Some distributions that are used in models for daily stock returns

Lennart Hoogerheide

Department of Econometrics and Data Science
School of Business and Economics
Vrije Universiteit Amsterdam

2023

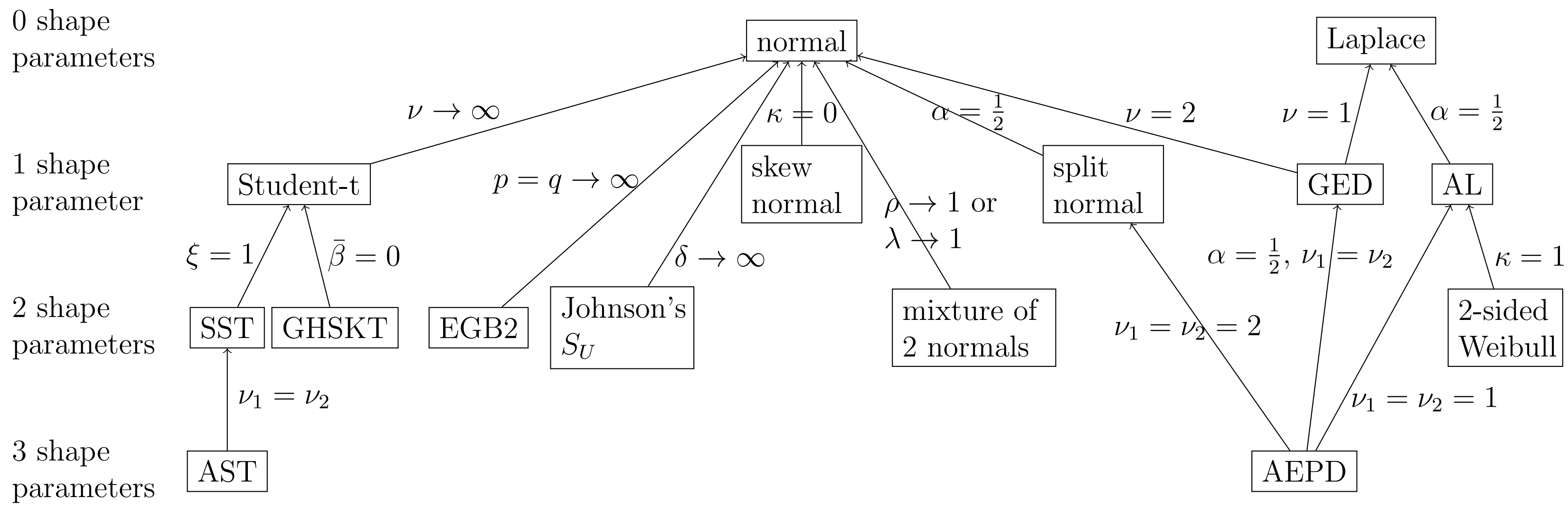
In this note we consider some of the distributions that are used for daily stock returns r_t. We assume that the conditional distribution of r_t (conditional upon the information available at time $t-1$) has mean μ_t (which may simply be 0) and variance h_t. In the notation we drop the dependence on $r_{t-1}, r_{t-2}, \ldots$ and the parameters, so that the probability density function is simply denoted by $p(r_t)$.

Note: as mentioned during the lecture, it is advised to make use of $f_t = \log(h_t)$ (the natural logarithm of the variance) in the framework of the Generalized Autoregressive Score (GAS) model, so that we need the derivative

$$\nabla_t = \frac{\partial \log p(r_t)}{\partial f_t} = \frac{\partial \log p(r_t)}{\partial \log(h_t)} = \frac{\partial \log p(r_t)}{\partial h_t} \frac{\partial h_t}{\partial \log(h_t)} = \frac{\partial \log p(r_t)}{\partial h_t} h_t.$$

In that case one can make use of scaling factor $S_t = 1$, so that $s_t = S_t \cdot \nabla_t = \nabla_t$.

The distributions have 0, 1, 2 or 3 shape parameters. The distributions with fewer parameters are often restricted versions of distributions with more parameters, as depicted in the overview on the next page.



GED = Generalized Error Distribution

AL = asymmetric Laplace

SST = skewed Student-t

GHSKT = Generalized Hyperbolic Skewed Student-t

EGB2 = Exponential Generalized Beta distribution of the second kind

AST = Asymmetric Student-t

AEPD = Asymmetric Exponential Power Distribution

1 Distributions with no shape parameters

1.1 Normal/Gaussian distribution

The normal distribution (or Gaussian distribution; Gauss (1809)) has probability density function

$$p(r_t) = (2\pi h_t)^{-1/2} \exp\left(-\frac{(r_t - \mu_t)^2}{2h_t}\right),$$

with (natural) logarithm

$$\log p(r_t) = -\frac{1}{2}\log(2\pi h_t) - \frac{(r_t - \mu_t)^2}{2h_t}.$$

This distribution has skewness 0 and kurtosis 3. The thin tails of the normal distribution (corresponding to the relatively small kurtosis) imply that this distribution is typically not a wise choice for daily returns, especially if one is interested in the far/deep tails (for example, for a 99% or 99.5% Value at Risk).

1.2 Laplace distribution

The Laplace distribution of Laplace (1774) has probability density function

$$p(r_t) = (2h_t)^{-1/2} \exp\left(-\frac{|r_t - \mu_t|}{\sqrt{h_t/2}}\right),$$

with (natural) logarithm

$$\log p(r_t) = -\frac{1}{2}\log(2h_t) - \frac{|r_t - \mu_t|}{\sqrt{h_t/2}}.$$

This distribution has skewness 0 and kurtosis 6. If $\mu_t = 0$, then the positive half is exactly an exponential density scaled by $\frac{1}{2}$; therefore it is also called the double exponential distribution.

The fat tails of the distribution imply that it can be a wise choice for daily returns, as long as the empirical distribution of the daily returns is (almost) symmetric.

2 Distributions with one shape parameter

2.1 Student-t distribution

The Student-t distribution (Gosset, 1908) with degrees of freedom parameter $\nu > 2$ has probability density function

$$p(r_t) = \frac{\Gamma(\frac{\nu+1}{2})}{\Gamma(\frac{\nu}{2})}((\nu-2)\pi h_t)^{-1/2}\left(1+\frac{(r_t-\mu_t)^2}{(\nu-2)h_t}\right)^{-\frac{\nu+1}{2}}$$

with natural logarithm

$$\log p(r_t) = \log\Gamma\left(\frac{\nu+1}{2}\right) - \log\Gamma\left(\frac{\nu}{2}\right) - \frac{1}{2}\log((\nu-2)\pi h_t) - \frac{\nu+1}{2}\log\left(1+\frac{(r_t-\mu_t)^2}{(\nu-2)h_t}\right).$$

This distribution has skewness 0 for $\nu > 3$, the skewness is undefined for $2 < \nu \leq 3$. The kurtosis is $3+\frac{6}{\nu-4}$ for $\nu > 4$, the kurtosis is ∞ for $2 < \nu \leq 4$. The fat tails of the Student-t distribution imply that this distribution can be a wise choice for daily returns, as long as the empirical distribution of the daily returns is (almost) symmetric. For $\nu \to \infty$ the distribution tends to the normal distribution.

