

# Analytic Loss Distributional Approach Models for Operational Risk from the $\alpha$ -Stable Doubly Stochastic Compound Processes and Implications for Capital Allocation.

Gareth W. Peters<sup>1</sup> Pavel . Shevchenko<sup>2</sup> Mark . Young<sup>3</sup> Wendy . Yip<sup>3</sup>

- 
- 
- <sup>1</sup> *UNSW Mathematics and Statistics Department, Sydney, 2052, Australia;  
email: garethpeters@unsw.edu.au  
(Corresponding Author)*
- <sup>2</sup> *CSIRO CMIS Sydney, NSW, Australia*
- <sup>3</sup> *Risk Services, Deloitte Touche Tohmatsu, Sydney, NSW, Australia*

arXiv:1102.3582v1 [q-fin.RM] 17 Feb 2011

# Analytic Loss Distributional Approach Models for Operational Risk from the $\alpha$ -Stable Doubly Stochastic Compound Processes and Implications for Capital Allocation.

Gareth W. Peters<sup>1</sup> Pavel . Shevchenko<sup>2</sup> Mark . Young<sup>3</sup> Wendy . Yip<sup>3</sup>

---

## Abstract

Under the Basel II standards, the Operational Risk (OpRisk) advanced measurement approach is not prescriptive regarding the class of statistical model utilised to undertake capital estimation. It has however become well accepted to utilise a Loss Distributional Approach (LDA) paradigm to model the individual OpRisk loss process corresponding to the Basel II Business line/event type. In this paper we derive a novel class of doubly stochastic  $\alpha$ -stable family LDA models. These models provide the ability to capture the heavy tailed loss process typical of OpRisk whilst also providing analytic expressions for the compound process annual loss density and distributions as well as the aggregated compound process annual loss models. In particular we develop models of the annual loss process in two scenarios. The first scenario considers the loss process with a stochastic intensity parameter, resulting in an inhomogeneous compound Poisson processes annually. The resulting arrival process of losses under such a model will have independent counts over increments within the year. The second scenario considers discretization of the annual loss process into monthly increments with dependent time increments as captured by a Binomial process with a stochastic probability of success changing annually. Each of these models will be coupled under an LDA framework with heavy-tailed severity models comprised of  $\alpha$ -stable severities for the loss amounts per loss event. In this paper we will derive analytic results for the annual loss distribution density and distribution under each of these models and study their properties.

*Key words:* Operational Risk, Loss Distributional Approach, Doubly stochastic Poisson Process,  $\alpha$ -Stable, Basel II, Solvency II.

---

## 1. Motivation

The modelling of operational risk (OpRisk) has taken a prominent place in financial quantitative measurement, this has occurred as a result of Basel II / Basel III regulatory requirements. As a result Operational risk (OpRisk) has become increasingly important to the banking industry to address these regulatory standards in light of globalization, complex financial products and changes in information technology, combined with a growing number of high-profile operational loss events worldwide.

There was no widely accepted definition of OpRisk when the Basel Committee on Banking Supervision (BCBS) began discussions on OpRisk management at the end of the 1990s; see BCBS (1998). Often, OpRisk was defined as any risk not categorised as market or credit risk. Some banks defined it as the risk of loss arising from various types of human or technical error. Some earlier definitions can be found in a 1997 survey conducted by the British Bankers Association (BBA). In January 2001, the Basel Committee on Banking Supervision issued a proposal for a New Basel Capital Accord (referred to as Basel II) where OpRisk was formally defined as a new category of risk, in addition to market and credit risks, attracting a capital charge. In the working paper BCBS (September 2001) on the regulatory treatment of OpRisk and in the revised Basel II framework BCBS (2004), the following definition of OpRisk was adopted. *“Operational risk is defined as the risk of loss resulting from inadequate or failed internal processes, people and systems or from external events. This definition includes legal risk but excludes strategic and reputational risk.”* This definition did not change in the latest version of Basel II framework, BCBS (2006, p. 144). The International Actuarial Association, IAA (2004), has adopted the same definition of operational risk in the capital requirements for insurance companies.

So OpRisk is indeed a broad category. The BCBS gives a further classification into seven event types of OpRisk (BCBS, 2006, Annex 9): Internal Fraud; External Fraud; Employment Practices and Workplace Safety; Clients, Products and Business Practices; Damage to Physical Assets; Business Disruption and System Failure; Execution, Delivery and Process Management; which serves to further illustrate the disparate nature of events in this class. Reputational and strategic risk do not fall under the OpRisk umbrella, and market and credit risks are treated separately, but almost any other event that may result in a loss to a bank, including legal action, may be termed OpRisk.

Basel II considers three pillars, which, by their very nature, emphasize the importance of assessing, modelling and understanding OpRisk profiles. These three pillars are *minimum capital requirements* (refining and enhancing risk modelling frameworks), *supervisory review* of an institution’s capital adequacy and internal assessment processes and *market discipline*, which deals with disclosure of information. Since this time, the discipline of OpRisk and its quantification have grown in prominence in the financial sector.

To illustrate just how significant OpRisk can be to a financial institution, one only needs to consider OpRisk related events such as the 1995 Barings Bank loss of around 1.3 billion GBP; the 2001 Enron loss of around 2.2 billion USD; the 2004 National Australia Bank loss of 360m AUD; and the more

recent Societe Generale loss of 4.9 billion Euro. Each of which demonstrates the severity of loss process that should be modelled by OpRisk statistical models, providing strong motivation for heavy-tailed loss process models such as those involving infinite mean and variance as captured by the family of  $\alpha$ -stable models considered in this paper.

The impact that such significant losses have had on the financial industry and its perceived stability combined with the Basel II regulatory requirements have significantly changed the view, that financial institutions have regarding OpRisk. Under the three pillars of the Basel II agreement, set out in the framework, internationally active banks are required to set aside capital reserves against risk, to implement risk management frameworks and processes for their continual review, and to adhere to certain disclosure requirements.

Whilst many OpRisk events occur frequently and with low impact (indeed, are ‘expected losses’), others are rare, and their impact may be as extreme as the total collapse of the bank. The modelling and development of methodology to capture, classify and understand properties of operational losses is a new research area in the banking and finance sector.

There are three broad approaches, that a bank may use to calculate its minimal capital reserve, as specified in the first pillar of the Basel II agreement. They are known as Basic Indicator Approach, Alternative Standardized Approach and Advanced Measurement Approach (AMA). In this paper the approach considered is the AMA. AMA is of interest since it is the most advanced framework with regards to statistical modelling.

A bank adopting the AMA must develop a comprehensive internal risk quantification system. This approach is the most flexible from a quantitative perspective, as banks may use a variety of methods and models, they believe are most suitable for their operating environment and culture, provided they can convince the local regulator, [BCBS, 2006, p150-2]. The key quantitative criteria is that a bank’s models must sufficiently account for potentially high-impact rare events. In this paper we consider the idea of the loss distribution approach (LDA) which involves modelling the severity and frequency distributions over a predetermined time horizon, typically annual as specified in the APS115 section on soundness standards.

The fitting of frequency and severity distributions, as opposed to simply fitting a single parametric annual loss distribution, involves making the mathematical choice of working with compound distributions. This would seem to complicate the matter, since it is well known, that for most situations, analytical expressions for the distribution of a compound random variable are not attainable. However, the special classes of  $\alpha$ -Stable models developed in this paper overcome this complication.

Typically, the reason for modelling severity and frequency distributions separately then constructing a compound process because some factors affect the frequency and others may affect the severity, see[13]. Some of the key points relating to why this is important in most practical settings are that the expected

number of operational losses will change as the company grows. Typically growth needs to be accounted for in forecasting the number of OpRisk losses in future years, based on previous years. This can easily be understood, when modelling is performed for frequency and severity separately. Economic inflationary effects can be directly factored into size of losses through scaling of the severity distribution. Insurance and the impacts of altering policy limits and excesses are easily understood by directly altering severity distributions. Changing recording thresholds for loss events and the impact this will have on the number of losses required to be recorded is transparent.

The most popular choices for frequency distributions are Poisson, binomial and negative binomial. The typical choices of severity distribution include exponential, Weibull, lognormal, generalized Pareto, the  $g$ -and- $h$  family of distributions [6], [14] and recently the  $\alpha$ -Stable family [20].

The most important processes to model accurately are those, which have relatively infrequent losses. However, when these losses do occur they are distributed as a very heavy-tailed severity distribution. In particular we focus our analysis on the scenarios involving heavy-tailed severity models in the rare-event extreme consequence context. Thereby providing analysis of the loss process most likely to have significant consequences on a financial institution, those which may lead to ruin. This involves introducing to OpRisk modeling an important family of severity models, utilized in insurance claims reserving in [1], given by the  $\alpha$ -stable severity model. This family of severity model are flexible enough to incorporate light-tailed Gaussian loss models through to infinite mean, infinite variance severity loss models such as the Cauchy model.

There are many approaches, which can be used to fit and incorporate expert opinion / scenario analysis for these parametric distributions and the approach adopted by a bank will depend on the data source being modelled and how much confidence one has in the data source. After the best-fitting models are selected, these are combined to produce a compound process for the annual loss distribution. From this compound process, VaR and capital estimates may be derived.

Once compound processes have been fitted for each business unit and risk type, the next step in the process is to aggregate these annual loss random variables for each individual {business line-event type} combination, and thus to obtain the institution-wide annual loss distribution. This paper will not address the issues associated with correlation and dependence modelling. For more information on typical approaches to introducing correlation in an aggregation process, including copula methods, correlation of frequency, severity or annual losses, see [5] and [16]. In the next section we present the details of the LDA modelling framework adopted in this paper.

## **2. Loss Distributional Approach Model Specifications**

OpRisk LDA models are discussed widely in the literature; see e.g. [5], [4] [9], [22]. Under the LDA Basel II requirements, banks should quantify distributions for frequency and severity of OpRisk for each

business line and event type over a one-year time horizon. In this section we begin by presenting a generic LDA framework before presenting specific models that we will develop under the compound process  $\alpha$ -Stable family. The standard LDA Basel II structures, involving an annual loss in a risk cell (business line/event type) modeled as a compound random variable,

$$Z_t^{(j)} = \sum_{s=1}^{N_t^{(j)}} X_s^{(j)}(t). \quad (2.1)$$

Here  $t = 1, 2, \dots, T, T + 1$  in our framework is discrete time (in annual units) with  $T + 1$  corresponding to the next year. The upper script  $j$  is used to identify the risk cell. The annual number of events  $N_t^{(j)}$  is a random variable distributed according to a frequency counting distribution  $P^{(j)}(\cdot)$ , typically Poisson. The severities in year  $t$  are represented by random variables  $X_s^{(j)}(t)$ ,  $s \geq 1$ , distributed according to a severity distribution  $F^{(j)}(\cdot)$ , typically lognormal. Severities represent actual loss amounts per event.

To reflect both the nature of OpRisk data with extreme but rare events, the severity models selected for this analysis are chosen to exhibit extreme heavy-tails, with particular interest in distributions with infinite mean/variance. To accomplish this we consider the family of  $\alpha$ -Stable severity distributions. These particular models have been proposed as suitable models for insurance claims modeling and finance in previous papers, such as [7], [11], [8], [19] and [17].

The study of the distribution of the annual loss process is one of the classical problems in risk theory. Closed-form solutions are typically not available for the distributions used in OpRisk. However with modern computer processing power, these distributions can be calculated virtually exactly using numerical algorithms, see approaches in [15]. The easiest to implement is the Monte Carlo method. However, because it is typically slow, Panjer recursion and Fourier inversion techniques are widely used. Both have a long history, but their applications to computing very high quantiles of the compound distribution functions with high frequencies and heavy tails are only recent developments and various pitfalls exist. In this paper the class of OpRisk models developed will be proven to admit closed form analytic solutions to the density and distribution functions of the annual loss and of the total bank's loss, for independent risks, in year  $t$  which is calculated as

$$Z_t = \sum_{j=1}^J Z_t^{(j)}, \quad (2.2)$$

where formally for OpRisk under the Basel II requirements  $J = 56$  (seven event types times eight business lines). However, this may differ depending on the financial institution and type of problem. In this paper we will drop the index  $j$  for the risk cell unless required.

