A Simulation Comparison of Quantile Approximation Techniques for Compound Distributions popular in Operational Risk

PJ de Jongh[1], T de Wet [2], K Panman,[3] H Raubenheimer [4]

Abstract

Many banks currently use the Loss Distribution Approach (LDA) for estimating economic and regulatory capital for operational risk under Basel's Advanced Measurement Approach. The LDA requires, amongst others, the modelling of the aggregate loss distribution in each operational risk category (ORC). The aggregate loss distribution is a compound distribution resulting from a random sum of losses, where the losses are distributed according to some severity distribution and the number (of losses) distributed according to some frequency distribution. In order to estimate the economic or regulatory capital in a particular ORC, an extreme quantile of the aggregate loss distribution has to be estimated from the fitted severity and frequency distributions. Since a closed-form expression for the quantiles of the resulting estimated compound distribution does not exist, the quantile is usually approximated by using brute force Monte Carlo simulation which is computationally intensive. However, a number of numerical approximation techniques have been proposed to lessen the computational burden. Such techniques include Panjer recursion, the fast Fourier transform, and different orders of both the single-loss approximation and perturbative approximation. The objective of this paper is to compare these methods in terms of their practical usefulness and potential applicability in an operational risk context. We find that the second order perturbative approximation, a closed-form approximation, performs very well at the extreme quantiles and over a wide range of distributions and is very easy to implement. This approximation can then be used as an input to the recursive fast Fourier algorithm to gain further improvements at the less extreme quantiles.

Keywords: *Compound distribution, quantile approximation, loss distribution approach, operational risk, single-loss approximation, perturbative approximation.*

1. Introduction

Under Basel III, banks are allowed to use the Advanced Measurement Approach (AMA) to calculate economic or regulatory capital for operational risk. Currently many banks use the Loss Distribution Approach (LDA) when implementing the AMA. The LDA has been studied and researched extensively and although many authors pointed out deficiencies in this approach (see e.g. Cope et al. 2009 and

[1] Professor and Director of the Centre for BMI, NWU, South Africa
[2] Extra-ordinary Professor at the Centre for BMI, NWU, South Africa.
[3] PhD Student at the Centre for BMI, NWU, South Africa.
[4] Associate Professor at the Centre for BMI, NWU, South Africa.

Embrechts and Hofert 2011), the approach remains being used in practice. One of the important building blocks of the LDA is to estimate the aggregate loss distribution in each operational risk category (ORC). In particular, an extreme quantile (0.999 and higher) needs to be estimated to determine the regulatory or economic capital in that ORC. This has been shown to be notoriously difficult and typically the estimates are inaccurate (see e.g. Cope et al. 2009 and Neslehová et al. 2006).

The aggregate loss distribution is a compound distribution of a random sum of losses, where the number of losses follows a certain frequency distribution, usually assumed to be Poisson or Negative Binomial, and the individual losses a severity distribution which can take on a number of forms, depending on the characteristics of the losses observed in a particular ORC. Popular distributions are the extreme value distributions, which include, amongst others, the Pareto, Burr, Nig, Lognormal and LogNig (see e.g. Cope et al., 2009). Because of the extreme nature of operational risk losses, these distributions are often, when fitted to a real data set, of the infinite-mean type. Infinite mean distributions present a number of difficulties when used in practice and is discussed extensively in Neslehová et al. (2006). For instance, infinite mean models can lead to ridiculously large capital estimates and the modeller is warned to handle such cases with care together with good judgement of the observed data. Neslehová et al. (2006) discuss several cases of underestimation and overestimation of high quantiles and is a paper that must be read by operational risk modellers and decision makers.

As stated previously, the aggregate loss distribution is a compound distribution of which the quantiles can typically not be calculated exactly and have to be approximated in some way. In practice, most banks use brute force Monte Carlo (MC) simulation methods to approximate these extreme quantiles. Depending on the accuracy required, these MC simulations are computer intensive and even utilising today's computer power can become impractical when implemented. Of course, variance reduction techniques may be used to lessen the computational burden (see Asmussen and Kroese, 2006). However, a number of numerical approximation techniques have been proposed to overcome this difficulty.

Numerical approximation recursive techniques that can be used to approximate the quantile of the compound distribution include the Panjer recursion (PR) algorithm, introduced by Panjer (1981), and techniques using Fourier inversion and the fast Fourier transform (FFT), proposed by Heckman and Meyers (1983), as well as Feilmeier and Bertram (1987). Grübel and Hermesmeier (1999 have investigated the propagation of discretization errors through compounding and established an improved FFT-based procedure using an exponential change of measure also referred to as exponential tilting. The latter contribution is substantial since it essentially eliminates the so called aliasing error, which is the fundamental deficit that arises through the use of the discrete Fourier transform. Embrechts and Frei (2009) evaluated and compared Panjer recursion and FFT (with exponential tilting) and noted two main advantages of FFT, namely that it works with arbitrary frequency distributions and is much more efficient. They further mention that, in situations where both Panjer recursion and FFT are equally applicable, and where one has to deal with a large number of lattice points, the FFT with suitable tilting is favoured because it is numerically cheaper. The above-mentioned recursive techniques require several input parameters that have to be selected carefully to ensure convergence and are computationally intensive.

Recently, closed-form approximation procedures have been introduced which include the single-loss approximation and the perturbative approximation. Böcker and Klüppelberg (2005) derived a simple single-loss approximation that provides closed-form Value-at-Risk (VaR) estimates for the class of sub-exponential distributions (see e.g. Foss et al., 2011), which, according to Fasen and Klüppelberg (2006), contain most of the popular heavy-tailed distributions in practical use today. This makes the single-loss approximation method very popular amongst practitioners and a number of academics have done research on improving the method. However, the single-loss approximation method only performs accurately if certain conditions hold. Generally, when distributions are heavy-tailed and Poisson intensities are not too high, the approximation works well. Mignola and Ugoccioni (2006) showed that the single-loss approximation method can lead to an underestimation of the true VaR especially for light-tailed loss severity distributions (this is confirmed by our study). Böcker and Sprittulla (2006) constructed a formula that adds a multiple of the mean to the original formula as an improvement and Sahay et al. (2007) tested the single-loss approximation for a Generalised Pareto distribution tail. Degen (2010) provided an analytical framework to analyse the accuracy of the single-loss approximation on the basis of the relative approximation error and derived a second order single-loss approximation for the class of sub-exponential distributions, based on the work of Omey and Willekens (1987). Hess (2011) evaluated the single-loss approximation method for the SAS OpRisk Global data and reported a very good approximation performance of the single-loss approximation method. According to them, the results of this method are even more reliable than the estimates obtained from a 1 million loss Monte Carlo simulation. Hannah and Puza (2015) extend the SLA adding another correction term to account for the effects of the two largest losses (over and above the contribution of the single largest loss). The proposed approximations and extensions are evaluated by way of a simulation study involving a selection of sub-exponential distributions.

Hernandez et al. (2013, 2014) derived closed-form approximations for high percentiles of the aggregate distribution based on a perturbative expansion, which they claim improve on the above-mentioned methods under certain conditions. The perturbative series introduced differs from previous approximations in that the terms in the series are expressed as a function of the moments of the right truncated distribution for the individual random variables in the sum. These censored moments exist even when the moments of the original distribution (without truncation) diverge. Consequently the same expression is valid for both the infinite and finite mean cases. For high percentiles the perturbative expansion provides a sequence of approximations that, up to certain order, has increasing quality as more terms are included. Beyond that order the convergence of the series deteriorates. Peters et al. (2013) provide a tutorial overview of heavy-tailed loss process modelling in operational risk under Basel III, with discussion on the implications of such tail assumptions for the severity model in a LDA structure. In particular, they consider three key families of risk measures and their equivalent second order asymptotic approximations: Value-at-Risk; Expected Shortfall (ES) and the Spectral Risk Measure. As a last note one should keep in mind that the closed-form approximation methods are computationally very efficient when compared to the recursive methods. A practical limitation of the closed-form approximation methods is that, while these methods are very accurate in estimating very high quantiles of a compound distribution, they cannot be used to approximate across all quantiles, and thus, to derive the full CDF of a compound distribution. The full CDF of a compound distribution is required if an institution considers modeling diversification benefit between different operational risk categories (units of measure) or between different risk types (credit, market, operational). Unlike closed-form approximation methods, both the Monte Carlo and recursive methods can be used to derive the full CDF of a compound distribution.

In this paper we conduct a Monte Carlo simulation study to investigate the performance of the closed-form approximation methods using the Poisson as frequency distribution and the Burr, Lognormal and LogNig as severity distributions. In order to limit the scope of the study, we decided to only focus on the Poisson distribution, because of its popularity in practice and the fact that the estimated VaR depends critically on the choice of severity distribution rather than the choice of frequency distribution (see Cope et al. 2009). In the same vein, we will only consider one risk measure namely VaR since it is still the most popular risk measure in practice. As far as severity distributions are concerned, the Lognormal and Burr distributions, together with mixtures of them with the Generalised Pareto distribution, are amongst the most popular models used in practice for modelling single-loss severities (see e.g. de Jongh et al. 2015 and Hess 2011). Another popular distribution in financial applications is the normal inverse Gaussian (NIG) distribution (see e.g. Venter and de Jongh, 2002) having semi-heavy tails. Because loss data in operational risk are typically heavy-tailed we decided to include the LogNig, a heavy-tailed distribution obtained by ‘exponentiating’ the NIG in the same way as the Lognormal is obtained by ‘exponentiating’ the normal distribution. A possible alternative is the Logphase distribution (see Ahn et al., 2012) that has the disadvantage that the distribution function is difficult to calculate numerically, which is not the case for the LogNig. From an Extreme Value Theory (EVT) point of view, a further motivation for using the LogNig is that this distribution provides an alternative to the Burr because its second order parameter $\rho = 0$ whereas the Burr has second order parameter $\rho < 0$. Of course many other choices of severity distributions are possible (e.g. the g-and-h distribution, see Degen et al., 2007), but we decided to limit the scope of our study. Due to their simplicity and ease of calculation, our focus will be on the closed-form approximations, but a comparison of one of the recursive methods with the best performing closed-form approximation method is included.

