

# Operational Risk Modelling in Insurance and Banking

**Ognjen Vukovic**

Department for Finance, University of Liechtenstein, Vaduz, Liechtenstein  
Email: [ognjen.vukovic@uni.li](mailto:ognjen.vukovic@uni.li), [oggyvukovich@gmail.com](mailto:oggyvukovich@gmail.com)

Received 20 June 2015; accepted 29 August 2015; published 1 September 2015

Copyright © 2015 by author and Scientific Research Publishing Inc.  
This work is licensed under the Creative Commons Attribution International License (CC BY).  
<http://creativecommons.org/licenses/by/4.0/>



Open Access

---

## Abstract

The author of the presented paper is trying to develop and implement the model that can mimic the state of the art models of operational risk in insurance. It implements generalized Pareto distribution and Monte Carlo simulation and tries to mimic and construct operational risk models in insurance. At the same time, it compares lognormal, Weibull and loglogistic distribution and their application in insurance industry. It is known that operational risk models in insurance are characterized by extreme tails, therefore the following analysis should be conducted: the body of distribution should be analyzed separately from the tail of the distribution. Afterwards the convolution method can be used to put together the annual loss distribution by combining the body and tail of the distribution. Monte Carlo method of convolution is utilized. Loss frequency in operational risk in insurance and overall loss distribution based on copula function, in that manner using student-t copula and Monte Carlo method are analysed. The aforementioned approach represents another aspect of observing operational risk models in insurance. This paper introduces: 1) Tools needed for operational risk models; 2) Application of R code in operational risk modeling; 3) Distributions used in operational risk models, specializing in insurance; 4) Construction of operational risk models.

## Keywords

Insurance, Operational Risk, Monte Carlo, Statistical Distributions, Modeling, Copula, Convolution, Loss Frequency, Severity

---

## 1. Introduction

Operational risk is defined according to Basel II (Nicolas & Firzli, 2011) as well as according to European Solvency II which adopted for insurance industry is defined in the following way (Nicolas & Firzli, 2011).

Operational risk is the risk of change in value caused by the fact that actual losses, incurred for inadequate or failed internal processes, people and systems, or from external events (including legal risk), and differs from the

expected losses.

In order to analyse operational risk in insurance, Solvency II Directive (Mittnik, 2011) must be discussed. Solvency II Directive is an EU Directive that codifies and harmonises the EU insurance regulation. This concerns the amount of capital that EU insurance companies must hold to reduce the risk of insolvency. Solvency II (Mittnik, 2011) is called “Basel for insurers” (Nicolas & Firzli, 2011). Solvency II is somewhat similar to the banking regulations of Basel II. The proposed Solvency II framework has three main areas (pillars) (Accords, 2006):

- Pillar 1—quantitative requirements
- Pillar 2—requirements for the governance and risk management of insurers and their supervision
- Pillar 3—disclosure and transparency requirements

In order to analyse the operational risk in insurance (CEA—Groupe Consultatif, 2005), the attention towards pillar 2 and pillar 1 should be directed. Ernst & Young reports that most of the insurance companies in 2016 will manage to implement the Solvency II requirement until January 2016 (PWC Financial Services Regulatory Practice, 2014). One of the major new features of Solvency II is that insurance companies must now devote a portion of their equity to covering their exposure to operational risks. There are approaches to calculate the capital requirement: a standard and more advance approach. The advanced approach uses an internal model of risk that corresponds to the company’s real situation. Quantitative impact study (QIS 5) has encouraged insurance companies to adopt the internal model by structuring the standard approach such that it uses up much more equity (Solvency—European Commission, 2012). In order to analyse the operational risk in this frame, the following assumption will be made: risk will be divided between frequency and severity risk (Doerig, 2000). They will be modeled by Loss Distribution approach (Power, 2005). Severity risk represents the risk of large but rare losses. Bayesian networks are used to model severity risk.

### 1.1. Bayesian Networks

Bayesian networks are defined as a probabilistic graphical model that represents a set of random variables and their conditional dependencies via a directed acyclic graph (DAG). For example, a Bayesian network could represent the probabilistic relationships between the cause and result. If we are aware of the causes, the probabilities of results can be calculated.

In order to define the Bayesian network, the following definition will be used.  $X$  is a Bayesian network with respect to  $G$  if its joint probability density function (with respect to a product measure) can be written as a product of the individual density functions, conditional on their parent variables:

$$p(x) = \prod_{v \in V} p(x_v | x_{pa(v)}) \quad (1)$$

where  $pa(v)$  is the set of parents of  $v$ .

The probability of any member of a joint distribution can be calculated from the conditional probabilities using the chain rule, taking into consideration at the same topological ordering of  $X$ .

$$P(X_1 = x_1, \dots, X_n = x_n) = \prod_{v=1}^n P(X_v = x_v | X_{v+1} = x_{v+1}, \dots, X_n = x_n) \quad (2)$$

Equation (2) can be written as:

$$P(X_1 = x_1, \dots, X_n = x_n) = \prod_{v=1}^n P(X_v = x_v | X_j = x_j) \quad (3)$$

(for each  $X_j$  which is a parent of  $X_v$ ).

$X$  is a Bayesian network with respect to  $G$  if it satisfies the local Markov property, each variable is conditionally independent of its non-descendants given its parent variables.

$$X_v \perp\!\!\!\perp X_{V \setminus de(v)} \mid X_{pa(v)} \quad \text{for all } v \in V \quad (4)$$

where  $de(v)$  is the set of descendants and  $V \setminus de(v)$  is the set of non-descendants of  $v$ .