The approach we consider in this paper utilizes the representation of the annual loss process for a given business line and event type combination given by a convolution. In particular, it is well known that the density and distribution of the sum of two independent continuous random variables  $Y_1 \sim F_1(\cdot)$  and  $Y_2 \sim F_2(\cdot)$  with densities  $f_1(\cdot)$  and  $f_2(\cdot)$  respectively, can be calculated via convolution as

$$f_{Y_1+Y_2}(y) = (f_1 \star f_2)(y) = \int f_2(y - y_1)f_1(y_1)dy_1$$

and

$$F_{Y_1+Y_2}(y) = (F_1 \star F_2)(y) = \int F_2(y - y_1) f_1(y_1) dy_1$$

respectively. Here the notation  $f_1 \star f_2$  denotes the convolution of  $f_1$  and  $f_2$  functions as above and the notation  $Y \sim F(y)$  means a random variable has a distribution function  $F(y)$ . Thus the distribution of the aggregated loss in Equation 2.1 can be calculated via convolutions as

$$\begin{aligned} F(z) &= Pr [Z \leq z] = \sum_{n=0}^{\infty} Pr [Z \leq z | N = n] Pr [N = n] \\ &= \sum_{n=0}^{\infty} p_n F^{(n)\star}(z). \end{aligned} \tag{2.3}$$

Here,  $F^{(n)\star}(z) = Pr [X_1 + X_2 + \dots + X_n \leq z]$  is the  $n$ -th convolution of  $F(\cdot)$  calculated recursively as

$$F^{(n)\star}(z) = \int_0^z F^{(n-1)\star}(z-x) f(x) dx$$

with  $F^{(0)\star}(z) = 1$  if  $z \geq 0$  and zero otherwise.

Next, we present a special family of statistical models for the severity distribution which are well known to be closed under convolution, meaning the analytic solution to the recursive convolution integrals above are known in closed analytic parameteric form for both the distribution and densities.

### 2.1. $\alpha$ -Stable Severity Distribution Models

Considered as generalizations of the Gaussian distribution,  $\alpha$ -Stable models are defined as the class of location-scale distributions which are closed under convolutions. As in [20] we restrict to the class of truncated  $\alpha$ -stable models to ensure we only work with non-negative loss processes. In an OpRisk context,  $\alpha$ -stable distributions possess several useful properties, including infinite mean and infinite variance, skewness and heavy tails [23] and [21].

We assume the  $i$ -th loss of the  $j$ -th risk process in year  $t$  in a risk cell is a random variable with  $\alpha$ -stable distribution, denoted by  $X_i^{(j)}(t) \sim S_\alpha(x; \beta, \gamma, \delta, 0)$ . Where,  $S_\alpha(x; \beta, \gamma, \delta, 0)$  denotes the univariate four parameter stable distribution family under parameterization  $S(0)$ , see algorithm in the Appendix and details contained in [17].

The univariate  $\alpha$ -stable distribution we consider is specified by four parameters:  $\alpha \in (0, 2]$  determining the rate of tail decay;  $\beta \in [-1, 1]$  determining the degree and sign of asymmetry (skewness);  $\gamma > 0$  the scale (under some parameterizations); and  $\delta \in \mathbb{R}$  the location. The parameter  $\alpha$  is termed the characteristic exponent, with small and large  $\alpha$  implying heavy and light tails respectively. Gaussian ( $\alpha = 2, \beta = 0$ ), Cauchy ( $\alpha = 1, \beta = 0$ ) and Levy ( $\alpha = 0.5, \beta = 1$ ) distributions provide the only analytically tractable sub-members of this family. Except these special cases, in general  $\alpha$ -stable models admit no closed-form expression for the density which can be evaluated point-wise, inference typically proceeds via the characteristic function, see discussions in [17]. However, intractable to evaluate point-wise, importantly

for OpRisk applications, simulation of random variates is very efficient, see [3] and the algorithm provided in Appendix 8.

From [12], a random variable  $X$  is said to have a stable distribution,  $S(\alpha, \beta, \gamma, \delta; 0)$ , if its characteristic function has the following form:

$$E[\exp(i\theta X)] = \begin{cases} \exp\{-\gamma^\alpha |\theta|^\alpha (1 + i\beta(\text{sign}(\theta)) \tan(\frac{\pi\alpha}{2})(|\gamma\theta|^{1-\alpha} - 1)) + i\delta\theta\} & \text{if } \alpha \neq 1 \\ \exp\{-\gamma|\theta|(1 + i\beta(\frac{2}{\pi})(\text{sign}(\theta))\ln(\gamma|\theta|)) + i\delta\theta\} & \text{if } \alpha = 1. \end{cases}$$

In the following Lemmas we present some fundamental basic facts about  $\alpha$ -Stable random variables that will be required to establish the novel analytic closed form expressions for the annual loss LDA models we develop in this paper. These will be used to construct analytic exact binomial, negative binomial, Poisson and doubly stochastic Poisson mixture representations of the annual loss process for a bank under  $\alpha$ -stable severity models with the required positive support. This will be achieved by considering a special sub-family of  $\alpha$ -stable models.

In addition we will provide an analytic expression for the tail distribution of these models and the properties of the ES in special cases of this LDA model. Finally we will utilise known asymptotic results to comment on the institution wide capital aggregation results as a result of models containing dominating risk processes such as those contained in the  $\alpha$ -Stable family. These results will extend those developed for OpRisk insurance models recently in [20].

**Lemma 1** *If  $Y \sim S(\alpha, \beta, \gamma, \delta; 0)$ , then for any  $a \neq 0, b \in \mathfrak{R}$ , the transformation  $Z = aY + b$  is a scaled version of the  $\alpha$ -stable distribution. That is  $Z \sim S(\alpha, (\text{sign}(a)\beta, |a|\gamma, a\delta + b; 0)$ . In addition, the characteristic functions, densities and distribution functions are jointly continuous in all four parameters  $(\alpha, \beta, \gamma, \delta)$  and in  $x$ . These results follow from [21] and [12] Proposition 1.16.*

**Lemma 2** *If for all  $i \in \{1, \dots, N\}$  one has random variables  $X_i \sim S(\alpha, \beta_i, \gamma_i, \delta_i; 0)$  then the distribution of the linear combination, given  $N$ , is*

$$Z = \sum_{i=1}^N X_i \sim S(\alpha, \tilde{\beta}, \tilde{\gamma}, \tilde{\delta}; 0) \quad (2.4)$$

$$\tilde{\gamma}^\alpha = \sum_{i=1}^N \gamma_i^\alpha, \quad \tilde{\beta} = \frac{\sum_{i=1}^N \beta_i \gamma_i^\alpha}{\sum_{i=1}^N \gamma_i^\alpha}, \quad \tilde{\delta} = \begin{cases} \sum_{i=1}^N \delta_i + \tan \frac{\pi\alpha}{2} \left( \tilde{\beta} \tilde{\gamma} - \sum_{i=1}^N \beta_j \gamma_j \right) & \text{if } \alpha \neq 1 \\ \sum_{i=1}^N \delta_i + \frac{2}{\pi} \left( \tilde{\beta} \tilde{\gamma} \log \tilde{\gamma} - \sum_{i=1}^N \beta_j \gamma_j \log \gamma_i \right) & \text{if } \alpha = 1 \end{cases}$$

*This result follows from ([21] Section 1.2, Property 1.2.1) and ([12], Proposition 1.17).*

The property in Lemma 2, of closure under convolution for random variables with identical  $\alpha$  parameter is lost under truncation to positive support. The implications of this are that in OpRisk the convolution property will in general be lost for general members of the truncated severity models ( $X > 0$ ). This is with one notable exception given by the sub-family of Levy distributions. Examples of loss distributions for this sub-family are illustrated in Figure 1 as a function of the scale parameter  $\gamma$ .

**Lemma 3** *If  $X \sim S(0.5, 1, \gamma, \delta; 0)$  this model specifies the sub-family of  $\alpha$ -stable models with positive*



real support  $x \in [\delta, \infty)$ . The density and distribution functions are analytic and given respectively, for  $\delta < x < \infty$ , by

$$f_X(x) = \sqrt{\frac{\gamma}{2\pi}} \frac{1}{(x - \delta)^{3/2}} \exp\left(-\frac{\gamma}{2(x - \delta)}\right), \quad F_X(x) = \operatorname{erfc}\left(\sqrt{\frac{\gamma}{2(x - \delta)}}\right).$$

The median is given by  $\tilde{\mu} = \frac{\tilde{\gamma}_n}{2} (\operatorname{erfc}^{-1}(0.5))^2$  and the mode  $M = \frac{\tilde{\gamma}_n}{3}$ , where  $\operatorname{erfc}(x) = 1 - \operatorname{erf}(x) = 1 - \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt$ . This result follows from [12] (Chapter 1. p.5).

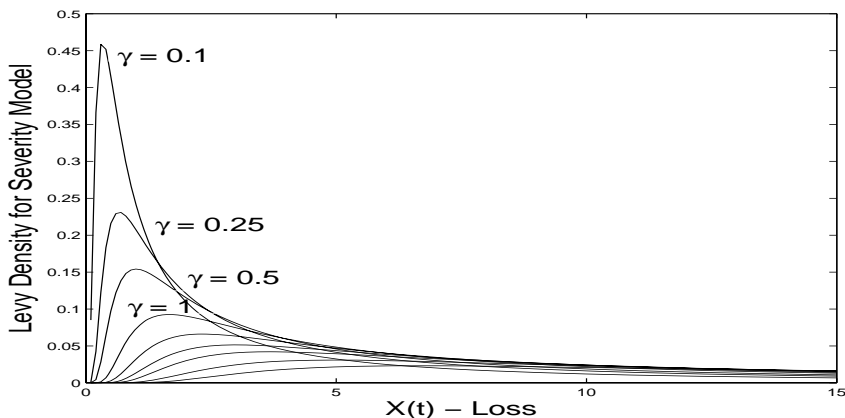


Figure 1: Example of the Levy severity model as a function of scale parameter, with location  $\delta = 0$ .

### 2.1.1. Estimation of $\alpha$ -stable severity models in OpRisk

From the perspective of application of such  $\alpha$ -stable models in OpRisk LDA settings, there are efficient parameter estimation routines available given OpRisk data; see a discussion on Bayesian approaches and review in [17]. There are also available numerical maximum likelihood based approaches, see [21] and the software package implementing such numerical procedures as discussed in [12]. In this section we discuss the practical perspective of utilising a Levy stable model for the severity distribution in an OpRisk context and how this would be justified in practical data analysis. Firstly, in an OpRisk LDA setting the severity distribution will only take positive support, hence the restriction of  $\beta = 1$  will provide this result by ensuring a positive skew that results in support only on the positive real line. We now discuss the appropriateness in practice of considering the Levy class of  $\alpha$ -stable models, which satisfy the setting in which  $\beta = 1$  and importantly admit closed form expressions for the density and distribution functions to be which we demonstrate in this paper results in closed form expressions for the annual loss distribution.

We illustrate this reasoning on the following example, where we consider  $N = 100$  losses  $X_{1:N}$  obtained from a  $\alpha$ -stable model  $S_{\alpha=0.7}(1, \gamma = 0.01, \delta = 0)$ . We fit general  $\alpha$ -stable models for an increasing number of annual losses  $n \in \{5, 10, 15, 100\}$ , the estimated MLE parameters and associated 95% C.I. are provided in Table 2.1.1. These results were obtained utilising the software available on John Nolans  $\alpha$ -stable website

<http://academic2.american.edu/~jpnolan/stable/stable.html>

MLE parameter estimates (95% C.I. [·])	N = 5	N = 10	N = 15	N = 100
$\alpha$	1.01 [0,2]	0.41 [0,2]	0.45 [0,2]	0.60 [0.46,0.74]
$\beta$	0.04 [-1,1]	0.99 [-1,1]	1.00 [-1,1]	1.00 [1.0,1.0]
$\gamma$	0.51 [-0.14,1.16]	0.14 [-0.23, 1.02]	0.12 [-0.18, 0.84]	0.01 [0.005,0.015]
$\delta$	-0.34 [0.31,0.99]	0.12 [-0.26, 0.83]	0.13 [-0.12, 0.56]	0.02 [0.016, 0.024]

Table 1: MLE parameter estimates of  $S_\alpha(\beta, \gamma, \delta)$  and 95% C.I.

