COMPUTING VALUE-AT-RISK AND EXPECTED SHORTFALL

IN OPERATIONAL RISK

by

Desire Issiaka Bakassa-Traore

A thesis

presented to Ryerson University

in partial fulfilment

of the requirements for the Degree of

Master of Science

in the program of

Applied Mathematics

Toronto, Ontario, Canada, 2015

AUTHOR'S DECLARATION FOR ELECTRONIC SUBMISSION OF A THESIS

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I authorize Ryerson University to lend this thesis to other institutions or individuals for the purpose of scholarly research.

I further authorize Ryerson University to reproduce this thesis by photocopying or by other means, in total or in part, at the request of other institutions or individuals for the purpose of scholarly research.

I understand that my thesis may be made electronically available to the public.

ABSTRACT

Operational Risk has become more popular in the past fifteen years. The Basel committee realized its importance and banks have to allocate more capital charge, yet this is still not enough. With these new rules, banks have put in place new procedures to compute their risk measures and allocate enough capital charge to avoid bankruptcy. The Basel committee under Basel II has proposed different approaches to compute risk measures for Operational Risk, namely the Basic Indicator Approach, the Advanced Measurement Approach and the Standardized Approach. In our research, we will study the case of Loss Distribution Approach, which has been discussed before, and will contribute to the field by using a heavy-tailed distributed severity: g-and-h distributed. Then, we will analyze and test some methods to compute the value-at-risk(VaR) and conditional value-at-risk or expected shortfall (CVaR).

ACKNOWLEDGMENTS

DEDICATION

To my father, Colonel Désiré Bakassa-Traoré (1951 - 2005)). A bright man who always

pushed me to go higher, sometimes the military ways but left this world too soon...

TABLE OF CONTENTS

vi

LIST OF TABLES

LIST OF FIGURES

Chapter 1

INTRODUCTION

The twentieth century and the beginning of the twenty-first century have been economically
and financially very hard for the world. These centuries have seen The Panic of 1907 with
bank failures, the Great Depression of 1929, the Subprime mortgage crisis of 2008.

Along with these crisis, the bad internal behaviors of some employees - such as Jérôme
Kerviel, a French trader who caused the loss of 50 Billions euros to Société Générale - have
caused bankruptcy to their financial institutions. As a result, banks were weakened in their
comfort and some regulations were put in place so that they have sufficient capital reserve
based on the risk structure. These were sometimes too strict.

In the 90's, Operational Risk, which was a type of risk neglected, started to monopolize
attention. The Basel Committee on Banking Supervision implemented more detailed regu-
lations in Operational Risk. Banks are allowed to explore advanced models but must submit
them to the regulator for approbation before being inserted to their daily tasks. This will
mitigate the excessive weight of capital required. (see Enhancing Bank Transparency - Pub-
lic disclosure and supervisory information that promote safety and soundness in banking
systems; Basel Committee on Banking Supervision - September 1998).

Nowadays, we notice an increasing interest toward Risk Management in the Financial In-
dustry, especially Operational Risk. We become more aware of its importance even though,

Type of Risk	Capital Charge	Weight
Credit	6%	75%
Market	0.4%	5%
Operational	1.6%	20%
Total	8%	100%

Table 1.1: Capital charge and weight of each Type of Risk

in most banks, people still give more attention to Market and Credit Risks. Table (1.1) illustrates the capital charge and weight allocated to Operational Risk compared to the other types of risks.

We also notice an increase of research in the field. Most of them are theoretical. Only recently financial institutions have incorporated analytical models based on some applied research. This is probably due to the fact that the models are not compliant with real case scenarios, or have not yet been studied deep enough. Either way, financial institutions do not necessarily give access to their full data for academic research, nor the data available is enough. Nonetheless, standard techniques such as Monte Carlo simulation, and Fourier Transform are being used for Operational Risk Management to compute risk measures.

In this thesis, we would like to approximate the Value-at-Risk (VaR) and the Expected Shortfall (ES) in Operational Risk. On top of Monte Carlo simulation, we will design a methodology to use other methods, which, we believe, could be less expensive in term of time.

For the thesis purpose and given the methods used, we will divide our analysis in three parts : On the first part, we will review some concept, definition and properties which will be useful to determine the VaR and CVaR. Then we will review methods and test them for single Operational Risk cell. On the second part, we will analyze the aggregate loss in two Operational Risk cells i.e two risk categories defined by the type of event and the business line by using other methods to compute the overall loss based on the dependence of the risk categories, introduced by a copula for severities and frequencies.

In this analysis, we will contribute to the Operation Risk research field by exploring the g-and-h distribution as an alternative to model severities. This distribution is heavy-tailed which results convenient to analyze extreme losses. On the other hand, it has no explicit density function expression thus we have faced some numerical challenges. Given the severity distribution, we will discuss the computation of VaR and CVaR using Monte Carlo approach and recursive convolution, Fast Fourier Transform and Panjer recurrence for a single cell. For two cells, we will model the dependence among cells via Copula and, again, compute the VaR and CVaR. Another contribution to the field is to explore the methods for parameter estimation.

Due to the lack of real data available, we will generate our own. We know that financial institutions do not only use their internal data and only use losses above a certain amount. Therefore, we will discuss the truncation and the mixture of internal and external data.

We will also discuss the parameter estimation based on the Methods of Moments and Maximum Likelihood Estimators. The structure of the thesis is as follow: Chapter 2 explains the

Loss Distribution Approach by giving a history as well as a definition of Operational Risk to make the concept clear to the reader, the general concept of the VaR and the Expected Shortfall. Also, we will provide the reader with some properties and simulations of univariate mixing g-and-h, Truncation and mixing internal with external data before exploring the computing methods for the VaR and the CVaR for a single cell using Monte Carlo, recursive convolution, Fast Fourier Transform and Panjer recurrence. Chapter 3 focuses on the computation of the same risk measures in aggregate loss by modeling the dependence among cells via copula. We will present the aggregate model, review the copula properties and present the results obtained using Clayton copula. Chapter 4 explores the parameter estimation methods: Method of Moments and MLE. Finally, Chapter 5 concludes our research.

<center>**Chapter 2**</center>

<center>**LOSS DISTRIBUTION APPROACH IN OR AND SINGLE CELL LOSS**</center>

<center>**VAR/CVAR**</center>

2.1 LDA: General concepts

Basel II is the second Basel Accord. It consists of regulations and laws issued by the Basel Committee on Banking Supervision. It proposes VaR as a risk measure to evaluate the capital charge. Due to the fact that VaR is a quantile, we need a distribution of Operational Losses. This distribution is called Loss Distribution Approach. Under this approach, the bank estimates, for each business line/risk type cell, the probability distributions of the severity (single event impact) and of the one year event frequency using its internal data. With these two distributions, the bank then computes the probability distribution of the aggregate operational loss. The total required capital is the sum of the Value-at-Risk of each business line and event type combination(see Frachot et al.(2001)).

LDA is a statistical approach which is very popular in actuarial sciences for computing aggregate loss distributions. Under the Loss Distribution Approach, the bank estimates, for each business line/risk type cell, the probability distribution functions of the single event impact and the event frequency for the next (one) year using its internal data, and computes the probability distribution function of the cumulative operational loss (see Frachot et al. (2001)). The LDA model for the total loss L_t on an interval [0,1], generally $t = 1year$ in a

<center>5</center>

Financial Institution can be formulated as:

$$L_t = \sum_{j=1}^{J} L_t^{(j)};$$

where $L_t^{(j)} = \sum_{i=1}^{N_t^{(j)}} X_i^{(j)}$ (the loss of a single cell j);

$j = 1, 2, .., J$

$N_t^{(j)} \sim \rho_{\theta_1}$ (Frequency of loss events - The number of times a loss occurs in a particular time frame.)

$X_i^{(j)} \sim \rho_{\theta_2}$ (Severity - The size of the losses)

LDA is very popular because it allows one hold less capital than with standard methods proposed by the Basel Committee. In general terms, LDA consists in estimating the parameters θ_1 and θ_2 and compute VaR_α i.e the Value-at-Risk such that $F_L(VaR_\alpha) = \alpha$, where F_L is the cumulative distribution function of the loss L.

2.2 Loss Model - OR definition and OR matrix

Operational Risk was introduced by the Basel Committee. It is a process including risk assessment, risk decision making and also risk of Loss due to bad internal procedures, it's a risk due to human error. The Basel II Committee defines operational risk as:

"The risk of loss resulting from inadequate or failed internal processes, people and systems

or from external events (see Basel Committee on Banking Supervision (2006, p. 204))".
It includes legal risk but not strategic and reputational risk. Therefore they have in place
a list of seven risk categories called event types:

(1) **Internal Fraud** - misappropriation of assets, tax evasion, intentional mismarking of positions, bribery,

(2) **External Fraud** - theft of information, hacking damage, third-party theft and forgery,

(3) **Employment Practices and Workplace Safety** - discrimination, workers compensation, employee health and safety

(4) **Clients, Products and Business Practice** - market manipulation, antitrust, improper trade, product defects, fiduciary breaches, account churning.

(5) **Damage to Physical Assets** - natural disasters, terrorism, vandalism.

(6) **Business Disruption and Systems Failures** - utility disruptions, software failures, hardware failure.

(7) **Execution, Delivery and Process Management** - data entry errors, accounting errors, failed mandatory reporting, negligent loss of client assets.

These seven event types permit to collect data for the eight business lines: Corporate Finance, Trading and Sales, Retail Banking, Commercial Banking, Payment and Settlement, Agency Services, Asset Management and Retail Brokerage. The combination in pairs of the seven event types and eight risk categories (56 cells) form the Operational Risk Matrix.

2.3 Univariate mixing g-and-h distribution

The g-and-h family of distributions was introduced by Tukey (1977), found application in Operational Risk in Moscadelli (2004) and Dutta and Perry (2007) and has been discussed by Degen et al. (2007) on the subadditivity of VaR for g-and-h random variables and showed that for reasonable g-and-h parameter values, superadditivity when estimating high quantiles. It allows the transformation of standard normal random variables with g and h parameters to generate non-normal distributions (see Kowalchuk and Headrick (2010).).

Definition 2.1. *A g-and-h distributed random variable X is a transformed normal variable Z according to*

$$X = A + B\frac{e^{gZ}-1}{g}e^{\frac{hZ^2}{2}} = A + Bk(Z) \tag{2.1}$$

with A and B determining the location and scale of the distribution, respectively, whereas g and $h > 0$ determine skewness and the kurtosis of the distribution.

We define the function $k(x)$ as:

$$k(x) = \frac{e^{gx}-1}{g}e^{\frac{hx^2}{2}}, x \in \mathbb{R} \tag{2.2}$$

Its derivative is:

$$k'(x) = \frac{1}{g}\left(((g+hx)e^{gx} - hx)e^{\frac{hx^2}{2}}\right) \tag{2.3}$$

Notice that $k(x)$ is a one-to-one function as its derivative k' is strictly positive for $g > 0$ or strictly negative for $g < 0$. Hence, the inverse function exists even though there is no analytic expression for it. Its probability density function (p.d.f.) can be computed by an

elementary approach:

$$f_X(x) = f_Z\left(\left(k^{-1}(x)\right)\right)\left|\left(k^{-1}\right)'(x)\right| \tag{2.4}$$

where f_Z is the probability density function (p.d.f.) of a standard normal distribution.