The aforementioned thing can also be expressed in terms similar to the first definition, as:

$$P(X_v = x_v | X_i = x_i) \tag{5}$$

For each  $X_i$  which is not a descendant of  $X_v$  is

$$P(X_v = x_v | X_j = x_j) \text{ For each } X_j \text{ which is a parent of } X_v.$$

Note that the set of parents is a subset of the set of non-descendants because the graph is acyclic.

### 1.2. Developing Bayesian Networks

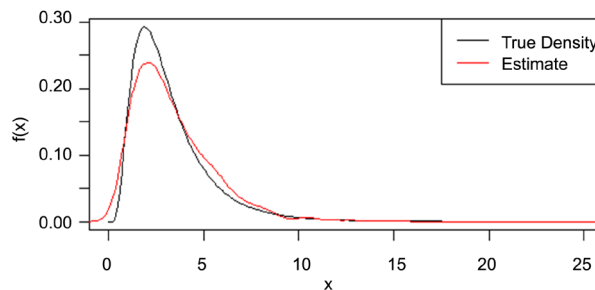
To develop a Bayesian network, we often first develop a *DAG* (directed acyclic graph)  $G$  such that we believe  $X$  satisfies the local Markov property with respect to  $G$ . Sometimes this is done by creating a casual *DAG*. We then ascertain the conditional probability distribution of each variable given its parents in  $G$ . In many cases and in particular in the case where the variables are discrete, we define the joint distribution of  $X$  to be the product of these conditional distributions, then  $X$  is a Bayesian network with respect to  $G$ .

In order to model loss severity, since it can be difficult to model operational risk losses of a risk class using only one probability distribution, we analyse severity at two levels: body and tail of the distribution delimited by a high threshold value  $u$ .

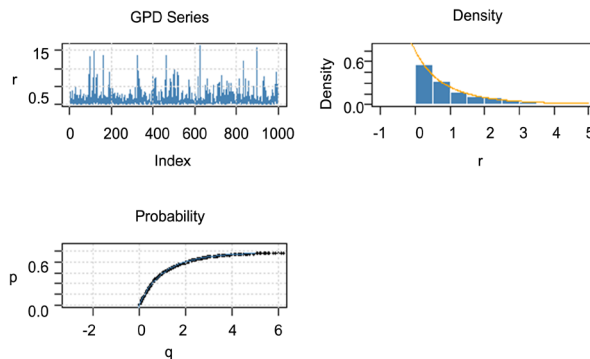
The central body can be modeled by using a parametric distribution. One of the distributions that can be used is lognormal. Body is usually estimated on internal data, since the sample size is sufficient below the threshold  $u$ .

**Figure 1** demonstrates the body of the distribution, it is assumed that the body of the distribution embraces lognormal value, therefore we assume the lognormal distribution. In order to analyse the tail, it can be modeled applying extreme value theory (Ebrechts, Kluppelberg, & Mikosch, 1997) the distribution above the threshold, we are talking about Generalized Pareto distribution (GPD). Tail is usually estimated on internal data integrated with external data and scenario generated data above  $u$ . Data are real scarce above threshold  $u$ .

Generalized Pareto distribution is used to capture the tail of severity distribution. **Figure 2** shows the Genera-



**Figure 1.** Lognormal distribution.



**Figure 2.** Generalized Pareto distribution.

lized Pareto distribution, its density as well as its cumulative function.

The commonly used approach to quantify operational risk is the Loss Distribution Approach (Embrechts, Kluppelberg, & Mikosch, 1997) where frequency and severity of operational risk losses are modeled separately. The yearly potential loss  $S$  is based on the sum of the yearly losses  $S_j$  related to the  $J$  risk classes.

The yearly potential loss  $S_j$  related to risk class  $j$ , is affected by two sources of uncertainty:

- The number of losses  $N_j$  in one year time horizon
- The impact of each single loss  $X_{ij}$

If the following procedure is to be applied to mathematical modeling of operational risk in insurance, the only thing that should be changed is the distribution that it is to be used. Therefore, the methods of copula and convolution will be explained as well as the possible distribution that could be applied in insurance.

### 1.3. Convolution

In mathematics, convolution is a mathematical operation on two functions  $f$  and  $g$ , producing a third function that is typically viewed as a modified version of one of the original functions, giving the area overlap between the two functions as a function of the amount that one of the original functions is translated (Bracewell, 1986; Damelin & Miller, 2011). Convolution is presented in the following manner:

It is defined as the integral of the product of the two functions after one is reversed and shifted. It is a particular kind of integral transform:

$$(f * g)(t) \stackrel{\text{def}}{=} \int_{-\infty}^{\infty} f(\tau)g(t-\tau)d\tau = \int_{-\infty}^{\infty} f(t-\tau)g(\tau)d\tau. \quad (6)$$

Although the symbol  $t$  is used above, it need not represent the time domain. However formula can be interpreted as a weighted average of the function  $f(\tau)$  at the moment  $t$  where the weighting is given by  $g(-\tau)$  simply shifted by amount  $t$ . As  $t$  changes, the weighting function emphasizes different part of the input function (Bracewell, 1986; Damelin & Miller, 2011).

