

RESEARCH

Open Access



pTAS distributions with application to risk management

Matthias Fischer^{1*} and Kevin Jakob²

*Correspondence:
Matthias.Fischer@fau.de
¹Department of Statistics &
Econometrics, FAU
Erlangen-Nürnberg, Erlangen,
Germany
Full list of author information is
available at the end of the article

Abstract

The family of positive tempered α -stable (pTAS) or sometimes also tempered one-sided α -stable distributions dates back to Tweedie (1984) and Hougaard (1986) who discussed it in the context of frailty distribution in life table methods for heterogenous populations. The pTAS family generalizes the well-known gamma distribution and allows for heavier tails depending on the parameter α . Because of this property, pTAS distributions appear to be useful in the context of risk management. Against this background, the contribution of his work is three-fold: Firstly, we summarize the properties of the pTAS family. Secondly, we describe its numerical implementation and illustrate the functions by means of R examples in the Appendix. Thirdly, we empirically demonstrate that this family can be successfully applied in risk management. Concretely, applications to credit and operational risk are given.

Keywords: Stable distributions, Positive distributions, Estimation, Operational risk, Credit risk

Mathematics Subject Classification: 44A10; 60E07; 91B28

1 Derivation and properties of the pTAS family

1.1 Evolution of the pTAS family

The distribution of a (non-degenerate) random variable X is called stable if there exist constants $a_n > 0$ and b_n such that, for any $n > 1$, if X_1, X_2, \dots are i.i.d. copies of X and $S_n \equiv \sum_{i=1}^n X_i$, then $S_n \stackrel{d}{=} a_n X + b_n$. The distribution is called strictly stable if $b_n = 0$. The normalizing constants are necessarily of the form $a_n = n^{1/\alpha}$ for some $\alpha \in (0, 2]$, where α is called the index or characteristic exponent of the distribution. We also say that X is α -stable if it is stable with index α and characteristic function

$$\varphi(t; \alpha, \beta, \gamma, \delta) = \begin{cases} \exp(-\gamma^\alpha |t|^\alpha (1 - i\beta \tan(\pi\alpha/2) \operatorname{sgn}(t)) + i\delta t), & \alpha \neq 1 \\ \exp(-\gamma |t| (1 - i\beta \frac{2}{\pi} \log(t) \operatorname{sgn}(t)) + i\delta t), & \alpha = 1 \end{cases}$$

where $\beta \in [-1, 1]$ (skewness parameter), $\delta \in \mathbb{R}$ (location parameter) and $\gamma > 0$ (scale parameter), briefly $X \sim S_\alpha(\beta, \gamma, \delta)$. Forth on, the focus is on positive or one-sided α -stable (pAS) variables, i.e. $X > 0$ a.s. which corresponds to these cases (parameter sets) where $0 < \alpha < 1$, $\beta = 1$ and $\delta \geq 0$. In this case and for $\delta = 0$, the Laplace transform is finite for all $s > 0$ and given by (see, e.g. Nolan (2003) or Janson (2011, Theorem 3.12))

$$\mathcal{L}_{\text{pAS}}(s; \alpha, \gamma) = \mathbb{E}(e^{-sX}) = \exp(-\xi s^\alpha), \quad \alpha \in (0, 1] \text{ and } \xi = \frac{\gamma^\alpha}{\cos(\pi\alpha/2)}. \quad (1.1)$$

The positive or one-sided tempered α -stable distribution has its origin in Tweedie (1984) and Hougaard (1986). The process of tempering (also exponential tilting or Esscher transformation, Esscher (1932)) shortens the right-hand tail of the stable distribution so that, in contrast with the stable case, moments of all orders exist. Technically, a new proper density g is defined by

$$g_Y(y; \alpha, \gamma, \theta) := f_{\text{pAS}}(y; \alpha, \gamma) \frac{\exp(-\theta y)}{\mathcal{L}_{\text{pAS}}(\theta; \alpha, \gamma)}, \quad \theta \geq 0. \quad (1.2)$$

The distribution of the corresponding random variable Y is called positive tempered α -stable (pTAS) distribution. Using (1.1), (1.2), its Laplace transformation derives as

$$\mathcal{L}_Y(s) = \exp\left(-\frac{\gamma^\alpha}{\cos(\pi\alpha/2)}((\theta + s)^\alpha - \theta^\alpha)\right). \quad (1.3)$$

1.2 Properties of pTAS distributions and some remarks

Hougaard (1986) proved that all moments exist if $\theta > 0$. In particular, using (1.3),

$$\mu = \mathbb{E}(Y) = \frac{\alpha\xi}{\theta^{1-\alpha}} \text{ and } \sigma^2 = \text{Var}(Y) = \frac{\alpha(1-\alpha)\xi}{\theta^{2-\alpha}},$$

where $v = \sigma/\mu$ denotes the coefficient of variation. Moreover, skewness and kurtosis, defined by the third and fourth standardized moments read as

$$\mathbb{S}(Y) = \frac{v(2-\alpha)}{1-\alpha} \quad \text{and} \quad \mathbb{K}(Y) = \frac{v^2(2-\alpha)(3-\alpha)}{(1-\alpha)^2} + 3. \quad (1.4)$$

In the relevant literature, different parameterizations appear, summarized in Table 1, below.

Figure 1 illustrates the possible combinations of skewness and kurtosis that can be modeled via a pTAS distribution (gray shaded area). One can see that the area of possible combinations is bounded by the gamma distribution from below whereas the possible combinations of the Lognormal distribution are covered by the pTAS family. Therefore, we think that the pTAS distribution can serve as a good alternative whenever data are assumed to follow a lognormal distribution in general but some more flexibility regarding the tail behavior is required. For comparison reasons, we also depict the Weibull distribution which is often used in operational risk models (see Section 3.2).

Hougaard (1986) showed that all pTAS densities are unimodal and proved the following scaling-property: $cX \sim \mathcal{P}(\alpha, c^\alpha \delta, \theta/c)$ when $X \sim \mathcal{P}(\alpha, \delta, \theta)$. Furthermore, the pTAS family is self-decomposable (i.e. if X follows a pTAS distribution then for each $\lambda \in (0, 1)$ there exists a random variable Y independent of X such that $\mathcal{L}(X) = \mathcal{L}(\lambda X + Y)$) and

Table 1 Different Parametrizations: $\mathcal{P}(K) \Leftrightarrow \mathcal{P}(T)$ with $\beta = \alpha$, $\alpha = \theta^\alpha \cdot \delta/(1-\alpha)$, $\lambda = (1-\alpha)/\delta$.

$\mathcal{P}(H) \Leftrightarrow \mathcal{P}(T) \Leftrightarrow \mathcal{P}(P)$ using $\delta = \mu \left(\frac{1-\alpha}{\mu v^2}\right)^{1-\alpha}$, $\gamma = \left[\frac{\mu \cos(\pi\alpha/2)}{\alpha}\right]^{\frac{1}{\alpha}} \left[\frac{1-\alpha}{\mu v^2}\right]^{\frac{1-\alpha}{\alpha}}$ and $\theta = \frac{1-\alpha}{\mu v^2}$

Reference	Parameter	Abbreviation
Hougaard (1986)	(α, δ, θ)	$\mathcal{P}(H)$
Tweedie (1984) ^a , Haas and Pigorsch (2009)	(α, γ, δ)	$\mathcal{P}(T)$
Palmer et al. (2008)	(α, μ, v)	$\mathcal{P}(P)$
Küchler and Tappe (2011)	(β, α, λ)	$\mathcal{P}(K)$

^aTweedie (1984) considered exponential dispersion models with variance function of the form $V(\mu) \propto \mu^p$ such models with $p > 2$ coincide with the pTAS family with $\alpha = (p-2)/(p-1)$, see Jørgensen (1987), chapter 4, for a detailed account