The paper is organised as follows. In Section 2 the Monte Carlo and closed-form numerical approximation techniques are described. Then, in Section 3, the simulation study is described and in Section 4 the results are provided, including a comparison with the FFT. Some concluding remarks and directions for future research are given in the last section.

2. Approximation techniques

In this section a brief overview is given of the standard Monte Carlo approximation, the single-loss approximation and the perturbative approximation techniques. Before we discuss these methods in detail let us state the problem in theoretical terms. Assume that random variables $X_1, \dots, X_N$ are independent and identically distributed according to some distribution G and that N follows the Poisson distribution with parameter λ. We know that if $N \sim Pois(\lambda)$ and $X_1, \dots, X_N \sim G$, then $S = \sum_{n=1}^{N} X_n \sim CoP(G, \lambda)$ where $CoP(G, \lambda)$ is the compound Poisson distribution with parameters λ and G. We are interested in obtaining the $100(1-\gamma)\%$ quantile of this distribution and since no closed-form solution exists, approximation methods are used. In operational risk terminology, the $X's$ represent the operational risk losses, G the severity distribution which has to be estimated from data from various sources such as internal bank data, external consortium data (e.g. the ORX 2008 data base) and scenario data, and $CoP(G, \lambda)$ the aggregate loss distribution. The $100(1-\gamma)\%$ quantile of the latter distribution is the so-called VaR and for calculating regulatory capital (RC) one is particularly interested in the 99.9% VaR, i.e. where $\gamma = 0.001$. In each ORC, the severity and frequency distributions of the annual losses are estimated and then convoluted by means of the random sums method in order to obtain an estimated aggregate loss distribution. The 99.9% VaR of the aggregate loss distribution is then used as estimate of the RC.

Monte Carlo approximation

As before, let the random variable N denote the annual number of loss events and the random variables $X_1, \dots, X_N$ denote the loss severities of these loss events. Then the annual aggregate loss is $S = \sum_{n=1}^{N} X_n$ and the distribution of S is the aggregate loss distribution of the ORC. This aggregate loss distribution and its VaR are difficult to calculate analytically and Monte Carlo (MC) simulation is often used to approximate it as follows:

i. Generate N distributed according to the assumed frequency distribution;
ii. Generate $X_1, \dots, X_N$ independent and identically distributed according to the severity distribution G and calculate $S = \sum_{n=1}^{N} X_n$;
iii. Repeat (i) and (ii) I times independently to obtain $S_i, i = 1,2, \dots, I$ and approximate the 99.9% VaR as $S_{([0.999*I]+1)}$ where $S_{(i)}$ denotes the i-th order statistic and $[k]$ the largest integer contained in k.

Note that three input items are required to perform it, namely the number of MC repetitions I as well as the frequency and loss severity distributions. The number of MC repetitions determines the accuracy of the approximation and the larger it is, the higher its accuracy. In principle infinitely many repetitions are required to get the exact true VaR. We always use one million repetitions in our simulations; while this may sound high, it should be kept in mind that a very high quantile (99.9%) is calculated which forces a large number of repetitions in order to gain accuracy. The Monte Carlo approximation method will subsequently be referred to as MC.

Single-loss approximation

A density function f belongs to the class of sub-exponential densities if it satisfies $\lim_{x\to\infty} f(x+y)/f(x) = 1$ for all $y \in R$ and $\lim_{x\to\infty} f^{*n}(x)/f(x) = n$ for all $n \geq 2$ (see e.g. Degen, 2010). Note that the sub-exponential class includes all regularly varying densities. A positive measurable function f is regularly varying with parameter β ($f \in RV_\beta$), if $\lim_{t\to\infty} f(tx)/f(t) = x^\beta$ for all $x > 0$. In the case of a probability density f by Karamata's Theorem, if $f \in RV_{-\frac{1}{\kappa}-1}$ then $\bar{F} \in RV_{-1/\kappa}$, with κ the extreme value or tail index of the distribution (see e.g. Embrechts et al. 1997) and $\bar{F}(x) = 1 - F(x)$.

Böcker and Klüppelberg (2005) derived a simple single-loss approximation that provides closed-form VaR estimates for the class of sub-exponential distributions. If G is the true underlying severity distribution function of the individual losses and λ the true annual frequency then the $100(1-\gamma)\%$ VaR of the compound loss distribution may be approximated by

$$G^{-1}(1-\gamma/\lambda). \tag{1}$$

This will be subsequently referred to as SLA.

For heavy-tailed distributions, Degen (2010), using second order exponentiality, derived an improved single-loss approximation. Assuming that $N \sim Pois(\lambda)$, G *has a finite mean* $(E(X) < \infty)$, and is regularly varying $\bar{G} \in RV_{-\frac{1}{\kappa}}$ with $\kappa > 0$, he showed that the $100(1-\gamma)\%$ VaR of the compound loss distribution may be approximated by

$$G^{-1}\left(1-\frac{\gamma}{\lambda}\right)+\lambda\mu, \tag{2}$$

where $\mu = E(X)$ is the finite mean of G.

For *infinite mean models* ($\mu = E(X) = \infty$) and $\bar{G} \in RV_{-\frac{1}{\kappa}}$ for $1 < \kappa < \infty$ the $100(1-\gamma)\%$ VaR of the compound loss distribution may be approximated by

$$G^{-1}\left(1-\frac{\gamma}{\lambda}\right)+\gamma G^{-1}\left(1-\frac{\gamma}{\lambda}\right)\frac{C_\kappa}{1-\frac{1}{\kappa}} \tag{3}$$

where $C_\kappa = (1-\kappa)\frac{\Gamma^2\left(1-\frac{1}{\kappa}\right)}{2\Gamma\left(1-\frac{2}{\kappa}\right)}$ and Γ the gamma function. For the special case $\kappa = 1$, the $100(1-\gamma)\%$ VaR of the compound loss distribution may be approximated by

$$G^{-1}\left(1-\frac{\gamma}{\lambda}\right)+\lambda\mu_G\left(G^{-1}\left(1-\frac{\gamma}{\lambda}\right)\right), \tag{4}$$

with $\mu_G(x) = \int_0^x \bar{G}(s)ds$. Note that when $\kappa = 2$ the ratio $\frac{\Gamma^2\left(1-\frac{1}{\kappa}\right)}{2\Gamma\left(1-\frac{2}{\kappa}\right)} = 0$.

For *semi-heavy-tailed* distributions with $\kappa = 0$, Degen showed that the $100(1-\gamma)\%$ VaR of the compound loss distribution may be approximated by (2). The approximation suggested by Degen will be referred to as SLAD.

Hannah and Puza (2015) extends the SLA adding another correction term to account for the effects of the two largest losses. Assuming that $N \sim Pois(\lambda)$, G *has a finite mean* $(E(X) < \infty)$, they showed that the $100(1-\gamma)\%$ VaR may be approximated by solving for c in

$$\gamma = \lambda\big(1-G(c-\lambda\mu)\big)+\frac{1}{2}\lambda^2\left(1-G\left(\frac{c-\lambda\mu}{2}\right)\right)^2. \tag{5}$$

Although this solution does not exist in closed-form, Hannah and Puza (2015) propose a solution by setting $c^{(0)} = G^{-1}\left(1-\frac{\gamma}{\lambda}\right)+\lambda\mu$ as an initial guess, and iterating according to

$$c^{(j)} = G^{-1}\left(1-\frac{1}{\lambda}\left[\gamma-\frac{1}{2}\lambda^2\left(1-G\left(\frac{c^{(j-1)}-\lambda\mu}{2}\right)\right)^2\right]\right)+\lambda\mu\,,\quad j = 1,2,\ldots.$$

Of course, an alternative would be to solve for c in (5) by using an appropriate root finder algorithm. In the case of *infinite mean models* ($\mu = E(X) = \infty$) and $\bar{G} \in RV_{-\frac{1}{\kappa}}$ we use Degen's approximation for $\lambda\mu \approx \gamma G^{-1}\left(1-\frac{\gamma}{\lambda}\right)\frac{C_\kappa}{1-\frac{1}{\kappa}}$ (for $1<\kappa<\infty$) and $\lambda\mu \approx \lambda\mu_G\left(G^{-1}\left(1-\frac{\gamma}{\lambda}\right)\right)$ (for $\kappa=1$) in equation (5). The approximation suggested by Hannah and Puza subsequently will be referred to as SLAH.

Perturbative approximation

Hernandez et al. (2014) introduced k-th order perturbative approximations for calculating the $100(1-\gamma)\%$ VaR of the compound loss distribution. Assume $X \sim G$ and the frequency distribution is $Pois(\lambda)$ then the 0th, 1st and 2nd order approximations (subsequently denoted by PA0, PA1 and PA2 respectively) are given by Q_0, $Q_0+Q_1$and $Q_0+Q_1+\frac{Q_2}{2}$ respectively, where

$$Q_0 = G^{-1}(\frac{\lambda+\ln(1-\gamma)}{\lambda})$$

$$Q_1 = (\lambda+\ln(1-\gamma))E(X|X<Q_0);$$

$$Q_2 = -\left(\lambda g(Q_0)+\frac{g'(Q_0)}{g(Q_0)}\right)(\lambda+\ln(1-\gamma))E(X^2|X<Q_0) - \lambda g(Q_0)\,Q_0^2$$

with g the density of G.

