

# Estimating Operational Risk Severities

– assessment of a semiparametric approach

---

Lund Institute of Technology

Viktor Davidsson & Anton Forssell

May 26, 2014



# Abstract

In this thesis a new method of modeling a severity distribution, proposed by Bolancé, Guillén, Gustafsson and Nielsen (2012), used for calculating the minimum capital requirement concerning operational risk for financial institutions is evaluated. To estimate the severity distribution, a generalized Champernowne distribution is used to transform loss data, followed by kernel density estimation.

The method has been evaluated by manipulating frequency and size of losses in different domains for a benchmark dataset, with a particular focus on how changes in the body affect the whole distribution. The impact of these alterations is measured in change of Value-at-Risk.

The main conclusion from this assessment is that the method is fairly responsive to changes over the whole domain. It responds and incorporates small changes, yet stable when extreme events occur. Consequently, the method is suitable for modeling operational risk severities.

Key Words: *Operational Risk, Severity Distribution, Champernowne, Transformation, Kernel Density Estimation.*



# Acknowledgements

We would like to thank our supervisors at Lund University, Birger Nilsson and Karl Larsson, for valuable tips and insightful comments about the writing and structure of the report.

# Contents

<b>1</b>	<b>Introduction .....</b>	<b>1</b>
1.1	Background.....	1
1.2	Problem Statement and Purpose.....	2
1.3	Terms & Abbreviations .....	3
<b>2</b>	<b>Regulations and Implementation .....</b>	<b>5</b>
2.1	Basel Framework.....	5
2.2	Methods of Measuring Operational Risk.....	6
2.2.1	The Basic Indicator Approach .....	7
2.2.2	The Standardized Approach .....	7
2.2.3	The Advanced Measurement Approach .....	9
2.3	The Loss Distribution Approach .....	10
2.3.1	Categorization .....	11
2.4	Loss Density Estimation.....	12
2.5	The Generalized Champernowne Distribution.....	13
2.5.1	The Effects of Parameter $c$ .....	14
2.6	Kernel Density Estimation.....	17
2.6.1	Boundary Correction.....	19
2.7	Risk Measures.....	20
2.7.1	Value-at-Risk.....	20
<b>3</b>	<b>Method.....</b>	<b>21</b>
3.1	Estimating Champernowne Parameters .....	21
3.2	Transformation .....	23

3.3	Kernel Density Estimation.....	24
3.3.1	Classical KDE with Boundary Correction.....	24
3.4	Back-transformation .....	26
<b>4</b>	<b>Sensitivity Analysis.....</b>	<b>29</b>
4.1	Benchmark .....	29
4.2	Body and Tail Manipulation.....	30
4.3	Modified Body.....	31
4.4	Manipulating Interval Data.....	32
4.4.1	Change Size of Losses.....	34
4.4.2	Increase Frequency of Losses.....	35
4.5	Repeated Ten-multiplier Stress Test.....	35
4.6	Probability Assignment.....	35
4.7	Bandwidth Selection .....	36
<b>5</b>	<b>Results.....</b>	<b>37</b>
5.1	Body and Tail Manipulation.....	37
5.2	Modified Body.....	40
5.3	Manipulating Interval Data.....	42
5.3.1	Change in Size of Losses.....	43
5.3.2	Increase Frequency of Losses.....	45
5.4	Repeated Ten-multiplier Stress Test.....	49
5.5	Probability Assignment.....	50
5.6	Bandwidth Selection .....	51
<b>6</b>	<b>Conclusions and Further Research.....</b>	<b>53</b>
6.1	Further Research.....	54
<b>7</b>	<b>References .....</b>	<b>55</b>
	<b>Appendix.....</b>	<b>57</b>





# 1 Introduction

## 1.1 Background

During the last couple of years, operational risk has been a highly debated subject and a popular research topic. How does one predict a natural disaster, a terrorist attack or a bank robbery? These are questions large corporations, institutional agencies, banks and insurance companies face on a daily basis. The complexity of its structure, lack of knowledge and data regarding extreme events are some of the reasons why operational risk is so challenging.

The Basel Committee on Banking Supervision (BCBS) defines operational risk as: “the risk of direct or indirect loss resulting from inadequate or failed internal processes, people and systems or from external events”. This definition is very broad and is commonly referred to as the risk of doing business for any company. Knowing about this risk, the incentives for understanding operational risk are many. By modeling the operational risk accurately, an adequate amount of capital can be held to prevent a possible default for financial institutions. There are three approaches to do this, the first two, Basic Indicator Approach and Standardized Approach are directly linked to the bank’s revenue, while the third, Advanced Measurement Approach (AMA) allows financial institutions to develop their own model as long as it fulfills certain criteria.

One common approach to predict or model the operational risk at a financial institution in the AMA, is to use an actuarial approach called the Loss Distribution Approach (LDA). Basically, a frequency and severity distribution are calculated separately before they are merged into one to find an appropriate loss distribution. The frequency distribution represents how often a loss occurs and is relatively simple to estimate. BCBS (2009) states that the Poisson and the negative binomial distribution are two of the most common discrete parametric distributions for modeling frequency. The severity distribution on the other hand is extremely hard to estimate. A common approach has been to separate the severity distribution in two parts. Typically, a parametric distribution is used for the main body of data and Extreme Value Theory (EVT) is used to capture tail events. However, by separating the severity distribution at a

certain threshold you might get a flawed outcome because of the uncertainty in determining threshold level.

Bolancé, Guillén, Gustafsson and Nielsen (2012) propose a new method of practice in modeling the severity distribution. Instead of using EVT, the proposed method considers the whole range of operational loss data, avoiding the problem with choosing a threshold. In short, they use the generalized Champernowne distribution to transform the data followed by kernel density estimation to obtain a severity distribution. The method has proven to estimate operational loss distribution data well. However, no studies regarding its sensitivity have been published yet. Therefore, in this thesis the authors will conduct a thorough sensitivity analysis of the proposed approach.

## **1.2 Problem Statement and Purpose**

Quantification of operational risk is very difficult. This is mainly due to the sparsely available data that is generally both heavy-tailed and has a large variance. Recent studies, like the one Bolancé et al. (2012) proposes, have shown that a possible solution to overcome issues of estimating a severity distribution is to transform the data with the generalized Champernowne distribution followed by kernel density estimation to capture the severity of operational risk losses. Due to the fact that it is a relatively new method of modeling operational risk, not many sensitivity assessments of the method have been conducted. Therefore, the aim of this thesis is to evaluate the sensitivity of the model.

The reader will see how different scenarios affect the final outcome and how this can be controlled, especially how changes in the body affect the whole distribution. This is because extreme events are very hard to prevent while smaller losses, occurring in the everyday business are manageable. The scenarios should give an indicator about which types of events have the most impact. Furthermore, important parameters and assumptions will be highlighted.

### 1.3 Terms & Abbreviations

AMA	Advanced Measurement Approach
BCBS	Basel Committee on Banking Supervision
BIA	Basic Indicator Approach
CDF	Cumulative Distribution Function
EVT	Extreme Value Theory
GCD	Generalized Champernowne Distribution
KDE	Kernel density estimation
LDA	Loss Distribution Approach
PDF	Probability Density Function
SA	Standardized Approach
VaR	Value-at-Risk



## 2 Regulations and Implementation

In this chapter, a brief history of the origin the Basel framework proposed by BCBS will be described. In the Basel II Capital Accord, proposed by the BCBS in 2004, it is suggested that financial institutions should allocate adequate capital for credit-, market- and operational risk. The Swedish authority supervising financial institutions and how these comply with existing regulations is Finansinspektionen. The reader will be guided through the essentials regarding operational risk and the allowed methods of implementing the framework.

Furthermore, this chapter will explain the Loss Distribution Approach commonly used for calculating minimum capital requirement due to operational risk in the Advanced Measurement Approach. The concepts and theory from Bolancé et al. (2012) that is needed for understanding their method of estimating a severity distribution will be described. Also, a brief description of the risk measure Value-at-Risk (VaR) will be provided.

### 2.1 Basel Framework

In BCBS (2013) a short history of the Basel Committee is described. In the mid 1970's a couple of huge banks defaulted due to some extreme foreign exchange exposure. In response to the disruption and instability in the international financial markets, the central governors of the G10 countries established the Committee on Banking Regulations and Supervisory Practices, which was later renamed as the Basel Committee on Banking Supervision (BCBS). BCBS was designed as a forum for regular cooperation between its member countries on banking supervisory matters. Its aim was and is to enhance financial stability by improving supervisory know-how and the quality of banking supervision worldwide.

In July 1988, a consultative paper about capital measurement system commonly referred to as the Basel Capital Accord or Basel I was published. The primary focus of the paper is credit risk and risk-weighted assets. Basically, the Accord called for a minimum capital ratio of capital to risk-weighted assets of 8 % to be implemented by the

end of 1992. The framework was introduced not only in member countries but also in most other countries with active international banks.

Capital adequacy was the main focus of the Committee and by June 2004 the new Revised Capital Framework was published. This paper is generally known as Basel II and comprises of three pillars:

- minimum capital requirements, which sought to develop and expand the standardized rules set out in the 1988 Accord;
- supervisory review of an institution's capital adequacy and internal assessment process; and
- effective use of disclosure as a lever to strengthen market discipline and encourage sound banking practices.

Basel II is a comprehensive framework which was designed to improve the way regulatory capital requirements reflect underlying risks and to better address the financial innovations that had emerged in recent years. The changes aimed at rewarding and encouraging continued improvements in risk management.

In the first pillar, the Basel II framework suggests that a financial institution should allocate adequate capital for credit-, market- and operational risk. Various sophisticated models for calculating minimum capital requirements are proposed. Since this thesis is focused on operational risk, the details of the calculations of operational capital requirement will be explained in detail in the next section.

Basel III is the latest framework available. The financial crisis that started in 2008 showed that Basel II did not capture the risks for banks sufficiently. The overall objective with the new framework is to strengthen the ability of banks to withstand losses and to reduce the likelihood of new financial crises. Basel III implies that banks must hold more capital of better quality and that new requirements on bank liquidity are introduced. The various elements of Basel III will gradually be phased in over the coming years. It does not directly mention operational risk or comprise any structural changes in the methods used to calculate the minimum capital requirements for operational risk i.e. the regulations introduced in Basel II are still valid. Basel III will therefore not be discussed further in this thesis.

## **2.2 Methods of Measuring Operational Risk**

The framework presented by BCBS (2004) and BCBS (2011) for calculating operational risk is described below. The Basel framework suggests three different approaches

to calculate the capital requirement for operational risk, namely the Basic Indicator Approach, the Standardized Approach and the Advanced Measurement Approach. The Basic Indicator Approach and the Standardized Approach are directly linked to the bank's revenue, while the Advanced Measurement Approach allows financial institutions to develop their own model as long as it fulfills certain criteria.

### 2.2.1 The Basic Indicator Approach

Banks using the Basic Indicator Approach (BIA) must hold capital for operational risk equal to the average over the previous three years of a fixed percentage ( $\alpha$ ) of positive annual gross income. Figures for any year in which annual gross income is negative or zero should be excluded from both the numerator and denominator when calculating the capital average. The capital requirement according to the BIA may be expressed as:

$$K_{BIA} = \frac{\sum_{i=1}^n (GI_i \cdot \alpha)}{n} \quad (2.1)$$

where:

- $K_{BIA}$  = the capital requirement under the Basic Indicator Approach.
- $GI_i$  = annual gross income, where positive for year  $i$ .
- $n$  = number of the previous three years for which gross income is positive.
- $\alpha$  = a percentage set by the Committee, relating the industry-wide level of required capital to the industry-wide level of the indicator. The suggested percentage is 15%.

The BIA can be used by all banks, no specific criteria have to be met to use it.

### 2.2.2 The Standardized Approach

For banks with more than one business line, the Standardized Approach (SA) can be used to calculate its capital requirement. In the SA, banks' activities are divided into eight business lines: corporate finance, trading & sales, retail banking, commercial banking, payment & settlement, agency services, asset management, and retail brokerage.

Within each business line, gross income is a broad indicator that serves as a proxy for the scale of business operations and thus the likely scale of operational risk exposure within each of these business lines. The capital requirement for each business line is calculated by multiplying gross income by a factor ( $\beta$ ) assigned to that business line.

The factor beta serves as a proxy for the industry-wide relationship between the operational risk loss experience for a given business line and the aggregate level of gross income for that business line. It should be noted that in the SA gross income is measured for each business line, not the whole institution, e.g. in corporate finance, the indicator is the gross income generated in the corporate finance business line.

The total capital requirement is calculated as the three-year average of the simple summation of the regulatory capital requirements across each of the business lines in each year. In any given year, negative capital requirements (resulting from negative gross income) in any business line may offset positive capital requirements in other business lines without limit. However, where the aggregate capital requirement across all business lines within a given year is negative, then the input to the numerator for that year will be zero. The total capital requirement according to the Standardized Approach may be expressed as:

$$K_{SA} = \frac{\sum_{i=1}^3 \max(\sum_{j=1}^8 GI_j \cdot \beta_j, 0)}{3} \quad (2.2)$$

where:

- $K_{SA}$  = the capital requirement under the Standardized Approach.
- $GI_j$  = annual gross income in a given year for each of the business lines.
- $\beta_j$  = a fixed percentage, set by the Committee, relating the level of required capital to the level of the gross income for each of the business lines. The values of  $\beta$  are shown in Table 2.1.

Business Line	$\beta_j$	Value of $\beta_j$
Corporate finance	$\beta_1$	18 %
Trading and sales	$\beta_2$	18 %
Retail banking	$\beta_3$	12 %
Commercial banking	$\beta_4$	15 %
Payment and settlement	$\beta_5$	18 %
Agency services	$\beta_6$	15 %
Asset management	$\beta_7$	12 %
Retail brokerage	$\beta_8$	12 %

Table 2.1 Values of  $\beta$  for proposed business lines in Basel II.



In order to qualify for use of the SA, a bank must satisfy a number of minimum requirements. Senior management must be actively involved in the oversight of the operational risk management framework; it has an operational risk management system that is conceptually sound; and it has sufficient resources in the use of the approach in the major business lines as well as in the control area. The SA is subject to consent by Finansinspektionen in Sweden.

### **2.2.3 The Advanced Measurement Approach**

The Advanced Measurement Approach (AMA) is the most sophisticated method of calculating a bank's capital requirement for operational risk. Supervisory guidelines on how to use AMA are explained in BCBS (2011). Under the AMA, the regulatory capital requirement will equal the risk measure generated by the bank's internal operational risk measurement system using quantitative and qualitative criteria.