Notation: $X \sim g - and - h(A, B, g, h)$

Remark 2.2. *The parameter g controls the skewness whereas h governs the heavy-tailedness*

of a distribution. This was shown using MatLab in graphs 2.1 and 2.2

When g = 0. Equation (2.1), interpreted as the limit $g \to 0$ is given by:

$$X = A + BZe^{\frac{hZ^2}{2}} \tag{2.5}$$

and is called the h-distribution.

Although, when h = 0, we have:

$$X = A + B\frac{e^{gZ} - 1}{g} \tag{2.6}$$

and is called the g-distribution. This distribution is a scaled log-normal distribution (see

Johnson and Kotz (1970)).

In the next lemma we summarize the cumulative distribution function (c.d.f.), the probability density function (p.d.f.) and the quantile function of the g-and-h distribution.

Lemma 2.3. *Let $X \sim g - and - h(A, B, g, h)$ be a random variable then their c.d.f, p.d.f.*

and quantile function are given respectively by:

$$F_X(y) \quad = \quad \Phi(k^{-1}(\frac{y-A}{B})) \tag{2.7}$$

$$f_X(y) \quad = \quad \frac{1}{B\sqrt{2\pi}} e^{-\frac{(k^{-1}(\frac{y-A}{B}))^2}{2}} \left[k'(k^{-1}(\frac{y-A}{B})) \right]^{-1} \tag{2.8}$$

$$Q_X(y) \quad = \quad F_X^{-1}(y) = k \circ \Phi^{-1}(\frac{y-A}{B}) = k \circ Q_Z(\frac{y-A}{B}) \tag{2.9}$$

where Φ, f_Z and Q_Z are the c.d.f., p.d.f. and the quantile function of a standard normal distribution.

Proof. From equation (3.8), noting that $k(.)$ is a strictly increasing function:

$$F_X(y) \quad = \quad P\left(k(Z) < \frac{y-A}{B} \right) = \Phi\left(k^{-1}\left(\frac{y-A}{B}\right) \right) \tag{2.10}$$

After differentiating with respect to y we have a density given equation (2.8). The quantile function follows computing its inverse.

$\square$

On the other hand, moments of a g-and-h distribution have been obtained, see example Iglewicz and Martinez (1984). For $g > 0$ and $0 < h < \frac{1}{n}$ the n-th moment of the g-and-h is finite and given by (see Dutta and Perry (2007, Appendix D).):

$$E(X^n) \quad = \quad \sum_{l=0}^{n} \binom{n}{l} A^{n-l} B^l \frac{\sum_{r=0}^{n}(-1)^r \binom{l}{r} exp\frac{[(l-r)g]^2}{2(1-lh)}}{g^l\sqrt{1-lh}} \tag{2.11}$$

The next result is about the probability of a sum of independent random variables with g-and-h distribution. It will be needed in computing the p.d.f of the sum of severities.

Lemma 2.4. *Let $X_1 \sim g - and - h(A, B, g, h)$ and $X_2 \sim g - and - h(A, B, g, h)$ independent random variables, then:*

i)

$$F_{X_1+X_2}(y) = F_{Y_1+Y_2}(\frac{y - 2A}{B}) \qquad (2.12)$$

where Y_1, Y_2 are two independent random variables $Y_i \sim g - and - h(0, 1, g, h)$, $i = 1, 2$.

ii)

$$
\begin{aligned}
F_{Y_1+Y_2}(y) &= \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \Phi(v(y, z)) e^{-\frac{z^2}{2}} dz = E(\Phi(v(y, Z))) \\
&= \int_0^1 \Phi(v(y, z)) dz \qquad (2.13)
\end{aligned}
$$

where $v = v(y, z)$ solves:

$$(e^{gv} - 1)e^{\frac{hv^2}{2}} - g(y - k(z)) = 0 \qquad (2.14)$$

iii) *The p.d.f. is given by:*

$$f_{Y_1+Y_2}(y) = \frac{1}{2\pi g} \int_{\mathbb{R}} e^{-\frac{1}{2}((1+h)v_1^2(y,z)+z^2)} [(g + hv_1(y, z))e^{gv_1(y,z)} - hv_1(y, z)]^{-1} dz \quad (2.15)$$

Proof.

$$
\begin{aligned}
F_{X_1+X_2}(y) &= \mathbb{P}(X_1 + X_2 < y) \\
&= \mathbb{P}(A + B(k(Z_1)) + A + B(k(Z_2)) < y) \\
&= \mathbb{P}(k(Z_1) + k(Z_2) < \frac{y - 2A}{B}) \\
&= F_{Y_1+Y_2}(\frac{y - 2A}{B})
\end{aligned}
$$

11

By the convolution product formula and equation (2.4):

$$F_{Y_1+Y_2}(y) \quad = \quad \int_{\mathbb{R}} F_{Y_1}(y-x)dF_{Y_2}(x) = \int_{\mathbb{R}} F_{Y_1}(y-x)f_{Y_2}(x)dx$$

$$= \quad \int_{\mathbb{R}} \Phi(k^{-1}(y-x))f_Z(k^{-1}(x))(k'(k^{-1}(x)))^{-1}dx$$

Taking the change of variable $z = k^{-1}(x)$ we have:

$$\int_{\mathbb{R}} \Phi(k^{-1}(y-x))f_Z(k^{-1}(x))(k'(k^{-1}(x)))^{-1}dx = \int_{\mathbb{R}} \Phi(k^{-1}(y-k(z)))f_Z(z)dz$$

$$= \quad \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \Phi(k^{-1}(y-k(z)))e^{-\frac{z^2}{2}}dz$$

Now, for a function u and random variable $Y \sim g-and-h(0,1,g,h)$ such that $E(u(Y)) < \infty$,

after the change of variable $y = F_X(x)$ we note that:

$$Eu(Y) \quad = \quad \int_{\mathbb{R}} u(x)f_Y(x)dx = \int_0^1 u(Q_Y(y))dy = \int_0^1 u(k(Q_Z(y)))dy$$

But specializing the previous result for $g(x) = F_{Y_1}(y-x)$

$$F_{Y_1+Y_2}(y) \quad = \quad P(Y_1 + Y_2 < y) = E(\mathbb{1}_{[Y_1+Y_2<y]} = E(E(\mathbb{1}_{[Y_1+y_2<y]}/Y_2))$$

$$= \quad E(F_{Y_1}(y-Y_2)) = \int_0^1 F_{Y_1}(y-k(Q_Z(y)))dz$$

$$= \quad \int_0^1 \Phi(k^{-1}(y-k(Q_Z(y))))dz = \int_0^1 \Phi(v(y,z))dz$$

Equation (2.15) is obtained in a similar way. Namely:

$$f_{Y_1+Y_2}(y) \quad = \quad \int_{\mathbb{R}} f_{Y_1}(y-x)f_{Y_2}(x)dx = \int_{\mathbb{R}} f_Z(k^{-1}(y-x))(k'(k^{-1}(y-x)))^{-1}f_Z(k^{-1}(x))(k'(k^{-1}(x)))^{-1}dx$$

$$= \quad \int_{\mathbb{R}} f_Z(k^{-1}(y-k(z)))(k'(k^{-1}(y-k(z))))^{-1}f_Z(z)dz$$

$$= \quad \frac{1}{2\pi g} \int_{\mathbb{R}} e^{-\frac{1}{2}((1+h)v_1^2(y,z)+z^2)}[(g+hv_1(y,z))e^{gv_1(y,z)} - hv_1(y,z)]^{-1}dz$$

$\square$

Remark 2.5. *The function k^{-1} is not explicitly available. Zeros of k need to be computed for different values of the argument. The algorithm is implemented in MatLab. Figure 2.1 shows density functions of a g-and-h probability distribution with parameters $A = 0, B = 1, h = 0.25$ and $g = 1$(blue line), $g = 2$(green line) and $g = 3$ (red line). As we can observe g is a parameter capturing the asymmetry or skewness of the data.*
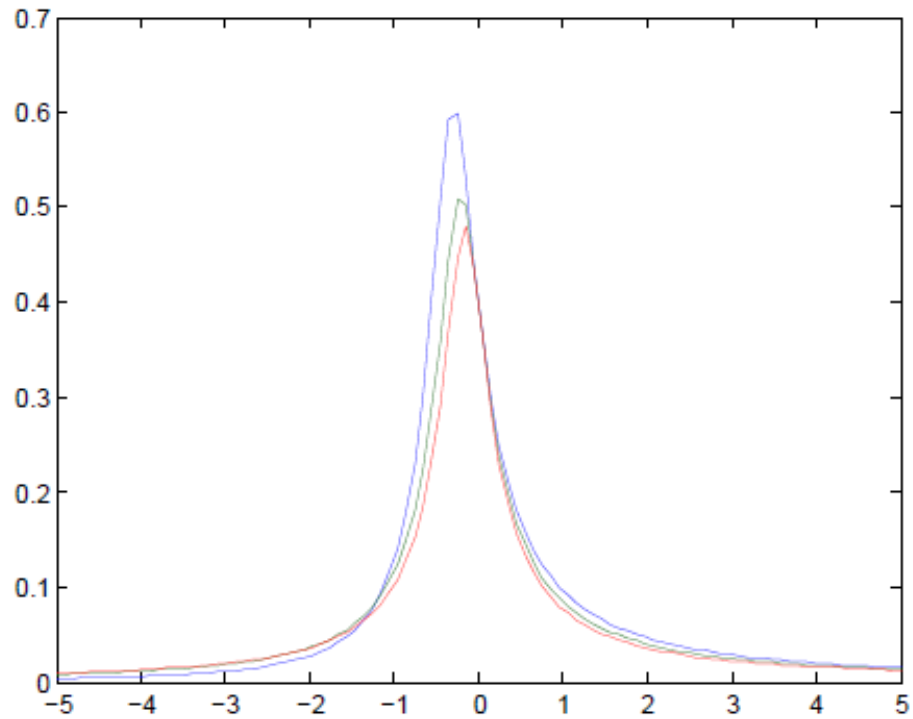
Figure 2.1: Variation of g-and-h density for different values of parameter g

In figure 2.2 the graph of a g-and-h p.d.f. is shown for different values of the parameter h. The curves provide the shapes for the functions with parameters $A = 0, B = 1, g = 2$ and

13

$h = 0.1$(blue line), $h = 0.3$(green line) and $h = 0.3$ (red line). It illustrates how the parameter h captures the thickness in the tail of the distribution.
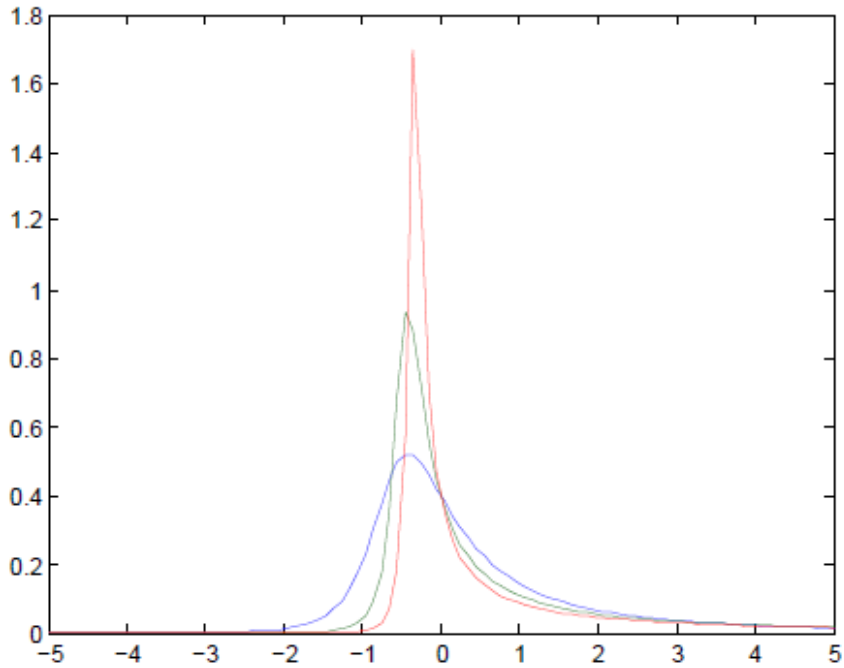
Figure 2.2: Variation of g-and-h density for different values of parameter h

Note that the code to retrieve graphs 2.1 and 2.2 is in Appendix A (densitydependgh.m)

Given a random variable L representing the loss in a single cell the value-at-risk at level α, denoted by VaR_α, solves $F_L(VaR_\alpha) = \alpha$.