3. Methodology for evaluation of approximation methods

In order to test the accuracy of the above-mentioned closed-form approximation techniques we designed a Monte Carlo (MC) study. As stated before, we assumed a Poisson frequency distribution throughout and selected three severity distributions, namely the Lognormal (LogN), Log Normal Inverse Gaussian (LogNig) and the Burr distribution. The densities of these distributions are given below as well as the parameter sets that were used to generate the compound distributions.

The Burr distribution

The three parameter Burr type XII distribution function is given by

$$Burr(x;\eta,\tau,\alpha) = 1-\ (1+(x/\eta)^\tau)^{-\alpha}\text{, for } x>0 \qquad (6)$$

with parameters $\eta,\tau,\alpha>0$ (see e.g. Beirlant et al., 2004). Here η is a scale parameter and τ and α shape parameters. Note the extreme value index of the Burr distribution is given by $EVI=\kappa=1/\tau\alpha$ and that heavy-tailed distributions have a positive EVI and larger EVI implies heavier tails. This follows (also) from the fact that for positive EVI the Burr distribution belongs to the Pareto-type class of distributions, having a distribution function of the form $1-F(x)=x^{-1/\kappa}\ell_F(x)$, with $\ell_F(x)$ a slowly varying function at infinity (see e.g. Embrechts et al., 1997). For Pareto-type, when the $EVI\geq 1$ the expected value does not exist, and when $EVI>0.5$, the variance is infinite. Note also that the Burr distribution is regularly varying with index $-\tau\alpha$ and therefore belongs to the class of sub-exponential distributions (see Fasen and Kluppelberg, 2006). The density of the Burr type XII is given by

$$burr(x;\ \eta,\tau,\alpha) = \frac{\tau\alpha}{\eta^{\tau}} x^{\tau-1}\left[1+\left(\frac{x}{\eta}\right)^{\tau}\right]^{-(\alpha+1)},\ \text{for } x>0 \qquad (7)$$

with parameters $\eta, \tau, \alpha > 0$ and the second order parameter is $\rho = -\frac{1}{\alpha}$.

The Lognormal distribution

The two parameter Lognormal distribution function is given by

$$Lognor(x;\ \mu,\sigma) = \frac{1}{2} + \frac{1}{2}\mathrm{erf}\left(\frac{ln(x)-\mu}{\sqrt{2}\sigma}\right) = \Phi\left(\frac{\ln(x)-\mu}{\sigma}\right),\ \text{for } x>0 \qquad (8)$$

with parameters $-\infty < \mu < \infty$ and $\sigma > 0$ where μ is a location parameter and σ a scale parameter. Note that $\Phi(.)$ denotes the standard normal distribution function. The extreme value index of the lognormal is $EVI = \kappa = 0$ and the second order parameter $\rho = 0$. The lognormal distribution is a semi-heavy tailed distribution from the class of sub-exponential distributions. The density of the lognormal is given by

$$lognor(x;\ \mu,\sigma) = \frac{1}{x\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{\ln(x)-\mu}{\sigma}\right)^2},\ \text{for } x>0 \qquad (9)$$

with parameters μ and σ .

The LogNig distribution

Let $Y = e^X$ where $X \sim$ NIG$(x;\alpha,\beta,\mu,\delta)$ and NIG is the four parameter normal inverse Gaussian distribution, then $Y \sim lognig(y;\ \alpha,\beta,\mu,\delta)$ is the four parameter log normal inverse Guassian (LogNig) distribution. Note that the distribution function $Lognig(y;\ \alpha,\beta,\mu,\delta)$ does not exist in an explicit form. Here $\alpha > 0$ is a parameter that controls tail heaviness, β $(0 < |\beta| < \alpha)$ a parameter that controls asymmetry, μ $(-\infty < \mu < \infty)$ is a location and $\delta > 0$ a scale parameter. The extreme value index is $EVI = \kappa = \frac{1}{\alpha-\beta}$ and the second order parameter $\rho = 0$. Note that when $EVI > 1$ the mean does not exist, but unlike the Burr, the mean does exist when $EVI = 1$. The density function of the LogNig follows immediately from that of the NIG and is given by

$$\begin{aligned} lognig(y;\ \alpha,\beta,\mu,\delta) &= \frac{1}{y} NIG(\alpha,\beta,\mu,\delta)(\ln(y)) \\ &= \frac{1}{y}\frac{\alpha e^{\xi+\beta(\ln(y)-\mu)} K_1\left(\alpha\delta q\left(\frac{\ln(y)-\mu}{\delta}\right)\right)}{\pi q\left(\frac{\ln(y)-\mu}{\delta}\right)} \end{aligned} \qquad (10)$$

where $q(k) = \sqrt{1+k^2}$, $\xi = \delta\sqrt{\alpha^2-\beta^2}$ and K_1 the modified Bessel function of the third order and index one. See e.g. Abramowitz and Stegun (1972).

Parameter sets for the simulation study

The parameter values considered in the simulation study are given in Tables 1-3 below for each of the three severity distributions. For each of the parameter sets, and for Poisson intensities $\lambda = 10, 20, 50, 100, 200, 500$ and probability levels $\gamma = 0.001, 0.005, 0.01, 0.025, 0.05$ the closed-form approximations were calculated (i.e. SLA, SLAD, SLAH, PA0, PA1 and PA2) and the results then compared with the corresponding MC simulation results. This consisted of 1000 VaR estimates, where each VaR estimate was based on 1 000 000 MC repetitions. The median, mean, minimum, maximum, 5% and 95% percentiles of the 1000 MC VaR estimates were then calculated.

Table 1: Burr parameter sets selected for the Monte Carlo study

Burr	η	τ	α	κ	ρ	$E(X)$
Burr1	1	0.6	5	0.333333	-0.2	$< \infty$
Burr2	1	4	0.5	0.5	-2	$< \infty$
Burr3	1	2	1	0.5	-1	$< \infty$
Burr4	1	1	2	0.5	-0.5	$< \infty$
Burr5	1	0.2	10	0.5	-0.1	$< \infty$
Burr6	1	0.6	2	0.833333	-0.5	$< \infty$
Burr7	1	2	0.5	1	-2	∞
Burr8	1	1	1	1	-1	∞
Burr9	1	0.5	2	1	-0.5	∞
Burr10	1	0.1	10	1	-0.1	∞
Burr11	1	0.5	1.5	1.333333	-0.66667	∞
Burr12	1	1.333333	0.5	1.5	-2	∞
Burr13	1	0.666667	1	1.5	-1	∞
Burr14	1	0.333333	2	1.5	-0.5	∞
Burr15	1	0.066667	10	1.5	-0.1	∞
Burr16	1	1.8	0.3	1.851852	-3.33333	∞
Burr17	1	1	0.5	2	-2	∞
Burr18	1	0.5	1	2	-1	∞
Burr19	1	0.25	2	2	-0.5	∞
Burr20	1	0.05	10	2	-0.1	∞
Burr21	1	2.5	0.17	2.352941	-5.88235	∞
Burr22	1	0.8	0.5	2.5	-2	∞
Burr23	1	0.4	1	2.5	-1	∞
Burr24	1	0.2	2	2.5	-0.5	∞
Burr25	1	0.04	10	2.5	-0.1	∞

Table 2: Lognormal parameter sets selected for the Monte Carlo study

Lognormal	μ	σ	κ	ρ	$E(X)$
LNor1	0	1	0	0	$< \infty$
LNor2	0	2	0	0	$< \infty$
LNor3	0	3	0	0	$< \infty$
LNor4	0	4	0	0	$< \infty$
LNor5	0	5	0	0	$< \infty$
LNor6	0	6	0	0	$< \infty$

Table 3: LogNig parameter sets selected for the Monte Carlo study

LogNig	α	β	μ	δ	κ	ρ	$E(X)$
LNig1	3	1	1	1	0.5	0	$< \infty$
LNig2	2	1	1	1	1	0	$< \infty$
LNig3	1	0.5	1	1	2	0	∞

Suppose V represents the value obtained for a particular approximation technique in a particular parameter set, intensity and probability level combination, then the following measures were used to compare the performance of the approximation techniques relative to the MC distribution:

- $ARE = |V - median|/median$, the absolute relative deviation from the MC median expressed as a percentage of the MC median,
- $CI = I(5\%\ MC\ percentile\ < V <\ 95\%\ MC\ percentile)$ an indicator which is one if V is included in the MC 90% confidence interval.

This was done separately for the distributions having finite and infinite means. For ease of graphical presentation we replace the probability levels $\gamma = 0.001, 0.005, 0.01, 0.025, 0.05$ by their logodds ($\ln(\frac{1-\gamma}{\gamma})$ = 6.9, 5.3, 4.6, 3.7, 2.9) and subsequently, where applicable, the horizontal axis of graphs are constructed using the logodds rather than the probability scale.

4. Results

The performance of the approximation techniques are discussed separately for distributions having finite and infinite means.

Finite mean distributions

In Table 4 below the results of the smallest ARE and percentage inclusions in the 90% MC CI over all the distributions having finite means and choices of intensities are given. It was found that the effect of intensities is not dramatic (as will be seen later) and it was included to simplify interpretation. In order to understand the figures in the table, consider a probability level of 0.001 and the PA1 approximation, then in 11.9% of the cases considered PA1 had the smallest ARE and the approximation was included in the MC IC in 75% of the cases considered.

