

A comparison of VaR estimates in risk management

K Panman
20699506

Thesis submitted for the degree *Philosophiae Doctor* in
Risk Analysis at the Potchefstroom Campus of the North-
West University

Promoter: Prof. T. de Wet

Co-promoter: Prof. P.J. de Jongh

December 2015

Abstract

In risk management one is often concerned with estimating high quantiles of the underlying distribution of a given sample of data. This is known to be challenging, especially if the quantile of interest is extremely high. Estimating such an extreme quantile becomes even more challenging when the sample size is very small. This is the case when one is interested in the measurement of operational risk within a bank. Even though the focus of this thesis is on operational risk quantification, the core principles illustrated in this thesis can be applied to various fields.

The most popular approach in the quantification of operational risk is the so-called advanced measurement approach (AMA). This approach allows banks to use internal models to measure its operational risk. A widely used model under the AMA is the loss distribution approach (LDA). Essentially the LDA makes use of both the frequency and severity distributions to obtain the aggregate distribution where the risk modeller is eventually interested in various quantiles of the aggregate distribution. From the literature it is clear that the frequency has a minimal impact on the aggregate distribution and that the latter is mainly influenced by the severity distribution (see e.g. BCBS, 2011). For this reason the majority of this thesis focusses on modelling the severities and in particular, on the estimation of extreme quantiles of the severity distribution. There are two main approaches in modelling the severities under the AMA. These are the so-called spliced distribution approach and the full-parametric approach. The first objective of this thesis is to investigate which quantile estimation procedure performs uniformly the best in an operational risk context under the spliced distribution approach. The second objective is to investigate the quantile estimation procedures under the full-parametric approach. Attending to both the first and second objectives will lead to sufficient estimates of the severity distribution. Given sufficient estimates of both the frequency and severity distributions, it is unclear from the literature which procedure is optimal when convoluting both the frequency and severity distributions in order to approximate the aggregate distribution. To this end the third objective of this thesis is to investigate the most widely used approximation techniques. The three objectives of this thesis are discussed in the following three paragraphs respectively.

The basic idea underlying the spliced distribution approach is to estimate extreme quantiles of the severity distribution by focusing on extreme observations only. A natural approach for this purpose is to make use of extreme value theory (EVT). A well-known quantile estimator that was derived under EVT is the so-called Weissman quantile estimator. Alternative quantile estimators are desired since the Weissman quantile estimator relies heavily on its

asymptotic properties and, as it was mentioned previously, the sample sizes often encountered in operational risk are relatively small. A new quantile estimator was investigated and is discussed in the thesis. The idea of the new estimator is that a lower quantile (a quantile at a lower probability level) is rather estimated which is then extrapolated to the desired quantile using a multiplying factor. This estimator is referred to as multipliers throughout this thesis. The multipliers proved to perform uniformly the best under this approach in an operational risk context.

The full-parametric approach consists of estimating extreme quantiles by fitting one class of distributions to the entire sample of data. Popular choices of these distribution classes are the Burr, LogNormal and normal inverse Gaussian (NIG) distributions. This approach enables one to model the entire severity distribution where each observation is used (as opposed to the spliced distribution approach using only extreme observations) to inform the fitting of the statistical distributions. It is well known that the data in an operational risk context is often heavy-tailed and that in some cases is even too heavy-tailed to model using the distributions popular in operational risk. A natural alternative to these distributions is to use its *logged* version. An example of such a *logged* version is the LogNormal distribution being the *logged* version of the Normal distribution. Since the NIG distribution is a parameter-rich distribution (in the sense that it has four parameters making it a very flexible distribution) it was decided to investigate its *logged* version, referred to as the log normal inverse Gaussian (LNIG) distribution. Although the LNIG distribution outperformed some of the widely used distributions (especially when aiming to reduce the bias of its quantile estimates) it did not prove to be the optimal distribution to use when following the full-parametric approach.

A number of techniques exist for approximating the quantiles of the aggregate distribution. These are Monte Carlo simulation, Panjer recursion, fast Fourier transforms, single loss approximations and perturbative approximations. Even though these techniques are discussed extensively in the literature, there is no study in the literature that widely compares them in an operational risk context. To this end it was decided to conduct an extensive comparison (theoretical and numerical) of these approximation techniques in order to determine which procedure performs uniformly the best. It was found that the second order perturbative approximation performed best for approximating extreme quantiles of aggregate distributions typically used in operational risk.

Keywords: Quantile estimation, extreme quantiles, small samples, heavy-tailed distributions, operational risk, loss distribution approach.

Opsomming

Die analis in risikobestuur is tipies geïnteresseerd in die beraming van hoë kwantiele van die onderliggende verdeling van 'n gegewe steekproef. Dit is bekend dat hierdie beraming baie uitdagend kan wees, veral as die kwantiel van belang baie ekstreem is. Die beraming van so 'n ekstreme kwantiel raak nog moeiliker wanneer die steekproefgrootte baie klein is. Hierdie is die geval wanneer mens geïnteresseerd is in die meting van 'n bank se operasionele risiko. Alhoewel die fokus van hierdie tesis op die kwantifisering van operasionele risiko is, kan die beginsels wat in die tesis geïllustreer word in verskeie velde toegepas word.

The mees bekende benadering in die kwantifisering van operasionele risiko is the sogenaamde *advanced measurement approach* (AMA). Hierdie benadering laat banke toe om van interne modelle gebruik te maak om operasionele risiko te meet. 'n Wyd gebruikte benadering onder die AMA is die sogenaamde *loss distribution approach* (LDA). Die LDA maak gebruik van beide die frekwensie verdeling (die spreiding van die aantal verliese) en die erns verdeling (die spreiding van die verliese se monetêre waardes) om sodoende die gesommeerde verdeling (die spreiding van die totale verlies) vas te stel waar die analis op die ou end belang stel in verskeie kwantiele van die gesommeerde verdeling. Vanuit die literatuur is dit duidelik dat die frekwensie verdeling 'n minimale impak het op die gesommeerde verdeling en dat die laasgenoemde hoofsaaklik van die erns verdeling af hang (sien bv. BCBS, 2011). Om hierdie rede fokus die tesis op die beraming van die erns verdeling en in besonder op die beraming van die erns verdeling se ekstreem kwantiele. Daar is hoofsaaklik twee benaderings onder die AMA om die erns verdeling te beraam. Die eerste benadering staan beken as die *spliced distribution approach* (SDA) en die tweede benadering word na verwys as die *full-parametric approach* (FPA). Die eerste doelwit van hierdie tesis is om vas te stel watter beramingstegniek uniform die beste presteer onder die SDA in 'n operasionele risiko konteks. Die tweede doelwit is om vas te stel watter tegniek die beste vaar onder die FPA. Die suksesvolle behaling van beide die eerste en tweede doelwitte sal beter beramers van die erns verdeling oplewer. Geges dat voldoende beramers van beide die frekwensie en erns verdelings beskikbaar is, bly daar steeds heelwat onsekerheid in die literatuur oor watter tegniek optimaal is om die gesommeerde verdeling te beraam (laat die tegnieke wat gebruik word om die gesommeerde verdeling te beraam, deur beide die frekwensie en erns verdelings te inkorporeer, na verwys word as die *benaderingstegnieke*). Om hierdie rede is die derde doelwit van die tesis om alle populêre benaderingstegnieke te ondersoek om sodoende vas te stel watter benaderingstegniek die beste werk. Al drie die doelwitte in hierdie tesis word in die volgende drie paragrawe afsonderlik bespreek.

Die basiese idee van die SDA is om die ekstreem kwantiele van die erns verdeling te beraam deur slegs van die ekstreme observasies gebruik te maak. 'n Natuurlike manier om slegs die ekstreme observasies te gebruik is om gebruik te maak van ekstreem waarde teorie (EWT). 'n Wyd gebruikte kwantielberamer uit die EWT is die sogenaamde Weissman beramer. Alternatiewe kwantielberamers word vereis omrede die Weissman beramer heelwat staat maak op sy asimptotiese eienskappe en, soos voorheen genoem, is die steekproefgroottes in operasionele risiko gewoonlik relatief klein. 'n Nuwe beramer was ondersoek en word in diepte bespreek in die tesis. Die basiese idee van die nuwe beramer is dat 'n laer kwantiel ('n kwantiel by 'n laer waarskynlikheidsvlak) eerder beraam word waar die laer kwantiel dan met 'n vermenigvuldigende faktor aangepas word om sodoende 'n beramer van die gewenste kwantiel te kry. Hierdie beramer word in die tesis na verwys as multipliers (vermenigvuldigers). Die vermenigvuldigers het uniform die beste presteer onder die SDA in 'n operasionele risiko-konteks.

Die FPA bestaan uit die beraming van ekstreem kwantiele deur een klas van verdelings op die hele steekproef te pas. Populêre verdelings hiervoor sluit in die Burr, *LogNormal* en *normal inverse Gaussian* (NIG) verdelings. Die FPA laat die analis toe om die hele erns verdeling te modelleer deur elke observasie te gebruik (vergeleke met die SDA waar slegs ekstreme observasies gebruik word) om die passing van die statistiese verdelings in te lig. Dit is alombekend dat data in die operasionele risiko omgewing swaar sterte het en in sommige gevalle is die data se sterte té swaar om selfs met die populêre verdelings gemodelleer te word. 'n Natuurlike alternatief vir hierdie verdelings is om die *logged* weergawe daarvan te gebruik. 'n Voorbeeld van so 'n *logged* weergawe is die *LogNormal* verdeling wat die *logged* weergawe is van die *Normal* verdeling. Aangesien die NIG verdeling 'n parameter-ryke verdeling is (in die sin dat hy vier parameters het wat hom buigbaar maak) was daar besluit om die NIG verdeling se *logged* weergawe te ondersoek. Hierdie verdeling staan bekend as die *log normal inverse Gaussian* (LNIG) verdeling. Alhoewel die LNIG verdeling sommige van die populêre verdelings geklop het in terme van sy numeriese resultate (veral wanneer die doel is om die sydigheid van die kwantielberamers te minimeer) was daar gevind dat die LNIG verdeling nie die optimale verdeling is om die die erns verdeling te beraam onder die FPA.

'n Aantal tegnieke om die kwantiele van die gesommeerde verdeling te benader word in die literatuur bespreek. Hierdie benaderingstegnieke sluit in *Monte Carlo simulation*, *Panjer recursion*, *fast Fourier transform*, *single loss approximation* en *perturbative approximation*. Alhoewel hierdie benaderingstegnieke in diepte bespreek word in die literatuur is daar geen omvattende studie wat elkeen met mekaar vergelyk in 'n operasionele risiko konteks nie. Dit

is om hierdie rede dat 'n vergelykende (teoretiese en numeriese) studie gedoen was om te bepaal watter tegniek uniform die beste presteer. Daar was gevind dat die tweede orde *perturbative approximation* tegniek die beste vaar in die beraming van ekstreem kwantiele van gesommeerde verdelings wat tipies in operasionele risiko teëgekom word.

Sleutelwoorde: Kwantielberaming, ekstreem kwantiele, klein steekproewe, swaargestert verdelings, operasionele risiko, *loss distribution approach*.

Acknowledgements

First of all I would like to thank my academic supervisor, Professor Tertius de Wet, for his excellent guidance during the course of this study. Professor De Wet was a marvellous leader throughout this study and in my personal life. I regard Professor De Wet as a role model and inspiration. Thank you for always motivating me through the tough times, whether academic or personal. My utmost and sincerest gratitude goes to you and your lovely wife, Lynette de Wet.

I would like to thank my colleagues at the Centre for Business Mathematics and Informatics (BMI) who had a direct or indirect influence in this study. My academic co-supervisor, Professor Riaan de Jongh, who initially was not part of this study, but soon became influential and later played a crucial role in this study. I thank you for all your comments and input during this study. Also, Professor Helgard Raubenheimer, thank you for all the comments and the double checking of my simulation results. The last person that I would like to single out from the Centre for BMI is Professor Hennie Venter, thank you for initiating and stimulating the idea of the multipliers.

Doing a full-time Ph.D. is tough enough, but having some form of financial assistance just makes it easier. First of all I would like to thank the National Research Foundation (NRF) for its financial assistance throughout the first three years of this study. I would also like to thank the Centre for BMI for assisting financially, especially for financing the majority of my travelling and conference expenses. The financial assistance received from BANKSETA is also much appreciated which became handy during the finalisation of this study.

To my high school teacher, Elmarie Human, thank you for believing in me during the times that I wanted to give up. Thank you for pushing me into the right direction regarding my high school subjects, especially with Mathematics, and thanks for all the time and support that you invested in me. Also, to my high school Mathematics teacher, Chrisma Grové, thank you for the time initially invested in helping me to understand the basics of Mathematics and later igniting my love for it. This study would not have been possible without the Mathematical foundation that you have laid.

I'll probably not hear the end of it if I do not thank my friends. I'll start with my housemates Armand de Bod, Louis Janse van Rensburg and Paul Aucamp. Thanks for all the support during this study, especially the first two years that we lived together. Thanks to Hermann Harris and Robin Rich for making my relocation to the Western Cape much easier and for

always being available when it was time to blow off some steam. Also, thank you Francois du Toit, Jurie Moolman, Dewaldt la Grange, Gerrie Jordaan, Angus van Aswegen and Tiaan Pieterse for always being there for me, even if it was just for the odd phone call. I apologise if I missed someone.

Last, but not least, I would like to thank my family. To my aunt, Estelle Pretorius, thank you for letting me stay at your home during the last part of this study. To my brother, Gordon Panman, my mother, Annalize Panman, and my father, Gordon Silas Panman, thanks for the way in which you brought me up. Thanks for teaching me what life is all about. Thank you for letting me make my own choices, and lastly, thank you for all you have done.

May God bless you all.

Table of Contents

ABSTRACT	I
OPSOMMING	III
ACKNOWLEDGEMENTS	VI
TABLE OF CONTENTS	VIII
LIST OF ABBREVIATIONS.....	XI
LIST OF FIGURES.....	XIII
LIST OF TABLES.....	XIV
NOTATION.....	XVI
CHAPTER 1: INTRODUCTION	1
1.1. OVERVIEW OF FINANCIAL RISK MANAGEMENT	1
1.1.1. REGULATORY ASPECTS OF FINANCIAL RISK MANAGEMENT.....	2
1.1.2. FINANCIAL RISK CATEGORIES	4
1.1.3. REGULATORY CAPITAL	4
1.1.4. RISK MEASURES	6
1.2. OPERATIONAL RISK MANAGEMENT	7
1.2.1. OPERATIONAL RISK.....	8
1.2.2. LOSS DISTRIBUTION APPROACH.....	8
1.2.3. AGGREGATE LOSS DISTRIBUTION	10
1.2.4. FREQUENCY AND SEVERITY DISTRIBUTIONS	11
1.3. PRACTICAL ISSUES IN OPERATIONAL RISK MODELLING.....	13
1.4. THESIS OBJECTIVES AND LAYOUT	15
CHAPTER 2: ‘EXISTING’ QUANTILE ESTIMATION TECHNIQUES FOR THE SEVERITY DISTRIBUTION	18
2.1. TRADITIONAL QUANTILE ESTIMATION TECHNIQUES	18
2.1.1. QUANTILE ESTIMATION USING ORDER STATISTICS.....	18
2.1.2. QUANTILE ESTIMATION USING A PARAMETRIC APPROACH.....	18
2.2. EXTREME VALUE THEORY	21
2.2.1. QUANTILE ESTIMATION USING FIRST ORDER EXTREME VALUE THEORY	28
2.2.2. QUANTILE ESTIMATION USING SECOND ORDER EXTREME VALUE THEORY.....	29

CHAPTER 3: ‘NEW’ QUANTILE ESTIMATION TECHNIQUES FOR THE SEVERITY DISTRIBUTION	32
3.1. QUANTILE ESTIMATION USING A MULTIPLYING FACTOR	32
3.1.1. A QUANTILE ESTIMATION TECHNIQUE PROPOSED BY THE BASEL COMMITTEE FOR MARKET RISK QUANTIFICATION.....	32
3.1.2. THE BASIC IDEA	32
3.1.3. DERIVING THE MULTIPLYING FACTORS USING SIMULATION	34
3.1.4. SOME ASYMPTOTIC RESULTS OF THE MULTIPLYING FACTORS.....	38
3.1.5. SECOND ORDER MULTIPLIERS.....	39
3.2. QUANTILE ESTIMATION BY FITTING THE LOGARITHM OF THE NORMAL INVERSE GAUSSIAN DISTRIBUTION.....	41
3.2.1. INTRODUCTION	41
3.2.2. THE NORMAL INVERSE GAUSSIAN DISTRIBUTION.....	42
3.2.2.1. INFERENCE ON THE NIG DISTRIBUTION	43
3.2.3. THE LOGARITHM OF THE NORMAL INVERSE GAUSSIAN DISTRIBUTION	47
3.2.3.1. SOME PROPERTIES OF THE LNIG DISTRIBUTION.....	48
 CHAPTER 4: A (NUMERICAL) COMPARISON OF QUANTILE ESTIMATION TECHNIQUES FOR THE SEVERITY DISTRIBUTION	 51
4.1. A SIMULATION STUDY USED TO COMPARE THE QUANTILE ESTIMATION TECHNIQUES WHICH RELY ON HIGH QUANTILES ONLY (THE EVT SIMULATION STUDY).....	52
4.1.1. SIMULATION DESIGN.....	52
4.1.2. SIMULATION RESULTS	56
4.1.2.1. FIRST ORDER RESULTS	56
4.1.2.2. SECOND ORDER RESULTS	75
4.1.3. SUMMARY OF THE EVT SIMULATION RESULTS.....	94
4.2. A SIMULATION STUDY USED TO COMPARE THE QUANTILE ESTIMATION TECHNIQUES BASED ON THE FULL PARAMETRIC APPROACH (THE FULL-PARAMETRIC SIMULATION STUDY).....	96
4.2.1. SIMULATION DESIGN.....	96
4.2.2. SIMULATION RESULTS	98
4.2.3. COMPUTATIONAL CHALLENGES.....	103
4.2.4. CONCLUDING REMARKS ON THE FULL-PARAMETRIC APPROACH	103
4.3. AN EMPIRICAL STUDY OF THE TECHNIQUES USED TO ESTIMATE THE SEVERITY DISTRIBUTION’S QUANTILES.....	105

CHAPTER 5: QUANTILE ESTIMATION OF THE COMPOUND DISTRIBUTION.	111
5.1. POPULAR APPROXIMATION TECHNIQUES.....	113
5.1.1. MONTE CARLO	114
5.1.2. PANJER RECURSION.....	115
5.1.3. FAST FOURIER TRANSFORMATION.....	118
5.2. CLOSED-FORM APPROXIMATION TECHNIQUES.....	121
5.2.1. OVERVIEW OF THE CLOSED-FORM APPROXIMATION TECHNIQUES.....	121
5.2.2. SINGLE-LOSS APPROXIMATION (SLA) UP TO SECOND ORDER	122
5.2.3. PERTURBATIVE APPROXIMATION (PA) UP TO SECOND ORDER	125
5.3. CONCLUDING REMARKS	126
 CHAPTER 6: A (NUMERICAL) COMPARISON OF QUANTILE APPROXIMATION TECHNIQUES FOR THE COMPOUND DISTRIBUTION (THE MC SIMULATION STUDY).....	 128
6.1. SIMULATION DESIGN.....	128
6.2. SIMULATION RESULTS.....	131
6.2.1. FINITE MEAN DISTRIBUTIONS	131
6.2.2. INFINITE MEAN DISTRIBUTIONS.....	136
6.3. SUMMARY OF THE APPROXIMATION TECHNIQUES.....	138
 CHAPTER 7: CONCLUSIONS	 139
7.1. CONCLUDING REMARKS	139
7.2. THESIS CONTRIBUTIONS.....	142
7.3. FURTHER RESEARCH	143
 REFERENCES.....	 146
 APPENDIX A: DISTRIBUTIONS USED THROUGHOUT THIS THESIS.....	 153
APPENDIX B: APPROXIMATING THE MODIFIED BESSEL FUNCTION	154
APPENDIX C: MAXIMUM LIKELIHOOD ESTIMATION.....	157

List of abbreviations

AMA	: Advanced measurement approach
BCBS	: Basel Committee on Banking Supervision
BIA	: Basic indicator approach
BM	: Block maxima
CR	: Credit risk
DFT	: Discrete Fourier transformation
ES	: Expected shortfall
EVI	: Extreme value index
EVT	: Extreme value theory
FFT	: Fast Fourier transformation
FRM	: Financial risk management
FT	: Fisher-Tippett
GEV	: Generalised extreme value
GPD	: Generalised Pareto distribution
IDFT	: Inverse discrete Fourier transform
i.i.d.	: Independent and identically distributed
LDA	: Loss distribution approach
LNIG	: Logarithm normal inverse Gaussian
MBF	: Modified Bessel function
MC	: Monte Carlo
MEF	: Mean excess function
MLE	: Maximum likelihood estimation
MOP	: Mean-of-order- p
MR	: Market risk
MSE	: Mean squared error
NIG	: Normal inverse Gaussian
OR	: Operational risk
ORC	: Operational risk category
ORM	: Operational risk management
PA	: Perturbative approximation
PBdH	: Pickands-Balkema-deHaan
PL	: Profit/loss
POT	: Peaks over threshold
PR	: Panjer recursion
QQ-plot	: Quantile-Quantile plot

RBMOP	: Reduced bias mean-of-order- p
RC	: Regulatory capital
RMSE	: Root mean squared error
SA	: Standardised approach
SAS	: Software used for simulation and numerical purposes
SLA	: Single loss approximation
TBE	: Tail based estimation
TQF	: Tail quantile function
VaR	: Value-at-Risk

List of Figures

Figure 1.1: Three Pillars from Basel III

Figure 1.2: Hypothetical operational risk loss distribution showing expected losses and unexpected losses at the 99.9th percentile (source: de Jongh et al., 2013)

Figure 2.1: Histogram of LogNormal data with $\mu = 0$ and $\sigma = 1$

Figure 2.2: Fitted density and histogram of LogNormal data with $\mu = 0$ and $\sigma = 1$

Figure 2.3: QQ-plot of LogNormal data with $\mu = 0$ and $\sigma = 1$

Figure 2.4: Artificial time series data from the LogNormal distribution with $\mu = 0$ and $\sigma = 1$

Figure 2.5: Mean excess plot of Pareto data with $\gamma = 1$

Figure 3.1: Quantiles of F

Figure 3.2: Distributions used to derive the multiplying factors

Figure 3.3: Graphical illustration of the NIG density (using its true parameters)

Figure 3.4: Graphical illustration of the NIG density (using its estimated parameters)

Figure 3.5: Graphical illustration of the true NIG density vs. estimated NIG density

Figure 4.1: A scatterplot of the distributions used for the EVT simulation study

Figure 4.2: Scatterplot of the distributions used in the full-parametric simulation study

Figure 6.1: Relative error plots for the (very) short-tailed LNor1 distribution (EVI=0)

Figure 6.2: Relative error plots for the semi-heavy tailed Burr3 distribution (EVI=0.5)

Figure 6.3: Relative error plots for the heavy-tailed LNig2 distribution (EVI=1)

Figure 6.4: Relative error plots for all finite mean distributions at probability level 0.999

Figure 6.5: Relative error plots for the short-tailed LNor1 distribution (EVI=0)

Figure 6.6: Relative error plots for the semi-heavy tailed Burr3 distribution (EVI=0.5)

Figure 6.7: Relative error plots for the heavy-tailed LNig2 distribution (EVI=1)

Figure 6.8: Relative error plots for the Burr12 distribution (EVI=1.5, Rho=-2)

Figure 6.9: Relative error plots for the Burr13 distribution (EVI=1.5, Rho=-1)

Figure 6.10: Relative error plots for the Burr15 distribution (EVI=1.5, Rho=-0.1)

Figure 6.11: Relative error plots for all infinite mean distributions at probability level 0.999

List of Tables

Table 3.1: Distributions used to derive the multiplying factors

Table 3.2: Initial multipliers used during the development simulation study

Table 3.3: Multipliers obtained during the development simulation study

Table 3.4: Asymptotic results of the MinMax multipliers

Table 3.5: Second order multipliers obtained during the development simulation study

Table 4.1: A list of the distributions used for the extreme value theory simulation study

Table 4.2: Root mean squared error results of the first order EVT simulation study for $p_0 = 0.999$

Table 4.3: Absolute bias results of the first order EVT simulation study for $p_0 = 0.999$

Table 4.4: Root mean squared error results of the first order EVT simulation study for $p_0 = 0.995$

Table 4.5: Absolute bias results of the first order EVT simulation study for $p_0 = 0.995$

Table 4.6: Root mean squared error results of the first order EVT simulation study for $p_0 = 0.99$

Table 4.7: Absolute bias results of the first order EVT simulation study for $p_0 = 0.99$

Table 4.8: Root mean squared error results of the second order EVT simulation study for $p_0 = 0.999$

Table 4.9: Absolute bias results of the second order EVT simulation study for $p_0 = 0.999$

Table 4.10: Root mean squared error results of the second order EVT simulation study for $p_0 = 0.995$

Table 4.11: Absolute bias results of the second order EVT simulation study for $p_0 = 0.995$

Table 4.12: Root mean squared error results of the second order EVT simulation study for $p_0 = 0.99$

Table 4.13: Absolute bias results of the second order EVT simulation study for $p_0 = 0.99$

Table 4.14: List of distributions used to generate data for the full-parametric simulation study

Table 4.15: RMSE results of the full-parametric quantile estimators for $p = 0.999$

Table 4.16: |BiasA| results of the full-parametric quantile estimators for $p = 0.999$

Table 4.17: RMSE results of the full-parametric quantile estimators for $p = 0.995$

Table 4.18: |BiasA| results of the full-parametric quantile estimators for $p = 0.995$

Table 4.19: RMSE results of the full-parametric quantile estimators for $p = 0.99$

Table 4.20: |BiasA| results of the full-parametric quantile estimators for $p = 0.99$

Table 4.21: Alternative distributions considered that had too many computational challenges

Table 4.22: Estimates of the empirical data's EVIs

Table 4.23: Quantile estimates relative to the true quantile of Sample1

Table 4.24: Quantile estimates relative to the true quantile of Sample2

Table 4.25: Quantile estimates of Sample1

Table 4.26: Quantile estimates of Sample2

Table 6.1: Burr parameter sets selected for the Monte Carlo simulation study

Table 6.2: LogNormal parameter sets selected for the Monte Carlo simulation study

Table 6.3: LogNIG parameter sets selected for the Monte Carlo simulation study

Table 6.4: Smallest ARE and percentage inclusions in the MC CI over all finite mean distributions

Table 6.5: Smallest ARE and percentage inclusions in the MC CI over selected distributions

Table 6.6: Smallest ARE and percentage inclusions in the MC CI over infinite mean distributions

Table A.1: Quantiles of the LogNormal distributions

Table A.2: Quantiles of the Burr distributions

Table A.3: Quantiles of the LogNIG distributions

Notation

$X_1, X_2, \dots, X_n$	: A sample of size n .
$X_{1,n} \leq X_{2,n} \leq \dots \leq X_{n,n}$	: Ordered version of $X_1, X_2, \dots, X_n$.
F	: A generic distribution function.
F^{-1}	: The inverse of the distribution function F .
q_p	: The quantile function of F evaluated in p , i.e. $F^{-1}(p)$.
$\hat{F}$	: An estimate of the generic distribution function F .
$\hat{F}^{-1}$	: The inverse of the estimated distribution function $\hat{F}$.
$\hat{q}_p$	: An estimate of the quantile function of F evaluated in p , i.e. $\hat{F}^{-1}(p)$.
$MSE(\hat{q}_p)$	: The mean squared error of the quantile function estimate, i.e. $\frac{1}{N} \sum_{i=1}^N (\hat{q}_{p,i} - q_p)^2$.
$RMSE(\hat{q}_p)$	: The root mean squared error of the quantile function estimate, i.e. $\sqrt{MSE(\hat{q}_p)}$.
$MAD(\hat{q}_p)$	: The median absolute deviation of the quantile function estimate, i.e. $med_i \hat{q}_{p,i} - q_p $.
$BiasA(\hat{q}_p)$	: The bias of the quantile function estimate using the average of the estimates, i.e. $\frac{1}{N} \sum_{i=1}^N \hat{q}_{p,i} - q_p$.
$ BiasA $	: The absolute value of the bias of the quantile function estimate using the average of the estimates, i.e. $ BiasA(\hat{q}_p) $.

$BiasM(\hat{q}_p)$

: The bias of the quantile function estimate using the median of the estimates, i.e. $med_i \hat{q}_{p,i} - q_p$.

Chapter 1: Introduction

This thesis will be concerned with the estimation of regulatory capital (RC) in operational risk management (ORM) and in particular with the estimation of the 99.9% Value-at-Risk (VaR), or equivalently, the 99.9% quantile of the annual distribution of aggregate losses in a specific operational risk category (ORC). Because the reader may be unfamiliar with the above-mentioned concepts, the first chapter is devoted to providing the necessary background to risk management and to motivate why the research in this thesis is important for the banking and insurance industry. The reader will be introduced to financial risk management (FRM) and regulations, and in particular to the concept of RC and risk measures. Then the focus will be on ORM and how RC can be calculated using the loss distribution approach (LDA). In the process some of the important research issues and motivation for the research that was conducted in this thesis will be highlighted.

The layout of this chapter is as follows. In Section 1.1 an overview of FRM will be given. Then, in Section 1.2, ORM will be discussed in more detail and, in Section 1.3 the practical issues of operational risk modelling that will constitute the focus of this thesis will be highlighted. In Section 1.4 the objectives and layout of the thesis will be given.

1.1. Overview of financial risk management

FRM has become increasingly important during the last number of decades (see e.g. Crouhy et al., 2000). As far back as 1988 the first Basel Accord was drafted to address poor risk management practices in the banking industry. Subsequently the Basel Accord was updated several times, first Basel I (in 1988), then Basel II (in 2004) and Basel III (in 2010) and currently Basel IV is in the making. The Basel Accord comprises of guidelines and best practices that banks need to follow in order to ensure good risk management practices. Similarly Solvency I and Solvency II address risk management in the insurance industry, but the focus of this thesis will mainly be on banks and therefore on the current set of guidelines as contained in Basel III. The Basel Accords are typically adopted by regulators of countries (central banks, in South Africa the Reserve Bank) who enforce banks to adhere to specific Basel guidelines. Basel III provides best practice guidelines for all the risk categories that a bank face, namely fluctuations in markets and investments (market risk), fluctuations in the credit quality of individual clients (retail credit risk) and of corporate clients (wholesale credit risk), and operational risks caused by failed processes, human behaviour (e.g. fraud) or external events (e.g. earthquakes). Of these risk categories, credit risk is the most important

risk encountered in a bank as it arises from the banks' primary business operations. As stated earlier, the Basel Accord is concerned with providing regulators with the necessary guidelines to ensure good risk management practices of banks. In particular, banks are forced by the regulators to hold capital (referred to as regulatory capital and abbreviated as RC) as a buffer against losses that could occur in each of the risk categories. RC calculations are typically based on standardised (for smaller banks) or advanced approaches (for bigger banks). Central to the calculation of RC is the construction of a one-year-ahead annual profit/loss (PL) distribution of each main risk category. Once the PL distribution is constructed, then the RC is calculated as the difference between some risk measure (usually the so-called VaR) and expected PL. The construction of the PL distribution is not easy and furthermore should also be stressed to cater for adverse economic conditions.

The above paragraph contains an overview of some of the essential components of risk management relevant for this thesis. In the next subsections, more detail will be provided on each of these components. In Section 1.1.1 more detail will be given on the Basel Accord and how regulatory aspects pertain to banks with specific reference to Basel III. Then, in Section 1.1.2, the various risk categories will be discussed in more detail. The calculation of RC through the construction of a PL distribution and by using risk measures will be discussed in Section 1.1.3. The last subsection, Section 1.1.4, will contain a discussion of various risk measures and their properties.

1.1.1. Regulatory aspects of financial risk management

The Basel Committee on Banking Supervision (BCBS) is an independent committee and acts as regulatory authority for the banking industry. The main objective of the BCBS is to develop a framework to strengthen the soundness and stability of the international banking system. This is done by setting international standards that globally and nationally active banks should comply with by improving their FRM processes. Adhering to these standards, the banks' risk exposure will be 'minimized' from the regulatory authorities' point of view. This is important as the national reserve bank, i.e. the South African Reserve Bank in South Africa, acts as guarantor for the local trading banks.

The BCBS supplies banks with three main guidelines (referred to as Pillars). These include:

- Pillar I: Minimum capital requirements,
- Pillar II: Supervisory review, and
- Pillar III: Market discipline.

In this thesis the focus will be on Pillar I and therefore the determination of minimal capital requirements. Figure 1.1 provides a graphical representation of the main building blocks of Pillar I. The minimal capital requirement (also referred to as the RC requirement) is determined by calculating the RC for each of the main risk categories (credit risk, market risk and operational risk) and then the total regulatory capital is determined by risk weighting the RC for each risk class as follows (see e.g. BCBS, 2006):

$$RC = C^{CR} + 12.5(C^{MR}) + 12.5(C^{OR}) \quad (1.1)$$

where RC is the total regulatory capital and C^{CR} , C^{MR} and C^{OR} the RC charge for respectively credit risk, market risk and operational risk. According to BCBS (2006), the capital charge for market risk and operational risk are both multiplied by 12.5 (which is the reciprocal of the minimum capital ratio of 8%) in order to “broadly maintain the aggregate level of minimum capital requirements”. Note that, as indicated in Figure 1.1, banks are allowed to follow different approaches to calculate regulatory capital. In this thesis the focus will be primarily on the advanced measurement approach (AMA) of operational risk which will be discussed later in this chapter.

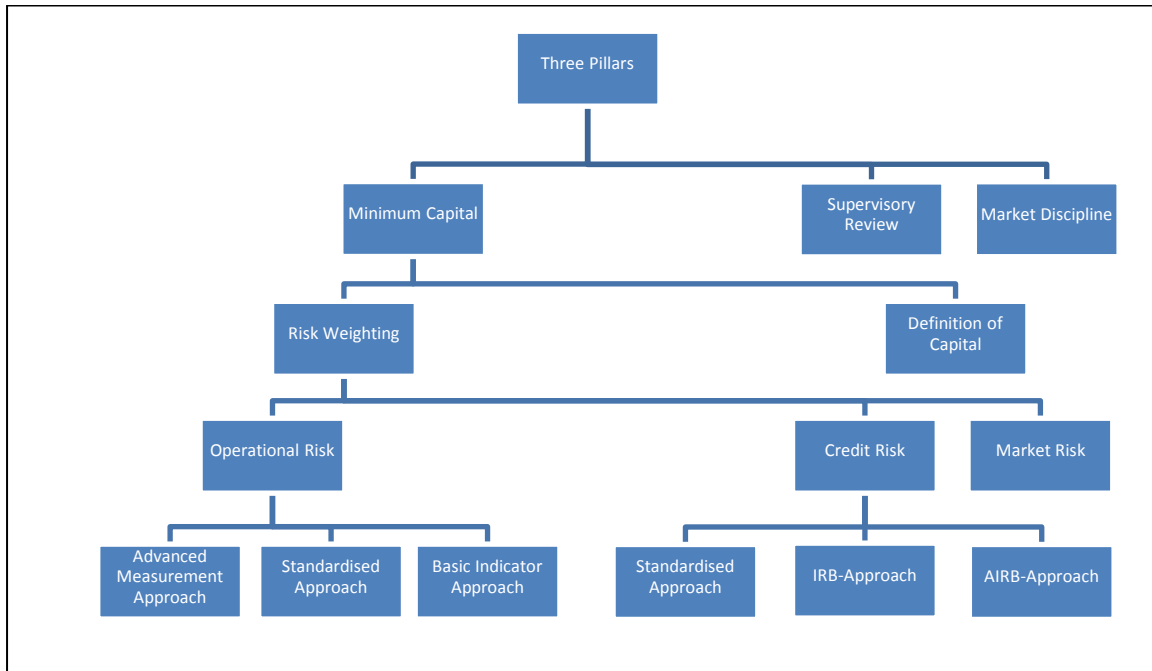


Figure 1.1: Three Pillars from Basel III

A brief description of the three main risk categories will be provided in the next section, followed by a detailed discussion on the concept of RC.

1.1.2. Financial risk categories

As mentioned previously the three main risk categories that banks face are credit risk (CR), market risk (MR) and operational risk (OR). The following definitions of the different risk categories are taken from the literature.

Credit risk is most simply defined as *“the potential that a bank borrower or counterparty will fail to meet its obligations in accordance with agreed terms”*, see e.g. BCBS (2000).

According to the BCBS (2006), market risk is defined *“as the risk of losses in on- and off-balance-sheet positions arising from movements in market prices. The risks subject to this requirement are (1) the risks pertaining to interest rate related instruments and equities in the trading book and (2) foreign exchange risk and commodities risk throughout the bank”*.

Remark: Both credit risk and market risk are extensively discussed in the literature and will therefore not be elaborated on, but the definitions thereof are given to provide the reader with an overall view of the main risk categories.

Operational risk is most widely defined as the *“risk of direct or indirect loss resulting from inadequate or failed internal processes, people and systems or from external events”*, see e.g. BCBS (2006). This risk category is the focus of this thesis and is discussed in more detail in Section 1.2.

1.1.3. Regulatory capital

The term regulatory capital (RC) is used to define the total capital that financial institutions should hold, which in addition to being a requirement for regulatory purposes, is also used to act as a buffer for any unexpected losses. In this section the RC concept will be illustrated from an operational risk perspective, but it also applies to the other risk types as well. Some of the typical operational risk types and associated events that may occur are discussed in Section 1.2. As with other risk types the losses are usually modelled by a loss density as indicated in Figure 1.2 below (see e.g. de Jongh et al., 2013). The bulk of operational risk loss data is typically centred in the middle of the density while extreme losses occur in the right tail of the density. The first type of losses is in operational risk terminology usually referred to as the ‘expected’ losses and those losses are usually observed in the ‘body’ of the distribution. The expected losses typically have a high probability of occurring, but having

a medium or low impact. Losses occurring in the tail of the distribution (away from the body of the distribution) are typically referred to as ‘unexpected’ losses, i.e. those losses having a low probability of occurrence but having high impact.

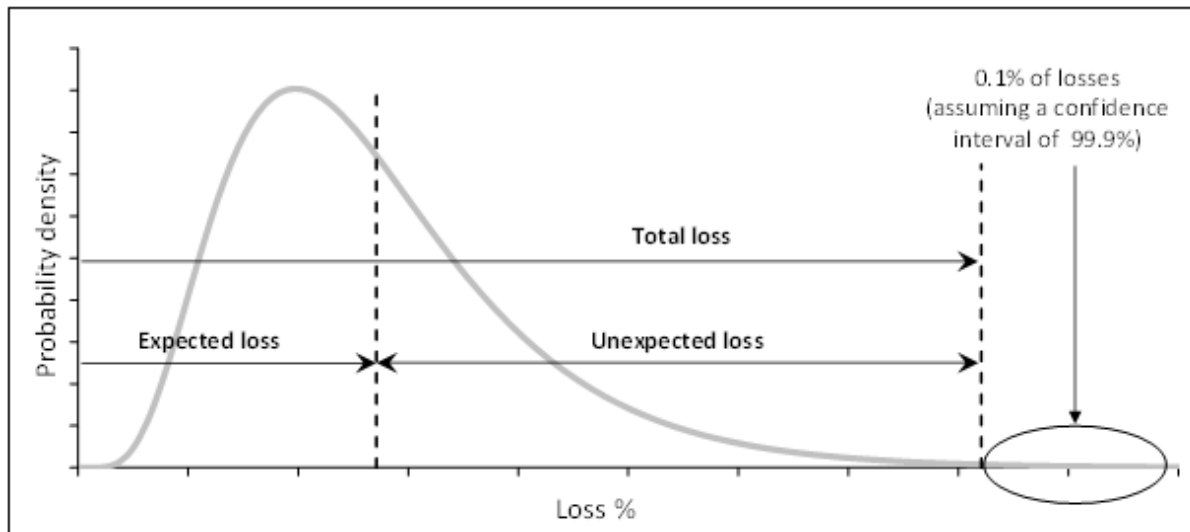


Figure 1.2: Hypothetical operational risk loss distribution showing expected losses and unexpected losses at the 99.9th percentile (source: de Jongh et al., 2013)

Regulatory capital is defined as the amount of capital to hold that will guard the bank against the unexpected losses. This quantity is then defined as the difference between the Value-at-Risk (the 99.9% quantile as specified for operational risk by the regulator) and the expected loss as shown in Figure 1.2. The Value-at-Risk or VaR is a quantile selected in the right tail of the loss density. Although VaR is widely used in practice it is not a coherent measure (see Artzner et al., 1999) and many alternative risk measures have been proposed in the literature (see e.g. McNeil et al., 2015). If regulatory capital guard against the unexpected losses one might ask what is done to protect against the expected losses. Expected losses are usually covered by financial institutions through capital provision. Regulatory capital for credit and market risk is determined in a similar way (to operational risk) although it involves the construction of a profit/loss distribution and not a loss distribution only. Operational risk differs from credit and market risk in that profit cannot arise from operational risk and therefore operational risk is modelled by a loss distribution instead of a profit/loss distribution as is the case of credit and market risk.

1.1.4. Risk measures

Risk measures are used to assist risk managers in the quantification of risk by calculating, amongst others regulatory capital. As discussed in the previous section, it is mostly used to calculate the capital charge for a given risk category. Various risk measures exist such as Value-at-Risk (VaR), expected shortfall (ES), spectral measures and expectiles (see e.g. McNeil et al., 2015). The reason for the proliferation of risk measures is due to the number of desirable properties that have been defined for risk measures such as coherence and elicitability (see e.g. McNeil et al., 2015). In this section VaR and ES will be defined as these are the most popular measures that are currently used in practice.

The definitions for VaR and ES are as follows:

Let X be a random variable with distribution function F_X and $\alpha \in (0,1)$ some probability level. In a FRM context F_X is usually an annual profit/loss distribution where the losses are represented by larger values of X .

Then the $100(\alpha)\%$ VaR is defined as

$$VaR_\alpha(X) = F_X^{-1}(\alpha) = \inf\{x \in \mathbb{R}: F_X(x) \geq \alpha\}, \quad (1.2)$$

with F_X^{-1} the inverse of F_X , $\mathbb{R}$ denoting the real space and $\inf$ the infimum.

The VaR is therefore the $100(\alpha)\%$ percentile of the distribution F_X . In other words the $100(\alpha)\%$ VaR may be interpreted as the loss value that will be exceeded with probability $(1 - \alpha)$. Since the loss distribution is typically an annual loss distribution the 99.9% VaR is associated with the worst loss in 1000 years. In an operational risk context Basel III requires banks to estimate the 99.9% VaR. The regulatory capital is then determined as $VaR_\alpha(X) - E(X)$ with $\alpha = 0.999$.

The $100(\alpha)\%$ ES is defined as

$$ES_\alpha(X) = \mathbb{E}[X|X > VaR_\alpha(X)],$$

in other words, the expected value of all losses greater than VaR.

This representation is made more precise by observing that for a continuous random variable X one has:

$$ES_{\alpha}(X) = \frac{1}{1-\alpha} \int_{\alpha}^1 VaR_{\varepsilon}(X) d\varepsilon, \quad 0 < \alpha < 1. \quad (1.3)$$

Should ES be used as risk measure the regulatory capital is determined as $ES_{\alpha}(X) - E(X)$ with $\alpha = 0.999$. The focus in this thesis will be primarily on VaR since it is still the risk measure of choice in practice.

This section highlighted the main building blocks required for calculating regulatory capital of a bank. An overview was given of the most relevant regulatory aspects, as well as the main financial risk categories of concern. A formal definition of total regulatory capital was given in Equation (1.1) where the calculation of regulatory capital for each risk type has been outlined in the previous two sections. In the next section the focus will be on operational risk management.

1.2. Operational risk management

The Basel II definition of operational risk is the risk of loss resulting from inadequate or failed internal processes, people and systems or from external events (BCBS, 2006). This definition excludes strategic and reputational risk, but includes legal risk. Note again that operational risk typically deals with losses only, unlike market and credit risk which consider the upside (profit) as well. Most of the operational losses encountered in practice are frequent and relatively small, however, of real concern to regulators and risk officers are the less frequent/high-impact losses. Types of operational risks are discussed in most textbooks (see e.g. Chernobai et al., 2007 and Bessis, 2011) but include fraud (internal and external), rogue trading, robbery and theft, errors in legal documents, IT disruptions, principal agent risk, and external (black swan) events. An example of an unpredictable, considerable-impact operational risk event is the terrorist attacks in the US in September 2001. Such low probability/high impact events are referred to as *black swan* events, i.e. rare events but ones whose impact on financial markets can lead to extremely high losses. These losses place considerable emphasis on the effective determination of regulatory capital by financial companies and are of paramount concern to regulators in their attempt to stabilise the international financial system. For a detailed account of the type of operational risk loss events the reader is referred to de Jongh et al. (2013). It should be clear that the

management and mitigation of these risk types involve a strong (qualitative) management component. This will involve setting up the necessary management processes to minimise the likelihood of these events occurring and mitigating its impact, should it occur. The latter will not be the focus of this thesis and it will be assumed that all the necessary management processes are in place. In the next sections the focus will be on the calculation of regulatory capital for operational risk and to highlight some of the issues.

1.2.1. Operational risk

In Figure 1.1 three approaches to modelling the minimum capital requirements for OR were mentioned. These include the basic indicator approach (BIA), standardised approach (SA), and advanced measurement approach (AMA). The BIA and SA suggest simple formulae for calculating the regulatory capital. The AMA on the other hand allows financial institutions to use internal models to estimate the operational risk exposure. Many of the bigger banks opt for the AMA approach in order to convince the regulator that the bank require less capital under Pillar I than what the BIA and SA approaches suggest. In this process a “bank using the advanced measurement approach must demonstrate to the satisfaction of its regulator that its systems for managing and measuring operational risk meet established standards, including producing an estimate of operational risk exposure that meets a one-year, 99.9th percentile soundness standard” (see e.g. Cope et al. 2009). Generally, most banks that opt for the AMA use the loss distribution approach (LDA) which is discussed next.

1.2.2. Loss distribution approach

This section focusses on the loss distribution approach (LDA) and how it can be used to calculate regulatory capital. The loss distribution approach is primarily concerned with modelling an aggregate loss distribution in each operational risk category (ORC). The ORC is based on a combination of a business line and operational risk event type and examples will be given later in this section. The aggregate loss distribution in each ORC is constructed as the convolution of an underlying frequency and severity distribution, both of which are estimated from loss data. Sometimes the internal loss data from a bank is augmented with external loss data (e.g. from the ORX database) as well as scenario data (see e.g. de Jongh et al., 2015). The construction of the aggregate loss distribution is discussed in detail in Section 1.2.3 below. After constructing the aggregate loss distribution in each ORC, the overall aggregate loss distribution over all ORC's is formed by assuming some dependence

structure (e.g. some kind of copula). Note that if the regulatory capital estimated in each ORC is summed over all ORC's the resulting estimate assumes perfect dependence between ORC's. On the other hand if all the loss data of the ORC's are pooled and only one aggregate loss distribution constructed from which the regulatory capital is determined, then no dependence is assumed. As shown by Cope et al. (2009) the case of no dependence assumptions will result in total diversification and a much smaller regulatory capital estimate compared to when perfect dependence is assumed. The focus of this thesis will primarily be on the construction of reliable regulatory capital estimates within an ORC. According to Cope et al. (2009) the overall regulatory capital estimate is more sensitive to the individual regulatory capital estimates within an ORC as opposed to the variation in dependence assumptions.

As stated previously, each ORC consists of a combination of business line and event type. The following choices of business lines (BL) are generally recommended (BCBS, 2011):

- Corporate finance (CF),
- Trading and sales (TS),
- Retail banking (RB),
- Commercial banking (CB),
- Payment and settlement (PS),
- Agency and custody (AC),
- Asset management (AM), and
- Retail brokerage (RB).

The following choices of risk types (RT) are generally recommended (BCBS, 2011):

- Internal fraud (IF),
- External fraud (EF),
- Employment practices and workplace safety (EPWS),
- Client, products and business practices (CPBS),
- Damage to physical assets (DPA),
- Business disruption and system failures (BDSF), and
- Execution, delivery and process management (EDPM).

To explain the concept of the LDA in generic terms, consider a bank that has $i = 1, 2, \dots, I$ different business lines and $j = 1, 2, \dots, J$ different risk event types. In total there are then IJ ORC's. Let BL_i refer to the i^{th} business line and RT_j to the j^{th} risk type, then $ORC_{i,j}$ denotes the ORC that corresponds to BL_i and RT_j . The bank will experience operational losses, denoted by $X_{i,j}$, within each $ORC_{i,j}$. Denote the distribution of these losses by $F_{i,j}$ (which is

referred to as the severity distribution), thus $X_{i,j} \sim F_{i,j}$. Furthermore, denote the number of losses within $ORC_{i,j}$, arising within a specific time period, by $N_{i,j}$. Denote the distribution of $N_{i,j}$ by $B_{i,j}$ (referred to as the frequency distribution), thus $N_{i,j} \sim B_{i,j}$. Here, $\{X_{i,j,n}\}_{n=1}^{N_{i,j}}$ are i.i.d. random variables. Furthermore, denote the total (aggregate) loss by $L_{i,j} = \sum_{n=1}^{N_{i,j}} X_{i,j,n}$ having distribution $G_{i,j}$ (referred to as the aggregate loss distribution), thus $L_{i,j} \sim G_{i,j}$.

Within each ORC (and dropping the subscripts i and j), the aggregate loss distribution G is obtained by convoluting the frequency distribution B with the severity distribution F . Regulatory capital within each ORC may now be calculated by determining the appropriate quantile or VaR of G . In the following section methods for estimating the VaR of the aggregate loss distribution will be considered and then, in the following two sections, popular choices for the frequency and severity distributions will be given.

1.2.3. Aggregate loss distribution

The aggregate loss distribution (also referred to as the compound distribution) is a combination of the frequency and severity distributions. Financial institutions can approximate the compound distribution once the frequency and severity distributions have been estimated. It is well known that the compound distribution cannot be written in explicit formulae and should therefore be approximated. This can be done using several methods. The most widely used technique to approximate the compound distribution in practice is through Monte Carlo (MC) simulation. Other approximation techniques include what is known as Panjer recursion (PR, see e.g. Panjer, 2006) and fast Fourier transformation (FFT, see e.g. Embrechts and Frei, 2010). Both PR and FFT have been discussed extensively in the literature. Closed-form approximation techniques for approximating the extreme quantiles of the compound distribution, such as the single loss approximation (SLA, see e.g. Böcker and Klüppelberg, 2005 and Degen, 2010) and perturbative approximations (PA, see e.g. Hernandez et al., 2014), were recently proposed in the literature. These approximation techniques are discussed in much more detail in Chapters 5 and 6. Irrespective of the methodology used to approximate the compound distribution, it is obvious that the compound distribution is constructed from both the severity and frequency distributions. These distributions are discussed next.

1.2.4. Frequency and severity distributions

BCBS (2011) states the following regarding the distributions. “Distributional assumptions underpin most, if not all, operational risk modelling approaches and are generally made for both the frequency and severity of operational risk loss event”. Furthermore, it is stated that “Given the continuing evolution of analytical approaches for operational risk, the Committee does not specify the approach or distributional assumptions used to generate the operational risk measure for regulatory capital purposes. However, a bank must be able to demonstrate that its approach captures potentially severe ‘tail’ loss events”. BCBS (2011) further states that “while it is common for banks to use the Poisson distribution for estimating frequency, there are significant differences in the way banks model severity, including the choice of severity distribution”. It is therefore important that the ‘correct’ (or most accurate) distributional assumptions are made. According to BCBS (2011), the Poisson distribution is most widely used to model the frequency distribution and is discussed next.

Let N be a discrete random variable which represents the number of losses arising within a specific ORC. N is said to have a Poisson distribution if the probability mass function of N is given by

$$b(k; \lambda) = P(N = k) = \frac{\lambda^k e^{-\lambda}}{k!}, \text{ for } k = 0, 1, 2, \dots \text{ and } \lambda > 0.$$

Here, $\lambda = E[N] = Var(N)$ is often referred to as the intensity of the distribution. The intensity of the frequency distribution can easily be estimated using the average number of losses within a specified time frame, which is usually one year in ORM. An alternative distribution for modelling the frequency of losses is the negative binomial distribution. Most banks use the Poisson distribution only although some banks use both. According to Cope et al. (2009) the estimated regulatory capital is more sensitive to the choice of severity distribution than the choice of frequency distribution. The Poisson distribution is used as frequency distribution throughout this thesis.