The largest number of observations considered in this example with  $N = 100$  is already practically a very large number of observations in the context of OpRisk in which heavy-tailed severity models are suitable and we see in this example that even with 100 observed losses, fitting the heavy tailed  $\alpha$ -stable model estimates accurately the parameters  $\beta$ ,  $\gamma$  and  $\delta$ . However, estimation of the tail index, in this class of model represented by  $\alpha$ , with precision is a well known challenging problem with many classes of heavy tailed severity model. We demonstrate this here, where the tail index 95% C.I remains wide even with  $N = 100$  observations. Therefore, when working with heavy-tailed models in the  $\alpha$ -stable family for an OpRisk setting it is sensible to consider selecting the sub-class of  $\alpha$ -stable models which admit a parametric form, this occurs for the Levy sub-family of  $\alpha$ -stable models as presented in Lemma 3. We will now consider the LDA framework under this sub-class of severity model and demonstrate in such classes we can obtain closed form expressions for the annual loss distribution in several settings.

Next we develop novel LDA models utilising the sub-family of  $\alpha$ -Stable Levy family severity models. We begin by considering discrete time loss processes in which the total number of losses is assessed in fixed intervals of time, such as quarterly or monthly time increments. In such cases we may wish to consider non-independent counts per month and so can consider the annual loss distribution modelled by a frequency distribution with binomial distribution or with negative binomial distribution, depending on whether we believe the variance in the number of losses exceeds the mean number of losses or is less than the mean number of losses. In the continuous arrival process setting, we consider annual losses to arrive according to independent increments of time according to an exponential arrival process, resulting in a Poisson frequency distribution. Clearly under a Poisson based model we are implicitly assuming the mean number of losses per year is equal to the variance in the number of annual losses, when the rate parameter represents the average number of annual losses. We will then extend this Poisson family of models to the doubly stochastic family of frequency distributions and derive closed form expressions for the annual loss distribution in such cases.

## 2.2. binomial and negative binomial $\alpha$ -Stable LDA Analytic Models

In this section we consider the binomial and negative binomial models where we assume that the random variable for the total number of events per year are first given by standard binomial and negative binomial Processes, before extending these results to doubly stochastic processes. For the binomial proces we obtain the following result in Theorem 1.

**Theorem 1** *The distribution of the annual loss process  $Z$  represented by a compound process model with LDA structure in which the frequency is  $N(t) \sim Bi(M, p)$  and the severity model*

*$X_i(t) \sim S(0.5, 1, \gamma, \delta; 0)$ , then the exact density of the annual loss process can be expressed analytically as a mixture density comprised of  $\alpha$ -stable components with binomial mixing weights for  $N_t > 0$ ,*

$$f_Z(z) = \sum_{n=1}^M C_n^M (p)^n (1-p)^{M-n} \left[ \sqrt{\frac{\tilde{\gamma}_n}{2\pi}} \frac{1}{(z - \tilde{\delta}_n)^{3/2}} \exp\left(-\frac{\tilde{\gamma}_n}{2(z - \tilde{\delta}_n)}\right) \right] \times \mathbb{I}[\tilde{\delta}_n < z < \infty] \quad (2.5)$$

with

$$\tilde{\gamma}_n^{0.5} = \sum_{i=1}^n \gamma_i^{0.5} = n|\gamma|^{0.5}, \quad \tilde{\beta}_n = 1 \quad \tilde{\delta}_n = \sum_{i=1}^n \delta_i + \tan \frac{\pi}{4} \left( \tilde{\gamma}_n - \sum_{j=1}^n \gamma_j \right) = n\delta + \tan \frac{\pi}{4} (n^2|\gamma| - n\gamma),$$

and  $F_Z(0) = \Pr(N_t = 0) = (1-p)^M$  for  $N = 0$ . The exact form of the annual loss cumulative distribution function is also expressible in closed-form,

$$\Pr(Z < z) = F_Z(z) = \sum_{n=1}^M C_n^M (p)^n (1-p)^{M-n} \operatorname{erfc} \left( \sqrt{\frac{\tilde{\gamma}_n}{2(z - \tilde{\delta}_n)}} \right) \times \mathbb{I}[\tilde{\delta}_n < z < \infty] + (1-p)^M. \quad (2.6)$$

This result follows directly from application of Lemma 1, Lemma 2 and Lemma 3.

For the doubly stochastic binomial-Beta proces we can state the following result in Theorem 2.

**Theorem 2** *The distribution of the annual loss process  $Z$  represented by a compound process model with LDA structure in which the frequency is  $N(t) \sim Bi(M, p)$ ,  $p \sim Be(\alpha, \beta)$  and the severity model  $X_i(t) \sim S(0.5, 1, \gamma, \delta; 0)$ , then the exact density of the annual loss process can be expressed analytically as a mixture density comprised of  $\alpha$ -stable components with Beta mixing weights for  $N_t > 0$ ,*

$$\begin{aligned} f_Z(z) &= \sum_{n=1}^M \int_0^1 C_n^M (p)^n (1-p)^{M-n} \frac{1}{B(\alpha, \beta)} p^{\alpha-1} (1-p)^{\beta-1} dp \\ &\quad \left[ \sqrt{\frac{\tilde{\gamma}_n}{2\pi}} \frac{1}{(z - \tilde{\delta}_n)^{3/2}} \exp\left(-\frac{\tilde{\gamma}_n}{2(z - \tilde{\delta}_n)}\right) \right] \times \mathbb{I}[\tilde{\delta}_n < z < \infty] \\ &= \sum_{n=1}^M \frac{M!B(\alpha+n, \beta+M-n)}{(M-n)!n!B(\alpha, \beta)} \left[ \sqrt{\frac{\tilde{\gamma}_n}{2\pi}} \frac{1}{(z - \tilde{\delta}_n)^{3/2}} \exp\left(-\frac{\tilde{\gamma}_n}{2(z - \tilde{\delta}_n)}\right) \right] \times \mathbb{I}[\tilde{\delta}_n < z < \infty] \end{aligned} \quad (2.7)$$

with

$$\widetilde{\gamma}_n^{0.5} = \sum_{i=1}^n \gamma_i^{0.5} = n|\gamma|^{0.5}, \quad \widetilde{\beta}_n = 1 \quad \widetilde{\delta}_n = \sum_{i=1}^n \delta_i + \tan \frac{\pi}{4} \left( \widetilde{\gamma}_n - \sum_{j=1}^n \gamma_j \right) = n\delta + \tan \frac{\pi}{4} (n^2|\gamma| - n\gamma),$$

and  $F_Z(0) = \Pr(N_t = 0) = \frac{B(\alpha, \beta + M)}{B(\alpha, \beta)}$  for  $N = 0$ . Note we also denote the standard Beta function by  $B(x, y) = \int_0^1 t^{x-1}(1-t)^{y-1} dt$ . The exact form of the annual loss cumulative distribution function is also expressible in closed-form,

$$\Pr(Z < z) = F_Z(z) = \sum_{n=1}^M \frac{M!B(\alpha + n, \beta + M - n)}{(M-n)!n!B(\alpha, \beta)} \operatorname{erfc} \left( \sqrt{\frac{\widetilde{\gamma}_n}{2(z - \widetilde{\delta}_n)}} \right) \times \mathbb{I} \left[ \widetilde{\delta}_n < z < \infty \right] + \frac{B(\alpha, \beta + M)}{B(\alpha, \beta)}. \quad (2.8)$$

**Proof** The result for the convolution between the stable random variables, conditional on  $n$ , follows from application of Lemma 1, Lemma 2 and Lemma 3. The mixing weight of the compound process for the binomial-Beta distribution is derived by considering Bayes Theorem and the conjugacy property of the binomial-Beta model from which we know  $p(n|n) = \operatorname{Be}(p; \alpha + n, \beta + M - n)$ , hence

$$\begin{aligned} \Pr(n) &= \frac{p(n|M, p)p(p)}{p(p|n)} = \frac{\operatorname{Bi}(n; M, p)\operatorname{Be}(p; \alpha, \beta)}{\operatorname{Be}(p; \alpha + n, \beta + M - n)} = \frac{\frac{M!}{(M-n)!n!} p^n (1-p)^{(M-n)} \frac{p^{(\alpha-1)}(1-p)^{(\beta-1)}}{B(\alpha, \beta)}}{\frac{p^{(\alpha+n-1)}(1-p)^{(\beta+M-n-1)}}{B(\alpha+n, \beta+M-n)}} \\ &= \frac{M!B(\alpha + n, \beta + M - n)}{(M-n)!n!B(\alpha, \beta)}. \quad \square \end{aligned} \quad (2.9)$$

For the negative binomial proces we obtain the following result in Theorem 3.

**Theorem 3** *The distribution of the annual loss process  $Z$  represented by a compound process model with LDA structure in which the frequency is  $N(t) \sim NB(r, p)$  and the severity model*

*$X_i(t) \sim S(0.5, 1, \gamma, \delta; 0)$ , then the exact density of the annual loss process can be expressed analytically as a mixture density comprised of  $\alpha$ -stable components with negative binomial mixing weights for  $N_t > 0$ ,*

$$f_Z(z) = \sum_{n=1}^{\infty} C_n^{n+r-1} (1-p)^{(n)} (p)^r \left[ \sqrt{\frac{\widetilde{\gamma}_n}{2\pi}} \frac{1}{(z - \widetilde{\delta}_n)^{3/2}} \exp \left( -\frac{\widetilde{\gamma}_n}{2(z - \widetilde{\delta}_n)} \right) \right] \times \mathbb{I} \left[ \widetilde{\delta}_n < z < \infty \right] \quad (2.10)$$

with

$$\widetilde{\gamma}_n^{0.5} = \sum_{i=1}^n \gamma_i^{0.5} = n|\gamma|^{0.5}, \quad \widetilde{\beta}_n = 1 \quad \widetilde{\delta}_n = \sum_{i=1}^n \delta_i + \tan \frac{\pi}{4} \left( \widetilde{\gamma}_n - \sum_{j=1}^n \gamma_j \right) = n\delta + \tan \frac{\pi}{4} (n^2|\gamma| - n\gamma),$$

and  $F_Z(0) = \Pr(N_t = 0) = (1-p)^n$  for  $N = 0$ . The exact form of the annual loss cumulative distribution function is also expressible in closed-form,

$$\Pr(Z < z) = F_Z(z) = \sum_{n=1}^{\infty} C_n^{n+r-1} (1-p)^r (p)^n \operatorname{erfc} \left( \sqrt{\frac{\widetilde{\gamma}_n}{2(z - \widetilde{\delta}_n)}} \right) \times \mathbb{I} \left[ \widetilde{\delta}_n < z < \infty \right] + (1-p)^r. \quad (2.11)$$

For the doubly stochastic negative binomial-Beta proces we can state the following result in Theorem 4.

**Theorem 4** *The distribution of the annual loss process  $Z$  represented by a compound process model with LDA structure in which the frequency is  $N(t) \sim NB(r, p)$ ,  $p \sim Be(\alpha, \beta)$  and the severity model  $X_i(t) \sim S(0.5, 1, \gamma, \delta; 0)$ , then the exact density of the annual loss process can be expressed analytically as a mixture density comprised of  $\alpha$ -stable components with Beta mixing weights for  $N_t > 0$ ,*

$$\begin{aligned}
f_Z(z) &= \sum_{n=1}^{\infty} \int_0^1 C_n^{n+r-1} (1-p)^r (p)^n \frac{1}{B(\alpha, \beta)} p^{\alpha-1} (1-p)^{\beta-1} dp \\
&\quad \left[ \sqrt{\frac{\tilde{\gamma}_n}{2\pi}} \frac{1}{(z - \tilde{\delta}_n)^{3/2}} \exp\left(-\frac{\tilde{\gamma}_n}{2(z - \tilde{\delta}_n)}\right) \right] \times \mathbb{I}[\tilde{\delta}_n < z < \infty] \\
&= \sum_{n=1}^{\infty} \frac{(n+r-1)! p^{(\alpha+r-1)} B(\alpha+rn, \beta+n)}{(r-1)! n! p^{(\alpha+rn-1)} B(\alpha, \beta)} \left[ \sqrt{\frac{\tilde{\gamma}_n}{2\pi}} \frac{1}{(z - \tilde{\delta}_n)^{3/2}} \exp\left(-\frac{\tilde{\gamma}_n}{2(z - \tilde{\delta}_n)}\right) \right] \times \mathbb{I}[\tilde{\delta}_n < z < \infty]
\end{aligned} \tag{2.12}$$

with

$$\tilde{\gamma}_n^{0.5} = \sum_{i=1}^n \gamma_i^{0.5} = n|\gamma|^{0.5}, \quad \tilde{\beta}_n = 1 \quad \tilde{\delta}_n = \sum_{i=1}^n \delta_i + \tan \frac{\pi}{4} \left( \tilde{\gamma}_n - \sum_{j=1}^n \gamma_j \right) = n\delta + \tan \frac{\pi}{4} (n^2|\gamma| - n\gamma),$$

and  $F_Z(0) = Pr(N_t = 0) = \frac{p^{(\alpha+r-1)}}{p^{(\alpha-1)}}$  for  $N = 0$ . The exact form of the annual loss cumulative distribution function is also expressible in closed-form,

$$\begin{aligned}
Pr(Z < z) = F_Z(z) &= \sum_{n=1}^{\infty} \frac{(n+r-1)! p^{(\alpha+r-1)} B(\alpha+rn, \beta+n)}{(r-1)! n! p^{(\alpha+rn-1)} B(\alpha, \beta)} \operatorname{erfc} \left( \sqrt{\frac{\tilde{\gamma}_n}{2(z - \tilde{\delta}_n)}} \right) \times \mathbb{I}[\tilde{\delta}_n < z < \infty] \\
&\quad + \frac{p^{(\alpha+r-1)}}{p^{(\alpha-1)}}.
\end{aligned} \tag{2.13}$$