Table 4: Smallest ARE and percentage inclusions in the MC CI over all finite mean distributions

Probability	ARE						MC CI					
	SLA	SLAD	SLAH	PA0	PA1	PA2	SLA	SLAD	SLAH	PA0	PA1	PA2
0.001	2.4%	8.3%	0.0%	2.4%	11.9%	75.0%	44.0%	72.6%	73.8%	41.7%	75.0%	91.7%
0.005	0.0%	0.0%	1.2%	0.0%	8.3%	90.5%	16.7%	14.3%	6.0%	16.7%	48.8%	71.4%
0.01	0.0%	0.0%	1.2%	0.0%	8.3%	90.5%	11.9%	2.4%	3.6%	8.3%	38.1%	59.5%
0.025	0.0%	0.0%	3.6%	0.0%	4.8%	91.7%	6.0%	0.0%	0.0%	0.0%	25.0%	57.1%
0.05	0.0%	7.1%	27.4%	0.0%	2.4%	63.1%	3.6%	1.2%	0.0%	0.0%	19.0%	35.7%
Mean	0.5%	3.1%	6.7%	0.5%	7.1%	82.1%	16.4%	18.1%	16.7%	13.3%	41.2%	63.1%

The results show that the PA2 approximation performs very well in terms of both performance measures. As far as smallest ARE is concerned, PA2 dominates the other approximation methods at probability levels 0.005, 0.01 and 0.025 and is still better at 0.001, where PA1 is its closest contender, and at 0.05, where SLAH is its closest contender. The percentage inclusions in the MC CI is very good at 0.001 (almost 92%) but deteriorates (as do the other methods) as the probability level increases to 0.05.

In order to study the performance of the approximation techniques in more detail, we present in Figures 1-3 below some of the more interesting finite mean distributions, viz LNor1, Burr3 and LNig2. In each case the relative error is plotted for each approximation technique against the log-odds of the probability levels and this is done for intensities 20, 100 and 500. The 5% and 95% MC confidence bands are also depicted as lines in each plot. In all cases PA0 and SLA perform poorly, while the other measures struggle at the higher probability levels, especially at 0.05. For LNor1 ($EVI = 0$) the performance of the approximation methods are particularly poor and it is only PA2 that performs reasonably, but only at the lower probability levels. The performance of the approximations (SLAD, SLAH, PA1 and PA2) improves significantly for Burr3 and LNig2 with PA2 clearly the preferred one.

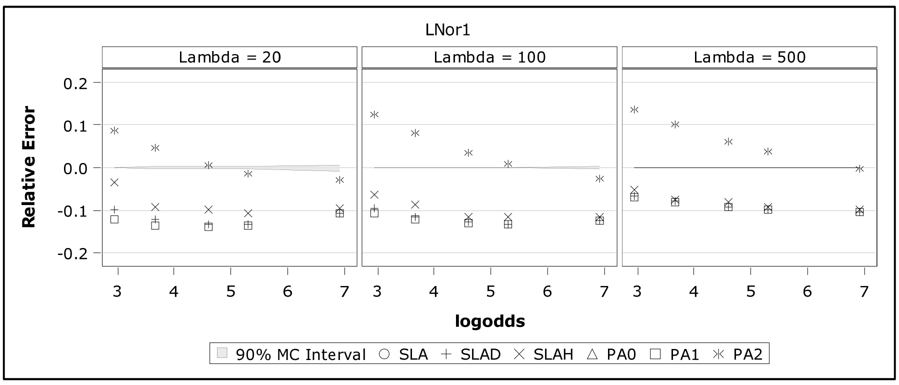

Figure 1: Relative error plots for the (very) short-tailed LNor1 distribution (EVI=0).

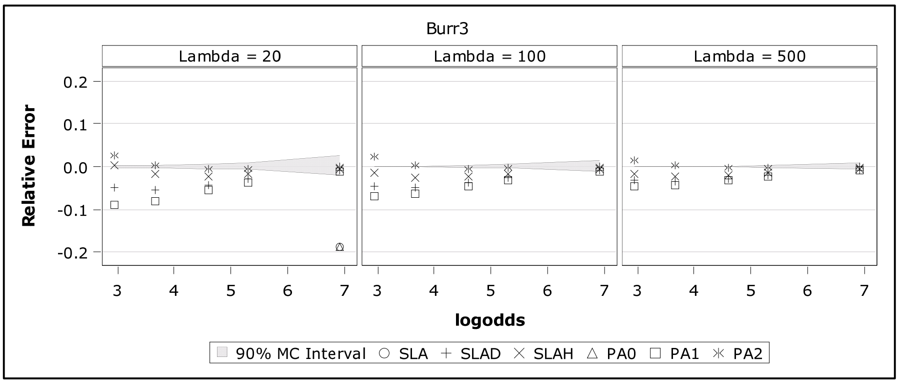

Figure 2: Relative error plots for the semi-heavy tailed Burr3 distribution (EVI=0.5)

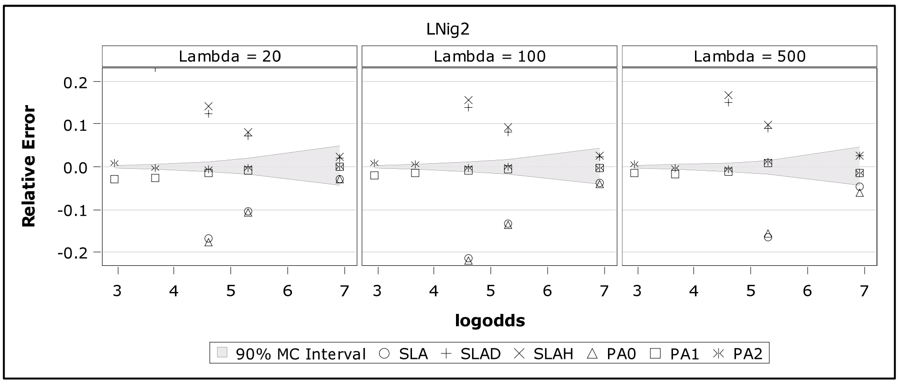

Figure 3: Relative error plots for the heavy tailed LNig2 distribution (EVI=1)

From an operational risk perspective, we are especially interested in the performance of the approximation methods at a probability level of 0.001. To this end we constructed Figure 4 below, which shows the relative error and approximation techniques for all finite mean distributions. It should be clear that, with the exception LNor1, the performance of the second order approximations is very good for all finite mean distributions considered. Given the performance of PA2 as observed in Figure 4 and Table 4, PA2 is the preferred choice.

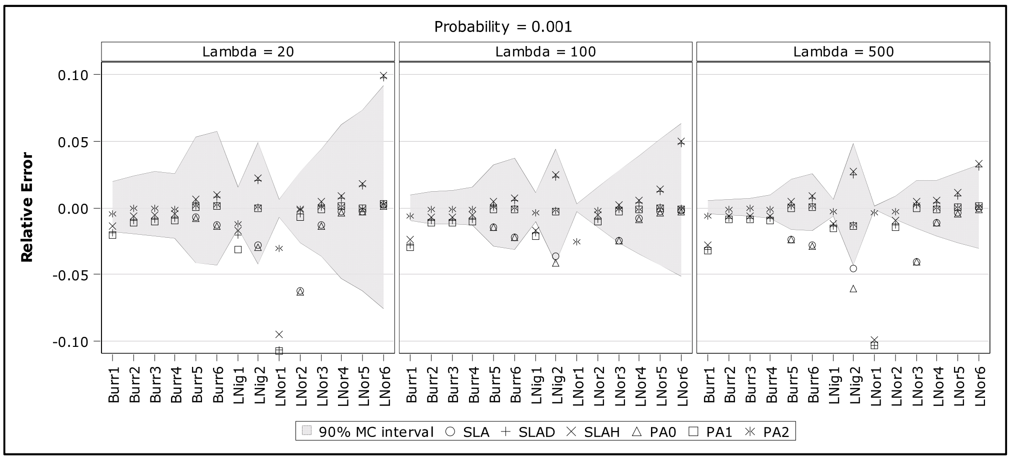

Figure 4: Relative error plots for all finite mean distributions at probability level 0.001.

Since it is difficult to compare numerically the relative performance of all the approximation methods as depicted in Figure 4 we provide the numerical relative errors of all methods as well as the Monte Carlo lower (PCLT5) and upper (PCLT95) percentiles in Table A1 of Appendix A.

As mentioned previously, other approximation techniques to consider include Panjer recursion and the fast Fourier transform of which the latter is often preferred (see e.g. Embrechts and Frei 2009). Both these techniques require the input of a numerical expert necessitated by the required intelligent selection of input parameters. Therefore application of these recursive approximation techniques is time consuming as opposed to the closed-form solutions.

Assuming the notation of Embrechts and Frei (2009), four input parameters (M, h, k, θ)need to be specified for the FFT algorithm. In order to automate the selection of input parameters we applied a rule of thumb in the following way: Take $k = 24$ and solve for h in $k = int(\frac{\ln(1.1PA2/h)}{\ln(2)})$ then take $M = 2^k$ and $\theta = M/20$. We found that this rule of thumb provided good choices for approximating the cases considered here. We computed the FFT (with tilting) for all distributions, intensities and quantiles. In Table 5 below the results of the smallest ARE and percentage inclusions in the 90% MC CI over all the distributions and all intensities are given.

Table 5: Smallest ARE and percentage inclusions in the MC CI over all finite mean distributions for FFT and PA2

Probability	ARE		MC CI	
	FFT	PA2	FFT	PA2
0.001	72.6%	27.4%	100.0%	91.7%
0.005	75.0%	25.0%	97.6%	71.4%
0.01	88.1%	11.9%	97.6%	59.5%
0.025	96.4%	3.6%	97.6%	57.1%
0.05	94.0%	6.0%	96.4%	35.7%
Mean	85.2%	14.8%	97.9%	63.1%

For the distributions considered the FFT performs very well, and outperforms PA2 at all probability levels. In order to study the performance in more detail, the relative error graphs for LNor1, Burr3 and LNig2 are depicted in Figures 5-7 below. The excellent performance of FFT is evident in all three figures. PA2 and FFT are close contenders at a probability level of 0.001 and it is noticeable that the performance of PA2 relative to FFT is much better for Burr3 and LNig2 than for the short-tailed LNor1.