If during the numerical optimization of the loglikelihood the parameter ν tends to its lower bound 2, then this may indicate that a distribution with fatter tails is needed. In this case one can make use of the following Student-t distribution with degrees of freedom parameter $\nu > 0$ (allowing $0 < \nu \leq 2$) and probability density function

$$p(r_t) = \frac{\Gamma(\frac{\nu+1}{2})}{\Gamma(\frac{\nu}{2})}(\nu\pi s_t)^{-1/2}\left(1+\frac{(r_t-\mu_t)^2}{\nu s_t}\right)^{-\frac{\nu+1}{2}}$$

with natural logarithm

$$\log p(r_t) = \log\Gamma\left(\frac{\nu+1}{2}\right) - \log\Gamma\left(\frac{\nu}{2}\right) - \frac{1}{2}\log(\nu\pi s_t) - \frac{\nu+1}{2}\log\left(1+\frac{(r_t-\mu_t)^2}{\nu s_t}\right).$$

If $1 < \nu \leq 2$, then the mean is μ_t and the variance infinite, although s_t still reflects the 'scale' of the distribution (for example, in terms of the width of prediction intervals). If $0 < \nu \leq 1$, the mean and variance are undefined, although μ_t and s_t still reflect the 'location' and 'scale' of the distribution. If $\nu > 2$, then this is simply the abovementioned Student-t distribution with mean μ_t and variance $h_t = \frac{\nu}{\nu-2}s_t$.

2.2 Generalized Error Distribution (GED)

The Generalized Error Distribution (GED[1]) with shape parameter $\nu > 0$ has probability density function

$$p(r_t) = \left(2^{(1+(1/\nu))}\Gamma(1/\nu)\lambda\right)^{-1} \; h_t^{-1/2} \; \nu \exp\left(-\frac{1}{2}\left|\frac{r_t - \mu_t}{\lambda\, h_t^{1/2}}\right|^{\nu}\right)$$

with

$$\lambda = \left(\frac{\Gamma(1/\nu)}{2^{2/\nu}\Gamma(3/\nu)}\right)^{1/2}.$$

The natural logarithm is

$$\log p(r_t) = -\log\left(2^{(1+(1/\nu))}\Gamma(1/\nu)\lambda\right) - \frac{1}{2}\log(h_t) + \log(\nu) - \frac{1}{2}\left|\frac{r_t - \mu_t}{\lambda\, h_t^{1/2}}\right|^{\nu}.$$

Note: λ is not a separate parameter, it is simply a function of ν. Note: in case of the GED this is **not** called a degrees of freedom parameter. This distribution has skewness 0. The kurtosis is $\frac{\Gamma(1/\nu)\Gamma(5/\nu)}{(\Gamma(3/\nu))^2}$. The GED has fatter tails (and larger kurtosis) than the normal distribution if $0 < \nu < 2$. The GED is the normal distribution if $\nu = 2$. The GED has thinner tails (and smaller kurtosis) than the normal distribution if $\nu > 2$. The GED is the Laplace (double exponential) distribution if $\nu = 1$.

The fat tails of the GED (if $0 < \nu < 2$) imply that this distribution can be a wise choice for daily returns, as long as the empirical distribution of the daily returns is (almost) symmetric.

[1]The GED class was first proposed by Subbotin (1923); Box and Tiao (1973) called such a distribution the Exponential Power Distribution (EPD); it is also called the generalized normal distribution, Generalized Power Distribution, Generalized Laplace Distribution or Box-Tiao distribution.

2.3 Skew normal distribution

The skew normal distribution of O'Hagan and Leonard (1976) (which is a member of a class of skewed distributions; Azzalini (1985)) with asymmetry parameter κ has probability density function:

$$p(r_t) = \frac{\sqrt{2}}{\omega\sqrt{\pi h_t}} \times \exp\left(-\frac{1}{2}\left(\frac{\frac{r_t-\mu_t}{h_t^{1/2}} - \xi}{\omega}\right)^2\right) \times \Phi\left(\kappa\left(\frac{\frac{r_t-\mu_t}{h_t^{1/2}} - \xi}{\omega}\right)\right)$$

with

$$\begin{aligned}
\delta &= \frac{\kappa}{\sqrt{1+\kappa^2}} \\
\omega &= \left(1 - \frac{2\delta^2}{\pi}\right)^{-\frac{1}{2}} \\
\xi &= -\omega\delta\sqrt{\frac{2}{\pi}}.
\end{aligned}$$

$\Phi()$ is the CDF of the standard normal distribution. The natural logarithm is

$$\log p(r_t) = \frac{1}{2}\log(2) - \log(\omega) - \frac{1}{2}\log(\pi h_t) - \frac{1}{2}\left(\frac{\frac{r_t-\mu_t}{h_t^{1/2}} - \xi}{\omega}\right)^2 + \log\Phi\left(\kappa\left(\frac{\frac{r_t-\mu_t}{h_t^{1/2}} - \xi}{\omega}\right)\right).$$

Note: δ, ω and ξ are not separate parameters, these are simply functions of κ.

For $\kappa = 0$, this skew normal distribution (with $\delta = \xi = 0$, $\omega = 1$ and $\Phi(0) = \frac{1}{2}$) reduces to the symmetric normal distribution. For $\kappa > 0$ the distribution has a positive skewness. For $\kappa < 0$ the distribution has a negative skewness. For $\kappa \neq 0$ the distribution has kurtosis larger than 3.

An advantage of this distribution is the possible asymmetry. However, a disadvantage is that it does not allow for very fat tails.

2.4 Split normal distribution

The split normal distribution or two-piece normal distribution of Fechner (1897), Gibbons and Mylroie (1973) and John (1982) with asymmetry parameter α $(0 < \alpha < 1)$ has probability density function:

$$p(r_t) = \frac{s}{\sqrt{2\pi h_t}} \exp\left(-\frac{1}{2}\left(m + s\frac{r_t - \mu_t}{h_t^{1/2}}\right)^2 \times \left(\frac{1 - I_t}{2\alpha} + \frac{I_t}{2(1-\alpha)}\right)\right)$$

with

$$\begin{aligned} I_t &= \begin{cases} 1 & \text{if } m + s\frac{r_t - \mu_t}{h_t^{1/2}} > 0 \\ 0 & \text{if } m + s\frac{r_t - \mu_t}{h_t^{1/2}} \leq 0 \end{cases} \\ m &= \frac{2\sqrt{2}}{\sqrt{\pi}}\left(-\alpha^2 + (1-\alpha)^2\right) \\ s &= \left[4\left(\alpha^3 + (1-\alpha)^3\right) - m^2\right]^{1/2}. \end{aligned}$$

The natural logarithm is

$$\log p(r_t) = \log(s) - \frac{1}{2}\log(2\pi h_t) - \frac{1}{2}\left(m + s\frac{r_t - \mu_t}{h_t^{1/2}}\right)^2 \times \left(\frac{1 - I_t}{2\alpha} + \frac{I_t}{2(1-\alpha)}\right).$$

Note: m and s are not separate parameters, these are simply functions of α.