**Proof** The result for the convolution between the stable random variables, conditional on  $n$ , follows from application of Lemma 1, Lemma 2 and Lemma 3. The mixing weight of the compound process for the negative binomial-Beta distribution is derived by considering Bayes Theorem and the conjugacy property of the binomial-Beta model from which we know  $p(p|n) = Be(p; \alpha + rn, \beta + n)$ , hence

$$\begin{aligned}
\mathbb{P}r(n) &= \frac{p(n|r, p)p(p)}{p(p|n)} = \frac{NB(n; r, p)Be(p; \alpha, \beta)}{Be(p; \alpha + rn, \beta + n)} = \frac{C_n^{n+r-1} (1-p)^n (p)^r \frac{p^{(\alpha-1)}(1-p)^{(\beta-1)}}{B(\alpha, \beta)}}{\frac{p^{(\alpha+rn-1)}(1-p)^{(\beta+n-1)}}{B(\alpha+rn, \beta+n)}} \\
&= \frac{(n+r-1)! p^{(\alpha+r-1)} B(\alpha+rn, \beta+n)}{(r-1)! n! p^{(\alpha+rn-1)} B(\alpha, \beta)}. \quad \square
\end{aligned} \tag{2.14}$$

### 2.3. Poisson and Doubly Stochastic Poisson-Gamma - $\alpha$ -Stable LDA Analytic Models

In this section we begin by reviewing some results recently developed in [20] before extending these results to the class of doubly stochastic Poisson-Gamma- $\alpha$ -Stable LDA models. Hence, we begin with the standard Poisson process LDA model result given in Theorem 5.

**Theorem 5** *The distribution of the annual loss process  $Z$  represented by a compound process model with LDA structure in which the frequency is  $N(t) \sim Po(\lambda)$  and the severity model  $X_i(t) \sim S(0.5, 1, \gamma, \delta; 0)$ , then the exact density of the annual loss process can be expressed analytically as a mixture density comprised of  $\alpha$ -stable components with Poisson mixing weights for  $N_t > 0$ ,*

$$f_Z(z) = \sum_{n=1}^{\infty} \exp(-\lambda) \frac{\lambda^n}{n!} \left[ \sqrt{\frac{\tilde{\gamma}_n}{2\pi}} \frac{1}{(z - \tilde{\delta}_n)^{3/2}} \exp\left(-\frac{\tilde{\gamma}_n}{2(z - \tilde{\delta}_n)}\right) \right] \times \mathbb{I}[\tilde{\delta}_n < z < \infty] \quad (2.15)$$

with

$$\tilde{\gamma}_n^{0.5} = \sum_{i=1}^n \gamma_i^{0.5} = n|\gamma|^{0.5}, \quad \tilde{\beta}_n = 1 \quad \tilde{\delta}_n = \sum_{i=1}^n \delta_i + \tan \frac{\pi}{4} \left( \tilde{\gamma}_n - \sum_{j=1}^n \gamma_j \right) = n\delta + \tan \frac{\pi}{4} (n^2|\gamma| - n\gamma),$$

and  $F_Z(0) = Pr(N_t = 0) = \exp(-\lambda)$  for  $N = 0$ . The exact form of the annual loss cumulative distribution function is also expressible in closed-form,

$$Pr(Z < z) = F_Z(z) = \sum_{n=1}^{\infty} \exp(-\lambda) \frac{\lambda^n}{n!} \operatorname{erfc} \left( \sqrt{\frac{\tilde{\gamma}_n}{2(z - \tilde{\delta}_n)}} \right) \times \mathbb{I}[\tilde{\delta}_n < z < \infty] + \exp(-\lambda). \quad (2.16)$$

This result follows directly from application of Lemma 1, Lemma 2 and Lemma 3.

We can now present results for the doubly stochastic Poisson-Gamma process LDA model, given in Theorem 6.

**Theorem 6** *The distribution of the annual loss process  $Z$  represented by a doubly stochastic compound process model with LDA structure in which the frequency is  $N(t) \sim Po(\lambda)$  and  $\lambda \sim \Gamma(\alpha, \beta)$  and the severity model*

*$X_i(t) \sim S(0.5, 1, \gamma, \delta; 0)$ , then the exact density of the annual loss process can be expressed analytically as a mixture density comprised of  $\alpha$ -stable components with Poisson mixing weights for  $N_t > 0$ ,*

$$\begin{aligned} f_Z(z) &= \sum_{n=1}^{\infty} \left[ \int_0^{\infty} \exp(-\lambda) \frac{\lambda^n}{n!} \frac{\beta^\alpha}{\Gamma(\alpha)} \exp(-\beta\lambda) \lambda^{\alpha-1} d\lambda \right] \left[ \sqrt{\frac{\tilde{\gamma}_n}{2\pi}} \frac{1}{(z - \tilde{\delta}_n)^{3/2}} \exp\left(-\frac{\tilde{\gamma}_n}{2(z - \tilde{\delta}_n)}\right) \right] \times \mathbb{I}[\tilde{\delta}_n < z < \infty] \\ &= \sum_{n=1}^{\infty} \frac{(\alpha + n - 1)!}{(\alpha - 1)!n!} \left( \frac{\beta}{1 + \beta} \right)^\alpha \left( \frac{1}{1 + \beta} \right)^n \left[ \sqrt{\frac{\tilde{\gamma}_n}{2\pi}} \frac{1}{(z - \tilde{\delta}_n)^{3/2}} \exp\left(-\frac{\tilde{\gamma}_n}{2(z - \tilde{\delta}_n)}\right) \right] \times \mathbb{I}[\tilde{\delta}_n < z < \infty] \end{aligned} \quad (2.17)$$

with

$$\tilde{\gamma}_n^{0.5} = \sum_{i=1}^n \gamma_i^{0.5} = n|\gamma|^{0.5}, \quad \tilde{\beta}_n = 1 \quad \tilde{\delta}_n = \sum_{i=1}^n \delta_i + \tan \frac{\pi}{4} \left( \tilde{\gamma}_n - \sum_{j=1}^n \gamma_j \right) = n\delta + \tan \frac{\pi}{4} (n^2|\gamma| - n\gamma),$$

and  $F_Z(0) = Pr(N_t = 0) = \frac{(\beta+1)^\alpha}{\Gamma(\alpha)}$  for  $N = 0$ . The exact form of the annual loss cumulative distribution

function is also expressible in closed-form,

$$\begin{aligned} \Pr(Z < z) = F_Z(z) &= \sum_{n=1}^{\infty} \frac{(\alpha + n - 1)!}{(\alpha - 1)!n!} \left(\frac{\beta}{1 + \beta}\right)^{\alpha} \left(\frac{1}{1 + \beta}\right)^n \operatorname{erfc} \left( \sqrt{\frac{\tilde{\gamma}_n}{2(z - \tilde{\delta}_n)}} \right) \times \mathbb{I} \left[ \tilde{\delta}_n < z < \infty \right] \\ &+ \frac{(\beta + 1)^{\alpha}}{\Gamma(\alpha)}. \end{aligned} \quad (2.18)$$

Next we present the proof of this result.

**Proof** The result for the convolution between the stable random variables, conditional on  $n$ , follows from application of Lemma 1, Lemma 2 and Lemma 3. The mixing weight of the compound process for the Poisson-Gamma distribution is derived by considering Bayes Theorem and the conjugacy property of the Poisson-Gamma model from which we know  $p(\lambda|n) = \Gamma(\lambda; \alpha + n, \beta + 1)$ , hence

$$\begin{aligned} \mathbb{P}_r(n) &= \frac{p(n|\lambda)p(\lambda)}{p(\lambda|n)} = \frac{Po(n; \lambda)\Gamma(\lambda; \alpha, \beta)}{\Gamma(\lambda; \alpha + n, \beta + 1)} = \frac{\frac{\lambda^n e^{-\lambda}}{n!} \frac{\beta^{\alpha}}{\Gamma(\alpha)} \lambda^{\alpha-1} e^{-\beta\lambda}}{\frac{(\beta+1)^{\alpha+n}}{\Gamma(\alpha+n)} \lambda^{\alpha+n-1} e^{-(\beta+1)\lambda}} \\ &= \frac{\Gamma(\alpha + n)}{\Gamma(\alpha)n!} \frac{\beta^{\alpha}}{(1 + \beta)^{\alpha+n}} = \frac{(\alpha + n - 1)!}{(\alpha - 1)!n!} \left(\frac{\beta}{1 + \beta}\right)^{\alpha} \left(\frac{1}{1 + \beta}\right)^n. \end{aligned} \quad (2.19)$$

Note, this mixing weight is then a negative binomial probability with  $r = \alpha$  and  $p = \frac{1}{1+\beta}$  as presented in p.53 of [10].  $\square$

Next we will also study the truncation error in these settings for single loss processes. This will be followed by a consideration of some simulation studies comparing the properties of the capital estimation under the binomial, negative binomial and Poisson models and the doubly stochastic models developed.

### 3. Approximation of Infinite Sum Representations by Truncation

In this section we consider analytic expressions for the approximation error of the infinite sum LDA model representations in Theorems 3,4,5,6. To achieve this we will build on the results developed in [20], which utilised the properties in Lemma 4.

**Lemma 4** *Given a random variable for the severity of a loss  $X \sim S(\alpha, \beta, \gamma, \delta; 0)$  then as  $x \rightarrow \infty$  one can write the limiting tail distribution*

$$\begin{aligned} P(X > x) &\sim \gamma^{\alpha} c_{\alpha} (1 + \beta) x^{-\alpha}, \text{ as } x \rightarrow \infty \\ f_X(x|\alpha, \beta, \gamma, \delta; 0) &\sim \alpha \gamma^{\alpha} c_{\alpha} (1 + \beta) x^{-(\alpha+1)}, \text{ as } x \rightarrow \infty. \end{aligned} \quad (3.1)$$

where  $c_{\alpha} = \sin\left(\frac{\pi\alpha}{2}\right) \frac{\Gamma(\alpha)}{\pi}$ . This result follows from ([12], Theorem 1.12)

We may now use Lemma 4 to help determine a truncation index for the infinite sums in Theorems 3,4,5 and 6 as show in Theorem 7. To achieve this we will first make an approximation in which we consider the index of the summation to be a continuous variable, as demonstrated in for example [Section 3.1] of

[18]. In addition, in this paper we will work with the analytic compound process expression for the tail asymptotic, and ensure that after truncation, the error in the finite expansion of the tail asymptotic is bounded to a desired precision. More precisely, the asymptotic tail probability of the compound process will be established to have a desired approximation error under the truncation derived. This is not unique, and equally we could have chosen other distributional properties such as the median that was used in [20].

**Theorem 7** *Given the LDA structure in which the frequency is distributed generically according to  $N(t) \sim f(\cdot)$  and the severity model  $X_i(t) \sim S(0.5, 1, \gamma, \delta; 0)$ , the exact expressions for the tail asymptotic of the resulting infinite mixture is given by,*

$$Pr(Z(t) > z_q) \sim 2x^{-0.5} c_{0.5} \sum_{n=1}^{\infty} Pr(N(t) = n) \tilde{\gamma}_n^{0.5} = C_{Tail} \sum_{n=1}^{\infty} W_n$$

where  $z_q$  is an upper tail quantile of the annual loss distribution and it will be assumed for simplicity that  $\tilde{\gamma}_n^{0.5} = \sum_{i=1}^n \gamma_i^{0.5} = n|\gamma|^{0.5}$ . These results allows us to determine a unique maximum of  $\frac{d}{dn} \log(W_n) = 0$  corresponding to the term  $n$  with the maximum contribution to the LDA compound processes tail probability for each model. Note this can be determined trivially for each model by searching for a sign change in the function for each integer value of  $n$  and so is very efficient to solve since only requires integer search. The specified equations for each model are given as follows:

Poisson- $\alpha$ -Stable.