### 1.4. Copula

A copula is a multivariate probability distribution for which the marginal probability distribution of each variable is uniform. Copulas are used to describe the dependence between random variables.

One important theorem for copulas is Sklar's theorem. We will now define the copula:

Copula (Nelsen, 1999) is a multivariate distribution function,  $C$ , with marginal functions distributed uniformly in  $[0, 1]$  ( $U(0,1)$ ) such that

- 1)  $C : [0, 1]^n \rightarrow [0, 1]$
- 2)  $C$  has marginal functions  $C_i$  such that  $C_i(u) = C(1, \dots, 1, u, 1, \dots, 1) = u \quad \forall u \in [0, 1]$

Sklar's theorem (Sklar, 1959) provides the theoretical foundation for the application of copula. Sklar's theorem states that every multivariate cumulative distribution function

$$H(x_1, \dots, x_d) = \mathbb{P}[X_1 \leq x_1, \dots, X_d \leq x_d] \quad (7)$$

Of a random vector  $(X_1, X_2, \dots, X_d)$  with marginal  $F_i(x) = \mathbb{P}[X_i \leq x]$  and it can be written as:

$$H(x_1, \dots, x_d) = C(F_1(x_1), \dots, F_d(x_d)) \quad (8)$$

where  $C$  is copula.

The copula that will be used is Gaussian copula and Gumbel copula. They will be presented shortly.

The Gaussian copula:

Gaussian copula (Nelsen, 1999) is a distribution over the unit cube  $[0, 1]^d$ . It is constructed from a multivariate normal distribution over  $R^d$  using the probability integral form.

The Gaussian copula for  $R \in R^{d \times d}$ , it can be written with parameter matrix  $R$  can be written as:

$$C_R^{\text{Gauss}}(u) = \Phi_R(\Phi^{-1}(u_1), \dots, \Phi^{-1}(u_d)) \quad (9)$$

where  $\Phi^{-1}$  is the inverse cumulative distribution function of a standard normal and  $\Phi_R$  is the joint cumulative distribution function of a multivariate normal distribution with mean vector zero and covariance matrix equal to the correlation matrix  $R$ .

The density matrix can be written as

$$c_R^{\text{Gauss}}(\mathbf{u}) = \frac{1}{\sqrt{\det R}} \exp \left( -\frac{1}{2} \begin{pmatrix} \Phi^{-1}(u_1) \\ \vdots \\ \Phi^{-1}(u_d) \end{pmatrix}^T \cdot (R^{-1} - I) \cdot \begin{pmatrix} \Phi^{-1}(u_1) \\ \vdots \\ \Phi^{-1}(u_d) \end{pmatrix} \right) \quad (10)$$

where  $I$  is the identity matrix.

Other famous copulas are Archimedean copulas.

### 1.5. Archimedean Copulas

Archimedean copulas (Nelsen, 1999) are an associative class of copulas. Archimedean copulas are popular because they allow modeling dependence in arbitrarily high dimensions with only one parameter, governing the strength of dependence.

A copula  $C$  is called Archimedean if it admits the representation

$$C(u_1, \dots, u_d; \theta) = \psi^{[-1]}(\psi(u_1; \theta) + \dots + \psi(u_d; \theta); \theta) \quad (11)$$

where  $\psi: [0, 1] \times \Theta \rightarrow [0, \infty)$  is a continuous, strictly decreasing and convex function such that  $\psi(1; \theta) = 0$ .  $\theta$  is a parameter within some parameter space  $\Theta$ .  $\psi$  is the so-called generator function and  $\psi^{[-1]}$  is the pseudo-inverse defined by:

$$\psi^{[-1]}(t; \theta) = \begin{cases} \psi^{-1}(t; \theta) & \text{if } 0 \leq t \leq \psi(0; \theta) \\ 0 & \text{if } \psi(0; \theta) \leq t \leq \infty \end{cases} \quad (12)$$

Moreover, the above formula for  $C$  yields a copula for  $\psi^{-1}$  if and only if  $\psi^{-1}$  is  $d$ -monotone on  $[0, \infty)$ . That is, if it is  $d-2$  times differentiable and the derivatives satisfy

$$(-1)^k \psi^{-1, (k)}(t; \theta) \geq 0 \quad (13)$$

for all  $t \geq 0$  and  $k = 0, 1, \dots, d-2$  and  $(-1)^{d-2} \psi^{-1, (d-2)}(t; \theta)$  is nonincreasing and convex.

The copula that we will also introduce is Gumbel copula.

It has the following bivariate form:

$$C_\theta(u, v) = \exp \left( -\left( (-\log(u))^\theta + (-\log(v))^\theta \right)^{1/\theta} \right) \quad (14)$$

After giving the definition of two most important techniques of aggregating operational risk, we will introduce the distributions that we will be using as well as the definition of value at risk and expected shortfall.