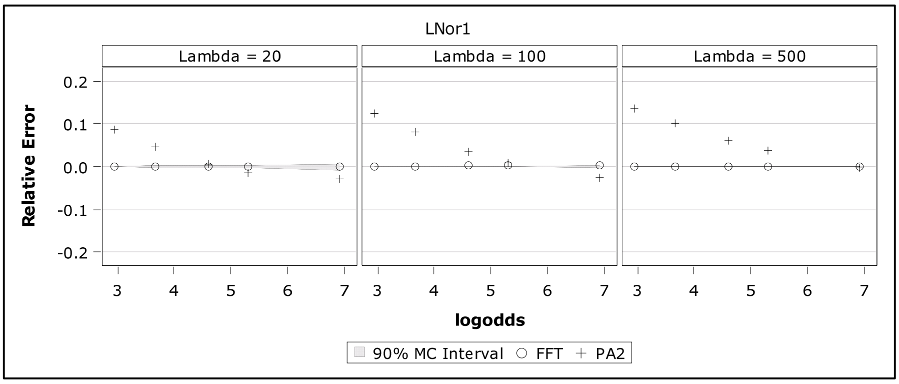

Figure 5: Relative error plots for the short-tailed LNor1 distribution (EVI=0)

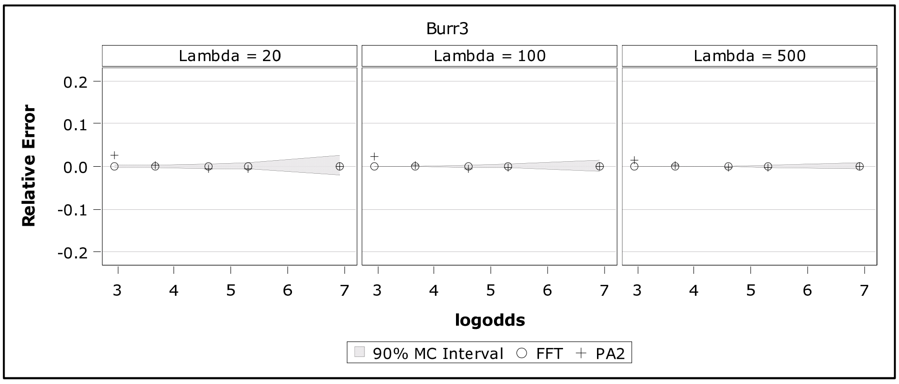

Figure 6: Relative error plots for the semi-heavy tailed Burr3 distribution (EVI=0.5)

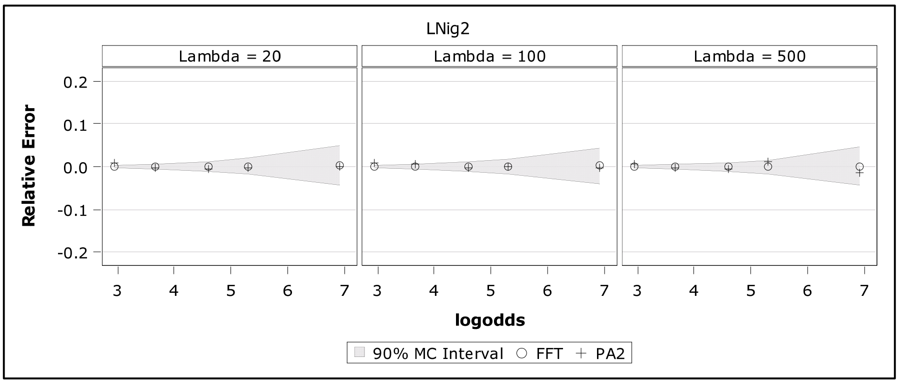

Figure 7: Relative error plots for the heavy tailed LNig2 distribution (EVI=1)

The results for the other distributions confirm that PA2 only really struggles to approximate LNor1, and is a close contender for the others, especially for distributions with higher EVI's and at a 0.001 probability level. Based on the results obtained here we recommend that the FFT be used as approximation technique using the rule of thumb provided to select input parameters. However, at a probability level of 0.001, PA2 is a good alternative.

Infinite mean distributions

In Table 6 below the results of the smallest ARE and percentage inclusions in the 90% MC CI over all the distributions having infinite means and over all intensities are given. The performance of PA2 is remarkable since it obtain 100% inclusion in the MC CI for all probability levels, except 0.05 where it achieves the best performance (99.2%). It is interesting that as far as smallest ARE is concerned, PA2 is challenged by all approximation methods at the lower probability levels, and especially at 0.001. This is to be expected since the asymptotic theory suggests that all approximations should perform well for large EVI and large intensities at high quantiles.

Table 6: Smallest ARE and percentage inclusions in the MC CI over infinite mean distributions

Probability	ARE						MC CI					
	SLA	SLAD	SLAH	PA0	PA1	PA2	SLA	SLAD	SLAH	PA0	PA1	PA2
0.001	11.9%	10.3%	23.0%	32.5%	16.7%	11.1%	90.5%	89.7%	89.7%	100%	100%	100%
0.005	10.3%	9.5%	0.8%	15.1%	26.2%	44.4%	66.7%	84.9%	81.7%	80.2%	100%	100%
0.01	5.6%	8.7%	0.8%	2.4%	41.3%	46.8%	61.1%	78.6%	51.6%	69.8%	100.%	100%
0.025	0.0%	3.2%	1.6%	0.0%	27.8%	67.5%	42.1%	45.2%	5.6%	31.0%	84.1%	100%
0.05	1.6%	5.6%	0.0%	0.0%	15.1%	79.4%	23.8%	38.1%	2.4%	9.5%	79.4%	99.2%
Mean	5.9%	7.5%	5.2%	10.0%	25.4%	49.8%	56.8%	67.3%	46.2%	58.1%	92.7%	99.8%

In order to study the quality of the approximations in more detail, we selected the Burr12, Burr13 and Burr15 distributions having the same EVI (1.5) but having different ρ parameters. The results are depicted in Figure 8, Figure 9 and Figure 10 for Burr12, Burr13 and Burr15 respectively. The results show that PA1 and PA2 are almost 'spot-on' in all cases, while SLA, SLAD and SLAH tend to overestimate slightly and PA0 underestimates slightly at especially the higher probability levels. Note how the MC CI increases as the probability level decreases.

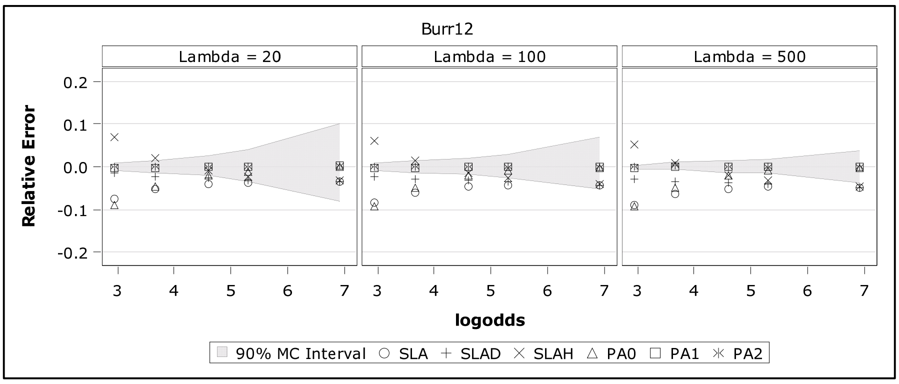

Figure 8: Relative error plots for the Burr12 distribution (EVI=1.5, ρ=-2)

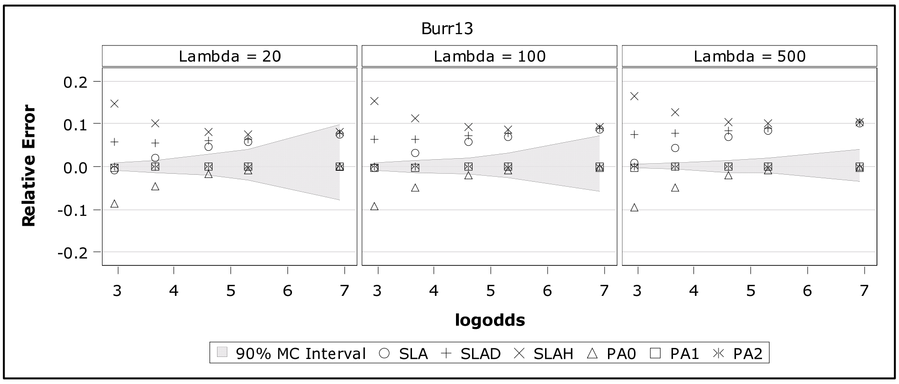

Figure 9: Relative error plots for the Burr13 distribution (EVI=1.5, ρ=-1)

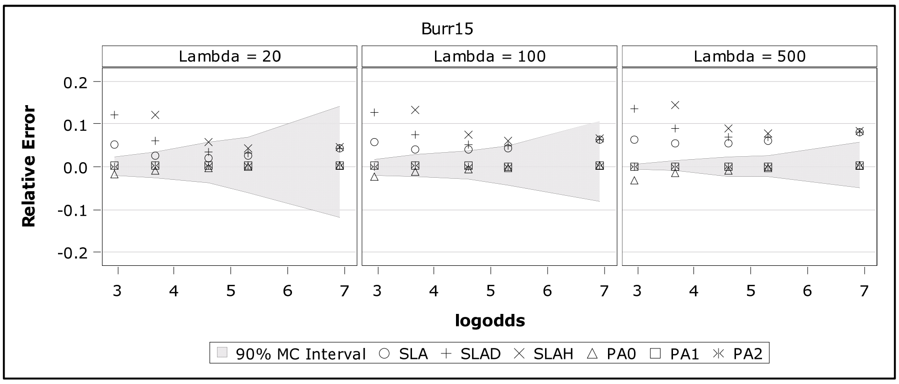

Figure 10: Relative error plots for the Burr15 distribution (EVI=1.5, Rho=-0.1)

As stated before, we are especially interested in the performance of the approximation methods at a probability level of 0.001. To this end we constructed Figure 11 below, which shows the relative error and performance of the approximation techniques for all infinite mean distributions.

