their own motion dynamics, such as in ARMA-GARCH models.

Implemented Distributions

Normal Distribution (μ, σ)

The Normal distribution is a spherical distribution described completely by it first two moments, the mean and variance. Formally, the random variable x is said to be normally distributed with mean μ and variance σ^2 (both of which may be time varying), with density given by,

$$f\left(x\right) = \frac{e^{\frac{-0.5\left(x-\mu\right)^2}{\sigma^2}}}{\sigma\sqrt{2\pi}}.$$

where $\mu \in \mathbb{R}$ and $\sigma \in \mathbb{R}_{>0}$.

Student's t Distribution (μ, σ, ν)

The Student's t distribution is described completely by a shape parameter ν, but for standardization we proceed by using its 3 parameter representation as follows:

$$f\left(x\right) = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\beta\nu\pi}\Gamma\left(\frac{\nu}{2}\right)}\left(1 + \frac{\left(x-\alpha\right)^2}{\beta\nu}\right)^{-\left(\frac{\nu+1}{2}\right)}$$

where α, β, and ν are the location, scale and shape parameters respectively, and Γ is the Gamma function. Similar to the GED distribution described later, this is a unimodal and symmetric distribution where the location parameter α is the mean (and mode) of the distribution while the variance is:

$$Var\left(x\right) = \frac{\beta\nu}{\left(\nu-2\right)}.$$

For the purposes of standardization we require that:

$$Var(x) = \frac{\beta\nu}{\left(\nu-2\right)} = 1$$

$$\therefore \beta = \frac{\nu-2}{\nu}$$

Substituting $\frac{\left(\nu-2\right)}{\nu}$ we obtain the standardized Student's distribution:

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

The Student distribution has zero skewness and excess kurtosis equal to $6/\left(\nu-4\right)$ for $\nu > 4$.

Generalized Error Distribution (μ, σ, ν)

The Generalized Error Distribution is a 3 parameter distribution belonging to the exponential family with conditional density given by,

$$f\left(x\right) = \frac{\nu e^{-0.5\left|\frac{x-\alpha}{\beta}\right|^\nu}}{2^{1+\nu^{-1}}\beta\Gamma\left(\nu^{-1}\right)}$$

with α, β and ν representing the location, scale and shape parameters. Since the distribution is symmetric and unimodal the location parameter is also the mode, median and mean of the distribution (i.e. μ). By symmetry, all odd moments beyond the mean are zero. The variance and kurtosis are given by,

$$Var\left(x\right) = \beta^2 2^{2/\nu}\frac{\Gamma\left(3\nu^{-1}\right)}{\Gamma\left(\nu^{-1}\right)}$$

$$Ku\left(x\right) = \frac{\Gamma\left(5\nu^{-1}\right)\Gamma\left(\nu^{-1}\right)}{\Gamma\left(3\nu^{-1}\right)\Gamma\left(3\nu^{-1}\right)}$$

As ν decreases the density gets flatter and flatter while in the limit as $\nu \to \infty$, the distribution tends towards the uniform. Special cases are the Normal when $\nu = 2$, the Laplace when $\nu = 1$. Standardization is simple and involves re-scaling the density to have unit standard deviation:

$$Var\left(x\right) = \beta^2 2^{2/\nu} \frac{\Gamma\left(3\nu^{-1}\right)}{\Gamma\left(\nu^{-1}\right)} = 1$$

$$\therefore \beta = \left[2^{-2/\nu} \frac{\Gamma\left(\nu^{-1}\right)}{\Gamma\left(3\nu^{-1}\right)}\right]^{0.5}$$

Finally, substituting into the scaled density of z:

$$f\left(\frac{x-\mu}{\sigma}\right) = \frac{1}{\sigma} f\left(z\right) = \frac{1}{\sigma} \frac{\nu e^{-0.5\left|\sqrt{2^{-2/\nu} \frac{\Gamma\left(\nu^{-1}\right)}{\Gamma\left(3\nu^{-1}\right)}} z\right|^{\nu}}}{\sqrt{2^{-2/\nu} \frac{\Gamma\left(\nu^{-1}\right)}{\Gamma\left(3\nu^{-1}\right)}} 2^{1+\nu^{-1}} \Gamma\left(\nu^{-1}\right)}$$

Skewed Distributions by Inverse Scale Factors

(Fernandez and Steel 1998) proposed introducing skewness into unimodal and symmetric distributions by introducing inverse scale factors in the positive and negative real half lines.