As stated previously the modelling of the severity distribution should be done carefully, because two different distributions fitted to the same data set could lead to very different estimates of regulatory capital. Dutta & Perry (2007) proposed the following criteria when considering a certain class of distributions when estimating the severity distribution (apart from statistical goodness-of-fit): The model should be realistic, well specified, flexible and simple. BCBS (2011) proposed using statistical tools to examine the statistical properties of each ORC’s data which include scatter plots, empirical distribution plots, histograms, P-P and Q-Q plots, and mean excess plots. Examining operational risk data using these tools will

give the risk manager a sense of possible distributions to consider when estimating the severity distribution. Operational risk data generally has positive skewness and has medium to heavy tails (see e.g. Rachev, 2006). Generally banks use two approaches to model the severity distribution. One approach is to select an appropriate distribution from a wide class of semi to heavy tailed distributions (e.g. the Burr, g-and-h, NIG, LogNIG, LogNormal, LogPhase) and the other to use a so-called spliced distribution where the body and tail of the severity distribution are modelled separately. For example the body could be modelled by using the empirical distribution or fitting the Burr distribution and then use the Generalised Pareto distribution (GPD) to model the tails (see e.g. de Jongh et al., 2013). Since the GPD is motivated from extreme value theory (which will be discussed in more detail in Chapter 2) it is a very popular distribution to model the tail of the severity distribution. According to BCBS (2011) 50% of banks using the AMA apply the spliced distribution approach. The distribution functions of some of the popular severity distributions, namely the Burr and the log normal (LogNormal) are given next.

Note that some of the terms used below will be elaborated upon later. These include the extreme value index, the second order parameter and slowly varying functions, which will be defined in more detail in Chapter 2.

The Burr distribution

The three parameter Burr type XII distribution function is given by

$$Burr(x; \eta, \tau, \kappa) = 1 - (1 + (x/\eta)^\tau)^{-\kappa}, \text{ for } x > 0 \quad (1.4)$$

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

$$burr(x; \eta, \tau, \kappa) = \frac{\tau\kappa}{\eta^\tau} x^{\tau-1} \left[1 + \left(\frac{x}{\eta} \right)^\tau \right]^{-(\kappa+1)}, \text{ for } x > 0 \quad (1.5)$$

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

The LogNormal distribution

The two parameter LogNormal distribution function is given by

$$\text{Lognor}(x; \mu, \sigma) = \frac{1}{2} + \frac{1}{2} \text{erf} \left(\frac{\log(x) - \mu}{\sqrt{2}\sigma} \right) = \Phi \left(\frac{\log(x) - \mu}{\sigma} \right), \text{ for } x > 0 \quad (1.6)$$

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

$$\text{lognor}(x; \mu, \sigma) = \frac{1}{x\sigma\sqrt{2\pi}} e^{-\frac{1}{2} \left(\frac{\log(x) - \mu}{\sigma} \right)^2}, \text{ for } x > 0 \quad (1.7)$$

with parameters μ and σ .

1.3. Practical issues in operational risk modelling

The LDA approach has been discussed in detail in the literature and deficiencies noted, especially by Cope et al. (2009), Embrechts & Hofert (2011) and Nešlehová et al. (2006). From the literature it is clear that many banks globally are facing the same issues in their calculation of operational risk capital using the loss distribution approach (LDA). Three of the main issues are firstly; the dependency structure used for aggregation of RC estimates among different operational risk categories, secondly; the choice of the ORCs (i.e. granularity) and lastly; the choice of severity distributions fitted to heavy-tailed loss data. The first issue of modelling the dependence structure falls outside the scope of this study, but note that extensive research is being conducted on this topic. The popular approach for addressing this issue is by way of copulas. The second issue relates to the choice of ORC which also falls outside the scope of this study. From the literature it is clear that the basic idea here is to choose the business lines and risk type combination in such a way that the ORCs contain sufficient and homogeneous loss data. Choosing too few ORCs will lead to heterogeneity of the loss data whereas too many ORCs will lead to scarcity of data within

each ORC which in turn causes other modelling difficulties. Finally, the heaviness of the tail of the individual loss distribution has much more to do with a reduction in capital charges than the diversification benefit obtained through more realistic dependency modelling options and construction of ORC's (see e.g. Cope et al. 2009). For this reason the focus of this study is on the third issue, i.e. the estimation of regulatory capital within an ORC.

It is well known that the distribution of the loss data encountered in operational risk is typically heavy-tailed. Also, as stated before, the regulatory requirement of estimating the 99,9% quantile or VaR of the annual aggregate loss distribution is equivalent to a one in a 1000 year loss. However, banks have accumulated at most 10 years of historical loss data. So, if only historical data is used, the estimation of the VaR requires extrapolation far beyond the observed data. Cope et al. (2009) analysed historical loss data and found that the estimate of regulatory capital (or VaR) is almost exclusively determined by the tail of the fitted severity distribution and therefore the extreme losses observed. Böcker and Klüppelberg (2005) proved that for a wide class of heavy-tailed distributions (the sub-exponential class), the extreme quantiles of the aggregate annual loss distribution may be approximated by a higher quantile of the underlying severity distribution. Therefore it should be clear that the tail of the fitted severity distribution needs to be modelled very accurately in order to estimate regulatory capital efficiently. Using only historical data, and for that matter less than 10 years of data, makes this task almost impossible. For example, Cope et al. (2009) found that two severity distributions may both fit the data very well in terms of goodness-of-fit statistics, yet may provide capital estimates that differ widely. They also found that regulatory capital estimates are extremely sensitive to high severity low frequency losses and concludes that the challenges to validate operational risk models are considerable and that the extrapolation problem demonstrates how volatile and unreliable estimates of high quantiles for total losses can be. In particular Cope et al. (2009) suggested that research should be done on rather estimating a lower quantile of the severity distribution (e.g. the 90% VaR or equivalently the one in 10 year loss, which can be estimated with some accuracy) and then using a multiplier to obtain an estimate for the one in 1000 year loss. *This will be one of the topics that will be researched further in this thesis.* Another issue related to the estimation of regulatory capital is that the 99,9% VaR is usually obtained by brute force Monte Carlo simulation. For example, once the severity and frequency distributions have been fitted to the loss data, the annual aggregate loss distribution is estimated by a parametric bootstrap of the random sum of losses. This is computationally inefficient and very time consuming. It was mentioned earlier that Böcker and Klüppelberg (2005) showed that a very high quantile of the compound distribution (or aggregate loss distribution) may be approximated by an even higher quantile of the underlying severity

distribution. This first order approximation is referred to as the single loss approximation (SLA) and several authors have done research on this topic in order to improve the approximation. Degen (2010) suggested a second order approximation and recently Hernandez et al. (2014) proposed a perturbative approximation. *This thesis will also focus on comparing these and other methods with brute force Monte Carlo and investigate the computational advantages.*

Note that research related to the improvement of modelling the severity distribution is ongoing. Many authors have suggested that the internal bank data should be augmented with external data (e.g. the pooled ORX data base). This has its own problems since the external data should be scaled (see e.g. Baud et al., 2002) to bring it in line with the bank's internal data. Other authors have criticized the use of only historical data and suggested that the severity distribution should be made 'forward-looking' by including scenarios and expert opinion (see e.g. de Jongh et al., 2015).

1.4. Thesis objectives and layout

Generally, the financial institution would be interested in calculating the capital for OR as a whole. In addition to calculating OR capital, the OR management would also be interested in the capital charge for each individual ORC. This is calculated by use of various risk measures based on the aggregate distribution of each ORC. As mentioned previously, the underlying severity distributions have the greatest influence on the eventual OR capital charge above all others and should therefore be investigated.

As mentioned previously there are two main approaches to calculating the various risk measures of the severity distribution. One approach (full-parametric approach) is to model the severity distribution using one class of distributions. Ahn et al. (2012) proposed using the LogPh distribution when modelling the entire dataset using a single distribution. This distribution is of particular interest as the 'Log-class' of distributions are very popular in OR modelling, e.g. the LogNormal and LogGamma distributions. One major drawback of the LogPh distribution is that the estimation of its parameters can be challenging. Nonetheless, the proposal of the LogPh distribution led to the investigation of the Log Normal Inverse Gaussian (LogNIG) distribution which is a 'Logged' version of the popular Normal Inverse Gaussian (NIG) distribution (see e.g. Venter & de Jongh, 2002). The LogNIG distribution was investigated and produced fairly good results. The other approach (splice distribution approach) allows one to fit the empirical distribution (or some class of distributions) to the

body of the severities (loss data) and then modelling the tail of the severities (extreme loss data) using another class of distributions. It was mentioned that the GPD (supported by EVT) is generally used for modelling the tail part of the severity distribution. Also, estimating, for example, the 99.9% VaR is very difficult due to the uncertainty that arises when extrapolating to such an extreme percentile. Therefore, alternative methods that produce less volatile estimates are desired. This led to the introduction of a scaling method that extrapolates the capital estimate at a lower percentile, say the 95% VaR, to the desired 99.9% VaR. This scaling method is referred to as multipliers in the thesis. The multipliers were developed and investigated and produced very good results.

Moving to the compound distribution which is more complex as it is a combination of both the severity and frequency distributions. As mentioned previously, there are a number of approaches to approximating the compound distribution. Some of these approaches rely on simulation, such as the MC techniques. The choice of the number of runs in a MC simulation is very important. A larger number of runs is usually required for heavier tailed distributions than for medium to light tailed distributions implying that a large number of runs will be required when using MC techniques for approximating the compound distribution (or relevant quantities of the compound distribution). This can be problematic as MC simulation can be time-consuming for large number of runs. Other iterative procedures proposed in the literature are Panjer recursion (PR) and fast Fourier transforms (FFT). One major drawback of both PR and FFT is that experts on these techniques are required to determine the values of the parameters that are used within each of the techniques. Alternative (closed-form) techniques to approximating the compound distribution also exist. These are the single loss approximation (SLA) and perturbative approximation (PA) techniques. In the literature there is no comprehensive study that compares the relative performance of these approximation techniques altogether. Therefore, an extensive simulation comparison was conducted where the FFT and PA produced very good results. For a review and evaluation of these methods see de Jongh et al. (2016).

Clearly there are numerous ways to calculate the VaR for both the severity and compound distributions. Given this information, it remains appropriate to ask the following question. *What estimation technique performs uniformly the best when calculating the VaR?* This study aims to answer this question by investigating the various techniques proposed in the literature, as well as investigating other techniques proposed in the thesis. The layout of the thesis is given next.

Chapters 2 - 4 focus on estimation techniques based on the severity distribution. The estimation techniques proposed in the literature (such as those arising from extreme value theory) are discussed in Chapter 2 whereas the newly proposed methods from this study (including the use of the multipliers and the LogNIG distribution) are discussed in Chapter 3. A simulation study was conducted to compare all the appropriate techniques discussed in Chapters 2 and 3. The simulation design, its results and recommendations on these techniques are discussed in Chapter 4.

Chapters 5 and 6 focus on the proposed techniques for the compound loss distribution. A theoretical discussion of these techniques is given in Chapter 5. Another simulation study was conducted to compare the approximation techniques for the compound distribution. The findings of this simulation study are discussed in Chapter 6.

Remark: Recall the definitions of VaR and ES as $VaR_\alpha(X) = F_X^{-1}(\alpha)$ and $ES_\alpha(X) = \frac{1}{1-\alpha} \int_\alpha^1 VaR_\varepsilon(X) d\varepsilon$ respectively. Also, the p^{th} -quantile of a random variable X with cumulative distribution function F_X , denoted by $q_p(X)$, can be written as $q_p(X) = F_X^{-1}(p)$ if F_X is continuous with F_X^{-1} the inverse of F_X . Setting α and p equal to z leads to $ES_z(X) = \frac{1}{1-z} \int_z^1 VaR_\varepsilon(X) d\varepsilon$ with $VaR_z(X) = F_X^{-1}(z)$ and $F_X^{-1}(z) = q_z(X)$. Thus, the risk measures of interest can be calculated once the quantile function is available. Therefore, the terminology quantile will be used instead of VaR and/or ES.

Chapter 2: ‘Existing’ quantile estimation techniques for the severity distribution

Quantiles are traditionally estimated by using order statistics or parametric approaches based on popular distributions, e.g. the normal distribution. In financial risk management, a method for estimating quantiles for regulatory purposes was also proposed by the BCBS (Danielsson et al., 1998). Techniques arising from extreme value theory became popular more recently (Coles, 2001 and Beirlant et al., 2004). These techniques, including others, are discussed in this chapter.

Throughout this chapter $X_1, X_2, \dots, X_n$ denotes a sample of i.i.d. observations from some distribution F and $X_{1,n} \leq X_{2,n} \leq \dots \leq X_{n,n}$ the corresponding order statistics.

2.1. Traditional quantile estimation techniques

2.1.1. Quantile estimation using order statistics

An estimate of the p^{th} -quantile q_p can be taken as $\hat{q}_p^{OS} = X_{[pn],n}$ with $[k]$ the largest integer less than or equal to k . Here, $p \in (0,1)$ is some probability level.

Using order statistics to estimate high quantiles can be problematic when working with relatively small samples. Assume a risk manager has to estimate the 99%-quantile using a sample of 100 observations. In this case $p = 0.99$ and $n = 100$. The 99%-quantile estimate is $\hat{q}_{0.99}^{(OS)} = X_{99,100}$ when using order statistics. Here, $X_{99,100}$ is the second largest observation from the available data $X_{1,n} \leq X_{2,n} \leq \dots \leq X_{n,n}$. As mentioned in the previous chapter, data used in ORM is often heavy-tailed which in turn means that $X_{99,100}$ may be very volatile. Thus, the quantile estimate will also not be very stable in this case. Obtaining stable quantile estimates are even more difficult for smaller samples and/or larger probability levels which are usually the case in ORM practice.

2.1.2. Quantile estimation using a parametric approach

Assume that the sample $X_1, X_2, \dots, X_n$ was obtained from a distribution with distribution function $F(\cdot; \underline{\theta})$, with F known and $\underline{\theta}$ unknown. One can easily estimate the parameters $\underline{\theta}$ as

$\hat{\theta}$ using any well-known estimation procedure, for example maximum likelihood estimation. An estimate of the p^{th} -quantile q_p can then be obtained as $\hat{q}_p^{(PAR)} = F^{-1}(p; \hat{\theta})$. Once again, $p \in (0,1)$ is some probability level and F^{-1} is defined in the usual way as

$$F^{-1}(u) = \inf\{x: F(x) \geq u\}.$$

Using a parametric approach may, however, also cause problems when working with small samples and high probability levels. The following example illustrates this statement.

Generate $n = 100$ observations $X_1, X_2, \dots, X_{100}$ from $F(\cdot; \mu, \sigma)$ where F is the LogNormal distribution with $\mu = 0$ and $\sigma = 1$ the location and scale parameters respectively. Figure 2.1 gives a histogram of the generated data.

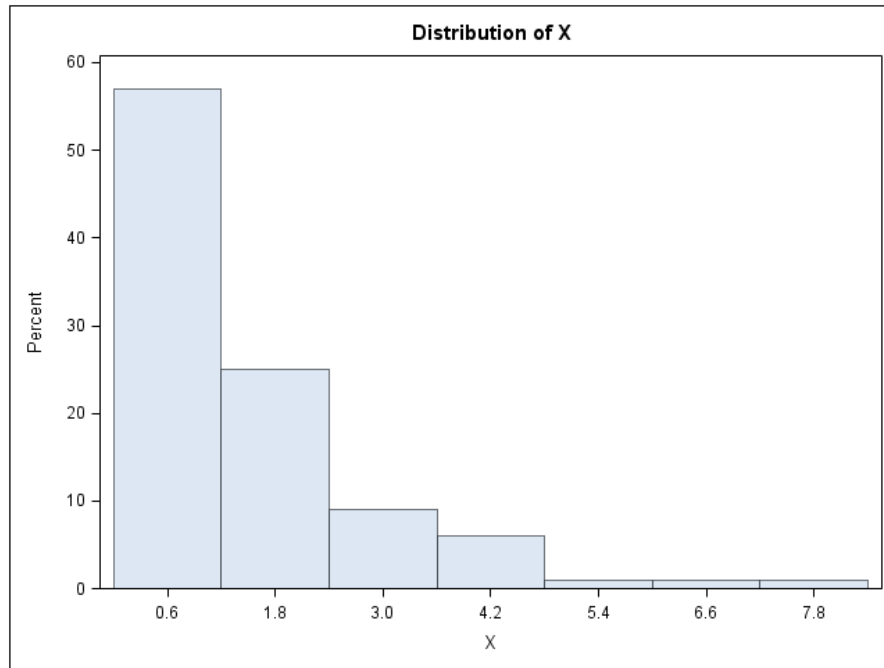


Figure 2.1: Histogram of LogNormal data with $\mu = 0$ and $\sigma = 1$

The LogNormal distribution was fitted to the sample using maximum likelihood. The density of the fitted distribution was then superimposed on the histogram which illustrates graphically how the fitted distribution differs from the actual data. This is given in Figure 2.2. A quantile-quantile plot (QQ-plot) was also constructed to investigate how well the fitted distribution represents the actual data, especially in the upper tail. This is given in Figure 2.3.

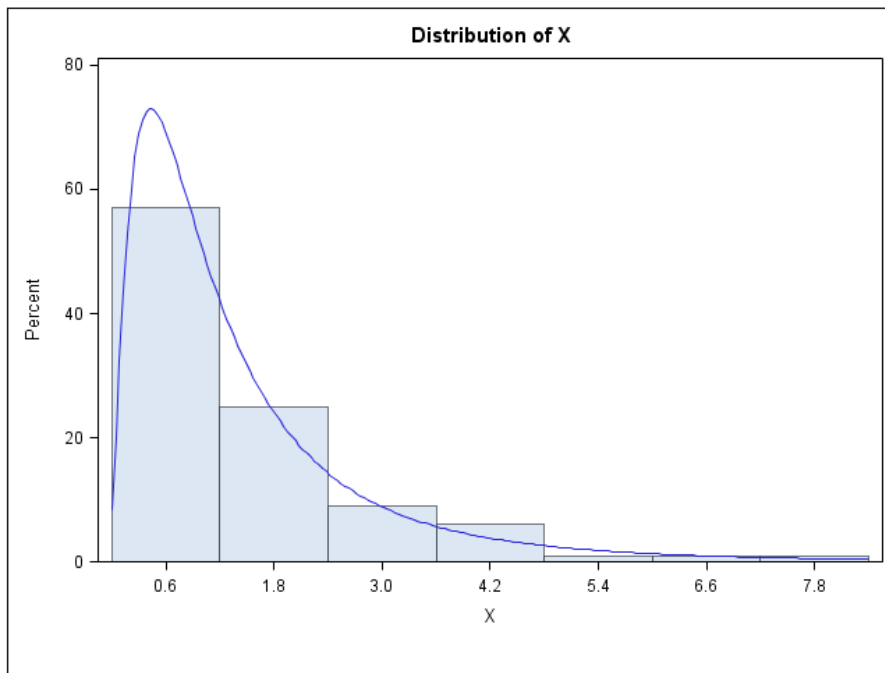


Figure 2.2: Fitted density and histogram of LogNormal data with $\mu = 0$ and $\sigma = 1$

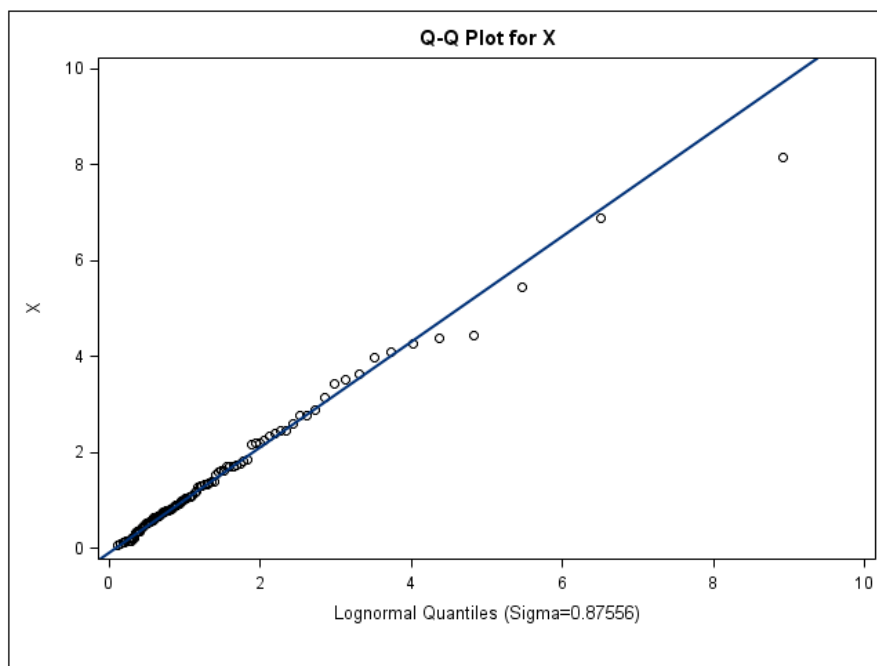


Figure 2.3: QQ-plot of LogNormal data with $\mu = 0$ and $\sigma = 1$

It is not clear from Figure 2.2 whether or not the fitted distribution represents the actual data in the upper tail. In Figure 2.3 a QQ-plot is given of the data together with the expected 45 degree line. Goodness-of-fit is measured by the closeness of the points to this line. Observe that although the lower values lie closely to the straight line, the largest observation deviates substantially from this line with the previous four observations also deviating a fair amount from the line. This means that quantile estimates can become unstable when p becomes too

close to 1. This raises concern for a risk manager as this is the region that is of particular interest. This problem is amplified when working with even smaller samples.

2.2. Extreme value theory

It was mentioned in the previous section that quantile estimates based on order statistics can be unreliable due to the variability in using a single very large order statistic. Also, the parametric approach relating to the entire sample discussed in the previous section does not necessarily model the tail accurately. These shortfalls cause concern for a risk manager as quantiles in the tail are typically of interest. Other quantile estimation techniques addressing these issues should therefore be investigated. One such technique is derived directly from extreme value theory and is discussed next.

There are two approaches to addressing the issues of variability in the tail of small samples. These are:

1. The block maxima (BM) approach whereby the sample is divided into subsamples, known as blocks. The block maxima are then used to estimate the desired extreme quantile, and
2. The peaks over threshold (POT) approach whereby the observations exceeding some threshold are used to estimate the high quantiles.

Both these approaches translate the non-parametric problem of estimating the quantiles of an unknown distribution F into a parametric problem, using asymptotic results. The first approach uses the Fisher-Tippett (FT) Theorem to achieve this and the second approach uses the Pickands-Balkema-deHaan (PBdH) Theorem for it. A necessary and sufficient condition for these theorems to hold will be given in the remarks following the formulation of the theorems.

Fisher-Tippett Theorem:

Let $M_n = \max(X_1, X_2, \dots, X_n)$. If a sequence of pairs of real numbers (a_n, b_n) exists such that each $a_n > 0$ and for all x

$$\lim_{n \rightarrow \infty} P\left(\frac{M_n - b_n}{a_n} \leq x\right) = G(x),$$

with G a non-degenerate distribution function, then the limit distribution G belongs to the Generalized Extreme Value (GEV) family of distributions, given by

$$G_\gamma(x) = \begin{cases} \exp\left(-(1 + \gamma x)^{-\frac{1}{\gamma}}\right) & \text{if } \gamma \neq 0 \\ \exp(-e^{-x}) & \text{if } \gamma = 0 \end{cases}$$

with $1 + \gamma x > 0$.

(See e.g. Fisher & Tippett, 1928 and Gnedenko, 1943.)

Remark 1: The distribution function G_γ is a parametric distribution. Denote the set of all distributions that satisfy the Fisher-Tippett Theorem by $\mathcal{D}(G_\gamma)$. A generic distribution in $\mathcal{D}(G_\gamma)$ will be denoted by F .

Remark 2: The parameter γ is called the EVI and is an indication of the tail type of the underlying F . Three sub-classes are distinguished based on the sign of γ . $\gamma > 0$ is associated with the Fréchet-Pareto class having power function tails, $\gamma = 0$ is associated with the Gumbel class consisting of distributions having exponential type tails and $\gamma < 0$ consisting of distributions having a finite right-hand end point.

Remark 3: The Fisher-Tippett Theorem forms the basis of extreme value theory. However, the block maxima method will not be used in this thesis. See Beirlant et al. (2004) for further details.

A more attractive and more modern approach is the POT method. This is based on the Pickands-Balkema-deHaan Theorem which will now be discussed.

Let $X \sim F$ and let F_u denote the conditional distribution function of $X - u$, given that $X > u$. Furthermore, let

$$H(x; \sigma, \gamma) = \left(1 + \frac{\gamma x}{\sigma}\right)^{-\frac{1}{\gamma}}, \quad x, \gamma, \sigma > 0,$$

denote the Generalized Pareto Distribution (GPD).

Pickands-Balkema-deHaan Theorem:

A positive function $\beta(u)$ exists such that

$$\lim_{u \rightarrow \infty} \sup_{x > 0} |F_u(x) - H(x; \beta(u), \gamma)| = 0$$

if and only if $F \in \mathcal{D}(G_\gamma)$.

(See Balkema & de Haan, 1974, Pickands, 1975 and McNeil et al., 2005.)

Remark 1: The distribution function H is a parametric distribution. In what follows only the POT method will be used. The POT approach is based on this Theorem. A discussion of its implication will be given after the following remarks.

Remark 2: From the Pickands-Balkema-deHaan Theorem $\bar{F}(y) \approx \bar{F}(u)\bar{H}(y - u; \sigma, \gamma)$ and if u is chosen such that $\hat{\bar{F}}(u) = \frac{k}{n}$, with k the number of exceedances above u and both σ, γ are estimated using e.g. maximum likelihood, then the quantiles of F can be approximated by inverting $\bar{F}(u)\bar{H}(y - u; \sigma, \gamma)$. However, in the special case of $\gamma > 0$, that an operational risk manager is typically interested in, more powerful methods exist for this (see below).

Remark 3: The choice of u is an important consideration. The threshold is often taken as the $(k + 1)^{th}$ -largest observation implying that the sample of exceedances has size k . The choice of u then translates into the choice of k . Note also that this means that the k largest order statistics are used in the analyses. Typically k will depend on n and grow with n . The choice of k will be discussed further in the simulation study of Chapter 4.

Remark 4: Denote $Q(u) = F^{-1}(u)$ and define the tail quantile function (TQF) by $U(y) = Q\left(1 - \frac{1}{y}\right)$, $y > 1$.

Assume that a positive function a exists such that for $t > 0$

$$\lim_{x \rightarrow \infty} \frac{U(tx) - U(x)}{a(x)} = h_\gamma(t) = \frac{t^\gamma - 1}{\gamma}. \quad (\mathcal{C}_\gamma(a))$$

Then $F \in \mathcal{D}(G_\gamma)$ if and only if $(\mathcal{C}_\gamma(a))$ holds.

(See e.g. de Haan & Ferreira, 2006.)

Remark 5: Note that $(\mathcal{C}_\gamma(a))$ is a necessary and sufficient condition for both the Fisher-Tippett and Pickands-Balkema-deHaan Theorems.

Remark 6: The following analogy can be drawn between the Central Limit Theory (CLT) and the Extreme Value Theory (EVT). The CLT, via the Central Limit Theorem, translates a non-parametric problem into a parametric one by obtaining a limit distribution of the sum/average of the observed random variables. Similarly, the non-parametric problem is translated via the Fisher-Tippett Theorem or the Pickands-Balkema-deHaan Theorem to a parametric one, where a limit distribution is obtained for the extreme of the observed random variables or the conditional exceedance distribution.

Consider the following numerical example to illustrate the POT method. Generate 100 i.i.d. random variables $X_1, X_2, \dots, X_{100}$ from the LogNormal distribution with $\mu = 0$ and $\sigma = 1$. Devide these 100 observations into 10 different sub-samples, denoted by $\underline{X}_{s=1}, \underline{X}_{s=2}, \dots, \underline{X}_{s=10}$, where each sub-sample has an equal number (10) of observations. For each sub-sample s generate a uniform random variable $t \sim U(s-1, s)$ for each observation within that sub-sample, leading to $t_1, t_2, \dots, t_{100}$. Constructing a scatterplot of t against X leads to the time series illustrated in Figure 2.4 below.

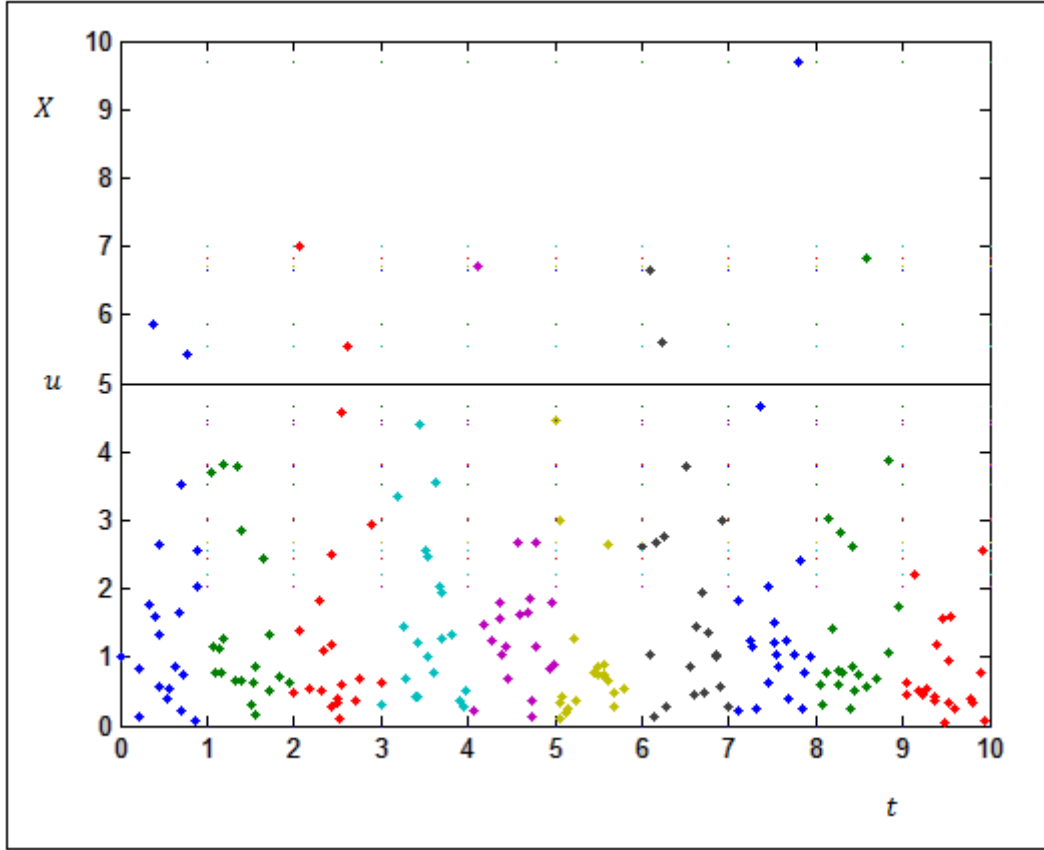


Figure 2.4: Artificial time series data from the LogNormal distribution with $\mu = 0$ and $\sigma = 1$

According to the Pickands-Balkema-deHaan Theorem the observations exceeding the threshold level u follows the GPD if and only if $F \in \mathcal{D}(G_\gamma)$ for u sufficiently large, with F taken as the $\text{Lognor}(x; 0, 1)$ distribution and $u = 5$ in the example depicted in Figure 2.4.

The study in this thesis focuses on heavy-tailed distributions. The class of heavy-tailed distributions that will be considered is the so-called Pareto-type distributions. These distributions $F \in \mathcal{D}(G_\gamma)$ with $\gamma > 0$ can be written as

$$1 - F(x) = x^{-\frac{1}{\gamma}} \ell_F(x) \quad (2.1)$$

where ℓ_F is a slowly varying function, i.e.

$$\lim_{t \rightarrow \infty} \frac{\ell_F(tx)}{\ell_F(x)} = 1, \text{ all } t > 0.$$

Distributions satisfying (2.1) are said to be of Pareto-type. If (2.1) holds, then it can be shown, using so-called de Bruyn conjugates, that the corresponding tail quantile function is of the form

$$U(y) = y^\gamma \ell_U(y) \quad (2.2)$$

with ℓ_U also slowly varying (see e.g. Beirlant et al., 2004).

The parameter γ is a first order parameter, called the extreme value index (EVI). A second order parameter is also considered which is discussed later. The best known and most often used estimator of the EVI γ , introduced by Hill (1975) (Beirlant et al., 2004; Coles, 2001), is defined as

$$\hat{\gamma}_k^{(Hill)} = \frac{1}{k} \sum_{i=1}^k \log X_{n-i+1,n} - \log X_{n-k,n} \quad (2.3)$$

for some $k = 1, 2, \dots, n - 1$.

Note that $k + 1$, the number of order statistics used in $\hat{\gamma}_k^{(Hill)}$, corresponds to the choice of $u = X_{n-k,n}$, leading to k exceedances in the POT approach.

Various other methods to estimate γ have been proposed in the literature. Some of these are a robust estimator proposed by Beran & Schell (2012), a generalized median estimator proposed by Brazauskas & Serfling (2000), the generalized Pickands estimator (Segers, 2005) and a local moment estimator proposed by Goegebeur et al. (2013). See Beirlant et al. (2004) for further discussion of alternative estimators.

Remark: The estimator proposed by Beran and Schell (2012) was investigated in the initial stages of this study due to its robustness. It was found that it does not compete with the Hill (1975) estimator and was therefore not considered further.

Different ways of choosing the value of k have been proposed in the literature. One such well-known approach is based on the mean excess function (MEF) as follows. For a random variable X , its mean excess function for some threshold u is given by

$$e(u) = E[X - u | X > u].$$

An estimator of $e(u)$ is given by

$$\hat{e}_n(u) = \frac{\sum_{i=1}^n X_i I_{[X_i > u]}}{\sum_{i=1}^n I_{[X_i > u]}} - u$$

with

$$I_{[X_i > u]} = \begin{cases} 1, & X_i > u \\ 0, & \text{otherwise} \end{cases}$$

(Beirlant et al., 2004).

The empirical MEF $\hat{e}_n(u)$ can be plotted against the values of $u = X_{n-k,n}$ with $k = 1, 2, \dots, n - 1$, and k is chosen as the largest value where the plot is still reasonably stable.

As an example, generate $n = 100$ observations $X_1, X_2, \dots, X_n$ from $F(\cdot; \alpha)$ where F is the Pareto distribution with $\gamma = 1$. The values of the empirical MEF can be calculated and used to obtain the mean excess plot (as in Figure 2.5).

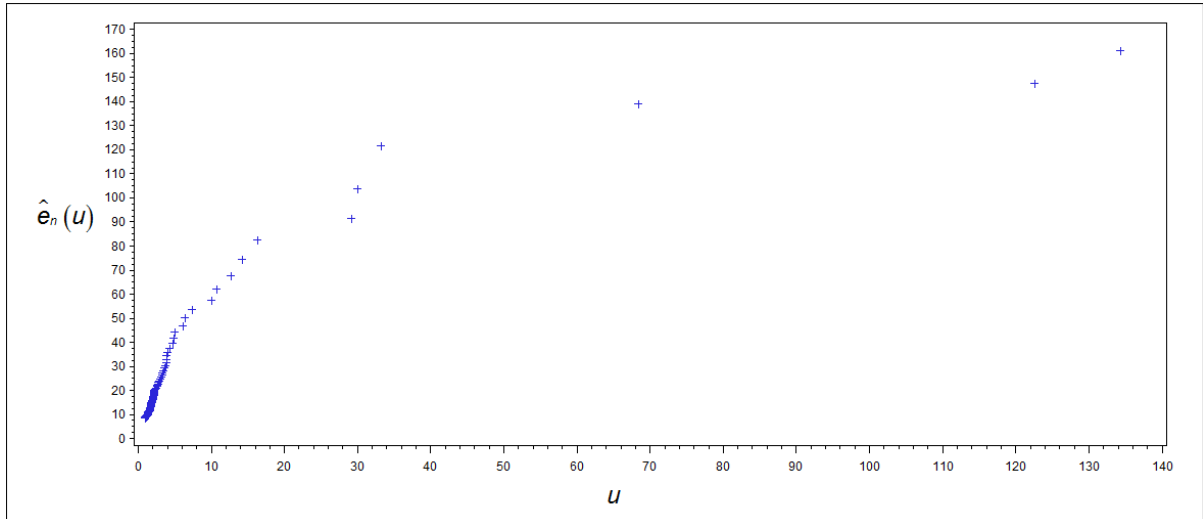


Figure 2.5: Mean excess plot of Pareto data with $\gamma = 1$

It is observed from Figure 2.5 that the mean excesses become unstable for values of $u > \pm 17$. This choice of u can then be used to determine the ‘appropriate’ value of k . Note that this approach allows subjectivity as the decisions made rely on visual inspection of the mean excess plot only. Caution should be taken when using the MEF for the purpose of determining the value of k .

Some other proposed methods for choosing the value of k are given in Beirlant et al. (2004), p. 123. Furthermore, another approach is also proposed in Chapter 4.

2.2.1. Quantile estimation using first order extreme value theory

The basic condition $(\mathcal{C}_\gamma(a))$ can be considered as a first order condition on the tail quantile function $U(y)$. For the Fréchet-Pareto case ($\gamma > 0$) this led to so-called Pareto-type distributions, with the tail quantile functions of the form in (2.2), i.e. $U(y) = y^\gamma \ell_U(y)$, with ℓ_U a slowly varying function. This further led to the well-known Hill estimator for γ . This form of the tail quantile function is also the basis for the well-known Weissman (first order) quantile estimator (Weissman, 1978), which is now briefly described.

Using (2.2), write

$$\frac{U(xu)}{U(x)} = u^\gamma \frac{\ell_U(xu)}{\ell_U(x)} \approx u^\gamma \Rightarrow U(xu) \approx U(x)u^\gamma, \text{ for } x \text{ large.}$$

Choose

$$xu = \frac{1}{\alpha}, x = \frac{n+1}{k+1}, \text{ giving } u = \frac{k+1}{\alpha(n+1)},$$

then for n large

$$Q(1 - \alpha) = U\left(\frac{1}{\alpha}\right) \approx U\left(\frac{n+1}{k+1}\right) \left(\frac{k+1}{\alpha(n+1)}\right)^\gamma.$$

Noting that with $U_{1,n}, U_{2,n}, \dots, U_{n,n}$ the order statistics from a $U(0,1)$ distribution, the probability integral transformation gives

$$X_{n-k,n} = U\left(\frac{1}{U_{k+1,n}}\right) \approx U\left(\frac{n+1}{k+1}\right),$$

where in the last step $U_{k+1,n}$ is approximated by its expected value $\frac{k+1}{n+1}$. Substituting in the above and estimating γ by the Hill estimator $H_{k,n}$, then leads to the Weissman estimator

$$\hat{Q}(1 - \alpha) \equiv \hat{Q}(p) = \hat{q}_{k,p}^{(W)} \equiv X_{n-k,n} \left(\frac{k+1}{(1-p)(n+1)} \right)^{H_{k,n}}. \quad (2.4)$$

See also Section 4.6.1 in Beirlant et al. (2004) for a discussion of this estimator.

Remark: Gomes et al. (2013) proposed the so-called mean-of-order- p (MOP) quantile estimator which is equivalent to the Weissman quantile estimator given in Equation (2.4) if the order of p is chosen as zero. This is convenient as the statistical software program ‘R’ has a package (dubbed ‘evt0’) readily available that can be used to apply the MOP quantile estimator (or equivalently the Weissman estimator).

2.2.2. Quantile estimation using second order extreme value theory

The discussion that follows is based on Sections 4.4 and 4.6.2 in Beirlant et al. (2004). The Weissman quantile estimator based on a first order condition on the tail quantile function, was discussed in Section 2.2.1. Due to the approximations used in deriving the estimator, it suffers from bias and in order to reduce this, a second order assumption is made. This assumption can be formulated in terms of the tail quantile function or the slowly varying function.

In order to state the assumption, the notion of a regularly varying function is needed. A function g is regularly varying of order ν if

$$\lim_{x \rightarrow \infty} \frac{g(tx)}{g(x)} = t^\nu, \quad t > 0.$$

This is denoted by $\mathcal{R}_\nu$. (See e.g. de Haan & Ferreira, 2006)

Second order assumption (in terms of the slowly varying function)

The slowly varying function ℓ_U satisfies a second order condition if $\log \ell_U \in C_\rho(b)$, for some $\rho < 0$ and $b \in \mathcal{R}_\rho$. Therefore, under this assumption

$$\lim_{x \rightarrow \infty} \frac{\log \ell_U(xu) - \log \ell_U(x)}{b(x)} = h_\rho(u) = \frac{u^\rho - 1}{\rho}.$$

Thus, if this assumption holds then for large x

$$\frac{\ell_U(xu)}{\ell_U(x)} = \exp[\log \ell_U(xu) - \log \ell_U(x)] \approx \exp[b(x)h_\rho(u)]. \quad (2.5)$$

Now, since $\frac{U(xu)}{U(x)} = u^\gamma \frac{\ell_U(xu)}{\ell_U(x)}$, substitution from (2.5) gives

$$\begin{aligned} U(xu) &= U(x)u^\gamma \frac{\ell_U(xu)}{\ell_U(x)} = U(x)u^\gamma \exp[\log \ell_U(xu) - \log \ell_U(x)] \\ &\approx U(x)u^\gamma \exp[b(x)h_\rho(u)] \\ &= U(x)u^\gamma \exp\left[b(x)\frac{1-u^\rho}{-\rho}\right]. \end{aligned}$$

As in the case of the Weissman estimator, choose $xu = \frac{1}{\alpha}$, $x = \frac{n+1}{k+1}$, and thus $u = \frac{k+1}{\alpha(n+1)}$, and again using $X_{n-k,n} = U\left(\frac{1}{U_{k+1,n}}\right) \approx U\left(\frac{n+1}{k+1}\right)$, give for n large that

$$\begin{aligned} Q(1-\alpha) &= U\left(\frac{1}{\alpha}\right) \approx U\left(\frac{n+1}{k+1}\right) \left(\frac{k+1}{\alpha(n+1)}\right)^\gamma \exp\left[b\left(\frac{n+1}{k+1}\right) \frac{1 - \left(\frac{k+1}{\alpha(n+1)}\right)^\rho}{-\rho}\right] \\ &= X_{n-k,n} \left(\frac{k+1}{\alpha(n+1)}\right)^\gamma \exp\left[b_{n,k} \frac{1 - \left(\frac{k+1}{\alpha(n+1)}\right)^\rho}{-\rho}\right], \end{aligned}$$

with $b_{n,k} \equiv b\left(\frac{n+1}{k+1}\right)$. Replacing γ , ρ and $b_{n,k}$ with their respective estimators, gives a second order Weissman quantile estimator.

The relevant parameters can be estimated by maximum likelihood as follows. Under the second order assumption, it can be shown that in distribution

$$j \left(\log(X_{n-j+1,n}) - \log(X_{n-j,n}) \right) \approx \left(\gamma + \left(\frac{j}{k+1} \right)^{-\rho} b_{n,k} \right) E_j, \quad j = 1, 2, \dots, k,$$

with $\{E_j\}$ i.i.d. standard exponential random variables (see e.g. Beirlant et al., 2004). Using this, an approximate likelihood function can be formed from which maximum likelihood estimators can be obtained. Denoting the latter by $\hat{\gamma}$, $\hat{\rho}$, and $\hat{b}_{n,k}$ leads to the bias reduced second order Weissman estimator

$$\hat{Q}(1 - \alpha) \equiv \hat{Q}(p) = \hat{q}_{k,p}^{(W+)} \equiv X_{n-k,n} \left(\frac{k+1}{(1-p)(n+1)} \right)^{\hat{p}} \exp \left[\hat{b}_{n,k} \frac{1 - \left(\frac{k+1}{(1-p)(n+1)} \right)^{\hat{p}}}{-\hat{p}} \right]. \quad (2.6)$$

Remark 1: Note that the second order estimator has the additional term $\exp \left[\hat{b}_{n,k} \frac{1 - \left(\frac{k+1}{(1-p)(n+1)} \right)^{\hat{p}}}{-\hat{p}} \right]$ compared to the first order estimator. This leads to two extra parameters $b_{n,k}$ and ρ that need to be estimated. Thus, even though this second order quantile estimator reduces the bias, it increases the variance due to the additional parameter estimation.

Remark 2: Once again, the reader is referred to Beirlant et al. (2004) for more information regarding the second order quantile estimate, as well as estimates of the parameters used in (2.6), see e.g. Gomes et al., 2010, Berning, 2010 and de Wet et al., 2012.

In this chapter the traditional estimators for extreme quantiles were discussed, namely $\hat{q}^{(OS)}$ (using ordered statistics) and $\hat{q}^{(PAR)}$ (using a parametric approach). Extreme value theory was then introduced and it was shown how this theory leads to estimation of such quantiles, both first order $\hat{q}^{(W)}$ and second order $\hat{q}^{(W+)}$, the former depending on the well-known Hill estimator and the latter depending on maximum likelihood estimates of the second order parameters.

Remark 3: Gomes et al. (2013) proposed the so-called reduced-bias mean-of-order p (RBMOP) quantile estimator which is equivalent to the bias reduced second order Weissman quantile estimator given in Equation (2.6) if the order of p is chosen as zero. Once again, this is convenient as the statistical software program ‘R’ has a package ‘evt0’ readily available that can be used to apply the RBMOP quantile estimator (or equivalently the second order Weissman estimator).

Since none of the discussed estimators is uniformly best, a new estimator is proposed in the next chapter. The definition of this estimator has some commonality with a BCBS company specific risk adjusted proposal (see Section 3.1.1).

In Chapter 4 the performance of the various quantile estimators that were discussed, will be compared by means of a simulation study, as well as an empirical study.

Chapter 3: ‘New’ quantile estimation techniques for the severity distribution

In this chapter a new estimator for extreme quantiles is proposed and discussed. This estimator has a vague relationship to a BCBS estimator, but using results from EVT.

3.1. Quantile estimation using a multiplying factor

3.1.1. A quantile estimation technique proposed by the Basel Committee for market risk quantification

Assume the sample $X_1, X_2, \dots, X_n$ comes from a $N(\mu, \sigma^2)$ distribution. The parameters (μ, σ) can be estimated as $(\hat{\mu}, \hat{\sigma})$ using any well-known procedure such as maximum likelihood. An estimate of the p^{th} -quantile q_p can then be obtained as $\hat{q}^* = \hat{\mu} + \hat{\sigma}\Phi^{-1}(p)$ with Φ the $N(0,1)$ distribution function. The BCBS proposed that this estimate should be multiplied by R , with $3 \leq R \leq 4$ a company-specific risk factor (Danielsson et al., 1998). An adjusted estimate of q_p is then $\hat{q}_p^{(B)} = \hat{q}^* \times R$.

Assuming that the sample $X_1, X_2, \dots, X_n$ has a Normal distribution does not prove sound risk management practice as it is known that data in FRM are typically not symmetric. The BCBS therefore proposed multiplying the quantile estimate based on the Normal distribution by a factor R to compensate for the skewness of the actual tail.

Even though this proposal is for market risk quantification, it was used as motivator for a new quantile estimator, the multiplier, to be discussed next.

3.1.2. The basic idea

Consider a heavy-tailed distribution F and let $q_{p_i} = F^{-1}(p_i; \underline{\theta})$ denote the p_i^{th} -quantile with $i = 0, 1$. These quantiles are illustrated graphically in Figure 3.1.

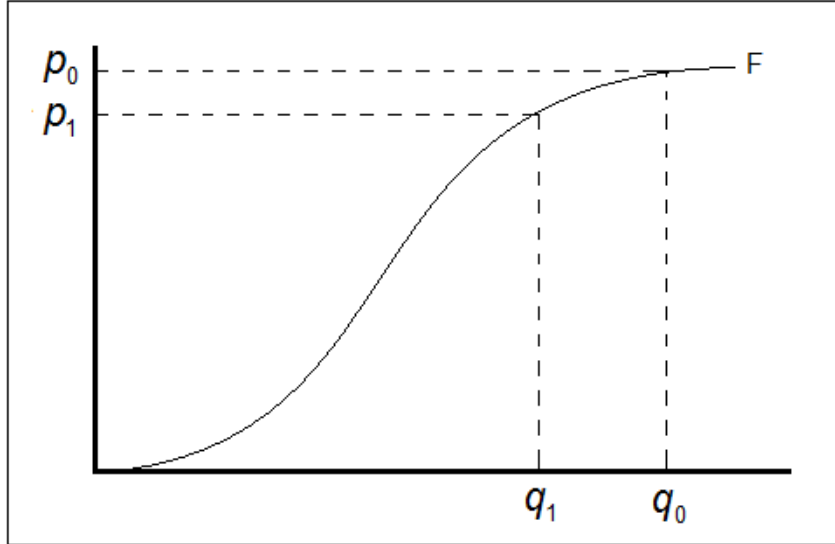


Figure 3.1: Quantiles of F

The basic idea of the multiplier method is the following. First note that $q_{p_1} < q_{p_0}$ for $p_1 < p_0$. Also, q_{p_1} can usually be estimated more reliably than q_{p_0} . Now, for a distribution F , write $q_{p_0} = C_F q_{p_1}$, with $C_F = \frac{q_{p_0}}{q_{p_1}}$. The idea is to estimate q_{p_1} from the available data and obtain through simulation a value C_* that ‘approximates’ C_F over a range of distributions $F \in \mathcal{F}$. C_* is referred to as the multiplier. q_{p_0} is then estimated by $\hat{q}_{p_0} = C_* \hat{q}_{p_1}$ with $\hat{q}_{p_1}$ an estimate of q_{p_1} .

The following two questions arise:

1. Which quantile estimation technique should be used to estimate the lower quantile q_{p_1} ?
2. Which value of the multiplier C_* should be used?

Financial risk managers are generally interested in quantiles for $p_0 \geq 0.99$. It is known that these quantiles are difficult to estimate, especially when working with small samples. It is realistic to assume that for small samples ($n \leq 200$), one can estimate a quantile more accurately for $p_1 \leq 0.95$ when using EVT. Thus, EVT is a natural choice for estimating the lower quantile q_{p_1} , making an assumption in respect of the first question.

A simulation study was conducted to answer the second question. The design of this simulation study is described in the next subsection.

3.1.3. Deriving the multiplying factors using simulation

A class $\mathcal{F}$ of heavy-tailed distributions was constructed and over this class a ‘best’ multiplier was obtained for each combination of the following; the sample size (n), lower probability (p_1) and higher probability (p_0).

The following values of the parameters were used in the simulation study.

- Sample sizes n of $n_1 = 20$, $n_2 = 50$, $n_3 = 100$, $n_4 = 150$ and $n_5 = 200$,
- Higher probabilities p_0 of $p_{0,1} = 0.99$, $p_{0,2} = 0.995$ and $p_{0,3} = 0.999$,
- Lower probabilities p_1 of $p_{1,1} = 0.95$, $p_{1,2} = 0.9$, $p_{1,3} = 0.85$ and $p_{1,4} = 0.8$,
- Number of threshold excesses k where this was taken as 20% of the sample size, i.e. $k = 0.2 \times n_i$ with $i = 1, 2, \dots, 5$,
- Distributions used that make up $\mathcal{F}$ are $F_1, F_2, \dots, F_{10}$, see Table 3.1 below,
- The quantiles $q_{p_{d,l},r} = F_r^{-1}(p_{d,l})$, with $r = 1, 2, \dots, 10$, $d = 0, 1$, $l = 1, 2, 3$ if $d = 0$ and $l = 1, 2, 3, 4$ if $d = 1$, and
- The multiplier $C_{r,p_0,a,p_1,b}^{(true)} = \frac{q_{p_0,a,r}}{q_{p_1,b,r}}$ for any choice of a and b for a given F_r .

Remark: As mentioned in the previous chapter, choosing the value of k still remains a major question in EVT. In this study the number of threshold excesses was fixed at 20% of the sample size. This decision was made after comparing the results of taking 10% and 20% of the sample size and also using $k = k_{opt}$, with k_{opt} determined by the method of Drees and Kaufmann (1998). The results obtained from all of these techniques remained relatively the same.

A class of distributions $\mathcal{F}$ was constructed having $0.2 \leq \gamma \leq 0.8$ and $-2 \leq \rho \leq -0.2$. This class was used to develop the multipliers for each combination of $(p_{0,a}, p_{1,b}, n_i)$. Table 3.1 contains the distributions that make up this class.