For $\alpha = \frac{1}{2}$, this split normal distribution (with $m = 0$ and $s = 1$) reduces to the symmetric normal distribution. For $\alpha < \frac{1}{2}$ the distribution has a positive skewness. For $\alpha > \frac{1}{2}$ the distribution has a negative skewness. For $\alpha \neq \frac{1}{2}$ the distribution has kurtosis larger than 3.

An advantage of this distribution is the possible asymmetry. However, a disadvantage is that it does not allow for very fat tails.

2.5 Asymmetric Laplace distribution

The asymmetric Laplace (AL) distribution of Chen et al. (2012) with asymmetry parameter α has probability density function:

$$p(r_t) = \frac{s}{\sqrt{h_t}} \times \exp\left(- \left| s \frac{r_t - \mu_t}{h_t^{1/2}} + m \right| \times \left(\frac{1 - I_t}{\alpha} + \frac{I_t}{1-\alpha} \right) \right)$$

with

$$\begin{aligned} I_t &= \begin{cases} 1 & \text{if } s\frac{r_t-\mu_t}{h_t^{1/2}} + m \geq 0 \\ 0 & \text{if } s\frac{r_t-\mu_t}{h_t^{1/2}} + m < 0 \end{cases} \\ m &= 1 - 2\alpha \\ s &= \sqrt{\alpha^2 + (1-\alpha)^2}. \end{aligned}$$

The natural logarithm is

$$\log p(r_t) = \log(s) - \frac{1}{2}\log(h_t) - \left| s \frac{r_t - \mu_t}{h_t^{1/2}} + m \right| \left(\frac{1 - I_t}{\alpha} + \frac{I_t}{1-\alpha} \right).$$

Note: m and s are not separate parameters, these are simply functions of α.

For $\alpha = \frac{1}{2}$, this asymmetric Laplace distribution (with $m = 0$ and $s = \sqrt{2}$) reduces to the symmetric Laplace distribution. For $\alpha < \frac{1}{2}$ the distribution has a positive skewness. For $\alpha > \frac{1}{2}$ the distribution has a negative skewness.

The fat tails and possible asymmetry imply that this distribution can be a wise choice for daily returns.

3 Distributions with two shape parameters

3.1 Skewed Student-t distribution (Fernandez and Steel, 1998)

The skewed Student-t distribution of Fernandez and Steel (1998) (rescaled by Lambert and Laurent (2000, 2001)) with $\nu > 2$ degrees of freedom and asymmetry parameter $\xi > 0$ has probability density function:

$$p(r_t) = \frac{\Gamma(\frac{\nu+1}{2})}{\Gamma(\frac{\nu}{2})}((\nu-2)\pi h_t)^{-1/2} \times s \times \left(\frac{2}{\xi + \frac{1}{\xi}}\right) \times \left(1 + \frac{\left(s\frac{r_t-\mu_t}{h_t^{1/2}} + m\right)^2}{\nu - 2}\xi^{-2I_t}\right)^{-\frac{\nu+1}{2}}$$

with

$$\begin{aligned}
I_t &= \begin{cases} 1 & \text{if } s\frac{r_t-\mu_t}{h_t^{1/2}} + m \geq 0 \\ \\ -1 & \text{if } s\frac{r_t-\mu_t}{h_t^{1/2}} + m < 0 \end{cases} \\
m &= \frac{\Gamma(\frac{\nu-1}{2})}{\Gamma(\frac{\nu}{2})} \times \sqrt{\frac{\nu-2}{\pi}} \times \left(\xi - \frac{1}{\xi}\right) \\
s &= \sqrt{\left(\xi^2 + \frac{1}{\xi^2} - 1\right) - m^2}.
\end{aligned}$$

The natural logarithm is

$$\begin{aligned}
\log p(r_t) &= \log\Gamma\left(\frac{\nu+1}{2}\right) - \log\Gamma\left(\frac{\nu}{2}\right) - \frac{1}{2}\log((\nu-2)\pi h_t) \\
&\quad + \log(s) + \log\left(\frac{2}{\xi + \frac{1}{\xi}}\right) \\
&\quad - \frac{\nu+1}{2}\log\left(1 + \frac{\left(s\frac{r_t-\mu_t}{h_t^{1/2}} + m\right)^2}{\nu-2}\xi^{-2I_t}\right).
\end{aligned}$$

Note: m and s are not separate parameters, these are simply functions of ν and ξ.

For $\xi = 1$, this skewed Student-t distribution (with $m = 0$ and $s = 1$) reduces to the symmetric Student-t distribution with degrees of freedom parameter ν. For $\xi > 1$ the distribution has a positive skewness (as long as $\nu > 3$). For $0 < \xi < 1$ the distribution has a negative skewness (as long as $\nu > 3$).[2]

The fat tails and possible asymmetry imply that this distribution can be a wise choice for daily returns.[3]

[2]If $\xi > 1$ ($0 < \xi < 1$) and $2 < \nu \leq 3$, then the distribution is still right-skewed (left-skewed), but the skewness is not defined then due to the fatness of the tails.

[3]The skewed Student-t distribution of Hansen (1994) with $\nu > 2$ degrees of freedom and asymmetry parameter λ ($-1 < \lambda < 1$) has probability density function:

$$p(r_t) = \frac{bc}{h_t^{1/2}} \left(1 + \frac{1}{\nu - 2} \left(\frac{b\frac{r_t - \mu_t}{h_t^{1/2}} + a}{1 + I_t \lambda} \right)^2 \right)^{-\frac{\nu+1}{2}}$$

with

$$\begin{aligned} I_t &= \begin{cases} 1 & \text{if } b\frac{r_t-\mu_t}{h_t^{1/2}} + a \geq 0 \\ -1 & \text{if } b\frac{r_t-\mu_t}{h_t^{1/2}} + a < 0 \end{cases} \\ c &= \frac{\Gamma(\frac{\nu+1}{2})}{\sqrt{\pi(\nu-2)}\Gamma(\frac{\nu}{2})} \\ a &= 4\lambda c \frac{\nu-2}{\nu-1} \\ b &= \sqrt{1 + 3\lambda^2 - a^2}. \end{aligned}$$

This is the same distribution as the skewed Student-t distribution of Fernandez and Steel (1998), with $\lambda = \frac{\xi^2 - 1}{1 + \xi^2}$ or equivalently $\xi = \sqrt{\frac{1+\lambda}{1-\lambda}}$.