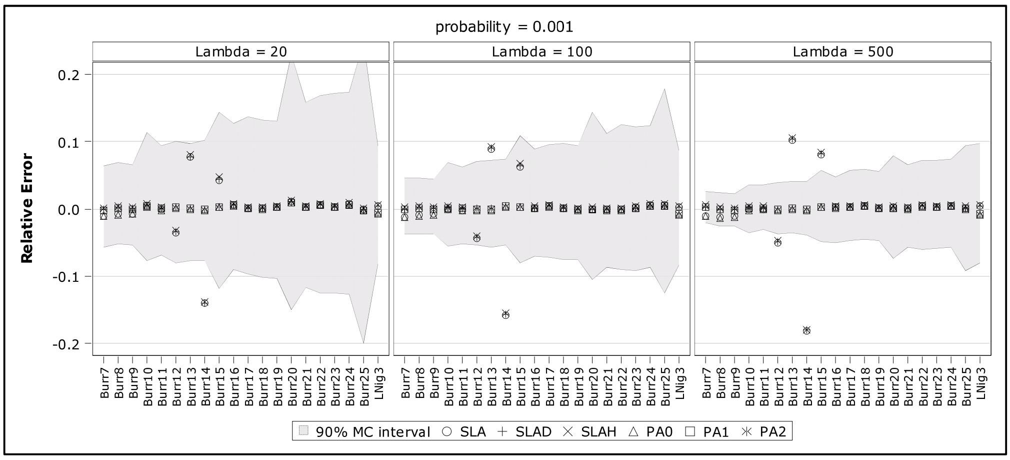

Figure 11: Relative error plots for all infinite mean distributions at probability level 0.001.

From Figure 11 it is clear that all the approximation methods, with the possible exception of SLA, SLAD and SLAH (for Burr12, Burr13, Burr14, and Burr15 respectively), perform very well.

As was the case for finite mean distributions, using the same rule of thumb we computed the FFT (with tilting) for all distributions, intensities and quantiles. In Table 7 below the results of the smallest ARE and percentage inclusions in the 90% MC CI over all the above-mentioned distributions and all intensities are given.

Taking all results into consideration it is clear that PA2 performs very well, and outperforms FFT at all probability levels. In order to study the performance in more detail, the relative error graphs for Burr12, Burr13 and Burr15 are depicted in Figure 12, Figure 13 and Figure 14 below. For these distributions the FFT struggles to compete with the PA2 approximation, which provides results almost identical to the median of the Monte Carlo approximation.

Table 7: Smallest ARE and percentage inclusions in the MC CI over all infinite mean distributions for FFT and PA2

Probability	ARE		MC CI	
	FFT	PA2	FFT	PA2
0.001	38.9%	61.1%	90.5%	100.0%
0.005	57.9%	42.1%	85.7%	100.0%
0.01	46.0%	54.0%	84.1%	100.0%
0.025	42.1%	57.9%	78.6%	100.0%
0.05	26.2%	73.8%	72.2%	99.2%
Mean	42.2%	57.8%	82.2%	99.8%

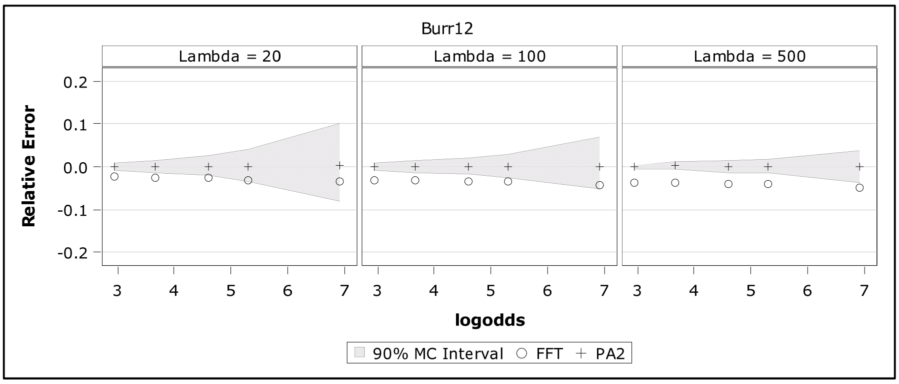

Figure 12: Relative error plots for the Burr12 distribution (EVI=1.5, Rho=-2)

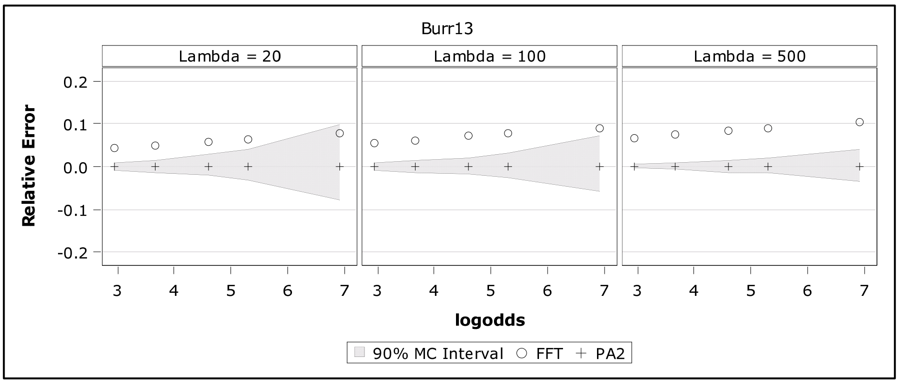

Figure 13: Relative error plots for the Burr13 distribution (EVI=1.5, Rho=-1)

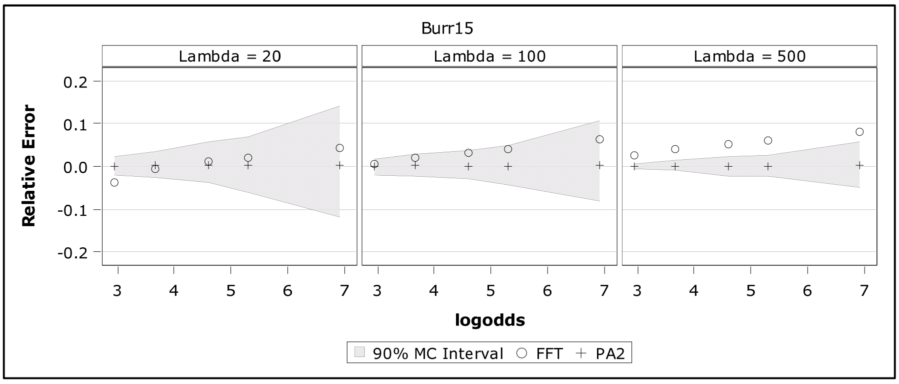

Figure 14: Relative error plots for the Burr15 distribution (EVI=1.5, Rho=-0.1)

Since we are especially interested in the performance of the approximation methods at a probability level of 0.001, we constructed Figure 15 below. The performance of FFT and PA2 are very similar with the FFT struggling at Burr12, 13, 14 and 15.

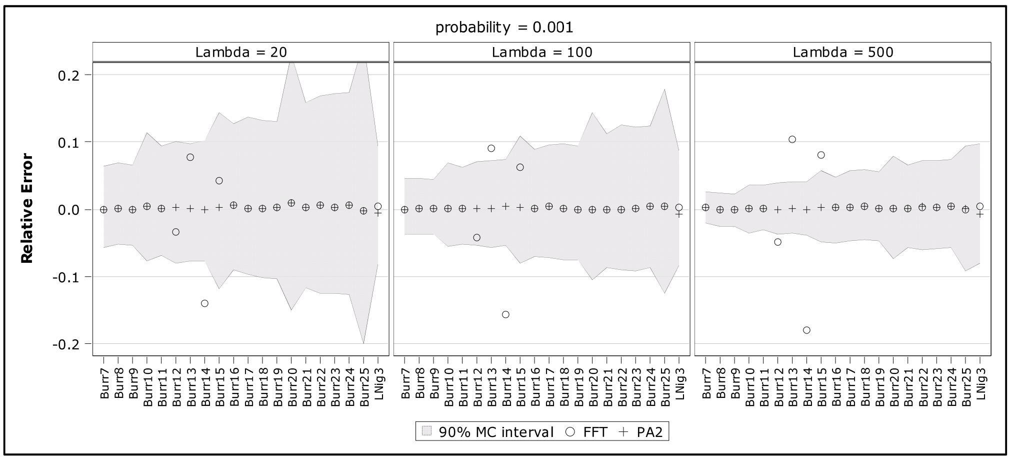

Figure 15: Relative error plots for all infinite mean distributions at probability level 0.001.

The performance of FFT could be due to the rule of thumb not providing optimal input parameters to the FFT algorithm. We have confirmed this by tweaking the tilting parameter (θ) for the FFT algorithm which yielded much better approximations. This illustrates the fact that the FFT approximation can be very good, but only when the inputs supplied are obtained with some knowledge of the correct answer. This is not the case in practice and in any case PA2 performs very well and is clearly the preferred method.

We conclude that for infinite mean distributions the PA2 approximation technique is the preferred choice at all distributions, quantiles and intensities considered.

Since it is difficult to compare numerically the relative performance of all the approximation methods as depicted in Figure 11 and Figure 15 we provide the numerical relative errors of all methods as well as the Monte Carlo lower (PCLT5) and upper (PCLT95) percentiles in Table A2 of Appendix A. The tables will help to better quantify relative errors associated with each method and answer a policy implication question: “By how much will operational risk capital be different if one of the perturbative methods are used instead of Monte Carlo?”