Given a skew parameter, ξ (when $\xi = 1$, the distribution is symmetric), the density of a random variable z can be represented as:

$$f\left(z|\xi\right) = \frac{2}{\xi + \xi^{-1}} \left[f\left(\xi z\right) H\left(-z\right) + f\left(\xi^{-1} z\right) H\left(z\right)\right]$$

where $\xi \in \mathbb{R}^+$ and $H(.)$ is the Heaviside function. The absolute moments, required for deriving the central moments, are generated from the following function:

$$M_r = 2 \int_0^\infty z^r f\left(z\right) dz.$$

The mean and variance are then defined as:

$$E\left(z\right) = M_1 \left(\xi - \xi^{-1}\right)$$
$$Var\left(z\right) = \left(M_2 - M_1^2\right)\left(\xi^2 + \xi^{-2}\right) + 2M_1^2 - M_2$$

The Normal, Student and GED distributions have skew variants which have been re-parameterized in terms of mean and standard deviation by making use of the moment conditions given above. The distributions available are:

- Skew Normal (μ, σ, ξ)
- Skew Student (μ, σ, ξ, ν)
- Skew GED (μ, σ, ξ, ν)

Johnson's SU Distribution (μ, σ, ξ, ν)

In the original parameterization of Johnson's SU distribution, the pdf is given by:

$$f\left(x, \xi, \lambda, \nu^*, \tau^*\right) = \frac{\tau^*}{\sigma \left(s^2 + 1\right)^{0.5} \sqrt{2\pi}} \exp\left[-\frac{1}{2} z^2\right]$$

for $-\infty < x < \infty$, $-\infty < \xi < \infty$, $\lambda > 0$, $-\infty < \nu^* < \infty$ and $\tau^* > 0$, where

$$z = \nu^* + \tau^* \sinh^{-1}(s) = \nu^* + \tau^* \log\left[s + \left(s^2 + 1\right)^{0.5}\right]$$

with $s = (x - \xi)/\lambda$.

Re-parameterizing the pdf in terms of the mean and standard deviation we set $\mu = \xi - \lambda\omega^{0.5}\sinh(\nu^*/\tau^*)$, $\sigma = \lambda/c$, $\nu = -\nu^*$ and $\tau = \tau^*$. Therefore:

$$f(x, \mu, \sigma, \nu, \tau) = \frac{\tau}{c\sigma\left(s^2 + 1\right)^{0.5}\sqrt{(2\pi)}} \exp\left[-\frac{1}{2}z^2\right]$$

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

$$z = -\nu + \tau\sinh^{-1}(s) = -\nu + \tau\log\left(s + \left(s^2 + 1\right)^{0.5}\right)$$

$$s = \frac{x - \mu + c\sigma\omega^{0.5}\sinh(\nu/\tau)}{c\sigma}$$

$$c = \left\{0.5\left(\omega - 1\right)\left[\omega\cosh\left(2\nu/\tau\right) + 1\right]\right\}^{-0.5}$$

$$\omega = \exp\left(1/\tau^2\right)$$

The re-parameterization is taken from (Rigby et al. 2019) Section 18.4.3 which also contains additional information on the properties of the ν and τ parameters. In our implementation, and to avoid confusion, ν in the originally presented parameterization is the skew parameter (ξ in our parameterization) and τ the shape parameter (ν in our parameterization).

The Generalized Hyperbolic Distribution (μ, σ, ξ, ν, λ)

In distributions where the expected moments are functions of all the parameters, it is not immediately obvious how to perform such a transformation. In the case of the Generalized Hyperbolic (GH) distribution, because of the existence of location and scale invariant parameterizations and the possibility of expressing the variance in terms of one of those, namely the (ζ, ρ), the task of standardizing and estimating the density can be broken down to one of estimating those 2 parameters, representing a combination of shape and skewness, followed by a series of transformation steps to demean, scale and then translate the parameters into the $(\alpha, \beta, \delta, \mu)$ parameterization for which standard formula exist for the likelihood function. The (ξ, χ) parameterization, which is a simple transformation of the (ζ, ρ), could also be used in the first step and then transformed into the latter before proceeding further. The only difference is the kind of 'immediate' inference one can make from the different parameterizations, each providing a different direct insight into the kind of dynamics produced and their place in the overall GH family particularly with regards to the limit cases.