### 1.6. Statistical Distributions

To analyse severity in that sense pertaining to body of the distribution, following distributions can be used, given their densities:

Lognormal distribution density (Hazewinkel, 2001):

$$f_x(x; \mu, \sigma) = \frac{1}{x\sigma\sqrt{2\pi}} e^{-\frac{(\ln x - \mu)^2}{2\sigma^2}}, \quad x > 0 \quad (15)$$

Weibull density distribution:

The probability density function of a Weibull random variable (Hazewinkel, 2001) is:

$$f(x; \lambda, k) = \begin{cases} \frac{k}{\lambda} \left( \frac{x}{\lambda} \right)^{k-1} e^{-(x/\lambda)^k} & x \geq 0, \\ 0 & x < 0, \end{cases} \quad (16)$$

where  $k > 0$  is the *shape parameter* and  $\lambda > 0$  is the *scale parameter* of the distribution. Its complementary cu-

mulative distribution function is a stretched exponential function. The Weibull distribution is related to a number of other probability distributions; in particular, it interpolates between the exponential distribution ( $k = 1$ ) and the Rayleigh distribution ( $k = 2$  and  $\lambda = \frac{\sqrt{2}}{k}$ )

Loglogistic probability density function is defined in the following way:

$$f(x; \alpha, \beta) = \frac{(\beta/\alpha)(x/\alpha)^{\beta-1}}{(1+(x/\alpha)^\beta)^2} \tag{17}$$

The parameter  $\alpha > 0$  is a scale parameter and is also the median of the distribution. The parameter  $\beta > 0$  is a shape parameter. The distribution is unimodal when  $\beta > 1$  and its dispersion decreases as  $\beta$  increases.

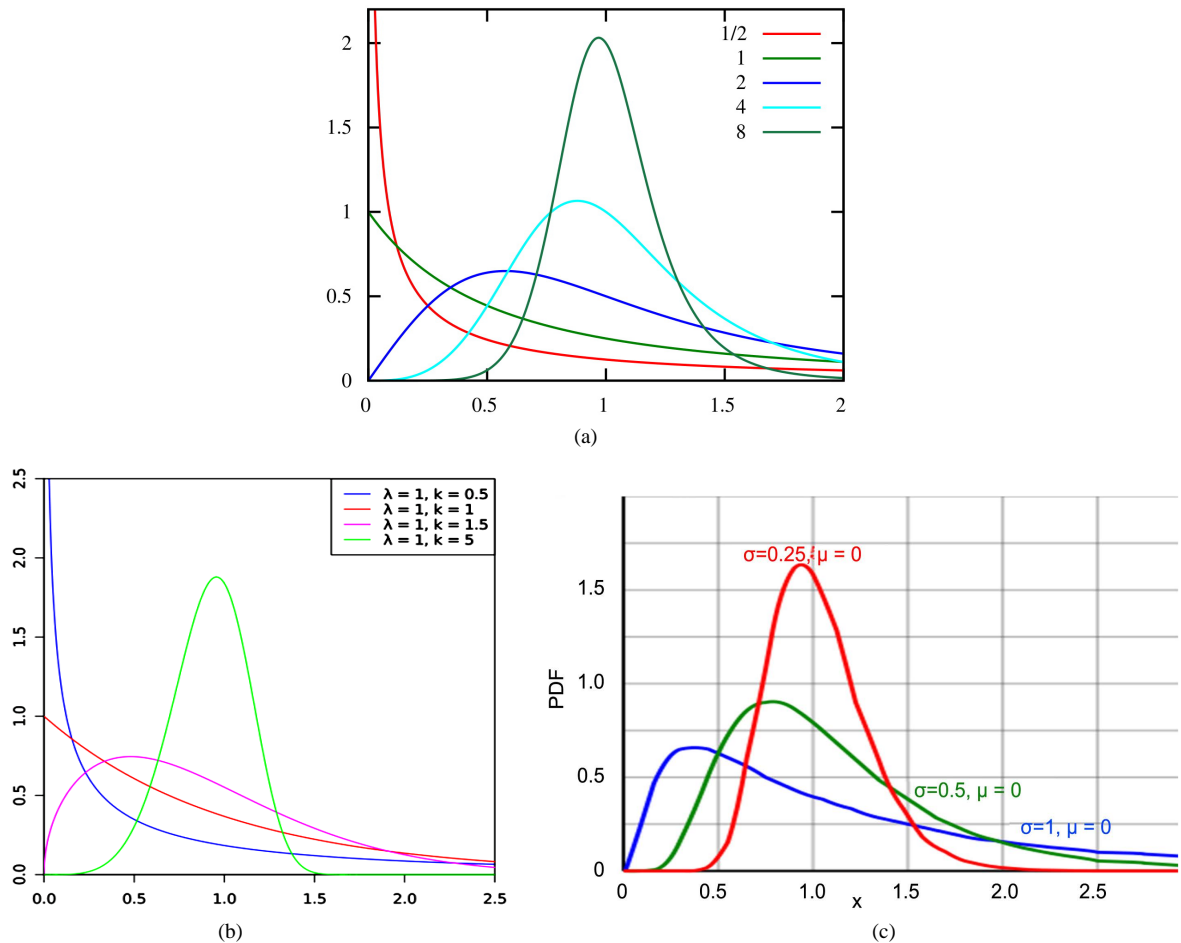
The following functions are shown in the graphs below:

**Figure 3** shows probability density functions of Loglogistic, Weibull and lognormal distributions.

At the same time, we will introduce Champernowne distribution (Hazewinkel, 2001) that we will also be using, its probability density function is given in the graph below:

$$f(y; \alpha, \lambda, y_0) = \frac{n}{\cosh[\alpha(y - y_0)] + \lambda}, \quad -\infty < y < \infty \tag{18}$$

where  $\alpha, \lambda, y_0$  are positive parameters, and  $n$  is the normalizing constant, which depends on the parameters.