5. Conclusions

We have analysed the performance of the most recently proposed approximation methods for the quantiles of the compound distribution by assuming a Poisson frequency distribution and several severity distributions. We found that the PA2 perturbative closed-form approximation method of Hernandez et al. (2014) performed very well in most cases considered. If the goal is to approximate an extreme quantile of the compound distribution (such as at a probability level of 0.001), then this method is recommended. However, for short-tailed distributions and at less extreme quantiles, the FFT method could be considered using the proposed rule of thumb for the choice of input parameters (which depends on PA2). In general, in an operational risk context, the PA2 approximation is recommended because of its accuracy, clear computational speed advantage and the fact that it requires no input parameters.

APPENDIX A

Table A1: Relative errors for all finite mean distributions at probability level 0.001

Lambda	Distribution	SLA	SLAD	SLAH	PA0	PA1	PA2	FFT	PCTL5	PCTL95
20	Burr1	-15.76%	-1.91%	-1.39%	-15.77%	-2.06%	-0.42%	0.04%	-1.76%	1.95%
20	Burr2	-21.53%	-0.95%	-0.64%	-21.55%	-1.13%	-0.08%	0.09%	-1.96%	2.42%
20	Burr3	-18.89%	-0.88%	-0.56%	-18.91%	-1.06%	-0.06%	0.12%	-2.12%	2.70%
20	Burr4	-13.14%	-0.77%	-0.43%	-13.16%	-0.96%	-0.14%	0.02%	-2.29%	2.54%
20	Burr5	-0.74%	0.40%	0.61%	-0.78%	0.02%	0.11%	0.11%	-4.16%	5.29%
20	Burr6	-1.34%	0.76%	0.97%	-1.38%	0.12%	0.22%	0.22%	-4.30%	5.72%

Lambda	Distribution	SLA	SLAD	SLAH	PA0	PA1	PA2	FFT	PCTL5	PCTL95
20	LNig1	-30.48%	-2.06%	-1.72%	-31.46%	-3.17%	-1.22%	0.11%	-1.58%	1.59%
20	LNig2	-2.82%	2.02%	2.23%	-3.00%	-0.05%	0.05%	0.18%	-4.24%	4.88%
20	LNor1	-46.62%	-10.66%	-9.53%	-46.63%	-10.74%	-3.04%	-0.02%	-0.73%	0.64%
20	LNor2	-6.26%	-0.48%	-0.13%	-6.29%	-0.67%	-0.11%	-0.01%	-2.59%	2.75%
20	LNor3	-1.32%	0.20%	0.43%	-1.35%	-0.12%	0.00%	0.01%	-3.63%	4.39%
20	LNor4	-0.33%	0.70%	0.90%	-0.38%	0.09%	0.13%	0.13%	-5.31%	6.24%
20	LNor5	-0.26%	1.65%	1.84%	-0.32%	-0.06%	-0.04%	-0.05%	-6.24%	7.30%
20	LNor6	0.19%	9.77%	9.95%	0.12%	0.29%	0.29%	0.29%	-7.59%	9.16%
100	Burr1	-32.85%	-2.81%	-2.34%	-32.86%	-2.93%	-0.64%	-0.01%	-0.98%	0.93%
100	Burr2	-37.58%	-0.98%	-0.74%	-37.60%	-1.12%	-0.14%	0.00%	-1.18%	1.22%
100	Burr3	-33.83%	-0.95%	-0.69%	-33.84%	-1.10%	-0.13%	0.02%	-1.17%	1.33%
100	Burr4	-24.72%	-0.84%	-0.54%	-24.74%	-1.01%	-0.12%	0.03%	-1.33%	1.59%
100	Burr5	-1.43%	0.22%	0.45%	-1.47%	-0.11%	0.01%	0.02%	-2.87%	3.21%
100	Burr6	-2.20%	0.50%	0.72%	-2.24%	-0.13%	-0.03%	0.05%	-3.15%	3.72%
100	LNig1	-50.75%	-1.91%	-1.67%	-50.89%	-2.14%	-0.39%	0.02%	-1.03%	1.11%
100	LNig2	-3.64%	2.29%	2.49%	-4.10%	-0.32%	-0.23%	0.19%	-4.13%	4.40%
100	LNor1	-73.61%	-12.46%	-11.66%	-73.61%	-12.50%	-2.54%	0.22%	-0.29%	0.28%
100	LNor2	-13.52%	-0.90%	-0.52%	-13.54%	-1.07%	-0.21%	-0.03%	-1.44%	1.56%
100	LNor3	-2.44%	0.00%	0.25%	-2.47%	-0.28%	-0.11%	-0.09%	-2.61%	2.83%
100	LNor4	-0.82%	0.34%	0.54%	-0.86%	-0.16%	-0.10%	-0.11%	-3.43%	3.85%
100	LNor5	-0.28%	1.18%	1.37%	-0.34%	0.00%	0.02%	0.04%	-4.33%	5.12%
100	LNor6	-0.23%	4.81%	5.00%	-0.30%	-0.09%	-0.08%	-0.08%	-5.13%	6.33%
500	Burr1	-56.82%	-3.12%	-2.78%	-56.83%	-3.19%	-0.65%	-0.02%	-0.48%	0.53%
500	Burr2	-57.07%	-0.78%	-0.61%	-57.08%	-0.88%	-0.09%	-0.01%	-0.55%	0.64%
500	Burr3	-52.99%	-0.78%	-0.60%	-53.00%	-0.89%	-0.08%	0.03%	-0.60%	0.73%
500	Burr4	-41.94%	-0.82%	-0.59%	-41.95%	-0.95%	-0.13%	0.00%	-0.80%	0.95%
500	Burr5	-2.39%	0.24%	0.49%	-2.42%	-0.06%	0.11%	0.19%	-1.65%	2.11%
500	Burr6	-2.81%	0.69%	0.90%	-2.85%	0.06%	0.16%	0.16%	-1.72%	2.53%
500	LNig1	-70.91%	-1.39%	-1.25%	-71.04%	-1.57%	-0.29%	0.00%	-0.63%	0.61%
500	LNig2	-4.60%	2.48%	2.69%	-6.09%	-1.40%	-1.32%	0.13%	-4.27%	4.78%
500	LNor1	-90.24%	-10.31%	-9.91%	-90.24%	-10.33%	-0.41%	0.08%	-0.12%	0.12%
500	LNor2	-27.69%	-1.30%	-0.93%	-27.70%	-1.43%	-0.25%	0.01%	-0.85%	0.87%
500	LNor3	-4.03%	0.21%	0.47%	-4.06%	-0.05%	0.19%	0.22%	-1.52%	2.07%
500	LNor4	-1.10%	0.34%	0.55%	-1.14%	-0.09%	-0.01%	-0.01%	-2.10%	2.06%
500	LNor5	-0.38%	0.91%	1.10%	-0.44%	0.02%	0.05%	0.05%	-2.62%	2.66%
500	LNor6	-0.07%	3.09%	3.28%	-0.13%	0.13%	0.15%	0.14%	-3.05%	3.25%

Table A2: Relative errors for all infinite mean distributions at probability level 0.001.

Lambda	Distribution	SLA	SLAD	SLAH	PA0	PA1	PA2	FFT	PCTL5	PCTL95
20	Burr7	-1.00%	0.00%	0.00%	-1.00%	0.00%	0.00%	-0.16%	-6.00%	6.00%
20	Burr8	-1.00%	0.00%	0.00%	-1.00%	0.00%	0.00%	0.08%	-5.00%	7.00%
20	Burr9	-1.00%	0.00%	0.00%	-1.00%	0.00%	0.00%	-0.04%	-5.00%	7.00%
20	Burr10	0.00%	1.00%	1.00%	0.00%	0.00%	0.00%	0.41%	-8.00%	11.00%