The package performs estimation using the (ζ, ρ) parameterization, after which a series of steps transform those parameters into the $(\alpha, \beta, \delta, \mu)$ while at the same time including the necessary recursive substitution of parameters in order to standardize the resulting distribution.

Consider the standardized Generalized Hyperbolic Distribution. Let ε_t be a r.v. with mean (0) and variance (σ^2) distributed as $\mathrm{GH}(\zeta, \rho)$, and let z be a scaled version of the r.v. ε with variance (1) and also distributed as $\mathrm{GH}(\zeta, \rho)$ (the parameters ζ and ρ do not change as a result of being location and scale invariant). The density $f(.)$ of z can be expressed as

$$f(\frac{\varepsilon_t}{\sigma}; \zeta, \rho) = \frac{1}{\sigma}f_t(z; \zeta, \rho) = \frac{1}{\sigma}f_t(z; \tilde{\alpha}, \tilde{\beta}, \tilde{\delta}, \tilde{\mu}),$$

where we make use of the $(\alpha, \beta, \delta, \mu)$ parameterization since we can only naturally express the density in that parameterization. The steps to transforming from the (ζ, ρ) to the $(\alpha, \beta, \delta, \mu)$ parameterization, while at the same time standardizing for zero mean and unit variance are given henceforth. Let

4

$$\zeta = \delta\sqrt{\alpha^2 - \beta^2}$$

$$\rho = \frac{\beta}{\alpha},$$

which after some substitution may be also written in terms of α and β as,

$$\alpha = \frac{\zeta}{\delta\sqrt{(1-\rho^2)}},$$

$$\beta = \alpha\rho.$$

For standardization we require that,

$$E(X) = \mu + \frac{\beta\delta}{\sqrt{\alpha^2 - \beta^2}}\frac{K_{\lambda+1}(\zeta)}{K_\lambda(\zeta)} = \mu + \frac{\beta\delta^2}{\zeta}\frac{K_{\lambda+1}(\zeta)}{K_\lambda(\zeta)} = 0$$

$$\therefore \mu = -\frac{\beta\delta^2}{\zeta}\frac{K_{\lambda+1}(\zeta)}{K_\lambda(\zeta)}$$

$$Var(X) = \delta^2\left(\frac{K_{\lambda+1}(\zeta)}{\zeta K_\lambda(\zeta)} + \frac{\beta^2}{\alpha^2 - \beta^2}\left(\frac{K_{\lambda+2}(\zeta)}{K_\lambda(\zeta)} - \left(\frac{K_{\lambda+1}(\zeta)}{K_\lambda(\zeta)}\right)^2\right)\right) = 1$$

$$\therefore \delta = \left(\frac{K_{\lambda+1}(\zeta)}{\zeta K_\lambda(\zeta)} + \frac{\beta^2}{\alpha^2 - \beta^2}\left(\frac{K_{\lambda+2}(\zeta)}{K_\lambda(\zeta)} - \left(\frac{K_{\lambda+1}(\zeta)}{K_\lambda(\zeta)}\right)^2\right)\right)^{-0.5}$$

Since we can express, $\beta^2/\left(\alpha^2 - \beta^2\right)$ as,

$$\frac{\beta^2}{\alpha^2 - \beta^2} = \frac{\alpha^2\rho^2}{a^2 - \alpha^2\rho^2} = \frac{\alpha^2\rho^2}{a^2(1-\rho^2)} = \frac{\rho^2}{(1-\rho^2)},$$

then we can re-write the formula for δ in terms of the estimated parameters $\hat{\zeta}$ and $\hat{\rho}$ as,

$$\delta = \left(\frac{K_{\lambda+1}\left(\hat{\zeta}\right)}{\hat{\zeta}K_\lambda\left(\hat{\zeta}\right)} + \frac{\hat{\rho}^2}{(1-\hat{\rho}^2)}\left(\frac{K_{\lambda+2}\left(\hat{\zeta}\right)}{K_\lambda\left(\hat{\zeta}\right)} - \left(\frac{K_{\lambda+1}\left(\hat{\zeta}\right)}{K_\lambda\left(\hat{\zeta}\right)}\right)^2\right)\right)^{-0.5}$$