There is no exact outline specifying how this model should be built and every bank has the right to develop its own. However, use of the AMA is subject to supervisory approval, i.e. in Sweden Finansinspektionen must give its consent. Just as in the use of SA, a bank must satisfy a number of minimum requirements. Senior management must be actively involved in the oversight of the operational risk management framework; it has an operational risk management system that is conceptually sound; and it has sufficient resources in the use of the approach in the major business lines as well as in the control area. The model must include credible, systematic, transparent and verifiable approaches for weighting internal operational loss data, external operational loss data, scenario analysis and internal control system.

Internal loss data is operational losses that the bank has observed internally. External operational loss data are losses that other banks have observed. There are both publicly available data and more comprehensive data from different consortiums. There is a problem with scaling external data due to the lack of knowledge of origination that has to be accounted for. External data is considered needed because of the general lack of data within the internal loss data. Scenario analysis should be performed as a complement. It should consist of both expert opinions on the bank's exposure to severe losses and risk measures from loss probability distributions. With internal control system, it should be stated how the internal risk assessment methodology captures key factors that are important for the bank's operational risk profile.

The most important quantitative criteria that the method has to satisfy are:

- Use a 99,9 % confidence interval.
- Time horizon of one year.
- Risk measure is Value-at-Risk.
- Correlations must be described if used.

In addition to the categorization of losses into business lines used in the SA, the Basel framework advocates the use of categorization of events. These event types are 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.

Modeling operational risk with AMA is commonly done with two different methods. The first method is called the Loss Distribution Approach (LDA) and the second is based on Extreme Value Theory (EVT). LDA tries to overcome the barrier with scarce data by analyzing each loss while EVT on the other hand focuses on the extreme observations. Note that EVT can be used as a partial method in the LDA to find a distribution. This thesis focuses on the LDA which will be described in the next section.

## **2.3 The Loss Distribution Approach**

LDA is a very popular statistical approach for calculating aggregate loss distributions. There are no fixed rules on how to implement the LDA from the AMA, but in order to understand the basics of LDA, some concepts have to be defined. To model the capital requirement, one should use a horizon of one year according to BCBS (2011). Modeling of the aggregate loss distribution is done in two steps. First, a loss frequency distribution is defined as the number of losses that are expected to occur in one year. Secondly, the severity for each loss has to be estimated. The focus on this thesis is about the tail behavior of a certain severity loss distribution and will be described in depth throughout following sections.

To obtain the desired final distribution, the frequency and severity has to be compounded into one distribution. According to Frachot, Georges and Roncalli (2001), there is no general analytical expression of the compound distribution. To compute the loss distribution, numerical algorithms are required. The most commonly used approaches are the Monte Carlo method, Panjer's recursive approach and the inverse of the characteristic function.

The Monte Carlo method comes in many shapes. However, in the LDA it could be explained relatively simple. Firstly, a number of losses are simulated from the frequency distribution which represents the number of losses that could be expected during one year. Thereafter, the size for each loss is drawn from the severity distribution. These extractions are then aggregated to form a potential one year loss. This approach is then repeated a large number of times, resulting in a new empirical distribution as seen in Figure 2.1. The obtained empirical distribution serves as an approximation of the true distribution. The final step in LDA is to calculate the desired percentile or the Value-at-Risk (VaR) with confidence level  $\alpha$ .

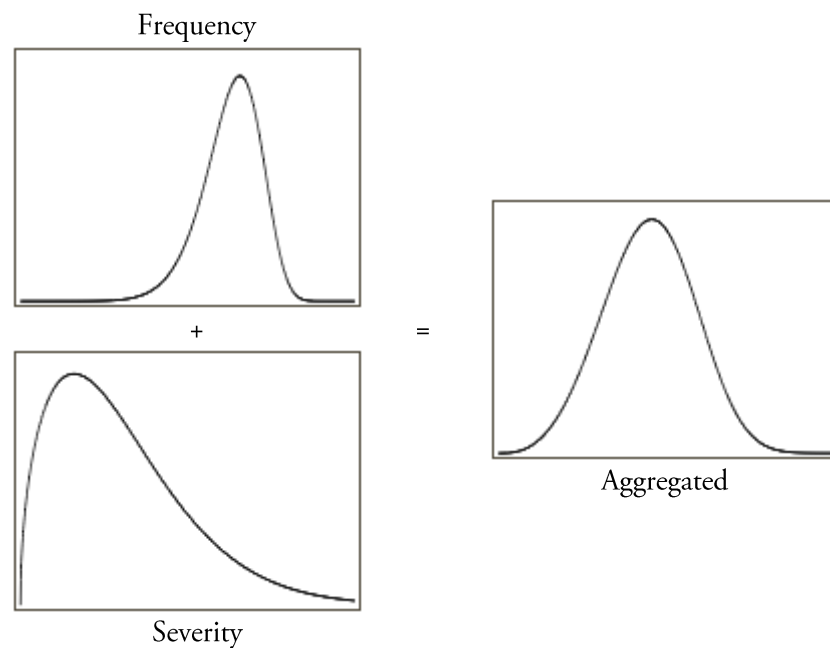


Figure 2.1 *Illustrative example of loss distribution approach.*

### 2.3.1 Categorization

As described in Section 2.2.3, the Basel Committee advocates the use of categorization into eight business lines and seven event types, usually shown in a matrix (see Table 2.2). Each cell in the matrix contains the number of losses as well as the size for each combination of business line and event type.

<b>Business Line</b>	<b>Event Type</b>	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
Corporate finance								
Trading & Sales								
Retail Banking								
Commercial Banking								
Payments & Settlements								
Agency Services								
Asset Management								
Retail Brokerage								

**Table 2.2** *The LDA Operational Risk Matrix.*

If no dependency in the matrix is assumed, it is straightforward to calculate the aggregate loss distribution and the 99,9 % VaR. One simply summarizes the outcome from each cell to achieve the capital requirement. However, in most cases, there will be some sort of dependency. The cell dependence structure is usually modeled and captured by copula techniques suggested by Di Clemente and Romano (2003). Furthermore, some cells might contain so few losses that it is impossible to draw any conclusions from them. It is because of this scarcity of internal data that external data is needed. These problems are not within the scope of this thesis; hence, it will not be discussed further.

## 2.4 Loss Density Estimation

As in most cases when one is looking for any distribution, not least a severity distribution, an empirical dataset of observations is a good starting point. The basic concept is to find a distribution that assign a probability to each given loss amount, i.e. finding a probability density function (pdf).

There are two common approaches of estimating this density, parametric and nonparametric. The parametric approach assumes that the data comes from a known parametric family of distributions, for example the normal distribution. In order to estimate the density function, if assumed to be normally distributed, it is sufficient to find

estimates of the mean and variance of the data and substitute these into the formula for normal density. However, forcing the data into a known distribution might not result in a good fit. The nonparametric approach, explained in Silverman (1986), is less rigid in its assumptions as it does not require any distributional assumptions. The complexity of the density function does only depend upon the sample and typically a bandwidth within the nonparametric method that is used. This thesis will make use of a combination of both the parametric and nonparametric method, which is commonly referred to as a semiparametric approach. The method, suggested in Bolancé et al. (2012), consists of first fitting the data to a parametric Champernowne distribution and then estimating the density function with the help of kernel density estimation.

## 2.5 The Generalized Champernowne Distribution

The original Champernowne distribution was first proposed by D. G. Champernowne in 1937 and further developed in his article “The graduation of income distributions” from 1952. The distribution was developed in order to describe the logarithm of income and is a generalization of the logistic distribution. The probability density function of the original Champernowne distribution is:

$$f(x) = \frac{c_*}{x \left( \frac{1}{2} \left( \frac{x}{M} \right)^{-\alpha} + \lambda + \frac{1}{2} \left( \frac{x}{M} \right)^{\alpha} \right)} \quad x \geq 0, \quad (2.3)$$

where  $c_*$  is a normalizing constant and  $\alpha, M$  and  $\lambda$  are positive parameters. When  $\lambda = 1$  and  $c_* = \frac{1}{2}\alpha$  the density of the original Champernowne distribution is called the Champernowne distribution with pdf:

$$f_{\alpha, M}(x) = \frac{\alpha M^{\alpha} x^{\alpha-1}}{(x^{\alpha} + M^{\alpha})^2} \quad (2.4)$$

and cdf:

$$F_{\alpha, M}(x) = \frac{x^{\alpha}}{x^{\alpha} + M^{\alpha}}. \quad (2.5)$$

The Champernowne distribution converges to a Pareto distribution in the tail but looks like a lognormal distribution near 0 when  $\alpha > 1$ . If  $\alpha \neq 1$  the density of the distribution is either 0 or infinity when evaluated at 0. Buch-Larsen, Nielsen, Guillén, and Bolancé (2005) presents the generalized Champernowne distribution (GCD) in order to overcome this inflexibility near 0. The GCD includes a new parameter  $c$  which ensures the possibility of a positive finite value at 0 for all  $\alpha$ . The generalized Champernowne cdf is defined as:

$$T_{\alpha,M,c}(x) = \frac{(x+c)^\alpha - c^\alpha}{(x+c)^\alpha + (M+c)^\alpha - 2c^\alpha} \quad x \geq 0, \quad (2.6)$$

where  $\alpha > 0$ ,  $M > 0$  and  $c \geq 0$ . The density of the GCD is:

$$t_{\alpha,M,c}(x) = \frac{\alpha(x+c)^{\alpha-1}((M+c)^\alpha - c^\alpha)}{((x+c)^\alpha + (M+c)^\alpha - 2c^\alpha)^2} \quad x \geq 0. \quad (2.7)$$

The GCD does, analogous to the Champernowne distribution, converge to a Pareto distribution in the tail. That means:

$$t_{\alpha,M,c}(x) \rightarrow \frac{\alpha(((M+c)^\alpha - c^\alpha)^{1/\alpha})^\alpha}{x^{\alpha+1}} \quad \text{as } x \rightarrow \infty. \quad (2.8)$$

### 2.5.1 The Effects of Parameter c

The parameter  $c$  in the GCD shows both scale parameter properties and density effects. However, most of these effects depend on the value of parameter  $\alpha$ , and if  $\alpha = 1$  then  $c$  has no effect at all. The effects of parameter  $c$  are summarized in Table 2.3.

c	$\alpha < 1$	$\alpha > 1$	$\alpha = 1$
Increasing c	Larger derivate of cdf	Smaller derivative of cdf	No effect
Decreasing c	Smaller derivative of cdf	Larger derivate of cdf	
c > 0	Lighter tails	Heavier tails	
	Positive finite density at 0		
	—	Shifts mode to the left	

Table 2.3 Compilation of parameter c effects.

The scale parameter properties depend on whether  $\alpha$  is less or greater than one. For an increasing parameter  $c$  when  $\alpha < 1$  the derivative of the cdf gets larger, and when  $\alpha > 1$  the derivative of the cdf gets smaller. For a decreasing  $c$  the opposite is true.

The parameter  $c$  has three different effects on the density when  $c$  is positive. Firstly,  $c$  alters the density in the tail. When  $\alpha < 1$  the density gets lighter tails, and  $\alpha > 1$  results in heavier tails. Secondly,  $c$  ensures that there is a positive finite density at 0, i.e.:

$$0 < t_{\alpha, M, c}(0) = \frac{\alpha c^{\alpha-1}}{(M+c)^\alpha - c^\alpha} < \infty \quad \text{when } c > 0. \quad (2.9)$$

Thirdly, when  $\alpha > 1$  the density has a mode and  $c$  shifts the mode to the left.

Figure 2.2 on the following page provides a graphical interpretation of some of the effects. In a) and b),  $\alpha$  is fixed at a value less than one. In the interval  $[0, M)$  the cdf is larger for  $c = 0$  than for  $c > 0$  while in the interval  $(M, \infty]$  the opposite is true. In the pdf it is clear that when  $c > 0$  there is a lighter tail. In c) and d)  $\alpha$  is fixed at one and it is apparent that the value of  $c$  has no effect in this case. In e) and f)  $\alpha$  is fixed at a value larger than one. The effects of parameter  $c$  in this case are the contrary of those seen in a) and b). In the interval  $[0, M)$  the cdf is larger for  $c > 0$  than for  $c = 0$  and in the interval  $(M, \infty]$  it is the opposite. In the pdf there can be observed a shift of the mode to the left when  $c > 0$ .

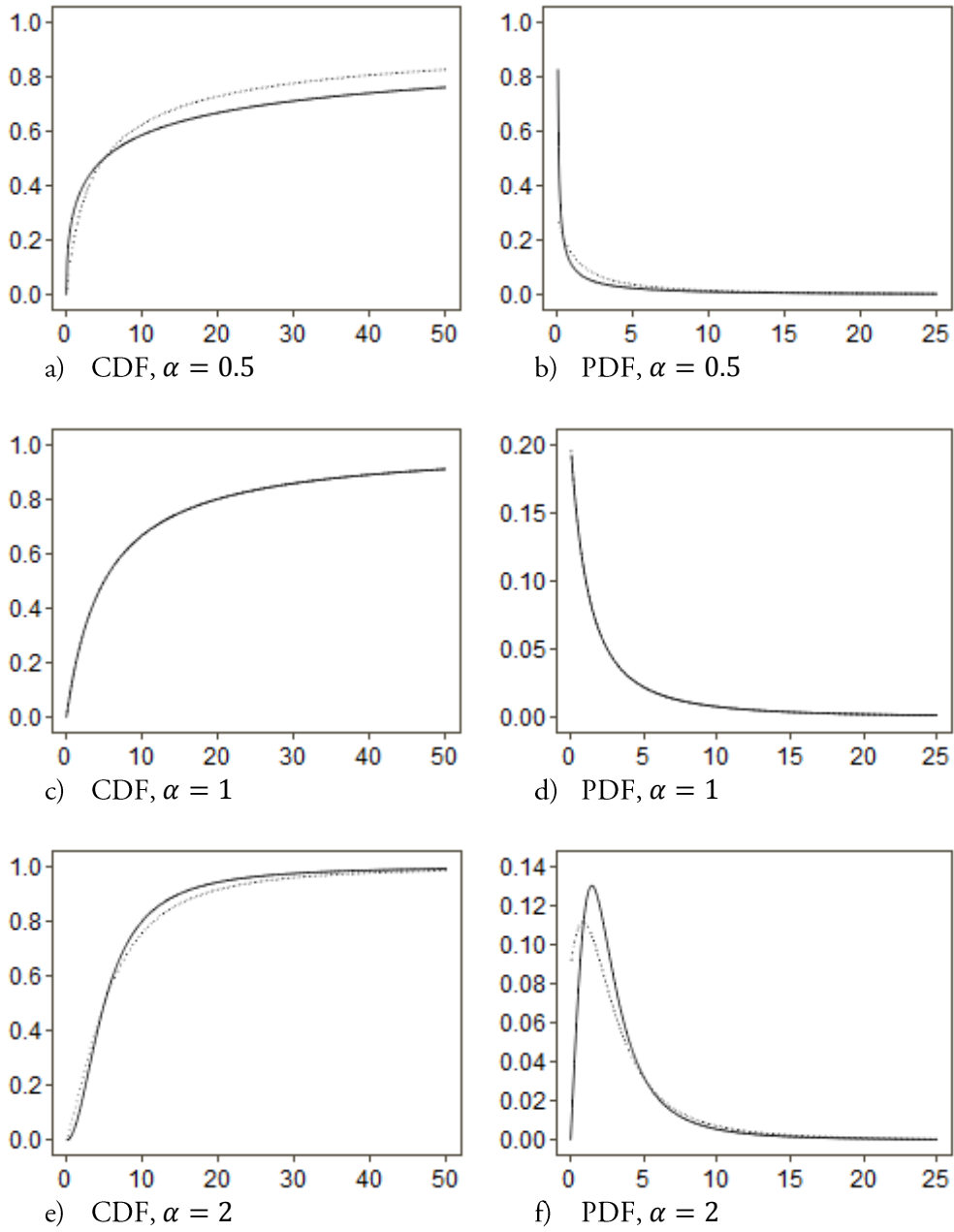


Figure 2.2 The effect of parameter  $c$ . The GCD plotted for different values of  $\alpha$  and  $c$ , while  $M = 5$ .  $c = 0$  in the solid line and,  $c = 2$  in the dashed line.