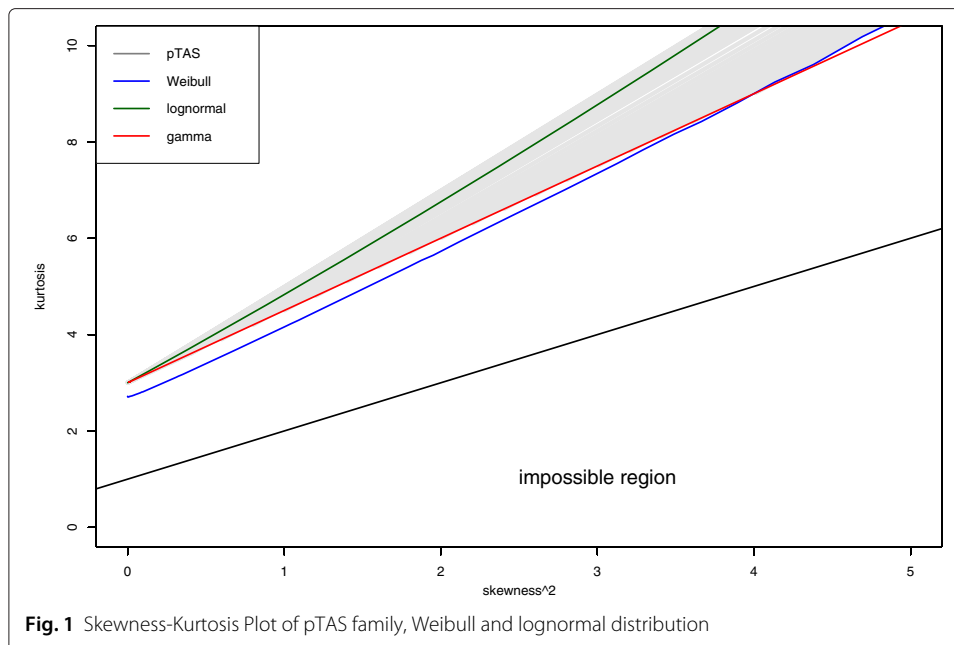


Fig. 1 Skewness-Kurtosis Plot of pTAS family, Weibull and lognormal distribution

infinitely divisible (i.e. for all $n \in \mathbb{N}$ exists a sequence of iid random variables $(X_j^{1/n})_{j=1}^n$ such that $X \stackrel{d}{=} X_1^{1/n} + \dots + X_n^{1/n}$. In this case, the characteristic function reads

$$\varphi_X(u) = \mathbb{E}(e^{iuX}) = \exp\left(i\gamma u - 0.5\sigma^2 u^2 + \int_{-\infty}^{\infty} (e^{iux} - 1 - iux1_{|x|<1}) \nu(dx)\right),$$

where $\gamma \in \mathbb{R}, \sigma \geq 0$ and ν denotes the Lévy measure. The triplet (γ, σ^2, ν) is called the Lévy triplet. If the Lévy measure $\nu(dx)$ has a density $k(x)$, then this density is called a Lévy density. Barndorff-Nielsen and Shephard (2001) discuss the Lévy-density of the pTAS distribution. After a suitable re-parametrization, one obtains

$$k(x) = \frac{\gamma 2^\alpha \alpha \exp(-\theta x)}{\Gamma(1-\alpha)x^{-(1+\alpha)}}, \quad x > 0.$$

for $0 < \alpha < 1, \theta > 0, \gamma > 0$. According to Lemma 4 in Küchler and Tappe (2011) it also holds that

$$\varphi_X(u) = \exp\left(\alpha \Gamma(-\beta) [(\lambda - iu)^\beta - \lambda^\beta]\right).$$

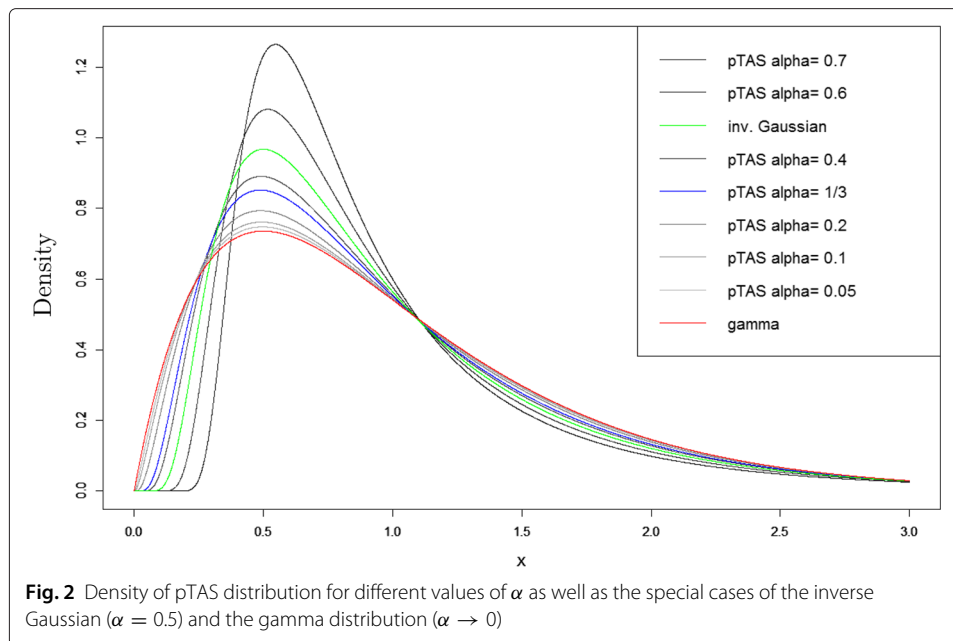
The pTAS density reduces to a simple closed form in three special cases. At first, letting $\alpha \rightarrow 0$, the gamma distribution with scale $\alpha_\Gamma = \frac{\theta^\alpha \delta}{1-\alpha}$, shape $\beta_\Gamma = \frac{1-\alpha}{\theta}$ is recovered. Secondly, setting $\alpha = 1/2$, the inverse Gaussian distribution with $\lambda_{IG} = \frac{\delta^2 \theta^{2\alpha-1}}{(1-\alpha)}$ arises. Finally, another special case arises when $\alpha = 1/3$ where the probability density function (pdf) can be expressed in closed form:

$$f_{\alpha=1/3}(y; \delta, \theta) = \frac{\sqrt{3}}{\pi} \left(\frac{\delta}{y}\right)^{1.5} K_{1/3}\left(2\sqrt{\delta^3 y^{-1}} \exp(3\delta\theta^{1/3} - \theta y)\right)$$

where $K_\lambda(x)$ denotes the modified Bessel function of the third kind with index λ .

The special cases as well as some density functions for further values of α are shown in Fig. 2.

The pTAS family can be considered as an alternative to the generalized inverse Gaussian (GIG) family (see Koudou and Ley (2014) for a review of this family) which itself arises



from the inverse Gaussian (IG) distribution by exponential tilting. Barndorff-Nielsen and Shephard (2001) introduce the four-parameter, so-called modified stable (MS) distributions which nests both GIG family and pTAS family (setting $\kappa = 0.5$ and $\nu = -\kappa$ in their notation), where ν denotes the fourth (additional) parameter.

2 Implementation issues

This chapter describes numerical algorithms which can be used for the standard functions (e.g. density and distribution function, random number generation and quantile function) of a pTAS random variable. The algorithms are described independently of a specific programming language. However, we give some examples how to use the pTAS R-package, which contains the described functions, in Appendix 4. The R-package is not yet published but can be obtained from the authors by request.

2.1 Density and distribution function

Since the pdf and cumulative distribution function (cdf) of the pTAS family are in general not available in closed form, but only via the Laplace transform, we have to use numerical inversion techniques to calculate values of the density or distribution function. Therefore, we follow a general approach given by Abate et al. (2000), which will be described briefly in this section. For further details please refer to the mentioned literature.