Table 3.1: Distributions used to derive the multiplying factors

F_r	Distribution	γ	ρ	$q_{p_{0,1},r}$	$q_{p_{0,2},r}$	$q_{p_{0,3},r}$
F_1	<i>BurrXII</i> (1,5,1)	0.2	-1	2.506842	2.882509	3.980275
F_2	<i>BurrXII</i> (1,5,0.5)	0.4	-2	6.309447	8.32549	15.84893
F_3	<i>BurrXII</i> (1,2.5,0.5)	0.8	-2	39.80913	69.31379	251.1885
F_4	<i>BurrXII</i> (1,0.25,5)	0.8	-0.2	5.224884	12.63612	78.97501
F_5	<i>BurrXII</i> (1,1.25,1)	0.8	-1	39.49191	69.03709	250.9877
F_6	<i>BurrXII</i> (1,3,2/3)	0.5	-1.5	9.882685	13.95526	31.08368
F_7	<i>BurrXII</i> (1,6,2/3)	0.25	-1.5	3.143674	3.735673	5.575274
F_8	<i>GPD</i> (0.25,0.2)	0.2	-0.2	1.889858	2.35675	3.72634
F_9	<i>GPD</i> (0.25,0.4)	0.4	-0.4	3.318483	4.578458	9.280583
F_{10}	$ t_2 $	0.5	-1	13.92911	19.84969	44.65425

Figure 3.2 illustrates how the distributions in this class are distributed with regards to their values of γ and ρ .

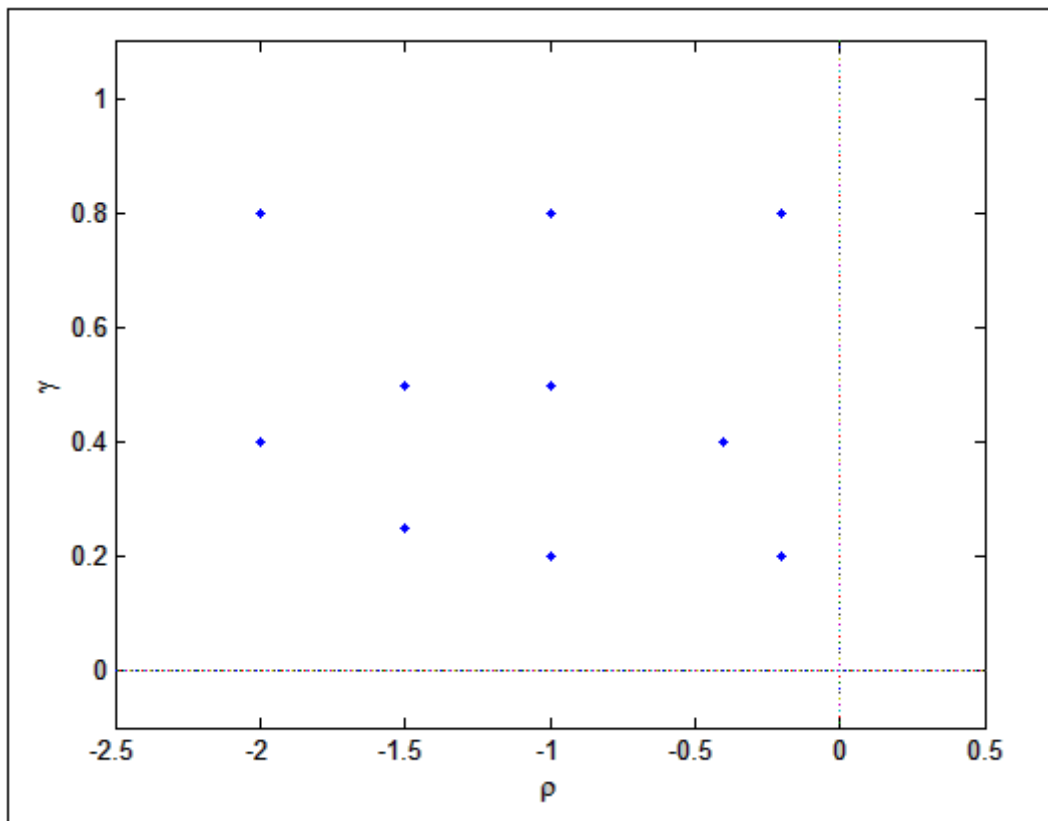


Figure 3.2: Distributions used to derive the multiplying factors

For any F_r from this class, its corresponding multiplier can be calculated at the choices of p_0 and p_1 considered. These are then taken as the true multiplier values $C_{r,p_{0,a},p_{1,b}}^{(true)}$, as given in Table 3.2. In that table, $p_{0,a}$ can take on the values 0.99, 0.995 and 0.999 for $a = 1, a = 2$

and $a = 3$ respectively, and $p_{1,b}$ can take on the values 0.95, 0.9, 0.85 and 0.8 for $b = 1$, $b = 2$, $b = 3$ and $b = 4$ respectively.

Table 3.2: Initial multipliers used during the development simulation study

Distribution	High quantile	Low quantile	Multiplier
F_1	$q_{p_{0,a},1}$	$q_{p_{1,b},1}$	$C_{1,p_{0,a},p_{1,b}}^{(true)}$
F_2	$q_{p_{0,a},2}$	$q_{p_{1,b},2}$	$C_{2,p_{0,a},p_{1,b}}^{(true)}$
.	.	.	.
F_{10}	$q_{p_{0,a},10}$	$q_{p_{1,b},10}$	$C_{10,p_{0,a},p_{1,b}}^{(true)}$

The following algorithm was implemented to obtain the ‘best’ multiplier over the class of distributions $\mathcal{F}$ at the different sample sizes n_i that can take on the values 20, 50, 100, 150 and 200 for $i = 1, i = 2, i = 3, i = 4$ and $i = 5$ respectively.

Algorithm 3.1: Do the following for each $(p_{0,a}, p_{1,b}, n_i)$ combination:

1. Generate a sample of size n_i from F_r where $r = 1$.
2. Calculate the high quantile $q_{p_{0,a},r}$.
3. Estimate the low quantile $q_{p_{1,b},r}$ using EVT as $\hat{q}_{k,p_{1,b},r}^{(W)}$ from equation (2.4) with $k = 0.2 \times n_i$.
4. For $s = 1, 2, \dots, 10$, let $\hat{q}_{p_{0,a},r}^{(M_s)} = C_{s,p_{0,a},p_{1,b}}^{(true)} \hat{q}_{k,p_{1,b},r}^{(W)}$ denote the multiplier estimate of the high quantile $q_{p_{0,a},r}$, when using the multiplier from F_s .
5. Calculate the squared error as $SE_{r,s} = \left(\hat{q}_{p_{0,a},r}^{(M_s)} - q_{p_{0,a},r} \right)^2$ for each $s = 1, 2, \dots, 10$.
6. Repeat steps 1-5 for each $r = 2, \dots, 10$.
7. Repeat steps 1-6 $N = 500$ times and calculate the mean squared error (MSE) as

$$MSE_{r,s} = \frac{1}{N} \sum_{t=1}^N SE_{r,s,t}.$$

8. For each combination of $(p_{0,a}, p_{1,b}, n_i)$, calculate $C_{p_{0,a},p_{1,b},n_i}^{(MM)}$ such that $\min_s \max_r MSE_{r,s}$ is attained.
9. End.

Remark: The $\min_s \text{ave}_r MSE_{r,s}$ argument was also considered in Step 8 where $C_{p_{0,a},p_{1,b},n_i}^{(MA)} = \min_s \text{ave}_r MSE_{r,s}$. See below for discussion of the choice between *MinMax* and *MinAve*.

Table 3.3 is a summary of the different multipliers obtained during the development of the multipliers using both *MinMax* and *MinAve* as criteria for the ‘best’ choice. Since the EVT quantile estimator of quantiles proposed by Gomes et al. (2013) are readily available in the software package ‘evt0’ in *R*, it was decided to use this estimator to estimate the lower quantiles $\hat{q}_{k,p_1,b,r}^{(W)}$ when developing the multipliers.

Table 3.3: Multipliers obtained during the development simulation study

p_0	n	$C_{p_{0,a},0.95,n}^{(MM)}$	$C_{p_{0,a},0.95,n}^{(MA)}$	$C_{p_{0,a},0.9,n}^{(MM)}$	$C_{p_{0,a},0.9,n}^{(MA)}$	$C_{p_{0,a},0.85,n}^{(MM)}$	$C_{p_{0,a},0.85,n}^{(MA)}$	$C_{p_{0,a},0.8,n}^{(MM)}$	$C_{p_{0,a},0.8,n}^{(MA)}$
0.99	20	2.294	1.905	3.694	3.694	8.807	5.024	11.166	6.566
0.99	50	2.385	2.385	6.335	3.694	8.807	5.024	11.166	6.566
0.99	100	3.627	3.627	6.335	6.335	8.807	5.024	11.166	6.566
0.99	150	3.627	3.627	6.335	6.335	8.807	5.024	11.166	6.566
0.99	200	3.627	3.627	6.335	6.335	8.807	5.024	11.166	6.566
0.995	20	2.513	2.513	11.030	4.845	15.334	7.160	19.442	9.357
0.995	50	3.399	3.399	11.030	11.030	15.334	15.334	19.442	9.357
0.995	100	6.316	6.316	11.030	11.030	15.334	15.334	19.442	9.357
0.995	150	6.316	6.316	11.030	11.030	15.334	15.334	19.442	9.357
0.995	200	6.316	6.316	11.030	11.030	15.334	15.334	19.442	9.357
0.999	20	6.416	6.416	43.277	11.841	55.568	55.568	70.456	21.050
0.999	50	22.888	22.888	39.971	39.971	55.568	55.568	70.456	21.050
0.999	100	22.888	22.888	39.971	39.971	55.568	55.568	70.456	70.456
0.999	150	22.888	22.888	39.971	39.971	55.568	55.568	70.456	70.456
0.999	200	22.888	22.888	39.971	39.971	55.568	55.568	70.456	70.456

In the applications that the method was developed for, underestimation is more negative than overestimation. For this reason it was decided to use the multipliers based on the *MinMax* criterion $C_{p_{0,a},p_{1,b},n_i}^{(MM)}$ since these gave the larger values. Also, these multipliers are referred to as the empirical multipliers, and denoted by $C_{p_{0,a},p_{1,b},n_i}^{(EM)}$, for the remainder of this thesis.

Determining the relative performance of these multipliers was done by comparing their ability to estimate the high quantiles using out-of-sample data. This was done during a second simulation study which is discussed in Chapter 4.

3.1.4. Some asymptotic results of the multiplying factors

Suppose the true underlying distribution is F_r . Then the quantile of interest is $q_{p_{0,a},r}$, given by

$$q_{p_{0,a},r} = C_{r,p_{0,a},p_{1,b}}^{(true)} q_{p_{1,b},r}.$$

The error of estimation is

$$error_r = \hat{q}_{p_{0,a},r} - q_{p_{0,a},r} = C_{p_{0,a},p_{1,b},n_i}^{(MM)} \hat{q}_{p_{1,b},r} - C_{r,p_{0,a},p_{1,b}}^{(true)} q_{p_{1,b},r}.$$

Since $\hat{q}_{p_{1,b},r}$ is typically consistent, i.e.

$$\hat{q}_{p_{1,b},r} \xrightarrow{P} q_{p_{1,b},r},$$

as $n \rightarrow \infty$, the error of estimation can be approximated for large n by

$$error_r \approx q_{p_{1,b},r} \left(C_{p_{0,a},p_{1,b},n_i}^{(MM)} - C_{r,p_{0,a},p_{1,b}}^{(true)} \right).$$

Returning to the approach based on Table 3.3, an ‘asymptotic’ choice of the best multiplier can be carried out as follows. The squared error of using distribution F_s when F_r is the true distribution, is given by

$$error_{r,s,\infty}^2 \approx \left[q_{p_{1,b},r} \left(C_{s,p_{0,a},p_{1,b}}^{(true)} - C_{r,p_{0,a},p_{1,b}}^{(true)} \right) \right]^2 = (q_{p_{1,b},r})^2 \left(C_{s,p_{0,a},p_{1,b}}^{(true)} - C_{r,p_{0,a},p_{1,b}}^{(true)} \right)^2.$$

Applying the *MinMax* criterion to these, i.e. $\min_s \max_r error_{r,s,\infty}^2$, gives the asymptotic multipliers, denoted by $C_{p_{0,a},p_{1,b},\infty}^{(AM)}$, which make up the last column of Table 3.4. The other columns contain the finite sample multiplier values $C_{p_{0,a},p_{1,b},n_i}^{(MM)}$ presented in Table 3.3. Note that the finite sample values are positively correlated with the asymptotic values.

Table 3.4: Asymptotic results of the MinMax multipliers

		n					
p_0	p_1	20	50	100	150	200	∞
0.99	0.95	2.294	2.385	3.627	3.627	3.627	3.627
	0.9	3.694	6.335	6.335	6.335	6.335	6.335
	0.85	8.807	8.807	8.807	8.807	8.807	7.856
	0.8	11.166	11.166	11.166	11.166	11.166	11.166
0.995	0.95	2.513	3.399	6.316	6.316	6.316	6.316
	0.9	11.03	11.03	11.03	11.03	11.03	11.03
	0.85	15.334	15.334	15.334	15.334	15.334	11.284
	0.8	19.442	19.442	19.442	19.442	19.442	19.442
0.999	0.95	6.416	22.888	22.888	22.888	22.888	22.888
	0.9	43.277	39.971	39.971	39.971	39.971	39.971
	0.85	55.568	55.568	55.568	55.568	55.568	55.568
	0.8	70.456	70.456	70.456	70.456	70.456	54.608

Remark: Other asymptotic multipliers were also investigated, such as those based on minimizing the absolute relative error, but these did not produce good results.

3.1.5. Second order multipliers

Second order multipliers were also considered in order to reduce bias. To develop second order multipliers one should decide on which second order quantile estimator to use for estimating the lower quantile.

Since the second order EVT quantile estimator of quantiles proposed by Gomes et al. (2013) are readily available in the software package ‘evt0’ in R, it was decided to use this to estimate the lower quantiles when developing the second order multipliers. The second order multipliers are presented in Table 3.5. These multipliers were also developed using Algorithm 3.1, but the estimator used for the lower quantile $\hat{q}_{k,p_1,b,r}^{(W)}$ is now replaced by $\hat{q}_{k,p_1,b,r}^{(W+)}$.

Table 3.5: Second order multipliers obtained during the development simulation study

p_0	n	$C_{p_0,a,0.95,n}^{(MM+)}$	$C_{p_0,a,0.95,n}^{(MA+)}$	$C_{p_0,a,0.9,n}^{(MM+)}$	$C_{p_0,a,0.9,n}^{(MA+)}$	$C_{p_0,a,0.85,n}^{(MM+)}$	$C_{p_0,a,0.85,n}^{(MA+)}$	$C_{p_0,a,0.8,n}^{(MM+)}$	$C_{p_0,a,0.8,n}^{(MA+)}$
0.99	20	1.905	1.905	6.335	3.694	8.807	5.024	11.166	6.566
0.99	50	3.627	2.385	6.335	6.335	8.807	5.024	11.166	6.566
0.99	100	3.627	3.627	6.335	6.335	8.807	8.807	11.166	6.566
0.99	150	3.627	3.627	6.335	6.335	8.807	8.807	11.166	6.566
0.99	200	3.745	3.627	6.335	6.335	8.807	8.807	11.166	6.566
0.995	20	3.399	3.399	11.030	4.845	15.334	7.160	19.442	9.357
0.995	50	6.316	6.316	11.030	11.030	15.334	15.334	19.442	9.357
0.995	100	6.316	6.316	11.030	11.030	15.334	15.334	19.442	9.357
0.995	150	6.316	6.316	11.030	11.030	15.334	15.334	19.442	9.357
0.995	200	6.548	6.316	11.030	11.030	15.334	15.334	19.442	9.357
0.999	20	7.029	7.029	43.277	39.971	55.568	55.568	70.456	21.050
0.999	50	22.888	22.888	39.971	39.971	55.568	55.568	70.456	21.050
0.999	100	22.888	22.888	39.971	39.971	55.568	55.568	70.456	70.456
0.999	150	23.804	23.804	39.971	39.971	55.568	55.568	70.456	70.456
0.999	200	22.888	22.888	39.971	39.971	55.568	55.568	70.456	70.456

Once again, the multipliers based on the *MinMax* criterion $C_{p_0,a,p_{1,b},n_i}^{(MM+)}$ produced larger values compared to the *MinAve* criterion. These multipliers will be referred to as the second order empirical multipliers, denoted by $C_{p_0,a,p_{1,b},n_i}^{(EM+)}$, for the remainder of this thesis. The performance of both the first and second order multipliers (empirical and asymptotic) was tested using simulation which is discussed in Chapter 4.

Remark: Other bias-reducing multipliers were also investigated where the empirical multipliers were adjusted with the resulting bias, but these did not produce improved results.

The quantile estimators discussed in this section are similar to those of EVT in that these estimators rely on the observations in the tail only. As mentioned in Chapter 2, the full-parametric approach includes fitting a single distribution to the entire dataset. The logarithm of the normal inverse Gaussian distribution can be used for this purpose and is discussed in the following section.

3.2. Quantile estimation by fitting the logarithm of the normal inverse Gaussian distribution

3.2.1. Introduction

It is important, from a risk manager's point of view, that the tail part of the data is well modelled (to calculate the VaR) when fitting a distribution to data. Modelling the body part of the data is also important when, for example the expected loss needs to be estimated. Recall from Chapter 1 that the expected loss can be estimated using the expected value $E(X)$ of the underlying distribution. Estimating $E(X)$ accurately can only be done when the underlying distribution is modelled accurately. As mentioned in Chapter 1, fitting various distributions effectively to data, especially small datasets, can be difficult. It is therefore important to investigate a wide range of distributions to determine which may possibly pose as viable options.

Various distributions that could possibly be used to model losses in an OR environment were recently discussed in the literature. Ahn et al. (2012) discusses the class of Log phase-type (LogPh) distributions for this purpose. The g-and-h class of distributions is also discussed by Degen et al. (2007).

Another distribution that received much attention recently is the normal inverse Gaussian (NIG) distribution. The NIG distribution was introduced by Barndorff-Nielsen (1997) and proved to be a promising candidate for modelling skew data with moderately heavy tails. Extensive work regarding the NIG distribution was done by e.g. Venter & De Jongh (2002), Lillestøl (2000), Godin et al. (2009) and Karlis (2002). The multivariate NIG distribution was also discussed by Chang & Hung (2010).

Motivated by the NIG's flexibility, having four parameters, and its popularity in the financial literature as well as Ahn et al. (2012) recently proposing the LogPh distribution, led to the proposal of the LogNIG (LNIG) distribution in this study. Before introducing the LNIG the NIG distribution is discussed in the following subsection.

3.2.2. The normal inverse Gaussian distribution

The normal inverse Gaussian (NIG) distribution is a mixture of the normal and inverse Gaussian distributions. Furthermore, the NIG distribution is a member of the generalized hyperbolic class of distributions. It is usually defined by using the Barndorff-Nielsen parameterization. This is denoted by $NIG(\alpha, \beta, \mu, \delta)$ where α is known as the steepness parameter, β represents the asymmetry of the distribution and μ and δ is the location and scale parameters respectively. The parameter space can be summarized as $0 < |\beta| < \alpha$, $\mu \in \mathbb{R}$, $\delta > 0$. The Cauchy distribution is obtained when $\alpha \rightarrow 0$ whereas the Normal distribution is obtained when $\alpha \rightarrow \infty$ and $\delta \rightarrow \infty$ so that $\frac{\delta}{\alpha} \rightarrow \sigma^2$. The case where $\beta = 0$ corresponds to a symmetric distribution.

Let $\xi = \sqrt{\alpha^2 - \beta^2}$ and $(z) = \sqrt{1 + z^2}$, then the NIG density function can be defined as

$$g(x; \alpha, \beta, \mu, \delta) = \frac{\alpha \exp(\xi \delta + \beta(x - \mu)) K_1\left(\alpha \delta f\left(\frac{x - \mu}{\delta}\right)\right)}{\pi f\left(\frac{x - \mu}{\delta}\right)}$$

where K_1 is the modified Bessel function (MBF) of the third order and index one (also known as the MacDonald function). Abramowitz & Stegun (1972) discusses the MBF in more detail. The reader is referred to Appendix B for polynomial approximations of the MBF extracted from Abramowitz & Stegun (1972).

The NIG distribution's moment-generating function can be derived as

$$M_x(u) = \exp\left(u\mu + \delta\left(\xi - \sqrt{\alpha^2 - (\beta + u)^2}\right)\right).$$

This function can be used to derive the following moments:

- $E[X] = \kappa_1 = \mu + \delta \frac{\beta}{\xi}$
- $Var[X] = \kappa_2 = \delta \frac{\alpha^2}{\xi^3}$
- $Skewness = 3 \frac{\beta}{\alpha(\delta\xi)^{\frac{1}{2}}}$
- $Kurtosis = 3 \left(1 + 4 \left(\frac{\beta}{\alpha}\right)^2\right) \frac{1}{\delta\xi}$

The NIG distribution also has the following attractive properties:

- If $X \sim NIG(\alpha, \beta, \mu, \delta)$ then $Y = kX \sim NIG(k^{-1}\alpha, k^{-1}\beta, k\mu, k\delta)$.
- If $X_1 \sim NIG(\alpha, \beta, \mu_1, \delta_1)$ and $X_2 \sim NIG(\alpha, \beta, \mu_2, \delta_2)$ are independent, then the sum $= X_1 + X_2 \sim NIG(\alpha, \beta, \mu_1 + \mu_2, \delta_1 + \delta_2)$.

3.2.2.1. Inference on the NIG distribution

Maximum likelihood estimation can be used to estimate the parameters of the NIG distribution. Other estimation procedures include the following:

- The 'hyp' program developed by Blaesild & Sørensen (1992),
- A similar program developed by Eberlein et al. (1998),
- Bayesian methods using the EM-algorithm as shown by Karlis (2002), and
- Bayesian methods using Markov chain Monte Carlo methods as shown by Lillestøl (2001).

A suggested method by Venter & De Jongh (2002) involves the approximate expression for the tail of the NIG density given as

$$g^*(x; \alpha, \beta, \mu, \delta): \begin{cases} \sim C|x|^{-\frac{3}{2}}e^{bx} & x \rightarrow -\infty \\ \sim C|x|^{-\frac{3}{2}}e^{-ax} & x \rightarrow \infty \end{cases}$$

where $a = \alpha - \beta$ and $b = \alpha + \beta$. The above mentioned expression was derived from

$$g^*(x; \alpha, \beta, \mu, \delta) \sim C|x|^{-\frac{3}{2}}\exp(-\alpha|x| + \beta x) \text{ as } |x| \rightarrow \infty$$

by introducing a and b . This follows from the well-known asymptotic relation

$$K_1(z) \sim \sqrt{\frac{\pi}{2}} z^{-\frac{1}{2}} \exp(-z) \text{ as } z \rightarrow \infty.$$

The following expressions were derived which can be used as approximations for a and b respectively:

$$a \sim \frac{1}{2} \frac{Q + E[X|X > Q]}{E[X^2|X > Q] - QE[X|X > Q]},$$

with Q an estimate of the p^{th} -quantile and

$$b \sim -\frac{1}{2} \frac{Q' + E[X|X < Q']}{E[X^2|X < Q'] - Q'E[X|X < Q']},$$

with Q' an estimate of the $(1 - p)^{th}$ -quantile.

Consider a dataset $X_1, X_2, \dots, X_n$ and a large probability, say p . Denote the order statistics by $X_{1,n} \leq X_{2,n} \leq \dots \leq X_{n,n}$. Q can then be estimated as $Q = X_{[np],n}$, whereas Q' can be estimated as $Q' = X_{[n(1-p)],n}$. Here, $[k]$ is the smallest integer in k . Other quantile estimation procedures include those arising from EVT and the multipliers. $E[X|X > Q]$ can be estimated as the mean of all the order statistics larger than Q , whereas $E[X^2|X > Q]$ is the mean of the squares of all the order statistics larger than Q . The same procedure can be applied when estimating $E[X|X < Q']$ and $E[X^2|X < Q']$, but now consider the order statistics smaller than Q' . This procedure was referred to as Tail-Based-Estimation (TBE) by Lillestöl (2002). The pair (a, b) may be used instead of (α, β) as shape parameters. Venter & De Jongh (2002) suggest using $(a, b, \kappa_1, \kappa_2)$ as starting values when optimizing the parameter estimation procedure.

The TBE proposed by Venter & De Jongh (2002) is illustrated next by using a numerical example. A sample of $n = 100$ variables (denoted by $X_1, X_2, \dots, X_n$) were randomly generated from a $NIG(\alpha, \beta, \mu, \delta)$ distribution with $\alpha = 2$, $\beta = 1$, $\mu = 1$ and $\delta = 1$. The density of the $NIG(\alpha, \beta, \mu, \delta)$ distribution (denoted by $nig(\alpha, \beta, \mu, \delta)$) is illustrated graphically in Figure 3.3 for this case. The NIG distribution was fitted to $X_1, X_2, \dots, X_n$ using a procedure in SAS called 'Proc NLP'. This procedure produced parameter estimates $(\hat{\alpha}, \hat{\beta}, \hat{\mu}, \hat{\delta})$. The fitted density $nig(\hat{\alpha}, \hat{\beta}, \hat{\mu}, \hat{\delta})$ is illustrated graphically in Figure 3.4.

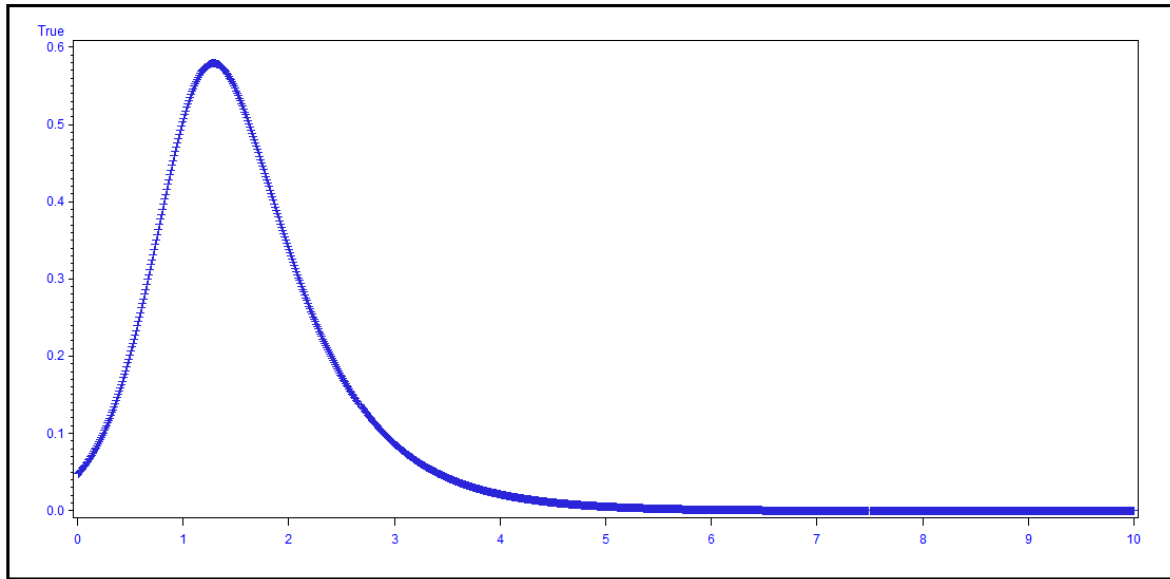


Figure 3.3: Graphical illustration of the NIG density (using its true parameters)

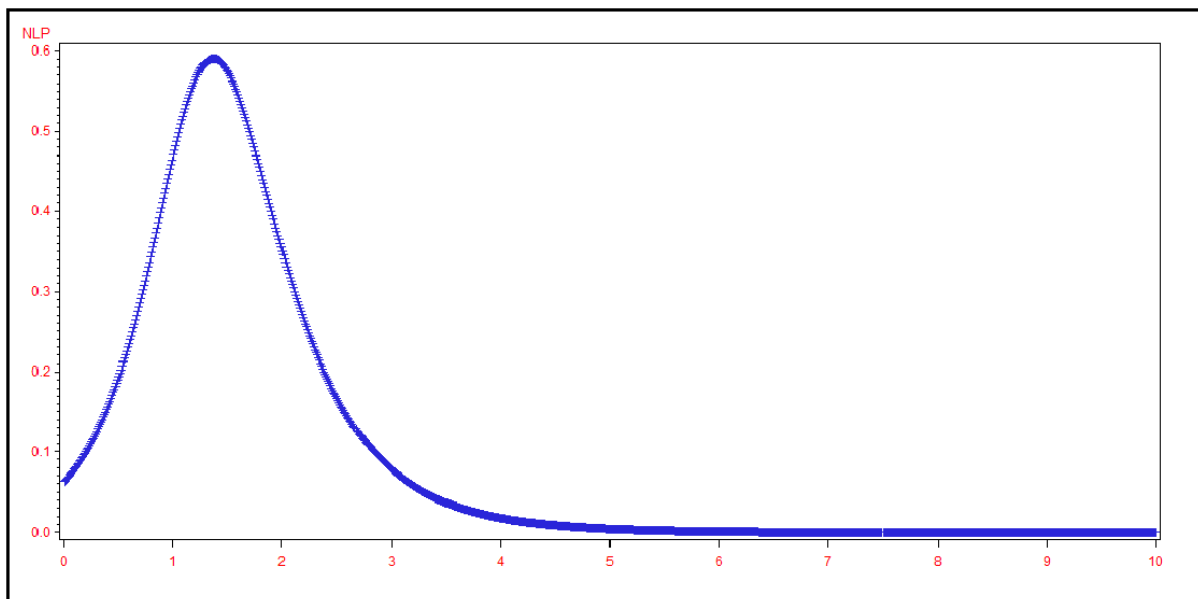


Figure 3.4: Graphical illustration of the NIG density (using its estimated parameters)

Combining the plots of the density functions in Figure 3.3 and Figure 3.4 will graphically illustrate how well the estimated distribution fits the true distribution. This is illustrated in Figure 3.5. Here, the true NIG density is represented by the blue curve whereas the red curve represents the fitted density. Also, the scales of both the true and fitted density values are on the left-hand and right-hand sides respectively. It is clear from Figure 3.5 that the estimated (fitted) density fits the true underlying density very well, given the small sample size used.

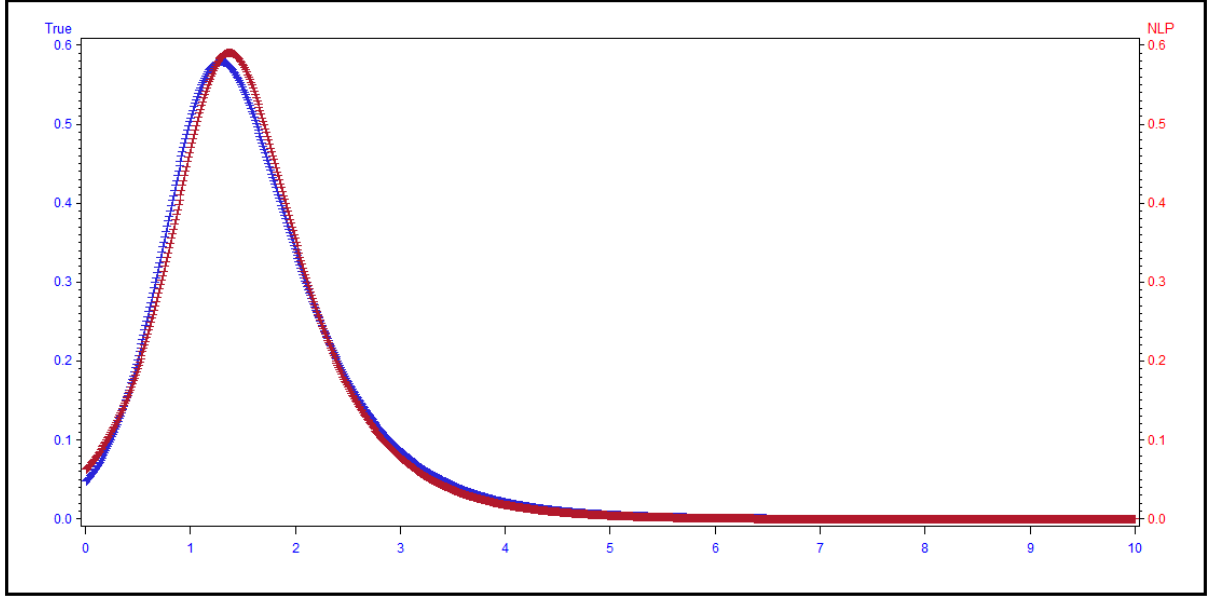


Figure 3.5: Graphical illustration of the true NIG density vs. estimated NIG density

Barndorff-Nielsen et al. (1985) introduced other shape parameters ψ and χ for the steepness and asymmetry parameters respectively defined as

$$\psi = \left(1 + \delta\sqrt{\alpha^2 - \beta^2}\right)^{-\frac{1}{2}} = (1 + \delta\xi)^{-\frac{1}{2}} \text{ and } \chi = \psi \frac{\beta}{\alpha}.$$

The pair (ψ, χ) may also be used instead of (α, β) as shape parameters.

The focus of this study remains on the estimation of extreme quantiles. This motivates the use of the TBE of the NIG distribution's parameters as proposed by Venter & De Jongh (2002). Eventually, the p^{th} -quantile can be estimated using the NIG distribution by inverting its distribution function and solving

$$G(x; \hat{\alpha}, \hat{\beta}, \hat{\mu}, \hat{\delta}) = p$$

leading to the quantile estimate

$$\hat{q}_p^{(NIG)} = G^{-1}(p; \hat{\alpha}, \hat{\beta}, \hat{\mu}, \hat{\delta}).$$

Although the NIG distribution possesses attractive modelling/statistical properties, it can be argued that the NIG distribution can often not be used in an OR environment due to the heaviness of OR loss data. A natural way to obtain heavier tails from a distribution is by

using its logarithm version. A well-known example of this is the LogNormal distribution obtained from the Normal distribution (the LogNormal distribution is a ‘heavier-tailed version’ of the Normal distribution). This conversion between the Normal distribution and its logarithm counterpart can also be applied to any other monotone increasing distribution function. One such distribution is the NIG distribution with the logarithm of the NIG (denoted by LNIG) distribution as its heavier tailed version. The LNIG distribution is discussed in the following subsection.

Remark: For clarity note that the term “log version” is a misnomer since it is actually an “exponential version”, taking the exponential transformation of the original random variable.

3.2.3. The logarithm of the normal inverse Gaussian distribution

Based on the above discussion a LNIG random variable is defined as $Y = e^X$, with X a NIG random variable. In this section a number of properties of the LNIG distribution is discussed. Before these properties are given, it is briefly indicated how LNIG based quantiles can be estimated.

Note that if $Y = e^X$ then the corresponding quantiles have the same relationship, i.e. $q_p^{(Y)} = e^{q_p^{(X)}}$. Also, if $Y = e^X$, then $X = \log(Y)$. Thus, given a sample $Y_1, Y_2, \dots, Y_n$ from a LNIG distribution, $X_1 = \log(Y_1), X_2 = \log(Y_2), \dots, X_n = \log(Y_n)$ will be a sample from a NIG distribution. Estimating the NIG quantiles by $\hat{q}_p^{(X)}$, the corresponding LNIG quantiles are estimated by $\hat{q}_p^{(Y)} = e^{\hat{q}_p^{(X)}}$. The resulting quantile estimator in Section 2.1.2, $\hat{q}_p^{(PAR)}$, will be used to estimate $\hat{q}_p^{(X)}$.

In the next subsection some mathematical properties of the LNIG distribution are discussed. The numerical properties of the LNIG distribution are discussed in Chapter 4.

3.2.3.1. Some properties of the LNIG distribution

Let $X \sim NIG(\alpha, \beta, \mu, \delta)$ with distribution and density functions denoted by F_X and f_X respectively. Recall the NIG density function (with slight change in notation) as

$$f_X(x) = \frac{\alpha \exp(\xi + \beta(x - \mu)) K_1\left(\alpha \delta q\left(\frac{x - \mu}{\delta}\right)\right)}{\pi q\left(\frac{x - \mu}{\delta}\right)}$$

with $\xi = \delta\sqrt{\alpha^2 - \beta^2}$ and $q(z) = \sqrt{1 + z^2}$. It is well known that the Bessel function can be approximated as $K_1(z) \approx \sqrt{\frac{\pi}{2}} z^{-\frac{1}{2}} \exp(-z)$ for large z . The NIG density can therefore be approximated by

$$f_X(x) \approx \frac{\alpha \exp(\xi + \beta(x - \mu)) \sqrt{\frac{\pi}{2}} \left(\alpha \delta q\left(\frac{x - \mu}{\delta}\right)\right)^{-\frac{1}{2}} \exp\left(-\alpha \delta q\left(\frac{x - \mu}{\delta}\right)\right)}{\pi q\left(\frac{x - \mu}{\delta}\right)} \text{ for large } x.$$

Simplifying the approximation of $f_X(x)$ above produces the following approximation of the NIG density function:

$$f_X(x) \approx C_{NIG} x^{-\frac{3}{2}} \exp[-x(\alpha - \beta)] \text{ for large } x, \quad (3.1)$$

with C_{NIG} some constant.

From this it is clear that the NIG distribution has semi-heavy tails, having a power function component as well as an exponential component.

It follows directly from $Y = e^X$, that

$$F_Y(y) = F_X(\log(y))$$

and

$$f_Y(y) = y^{-1} f_X(\log(y)).$$

Note that $y \rightarrow \infty \Rightarrow \log(y) \rightarrow \infty$. From (3.1) it then follows that:

$$\begin{aligned} f_Y(y) &= y^{-1} f_X(\log(y)) \approx y^{-1} C_{NIG}(\log(y))^{-\frac{3}{2}} \exp[-\log(y)(\alpha - \beta)], \text{ for large } y, \\ &= C_{NIG} y^{-(1+\alpha-\beta)} (\log(y))^{-\frac{3}{2}}. \end{aligned}$$

This last expression can be written as

$$f_Y(y) = y^{-(1+\alpha-\beta)} \ell_Y(y), \text{ as } y \rightarrow \infty,$$

with $\ell_Y(y)$ a slowly varying function.

Using Karamata's Theorem, the survival function of F_Y can then be written as

$$\bar{F}_Y(y) = \frac{1}{\alpha-\beta} y^{-(\alpha-\beta)} \ell_Y(y), \text{ } y \rightarrow \infty,$$

(see e.g. de Haan & Ferreira, 2006).

Thus, $\bar{F}_Y$ is regular varying of order $-(\alpha - \beta)$, written as $\bar{F}_Y \in \mathcal{R}_{-(\alpha-\beta)}$. This implies that the LNIG distribution falls within the Frechét-Pareto class of heavy-tailed distributions with EVI $\gamma = \frac{1}{\alpha-\beta}$. Furthermore, it can also be shown that the LNIG distribution satisfies the second order condition mentioned in the previous chapter with second order parameter $\rho = 0$.

Also, its MEF is approximately given by $e(v) \approx \frac{v}{\alpha-\beta-1}$, $v \rightarrow \infty$. Therefore the MEF is approximately linear which is similar to that of the GPD.

The following summarizes the useful properties of the LNIG distribution discussed in this section:

- The EVI of the LNIG distribution is $\gamma = \frac{1}{\alpha-\beta}$,
- The second order parameter $\rho = 0$, and
- The mean excess function $e(v) \approx \frac{v}{\alpha-\beta-1}$.

The LNIG distribution is a 'rich' family of distributions, having four parameters and possessing attractive properties (from an EVT point of view) when considering its use to

model the severity distribution. This distributional family enables one to model heavy-tailed data using the entire sample as opposed to using only extreme values.

A simulation study was conducted to determine the accuracy of the quantile estimation techniques discussed in this chapter. These techniques are compared to one another, as well as to the most widely used techniques discussed in Chapter 3. The simulation design and simulation results, together with recommendations on the application of the various estimation techniques are given in the following chapter.

Chapter 4: A (numerical) comparison of quantile estimation techniques for the severity distribution

It was mentioned in Chapter 1 that there are two main approaches to estimating the severity distribution's extreme quantiles, i.e. the full-parametric approach and the approach relating to the use of high quantiles only. The full-parametric approach consists of modelling the underlying distribution by fitting a statistical distribution to the entire sample of data. The latter of the two approaches consist of modelling the underlying distribution by fitting a statistical distribution to the tail-part of the sample only. It is for this reason that the quantile estimation techniques discussed in both Chapters 2 and 3 will be compared separately based on which one of the two main approaches they fall under. The various quantile estimation techniques are compared using results obtained from both simulation and empirical studies.

The first simulation study consists of a comparison of those quantile estimation techniques that are based on the tail-part only. The following quantile estimation techniques were considered for this simulation study which is discussed in Section 4.1:

- The first order EVT quantile estimator (Equation 2.4) which was discussed in Section 2.2.1 (referred to as W),
- The first order empirical multiplier quantile estimator (Table 3.3) derived in Section 3.1.3 using first order EVT estimator for the lower quantile (referred to as EM),
- The first order asymptotic multiplier quantile estimator (Table 3.4) derived in Section 3.1.4 using first order EVT estimator for the lower quantile (referred to as AM),
- The second order EVT quantile estimator (Equation 2.6) which was discussed in Section 2.2.2 where a second order assumption was made (referred to as W+),
- The second order empirical multiplier quantile estimator (Table 3.5) derived in Section 3.1.5 using the second order EVT estimator for the lower quantile (referred to as EM+), and
- The second order asymptotic multiplier quantile estimator (Table 3.4) derived in Section 3.1.4 using the second order EVT estimator for the lower quantile (referred to as AM+).

The second simulation study involves comparing those quantile estimation techniques that are based on the full-parametric approach. The following distributions were considered for this simulation study which is discussed in Section 4.2:

- Some popular distributions used in operational risk modelling such as the Burr, LogNormal and LogGamma, and

- The newly proposed LogNIG distribution discussed in Section 3.2.

Lastly, the most promising quantile estimation techniques are compared using an empirical study. The empirical study is discussed in Section 4.3.

It was mentioned that the estimation methodologies discussed in Section 2 are currently considered best practice when estimating extreme quantiles of the severity distribution. Therefore, these estimation techniques will be considered as the benchmark when assessing the newly proposed estimation techniques' (Section 3) ability to estimate the desired quantiles.

4.1. A simulation study used to compare the quantile estimation techniques which rely on high quantiles only (The EVT simulation study)

4.1.1. Simulation design

The multiplier method's ability to estimate high quantiles can be compared to that of the EVT techniques. These techniques were tested on data generated from eight distributions (denoted by $G_1, G_2, \dots, G_8$) similar to those used to derive the various multipliers in Section 3.1. The distributions used were subjectively divided into four different groups by considering the values of both γ (first order parameter) and ρ (second order parameter). These four groups are denoted by $\underline{G}^{(1)}$, $\underline{G}^{(2)}$, $\underline{G}^{(3)}$ and $\underline{G}^{(4)}$ respectively with $(G_1, G_2) \in \underline{G}^{(1)}$, $(G_3, G_4) \in \underline{G}^{(2)}$, $(G_5, G_6) \in \underline{G}^{(3)}$ and $(G_7, G_8) \in \underline{G}^{(4)}$. Table 4.1 illustrates the distributions that make up $(\underline{G}^{(1)}, \underline{G}^{(2)}, \underline{G}^{(3)}, \underline{G}^{(4)}) \in \underline{G}$ and Figure 4.1 illustrates how these distributions are distributed with regards to their values of γ and ρ .

Table 4.1: A list of the distributions used for the extreme value theory simulation study

Reference G_v	Group	Distribution	γ	ρ	$q_{0,1,v}$	$q_{0,2,v}$	$q_{0,3,v}$
G_1	$\underline{G}^{(1)}$	<i>BurrXII</i> (1,1.5,1)	0.67	-1	21.400	34.085	99.933
G_2	$\underline{G}^{(1)}$	<i>BurrXII</i> (1,2,0.5)	1	-2	99.995	199.997	999.999
G_3	$\underline{G}^{(2)}$	<i>BurrXII</i> (1,0.375, 10/3)	0.8	-0.3	18.407	37.718	175.458
G_4	$\underline{G}^{(2)}$	<i>BurrXII</i> (1,0.4,5)	0.5	-0.2	2.811	4.881	15.344
G_5	$\underline{G}^{(3)}$	<i>BurrXII</i> (1,1.6,2.5)	0.25	-0.4	2.839	3.472	5.399
G_6	$\underline{G}^{(3)}$	<i>BurrXII</i> (1,2.67,1.25)	0.3	-0.8	3.937	4.865	7.911
G_7	$\underline{G}^{(4)}$	<i>BurrXII</i> (1,3.5,0.5)	0.57	-2	13.895	20.648	51.795
G_8	$\underline{G}^{(4)}$	<i>BurrXII</i> (1,8,0.5)	0.25	-2	3.162	3.761	5.623

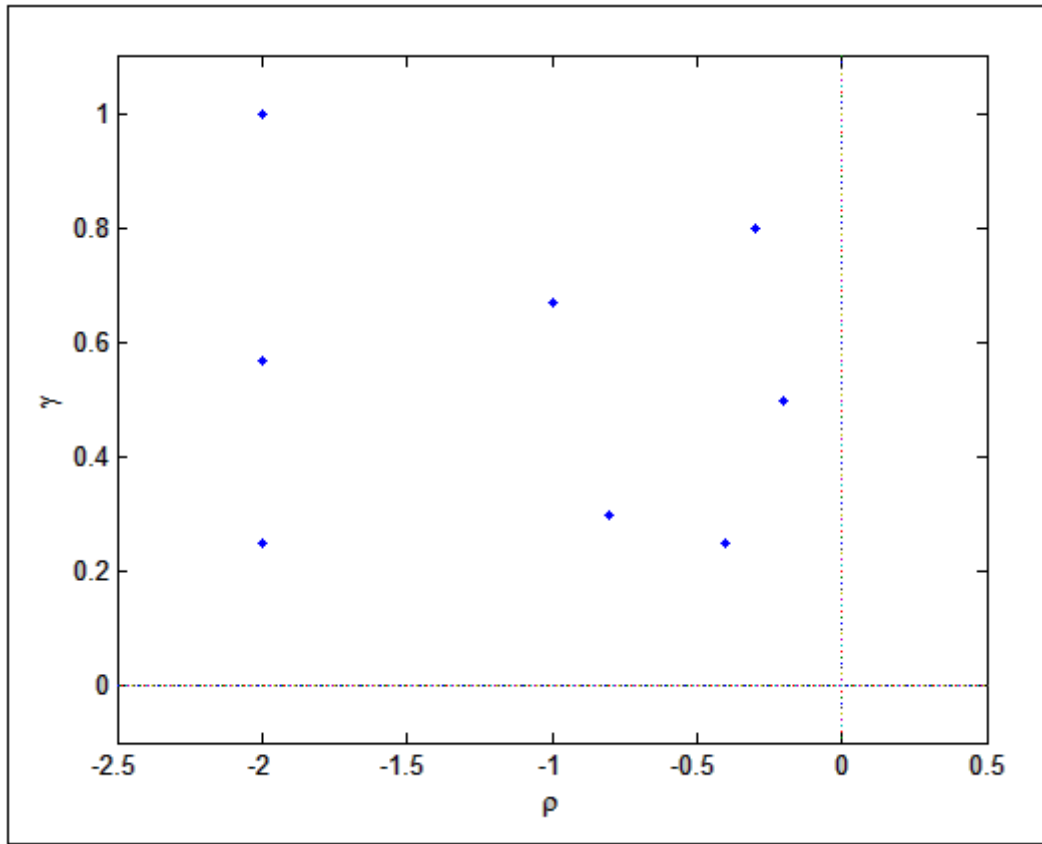


Figure 4.1: A scatterplot of the distributions used for the EVT simulation study

The probability levels of the quantiles of interest for this simulation are:

- $p_{0,1} = 0.999$,
- $p_{0,2} = 0.995$, and
- $p_{0,3} = 0.99$.

The sample sizes that were considered are:

- $n_1 = 20$,
- $n_2 = 50$,
- $n_3 = 100$,
- $n_4 = 150$, and
- $n_5 = 200$.

The following algorithm (4.1) was used for the simulation.

Algorithm 4.1:

1. Let $a, i, v = 1$ and generate a sample of size n_i from G_v (the index a will be used in the next step).
2. Calculate the high quantile $q_{p_0,a,v} = G_v^{-1}(p_{0,a})$ and denote it by $q_{p_0,v}$ (the subscript a was dropped).
3. Estimate the low quantile $q_{p_1,b,v}$ for each $b = 1,2,3,4$ as follows and denote it by $q_{p_1,v}$ (the subscript b was dropped):
 - 3.1. First order EVT as $\hat{q}_{p_1,v}^{(W)}$ (see Equation 2.4), and
 - 3.2. Second order EVT as $\hat{q}_{p_1,v}^{(W+)}$ (see Equation 2.6).
4. Estimate the high quantile $q_{p_0,v}$ (as in Step 2) using:
 - 4.1. First order:
 - 4.1.1. EVT as $\hat{q}_{p_0,v}^{(W)}$,
 - 4.1.2. Empirical multipliers as $\hat{q}_{p_0,p_1,n,v}^{(EM)} = C_{p_0,p_1,n}^{(EM)} \hat{q}_{p_1,v}^{(W)}$ with $C_{p_0,p_1,n}^{(EM)}$ as in Table 3.3 and $\hat{q}_{p_1,v}^{(W)}$ calculated in Step 3.1,
 - 4.1.3. Asymptotic multipliers as $\hat{q}_{p_0,p_1,v}^{(AM)} = C_{p_0,p_1,\infty}^{(AM)} \hat{q}_{p_1,v}^{(W)}$ with $C_{p_0,p_1,\infty}^{(AM)}$ as in Table 3.4 and $\hat{q}_{p_1,v}^{(W)}$ calculated in Step 3.1,
 - 4.2. Second order:
 - 4.2.1. EVT as $\hat{q}_{p_0,v}^{(W+)}$,
 - 4.2.2. Empirical multipliers as $\hat{q}_{p_0,p_1,n,v}^{(EM+)} = C_{p_0,p_1,n}^{(EM+)} \hat{q}_{p_1,v}^{(W+)}$ with $C_{p_0,p_1,n}^{(EM+)}$ as in Table 3.5 and $\hat{q}_{p_1,v}^{(W+)}$ calculated in Step 3.2,
 - 4.2.3. Asymptotic multipliers as $\hat{q}_{p_0,p_1,v}^{(AM+)} = C_{p_0,p_1,\infty}^{(AM+)} \hat{q}_{p_1,v}^{(W+)}$ with $C_{p_0,p_1,\infty}^{(AM+)}$ as in Table 3.4 and $\hat{q}_{p_1,v}^{(W+)}$ calculated in Step 3.2,
5. Let $\hat{q}^*$ denote any of the estimates in Step 4 (this will be used in Step 6).
6. Repeat steps 1-5 $N = 10,000$ times to obtain the following:

- 6.1. Mean squared error as $MSE_v^* = \frac{1}{N} \sum_{n=1}^N (\hat{q}_n^* - q_{p_0, v})^2$,
- 6.2. Root mean squared error as $RMSE_v^* = \sqrt{MSE_v^*}$,
- 6.3. Bias using the average of the estimates as $BiasA_v^* = \frac{1}{N} \sum_{n=1}^N \hat{q}_n^* - q_{p_0, v}$, and
- 6.4. Absolute bias as $|BiasA_v^*|$.
7. Repeat steps 1-6 for each combination of (a, i, v) .
8. End.

Remark: Algorithm 4.1 was applied using the statistical software ‘R’. Furthermore, a package ‘evt0’ was used for the calculation of the EVT quantile estimators. A function ‘mop.q’ in the ‘evt0’ package can be used to calculate the EVT quantile estimates. This function uses the mean-of-order- p class of quantile estimators proposed by Gomes et al. (2013). Note that this function uses as input, amongst other, the method used to estimate the EVI which can either be the standard first order mean-of-order- p (MOP) method or the second order reduced bias mean-of-order- p (RBMOP) method. The quantile estimator using the MOP method is equivalent to the first order Weissman quantile estimator (Equation 2.4) if the order p is chosen $p = 0$. Furthermore, the quantile estimator using the RBMOP method is equivalent to the second order Weissman quantile estimator (Equation 2.6) if the order p is chosen $p = 0$. Lastly, the choice of k used in the calculation of the Hill estimate was taken as $k = 0.2 \times n$ for all values of n considered.

The results of the simulation study are presented in the next section.

4.1.2. Simulation results

The results based on first order EVT (Step 4.1 from Algorithm 4.1) only are presented in the next section (4.1.2.1), followed by the results based on second order EVT (Step 4.2 from Algorithm 4.1) in Section 4.1.2.2. A final summary of the results are given in Section 4.1.3 in which conclusions are given on the performance of the various estimators.