Under this model we have that  $W_n = |\gamma|^{0.5} \exp(-\lambda) \frac{\lambda^n}{(n-1)!}$  which results in the function  $\frac{d}{dn} \log(W_n) = 0$  producing  $\log(\lambda) - \log(n) - \frac{1}{2n} = 0$ .

Doubly Stochastic Poisson-Gamma- $\alpha$ -Stable.

Under this model we have that  $W_n = |\gamma|^{0.5} \frac{(\alpha+n-1)!}{(\alpha-1)!n!} \left(\frac{\beta}{1+\beta}\right)^\alpha \left(\frac{1}{1+\beta}\right)^n$  which results in the function  $\frac{d}{dn} \log(W_n) = 0$  producing  $\log(n + \alpha - 1) + \frac{n}{(n+\alpha-1)} + \frac{(\alpha-0.5)}{(n+\alpha-1)} - \log(n) - 1 - \frac{1}{2n} - \log(\beta + 1) = 0$ .

negative binomial- $\alpha$ -Stable.

Under this model we have that  $W_n = |\gamma|^{0.5} \frac{(n+r-1)!}{(n+r-1-n)!(n)!} (1-p)^r (p)^n$  which results in the function  $\frac{d}{dn} \log(W_n) = 0$  producing  $\log(n + r - 1) + \frac{n}{n+r-0.5} + \frac{r-0.5}{n+r-1} - 1 - \log(n) - \frac{1}{2n} + \log(p) = 0$ .

Doubly Stochastic negative binomial- Beta- $\alpha$ -Stable.

Under this model we have that  $W_n = |\gamma|^{0.5} \frac{(n+r-1)! p^{(\alpha+r-1)} B(\alpha+rn, \beta+n)}{(r-1)! n! p^{(\alpha+r-1)} B(\alpha, \beta)}$  which results in the function  $\frac{d}{dn} \log(W_n) = 0$  producing,

$$\begin{aligned} & \log(n+r) + \frac{(n+r+0.5)}{(n+r)} + r \log(\alpha+rn) + \frac{r(\alpha+rn-0.5)}{(\alpha+rn)} + \log(\beta+n) + \frac{(\beta+n-0.5)}{(\beta+n)} \\ & - (r+1) \log(\alpha+rn+\beta+n) - \frac{(r+1)(\alpha+rn+\beta+n-0.5)}{(\alpha+rn+\beta+n)} - \log(n+1) - \frac{(n+1+0.5)}{(n+1)} - r \log(p) = 0. \end{aligned} \tag{3.2}$$

We can now find  $W_L < W_0 < W_U$  such that the truncated sum approximation is sufficiently accurate for use in the expression of the annual loss distribution of the compound process. Hence, we can bound the number of terms in the mixture representations to not change by a precision amount  $\epsilon$ . This is expressed



analytically as

$$f_Z(z) = \sum_{n=N_L}^{N_U} Pr(N_t = n) \left[ \sqrt{\frac{\tilde{\gamma}_n}{2\pi}} \frac{1}{(x - \tilde{\delta}_n)^{3/2}} \exp\left(-\frac{\tilde{\gamma}_n}{2(x - \tilde{\delta}_n)}\right) \right] \quad (3.3)$$

with  $N_L$  and  $N_U$  are determined from  $\max(W_L \leq e^{-37}W_0, 1)$  and  $W_U \leq e^{-37}W_0$ . The derivation of these results are found in Appendix 9. We will demonstrate simulated examples for these results on the truncation in the following results and discussion section.

#### 4. Simulation Studies

In this section we undertake several simulation studies, for all studies we will utilise the parameter settings in Table 4. In addition, the calculated values for  $N_L$  and  $N_U$  are provided for each of these models, utilising the results derived in Theorem 7 in Section 3. For each of the frequency distributions, we considered two settings, low frequency and high frequency. In addition, all simulations were completed in Matlab with vectorised code on an SGI cluster in which the compute node utilised was an Altix XE320 with two Intel Xeon X5472 (quad core 3.0GHz) CPU's, the source code available at:

`\protect\vrule width0pt\protect\href{http://web.maths.unsw.edu.au/~peterga/index_files/QRSLab/Te`

LDA Model	Frequency Parameters				Severity Parameters		Truncation	
	$\lambda$	n	p	r	$\gamma$	$\delta$	$N_L$	$N_U$
Low Frequency Example								
binomial-Levy	-	12	0.1	-	0.01	0	0	M
binomial-Beta-Levy	-	12	Be(1,5)	-	0.01	0	0	M
negative binomial-Levy	-	-	0.1	2	0.01	0	0	19
negative binomial-Beta-Levy	-	-	Be(1,5)	2	0.01	0	0	20
Poisson-Levy	0.1	-	-	-	0.01	0	0	11
Poisson-Gamma-Levy	Ga(1,0.1)	-	-	-	0.01	0	0	171
High Frequency Example								
binomial-Levy	-	12	0.6	-	0.01	0	0	M
binomial-Beta-Levy	-	12	Be(5,1)	-	0.01	0	0	M
negative binomial-Levy	-	-	0.6	10	0.01	0	0	120
negative binomial-Beta-Levy	-	-	Be(5,1)	10	0.01	0	0	169
Poisson-Levy	10	-	-	-	0.01	0	0	49
Poisson-Gamma-Levy	Ga(1,10)	-	-	-	0.01	0	0	171

Table 2: Parameter settings for the models utilised in the simulations.

#### 4.1. Analysis of Model Estimation under Exact Truncated Solutions versus Monte Carlo

In this section we first illustrate that the derived annual loss compound process distributions in Theorems 1 to Theorem 6 are correct. This is achieved by undertaking a simulation study in which we simulate from each of the models developed first via a standard Monte Carlo procedure, utilising  $T = 1,000,000$  annual simulated years for each LDA model. We record the simulation time as well as the estimated Monte Carlo error obtained by blocking the simulated years into sets of 50,000 samples, used to calculate the standard error in estimates obtained from the simulated annual losses. We then estimate the empirical distribution function and plot this along with the estimated standard error from the Monte Carlo simulation. Next we evaluate each of the models using truncations of the exact expressions derived in Theorem 1 through to Theorem 6, where we utilise  $N_L = 1$  and  $N_U = 1,000$ . The evaluations are done at 200 points from  $Z = 1$  to  $Z = 200$  which for the parameter settings in Table 4 covered sufficient range for the evaluation to include quantiles representing regulatory capital VaR's at 99.5%. We then plot the evaluated distribution functions versus the Monte Carlo estimated empirical distributions functions for each of the compound process models.

In Figure 2 we present the results of this simulation for the Annual loss compound process with Levy severity distribution and the binomial and the doubly stochastic binomial-Beta models derived in Theorem 1 and Theorem 2. The distribution functions are presented for both the low and high frequency examples.

In Figure 3 we present the results of this simulation for the Annual loss compound process with Levy severity distribution and the negative binomial and the doubly stochastic negative binomial-Beta models derived in Theorem 5 and Theorem 6. The distribution functions are presented for both the low and high frequency examples.

In Figure 4 we present the results of this simulation for the Annual loss compound process with Levy severity distribution and the Poisson and the doubly stochastic Poisson-Gamma models derived in Theorem 5 and Theorem 6. The distribution functions are presented for both the low and high frequency examples.

### 5. Analysis of Truncation Error and Computational Saving versus Monte Carlo

Monte Carlo techniques are widely used in the generation of compound processes and especially in the estimation of the annual loss distribution and the associated risk measures of Value at Risk (VaR) and Expected Shortfall (ES). In this section we first study the truncation error when applied to the models derived in Theorem 1 through to Theorem 6 and then we demonstrate the accuracy of the derived models under the estimated truncations  $N_L$  and  $N_U$ . Finally, we compared the simulation times for each method when we very conservatively set  $N_L = 1$  and  $N_U = 1,000$ , well in excess of the required estimates truncations for these examples as shown below.

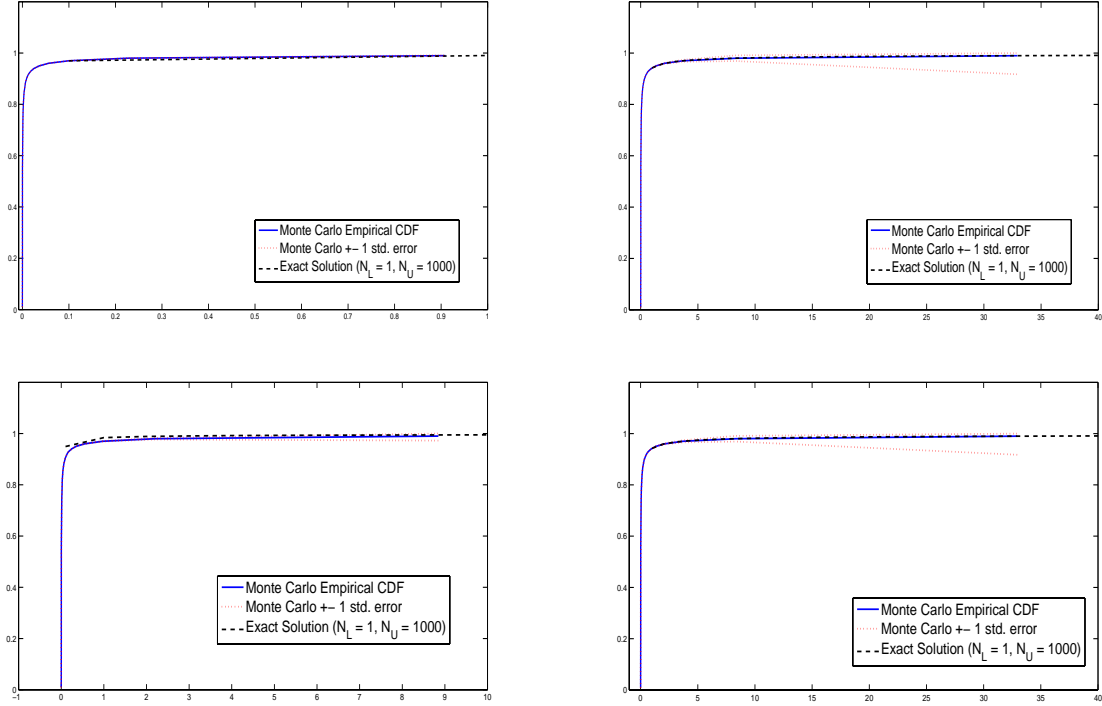


Figure 2: Comparison of empirical CDF from  $T = 1,000,000$  simulated annual years (dashed lines), versus Truncated  $N_U = 1000$  Exact Expressions (solid line). **Top Left** - Low frequency Poisson-Levy model **Top Right** - High frequency Poisson-Levy model **Bottom Left** - Low frequency Poisson-Gamma-Levy model **Bottom Right** - High frequency Poisson-Gamma-Levy model

We demonstrate the estimated CDF and MSE for different levels of truncation  $N_U \in \{1, 2, \dots, 1,000\}$  in Figure 5 for the binomial and doubly stochastic binomial-Beta examples. Figure 6 shows the negative binomial and doubly stochastic negative binomial-Beta examples and Figure 7 shows the Poisson and doubly stochastic Poisson-Gamma examples.

These results demonstrate that in all cases, the mean squared error falls rapidly for small numbers of upper truncation index  $N_U$ . This confirms the utility of utilising these truncations, which are extremely efficient simulation wise, where in Table 3 we see that even when using an excessive number of terms  $N_U = 1,000$  we make very significant computational savings and accuracy improvements over the Monte Carlo equivalent simulations.

Next we present a general result which provides a closed form distributional result for the combined loss process for multiple risk processes, each under one of the LDA model structures presented.