## 2.6 Kernel Density Estimation

Kernel density estimation (KDE) is a nonparametric method designed to estimate the pdf of a random variable, described in Silverman (1986) and Zucchini (2003). The estimator use empirical data to create an estimation of the pdf at every point in the domain of the random variable. The influence of the empirical data will vary, giving more influence to nearby data and less to data further away from the evaluation point. How the estimator does this weighting of data depends on the choice of both kernel function and bandwidth.

For a random sample of  $n$  independent and identically distributed loss observations  $X_1, X_2, \dots, X_n$  for a random variable  $X$ , the kernel density estimator of the density function  $f$  is defined as:

$$\hat{f}(x) = \frac{1}{nb} \sum_{i=1}^n K\left(\frac{x - X_i}{b}\right), \quad (2.10)$$

where  $K(\cdot)$  is the kernel function and  $b$  is the bandwidth parameter. The kernel function determines the shape of the weighting function and the bandwidth determines the width of it.

The kernel is a function of a single variable and it must integrate to one. Usually the kernel is a symmetric pdf with zero mean. In Table 2.4 three of the most common kernel functions are listed. The kernel determines the local behavior of the estimator, giving weights to each sample in the data. If for example an unbounded kernel such as the Gaussian kernel is used, data far away from the evaluation point will still have influence on the estimated density in this point. Conversely, if a bounded kernel such as the Epanechnikov is used, data outside its domain will not give any influence at all to the estimated density. The Epanechnikov kernel has been proven optimal in a minimum variance sense and is described in detail in Epanechnikov (1969).

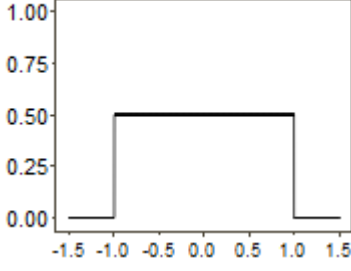
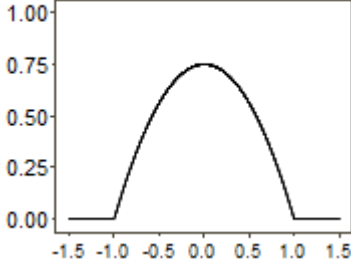
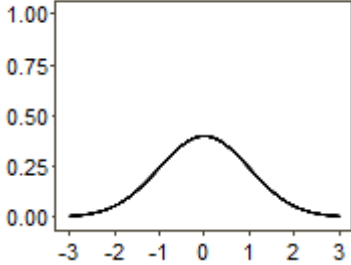
Kernel	$K(u)$	Plot
Uniform	$\begin{cases} 1/2 & \text{for }  u  \leq 1 \\ 0 & \text{otherwise} \end{cases}$	
Epanechnikov	$\begin{cases} (3/4)(1 - u^2) & \text{for }  u  \leq 1 \\ 0 & \text{otherwise} \end{cases}$	
Gaussian	$\frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}u^2}$ for $ u  \leq \infty$	

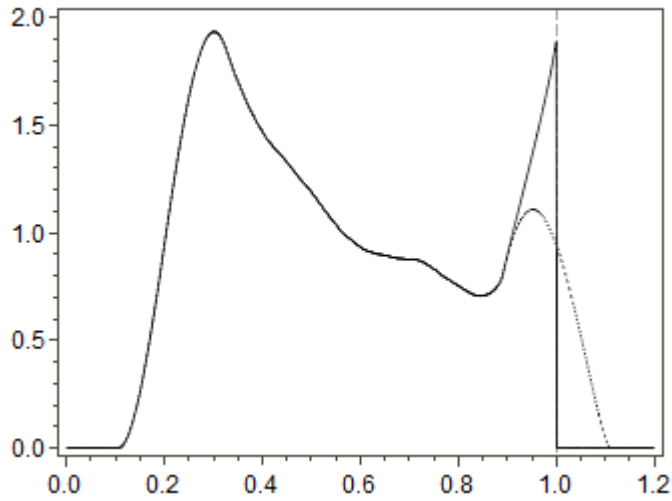
Table 2.4 Three commonly used kernel functions.

The choice of  $b$  controls the width of the kernel, determining how wide probability mass that is given in the vicinity of each data point. Thus, parameter  $b$  has a smoothing effect on the KDE. For a smaller choice of  $b$  the resulting KDE will exhibit larger fluctuations because each data point only will distribute probability mass close to itself. On the contrary, when  $b$  is large the fluctuations decrease and the details in the KDE gets obscured.

A disadvantage with KDE described in Bolancé et al. (2012), is that it does not manage to estimate heavy-tailed distributions very well. With the presence of some very large observations, which is generally the case with loss data, the dispersion will be large and the bandwidth parameter will in general also be big. As a consequence, the estimated density will be underestimated in the part where the most mass is located, i.e. the body. If one tries to solve this problem by using a smaller bandwidth parameter another problem arises, the resulting density will have a bumpy shape in the tail due to the large observations.

### 2.6.1 Boundary Correction

The classical KDE has unbounded support. It has no knowledge about boundaries and allowed intervals in its original form. In the region near the boundary, data overflows which make the estimator poor. The consequence of this would be a flawed probability density function that does not integrate to one as discussed in Jones (1993). When using a transformation method as discussed in this thesis, boundary correction is needed because data are transformed into the  $[0,1]$  interval.



**Figure 2.3** *Illustrative example of boundary correction. The dotted line is without correction term and the solid line is with correction term.*

Jones (1993), Gasser and Müller (1979) and Zhang and Karunamuni (1998) suggests a number of different boundary correction methods. All methods have proven to have some success, however, with different precision. As in most situations the more complex the model, the harder it is to understand its features. For the purpose of this thesis the most obvious method is used, i.e. forcing integration of each kernel that overflows the boundary to unity within the support by renormalization, as seen in Figure 2.3. The kernel density estimator of the density function  $f$  with a boundary correction term  $a_{kl}(y, b)$  and  $y \in [0,1]$ , is defined as:

$$\hat{f}(y) = \frac{1}{nba_{kl}(y, b)} \sum_{i=1}^n K\left(\frac{y - Y_i}{b}\right), \quad (2.11)$$

The function above is simply the classical KDE, stated in equation 2.10 with an added correction term.

The kernel density estimator with asymptotic properties around boundaries is defined in Bolancé et al. (2012) as:

$$a_{kl}(y, b) = \int_{\max\{-1, \frac{y-1}{b}\}}^{\min\{1, \frac{y}{b}\}} u^k K(u)^l du, \quad (2.12)$$

for  $y \in [0, 1]$ .

Note that the limits of the integral will be -1 and 1 when  $y$  is at least one bandwidth away from the boundaries. At the boundary points belonging to the intervals  $[0, b)$  and  $(1 - b, 1]$ , the limits of the integral will change and the boundary correction term will take nontrivial values.

## 2.7 Risk Measures

### 2.7.1 Value-at-Risk

The measurement Value-at-Risk (VaR) is widely used by mathematicians and financial institutions. Furthermore, it is the underlying risk measure used throughout the Basel framework. VaR measures the worst expected loss under normal conditions over a specific time interval at a given confidence level. Booth et al. (2005) defines VaR as:

$$\int_{-\infty}^{\text{VaR}_\alpha} f(v) dv = 1 - \alpha = P(v \leq \text{VaR}_\alpha) \quad (2.13)$$

where  $f(v)$  is the general density function,  $\alpha$  is the VaR confidence level and  $\text{VaR}_\alpha$  is the Value-at-Risk.  $\text{VaR}_\alpha$  is defined so that the probability of obtaining a loss less than  $\text{VaR}_\alpha$  is equal to one minus  $\alpha$ .

## 3 Method

In this chapter the suggested method, presented in Bolancé et al. (2012), for finding a desired severity distribution is described in detail. The reader will be guided through the different steps required to perform the calculations. The GCD was presented in Chapter 2 where a parametric approach is used to fit the density as a start. For the purpose of explanation we let  $X_i, i = 1, \dots, n$  be independent observations obtained from a dataset which illustrates operational losses.

The method is based on kernel density estimation. Due to the problem with KDE on heavy-tailed data stated in Section 2.6, a transformation is combined with the KDE. First the data is transformed into the  $[0,1]$  interval with the GCD. Thereafter the density of the transformed data is estimated with KDE. Finally the estimator of the original density is obtained by transforming the estimator of transformed data back to the original scale with the help of a reverse transformation. The method is called semi-parametric because we use a parametric transformation function and then a nonparametric kernel estimation.

According to Bolancé et al. (2012), existing results based on simulation studies have shown that the method is able to estimate all three possible kinds of tails, namely, the Fréchet type, the Weibull type, and the Gumbel type known from EVT. This makes the semiparametric transformation approach method extremely useful for operational risk analysis.

### 3.1 Estimating Champernowne Parameters

The first step in the suggested method is to calculate the parameters  $\hat{\alpha}, \hat{M}$  and  $\hat{c}$  of the generalized Champernowne distribution to obtain the transformation function. Buch-Larsen et al. (2005) suggests that the GCD parameters should be calculated step by step, beginning with estimating  $M$  and then continuing with the other parameters. For the purpose of visualization Figure 3.1 shows a histogram of the initial data.

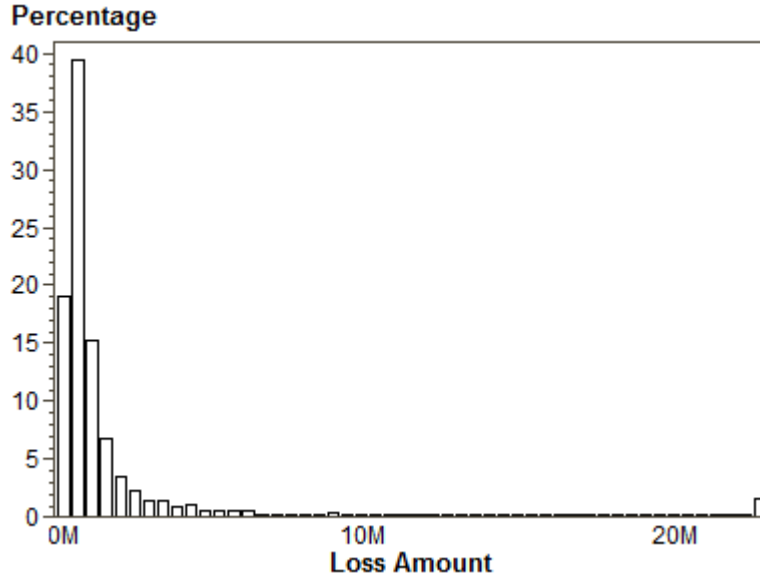


Figure 3.1 Histogram of benchmark dataset.

As seen in the above histogram, the data is extremely heavy-tailed. The last bar is cumulative which indicates that the tail is very long.

For the Champernowne distribution it holds that  $F_{\alpha,M}(M) = 0,5$ . The same holds for the GCD where  $T_{\alpha,M,c}(M) = 0,5$ . This result suggests that  $M$  can be estimated as the median of the dataset of losses which is advantageous due to its simplicity.

When  $M$  has been acquired the next step is to estimate the parameters  $(\alpha, c)$ . A straightforward and relatively simple way is to maximize the following log-likelihood function:

$$\begin{aligned}
 l(\alpha, c) = & n \log \alpha + n \log((M + c)^\alpha - c^\alpha) + (\alpha - 1) \sum_{i=1}^n \log(X_i + c) \\
 & - 2 \sum_{i=1}^n \log((X_i + c)^\alpha + (M + c)^\alpha - 2c^\alpha).
 \end{aligned} \tag{3.1}$$

When  $M$  is fixed the log-likelihood function is concave and has a maximum. The optimization is done numerical with the Newton-Rhapson method. Our estimated parameters  $\hat{\alpha}$ ,  $\hat{M}$  and  $\hat{c}$  are stored for further use throughout the process. The parametric estimation of our dataset is shown below in Figure 3.2.

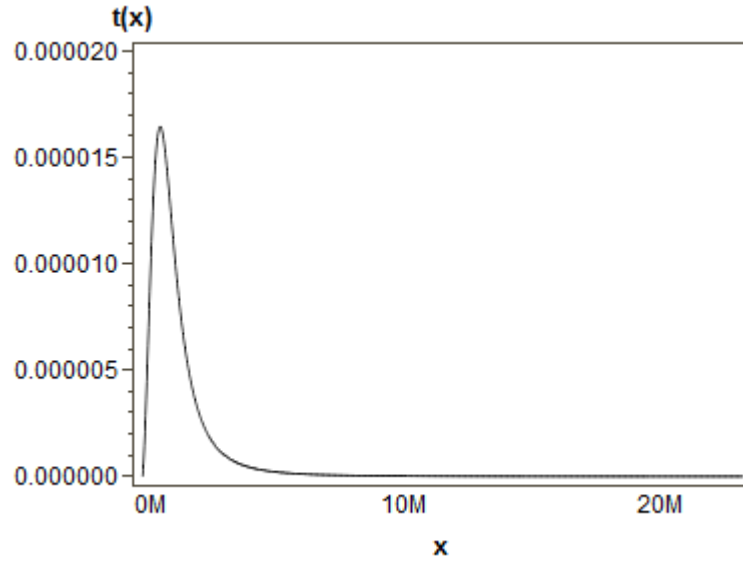


Figure 3.2 Parametric estimation of benchmark dataset.