**Figure 3.** Loglogistic, Weibull and lognormal probability density functions respectively.

When obtaining copula, Monte Carlo method will be used, so we will introduce it shortly.

One is interested in the expectation of a response function  $g: R^d \rightarrow R$  applied to some random vector  $(X_1, \dots, X_d)$ . If we denoted the cumulative distribution function of this random vector with  $H$ , the quantity of interest can thus be written as:

$$\mathbb{E}[g(X_1, \dots, X_d)] = \int_{\mathbb{R}^d} g(x_1, \dots, x_d) dH(x_1, \dots, x_d) \quad (19)$$

If  $H$  is given by a copula model, example:

$$H(x_1, \dots, x_d) = C(F_1(x_1), \dots, F_d(x_d)) \quad (20)$$

Then expectation can be written as:

$$\mathbb{E}[g(X_1, \dots, X_d)] = \int_{[0,1]^d} g(F_1^{-1}(u_1), \dots, F_d^{-1}(u_d)) dC(u_1, \dots, u_d). \quad (21)$$

In case the copula  $C$  is absolutely continuous,  $C$  has density  $c$ , the equation can be written as

$$\mathbb{E}[g(X_1, \dots, X_d)] = \int_{[0,1]^d} g(F_1^{-1}(u_1), \dots, F_d^{-1}(u_d)) c(u_1, \dots, u_d) du_1 \dots du_d \quad (22)$$

If copula and margins are known (or if they have been estimated), then the following Monte Carlo algorithm can be used:

- 1) Draw a sample  $(U_1^k, \dots, U_d^k) \sim C, (k = 1, \dots, n)$  of size  $n$  from the copula  $C$
- 2) By applying the inverse marginal cdf's, produce a sample of  $(X_1, \dots, X_d)$  by setting
 
$$(X_1^k, \dots, X_d^k) = (F_1^{-1}(U_1^k), \dots, F_d^{-1}(U_d^k)) \sim H, (k = 1, \dots, n)$$
- 3) Approximate  $\mathbb{E}[g(X_1, \dots, X_d)]$  by its empirical value:

$$\mathbb{E}[g(X_1, \dots, X_d)] \approx \frac{1}{n} \sum_{k=1}^n g(X_1^k, \dots, X_d^k) \quad (23)$$

## 1.7. VAR and Expected Shortfall

After having introduced all the necessary tools for operational risk model in insurance, we will just provide the VAR and expected shortfall definition (Longin, 1997) and then we are going to present the modeling results.

$$\text{VaR}_\alpha(L) = \inf \{l \in \mathbb{R} : P(L > l) \leq 1 - \alpha\} = \inf \{l \in \mathbb{R} : F_L(l) \geq \alpha\} \quad (24)$$

Confidence level  $\alpha \in (0, 1)$ , the VaR (Longin, 1997) of the portfolio at the confidence level  $\alpha$  is given by the smallest number  $l$  such that the probability that the loss  $L$  exceeds  $l$  is at most  $(1 - \alpha)$ . Mathematically, if  $L$  is the loss of a portfolio, then  $\text{VaR}_\alpha(L)$  is the level  $\alpha$  quantile.

In order to model tails, Generalized Pareto distribution will be used.

The standard cumulative distribution function (cdf) of the GPD is defined by

$$F_\xi(z) = \begin{cases} 1 - (1 + \xi z)^{-1/\xi} & \text{for } \xi \neq 0, \\ 1 - e^{-z} & \text{for } \xi = 0. \end{cases} \quad (25)$$

where the support is  $z \geq 0$  for  $\xi \geq 0$  and  $0 \leq z \leq -1/\xi$  for  $\xi < 0$ .

$$f_\xi(z) = \begin{cases} (\xi z + 1)^{-\xi-1} & \text{for } \xi \neq 0, \\ e^{-z} & \text{for } \xi = 0. \end{cases} \quad (26)$$

### Expected shortfall

Expected shortfall is defined as

$$ES_\alpha = \frac{1}{\alpha} \int_0^\alpha \text{VaR}_\gamma(X) d\gamma \quad (27)$$

Or equivalently can be written as:

$$ES_\alpha = -\frac{1}{\alpha} \left( E \left[ X 1_{\{X \leq x_\alpha\}} \right] + x_\alpha (\alpha - P[X \leq x_\alpha]) \right) \quad (28)$$

where

$$x_\alpha = \inf \{x \in \mathbb{R} : P(X \leq x) \geq \alpha\} \quad (29)$$

is the lower  $\alpha$ -quantile and  $1_A(x) = \begin{cases} 1 & \text{if } x \in A \\ 0 & \text{else} \end{cases}$  is the indicator function.

## 2. Experimental Results

After having introduced all the necessary tools, the experimental results will be presented. As we are considering operational risk models in insurance, we will use the following tools. The procedure is the following. As we don't have enough data, we will use random number generator, let it simulate the events, but at the same defining that extreme events don't occur frequently. Afterwards we will implement the distribution for body and tail of the severity of distribution. Frequency of events was modeled by using Poisson distribution. A standard choice to estimate annual frequency of operational risk loss is performed by using Poisson distribution. On the basis of extreme value theory, the distribution function of loss data above a high threshold  $u$  is supposed to follow a Generalized Pareto distribution. Lognormal distribution will be used for modeling body of severity distribution. Afterwards, we use the convolution method using frequency and severity distribution and we obtain annual loss distribution. Convolution is performed by using Monte-Carlo method.