## 6. Multiple Risk Process Aggregation

In this section we demonstrate how to obtain the distribution of the combined loss processes under each of the models developed by presenting a general closed form result for the aggregation of two risk processes

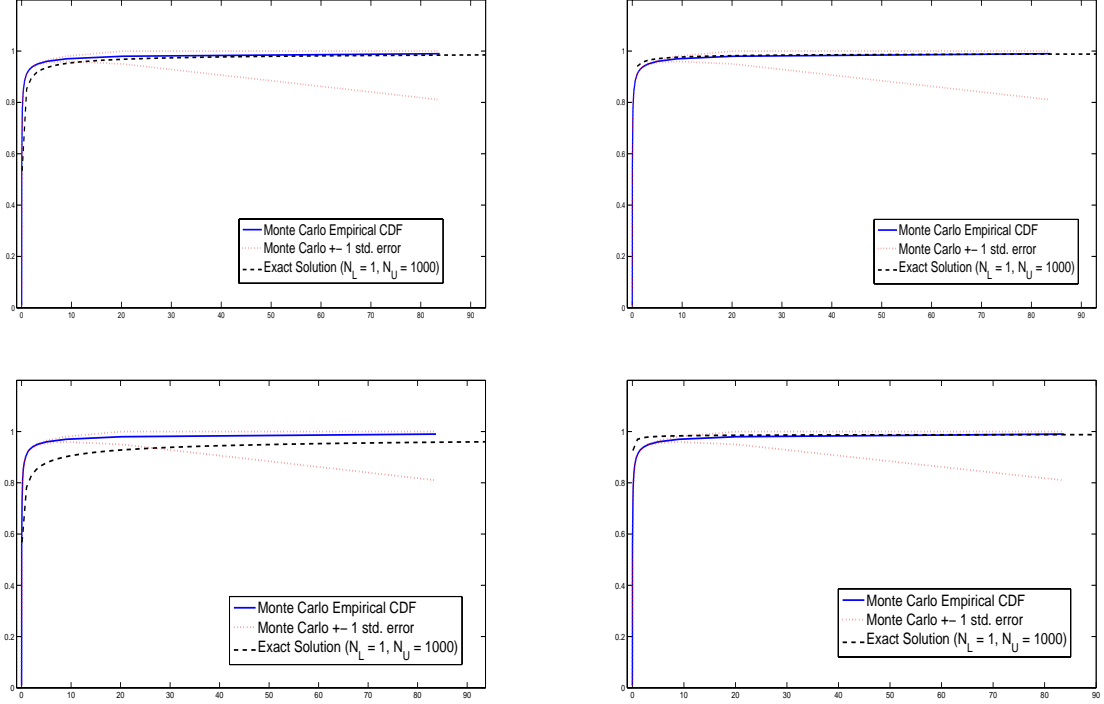


Figure 3: Comparison of epirical CDF from  $T = 1,000,000$  simulated annual years (dashed lines), versus Truncated  $N_U = 1000$  Exact Expressions (solid line). **Top Left** - Low frequency Poisson-Levy model **Top Right** - High frequency Poisson-Levy model **Bottom Left** - Low frequency Poisson-Gamma-Levy model **Bottom Right** - High frequency Poisson-Gamma-Levy model

which is trivially extended to any number of risk processes. Then we illustrate this result with an example using two Poisson processes LDA models.

**Theorem 8** *The distribution of the annual loss process represented by multiple risks, eg.  $j \in \{1, 2\}$ , in a LDA compound process structure in which the frequency is  $N^{(j)}(t) \sim F(\theta^{(j)})$  and the severity model  $X_i^{(j)}(t) \sim S(0.5, 1, \gamma^{(j)}, \delta^{(j)}; 0)$ , can be expressed analytically as a mixture distribution comprised of  $\alpha$ -stable components with mixing weights specified by the chosen frequency distribution. Define the total annual loss random variable as,*

$$Z_t = Z_t^{(1)} + Z_t^{(2)} = \sum_{i=1}^{N^{(1)}(t)} X_i^{(1)}(t) + \sum_{j=1}^{N^{(2)}(t)} X_j^{(2)}(t). \quad (6.1)$$

*Furthermore, if each of these loss processes has an insurance mitigation applied under an ILP policy with top cover limits,  $\delta^{(1)} \geq TCL^{(1)}$  and  $\delta^{(2)} \geq TCL^{(2)}$ , then the closed-form analytic expression for the*

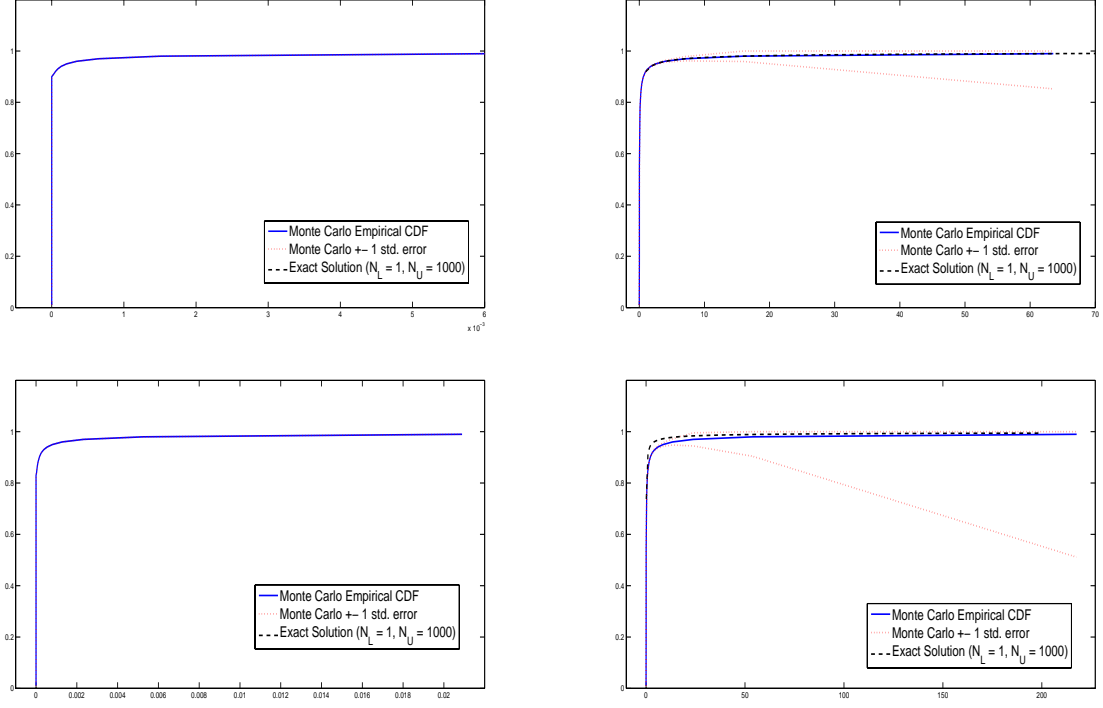


Figure 4: Comparison of empirical CDF from  $T = 1,000,000$  simulated annual years (dashed lines), versus Truncated  $N_U = 1000$  Exact Expressions (solid line). **Top Left** - Low frequency Poisson-Levy model **Top Right** - High frequency Poisson-Levy model **Bottom Left** - Low frequency Poisson-Gamma-Levy model **Bottom Right** - High frequency Poisson-Gamma-Levy model

annual loss density for  $N_t^{(1)} + N_t^{(2)} > 0$  is given by

$$\begin{aligned}
 f_{Z_T}(z) &= \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} Pr(N_t^{(1)} = m) Pr(N_t^{(2)} = n) \left[ \sqrt{\frac{\tilde{\gamma}_{nm}}{2\pi}} \frac{1}{(z - \tilde{\delta}_{nm})^{3/2}} \exp\left(-\frac{\tilde{\gamma}_{nm}}{2(z - \tilde{\delta}_{nm})}\right) \right] \times \mathbb{I}[z > \tilde{\delta}_{nm}] \\
 \tilde{\gamma}_{nm}^{0.5} &= \sum_{i=1}^n |\gamma_i^{(1)}|^{0.5} + \sum_{j=1}^m |\gamma_j^{(2)}|^{0.5} = n|\gamma^{(1)}|^{0.5} + m|\gamma^{(2)}|^{0.5}, \quad \tilde{\beta}_{nm} = 1 \\
 \tilde{\delta}_{nm} &= \sum_{i=1}^n \delta_i^{(1)} + \sum_{j=1}^m \delta_j^{(2)} + \tan \frac{\pi}{4} \left( \tilde{\gamma}_{nm} - \sum_{j=1}^n \gamma_j^{(1)} - \sum_{k=1}^m \gamma_k^{(2)} \right) \\
 &= n\delta^{(1)} + m\delta^{(2)} + \tan \frac{\pi}{4} \left( (n|\gamma^{(1)}|^{0.5} + m|\gamma^{(2)}|^{0.5})^2 - n\gamma^{(1)} - m\gamma^{(2)} \right),
 \end{aligned} \tag{6.2}$$

and  $f_{Z_T}(0) = Pr(N_t^{(1)} + N_t^{(2)} = 0)$ . The exact form of the insurance mitigated annual loss cumulative distribution function is also expressible in closed-form,

$$\begin{aligned}
 Pr(Z_T < z) &= F_{Z_T}(z) = \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} Pr(N_t^{(1)} = m) Pr(N_t^{(2)} = n) \operatorname{erfc} \left( \sqrt{\frac{\tilde{\gamma}_{nm}}{2(z - \tilde{\delta}_{nm})}} \right) \times \mathbb{I}[z > 0] \\
 &+ Pr(N_t^{(1)} + N_t^{(2)} = 0).
 \end{aligned} \tag{6.3}$$

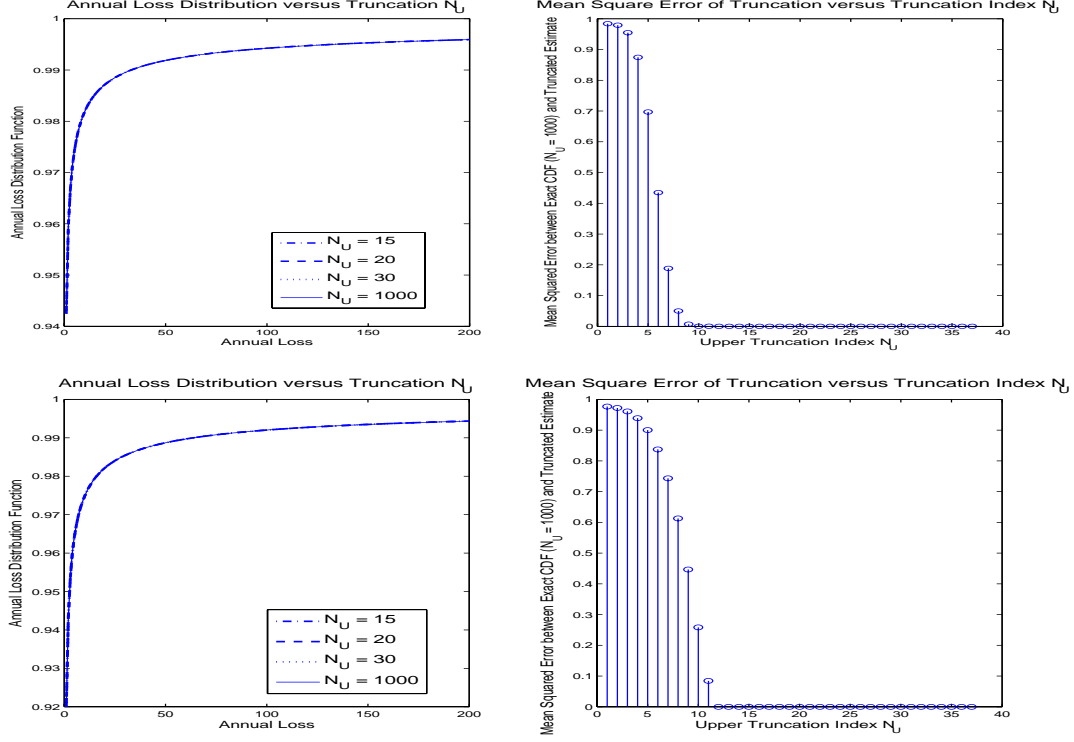


Figure 5: Comparison of high frequency Poisson-Levy LDA model empirical CDF versus Truncation Threshold  $N_U$  - Exact Expression ( $N_U = 1000$ ) (solid line). **Left** - Estimated CDF versus  $N_U$  **Right** - Mean Square Error versus  $N_U$