3.2 Generalized Hyperbolic Skewed Student-t (GHSKT) distribution

The Generalized Hyperbolic Skewed Student-t (GHSKT) distribution of Catania and Grassi (2017) with degrees of freedom parameter $\nu > 4$ and asymmetry parameter $\bar{\beta}$ (with $\bar{\beta} \neq 0$) has probability density function:

$$
\begin{aligned}
p(r_t) \quad = \quad & \frac{2^{\frac{1-\nu}{2}} \delta^{\nu} |\beta|^{\frac{\nu+1}{2}}}{h_t^{1/2} \Gamma(\frac{\nu}{2}) \sqrt{\pi}} \times K_{\frac{\nu+1}{2}} \left(\sqrt{\beta^2 \left(\delta^2 + \left(\frac{r_t - \mu_t}{h_t^{1/2}} - m \right)^2 \right)} \right) \times \\
& \exp \left(\beta \left(\frac{r_t - \mu_t}{h_t^{1/2}} - m \right) \right) \times \left(\delta^2 + \left(\frac{r_t - \mu_t}{h_t^{1/2}} - m \right)^2 \right)^{-\frac{\nu+1}{4}}
\end{aligned}
$$

with

$$
\begin{aligned}
\delta \quad &= \quad \left(\frac{2\bar{\beta}^2}{(\nu-2)^2(\nu-4)} + \frac{1}{\nu-2} \right)^{-\frac{1}{2}} \\
\beta \quad &= \quad \frac{\bar{\beta}}{\delta} \\
m \quad &= \quad -\frac{\beta\delta^2}{\nu-2}.
\end{aligned}
$$

$K_a(x)$ is the modified Bessel function of the third kind (also called the modified Bessel function of the second kind, to make matters more confusing), evaluated in x with order a (Abramowitz and Stegun, 1964). The natural logarithm is

$$
\begin{aligned}
\log p(r_t) \quad = \quad & \frac{1-\nu}{2} \log(2) + \nu \log(\delta) + \frac{\nu+1}{2} \log|\beta| - \log\Gamma\left(\frac{\nu}{2}\right) - \frac{1}{2} \log(\pi h_t) \\
& + \log K_{\frac{\nu+1}{2}} \left(\sqrt{\beta^2 \left(\delta^2 + \left(\frac{r_t - \mu_t}{h_t^{1/2}} - m \right)^2 \right)} \right) \\
& + \beta \left(\frac{r_t - \mu_t}{h_t^{1/2}} - m \right) - \frac{\nu+1}{4} \log \left(\delta^2 + \left(\frac{r_t - \mu_t}{h_t^{1/2}} - m \right)^2 \right).
\end{aligned}
$$

Note: δ, β and m are not separate parameters, these are simply functions of ν and $\bar{\beta}$.

For $\bar{\beta} \to 0$, the GHSKT distribution (with $\tilde{\mu} \to 0$ and $\beta \to 0$) tends to the symmetric Student-t distribution with degrees of freedom parameter ν. For $\bar{\beta} > 0$ the distribution has a positive skewness (as long as $\nu > 6$). For $\bar{\beta} < 0$ the distribution has a negative skewness (as long as $\nu > 6$).[4]

The fat tails and possible asymmetry imply that this distribution can be a wise choice for daily returns.

[4]If $\bar{\beta} > 0$ ($\bar{\beta} < 0$) and $4 < \nu \leq 6$, then the distribution is still right-skewed (left-skewed), but the skewness is not defined then due to the fatness of the tails.

If during the numerical optimization of the loglikelihood (where one should either use initial value $\bar{\beta} \neq 0$ or define the density to be equal to the Student-t density if $\bar{\beta} = 0$, in order to prevent a 'crash' due to evaluating the loglikelihood for $\bar{\beta} = 0$) the parameter ν tends to its lower bound 4, then this may indicate that a distribution with fatter tails is needed. In this case one can make use of the following GHSKT distribution of Aas and Haff (2006), with degrees of freedom parameter $\nu > 0$ (allowing $0 < \nu \leq 4$) and probability density function

$$\begin{aligned} p(r_t) &= \frac{2^{\frac{1-\nu}{2}} \delta_t^{\nu} |\beta|^{\frac{\nu+1}{2}}}{\Gamma(\frac{\nu}{2})\sqrt{\pi}} \times K_{\frac{\nu+1}{2}} \left(\sqrt{\beta^2 \left(\delta_t^2 + (r_t - \mu_t)^2 \right)} \right) \times \\ &\quad \exp\left(\beta(r_t - \mu_t)\right) \times \left(\delta_t^2 + (r_t - \mu_t)^2 \right)^{-\frac{\nu+1}{4}} \end{aligned}$$

with natural logarithm

$$\begin{aligned} \log p(r_t) &= \frac{1-\nu}{2} \log(2) + \nu \log(\delta_t) + \frac{\nu+1}{2} \log|\beta| - \log\Gamma\left(\frac{\nu}{2}\right) - \frac{1}{2}\log(\pi) \\ &\quad + \log K_{\frac{\nu+1}{2}} \left(\sqrt{\beta^2 \left(\delta_t^2 + (r_t - \mu_t)^2 \right)} \right) \\ &\quad + \beta\left(r_t - \mu_t\right) - \frac{\nu+1}{4} \log\left(\delta_t^2 + (r_t - \mu_t)^2 \right). \end{aligned}$$

If $2 < \nu \leq 4$, then the variance is infinite. If $0 < \nu \leq 2$, the mean and variance are undefined. Still μ_t and δ_t affect the 'location' and 'scale' of the distribution (for example, in terms of the location and width of prediction intervals). For $\bar{\beta} < 0$ ($\bar{\beta} > 0$) this distribution allows for an extremely fat left (right) tail. If $\nu > 4$, then this is the abovementioned GHSKT distribution (with a different parametrization).

3.3 Exponential Generalized Beta distribution of the second kind (EGB2)

The Exponential Generalized Beta distribution of the second kind (EGB2; McDonald (1991) extended the GB2 distribution of McDonald (1984)) with shape parameters $p > 0$ and $q > 0$ has probability density function:

$$p(r_t) = \frac{\sqrt{\Omega}\exp\left(p\left(\sqrt{\Omega}\frac{r_t-\mu_t}{\sqrt{h_t}} + \Delta\right)\right)}{\sqrt{h_t}B(p,q)\left(1+\exp\left(\sqrt{\Omega}\frac{r_t-\mu_t}{\sqrt{h_t}} + \Delta\right)\right)^{p+q}}$$

with

$$\begin{aligned} \Delta &= \psi(p) - \psi(q) \\ \Omega &= \psi'(p) + \psi'(q) \end{aligned}$$

where $\psi()$ is the digamma function (the first order derivative of the logarithm of the Gamma function $\log\Gamma(.)$) and $\psi'()$ is the trigamma function (the second order derivative of the logarithm of the Gamma function $\log\Gamma(.)$). $B(p,q)$ is the Beta function $B(p,q) = \frac{\Gamma(p)\Gamma(q)}{\Gamma(p+q)}$.

The natural logarithm is

$$\begin{aligned} \log p(r_t) &= \frac{1}{2}\log(\Omega) + p\left(\sqrt{\Omega}\frac{r_t-\mu_t}{\sqrt{h_t}} + \Delta\right) \\ &\quad -\frac{1}{2}\log(h_t) - \log B(p,q) - (p+q)\log\left(1+\exp\left(\sqrt{\Omega}\frac{r_t-\mu_t}{\sqrt{h_t}} + \Delta\right)\right). \end{aligned}$$

Note: Δ and Ω are not separate parameters, these are simply functions of p and q.