The conditional value-at-risk or expected shortfall at level α, denoted by $CVaR(\alpha)$ or

$ES(\alpha)$ is defined as

$$ES(\alpha) = E(L/L > VaR_\alpha) = \frac{1}{1-\alpha} \int_\alpha^1 VaR_\beta d\beta \qquad (2.16)$$

Notice that the conditional value-at-risk represents the average loss in a single cell given that the loss surpasses the value-at-risk (see for example Pavel V. Shevchenko, 2011).

We denote by F_X^{*k} the convolution of function F_X with itself k times. Under a model with Poissonian frequencies and g-and-h severities described by a random variable X, by conditioning on the number of loss events, the c.d.f. of the loss L is given by:

$$F_L(y) \quad = \quad \sum_{k=0}^{+\infty} \frac{e^{-\lambda}\lambda^k}{k!} F_X^{*k} \qquad (2.17)$$

where $F_X^{*0} := \delta_0$ and $F_X^{*1} = F_X$, i.e. the dirac delta at zero and the c.d.f. of X respectively. Unfortunately this equation does not have an explicit solution. Numerical methods are needed. In forthcoming sections we study the implementation of different approaches.

2.4 Truncating and mixing internal and external data

Due to the characteristic of the data available, modeling the severity by a standard g-and-h distribution has limited application. Two modifications are required. On one hand, losses below certain threshold are not usually reported by banks and another financial institutions, it suggests the introduction of a truncated distribution.

On the other hand, in order to increase the amount of data available, banks use a combination of internal and external information, with possible different probability distributions. It leads to the use of mixing distributions to model losses. Here we assume frequencies

come from two independent Poisson distributions, in general with different parameters, and severities come from a mixing of independent g-and-h probability distributions with different parameters and truncation levels.

We begin by tackling both problems separately, then we consider both, truncation and mixing together.

For a random variable X we define its truncated c.d.f. with truncation level T as:

$$F_{X,T}(y) = P(X < y/X > T), y \in \mathbb{R} \tag{2.18}$$

The respective p.d.f. and quantile functions are denoted by $f_{X,T}$ and $Q_{X,T}$.

Results from the previous section can easily adapted by noticing that a truncated distribution is nothing but a conditional one on the set $[X > T]$, where T is the truncation level.

Proposition 2.6. *Let* $X \sim g - and - h(A, B, g, h)$ *truncated at* T *then:*

$$F_{X,T}(y) = \frac{\Phi\left(k^{-1}\left(\frac{y-A}{B}\right)\right) - \Phi\left(k^{-1}\left(\frac{T-A}{B}\right)\right)}{1 - \Phi\left(k^{-1}\left(\frac{T-A}{B}\right)\right)} \mathbb{1}_{[T,+\infty)}(y) \tag{2.19}$$

$$Q_{X,T}(y) = F_X^{-1}(y) = k \circ Q_Z((1 - F_Y(\frac{y-A}{B})y + F_Y(\frac{y-A}{B}))) \tag{2.20}$$

where $F_{X,T}(y)$ *and* $Q_{X,T}$ *are the c.d.f. and the quantile function of a standard normal distribution respectively, and*

$$f_{X,T}(y) = \frac{1}{B\sqrt{2\pi}(1 - \Phi(k^{-1}(\frac{T-A}{B})))} e^{-\frac{(k^{-1}(\frac{y-A}{B}))^2}{2}} (k'(k^{-1}(\frac{y-A}{B})))^{-1} \mathbb{1}_{[T,+\infty]}(y) \tag{2.21}$$

is its p.d.f.

16

Proof.

$$
\begin{aligned}
F_{X,T}(y) \quad &= \quad P(X < y / X > T) = \frac{P[X < y] \cap [X > T]}{P(X > T)} \\
&= \quad \frac{F_X(y) - P(X < y, X \le T)}{1 - F_X(T)} = \frac{F_X(y) - F_X(T)}{1 - F_X(T)}
\end{aligned}
$$

where the last equality holds for $y > T$ and is equal to zero otherwise.

Combining equation (2.7) in Lemma 1 with the expression above we have equation (2.19).

Differentiating the latter leads to equation (2.21).

Finally, we notice that by definition the value $x = Q_{X,T}(y)$ solves $F_{Y,T}(x) = y$. Then, for $y > T$:

$$
y = \frac{F_X(x) - F_X(T)}{1 - F_X(T)}
$$

or equivalently $(1 - F_X(T))y + F_X(T) = F_X(x)$, then:

$$
Q_{X,T}(y) = F_X^{-1}((1-F_X(T))y+F_X(T)) = Q_X((1-F_X(T))y+F_X(T)) = k \circ Q_Z((1-F_X(T))y+F_X(T))
$$

$\square$

The truncated moments are obtained in the next proposition:

Proposition 2.7. *Let* $X \sim g - and - h(A, B, g, h)$ *and* u *a real valued function such that* $E(u(X)) < +\infty$. *Then:*

$$
E(u(X)/X > T) = \frac{1}{B^2 \sqrt{2\pi}(1 - \Phi(k^{-1}(\frac{T-A}{B})))} \int_{k^{-1}(\frac{T-A}{B})}^{+\infty} u(Bk(z) + A)e^{\frac{-z^2}{2}} dz \qquad (2.22)
$$

In particular for $h < \frac{1}{n}$:

$$
E(X^n/X > T) = \frac{1}{B^2(1 - \Phi(k^{-1}(\frac{T-A}{B})))} \sum_{l=0}^{n} \binom{n}{l} \frac{B^l A^{n-l}}{g^l}(1 - hl)^{-\frac{1}{2}} I_l(T, A, B, g, h) \qquad (2.23)
$$

17

where:

$$I_l(T, A, B, g, h) \quad = \quad J_l(T, A, B, g, h) - K_l(T, A, B, g, h) \tag{2.24}$$

$$J_l(T, A, B, g, h) \quad = \quad \sum_{m=0}^{l} (-1)^m e^{\frac{m^2}{2(1-hl)}} \tag{2.25}$$

$$K_l(T, A, B, g, h) \quad = \quad \sum_{m=0}^{l} (-1)^m e^{\frac{m^2}{2(1-hl)}} \Phi\left(\frac{A-T}{B} - \frac{m}{1-hl}\right)(1-hl)^{\frac{1}{2}} \tag{2.26}$$

Proof.

$$E(u(X)/X > T) \quad = \quad \int_T^{+\infty} u(x) f_{X,T}(x) dx = \frac{1}{1 - F_X(T)} \int_T^{+\infty} u(x) f_X(x) dx \tag{2.27}$$

$$= \quad \frac{1}{(1 - \Phi(k^{-1}(\frac{T-A}{B})))B\sqrt{2\pi}} \int_T^{+\infty} u(x) e^{-\frac{(k^{-1}(\frac{x-A}{B}))^2}{2}} (k'(k^{-1}(\frac{x-A}{B}))) \tag{2.28}$$

After the changes of variables $u = \frac{x-A}{B}$ and $z = k^{-1}(u)$ we have:

$$E(u(X)/X > T) \quad = \quad \frac{1}{(1 - \Phi(k^{-1}(\frac{T-A}{B})))B^2\sqrt{2\pi}} \int_{\frac{T-A}{B}}^{+\infty} u(Bu + A) e^{-\frac{k^{-1}(u)^2}{2}} (k'(k^{-1}(u))) \tag{2.29}$$

$$= \quad \frac{1}{(1 - \Phi(k^{-1}(\frac{T-A}{B})))B^2\sqrt{2\pi}} \int_{k^{-1}(\frac{T-A}{B})}^{+\infty} u(Bk(z) + A) e^{-\frac{z^2}{2}} dz \tag{2.30}$$

We specialize for $u(x) = x^n$.

$$E(X^n/X > T) = \frac{1}{(1 - \Phi(\frac{T-A}{B}))B^2} \int_{\frac{T-A}{B}}^{+\infty} (Bu + A)^n f_Z(k^{-1}(u))(k'(k^{-1}(u)))^{-1} du \qquad (2.31)$$

$$= \frac{1}{(1 - \Phi(k^{-1}(\frac{T-A}{B})))B^2} \int_{k^{-1}(\frac{T-A}{B})}^{+\infty} (Bk(z) + A)^n f_Z(z) dz \qquad (2.32)$$

$$= \frac{1}{(1 - \Phi(k^{-1}(\frac{T-A}{B})))B^2} \sum_{l=0}^{n} \binom{n}{l} B^l A^{n-l} \int_{k^{-1}(\frac{T-A}{B})}^{+\infty} k(z)^l f_Z(z) dz \qquad (2.33)$$

$$= \frac{1}{(1 - \Phi(k^{-1}(\frac{T-A}{B})))B^2 \sqrt{2\pi}} \sum_{l=0}^{n} \binom{n}{l} B^l A^{n-l} \int_{k^{-1}(\frac{T-A}{B})}^{+\infty} k(z)^l e^{-\frac{z^2}{2}} dz \qquad (2.34)$$

$$= \frac{1}{(1 - \Phi(k^{-1}(\frac{T-A}{B}))B^2 \sqrt{2\pi}} \sum_{l=0}^{n} \binom{n}{l} B^l A^{n-l} g^{-l} \int_{k^{-1}(\frac{T-A}{B})}^{+\infty} (e^{gz} - 1)^l e^{\frac{hlz^2}{2}} e^{-\frac{z^2}{2}} \qquad (2.35)$$

$$\qquad (2.36)$$

Now, expanding again:

$$\int_{k^{-1}(\frac{T-A}{B})}^{+\infty} (e^{gz} - 1)^l e^{\frac{hlz^2}{2}} e^{-\frac{z^2}{2}} dz \tag{2.37}$$

$$= \sum_{m=0}^{l} \binom{l}{m} \int_{k^{-1}(\frac{T-A}{B})}^{+\infty} (-1)^{l-m} e^{mgz} e^{\frac{hlz^2}{2}} e^{-\frac{z^2}{2}} dz \tag{2.38}$$

$$= \sum_{m=0}^{l} \binom{l}{m} (-1)^{l-m} \int_{k^{-1}(\frac{T-A}{B})}^{+\infty} e^{-\frac{1}{2}((1-hl)z^2 - 2mgz)} dz \tag{2.39}$$

$$= \sum_{m=0}^{l} \binom{l}{m} (-1)^{l-m} e^{\frac{(mg)^2}{2(1-hl)}} \int_{k^{-1}(\frac{T-A}{B})}^{+\infty} e^{-\frac{1}{2}(1-hl)(z^2 - 2\frac{mg}{(1-hl)}z + (\frac{mg}{1-hl})^2)} dz \tag{2.40}$$

$$= \sum_{m=0}^{l} \binom{l}{m} (-1)^{l-m} e^{\frac{(mg)^2}{2(1-hl)}} \int_{k^{-1}(\frac{T-A}{B})}^{+\infty} e^{-\frac{1}{2}(1-hl)(z - \frac{mg}{(1-hl)})^2} dz \tag{2.41}$$

$$= \sqrt{2\pi} \sum_{m=0}^{l} \binom{l}{m} (-1)^{l-m} (1-hl)^{-\frac{1}{2}} e^{\frac{(mg)^2}{2(1-hl)}} \Phi\left(\left(k^{-1}\left(\frac{T-A}{B}\right) - \frac{mg}{(1-hl)}\right)(1-hl)^{\frac{1}{2}}\right) \tag{2.42}$$

where the last equality occurs after the change of variable $y = (1-hl)^{\frac{1}{2}}(z - \frac{mg}{(1-hl)})$ $\qquad\square$

Next, consider the case of mixing g-and-h distributions, namely, let $Y_i \sim g - and - h(A_i, B_i, g_i, h_i), i = 1, 2.$ be independent random variables and independent also of a random variable $U \sim U(0,1)$. A mixing g-and-h random variable defined as:

$$Y = \mathbb{1}_{[U<p]}Y_1 + \mathbb{1}_{[U\geq p]}Y_2 \tag{2.43}$$

where p is the (known) proportion of internal data.