Before the results are presented it is worth mentioning that it is difficult to obtain an overall view of the comparative performance of estimators based on the raw values. Use is therefore made of relative values whereby each estimator's performance is taken relative to the best one. This leads to all values being in the interval $(0,1]$. This can be written mathematically as

$$0 < \frac{Measure_{p_0,n,r}}{Measure_{p_0,n,r,*}} \leq 1 \text{ with } Measure_{p_0,n,r} = \min_* Measure_{p_0,n,r,*}$$

where $*$ denotes the quantile estimators. Note that the value 0 should theoretically be excluded, but instead (as will be seen in the following tables) is often encountered due to rounding errors. In addition, the value 0 indicates that the corresponding estimator performs relatively very poor for the given case. Furthermore, conclusions for both the first and second order results are grouped according to the choice of distributional group and sample size.

4.1.2.1. First order results

The results from the first order EVT simulation study are presented in this section. The following three quantile estimators are considered throughout this section:

- $\hat{q}_{p_0,r}^{(W)}$: First order EVT calculated in Step 4.1.1 in Algorithm 4.1,
- $\hat{q}_{p_0,p_1,n,r}^{(EM)} = C_{p_0,p_1,n}^{(EM)} \times \hat{q}_{p_1,r}^{(W)}$: First order multiplier using the **empirical** $C_{p_0,p_1,n}^{(EM)}$ calculated in Step 4.1.2 in Algorithm 4.1, and
- $\hat{q}_{p_0,p_1,r}^{(AM)} = C_{p_0,p_1,\infty}^{(AM)} \times \hat{q}_{p_1,r}^{(W)}$: First order multiplier using the **asymptotic** $C_{p_0,p_1,\infty}^{(AM)}$ calculated in Step 4.1.3 in Algorithm 4.1.

The results are presented separately for $p_0 = 0.999$, $p_0 = 0.995$ and $p_0 = 0.99$, starting with $p_0 = 0.999$ first.

Results of the first order EVT simulation study for $p_0 = 0.999$

Tables 4.2 and 4.3 contain the **RMSE** (Step 6.2 in Algorithm 4.1) and **absolute bias** (Step 6.4 in Algorithm 4.1) results respectively for $p_0 = 0.999$ in the **first order EVT** environment.

Table 4.2: Root mean squared error results of the first order EVT simulation study for $p_0 = 0.999$

p_0	n	r	$\hat{q}_{p_0,r}^{(W)}$	$\hat{q}_{p_0,0.95,n,r}^{(EM)}$	$\hat{q}_{p_0,0.9,n,r}^{(EM)}$	$\hat{q}_{p_0,0.85,n,r}^{(EM)}$	$\hat{q}_{p_0,0.8,n,r}^{(EM)}$	$\hat{q}_{p_0,0.95,r}^{(AM)}$	$\hat{q}_{p_0,0.9,r}^{(AM)}$	$\hat{q}_{p_0,0.85,r}^{(AM)}$	$\hat{q}_{p_0,0.8,r}^{(AM)}$
0.999	20	1	0.00	1.00	0.53	0.67	0.65	0.38	0.61	0.67	0.65
0.999	20	2	0.00	0.73	1.00	0.95	0.93	0.68	0.97	0.95	0.93
0.999	20	3	0.00	0.88	1.00	0.93	0.91	0.63	0.99	0.93	0.91
0.999	20	4	0.00	0.83	0.98	1.00	0.95	0.42	1.00	1.00	0.95
0.999	20	5	0.20	1.00	0.13	0.12	0.10	0.18	0.14	0.12	0.10
0.999	20	6	0.24	1.00	0.12	0.11	0.09	0.17	0.13	0.11	0.09
0.999	20	7	0.00	1.00	0.23	0.24	0.22	0.23	0.25	0.24	0.22
0.999	20	8	0.97	1.00	0.12	0.10	0.09	0.19	0.13	0.10	0.09
Average			0.18	0.93	0.51	0.51	0.49	0.36	0.53	0.51	0.49
Group Maximum			0.18	0.93				0.53			
0.999	50	1	0.07	0.80	1.00	1.00	0.96	0.86	1.00	1.00	0.96
0.999	50	2	0.01	1.00	0.94	0.89	0.87	0.98	0.94	0.89	0.87
0.999	50	3	0.00	1.00	0.85	0.80	0.77	0.99	0.85	0.80	0.77
0.999	50	4	0.00	0.81	1.00	0.87	0.79	0.84	1.00	0.87	0.79
0.999	50	5	1.00	0.28	0.22	0.18	0.16	0.29	0.22	0.18	0.16
0.999	50	6	1.00	0.18	0.13	0.11	0.10	0.19	0.13	0.11	0.10
0.999	50	7	0.39	0.94	0.87	0.78	0.71	1.00	0.87	0.78	0.71
0.999	50	8	1.00	0.08	0.05	0.04	0.03	0.08	0.05	0.04	0.03
Average			0.43	0.64	0.63	0.58	0.55	0.65	0.63	0.58	0.55
Group Maximum			0.43	0.64				0.65			
0.999	100	1	0.28	0.97	1.00	0.97	0.92	0.97	1.00	0.97	0.92
0.999	100	2	0.08	1.00	0.91	0.87	0.84	1.00	0.91	0.87	0.84
0.999	100	3	0.00	1.00	0.83	0.78	0.76	1.00	0.83	0.78	0.76
0.999	100	4	0.01	1.00	0.88	0.73	0.66	1.00	0.88	0.73	0.66
0.999	100	5	1.00	0.18	0.13	0.11	0.10	0.18	0.13	0.11	0.10
0.999	100	6	1.00	0.11	0.08	0.07	0.06	0.11	0.08	0.07	0.06
0.999	100	7	1.00	0.81	0.66	0.58	0.53	0.81	0.66	0.58	0.53
0.999	100	8	1.00	0.05	0.03	0.02	0.02	0.05	0.03	0.02	0.02
Average			0.55	0.64	0.57	0.52	0.49	0.64	0.57	0.52	0.49
Group Maximum			0.55	0.64				0.64			

p_0	n	r	$\hat{q}_{p_0,r}^{(W)}$	$\hat{q}_{p_0,0.95,n,r}^{(EM)}$	$\hat{q}_{p_0,0.9,n,r}^{(EM)}$	$\hat{q}_{p_0,0.85,n,r}^{(EM)}$	$\hat{q}_{p_0,0.8,n,r}^{(EM)}$	$\hat{q}_{p_0,0.95,r}^{(AM)}$	$\hat{q}_{p_0,0.9,r}^{(AM)}$	$\hat{q}_{p_0,0.85,r}^{(AM)}$	$\hat{q}_{p_0,0.8,r}^{(AM)}$
0.999	150	1	0.43	1.00	1.00	0.96	0.92	1.00	1.00	0.96	0.92
0.999	150	2	0.16	1.00	0.90	0.86	0.83	1.00	0.90	0.86	0.83
0.999	150	3	0.01	1.00	0.83	0.78	0.76	1.00	0.83	0.78	0.76
0.999	150	4	0.01	1.00	0.74	0.60	0.54	1.00	0.74	0.60	0.54
0.999	150	5	1.00	0.15	0.11	0.09	0.08	0.15	0.11	0.09	0.08
0.999	150	6	1.00	0.09	0.06	0.05	0.04	0.09	0.06	0.05	0.04
0.999	150	7	1.00	0.55	0.45	0.39	0.35	0.55	0.45	0.39	0.35
0.999	150	8	1.00	0.04	0.02	0.02	0.02	0.04	0.02	0.02	0.02
Average			0.58	0.60	0.51	0.47	0.44	0.60	0.51	0.47	0.44
Group Maximum			0.58	0.60				0.60			
0.999	200	1	0.57	1.00	0.97	0.92	0.88	1.00	0.97	0.92	0.88
0.999	200	2	0.29	1.00	0.90	0.86	0.83	1.00	0.90	0.86	0.83
0.999	200	3	0.02	1.00	0.83	0.78	0.75	1.00	0.83	0.78	0.75
0.999	200	4	0.02	1.00	0.65	0.53	0.47	1.00	0.65	0.53	0.47
0.999	200	5	1.00	0.13	0.10	0.08	0.07	0.13	0.10	0.08	0.07
0.999	200	6	1.00	0.08	0.05	0.04	0.04	0.08	0.05	0.04	0.04
0.999	200	7	1.00	0.46	0.37	0.32	0.29	0.46	0.37	0.32	0.29
0.999	200	8	1.00	0.03	0.02	0.02	0.01	0.03	0.02	0.02	0.01
Average			0.61	0.59	0.49	0.44	0.42	0.59	0.49	0.44	0.42
Group Maximum			0.61	0.59				0.59			

The following conclusions are drawn from **Table 4.2** in terms of the relative performance of the estimators over the range of distributions and sample sizes considered. Note that this table contains the **RMSE** results for the quantile with $p_0 = 0.999$.

- Distribution view

- For $\underline{G}$, i.e. distributions with $0.2 \leq \gamma \leq 1$ and $-2 \leq \rho \leq -0.2$:
 - For $n \leq 150$ the preferred choice is the empirical multiplier $\hat{q}_{p_0,0.95,n,r}^{(EM)}$.
 - For $n \geq 200$ the preferred choice is the EVT estimator $\hat{q}_{p_0,r}^{(MOP)}$.
- For $\underline{G}_1$, i.e. distributions with $0.67 \leq \gamma \leq 1$ and $-2 \leq \rho \leq -1$:
 - For any n the preferred choice is the empirical multiplier $\hat{q}_{p_0,0.95,n,r}^{(EM)}$.
- For $\underline{G}_2$, i.e. distributions with $0.5 \leq \gamma \leq 0.8$ and $-0.3 \leq \rho \leq -0.2$:
 - For any n the preferred choice is the empirical multiplier $\hat{q}_{p_0,0.95,n,r}^{(EM)}$.
- For $\underline{G}_3$, i.e. distributions with $0.25 \leq \gamma \leq 0.3$ and $-0.8 \leq \rho \leq -0.4$:
 - For $n \leq 20$ the preferred choice is the empirical multiplier $\hat{q}_{p_0,0.95,n,r}^{(EM)}$.
 - For $n \geq 50$ the preferred choice is the EVT estimator $\hat{q}_{p_0,r}^{(MOP)}$.
- For $\underline{G}_4$, i.e. distributions with $0.25 \leq \gamma \leq 0.57$ and $\rho = -2$:

- For $n \leq 20$ the preferred choice is the empirical multiplier $\hat{q}_{p_0, 0.95, n, r}^{(EM)}$.
 - For $n \geq 50$ the preferred choice is the EVT estimator $\hat{q}_{p_0, r}^{(MOP)}$.
- Sample size view
 - For smaller sample sizes, i.e. $20 \leq n \leq 100$:
 - For $\underline{G}$ the preferred choice is the empirical multiplier $\hat{q}_{p_0, 0.95, n, r}^{(EM)}$.
 - For $\underline{G}^{(1)}$ the preferred choice is the empirical multiplier $\hat{q}_{p_0, 0.95, n, r}^{(EM)}$.
 - For $\underline{G}^{(2)}$ the preferred choice is the empirical multiplier $\hat{q}_{p_0, 0.95, n, r}^{(EM)}$.
 - For $\underline{G}^{(3)}$ the preferred choice is the empirical multiplier $\hat{q}_{p_0, 0.95, n, r}^{(EM)}$ for $n \leq 20$, otherwise the EVT estimator $\hat{q}_{p_0, r}^{(MOP)}$.
 - For $\underline{G}^{(4)}$ the preferred choice is the empirical multiplier $\hat{q}_{p_0, 0.95, n, r}^{(EM)}$ for $n \leq 20$, otherwise the EVT estimator $\hat{q}_{p_0, r}^{(MOP)}$.
 - For larger sample sizes, i.e. $100 \leq n \leq 200$:
 - For $\underline{G}$ the preferred choice is the empirical multiplier $\hat{q}_{p_0, 0.95, n, r}^{(EM)}$ or the EVT estimator $\hat{q}_{p_0, r}^{(MOP)}$.
 - For $\underline{G}^{(1)}$ the preferred choice is the empirical multiplier $\hat{q}_{p_0, 0.95, n, r}^{(EM)}$.
 - For $\underline{G}^{(2)}$ the preferred choice is the empirical multiplier $\hat{q}_{p_0, 0.95, n, r}^{(EM)}$.
 - For $\underline{G}^{(3)}$ the preferred choice is the EVT estimator $\hat{q}_{p_0, r}^{(MOP)}$.
 - For $\underline{G}^{(4)}$ the preferred choice is the EVT estimator $\hat{q}_{p_0, r}^{(MOP)}$.

Table 4.3: Absolute bias results of the first order EVT simulation study for $p_0 = 0.999$

p_0	n	r	$\hat{q}_{p_0,r}^{(W)}$	$\hat{q}_{p_0,0.95,n,r}^{(EM)}$	$\hat{q}_{p_0,0.9,n,r}^{(EM)}$	$\hat{q}_{p_0,0.85,n,r}^{(EM)}$	$\hat{q}_{p_0,0.8,n,r}^{(EM)}$	$\hat{q}_{p_0,0.95,r}^{(AM)}$	$\hat{q}_{p_0,0.9,r}^{(AM)}$	$\hat{q}_{p_0,0.85,r}^{(AM)}$	$\hat{q}_{p_0,0.8,r}^{(AM)}$
0.999	20	1	0.02	1.00	0.59	0.71	0.68	0.59	0.70	0.71	0.68
0.999	20	2	0.00	0.47	0.72	0.62	0.60	1.00	0.67	0.62	0.60
0.999	20	3	0.00	0.52	0.59	0.52	0.51	1.00	0.57	0.52	0.51
0.999	20	4	0.00	0.26	0.56	0.35	0.31	1.00	0.48	0.35	0.31
0.999	20	5	0.53	1.00	0.11	0.10	0.09	0.16	0.12	0.10	0.09
0.999	20	6	0.87	1.00	0.10	0.09	0.08	0.16	0.11	0.09	0.08
0.999	20	7	0.03	1.00	0.13	0.13	0.12	0.16	0.14	0.13	0.12
0.999	20	8	1.00	0.23	0.03	0.02	0.02	0.04	0.03	0.02	0.02
Average			0.31	0.69	0.35	0.32	0.30	0.51	0.36	0.32	0.30
Group Maximum			0.31	0.69				0.51			
0.999	50	1	0.31	0.91	0.99	0.95	0.90	1.00	0.99	0.95	0.90
0.999	50	2	0.07	1.00	0.80	0.75	0.72	0.96	0.80	0.75	0.72
0.999	50	3	0.00	1.00	0.74	0.69	0.66	0.97	0.74	0.69	0.66
0.999	50	4	0.00	1.00	0.05	0.04	0.03	0.38	0.05	0.04	0.03
0.999	50	5	1.00	0.15	0.11	0.10	0.08	0.15	0.11	0.10	0.08
0.999	50	6	1.00	0.07	0.05	0.04	0.04	0.07	0.05	0.04	0.04
0.999	50	7	1.00	0.50	0.43	0.37	0.34	0.53	0.43	0.37	0.34
0.999	50	8	1.00	0.01	0.01	0.01	0.01	0.02	0.01	0.01	0.01
Average			0.55	0.58	0.40	0.37	0.35	0.51	0.40	0.37	0.35
Group Maximum			0.55	0.58				0.51			
0.999	100	1	0.72	1.00	0.94	0.89	0.84	1.00	0.94	0.89	0.84
0.999	100	2	0.35	1.00	0.87	0.82	0.79	1.00	0.87	0.82	0.79
0.999	100	3	0.02	1.00	0.80	0.75	0.73	1.00	0.80	0.75	0.73
0.999	100	4	0.01	1.00	0.25	0.20	0.17	1.00	0.25	0.20	0.17
0.999	100	5	1.00	0.12	0.09	0.07	0.06	0.12	0.09	0.07	0.06
0.999	100	6	1.00	0.05	0.04	0.03	0.03	0.05	0.04	0.03	0.03
0.999	100	7	1.00	0.21	0.16	0.14	0.13	0.21	0.16	0.14	0.13
0.999	100	8	1.00	0.01	0.00	0.00	0.00	0.01	0.00	0.00	0.00
Average			0.64	0.55	0.39	0.36	0.35	0.55	0.39	0.36	0.35
Group Maximum			0.64	0.55				0.55			

p_0	n	r	$\hat{q}_{p_0,r}^{(W)}$	$\hat{q}_{p_0,0.95,n,r}^{(EM)}$	$\hat{q}_{p_0,0.9,n,r}^{(EM)}$	$\hat{q}_{p_0,0.85,n,r}^{(EM)}$	$\hat{q}_{p_0,0.8,n,r}^{(EM)}$	$\hat{q}_{p_0,0.95,r}^{(AM)}$	$\hat{q}_{p_0,0.9,r}^{(AM)}$	$\hat{q}_{p_0,0.85,r}^{(AM)}$	$\hat{q}_{p_0,0.8,r}^{(AM)}$
0.999	150	1	0.93	1.00	0.93	0.88	0.84	1.00	0.93	0.88	0.84
0.999	150	2	0.63	1.00	0.88	0.83	0.80	1.00	0.88	0.83	0.80
0.999	150	3	0.04	1.00	0.81	0.76	0.74	1.00	0.81	0.76	0.74
0.999	150	4	0.01	1.00	0.30	0.23	0.20	1.00	0.30	0.23	0.20
0.999	150	5	1.00	0.11	0.08	0.07	0.06	0.11	0.08	0.07	0.06
0.999	150	6	1.00	0.05	0.03	0.03	0.02	0.05	0.03	0.03	0.02
0.999	150	7	1.00	0.15	0.12	0.10	0.09	0.15	0.12	0.10	0.09
0.999	150	8	1.00	0.01	0.00	0.00	0.00	0.01	0.00	0.00	0.00
Average			0.70	0.54	0.39	0.36	0.34	0.54	0.39	0.36	0.34
Group Maximum			0.70	0.54				0.54			
0.999	200	1	1.00	0.87	0.80	0.75	0.71	0.87	0.80	0.75	0.71
0.999	200	2	0.96	1.00	0.88	0.84	0.81	1.00	0.88	0.84	0.81
0.999	200	3	0.05	1.00	0.81	0.76	0.74	1.00	0.81	0.76	0.74
0.999	200	4	0.01	1.00	0.32	0.25	0.23	1.00	0.32	0.25	0.23
0.999	200	5	1.00	0.10	0.07	0.06	0.05	0.10	0.07	0.06	0.05
0.999	200	6	1.00	0.04	0.03	0.02	0.02	0.04	0.03	0.02	0.02
0.999	200	7	1.00	0.11	0.09	0.08	0.07	0.11	0.09	0.08	0.07
0.999	200	8	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Average			0.75	0.52	0.38	0.35	0.33	0.52	0.38	0.35	0.33
Group Maximum			0.75	0.52				0.52			

The following conclusions are drawn from Table 4.3 in terms of the relative performance of the estimators over the range of distributions and sample sizes considered. Note that this table contains the **absolute bias** results for the quantile with $p_0 = 0.999$.

- Distribution view

- For $\underline{G}$, i.e. distributions with $0.2 \leq \gamma \leq 1$ and $-2 \leq \rho \leq -0.2$:
 - For $n \leq 50$ the preferred choice is the empirical multiplier $\hat{q}_{p_0,0.95,n,r}^{(EM)}$.
 - For $n \geq 100$ the preferred choice is the EVT estimator $\hat{q}_{p_0,r}^{(MOP)}$.
- For $\underline{G}_1$, i.e. distributions with $0.67 \leq \gamma \leq 1$ and $-2 \leq \rho \leq -1$:
 - For any n the preferred choice is the empirical multiplier $\hat{q}_{p_0,0.95,n,r}^{(EM)}$.
- For $\underline{G}_2$, i.e. distributions with $0.5 \leq \gamma \leq 0.8$ and $-0.3 \leq \rho \leq -0.2$:
 - For any n the preferred choice is the empirical multiplier $\hat{q}_{p_0,0.95,n,r}^{(EM)}$.
- For $\underline{G}_3$, i.e. distributions with $0.25 \leq \gamma \leq 0.3$ and $-0.8 \leq \rho \leq -0.4$:
 - For $n \leq 20$ the preferred choice is the empirical multiplier $\hat{q}_{p_0,0.95,n,r}^{(EM)}$.
 - For $n \geq 50$ the preferred choice is the EVT estimator $\hat{q}_{p_0,r}^{(MOP)}$.
- For $\underline{G}_4$, i.e. distributions with $0.25 \leq \gamma \leq 0.57$ and $\rho = -2$:

- For $n \leq 20$ the preferred choice is the empirical multiplier $\hat{q}_{p_0,0.95,n,r}^{(EM)}$.
 - For $n \geq 50$ the preferred choice is the EVT estimator $\hat{q}_{p_0,r}^{(MOP)}$.
- Sample size view
 - For smaller sample sizes, i.e. $20 \leq n \leq 100$:
 - For $\underline{G}$ the preferred choice is the empirical multiplier $\hat{q}_{p_0,0.95,n,r}^{(EM)}$.
 - For $\underline{G}^{(1)}$ the preferred choice is the empirical multiplier $\hat{q}_{p_0,0.95,n,r}^{(EM)}$.
 - For $\underline{G}^{(2)}$ the preferred choice is the empirical multiplier $\hat{q}_{p_0,0.95,n,r}^{(EM)}$.
 - For $\underline{G}^{(3)}$ the preferred choice is the empirical multiplier $\hat{q}_{p_0,0.95,n,r}^{(EM)}$ for $n \leq 20$, otherwise the EVT estimator $\hat{q}_{p_0,r}^{(MOP)}$.
 - For $\underline{G}^{(4)}$ the preferred choice is the empirical multiplier $\hat{q}_{p_0,0.95,n,r}^{(EM)}$ for $n \leq 20$, otherwise the EVT estimator $\hat{q}_{p_0,r}^{(MOP)}$.
 - For larger sample sizes, i.e. $100 \leq n \leq 200$:
 - For $\underline{G}$ the preferred choice is the EVT estimator $\hat{q}_{p_0,r}^{(MOP)}$.
 - For $\underline{G}^{(1)}$ the preferred choice is the empirical multiplier $\hat{q}_{p_0,0.95,n,r}^{(EM)}$.
 - For $\underline{G}^{(2)}$ the preferred choice is the empirical multiplier $\hat{q}_{p_0,0.95,n,r}^{(EM)}$.
 - For $\underline{G}^{(3)}$ the preferred choice is the EVT estimator $\hat{q}_{p_0,r}^{(MOP)}$.
 - For $\underline{G}^{(4)}$ the preferred choice is the EVT estimator $\hat{q}_{p_0,r}^{(MOP)}$.

Note that the only difference in the conclusions made (for $p_0 = 0.001$) between minimizing the RMSE and absolute bias is the use of EVT only for larger sample sizes when the underlying distribution is unknown, i.e. $100 \leq n \leq 200$ with $\underline{G}$.

Results of the first order EVT simulation study for $p_0 = 0.995$

Tables 4.4 and 4.5 contain the **RMSE** (Step 6.2 in Algorithm 4.1) and **absolute bias** (Step 6.4 in Algorithm 4.1) results respectively for $p_0 = 0.995$ in the **first order EVT** environment.

Table 4.4: Root mean squared error results of the first order EVT simulation study for $p_0 = 0.995$

p_0	n	r	$\hat{q}_{p_0,r}^{(W)}$	$\hat{q}_{p_0,0.95,n,r}^{(EM)}$	$\hat{q}_{p_0,0.9,n,r}^{(EM)}$	$\hat{q}_{p_0,0.85,n,r}^{(EM)}$	$\hat{q}_{p_0,0.8,n,r}^{(EM)}$	$\hat{q}_{p_0,0.95,r}^{(AM)}$	$\hat{q}_{p_0,0.9,r}^{(AM)}$	$\hat{q}_{p_0,0.85,r}^{(AM)}$	$\hat{q}_{p_0,0.8,r}^{(AM)}$
0.995	20	1	0.02	1.00	0.76	0.97	0.94	0.51	0.86	0.97	0.94
0.995	20	2	0.00	0.48	0.94	1.00	0.98	0.27	0.93	1.00	0.98
0.995	20	3	0.00	0.73	1.00	0.93	0.90	0.45	0.99	0.93	0.90
0.995	20	4	0.00	0.80	1.00	0.96	0.92	0.51	1.00	0.96	0.92
0.995	20	5	0.38	1.00	0.22	0.20	0.17	0.32	0.24	0.20	0.17
0.995	20	6	0.46	1.00	0.20	0.18	0.15	0.31	0.22	0.18	0.15
0.995	20	7	0.06	1.00	0.41	0.43	0.39	0.41	0.47	0.43	0.39
0.995	20	8	1.00	0.83	0.16	0.13	0.11	0.27	0.17	0.13	0.11
Average			0.24	0.86	0.59	0.60	0.57	0.38	0.61	0.60	0.57
Group Maximum			0.24	0.86				0.61			
0.995	50	1	0.17	1.00	0.79	0.81	0.77	0.63	0.79	0.81	0.77
0.995	50	2	0.06	0.73	1.00	0.94	0.91	0.92	1.00	0.94	0.91
0.995	50	3	0.01	0.84	0.86	0.79	0.76	1.00	0.86	0.79	0.76
0.995	50	4	0.02	0.85	1.00	0.88	0.82	0.97	1.00	0.88	0.82
0.995	50	5	0.87	1.00	0.22	0.18	0.15	0.31	0.22	0.18	0.15
0.995	50	6	1.00	0.88	0.18	0.14	0.12	0.27	0.18	0.14	0.12
0.995	50	7	0.23	1.00	0.29	0.25	0.23	0.33	0.29	0.25	0.23
0.995	50	8	1.00	0.41	0.09	0.07	0.05	0.14	0.09	0.07	0.05
Average			0.42	0.84	0.55	0.51	0.47	0.57	0.55	0.51	0.47
Group Maximum			0.42	0.84				0.57			
0.995	100	1	0.42	0.83	1.00	0.98	0.92	0.90	1.00	0.98	0.92
0.995	100	2	0.17	1.00	0.89	0.83	0.79	0.97	0.89	0.83	0.79
0.995	100	3	0.04	1.00	0.77	0.71	0.68	0.98	0.77	0.71	0.68
0.995	100	4	0.07	1.00	0.77	0.66	0.61	0.99	0.77	0.66	0.61
0.995	100	5	1.00	0.23	0.17	0.14	0.12	0.24	0.17	0.14	0.12
0.995	100	6	1.00	0.17	0.12	0.09	0.08	0.18	0.12	0.09	0.08
0.995	100	7	1.00	0.82	0.70	0.60	0.53	0.88	0.70	0.60	0.53
0.995	100	8	1.00	0.09	0.06	0.04	0.03	0.09	0.06	0.04	0.03
Average			0.59	0.64	0.56	0.51	0.47	0.65	0.56	0.51	0.47
Group Maximum			0.59	0.64				0.65			

p_0	n	r	$\hat{q}_{p_0,r}^{(W)}$	$\hat{q}_{p_0,0.95,n,r}^{(EM)}$	$\hat{q}_{p_0,0.9,n,r}^{(EM)}$	$\hat{q}_{p_0,0.85,n,r}^{(EM)}$	$\hat{q}_{p_0,0.8,n,r}^{(EM)}$	$\hat{q}_{p_0,0.95,r}^{(AM)}$	$\hat{q}_{p_0,0.9,r}^{(AM)}$	$\hat{q}_{p_0,0.85,r}^{(AM)}$	$\hat{q}_{p_0,0.8,r}^{(AM)}$
0.995	150	1	0.54	0.97	1.00	0.95	0.89	0.97	1.00	0.95	0.89
0.995	150	2	0.31	1.00	0.87	0.81	0.77	1.00	0.87	0.81	0.77
0.995	150	3	0.09	1.00	0.78	0.71	0.69	1.00	0.78	0.71	0.69
0.995	150	4	0.09	1.00	0.70	0.59	0.54	1.00	0.70	0.59	0.54
0.995	150	5	1.00	0.21	0.14	0.12	0.10	0.21	0.14	0.12	0.10
0.995	150	6	1.00	0.14	0.09	0.07	0.06	0.14	0.09	0.07	0.06
0.995	150	7	1.00	0.67	0.52	0.44	0.39	0.67	0.52	0.44	0.39
0.995	150	8	1.00	0.07	0.04	0.03	0.03	0.07	0.04	0.03	0.03
Average			0.63	0.63	0.52	0.47	0.43	0.63	0.52	0.47	0.43
Group Maximum			0.63	0.63				0.63			
0.995	200	1	0.64	1.00	1.00	0.94	0.88	1.00	1.00	0.94	0.88
0.995	200	2	0.41	1.00	0.86	0.79	0.75	1.00	0.86	0.79	0.75
0.995	200	3	0.12	1.00	0.77	0.71	0.68	1.00	0.77	0.71	0.68
0.995	200	4	0.11	1.00	0.66	0.56	0.51	1.00	0.66	0.56	0.51
0.995	200	5	1.00	0.19	0.13	0.11	0.09	0.19	0.13	0.11	0.09
0.995	200	6	1.00	0.12	0.08	0.06	0.05	0.12	0.08	0.06	0.05
0.995	200	7	1.00	0.55	0.41	0.34	0.30	0.55	0.41	0.34	0.30
0.995	200	8	1.00	0.06	0.04	0.03	0.02	0.06	0.04	0.03	0.02
Average			0.66	0.61	0.49	0.44	0.41	0.61	0.49	0.44	0.41
Group Maximum			0.66	0.61				0.61			

The following conclusions are drawn from **Table 4.4** in terms of the relative performance of the estimators over the range of distributions and sample sizes considered. Note that this table contains the **RMSE** results for the quantile with $p_0 = 0.995$.

- Distribution view

- For $\underline{G}$, i.e. distributions with $0.2 \leq \gamma \leq 1$ and $-2 \leq \rho \leq -0.2$:
 - For $n \leq 150$ the preferred choice is the empirical multiplier $\hat{q}_{p_0,0.95,n,r}^{(EM)}$.
 - For $n \geq 200$ the preferred choice is the EVT estimator $\hat{q}_{p_0,r}^{(MOP)}$.
- For $\underline{G}_1$, i.e. distributions with $0.67 \leq \gamma \leq 1$ and $-2 \leq \rho \leq -1$:
 - For any n the preferred choice is the empirical multiplier $\hat{q}_{p_0,0.95,n,r}^{(EM)}$.
- For $\underline{G}_2$, i.e. distributions with $0.5 \leq \gamma \leq 0.8$ and $-0.3 \leq \rho \leq -0.2$:
 - For $n \leq 50$ the preferred choice is the empirical multiplier $\hat{q}_{p_0,0.9,n,r}^{(EM)}$.
 - For $n \geq 100$ the preferred choice is the empirical multiplier $\hat{q}_{p_0,0.95,n,r}^{(EM)}$.
- For $\underline{G}_3$, i.e. distributions with $0.25 \leq \gamma \leq 0.3$ and $-0.8 \leq \rho \leq -0.4$:
 - For $n \leq 50$ the preferred choice is the empirical multiplier $\hat{q}_{p_0,0.95,n,r}^{(EM)}$.

- For $n \geq 100$ use the EVT estimator $\hat{q}_{p_0, r}^{(MOP)}$.
- For $\underline{G}_4$, i.e. distributions with $0.25 \leq \gamma \leq 0.57$ and $\rho = -2$:
 - For $n \leq 50$ the preferred choice is the empirical multiplier $\hat{q}_{p_0, 0.95, n, r}^{(EM)}$.
 - For $n \geq 100$ the preferred choice is the EVT estimator $\hat{q}_{p_0, r}^{(MOP)}$.
- Sample size view
 - For smaller sample sizes, i.e. $20 \leq n \leq 100$:
 - For $\underline{G}$ the preferred choice is the empirical multiplier $\hat{q}_{p_0, 0.95, n, r}^{(EM)}$.
 - For $\underline{G}^{(1)}$ the preferred choice is the empirical multiplier $\hat{q}_{p_0, 0.95, n, r}^{(EM)}$.
 - For $\underline{G}^{(2)}$ the preferred choice is the empirical multiplier $\hat{q}_{p_0, 0.9, n, r}^{(EM)}$ for $n \leq 50$, otherwise $\hat{q}_{p_0, 0.95, n, r}^{(EM)}$.
 - For $\underline{G}^{(3)}$ the preferred choice is the empirical multiplier $\hat{q}_{p_0, 0.95, n, r}^{(EM)}$ for $n \leq 50$, otherwise the EVT estimator $\hat{q}_{p_0, r}^{(MOP)}$.
 - For $\underline{G}^{(4)}$ the preferred choice is the empirical multiplier $\hat{q}_{p_0, 0.95, n, r}^{(EM)}$ for $n \leq 50$, otherwise the EVT estimator $\hat{q}_{p_0, r}^{(MOP)}$.
 - For larger sample sizes, i.e. $100 \leq n \leq 200$:
 - For $\underline{G}$ the preferred choice is the empirical multiplier $\hat{q}_{p_0, 0.95, n, r}^{(EM)}$ or the EVT estimator $\hat{q}_{p_0, r}^{(MOP)}$.
 - For $\underline{G}^{(1)}$ the preferred choice is the empirical multiplier $\hat{q}_{p_0, 0.95, n, r}^{(EM)}$.
 - For $\underline{G}^{(2)}$ the preferred choice is the empirical multiplier $\hat{q}_{p_0, 0.95, n, r}^{(EM)}$.
 - For $\underline{G}^{(3)}$ the preferred choice is the EVT estimator $\hat{q}_{p_0, r}^{(MOP)}$.
 - For $\underline{G}^{(4)}$ the preferred choice is the EVT estimator $\hat{q}_{p_0, r}^{(MOP)}$.

Table 4.5: Absolute bias results of the first order EVT simulation study for $p_0 = 0.995$

p_0	n	r	$\hat{q}_{p_0,r}^{(W)}$	$\hat{q}_{p_0,0.95,n,r}^{(EM)}$	$\hat{q}_{p_0,0.9,n,r}^{(EM)}$	$\hat{q}_{p_0,0.85,n,r}^{(EM)}$	$\hat{q}_{p_0,0.8,n,r}^{(EM)}$	$\hat{q}_{p_0,0.95,r}^{(AM)}$	$\hat{q}_{p_0,0.9,r}^{(AM)}$	$\hat{q}_{p_0,0.85,r}^{(AM)}$	$\hat{q}_{p_0,0.8,r}^{(AM)}$
0.995	20	1	0.08	1.00	0.56	0.73	0.68	0.56	0.72	0.73	0.68
0.995	20	2	0.00	0.23	0.35	0.26	0.24	1.00	0.31	0.26	0.24
0.995	20	3	0.00	0.42	0.39	0.33	0.32	1.00	0.38	0.33	0.32
0.995	20	4	0.00	0.08	0.10	0.07	0.07	1.00	0.09	0.07	0.07
0.995	20	5	0.82	1.00	0.17	0.15	0.13	0.26	0.18	0.15	0.13
0.995	20	6	1.00	0.71	0.11	0.10	0.08	0.19	0.13	0.10	0.08
0.995	20	7	0.08	1.00	0.09	0.09	0.08	0.13	0.11	0.09	0.08
0.995	20	8	1.00	0.16	0.03	0.02	0.02	0.05	0.03	0.02	0.02
Average			0.37	0.58	0.22	0.22	0.20	0.52	0.24	0.22	0.20
Group Maximum			0.37	0.58				0.52			
0.995	50	1	0.42	1.00	0.82	0.77	0.71	0.83	0.82	0.77	0.71
0.995	50	2	0.20	0.48	0.68	0.61	0.58	1.00	0.68	0.61	0.58
0.995	50	3	0.04	0.67	0.66	0.60	0.58	1.00	0.66	0.60	0.58
0.995	50	4	0.03	0.32	0.38	0.31	0.29	1.00	0.38	0.31	0.29
0.995	50	5	1.00	0.64	0.12	0.10	0.09	0.18	0.12	0.10	0.09
0.995	50	6	1.00	0.33	0.06	0.05	0.04	0.09	0.06	0.05	0.04
0.995	50	7	0.32	1.00	0.10	0.09	0.08	0.14	0.10	0.09	0.08
0.995	50	8	1.00	0.06	0.01	0.01	0.01	0.02	0.01	0.01	0.01
Average			0.50	0.56	0.36	0.32	0.30	0.53	0.36	0.32	0.30
Group Maximum			0.50	0.56				0.53			
0.995	100	1	0.88	0.87	0.92	0.85	0.79	1.00	0.92	0.85	0.79
0.995	100	2	0.60	1.00	0.70	0.63	0.60	0.93	0.70	0.63	0.60
0.995	100	3	0.12	1.00	0.69	0.63	0.61	0.97	0.69	0.63	0.61
0.995	100	4	0.07	1.00	0.41	0.34	0.31	0.88	0.41	0.34	0.31
0.995	100	5	1.00	0.14	0.10	0.08	0.07	0.15	0.10	0.08	0.07
0.995	100	6	1.00	0.07	0.05	0.04	0.03	0.07	0.05	0.04	0.03
0.995	100	7	1.00	0.19	0.15	0.12	0.11	0.20	0.15	0.12	0.11
0.995	100	8	1.00	0.01	0.01	0.00	0.00	0.01	0.01	0.00	0.00
Average			0.71	0.53	0.38	0.34	0.32	0.53	0.38	0.34	0.32
Group Maximum			0.71	0.53				0.53			

p_0	n	r	$\hat{q}_{p_0,r}^{(W)}$	$\hat{q}_{p_0,0.95,n,r}^{(EM)}$	$\hat{q}_{p_0,0.9,n,r}^{(EM)}$	$\hat{q}_{p_0,0.85,n,r}^{(EM)}$	$\hat{q}_{p_0,0.8,n,r}^{(EM)}$	$\hat{q}_{p_0,0.95,r}^{(AM)}$	$\hat{q}_{p_0,0.9,r}^{(AM)}$	$\hat{q}_{p_0,0.85,r}^{(AM)}$	$\hat{q}_{p_0,0.8,r}^{(AM)}$
0.995	150	1	1.00	0.90	0.81	0.74	0.68	0.90	0.81	0.74	0.68
0.995	150	2	1.00	0.94	0.72	0.66	0.62	0.94	0.72	0.66	0.62
0.995	150	3	0.18	1.00	0.74	0.68	0.65	1.00	0.74	0.68	0.65
0.995	150	4	0.10	1.00	0.49	0.41	0.37	1.00	0.49	0.41	0.37
0.995	150	5	1.00	0.14	0.09	0.08	0.07	0.14	0.09	0.08	0.07
0.995	150	6	1.00	0.06	0.04	0.03	0.03	0.06	0.04	0.03	0.03
0.995	150	7	1.00	0.14	0.10	0.09	0.08	0.14	0.10	0.09	0.08
0.995	150	8	1.00	0.01	0.00	0.00	0.00	0.01	0.00	0.00	0.00
Average			0.79	0.52	0.38	0.34	0.31	0.52	0.38	0.34	0.31
Group Maximum			0.79	0.52				0.52			
0.995	200	1	1.00	0.79	0.70	0.64	0.59	0.79	0.70	0.64	0.59
0.995	200	2	1.00	0.68	0.53	0.48	0.46	0.68	0.53	0.48	0.46
0.995	200	3	0.21	1.00	0.74	0.68	0.65	1.00	0.74	0.68	0.65
0.995	200	4	0.12	1.00	0.51	0.42	0.39	1.00	0.51	0.42	0.39
0.995	200	5	1.00	0.13	0.09	0.07	0.06	0.13	0.09	0.07	0.06
0.995	200	6	1.00	0.06	0.04	0.03	0.03	0.06	0.04	0.03	0.03
0.995	200	7	1.00	0.10	0.07	0.06	0.05	0.10	0.07	0.06	0.05
0.995	200	8	1.00	0.01	0.00	0.00	0.00	0.01	0.00	0.00	0.00
Average			0.79	0.47	0.34	0.30	0.28	0.47	0.34	0.30	0.28
Group Maximum			0.79	0.47				0.47			

The following conclusions are drawn from **Table 4.5** in terms of the relative performance of the estimators over the range of distributions and sample sizes considered. Note that this table contains the **absolute bias** results for the quantile with $p_0 = 0.995$.

- Distribution view

- For $\underline{G}$, i.e. distributions with $0.2 \leq \gamma \leq 1$ and $-2 \leq \rho \leq -0.2$:
 - For $n \leq 50$ the preferred choice is the empirical multiplier $\hat{q}_{p_0,0.95,n,r}^{(EM)}$.
 - For $n \geq 100$ the preferred choice is the EVT estimator $\hat{q}_{p_0,r}^{(MOP)}$.
- For $\underline{G}_1$, i.e. distributions with $0.67 \leq \gamma \leq 1$ and $-2 \leq \rho \leq -1$:
 - For $n \leq 100$ the preferred choice is the asymptotic multiplier $\hat{q}_{p_0,0.95,n,r}^{(AM)}$.
 - For $n \geq 100$ the preferred choice is the EVT estimator $\hat{q}_{p_0,r}^{(MOP)}$.
- For $\underline{G}_2$, i.e. distributions with $0.5 \leq \gamma \leq 0.8$ and $-0.3 \leq \rho \leq -0.2$:
 - For $n \leq 50$ the preferred choice is the asymptotic multiplier $\hat{q}_{p_0,0.95,n,r}^{(AM)}$.
 - For $n \geq 100$ the preferred choice is the empirical multiplier $\hat{q}_{p_0,0.95,n,r}^{(EM)}$.
- For $\underline{G}_3$, i.e. distributions with $0.25 \leq \gamma \leq 0.3$ and $-0.8 \leq \rho \leq -0.4$:
 - For any n the preferred choice is the EVT estimator $\hat{q}_{p_0,r}^{(MOP)}$.

- For $\underline{G}_4$, i.e. distributions with $0.25 \leq \gamma \leq 0.57$ and $\rho = -2$:
 - For any n the preferred choice is the EVT estimator $\hat{q}_{p_0,r}^{(MOP)}$.
- Sample size view
 - For smaller sample sizes, i.e. $20 \leq n \leq 100$:
 - For $\underline{G}$ the preferred choice is the asymptotic multiplier $\hat{q}_{p_0,0.95,n,r}^{(AM)}$ for $n \leq 50$, otherwise use the EVT estimator $\hat{q}_{p_0,r}^{(MOP)}$.
 - For $\underline{G}^{(1)}$ the preferred choice is the asymptotic multiplier $\hat{q}_{p_0,0.95,n,r}^{(AM)}$.
 - For $\underline{G}^{(2)}$ the preferred choice is the asymptotic multiplier $\hat{q}_{p_0,0.95,n,r}^{(AM)}$.
 - For $\underline{G}^{(3)}$ the preferred choice is the EVT estimator $\hat{q}_{p_0,r}^{(MOP)}$.
 - For $\underline{G}^{(4)}$ the preferred choice is the EVT estimator $\hat{q}_{p_0,r}^{(MOP)}$.
 - For larger sample sizes, i.e. $100 \leq n \leq 200$:
 - For $\underline{G}$ the preferred choice is the EVT estimator $\hat{q}_{p_0,r}^{(MOP)}$.
 - For $\underline{G}^{(1)}$ the preferred choice is the EVT estimator $\hat{q}_{p_0,r}^{(MOP)}$.
 - For $\underline{G}^{(2)}$ the preferred choice is the empirical multiplier $\hat{q}_{p_0,0.95,n,r}^{(EM)}$.
 - For $\underline{G}^{(3)}$ the preferred choice is the EVT estimator $\hat{q}_{p_0,r}^{(MOP)}$.
 - For $\underline{G}^{(4)}$ the preferred choice is the EVT estimator $\hat{q}_{p_0,r}^{(MOP)}$.

Results of the first order EVT simulation study for $p_0 = 0.99$

Tables 4.6 and 4.7 contain the **RMSE** (Step 6.2 in Algorithm 4.1) and **absolute bias** (Step 6.4 in Algorithm 4.1) results respectively for $p_0 = 0.99$ in the **first order EVT** environment.

Table 4.6: Root mean squared error results of the first order EVT simulation study for $p_0 = 0.99$

p_0	n	r	$\hat{q}_{p_0,r}^{(W)}$	$\hat{q}_{p_0,0.95,n,r}^{(EM)}$	$\hat{q}_{p_0,0.9,n,r}^{(EM)}$	$\hat{q}_{p_0,0.85,n,r}^{(EM)}$	$\hat{q}_{p_0,0.8,n,r}^{(EM)}$	$\hat{q}_{p_0,0.95,r}^{(AM)}$	$\hat{q}_{p_0,0.9,r}^{(AM)}$	$\hat{q}_{p_0,0.85,r}^{(AM)}$	$\hat{q}_{p_0,0.8,r}^{(AM)}$
0.99	20	1	0.06	0.88	0.87	1.00	0.97	0.49	0.87	1.00	0.97
0.99	20	2	0.00	0.53	0.91	1.00	0.98	0.31	0.91	1.00	0.98
0.99	20	3	0.00	0.58	1.00	0.95	0.91	0.33	1.00	0.95	0.91
0.99	20	4	0.00	0.79	1.00	0.95	0.91	0.51	1.00	0.95	0.91
0.99	20	5	0.34	1.00	0.19	0.15	0.13	0.27	0.19	0.15	0.13
0.99	20	6	0.39	1.00	0.18	0.14	0.12	0.26	0.18	0.14	0.12
0.99	20	7	0.14	1.00	0.57	0.53	0.47	0.48	0.57	0.53	0.47
0.99	20	8	0.67	1.00	0.14	0.10	0.08	0.23	0.14	0.10	0.08
Average			0.20	0.85	0.61	0.60	0.57	0.36	0.61	0.60	0.57
Group Maximum			0.20	0.85				0.61			
0.99	50	1	0.23	1.00	0.93	0.97	0.92	0.68	0.93	0.97	0.92
0.99	50	2	0.13	0.75	1.00	0.95	0.90	0.80	1.00	0.95	0.90
0.99	50	3	0.03	0.94	0.93	0.84	0.80	1.00	0.93	0.84	0.80
0.99	50	4	0.08	0.93	1.00	0.88	0.82	0.95	1.00	0.88	0.82
0.99	50	5	0.87	1.00	0.24	0.19	0.16	0.36	0.24	0.19	0.16
0.99	50	6	1.00	0.91	0.21	0.16	0.13	0.33	0.21	0.16	0.13
0.99	50	7	0.34	1.00	0.40	0.35	0.30	0.46	0.40	0.35	0.30
0.99	50	8	1.00	0.48	0.11	0.08	0.07	0.19	0.11	0.08	0.07
Average			0.46	0.88	0.60	0.55	0.51	0.60	0.60	0.55	0.51
Group Maximum			0.46	0.88				0.60			
0.99	100	1	0.47	0.80	1.00	0.99	0.92	0.86	1.00	0.99	0.92
0.99	100	2	0.30	1.00	0.95	0.87	0.82	0.98	0.95	0.87	0.82
0.99	100	3	0.14	1.00	0.75	0.68	0.64	0.98	0.75	0.68	0.64
0.99	100	4	0.16	1.00	0.76	0.65	0.60	0.99	0.76	0.65	0.60
0.99	100	5	1.00	0.28	0.19	0.15	0.13	0.29	0.19	0.15	0.13
0.99	100	6	1.00	0.21	0.14	0.11	0.09	0.23	0.14	0.11	0.09
0.99	100	7	1.00	0.84	0.70	0.58	0.51	0.90	0.70	0.58	0.51
0.99	100	8	1.00	0.12	0.07	0.05	0.04	0.13	0.07	0.05	0.04
Average			0.63	0.66	0.57	0.51	0.47	0.67	0.57	0.51	0.47
Group Maximum			0.63	0.66				0.67			

p_0	n	r	$\hat{q}_{p_0,r}^{(W)}$	$\hat{q}_{p_0,0.95,n,r}^{(EM)}$	$\hat{q}_{p_0,0.9,n,r}^{(EM)}$	$\hat{q}_{p_0,0.85,n,r}^{(EM)}$	$\hat{q}_{p_0,0.8,n,r}^{(EM)}$	$\hat{q}_{p_0,0.95,r}^{(AM)}$	$\hat{q}_{p_0,0.9,r}^{(AM)}$	$\hat{q}_{p_0,0.85,r}^{(AM)}$	$\hat{q}_{p_0,0.8,r}^{(AM)}$
0.99	150	1	0.58	0.93	1.00	0.95	0.88	0.93	1.00	0.95	0.88
0.99	150	2	0.42	1.00	0.88	0.80	0.75	1.00	0.88	0.80	0.75
0.99	150	3	0.20	1.00	0.74	0.67	0.63	1.00	0.74	0.67	0.63
0.99	150	4	0.20	1.00	0.70	0.59	0.55	1.00	0.70	0.59	0.55
0.99	150	5	1.00	0.25	0.16	0.12	0.11	0.25	0.16	0.12	0.11
0.99	150	6	1.00	0.19	0.11	0.09	0.07	0.19	0.11	0.09	0.07
0.99	150	7	1.00	0.74	0.54	0.44	0.38	0.74	0.54	0.44	0.38
0.99	150	8	1.00	0.10	0.06	0.04	0.03	0.10	0.06	0.04	0.03
Average			0.68	0.65	0.52	0.46	0.43	0.65	0.52	0.46	0.43
Group Maximum			0.68	0.65				0.65			
0.99	200	1	0.66	0.88	1.00	0.94	0.87	0.97	1.00	0.94	0.87
0.99	200	2	0.47	1.00	0.81	0.72	0.68	0.96	0.81	0.72	0.68
0.99	200	3	0.25	1.00	0.70	0.63	0.60	0.97	0.70	0.63	0.60
0.99	200	4	0.22	1.00	0.64	0.54	0.49	0.96	0.64	0.54	0.49
0.99	200	5	1.00	0.21	0.14	0.11	0.09	0.22	0.14	0.11	0.09
0.99	200	6	1.00	0.15	0.10	0.07	0.06	0.16	0.10	0.07	0.06
0.99	200	7	1.00	0.58	0.45	0.37	0.32	0.63	0.45	0.37	0.32
0.99	200	8	1.00	0.09	0.05	0.04	0.03	0.09	0.05	0.04	0.03
Average			0.70	0.61	0.49	0.43	0.39	0.62	0.49	0.43	0.39
Group Maximum			0.70	0.61				0.62			

The following conclusions are drawn from **Table 4.6** in terms of the relative performance of the estimators over the range of distributions and sample sizes considered. Note that this table contains the **RMSE** results for the quantile with $p_0 = 0.99$.

- Distribution view
 - For $\underline{G}$, i.e. distributions with $0.2 \leq \gamma \leq 1$ and $-2 \leq \rho \leq -0.2$:
 - For $n \leq 100$ the preferred choice is the empirical multiplier $\hat{q}_{p_0,0.95,n,r}^{(EM)}$.
 - For $n \geq 150$ the preferred choice is the EVT estimator $\hat{q}_{p_0,r}^{(MOP)}$.
 - For $\underline{G}_1$, i.e. distributions with $0.67 \leq \gamma \leq 1$ and $-2 \leq \rho \leq -1$:
 - For $n \leq 100$ the preferred choice is the empirical multiplier $\hat{q}_{p_0,0.9,n,r}^{(EM)}$.
 - For $n \geq 150$ the preferred choice is the empirical multiplier $\hat{q}_{p_0,0.95,n,r}^{(EM)}$.
 - For $\underline{G}_2$, i.e. distributions with $0.5 \leq \gamma \leq 0.8$ and $-0.3 \leq \rho \leq -0.2$:
 - For $n \leq 50$ the preferred choice is the empirical multiplier $\hat{q}_{p_0,0.9,n,r}^{(EM)}$.
 - For $n \geq 100$ the preferred choice is the empirical multiplier $\hat{q}_{p_0,0.95,n,r}^{(EM)}$.
 - For $\underline{G}_3$, i.e. distributions with $0.25 \leq \gamma \leq 0.3$ and $-0.8 \leq \rho \leq -0.4$:
 - For $n \leq 50$ the preferred choice is the empirical multiplier $\hat{q}_{p_0,0.95,n,r}^{(EM)}$.