Starting from the Bromwich integral, the value of the pdf f at $t > 0$ can be recovered from the Laplace transform $\hat{f} := \mathcal{L}(f)$ by¹

$$f(t) = \frac{1}{2\pi i} \int_{b-i\infty}^{b+i\infty} e^{st} \hat{f}(s) ds = \frac{2e^{bt}}{\pi} \int_0^{\infty} \operatorname{Re} \left(\hat{f}(b + iu) \right) \cos(ut) du, \quad (2.1)$$

where $i = \sqrt{-1}$ and b is a real number greater than all singularities of \hat{f} . The second transformation is achieved via a change of variables.

Approximating the integral in (2.1) by the trapezoidal rule with step size $h > 0$, we get

$$f(t) \approx \frac{he^{bt}}{\pi} \left(\hat{f}(b) + 2 \sum_{k=1}^{\infty} \operatorname{Re} \left(\hat{f}(b + ikh) \right) \cos(kht) \right). \quad (2.2)$$

This equation gives rise to three kinds of error, a discretization error caused by the step size h , a truncation error and a round-off error because of the numerical operations performed by calculating the series.

The discretization error can be reduced by interpreting (2.1) as a Fourier series of a function $g(t) := e^{-bt}f(t)$, whereby the coefficients of the series itself can be expressed via the values of the Laplace transform. The error then can be reduced by increasing the ratio b/h . The round-off error can be controlled by choosing $h = \frac{\pi}{lt}$ and $b = \frac{A}{2lt}$ for some $l \in \mathbb{N} > 0$ and $A = \frac{2l}{2l+1} m \ln 10$, where in turn $m > 0$ should be chosen such that 10^{-m} is close to the machine precision. In order to improve the truncation error, the so called Euler summation technique can be applied, which gives us another parameter $n \in \mathbb{N}$ determining the number of coefficients being (fully) respected for the summation. Similarly to Abate et al. (2000) we recommend the following default values: $l = 1$, $n = 38$ and $m = 11$.

Finally, the pdf is approximated via

$$f(t) \approx \frac{\exp\left(\frac{A}{2lt}\right)}{2lt} \sum_{k=1}^m \binom{m}{k} 2^{-m} S_{n+k}, \quad (2.3)$$

with

$$\begin{aligned} S_0 &= \hat{f}\left(\frac{A}{2lt}\right) + 2 \sum_{j=1}^l \operatorname{Re} \left[\hat{f}\left(\frac{A}{2lt} + \frac{ij\pi}{lt}\right) \exp\left(\frac{ij\pi}{l}\right) \right] \quad \text{and} \\ S_k &= S_{k-1} + (-1)^k 2 \sum_{j=1}^l \operatorname{Re} \left[\hat{f}\left(\frac{A}{2lt} + \frac{ij\pi}{lt} + \frac{ik\pi}{t}\right) \exp\left(\frac{ij\pi}{l}\right) \right] \quad k \in \mathbb{N}_{>0}. \end{aligned} \quad (2.4)$$

Using the definition of the Laplace transform and partial integration it is easy to see that for $F(t) = \int_0^\infty f(t) dt$

$$\hat{F}(s) := \mathcal{L}(F)(s) = \mathcal{L}(f)(s)/s. \quad (2.5)$$

Therefore, with just a minor modification, the same method used to approximate the pdf can be used to approximate the cdf of a pTAS distribution.

2.2 Quantiles and random numbers

To calculate the quantile for a given level $p \in (0, 1)$, Ridout (2008) proposed a modified version of Newton's method to invert the cdf. Because, especially in the upper and lower tail of the distribution, the pdf may be close to zero, which may lead to an iteration step outside a given interval $[t_{\min}, t_{\max}]$. In this case, the Newton step is replaced by a bisection method.

Given a tolerance $\epsilon > 0$, a maximum number of iterations $N_{\max} \in \mathbb{N}$ and a closed interval $[t_{\min}, t_{\max}]$ such that $\exists t \in [t_{\min}, t_{\max}]: |F(t) - p| \leq \epsilon$ the following algorithm can be used to calculate the p -quantile of a pTAS distribution:

In order to determine the initial values of the lower and upper bounds, we propose to store some values for t and $F(t)$ on a grid $t_{i=1, \dots, N}$ depending on mean and standard

Algorithm 1 Inversion of the CDF

Set initial values

 $n \leftarrow 1$ $t_1 \leftarrow (t_{\min} + t_{\max}) / 2$ While $n \leq N$ and $|F(t) - p| > \epsilon$ $t_{n+1} \leftarrow t_n - (F(t_n) - p) / f(t_n)$

Newton step

if $t_{n+1} > t_{\max}$ or $t_{n+1} < t_{\min}$

bisection method

 $t_{n+1} \leftarrow (t_{\min} + t_{\max}) / 2$ if $F(t_{n+1}) > p$

update lower/upper bounds

 $t_{\max} \leftarrow t_{n+1}$ if $F(t_{n+1}) < p$ $t_{\min} \leftarrow t_{n+1}$ $n \leftarrow n + 1$ if $n > N$ *Warning: Iteration not successful.*

else

return t_n

deviation. Using precalculated values for t_{\min} and t_{\max} increases performance especially when quantile transformations should be performed very often.

If multiple quantiles should be calculated at the same time for different probabilities $p_{m=1,\dots,M}$, the probabilities should be sorted in ascending order before the transformation. Assuming that $p_1 \leq \dots \leq p_M$ we can apply Algorithm 1 on the first element p_1 and use the resulting value $t_1 = F^{-1}(p_1)$ as starting point (and also as lower bound) for the next element p_2 .

Algorithm 2 Multiple Inversion of CDF

Choose initial bounds (from precalculated grid):

 $t_{\min} > 0 : F(t_{\min}) < \min_{m=1,\dots,M} p_m$ $t_{\max} > 0 : F(t_{\max}) > \max_{m=1,\dots,M} p_m$

Calculate first quantile via Algorithm 1:

 $t_1 = F^{-1}(p_1)$ For $m = 2, \dots, M$ $t_m \leftarrow F^{-1}(p_m)$ # via Algorithm 1 with starting value t_{m-1} $t_{\min} \leftarrow t_m$

For random number generation the inverse probability integral transform method can be applied very easily by using Algorithm 2 based on sorted uniform drawings.

2.3 Estimating parameters

Given a series of N positive observations x_1, \dots, x_N a pTAS distribution can be fitted via various methods. First of all, a numerical maximum likelihood estimation (MLE) can be used. Within the package (see appendix A) we use the *mle()* function of the R-package *stats4* which in turn uses the *optim()* optimizer of R in order to perform a maximum