Denote by

$$k_i(x) = \frac{(e^{g_i x} - 1)}{g_i} e^{\frac{h_i}{2}x^2}, i = 1, 2 \tag{2.44}$$

Their c.d.f., p.d.f. and quantile function are respectively:

$$
\begin{aligned}
F_Y(y) &= \mathbb{P}(\mathbb{1}_{[U<p]}Y_1 + \mathbb{1}_{[U\geq p]}Y_2 < y) \\[2mm]
&= \mathbb{P}(\mathbb{1}_{[U<p]}Y_1|U < p) + \mathbb{P}(\mathbb{1}_{[U\geq p]}Y_2|U \geq p) \\[2mm]
&= \mathbb{P}(Y_1 < y)\mathbb{P}(U < p) + \mathbb{P}(Y_2 < y)\mathbb{P}(U \geq p) \\[2mm]
&= pF_{Y_1}(y) + (1-p)F_{Y_2}(y)
\end{aligned}
$$

By Lemma 2.3, we have:

$$
F_Y(y) = p\Phi(k_1^{-1}(\frac{y-A_1}{B_1})) + (1-p)\Phi(k_2^{-1}(\frac{y-A_2}{B_2})) \tag{2.45}
$$

$$
f_Y(y) = pf_{Y_1}(y) + (1-p)f_{Y_2}(y) = \frac{p}{\sqrt{2\pi}B_1}e^{\frac{(k_1^{-1}(\frac{y-A_1}{B_1}))^2}{2}}(k_1'k_1^{-1}(\frac{y-A_1}{B_1}))^{-1}
$$

$$
+ \frac{1-p}{\sqrt{2\pi}B_2}e^{\frac{(k_2^{-1}(\frac{y-A_2}{B_2}))^2}{2}}(k_2'k_2^{-1}(\frac{y-A_2}{B_2}))^{-1} \tag{2.46}
$$

$$
Q_Y(y) = pk_1 \circ Q_Z(\frac{y-A_1}{B_1}) + (1-p)k_2 \circ Q_Z(\frac{y-A_2}{B_2}) \tag{2.47}
$$

In figure 2.3 the graph of the empirical density of a mixture of g-and-h functions with parameters $A_1 = 0, B_1 = 1, g_1 = 2, h_1 = 2, A_2 = 5, B_2 = 1, g_2 = 4, h_2 = 0.2, p = 0.5$ is shown. The density is computed using a Gaussian kernel. We observe a bimodal probability density function in agreement with the theory.
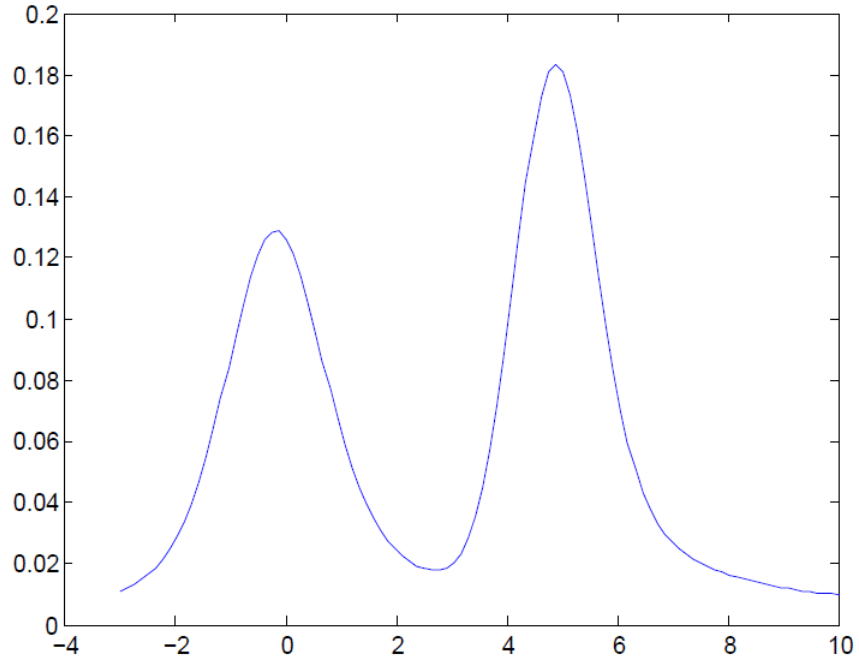
Figure 2.3: Density function of g-and-h mixing internal and external data

We have the following proposition for the probability distribution of the sum of two mixing g-and-h random variables.

Proposition 2.8. *Let Y_1 and Y_2 be two independent mixing g-and-h random variables given*

by equation (2.43) then:

$$
\begin{aligned}
F_{Y_1+Y_2}(y) \quad &= \quad \frac{p^2}{2\pi} \int_{\mathbb{R}} \Phi(vm_1(y,z)) e^{\frac{-z^2}{2}} dz \\[2mm]
&+ \quad \frac{p(1-p)B_2}{2\pi B_1} \int_{\mathbb{R}} \Phi(um_2(y,z)) e^{\frac{-z^2}{2}} dz \\[2mm]
&+ \quad \frac{p(1-p)B_1}{2\pi B_2} \int_{\mathbb{R}} \Phi(um_1(y,z)) e^{\frac{-z^2}{2}} dz \\[2mm]
&+ \quad \frac{(1-p)^2}{2\pi} \int_{\mathbb{R}} \Phi(vm_2(y,z)) e^{\frac{-z^2}{2}} dz \qquad (2.48)
\end{aligned}
$$

where $vm_i = vm_i(y,z), i = 1,2$ solves:

$$
(e^{g_i vm_i} - 1) e^{\frac{h_i vm_i^2}{2}} - g_i\left(k_i(z) + \frac{2A_i - y}{B_i}\right) = 0, i = 1,2
$$

and $um_i = um_i(y,z), i = 1,2$ solves:

$$
(e^{g_i um_i} - 1) e^{\frac{h_i um_i^2}{2}} - g_i\left(k_{3-i}(z) + \frac{A_1 + A_2 - y}{B_i}\right) = 0, i = 1,2
$$

and

$$
\begin{aligned}
f_{Y_1+Y_2}(y) \quad &= \quad \frac{p^2}{2\pi} \int_{\mathbb{R}} \frac{\partial vm_1(y,z)}{\partial y} \frac{1}{\sqrt{2\pi}} e^{\frac{-1}{2}(vm_1(y,z)^2)} e^{\frac{-z^2}{2}} dz \\[2mm]
&+ \quad \frac{p(1-p)B_2}{2\pi B_1} \int_{\mathbb{R}} \frac{\partial um_1(y,z)}{\partial y} \frac{1}{\sqrt{2\pi}} e^{\frac{-1}{2}(um_1(y,z)^2)} e^{\frac{-z^2}{2}} dz \\[2mm]
&+ \quad \frac{p(1-p)B_1}{2\pi B_2} \int_{\mathbb{R}} \frac{\partial um_2(y,z)}{\partial y} \frac{1}{\sqrt{2\pi}} e^{\frac{-1}{2}(um_2(y,z)^2)} e^{\frac{-z^2}{2}} dz \\[2mm]
&+ \quad \frac{(1-p)^2}{2\pi} \int_{\mathbb{R}} \frac{\partial vm_2(y,z)}{\partial y} \frac{1}{\sqrt{2\pi}} e^{\frac{-1}{2}(vm_2(y,z)^2)} e^{\frac{-z^2}{2}} dz \qquad (2.49)
\end{aligned}
$$

$$
(2.50)
$$

Proof.

$$F_{Y_1+Y_2}(y) = \int_{\mathbb{R}} F_{Y_1}(y-x) f_{Y_2} dx \tag{2.51}$$

$$= \int_{\mathbb{R}} \left(p\Phi\left(k_1^{-1}\left(\frac{y-x-A_1}{B_1}\right)\right) + (1-p)\Phi\left(k_2^{-1}\left(\frac{y-x-A_2}{B_2}\right)\right) \right) \tag{2.52}$$

$$\left(\frac{p}{\sqrt{2\pi}B_1} e^{\frac{\left(k_1^{-1}\left(\frac{x-A_1}{B_1}\right)\right)^2}{2}} \left(k_1' k_1^{-1}\left(\frac{x-A_1}{B_1}\right)\right)^{-1} \right. \tag{2.53}$$

$$\left. + \frac{1-p}{\sqrt{2\pi}B_2} e^{\frac{\left(k_2^{-1}\left(\frac{x-A_2}{B_2}\right)\right)^2}{2}} \left(k_2' k_2^{-1}\left(\frac{x-A_2}{B_2}\right)\right)^{-1} \right) dx \tag{2.54}$$

$$= \frac{p^2}{\sqrt{2\pi}B_1} \int_{\mathbb{R}} \Phi\left(k_1^{-1}\left(\frac{y-x-A_1}{B_1}\right)\right) e^{\frac{(k_1^{-1}(\frac{x-A_1}{B_1}))^2}{2}} \left(k_1' k_1^{-1}\left(\frac{x-A_1}{B_1}\right)\right)^{-1} dx \tag{2.55}$$

$$+ \frac{p(1-p)}{\sqrt{2\pi}B_1} \int_{\mathbb{R}} \Phi\left(k_2^{-1}\left(\frac{y-x-A_2}{B_2}\right)\right) e^{\frac{(k_1^{-1}(\frac{x-A_1}{B_1}))^2}{2}} \left(k_1' k_1^{-1}(\frac{x-A_1}{B_1})\right)^{-1} dx \tag{2.56}$$

$$+ \frac{p(1-p)}{\sqrt{2\pi}B_2} \int_{\mathbb{R}} \Phi\left(k_1^{-1}(\frac{y-x-A_1}{B_1})\right) e^{\frac{(k_2^{-1}(\frac{x-A_2}{B_2}))^2}{2}} \left(k_2' k_2^{-1}(\frac{x-A_2}{B_2})\right)^{-1} dx \tag{2.57}$$

$$+ \frac{(1-p)^2}{\sqrt{2\pi}B_2} \int_{\mathbb{R}} \Phi\left(k_2^{-1}(\frac{y-x-A_2}{B_2})\right) e^{\frac{(k_2^{-1}(\frac{x-A_2}{B_2}))^2}{2}} \left(k_2' k_2^{-1}(\frac{x-A_2}{B_2})\right)^{-1} dx \tag{2.58}$$

$$= I_1 + I_2 + I_3 + I_4 \tag{2.59}$$

where I_j corresponds with the j-th term in the previous expression.