Transforming into the $(\tilde{\alpha}, \tilde{\beta}, \tilde{\delta}, \tilde{\mu})$ parameterization proceeds by first substituting the above value of δ into the equation for α and simplifying:

$$\tilde{\alpha} = \frac{\hat{\zeta}\left(\frac{\mathrm{K}_{\lambda+1}(\hat{\zeta})}{\hat{\zeta}\mathrm{K}_\lambda(\hat{\zeta})} + \frac{\hat{\rho}^2\left(\frac{\mathrm{K}_{\lambda+2}(\hat{\zeta})}{\mathrm{K}_\lambda(\hat{\zeta})} - \frac{(\mathrm{K}_{\lambda+1}(\hat{\zeta}))^2}{(\mathrm{K}_\lambda(\hat{\zeta}))^2}\right)}{(1-\hat{\rho}^2)}\right)^{0.5}}{\sqrt{(1-\hat{\rho}^2)}},$$

$$= \frac{\left(\frac{\hat{\zeta}\mathrm{K}_{\lambda+1}(\hat{\zeta})}{\mathrm{K}_\lambda(\hat{\zeta})} + \frac{\hat{\zeta}^2\hat{\rho}^2\left(\frac{\mathrm{K}_{\lambda+2}(\hat{\zeta})}{\mathrm{K}_\lambda(\hat{\zeta})} - \frac{(\mathrm{K}_{\lambda+1}(\hat{\zeta}))^2}{(\mathrm{K}_\lambda(\hat{\zeta}))^2}\right)}{(1-\hat{\rho}^2)}\right)^{0.5}}{\sqrt{(1-\hat{\rho}^2)}},$$

$$= \left(\frac{\frac{\hat{\zeta}\mathrm{K}_{\lambda+1}(\hat{\zeta})}{\mathrm{K}_\lambda(\hat{\zeta})}}{(1-\hat{\rho}^2)} + \frac{\hat{\zeta}^2\hat{\rho}^2\left(\frac{\mathrm{K}_{\lambda+2}(\hat{\zeta})}{\mathrm{K}_{\lambda+1}(\hat{\zeta})}\frac{\mathrm{K}_{\lambda+1}(\hat{\zeta})}{\mathrm{K}_\lambda(\hat{\zeta})} - \frac{(\mathrm{K}_{\lambda+1}(\hat{\zeta}))^2}{(\mathrm{K}_\lambda(\hat{\zeta}))^2}\right)}{(1-\hat{\rho}^2)^2}\right)^{0.5},$$

$$= \left(\frac{\frac{\hat{\zeta}\mathrm{K}_{\lambda+1}(\hat{\zeta})}{\mathrm{K}_\lambda(\hat{\zeta})}}{(1-\hat{\rho}^2)}\left(1 + \frac{\hat{\zeta}\hat{\rho}^2\left(\frac{\mathrm{K}_{\lambda+2}(\hat{\zeta})}{\mathrm{K}_{\lambda+1}(\hat{\zeta})} - \frac{\mathrm{K}_{\lambda+1}(\hat{\zeta})}{\mathrm{K}_\lambda(\hat{\zeta})}\right)}{(1-\hat{\rho}^2)}\right)\right)^{0.5}.$$

Finally, the rest of the parameters are derived recursively from $\tilde{\alpha}$ and the previous results,

$$\tilde{\beta} = \tilde{\alpha}\hat{\rho},$$

$$\tilde{\delta} = \frac{\hat{\zeta}}{\tilde{\alpha}\sqrt{1-\hat{\rho}^2}},$$

$$\tilde{\mu} = \frac{-\tilde{\beta}\tilde{\delta}^2 K_{\lambda+1}\left(\hat{\zeta}\right)}{\hat{\zeta}K_\lambda\left(\hat{\zeta}\right)}.$$

For the use of the (ξ, χ) parameterization in estimation, the additional preliminary steps of converting to the (ζ, ρ) are,

$$\zeta = \frac{1}{\hat{\xi}^2} - 1,$$

$$\rho = \frac{\hat{\chi}}{\hat{\xi}}.$$

In our notation, ρ is the skew parameter (ξ) and ζ the shape parameter (ν).