Lambda	Distribution	SLA	SLAD	SLAH	PA0	PA1	PA2	FFT	PCTL5	PCTL95
20	Burr11	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.10%	-7.00%	9.00%
20	Burr12	-4.00%	-3.00%	-3.00%	0.00%	0.00%	0.00%	-3.38%	-8.00%	10.00%
20	Burr13	8.00%	8.00%	8.00%	0.00%	0.00%	0.00%	7.79%	-8.00%	10.00%
20	Burr14	-14.00%	-14.00%	-14.00%	0.00%	0.00%	0.00%	-13.97%	-8.00%	10.00%
20	Burr15	4.00%	4.00%	5.00%	0.00%	0.00%	0.00%	4.25%	-12.00%	14.00%
20	Burr16	1.00%	1.00%	1.00%	0.00%	1.00%	1.00%	0.61%	-9.00%	13.00%
20	Burr17	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.12%	-10.00%	14.00%
20	Burr18	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.07%	-10.00%	13.00%
20	Burr19	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.28%	-10.00%	13.00%
20	Burr20	1.00%	1.00%	1.00%	1.00%	1.00%	1.00%	0.90%	-15.00%	23.00%
20	Burr21	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.28%	-12.00%	16.00%
20	Burr22	1.00%	1.00%	1.00%	1.00%	1.00%	1.00%	0.57%	-13.00%	17.00%
20	Burr23	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.24%	-12.00%	17.00%
20	Burr24	1.00%	1.00%	1.00%	1.00%	1.00%	1.00%	0.63%	-13.00%	17.00%
20	Burr25	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	-0.25%	-20.00%	25.00%
20	LNig3	0.00%	0.00%	1.00%	-1.00%	-1.00%	-1.00%	0.45%	-8.00%	9.00%
100	Burr7	-1.00%	0.00%	0.00%	-1.00%	0.00%	0.00%	-0.11%	-4.00%	5.00%
100	Burr8	-1.00%	0.00%	0.00%	-1.00%	0.00%	0.00%	0.07%	-4.00%	5.00%
100	Burr9	-1.00%	0.00%	0.00%	-1.00%	0.00%	0.00%	0.03%	-4.00%	4.00%
100	Burr10	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.06%	-6.00%	7.00%
100	Burr11	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.04%	-5.00%	6.00%
100	Burr12	-4.00%	-4.00%	-4.00%	0.00%	0.00%	0.00%	-4.23%	-5.00%	7.00%
100	Burr13	9.00%	9.00%	9.00%	0.00%	0.00%	0.00%	8.97%	-6.00%	7.00%
100	Burr14	-16.00%	-16.00%	-16.00%	0.00%	0.00%	0.00%	-15.68%	-5.00%	7.00%
100	Burr15	6.00%	6.00%	7.00%	0.00%	0.00%	0.00%	6.25%	-8.00%	11.00%
100	Burr16	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.16%	-7.00%	9.00%
100	Burr17	0.00%	0.00%	1.00%	0.00%	0.00%	0.00%	0.41%	-7.00%	10.00%
100	Burr18	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.12%	-8.00%	10.00%
100	Burr19	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	-0.08%	-8.00%	9.00%
100	Burr20	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	-0.12%	-11.00%	14.00%
100	Burr21	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	-0.16%	-9.00%	11.00%
100	Burr22	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	-0.11%	-9.00%	12.00%
100	Burr23	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.14%	-9.00%	12.00%
100	Burr24	1.00%	0.00%	1.00%	0.00%	0.00%	0.00%	0.49%	-9.00%	12.00%
100	Burr25	1.00%	0.00%	1.00%	0.00%	0.00%	0.00%	0.41%	-12.00%	18.00%
100	LNig3	0.00%	0.00%	0.00%	-1.00%	-1.00%	-1.00%	0.28%	-8.00%	9.00%
500	Burr7	-1.00%	0.00%	1.00%	-1.00%	0.00%	0.00%	0.21%	-2.00%	3.00%
500	Burr8	-1.00%	0.00%	0.00%	-1.00%	0.00%	0.00%	-0.11%	-3.00%	2.00%
500	Burr9	-1.00%	0.00%	0.00%	-1.00%	0.00%	0.00%	-0.16%	-3.00%	2.00%
500	Burr10	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.01%	-4.00%	4.00%
500	Burr11	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.15%	-3.00%	4.00%
500	Burr12	-5.00%	-5.00%	-5.00%	0.00%	0.00%	0.00%	-4.84%	-4.00%	4.00%
500	Burr13	10.00%	10.00%	11.00%	0.00%	0.00%	0.00%	10.33%	-4.00%	4.00%

Lambda	Distribution	SLA	SLAD	SLAH	PA0	PA1	PA2	FFT	PCTL5	PCTL95
500	Burr14	-18.00%	-18.00%	-18.00%	0.00%	0.00%	0.00%	-18.08%	-4.00%	4.00%
500	Burr15	8.00%	8.00%	8.00%	0.00%	0.00%	0.00%	8.07%	-5.00%	6.00%
500	Burr16	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.19%	-5.00%	5.00%
500	Burr17	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.28%	-5.00%	6.00%
500	Burr18	0.00%	0.00%	1.00%	0.00%	0.00%	0.00%	0.44%	-5.00%	6.00%
500	Burr19	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.12%	-5.00%	6.00%
500	Burr20	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.14%	-7.00%	8.00%
500	Burr21	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.07%	-6.00%	6.00%
500	Burr22	0.00%	0.00%	1.00%	0.00%	0.00%	0.00%	0.28%	-6.00%	7.00%
500	Burr23	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.33%	-6.00%	7.00%
500	Burr24	0.00%	0.00%	1.00%	0.00%	0.00%	0.00%	0.38%	-6.00%	7.00%
500	Burr25	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	-0.05%	-9.00%	9.00%
500	LNig3	0.00%	0.00%	1.00%	-1.00%	-1.00%	-1.00%	0.41%	-8.00%	10.00%

DECLARATION OF INTEREST

The authors acknowledge grants received from the National Research Foundation, the Department of Science and Technology and the Department of Trade and Industry of South Africa. Any opinions, findings and conclusions or recommendations expressed in this material are those of the authors, and therefore the National Research Foundation does not accept any liability in regard to them.

ACKNOWLEDGEMENTS

The authors thank the referee for comments that improved the presentation of the paper.

REFERENCES

Abramowitz, M. and Stegun, I.A. (1972). Handbook of Mathematical Functions. *Dover Publications*.

Ahn, S., Kim, J.H.T., Ramaswami, V. (2012). A new class of models for heavy tailed distributions in finance and insurance risk. *Insurance: Mathematics and Economics*, 51, 43-52.

Asmussen, S. and Kroese, D.P. (2006). Improved algorithms for rare event simulation with heavy tails. *Advances in Applied Probability*, 38(2): 545-558.

Beirlant, J., Goegebeur, Y., Segers, J. and Teugels, J. (2004). Statistics of Extremes: Theory and Applications. *John Wiley and Sons*.

Böcker, K., and Klüppelberg, C. (2005). Operational VaR: a closed-form approximation. *Risk* 18(12), 90–93.

Böcker, K. and Sprittulla, J. (2006). Operational VaR: meaningful means. *Risk* 19(12), 96-98.

Cope, E.W., Mignola, G. and Ugoccioni, R. (2009). Challenges and pitfalls in measuring operational risk from loss data. *The Journal of Operational Risk*, 4(4), 3-27.

de Jongh, P.J., de Wet, T., Raubenheimer, H. and Venter, J.H. (2015). Combining scenario and historical data in the loss distribution approach: A new procedure that incorporates measures of agreement between scenarios and historical data. *Journal of Operational Risk*, 10(1), 1- 31.

Degen, M. (2010). The calculation of minimum regulatory capital using single-loss approximations. *The Journal of Operational Risk,* 5(4), 3–17.

Degen, M., Embrechts, P., Lambrigger, D.D. (2007). The quantitative modeling of operational risk: between g-and-h and EVT. *Astin Bulletin* 37(2), 265-291

Embrechts, P., Kluppelberg, C. and Mikosch, T. (1997). Modelling extremal events for Insurance and Finance. Springer.

Embrechts, P. and Frei, M. (2009). Panjer recursion versus FFT for compound distributions. Mathematical Methods of Operations Research, 69, 497–508.

Embrechts, P. and Hofert, M. (2011). Practices and issues in operational risk modelling under Basel II, *Lithuanian Mathematical Journal,* 51(2), 180–193

Fasen, V. and Kluppelberg, C. (2006). Large insurance losses distributions. In: Everitt, B. and Melnick, E. (Eds.) *Encyclopedia of Quantitative Risk Assessment*, pp. 961-969.Wiley, Chichester.

Feilmeier, M. and Bertram, J. (1987). Anwendung numerischer Methoden in der Risikotheorie. Verlag Versicherungswirtschaft E.V., Karlsruhe.

Foss, S.; Korshunov, D. and Zachary, S. (2011). An Introduction to Heavy-Tailed and Subexponential Distributions. *Springer Series in Operations Research and Financial Engineering*. Springer.

Grübel, R. and Hermesmeier, R. (1999). Computation of compound distributions I: Aliasing errors and exponential tilting. *ASTIN Bulletin,* 29(2), 197-214.

Hannah, L. and Puza, P. (2015). Approximations of value-at-risk as an extreme quantile of a random sum of heavy-tailed random variables. *The Journal of Operational Risk*, 10(2), 1-21.

Heckman and Meyers (1983). The calculation of aggregate loss distributions from claim severity and claim count distributions. *Proceedings of the Casualty Actuarial Society*, LXX:22-61.

Hernández, L., Tejero, J., Suárez, A. and Carillo-Menéndez, S. (2013). Percentiles of sums of heavy-tailed random variables: Beyond the single-loss approximation. *Statistics and Computing* (February), 1-21.

Hernández, L., Tejero, J., Suárez, A. and Carillo-Menéndez, S. (2014). Closed-form approximations for operational value-at-risk. *The Journal of Operational Risk,* 8(4), 39-54.

Hess, C. 2011. Can the single-loss approximation method compete with the standard Monte Carlo simulation technique? *The Journal of Operational Risk* 6(2), 31-43.

Hess, K. T., Liewald, A., and Schmidt, K. D. (2002). An extension of Panjer's recursion. ASTIN Bulletin, 32(2):283-297.

Mignola, G. and Ugoccioni, R. (2006), Sources of uncertainty in modeling operational risk losses, *Journal of Operational Risk* 1(2), 33-50.

Nešlehová, J., Embrechts, P., and Chavez-Demoulin, V. (2006). Infinite mean models and the LDA for operational risk. *The Journal of Operational Risk*, 1(1), 3–25.

Omey, E. and Willekens, E. (1987). Second-order behaviour of distributions subordinate to a distribution with finite mean. *Communications in Statistics: Stochastic Models*, 3(3): 311-342.

ORX (2008). Dependence and Correlation Analysis. ORX Analytic Agent: IBM Research Report to ORX Analytics Working Group.

Panjer, H.H. (1981). Recursive evaluation of a family of compound distributions. *ASTIN Bulletin* 12(1), 22-26.

Panjer, H. H. and Wang, S. (1993). On the stability of recursive formulas. *ASTIN Bulletin*, 23 (2), 227-258.

Peters, G.W., Targino,R.S. and Shevchenko,P.V. (2013). Understanding Operational Risk Capital Approximations: First and Second Orders, preprint submitted to Elsevier, http://arxiv.org/pdf/1303.2910.pdf [Date Accessed: 20 November 2013]

Sahay, A., Wan, Z., and Keller, B. (2007). Operational risk capital: asymptotics in the case of heavy-tailed severity. *The Journal of Operational Risk,* 2(2), 61–72.

Venter, J.H. and de Jongh, P.J. (2002). Risk Estimation using the Normal Inverse Gaussian Distribution, *The Journal of Risk*, 4(2), 1-23.