### 3.2 Transformation

Next step in the process is to transform the desired dataset into the  $[0,1]$  interval. As described in Chapter 2, the dataset  $X_i$  is transformed with the transformation function  $T_{\alpha,M,c}(\cdot)$  so that:

$$Y_i = T_{\hat{\alpha},\hat{M},\hat{c}}(X_i) \quad (3.2)$$

where  $i = 1, \dots, n$ .

The parameter estimation is designed to make the transformed data as close to a uniform distribution as possible. However, even if our transformed dataset does not look like a uniform distribution it does still contain the right characteristics. Figure 3.3 shows a histogram of the transformed dataset.

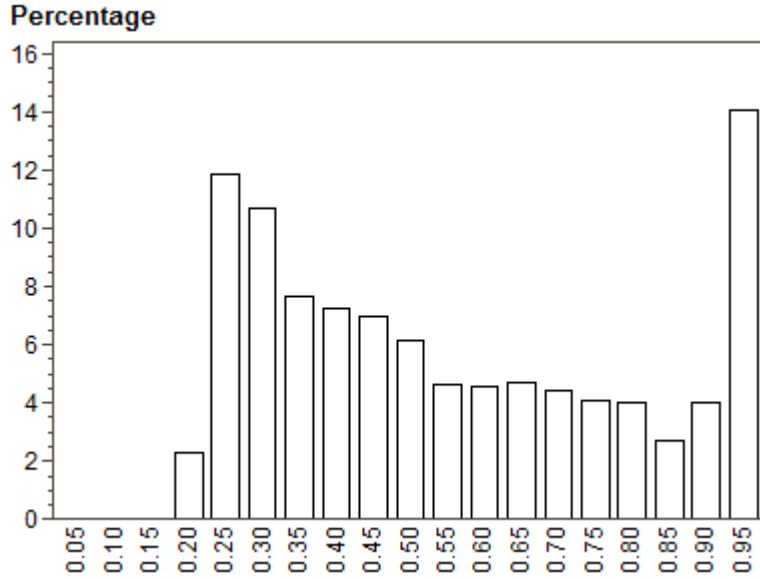


Figure 3.3 Histogram of transformed benchmark dataset.

### 3.3 Kernel Density Estimation

Now that the data has been transformed into the  $[0,1]$  interval we continue with classical KDE where we add a boundary correction term since the data is now bounded. In order to perform the KDE we also have to choose a kernel and an appropriate bandwidth.

#### 3.3.1 Classical KDE with Boundary Correction

The KDE is calculated with the transformed data,  $Y_i, i = 1, \dots, n$ , also, we include the boundary correction term from Section 2.6.1 giving us the following KDE:

$$\hat{f}_{transformed}(y) = \frac{1}{nba_{01}(y, b)} \sum_{i=1}^n K\left(\frac{y - Y_i}{b}\right). \quad (3.3)$$

For simplicity we choose to use the easier boundary correction term  $a_{01}(y, b)$ , defined as:

$$a_{01}(y, b) = \int_{\max\{-1, \frac{y-1}{b}\}}^{\min\{1, \frac{y}{b}\}} K(u) du. \quad (3.4)$$



The boundary correction allows the KDE to almost reach unit integration. However, as it does not quite reach unity and it is the tail we are interested in, we decided to correct for this by dividing the KDE with its integral over the  $[0,1]$  interval. In Bolland et al. (2012), this problem is not addressed. By doing the renormalization we maintain the shape of the pdf, but shift the curve slightly higher. This ensures a pdf where we will be able to examine high confidence levels in the tail.

The kernel we choose to work with is the Epanechnikov because it is the most efficient one according to Zucchini (2003). Furthermore, as the Epanechnikov is bounded, observations in the body do not have a big impact at the tail.

The bandwidth is calculated in line with Silverman's rule-of-thumb as proposed in Gustafsson (2006). Silverman's rule-of-thumb bandwidth aims to minimize the mean integrated square error and is calculated as follows:

$$b = \hat{\sigma} \left( \frac{40\sqrt{\pi}}{n} \right)^{1/5}, \quad (3.5)$$

where  $\hat{\sigma}$  is the empirical standard deviation of the transformed losses. Note that the bandwidth gets smaller as the number of observations increases.

The estimated classical kernel density estimator of the transformed dataset is shown in Figure 3.4 below.

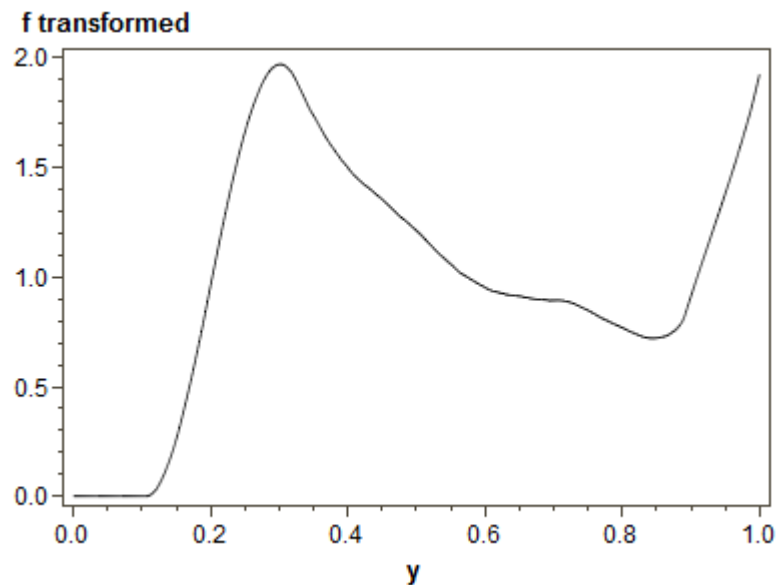


Figure 3.4 *Transformed kernel density of benchmark dataset.*

### 3.4 Back-transformation

We now have a pdf that models the loss distribution in a transformed scale. In order to interpret this pdf properly we need to convert it. Therefore the final step is to perform a back-transformation from the  $[0,1]$  interval to the original scale.

The relationship between two densities  $f_A(a)$  and  $f_B(b)$ , where  $a$  is a transformation of  $b$  such that  $a = t(b)$  where  $t$  is an invertible and differentiable function, is the following:

$$f_B(b) = f_A(t^{-1}(b)) \left| \frac{dt^{-1}(b)}{db} \right|. \quad (3.6)$$

In our case  $x$  can be written as a transform of  $y$  such that  $x = T_{\hat{\alpha}, \hat{M}, \hat{c}}^{-1}(y)$ . This results in the following relationship:

$$\begin{aligned} \hat{f}(x) &= \hat{f}_{transformed}(T_{\hat{\alpha}, \hat{M}, \hat{c}}(x)) \left| \frac{dT_{\hat{\alpha}, \hat{M}, \hat{c}}(x)}{dx} \right| \\ &= \hat{f}_{transformed}(y) T'_{\hat{\alpha}, \hat{M}, \hat{c}}(x), \end{aligned} \quad (3.7)$$

where  $T'_{\hat{\alpha}, \hat{M}, \hat{c}}(x)$  is the derivative of the GCD and equals:

$$T'_{\hat{\alpha}, \hat{M}, \hat{c}}(x) = \frac{a(x+c)^{\alpha-1}((M+c)^\alpha - c^\alpha)}{((x+c)^\alpha + (M+c)^\alpha - 2c^\alpha)^2}. \quad (3.8)$$

Thus, the semiparametric transformation can be summarized as:

$$\hat{f}(x) = \frac{T'_{\hat{\alpha}, \hat{M}, \hat{c}}(x)}{nba_{01}(T_{\hat{\alpha}, \hat{M}, \hat{c}}(x), b)} \sum_{i=1}^n K\left(\frac{T_{\hat{\alpha}, \hat{M}, \hat{c}}(x) - T_{\hat{\alpha}, \hat{M}, \hat{c}}(X_i)}{b}\right). \quad (3.9)$$

The results from performing all steps stated above are shown in Figure 3.5 below.

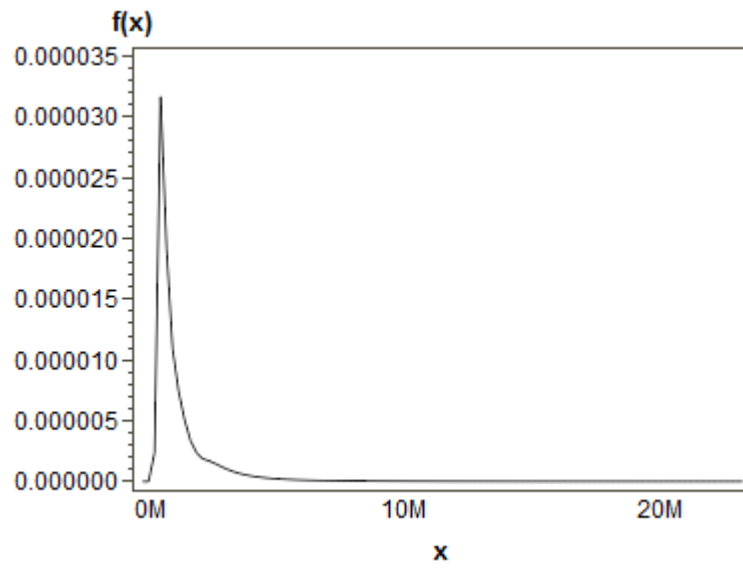


Figure 3.5 Back-transformed data which also is the final distribution.



## 4 Sensitivity Analysis

The method proposed to estimate the severity distribution as seen in Chapter 3 has proven to capture operational losses very well. Bolancé et al. (2012) showed this by comparing the semiparametric method with other parametric methods, e.g. fitting operational loss data to distributions like Weibull and Log-normal etc. In this chapter we will describe our methods for finding relationships and highlight these. Furthermore, we are interested in the sensitivity of changes in the data and what impact it will have on the final outcome.

Operational losses are very hard to predict and they will occur no matter how hard the regulations are. Extreme events or extreme losses are rare but account for a huge mass of the total sum of losses. The actions a financial institution imposes will never eliminate these extreme losses. However, they might be able to control the losses occurring on a daily basis. These losses are commonly more frequent and smaller in size. The sensitivity analysis will therefore focus on how changes in the body affect the estimated severity distribution. Throughout the analysis the changes will be measured in Value-at-Risk at a 99,97 % confidence level.

### 4.1 Benchmark

The simulations are based on a benchmark dataset resembling a typical pattern of operational losses for financial institutions. Throughout this thesis, losses will be presented in relation to the median due to the fact that actual loss size is of no importance. There are 5000 losses in the dataset and the mean is approximately four times the median, indicating we indeed have a heavy-tailed dataset. For convenience we show the histogram of the benchmark data, as seen in Figure 3.1, once more in Figure 4.1 below.

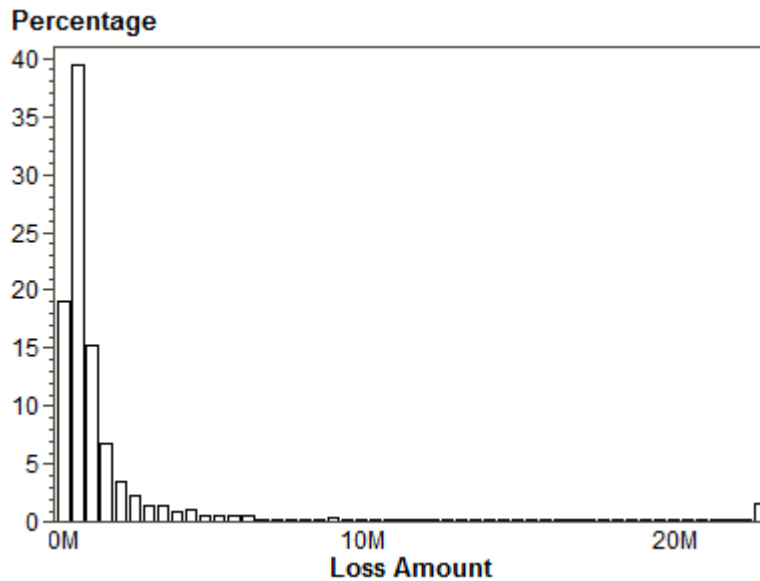


Figure 4.1 Histogram of benchmark dataset.

The dataset will not contain data for every combination of business line and event type that were explained in Section 2.3.1. However, for the purpose of this thesis, the lack of combinations will not be a problem since an aggregated dataset will be used to perform the calculations. When using an aggregated dataset, the complexity of handling dependencies between different business lines and event types and the problem concerning how to incorporate external data will not be an issue.

## 4.2 Body and Tail Manipulation

First we would like to get a decent picture of how changes in the body and/or tail affect the outcome. Initially, to control the data we split it into two categories, a body part and a tail part. Thereafter, we choose a cut-off point from where we fix the tail and generate datasets similar to our benchmark body. The generated datasets are processed through the transformation and finally a 99,97 % VaR is calculated. Then we proceed with systematically multiplying each of the two parts with a factor and combining them. Figure 4.2 below shows how the data is divided between body and tail.

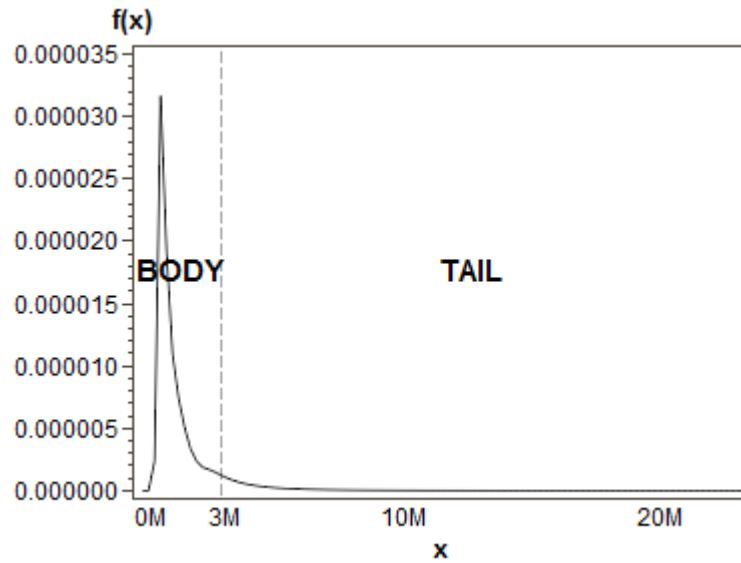


Figure 4.2 Severity distribution divided into a body and tail part.