The distribution is symmetric if $p = q$. If $p = q \to \infty$, then the distribution tends to a normal distribution. The skewness is positive if $p > q$. The skewness is negative if $p < q$. The skewness lies between -2 and 2. The kurtosis can take values up to 9.[5]

The fat tails (with kurtosis up to 9) and possible asymmetry imply that this distribution can be a wise choice for daily returns (as long as one does not need an extremely fat-tailed distribution).

[5] The skewness is equal to

$$\frac{\psi''(p) - \psi''(q)}{(\psi'(p) + \psi'(q))^{3/2}}.$$

The kurtosis is equal to

$$3 + \frac{\psi'''(p) + \psi'''(q)}{(\psi'(p) + \psi'(q))^2},$$

where $\psi''()$ and $\psi'''()$ are the third and fourth order derivatives of the logarithm of the Gamma function $\log\Gamma()$.

3.4 Johnson's S_U distribution

Johnson's S_U distribution (Johnson, 1949) with asymmetry parameter γ and kurtosis parameter $\delta > 0$ has probability density function:

$$p(r_t) = \frac{\delta}{\lambda\sqrt{2\pi h_t}} \times \left(1 + \left(\frac{\frac{r_t-\mu_t}{h_t^{1/2}} - \xi}{\lambda}\right)^2\right)^{-1/2} \times \exp\left(-\frac{1}{2}\left(\gamma + \delta\sinh^{-1}\left(\frac{\frac{r_t-\mu_t}{h_t^{1/2}} - \xi}{\lambda}\right)\right)^2\right)$$

with

$$\begin{aligned} \lambda &= \sqrt{2} \times \left(\exp(\delta^{-2}) - 1\right)^{-1/2} \times \left(\exp(\delta^{-2})\cosh\left(\frac{2\gamma}{\delta}\right) + 1\right)^{-1/2} \\ \xi &= \lambda\exp\left(\frac{\delta^{-2}}{2}\right)\sinh\left(\frac{\gamma}{\delta}\right). \end{aligned}$$

The hyperbolic sine, hyperbolic cosine and inverse hyperbolic sine are given by

$$\begin{aligned} \sinh(x) &= \frac{\exp(x) - \exp(-x)}{2} \\ \cosh(x) &= \frac{\exp(x) + \exp(-x)}{2} \\ \sinh^{-1}(x) &= \log\left(x + \sqrt{x^2+1}\right). \end{aligned}$$

The natural logarithm is

$$\begin{aligned} \log p(r_t) &= \log(\delta) - \log(\lambda) - \frac{1}{2}\log(2\pi h_t) - \frac{1}{2}\left(1 + \left(\frac{\frac{r_t-\mu_t}{h_t^{1/2}} - \xi}{\lambda}\right)^2\right) \\ &\quad -\frac{1}{2}\left(\gamma + \delta\sinh^{-1}\left(\frac{\frac{r_t-\mu_t}{h_t^{1/2}} - \xi}{\lambda}\right)\right)^2. \end{aligned}$$

Note: λ and ξ are not separate parameters, these are simply functions of γ and δ.

For $\gamma = 0$, this distribution is symmetric. For $\gamma > 0$, the distribution has a negative (!) skewness. For $\gamma < 0$, the distribution has a positive (!) skewness. The kurtosis decreases for increasing values of δ; for $\delta \to \infty$, the distribution tends to a normal distribution.

The fat tails and possible asymmetry imply that this distribution can be a wise choice for daily returns.

3.5 Mixture of two normal distributions (Bai et al., 2001)

The mixture of two normal distributions of Bai et al. (2001, 2003) with two parameters ρ ($\frac{1}{2} \leq \rho < 1$) and λ ($0 < \lambda < 1$) is given by:

$$\frac{r_t - \mu_t}{h_t^{1/2}} \sim \begin{cases} N(0, \sigma^2) & \text{with probability } \rho, \\ N(0, \frac{\sigma^2}{\lambda}) & \text{with probability } 1-\rho, \end{cases}$$

or

$$r_t \sim \begin{cases} N(\mu_t, h_t\sigma^2) & \text{with probability } \rho, \\ N(\mu_t, \frac{h_t\sigma^2}{\lambda}) & \text{with probability } 1-\rho \end{cases}$$

with

$$\sigma^2 = \frac{1}{\rho + \frac{1-\rho}{\lambda}}$$

so that $\text{var}(\frac{r_t - \mu_t}{h_t^{1/2}}) = \rho\sigma^2 + (1-\rho)\frac{\sigma^2}{\lambda} = \frac{\rho + \frac{1-\rho}{\lambda}}{\rho + \frac{1-\rho}{\lambda}} = 1$ and $\text{var}(r_t) = h_t$. The probability density function is

$$p(r_t) \;=\; \rho \; \left(2\pi h_t \sigma^2\right)^{-1/2} \exp\left(-\frac{(r_t-\mu_t)^2}{2h_t\sigma^2}\right) + (1-\rho) \; \left(\frac{2\pi h_t \sigma^2}{\lambda}\right)^{-1/2} \exp\left(-\frac{(r_t-\mu_t)^2}{2h_t\frac{\sigma^2}{\lambda}}\right).$$

The natural logarithm is

$$\log p(r_t) \;=\; \log\left[\rho \; \left(2\pi h_t \sigma^2\right)^{-1/2} \exp\left(-\frac{(r_t-\mu_t)^2}{2h_t\sigma^2}\right) + (1-\rho) \; \left(\frac{2\pi h_t \sigma^2}{\lambda}\right)^{-1/2} \exp\left(-\frac{(r_t-\mu_t)^2}{2h_t\frac{\sigma^2}{\lambda}}\right)\right].$$

Note: σ^2 is not a separate parameter, this is simply a function of ρ and λ.

There is one state (or component or regime) with probability ρ with small variance $h_t\sigma^2 < h_t$ (since $\sigma^2 < 1$) and one state with probability $1-\rho$ with large variance $h_t\frac{\sigma^2}{\lambda} > h_t$, where the state with the small variance has at least as much probability as the state with the large variance (since $\frac{1}{2} \leq \rho < 1$). $\frac{1}{\lambda} > 1$ is the ratio of the large and small variance.

This distribution is symmetric and has a higher kurtosis than the normal distribution. The distribution tends to the normal distribution if $\lambda \to 1$ (when the difference between the variances disappears) and/or $\rho \to 1$ (when the probability of the state with the high variance tends to 0).

This distribution can be a wise choice for daily returns, as long as the empirical distribution of the daily returns is (almost) symmetric.