Following steps are performed:

- 1) Extract one random number  $(n_j)$  from frequency distribution
- 2) Extract  $n_j$  random numbers  $(x_{ij})$  from severity distribution
- 3) Obtain a possible figure of yearly op. loss  $s_j = \sum_j x_{ij}$

Repeat  $n$  times and we obtain annual loss distribution. In order to obtain overall annual loss distribution, we use the following procedure:

- Extract one random value  $\bar{u} = (u_1, \dots, u_J)$  from copula function
- Extract  $n$  random numbers from loss distributions related to  $J$  risk classes
- Sum yearly losses to obtain overall annual loss distribution

If the real data is used, then to estimate the parameters maximum likelihood, method of moments, probability weighted method of moments should be used.

The simulation results are given below:

```

$distribution
[1] "logNormal"

$fit
$fit$par
[1] 2.500502 1.994844

$fit$value
[1] 33653.66

$fit$counts
function gradient
      11      11

$fit$convergence
[1] 0

$fit$message
[1] "CONVERGENCE: REL_REDUCTION_OF_F <= FACTR*EPSMCH"

```

The body of distribution is considered to be lognormal. The first fit value is the mean and the other one is standard deviation. The gradient of the function demonstrates where the function is 0.

In **Figure 4**, optimisation values are given for different theta 1 and theta 2 that represent optimal values.



Goodness of fit can also be performed, if the real data is used. **Figure 7** shows negative log-likelihood method for fitting log-normal distribution. In that direction, Kolmogorov-Smirnov and Anderson-Darling test statistics are used, but this is left for further research.

To estimate the tail of the severity distribution, we have to set a high threshold  $u$ : the parameters are estimated on excess over threshold using probability weighted moments, maximum likelihood method etc. Mean excess function is defined in the following way:

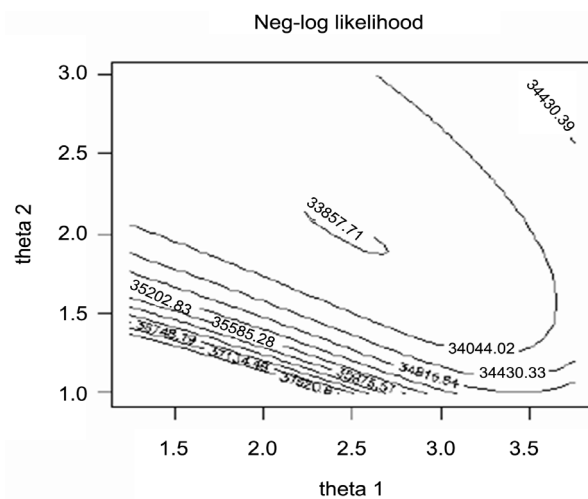
$$e(u) = E(x - u | X > u) \rightarrow e(u) = \frac{\beta + \xi u}{1 - \xi} \tag{30}$$

What should be noticed is that if the mean excess function for Generalized Pareto distribution is a linear function of  $u$ , if the empirical mean excess function is a straight line above a threshold, it is an indication that the excess over  $u$  follows a generalized Pareto distribution, which is shown in the **Figure 8**.

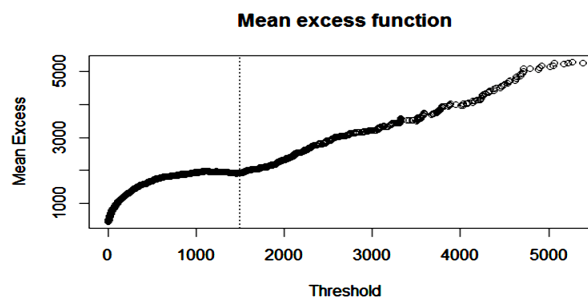
It is obvious that **Figure 5** demonstrates that the tail of severity distribution is following Generalized Pareto distribution. QQ plots of the Generalized Pareto distribution using maximum likelihood method and probability weighted method of moments.

Q-Q plots in **Figure 6** demonstrate that the tail of severity distribution is characterised by generalized Pareto distribution. Results are the following. Using the aforementioned methods, the following results are obtained:

xi	beta	
0.4878122	1025.8805247	for maximum likelihood method
xi	beta	
0.4533816	1049.2300744	for probability weighted method of moment



**Figure 4.** Neg-log likelihood for fitting log-normal distribution with random number generator.



**Figure 5.** Mean excess function.

To analyse the loss frequency, we assume Poisson process and obtain the following results:

**Figure 7** shows loss frequency distribution for Poisson case.

After having introduced the severity and loss frequency distribution, we use the convolution method to obtain annual loss distribution. The code will be given in the appendix, so that the calculation can be performed.

The annual loss distributions are obtained in **Figure 8**.

**Figure 8** shows annual loss distribution as a result of convolution method, it is the combination of lognormal and generalized Pareto distribution.