- For $n \geq 100$ the preferred choice is the EVT estimator $\hat{q}_{p_0, r}^{(MOP)}$.
- For $\underline{G}_4$, i.e. distributions with $0.25 \leq \gamma \leq 0.57$ and $\rho = -2$:
 - For $n \leq 50$ the preferred choice is the empirical multiplier $\hat{q}_{p_0, 0.95, n, r}^{(EM)}$.
 - For $n \geq 100$ the preferred choice is the EVT estimator $\hat{q}_{p_0, r}^{(MOP)}$.
- Sample size view
 - For smaller sample sizes, i.e. $20 \leq n \leq 100$:
 - For $\underline{G}$ the preferred choice is the empirical multiplier $\hat{q}_{p_0, 0.95, n, r}^{(EM)}$.
 - For $\underline{G}^{(1)}$ the preferred choice is the empirical multiplier $\hat{q}_{p_0, 0.95, n, r}^{(EM)}$.
 - For $\underline{G}^{(2)}$ the preferred choice is the empirical multiplier $\hat{q}_{p_0, 0.9, n, r}^{(EM)}$ for $n \leq 50$, otherwise use $\hat{q}_{p_0, 0.95, n, r}^{(EM)}$.
 - For $\underline{G}^{(3)}$ the preferred choice is the empirical multiplier $\hat{q}_{p_0, 0.95, n, r}^{(EM)}$ for $n \leq 50$, otherwise the EVT estimator $\hat{q}_{p_0, r}^{(MOP)}$.
 - For $\underline{G}^{(4)}$ the preferred choice is the empirical multiplier $\hat{q}_{p_0, 0.95, n, r}^{(EM)}$ for $n \leq 50$, otherwise the EVT estimator $\hat{q}_{p_0, r}^{(MOP)}$.
 - For larger sample sizes, i.e. $100 \leq n \leq 200$:
 - For $\underline{G}$ the preferred choice is the EVT estimator $\hat{q}_{p_0, r}^{(MOP)}$.
 - For $\underline{G}^{(1)}$ the preferred choice is the empirical multiplier $\hat{q}_{p_0, 0.95, n, r}^{(EM)}$.
 - For $\underline{G}^{(2)}$ the preferred choice is the empirical multiplier $\hat{q}_{p_0, 0.95, n, r}^{(EM)}$.
 - For $\underline{G}^{(3)}$ the preferred choice is the EVT estimator $\hat{q}_{p_0, r}^{(MOP)}$.
 - For $\underline{G}^{(4)}$ the preferred choice is the EVT estimator $\hat{q}_{p_0, r}^{(MOP)}$.

Table 4.7: Absolute bias results of the first order EVT simulation study for $p_0 = 0.99$

p_0	n	r	$\hat{q}_{p_0,r}^{(W)}$	$\hat{q}_{p_0,0.95,n,r}^{(EM)}$	$\hat{q}_{p_0,0.9,n,r}^{(EM)}$	$\hat{q}_{p_0,0.85,n,r}^{(EM)}$	$\hat{q}_{p_0,0.8,n,r}^{(EM)}$	$\hat{q}_{p_0,0.95,r}^{(AM)}$	$\hat{q}_{p_0,0.9,r}^{(AM)}$	$\hat{q}_{p_0,0.85,r}^{(AM)}$	$\hat{q}_{p_0,0.8,r}^{(AM)}$
0.99	20	1	0.17	0.76	0.99	1.00	0.92	0.72	0.99	1.00	0.92
0.99	20	2	0.00	0.10	0.16	0.12	0.11	1.00	0.16	0.12	0.11
0.99	20	3	0.00	0.24	0.20	0.17	0.16	1.00	0.20	0.17	0.16
0.99	20	4	0.00	0.06	0.07	0.05	0.05	1.00	0.07	0.05	0.05
0.99	20	5	0.24	1.00	0.05	0.04	0.03	0.07	0.05	0.04	0.03
0.99	20	6	0.48	1.00	0.06	0.04	0.04	0.09	0.06	0.04	0.04
0.99	20	7	0.34	1.00	0.32	0.27	0.23	0.40	0.32	0.27	0.23
0.99	20	8	1.00	0.31	0.03	0.02	0.02	0.05	0.03	0.02	0.02
Average			0.28	0.56	0.23	0.21	0.20	0.54	0.23	0.21	0.20
Group Maximum			0.28	0.56				0.54			
0.99	50	1	0.50	1.00	0.85	0.79	0.72	0.86	0.85	0.79	0.72
0.99	50	2	0.25	0.40	0.54	0.46	0.43	1.00	0.54	0.46	0.43
0.99	50	3	0.10	0.64	0.57	0.51	0.48	1.00	0.57	0.51	0.48
0.99	50	4	0.07	0.36	0.36	0.29	0.27	1.00	0.36	0.29	0.27
0.99	50	5	1.00	0.58	0.12	0.09	0.08	0.18	0.12	0.09	0.08
0.99	50	6	1.00	0.29	0.06	0.04	0.04	0.09	0.06	0.04	0.04
0.99	50	7	0.32	1.00	0.09	0.07	0.06	0.13	0.09	0.07	0.06
0.99	50	8	1.00	0.05	0.01	0.01	0.01	0.02	0.01	0.01	0.01
Average			0.53	0.54	0.33	0.28	0.26	0.53	0.33	0.28	0.26
Group Maximum			0.53	0.54				0.53			
0.99	100	1	0.95	0.85	0.89	0.80	0.73	1.00	0.89	0.80	0.73
0.99	100	2	0.75	1.00	0.59	0.52	0.48	0.90	0.59	0.52	0.48
0.99	100	3	0.25	1.00	0.62	0.56	0.53	0.96	0.62	0.56	0.53
0.99	100	4	0.16	1.00	0.42	0.35	0.32	0.89	0.42	0.35	0.32
0.99	100	5	1.00	0.14	0.10	0.08	0.06	0.15	0.10	0.08	0.06
0.99	100	6	1.00	0.07	0.04	0.03	0.03	0.07	0.04	0.03	0.03
0.99	100	7	1.00	0.18	0.13	0.11	0.09	0.20	0.13	0.11	0.09
0.99	100	8	1.00	0.01	0.01	0.01	0.00	0.01	0.01	0.01	0.00
Average			0.76	0.53	0.35	0.31	0.28	0.52	0.35	0.31	0.28
Group Maximum			0.76	0.53				0.52			

p_0	n	r	$\hat{q}_{p_0,r}^{(W)}$	$\hat{q}_{p_0,0.95,n,r}^{(EM)}$	$\hat{q}_{p_0,0.9,n,r}^{(EM)}$	$\hat{q}_{p_0,0.85,n,r}^{(EM)}$	$\hat{q}_{p_0,0.8,n,r}^{(EM)}$	$\hat{q}_{p_0,0.95,r}^{(AM)}$	$\hat{q}_{p_0,0.9,r}^{(AM)}$	$\hat{q}_{p_0,0.85,r}^{(AM)}$	$\hat{q}_{p_0,0.8,r}^{(AM)}$
0.99	150	1	1.00	0.82	0.70	0.62	0.56	0.82	0.70	0.62	0.56
0.99	150	2	1.00	0.74	0.51	0.44	0.41	0.74	0.51	0.44	0.41
0.99	150	3	0.35	1.00	0.67	0.60	0.57	1.00	0.67	0.60	0.57
0.99	150	4	0.23	1.00	0.50	0.41	0.38	1.00	0.50	0.41	0.38
0.99	150	5	1.00	0.14	0.09	0.07	0.06	0.14	0.09	0.07	0.06
0.99	150	6	1.00	0.07	0.04	0.03	0.03	0.07	0.04	0.03	0.03
0.99	150	7	1.00	0.14	0.09	0.08	0.07	0.14	0.09	0.08	0.07
0.99	150	8	1.00	0.01	0.01	0.00	0.00	0.01	0.01	0.00	0.00
Average			0.82	0.49	0.33	0.28	0.26	0.49	0.33	0.28	0.26
Group Maximum			0.82	0.49				0.49			
0.99	200	1	1.00	0.63	0.62	0.55	0.50	0.74	0.62	0.55	0.50
0.99	200	2	1.00	0.65	0.41	0.36	0.33	0.59	0.41	0.36	0.33
0.99	200	3	0.39	1.00	0.65	0.59	0.56	0.96	0.65	0.59	0.56
0.99	200	4	0.24	1.00	0.47	0.39	0.36	0.91	0.47	0.39	0.36
0.99	200	5	1.00	0.13	0.08	0.07	0.06	0.13	0.08	0.07	0.06
0.99	200	6	1.00	0.06	0.04	0.03	0.02	0.06	0.04	0.03	0.02
0.99	200	7	1.00	0.09	0.06	0.05	0.04	0.10	0.06	0.05	0.04
0.99	200	8	1.00	0.01	0.00	0.00	0.00	0.01	0.00	0.00	0.00
Average			0.83	0.44	0.29	0.25	0.23	0.44	0.29	0.25	0.23
Group Maximum			0.83	0.44				0.44			

The following conclusions are drawn from **Table 4.7** in terms of the relative performance of the estimators over the range of distributions and sample sizes considered. Note that this table contains the **absolute bias** results for the quantile with $p_0 = 0.99$.

- Distribution view
 - For $\underline{G}$, i.e. distributions with $0.2 \leq \gamma \leq 1$ and $-2 \leq \rho \leq -0.2$:
 - For $n \leq 50$ the preferred choice is the empirical multiplier $\hat{q}_{p_0,0.95,n,r}^{(EM)}$.
 - For $n \geq 100$ the preferred choice is the EVT estimator $\hat{q}_{p_0,r}^{(MOP)}$.
 - For $\underline{G}_1$, i.e. distributions with $0.67 \leq \gamma \leq 1$ and $-2 \leq \rho \leq -1$:
 - For $n \leq 100$ the preferred choice is the asymptotic multiplier $\hat{q}_{p_0,0.95,n,r}^{(AM)}$.
 - For $n \geq 100$ the preferred choice is the EVT estimator $\hat{q}_{p_0,r}^{(MOP)}$.
 - For $\underline{G}_2$, i.e. distributions with $0.5 \leq \gamma \leq 0.8$ and $-0.3 \leq \rho \leq -0.2$:
 - For $n \leq 50$ the preferred choice is the asymptotic multiplier $\hat{q}_{p_0,0.95,n,r}^{(AM)}$.
 - For $n \geq 100$ the preferred choice is the empirical multiplier $\hat{q}_{p_0,0.95,n,r}^{(EM)}$.
 - For $\underline{G}_3$, i.e. distributions with $0.25 \leq \gamma \leq 0.3$ and $-0.8 \leq \rho \leq -0.4$:

- For $n \leq 20$ the preferred choice is the empirical multiplier $\hat{q}_{p_0,0.95,n,r}^{(EM)}$.
 - For $n \geq 50$ the preferred choice is the EVT estimator $\hat{q}_{p_0,r}^{(MOP)}$.
- For $\underline{G}_4$, i.e. distributions with $0.25 \leq \gamma \leq 0.57$ and $\rho = -2$:
 - For $n \leq 20$ the preferred choice is the empirical multiplier $\hat{q}_{p_0,0.95,n,r}^{(EM)}$.
 - For $n \geq 50$ the preferred choice is the EVT estimator $\hat{q}_{p_0,r}^{(MOP)}$.
- Sample size view
 - For smaller sample sizes, i.e. $20 \leq n \leq 100$:
 - For $\underline{G}$ the preferred choice is the asymptotic multiplier $\hat{q}_{p_0,0.95,n,r}^{(AM)}$ for $n \leq 50$, otherwise the EVT estimator $\hat{q}_{p_0,r}^{(MOP)}$.
 - For $\underline{G}^{(1)}$ the preferred choice is the asymptotic multiplier $\hat{q}_{p_0,0.95,n,r}^{(AM)}$.
 - For $\underline{G}^{(2)}$ the preferred choice is the asymptotic multiplier $\hat{q}_{p_0,0.95,n,r}^{(AM)}$.
 - For $\underline{G}^{(3)}$ the preferred choice is the empirical multiplier $\hat{q}_{p_0,0.95,n,r}^{(EM)}$ for $n \leq 20$, otherwise the EVT estimator $\hat{q}_{p_0,r}^{(MOP)}$.
 - For $\underline{G}^{(4)}$ the preferred choice is the empirical multiplier $\hat{q}_{p_0,0.95,n,r}^{(EM)}$ for $n \leq 20$, otherwise the EVT estimator $\hat{q}_{p_0,r}^{(MOP)}$.
 - For larger sample sizes, i.e. $100 \leq n \leq 200$:
 - For $\underline{G}$ the preferred choice is the EVT estimator $\hat{q}_{p_0,r}^{(MOP)}$.
 - For $\underline{G}^{(1)}$ the preferred choice is the EVT estimator $\hat{q}_{p_0,r}^{(MOP)}$.
 - For $\underline{G}^{(2)}$ the preferred choice is the empirical multiplier $\hat{q}_{p_0,0.95,n,r}^{(EM)}$.
 - For $\underline{G}^{(3)}$ the preferred choice is the EVT estimator $\hat{q}_{p_0,r}^{(MOP)}$.
 - For $\underline{G}^{(4)}$ the preferred choice is the EVT estimator $\hat{q}_{p_0,r}^{(MOP)}$.

A final summary of the first order EVT simulation results are given in Section 4.1.3, but first the simulation results (from Steps 6.2 and 6.4 in Algorithm 4.1) obtained using the second order quantile estimators (from Step 4.2 in Algorithm 4.1) are presented in the following section.

4.1.2.2. Second order results

The results from the second order EVT simulation study are presented in this section. The following three quantile estimators are considered throughout this section:

- $\hat{q}_{p_0,r}^{(W+)}$: Second order EVT calculated in Step 4.2.1 in Algorithm 4.1,
- $\hat{q}_{p_0,p_1,n,r}^{(EM+)} = C_{p_0,p_1,n}^{(EM+)} \times \hat{q}_{p_1,r}^{(W+)}$: Second order multiplier using the **empirical** $C_{p_0,p_1,n}^{(EM+)}$ calculated in Step 4.2.2 in Algorithm 4.1, and
- $\hat{q}_{p_0,p_1,r}^{(AM+)} = C_{p_0,p_1,\infty}^{(AM)} \times \hat{q}_{p_1,r}^{(W+)}$: Second order multiplier using the **asymptotic** $C_{p_0,p_1,\infty}^{(AM)}$ calculated in Step 4.2.3 in Algorithm 4.1.

The second order EVT results are presented next.

Results of the second order EVT simulation study for $p_0 = 0.999$

Tables 4.8 and 4.9 contain the **RMSE** (Step 6.2 in Algorithm 4.1) and **absolute bias** (Step 6.4 in Algorithm 4.1) results respectively for $p_0 = 0.999$ in the **second order EVT** environment.

Table 4.8: Root mean squared error results of the second order EVT simulation study for $p_0 = 0.999$

p_0	n	r	$\hat{q}_{p_0,r}^{(W+)}$	$\hat{q}_{p_0,0.95,n,r}^{(EM+)}$	$\hat{q}_{p_0,0.9,n,r}^{(EM+)}$	$\hat{q}_{p_0,0.85,n,r}^{(EM+)}$	$\hat{q}_{p_0,0.8,n,r}^{(EM+)}$	$\hat{q}_{p_0,0.95,r}^{(AM+)}$	$\hat{q}_{p_0,0.9,r}^{(AM+)}$	$\hat{q}_{p_0,0.85,r}^{(AM+)}$	$\hat{q}_{p_0,0.8,r}^{(AM+)}$
0.999	20	1	0.00	1.00	0.65	0.74	0.68	0.49	0.74	0.74	0.68
0.999	20	2	0.00	0.00	0.00	0.01	1.00	0.00	0.00	0.01	1.00
0.999	20	3	0.00	0.90	1.00	0.96	0.94	0.90	0.99	0.96	0.94
0.999	20	4	0.00	0.79	1.00	0.97	0.94	0.71	1.00	0.97	0.94
0.999	20	5	0.30	1.00	0.11	0.10	0.08	0.16	0.12	0.10	0.08
0.999	20	6	0.28	1.00	0.11	0.09	0.08	0.16	0.11	0.09	0.08
0.999	20	7	0.00	0.22	0.87	1.00	0.89	0.06	0.96	1.00	0.89
0.999	20	8	0.01	1.00	0.12	0.10	0.08	0.19	0.13	0.10	0.08
Average			0.07	0.74	0.48	0.50	0.59	0.34	0.51	0.50	0.59
Group Maximum			0.07	0.74				0.59			
0.999	50	1	0.11	0.90	1.00	0.91	0.82	0.97	1.00	0.91	0.82
0.999	50	2	0.01	1.00	0.96	0.94	0.92	0.98	0.96	0.94	0.92
0.999	50	3	0.00	1.00	0.91	0.88	0.86	0.99	0.91	0.88	0.86
0.999	50	4	0.01	1.00	0.91	0.84	0.80	0.99	0.91	0.84	0.80
0.999	50	5	1.00	0.15	0.11	0.09	0.08	0.16	0.11	0.09	0.08
0.999	50	6	1.00	0.13	0.09	0.07	0.06	0.13	0.09	0.07	0.06
0.999	50	7	0.40	0.94	0.83	0.72	0.64	1.00	0.83	0.72	0.64
0.999	50	8	1.00	0.07	0.05	0.04	0.03	0.08	0.05	0.04	0.03
Average			0.44	0.65	0.61	0.56	0.53	0.66	0.61	0.56	0.53
Group Maximum			0.44	0.65				0.66			
0.999	100	1	0.32	1.00	0.91	0.80	0.72	1.00	0.91	0.80	0.72
0.999	100	2	0.10	1.00	0.94	0.92	0.90	1.00	0.94	0.92	0.90
0.999	100	3	0.03	1.00	0.92	0.89	0.87	1.00	0.92	0.89	0.87
0.999	100	4	0.05	1.00	0.85	0.78	0.74	1.00	0.85	0.78	0.74
0.999	100	5	1.00	0.09	0.06	0.05	0.04	0.09	0.06	0.05	0.04
0.999	100	6	1.00	0.08	0.06	0.05	0.04	0.08	0.06	0.05	0.04
0.999	100	7	1.00	0.80	0.63	0.54	0.48	0.80	0.63	0.54	0.48
0.999	100	8	1.00	0.05	0.03	0.02	0.02	0.05	0.03	0.02	0.02
Average			0.56	0.63	0.55	0.51	0.48	0.63	0.55	0.51	0.48
Group Maximum			0.56	0.63				0.63			

p_0	n	r	$\hat{q}_{p_0,r}^{(W+)}$	$\hat{q}_{p_0,0.95,n,r}^{(EM+)}$	$\hat{q}_{p_0,0.9,n,r}^{(EM+)}$	$\hat{q}_{p_0,0.85,n,r}^{(EM+)}$	$\hat{q}_{p_0,0.8,n,r}^{(EM+)}$	$\hat{q}_{p_0,0.95,r}^{(AM+)}$	$\hat{q}_{p_0,0.9,r}^{(AM+)}$	$\hat{q}_{p_0,0.85,r}^{(AM+)}$	$\hat{q}_{p_0,0.8,r}^{(AM+)}$
0.999	150	1	0.48	0.91	0.87	0.76	0.68	1.00	0.87	0.76	0.68
0.999	150	2	0.19	1.00	0.91	0.88	0.87	0.97	0.91	0.88	0.87
0.999	150	3	0.10	1.00	0.91	0.88	0.86	0.99	0.91	0.88	0.86
0.999	150	4	0.09	1.00	0.79	0.72	0.68	0.96	0.79	0.72	0.68
0.999	150	5	1.00	0.07	0.05	0.04	0.03	0.07	0.05	0.04	0.03
0.999	150	6	1.00	0.06	0.04	0.03	0.03	0.06	0.04	0.03	0.03
0.999	150	7	1.00	0.50	0.42	0.36	0.32	0.54	0.42	0.36	0.32
0.999	150	8	1.00	0.04	0.02	0.02	0.02	0.04	0.02	0.02	0.02
Average			0.61	0.57	0.50	0.46	0.44	0.58	0.50	0.46	0.44
Group Maximum			0.61	0.57				0.58			
0.999	200	1	0.62	0.90	0.83	0.71	0.64	1.00	0.83	0.71	0.64
0.999	200	2	0.34	1.00	0.91	0.88	0.86	0.97	0.91	0.88	0.86
0.999	200	3	0.18	1.00	0.91	0.88	0.86	0.99	0.91	0.88	0.86
0.999	200	4	0.14	1.00	0.77	0.70	0.67	0.95	0.77	0.70	0.67
0.999	200	5	1.00	0.05	0.04	0.03	0.03	0.06	0.04	0.03	0.03
0.999	200	6	1.00	0.05	0.04	0.03	0.02	0.05	0.04	0.03	0.02
0.999	200	7	1.00	0.43	0.36	0.30	0.27	0.46	0.36	0.30	0.27
0.999	200	8	1.00	0.03	0.02	0.02	0.01	0.03	0.02	0.02	0.01
Average			0.66	0.56	0.48	0.44	0.42	0.56	0.48	0.44	0.42
Group Maximum			0.66	0.56				0.56			

The following conclusions are drawn from **Table 4.8** in terms of the relative performance of the estimators over the range of distributions and sample sizes considered. Note that this table contains the **RMSE** results for the quantile with $p_0 = 0.999$.

- Distribution view
 - For $\underline{G}$, i.e. distributions with $0.2 \leq \gamma \leq 1$ and $-2 \leq \rho \leq -0.2$:
 - For $n \leq 100$ the preferred choice is the empirical multiplier $\hat{q}_{p_0,0.95,n,r}^{(EM+)}$.
 - For $n \geq 150$ the preferred choice is the EVT estimator $\hat{q}_{p_0,r}^{(RBMOP)}$.
 - For $\underline{G}_1$, i.e. distributions with $0.67 \leq \gamma \leq 1$ and $-2 \leq \rho \leq -1$:
 - For any n the preferred choice is the empirical multiplier $\hat{q}_{p_0,0.95,n,r}^{(EM+)}$.
 - For $\underline{G}_2$, i.e. distributions with $0.5 \leq \gamma \leq 0.8$ and $-0.3 \leq \rho \leq -0.2$:
 - For any n the preferred choice is the empirical multiplier $\hat{q}_{p_0,0.95,n,r}^{(EM+)}$.
 - For $\underline{G}_3$, i.e. distributions with $0.25 \leq \gamma \leq 0.3$ and $-0.8 \leq \rho \leq -0.4$:
 - For $n \leq 20$ the preferred choice is the empirical multiplier $\hat{q}_{p_0,0.95,n,r}^{(EM+)}$.
 - For $n \geq 50$ the preferred choice is the EVT estimator $\hat{q}_{p_0,r}^{(RBMOP)}$.

- For $\underline{G}_4$, i.e. distributions with $0.25 \leq \gamma \leq 0.57$ and $\rho = -2$:
 - For $n \leq 20$ the preferred choice is the empirical multiplier $\hat{q}_{p_0, 0.95, n, r}^{(EM+)}$.
 - For $n \geq 50$ the preferred choice is the EVT estimator $\hat{q}_{p_0, r}^{(RBMOP)}$.
- Sample size view
 - For smaller sample sizes, i.e. $20 \leq n \leq 100$:
 - For $\underline{G}$ the preferred choice is the empirical multiplier $\hat{q}_{p_0, 0.95, n, r}^{(EM+)}$.
 - For $\underline{G}^{(1)}$ the preferred choice is the empirical multiplier $\hat{q}_{p_0, 0.8, n, r}^{(EM+)}$ for $n \leq 20$, otherwise the empirical multiplier $\hat{q}_{p_0, 0.95, n, r}^{(EM+)}$.
 - For $\underline{G}^{(2)}$ the preferred choice is the empirical multiplier $\hat{q}_{p_0, 0.95, n, r}^{(EM+)}$.
 - For $\underline{G}^{(3)}$ the preferred choice is the empirical multiplier $\hat{q}_{p_0, 0.95, n, r}^{(EM+)}$ for $n \leq 20$, otherwise the EVT estimator $\hat{q}_{p_0, r}^{(RBMOP)}$.
 - For $\underline{G}^{(4)}$ the preferred choice is the empirical multiplier $\hat{q}_{p_0, 0.95, n, r}^{(EM+)}$ for $n \leq 20$, otherwise the EVT estimator $\hat{q}_{p_0, r}^{(RBMOP)}$.
 - For larger sample sizes, i.e. $100 \leq n \leq 200$:
 - For $\underline{G}$ the preferred choice is the empirical multiplier $\hat{q}_{p_0, 0.95, n, r}^{(EM+)}$ or the EVT estimator $\hat{q}_{p_0, r}^{(RBMOP)}$.
 - For $\underline{G}^{(1)}$ the preferred choice is the empirical multiplier $\hat{q}_{p_0, 0.95, n, r}^{(EM+)}$.
 - For $\underline{G}^{(2)}$ the preferred choice is the empirical multiplier $\hat{q}_{p_0, 0.95, n, r}^{(EM+)}$.
 - For $\underline{G}^{(3)}$ the preferred choice is the EVT estimator $\hat{q}_{p_0, r}^{(RBMOP)}$.
 - For $\underline{G}^{(4)}$ the preferred choice is the EVT estimator $\hat{q}_{p_0, r}^{(RBMOP)}$.

Table 4.9: Absolute bias results of the second order EVT simulation study for

$p_0 = 0.999$

p_0	n	r	$\hat{q}_{p_0,r}^{(W+)}$	$\hat{q}_{p_0,0.95,n,r}^{(EM+)}$	$\hat{q}_{p_0,0.9,n,r}^{(EM+)}$	$\hat{q}_{p_0,0.85,n,r}^{(EM+)}$	$\hat{q}_{p_0,0.8,n,r}^{(EM+)}$	$\hat{q}_{p_0,0.95,r}^{(AM+)}$	$\hat{q}_{p_0,0.9,r}^{(AM+)}$	$\hat{q}_{p_0,0.85,r}^{(AM+)}$	$\hat{q}_{p_0,0.8,r}^{(AM+)}$
0.999	20	1	0.02	0.95	0.78	0.85	0.75	1.00	0.97	0.85	0.75
0.999	20	2	0.00	0.00	0.00	0.58	1.00	0.00	0.00	0.58	1.00
0.999	20	3	0.00	0.74	0.84	0.79	0.78	1.00	0.83	0.79	0.78
0.999	20	4	0.00	0.33	0.59	0.47	0.45	1.00	0.54	0.47	0.45
0.999	20	5	1.00	0.87	0.08	0.07	0.06	0.12	0.09	0.07	0.06
0.999	20	6	1.00	0.59	0.05	0.05	0.04	0.08	0.06	0.05	0.04
0.999	20	7	0.00	1.00	0.20	0.19	0.17	0.20	0.22	0.19	0.17
0.999	20	8	1.00	0.92	0.10	0.08	0.07	0.17	0.11	0.08	0.07
Average			0.38	0.67	0.33	0.39	0.41	0.45	0.35	0.39	0.41
Group Maximum			0.38	0.67				0.45			
0.999	50	1	0.58	0.88	0.79	0.67	0.59	1.00	0.79	0.67	0.59
0.999	50	2	0.11	1.00	0.86	0.83	0.81	0.97	0.86	0.83	0.81
0.999	50	3	0.03	1.00	0.88	0.85	0.83	0.99	0.88	0.85	0.83
0.999	50	4	0.04	1.00	0.65	0.58	0.55	0.93	0.65	0.58	0.55
0.999	50	5	1.00	0.04	0.03	0.03	0.02	0.04	0.03	0.03	0.02
0.999	50	6	1.00	0.01	0.01	0.00	0.00	0.01	0.01	0.00	0.00
0.999	50	7	1.00	0.30	0.24	0.20	0.18	0.32	0.24	0.20	0.18
0.999	50	8	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Average			0.59	0.53	0.43	0.40	0.37	0.53	0.43	0.40	0.37
Group Maximum			0.59	0.53				0.53			
0.999	100	1	1.00	0.40	0.30	0.26	0.23	0.40	0.30	0.26	0.23
0.999	100	2	0.60	1.00	0.91	0.88	0.86	1.00	0.91	0.88	0.86
0.999	100	3	0.18	1.00	0.91	0.88	0.86	1.00	0.91	0.88	0.86
0.999	100	4	0.12	1.00	0.74	0.66	0.63	1.00	0.74	0.66	0.63
0.999	100	5	1.00	0.03	0.02	0.02	0.01	0.03	0.02	0.02	0.01
0.999	100	6	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.999	100	7	1.00	0.07	0.05	0.04	0.04	0.07	0.05	0.04	0.04
0.999	100	8	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Average			0.74	0.44	0.37	0.34	0.33	0.44	0.37	0.34	0.33
Group Maximum			0.74	0.44				0.44			

p_0	n	r	$\hat{q}_{p_0,r}^{(W+)}$	$\hat{q}_{p_0,0.95,n,r}^{(EM+)}$	$\hat{q}_{p_0,0.9,n,r}^{(EM+)}$	$\hat{q}_{p_0,0.85,n,r}^{(EM+)}$	$\hat{q}_{p_0,0.8,n,r}^{(EM+)}$	$\hat{q}_{p_0,0.95,r}^{(AM+)}$	$\hat{q}_{p_0,0.9,r}^{(AM+)}$	$\hat{q}_{p_0,0.85,r}^{(AM+)}$	$\hat{q}_{p_0,0.8,r}^{(AM+)}$
0.999	150	1	1.00	0.08	0.07	0.06	0.05	0.09	0.07	0.06	0.05
0.999	150	2	1.00	0.78	0.69	0.67	0.66	0.76	0.69	0.67	0.66
0.999	150	3	0.35	1.00	0.90	0.87	0.86	0.99	0.90	0.87	0.86
0.999	150	4	0.17	1.00	0.70	0.63	0.60	0.93	0.70	0.63	0.60
0.999	150	5	1.00	0.02	0.01	0.01	0.01	0.02	0.01	0.01	0.01
0.999	150	6	1.00	0.01	0.01	0.00	0.00	0.01	0.01	0.00	0.00
0.999	150	7	1.00	0.02	0.02	0.01	0.01	0.02	0.02	0.01	0.01
0.999	150	8	1.00	0.01	0.00	0.00	0.00	0.01	0.00	0.00	0.00
Average			0.82	0.36	0.30	0.28	0.27	0.35	0.30	0.28	0.27
Group Maximum			0.82	0.36				0.35			
0.999	200	1	1.00	0.11	0.09	0.08	0.07	0.12	0.09	0.08	0.07
0.999	200	2	1.00	0.40	0.35	0.34	0.33	0.38	0.35	0.34	0.33
0.999	200	3	0.46	1.00	0.91	0.87	0.86	0.99	0.91	0.87	0.86
0.999	200	4	0.23	1.00	0.72	0.65	0.61	0.94	0.72	0.65	0.61
0.999	200	5	1.00	0.02	0.01	0.01	0.01	0.02	0.01	0.01	0.01
0.999	200	6	1.00	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
0.999	200	7	1.00	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
0.999	200	8	1.00	0.01	0.00	0.00	0.00	0.01	0.00	0.00	0.00
Average			0.84	0.32	0.26	0.25	0.24	0.31	0.26	0.25	0.24
Group Maximum			0.84	0.32				0.31			

The following conclusions are drawn from **Table 4.9** in terms of the relative performance of the estimators over the range of distributions and sample sizes considered. Note that this table contains the **absolute bias** results for the quantile with $p_0 = 0.999$.

- Distribution view

- For $\underline{G}$, i.e. distributions with $0.2 \leq \gamma \leq 1$ and $-2 \leq \rho \leq -0.2$:
 - For $n \leq 20$ the preferred choice is the empirical multiplier $\hat{q}_{p_0,0.95,n,r}^{(EM+)}$.
 - For $n \geq 50$ the preferred choice is the EVT estimator $\hat{q}_{p_0,r}^{(RBMOP)}$.
- For $\underline{G}_1$, i.e. distributions with $0.67 \leq \gamma \leq 1$ and $-2 \leq \rho \leq -1$:
 - For $n \leq 50$ the preferred choice is the asymptotic multiplier $\hat{q}_{p_0,0.95,n,r}^{(AM+)}$.
 - For $n \geq 100$ the preferred choice is the EVT estimator $\hat{q}_{p_0,r}^{(RBMOP)}$.
- For $\underline{G}_2$, i.e. distributions with $0.5 \leq \gamma \leq 0.8$ and $-0.3 \leq \rho \leq -0.2$:
 - For $n \leq 50$ the preferred choice is the asymptotic multiplier $\hat{q}_{p_0,0.95,n,r}^{(AM+)}$.
 - For $n \geq 100$ the preferred choice is the empirical multiplier $\hat{q}_{p_0,0.95,n,r}^{(EM+)}$.
- For $\underline{G}_3$, i.e. distributions with $0.25 \leq \gamma \leq 0.3$ and $-0.8 \leq \rho \leq -0.4$:
 - For any n the preferred choice is the EVT estimator $\hat{q}_{p_0,r}^{(RBMOP)}$.

- For $\underline{G}_4$, i.e. distributions with $0.25 \leq \gamma \leq 0.57$ and $\rho = -2$:
 - For $n \leq 20$ the preferred choice is the empirical multiplier $\hat{q}_{p_0, 0.95, n, r}^{(EM+)}$.
 - For $n \geq 50$ the preferred choice is the EVT estimator $\hat{q}_{p_0, r}^{(RBMOP)}$.
- Sample size view
 - For smaller sample sizes, i.e. $20 \leq n \leq 100$:
 - For $\underline{G}$ the preferred choice is the empirical multiplier $\hat{q}_{p_0, 0.95, n, r}^{(EM+)}$ for $n \leq 20$, otherwise the EVT estimator $\hat{q}_{p_0, r}^{(RBMOP)}$.
 - For $\underline{G}^{(1)}$ the preferred choice is the asymptotic multiplier $\hat{q}_{p_0, 0.95, n, r}^{(AM+)}$.
 - For $\underline{G}^{(2)}$ the preferred choice is the empirical multiplier $\hat{q}_{p_0, 0.95, n, r}^{(EM+)}$.
 - For $\underline{G}^{(3)}$ the preferred choice is the EVT estimator $\hat{q}_{p_0, r}^{(RBMOP)}$.
 - For $\underline{G}^{(4)}$ the preferred choice is the empirical multiplier $\hat{q}_{p_0, 0.95, n, r}^{(EM+)}$ for $n \leq 20$, otherwise the EVT estimator $\hat{q}_{p_0, r}^{(RBMOP)}$.
 - For larger sample sizes, i.e. $100 \leq n \leq 200$:
 - For $\underline{G}$ the preferred choice is the EVT estimator $\hat{q}_{p_0, r}^{(RBMOP)}$.
 - For $\underline{G}^{(1)}$ the preferred choice is the EVT estimator $\hat{q}_{p_0, r}^{(RBMOP)}$.
 - For $\underline{G}^{(2)}$ the preferred choice is the empirical multiplier $\hat{q}_{p_0, 0.95, n, r}^{(EM+)}$.
 - For $\underline{G}^{(3)}$ the preferred choice is the EVT estimator $\hat{q}_{p_0, r}^{(RBMOP)}$.
 - For $\underline{G}^{(4)}$ the preferred choice is the EVT estimator $\hat{q}_{p_0, r}^{(RBMOP)}$.

Results of the second order EVT simulation study for $p_0 = 0.995$

Tables 4.10 and 4.11 contain the RMSE (Step 6.2 in Algorithm 4.1) and absolute bias (Step 6.4 in Algorithm 4.1) results respectively for $p_0 = 0.995$ in the second order EVT environment.

Table 4.10: Root mean squared error results of the second order EVT simulation study for $p_0 = 0.995$

p_0	n	r	$\hat{q}_{p_0,r}^{(W+)}$	$\hat{q}_{p_0,0.95,n,r}^{(EM+)}$	$\hat{q}_{p_0,0.9,n,r}^{(EM+)}$	$\hat{q}_{p_0,0.85,n,r}^{(EM+)}$	$\hat{q}_{p_0,0.8,n,r}^{(EM+)}$	$\hat{q}_{p_0,0.95,r}^{(AM+)}$	$\hat{q}_{p_0,0.9,r}^{(AM+)}$	$\hat{q}_{p_0,0.85,r}^{(AM+)}$	$\hat{q}_{p_0,0.8,r}^{(AM+)}$
0.995	20	1	0.00	0.00	0.18	1.00	0.95	0.00	0.18	1.00	0.95
0.995	20	2	0.00	0.01	0.52	1.00	1.00	0.01	0.52	1.00	1.00
0.995	20	3	0.00	0.92	1.00	0.96	0.94	0.83	1.00	0.96	0.94
0.995	20	4	0.00	0.87	1.00	0.97	0.94	0.84	1.00	0.97	0.94
0.995	20	5	0.49	1.00	0.19	0.15	0.13	0.28	0.19	0.15	0.13
0.995	20	6	0.04	1.00	0.21	0.16	0.14	0.31	0.21	0.16	0.14
0.995	20	7	0.00	0.79	1.00	0.91	0.79	0.39	1.00	0.91	0.79
0.995	20	8	0.01	1.00	0.27	0.21	0.17	0.37	0.27	0.21	0.17
Average			0.07	0.70	0.55	0.67	0.63	0.38	0.55	0.67	0.63
Group Maximum			0.07	0.70				0.67			
0.995	50	1	0.28	0.83	0.83	0.92	0.81	0.88	1.00	0.92	0.81
0.995	50	2	0.07	0.88	1.00	0.92	0.90	0.87	0.94	0.92	0.90
0.995	50	3	0.05	1.00	0.90	0.85	0.83	0.99	0.89	0.85	0.83
0.995	50	4	0.09	1.00	0.94	0.84	0.81	0.98	0.90	0.84	0.81
0.995	50	5	1.00	0.22	0.13	0.12	0.10	0.23	0.15	0.12	0.10
0.995	50	6	1.00	0.21	0.12	0.11	0.09	0.22	0.14	0.11	0.09
0.995	50	7	0.68	0.93	0.71	0.69	0.60	1.00	0.82	0.69	0.60
0.995	50	8	1.00	0.13	0.08	0.06	0.05	0.14	0.08	0.06	0.05
Average			0.52	0.65	0.59	0.56	0.52	0.66	0.61	0.56	0.52
Group Maximum			0.52	0.65				0.66			
0.995	100	1	0.47	0.92	1.00	0.88	0.76	1.00	1.00	0.88	0.76
0.995	100	2	0.20	1.00	0.93	0.89	0.86	0.97	0.93	0.89	0.86
0.995	100	3	0.24	1.00	0.88	0.84	0.82	0.99	0.88	0.84	0.82
0.995	100	4	0.30	1.00	0.84	0.78	0.75	0.97	0.84	0.78	0.75
0.995	100	5	1.00	0.15	0.10	0.08	0.07	0.16	0.10	0.08	0.07
0.995	100	6	1.00	0.14	0.09	0.07	0.06	0.15	0.09	0.07	0.06
0.995	100	7	1.00	0.84	0.69	0.57	0.49	0.91	0.69	0.57	0.49
0.995	100	8	1.00	0.09	0.06	0.04	0.04	0.09	0.06	0.04	0.04
Average			0.65	0.64	0.57	0.52	0.48	0.66	0.57	0.52	0.48
Group Maximum			0.65	0.64				0.66			

p_0	n	r	$\hat{q}_{p_0,r}^{(W+)}$	$\hat{q}_{p_0,0.95,n,r}^{(EM+)}$	$\hat{q}_{p_0,0.9,n,r}^{(EM+)}$	$\hat{q}_{p_0,0.85,n,r}^{(EM+)}$	$\hat{q}_{p_0,0.8,n,r}^{(EM+)}$	$\hat{q}_{p_0,0.95,r}^{(AM+)}$	$\hat{q}_{p_0,0.9,r}^{(AM+)}$	$\hat{q}_{p_0,0.85,r}^{(AM+)}$	$\hat{q}_{p_0,0.8,r}^{(AM+)}$
0.995	150	1	0.52	0.91	0.90	0.75	0.65	1.00	0.90	0.75	0.65
0.995	150	2	0.37	1.00	0.89	0.85	0.82	0.97	0.89	0.85	0.82
0.995	150	3	0.42	1.00	0.88	0.84	0.82	0.99	0.88	0.84	0.82
0.995	150	4	0.48	1.00	0.82	0.76	0.73	0.97	0.82	0.76	0.73
0.995	150	5	1.00	0.12	0.08	0.06	0.05	0.12	0.08	0.06	0.05
0.995	150	6	1.00	0.12	0.08	0.06	0.05	0.13	0.08	0.06	0.05
0.995	150	7	1.00	0.65	0.51	0.42	0.37	0.71	0.51	0.42	0.37
0.995	150	8	1.00	0.07	0.05	0.03	0.03	0.08	0.05	0.03	0.03
Average			0.72	0.61	0.53	0.47	0.44	0.62	0.53	0.47	0.44
Group Maximum			0.72	0.61				0.62			
0.995	200	1	0.58	0.89	0.84	0.69	0.59	1.00	0.84	0.69	0.59
0.995	200	2	0.50	1.00	0.88	0.83	0.80	0.96	0.88	0.83	0.80
0.995	200	3	0.56	1.00	0.88	0.84	0.82	0.99	0.88	0.84	0.82
0.995	200	4	0.63	1.00	0.82	0.76	0.73	0.97	0.82	0.76	0.73
0.995	200	5	1.00	0.10	0.07	0.05	0.04	0.11	0.07	0.05	0.04
0.995	200	6	1.00	0.10	0.07	0.05	0.04	0.11	0.07	0.05	0.04
0.995	200	7	1.00	0.54	0.42	0.34	0.30	0.59	0.42	0.34	0.30
0.995	200	8	1.00	0.06	0.04	0.03	0.03	0.07	0.04	0.03	0.03
Average			0.78	0.59	0.50	0.45	0.42	0.60	0.50	0.45	0.42
Group Maximum			0.78	0.59				0.60			

The following conclusions are drawn from **Table 4.10** in terms of the relative performance of the estimators over the range of distributions and sample sizes considered. Note that this table contains the **RMSE** results for the quantile with $p_0 = 0.995$.

- Distribution view
 - For $\underline{G}$, i.e. distributions with $0.2 \leq \gamma \leq 1$ and $-2 \leq \rho \leq -0.2$:
 - For $n \leq 100$ the preferred choice is the asymptotic multiplier $\hat{q}_{p_0,0.95,n,r}^{(AM+)}$.
 - For $n \geq 150$ the preferred choice is the EVT estimator $\hat{q}_{p_0,r}^{(RBMOP)}$.
 - For $\underline{G}_1$, i.e. distributions with $0.67 \leq \gamma \leq 1$ and $-2 \leq \rho \leq -1$:
 - For $n \leq 50$ the preferred choice is the asymptotic multiplier $\hat{q}_{p_0,0.85,n,r}^{(AM+)}$.
 - For $n \geq 100$ the preferred choice is the empirical multiplier $\hat{q}_{p_0,0.95,n,r}^{(EM+)}$.
 - For $\underline{G}_2$, i.e. distributions with $0.5 \leq \gamma \leq 0.8$ and $-0.3 \leq \rho \leq -0.2$:
 - For $n \leq 20$ the preferred choice is the empirical multiplier $\hat{q}_{p_0,0.9,n,r}^{(EM+)}$.
 - For $n \geq 50$ the preferred choice is the empirical multiplier $\hat{q}_{p_0,0.95,n,r}^{(EM+)}$.
 - For $\underline{G}_3$, i.e. distributions with $0.25 \leq \gamma \leq 0.3$ and $-0.8 \leq \rho \leq -0.4$:

- For $n \leq 20$ the preferred choice is the empirical multiplier $\hat{q}_{p_0, 0.95, n, r}^{(EM+)}$.
 - For $n \geq 50$ the preferred choice is the EVT estimator $\hat{q}_{p_0, r}^{(RBMOP)}$.
- For $\underline{G}_4$, i.e. distributions with $0.25 \leq \gamma \leq 0.57$ and $\rho = -2$:
 - For $n \leq 20$ the preferred choice is the empirical multiplier $\hat{q}_{p_0, 0.95, n, r}^{(EM+)}$.
 - For $n \geq 50$ the preferred choice is the EVT estimator $\hat{q}_{p_0, r}^{(RBMOP)}$.
- Sample size view
 - For smaller sample sizes, i.e. $20 \leq n \leq 100$:
 - For $\underline{G}$ the preferred choice is the asymptotic multiplier $\hat{q}_{p_0, 0.95, n, r}^{(AM+)}$.
 - For $\underline{G}^{(1)}$ the preferred choice is the asymptotic multiplier $\hat{q}_{p_0, 0.85, n, r}^{(AM+)}$ for $n \leq 20$, otherwise the asymptotic multiplier $\hat{q}_{p_0, 0.9, n, r}^{(AM+)}$.
 - For $\underline{G}^{(2)}$ the preferred choice is the empirical multiplier $\hat{q}_{p_0, 0.9, n, r}^{(EM+)}$ for $n \leq 20$, otherwise $\hat{q}_{p_0, 0.95, n, r}^{(EM+)}$.
 - For $\underline{G}^{(3)}$ the preferred choice is the empirical multiplier $\hat{q}_{p_0, 0.95, n, r}^{(EM+)}$ for $n \leq 20$, otherwise the EVT estimator $\hat{q}_{p_0, r}^{(RBMOP)}$.
 - For $\underline{G}^{(4)}$ the preferred choice is the empirical multiplier $\hat{q}_{p_0, 0.95, n, r}^{(EM+)}$ for $n \leq 20$, otherwise the EVT estimator $\hat{q}_{p_0, r}^{(RBMOP)}$.
 - For larger sample sizes, i.e. $100 \leq n \leq 200$:
 - For $\underline{G}$ the preferred choice is the EVT estimator $\hat{q}_{p_0, r}^{(RBMOP)}$.
 - For $\underline{G}^{(1)}$ the preferred choice is the empirical multiplier $\hat{q}_{p_0, 0.95, n, r}^{(EM+)}$.
 - For $\underline{G}^{(2)}$ the preferred choice is the empirical multiplier $\hat{q}_{p_0, 0.95, n, r}^{(EM+)}$.
 - For $\underline{G}^{(3)}$ the preferred choice is the EVT estimator $\hat{q}_{p_0, r}^{(RBMOP)}$.
 - For $\underline{G}^{(4)}$ the preferred choice is the EVT estimator $\hat{q}_{p_0, r}^{(RBMOP)}$.

Table 4.11: Absolute bias results of the second order EVT simulation study for $p_0 = 0.995$

p_0	n	r	$\hat{q}_{p_0,r}^{(W+)}$	$\hat{q}_{p_0,0.95,n,r}^{(EM+)}$	$\hat{q}_{p_0,0.9,n,r}^{(EM+)}$	$\hat{q}_{p_0,0.85,n,r}^{(EM+)}$	$\hat{q}_{p_0,0.8,n,r}^{(EM+)}$	$\hat{q}_{p_0,0.95,r}^{(AM+)}$	$\hat{q}_{p_0,0.9,r}^{(AM+)}$	$\hat{q}_{p_0,0.85,r}^{(AM+)}$	$\hat{q}_{p_0,0.8,r}^{(AM+)}$
0.995	20	1	0.00	0.11	1.00	0.91	0.75	0.06	1.00	0.91	0.75
0.995	20	2	0.00	0.41	1.00	0.87	0.85	0.23	1.00	0.87	0.85
0.995	20	3	0.00	0.74	0.76	0.72	0.70	1.00	0.76	0.72	0.70
0.995	20	4	0.03	0.54	0.67	0.61	0.59	1.00	0.67	0.61	0.59
0.995	20	5	1.00	0.42	0.06	0.04	0.04	0.09	0.06	0.04	0.04
0.995	20	6	1.00	0.64	0.09	0.07	0.06	0.14	0.09	0.07	0.06
0.995	20	7	0.00	1.00	0.19	0.15	0.13	0.24	0.19	0.15	0.13
0.995	20	8	0.33	1.00	0.18	0.13	0.11	0.29	0.18	0.13	0.11
Average			0.30	0.61	0.49	0.44	0.40	0.38	0.49	0.44	0.40
Group Maximum			0.30	0.61				0.49			
0.995	50	1	1.00	0.56	0.32	0.34	0.28	0.70	0.44	0.34	0.28
0.995	50	2	0.37	1.00	0.83	0.71	0.69	0.94	0.75	0.71	0.69
0.995	50	3	0.37	1.00	0.86	0.80	0.78	0.98	0.84	0.80	0.78
0.995	50	4	0.40	1.00	0.79	0.68	0.66	0.95	0.74	0.68	0.66
0.995	50	5	1.00	0.01	0.01	0.00	0.00	0.01	0.01	0.00	0.00
0.995	50	6	1.00	0.02	0.01	0.01	0.01	0.02	0.01	0.01	0.01
0.995	50	7	1.00	0.15	0.10	0.09	0.08	0.17	0.11	0.09	0.08
0.995	50	8	1.00	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
Average			0.77	0.47	0.36	0.33	0.31	0.47	0.37	0.33	0.31
Group Maximum			0.77	0.47				0.47			
0.995	100	1	1.00	0.14	0.11	0.08	0.07	0.17	0.11	0.08	0.07
0.995	100	2	1.00	0.66	0.52	0.49	0.47	0.62	0.52	0.49	0.47
0.995	100	3	1.00	0.83	0.72	0.68	0.66	0.82	0.72	0.68	0.66
0.995	100	4	1.00	0.98	0.75	0.69	0.66	0.93	0.75	0.69	0.66
0.995	100	5	1.00	0.01	0.00	0.00	0.00	0.01	0.00	0.00	0.00
0.995	100	6	1.00	0.03	0.02	0.01	0.01	0.03	0.02	0.01	0.01
0.995	100	7	1.00	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
0.995	100	8	1.00	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
Average			1.00	0.33	0.27	0.25	0.24	0.33	0.27	0.25	0.24
Group Maximum			1.00	0.33				0.33			

p_0	n	r	$\hat{q}_{p_0,r}^{(W+)}$	$\hat{q}_{p_0,0.95,n,r}^{(EM+)}$	$\hat{q}_{p_0,0.9,n,r}^{(EM+)}$	$\hat{q}_{p_0,0.85,n,r}^{(EM+)}$	$\hat{q}_{p_0,0.8,n,r}^{(EM+)}$	$\hat{q}_{p_0,0.95,r}^{(AM+)}$	$\hat{q}_{p_0,0.9,r}^{(AM+)}$	$\hat{q}_{p_0,0.85,r}^{(AM+)}$	$\hat{q}_{p_0,0.8,r}^{(AM+)}$
0.995	150	1	1.00	0.44	0.32	0.25	0.21	0.56	0.32	0.25	0.21
0.995	150	2	1.00	0.24	0.19	0.18	0.17	0.22	0.19	0.18	0.17
0.995	150	3	1.00	0.39	0.34	0.32	0.32	0.38	0.34	0.32	0.32
0.995	150	4	1.00	0.62	0.48	0.44	0.42	0.59	0.48	0.44	0.42
0.995	150	5	1.00	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
0.995	150	6	1.00	0.04	0.02	0.02	0.02	0.04	0.02	0.02	0.02
0.995	150	7	1.00	0.04	0.03	0.03	0.02	0.05	0.03	0.03	0.02
0.995	150	8	1.00	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
Average			1.00	0.22	0.18	0.16	0.15	0.23	0.18	0.16	0.15
Group Maximum			1.00	0.22				0.23			
0.995	200	1	1.00	0.61	0.43	0.33	0.28	0.77	0.43	0.33	0.28
0.995	200	2	1.00	0.09	0.07	0.07	0.06	0.08	0.07	0.07	0.06
0.995	200	3	1.00	0.25	0.22	0.21	0.21	0.25	0.22	0.21	0.21
0.995	200	4	1.00	0.42	0.33	0.30	0.29	0.40	0.33	0.30	0.29
0.995	200	5	1.00	0.02	0.01	0.01	0.01	0.02	0.01	0.01	0.01
0.995	200	6	1.00	0.04	0.03	0.02	0.02	0.04	0.03	0.02	0.02
0.995	200	7	1.00	0.07	0.05	0.04	0.04	0.08	0.05	0.04	0.04
0.995	200	8	1.00	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
Average			1.00	0.19	0.14	0.12	0.11	0.21	0.14	0.12	0.11
Group Maximum			1.00	0.19				0.21			

The following conclusions are drawn from **Table 4.11** in terms of the relative performance of the estimators over the range of distributions and sample sizes considered. Note that this table contains the **absolute bias** results for the quantile with $p_0 = 0.995$.

- Distribution view

- For $\underline{G}$, i.e. distributions with $0.2 \leq \gamma \leq 1$ and $-2 \leq \rho \leq -0.2$:
 - For $n \leq 20$ the preferred choice is the empirical multiplier $\hat{q}_{p_0,0.95,n,r}^{(EM+)}$.
 - For $n \geq 50$ the preferred choice is the EVT estimator $\hat{q}_{p_0,r}^{(RBMOP)}$.
- For $\underline{G}_1$, i.e. distributions with $0.67 \leq \gamma \leq 1$ and $-2 \leq \rho \leq -1$:
 - For $n \leq 50$ the preferred choice is the asymptotic multiplier $\hat{q}_{p_0,0.9,n,r}^{(AM+)}$.
 - For $n \geq 100$ the preferred choice is the EVT estimator $\hat{q}_{p_0,r}^{(RBMOP)}$.
- For $\underline{G}_2$, i.e. distributions with $0.5 \leq \gamma \leq 0.8$ and $-0.3 \leq \rho \leq -0.2$:
 - For $n \leq 50$ the preferred choice is the asymptotic multiplier $\hat{q}_{p_0,0.95,n,r}^{(AM+)}$.
 - For $n \geq 100$ the preferred choice is the EVT estimator $\hat{q}_{p_0,r}^{(RBMOP)}$.
- For $\underline{G}_3$, i.e. distributions with $0.25 \leq \gamma \leq 0.3$ and $-0.8 \leq \rho \leq -0.4$:
 - For any n the preferred choice is the EVT estimator $\hat{q}_{p_0,r}^{(RBMOP)}$.