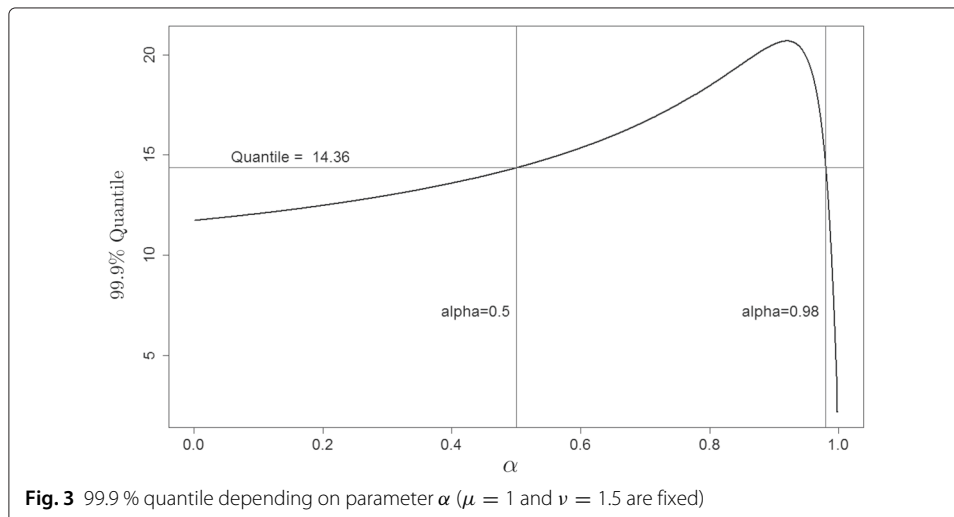
likelihood estimation. In particular, for observations t_1, \dots, t_N the parameters are determined by

$$(\alpha^*, \mu^*, \sigma^*) = \operatorname{argmin}_{\alpha, \mu, \sigma} - \sum_{n=1}^N \log f_{(\alpha, \mu, \sigma)}(t_n),$$

whereas for the calculation of the pdf $f_{(\alpha, \mu, \sigma)}(t)$ the algorithm described in 2.1 is used. As starting values we use the empirical mean and variance and a value of 0.5 for α . In order to ensure that the parameters stay in their specific range, we use the L-BFGS-B method with box constraints. Alternatively, distribution parameters can be estimated based on mean and variance, which determine parameters μ and ν regarding the \mathcal{P}_p parametrization and either the skewness, the kurtosis or a quantile to estimate parameter α . Given a skewness or kurtosis value (besides mean and variance), parameter α can be calculated by using Eq. (1.4). Please note, that the minimal skewness for a pTAS distribution equals 2ν . Please also note, that if α should be estimated based on a given kurtosis value the solution may not be unique (see Eq. 1.4).

A similar problem occurs if α should be estimated based on a given quantile t^* for probability p^* . Therefore, we propose to start with a grid $\alpha_{i=1, \dots, N-1}$ (e.g. $\alpha_i = \frac{i}{N}$ for $i = 1, \dots, N-1$ for some $N \in \mathbb{N}$) and calculate quantiles $t_i = F_{\mathcal{P}_p(\alpha_i, \mu, \nu)}^{-1}(p^*)$ in order to determine the number of possible solutions as well as lower and upper bounds for each of them. Given a lower and upper bound for each possible solution, a standard root finding method (e.g. bisection method) can be applied. In order to choose a proper value of N , we recommend to visualize the problem at first with the help of a plot like the one given in Fig. 3, which shows an example containing two possible solutions for α . In this case, we fixed the parameters μ and ν based on mean and variance and want to find a solution for α such that $F_{\mathcal{P}_p(\alpha, \mu, \nu)}(14.36) = 0.999$. As Fig. 3 shows, we get the two solutions $\alpha_1 = 0.5$ and $\alpha_2 = 0.98$.

In cases with no unique solution, the resulting parameters may lead to very different distributions (e.g. with different tails) which, in the context of risk management, can cause different risk figures. Therefore, we recommend to choose the estimation method with respect to the proposed area of application and to compare estimation results between



different methods if they are not unique. Otherwise, using inappropriate parameter estimators may lead to an increase of model risk.

3 Application to risk management

After showing how the pTAS distribution can be implemented and parameters can be estimated, we provide two extensive risk management applications of the pTAS family.

3.1 Credit risk

Our first application deals with the quantification of credit risk. Given a particular credit portfolio of N counterparties, financial institutions use credit portfolio models to estimate the distribution of the portfolio loss over a fixed time horizon (usually one year) due to counterparties' defaults². For each counterparty we denote the exposure at default by $EAD_i > 0$, the loss given default (as fraction of EAD_i) by $LGD_i \in [0, 1]$ and the probability of default by $PD_i \in (0, 1)$. The portfolio loss L then reads as

$$L = \sum_{i=1}^N EAD_i \cdot LGD_i \cdot D_i,$$

where D_i is a Bernoulli distributed random number with probability PD_i representing the default ($D_i = 1$) or the survival of counterparty i .

Given these counterparty specific information, a crucial task of a credit portfolio model is to model the changes of PD_i over the specified time horizon by taking into consideration systematic influences due to country or business dependencies as well as idiosyncratic changes. For our example we use the CreditRisk⁺ model, which is available via the GCPM R-package on CRAN. We give a short introduction on those model parts, which are necessary for this example. For detailed information please refer to Credit Suisse First Boston International (1997) or Gundlach and Lehrbass (2004).

Within the CreditRisk⁺ model the possible change of counterparty i 's PD is modeled by the conditional \overline{PD}_i , which is given by

$$\overline{PD}_i = PD_i \left(w_{i,0} + \sum_{k=1}^K w_{i,k} S_k \right),$$

where $w_{i,k} \in [0, 1]$ represent the affiliation to one or multiple out of K sectors (business-country combinations). In contrast, $w_{i,0} = 1 - \sum_{k=1}^K w_{i,k} \geq 0$ represent the idiosyncratic component. The variables $S_k > 0$ (so-called sector variables) represent the economic situation of sector $k = 1, \dots, K$. An economic boom is associated with values $S_k < 1$, which yields to a conditional PD below the unconditional one (i.e. $\overline{PD}_i < PD_i$), whereas a recession is expressed by values $S_k > 1$, which increases counterparties' PD. Within the original framework of Credit Suisse First Boston International (1997), the sector variables are modeled via a gamma distribution, which insures that (together with some additional assumptions) the portfolio loss distribution can be calculated analytically³.

As mentioned above, the gamma distribution was chosen for performance issues. From an economic point of view, this assumption may be questionable. Instead of a gamma distribution, one can also use a lognormal distribution, which is more heavy tailed and therefore more conservative compared to a gamma distribution. However, the lognormal distribution also possesses only two parameters. Therefore, using mean and variance⁴ to

parametrize the distribution of S_k the heaviness of the tail can not be controlled explicitly. If we use a pTAS distribution instead, we can fit the sector distribution to mean and variance of observed PD changes and still have control over the tail via parameter α .

For our example, we use portfolio of 5,000 counterparties belonging to ten different sectors. The data used to estimate the sector distributions are monthly PD values over 10 years, which are estimated via a Merton type model (see Merton (1974)) from marked data (i.e. stock prices and liabilities) for over 20,000 corporations and aggregated on sector level. The portfolio as well as the underlying data are explained in more detail in Fischer and Jakob (2015) and Jakob and Fischer (2014). We estimate pTAS distributions for each sector by using MLE and the empirically observed skewness and compare the results on portfolio risk figures with the original setting (i.e. gamma distributions parametrized via mean and variance) and a framework using the lognormal distribution and the Weibull distribution. Because of the relatively small number of observations (10 years of monthly data), we do not use the kurtosis or quantile estimation methods in this example. For the Monte Carlo simulation of the CreditRisk⁺ model the GCPM R-package is used. The dependency between sector variables are modeled via a Student-t copula with 3.8 degrees of freedom, estimated using a maximum likelihood approach⁵.

Table 2 shows the empirical skewness of each sector as well as the estimated parameter α and the skewness of a Gamma, a lognormal and a Weibull distribution parametrized only by mean and variance. As already indicated by Fig. 1, the lognormal distribution possesses a skewness and kurtosis similar to a pTAS distribution with higher values of α . By looking at Table 2 we can see, that the observed skewness is often underestimated by the lognormal distribution, which is also confirmed by Fig. 4, where the risk figures of the lognormal framework are slightly below those of the pTAS(skewness) framework. However, this may only be the case if skewness is in a suitable range. By using a pTAS distribution one always has greater flexibility to account for semi-heavy (or not so heavy) tailed distributions compared to a lognormal or a gamma distribution. In addition, Fig. 5 shows the empirical observations of the sector variables S_k together with the densities of the fitted distributions exemplarily for Sector 4. The Figure shows that a pTAS distribution estimated via a MLE approach fits the data much better compared to all other distributions. The pTAS distribution estimated via a MLE approach also possesses the heaviest tail of all presented competitors, which in turn causes significant higher risk figures as shown by Fig. 4.