After we combine the change of variables $u = \frac{x-A_i}{B_i}$ and $z = k_i^{-1}(u), i = 1, 2$ depending on the integrals above:

$$I_1 = \frac{p^2}{2\pi} \int_{\mathbb{R}} \Phi\left(k_1^{-1}\left(\frac{y-2A_1}{B_1} - k_1(z)\right)\right) e^{-\frac{z^2}{2}} dz \tag{2.60}$$

$$I_2 = \frac{p(1-p)B_2}{2\pi B_1} \int_{\mathbb{R}} \Phi\left(k_1^{-1}\left(\frac{y - B_2 k_2(z) - A_2 - A_1}{B_2}\right)\right) e^{-\frac{z^2}{2}} dz \tag{2.61}$$

$$I_3 = \frac{p(1-p)B_1}{2\pi B_2} \int_{\mathbb{R}} \Phi\left(k_2^{-1}\left(\frac{y - B_1 k_1(z) - A_1 - A_2}{B_2}\right)\right) e^{-\frac{z^2}{2}} dz \tag{2.62}$$

$$I_4 = \frac{(1-p)^2}{2\pi} \int_{\mathbb{R}} \Phi\left(k_2^{-1}\left(\frac{y-2A_2}{B_2} - k_2(z)\right)\right) e^{-\frac{z^2}{2}} dz \tag{2.63}$$

therefore, we have (2.48) and by differentiating we get (2.50). □

24

Finally we tackle the more general case of truncated mixing g-and-h severities with a similar procedure.

Let $Y_{T_i} \sim g-and-h(A_i, B_i, g_i, h_i, T_i), i = 1, 2$ be truncated independent g-and-h random variables with respective truncation levels T_1 and T_2, and

$$Y = \mathbb{1}_{[U<p]}Y_{T_1} + \mathbb{1}_{[U\geq p]}Y_{T_2} \tag{2.64}$$

The probability distribution of the sum of mixing g-and-h is given in the next theorem.

Theorem 2.9. *Let Y_1 and Y_2 be independent g-and-h random variables with parameters $(A_i, B_i, g_i, h_i), i = 1, 2$ satisfying equation:*

$$Y_i = \mathbb{1}_{[U_i<p]}Y_{T_1} + \mathbb{1}_{[U_i\geq p]}Y_{T_2}, i = 1, 2. \tag{2.65}$$

where U_1 and U_2 are independent random variables, and independent of Y_{T_i} also, uniformly distributed on $[0, 1]$, then:

$$
\begin{aligned}
F_{Y_1+Y_2}(y) &= E_1 \int_{k_1^{-1}(\frac{T_2-A_1}{B_1})}^{k_1^{-1}(\frac{y-T_1-A_1}{B_1})} (\Phi(vm_1(y,z))e^{-\frac{z^2}{2}} dz \\
&- E_1\Phi(k_1^{-1}(\frac{T_1-A_1}{B_1})))\sqrt{2\pi}(\Phi(k_1^{-1}(\frac{y-T_1-A_1}{B_1})) - \Phi(k_1^{-1}(\frac{T_2-A_1}{B_1})) \\
&+ E_2 \int_{k_2^{-1}(\frac{T_2-A_2}{B_2})}^{k_2^{-1}(\frac{y-T_1-A_2}{B_2})} (\Phi(um_2(y,z))e^{-\frac{z^2}{2}} dz \\
&- \sqrt{2\pi}E_2\Phi(k_1^{-1}(\frac{T_1-A_1}{B_1})))(\Phi(k_2^{-1}(\frac{y-T_1-A_2}{B_2})) - \Phi(k_2^{-1}(\frac{T_2-A_2}{B_2}))) \\
&+ E_3 \int_{k_2^{-1}(\frac{T_2-A_2}{B_2})}^{k_2^{-1}(\frac{y-T_1-A_2}{B_2})} (\Phi(um_1(y,z))e^{-\frac{z^2}{2}} dz \\
&- E_3\Phi(k_2^{-1}(\frac{T_2-A_2}{B_2}))\sqrt{2\pi}(\Phi(k_2^{-1}(\frac{y-T_1-A_2}{B_2})) - \Phi(k_2^{-1}(\frac{T_2-A_2}{B_2}))) \\
&+ E_4 \int_{k_2^{-1}(\frac{T_2-A_2}{B_2})}^{k_2^{-1}(\frac{y-T_1-A_2}{B_2})} (\Phi(vm_2(y,z))e^{-\frac{z^2}{2}} dz \\
&- E_4\Phi(k_2^{-1}(\frac{T_2-A_2}{B_2}))\sqrt{2\pi}(\Phi(k_2^{-1}(\frac{y-T_1-A_2}{B_2})) - \Phi(k_2^{-1}(\frac{T_2-A_2}{B_2}))) \tag{2.66}
\end{aligned}
$$

where:

$$E_1 = \frac{p^2}{2\pi(1-\Phi(k_1^{-1}(\frac{T_1-A_1}{B_1})))^2} \tag{2.67}$$

$$E_2 = \frac{p(1-p)B_2}{2\pi B_1(1-\Phi(k_2^{-1}(\frac{T_2-A_2}{B_2})))(1-\Phi(k_1^{-1}(\frac{T_1-A_1}{B_1})))} \tag{2.68}$$

$$E_3 = \frac{p(1-p)B_1}{2\pi B_2(1-\Phi(k_1^{-1}(\frac{T_1-A_1}{B_1})))(1-\Phi(k_2^{-1}(\frac{T-A_2}{B_2})))} \tag{2.69}$$

$$E_4 = \frac{(1-p)^2}{2\pi(1-\Phi(k_2^{-1}(\frac{T_2-A_2}{B_2})))^2} \tag{2.70}$$

Proof. As in proposition 2.6 we can write:

$$F_{Y_{T_i}}(y) = \frac{\Phi(k_i^{-1}(\frac{y-A_i}{B_i})) - \Phi(k_i^{-1}(\frac{T-A_i}{B_i}))}{1-\Phi(k_i^{-1}(\frac{T-A_i}{B_i}))} \mathbb{1}_{[T_i,+\infty]}(y) \tag{2.71}$$

$$Q_{Y_{T_i}}(y) = F_{XT_i}^{-1}(y) = k_i \circ Q_Z((1-F_Y(\frac{y-A_i}{B_i})y + F_Y(\frac{y-A_i}{B_i}))) \tag{2.72}$$

where Q_Z is the quantile function of a standard normal distribution.

$$f_{Y_{T_i}}(y) = \frac{1}{(1-\Phi(k_i^{-1}(\frac{T-A_i}{B_i})))B_i\sqrt{2\pi}} e^{-\frac{(k_i^{-1}(\frac{y-A_i}{B_i}))^2}{2}} (k_i'(k_i^{-1}(\frac{y-A_i}{B_i})))^{-1}\mathbb{1}_{[T_i,+\infty]} \tag{2.73}$$

Combining proposition 2.49 for $Y = \mathbb{1}_{[U<p]}YT_1 + \mathbb{1}_{[U\geq p]}YT_2$ it gives:

$$\begin{aligned} F_Y(y) &= pF_{Y_{T_1}}(y) + (1-p)F_{Y_{T_2}}(y) = p(\frac{\Phi(k_1^{-1}(\frac{y-A_1}{B_1})) - \Phi(k_1^{-1}(\frac{T_1-A_1}{B_1}))}{1-\Phi(k_1^{-1}(\frac{T_1-A_1}{B_1}))})\mathbb{1}_{[T_1,+\infty]}(y) \\ &+ (1-p)(\frac{\Phi(k_2^{-1}(\frac{y-A_2}{B_2})) - \Phi(k_2^{-1}(\frac{T_2-A_2}{B_2}))}{1-\Phi(k_2^{-1}(\frac{T_2-A_2}{B_2}))})\mathbb{1}_{[T_2,+\infty]}(y)) \\ f_Y(y) &= p\frac{1}{(1-\Phi(k_1^{-1}(\frac{T_1-A_1}{B_1})))B_1\sqrt{2\pi}} e^{-\frac{(k_1^{-1}(\frac{y-A_1}{B_1}))^2}{2}} (k_1'(k_1^{-1}(\frac{y-A_1}{B_1})))^{-1}\mathbb{1}_{[T_1,+\infty]}(y) \\ &+ (1-p)\frac{1}{(1-\Phi(k_2^{-1}(\frac{T_2-A_2}{B_2})))B_2\sqrt{2\pi}} e^{-\frac{(k_2^{-1}(\frac{y-A_2}{B_2}))^2}{2}} (k_2'(k_2^{-1}(\frac{y-A_2}{B_2})))^{-1}\mathbb{1}_{[T_2,+\infty]}(y) \end{aligned}$$

$$\tag{2.74}$$

26

Then for $y > T_1 + T_2$:

$$F_{Y_1+Y_2}(y) \;=\; \int_{\mathbb{R}} F_{Y_1}(y-x) f_{Y_2}(x)\,dx \tag{2.75}$$

$$= \int_{T_2}^{y-T_1}\left(p\Big(\frac{\Phi(k_1^{-1}(\frac{y-x-A_1}{B_1})) - \Phi(k_1^{-1}(\frac{T_1-A_1}{B_1}))}{1-\Phi(k_1^{-1}(\frac{T-A_1}{B_1}))}\Big)\right. \tag{2.76}$$

$$+\; (1-p)\Big(\frac{\Phi(k_2^{-1}(\frac{y-x-A_2}{B_2})) - \Phi(k_2^{-1}(\frac{T-A_2}{B_2}))}{1-\Phi(k_2^{-1}(\frac{T_2-A_2}{B_2}))}\Big) \tag{2.77}$$

$$\Big(p\frac{1}{(1-\Phi(k_1^{-1}(\frac{T-A_1}{B_1})))B_1\sqrt{2\pi}}e^{-\frac{(k_1^{-1}(\frac{y-A_1}{B_1}))^2}{2}}(k_1'(k_1^{-1}(\frac{y-A_1}{B_1})))^{-1}\mathbb{1}_{[T_1,+\infty]} \tag{2.78}$$

$$+\;(1-p)\frac{1}{(1-\Phi(k_2^{-1}(\frac{T_2-A_2}{B_2})))B_2\sqrt{2\pi}}e^{-\frac{(k_2^{-1}(\frac{y-A_2}{B_2}))^2}{2}}(k_2'(k_2^{-1}(\frac{y-A_2}{B_2})))^{-1}\mathbb{1}_{[T_2,+\infty]}(x)\Big)\,dx \tag{2.79}$$

$$=\; I_1 + I_2 + I_3 + I_4 \tag{2.80}$$

After we combine the change of variables $u = \frac{x-A_i}{B_i}$ and $z = k_i^{-1}(u), i = 1,2$ depending on the

integrals above:

$$I_1 = E_1 \int_{\frac{T_2-A_1}{B_1}}^{\frac{y-T_1-A_1}{B_1}} (\Phi(k_1^{-1}(\frac{y-2A_1}{B_1}-k(z)))-\Phi(k_1^{-1}(\frac{T_1-A_1}{B_1})))e^{-\frac{z^2}{2}}\,dz \tag{2.81}$$

$$= E_1 \int_{k_1^{-1}(\frac{T_2-A_1}{B_1})}^{k_1^{-1}(\frac{y-T_1-A_1}{B_1})} (\Phi(k_1^{-1}(\frac{y-2A_1}{B_1}-k(z)))e^{-\frac{z^2}{2}}\,dz \tag{2.82}$$