- For $\underline{G}_4$, i.e. distributions with $0.25 \leq \gamma \leq 0.57$ and $\rho = -2$:
 - For $n \leq 20$ the preferred choice is the empirical multiplier $\hat{q}_{p_0, 0.95, n, r}^{(EM+)}$.
 - For $n \geq 50$ the preferred choice is the EVT estimator $\hat{q}_{p_0, r}^{(RBMOP)}$.
- Sample size view
 - For smaller sample sizes, i.e. $20 \leq n \leq 100$:
 - For $\underline{G}$ the preferred choice is the empirical multiplier $\hat{q}_{p_0, 0.95, n, r}^{(EM+)}$ for $n \leq 20$, otherwise the EVT estimator $\hat{q}_{p_0, r}^{(RBMOP)}$.
 - For $\underline{G}^{(1)}$ the preferred choice is the empirical multiplier $\hat{q}_{p_0, 0.9, n, r}^{(EM+)}$ for $n \leq 20$, otherwise the EVT estimator $\hat{q}_{p_0, r}^{(RBMOP)}$.
 - For $\underline{G}^{(2)}$ the preferred choice is the asymptotic multiplier $\hat{q}_{p_0, 0.95, n, r}^{(AM+)}$.
 - For $\underline{G}^{(3)}$ the preferred choice is the EVT estimator $\hat{q}_{p_0, r}^{(RBMOP)}$.
 - For $\underline{G}^{(4)}$ the preferred choice is the empirical multiplier $\hat{q}_{p_0, 0.95, n, r}^{(EM+)}$ for $n \leq 20$, otherwise the EVT estimator $\hat{q}_{p_0, r}^{(RBMOP)}$.
 - For larger sample sizes, i.e. $100 \leq n \leq 200$:
 - For $\underline{G}$ the preferred choice is the EVT estimator $\hat{q}_{p_0, r}^{(RBMOP)}$.
 - For $\underline{G}^{(1)}$ the preferred choice is the EVT estimator $\hat{q}_{p_0, r}^{(RBMOP)}$.
 - For $\underline{G}^{(2)}$ the preferred choice is the EVT estimator $\hat{q}_{p_0, r}^{(RBMOP)}$.
 - For $\underline{G}^{(3)}$ the preferred choice is the EVT estimator $\hat{q}_{p_0, r}^{(RBMOP)}$.
 - For $\underline{G}^{(4)}$ the preferred choice is the EVT estimator $\hat{q}_{p_0, r}^{(RBMOP)}$.

Results of the second order EVT simulation study for $p_0 = 0.99$

Tables 4.12 and 4.13 contain the RMSE (Step 6.2 in Algorithm 4.1) and absolute bias (Step 6.4 in Algorithm 4.1) results respectively for $p_0 = 0.99$ in the second order EVT environment.

Table 4.12: Root mean squared error results of the second order EVT simulation study for $p_0 = 0.99$

p_0	n	r	$\hat{q}_{p_0,r}^{(W+)}$	$\hat{q}_{p_0,0.95,n,r}^{(EM+)}$	$\hat{q}_{p_0,0.9,n,r}^{(EM+)}$	$\hat{q}_{p_0,0.85,n,r}^{(EM+)}$	$\hat{q}_{p_0,0.8,n,r}^{(EM+)}$	$\hat{q}_{p_0,0.95,r}^{(AM+)}$	$\hat{q}_{p_0,0.9,r}^{(AM+)}$	$\hat{q}_{p_0,0.85,r}^{(AM+)}$	$\hat{q}_{p_0,0.8,r}^{(AM+)}$
0.99	20	1	0.06	0.79	0.84	0.81	0.91	0.56	0.94	1.00	0.91
0.99	20	2	0.00	0.00	0.00	0.64	1.00	0.00	0.00	0.67	1.00
0.99	20	3	0.00	0.81	1.00	0.97	0.93	0.55	0.99	0.95	0.93
0.99	20	4	0.01	0.83	1.00	0.97	0.92	0.84	0.98	0.95	0.92
0.99	20	5	0.48	1.00	0.17	0.12	0.12	0.30	0.19	0.15	0.12
0.99	20	6	0.42	1.00	0.14	0.10	0.10	0.26	0.16	0.12	0.10
0.99	20	7	0.00	0.00	0.01	0.76	1.00	0.00	0.01	0.91	1.00
0.99	20	8	0.00	0.00	0.00	0.16	1.00	0.00	0.00	0.18	1.00
Average			0.12	0.55	0.40	0.57	0.75	0.31	0.41	0.62	0.75
Group Maximum			0.12	0.75				0.75			
0.99	50	1	0.28	0.83	1.00	0.95	0.83	0.79	1.00	0.95	0.83
0.99	50	2	0.15	0.74	1.00	0.98	0.96	0.84	1.00	0.98	0.96
0.99	50	3	0.12	0.88	0.90	0.85	0.83	1.00	0.90	0.85	0.83
0.99	50	4	0.29	0.82	0.91	0.86	0.82	1.00	0.91	0.86	0.82
0.99	50	5	0.96	1.00	0.18	0.14	0.11	0.31	0.18	0.14	0.11
0.99	50	6	1.00	0.91	0.17	0.13	0.10	0.30	0.17	0.13	0.10
0.99	50	7	0.38	1.00	0.44	0.36	0.31	0.53	0.44	0.36	0.31
0.99	50	8	1.00	0.51	0.11	0.08	0.06	0.20	0.11	0.08	0.06
Average			0.52	0.84	0.59	0.54	0.50	0.62	0.59	0.54	0.50
Group Maximum			0.52	0.84				0.62			
0.99	100	1	0.49	0.86	1.00	0.88	0.75	0.91	1.00	0.88	0.75
0.99	100	2	0.36	1.00	0.97	0.93	0.89	0.98	0.97	0.93	0.89
0.99	100	3	0.53	1.00	0.86	0.81	0.79	0.99	0.86	0.81	0.79
0.99	100	4	0.61	1.00	0.84	0.78	0.75	0.97	0.84	0.78	0.75
0.99	100	5	1.00	0.22	0.13	0.10	0.08	0.23	0.13	0.10	0.08
0.99	100	6	1.00	0.21	0.12	0.09	0.08	0.22	0.12	0.09	0.08
0.99	100	7	1.00	0.89	0.70	0.56	0.48	0.97	0.70	0.56	0.48
0.99	100	8	1.00	0.13	0.08	0.06	0.05	0.14	0.08	0.06	0.05
Average			0.75	0.66	0.59	0.53	0.48	0.68	0.59	0.53	0.48
Group Maximum			0.75	0.66				0.68			

p_0	n	r	$\hat{q}_{p_0,r}^{(W+)}$	$\hat{q}_{p_0,0.95,n,r}^{(EM+)}$	$\hat{q}_{p_0,0.9,n,r}^{(EM+)}$	$\hat{q}_{p_0,0.85,n,r}^{(EM+)}$	$\hat{q}_{p_0,0.8,n,r}^{(EM+)}$	$\hat{q}_{p_0,0.95,r}^{(AM+)}$	$\hat{q}_{p_0,0.9,r}^{(AM+)}$	$\hat{q}_{p_0,0.85,r}^{(AM+)}$	$\hat{q}_{p_0,0.8,r}^{(AM+)}$
0.99	150	1	0.56	1.00	1.00	0.83	0.69	1.00	1.00	0.83	0.69
0.99	150	2	0.52	1.00	0.94	0.88	0.85	1.00	0.94	0.88	0.85
0.99	150	3	0.86	1.00	0.87	0.82	0.80	1.00	0.87	0.82	0.80
0.99	150	4	0.89	1.00	0.85	0.79	0.76	1.00	0.85	0.79	0.76
0.99	150	5	1.00	0.19	0.11	0.08	0.07	0.19	0.11	0.08	0.07
0.99	150	6	1.00	0.19	0.10	0.08	0.06	0.19	0.10	0.08	0.06
0.99	150	7	1.00	0.81	0.56	0.44	0.37	0.81	0.56	0.44	0.37
0.99	150	8	1.00	0.11	0.06	0.05	0.04	0.11	0.06	0.05	0.04
Average			0.85	0.66	0.56	0.50	0.45	0.66	0.56	0.50	0.45
Group Maximum			0.85	0.66				0.66			
0.99	200	1	0.57	1.00	0.93	0.74	0.61	1.00	0.93	0.74	0.61
0.99	200	2	0.60	1.00	0.91	0.85	0.81	1.00	0.91	0.85	0.81
0.99	200	3	1.00	0.84	0.73	0.69	0.67	0.84	0.73	0.69	0.67
0.99	200	4	1.00	0.90	0.77	0.71	0.68	0.90	0.77	0.71	0.68
0.99	200	5	1.00	0.17	0.10	0.07	0.06	0.17	0.10	0.07	0.06
0.99	200	6	1.00	0.17	0.09	0.07	0.06	0.17	0.09	0.07	0.06
0.99	200	7	1.00	0.72	0.48	0.38	0.32	0.72	0.48	0.38	0.32
0.99	200	8	1.00	0.10	0.06	0.04	0.03	0.10	0.06	0.04	0.03
Average			0.90	0.61	0.51	0.44	0.41	0.61	0.51	0.44	0.41
Group Maximum			0.90	0.61				0.61			

The following conclusions are drawn from **Table 4.12** in terms of the relative performance of the estimators over the range of distributions and sample sizes considered. Note that this table contains the **RMSE** results for the quantile with $p_0 = 0.99$.

- Distribution view
 - For $\underline{G}$, i.e. distributions with $0.2 \leq \gamma \leq 1$ and $-2 \leq \rho \leq -0.2$:
 - For $n \leq 50$ the preferred choice is the empirical multiplier $\hat{q}_{p_0,0.95,n,r}^{(EM+)}$.
 - For $n \geq 100$ the preferred choice is the EVT estimator $\hat{q}_{p_0,r}^{(RBMOP)}$.
 - For $\underline{G}_1$, i.e. distributions with $0.67 \leq \gamma \leq 1$ and $-2 \leq \rho \leq -1$:
 - For $n \leq 50$ the preferred choice is the asymptotic multiplier $\hat{q}_{p_0,0.9,n,r}^{(AM+)}$.
 - For $n \geq 100$ the preferred choice is the empirical multiplier $\hat{q}_{p_0,0.95,n,r}^{(EM+)}$.
 - For $\underline{G}_2$, i.e. distributions with $0.5 \leq \gamma \leq 0.8$ and $-0.3 \leq \rho \leq -0.2$:
 - For $n \leq 150$ the preferred choice is the empirical multiplier $\hat{q}_{p_0,0.95,n,r}^{(EM+)}$.
 - For $n \geq 200$ the preferred choice is the EVT estimator $\hat{q}_{p_0,r}^{(RBMOP)}$.
 - For $\underline{G}_3$, i.e. distributions with $0.25 \leq \gamma \leq 0.3$ and $-0.8 \leq \rho \leq -0.4$:
 - For $n \leq 50$ the preferred choice is the empirical multiplier $\hat{q}_{p_0,0.95,n,r}^{(EM+)}$.

- For $n \geq 100$ the preferred choice is the EVT estimator $\hat{q}_{p_0,r}^{(RBMOP)}$.
- For $\underline{G}_4$, i.e. distributions with $0.25 \leq \gamma \leq 0.57$ and $\rho = -2$:
 - For $n \leq 50$ the preferred choice is the empirical multiplier $\hat{q}_{p_0,0.8,n,r}^{(EM+)}$.
 - For $n \geq 100$ the preferred choice is the EVT estimator $\hat{q}_{p_0,r}^{(RBMOP)}$.
- Sample size view
 - For smaller sample sizes, i.e. $20 \leq n \leq 100$:
 - For $\underline{G}$ the preferred choice is the empirical multiplier $\hat{q}_{p_0,0.95,n,r}^{(EM+)}$.
 - For $\underline{G}^{(1)}$ the preferred choice is the asymptotic multiplier $\hat{q}_{p_0,0.85,n,r}^{(AM+)}$ for $n \leq 20$, otherwise the asymptotic multiplier $\hat{q}_{p_0,0.9,n,r}^{(AM+)}$.
 - For $\underline{G}^{(2)}$ the preferred choice is the empirical multiplier $\hat{q}_{p_0,0.9,n,r}^{(EM+)}$ for $n \leq 20$, otherwise $\hat{q}_{p_0,0.95,n,r}^{(EM+)}$.
 - For $\underline{G}^{(3)}$ the preferred choice is the empirical multiplier $\hat{q}_{p_0,0.95,n,r}^{(EM+)}$ for $n \leq 20$, otherwise the EVT estimator $\hat{q}_{p_0,r}^{(RBMOP)}$.
 - For $\underline{G}^{(4)}$ the preferred choice is the empirical multiplier $\hat{q}_{p_0,0.8,n,r}^{(EM+)}$ for $n \leq 20$, otherwise the EVT estimator $\hat{q}_{p_0,r}^{(RBMOP)}$.
 - For larger sample sizes, i.e. $100 \leq n \leq 200$:
 - For $\underline{G}$ the preferred choice is the EVT estimator $\hat{q}_{p_0,r}^{(RBMOP)}$.
 - For $\underline{G}^{(1)}$ the preferred choice is the empirical multiplier $\hat{q}_{p_0,0.95,n,r}^{(EM+)}$.
 - For $\underline{G}^{(2)}$ the preferred choice is the empirical multiplier $\hat{q}_{p_0,0.95,n,r}^{(EM+)}$.
 - For $\underline{G}^{(3)}$ the preferred choice is the EVT estimator $\hat{q}_{p_0,r}^{(RBMOP)}$.
 - For $\underline{G}^{(4)}$ the preferred choice is the EVT estimator $\hat{q}_{p_0,r}^{(RBMOP)}$.

Table 4.13: Absolute bias results of the second order EVT simulation study for

$p_0 = 0.99$

p_0	n	r	$\hat{q}_{p_0,r}^{(W+)}$	$\hat{q}_{p_0,0.95,n,r}^{(EM+)}$	$\hat{q}_{p_0,0.9,n,r}^{(EM+)}$	$\hat{q}_{p_0,0.85,n,r}^{(EM+)}$	$\hat{q}_{p_0,0.8,n,r}^{(EM+)}$	$\hat{q}_{p_0,0.95,r}^{(AM+)}$	$\hat{q}_{p_0,0.9,r}^{(AM+)}$	$\hat{q}_{p_0,0.85,r}^{(AM+)}$	$\hat{q}_{p_0,0.8,r}^{(AM+)}$
0.99	20	1	0.17	0.24	0.55	0.37	0.48	1.00	0.88	0.63	0.48
0.99	20	2	0.00	0.00	0.03	1.00	0.80	0.00	0.03	0.84	0.80
0.99	20	3	0.01	0.62	0.70	0.65	0.61	1.00	0.68	0.63	0.61
0.99	20	4	0.11	0.51	0.72	0.67	0.60	1.00	0.68	0.62	0.60
0.99	20	5	1.00	0.10	0.01	0.01	0.01	0.02	0.01	0.01	0.01
0.99	20	6	1.00	0.11	0.01	0.01	0.01	0.02	0.01	0.01	0.01
0.99	20	7	0.00	0.00	0.28	0.79	0.84	0.00	0.30	1.00	0.84
0.99	20	8	0.00	0.00	0.00	0.87	0.86	0.00	0.00	1.00	0.86
Average			0.29	0.20	0.29	0.54	0.53	0.38	0.32	0.59	0.53
Group Maximum			0.29	0.54				0.59			
0.99	50	1	1.00	0.11	0.22	0.15	0.12	0.51	0.22	0.15	0.12
0.99	50	2	0.67	0.53	0.72	0.66	0.63	1.00	0.72	0.66	0.63
0.99	50	3	1.00	0.78	0.78	0.73	0.71	0.95	0.78	0.73	0.71
0.99	50	4	1.00	0.46	0.53	0.49	0.47	0.69	0.53	0.49	0.47
0.99	50	5	1.00	0.24	0.03	0.02	0.02	0.06	0.03	0.02	0.02
0.99	50	6	1.00	0.26	0.04	0.03	0.02	0.07	0.04	0.03	0.02
0.99	50	7	1.00	0.27	0.05	0.04	0.03	0.09	0.05	0.04	0.03
0.99	50	8	1.00	0.08	0.02	0.01	0.01	0.03	0.02	0.01	0.01
Average			0.96	0.34	0.30	0.27	0.25	0.42	0.30	0.27	0.25
Group Maximum			0.96	0.34				0.42			
0.99	100	1	0.91	0.66	0.41	0.28	0.22	1.00	0.41	0.28	0.22
0.99	100	2	1.00	0.28	0.20	0.18	0.18	0.26	0.20	0.18	0.18
0.99	100	3	1.00	0.11	0.09	0.08	0.08	0.11	0.09	0.08	0.08
0.99	100	4	1.00	0.08	0.06	0.05	0.05	0.07	0.06	0.05	0.05
0.99	100	5	1.00	0.06	0.04	0.03	0.02	0.07	0.04	0.03	0.02
0.99	100	6	1.00	0.07	0.04	0.03	0.02	0.07	0.04	0.03	0.02
0.99	100	7	1.00	0.06	0.04	0.03	0.03	0.06	0.04	0.03	0.03
0.99	100	8	1.00	0.02	0.01	0.01	0.01	0.02	0.01	0.01	0.01
Average			0.99	0.17	0.11	0.09	0.08	0.21	0.11	0.09	0.08
Group Maximum			0.99	0.17				0.21			

p_0	n	r	$\hat{q}_{p_0,r}^{(W+)}$	$\hat{q}_{p_0,0.95,n,r}^{(EM+)}$	$\hat{q}_{p_0,0.9,n,r}^{(EM+)}$	$\hat{q}_{p_0,0.85,n,r}^{(EM+)}$	$\hat{q}_{p_0,0.8,n,r}^{(EM+)}$	$\hat{q}_{p_0,0.95,r}^{(AM+)}$	$\hat{q}_{p_0,0.9,r}^{(AM+)}$	$\hat{q}_{p_0,0.85,r}^{(AM+)}$	$\hat{q}_{p_0,0.8,r}^{(AM+)}$
0.99	150	1	0.48	1.00	0.35	0.24	0.19	1.00	0.35	0.24	0.19
0.99	150	2	1.00	0.06	0.05	0.04	0.04	0.06	0.05	0.04	0.04
0.99	150	3	1.00	0.10	0.09	0.08	0.08	0.10	0.09	0.08	0.08
0.99	150	4	1.00	0.10	0.08	0.07	0.07	0.10	0.08	0.07	0.07
0.99	150	5	1.00	0.08	0.04	0.03	0.03	0.08	0.04	0.03	0.03
0.99	150	6	1.00	0.08	0.04	0.03	0.03	0.08	0.04	0.03	0.03
0.99	150	7	1.00	0.11	0.06	0.05	0.04	0.11	0.06	0.05	0.04
0.99	150	8	1.00	0.03	0.01	0.01	0.01	0.03	0.01	0.01	0.01
Average			0.94	0.19	0.09	0.07	0.06	0.19	0.09	0.07	0.06
Group Maximum			0.94	0.19				0.19			
0.99	200	1	0.41	1.00	0.33	0.23	0.18	1.00	0.33	0.23	0.18
0.99	200	2	1.00	0.08	0.06	0.06	0.05	0.08	0.06	0.06	0.05
0.99	200	3	1.00	0.19	0.16	0.15	0.15	0.19	0.16	0.15	0.15
0.99	200	4	1.00	0.19	0.16	0.15	0.14	0.19	0.16	0.15	0.14
0.99	200	5	1.00	0.08	0.05	0.03	0.03	0.08	0.05	0.03	0.03
0.99	200	6	1.00	0.09	0.05	0.03	0.03	0.09	0.05	0.03	0.03
0.99	200	7	1.00	0.15	0.09	0.07	0.06	0.15	0.09	0.07	0.06
0.99	200	8	1.00	0.03	0.01	0.01	0.01	0.03	0.01	0.01	0.01
Average			0.93	0.22	0.11	0.09	0.08	0.22	0.11	0.09	0.08
Group Maximum			0.93	0.22				0.22			

The following conclusions are drawn from **Table 4.13** in terms of the relative performance of the estimators over the range of distributions and sample sizes considered. Note that this table contains the **absolute bias** results for the quantile with $p_0 = 0.99$.

- Distribution view
 - For $\underline{G}$, i.e. distributions with $0.2 \leq \gamma \leq 1$ and $-2 \leq \rho \leq -0.2$:
 - For $n \leq 20$ the preferred choice is the asymptotic multiplier $\hat{q}_{p_0,0.85,n,r}^{(AM+)}$.
 - For $n \geq 50$ the preferred choice is the EVT estimator $\hat{q}_{p_0,r}^{(RBMOP)}$.
 - For $\underline{G}_1$, i.e. distributions with $0.67 \leq \gamma \leq 1$ and $-2 \leq \rho \leq -1$:
 - For $n \leq 20$ the preferred choice is the asymptotic multiplier $\hat{q}_{p_0,0.85,n,r}^{(AM+)}$.
 - For $n \geq 50$ the preferred choice is the EVT estimator $\hat{q}_{p_0,r}^{(RBMOP)}$.
 - For $\underline{G}_2$, i.e. distributions with $0.5 \leq \gamma \leq 0.8$ and $-0.3 \leq \rho \leq -0.2$:
 - For $n \leq 20$ the preferred choice is the asymptotic multiplier $\hat{q}_{p_0,0.95,n,r}^{(AM+)}$.
 - For $n \geq 50$ the preferred choice is the EVT estimator $\hat{q}_{p_0,r}^{(RBMOP)}$.
 - For $\underline{G}_3$, i.e. distributions with $0.25 \leq \gamma \leq 0.3$ and $-0.8 \leq \rho \leq -0.4$:

- For any n the preferred choice is the EVT estimator $\hat{q}_{p_0,r}^{(RBMOP)}$.
- For $\underline{G}_4$, i.e. distributions with $0.25 \leq \gamma \leq 0.57$ and $\rho = -2$:
 - For $n \leq 20$ the preferred choice is the asymptotic multiplier $\hat{q}_{p_0,0.85,n,r}^{(AM+)}$.
 - For $n \geq 50$ the preferred choice is the EVT estimator $\hat{q}_{p_0,r}^{(RBMOP)}$.
- Sample size view
 - For smaller sample sizes, i.e. $20 \leq n \leq 100$:
 - For $\underline{G}$ the preferred choice is the empirical multiplier $\hat{q}_{p_0,0.85,n,r}^{(AM+)}$ for $n \leq 20$, otherwise the EVT estimator $\hat{q}_{p_0,r}^{(RBMOP)}$.
 - For $\underline{G}^{(1)}$ the preferred choice is the empirical multiplier $\hat{q}_{p_0,0.85,n,r}^{(EM+)}$ for $n \leq 20$, otherwise the EVT estimator $\hat{q}_{p_0,r}^{(RBMOP)}$.
 - For $\underline{G}^{(2)}$ the preferred choice is the asymptotic multiplier $\hat{q}_{p_0,0.95,n,r}^{(AM+)}$ for $n \leq 20$, otherwise the EVT estimator $\hat{q}_{p_0,r}^{(RBMOP)}$.
 - For $\underline{G}^{(3)}$ the preferred choice is the EVT estimator $\hat{q}_{p_0,r}^{(RBMOP)}$.
 - For $\underline{G}^{(4)}$ the preferred choice is the asymptotic multiplier $\hat{q}_{p_0,0.85,n,r}^{(AM+)}$ for $n \leq 20$, otherwise the EVT estimator $\hat{q}_{p_0,r}^{(RBMOP)}$.
 - For larger sample sizes, i.e. $100 \leq n \leq 200$:
 - For $\underline{G}$ the preferred choice is the EVT estimator $\hat{q}_{p_0,r}^{(RBMOP)}$.
 - For $\underline{G}^{(1)}$ the preferred choice is the EVT estimator $\hat{q}_{p_0,r}^{(RBMOP)}$.
 - For $\underline{G}^{(2)}$ the preferred choice is the EVT estimator $\hat{q}_{p_0,r}^{(RBMOP)}$.
 - For $\underline{G}^{(3)}$ the preferred choice is the EVT estimator $\hat{q}_{p_0,r}^{(RBMOP)}$.
 - For $\underline{G}^{(4)}$ the preferred choice is the EVT estimator $\hat{q}_{p_0,r}^{(RBMOP)}$.

The results presented in both this and the previous sections are tedious, but these provide important information when attempting to understand the behaviour of the various quantile estimators. A summary of these results are given in the next section where the overall performance of the different estimators is highlighted.

4.1.3. Summary of the EVT simulation results

A summary of the various quantile estimators is given in this section. This is made after an extensive investigation of the results presented in the previous sections. Furthermore, the summary is made taking different practical applications of the estimators into consideration in the sense that usually an OR analyst would know the values of both the required p_0 and n . Thus, the performance of the various estimators are summarized for different values of p_0 and n . In addition, the results are compiled separately for first and second order EVT.

It was found that in the first order EVT environment the empirical multipliers (and in particular those using $p_1 = 0.95$ for the lower quantile) generally perform best for estimating the extreme quantiles ($p_0 = 0.999$), especially for the smaller sample sizes of 20, 50 and 100. As expected, the first order EVT quantile estimator performed better than the multipliers for larger sample sizes (150 and 200) due to its asymptotic nature/properties. Also, the empirical multipliers outperformed the EVT estimators for the underlying distributions contained within both groups $\underline{G}_1$ and $\underline{G}_2$, i.e. distributions having larger EVI's which in turn relates to distributions with heavier tails.

Remaining within the first order EVT environment, but moving to the $p_0 = 0.995$ quantiles it can be noted that the empirical multipliers (and in particular those using $p_1 = 0.95$ for the lower quantile) generally still perform best for estimating the high quantiles, especially for the smaller sample sizes of 20, 50 and 100. Again, as expected, the first order EVT quantile estimator performed better than the multipliers for larger sample sizes (150 and 200). In contrast to the case of the $p_0 = 0.999$ quantiles, the empirical multipliers outperformed the EVT estimators for the underlying distributions contained within group $\underline{G}_2$ only, i.e. distributions having larger EVI's. The conclusions drawn for the $p_0 = 0.99$ case are similar to those of the $p_0 = 0.995$ case.

Moving to the second order EVT environment it was found that the empirical multipliers (and in particular those using $p_1 = 0.95$ for the lower quantile) generally perform best for estimating the extreme quantiles ($p_0 = 0.999$), especially for the smaller sample sizes of 20, 50 and 100. Note that although the empirical multipliers generally outperformed the asymptotic multipliers, the results are roughly similar in many cases as the multiplying factor used for both classes of multiplier estimators are the same. As expected, the second order EVT quantile estimator performed better than the multipliers for larger sample sizes (150 and 200) due to its asymptotic nature/properties. Also, the empirical multipliers outperformed the

EVT estimators for the underlying distributions contained within both groups $\underline{G}_1$ and $\underline{G}_2$, i.e. distributions having larger EVI's which in turn relates to distributions with heavier tails. In addition, it is worth mentioning that the EVT estimator outperformed the multiplier estimators regarding the bias. This is expected as the second order EVT estimators are developed with the aim to reduce bias in its estimates. The results for $p_0 = 0.995$ and $p_0 = 0.99$ are very similar to those of $p_0 = 0.999$.

These conclusions were drawn from the results obtained from the simulation study. One should of course be careful in generalizing these results, as is always the case in such a simulation study. However, the distributions in $\underline{G}$ cover a wide range of (γ, ρ) values that are often encountered in practice.

Remark: Even though second order EVT estimators should perform better than first order EVT estimators, the same amount of attention is given to the latter due to its relative simplicity and popularity in practice.

4.2. A simulation study used to compare the quantile estimation techniques based on the full-parametric approach (The full-parametric simulation study)

In this section the LogNIG distribution's ability to estimate extreme quantiles is tested by comparing it with other related distributions. The idea of this simulation study is to generate data from some predetermined distribution to which the distributions considered are then fitted. The fitted distributions are used to estimate the desired quantiles which in turn are used to calculate various statistical measures, together with their standard errors. The values of the statistical measures are used to compare the various distributions' ability to estimate the quantiles.

4.2.1. Simulation design

The distributions used to test the LogNIG distribution are the Burr, LogGamma and LogNormal distributions. It was decided to use the Burr distribution since this distribution is widely used to model the severities often encountered in OR. Furthermore, both the LogGamma and LogNormal distributions were also included in this study to investigate the "Log" class of distributions. Initially the LogPh distribution proposed by Ahn et al. (2012) was also considered for this study, but was omitted due to the complexity of fitting the distribution to small sampled data. Further research on this distribution is suggested.

Data was generated from the distributions listed in Table 4.14. These distributions are also illustrated graphically in Figure 4.2 by use of a scatterplot of the distributions' γ and ρ . The following explanations are given to understand the content of Table 4.14. The first column (named r) contains the distribution reference number that comes in handy when interpreting the results. The second column (named group) indicates which distribution is used (the reader is referred to Table A.1, Table A.2 and Table A.3 in Appendix A for more on these groups). The third column (named distribution) contains the actual distribution, together with its parameters, used to generate the data. Lastly, the fourth (named γ) and fifth (named ρ) columns contain the EVI and second order parameter of the corresponding distribution respectively. Columns 6, 7, 8 and 9 contain the true quantiles at probability levels 0.95, 0.99, 0.995 and 0.999 respectively. It can be seen from Table 4.14 that data was generated from various Burr distributions only. Initially, other distributions such as the LogNIG and LogN were also considered for generating the sample data, but were omitted from the study as various computational challenges arose which will be discussed in Section 4.2.3.

Table 4.14: List of distributions used to generate data for the full-parametric simulation study

r	Group	Distribution	γ	ρ	q_{95}	q_{99}	$q_{99.5}$	$q_{99.9}$
1	F1	Burr(1,5,1)	0.2	-1	1.80	2.51	2.88	3.98
2	G6	Burr(1,1.6,2.5)	0.25	-0.4	1.69	2.84	3.47	5.40
3	F2	Burr(1,5,0.5)	0.4	-2	3.31	6.31	8.33	15.85
4	F6	Burr(1,3,2/3)	0.5	-1.5	4.46	10.00	14.14	31.62
5	No group	Burr(1,2,1)	0.5	-1	4.36	9.95	14.11	31.61
6	No group	Burr(1,1,2)	0.5	-0.5	3.47	9.00	13.14	30.62
7	F3	Burr(1,2.5,0.5)	0.8	-2	10.98	39.81	69.31	251.19
8	F5	Burr(1,1.25,1)	0.8	-1	10.54	39.49	69.04	250.99

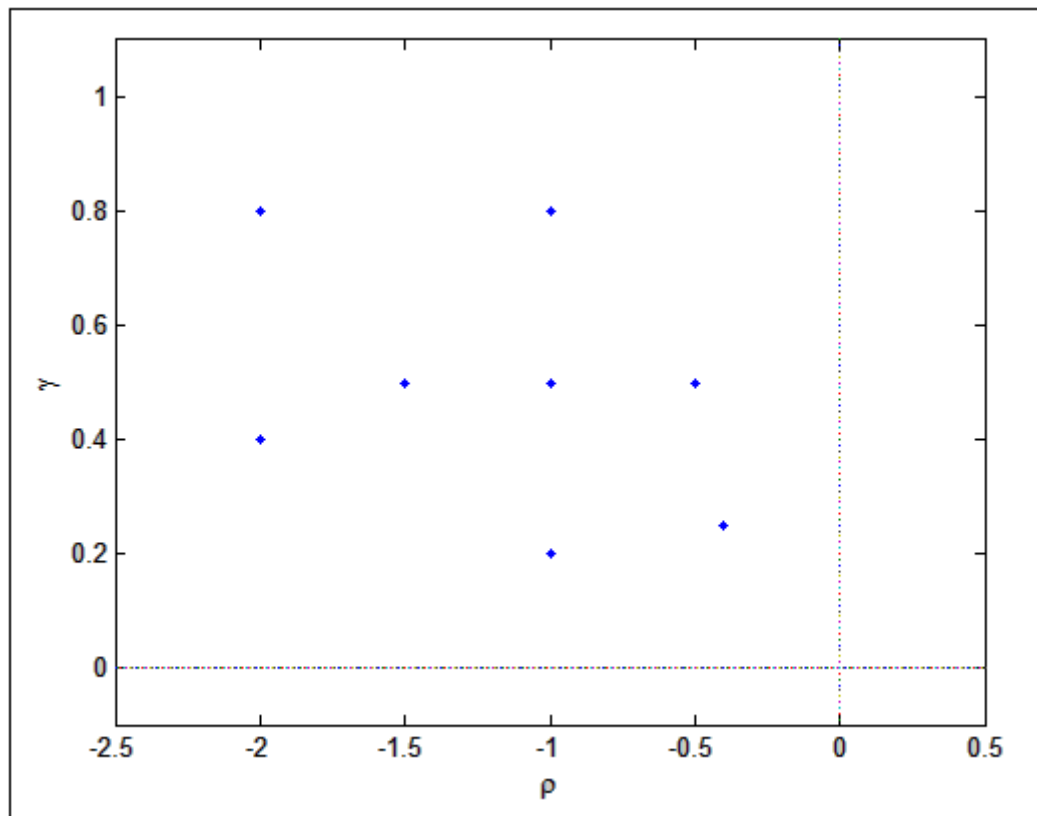


Figure 4.2: Scatterplot of the distributions used in the full-parametric simulation study

The sample sizes considered were 200 and 500. Initially a sample size of 100 was also considered, but it was decided to exclude it from the study as fitting issues arose even at the 200 level. On the other end of the scale, a sample size of 1000 was also considered but it was felt that the 1000 level does not conform to the typical problem area in operational risk regarding sample size.

The distributions fitted to the simulated data (in order to estimate the high quantiles) are the Burr, LogNIG, LogGamma and LogNormal distributions. The probability levels of the quantiles estimated are 0.999, 0.995 and 0.99. It is well known that these are typical quantile

levels of concern in operational risk management. Lastly, the statistical measures used to compare the various quantile estimates are the root mean squared error (RMSE) and the absolute value of bias, denoted by $|\text{BiasA}|$ where the expected value of the estimate is calculated using the average (BiasA). The mean squared error was also calculated but is not presented in this thesis as the RMSE is rather presented which is of the same order as the $|\text{BiasA}|$ measure. The statistical measures were calculated using $N = 10,000$ quantile estimates. In addition, the standard errors of the statistical measure were also calculated using the N estimates, but these are not presented in this thesis.

The aim of the simulation study is to compare the LogNIG distribution with other distributions when used to estimate extreme quantiles. The distributions used as competitors are the Burr, LogGamma and LogNormal distributions. The following section (4.2.2) discusses the simulation results.

4.2.2. Simulation results

The simulation results are presented in this section. It can be expected that the Burr distribution will perform the best due to the data being generated using the Burr distribution. Thus, the Burr distribution will be used as the benchmark when comparing the results. Before presenting the results it is worth mentioning that, as was the case for the EVT simulation study, it is difficult to compare the performance of the various quantile estimators using the raw results. For this reason the results are presented using an efficiency measure of the results where the results are transformed into relative measures of the best performing estimator.

Table 4.15 and Table 4.16 contain the **RMSE** and **absolute bias** results of the $p = 0.999$ quantile estimates respectively.

Table 4.15: RMSE results of the full-parametric quantile estimators for $p = 0.999$

p	n	r	Burr	LogNIG	LogGamma	LogN
0.999	200	1	1.00	0.72	0.79	0.78
0.999	200	2	0.62	0.53	1.00	0.44
0.999	200	3	1.00	0.46	0.53	0.51
0.999	200	4	1.00	0.50	0.79	0.72
0.999	200	5	0.78	0.44	1.00	0.83
0.999	200	6	0.45	0.31	1.00	0.70
0.999	200	7	0.96	0.25	1.00	0.89
0.999	200	8	0.38	0.18	1.00	0.68
Average over $n = 200$			0.77	0.43	0.89	0.70
0.999	500	1	1.00	0.76	0.49	0.48
0.999	500	2	0.80	0.68	1.00	0.37
0.999	500	3	1.00	0.51	0.32	0.31
0.999	500	4	1.00	0.59	0.45	0.41
0.999	500	5	1.00	0.56	0.76	0.62
0.999	500	6	0.54	0.36	1.00	0.54
0.999	500	7	1.00	0.41	0.55	0.49
0.999	500	8	0.64	0.41	1.00	0.53
Average over $n = 500$			0.87	0.53	0.70	0.47

Table 4.16: |BiasA| results of the full-parametric quantile estimators for $p = 0.999$

p	n	r	Burr	LogNIG	LogGamma	LogN
0.999	200	1	0.42	1.00	0.01	0.01
0.999	200	2	0.95	1.00	0.31	0.10
0.999	200	3	1.00	0.17	0.05	0.05
0.999	200	4	1.00	0.33	0.11	0.10
0.999	200	5	1.00	0.57	0.19	0.15
0.999	200	6	0.25	0.21	1.00	0.12
0.999	200	7	1.00	0.20	0.22	0.19
0.999	200	8	0.87	0.45	1.00	0.36
Average over $n = 200$			0.81	0.49	0.36	0.14
0.999	500	1	0.10	1.00	0.00	0.00
0.999	500	2	1.00	0.83	0.08	0.02
0.999	500	3	1.00	0.10	0.02	0.02
0.999	500	4	1.00	0.22	0.04	0.03
0.999	500	5	1.00	0.57	0.08	0.06
0.999	500	6	0.35	0.26	1.00	0.06
0.999	500	7	1.00	0.16	0.07	0.07
0.999	500	8	1.00	0.59	0.34	0.14
Average over $n = 500$			0.81	0.47	0.20	0.05

Table 4.17 and Table 4.18 contain the **RMSE** and **absolute bias** results of the $p = 0.995$ quantile estimates respectively.

Table 4.17: RMSE results of the full-parametric quantile estimators for $p = 0.995$

p	n	r	Burr	LogNIG	LogGamma	LogN
0.995	200	1	1.00	0.86	0.95	0.96
0.995	200	2	0.86	0.81	1.00	0.43
0.995	200	3	1.00	0.67	0.54	0.52
0.995	200	4	1.00	0.73	0.77	0.71
0.995	200	5	0.78	0.64	1.00	0.89
0.995	200	6	0.59	0.54	1.00	0.49
0.995	200	7	1.00	0.57	0.87	0.76
0.995	200	8	0.53	0.39	1.00	0.80
Average over $n = 200$			0.85	0.65	0.89	0.70
0.995	500	1	1.00	0.86	0.63	0.63
0.995	500	2	1.00	0.93	0.90	0.34
0.995	500	3	1.00	0.69	0.34	0.33
0.995	500	4	1.00	0.76	0.47	0.43
0.995	500	5	1.00	0.85	0.85	0.73
0.995	500	6	0.65	0.19	1.00	0.36
0.995	500	7	1.00	0.65	0.50	0.43
0.995	500	8	0.70	0.48	1.00	0.59
Average over $n = 500$			0.92	0.68	0.71	0.48

Table 4.18: |BiasA| results of the full-parametric quantile estimators for $p = 0.995$

p	n	r	Burr	LogNIG	LogGamma	LogN
0.995	200	1	1.00	0.03	0.00	0.00
0.995	200	2	0.70	1.00	0.10	0.03
0.995	200	3	1.00	0.28	0.04	0.04
0.995	200	4	1.00	0.63	0.08	0.07
0.995	200	5	0.86	1.00	0.12	0.10
0.995	200	6	0.34	0.40	1.00	0.05
0.995	200	7	1.00	0.39	0.14	0.12
0.995	200	8	1.00	0.84	0.87	0.30
Average over $n = 200$			0.86	0.57	0.29	0.09
0.995	500	1	1.00	0.06	0.00	0.00
0.995	500	2	0.25	1.00	0.00	0.00
0.995	500	3	1.00	0.15	0.01	0.01
0.995	500	4	1.00	0.39	0.02	0.02
0.995	500	5	0.82	1.00	0.05	0.04
0.995	500	6	0.50	0.36	1.00	0.02
0.995	500	7	1.00	0.26	0.05	0.04
0.995	500	8	1.00	0.85	0.26	0.10
Average over $n = 500$			0.82	0.51	0.17	0.03

Table 4.19 and Table 4.20 contain the **RMSE** and **absolute bias** results of the $p = 0.99$ quantile estimates respectively.

Table 4.19: RMSE results of the full-parametric quantile estimators for $p = 0.99$

p	n	r	Burr	LogNIG	LogGamma	LogN
0.99	200	1	0.90	0.83	0.98	1.00
0.99	200	2	0.99	0.97	1.00	0.44
0.99	200	3	1.00	0.78	0.58	0.56
0.99	200	4	1.00	0.83	0.81	0.76
0.99	200	5	0.76	0.68	1.00	0.95
0.99	200	6	0.67	0.65	1.00	0.45
0.99	200	7	1.00	0.72	0.85	0.75
0.99	200	8	0.60	0.43	1.00	0.93
Average over $n = 200$			0.87	0.74	0.90	0.73
0.99	500	1	1.00	0.91	0.76	0.78
0.99	500	2	1.00	0.96	0.79	0.30
0.99	500	3	1.00	0.79	0.36	0.36
0.99	500	4	1.00	0.85	0.50	0.47
0.99	500	5	1.00	0.91	0.94	0.86
0.99	500	6	0.72	0.68	1.00	0.32
0.99	500	7	1.00	0.77	0.50	0.44
0.99	500	8	0.72	0.40	1.00	0.70
Average over $n = 500$			0.93	0.78	0.73	0.53

Table 4.20: |BiasA| results of the full-parametric quantile estimators for $p = 0.99$

p	n	r	Burr	LogNIG	LogGamma	LogN
0.99	200	1	1.00	0.11	0.01	0.01
0.99	200	2	0.63	1.00	0.06	0.02
0.99	200	3	1.00	0.45	0.03	0.03
0.99	200	4	0.97	1.00	0.07	0.06
0.99	200	5	0.56	1.00	0.07	0.06
0.99	200	6	0.41	0.57	1.00	0.03
0.99	200	7	1.00	0.58	0.12	0.10
0.99	200	8	0.99	1.00	0.81	0.29
Average over $n = 200$			0.82	0.71	0.27	0.08
0.99	500	1	1.00	0.05	0.00	0.00
0.99	500	2	1.00	0.17	0.00	0.00
0.99	500	3	1.00	0.23	0.01	0.01
0.99	500	4	1.00	0.77	0.02	0.02
0.99	500	5	0.46	1.00	0.02	0.02
0.99	500	6	0.64	0.76	1.00	0.02
0.99	500	7	1.00	0.40	0.04	0.03
0.99	500	8	1.00	0.95	0.24	0.10
Average over $n = 500$			0.89	0.54	0.17	0.03

As expected, the Burr distribution performed overall the best at estimating the extreme quantiles considered. This is mainly due to the data being generated from various Burr distributions. More specific results revealed that the LogGamma outperformed all of the other full-parametric quantile estimators (including the Burr distribution) when minimizing the RMSE (of all the quantiles considered, i.e. all p 's) for sample size $n = 200$.

The conclusions drawn from Table 4.15 to Table 4.20 are summarized according to the statistical measure of interest, i.e. a first group summarizing the RMSE results (Table 4.15, Table 4.17 and Table 4.19) and a second group summarizing the absolute bias results (Table 4.16, Table 4.18 and Table 4.20). In addition to separating the results into two groups, the Burr distribution was excluded from the analysis in order to get an unbiased view of the results.

When aiming to minimize the RMSE the LogGamma distribution is the preferred choice for all values of n and p , apart from the case of $p = 0.99$ where the LogNIG is preferred for $n = 500$. Also, the LogGamma appears to perform especially good (generally even better than the Burr distribution) when the sample data was generated from distributions $4 \leq r \leq 8$, i.e. heavier tailed Burr distributions.

When aiming to minimize the $|\text{BiasA}|$ the LogNIG distribution is the preferred choice for all values of n and p . These general results also hold when focusing on more specific results, i.e. for each combination of n and p the LogNIG performed best, apart from the Burr distribution. In addition, the LogNIG distribution outperformed the other “Log-type” distributions when the sample data was generated from all of the distributions other than $r = 6$, where the LogGamma performed best. Note that the distribution $r = 6$ is one of the heavier tailed distributions when compared to the other distributions used to generate the data. The tail of distribution $r = 6$ is considered relatively heavier in the sense that $\gamma = 0.5$, but the second order parameter $\rho = -0.5$ which is relatively close to zero.

A wider set of simulation parameters was originally investigated but the scope of the simulation study was reduced due to various computational challenges. These challenges are discussed next.

4.2.3. Computational challenges

It is worth mentioning that a number of other distributions were initially also considered as distributions to be used to generate the sample data. These distributions are listed in Table 4.21. The distributions listed in this table were eventually omitted from this study due to the computational challenges that it presented.

Table 4.21: Alternative distributions considered that had too many computational challenges

r	Group	Distribution	γ	ρ	q_{95}	q_{99}	$q_{99.5}$	$q_{99.9}$
9	No group	LogN(0,2)	0	0	26.84	104.87	172.72	483.22
10	No group	LogNIG(3,1,1,1)	0.5	0	11.68	23.51	31.73	63.94
11	F4	Burr(1,0.25,5)	0.8	-0.2	0.45	5.23	12.64	78.98
12	No group	LogNIG(2,1,1,1)	1	0	23.77	79.51	136.39	496.79
13	No group	LogNIG(1,0.5,1,1)	2	0	44.95	347.51	910.42	9830.90

The following explains the computational issues using distribution $r = 11$ as an example. It was mentioned that $N \times B = 1000 \times 10 = 10\,000$ repetitions will be used during the simulation study which in turn means that 10 000 quantile estimates for each case will be used. One therefore needs 10 000 fitted distributions to calculate the 10 000 quantile estimates. It was found that either the Burr distribution or the LogNIG distribution could not be fitted to the simulated data (due to parameters that do not converge for MLE) for roughly 2/3 of the samples generated. Therefore, roughly 30 000 samples are required to obtain the desired 10 000 quantile estimates (at a specific probability level). It was decided to exclude the distributions $9 \leq r \leq 13$ from this study due to this reason. After investigating the reason for the convergence issues it was realized that the distributions that produced these challenges all have a common property, which is that their second order parameter ρ is either equal to or very close to zero. It is well-known that distributions possessing this property can be challenging to work with (see e.g. Beirlant et al., 2004).

4.2.4. Concluding remarks on the full-parametric approach

A simulation study was conducted to compare the performance of the LogNIG distribution to (mainly) three different distributions when used to estimate various quantiles. The competing distributions are the Burr, LogGamma and LogNormal distributions. The quantiles of interest are those with probability $p = 0.999, 0.995, 0.99$. Samples of two different sizes $n = 200, 500$ were generated from eight different Burr distributions $r = 1, 2, 3, 4, 5, 6, 7, 8$ (see Table 4.14) with ranging values of $0.2 \leq \gamma \leq 0.8$ and $-2 \leq \rho \leq -0.4$. The quantile estimates were used to calculate various statistical measures, in particular the RMSE and absolute bias, to compare

the relative performance of the four statistical distributions considered. The results proved that the Burr distribution performed best from an overall point of view. This was expected as all of the distributions used to generate the sample data were from the Burr class of distributions.

When excluding the Burr distribution from the analysis it was found/shown that generally the LogGamma performed best when the aim is to minimize the RMSE. The LogNIG outperformed the LogGamma for $p = 0.99$ and $n = 500$. In addition, the LogNIG generally outperformed both the LogGamma and LogNormal when aiming to reduce/minimize the absolute bias. The only exception regarding the absolute bias result is that the LogGamma performed best for data being generated from $r = 6$, which is a Burr distribution having $\rho = -0.5$.

Future research regarding the LogNIG distribution includes a comparison of the LogNIG distribution's numerical properties to that of more recently proposed distributions, such as the LogPhase type of distributions proposed by Ahn et al. (2012) and the g-and-h distribution proposed by Degen et al. (2007). In addition, the possibility of using the LogNIG distribution to estimate the lower quantile of the multiplier quantile estimator seems viable. This is due to the LogNIG distribution's relatively good performance at $p = 0.99$ for the sample size $n = 500$.

4.3. An empirical study of the techniques used to estimate the severity distribution's quantiles

An empirical analysis was also conducted to measure the performance of the newly proposed quantile estimators' ability to estimate extreme quantiles. Both the multiplier and LogNIG distribution was compared to their respective peers, i.e. the EVT quantile estimator proposed by Weissman and the Burr distribution. It is worth mentioning that it is difficult to conduct an empirical study on actual OR loss data as OR data samples are typically very small. In addition to the generally small samples, OR data is also heavy-tailed. Combining these two properties of OR loss data leads to questionable assumptions of the true quantile to be estimated. It is for this reason that insurance data will rather be used which can be assumed to have similar properties to OR loss data regarding the tail behaviour. Two samples of insurance claims data obtained from Berning (2010), which will be referred to as Sample1 and Sample2, were used to assess the newly proposed quantile estimators. The quantiles of interest are $p = 0.99, 0.995, 0.999$.

Let the observations in Sample1 be ordered as $X_{1,N} \leq X_{2,N} \leq \dots \leq X_{N,N}$ with $N = 18198$ the sample size. Due to the large size of the sample it can be assumed that the ordered statistic $X_{[pN],N}$ with $[k]$ the largest integer in k is the true p^{th} -quantile of the underlying distribution. The same assumption is valid for Sample2 $Y_{1,M} \leq Y_{2,M} \leq \dots \leq Y_{M,M}$ with $M = 14197$.

The EVI (γ) was estimated for both samples using a list of estimators. The first is the Hill (1975) estimator discussed in Section 2.2 where the number of exceedances was taken at different percentages of the sample size. These percentages are 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45% and 50%. Both the perturbed Pareto distribution (PPD) (see e.g. Berning, 2010) and the robust technique proposed by Beran and Schell (2012) were also used to estimate the EVI. Lastly, both the Burr and LogNIG distributions were also used to estimate the EVI. Table 4.22 illustrates the EVI estimates obtained from the Hill estimator using all the listed percentages, as well as estimated values of the EVI using the PPD, Burr and LogNIG distribution. Note that the robust estimates are not presented as these varied widely for different choice of the required parameters. It can be noted from Table 4.22 that the EVI estimates obtained from the fitted distributions (PPD, Burr and LogNIG) are on the lower end of the range of EVI estimates compared to the Hill estimator. This is the case for both samples.

Table 4.22: Estimates of the empirical data's EVIs

EVI estimator	Sample1	Sample2
Hill (5%)	0.91	0.93
Hill (10%)	0.93	0.98
Hill (15%)	0.94	0.99
Hill (20%)	0.96	1.01
Hill (25%)	1.00	1.05
Hill (30%)	1.01	1.07
Hill (35%)	1.04	1.11
Hill (40%)	1.08	1.16
Hill (45%)	1.13	1.22
Hill (50%)	1.17	1.27
PPD	0.81	0.80
Burr	0.74	0.81
LNIG	0.95	0.73

A long list of quantile estimators was considered to estimate the true quantiles. The following summarizes the estimators used for this empirical study:

- The Burr distribution, referred to as **Burr**,
- The empirical multipliers where the lower quantile ($p = 0.95$) was estimated using the Burr distribution, referred to as **Multiplier (Burr)**,
- The LogNIG distribution, referred to as **LNIG**,
- The empirical multipliers where the lower quantile ($p = 0.95$) was estimated using the LogNIG distribution, referred to as **Multiplier (LNIG)**,
- Weissman's quantile estimator with the number of exceedances takes as percentages of the sample size. The percentages considered were 10%, 20%, 30%, 40% and 50%. These will be referred to as **EVT (10%)**, **EVT (20%)**, **EVT (30%)**, **EVT (40%)** and **EVT (50%)** respectively, and
- The empirical multipliers where the lower quantile ($p = 0.95$) was estimated using Weissman's quantile estimator with the number of exceedances considered were once again taken as a percentage of the sample size. The percentages of 10%, 20%, 30%, 40% and 50% were also considered for this. These will be referred to as **Multi (10%)**, **Multi (20%)**, **Multi (30%)**, **Multi (40%)** and **Multi (50%)** respectively.

Table 4.23 and Table 4.24 contain the quantile estimate values relative to the true quantiles of Sample1 and Sample2 respectively, whereas the raw values of the quantile estimates obtained for Sample1 and Sample2 are given in Table 4.25 and Table 4.26 respectively. These tables are presented and discussed next.

As mentioned, Table 4.23 and Table 4.24 contain the quantile estimate values relative to the true quantiles of Sample1 and Sample2 respectively. This can be written mathematically as follows. Let q_p denote the assumed true quantile and $\hat{q}_p^*$ any of its estimates, then the values presented in Table 4.23 and Table 4.24 are calculated as $\frac{\hat{q}_p^*}{q_p}$. These values can be interpreted using 1.03 and 0.94 as examples. A value of 1.03 implies that the corresponding estimator over-estimated the true quantile with 3%, whereas a value of 0.94 implies under-estimation of 6%. Furthermore, note that values closest to 1 imply better estimates. In these tables, as well as the tables that follow in this section, the three best and worst performing estimators are indicated in brackets as follows:

- (1), (2) and (3) indicate the best, second best and third best where best implies that the quantile estimate is the closest/nearest to the true quantile, and
- (-1), (-2) and (-3) indicate the worst, second worst and third worst where worst implies that the quantile estimate is the furthest from the true quantile.