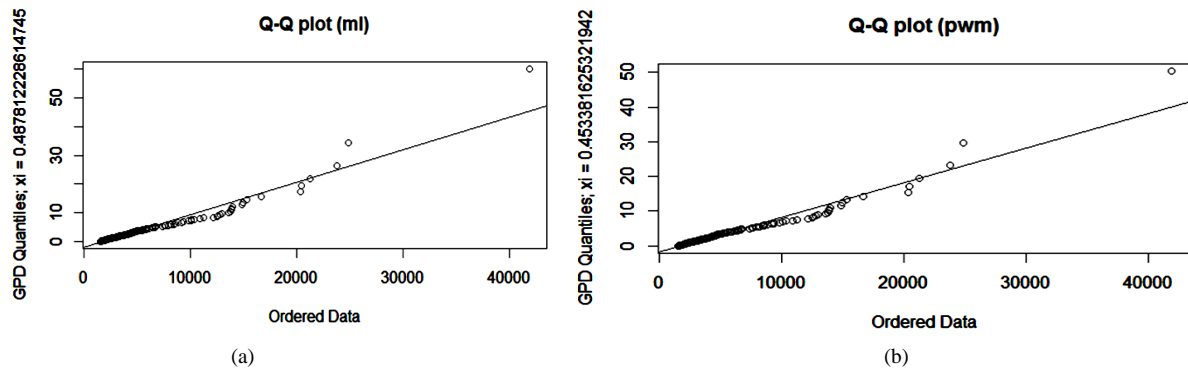
It is the combination of lognormal and generalized Pareto distribution, therefore because they differ and vary from sample to sample, the copula method should be used to obtain overall annual loss distribution (Hazewinkel, 2001).

Using Student t-copula, overall loss distribution is calculated and code that can be replicated is given, together with the VaR (Longin, 1997) results which are shown in **Figure 9**.

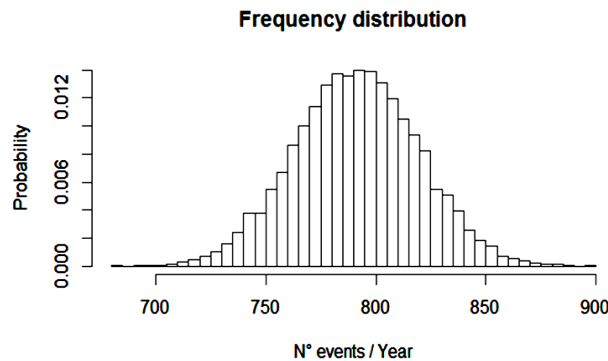
The code given in appendix demonstrates how to calculate VaR and copula for overall loss distribution. At the same time, some results are given for Value at Risk, expected loss and capital requirements. The given calculations and results demonstrate how to model risk in insurance by using lognormal and generalized Pareto distribution. At the same time distribution can be changed depending of the situation. This approach introduces modeling of operational risk in insurance and financial institutions and represents a review of the state of art techniques.

### 3. Conclusion

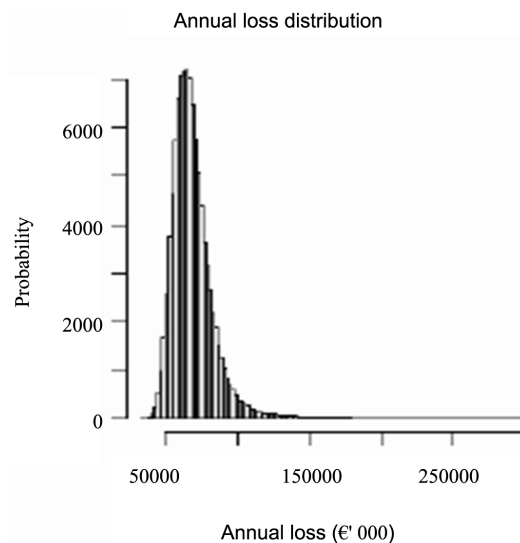
This paper introduces the state of the art techniques in operational risk modeling. It begins by presenting Solvency II and Basel II criteria. Afterwards it introduces statistical and mathematical techniques. Consequently, it presents modeling techniques by introducing Generalized Pareto distribution which is used in operational risk models in insurance and banking. The aforementioned paper represents a review in operational risk models in



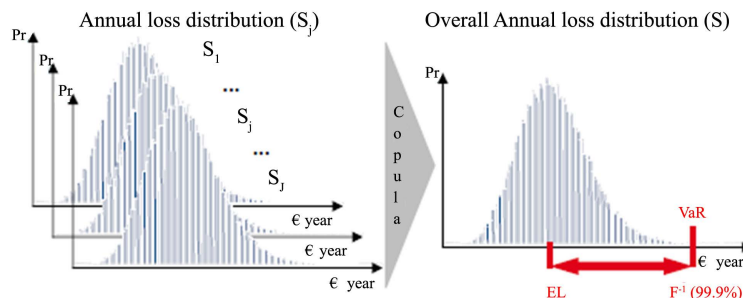
**Figure 6.** QQ plots with maximum likelihood method and probability weighted moments.



**Figure 7.** Loss frequency distribution Results: lambda Sample 532 > lambda 791.7354.



**Figure 8.** Annual loss distribution.



**Figure 9.** Copula method for obtaining overall annual loss distribution.

insurance. It introduces techniques that can be used in modeling operational risk in insurance and banking, and it can be immediately applied. Hopefully, this review will provide further research in risk models and provide new ideas that can be applied to operational risk in whole.

### Acknowledgements and Declaration of Interest Section

The authors report no conflicts of interest. The authors alone are responsible for the content and writing of the paper. I would like to thank my family for the immense support.