The VaR is considerably higher

Table 2 Skewness values and estimated parameters for different sector distributions

		Sector									
		S_1	S_2	S_3	S_4	S_5	S_6	S_7	S_8	S_9	S_{10}
Skewness	Empirical	2.59	1.74	2.24	2.10	1.74	2.46	2.44	2.16	1.84	2.57
	Weibull	0.98	0.52	0.84	0.63	0.37	0.93	1.05	0.70	0.62	1.54
	Gamma	1.29	0.97	1.19	1.04	0.87	1.25	1.34	1.10	1.04	1.69
	lognormal	2.21	1.56	2.00	1.71	1.38	2.13	2.31	1.81	1.69	3.13
	pTAS (skewness)	2.59	1.74	2.24	2.10	1.74	2.46	2.44	2.16	1.84	2.57
	pTAS (MLE)	4.57	2.92	3.68	3.73	2.74	4.73	5.62	3.43	3.43	5.68
α	pTAS (skewness)	0.67	0.62	0.64	0.67	0.67	0.66	0.62	0.66	0.61	0.51
	pTAS (MLE)	0.84	0.80	0.81	0.84	0.81	0.85	0.86	0.81	0.82	0.83

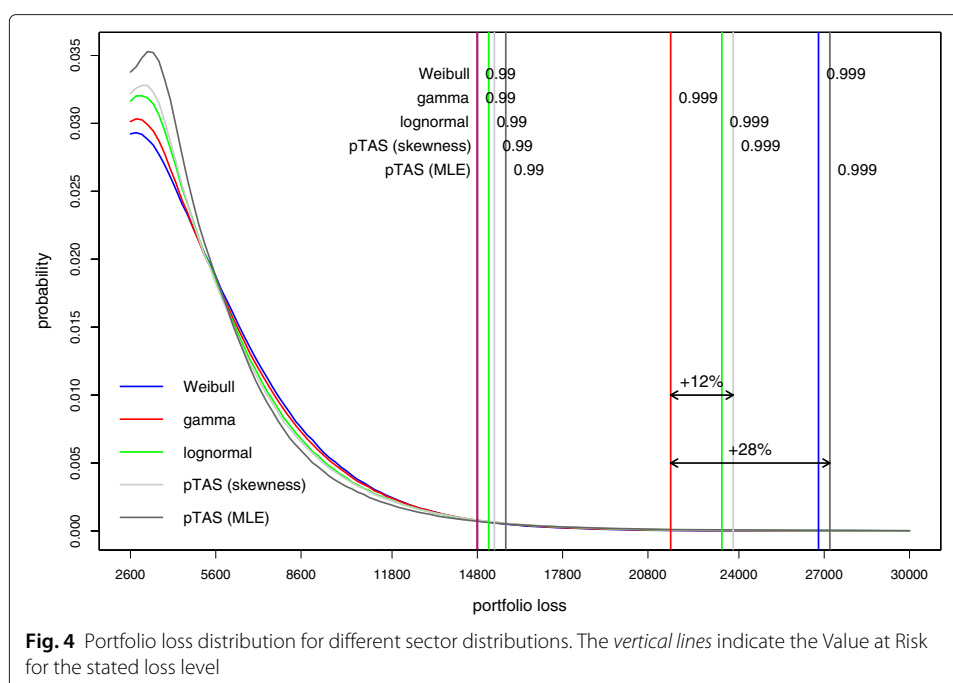
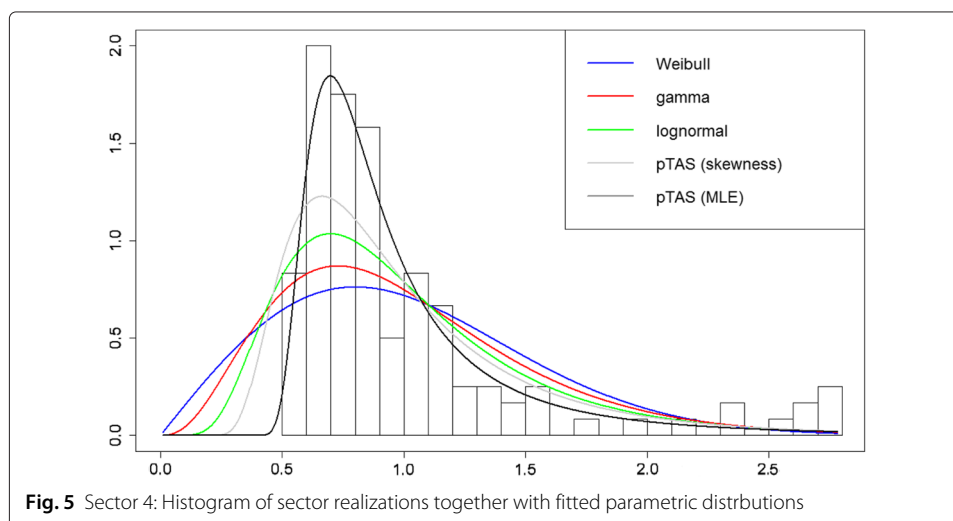


Figure 4 shows the pdf of the portfolio loss distribution together with vertical lines indicating the Value at Risk (VaR) for level τ , which is simply the quantile of the portfolio loss distribution. The VaR is considerably higher if we use a pTAS distribution which considers the skewness or which is estimated via MLE for the sector distribution compared to the standard case of a gamma distribution, which only accounts for mean and variance. Because financial institutions typically use higher values for τ (e.g. $\tau = 0.999$) to calculate the economic capital which is necessary to cover unexpected losses, the use of a simple gamma distribution may imply a significant amount of model risk. In our case, the $\text{VaR}_{0.999}$ rises by around 12 % if we use pTAS distributions based on skewness and up to 28 % if we use MLE. Please note, that the effect in general depends on the sectors (i.e.



business lines and countries) of the portfolio as well as the data used for estimating the sector distributions.

3.2 Operational risk

A financial institution is exposed to different risks such as credit risk and market risk, for instance, and is required to put aside a capital buffer against unexpected losses. With the implementation of Basel II recommendations a capital requirement for operational risk (OpRisk) was set under regulation, too. There are many definitions of operational risk and many institutions have adopted their own definitions which better reflects their business model and strategy. The Basel Committee defines operational risk *as the risk of loss resulting from inadequate or failed internal processes, people and systems or from external events*. It becomes obvious that OpRisk is a very broad concept and may include anything from bank robberies, unauthorized trading, failures in the internal systems to terrorist attacks and natural catastrophies. Commonly, the loss data are organized according to seven official Basel II defined event types (Internal Fraud, External Fraud, Employment Practices & Workplace Safety, Clients, Products & Business Practice, Damage to Physical Assets, Business Disruption & System Failures, Execution, Delivery & Process Management) and eight business lines (Corporate Finance, Trading & Sales, Retail Banking, Commercial Banking, Payment & Settlement, Agency Services, Asset Management, Retail Brokerage) which implies 56 categories. In order to derive the OpRisk loss distribution, one typically derives the loss distribution for each relevant category in the first step⁶ and aggregates the individual loss distributions to the overall loss distribution, in a second step. For reason of simplicity, we assume in the sequel that the bank operates on one business line and is faced to one event type only. In this case, the loss distribution reduces to a compound sum of the type

$$L = L_1 + \dots + L_N$$

with random variables N (number of loss in the next year) and corresponding loss severities L_1, \dots, L_N for each loss. For reasons of simplicity we assume that N and L_1, \dots, L_N are mutually independent (see, e.g. Neslehova et al. (2006) for an inclusion of dependence). Once these variables are specified and estimated within a parametric setting and for a given loss data set, the loss distribution can be easily simulated within a Monte Carlo framework. Typically, the number of losses are assumed to follow a Poisson distribution with parameter $\lambda > 0$ which corresponds to the expected number of losses within the next period/year. In contrast, specifying the severity distribution is more complex to some extent. To avoid the problems in fitting a single parametric distribution with one or two parameters to the whole data set, one often uses a compound severity distribution, see El Adlouni et al. (2011). This involves dividing the severity distribution into the so-called body and the so-called tail by a threshold, and presuming a different distribution for each part. Both distributions are then combined into a single severity distribution which is commonly termed as a compound distribution. Losses below and above the threshold are referred to as low-severity and high-severity losses. Once, the parametric model(s) for the severity distribution has been selected, the estimation of unknown parameters can be done using the method of moments (MM), Maximum Likelihood (ML, used in the empirical part) or Ordinary Least Square (OLS). In order to compare the goodness-of-fit (GoF), graphical tools like quantile-quantile-plots or suitable GoF-test should be used.