$$- E_1\Phi(k_1^{-1}(\frac{T_1-A_1}{B_1}))) \int_{k_1^{-1}(\frac{T_2-A_1}{B_1})}^{k_1^{-1}(\frac{y-T_1-A_1}{B_1})} e^{-\frac{z^2}{2}}\,dz \tag{2.83}$$

$$= E_1 \int_{k_1^{-1}(\frac{T_2-A_1}{B_1})}^{k_1^{-1}(\frac{y-T_1-A_1}{B_1})} (\Phi(k_1^{-1}(\frac{y-2A_1}{B_1}-k(z)))e^{-\frac{z^2}{2}}\,dz \tag{2.84}$$

$$- E_1\Phi(k_1^{-1}(\frac{T_1-A_1}{B_1})))\sqrt{2\pi}(\Phi(k_1^{-1}(\frac{y-T_1-A_1}{B_1}))-\Phi(k_1^{-1}(\frac{T_2-A_1}{B_1}))) \tag{2.85}$$

$$I_2 = B_1 E_2 \int_{T_2}^{y-T_1} (\Phi(k_1^{-1}(\frac{y-x-A_1}{B_1}))-\Phi(k_1^{-1}(\frac{T_1-A_1}{B_1})))e^{-\frac{(k_2^{-1}(\frac{y-A_2}{B_2}))^2}{2}}(k_2'(k_2^{-1}(\frac{x-A_2}{B_2}))^{-1} \tag{2.86}$$

$$= E_2 \int_{k_2^{-1}(\frac{T_2-A_2}{B_2})}^{k_2^{-1}(\frac{y-T_1-A_2}{B_2})} (\Phi(k_1^{-1}(\frac{y-B_2u-A_1-A_2}{B_2}))-\Phi(k_1^{-1}(\frac{T_1-A_1}{B_1})))e^{-\frac{z^2}{2}}\,dz \tag{2.87}$$

$$= E_2 \int_{k_2^{-1}(\frac{T_2-A_2}{B_2})}^{k_2^{-1}(\frac{y-T_1-A_2}{B_2})} (\Phi(k_1^{-1}(\frac{y-B_2u-A_1-A_2}{B_2}))e^{-\frac{z^2}{2}}\,dz \tag{2.88}$$

$$- \sqrt{2\pi}E_2\Phi(k_1^{-1}(\frac{T_1-A_1}{B_1})))(\Phi(k_2^{-1}(\frac{y-T_1-A_2}{B_2}))-\Phi(k_2^{-1}(\frac{T_2-A_2}{B_2}))) \tag{2.89}$$

$$I_3 = B_2 E_3 \int_{T_2}^{y-T_1} (\Phi(k_2^{-1}(\frac{y-x-A_2}{B_2}))-\Phi(k_2^{-1}(\frac{T_2-A_2}{B_2})))e^{-\frac{(k_1^{-1}(\frac{y-A_1}{B_1}))^2}{2}}(k_1'(k_1^{-1}(\frac{x-A_1}{B_1}))^{-1} \tag{2.90}$$

$$= E_3 \int_{k_2^{-1}(\frac{T_2-A_2}{B_2})}^{k_2^{-1}(\frac{y-T_1-A_2}{B_2})} (\Phi(k_2^{-1}(\frac{y-x-A_2}{B_2}))-\Phi(k_2^{-1}(\frac{T_2-A_2}{B_2})))e^{-\frac{z^2}{2}}\,dz \tag{2.91}$$

$$= E_3 \int_{k_2^{-1}(\frac{T_2-A_2}{B_2})}^{k_2^{-1}(\frac{y-T_1-A_2}{B_2})} (\Phi(k_2^{-1}(\frac{y-B_1k_1(z)-A_1-A_2}{B_2}))e^{-\frac{z^2}{2}}\,dz \tag{2.92}$$

$$- E_3\Phi(k_2^{-1}(\frac{T_2-A_2}{B_2}))\sqrt{2\pi}(\Phi(k_2^{-1}(\frac{y-T_1-A_2}{B_2}))-\Phi(k_2^{-1}(\frac{T_2-A_2}{B_2}))) \tag{2.93}$$

$$I_4 = B_2 E_4 \int_{T_2}^{y-T_1} (\Phi(k_2^{-1}(\frac{y-x-A_2}{B_2}))-\Phi(k_2^{-1}(\frac{T_2-A_2}{B_2})))e^{-\frac{(k_2^{-1}(\frac{y-A_2}{B_2}))^2}{2}}(k_2'(k_2^{-1}(\frac{x-A_2}{B_2}))^{-1} \tag{2.94}$$

$$= E_4 \int_{k_2^{-1}(\frac{T_2-A_2}{B_2})}^{k_2^{-1}(\frac{y-T_1-A_2}{B_2})} (\Phi(k_2^{-1}(\frac{y-B_2k_2(z)-A_1-A_2}{B_2}))-\Phi(k_2^{-1}(\frac{T_2-A_2}{B_2})))e^{-\frac{z^2}{2}}\,dz \tag{2.95}$$

$$= E_4 \int_{k_2^{-1}(\frac{T_2-A_2}{B_2})}^{k_2^{-1}(\frac{y-T_1-A_2}{B_2})} (\Phi(k_2^{-1}(\frac{y-B_2k_2(z)-A_1-A_2}{B_2}))e^{-\frac{z^2}{2}}\,dz \tag{2.96}$$

$$- E_4\Phi(k_2^{-1}(\frac{T_2-A_2}{B_2}))\sqrt{2\pi}(\Phi(k_2^{-1}(\frac{y-T_1-A_2}{B_2}))-\Phi(k_2^{-1}(\frac{T_2-A_2}{B_2}))) \tag{2.97}$$

28

therefore, we have (2.66). □

2.5   Monte Carlo VaR and CVaR

We aim to simulate the loss process:

$$L_t = \sum_{k=0}^{N_t} Y_k \qquad (2.98)$$

where $(N_t)_{t\geq 0}$ is a Poisson process with intensity $\lambda > 0$ describing the frequency process or number of losses on $[0,t]$ and $(Y_i)_{i\in\mathbb{N}}$ is a sequence of independent random variables with common g-and-h probability distribution and parameters A, B, g and h. Both random elements are independent. For notational simplicity we drop the dependence on cell i and fix the time period , typically a year, dropping the time dependency in equation (2.98). Hence we write N and L instead of N_t and L_t.

Simulation of severities, i.e. numbers coming from a g-and-h distribution is a matter of evaluating a non-linear function of standard normal random variables. Algorithms to simulate the latter as well as frequencies given by a Poisson distribution, are widely known. By combining both we will be able to simulate losses in a single cell. The algorithm goes in the following lines:

Algorithm 2.10.   (1) *Inputs: A,B,g,h,λ (parameters), n(number of simulated losses).*

(2) *Generate n Poisson random variables with parameter λ given by a vector N.*

(3) *Generate $N \times n$ standard normal random numbers given by matrix Z.*

(4) *Generate $N \times n$ g-and-h random numbers given by $Y = k(Z)$.*

29

(5) *Compute the losses $L_i = \sum_{k=0}^{N_i} Y_k$, $i = 1, 2, \ldots, n$.*

In Figure 2.4 both the theoretical (blue line) and empirical (green line) g-and-h density functions are shown. The parameters are $A = 0, b = 1, g = 2, h = 0.25$. The empirical p.d.f. is obtained by simulating $n = 10000$ g-and-h numbers according to the previous algorithm with the same parameters via a Gaussian kernel. We observe a reasonable agreement between both.
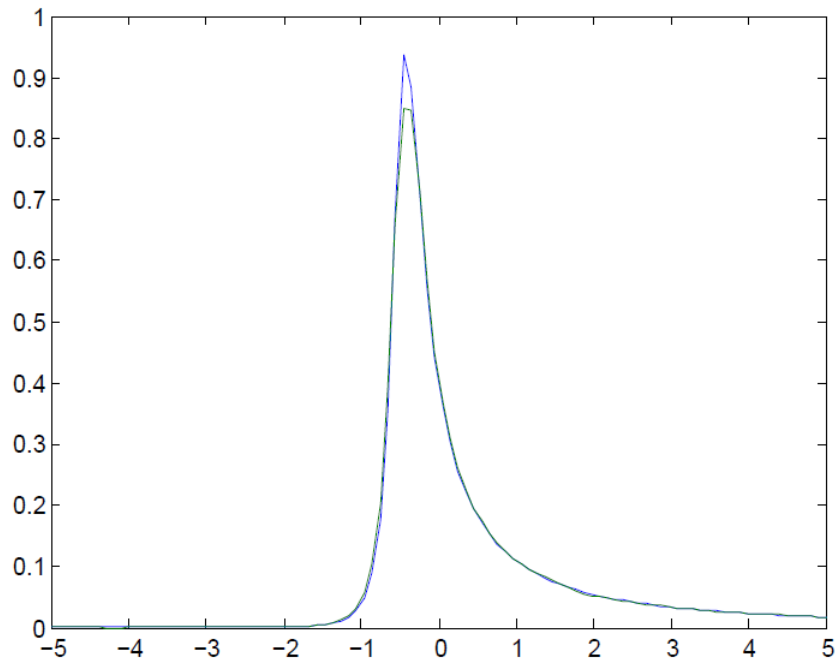
Figure 2.4: Empirical VS Theoretical g-and-h density functions

Next we show in figure 2.5 the empirical density of the loss distribution (green line) in

comparison with a normal density with the same mean and variance (blue line). Parameters for severity are $A = 100000$, $B = 1$, $g = 2$ and $h = 0.25$.

The empirical p.d.f. of the loss random variable, conditionally on having observed at least one loss event and obtained also by a Gaussian kernel, is observed to have thicker tail than the normal distribution. Frequency of losses is assumed to follow a Poisson distribution with an average $\lambda = 200$ events per year. Notice that actually the loss distribution is not absolutely continuous, it has a positive mass at zero.
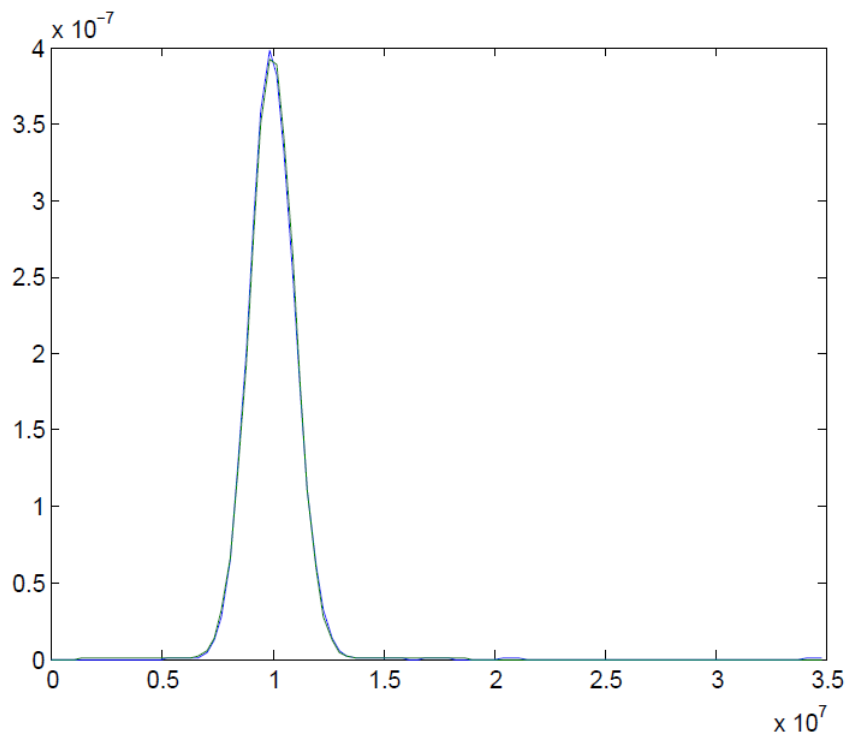
Figure 2.5: Empirical g-and-h VS normal density functions

λ / measure(L)	Mean(L)	STD(L)	Skewness(L)	Excess of kurtosis(L)	Median(L)
$\lambda = 100$	10,003,000	999,980	0.1021	3.0109	10,000,000
$\lambda = 200$	19,999,086	1,415,281	0.0727	3.0036	20,000,752
$\lambda = 300$	30,001,299	1,732,590	0.0562	3.0096	30,001,280

Table 2.1: Statistical Measures of the Losses by Monte Carlo

In table 2.1 we present a summary with relevant statistical information about empirical losses on a single cell of the Operational risk Matrix, obtained by Monte Carlo simulation and assuming a Poisson probability distribution for the frequency with λ = 200 events per year and a g-and-h severity distribution with parameters A = 100000, B = 1, g = 2 and h = 0.25. We observe that asymmetry and excess of kurtosis are translated from the severity to the loss probability distribution formation itself.