3.6 Two-sided Weibull distribution

The Two-sided Weibull distribution of Chen and Gerlach (2013) with skewness parameter α ($0 < \alpha < 1$) and kurtosis parameter $\kappa > 0$ has probability density function:

$$\begin{aligned} p(r_t) &= \frac{s}{h_t^{1/2}} \left[(1 - I_t) \left(\frac{-\left(m + s\frac{r_t - \mu_t}{h_t^{1/2}}\right)}{\alpha\kappa} \right)^{\kappa - 1} \exp\left(- \frac{-\left(m + s\frac{r_t - \mu_t}{h_t^{1/2}}\right)}{\alpha\kappa} \right) \right. \\ &\quad \left. + I_t \left(\frac{m + s\frac{r_t - \mu_t}{h_t^{1/2}}}{(1 - \alpha)\kappa} \right)^{\kappa - 1} \exp\left(- \frac{m + s\frac{r_t - \mu_t}{h_t^{1/2}}}{(1 - \alpha)\kappa} \right) \right] \end{aligned}$$

with

$$\begin{aligned} I_t &= \begin{cases} 1 & \text{if } m + s\frac{r_t - \mu_t}{h_t^{1/2}} > 0 \\ 0 & \text{if } m + s\frac{r_t - \mu_t}{h_t^{1/2}} \leq 0 \end{cases} \\ m &= \left(-\alpha^2 + (1 - \alpha)^2\right) \kappa \, \Gamma\left(1 + \frac{1}{\kappa}\right) \\ s &= \left[(\alpha^3 + (1 - \alpha)^3)\kappa^2 \Gamma\left(1 + \frac{2}{\kappa}\right) - m^2 \right]^{1/2}. \end{aligned}$$

The natural logarithm is

$$\begin{aligned} \log p(r_t) &= \log(s) - \frac{1}{2}\log(h_t) \\ &\quad + (1 - I_t) \left((\kappa - 1) \log\left(\frac{-\left(m + s\frac{r_t - \mu_t}{h_t^{1/2}}\right)}{\alpha\kappa} \right) - \frac{-\left(m + s\frac{r_t - \mu_t}{h_t^{1/2}}\right)}{\alpha\kappa} \right) \\ &\quad + I_t \left((\kappa - 1) \log\left(\frac{m + s\frac{r_t - \mu_t}{h_t^{1/2}}}{(1 - \alpha)\kappa} \right) - \frac{m + s\frac{r_t - \mu_t}{h_t^{1/2}}}{(1 - \alpha)\kappa} \right). \end{aligned}$$

Note: m and s are not separate parameters, these are simply functions of α and κ.

The definition of this Two-sided Weibull distribution is that

$$\begin{cases} -(m + s\frac{r_t - \mu_t}{h_t^{1/2}}) & \sim \ \text{Weibull}(\lambda_1, \kappa) \ \text{ with probability } \alpha \\ m + s\frac{r_t - \mu_t}{h_t^{1/2}} & \sim \ \text{Weibull}(\lambda_2, \kappa) \ \text{ with probability } 1 - \alpha \end{cases}$$

with $\lambda_1 = \alpha\kappa$, $\lambda_2 = (1 - \alpha)\kappa = \kappa - \lambda_1$. The Weibull distribution, introduced by Weibull (1951), is a continuous distribution, where Weibull distributed random variables take non-negative values. It is a special case of an extreme value distribution and of the generalised gamma distribution; the exponential distribution is a special case of the Weibull distribution.

The smaller the value of κ, the fatter the tails of the distribution, where $\kappa = 1$ corresponds to Laplace tails. For $\alpha < \frac{1}{2}$, $\alpha = \frac{1}{2}$ and $\alpha > \frac{1}{2}$ the distribution is right-skewed

(!), symmetric and left-skewed. α is equal to the probability of stemming from the 'left part' of the distribution, $\Pr(m + s\frac{r_t-\mu_t}{h_t^{1/2}} \leq 0)$. For $\kappa \leq 1$ the distribution has one mode at $m + s\frac{r_t-\mu_t}{h_t^{1/2}}$. In that case $\alpha < \frac{1}{2}$ means that the mode is smaller than the median (a phenomenon typically observed for right-skewed distributions), whereas $\alpha > \frac{1}{2}$ means that the mode is larger than the median (a phenomenon typically observed for left-skewed distributions). For $\kappa > 1$ the distribution is bi-modal, having one mode at a value smaller than $m + s\frac{r_t-\mu_t}{h_t^{1/2}}$ and one mode at a value larger than $m + s\frac{r_t-\mu_t}{h_t^{1/2}}$. This may not be a good fit in the centre of the return distribution, but the tails (and thus risk measures such as the Value-at-Risk and Expected Shortfall) could still be estimated accurately.

The skewness is in the range $[-2.4, 2.4]$, and the kurtosis is in the range $[2.5, 11.5]$, so that this distribution has a considerable degree of flexibility in the shapes and tail properties.

For $\kappa = 1$ this distribution reduces to the Asymmetric Laplace (AL) distribution of Lu et al. (2010) and Chen et al. (2012). For $\kappa = 1$ and $\alpha = \frac{1}{2}$ this distribution reduces to the Laplace distribution.

The fat tails and possible asymmetry imply that this distribution can be a wise choice for daily returns.

4 Distributions with three shape parameters

4.1 Asymmetric Student-t (AST) distribution

The Asymmetric Student-t (AST) distribution of Zhu and Galbraith (2010) with so-called skewness parameter α ($0 < \alpha < 1$), left tail parameter $\nu_1 > 2$ and right tail parameter $\nu_2 > 2$ has probability density function:

$$p(r_t) = \frac{s\,B}{h_t^{1/2}} \left((1-I_t)\left(1+\frac{1}{\nu_1}\left(\frac{m+s\frac{r_t-\mu_t}{h_t^{1/2}}}{2\alpha^*}\right)^2\right)^{-\frac{\nu_1+1}{2}} + I_t\left(1+\frac{1}{\nu_2}\left(\frac{m+s\frac{r_t-\mu_t}{h_t^{1/2}}}{2(1-\alpha^*)}\right)^2\right)^{-\frac{\nu_2+1}{2}} \right)$$

with

$$\begin{aligned}
I_t &= \begin{cases} 1 & \text{if } m+s\frac{r_t-\mu_t}{h_t^{1/2}} > 0 \\ 0 & \text{if } m+s\frac{r_t-\mu_t}{h_t^{1/2}} \le 0 \end{cases} \\
K(\nu) &= \frac{\Gamma(\frac{\nu+1}{2})}{\sqrt{\pi\nu}\,\Gamma(\frac{\nu}{2})} \\
B &= \alpha K(\nu_1) + (1-\alpha)K(\nu_2) \\
\alpha^* &= \frac{\alpha K(\nu_1)}{B} \\
m &= 4B\left(-\alpha^{*2}\frac{\nu_1}{\nu_1-1} + (1-\alpha^*)^2\frac{\nu_2}{\nu_2-1}\right) \\
s &= \left[4\left(\alpha\alpha^{*2}\frac{\nu_1}{\nu_1-2} + (1-\alpha)(1-\alpha^*)^2\frac{\nu_2}{\nu_2-2}\right) - m^2\right]^{1/2}.
\end{aligned}$$

The natural logarithm is

$$\begin{aligned}
\log p(r_t) = {} & \log(s) + \log(B) - \frac{1}{2}\log(h_t) - (1-I_t)\,\frac{\nu_1+1}{2}\log\left(1+\frac{1}{\nu_1}\left(\frac{m+s\frac{r_t-\mu_t}{h_t^{1/2}}}{2\alpha^*}\right)^2\right) \\
& -I_t\,\frac{\nu_2+1}{2}\log\left(1+\frac{1}{\nu_2}\left(\frac{m+s\frac{r_t-\mu_t}{h_t^{1/2}}}{2(1-\alpha^*)}\right)^2\right).
\end{aligned}$$

Note: B, α^*, m and s are not separate parameters, these are simply functions of α, ν_1 and ν_2. ν_1 and ν_2 control the fatness of the left tail and right tail, respectively, where a lower value means a fatter tail. α can be roughly interpreted as the parameter reflecting the asymmetry of the central part of the distribution, where $\alpha = \frac{1}{2}$ reflects a symmetric shape, and where $\alpha < \frac{1}{2}$ and $\alpha > \frac{1}{2}$ reflect right-skewed (!) and left-skewed shapes, respectively.