### 4.3 Modified Body

In the next step we further investigate how modifications to the body alter the estimated severity distribution. This is done by replacing the actual body with generated datasets that have similar features as the benchmark dataset. Once again we divide the data between body and tail, as seen in Figure 4.2.

Since a randomly generated dataset has an overwhelming possibility of being non-meaningful, we choose the Weibull distribution as an approximation of body data. This is because its features are similar to our benchmark up to a certain point. A fixed number of losses are generated systematically in an interval that is supposed to reflect body data. Both the median and standard deviation in our generated dataset are shifted from low to high. The generated dataset is then connected with the tail of actual losses for consistency and because we want to evaluate what changes in body data does to the outcome which is measured in VaR.

Some might argue that this principle of generating datasets are in some way similar to modeling operational risk with the more classical and commonly used EVT where a threshold is chosen to separate body and tail. However, that is not correct because it is only the input dataset that is generated with body and tail separated, the final outcome does still depend on both the body and tail.

## 4.4 Manipulating Interval Data

The approach of generating datasets should give an overall perception of how changing the entire body could affect the outcome. Now, we want to find out how sensitive the model is for changes in smaller subintervals, covering the whole loss domain. We will use the benchmark dataset and divide the domain into three parts; which we will call body, tail and extreme events. We will use smaller intervals in the body part (Figure 4.3), continuing with larger intervals in the tail part (Figure 4.4) and finally much larger intervals in the extreme events part (Figure 4.5). This is due to the relationship of number of losses in each region, with the majority of losses in the body and least number of losses in the extreme events. It is worth mentioning that the number of observed losses in the benchmark dataset within each interval will vary. For convenience, the sectioning has been done in regards to the median. This will give the reader a good perception of the different domains.

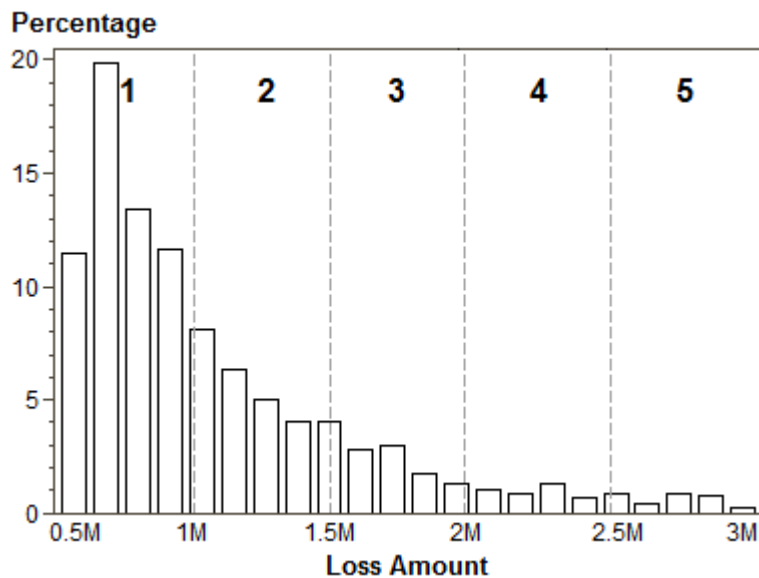


Figure 4.3 Histogram of the body part, interval one to five.

Below follows a table describing the range, the relative loss mass and the relative number of losses of each interval in the body.



Interval	1	2	3	4	5
Interval size	0,5M-1M	1M-1,5M	1,5M-2M	2M-2,5M	2,5M-3M
Relative loss mass	8,8%	5,8%	3,7%	2,0%	1,7%
Relative number of losses	50,0%	20,6%	9,3%	3,7%	2,7%

Table 4.1 Descriptive statistics of the body part.

From Table 4.1 above, we can see that the first interval contains 50 % of all the losses in number, but only 8,8 % relative to the total mass.

Next part is the main tail. The size of each interval has been increased from half a median to four medians. This section is also divided into five different parts as seen in Figure 4.4 below.

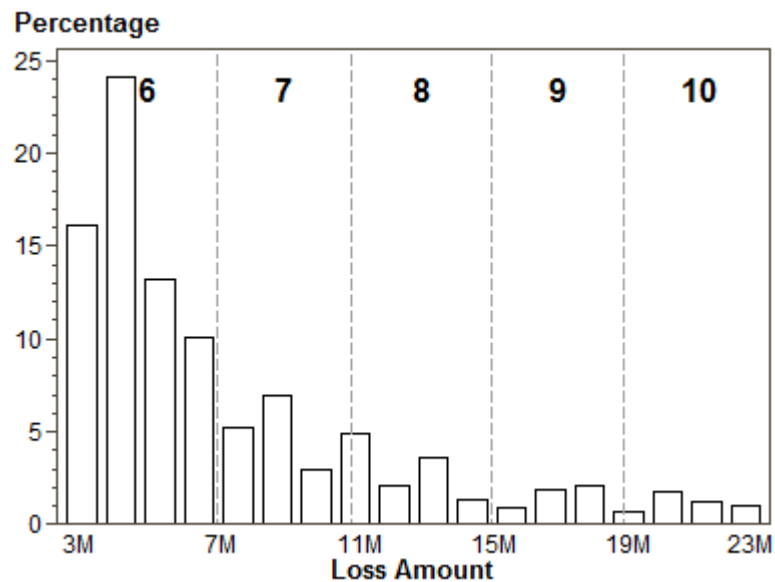


Figure 4.4 Histogram of the tail part, interval six to ten.

Each interval is eight times bigger than in the body part. It can be seen in Table 4.2 below that the relative loss mass varies from 2,7 % to 8 % and the relative number of losses varies from 0,5 % to 7,6 %.

Interval	6	7	8	9	10
Interval size	3M-7M	7M-11M	11M-15M	15M-19M	19M-23M
Relative loss mass	8,0%	4,5%	3,5%	2,4%	2,7%
Relative number of losses	7,6%	2,2%	1,2%	0,6%	0,5%

Table 4.2 Descriptive statistics of the main tail part.

The last part of the sectioning is the extreme events. The intervals used are much bigger, namely 600 times the median, see Figure 4.5. This indicates that the dataset is extremely heavy tailed.

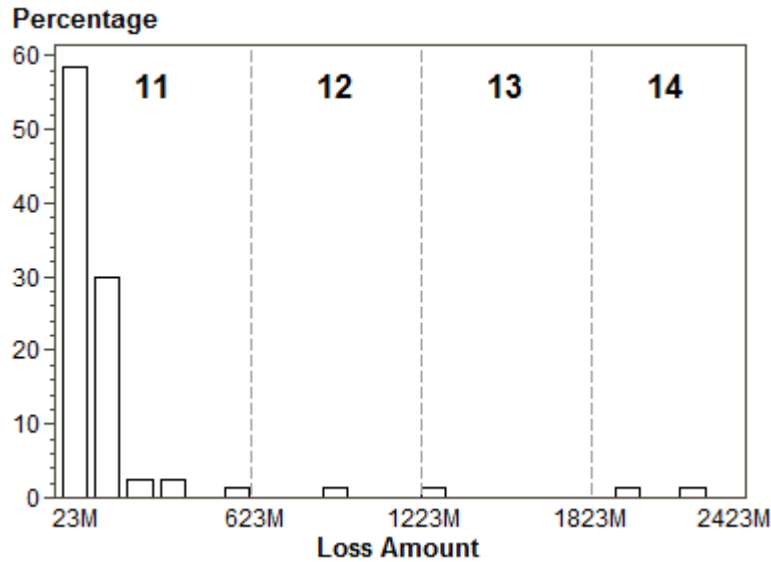


Figure 4.5 Histogram of the tail part, interval 11 to 14.

In Table 4.3 below, it can be noted that even though the relative number of losses in each of these intervals is very low, they constitute a large relative loss mass.

Interval	11	12	13	14
Interval size	23M-623M	623M-1223M	1223M-1823M	1823M-2423M
Relative loss mass	26,2%	4,7%	6,2%	20,0%
Relative number of losses	1,5%	0,02%	0,02%	0,04%

Table 4.3 Descriptive statistics of the extreme events part.

To test the sensitivity in each interval we will use two different approaches. The first approach is to increase the size of the losses in the interval and the second approach is to increase loss frequency in the interval.

#### 4.4.1 Change Size of Losses

When changing the size of losses we systematically multiply all the losses within an interval with a factor, while keeping the other intervals unchanged. The multiplication factor in this test will vary from 0,5 to 1,5. Intuitively, this could be thought of as a

way of moving losses from one interval into another and is done to see what parts of the distribution are the most sensitive. By using this technique the total number of loss observations will remain constant.

#### **4.4.2 Increase Frequency of Losses**

We use two methods of increasing the frequency of losses within each interval. In the first method we systematically add a fixed number of losses to an interval while keeping the other intervals constant. The size of the added losses is the mean size of the corresponding interval for simplicity. The effect of adding a fixed number of losses to each interval will have different impact depending on the number of losses already in the interval.

In the second method, instead of adding a fixed number of losses we let the number of added losses depend on how many losses the interval already contains, i.e. we increase the number of losses relative to the actual quantity. We separate these two tests because there is a difference in number of losses in each interval and it is desired to see both what a single loss does to the final outcome as well as a percentage increase of losses in each interval. Just as in the test described in Section 4.4.1, it is done to see what parts of the distribution are the most sensitive.

#### **4.5 Repeated Ten-multiplier Stress Test**

As an additional exercise in the sensitivity analysis, the benchmark dataset is stressed systematically. Basically, different fictive losses which alter in number and size are added. The number of losses that are added varies from one to a thousand and the size of these losses varies from 1M to 100 000M, both increasing with a ten multiplier from the initial value. Each of these tests is done in isolation from each other.

#### **4.6 Probability Assignment**

Many of the tests proposed above consist of extreme and perhaps improbable situations. For that reason we pursue a test that gives us some indication of the probability that different scenarios occur.

The test will assume that the loss structure in the benchmark dataset is constant, i.e. there is an equal probability of each loss in the dataset to materialize again. Following that logic, a large number of new datasets will be created by drawing losses from the benchmark dataset. The new datasets will have the same number of losses as the benchmark dataset, but the concentration of losses will differ slightly from each other.

Finally, the new datasets will be categorized differently depending on their appearance with the two-sample Kolmogorov-Smirnov test.

## **4.7 Bandwidth Selection**

Apart from testing how the model reacts when manipulating the data, we want to test how stable the bandwidth assumption is. The choice of bandwidth can be tested by computing the VaR of the benchmark dataset for different selections of bandwidth. The bandwidths we test are the benchmark dataset bandwidth according to Silverman's rule-of-thumb multiplied by a factor from 0,8 to 1,2.

## 5 Results

The following sections will present the results obtained from the tests specified in Chapter 4. The results presented will mainly be displayed through change in VaR, where the initial dataset and its describing statistics will be the benchmark. The data will be presented in relation to the median, i.e. if referred to 2M, it stands for a value twice the size of the median.

### 5.1 Body and Tail Manipulation

As an initial test, we wanted to see what would happen to the VaR when we changed the size of either the body or tail. The figure below presents how the change in VaR is affected by different combinations of factors, which are multiplied to the tail and body part described in Figure 4.2 respectively.

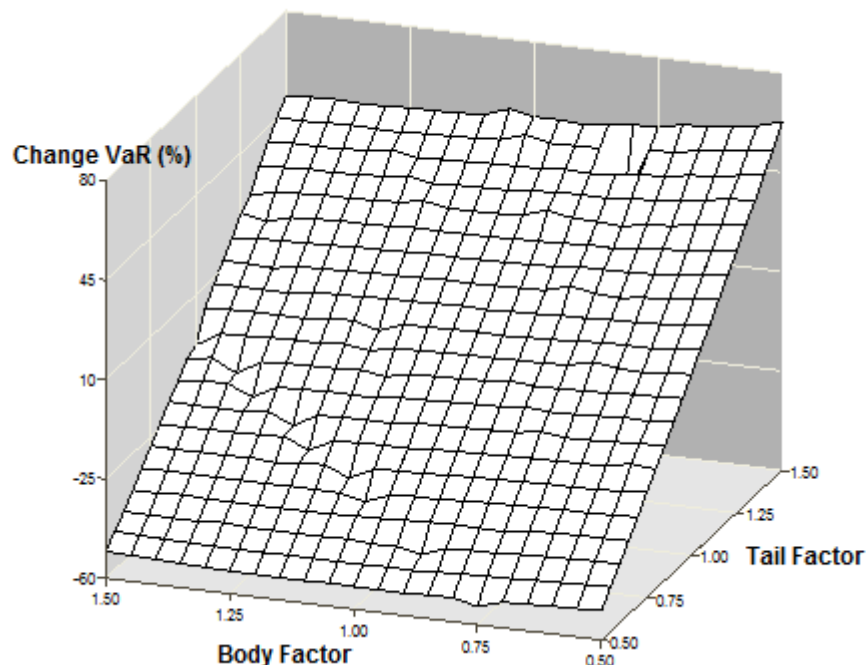


Figure 5.1 Benchmark dataset divided in body and tail, then multiplied with a factor.

From Figure 5.1 we can see that there is an increase in VaR when the tail is multiplied with an increasing factor. Whereas, when the body is multiplied with an increasing factor, the VaR is decreasing. However, the impact from change in tail size is much greater than change in body size.

The highest VaR is obtained when the tail factor is high and the body factor low, resulting in a 62 % increase. Conversely, the lowest VaR comes from the combination of a low tail factor and a high body factor, which gives a 52 % decrease.

The relationship between the two factors and change in VaR can be explained with the help of following figure.