### 6.1. Example 1: Poisson- $\alpha$ -Stable Bivariate Risk Aggregation

In the case of two Poisson process LDA models one would obtain the result given by

$$\begin{aligned}
 f_{Z_T}(z) &= \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} \exp(-\lambda^{(1)} - \lambda^{(2)}) \frac{(\lambda^{(1)})^n (\lambda^{(2)})^m}{n!m!} \left[ \sqrt{\frac{\tilde{\gamma}_{nm}}{2\pi}} \frac{1}{(z - \tilde{\delta}_{nm})^{3/2}} \exp\left(-\frac{\tilde{\gamma}_{nm}}{2(z - \tilde{\delta}_{nm})}\right) \right] \times \mathbb{I}[z > \tilde{\delta}_{nm}] \\
 \tilde{\gamma}_{nm}^{0.5} &= \sum_{i=1}^n |\gamma_i^{(1)}|^{0.5} + \sum_{j=1}^m |\gamma_j^{(2)}|^{0.5} = n|\gamma^{(1)}|^{0.5} + m|\gamma^{(2)}|^{0.5}, \quad \tilde{\beta}_{nm} = 1 \\
 \tilde{\delta}_{nm} &= \sum_{i=1}^n \delta_i^{(1)} + \sum_{j=1}^m \delta_j^{(2)} + \tan \frac{\pi}{4} \left( \tilde{\gamma}_{nm} - \sum_{j=1}^n \gamma_j^{(1)} - \sum_{k=1}^m \gamma_k^{(2)} \right) \\
 &= n\delta^{(1)} + m\delta^{(2)} + \tan \frac{\pi}{4} \left( \left( n|\gamma^{(1)}|^{0.5} + m|\gamma^{(2)}|^{0.5} \right)^2 - n\gamma^{(1)} - m\gamma^{(2)} \right),
 \end{aligned} \tag{6.4}$$

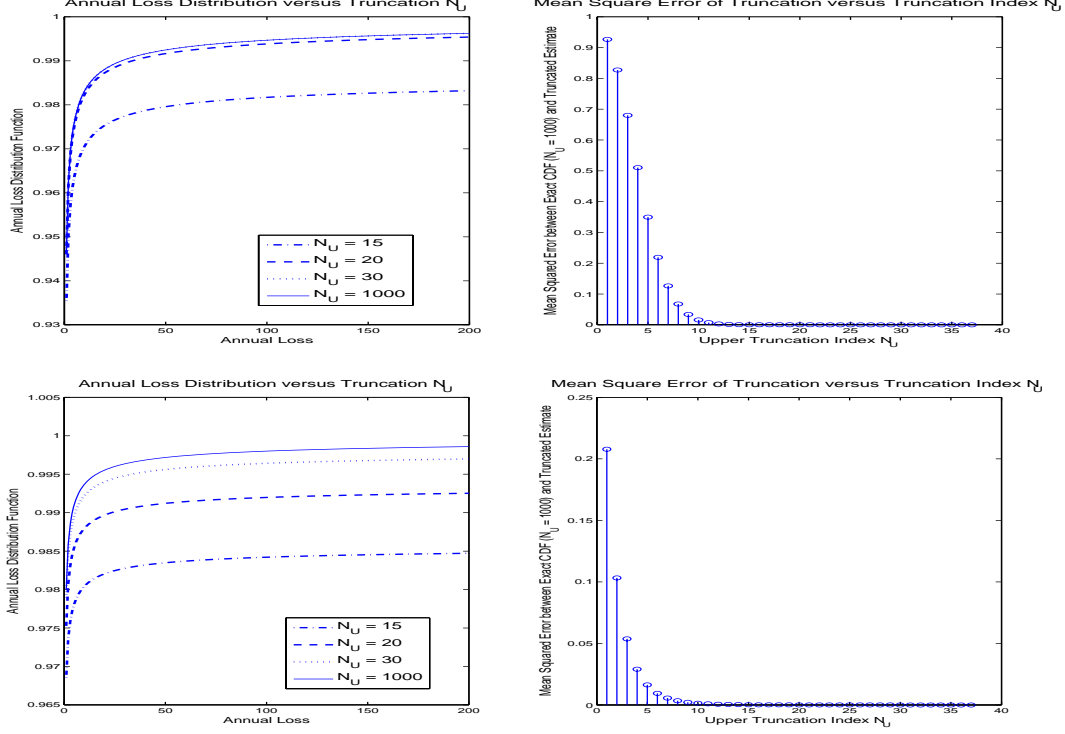


Figure 6: Comparison of high frequency Poisson-Levy LDA model empirical CDF versus Truncation Threshold  $N_U$  - Exact Expression ( $N_U = 1000$ ) (solid line). **Left** - Estimated CDF versus  $N_U$  **Right** - Mean Square Error versus  $N_U$

and  $f_{Z_T}(0) = \Pr(N_t^{(1)} + N_t^{(2)} = 0) = \exp(-\lambda_1 - \lambda_2)$ . The exact form of the insurance mitigated annual loss cumulative distribution function is also expressible in closed-form,

$$\Pr(Z_T < z) = F_{Z_T}(z) = \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} \exp(-\lambda^{(1)} - \lambda^{(2)}) \frac{(\lambda^{(1)})^n (\lambda^{(2)})^m}{n!m!} \operatorname{erfc} \left( \sqrt{\frac{\tilde{\gamma}_{nm}}{2(z - \tilde{\delta}_{nm})}} \right) \times \mathbb{I}[z > 0] + \exp(-\lambda_1 - \lambda_2). \quad (6.5)$$

## 7. Conclusions

This paper provided novel closed form analytic results for the several interesting classes of OpRisk LDA models based on the family of doubly stochastic  $\alpha$ -stable LDA models. Thereby providing models with the ability to capture the heavy tailed loss process typical of OpRisk whilst also providing analytic expressions for the compound process annual loss density and distributions as well as the aggregated compound process annual loss models.

Models for the annual loss process were developed in two scenarios, the first based on a loss process treated as continuous in time throughout the year with a stochastic intensity parameter, resulting in an inhomogeneous compound Poisson processes annually. The second scenario considered discretization

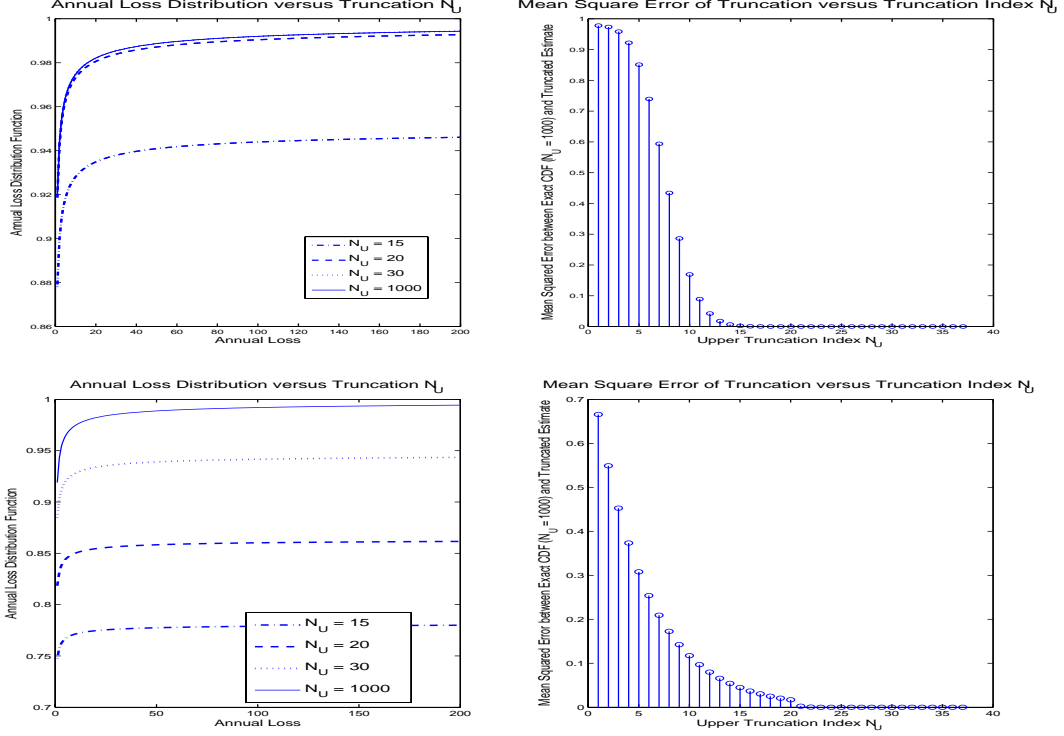


Figure 7: Comparison of high frequency Poisson-Levy LDA model empirical CDF versus Truncation Threshold  $N_U$  - Exact Expression ( $N_U = 1000$ ) (solid line). **Left** - Estimated CDF versus  $N_U$  **Right** - Mean Square Error versus  $N_U$

of the annual loss process into monthly increments with dependent time increments as captured by a binomial process with a stochastic probability of success changing annually.

The resulting LDA annual loss distribution exact representations derived were then analysed and compared to a standard Monte Carlo estimates obtained after simulating 1,000,000 annual years. The results demonstrated how accurate the exact expressions were in both low and high frequency scenarios even with a small number of terms in the expansions. In addition, we studied the mean square error of the stated distribution functions in each of the six models developed as a function of the truncation.

Finally, we demonstrated a novel multiple risk result showing how these results could be generalized to the multiple risk process setting to allow one to obtain expressions for the annual loss process of an institutions based on individual loss processes from one of the six doubly stochastic families of models developed.

## References

- [1] Adler, R.J., Feldman, R.E. and Taqqu, M.S. (1998). A practical guide to heavy tails: statistical techniques and applications. Birkhauser



Model	Exact Truncated ( $N_L = 1, N_U = 1,000$ )	Monte Carlo ( $T = 1,000,000$ )
Low Frequency Examples		
binomial-Levy	6.1 (sec)	10 (min)
binomial-Beta-Levy	1.2 (sec)	8 (min)
negative binomial-Levy	4.1 (sec)	4 (min)
negative binomial-Beta-Levy	1.2 (sec)	3 (min)
Poisson-Levy	1.2 (sec)	8 (min)
Poisson-Gamma-Levy	1.2 (sec)	7 (min)
High Frequency Examples		
binomial-Levy	6.1 (sec)	7 (min)
binomial-Beta-Levy	1.2 (sec)	7 (min)
negative binomial-Levy	4.2 (sec)	7 (min)
negative binomial-Beta-Levy	1.2 (sec)	5 (min)
Poisson-Levy	1.2 (sec)	10 (min)
Poisson-Gamma-Levy	1.2 (sec)	6 (min)

Table 3: Simulation times for exact solutions versus Monte Carlo estimates.

- [2] BIS (June, 2006). *International Convergence of Capital Measurement and Capital Standards*. Basel Committee on Banking Supervision, Bank for International Settlements, Basel. [www.bis.org](http://www.bis.org).
- [3] Chambers, J.M., Mallows, C.L. and Stuck, BW, (1976). A method for simulating stable random variables. *Journal of the American Statistical Association*, **71**(354), 340-344.
- [4] Chavez-Demoulin V., Embrechts P. and Nešlehová J. (2006). Quantitative models for Operational Risk: Extremes, dependence and aggregation. *Journal of Banking and Finance* **30**(10), 2635-2658.
- [5] Cruz M. (ed.) (2004). *Operational Risk Modeling and Analysis: Theory and Practice*. Risk Books, London.
- [6] Dutta K. and J. Perry (2006). A tale of tails: An empirical analysis of loss distribution models for estimating operational risk capital. Federal Reserve Bank of Boston, Working Papers No. 06-13.
- [7] Embrechts, P. and Schmidli, H. (1994). Modeling of extremal events in insurance and finance. *Mathematical Methods of Operations Research*, **39**(1), 1-34.
- [8] Fama, E.F. and Roll, R. (1968). Some properties of symmetric stable distributions. *Journal of the American Statistical Association*, **63**(323), 817-836.
- [9] Frachot A., Moudoulaud O. and Roncalli T. (2004). Loss distribution approach in practice. *The Basel Handbook: A Guide for Financial Practitioners*. Ong M. (ed), Risk Books.
- [10] Gelman A., Carlin J.B., Stern H.S. and Rubin D.B. (2004). Bayesian data analysis, CRC press.