The skewness is affected by all shape parameters α, ν_1 and ν_2. The AST density is continuous and unimodal with mode at $r_t = \mu_t - h^{1/2}\frac{m}{s}$. The mode is larger than the mean μ_t (a phenomenon typically observed for left-skewed distributions) if $-\alpha^{*2}\frac{\nu_1}{\nu_1-1} +$

$(1-\alpha^*)^2\frac{\nu_2}{\nu_2-1} < 0$, which may be the case for $\alpha > \frac{1}{2}$ and/or $\nu_1 < \nu_2$. The mode is smaller than the mean μ_t (a phenomenon typically observed for right-skewed distributions) if $-\alpha^{*2}\frac{\nu_1}{\nu_1-1} + (1-\alpha^*)^2\frac{\nu_2}{\nu_2-1} > 0$, which may be the case for $\alpha < \frac{1}{2}$ and/or $\nu_1 > \nu_2$.

For $\nu_1 = \nu_2 = \nu$ the AST distribution (with $B = K(\nu) = \frac{\Gamma(\frac{\nu+1}{2})}{\sqrt{\pi\nu}\,\Gamma(\frac{\nu}{2})}$, $\alpha^* = \alpha$) reduces to the skewed Student-t (SST) distribution of Hansen (1994) with $\alpha = \frac{1-\lambda}{2}$ $(\lambda = 1-2\alpha)$.

For $\alpha = \frac{1}{2}$ and $\nu_1 = \nu_2 = \nu$ the AST distribution (with $B = K(\nu) = \frac{\Gamma(\frac{\nu+1}{2})}{\sqrt{\pi\nu}\,\Gamma(\frac{\nu}{2})}$, $\alpha^* = \frac{1}{2}$, $m = 0$, $s = \sqrt{\frac{\nu}{\nu-2}}$) reduces to the symmetric Student-t distribution.

The fat tails and possible asymmetry imply that this distribution can be a wise choice for daily returns.

4.2 Asymmetric Exponential Power Distribution (AEPD)

The Asymmetric Exponential Power Distribution (AEPD) of Zhu and Zinde-Walsh (2009) with so-called skewness parameter α ($0 < \alpha < 1$), left tail parameter $\nu_1 > 0$ and right tail parameter $\nu_2 > 0$ has probability density function:

$$p(r_t) = \frac{s\,B}{h_t^{1/2}} \left((1-I_t) \exp\left(-\frac{1}{\nu_1} \left| \frac{m + s\frac{r_t-\mu_t}{h_t^{1/2}}}{2\alpha^*} \right|^{\nu_1} \right) + I_t \exp\left(-\frac{1}{\nu_2} \left| \frac{m + s\frac{r_t-\mu_t}{h_t^{1/2}}}{2(1-\alpha^*)} \right|^{\nu_2} \right) \right)$$

with

$$\begin{aligned}
I_t &= \begin{cases} 1 & \text{if } m + s\frac{r_t-\mu_t}{h_t^{1/2}} > 0 \\[2ex] 0 & \text{if } m + s\frac{r_t-\mu_t}{h_t^{1/2}} \leq 0 \end{cases} \\
K_{EP}(\nu) &= \frac{1}{2\,\nu^{1/\nu}\,\Gamma\left(1+\frac{1}{\nu}\right)} \\
B &= \alpha K_{EP}(\nu_1) + (1-\alpha) K_{EP}(\nu_2) \\
\alpha^* &= \frac{\alpha K_{EP}(\nu_1)}{B} \\
m &= \frac{1}{B}\left(-\alpha^2 \frac{\nu_1 \Gamma(2/\nu_1)}{(\Gamma(1/\nu_1))^2} + (1-\alpha)^2 \frac{\nu_2 \Gamma(2/\nu_2)}{(\Gamma(1/\nu_2))^2} \right) \\
s &= \left[\frac{\alpha^3}{B^2} \frac{\nu_1^2 \Gamma(3/\nu_1)}{(\Gamma(1/\nu_1))^3} + \frac{(1-\alpha)^3}{B^2} \frac{\nu_2^2 \Gamma(3/\nu_2)}{(\Gamma(1/\nu_2))^3} - m^2 \right]^{1/2}.
\end{aligned}$$

The natural logarithm is

$$\log p(r_t) = \log(s) + \log(B) - \frac{1}{2}\log(h_t) - (1-I_t)\,\frac{1}{\nu_1} \left| \frac{m + s\frac{r_t-\mu_t}{h_t^{1/2}}}{2\alpha^*} \right|^{\nu_1} - I_t\,\frac{1}{\nu_2} \left| \frac{m + s\frac{r_t-\mu_t}{h_t^{1/2}}}{2(1-\alpha^*)} \right|^{\nu_2}.$$

Note: B, α^*, m and s are not separate parameters, these are simply functions of α, ν_1 and ν_2. ν_1 and ν_2 control the fatness of the left tail and right tail, respectively, where a lower value means a fatter tail. $\nu = 2$ and $\nu = 1$ correspond to the tail of a normal and Laplace distribution, respectively. α can be roughly interpreted as the parameter reflecting the asymmetry of the central part of the distribution, where $\alpha = \frac{1}{2}$ reflects a symmetric shape, and where $\alpha < \frac{1}{2}$ and $\alpha > \frac{1}{2}$ reflect right-skewed (!) and left-skewed shapes, respectively.

The skewness is affected by all shape parameters α, ν_1 and ν_2. The AEPD density is continuous and unimodal with mode at $r_t = \mu_t - h^{1/2}\frac{m}{s}$. The mode is larger than the mean μ_t (a phenomenon typically observed for left-skewed distributions) if $-\alpha^2 \frac{\nu_1\Gamma(2/\nu_1)}{(\Gamma(1/\nu_1))^2} + (1-\alpha)^2 \frac{\nu_2\Gamma(2/\nu_2)}{(\Gamma(1/\nu_2))^2} < 0$, which may be the case for $\alpha > \frac{1}{2}$ and/or $\nu_1 < \nu_2$. The mode is smaller than the mean μ_t (a phenomenon typically observed for right-skewed distributions) if $-\alpha^2 \frac{\nu_1\Gamma(2/\nu_1)}{(\Gamma(1/\nu_1))^2} + (1-\alpha)^2 \frac{\nu_2\Gamma(2/\nu_2)}{(\Gamma(1/\nu_2))^2} > 0$, which may be the case for $\alpha < \frac{1}{2}$ and/or $\nu_1 > \nu_2$.