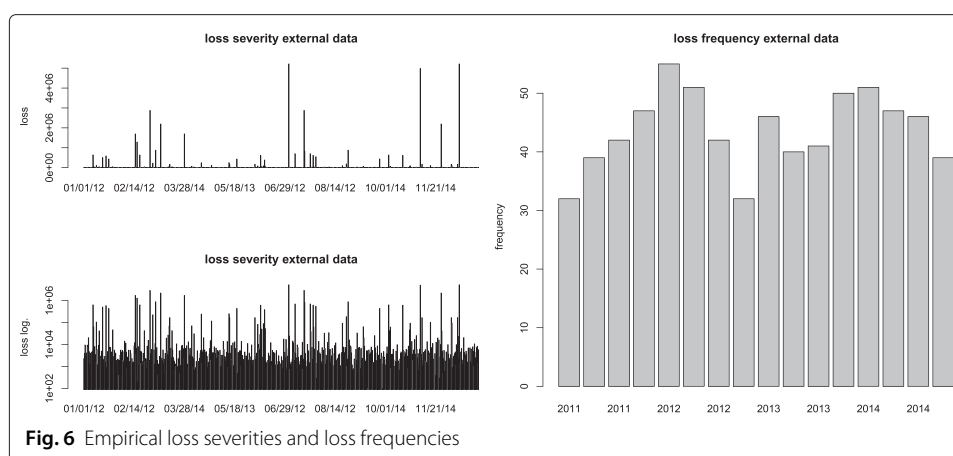


Fig. 6 Empirical loss severities and loss frequencies

The underlying data set is provided in the textbook of Bolancè et al. (2012), chapter 7. It contains 700 loss data from 2011 to 2014 which were multiplied by 1000 for technical reasons. The corresponding loss severities and loss frequencies in the course of time are depicted in Fig. 6.

In order to illustrate the OpRisk calculation, we assume that the number of losses is Poisson distributed and that we have a compound severity distribution, where the body is modelled by the empirical distribution function, whereas the tail is assumed to follow a lognormal, a Weibull, a pTAS, a gamma and a generalized gamma (see, e.g. Stacy (1962)) distribution, respectively. Table 3 summarizes the corresponding OpVaR (i.e. quantile of the loss distribution) for different confidence levels and each of the four distributions. Obviously, the Weibull (which is often applied in banking industry) and the pTAS tail model are close together, whereas the lognormal model implies significant higher OpVaR's. In contrast, both gamma and generalized gamma distribution produce lower OpVaR figures.

In order to compare the goodness-of-fit, both quantile-quantile plots (see Fig. 7) and goodness-of-fit statistics (Kolmogorov-Smirnov and Anderson-Darling, see Table 4 and Chernobay et al. (2015) for a detailed discussion) indicate that the pTAS distribution with $\hat{\alpha} = 0.3842$ seems to be the preferable choice. Although the generalized gamma distribution provides reasonable AD2UP and KS statistics, too, the user is in danger of being too progressive, taking into account that there are only a few number of observations and that the true parametric model is unknown⁷.

Finally, Fig. 7 also depicts the simulated portfolio loss for the OpRisk data set.

4 Conclusion

Within this article, we discussed the family of positive tempered α -stable distributions, which is a flexible distribution family and well suited to model both light and semi-heavy

Table 3 OpVaR for different confidences level and distributions

Percentile →	0.95	0.99	0.995	0.999	0.9995
lognormal	40,937,057	76,509,385	100,584,286	187,533,641	247,986,179
Weibull	37,542,382	47,459,865	51,335,798	60,440,362	64,322,549
pTAS	33,554,021	44,883,975	49,657,499	60,450,462	65,182,740
gamma	32,023,064	41,335,245	45,109,838	53,586,965	57,257,060
gen.gamma	24,836,520	31,636,053	34,333,082	40,596,749	43,033,552

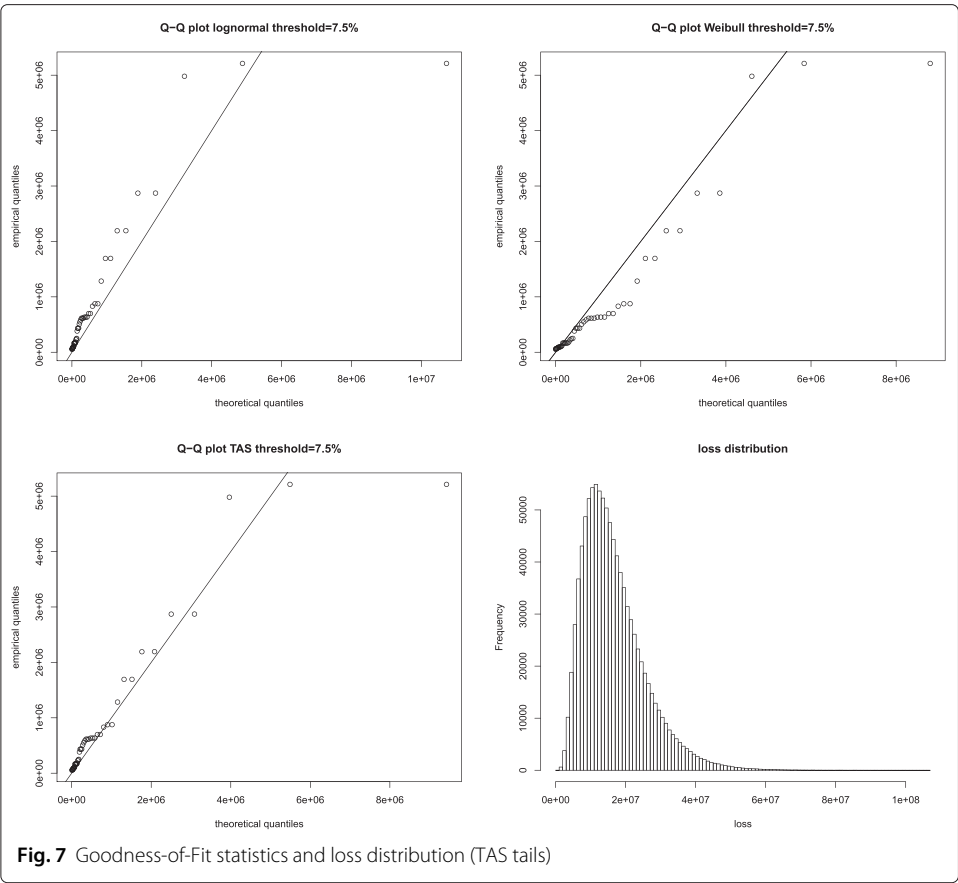


Fig. 7 Goodness-of-Fit statistics and loss distribution (TAS tails)

tailed data on the positive half-axis. Besides the derivation of the family and a summary of certain characteristics, which are relevant for practical applications, we provided an overview of the existing literature on this family. Furthermore, we introduced algorithms that can be used to implement the basic functionalities of the pTAS family (i.e. density and distribution function, calculation of quantiles and random numbers generation), which are also available via the pTAS R-package. By applying the pTAS distribution in the field of credit and operational risk we show that this distribution is more flexible and provides a better fit to empirical data compared to other competitors which are often used. Therefore, as in case of credit risk, the pTAS distribution can help to reduce the amount of model risk or as in case of operational risk, to reflect the risk more adequately which in turn helps banks to allocate economic capital more appropriately. Beyond the two given examples, the pTAS distribution may be also used to model (stock) returns and therefore being beneficial for market risk management as well.