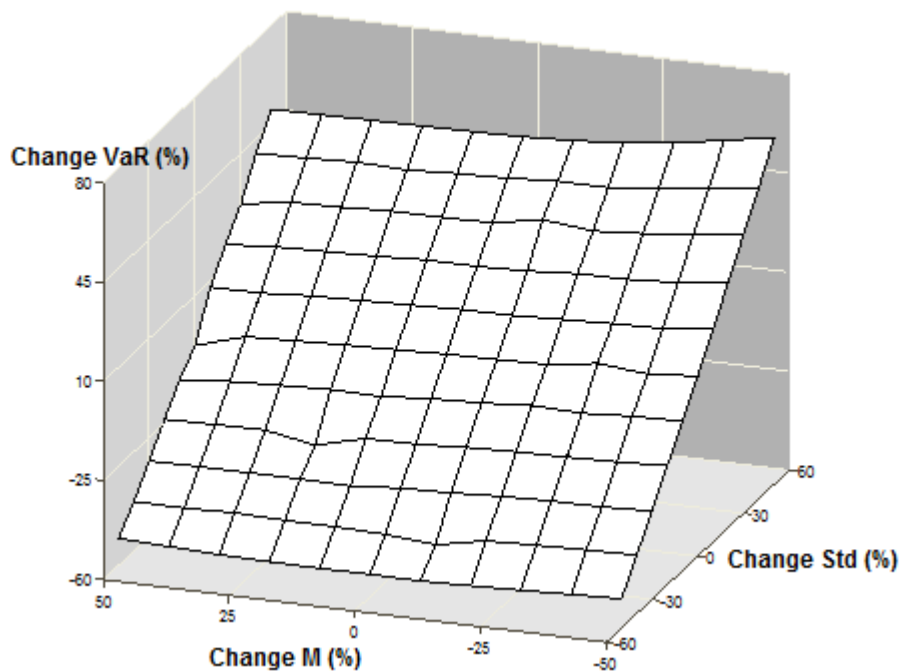


Figure 5.2 Change in VaR in relation to median and standard deviation.

Figure 5.2 presents the change in VaR in relation to the change in median and standard deviation. Figure 5.1 and Figure 5.2 are very similar in their shapes which might indicate a strong relationship between the body and median as well as the tail and standard deviation. It seems as the median and standard deviation are parameters that drives the change in VaR. This observation is further confirmed from Figure 5.3 and Figure 5.4 below.

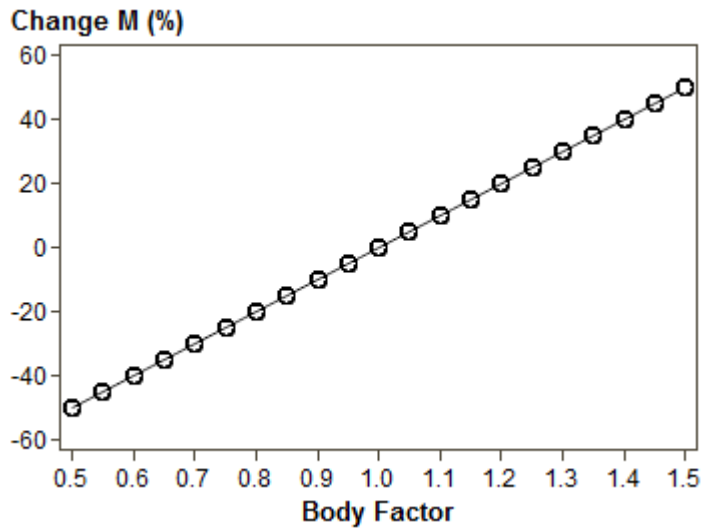


Figure 5.3 Change in median in relation to the body factor.

Figure 5.3 displays a linear relationship between the factor which is multiplied with the body data and the size of the median. Moreover, below is a figure which displays a linear relationship between the factor which is multiplied with the tail and the standard deviation.

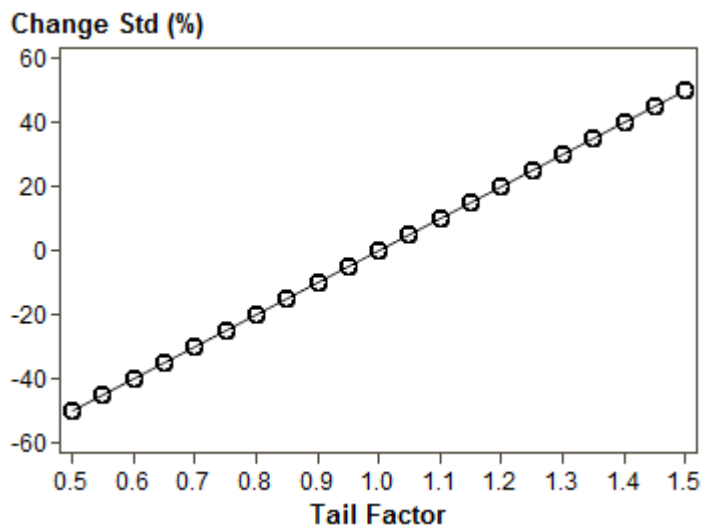


Figure 5.4 Change in standard deviation in relation to the tail factor.

In this test, the body and tail has been divided at a chosen point of 3M, which might seem arbitrary but is chosen graphically. If the separation between body and tail instead is done at 2,5M, the following result presented in Figure 5.5 is obtained.

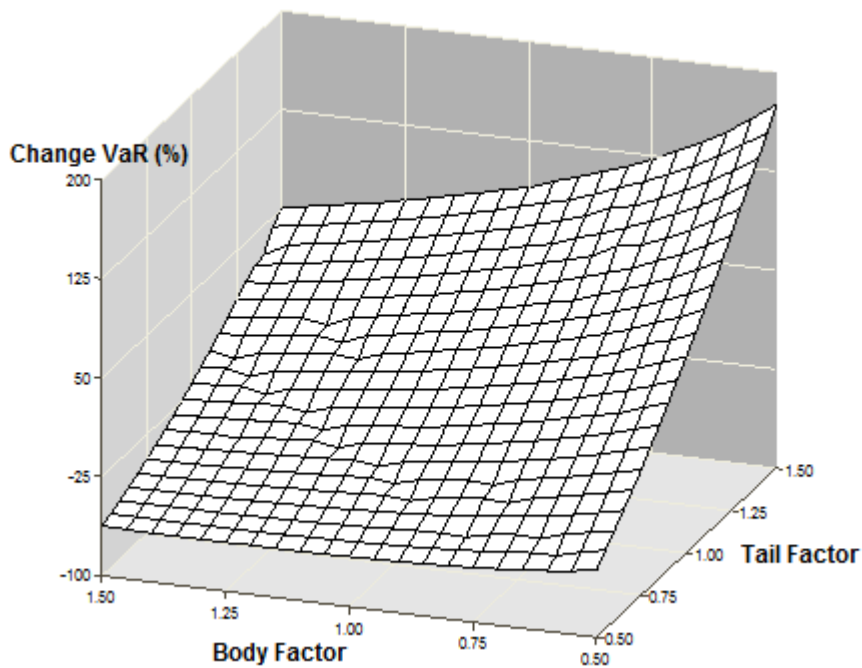


Figure 5.5 Factor multiplication with body defined as  $2,5M$ .

If Figure 5.1 and Figure 5.5 are compared, a significant difference can be seen. It appears that the body, or the median, has a greater impact when a smaller interval is used. Also, the effect from the standard deviation increases which comes naturally because the mass in the tail increases. Doing the exact same tests as described above, we can see that the maximum increase in VaR is 175 % and the biggest decrease in VaR is 63 %.

The purpose of this test was to get an overall perception of how the distribution behaves when body and tail is moved. The reader should consider the fact that the relative changes depend heavily on where the point of sectioning is done.

## 5.2 Modified Body

The results from modifying the body are presented below. Remember that the tail is kept constant as well as the number of observations. It is only the appearance of the body that is changed.

In the first test we choose to have a very narrow peak in the body, this peak is then shifted from a very low value to just before the tail starts.



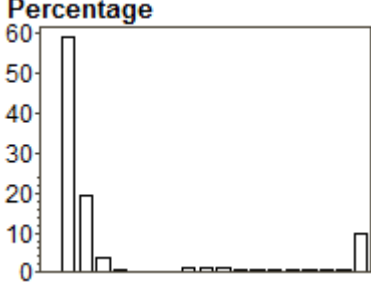
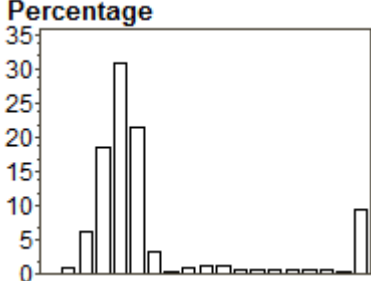
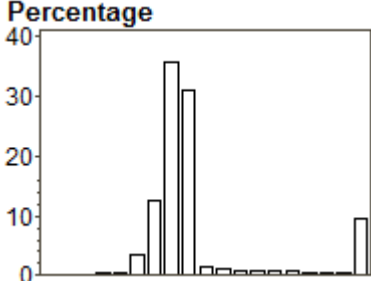
Plot	Change M (%)	Change Mean (%)	Change Std (%)	Change VaR (%)
	-24,09	-5,03	0,03	-28,66
	45,52	6,80	-0,04	-52,00
	118,83	20,77	-0,12	-69,24

Table 5.1 *Narrow peak in the body.*

As seen in Table 5.1 above, there is a negative change in VaR for all three input datasets. The finding that VaR is decreasing when the peak is moved closer to the tail is somewhat surprising. Except for the case where we have a very low peak, the median is increased compared to our benchmark. Furthermore, the standard deviation is changed from negative to positive in the same cases. Note that a small change in standard deviation might have a big impact on the outcome.

In the second test, the peak is still changed from a very low value to just before the tail starts. However, it has a wider appearance.

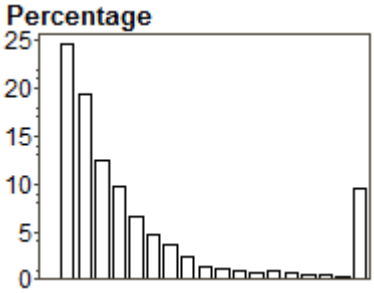
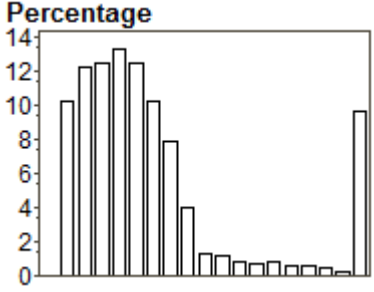
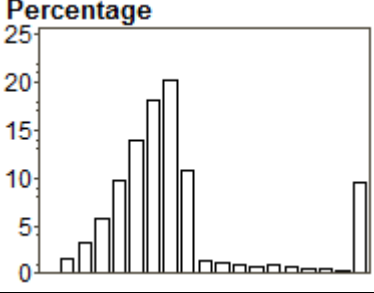
Plot	Change M (%)	Change Mean (%)	Change Std (%)	Change VaR (%)
	14,80	1,94	-0,01	-0,11
	52,71	7,15	-0,04	-14,86
	93,79	14,20	-0,08	-47,17

Table 5.2 *Wide peak in the body.*

In Table 5.2 above, it can be seen that the first case is very similar to our benchmark. The median has increased by almost 15 % and the standard deviation has decreased with 0,01 %. The other two cases confirm what could be seen in Table 5.1. A rising median and a decreasing standard deviation lowers the VaR significantly.

### 5.3 Manipulating Interval Data

As described in Chapter 4, the dataset is broken down into different intervals to better control changes. We will conduct each test or evaluation described in Section 4.4 systematically for each interval. The results shown in each figure are percentage change in VaR. Detailed results from each test can be found in Appendix.

### 5.3.1 Change in Size of Losses

In the first test we change the size of the losses for each interval by multiplying them with a factor. We begin with the first five intervals, corresponding to the body part.

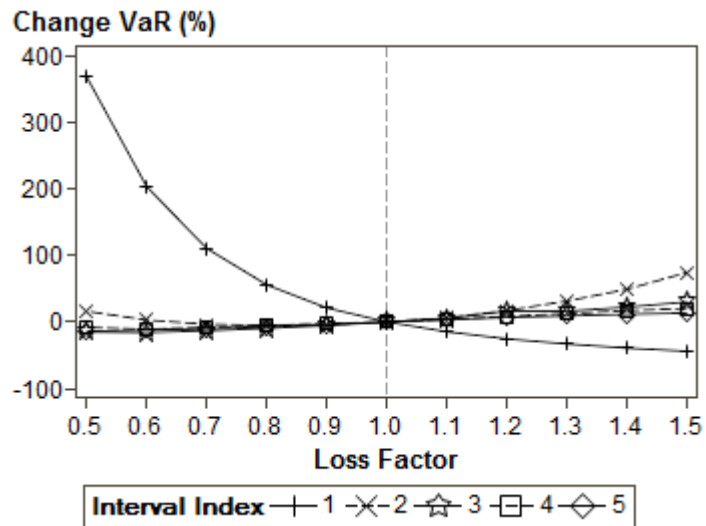


Figure 5.6 Results of changing the size of losses in the body part.

Interval one and two have according to Figure 5.6 the biggest impact on the VaR. When amplifying losses, interval one has a negative relationship to the VaR, i.e. it lowers when the factor is increasing. Interval two and upwards have a positive relationship which in this case means that while increasing the factor, the VaR increases. Conversely, when reducing the losses interval one has a great positive impact on VaR, while interval three to five has a slight negative effect on VaR. Interval two on the other hand has to start a negative impact, but as the loss factor decreases, VaR increases again. For detailed results, see Table A.1 in Appendix.

We continue with interval six to ten, the tail part. The results are presented in Figure 5.7 on the next page.

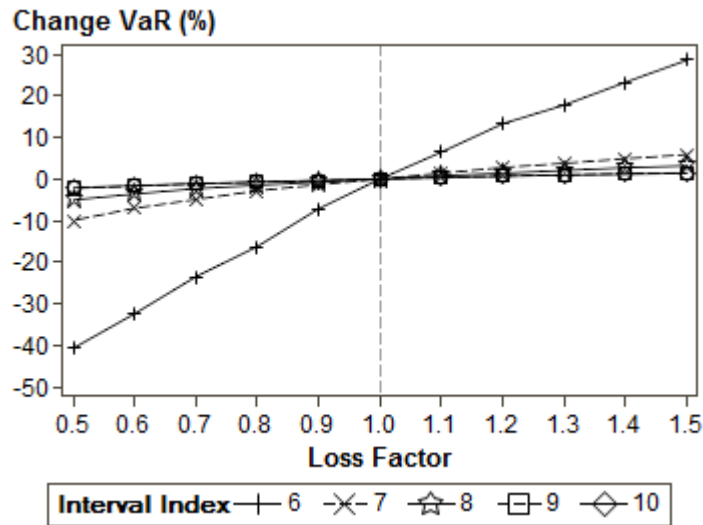


Figure 5.7 Results of changing the size of losses in the tail part.