Moreover, a quantile comparison between the empirical loss density obtained from a Gaussian kernel and the normal distribution is given in figure 2.6. Quantile levels range from 95% to 99.5%.
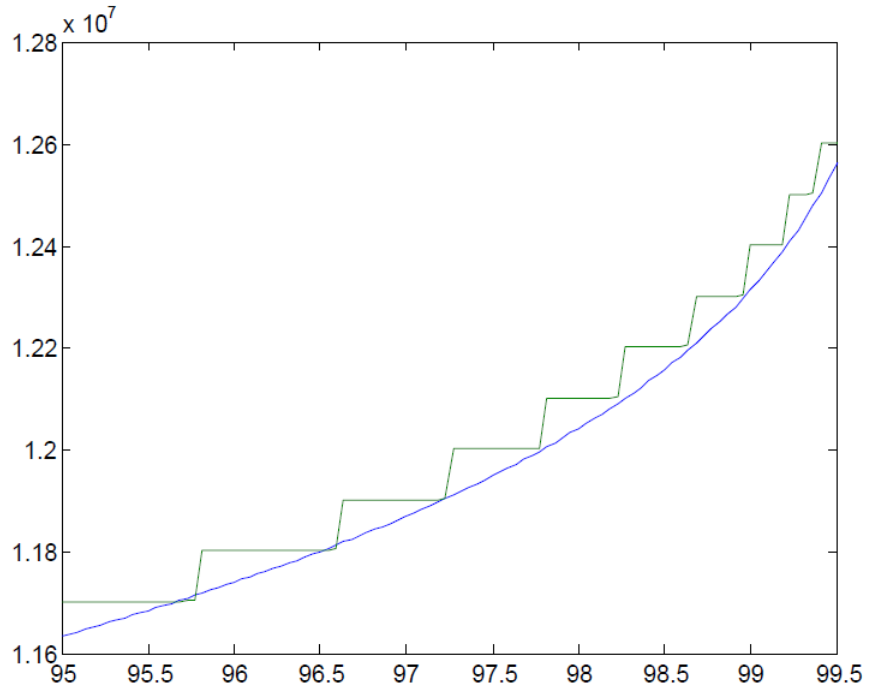
Figure 2.6: Quantile Comparison between g-and-h and normal distribution

In table 2.2 results for the Var and CVaR on a single cell are shown, using a model with Poisson frequencies of intensity $\lambda = 200$ and a range of values for the VaR/CVaR levels. VaR is computed by obtaining the correct percentile from the sequence of generated losses. CVaR is computed according to equation 2.16

In table 2.3 the dependence of VaR and CVaR on the parameters of the severity are shown.

measure/α	95%	97.5 %	99 %	99.5%
VaR	22,400,458	22,801,680	23,400,597	23,701,560
CVaR	22,975,101	23,372,236	23,852,866	24,174,057

Table 2.2: Dependence of VaR and CVaR with $\lambda = 200$ on Quantile α

measure/h	0.1	0.25	1	2
VaR	23,701,306	23,701,560	26,790,688	24,899,532,560
CVar	24,235,562	24,174,057	1.121×10^9	99.624×10^{15}

Table 2.3: Dependence of VaR and CVaR with $\lambda = 200$ on severity parameter h with $\alpha = 99.5\%$ (One can observe that $h \geq 1$ is unrealistic)

The algorithm to generate losses for a single cell, with severities coming from a mixed truncated g-and-h distribution, and frequencies representing external and internal losses requires minor modifications from Algorithm 2.10.

Notice that a Poisson process having intensity λ, counting two type of events representing internal and external losses, is equivalent to consider two independent Poisson processes with respective intensities $\lambda_1 = p\lambda$ and $\lambda_1 = (1-p)\lambda$, where p is the probability to have an internal event. We can compute separately the losses in cell i due to internal events, denoted by $L(i,1)$, from the losses due to external events, denoted by $L(i,2)$. It leads to the following algorithm:

Algorithm 2.11. (1) *Inputs: $A_1, B_1, g_1, h_1, \lambda, A_2, B_2, g_2, h_2, n, p.$*

(2) *Generate n Poisson random variables with parameter $\lambda_1 = p\lambda$ and $\lambda_1 = (1-p)\lambda$ given by vectors $N_i, i = 1, 2$.*

(3) *Generate $N_i \times n$ standard normal random numbers given by matrices $Z_i, i = 1, 2$.*

(4) *Generate $N_i \times n$ g-and-h random numbers given by $Y_i = k_i(Z), i = 1, 2$.*

(5) *Compute the losses $L(i, j) = \sum_{k=0}^{N_i(j)} Y_k^{(j)}$, $i = 1, 2$, $j = 1, 2, \ldots, n$.*

(6) *Compute the aggregate loss $L(1, j) + L(2, j)$.*

In figure 2.7 it can be observed the empirical density of the loss for a single cell. Parameters used in the simulation are $A_1 = 10000, B_1 = 1, h_1 = 0.25, g_1 = 2$ for the severities of internal data and $A_2 = 20000, B_2 = 1, h_2 = 0.3, g_2 = 2$ for external data, i.e. a distribution with heavier tails and more skewed. The frequency of loss events is a Poisson distribution with an average of $\lambda = 100$ events per year. The losses are chosen from external and internal data at 50%. It is compared with a normal distribution with the same mean and variance. Heavier tails can be observed.
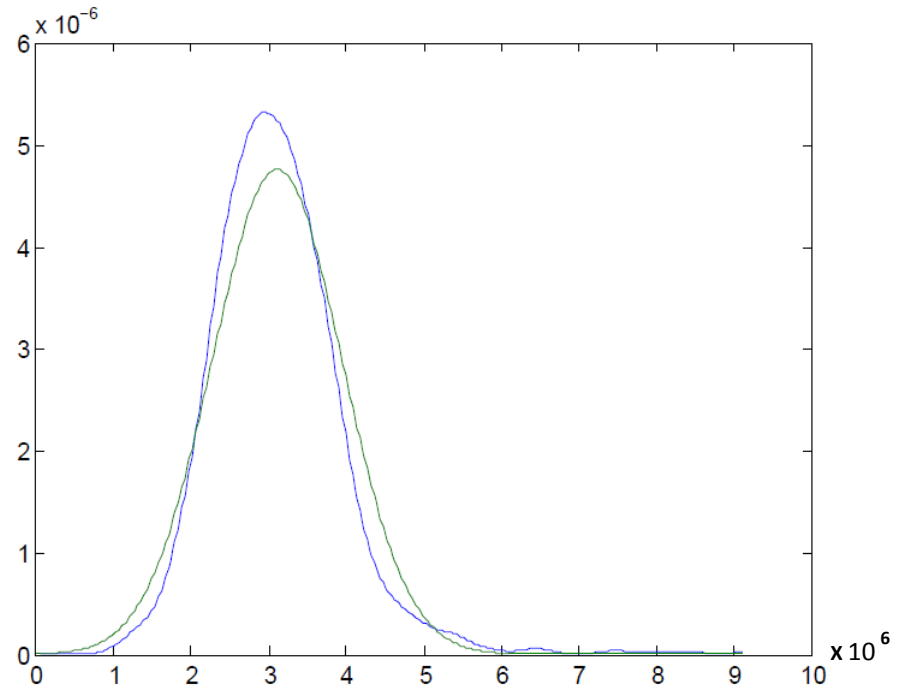
Figure 2.7: Mix internal and external g-and-h VS normal distribution

Table 2.4 shows the results of mixing internal and external data

In figure 2.8 percentiles of the loss probability distributions with g-and-h severities following a mixture of g-and-h distributions and mixed Poissonian frequencies with parameters as above are contrasted with the percentiles of an equivalent normal distribution (green line). Table 2.4, graphs 2.7 and 2.8are obtained using losssincelmix.m

36

measure/α	95%	97.5 %	99 %	99.5%
VaR	1,770,355	1,820,537	1,880,501	1,920,866
CVaR	1,837,688	1,886,774	1,944,731	1,981,561

Table 2.4: Dependence of VaR and CVaR on Quantile mixing internal and external data
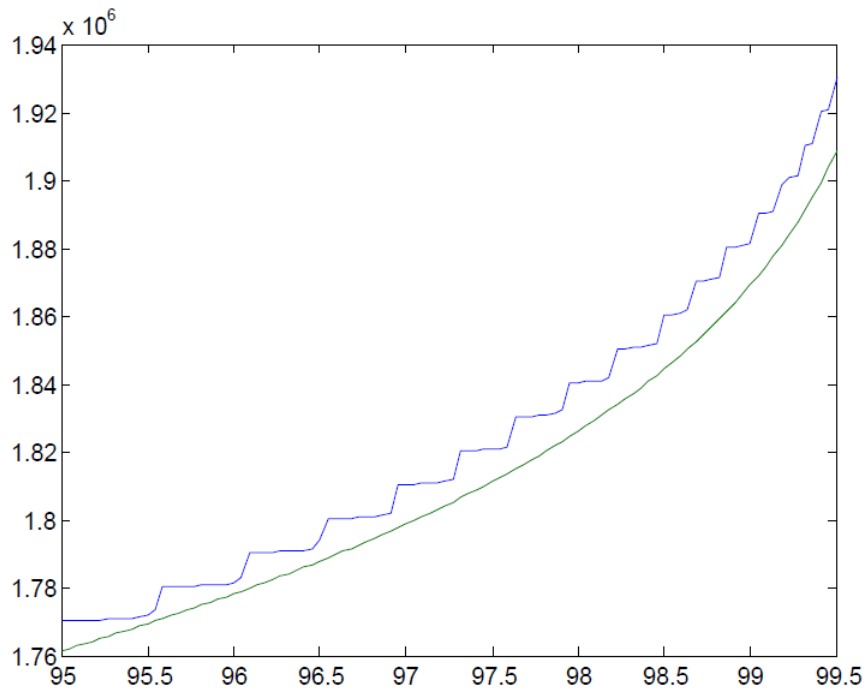
Figure 2.8: VaR comparison between g-and-h and normal distributions mixing internal and external data

2.6 Computing VaR/CVaR by recursive convolution

Alternatively to a Monte Carlo approach, the value-at-risk at level α for a single cell can be computed by solving numerically the equation $F_L(x) = \alpha$, where F_L verifies equation (2.17). The series need to be truncated at some value M.