Table 4.23: Quantile estimates relative to the true quantile of Sample1

Quantile	TRUE	Burr	LNIG	Multiplier (Burr)	Multiplier (LNIG)	EVT (10%)	EVT (20%)	EVT (30%)	EVT (40%)	EVT (50%)	Multi (10%)	Multi (20%)	Multi (30%)	Multi (40%)	Multi (50%)
95%	1.00	1.37	0.99			0.99	0.99	1.02	1.09	1.22					
99%	1.00	1.09	0.85	1.10	0.80	(2) 0.98	(3) 1.05	1.15	(-2) 1.37	(-1) 1.80	(-3) 0.80	0.80	0.82	0.88	(1) 0.99
99.5%	1.00	0.94	0.76	(1) 0.98	0.71	(3) 0.94	(2) 1.04	1.18	(-2) 1.47	(-1) 2.07	(-3) 0.71	0.71	0.73	0.78	0.87
99.9%	1.00	(1) 1.07	(1) 0.98	1.21	0.88	1.43	1.68	(-3) 2.04	(-2) 2.85	(-1) 4.68	0.88	0.88	0.90	(2) 0.97	1.08

From Table 4.23 it can be seen that the EVT quantile estimates with large values of k perform poorly. This is expected as EVT rely heavily on the asymptotic properties implying that the underlying theories become increasingly more valid as the value of k decreases. This statement is supported by the EVT quantile estimates with smaller values of k producing a number of best performers for $p = 0.99$ and $p = 0.995$. Both the Burr and LogNIG distributions produced accurate estimates for the quantile that is of particular interest (i.e. $p = 0.999$). The multiplier that uses EVT to estimate the lower quantile also produced a best performer for the most extreme quantile.

Table 4.24: Quantile estimates relative to the true quantile of Sample2

Quantile	TRUE	Burr	LNIG	Multiplier (Burr)	Multiplier (LNIG)	EVT (10%)	EVT (20%)	EVT (30%)	EVT (40%)	EVT (50%)	Multi (10%)	Multi (20%)	Multi (30%)	Multi (40%)	Multi (50%)
95%	1.00	2.38	1.01			0.97	0.97	1.01	1.10	1.25					
99%	1.00	(-2) 2.11	0.83	(-3) 1.90	0.81	(2) 1.02	(3) 1.10	1.24	1.56	(-1) 2.14	0.77	0.78	0.80	0.87	(1) 1.00
99.5%	1.00	(-2) 2.14	0.81	1.89	0.80	(3) 1.15	1.27	1.49	(-3) 2.00	(-1) 2.96	0.77	0.77	0.80	(2) 0.87	(1) 1.00
99.9%	1.00	2.55	0.85	2.18	(2) 0.93	1.77	2.07	(-3) 2.66	(-2) 4.13	(-1) 7.30	0.89	0.89	(3) 0.92	(1) 1.01	1.15

From Table 4.24 it can once again be seen that the EVT quantile estimates with large values of k perform poorly, as well as the EVT quantile estimates with smaller values of k produced a number of best performers for $p = 0.99$ and $p = 0.995$. In contrast to the results of Sample1 presented in Table 4.23, Sample2 did not produce promising results for both the Burr and LogNIG distributions. Focusing again on the main quantile of interest (i.e. $p = 0.999$), the multiplier that uses EVT to estimate the lower quantile produced two best performers for the most extreme quantile. In addition, the multiplier based on the LogNIG distribution also performed well.

In Table 4.25 and Table 4.26 the raw values of the quantile estimates of Sample1 and Sample2 are given respectively.

Table 4.25: Quantile estimates of Sample1

Quantile	TRUE	Burr	LNIG	Multiplier (Burr)	Multiplier (LNIG)	EVT (10%)	EVT (20%)	EVT (30%)	EVT (40%)	EVT (50%)	Multi (10%)	Multi (20%)	Multi (30%)	Multi (40%)	Multi (50%)
95%	61456	84004	60886			60709	61087	62592	66878	75104					
99%	275869	301512	234392	304716	220856	(2) 269291	(3) 288643	316965	(-2) 378419	(-1) 496886	(-3) 220216	221587	227044	242593	(1) 272430
99.5%	542354	510720	414466	(1) 530557	384544	(3) 511511	(2) 563395	637412	(-2) 798249	(-1) 1121162	(-3) 383430	385817	395318	422391	474342
99.9%	1585521	(3) 1702256	(1) 1558848	1922702	1393563	2268927	2662089	(-3) 3227860	(-2) 4516766	(-1) 7417613	1389526	1398177	1432606	(2) 1530716	1718984

Table 4.26: Quantile estimates of Sample2

Quantile	TRUE	Burr	LNIG	Multiplier (Burr)	Multiplier (LNIG)	EVT (10%)	EVT (20%)	EVT (30%)	EVT (40%)	EVT (50%)	Multi (10%)	Multi (20%)	Multi (30%)	Multi (40%)	Multi (50%)
95%	69545	165848	70317			67342	67795	70065	76389	87264					
99%	316739	(-2) 666949	263830	(-3) 601594	255065	(2) 324106	(3) 347033	392611	495272	(-1) 676626	244276	245920	254153	277094	(1) 316539
99.5%	552919	(-2) 1185883	448589	1047467	444108	(3) 637647	701123	824721	(-3) 1107824	(-1) 1634715	425323	428184	442519	(2) 482463	(1) 551143
99.9%	1737346	4424461	1472513	3795950	(2) 1609417	3068879	3588930	(-3) 4621344	(-2) 7182589	(-1) 12675269	1541341	1551710	(3) 1603660	(1) 1748413	1997306

Note that the interpretation of the results in Table 4.25 and Table 4.26 are exactly the same as the results of Table 4.23 and Table 4.24, but these tables are given for the reader interested in the actual quantile estimates obtained.

Chapter 5: Quantile estimation of the compound distribution

Under Basel III, banks are allowed to use the Advanced Measurement Approach (AMA) to calculate regulatory capital for operational risk. Currently many banks use the Loss Distribution Approach (LDA) when implementing the AMA. The LDA has been studied and researched extensively and although many authors pointed out deficiencies in this approach (see e.g. Cope et al., 2009 and Embrechts and Hofert, 2011), the approach remains being used in practice. One of the important building blocks of the LDA is to estimate the aggregate loss distribution in each operational risk category (ORC). In particular, an extreme quantile (0.999) needs to be estimated to determine the regulatory capital in that ORC. This has been shown to be notoriously difficult and typically the estimates are inaccurate (see e.g. Cope et al., 2009 and Nešlehová et al., 2006).

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

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

Numerical approximation recursive techniques that can be used to approximate the quantile of the compound distribution include the Panjer recursion (PR) algorithm, introduced by Panjer (1981), and techniques using Fourier inversion and the fast Fourier transform (FFT), proposed by Heckman & Meyers (1983), as well as Feilmeier & Bertram (1987). Grübel & Hermesmeier (1999) have investigated the propagation of discretization errors through compounding and established an improved FFT-based procedure using an exponential change of measure also referred to as exponential tilting. The latter contribution is substantial since it essentially eliminates the so-called aliasing error, which is the fundamental deficit that arises through the use of the discrete Fourier transform. Embrechts & Frei (2010) evaluated and compared PR and FFT (with exponential tilting) and noted two main advantages of FFT, namely that it works with arbitrary frequency distributions and is much more efficient. Embrechts & Frei (2010) further mention that, in situations where both PR and FFT are equally applicable, and where one has to deal with a large number of lattice points, the FFT with suitable tilting is favoured because it is numerically cheaper. The above-mentioned recursive techniques require several input parameters that have to be selected carefully to ensure convergence and are computationally intensive.

Recently, closed-form approximation procedures have been introduced which include the single loss approximation (SLA) and the perturbative approximation (PA). Böcker & Klüppelberg (2005) derived a simple SLA that provides closed-form quantile estimates for the class of sub-exponential distributions (see e.g. Foss et al., 2011), which, according to Fasen & Klüppelberg (2006), contain most of the popular heavy-tailed distributions in practical use today. This makes the SLA method very popular amongst practitioners and a number of academics have done research on improving the method. However, the SLA method only performs accurately if certain conditions hold. Generally, when distributions are heavy-tailed and Poisson intensities are not too high, the approximation works well. Mignola & Ugocioni (2006) showed that the SLA method can lead to an underestimation of the true quantiles, especially for light-tailed loss severity distributions (this is confirmed by the simulation study in Chapter 6). Böcker & Sprittulla (2006) constructed a formula that adds a multiple of the mean to the original formula as an improvement and Sahay et al. (2007) tested the SLA for a Generalised Pareto distribution tail. Degen (2010) provided an analytical framework to analyse the accuracy of the SLA on the basis of the relative approximation error and derived a second order SLA for the class of sub-exponential distributions, based on the work of Omey & Willekens (1986). Hess (2011) evaluated the SLA method for the SAS OpRisk Global data and reported a very good approximation performance of the SLA method. According to Hess (2011), the results of this method are even more reliable than the estimates obtained from a 1 million loss MC simulation. Hannah & Puza (2015) extend the

SLA adding another correction term to account for the effects of the two largest losses (over and above the contribution of the single largest loss).

Hernandez et al. (2014) derived closed-form approximations for high percentiles of the aggregate distribution based on a perturbative expansion, which it is claimed to improve on the above-mentioned methods under certain conditions. The perturbative series introduced differs from previous approximations in that the terms in the series are expressed as a function of the moments of the right truncated distribution for the individual random variables in the sum. These censored moments exist even when the moments of the original distribution (without truncation) diverge. Consequently the same expression is valid for both the infinite and finite mean cases. For high percentiles the perturbative expansion provides a sequence of approximations that, up to certain order, has increasing quality as more terms are included. Beyond that order the convergence of the series deteriorates. Peters et al. (2013) provide a tutorial overview of heavy-tailed loss process modelling in operational risk under Basel III, with discussion on the implications of such tail assumptions for the severity model in a LDA structure. In particular, Peters et al. (2013) consider three key families of risk measures and their equivalent second order asymptotic approximations: VaR, ES and the Spectral Risk Measure. As a last note one should keep in mind that the closed-form approximation methods are computationally very efficient when compared to the recursive methods.

The remainder of this chapter is organised as follows. In Section 5.1 the popular iterative approximation techniques (MC, PR and FFT) are discussed. Then, in Section 5.2, the closed-form approximation techniques (SLA and PA) are given in more detail. Concluding remarks on the theoretical aspects of the approximation techniques are given in Section 5.3. The approximation techniques' numerical properties are given in Chapter 6 which were obtained by way of simulation.

5.1. Popular approximation techniques

It was mentioned that the MC, PR and FFT techniques are well-established methods for approximating the compound distribution within the LDA framework. These approximation techniques are discussed in this section. In Section 5.1.1 the MC technique is discussed. This is followed by giving an overview of both the PR and FFT techniques in Section 5.1.2 and Section 5.1.3 respectively.

5.1.1. Monte Carlo

As stated previously, one of the important building blocks of the LDA is to estimate the aggregate loss distribution in a particular unit of measure or operational risk category (ORC). The p^{th} -quantile (with p large) of the aggregate loss distribution is then used as estimate of the regulatory capital (RC) in the particular ORC. This is done by first estimating the severity and frequency distributions of the annual losses and then by convoluting frequency and severities by means of the random sums method in order to obtain the compound distribution. Monte Carlo (MC) simulation is a well-established technique in the literature for this purpose. The following algorithm can be used to obtain MC approximations of the compound loss distribution's quantiles:

Monte Carlo algorithm:

1. Let $p \in (0,1)$ be some probability level.
2. Generate the number of losses N from the frequency distribution B .
3. Generate N losses X_n with $n = 1, 2, \dots, N$ from the severity distribution F .
4. Calculate the total (aggregate) loss $Y = \sum_{n=1}^N X_n$.
5. Repeat steps 2 - 4 M times to obtain M total losses Y_m with $m = 1, 2, \dots, M$.
6. Let $L = Y_{[pM],M}$ be the $(pM)^{th}$ loss with $Y_{1,M} \leq \dots \leq Y_{m,M} \leq \dots \leq Y_{M,M}$ and $[k]$ the largest integer in k .
7. Repeat steps 5 and 6 S times to obtain S of these quantiles L_s with $s = 1, 2, \dots, S$.
8. The p^{th} -quantile can then be approximated as $q_p^{(MC)} = E[L]$ with $E[L]$ some expected value of L .

Here, the choice of p (in Step 1) does not create any problems since a financial manager will typically know the probability level that is necessary. In the case of OR, the probability level will usually be $p = 0.999$ over a one-year time horizon.

Although the MC simulation technique is relatively easy to implement, it does not come without pitfalls/disadvantages. In order to apply this algorithm, one should know both the frequency and severity distributions (B and F in Steps 2 & 3 respectively). This is generally not true in practice. Estimating the frequency and severity distributions accurately is an essential part of this MC method. This can be done by fitting appropriate distributions to the sample data and choosing the distribution that fits the data the best based on various goodness-of-fit tests. Assuming that these distributions can be estimated accurately, the

next problem is choosing the values of M and S (in Steps 5 & 7). Hess (2011) discusses these choices. The idea is to choose M (together with S) large enough so that $L = Y_{[pM],M}$ (and therefore L_s) are not significantly volatile, but small enough to minimize the time necessary to calculate the quantile approximations L_s accurately. Implementing the MC simulation procedure can easily be done once solutions to these problems are obtained.

One major concern regarding this technique is the computational time necessary to obtain quantile approximations. It is easy to notice from the algorithm that this procedure will be time consuming, especially when the values of M and/or S are large. This can be problematic for financial risk managers when the regulatory capital amount should be updated on a regular basis.

As mentioned earlier, Panjer recursion and Fast Fourier transformation are two other well-known procedures used to approximate the compound distribution. The technique based on Panjer recursion is discussed in the next section, followed by the Fast Fourier transformation.

5.1.2. Panjer recursion

The procedure known as Panjer recursion (PR) was initially introduced by Panjer (1981). This technique can be applied to a certain class of frequency distributions only (Panjer class). This class is defined by

$$p_n = \left(a + \frac{b}{n}\right) p_{n-1}, n \geq 1.$$

Fortunately, the distributions often used in practice to model the frequency fall within this class. These include the Poisson, Binomial and negative Binomial distributions. For the Poisson distribution, $a = 0$ and $b = \lambda$ (in the Panjer class) with λ the intensity of the Poisson distribution.

Another restriction on the PR method is that only discrete severity distributions can be used. Although the distribution used in practice to model the severity are usually continuous, this does not create major problems as the severity distribution can be ‘discretized’ in that the distribution itself can be evaluated at fixed values only. Hess (2011) mentioned that this can be done by using central difference approximation. The density function of the discrete version of the severity distribution is given as

$$f_0 = F\left(\frac{\delta}{2}\right)$$

and

$$f_n = F\left(n\delta + \frac{\delta}{2}\right) - F\left(n\delta - \frac{\delta}{2}\right), \text{ with } \delta > 0 \text{ and } n = 1, 2, \dots$$

Given f_0 , f_n and p_n , the compound distribution then satisfies the following recursion:

$$g_n = \frac{1}{1-af_0} \sum_{j=1}^n \left(a + \frac{bj}{n}\right) f_j g_{n-j}, \quad n \geq 1$$

and

$$g_0 = \sum_{k=0}^{\infty} (f_0)^k p_k.$$

As mentioned, $a = 0$ and $b = \lambda$ for the $Poisson(\lambda)$ distribution so that g_n and g_0 can be simplified as

$$g_n = \sum_{j=1}^n \frac{\lambda^j}{n} f_j g_{n-j}, \quad n \geq 1$$

and

$$g_0 = e^{\lambda(f_0-1)}$$

respectively.

The following algorithm can be used to approximate the compound distribution's quantiles in the case when the Poisson distribution is used to model the frequency.

Panjer recursion algorithm:

1. Let:
 - 1.1. $p \in (0,1)$ denote the probability level,
 - 1.2. $\delta > 0$ denote the increment value,
 - 1.3. λ denote the intensity of the Poisson distribution, and
 - 1.4. F denote the severity distribution.
2. Calculate:
 - 2.1. $f_0 = F\left(\frac{\delta}{2}\right)$, and
 - 2.2. $g_0 = e^{\lambda(f_0-1)}$.
3. Set $G_0 = g_0$.
4. Calculate the following for $n = 1, 2, \dots$:
 - 4.1. $f_n = F\left(n\delta + \frac{\delta}{2}\right) - F\left(n\delta - \frac{\delta}{2}\right)$,
 - 4.2. $g_n = \sum_{j=1}^n \frac{\lambda j}{n} f_j g_{n-j}$,
 - 4.3. $G_n = G_{n-1} + g_n$,
 - 4.4. Stop if $G_n \geq p$.
5. Approximate the p^{th} -quantile of the compound distribution as $q_p^{(PR)} = n\delta$.

Let the true compound distribution be denoted by G , and its p^{th} -quantile by $q_p = G^{-1}(p)$. Furthermore, if $q_p^{(PR)}$ is calculated using the Panjer recursion algorithm presented above, then $q_p^{(PR)} \rightarrow q_p$ as $\delta \rightarrow 0$ with $\delta > 0$ the increment parameter. Choosing δ very small will provide accurate approximations of q_p , but this will be time consuming. In contrast, approximations of q_p can be calculated in a fraction of the time for larger values of δ , but these would/could be unreliable. Note that the size of δ is relative in terms of a couple of quantities, such as F and λ . Thus, there is an optimal balance between computational time and the accuracy of approximation when considering the choice of δ . The same problem exists for the Monte Carlo algorithm presented in the previous section where M and S should be chosen with caution.

Other procedures that provide accurate quantile approximations should therefore also be investigated. The Fast Fourier transformation procedure received a lot of attention in this regard. This technique is discussed in the following section.

5.1.3. Fast Fourier transformation

Fast Fourier transformation (FFT) became very popular as a result of research done by Cooley & Tukey (1965). Since then, a number of researchers focussed on improving the techniques based on FFT. These provided good alternatives to MC techniques regarding computational time. Considering heavy-tailed severities and probability level $p = 0.999$, Embrechts & Frei (2010) mention that the FFT procedure will then most definitely outperform (regarding computational time) the PR method when approximating the compound distribution's p^{th} -quantile. Therefore, the FFT procedure can be considered as a serious alternative to the MC (and PR) techniques, depending on its approximation accuracy.

Let F denote the relevant severity distribution with f its density function. F will typically be continuous in an OR environment. The density function f can be 'discretized' as $f = (f_0, f_1, \dots, f_{T-1})$ with T some truncation point. Note that $\sum_{j=0}^{T-1} f_j = 1$ indicates no truncation error. The discrete Fourier transform (DFT) of f is given by

$$\hat{f} = (\hat{f}_0, \hat{f}_1, \dots, \hat{f}_{T-1})$$

with

$$\hat{f}_j = \sum_{k=0}^{T-1} f_k e^{i2\pi jk/T}, j = 0, 1, \dots, T-1.$$

The discrete version of the original density function $f = (f_0, f_1, \dots, f_{T-1})$ can then be recovered by using the inverse discrete Fourier transform (IDFT) as

$$f_j = \frac{1}{T} \sum_{k=0}^{T-1} \hat{f}_k e^{-i2\pi jk/T}.$$

Grübel & Hermesmeier (1999) suggest using a tilted version of the discrete density function. This tilted version of the discrete density f , denoted by $f^{(\theta)}$, is obtained as

$$f_j^{(\theta)} = e^{-\theta j} f_j$$

with $\theta > 0$ and $j = 0, 1, \dots, T-1$. The DFT can be applied to $f^{(\theta)}$ in order to obtain $\hat{f}^{(\theta)}$, followed by the application of the IDFT to $\hat{f}^{(\theta)}$ which recovers $f^{(\theta)}$. 'Untilting' $f^{(\theta)}$ to obtain f is done by applying

$$f_j = e^{\theta j} f_j^{(\theta)}.$$

Now, let $\mathbb{P}_N$ denote the probability generating function of N and ϕ_K denote the characteristic function of K . The characteristic function of the aggregate loss $Y \sim G$ can be expressed in terms of the characteristic function of the individual loss severities $X \sim F$ (with $Y = \sum_{n=1}^N X_n$) as

$$\phi_Y(t) = \mathbb{P}_N(\phi_X(t)).$$

The compound distribution G (and its density g) can be obtained by applying the DFT (and IDFT) to $\mathbb{P}_N(\phi_X(t))$. A tilted version of g , denoted by $g^{(\theta)}$, can be obtained by

$$g_j^{(\theta)} = \mathbb{P}_N(e^{-\theta j} f_j) = \mathbb{P}_N(f_j^{(\theta)}), j = 0, 1, \dots, T-1.$$

Note that the probability generating function of N when using the *Poisson*(λ) distribution to model the frequency is $\mathbb{P}_N(t) = e^{\lambda(t-1)}$. Applying the DFT to $g^{(\theta)}$ produces $\hat{g}^{(\theta)}$. This is done as

$$\hat{g}_j^{(\theta)} = \mathbb{P}_N(\hat{f}_j^{(\theta)})$$

with $\hat{f}_j^{(\theta)} = \sum_{k=0}^{T-1} f_k^{(\theta)} e^{i2\pi j k l T}$. The IDFT can then be applied to $\hat{g}^{(\theta)}$ in order to recover $g^{(\theta)}$ as

$$g_j^{(\theta)} = \frac{1}{T} \sum_{k=0}^{T-1} \hat{g}_k^{(\theta)} e^{-i2\pi j k l T} = \frac{1}{T} \sum_{k=0}^{T-1} \mathbb{P}_N(\hat{f}_k^{(\theta)}) e^{-i2\pi j k l T}.$$

Finally, the discrete density of the compound distribution g is obtained by ‘untilting’ $g^{(\theta)}$ as

$$g_j = e^{\theta j} g_j^{(\theta)}.$$

The quantiles of the compound distribution can be approximated once its density is obtained (through approximation) using the FFT technique discussed above. Let G and g denote the approximated (using FFT) distribution and density functions respectively with $G(x) = \int_{-\infty}^x g(s) ds$, then the p^{th} -quantile of G is $q_p^{(FFT)} = G^{-1}(p)$. This can be solved by using numerical integration.

Fast Fourier transformation algorithm (with Poisson frequencies):

1. Let:
 - 1.1. $p \in (0,1)$ denote the probability level,
 - 1.2. T be some truncation point,
 - 1.3. $\theta > 0$ be some tilting factor,
 - 1.4. λ denote the intensity of the Poisson distribution,
 - 1.5. F denote the severity distribution, and
 - 1.6. $f = (f_0, f_1, \dots, f_{T-1})$ be the discrete version of the density function of F .
2. Calculate:
 - 2.1. $f^{(\theta)} = (f_0^{(\theta)}, f_1^{(\theta)}, \dots, f_{T-1}^{(\theta)})$ by tilting f using the tilting factor θ as $f_j^{(\theta)} = e^{-\theta j} f_j$,
 - 2.2. $\hat{f}^{(\theta)} = (\hat{f}_0^{(\theta)}, \hat{f}_1^{(\theta)}, \dots, \hat{f}_{T-1}^{(\theta)})$ by applying the discrete Fourier transform,
 - 2.3. $\hat{g}^{(\theta)} = (\hat{g}_0^{(\theta)}, \hat{g}_1^{(\theta)}, \dots, \hat{g}_{T-1}^{(\theta)})$ by using the probability generating function of $N \sim \text{Pois}(\lambda)$ as $\hat{g}_j^{(\theta)} = e^{\lambda(\hat{f}_j^{(\theta)} - 1)}$,
 - 2.4. $g^{(\theta)} = (g_0^{(\theta)}, g_1^{(\theta)}, \dots, g_{T-1}^{(\theta)})$ by applying the inverse discrete Fourier transform, and
 - 2.5. $g = (g_0, g_1, \dots, g_{T-1})$ by untilting $g^{(\theta)}$ using the tilting factor θ as $g_j = e^{\theta j} g_j^{(\theta)}$.
3. The p^{th} -quantile of the compound distribution can then be approximated as $q_p^{(FFT)} = G^{-1}(p)$ with G^{-1} the inverse of the compound distribution G that should be calculated numerically as $G(x) = \int_{-\infty}^x g(s) ds$.

Embrechts & Frei (2010) suggest that if the compound distribution is to be investigated on $\{0, 1, \dots, n\}$, then the truncation level T should be chosen much larger than n to obtain reasonable approximations. It is also mentioned that the tilting parameter $\theta \approx \frac{20}{T}$ would be efficient if T is chosen 'correctly'. Embrechts & Frei (2010) further states that $O(n \log(n))$ operations are required for the FFT, whereas PR requires $O(n^2)$ operations when approximating $g = (g_0, g_1, \dots, g_n)$. Although FFT outperforms PR (regarding computational time), it can still be time consuming when n is very large. Alternative approximation techniques are therefore required. Closed-form approximations are naturally considered in this regard and are discussed in the following section.

5.2. Closed-form approximation techniques

It was mentioned that the popular approximation techniques do not come without any pitfalls and for this reason alternative methods are required. Recently proposed closed-form approximation techniques appear to be viable options and these are discussed in this section. Background on the relevant closed-form approximation techniques are given in Section 5.2.1. Then, both the SLA and PA are given in Section 5.2.2 and Section 5.2.3 respectively.

5.2.1. Overview of the closed-form approximation techniques

Data in the OR environment is usually heavy-tailed where the concept of heavy-tails were defined in Chapter 3. That definition will help in identifying distributions that could be used to estimate the severity distributions. Another important definition for the purpose of discussing the content of this chapter is that of sub-exponential distributions, which is discussed next.

Sub-exponential distribution

Sub-exponential distributions are a special class of heavy-tailed distributions. A distribution F is sub-exponential if

$$\lim_{x \rightarrow \infty} \frac{\bar{F}^N(x)}{\bar{F}(x)} = \lim_{x \rightarrow \infty} \frac{1 - F^N(x)}{1 - F(x)} = \lim_{x \rightarrow \infty} \frac{P(X_1 + X_2 + \dots + X_N > x)}{P(X_1 > x)} = N, \forall N > 1. \quad (5.1)$$

Here, $X_1, X_2, \dots, X_N$ are i.i.d. random variables from F .

If the distribution of individual losses is sub-exponential, then large values of the aggregate loss from N events are dominated by the maximum loss in one of the N events.

It is known that

$$\begin{aligned} P[\max(X_1, X_2, \dots, X_N) > x] &= 1 - P[\max(X_1, X_2, \dots, X_N) \leq x] \\ &= 1 - F^N(x) = 1 - (1 - \bar{F}(x))^N \\ &\approx 1 - (1 - N\bar{F}(x)) \text{ (for large } x) \\ &= N\bar{F}(x). \end{aligned}$$

Using this in the definition, gives

$$\lim_{x \rightarrow \infty} \frac{P(X_1 + X_2 + \dots + X_N > x)}{P[\max(X_1, X_2, \dots, X_N) > x]} = 1, \forall N > 1.$$

A number of theoretical approximation techniques were developed/derived from these definitions. These include the single loss approximation (SLA) and perturbative approximation (PA) techniques.

The next section discusses the mathematics of the SLA methods, followed by a discussion of the PA techniques.

5.2.2. Single loss approximation (SLA) up to second order

Böcker and Klüppelberg (2005) introduced a 0^{th} order SLA technique. Let $X \sim F$ and $Y = \sum_{n=1}^N X_n$ with $Y \sim G$. Then, the following can be obtained from the definition of sub-exponential distributions:

$$\begin{aligned} P(X_1 + X_2 + \dots + X_N > x) &\underset{x \rightarrow \infty}{\approx} P(X_1 > x)N \\ &\Rightarrow 1 - G(x) \underset{x \rightarrow \infty}{\approx} [1 - F(x)]N. \end{aligned}$$

Recall the definition of the p^{th} -quantile q_p of the aggregate distribution G as

$$P(Y > q_p) = 1 - p \Rightarrow P(Y \leq q_p) = p \Rightarrow G(q_p) = p.$$

Substituting $G(q_p) = p$ into the definition of sub-exponential distributions leads to

$$1 - p \underset{x \rightarrow \infty}{\approx} [1 - q_p]N.$$

Solving for q_p leads to $q_p \underset{x \rightarrow \infty}{\approx} F^{-1}\left(1 - \frac{1-p}{N}\right)$.

Assuming the true underlying severity distribution is denoted by F and the frequency if $Poisson(\lambda)$ distributed, then the 0^{th} order single loss approximation (referred to as SLA) of $q_p \sim G$ is

$$\hat{q}_{p,G}^{(SLA)} = F^{-1} \left(1 - \frac{1-p}{\lambda} \right). \quad (5.2)$$

Böcker and Sprittulla (2006) introduced a 1st order SLA technique by adding a mean-correction to the SLA. Applying this in the case of sub-exponential distributions, gives

$$1 - G(x) \approx_{x \rightarrow \infty} [1 - F(x - (N - 1)\mu)] N.$$

Substituting x with q_p and solving for the latter, leads to

$$q_p = G^{-1}(p) \approx_{x \rightarrow \infty} F^{-1} \left(1 - \frac{1-p}{N} \right) + (N - 1)\mu.$$

This mean-corrected single loss approximation (SLAM) of q_p is given as follows when the severity distribution is denoted by F with μ its mean and the frequency distribution is $Poisson(\lambda)$:

$$\hat{q}_{p,G}^{(SLAM)} = F^{-1} \left(1 - \frac{1-p}{\lambda} \right) + (\lambda - 1)\mu.$$

Remark: Note that the mean-corrected single loss approximation (SLAM) can only be used when working with finite mean severity distributions as the mean is used in the approximation.

Degen (2010) derived an improved single loss approximation using second order exponentiality. When the frequency $Poisson(\lambda)$ distributed, and the frequency distribution F has a finite mean and is regularly varying with $\gamma \geq 0$, then the p^{th} -quantile of the compound loss distribution may be approximated by

$$F^{-1} \left(1 - \frac{1-p}{\lambda} \right) + \lambda\mu$$

with μ the finite mean of F . Note that this equation is very similar to the SLAM approximation, but here $\lambda\mu$ is added to the original SLA instead of $(\lambda - 1)\mu$ as in the SLAM approximation.

For infinite mean models with $\gamma > 1$, the p^{th} -quantile of the compound loss distribution may be approximated by

$$F^{-1}\left(1 - \frac{1-p}{\lambda}\right) + (1-p)F^{-1}\left(1 - \frac{1-p}{\lambda}\right) \frac{C_\gamma}{1 - \frac{1}{\gamma}}$$

with $C_\gamma = (1-\gamma) \frac{\Gamma^2(1-\frac{1}{\gamma})}{2\Gamma(1-\frac{2}{\gamma})}$ and Γ the gamma function. Note that when $\gamma = 2$ the ratio

$\frac{\Gamma^2(1-\frac{1}{\gamma})}{2\Gamma(1-\frac{2}{\gamma})} = 0$. For the special case $\gamma = 1$, the p^{th} -quantile of the compound loss distribution

may be approximated by

$$F^{-1}\left(1 - \frac{1-p}{\lambda}\right) + \lambda\mu_F\left(F^{-1}\left(1 - \frac{1-p}{\lambda}\right)\right)$$

with $\mu_F(x) = \int_0^x \bar{F}(s)ds$.

If the underlying severity distribution is denoted by F and the frequency is $Poisson(\lambda)$ distributed, then the p^{th} -quantile approximation of the compound distribution proposed by Degen (2010), referred to as SLAD, can therefore be given by

$$\hat{q}_{p,G}^{(SLAD)} = \begin{cases} F^{-1}\left(1 - \frac{1-p}{\lambda}\right) + \lambda\mu & \text{if } \mu < \infty \\ F^{-1}\left(1 - \frac{1-p}{\lambda}\right) + \lambda\mu_F\left(F^{-1}\left(1 - \frac{1-p}{\lambda}\right)\right) & \text{if } \mu = \infty, \gamma = 1 \\ F^{-1}\left(1 - \frac{1-p}{\lambda}\right) + (1-p)F^{-1}\left(1 - \frac{1-p}{\lambda}\right) \frac{C_\gamma}{1 - \frac{1}{\gamma}} & \text{if } \mu = \infty, \gamma > 1 \end{cases} \quad (5.3)$$

with μ , C_γ and μ_F as above.

Hannah and Puza (2015) extended the SLA by adding another correction term to account for the effects of the two largest losses. Once again, if the frequency is $Poisson(\lambda)$ distributed and the severity distribution F has a finite mean, then solving for c in

$$1 - p = \lambda(1 - F(c - \lambda\mu)) + \frac{1}{2}\lambda^2\left(1 - F\left(\frac{c - \lambda\mu}{2}\right)\right)^2 \quad (5.4)$$

leads to the p^{th} -quantile of the compound distribution. Although this solution does not exist in closed form, Hannah and Puza (2015) propose a solution by setting $c^{(0)} = F^{-1}\left(1 - \frac{1-p}{\lambda}\right) + \lambda\mu$ as an initial guess, and iterating according to

$$c^{(j)} = F^{-1}\left(1 - \frac{1}{\lambda}\left[1 - p - \frac{1}{2}\lambda^2\left(1 - F\left(\frac{c^{(j-1)} - \lambda\mu}{2}\right)\right)^2\right]\right) + \lambda\mu, j = 1, 2, \dots \quad (5.5)$$

with $\lambda\mu$ in (5.5) approximated using Degen's (2010) approximation thereof, i.e. $\lambda\mu \approx \lambda\mu_F\left(F^{-1}\left(1 - \frac{1-p}{\lambda}\right)\right)$ for $\mu = \infty$ with $\gamma = 1$ and $\lambda\mu \approx (1-p)F^{-1}\left(1 - \frac{1-p}{\lambda}\right)\frac{c_\gamma}{1-\frac{1}{\gamma}}$ for $\mu = \infty$ with $\gamma > 1$. The quantile approximation suggested by Hannah and Puza will subsequently be referred to as SLAH.

The perturbative approximation is discussed next.

5.2.3. Perturbative approximation (PA) up to second order

Inspired by the asymptotic relationship in Equation (5.1), Hernández et al. (2014) introduced a k^{th} -order perturbative approximation technique for calculating the p^{th} -quantile of compound distributions. Again, let $X \sim F$ and $Y = \sum_{n=1}^N X_{n,N}$ be the aggregate loss with $Y \sim G$. Split the aggregate loss $Y(\varepsilon) = X_{N,N} + \varepsilon \sum_{n=1}^{N-1} X_{n,N}$ with $X_{1,N} \leq X_{2,N} \leq \dots \leq X_{N,N}$.

The p^{th} -quantile of the aggregate loss $q_{p,G} = G^{-1}(p)$ is expressed as a formal power series in ε as

$$q_{p,G}(\varepsilon) = Q_0 + \sum_{k=1}^{\infty} \frac{1}{k!} Q_k \varepsilon^k.$$

The K^{th} -order perturbative approximation of $q_{p,G}$ is computed by taking the first $K + 1$ terms and setting $\varepsilon = 1$ as

$$q_{p,G}^{(K)} = Q_0 + \sum_{k=1}^K \frac{1}{k!} Q_k$$

leading to the following equations used for the 0^{th} , 1^{st} and 2^{nd} order perturbative approximations of $q_{p,G}$ respectively:

$$q_{p,G}^{(PA0)} = Q_0, \quad (5.6)$$

$$q_{p,G}^{(PA1)} = Q_0 + Q_1, \quad (5.7)$$

$$q_{p,G}^{(PA2)} = Q_0 + Q_1 + \frac{1}{2}Q_2. \quad (5.8)$$

If the frequency is $Poisson(\lambda)$ distributed, then

$$\begin{aligned} Q_0 &= F^{-1}\left(\frac{\lambda + \log(p)}{\lambda}\right), \\ Q_1 &= (\lambda + \log(p))E[X|X < Q_0], \\ Q_2 &= -\left(\lambda f(Q_0) + \frac{f'(Q_0)}{f(Q_0)}\right)(\lambda + \log(p))E[X^2|X < Q_0] - \lambda f(Q_0)Q_0^2. \end{aligned}$$

with f the probability density function of the severity distribution function F .

In the next section some concluding remarks are made on the approximation techniques discussed in this chapter. This is followed by a discussion of the simulation results used to compare these techniques numerically in Chapter 6.

5.3. Concluding remarks

Various techniques exist for approximating high quantiles of the compound distribution. Those discussed in this chapter are the following:

- Iterative procedures:
 - Monte Carlo,
 - Panjer recursion, and
 - Fast Fourier transformation.
- Closed-form procedures:
 - Single loss approximation, and
 - Perturbative approximation.

The iterative procedures generally produce accurate approximations if the number of iterations is sufficiently large. In theory, the actual (true) compound distribution can be obtained if an infinite number of iterations is used during these procedures. Choosing a sufficiently large number of iterations is often impractical as these procedures might take days to obtain the necessary approximations. The closed-form procedures provide good alternatives in this regard. These techniques can be used to obtain various quantities of the

compound distribution in seconds. Therefore the theoretical approximation techniques discussed in Section 5.2 may be viable alternatives to e.g. MC for practical purposes. These techniques, together with FFT, are investigated in the next chapter by comparing their approximation accuracy with one another. This is done by using the MC technique discussed in Section 5.1.1 as the benchmark.

Chapter 6: A (numerical) comparison of quantile approximation techniques for the compound distribution (The MC simulation study)

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

6.1. Simulation design

The parameter values considered in the simulation study are given in Table 6.1, Table 6.2 and Table 6.3 below for each of the three severity distributions respectively. For each of the parameter sets, and for Poisson intensities $\lambda = 10, 20, 50, 100, 200, 500$ and probability levels $p = 1 - \alpha = 0.999, 0.995, 0.99, 0.975, 0.95$ the closed-form approximations were calculated (i.e. SLA, SLAD, SLAH, PA0, PA1 and PA2) and the results then compared with the

corresponding MC simulation results. This consisted of 1,000 VaR estimates, where each VaR estimate was based on 1,000,000 MC repetitions. The minimum, 5% percentile, median, mean, 95% percentile and maximum of the 1,000 MC VaR estimates were then calculated.

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

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

Table 6.1: Burr parameter sets selected for the Monte Carlo simulation study

Burr	η	τ	κ	γ	ρ	$E[X]$
Burr1	1	0.6	5	0.33	-0.2	$< \infty$
Burr2	1	4	0.5	0.5	-2	$< \infty$
Burr3	1	2	1	0.5	-1	$< \infty$
Burr4	1	1	2	0.5	-0.5	$< \infty$
Burr5	1	0.2	10	0.5	-0.1	$< \infty$
Burr6	1	0.6	2	0.83	-0.5	$< \infty$
Burr7	1	2	0.5	1	-2	∞
Burr8	1	1	1	1	-1	∞
Burr9	1	0.5	2	1	-0.5	∞
Burr10	1	0.1	10	1	-0.1	∞
Burr11	1	0.5	1.5	1.33	-0.67	∞
Burr12	1	1.33	0.5	1.5	-2	∞
Burr13	1	0.67	1	1.5	-1	∞
Burr14	1	0.33	2	1.5	-0.5	∞
Burr15	1	0.07	10	1.5	-0.1	∞
Burr16	1	1.8	0.3	1.85	-3.33	∞
Burr17	1	1	0.5	2	-2	∞
Burr18	1	0.5	1	2	-1	∞
Burr19	1	0.25	2	2	-0.5	∞
Burr20	1	0.05	10	2	-0.1	∞
Burr21	1	2.5	0.17	2.35	-5.88	∞
Burr22	1	0.8	0.5	2.5	-2	∞
Burr23	1	0.4	1	2.5	-1	∞
Burr24	1	0.2	2	2.5	-0.5	∞
Burr25	1	0.04	10	2.5	-0.1	∞

Table 6.2: LogNormal parameter sets selected for the Monte Carlo simulation study

LogNormal	μ	σ	γ	ρ	$E[X]$
LNor1	0	1	0	0	$< \infty$
LNor2	0	2	0	0	$< \infty$
LNor3	0	3	0	0	$< \infty$
LNor4	0	4	0	0	$< \infty$
LNor5	0	5	0	0	$< \infty$
LNor6	0	6	0	0	$< \infty$

Table 6.3: LogNIG parameter sets selected for the Monte Carlo simulation study

LNIG	α	β	μ	δ	γ	ρ	$E[X]$
LNig1	3	1	1	1	0.5	0	$< \infty$
LNig2	2	1	1	1	1	0	$< \infty$
LNig3	1	0.5	1	1	2	0	∞

The results of the simulation study are discussed next.

6.2. Simulation results

The performance of the approximation techniques are discussed separately for distributions having finite and infinite means. The results for severity distributions with finite means are discussed in the next subsection, followed by the infinite mean severity distributions.

6.2.1. Finite mean distributions

In Table 6.4 below the results of the smallest ARE and percentage inclusions in the 90% MC CI over all the distributions having finite means and choices of intensities, λ , are given. It was found that the effect of intensities is not dramatic (as will be seen later) and it was included to simplify interpretation. In order to understand the figures in the table, consider a probability level of $p = 1 - \alpha = 0.999$ and the PA1 approximation, then in 11.9% of the cases considered PA1 had the smallest ARE and the approximation was included in the MC CI in 75% of the cases considered.

Table 6.4: Smallest ARE and percentage inclusions in the MC CI over all finite mean distributions

(p)	ARE						MC IC					
	SLA	SLAD	SLAH	PA0	PA1	PA2	SLA	SLAD	SLAH	PA0	PA1	PA2
0.999	2.4%	8.3%	0.0%	2.4%	11.9%	75.0%	44.0%	72.6%	73.8%	41.7%	75.0%	91.7%
0.995	0.0%	0.0%	1.2%	0.0%	8.3%	90.5%	16.7%	14.3%	6.0%	16.7%	48.8%	71.4%
0.99	0.0%	0.0%	1.2%	0.0%	8.3%	90.5%	11.9%	2.4%	3.6%	8.3%	38.1%	59.5%
0.975	0.0%	0.0%	3.6%	0.0%	4.8%	91.7%	6.0%	0.0%	0.0%	0.0%	25.0%	57.1%
0.95	0.0%	7.1%	27.4%	0.0%	2.4%	63.1%	3.6%	1.2%	0.0%	0.0%	19.0%	35.7%
Mean	0.5%	3.1%	6.7%	0.5%	7.1%	82.1%	16.4%	18.1%	16.7%	13.3%	41.2%	63.1%

The results show that the PA2 approximation performs very well in terms of both performance measures. As far as smallest ARE is concerned, PA2 dominates the other

approximation methods at probability levels 0.995, 0.99 and 0.975 and is still better at 0.999, where PA1 is its closest contender, and at 0.95, where SLAH is its closest contender. The percentage inclusions in the MC IC is very good at 0.999 (almost 92%) but deteriorates (as do the other methods) as the probability level decreases to 0.95.

In order to study the performance of the approximation techniques in more detail, Figure 6.1, Figure 6.2 and Figure 6.3 below illustrate some of the more interesting finite mean distributions, viz LNor1, Burr3 and LNig2. In each case the relative error is plotted for each approximation technique against the log-odds of the probability levels α and this is done for intensities 20, 100 and 500. The 5% and 95% MC confidence bands are also depicted as lines in each plot. In all cases PA0 and SLA perform poorly, while the other measures struggle at the lower probability levels, especially at 0.95.

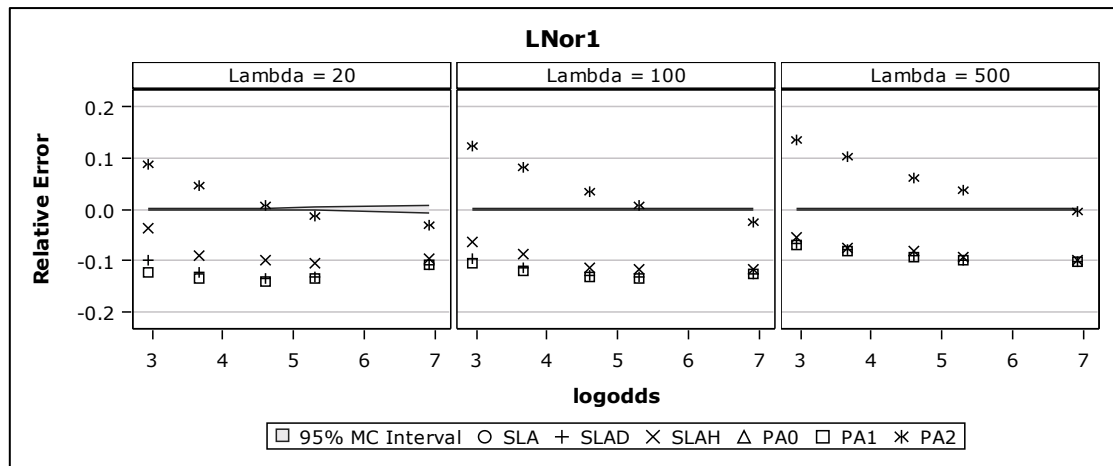


Figure 6.1: Relative error plots for the (very) short-tailed LNor1 distribution (EVI=0)

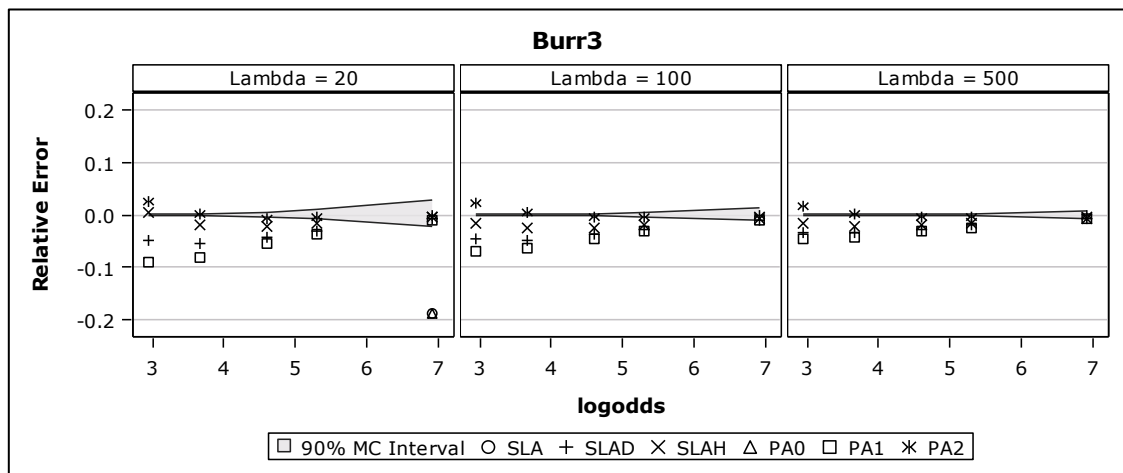


Figure 6.2: Relative error plots for the semi-heavy tailed Burr3 distribution (EVI=0.5)

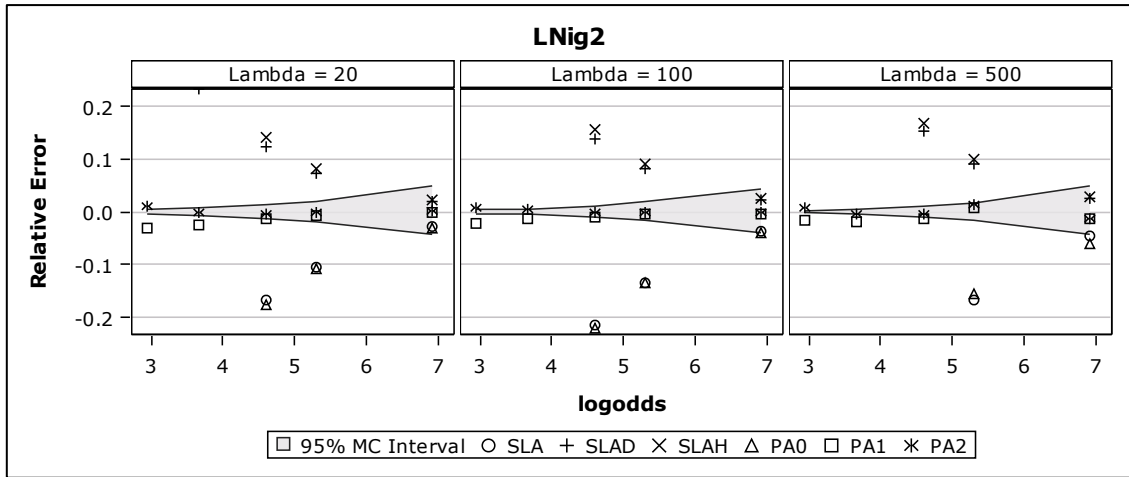


Figure 6.3: Relative error plots for the heavy-tailed LNig2 distribution (EVI=1)

For LNor1 ($\gamma = 0$) the performance of the approximation methods are particularly poor and it is only PA2 that performs reasonably, but only at the higher probability levels. The performance of the higher order approximations (SLAD, SLAH, PA1 and PA2) improves significantly for Burr3 and LNig2 with PA2 clearly the preferred one. From an operational risk perspective, the focus is especially on the performance of the approximation methods at a probability level of 0.999. To this end Figure 6.4 was constructed below, which shows the relative error and approximation techniques for all finite mean distributions. It should be clear that, with the exception LNor1, the performance of the second order approximations is very good for all finite mean distributions considered. Given the performance of PA2 as observed in Figure 6.4 and Table 6.4, PA2 is the preferred choice.

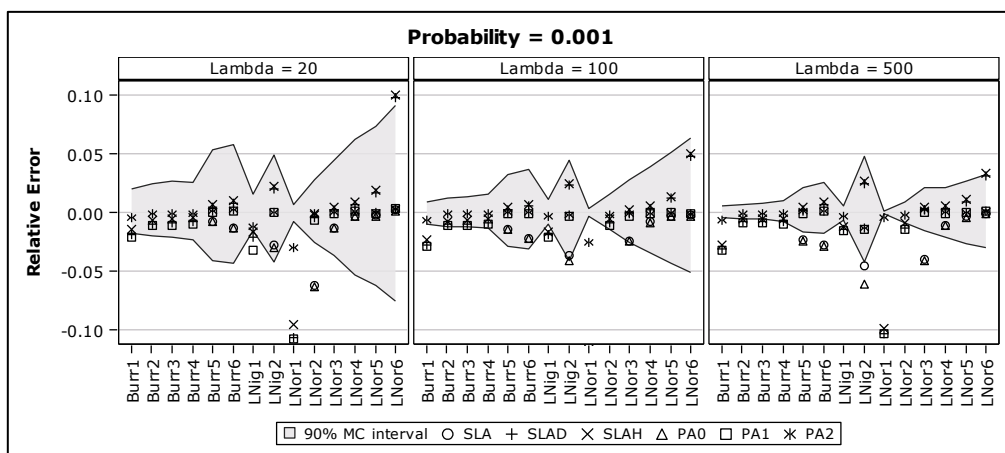


Figure 6.4: Relative error plots for all finite mean distributions at probability level 0.999

As mentioned previously, other approximation techniques to consider include Panjer recursion and the Fast Fourier transform of which the latter is often preferred (see e.g.

Embrechts & Frei, 2010). Both these techniques require the input of a numerical expert necessitated by the required intelligent selection of input parameters. Therefore application of these recursive approximation techniques is time consuming as opposed to the closed-form solutions. Assuming the notation of Embrechts & Frei (2010), four input parameters (M, h, k, θ) need to be specified for the FFT algorithm. In order to automate the selection of input parameters a rule of thumb was applied in the following way. Take $k = 24$ and solve for h in $k = \text{int}\left(\frac{\ln(1.1PA2/h)}{\ln(2)}\right)$ then take $M = 2^k$ and $\theta = M/20$. It was found that this rule of thumb provided good choices for approximating the cases considered here.

Below the FFT (with tilting) is compared with the PA2 approximation for all finite mean distributions considered. In Table 6.5 below the results of the smallest ARE and percentage inclusions in the 90% MC CI over all the above-mentioned distributions and all intensities are given.

Table 6.5: Smallest ARE and percentage inclusions in the MC CI over selected distributions

(p)	ARE		MC CI	
	FFT	PA2	FFT	PA2
0.999	72.6%	27.4%	100.0%	91.7%
0.995	75.0%	25.0%	97.6%	71.4%
0.99	88.1%	11.9%	97.6%	59.5%
0.975	96.4%	3.6%	97.6%	57.1%
0.95	94.0%	6.0%	96.4%	35.7%
Mean	85.2%	14.8%	97.9%	63.1%

For the distributions considered the FFT performs very well, and outperforms PA2 at all probability levels, except at 0.999. In order to study the performance in more detail, the relative error graphs for LNor1, Burr3 and LNig2 are depicted in Figure 6.5, Figure 6.6 and Figure 6.7 respectively.