Table 4 Goodness-of-fit statistics

Test statistic	AD2UP	<i>p</i> -value	KS	<i>p</i> -value
lognormal	2.6582	0.080	0.8235	0.060
Weibull	7.8001	0.024	1.7422	0.005
pTAS	2.7823	0.682	0.6585	0.739
gamma	3.1287	0.077	0.8612	0.197
gen.gamma	2.6755	0.167	0.5674	0.697

Endnotes

¹Please refer to Schiff (1999, Theorem 4.3) for further details.

²Besides counterparties' defaults also changes in a counterparties creditworthiness (so-called rating migrations) cause a portfolio loss. However, for reasons of simplicity we restrict this example to default risk only.

³In 1997, when the CreditRisk⁺ was published, this was a major advantage over simulative models, which use a Monte Carlo simulation to estimate the loss distribution. However, nowadays computers are much more powerful and Monte Carlo simulations are widely used.

⁴The parametrization based on mean and variance is a standard method within the CreditRisk⁺ model. In order to ensure that $\mathbb{E}(\overline{PD}_i) = PD_i$, which is a common assumption, we have the condition that $\mathbb{E}(S_k) = 1$ for all k .

⁵For readers interested in the topic of copulas within credit portfolio models, we refer to Jakob and Fischer (2014) and Fischer and Jakob (2015).

⁶The Basel Committee have prescribed guidance for three types of methods for the calculation of capital requirement for operational risk. Those are the Basic Indicator Approach, the Standard Approach and the Advanced Measurement Approach. The latter is the most sophisticated of the approaches and this is what this section is about.

⁷Please note that for the GoF-test for left-truncated samples the distribution of the statistic is not parameter-free, and the p -values and the critical values are obtained by means of Monte Carlo simulation. There remains a certain variation of the p -values - a direct comparison is not meaningful.

Appendix

Using the pTAS R-package

With the help of short examples we explain how the functions can be used within an R-session. We will not describe the several functions in all details and parameters. Please refer to the corresponding help pages within the package for more information.

Creating a pTAS distribution

The pTAS package uses an object oriented approach, which means that every distribution (e.g. with different parameters) is represented by a different object of class pTAS. Therefore, one can work with many different distributions at the same time (within the same workspace) without jeopardizing their consistency regarding distributional or numerical parameters.

Hence, the first step is to create a new pTAS distribution object.

```
library(pTAS)
## Loading required package: stats4
# creating pTAS object with parameters according to Palmer(2008)
MYpTAS = pTAS(param.P = c(0.5, 1, 0.75))
```

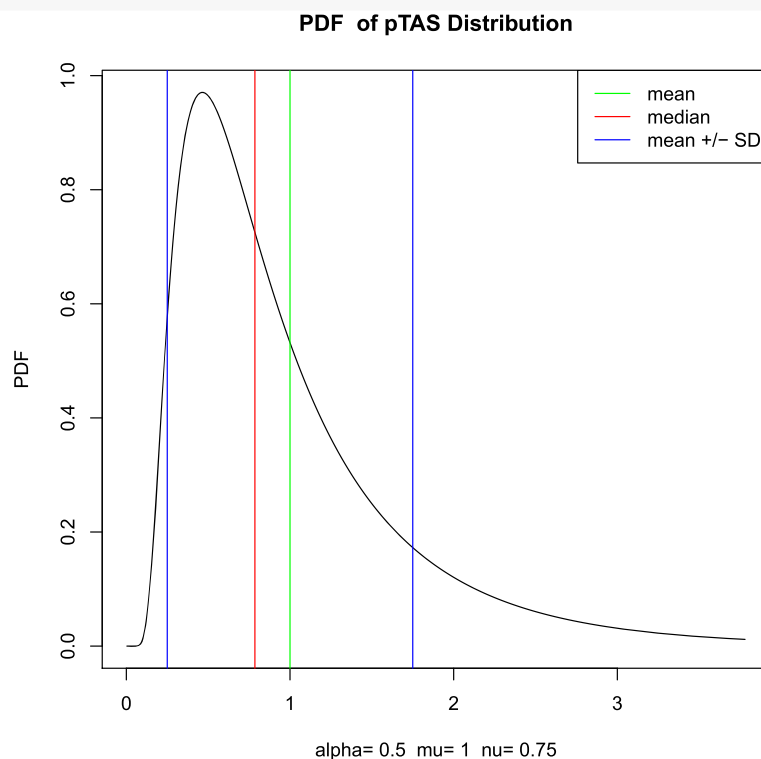
The MYpTAS object describes a pTAS distribution with parameterization $\mathcal{P}_P(\alpha, \gamma, \theta) = (0.5, 1, 1.5)$. The pTAS function automatically performs some plausibility checks on the given parameters and translates the given parameterization (i.e. according to Palmer et al.

(2008) in this case) to the other ones. In addition, distribution figures such as mean, variance, skewness and kurtosis are calculated.

```
# get parametrization according to Hougaard (1986)
param.H(MYpTAS)
##      alpha      delta      theta
## 0.5000000 0.9428090 0.8888889
# get mean and skewness
mean.pTAS(MYpTAS)
## [1] 1
skew.pTAS(MYpTAS)
## [1] 2.25
```

To obtain a first impression of the distribution, the density (and also the distribution function) can be shown by simply using the PLOT function.

```
# plot pdf (use additional argument cdf=TRUE to plot cdf)
plot(MYpTAS)
```



Density, distribution function, quantiles and random numbers

Density and distribution function are available via the standard R-notation (d.../ p...). The numerical parameters described in 2.1 can be set when the object is created via the pTAS function.

```
# calculate pdf and cdf
t = seq(0.5, 2.5, 0.5)
dpTAS(t, MYpTAS)
## [1] 0.96465947 0.53192304 0.24967324 0.12058243 0.06046488
ppTAS(t, MYpTAS)
## [1] 0.2547667 0.6340912 0.8205896 0.9089881 0.9524083
```

Quantiles and random numbers are generated via Algorithm 2.

```
# calculate quantiles
p = c(0.1, 0.25, 0.5, 0.75, 0.9)
qpTAS(p, MYpTAS)
## [1] 0.3355067 0.4950626 0.7855115 1.2644107 1.9291478
# draw random numbers
set.seed(1) # for reproducibility
rpTAS(5, MYpTAS)
## [1] 1.9935582 0.8937332 0.5111581 0.6264167 0.4452205
```

Estimation methods

For implementation issues, the MLE will be performed on the \mathcal{P}_p - parameterization always. For parameters α, μ, ν lower and upper bounds as well as fixed values can be specified. Especially for parameter α an appropriate upper bound may be helpful, because pdf calculations for parameters α close to 1 are numerically challenging. The `FIT_MLE` function uses the MLE function from the `STATS4` package, which in turn uses R's `OPTIM` optimizer.

```
# draw some random numbers
x = rpTAS(200, MYpTAS)
# perform MLE
MLE_pTAS = fit_mle(x)
# estimation results
param.P(MLE_pTAS)
##      alpha      mu      nu
## 0.5489056 0.9729183 0.6668250
# compared to original values
param.P(MYpTAS)
## alpha      mu      nu
## 0.50      1.00      0.75
```

Detailed information on the optimization results can be obtained via the `OPTIM_RES` function. This gives a list containing the number of iterations, the convergence result and the Hessian matrix for further calculations (e.g. to calculate confidence intervals). If the optimization did not converge properly, a warning is displayed automatically.

```
optim_res(MLE_pTAS)
## $par
##      alpha      mu      nu
## 0.5489056 0.9729183 0.6668250
##
## $value
## [1] 136.1362
##
## $counts
## function gradient
##      23      23
##
## $convergence
## [1] 0
##
## $message
## [1] "CONVERGENCE: REL_REDUCTION_OF_F <= FACTR*EPSMCH"
```



```
##
## $hessian
##      alpha      mu      nu
## alpha  567.3590  198.1225 -475.1069
## mu     198.1225  628.2451 -367.0736
## nu     -475.1069 -367.0736  880.2828
```

Alternatively, distribution parameters can be also estimated based on mean and variance, which determine parameters μ and ν regarding the \mathcal{P}_P parametrization and either the skewness, the kurtosis or a quantile to estimate parameter α . If α should be estimated based on a given kurtosis value, which possible has multiple solutions, a pTAS distribution with either the highest or the lowest value of possible α 's or a list containing all distributions will be returned.