The first observation that can be seen from amplifying the losses in each interval is a consistently increasing VaR, where interval six has the greatest impact. The impact diminishes for each consecutive interval, resulting in changes no greater than 6%. On the other hand, when reducing the losses the VaR is consistently decreasing. However, compared to interval one to five the impact is greatly reduced. Note that the range of change in VaR has decreased significantly in comparison. This could be explained by the number of losses in each interval. For detailed results, see Table A.2 in Appendix.

The last domain, the extreme events part is shown below.

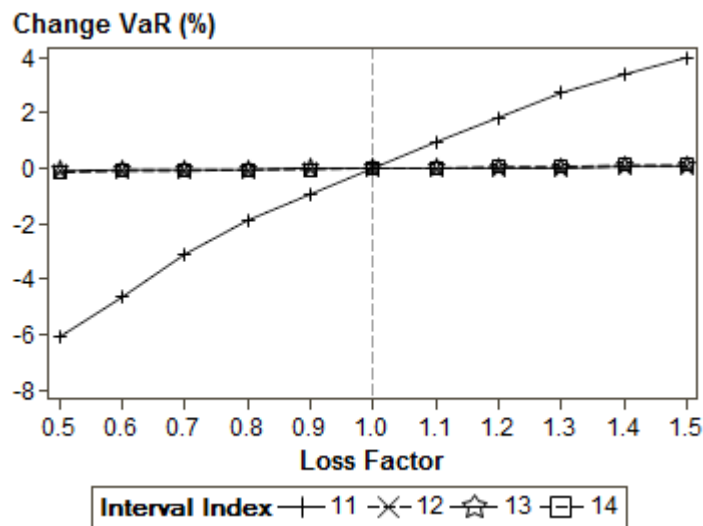
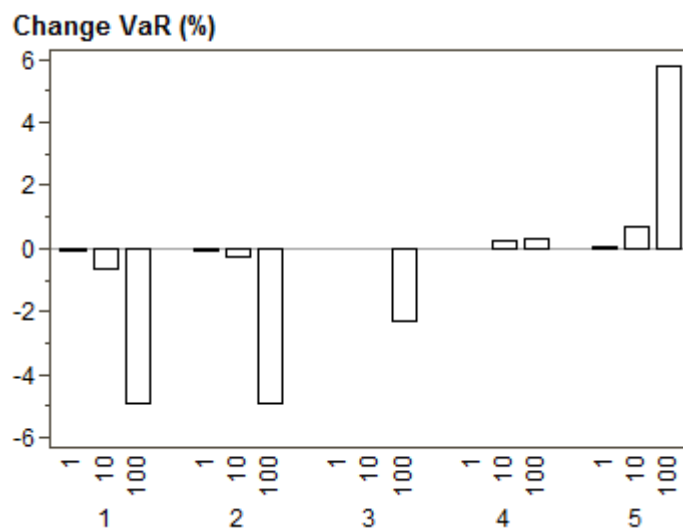


Figure 5.8 Results of changing the size of losses in the extreme events part.

Once again, the range of change in VaR has decreased significantly in comparison. It is only interval eleven that has some impact in this test which can be seen in Figure 5.8. Interval 12 to 14 has a negligible effect on the VaR. For detailed results, see Table A.3 in Appendix.

### 5.3.2 Increase Frequency of Losses

The second test described in Section 4.4, losses are added to the dataset resulting in a change of frequency. Initially, a fixed number of losses are added to each interval systematically, starting with the body part or interval one to five.



**Figure 5.9** Results of adding a fixed number of losses (1,10 and 100) to the benchmark dataset in the body part, interval one to five.

In the first three intervals adding losses have a negative impact on the VaR while interval four and five has a positive impact as seen in Figure 5.9 above. Interval four is the least sensitive interval and there is not much change when adding either one, ten or a hundred losses. Adding losses have a greater impact on VaR the further from interval four the losses are added. For detailed results, see Table A.4 in Appendix.

We continue with interval six to ten, the tail part. The results are presented in Figure 5.10 on the following page.

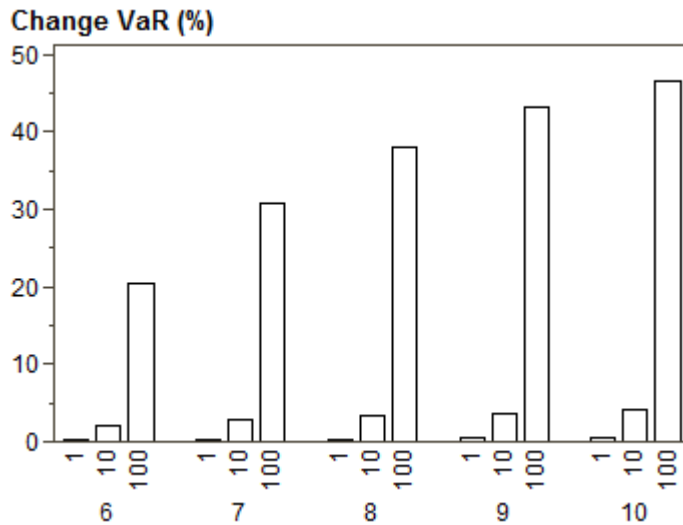


Figure 5.10 Results of adding a fixed number of losses (1,10 and 100) to the benchmark dataset in the tail part, interval six to ten.

When a fixed number of losses are added to each consecutive interval, the impact on VaR increases. It can be seen that there is a linear relationship between the number of losses added and the change in VaR, i.e. adding one loss in interval six gives a 0,2 % increase in VaR while adding ten losses in the same interval account for a 2 % increase. Furthermore, compared to interval one to five the impact is greatly increased. Note that the range of change in VaR has increased significantly in comparison. For detailed results, see Table A.5 in Appendix.

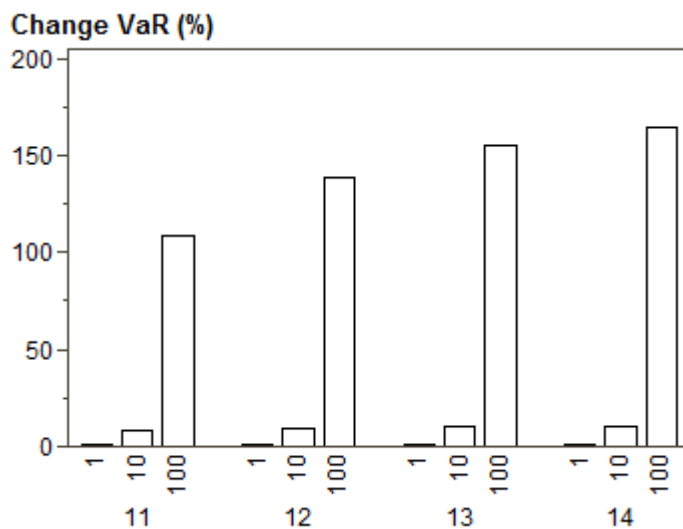
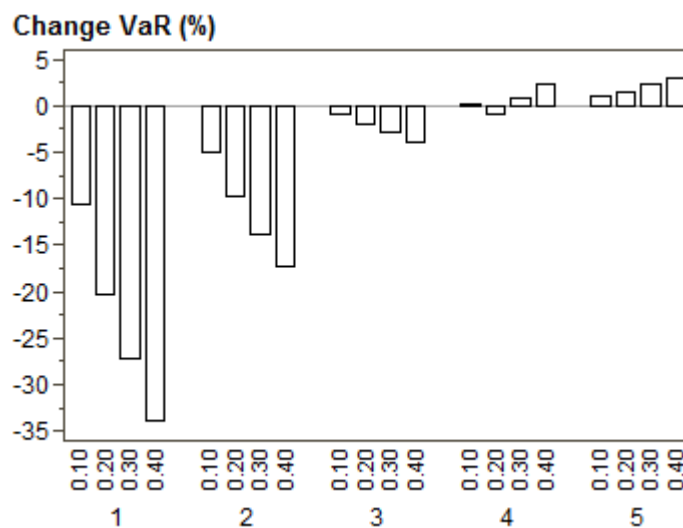


Figure 5.11 Results of adding a fixed number of losses (1,10 and 100) to the benchmark dataset in the extreme events part, interval 11 to 14.

Continuing with the extreme events part, results are presented in Figure 5.11. Consistently with the tail part in interval six to ten, VaR increases almost linear for each loss added to the respective interval. Moreover, the intervals have a bigger impact the further up in the distribution we get. That is, adding a bigger loss has more impact than a smaller loss. Also, note that the range of change in VaR has increased significantly compared to the body and tail part. For detailed results, see Table A.6 in Appendix.

The third test conducted is testing the effect of adding a relative frequency in each interval. Instead of adding a single loss, ten or a hundred, a percentage increase of losses added are imposed in the range from 10 % to 40 %. We start with the body part, interval one to five.



**Figure 5.12** Results of adding a variable number of losses (10 %, 20 %, 30 % and 40 %) to the benchmark dataset in the body part, interval 1 to 5.

As seen in Figure 5.12, once again, interval one, two and three have a negative impact on VaR and interval four is the least sensitive. Interval four to five are almost consistently increasing for each increase in frequency. The changes in the first two intervals are much greater than the others, mainly due to the large number of observations in each of them. By adding a relative number of losses, the greater significance the interval get for the final outcome. For example, by adding 10 % losses in interval one, it corresponds to 5 % of the total number of losses in the benchmark dataset. However, adding 10 % in interval five only accounts for approximately 0.3 % of the total number of losses added to the dataset. For detailed results, see Table A.7 in Appendix.

Next part of the sectioning is the tail part, interval six to ten.

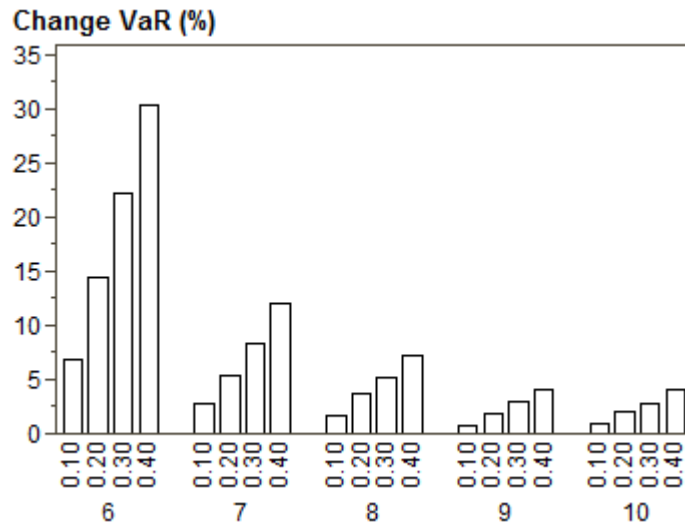


Figure 5.13 Results of adding a variable number of losses (10 %, 20 %, 30 % and 40 %) to the benchmark dataset in the tail part, interval 6 to 10.

As seen in Figure 5.13 above, VaR is increasing with the relative frequency. However, further up in the intervals, fewer observations there are. Hence, a less impact on the final outcome. For detailed results, see Table A.8 in Appendix.

This approach was not conducted for the extreme events part because of the low frequency of observations in each interval. Adding a relative number of losses would therefore be non-meaningful.



## 5.4 Repeated Ten-multiplier Stress Test

In Table 5.3 below, the change in the median, mean, standard deviation and VaR are displayed in percentage as an outcome from the number and size of losses added to the original dataset.

Number of losses added	Size of loss	Change			
		Change M (%)	Mean (%)	Change Std (%)	Change VaR (%)
1	1M	0,00	-0,02	-0,01	-0,06
1	10M	0,00	0,03	-0,01	0,33
1	100M	0,00	0,47	0,03	0,61
1	1000M	0,00	4,86	4,10	0,94
1	10 000M	0,00	48,76	207,48	1,22
1	100 000M	0,00	487,74	2810,48	1,55
10	1M	0,00	-0,16	-0,10	-0,50
10	10M	0,07	0,28	-0,10	2,88
10	100M	0,07	4,66	0,28	6,21
10	1000M	0,07	48,48	35,41	9,37
10	10 000M	0,07	486,64	823,17	12,42
10	100 000M	0,07	4868,27	9081,58	15,35
100	1M	0,00	-1,57	-1,03	-5,76
100	10M	1,82	2,73	-1,02	32,71
100	100M	1,82	45,74	2,66	79,43
100	1000M	1,82	475,82	200,53	141,02
100	10 000M	1,82	4776,63	2750,36	222,12
100	100 000M	1,82	47784,72	28397,56	328,49
1000	1M	0,00	-13,27	-9,03	-40,02
1000	10M	20,19	23,04	-8,96	935,98
1000	100M		N/A		
1000	1000M		N/A		
1000	10 000M		N/A		
1000	100 000M		N/A		

Table 5.3 Repeated ten-multiplier stress test.

It is worth noting that the model could not calculate a VaR when 1000 losses of 100M and higher were added. The model itself is able to estimate the distribution, but the VaR is calculated numerical and does not cover such a great increase. However, 1000 losses account for approximately 20 % of total number of losses in the dataset.

The results in Table 5.3 above tell us that the model reacts even when a single loss is added. Depending on the size of the loss added, the parameters will change increasing-

ly. The largest loss in the benchmark dataset is approximately 2400M which means that 10 000M and 100 000M are outside the original domain. A single loss outside the domain has naturally a bigger impact on the parameters. Although, a single loss that size does not alone control the whole distribution, which can be seen in the VaR change for these losses. The mean and standard deviation increases drastically, but the strength of this model is that it considers the whole range of losses, hence the slight change in VaR.

## 5.5 Probability Assignment

As described in Section 4.6, a large number of datasets were generated from the loss structure in the benchmark dataset and compared to each other. The generated datasets change in VaR in relation to the benchmark dataset is presented in Figure 5.14 below.