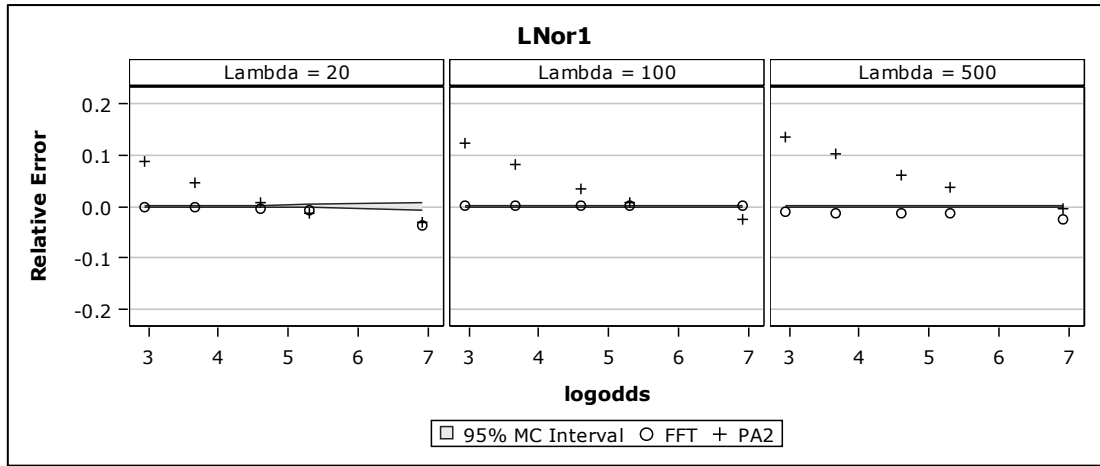


Figure 6.5: Relative error plots for the short-tailed LNor1 distribution ($EVI=0$)

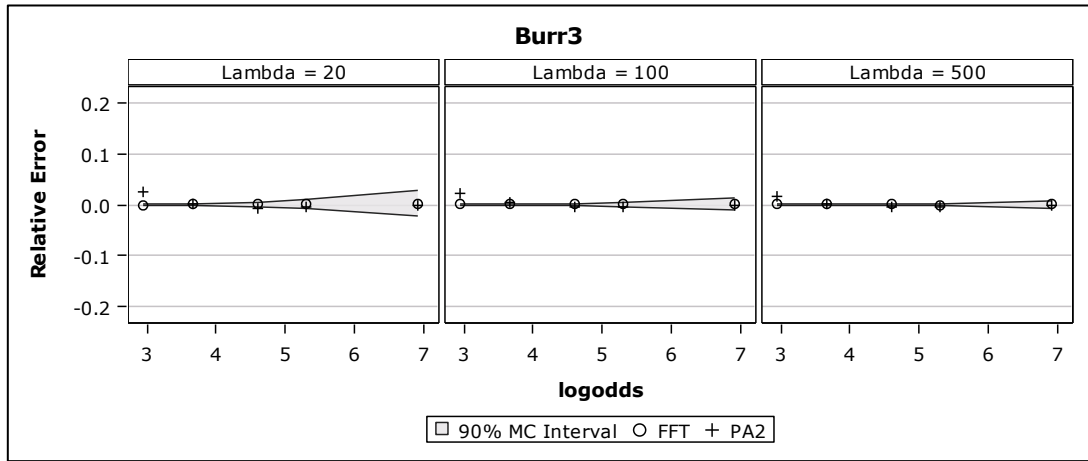


Figure 6.6: Relative error plots for the semi-heavy tailed Burr3 distribution ($EVI=0.5$)

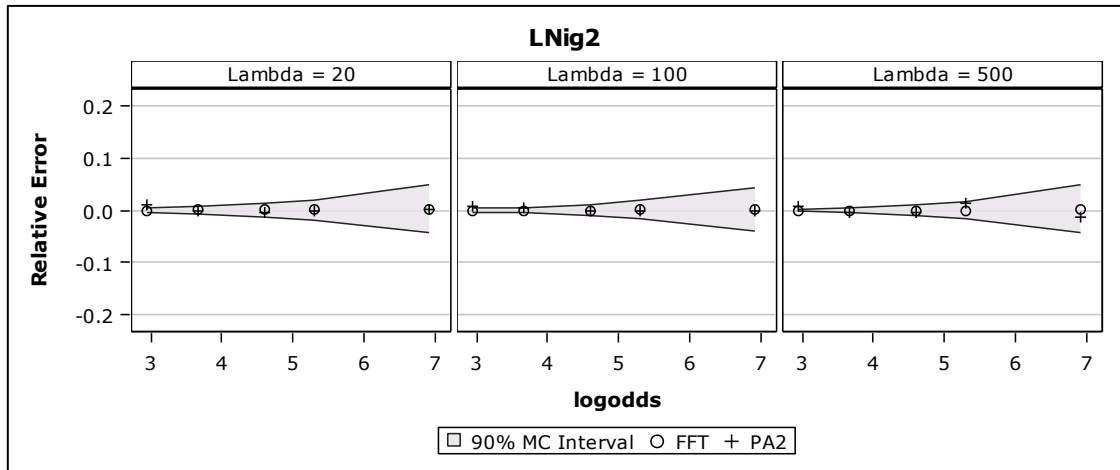


Figure 6.7: Relative error plots for the heavy-tailed LNig2 distribution ($EVI=1$)

The excellent performance of FFT is evident in all three figures. PA2 and FFT are close contenders at a probability level of 0.999 and it is noticeable that the performance of PA2 relative to FFT is much better for Burr3 and LNig2 than for the short-tailed LNor1. The

results for the other distributions confirm that PA2 only really struggles to approximate LNor1, and is a close contender for the others, especially for distributions with higher EVI's and at a 0.999 probability level.

The results of the infinite mean severity distributions are given next.

6.2.2. Infinite mean distributions

It is well-known that the FFT and Panjer approximation methods break down for infinite mean distributions with large EVIs, therefore these methods are excluded from this analysis. In Table 6.6 below the results of the smallest ARE and percentage inclusions in the 90% MC CI over all the distributions having infinite means and over all intensities are given. The performance of PA2 is remarkable since it obtain 100% inclusion in the MC CI for all probability levels, except 0.95 where it achieves the best performance (99.2%). It is interesting that as far as smallest ARE is concerned, PA2 is challenged by all approximation methods at the higher probability levels, and especially at 0.999. This is to be expected since the asymptotic theory suggests that all approximations should perform well for large EVI and large intensities at high quantiles.

Table 6.6: Smallest ARE and percentage inclusions in the MC CI over infinite mean distributions

(p)	ARE						MC CI					
	SLA	SLAD	SLAH	PA0	PA1	PA2	SLA	SLAD	SLAH	PA0	PA1	PA2
0.999	11.9%	10.3%	23.0%	32.5%	16.7%	11.1%	90.5%	89.7%	89.7%	100%	100%	100%
0.995	10.3%	9.5%	0.8%	15.1%	26.2%	44.4%	66.7%	84.9%	81.7%	80.2%	100%	100%
0.99	5.6%	8.7%	0.8%	2.4%	41.3%	46.8%	61.1%	78.6%	51.6%	69.8%	100%	100%
0.975	0.0%	3.2%	1.6%	0.0%	27.8%	67.5%	42.1%	45.2%	5.6%	31.0%	84.1%	100%
0.95	1.6%	5.6%	0.0%	0.0%	15.1%	79.4%	23.8%	38.1%	2.4%	9.5%	79.4%	99.2%
Mean	5.9%	7.5%	5.2%	10.0%	25.4%	49.8%	56.8%	67.3%	46.2%	58.1%	92.7%	99.8%

In order to study the quality of the approximations in more detail, it was decided to select the Burr12, Burr13 and Burr15 distributions having the same EVI ($\gamma = 1.5$) but having different values of the second order parameter ρ . The results are depicted in Figure 6.8, Figure 6.9 and Figure 6.10 for Burr12, Burr13 and Burr15 respectively. The results show that PA1 and PA2 is almost 'spot-on' in all cases, while SLA, SLAD and SLAH tend to overestimate slightly and PA0 underestimates slightly at especially the lower probability levels. Note how the MC CI increases as the probability level increases. As stated before, the probability level of interest when analysing the performance of the approximation methods is at 0.999. To this

end Figure 6.11 was constructed and is presented below, which shows the relative error and performance of the approximation techniques for all infinite mean distributions.

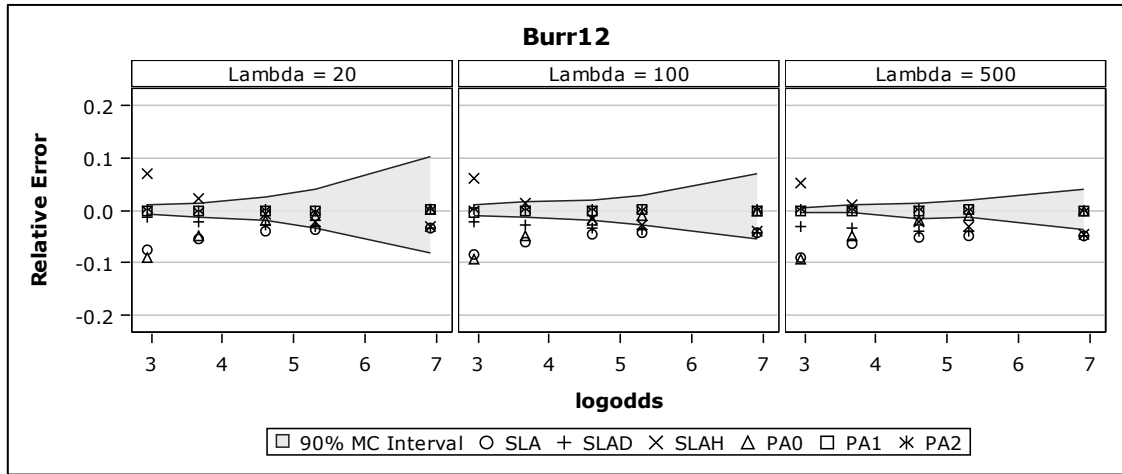


Figure 6.8: Relative error plots for the Burr12 distribution ($EVI=1.5$, $Rho=-2$)

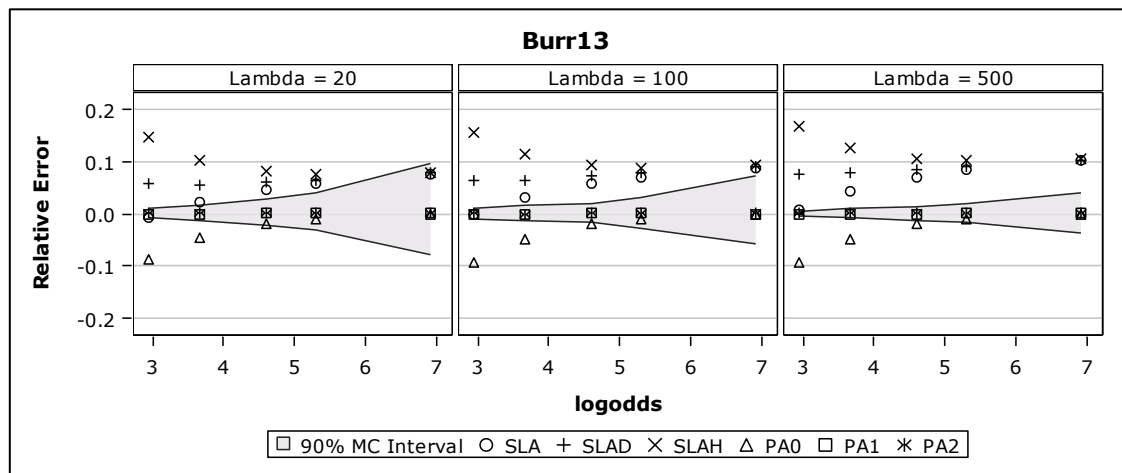


Figure 6.9: Relative error plots for the Burr13 distribution ($EVI=1.5$, $Rho=-1$)

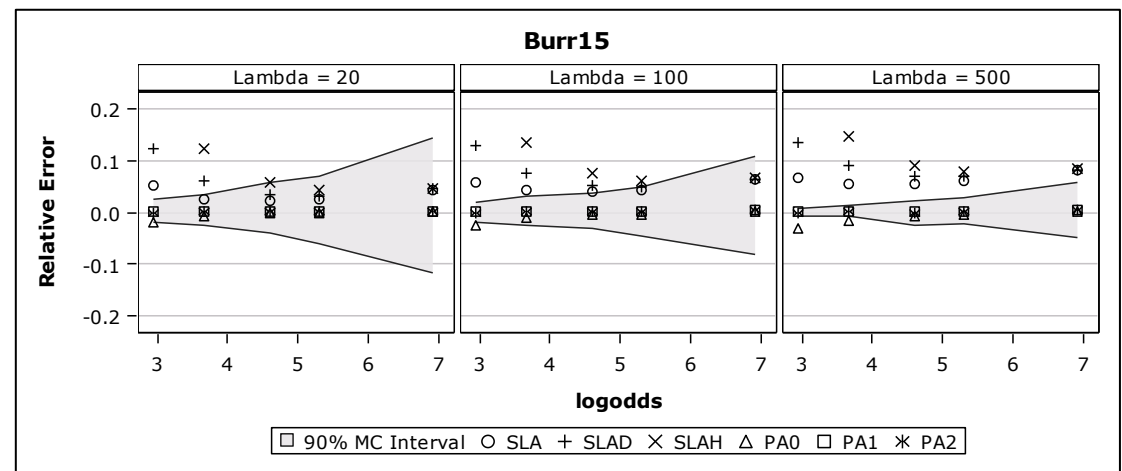


Figure 6.10: Relative error plots for the Burr15 distribution ($EVI=1.5$, $Rho=-0.1$)

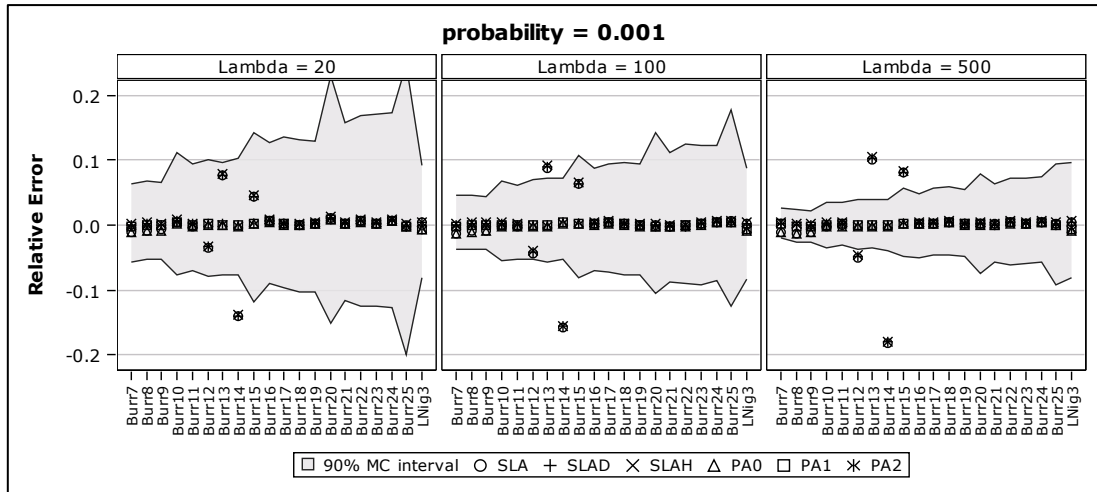


Figure 6.11: Relative error plots for all infinite mean distributions at probability level 0.999

From Figure 6.11 it is clear that all the approximation methods, with the possible exception of SLA, SLAD and SLAH (for Burr12, Burr13, Burr14 and Burr15), perform very well.

6.3. Summary of the approximation techniques

The performance of the most recently proposed approximation methods for the quantiles of the compound distribution was analysed. This was done by assuming a Poisson frequency distribution and several severity distributions. It was found that the second order perturbative closed-form approximation method of Hernandez et al. (2014) performed very well in most cases considered. If the goal is to approximate an extreme quantile of the compound distribution (such as at a probability level of 0.999), then this method is recommended. However, for short-tailed distributions and at less extreme quantiles, the FFT method could be considered using the proposed rule of thumb for the choice of input parameters (which depends on PA2). In general, in an operational risk context, the PA2 approximation is recommended because of its accuracy, clear computational speed advantage and the fact that it requires no input parameters.

Chapter 7: Conclusions

This final chapter is used to make some concluding remarks of this study in Section 7.1, discussing the main contributions of the thesis in Section 7.2, and lastly, to give a brief discussion in Section 7.3 of further research relevant to the topics covered in the thesis.

7.1. Concluding remarks

It was mentioned in Chapter 1 that operational risk (OR) is a relatively new risk category, compared to the well-established credit risk (CR) and market risk (MR) categories. All three of these risk categories are regulated locally by way of the BCBS guidelines for financial risk management. The latest published Basel accord, the BASEL III accord, focusses mostly on the management of OR. As is the case with CR and MR, OR is quantified using various risk measures. It was shown that the most widely used risk measures all rely heavily on the quantiles of the distribution underlying the loss data. Therefore, the focus of this thesis was on investigating different quantile estimators rather than the different risk measures.

According to BASEL III, banks can use the more sophisticated advanced measurement approach (AMA) to calculate its regulatory capital (which is equal to the 99.9th quantile of the underlying distribution), as opposed to the simpler approaches such as the basic indicator approach and the standardised approach. Most banks use the loss distribution approach (LDA) under the AMA. The LDA comprises of estimating the regulatory capital as an extremely high quantile of the aggregate distribution where the aggregate distribution is obtained by convoluting both the severity and frequency distributions.

The aggregate distribution cannot be obtained exactly and must therefore be approximated. Various well-known approximation techniques, such as Monte Carlo (MC) simulation, Panjer recursion (PR) and fast Fourier transformation (FFT), can be used to approximate this aggregate distribution. In addition, more recent approximation techniques were proposed, such as the single loss distribution (SLA) and perturbative approximation (PA). Both the SLA and PA were investigated with approximation orders up to two. Each of these approximation methodologies has their advantages and disadvantages. It was shown in Chapter 5 and Chapter 6 that the perturbative approximation (of order 2) performs generally the best for estimating regulatory capital (at the 99.9th probability level) when the underlying severity distribution has an infinite mean. This was also the case for quantiles with lower probability levels such as both the 99.5th and 99th. For underlying severity distributions with finite means

it was found that the well-known FFT approximation performs best. As was mentioned in Chapter 5, an experienced user of the FFT is required to determine the parameter values used in this technique. A suggestion was made in Chapter 6 where the second order PA can be used to inform the FFT user of the parameter values to be selected which helps in improving the accuracy of the FFT approximation.

As mentioned, the aggregate distribution is a combination of both the frequency and severity distributions. It is well-known that the frequency distribution does not have an adverse effect on the aggregate distribution and therefore the Poisson distribution was used throughout this thesis to model the frequency distribution. There are two approaches to modelling the severity distribution. The first approach is the full-parametric approach where a statistical distribution is fitted to the entire sample of loss data which is then used to estimate the desired high quantile. The second approach is referred to as the splice distribution approach where the sample is separated into two sub-samples using some splicing point. The one sub-sample comprises of the body-part of the original sample from which the empirical distribution is obtained and the other sub-sample consists of the tail-part of the original sample to which a statistical distribution is fitted. The distribution fitted to the tail-part of the original sample is then used to estimate the desired high quantile (this method was referred to as the EVT approach in the thesis). The conclusions drawn from both the full-parametric and EVT approaches are discussed next, starting with the full-parametric approach.

Distributions traditionally used in the full-parametric approach are the Burr, LogNormal, Pareto and Normal Inverse Gaussian (NIG) distributions. While many of the distributions used in practice are heavy-tailed, some of them are not, even though loss data in OR possesses heavy tails. One such distribution that does not have heavy tails is the NIG distribution. Despite the NIG distribution having semi-heavy tails, it remains popular in OR quantification due to its large number of parameters (four) which gives it the ability to be flexible. Although these distributions are widely used, none of them perform uniformly the best when used to estimate the high quantiles used in practice. It is for this reason that alternative distributions are desired. This led to considering the LogNIG distribution where the LogNIG was derived from the NIG distribution just as the LogNormal is derived from the Normal distribution. This 'exponentiation' of the NIG distribution to the LogNIG distribution gave rise to another parameter-rich distribution (as is the case of the NIG) in the form of the LogNIG distribution. In addition to having four parameters, the LogNIG distribution also has heavy tails. The LogNIG distribution was then compared to the most widely used distributions by assessing their numerical properties obtained from a simulation study. The distributions that were compared to the LogNIG are the Burr and LogNormal distributions,

together with the LogGamma where the LogGamma is derived from the Gamma distribution just as the LogNormal (and LogNIG) was derived from the Normal (and NIG) distribution. The reason for the inclusion of the LogGamma was to have at least two other 'Log' type of distributions. The LogPhase distribution proposed by Ahn et al. (2012) was also considered but was found to be somewhat difficult to fit to samples of data. It is worth mentioning that the LogPhase distribution will be considered in future research. Returning to the simulation study, the distributions considered were fitted to simulated data that was generated from Burr distributions only making the Burr the benchmark distribution for this simulation study. It was found that the Burr distribution generally outperformed all of the other distributions, which was expected. Analyzing the simulation results without the Burr distribution gave rise to the LogNIG distribution being superior when aiming to minimize the bias. The LogGamma performed overall very well in the case of reducing the root mean squared error.

As mentioned, the EVT approach consists (for the purpose of this study) mainly of estimating high quantiles by considering the tail-part of sample data only. Well-known quantile estimators under this approach are those derived from extreme value theory (EVT). These are both the first and second order Weissman quantile estimators where the second order Weissman estimator is a bias-reduced version of the first order Weissman estimator. Other quantile estimators derived from EVT also exist, but this study limited the EVT-based quantile estimators to just the well-known and widely used Weissman quantile estimator. The EVT-based estimators have two major drawbacks. The first is the choice of splicing point. The second drawback is that these estimators rely heavily on asymptotic assumptions where large samples are required, which is not the case in ORM. Other quantile estimators are therefore desired under this approach. This led to the development and investigation of what is called the multiplying factors (or multipliers). The idea of the multipliers is to estimate a quantile at a probability level lower than the desired level which is then extrapolated to the desired quantile by multiplying the lower quantile with an appropriate multiplying factor. A number of multiplier sets was developed using different methodologies. The first method relied on asymptotic properties where it was assumed that the lower quantile can be estimated perfectly. These multiplying factors were referred to as asymptotic multipliers. The second method consisted of developing the multiplying factors using simulation where the lower quantiles were estimated using the Weissman quantile estimators. Using, for example, the Weissman estimators for the lower quantiles would produce factors that are developed specifically in a practical setting, i.e. empirical multipliers. It was found that the multipliers (specifically the empirical multipliers) generally outperformed the Weissman quantile estimator for both the first and second order environments. Zooming in on the results revealed that, as expected, the Weissman quantile estimators performed better for the larger

sample sizes (100, 150 and 200) considered even though these are considered small samples in the EVT environment. It is well-known that loss data in ORM can typically be of sizes 50, 20 and even smaller than 20. This is the cases where the multipliers were far superior to the EVT-based quantile estimator. Although this study focused on small samples, it was shown in the severity distribution's empirical study in Section 4.3 that the multipliers can be beneficial for larger samples when working with real loss data. The multipliers loosen the problem associated with the choice of the splicing point when using the Weissman quantile estimator to estimate the lower quantile. This challenge can even be avoided by using, for example, the full-parametric approach to estimate the lower quantile.

This section was used to give concluding remarks regarding the findings of the study discussed in this thesis. The following section will highlight the contributions made by this thesis.

7.2. Thesis contributions

A summary of the three main contributions made by the thesis is given in this section. These are given in the following three paragraphs where each paragraph is dedicated to a specific contribution. The first paragraph deals with quantile approximation techniques of the severity distribution where the full-parametric approach is preferred. The second paragraph also addresses the estimation of the severity distribution's quantiles, but here the focus shifts to the splice distribution approach. The third and final paragraph highlights the thesis contributions regarding approximation techniques that relate to the aggregate distribution. Note that the problem of estimating extreme quantiles (such as the 99.9th, 99.5th and 99th) of both the severity and aggregate distributions are not unique to operational risk management only. Various fields of academia and practice dedicate much time to this topic.

Estimating high quantiles of the severity distribution under the full-parametric approach in ORM is traditionally done using heavy-tailed distributions such as the Burr, LogNormal and LogGamma distributions. These distributions were compared to the newly proposed LogNIG distribution using numerical results. It was found that the LogNIG distribution most certainly outperformed the LogNormal distribution in most cases. The LogNIG distribution also outperformed the LogGamma from a bias-reduction point of view. As expected, the LogNIG did not fair well against the Burr distribution. This is possibly because the sample data was generated from Burr distributions only.

Estimating high quantiles of the severity distribution under the EVT approach in ORM is traditionally done using the Weissman quantile estimator. Due to the heavy-tailed and small sample nature of operational risk loss data estimates of extreme quantiles can be very volatile. This is especially true when using quantile estimators derived from extreme value theory as these estimators rely mainly on asymptotic assumptions. The idea of multiplying a lower (and less volatile) quantile with some factor to obtain an estimate of the extreme quantile was investigated. This class of estimators was referred to as multipliers. It was found that the multipliers generally outperformed the Weissman quantile estimator for a wide range of parameters selected for the simulation study. The multipliers' ability to estimate the desired extreme quantiles increased greatly for small samples (which is the area of most concern for operational risk managers) when compared to that of the Weissman estimators.

Approximating extreme quantiles of the aggregate distribution under the LDA is traditionally done by banks using MC simulation, although other approximation techniques are discussed and proposed in the literature. All of these techniques have their advantages and disadvantages and it is for this reason that all of relevant techniques were investigated and compared with one another. This was done by comparing the techniques' abilities to approximate, amongst other, the 99.9th quantile of the aggregate loss distribution by way of simulation. Taking the numerical results of the simulation study into consideration, together with the pitfalls of each technique, it was found that the perturbative approximation of the second order is generally the best approximation technique to use when the underlying severity distribution has an infinite mean. When the underlying severity distribution has a finite mean, then the fast Fourier transformation approximation technique generally performs best where the second order perturbative approximation is used to inform the parameter values used in the fast Fourier transformation approximation.

7.3. Further research

Since the content of this thesis was presented separately according to the various approaches considered, the planned future research will also be discussed separately. Firstly, research pertaining to quantile estimation techniques of the severity distribution under the full-parametric approach will be discussed. This is followed by further research ideas regarding the extreme value theory approach to estimating extreme quantiles of the severity distribution. Lastly, ideas surrounding approximation techniques of the aggregate distribution's quantiles will be discussed.

Some promising distributions proposed for estimating the severity distribution's quantiles under the full-parametric approach are the LogPhase type of distributions proposed by Ahn et al. (2012) and the g-and-h distribution proposed by Degen et al. (2007). These will be compared numerically to the LogNIG distribution, as well as the distributions most widely used in practice.

Two different classes of quantile estimators relating to the EVT approach were discussed in this thesis. These are the Weissman quantile estimators derived from extreme value theory and the multipliers where the lower quantiles were estimated using both the first and second order Weissman quantile estimator. For the EVT-based estimators it was mentioned that the choice of k (the number of threshold exceedances) was varied as a percentage of the sample size, i.e. $k = \omega n$ with $\omega = 0.1, 0.15, 0.2, \dots$. Berning (2010) proposed using $k = 2n^{2/3}$ which could possibly improve the Weissman estimator's ability to estimate the extreme quantiles. An alternative to the Weissman quantile estimator is also an extreme value theory based estimator known as a mean-of-order p (MOP) estimator proposed by Gomes et al. (2013). A reduced-bias version of the MOP estimator was also proposed and is referred to as the RBMOP estimator. A number of ideas that could potentially increase the multipliers' performance is given next.

The empirical multipliers were developed using both the MinMax and MinAve of the mean squared error (MSE) statistical measure. Instead, applying these arguments to the median absolute deviation (MAD) rather than the MSE could possibly lead to different multiplying factors due to the MSE being highly influenced by the presence of outliers. Also, the Weissman quantile estimator was used to estimate the lower quantiles during the development of the empirical multipliers. Alternatively, estimating the lower quantile using the full-parametric approach (i.e. fitting, for example, the LogNIG distribution to the sample) could lead to an improvement in the multipliers' performance. This is possible since the lower quantile will generally be estimated more accurately under the full-parametric approach when compared to the EVT approach. Recall that the asymptotic multipliers were developed as a function of both the higher and lower quantiles, i.e. $C_{p_1, p_0}^{(AM)} = f(p_1, p_0)$ with f some function. The empirical multipliers added the sample size n as another factor affecting the multipliers, i.e. $C_{p_1, p_0, n}^{(EM)} = f(p_1, p_0, n)$. More advanced and sophisticated research regarding the multipliers would include the incorporation of the extreme value index γ as a factor of the multipliers' values, i.e. $C_{p_1, p_0, n, \gamma}^{(EM)} = f(p_1, p_0, n, \gamma)$. This proves to be promising as increasing values of the extreme value index implies heavier tails and consequently, larger ratios between the higher and lower quantiles, i.e. larger multiplying factors. One proposal

would be to develop separate sets of multipliers for distributions with EVIs falling into predetermined ranges. Furthermore, the multipliers will be investigated for cases where the data is not i.i.d. but rather depends on various covariates. Also, the multipliers' ability to estimate extreme quantiles in cases where data is censored will also be investigated in the future.

It was mentioned that the aggregate distribution relies mainly on the severity distribution. In practice, estimating the severity distribution still remains challenging, especially when working with small samples. A deviation in the estimate of the severity distribution from the true underlying severity distribution can have an adverse effect on the extreme quantiles of the eventual aggregate distribution, irrespective of the technique used to approximate the aggregate distribution. It is therefore proposed to investigate the use of the multipliers when approximating extreme quantiles of the aggregate distribution. Due to the computational intensity of such a simulation study, it is proposed to estimate the higher quantiles using the second order perturbative approximation when the true underlying severity has an infinite mean, and the fast Fourier transformation approximation (where the second order perturbative approximation informs the researcher of the parameter values used in the fast Fourier transform) when the severity distribution's mean is finite. Monte Carlo simulation can be used for the lower quantile with the number of simulations efficiently low.

References

ABRAMOWITZ, M. & STEGUN, I.A. 1972. *Handbook of mathematical functions*. Dover Publications, New York.

AHN, S., KIM, J.H.T. & RAMASWAMI, V. 2012. A new class of models for heavy tailed distributions in finance and insurance risk. *Insurance: Mathematics and Economics*, 51:43-52.

ARTZNER, P., DELBAEN, F., EBER, J.M. & HEATH, D. 1999. Coherent measures of risk. *Mathematical Finance*, 9(3):202-228.

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

BALKEMA, A.A. & DE HAAN, L. 1974. Residual life time at great age. *Annals of Probability*, 2:792-804.

BARNDORFF-NIELSEN, O.E., BLAESILD, P., JENSEN, J.L. & SORENSEN, M. 1985. The fascination of sand. *A celebration of Statistics*. Springer-verslag, 57-87. New York.

BARNDORFF-NIELSEN, O.E. 1997. Normal inverse Gaussian processes and stochastic volatility modelling. *Scandinavian Journal of Statistics*, 24:1-13.

BAUD, N., FRACHOT, A. & RONCALLI, T. 2002. Internal data, external data and consortium data for operational risk measurement: How to pool data properly? Working paper. Available at: <http://www.thierry-roncalli.com/download/oprisk-data-light-version.pdf>

BCBS. 2000. Principles for the Management of Credit Risk. September.

BCBS. 2006. International Convergence of Capital Measurement and Capital Standards: A Revised Framework Comprehensive Version. Available at: <http://www.bis.org/publ/bcbs128.pdf>

BCBS. 2011. Operational risk: supervisory guidelines for the advanced measurement approaches. *Bank for International Settlements*. Available at: <http://www.bis.org/publ/bcbs196.pdf>

BEIRLANT, J., GOEGEBEUR, Y., SEGERS, J., TEUGELS, J., DE WAAL, D. & FERRO, C. 2004. *Statistics of Extremes: Theory and Applications*. West Sussex: Wiley.

BERAN, J. & SCHELL, D. 2012. On robust tail index estimation. *Computational Statistics and Data Analysis*, 56:3430-3443.

BERNING, T.L. 2010. Improved estimation procedures for a positive extreme value index. Stellenbosch: SU. (Thesis - PhD), 240 p.

BESSIS, J. 2011. *Risk Management in Banking*, 3rd edition. John Willey & Sons. ISBN: 978-0-470-68985-1. December.

BLAESILD, P. & SÖRENSEN, M.K. 1992. 'hyp' - A computer program for analyzing data by means of the hyperbolic distribution. Research Report No 248, Department of Theoretical Statistics, University of Aarhus, Denmark.

BÖCKER, K. & KLÜPPELBERG, C. 2005. Operational VaR: a closed-form approximation. *Risk* 18(12):90-93.

BÖCKER, K. & SPRITTULLA, J. 2006. Operational VaR: meaningful means. *Risk* 19(12):96-98.

BRAZAUSKAS, V. & SERFLING, R. 2000. Robust estimation of tail parameters for two-parameter Pareto and exponential models via generalized quantile statistics. *Extremes* 3(3):231-249.

CHANG, Y. & HUNG, M. 2010. An EM algorithm for multivariate NIG distribution and its application to value-at-risk. *International Journal of Information and Management Sciences*, 21:265-283.

CHERNOBAI, A.S., RACHEV, S.T. & FABOZZI, F.J. 2007. *Operational Risk: A Guide to Basel II Capital Requirements, Models and Analysis*. John Willey & Sons. ISBN: 978-0-471-78051-9. July.

COLES, S. 2001. *An Introduction to Statistical Modeling of Extreme Values*. Springer. ISBN: 978-1-4471-3675-0.

COOLEY, J.W. & TUKEY, J.W. 1965. An algorithm for the machine calculation of complex Fourier series. *Mathematical Computation* 19:297-301.

COPE, E.W., MIGNOLA, G., ANTONINI, G. & UGOCCIONI, R. 2009. Challenges and pitfalls in measuring operational risk from loss data. *The Journal of Operational Risk*, 4(4):3-27.

CROUHY, M., GALAI, D. & MARK, R. 2000. *Risk Management*. McGraw-Hill.

DANIELSSON, J., HARTMANN, P. & DE VRIES, C.G. 1998. Extreme returns, value-at-risk and the Basle 'multiplication factor'. *Risk*, 11:101-103.

DE HAAN, L. & FERREIRA, A. 2006. *Extreme Value Theory: An Introduction*. Springer. ISBN 978-0-387-34471-3.

DE JONGH, E., DE JONGH, D., DE JONGH, P.J. & VAN VUUREN, G. 2013. A review of operational risk in banks and its role in the financial crisis. *South African Journal of Economic and Management Sciences*, 4:364-382.

DE JONGH, P.J., DE WET, T., PANMAN, K. & RAUBENHEIMER, H. 2016. A Simulation Comparison of Quantile Approximation Techniques for Compound Distributions popular in Operational Risk. Paper accepted for publish in the *Journal of Operational Risk*, March edition.

DE JONGH, P.J., DE WET, T., RAUBENHEIMER, H. & VENTER, H. 2015. Combining scenario and historical data in the loss distribution approach: A new procedure that incorporates measures of agreement between scenarios and historical data. *Journal of Operational Risk*, 10(1):1-31.

DE WET, T., GOEGGEBEUR, Y. & MUNCH, M. R. 2012. Asymptotically unbiased estimation of the second order tail parameter. *Statistics and probability letters*, 82(3):565-573.

DEGEN, M., EMBRECHTS, P. & LAMBRIGGER, D.D. 2007. The Quantitative Modeling of Operational Risk: Between g-and-h and EVT. Working paper. Available at: https://people.math.ethz.ch/~embrecht/ftp/g-and-h_May07.pdf

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

DREES, H. & KAUFMANN, E. 1998. Selecting the optimal sample fraction in univariate extreme value estimation. *Stochastic Processes and their Applications*, 75:149-172.

DUTTA, K. & PERRY, J. 2007. A tale of tails: an empirical analysis of loss distribution models for estimating operational risk capital. Federal reserve bank of Boston working paper no 6-13.

EBERLEIN, E., EHRET, A., LÜBKE, O., ÖZKAN, F., PRAUSE, K., RAIBLE, S., WIRTH, R. & WIESENDORFER ZAHN, M. 1998. Freiburg Financial Data Tools. Freiburg: Mathematische Stokastik, Universität Freiburg.

EMBRECHTS, P. & FREI, M. 2010. Panjer recursion versus FFT for compound distributions. *Mathematical Methods of Operations Research*, 69:497-508.

EMBRECHTS, P. & HOFERT, M. 2011. Practices and issues in operational risk modeling under Basel II. *Lithuanian Mathematical Journal*, 50(2):180-193.

EMBRECHTS, P., KLÜPPELBERG, C. & MIKOSCH, T. 1997. Modelling extremal events for Insurance and Finance. Springer.

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

FEILMEIER, M. & BERTRAM, J. 1987. Anwendung numerischer Methoden in der Risikotheorie. Verlag Versicherungswirtschaft.

FISHER, R. & TIPPETT, L. H. C. 1928. Limiting forms of the frequency distribution of largest or smallest member of a sample. *Proceedings of the Cambridge Philosophical Society*, 24:180-190.

FOSS, S., KORSHUNOV, D. & ZACHARY, S. 2011. *An Introduction to Heavy-Tailed and Subexponential Distributions*. Springer Series in Operations Research and Financial Engineering.

GNEDENKO, B. V. 1943. Sur la distribution limite du terme d'une série aléatoire. *Annals of Mathematics*, 44:423-453.

GODIN, F., MAYORAL, S. & MORALES, M. 2009. Contingent claim pricing using normal inverse Gaussian probability distortion operator.

GOEGEBEUR, Y., GUILLOU, A. & OSMANN, M. 2013. A local moment type estimator for the extreme value index in regression with random covariates. Preprint.

GOMES, M.I., RODRIGUES, L.H., PEREIRA, H. & PESTANA, D. 2010. Tail index and second-order parameters' semi-parametric estimation based on the log-excesses. *Journal of Statistical Computation and Simulation*, 80(6):653-666.

GOMES, M.I., BRILHANTE, M.F. & PESTANA, D. 2013. A Mean-of-order-p Class of Value-at-Risk Estimators. Preprint. Available at:

http://www.google.co.za/url?sa=t&rct=j&q=&esrc=s&frm=1&source=web&cd=1&ved=0ahUK EwiGhqylhcXJA hWFWxQKHekrD78QFggaMAA&url=http%3A%2F%2Fwww.ceaul.fc.ul.pt%2Fgetfile.asp%3Fwhere%3Dnotas%26id%3D365&usq=AFQjCNHAC-4G2gpi8YMiCRYNYOJIE_wHlw

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

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

HECKMAN, P.E. & MEYERS, G.G. 1983. The calculation of aggregate loss distributions from claim severity and claim count distributions. *Proceedings of the Casualty Actuarial Society*, LXX:22-61.

HERNÁNDEZ, L., TEJERO, J., SUÁREZ, A. & CARILLO-MENÉNDEZ, S. 2014. Closed-form approximations for operational value-at-risk. *Journal of Operational Risk* 8(4):39-54.

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

- HILL, B.M. 1975. A Simple General Approach to Inference About the Tail of a Distribution. *Annals of Statistics* 3(5):1163-1174.
- KARLIS, D. 2002. An EM type algorithm for maximum likelihood estimation for the Normal Inverse Gaussian distribution, *Statistics and Probability Letters*, 57:43-52.
- LILLESTÖL, J. 2000. Risk analysis and the NIG distribution. *The Journal of Risk*, 2:41-56.
- LILLESTÖL, J. 2001. Bayesian Estimation of NIG-parameters by Markov chain Monte Carlo Methods, Discussion paper 2001/3, Department of Finance and Management Science, The Norwegian School of Economics and Business Administration. Earlier version as Discussion paper 112/2000, Sonderforschungsbereich 373, Humboldt Universität zu Berlin.
- LILLESTÖL, J 2002. Some crude approximation, calibration and estimation procedures for NIG-variates, Discussion Papers, Interdisciplinary Research Project 373: Quantification and Simulation of Economic Processes, No. 2002,85, <http://nbnresolving.de/urn:nbn:de:kobv:11-10049614>
- MCNEIL, A.J., FREY, R. & EMBRECHTS, P. 2005. *Quantitative Risk Management: Concepts, Techniques and Tools*. Princeton University Press.
- MCNEIL, A.J., FREY, R. & EMBRECHTS, P. 2015. *Quantitative Risk Management: Concepts, Techniques and Tools, 2nd edition*. Princeton University Press.
- MIGNOLA, G. & UGOCCIONI R. 2006. Sources of uncertainty in modelling operational risk losses. *The Journal of Operational Risk* 1(2):33-50.
- NEŠLEHOVÁ, J., EMBRECHTS, P. & CHAVEZ-DEMOULIN, V. 2006. Infinite mean models and the LDA for operational risk. *The Journal of Operational Risk* 1(1):3-25.
- OMEY, E. & WILLEKENS, E. 1986. Second-order behaviour of distributions subordinate to a distribution with finite mean. *Communications in Statistics: Stochastic Models* 3(3):311-342.
- PANJER, H.H. 1981. Recursive evaluation of a family of compound distributions. *ASTIN Bulletin* 12(1):22-26.

PANJER, H.H. 2006. *Operational Risk: Modeling Analytics*. Wiley, Chichester.

PETERS, G.W., TARGINO, R.S. & SHEVCHENKO, P.V. 2013. Understanding Operational Risk Capital Approximations: First and Second Orders, preprint submitted to Elsevier, <http://arxiv.org/pdf/1303.2910.pdf> [Date Accessed: 20 November 2013].

PICKANDS, J.I. 1975. Statistical inference using extreme value order statistics. *Annals of Statistics*, 3:119-131.

RACHEV, S.T., CHERNOBAI, A. & MENN, C. 2006. Empirical Examination of Operational Loss Distributions. *Perspectives on Operations Research*, p. 379-401. DUV. ISBN 978-3-8350-9064-4.

SAHAY, A., WAN, Z., & KELLER, B. 2007. Operational risk capital: asymptotics in the case of heavy-tailed severity. *The Journal of Operational Risk* 2(2):61–72.

SEGERS, J. 2005. Generalized Pickands estimators for the extreme value index. *The Journal of Statistical Planning and Inference*, 128:381-396.

VENTER, J.H. & DE JONGH, P.J. 2002. Risk estimation using the Normal Inverse Gaussian distribution. *The Journal of Risk*, 4:1-23.

WEISSMAN, I. 1978. Estimation of Parameters and Large Quantiles Based on the k Largest Observations. *Journal of the American Statistical Association*, 73(364):812-815.

Appendix A: Distributions used throughout this thesis

The distributions, together with their parameter values, that were used throughout this thesis are listed in the following three tables. Table A.1, Table A.2 and Table A.3 list the LogNormal, Burr and LogNIG distributions respectively. The LogNIG distribution was discussed in more detail in Chapter 3. Various properties of these distributions such as the extreme value index, the second order parameter and the quantiles relevant to operational risk management are also given in these tables.

Table A.1: Quantiles of the LogNormal distributions

Group	Distribution	γ	ρ	q_{95}	q_{99}	$q_{99.5}$	$q_{99.9}$
No group	LogN(0,1)	0	0	5.18	10.240	13.142	21.982
No group	LogN(0,2)	0	0	26.835	104.867	172.718	483.216
No group	LogN(0,3)	0	0	139.012	1073.891	2269.893	10622.15
No group	LogN(0,4)	0	0	720.118	10997.15	29831.41	233498.1
No group	LogN(0,5)	0	0	3730.39	112616	392050.7	5132798
No group	LogN(0,6)	0	0	19324	1153242	5152414	1.13E+08

Table A.2: Quantiles of the Burr distributions

Group	Distribution	γ	ρ	q_{95}	q_{99}	$q_{99.5}$	$q_{99.9}$
No group	Burr(1,2,1)	0.50	-1	4.359	9.950	14.107	31.607
No group	Burr(1,1,2)	0.50	-0.5	3.472	9.000	13.142	30.623
F1	Burr(1,5,1)	0.20	-1	1.802	2.507	2.883	3.980
F2	Burr(1,5,0.5)	0.40	-2	3.313	6.309	8.325	15.849
F3	Burr(1,2.5,0.5)	0.80	-2	10.975	39.809	69.314	251.189
F4	Burr(1,0.25,5)	0.80	-0.2	0.453	5.225	12.636	78.975
F5	Burr(1,1.25,1)	0.80	-1	10.544	39.492	69.037	250.988
F6	Burr(1,3,2/3)	0.50	-1.5	4.455	9.997	14.140	31.622
F7	Burr(1,6,2/3)	0.25	-1.5	2.111	3.162	3.760	5.623
G2	Burr(1,1.5,1)	0.67	-1	7.120	21.400	34.085	99.933
G3	Burr(1,2,0.5)	1.00	-2	19.975	99.995	199.997	999.999
G4	Burr(1,0.375,10/3)	0.80	-0.3	2.726	18.407	37.718	175.458
G5	Burr(1,0.4,5)	0.50	-0.2	0.610	2.811	4.881	15.344
G6	Burr(1,1.6,2.5)	0.25	-0.4	1.690	2.839	3.472	5.399
G7	Burr(1,2.67,1.25)	0.30	-0.8	2.368	3.937	4.865	7.911
G8	Burr(1,3.5,0.5)	0.57	-2	5.535	13.895	20.648	51.795
G9	Burr(1,8,0.5)	0.25	-2	2.114	3.162	3.761	5.623

Table A.3: Quantiles of the LogNIG distributions

Group	Distribution	γ	ρ	q_{95}	q_{99}	$q_{99.5}$	$q_{99.9}$
No group	LogNIG(3,1,1,1)	0.5	0	11.675	23.507	31.733	63.938
No group	LogNIG(2,1,1,1)	1	0	23.767	79.5074	136.392	496.788
No group	LogNIG(1,0.5,1,1)	2	0	44.945	347.507	910.416	9830.896

Appendix B: Approximating the Modified Bessel Function

The modified Bessel function (MBF) of the first kind $I_{\pm v}(z)$ and of the third kind (also called MacDonald functions) $K_v(z)$ with index v are solutions to the differential equation

$$z^2 \frac{d^2 \omega}{dz^2} + z \frac{d\omega}{dz} - (z^2 + v^2)\omega = 0.$$

The function $I_v(z)$ can be written as

$$I_v(z) = \left(\frac{z}{2}\right)^v \sum_{k=0}^{\infty} \frac{\left(\frac{z^2}{4}\right)^k}{k! \Gamma(v + k + 1)}$$

and $K_v(z)$ satisfies

$$K_v(z) = \frac{\pi}{2} \frac{I_v(z) - I_{-v}(z)}{\sin(v\pi)},$$

where the right-hand side of this equation is replaced by its limiting value if v is an integer or zero.

The Bessel function $K_v(z)$ can also be written in integral form as

$$K_v(z) = \frac{1}{2} \int_0^{\infty} u^{v-1} \exp\left(-\frac{1}{2}z(u + u^{-1})\right) du, \quad x > 0.$$

Useful properties of the MBF include:

- $K_v(z) = K_{-v}(z)$
- $K_{v+1}(z) = \frac{2v}{z} K_v(z) + K_{v-1}(z)$
- $K_{\frac{1}{2}}(z) = \sqrt{\frac{\pi}{2}} z^{-\frac{1}{2}} \exp(-z)$
- $K'_v(z) = -\frac{v}{z} K_v(z) - K_{v-1}(z)$

Abramowitz & Stegun (1972) provide polynomial approximations for both the first and third kind MBF with $v = 0$ and $v = 1$. Equations B.1 to B.4 provide approximations for the first kind MBF whereas equations B.5 to B.8 are related to approximations regarding the third kind. These are given in the following sections.

Polynomial approximations for the modified Bessel function of first kind

Take $t = \frac{z}{3.75}$ in Equations B.1 to B.4 below.

Equation B.1:

For $-3.75 \leq z \leq 3.75$,

$$I_0(z) = 1 + 3.5156229t^2 + 3.0899424t^4 + 1.2067492t^6 \\ + 0.2659732t^8 + 0.360768t^{10} + 0.0045813t^{12} + \varepsilon$$

with $|\varepsilon| < 1.6 \times 10^{-7}$.

Equation B.2:

For $3.75 \leq z < \infty$,

$$\sqrt{z}e^{-z}I_0(z) = 0.39894228 + 0.01328592t^{-1} + 0.00225319t^{-2} \\ - 0.00157565t^{-3} + 0.0091628t^{-4} - 0.02057706t^{-5} \\ + 0.02635537t^{-6} - 0.01647633t^{-7} + 0.00392377t^{-8} + \varepsilon$$

with $|\varepsilon| < 1.9 \times 10^{-7}$.

Equation B.3:

For $-3.75 \leq z \leq 3.75$,

$$z^{-1}I_1(z) = 0.5 + 0.87890594t^2 + 0.51498869t^4 + 0.15084934t^6 \\ + 0.02658733t^8 + 0.00301532t^{10} + 0.00032411t^{12} + \varepsilon$$

with $|\varepsilon| < 8 \times 10^{-9}$.

Equation B.4:

For $3.75 \leq z < \infty$,

$$\sqrt{z}e^{-z}I_1(z) = 0.39894228 - 0.03988024t^{-1} - 0.00362018t^{-2} \\ + 0.00163801t^{-3} - 0.01031555t^{-4} + 0.02282967t^{-5} \\ - 0.02895312t^{-6} + 0.1787654t^{-7} - 0.00420059t^{-8} + \varepsilon$$

with $|\varepsilon| < 2.2 \times 10^{-7}$.

Polynomial approximations for the modified Bessel function of third kind

Equation B.5:

For $0 < z \leq 2$ and $z_2 = \frac{z}{2}$,

$$K_0(z) = -\ln(z_2)I_0(z) - 0.57721566 + 0.42278420z_2^2 + 0.23069756z_2^4 \\ + 0.03488590z_2^6 + 0.00262698z_2^8 + 0.00010750z_2^{10} + 0.00000740z_2^{12} + \varepsilon$$

with $|\varepsilon| < 1 \times 10^{-8}$.

Equation B.6:

For $2 \leq z < \infty$ and $z_3 = \frac{2}{z}$,

$$\sqrt{z}e^z K_0(z) = 1.25331414 - 0.07832358z_3 + 0.02189568z_3^2 \\ - 0.01062446z_3^3 + 0.00587872z_3^4 - 0.0025154z_3^5 + 0.00053208z_3^6 + \varepsilon$$

with $|\varepsilon| < 1.9 \times 10^{-7}$.

Equation B.7:

For $0 < z \leq 2$ and $z_2 = \frac{z}{2}$,

$$zK_1(z) = z \times \ln(z_2)I_1(z) + 1 + 0.15443144z_2^2 - 0.67278579z_2^4 \\ - 0.12156897z_2^6 - 0.01919402z_2^8 - 0.00110404z_2^{10} - 0.00004686z_2^{12} + \varepsilon$$

with $|\varepsilon| < 8 \times 10^{-9}$.

Equation B.8:

For $2 \leq z < \infty$ and $z_3 = \frac{2}{z}$,

$$\sqrt{z}e^z K_1(z) = 1.25331414 + 0.23498619z_3 - 0.0365562z_3^2 \\ + 0.01504268z_3^3 - 0.00780353z_3^4 + 0.00325614z_3^5 - 0.00068245z_3^6 + \varepsilon$$

with $|\varepsilon| < 2.2 \times 10^{-7}$.

Appendix C: Maximum likelihood estimation

Consider a random sample $X_1, X_2, \dots, X_n$ that was obtained from a distribution with distribution function $(x; \underline{\theta})$, with F known and $\underline{\theta}$ unknown. Furthermore, denote the density function of F by f . Maximum likelihood estimation (MLE) is perhaps the most well-known method to estimate the parameter(s) $\underline{\theta}$ of the distribution F . This is obtained as the value for $\underline{\theta}$ obtained from maximising the log likelihood $\ell(\underline{\theta}; \underline{X})$, i.e.

$$\hat{\underline{\theta}}^{MLE} = \arg \max_{\underline{\theta}} \ell(\underline{\theta}; \underline{X}) \equiv \arg \max_{\underline{\theta}} \sum_{i=1}^n f(X_i | \underline{\theta}).$$

The value of $\underline{\theta}$, denoted by $\hat{\underline{\theta}}^{MLE}$, that maximizes $\ell(\underline{\theta}; \underline{X})$ is termed the maximum likelihood estimate. An estimate of the p^{th} -quantile q_p can then be obtained as $\hat{q}_p^{MLE} = F^{-1}(p; \hat{\underline{\theta}}^{MLE})$.