### References

(2012). Solvency—European Commission. Ec.europa.eu.  
 Accords, B. (2006). Basel II: Revised International Capital Framework.  
 Bracewell, R. (1986). *The Fourier Transform and Its Applications* (2nd ed.). New York: McGraw-Hill.  
 CEA—Groupe Consultatif (2005). Solvency II Glossary—European Commission.  
 Damelin, S., & Miller, W. (2011). *The Mathematics of Signal Processing*. Cambridge: Cambridge University Press. <http://dx.doi.org/10.1017/CBO9781139003896>  
 Doerig, H. U. (2000). *Operational Risks in Financial Services: An Old Challenge in a New Environment*. Switzerland: Credit Suisse Group.  
 Embrechts, P., Kluppelberg, C., & Mikosch, T. (1997). *Modelling Extremal Events*. Berlin: Springer. <http://dx.doi.org/10.1007/978-3-642-33483-2>  
 Hazewinkel, M. (Ed.) (2001). *Probability Distribution. Encyclopedia of Mathematics*. Berlin: Springer.

- Longin, F. (1997). *From Value at Risk to Stress Testing: The Extreme Value Approach*. Ceressec Working Paper, Paris: ESSEC.
- Mittnik, S. (2011). *Solvency II Calibrations: Where Curiosity Meets Spuriousity*. Munich: Center for Quantitative Risk Analysis (CEQURA), Department of Statistics, University of Munich.
- Nelsen, R. B. (1999). *An Introduction to Copulas*. New York: Springer.  
<http://dx.doi.org/10.1007/978-1-4757-3076-0>
- Nicolas, M., & Firlzi, J. (2011). *A Critique of the Basel Committee on Banking Supervision*. Paris: Revue Analyse Financière.
- Power, M. (2005). The Invention of Operational Risk. *Review of International Political Economy*, 12, 577-599.
- PWC Financial Services Regulatory Practice (2014). Operational Risk Capital: Nowhere to Hide.  
<http://www.pwc.com/us/en/financial-services/regulatory-services/publications/operational-risk-capital.jhtml>
- Sklar, A. (1959). *Fonctions de répartition à n dimensions et leurs marges* (pp. 229-231). Paris: Publications de l'Institut de Statistique de l'Université de Paris.

## Appendix

### Convolution Method Code in R to Obtain Overall Annual Loss Distribution

```
install.packages('evir')
library(evir)
# Quantile function of lognormal-GPD severity distribution
qlnorm.gpd = function(p, theta, theta.gpd, u)
{
  Fu = plnorm(u, meanlog=theta[1], sdlog=theta[2])
  x = ifelse(p<Fu,
            qlnorm( p=p, meanlog=theta[1], sdlog=theta[2] ),
            qgpd( p=(p - Fu) / (1 - Fu) , xi=theta.gpd[1], mu=theta.gpd[2], beta=theta.gpd[3]) )
  return(x)
}
# Random sampling function of lognormal-GPD severity distribution
rlnorm.gpd = function(n, theta, theta.gpd, u)
{ r = qlnorm.gpd(runif(n), theta, theta.gpd, u) }

set.seed(1000)
nSim = 10000 # Number of simulated annual losses
H = 1500 # Threshold body-tail
lambda = 791.7354 # Parameter of Poisson body
theta1 = 2.5 # Parameter mu of lognormal (body)
theta2 = 2 # Parameter sigma of lognormal (body)
theta1.tail = 0.5 # Shape parameter of GPD (tail)
theta2.tail = H # Location parameter of GPD (tail)
theta3.tail = 1000 # Scale parameter of GPD (tail)
sj = rep(0,nSim) # Annual loss distribution inzialization
freq = rpois(nSim, lambda) # Random sampling from Poisson
for(i in 1:nSim) # Convolution with Monte Carlo method
sj[i] = sum(rlnorm.gpd(n=freq[i], theta=c(theta1,theta2), theta.gpd=c(theta1.tail, theta2.tail, theta3.tail), u=H))
sj[i]
```

## Code to Calculate Overall Loss Distribution with VAR Results

```

>library(QRM)
>set.seed(1000)
>nSim = 1000000 # Number of simulated overall annual losses
>s1 = rlnorm(n=nSim, meanlog=4.5, sdlog=2.3) # Loss distribution risk class 1
>s2 = rlnorm(n=nSim, meanlog=5, sdlog=2.5) # Loss distribution risk class 2
>VaR.s1 = quantile(s1, 0.999) # VaR risk class 1
>VaR.s2 = quantile(s2, 0.999) # VaR risk class 2
>corr = 0.6 # Correlation among risk classes
>corrMatrix = matrix(data=c(1,corr,corr,1), nrow=2) # correlation matrix
>dof = 5 # degrees of freedom
>simCopulaT = rcopula.t(n=nSim, df=dof, Sigma=corrMatrix) # Simulations from Student-t copula
>s = quantile(s1, simCopulaT[,1]) + quantile(s2, simCopulaT[,2]) # overall annual loss distribution
>VaR.s = quantile(s, 0.999)
>divEff = (VaR.s1+VaR.s-VaR.s)/(VaR.s1+VaR.s) # diversification effect
>EL = quantile(s, 0.5) # Expected loss
>capReq = VaR.s - EL # Capital requirement
>VaR.s1; VaR.s2; VaR.s; divEff; EL; capReq
 99.9%
111300.8
 99.9%
339577.2
 99.9%
410748.6
 99.9%
0.2131997
 50%
355.3662
 99.9%
410393.2

```