- [11] McNeil, A.J. and Frey, R. (2000). Estimation of tail-related risk measures for heteroscedastic financial time series: an extreme value approach. *Journal of empirical finance*, **7**(3-4), 271-300.
- [12] Nolan J. (2007). *Stable Distributions - Models for Heavy-Tailed Data*, Birkhauser.
- [13] Willmot, G. and Panjer, H. (1985). Difference Equation Approaches in Evaluation of Compound Distributions. *Insurance: Mathematics and Economics*, **6**, 195-202
- [14] Peters G. and Sisson S. (2006). Bayesian Inference Monte Carlo Sampling and Operational Risk. *Journal Of Operational Risk*, **1**, No. 3.
- [15] Peters, G.W., Johansen, A.M. and Doucet, A. (2007). Simulation of the Annual Loss Distribution in Operational Risk via Panjer Recursions and Volterra Integral Equations for Value at Risk and Expected Shortfall Estimation. *The Journal of Operational Risk*, **2**(3), 29-58.
- [16] Peters, G.W., Shevchenko, P.V. and Wüthrich, M.V. (2009). Dynamic operational risk: modeling dependence and combining different sources of information. *The Journal of Operational Risk*, **4**(2), 69-104.
- [17] Peters, GW and Sisson, SA and Fan, Y., (2010). Likelihood-free Bayesian inference for  $\alpha$ -stable models. *Computational Statistics and Data Analysis*, to appear special issue on financial modeling.
- [18] G. W. Peters, P. V. Shevchenko and M. V. Wthrich (2009). Model uncertainty in claims reserving within Tweedie's compound Poisson models. *ASTIN Bulletin* **39**(1), 1-33.
- [19] Peters, G.W., Kannan, B.B., Lasscock, B., Mellen, C. and Godsill, S. (2010). Bayesian Cointegrated Vector Autoregression models incorporating Alpha-stable noise for inter-day price movements via Approximate Bayesian Computation. *Quantitative Finance Papers*, *arXiv. org*.
- [20] G.W. Peters, A.D. Byrnes and P.V. Shevchenko (2011). Impact of insurance for operational risk: Is it worthwhile to insure or be insured for severe losses? *Insurance: Mathematics and Economics* **48**, 287-303.
- [21] Samorodnitsky, G. and Taqqu, M.S., (1994). Stable non-Gaussian random processes: stochastic models with infinite variance. *Chapman & Hall/CRC*.
- [22] P. V. Shevchenko (2010). Implementing loss distribution approach for operational risk. *Applied Stochastic Models in Business and Industry* **26**(3), 277-307.
- [23] Zolotarev, V.M. and McFaden, H.H. (1986). One-dimensional stable distributions. *American Mathematical Society, Providence, RI*.

## 8. Appendix 1

Simulation of a univariate random variable from a general  $\alpha$ -stable model can be achieved based on the following steps.

1. Sample  $W \sim \text{Exp}(1)$
2. Sample  $U \sim \text{Uniform} \left[ \frac{-\pi}{2}, \frac{\pi}{2} \right]$
3. Apply transformation to obtain sample  $\bar{y}$  split into a composite:

$$(a) \quad \bar{y} = S_{\alpha,\beta} \frac{\sin \alpha(u+B_{\alpha,\beta})}{(\cos u)^{\frac{1}{\alpha}}} \left[ \frac{\cos(u-\alpha(u+B_{\alpha,\beta}))}{w} \right]^{\frac{1-\alpha}{\alpha}} \quad \text{if } \alpha \neq 1$$

$$(b) \quad \bar{y} = \frac{2}{\pi} \left[ \left( \frac{\pi}{2} + \beta u \right) \tan u - \beta \ln \frac{\frac{\pi}{2} w \cos u}{\frac{\pi}{2} + \beta u} \right] \quad \text{if } \alpha = 1$$

where  $S_{\alpha,\beta} = (1 + \beta^2 \tan^2(\frac{\pi\alpha}{2}))^{\frac{1}{\alpha}}$  and  $B_{\alpha,\beta} = \frac{1}{\alpha} \arctan(\beta \tan(\frac{\pi\alpha}{2}))$ . In this setting  $\bar{y}$  will be distributed by  $\mathcal{S}_\alpha(\beta, 1, 0)$

4. Apply the transformation to obtain sample  $y = \gamma \bar{y} + \delta$  with parameters  $\mathcal{S}_\alpha(\beta, \gamma, \delta)$

## 9. Appendix 2

**Proof of Theorem 7** Given the LDA structure in which the frequency is distributed generically according to  $N(t) \sim f(\cdot)$  and the severity model  $X_i(t) \sim S(0.5, 1, \gamma, \delta; 0)$ , the exact expressions for the tail asymptotic of the resulting infinite mixture is given by,

$$\Pr(Z(t) > z_q) \sim 2x^{-0.5} c_{0.5} \sum_{n=1}^{\infty} \Pr(N(t) = n) \tilde{\gamma}_n^{0.5} = C_{\text{Tail}} \sum_{n=1}^{\infty} W_n$$

where  $z_q$  is an upper tail quantile of the annual loss distribution and it will be assumed for simplicity that  $\tilde{\gamma}_n^{0.5} = \sum_{i=1}^n \gamma_i^{0.5} = n|\gamma|^{0.5}$ . These results allows us to determine a unique maximum of  $\frac{d}{dn} \log(W_n) = 0$  corresponding to the term  $n$  with the maximum contribution to the LDA compound processes tail probability for each model as follows:

### Poisson- $\alpha$ -Stable.

Under this model we have  $N(t) \sim Po(\lambda)$  which results in the term  $W_n = |\gamma|^{0.5} \exp(-\lambda) \frac{\lambda^n}{(n-1)!}$ . Now we approximate this function as being continuous in  $n$  and we differentiate and equate to zero,  $\frac{d}{dn} \log(W_n) = 0$ , to find the mode of the terms in the series being summed. We first apply Stirling's approximation to  $n! = \Gamma(n+1) \approx \sqrt{2\pi n} n^{n+0.5} e^{-n}$  to give

$$\begin{aligned} \log W_n &= 0.5 \log(|\gamma|) - \lambda + n \log(\lambda) - \log((n-1)!) \\ &= 0.5 \log(|\gamma|) - \lambda + n \log(\lambda) - \log(\Gamma(n)) \\ &\approx 0.5 \log(|\gamma|) - \lambda + n \log(\lambda) - (0.5 \log(2\pi) + (n+0.5) \log(n) - n) \\ \Rightarrow \quad \frac{d}{dn} [\log W_n] &= \log(\lambda) - \log(n) - \frac{1}{2n} \\ \Rightarrow \quad n \text{ large } \frac{d}{dn} [\log W_n] &\approx \log(\lambda) - \log(n). \end{aligned}$$

Equating this to zero and solving for  $n$ , tells us that if  $\lambda \geq 1$ , then the maximum occurs at approx  $n = \lambda$ .

If  $\lambda < 1$ , then max occurs at approx  $n = 1$ .  $\square$

#### Doubly Stochastic Poisson-Gamma- $\alpha$ -Stable.

Under this model we have  $N(t) \sim Po(\lambda)$  and  $\lambda \sim \Gamma(\alpha, \beta)$  which results in the term  $W_n = |\gamma|^{0.5} \frac{(\alpha+n-1)!}{(\alpha-1)!n!} \left(\frac{\beta}{1+\beta}\right)^\alpha \left(\frac{1}{1+\beta}\right)^n$ .

Now we approximate this function as being continuous in  $n$  and we differentiate and equate to zero,  $\frac{d}{dn} \log(W_n) = 0$ , to find the mode of the terms in the series being summed. Again we apply Stirling's approximation  $\Gamma(n+1) \approx \sqrt{2\pi}n^{n+0.5}e^{-n}$  to give

$$\begin{aligned} \log W_n &= 0.5 \log(|\gamma|) + \log(\Gamma(\alpha+n)) - \log(\Gamma(\alpha)) - \log(\Gamma(n+1)) + \alpha \log(\beta) - (\alpha+n) \log(\beta+1) \\ &\approx 0.5 \log(|\gamma|) + 0.5 \log(2\pi) + (n+\alpha-0.5) \log(n+\alpha-1) - \alpha + 1 - \log(\Gamma(\alpha)) - 0.5 \log(2\pi) \\ &\quad - (n+0.5) \log(n) + \alpha \log(\beta) - (\alpha+n) \log(\beta+1) \\ \Rightarrow \frac{d}{dn} [\log W_n] &= \log(n+\alpha-1) + \frac{n}{(n+\alpha-1)} + \frac{(\alpha-0.5)}{(n+\alpha-1)} - \log(n) - 1 - \frac{1}{2n} - \log(\beta+1) \end{aligned}$$

Equating this to zero and solving for  $n$ , ie. solving for the roots  $\log(2) - \log(n) - \frac{1}{2n} = 0$  can be done by an explicit search, which produces  $n = 2$ .  $\square$

#### negative binomial- $\alpha$ -Stable.

Under this model we have  $N(t) \sim NB(r, p)$  which results in the term

$W_n = |\gamma|^{0.5} \frac{(n+r-1)!}{(n+r-1-n)!(n)!} (1-p)^r (p)^n$ . Now we approximate this function as being continuous in  $n$  and we differentiate and equate to zero,  $\frac{d}{dn} \log(W_n) = 0$ , to find the mode of the terms in the series being summed. We first apply Stirling's approximation to  $n! = \Gamma(n+1) \approx \sqrt{2\pi}n^{n+0.5}e^{-n}$  to give

$$\begin{aligned} \log W_n &= 0.5 \log(|\gamma|) + \log((n+r-1)!) - \log((r-1)!) - \log((n)!) + r \log(1-p) + n \log(p) \\ &\approx 0.5 \log(|\gamma|) + 0.5 \log(2\pi) + (n+r-0.5) \log(n+r-1) - (n+r-1) - \log((r-1)!) \\ &\quad - 0.5 \log(2\pi) - (n+0.5) \log(n) + n+r \log(1-p) + n \log(p) \\ \Rightarrow \frac{d}{dn} [\log W_n] &= \log(n+r-1) + \frac{n}{n+r-0.5} + \frac{r-0.5}{n+r-1} - 1 - \log(n) - \frac{1}{2n} + \log(p) \end{aligned}$$

Then for a given set of values  $r$  and  $p$  in the model we can equate to zero this expression and solve for  $n$ .  $\square$

#### Doubly Stochastic negative binomial- Beta- $\alpha$ -Stable.

Under this model we have  $N(t) \sim NB(r, p)$  and  $p \sim Be(\alpha, \beta)$  which results in the term  $W_n = |\gamma|^{0.5} \frac{(n+r-1)!p^{(\alpha+r-1)}B(\alpha+rn, \beta+n)}{(r-1)!n!p^{(\alpha+r-1)}B(\alpha, \beta)}$ . We will now utilise the well known identity for the Beta function with respect to Gamma functions and a Stirling Approximation given by

$$B(x, y) \approx \sqrt{2\pi} \frac{x^{x-0.5}y^{y-0.5}}{(x+y)^{x+y-0.5}}$$

Now we approximate this function as being continuous in  $n$  and we differentiate and equate to zero,

$\frac{d}{dn} \log(W_n) = 0$ , to find the mode of the terms in the series being summed.

$$\begin{aligned}
\log W_n &= 0.5 \log(|\gamma|) + \log(\Gamma(n+r)) + (\alpha+r-1) \log(p) + \log(B(\alpha+rn, \beta+n)) \\
&\quad - \log(\Gamma(r)) - \log(\Gamma(n+1)) - (\alpha+rn-1) \log(p) - \log(B(\alpha, \beta)) \\
&\approx 0.5 \log(|\gamma|) + 0.5 \log(2\pi) + (n+r+0.5) \log(n+r) - n-r + (\alpha+r-1) \log(p) \\
&\quad + 0.5 \log(2\pi) + (\alpha+rn-0.5) \log(\alpha+rn) + (\beta+n-0.5) \log(\beta+n) \\
&\quad - (\alpha+rn+\beta+n-0.5) \log(\alpha+rn+\beta+n) - \log(\Gamma(r)) - 0.5 \log(2\pi) \\
&\quad - (n+1+0.5) \log(n+1) + n+1 - (\alpha+rn-1) \log(p) - \log(B(\alpha, \beta)) \\
\Rightarrow \frac{d}{dn} [\log W_n] &= \log(n+r) + \frac{(n+r+0.5)}{(n+r)} + r \log(\alpha+rn) + \frac{r(\alpha+rn-0.5)}{(\alpha+rn)} + \log(\beta+n) \\
&\quad + \frac{(\beta+n-0.5)}{(\beta+n)} - (r+1) \log(\alpha+rn+\beta+n) - \frac{(r+1)(\alpha+rn+\beta+n-0.5)}{(\alpha+rn+\beta+n)} \\
&\quad - \log(n+1) - \frac{(n+1+0.5)}{(n+1)} - r \log(p)
\end{aligned}$$

Then for a given set of values  $r$  and  $p$  in the model we can equate to zero this expression and solve for  $n$ .

□