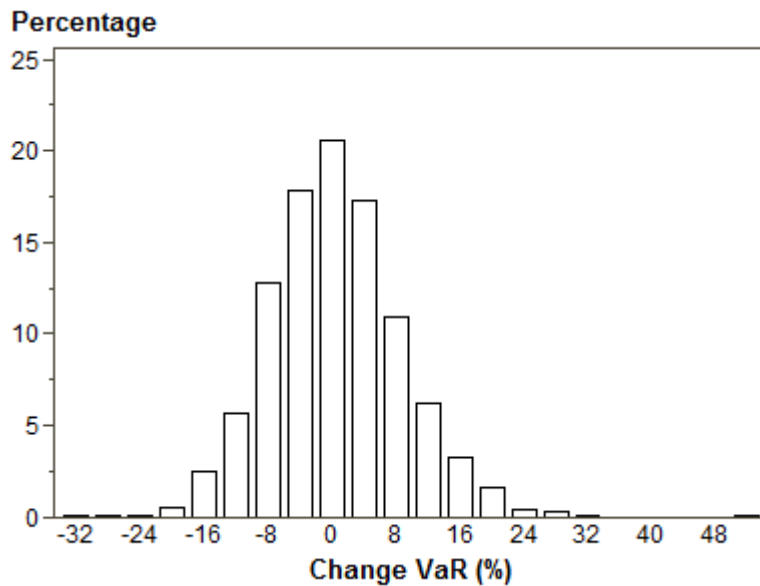


Figure 5.14 Histogram of the change in VaR in datasets generated from the benchmark dataset.

The generated datasets could be categorized into five different distributions, as can be seen in Table 5.4 on the following page.

Distribution Category	Probability (%)	Change M (%)	Change Mean (%)	Change Std (%)	Change VaR (%)
1	99,49	0,00	0,00	0,00	0,00
2	0,25	-1,32	8,12	15,87	-1,27
3	0,22	2,64	5,37	5,90	-2,44
4	0,02	1,46	35,06	46,36	30,88
5	0,02	-2,39	-18,00	-24,09	-25,17

Table 5.4 Results of categorizing datasets generated from the benchmark dataset.

Even though a fairly low confidence level at 90 % was chosen for the two-sample Kolmogorov-Smirnov test, there is an overwhelming majority of distributions that are categorized as the benchmark dataset itself. 99,49 % of the datasets were categorized as distribution category one, which corresponds to the benchmark dataset. The other distribution categories had a change in VaR ranging from -25 % to +31 %, all with very low probabilities. This indicates that with the loss structure of the benchmark dataset, the VaR is fairly stable according to the proposed model.

## 5.6 Bandwidth Selection

From equation 3.5 we know that our choice of bandwidth, Silverman's rule-of-thumb, depends on the standard deviation and the number of losses in the dataset. The bandwidth decreases with an increasing number of observations in the dataset and increases with a rising standard deviation. The figure below shows change in VaR when the benchmark dataset is kept constant and the bandwidth is changed, i.e. the original bandwidth is multiplied with a factor.

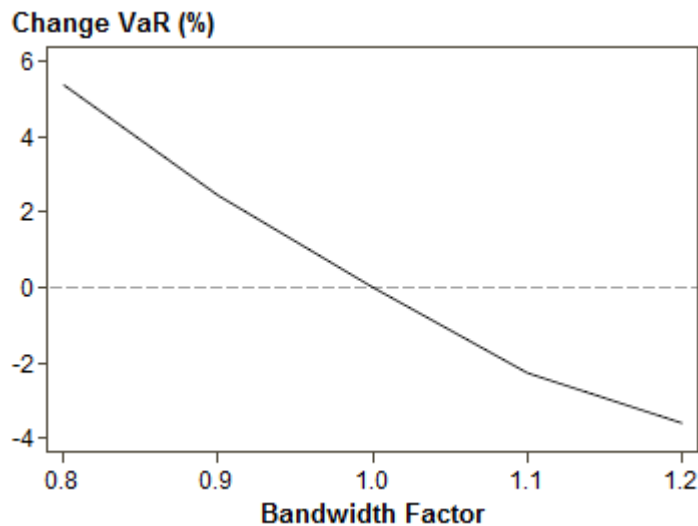


Figure 5.15 Bandwidth impact on the change in VaR.

From Figure 5.15, it can be seen that the choice of bandwidth has an impact on the tail behavior. When a smaller bandwidth is used, the VaR increases. Conversely, when a higher bandwidth is used, the VaR decreases.

As an additional test, we chose to double our benchmark dataset, i.e. an identical dataset was appended to the original dataset. Because it was an identical dataset, the standard deviation was kept almost constant. However, the number of observations doubled. By doing this, the bandwidth was reduced by 13 % and the change in VaR increased with almost 3 %.

## 6 Conclusions and Further Research

Modeling operational risk with the method proposed in Bolancé et al. (2012) seems promising. We have seen that the body does indeed have an impact on the whole severity distribution, enabling a financial institution to link everyday operational risk management with capital requirement imposed by BCBS. As far as we can see, the method is definitely a viable alternative to the more commonly used EVT.

A single loss added has an impact in the whole distribution, it is small but noticeable. All events will subsequently influence the tail of the severity distribution and in the extent also the capital requirement. Added losses have the least effect in interval three and four. Losses added below these intervals have a negative effect on VaR which intuitively makes sense since the loss structure changes slightly, i.e. there is a higher probability for a small loss. Above interval four, losses added have an increasing positive effect. Not surprisingly, losses added in the extreme events part have the most impact.

It should be noted that the frequency is also changed when adding losses. Hence, when calculating the capital requirement with the LDA, we will have a higher probability of losses occurring, which means that even though we have a decrease in VaR we can still obtain a higher capital requirement.

We have seen that even when we add losses that are much bigger than the highest loss in the benchmark dataset, the model is still able to capture the severity distribution. It is responsive, but not overly volatile. When the number of losses added constitutes a substantial mass of the total number of losses, the VaR increases heavily.

Moreover, the bandwidth parameter was evaluated with the conclusion that it does have an effect on the distribution. When the bandwidth was increased the VaR decreased and conversely when the bandwidth decreased the VaR increased. As the Silverman's rule-of-thumb is used, the bandwidth is fairly stable even when the benchmark dataset was doubled in size, which can be seen as a rather large change imposed on the dataset. Since the bandwidth seems not to be very volatile, it will only have a minor effect in the severity distribution.

From the probability assignment it can be seen that if the current loss structure of the benchmark dataset is assumed as stable, the proposed model does not deviate significantly in the VaR. However, it is still responsive to changes in the dataset, implying that the model enable small changes to have an impact on the severity distribution, facilitating everyday operational risk management.

## 6.1 Further Research

The focus of this thesis has been entirely on the severity distribution used in the LDA. To compute a proper minimum capital requirement, which is the final procedure, there are a range of factors that needs to be considered. First of all, the operational loss data has to be divided into the LDA Operational Risk Matrix described above. When doing this, the probability of having a satisfactory number of losses in each cell is very low; hence, external data needs to be added. The question is then, how can we add external data to internal data given that external data often contains large number of losses that might not be valid for another financial institution. Scaling the data will then be necessary because it will most definitely have a huge impact of the final outcome. However, this is easier said than done, appropriate scaling of the data can be a big issue.

Furthermore, many of the existing losses are not, as assumed in this thesis, independent of each other. Dependencies are a tricky question that needs to be considered. There will be dependencies that might have an impact. This does not necessarily mean that the minimum capital requirement will increase by including dependencies.

As a final suggestion of concerns to further investigate, we suggest that a more advanced boundary correction method is evaluated than the one presented in this thesis. In the KDE where the data overflows the boundaries, it needs to be corrected for. By doing the simplest correction, further compensation is necessary as described in this thesis. A more sophisticated method could give a better result.

## 7 References

- Basel Committee on Banking Supervision, (2013), *A brief history of the Basel Committee*, Available at <https://www.bis.org/bcbs/history.htm>.
- Basel Committee on Banking Supervision, (2011), *Operational Risk- Supervisory Guidelines for the Advanced Measurement Approaches*, Available at <http://www.bis.org/>.
- Basel Committee on Banking Supervision, (2009), *Observed range of practice in key elements of Advanced Measurement Approaches (AMA)*, Available at <http://www.bis.org/>.
- Basel Committee on Banking Supervision, (2006), *The First Pillar – Minimum Capital Requirements*, Available at <http://www.bis.org/>.
- Booth, P., Chadburn, R., Haberman, S., James, D., Khorasanee, Z., Plumb, R.H. and Rickayzen, B., (2005), *Modern Actuarial Theory and Practice*, Chapman & Hall/CRC, 2<sup>nd</sup> edition.
- Buch-Larsen, T., Nielsen, J.P., Guillén, M., and Bolancé, C., (2005), *Kernel density estimation for heavy-tailed distributions using the Champernowne transformation*, *Statistics*, 39(6): page 503–518.
- Champernowne, D. G., (1952), *The graduation of income distributions*, *Econometrica*, 20: page 591–614.
- Di Clemente, A. and Romano, C., (2003), *A Copula-Extreme Value Approach for Modelling Operational Risk*, Risk Books.
- Epanechnikov, V.A., (1969), *Non-parametric estimation of a multivariate probability density*, *Theory of Probability and its Applications*, 14: page 153–158.
- Frachot, A., Georges, P. and Roncalli, T., (2001), *Loss Distribution Approach for operational risk*, Working paper, Groupe de Recherche Opérationnelle, Crédit Lyonnais.
- Gasser, T. and Müller, H.G., (1979), *Kernel Estimation of Regression Functions*. In *Smoothing Techniques for Curve Estimation*, Springer-Verlag.

Gustafsson, J., (2006), *Modelling operational risk severities with kernel density estimation using the Champernowne transformation*, Working paper.

Jones, M.C., (1993), *Simple Boundary Correction for Kernel Density Estimation*, *Statistics and Computing*, 3: page 135-146.

Silverman, B.W., (1986), *Density estimation for statistics and data analysis*, Chapman & Hall.

Zhang, S. and Karunamuni, R.J., (1998), *On Kernel Density Estimation Near Endpoints*, *Journal of Statistical Planning and Inference*, 70: page 301-316.

Zucchini, W., (2003), *Applied Smoothing Techniques – Part 1: Kernel Density Estimation*, Technical Report.



# Appendix

Loss Factor	Interval					
		1	2	3	4	5
0,5		367,35	15,96	-14,86	-8,15	-14,52
0,6		202,66	2,88	-15,80	-10,92	-11,81
0,7		110,20	-3,44	-13,14	-8,09	-9,92
0,8		55,54	-5,21	-9,48	-5,43	-7,21
0,9		21,01	-3,77	-5,16	-2,49	-3,88
1,0		0,00	0,00	0,00	0,00	0,00
1,1		-14,69	6,21	5,32	3,60	3,60
1,2		-25,50	17,07	16,24	8,26	6,82
1,3		-32,98	31,32	16,02	12,58	9,20
1,4		-39,19	49,06	22,34	17,02	11,47
1,5		-43,96	74,28	29,93	20,18	13,30

Table A.1 Results of changing the size of losses in the body part. Displayed in percentage change in VaR.

Loss Factor	Interval					
		6	7	8	9	10
0,5		-40,74	-9,98	-5,04	-2,22	-2,11
0,6		-32,43	-7,04	-3,60	-1,72	-1,61
0,7		-23,50	-4,93	-2,33	-1,16	-1,11
0,8		-16,24	-2,83	-1,50	-0,78	-0,72
0,9		-7,15	-1,39	-0,72	-0,39	-0,33
1,0		0,00	0,00	0,00	0,00	0,00
1,1		6,49	1,39	0,78	0,39	0,33
1,2		13,30	2,77	1,39	0,72	0,67
1,3		17,79	3,77	2,05	1,00	0,94
1,4		23,28	4,88	2,72	1,27	1,16
1,5		28,60	5,88	3,16	1,50	1,39

Table A.2 Results of changing the size of losses in the tail part. Displayed in percentage change in VaR.

Loss Factor	Interval				
		11	12	13	14
0,5		-6,10	-0,11	-0,11	-0,17
0,6		-4,66	-0,06	-0,06	-0,11
0,7		-3,10	-0,06	-0,06	-0,11
0,8		-1,88	-0,06	-0,06	-0,06
0,9		-0,94	0,00	0,00	-0,06
1,0		0,00	0,00	0,00	0,00
1,1		0,94	0,00	0,00	0,00
1,2		1,83	0,00	0,00	0,06
1,3		2,72	0,00	0,00	0,06
1,4		3,38	0,06	0,06	0,11
1,5		3,99	0,06	0,06	0,11

Table A.3 Results of changing the size of losses in the extreme events part. Displayed in percentage change in VaR.

Number of losses added	Interval					
		1	2	3	4	5
1		-0,06	-0,06	0,00	0,00	0,06
10		-0,67	-0,28	0,00	0,28	0,72
100		-4,93	-4,93	-2,33	0,33	576,5

Table A.4 Results of adding a fixed number of losses (1,10 and 100) to the benchmark dataset in the body part, interval one to five. Displayed in percentage change in VaR.

Number of losses added	Interval					
		6	7	8	9	10
1		0,22	0,28	0,33	0,39	0,39
10		2,00	2,77	3,33	3,71	3,99
100		20,29	30,76	37,97	43,18	46,56

Table A.5 Results of adding a fixed number of losses (1,10 and 100) to the benchmark dataset in the tail part, interval six to ten. Displayed in percentage change in VaR.

Number of losses added	Interval				
		11	12	13	14
1		0,78	0,94	1,00	1,05
10		7,71	9,20	9,98	10,25
100		108,37	138,19	154,66	164,91

Table A.6 Results of adding a fixed number of losses (1,10 and 100) to the benchmark dataset in the extreme events part, interval 11 to 14. Displayed in percentage change in VaR.

Relative frequency added	Interval					
		1	2	3	4	5
0,1		-10,70	-4,93	-0,89	0,17	0,94
0,2		-20,23	-9,70	-2,00	-0,94	1,55
0,3		-27,16	-13,86	-2,94	0,72	2,33
0,4		-33,76	-17,18	-3,99	2,27	2,94

**Table A.7** Results of adding a variable number of losses (10 %, 20 %, 30 % and 40 %) to the benchmark dataset in the body part, interval 1 to 5. Displayed in percentage change in VaR.

Relative frequency added	Interval					
		6	7	8	9	10
0,1		6,87	2,77	1,66	0,78	0,83
0,2		14,41	5,27	3,60	1,83	2,00
0,3		22,12	8,37	5,16	2,99	2,83
0,4		30,21	11,97	7,15	4,05	3,99

**Table A.8** Results of adding a variable number of losses (10 %, 20 %, 30 % and 40 %) to the benchmark dataset in the tail part, interval 6 to 10. Displayed in percentage change in VaR.