The probability that r_t is smaller than the mode is equal to α, $\Pr(r_t \leq \mu_t - h^{1/2}\frac{m}{s}) = \alpha$. So according to the skewness measure of Arnold and Groeneveld (1995), $1-2$ CDF(mode), the AEPD density is skewed to the left for $\alpha > \frac{1}{2}$ and to the right for $\alpha < \frac{1}{2}$. The mode is larger than the median (a phenomenon typically observed for left-skewed distributions) if $\alpha > \frac{1}{2}$. The mode is smaller than the median (a phenomenon typically observed for right-skewed distributions) if $\alpha < \frac{1}{2}$.

For $\nu_1 = \nu_2 = \nu$ the AEPD (with $B = K_{EP}(\nu) = \frac{1}{2\,\nu^{1/\nu}\,\Gamma\left(1+\frac{1}{\nu}\right)}$, $\alpha^* = \alpha$) reduces to the Skewed Exponential Power Distribution (SEPD) which is equivalent to the distributions of Fernandez et al. (1995), Komunjer (2007) and Theodossiou (2015). The skewed Laplace distribution and skewed normal distribution (which is a different distribution than the skew normal distribution of O'Hagan and Leonard (1976)) are special cases of the SEPD, respectively, with $\nu = 1$ and $\nu = 2$.

For $\alpha = \frac{1}{2}$ and $\nu_1 = \nu_2 = \nu$ the AEPD (with $B = K_{EP}(\nu) = \frac{1}{2\,\nu^{1/\nu}\,\Gamma\left(1+\frac{1}{\nu}\right)}$, $\alpha^* = \frac{1}{2}$, $m = 0$, $s = \nu^{1/\nu}\sqrt{\frac{\Gamma(3/\nu)}{\Gamma(1/\nu)}}$) reduces to the Exponential Power Distribution (EPD) or Generalized Error Distribution (GED) with shape parameter ν (where in the density in the exponent $-\frac{s^\nu}{\nu}$ of the AEPD is equal to $-\frac{1}{2\lambda^\nu}$ of the GED, and where 'scaling factor' sB of the AEPD is equal to $\left(2^{(1+(1/\nu))}\Gamma(1/\nu)\lambda\right)^{-1}\nu$ of the GED). So for $\alpha = \frac{1}{2}$ and $\nu_1 = \nu_2 = 2$ the AEPD reduces to the normal distribution, and for $\alpha = \frac{1}{2}$ and $\nu_1 = \nu_2 = 1$ the AEPD reduces to the Laplace distribution.

The fat tails and possible asymmetry imply that this distribution can be a wise choice for daily returns.

References

Aas K, Haff IH. 2006. The Generalized Hyperbolic Skew Student's t-distribution. *Journal of Financial Econometrics* **4**: 275–309.

Abramowitz M, Stegun IA. 1964. *Handbook of Mathematical Functions: with Formulas, Graphs, and Mathematical Tables*, volume 55. Courier Corporation.

Arnold BC, Groeneveld RA. 1995. Measuring skewness with respect to the mode. *The American Statistician* **49**: 34–38.

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

Bai X, Russell J, Tiao GC. 2001. Beyond Merton's utopia (I): effects of non-normality and dependence on the precision of the variance estimates using high-frequency financial data. Technical report, GSB, University of Chicago.

Bai X, Russell J, Tiao GC. 2003. Kurtosis of GARCH and stochastic volatility models with non-normal innovations. *Journal of Econometrics* **114**: 349–360.

Box GEP, Tiao GC. 1973. *Bayesian Inference in Statistical Analysis*. Addison Wesley.

Catania L, Grassi S. 2017. Modelling crypto-currencies financial time-series. working paper, available at SSRN 3028486.

Chen Q, Gerlach RH. 2013. The two-sided Weibull distribution and forecasting financial tail risk. *International Journal of Forecasting* **29**: 527–540.

Chen Q, Gerlach RH, Lu Z. 2012. Bayesian Value-at-Risk and expected shortfall forecasting via the asymmetric Laplace distribution. *Computational Statistics and Data Analysis* **56**: 3498–3516.

Fechner GT. 1897. *Kollectivmasslehre*. Engelmann, Leipzig. Ed. Lipps, G.F.

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

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

Gauss CF. 1809. Theoria motvs corporvm coelestivm in sectionibvs conicis Solem ambientivm.

Gibbons JF, Mylroie S. 1973. Estimation of impurity profiles in ion-implanted amorphous targets using joined half-gaussian distributions. *Applied Physics Letters* **22**: 568–569.

Gosset WS. 1908. The probable error of a mean. *Biometrika* **6**: 1–25.

Hansen BE. 1994. Autoregressive conditional density estimation. *International Economic Review* **35**: 705–730.

John S. 1982. The three-parameter two-piece normal family of distributions and its fitting. *Communications in Statistics - Theory and Methods* **11**: 879–885.

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

Komunjer I. 2007. Asymmetric power distribution: Theory and applications to risk measurement. *Journal of Applied Econometrics* **22**: 891–921.

Lambert P, Laurent S. 2000. Modelling skewness dynamics in series of financial data. Technical report, Institut de Statistique, Louvain-la-Neuve.

Lambert P, Laurent S. 2001. Modelling financial time series using GARCH-type models and a skewed student density. Technical report, Université de Liège.

Laplace PS. 1774. Mémoire sur la probabilité des causes par les évènements. *Mémoires de l'Academie Royale des Sciences Presentés par Divers Savan* **6**: 621–656.

Lu Z, Huang H, Gerlach RH. 2010. Estimating Value at Risk: from JP Morgan's standard-EWMA to skewed-EWMA forecasting. Technical report, University of Sydney.

McDonald JB. 1984. Some generalized functions for the size distribution of income. *Econometrica* **52**: 647–663.

McDonald JB. 1991. Parametric models for partially adaptive estimation with skewed and leptokurtic residuals. *Economics Letters* **37**: 273–278.

O'Hagan A, Leonard T. 1976. Bayes estimation subject to uncertainty about parameter constraints. *Biometrika* **63**: 201–203.

Subbotin MT. 1923. On the law of frequency of errors. *Matematicheskii Sbornik* **31**: 296–301.

Theodossiou P. 2015. Skewed Generalized Error Distribution of financial assets and option pricing. *Multinational Finance Journal* **19**: 223–266. Available at SSRN 219679 since 2000.

Weibull W. 1951. A statistical distribution function of wide applicability. *Journal of Applied Mechanics* **18**: 293–297.

Zhu D, Galbraith JW. 2010. A generalized asymmetric Student-t distribution with application to financial econometrics. *Journal of Econometrics* **157**: 297–305.

Zhu D, Zinde-Walsh V. 2009. Properties and estimation of asymmetric exponential power distribution. *Journal of Econometrics* **148**: 86–99.