If α should be estimated based on a given quantile, an grid search method as described in section 2.3 is applied. For given values $t^* > 0$ and $p^* \in (0, 1)$ the algorithm terminates if $|F_{\mathcal{P}_P(\alpha, \mu, \nu)}(t^*) - p^*| < \epsilon$ for a prespecified $\epsilon > 0$, whereas values for μ and ν are fixed based on mean and variance. Similar to the estimation based on kurtosis, the solution for α (if one exists) may be not unique. Again, one can determine which distribution should be returned with an additional argument (see example below).

```
# estimation based on skewness/kurtosis
skew_pTAS = fit_moments(mean = 1, var = 0.5, skew = 3)
kurt_pTAS = fit_moments(mean = 1, var = 0.5, kurt = 10)
## alpha1: 0.4757636 alpha2: 1.293467
# estimation based on quantile
quant_pTAS_1 = fit_quantile(mean = 1, var = 1, x = 5, prob = 0.99)
## Warning in fit_quantile(mean = 1, var = 1, x = 5, prob =
0.99):
Solution for alpha is not unique, choosing solution with
highest alpha.
param.P(quant_pTAS_1)
##      alpha      mu      nu
## 0.8898322 1.0000000 1.0000000
quant_pTAS_2 = fit_quantile(mean = 1, var = 1, x = 5, prob = 0.99,
which.alpha = "lowest") # return lower alpha
## Warning in fit_quantile(mean = 1, var = 1, x = 5, prob =
0.99,...):
Solution for alpha is not unique, choosing solution with
lowest alpha.
param.P(quant_pTAS_2)
##      alpha      mu      nu
## 0.5147693 1.0000000 1.0000000
```

Acknowledgement

The authors thank two anonymous referees and an associate editor for their helpful comments and suggestions which significantly improved the presentation of the paper.

Authors' contributions

MF carried out the derivation and general properties of the pTAS family as well as the application to operational risk. KJ addressed issues regarding the implementation and the credit risk example. The manuscript was drafted together. All authors read and approved the final manuscript.

Competing interests

The authors declare that they have no competing interests.

Author details

¹Department of Statistics & Econometrics, FAU Erlangen-Nürnberg, Erlangen, Germany. ²Department of Statistics, University of Augsburg, Augsburg, Germany.

Received: 2 October 2015 Accepted: 16 June 2016

Published online: 18 July 2016

References

- Abate, J, Choudhury, GL, Whitt, W: An introduction to numerical transform inversion and its application to probability models. In: *Computational Probability*, pp. 257–323. Springer Science+Business Media, New York (2000)
- Barndorff-Nielsen, OE, Shephard, N: Normal modified stable processes. *Economics papers 2001-w6*. University of Oxford, Oxford (2001)
- Bolancé, C, Guillén, M, Gustafsson, J, Nielsen, JP: *Quantitative Models for Operational Risk: Extremes, Dependence and Aggregation*. Chapman & Hall/CRC Finance Series, New York (2012)
- Chernobay, A, Rachev, S, Fabozzi, F: Composites goodness-of-fit tests for left-truncated loss samples. In: Lee, CF, Lee, J (eds.) *Handbook of Financial Econometrics and Statistics*, pp. 575–596. Springer, (2015)
- Credit Suisse First Boston International: *CreditRisk+ A Credit Risk Management Framework* (1997). <http://www.csfb.com/institutional/research/assets/creditrisk.pdf>. Accessed 27 Jan 2014
- El Adlouni, S, Ezzahid, E, Moutazzim, J: Mixed distributions for loss severity modelling with zeros in the operational risk losses. *Int. J. Appl. Math. Stat.* **11**(21), 96–109 (2011)
- Esscher, F: On the probability function in the collective theory of risk. *Scand. Actuarial J.* **1932**(3), 175–195 (1932). doi:10.1080/03461238.1932.10405883
- Fischer, M, Jakob, K: Copula-Specific Credit Portfolio Modeling. In: Glau, K, Scherer, M, Zagst, R (eds.) *Innovations in Quantitative Risk Management*. Springer Proceedings in Mathematics & Statistics, pp. 129–145. Springer International Publishing, (2015)
- Gundlach, M, Lehrbass, F: *CreditRisk+ in the Banking Industry*. Springer Finance. Springer, Berlin Heidelberg (2004)
- Haas, M, Pigorsch, C: Financial economics: Fat-tailed distributions. In: *Encyclopedia of Complexity and Systems Science* vol. 4, pp. 3404–3435. Springer, New York (2009)
- Hougaard, P: Survival models for heterogeneous populations derived from stable distributions. *Biometrika.* **73**(2), 387–396 (1986)
- Jakob, K, Fischer, M: Quantifying the impact of different copulas in a generalized CreditRisk+ framework An empirical study. *Depend. Model.* **2**, 1–21 (2014)
- Janson, S: Stable distributions. arXiv preprint arXiv:1112.0220 (2011). Accessed 2015-08-16
- Jørgensen, B: Exponential dispersion models (with discussion). *J. R. Stat. Soc. Series B.* **49**(2), 127–162 (1987)
- Koudou, AE, Ley, C: Characterizations of gig laws: A survey. *Probab. Surv.* **11**, 161–176 (2014)
- Küchler, U, Tappe, S: *Tempered stable distributions and applications to financial mathematics* (2011)
- Merton, RC: On the pricing of corporate debt: The risk structure of interest rates. *J. Finance.* **29**(2), 449–470 (1974)
- Neslehova, J, Embrechts, P, Chavez-Demoulin, V: Infinite mean models and the LDA for operational risk. *Journal of Operational Risk.* **1**(1), 3–25 (2006)
- Nolan, J: *Stable Distributions: Models for Heavy-tailed Data*. Birkhauser, Boston (2003). In progress available via: <http://fs2.american.edu/jpnolan/www/stable/chap1.pdf>
- Palmer, KJ, Ridout, MS, Morgan, BJ: Modelling cell generation times by using the tempered stable distribution. *J. R. Stat. Soc. Series C (Appl. Stat.)* **57**(4), 379–397 (2008)
- Ridout, MS: Generating random numbers from a distribution specified by its Laplace transform. *Stat. Comput.* **19**(4), 439–450 (2008). doi:10.1007/s11222-008-9103-x
- Schiff, JL: *The Laplace Transform: Theory and Applications*. Springer-Verlag, New York (1999)
- Stacy, EW: A generalization of the gamma distribution. *Ann. Math. Stat.* **3**(33), 1187–1192 (1962)
- Tweedie, MCK: An index which distinguishes between some important exponential families. In: *Statistics: Applications and New Directions: Proc, Indian Statistical Institute Golden Jubilee International conference*, pp. 579–604, (1984)