Special cases of interest are the Hyperbolic distribution $(\lambda = 1)$ and the Normal Inverse Gaussian $(\lambda = -0.5)$. The latter is included in the package as a separate case to make use of some of it's unique properties and enable faster computation, whilst other cases can be achieved by fixing the λ parameter post specification.

The Generalized Hyperbolic Skew Student Distribution (μ, σ, ξ, ν)

The Generalized Hyperbolic (GH) Skew-Student distribution was popularized by (Aas and Haff 2006) because of its uniqueness in the GH family for having one tail with polynomial and one with exponential behavior. This distribution is a limiting case of the GH when $\alpha \to |\beta|$ and $\lambda = -\nu/2$, where ν is the shape parameter of the Student distribution.

The domain of variation of the parameters is $\beta \in \mathbb{R}$ and $\nu > 0$, but for the variance to be finite $\nu > 4$, while for the existence of skewness and kurtosis, $\nu > 6$ and $\nu > 8$ respectively. The density of the random variable x is then given by:

$$f\left(x\right) = \frac{2^{(1-\nu)/2}\delta^\nu|\beta|^{(\nu+1)/2}K_{(\nu+1)/2}\left(\sqrt{\beta^2\left(\delta^2 + (x-\mu)^2\right)}\right)\exp\left(\beta\left(x-\mu\right)\right)}{\Gamma\left(\nu/2\right)\sqrt{\pi}\left(\sqrt{\delta^2 + (x-\mu)^2}\right)^{(\nu+1)/2}}$$

To standardize the distribution to have zero mean and unit variance, I make use of the first two moment conditions for the distribution which are:

$$E\left(x\right) = \mu + \frac{\beta\delta^2}{\nu - 2}$$

$$Var\left(x\right) = \frac{2\beta^2\delta^4}{\left(\nu - 2\right)^2\left(\nu - 4\right)} + \frac{\delta^2}{\nu - 2}$$

We require that $Var(x) = 1$, thus:

$$\delta = \left(\frac{2\bar{\beta}^2}{\left(\nu - 2\right)^2\left(\nu - 4\right)} + \frac{1}{\nu - 2}\right)^{-1/2}$$

where I have made use of the 4th parameterization of the GH distribution given in (Prause 1999) where $\hat{\beta} = \beta\delta$. The location parameter is then rescaled by substituting into the first moment formula δ so that it has zero mean:

$$\bar{\mu} = -\frac{\beta\delta^2}{\nu - 2}$$

Therefore, we model the GH Skew-Student using the location-scale invariant parameterization $(\bar{\beta}, \nu)$ and then translate the parameters into the usual GH distribution's $(\alpha, \beta, \delta, \mu)$, setting $\alpha = abs(\beta) + 1e - 12$. In our notation, $\bar{\beta}$ is the skew parameter (ξ) and ν the shape parameter.

References

Aas, K., and I.H. Haff. 2006. "The Generalized Hyperbolic Skew Student's t-Distribution." *Journal of Financial Econometrics* 4 (2): 275–309.

Barndorff-Nielsen, Ole. 1977. "Exponentially Decreasing Distributions for the Logarithm of Particle Size." *Proceedings of the Royal Society of London. A. Mathematical and Physical Sciences* 353 (1674): 401–19.

Fernandez, C., and M.F. Steel. 1998. "On Bayesian Modeling of Fat Tails and Skewness." *Journal of the American Statistical Association* 93 (441): 359–71.

Johnson, Norman L. 1949. "Systems of Frequency Curves Generated by Methods of Translation." *Biometrika* 36 (1/2): 149–76.

Prause, K. 1999. "The Generalized Hyperbolic Model: Estimation, Financial Derivatives, and Risk Measures." PhD thesis, University of Freiburg.

Rigby, Robert A, Mikis D Stasinopoulos, Gillian Z Heller, and Fernanda De Bastiani. 2019. *Distributions for Modeling Location, Scale, and Shape: Using GAMLSS in r.* CRC press.