Convolution products can be computed recursively following a similar approach than in Lemma 2.4.

More precisely, for the case when severities are given by a random variable with g-and-h probability distribution and $k \geq 3$ the recursion formula follows as:

$$
\begin{aligned}
F_X^{*k}(y) &= \int_{\mathbb{R}} F_X^{*(k-1)}(y-x)dF_Y(x) = \int_{\mathbb{R}} F_X^{*(k-1)}(y-x)f_X(x)dx \\
&= \int_{\mathbb{R}} F_X^{*(k-1)}(y-x)f_Z(k^{-1}(x))(k'(k^{-1}(x)))^{-1}dx \\
&= \int_{\mathbb{R}} F_X^{*(k-1)}(y-k(z))f_Z(z)dz \\
&= \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} F^{*(k-1)}(y-k(z))e^{-\frac{z^2}{2}}dz
\end{aligned}
\tag{2.99}
$$

First, we compute the convolution product of two independent g-and-h probability distributions, cumulative and density functions, according to formulas (2.12), (2.13) and (2.15). The integrals are approximately calculated by the trapezoid rule. It requires to solve for the inverse of function k at several points determined by the truncation interval over the integral.

In figures 2.9 and 2.10 the respective two fold cumulative and density convolutions functions are shown. The parameters of the distribution are $A = 0, B = 1, g = 2$ and $h = 0.25$. Interval in the trapezoid rule is $[-3, 3]$ while a number of 100 points were used.

38
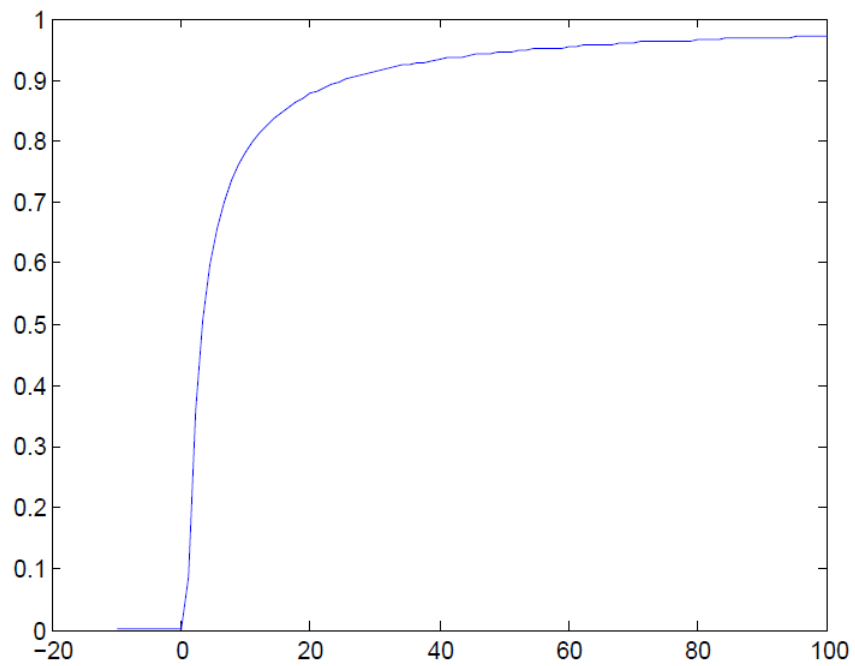
Figure 2.9: Two fold cumulative convolution function
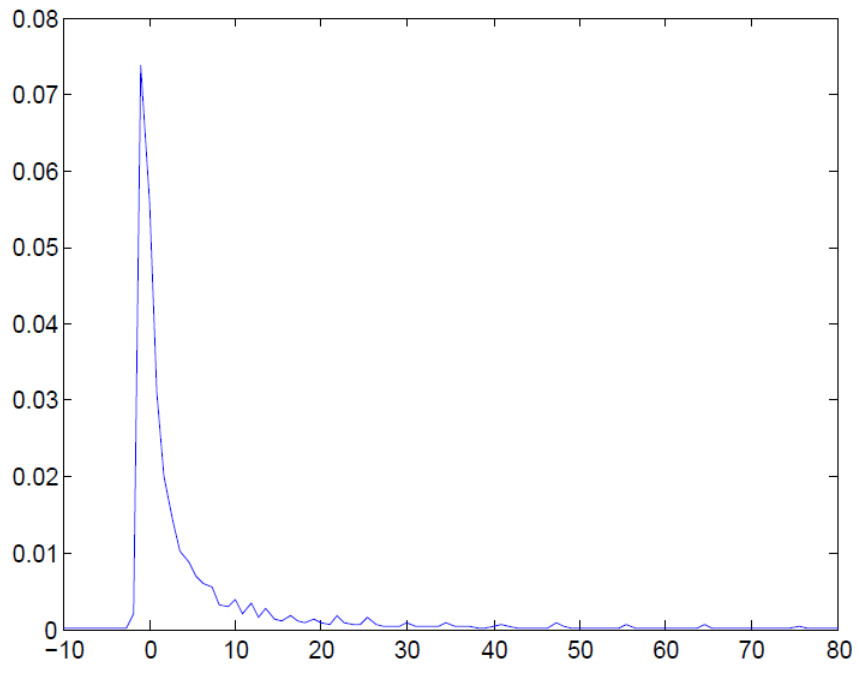
Figure 2.10: Two fold density convolution function

The algorithm to compute the VaR goes in the following lines:

Algorithm 2.12. (1) *Inputs: parameters:A,B,g,h, truncation levels: my,My,mz,Mz,TL, points in trapezoid:M.*

(2) *Compute the two-fold c.d.f. convolution F^{*2}.*

(3) *Compute the k-th fold c.d.f.convolution F^{*k} from formula (2.99), $k = 3,\ldots,TL$.*

(4) *Compute F_L.*

(5) *Solve $F_L(x) = \alpha$ to compute VaR_α.*

(6) *Compute the conditional VaR, $ES_\alpha = \frac{1}{1-\alpha}\int_\alpha^1 VaR_\beta d\beta$.*

When combining external and internal data with truncated values, the algorithm is similar, but now we use equation (2.74) for the p.d.f. Hence:

$$
\begin{aligned}
F_X^{*k}(y) \quad = \quad & \Big[p\frac{1}{(1-\Phi(k_1^{-1}(\frac{T_1-A_1}{B_1})))B_1\sqrt{2\pi}} \\
& \cdot \int_{\mathbb{R}} F_X^{*(k-1)}(y-x)e^{-\frac{(k_1^{-1}(\frac{x-A_1}{B_1}))^2}{2}}(k_1'(k_1^{-1}(\frac{x-A_1}{B_1})))^{-1}\mathbb{1}_{[T_1,+\infty]}(x)dx \Big] \\
+ \quad & \Big[(1-p)\frac{1}{(1-\Phi(k_2^{-1}(\frac{T_2-A_2}{B_2})))B_2\sqrt{2\pi}} \\
& \cdot \int_{\mathbb{R}} F_X^{*(k-1)}(y-x)e^{-\frac{(k_2^{-1}(\frac{x-A_2}{B_2}))^2}{2}}(k_2'(k_2^{-1}(\frac{x-A_2}{B_2})))^{-1}\mathbb{1}_{[T_2,+\infty]}(x)dx \Big]
\end{aligned}
$$

Although we have not implemented numerically, the algorithm seems to require a considerable computational effort.

2.7 Computing VaR/CVaR by Fast Fourier Transform

Here, we will study a methodology to compute VaR/CVaR via FFT. Note that the characteristic function of the loss process driven by the model described by equation (2.98), denoted by $\varphi_{L_t}(u) = E(e^{iuL_t})$ is computed as follows:

$$
\begin{aligned}
\varphi_{L_t}(u) &= E\left[\left(Ee^{iuL_t}/N_t\right)\right] = E(\prod_{k=0}^{N_t} E\left(e^{iuY_k}\right) = E(E\left[(e^{iuY_k})^{N_t}\right] \\
&= E(\varphi_Y(u))^{N_t}) = E(e^{\log(\varphi_Y(u))N_t}) = E(e^{i(-i\log(\varphi_Y(u)))N_t}) = \varphi_{N_t}(-i\log(\varphi_Y(u))) \\
&= e^{\lambda t(\varphi_Y(u)-1)}
\end{aligned}
$$

where $\varphi_Y(u)$ is the characteristic function of a random variable Y with probability distribution $g - and - h(A, B, g, h)$.

On the other hand the characteristic function of the severity Y following a g-and-h distribution can be computed as:

$$
\begin{aligned}
\varphi_Y(u) &= e^{iuA}\varphi_Y(uB) = \frac{e^{iuA}}{B}\int_{\mathbb{R}} e^{iuBx} f_Y(x)dx \\
&= \frac{e^{iuA}}{B}\int_{\mathbb{R}} e^{iuBx} f_Z(k^{-1}(x))(k'(k^{-1}(x))^{-1}dx \\
&= \frac{e^{iuA}}{B}\int_{\mathbb{R}} e^{iuBk(z)} f_Z(z)dz \quad\quad (2.100)
\end{aligned}
$$

By inversion Fourier formula, for $x, y > 0$:

$$
\begin{aligned}
F_{L_t}(x) - F_{L_t}(y) &= \frac{1}{2\pi} \lim_{T\to+\infty} \int_0^T \frac{e^{-ixu} - e^{-iyu}}{iu} \varphi_{L_t}(u)du \\
&= \frac{1}{2\pi} \lim_{T\to+\infty} \int_0^T \frac{e^{-ixu} - e^{-iyu}}{iu} e^{\lambda t(\varphi_Y(u)-1)} du \quad\quad (2.101) \\
f_{L_t}(x) &= \frac{1}{2\pi} \int_0^{+\infty} e^{-ixu} \varphi_{L_t}(u)du = \frac{1}{2\pi} \int_0^{+\infty} e^{-ixu} e^{\lambda t(\varphi_Y(u)-1)} du \quad (2.102)
\end{aligned}
$$

where the density is understood conditionally on the loss being greater than zero.

Efficient methods to compute the integrals above are based on the Fast Fourier Transform, see for example Bailey and Swarztrauber (1993) for details. Moreover, its efficiency can me improved with the use of a Fractionary Fourier Transform (FRFT). The latter has been applied by Chourdakis (2004) in connection to option prices. The method FRFT offers a great advantage over FFT when the c.d.f. is calculated repetitively over only a small interval, as it is the case when we try to compute the VaR solving the equation $F_{L_t}(x) = \alpha$. FRFT provides more flexibility in choosing the grid about which the function is to be evaluated.

Our approach is based on a double approximation. First, the characteristic function of the loss has to be computed from the integrals in expression (2.100) then approximated by (2.101). It goes in the following lines:

$$\frac{e^{iuA}}{B} \int_{\mathbb{R}} e^{iuBk(z)} f_Z(z) dz \simeq \frac{e^{iuA}}{B} \int_{-M_1}^{M_1} e^{iuBk(z)} f_Z(z) dz \qquad (2.103)$$

$$= \frac{e^{iuA}}{B} \int_{-M_1}^{M_1} e^{iuz} e^{iuB(k(z)-\frac{z}{B})} f_Z(z) dz = \int_{-M_1}^{M_1} e^{iuz} h_1(z) dz \qquad (2.104)$$

$$= \int_{-M_1}^{0} e^{iuz} h_1(z) dz + \int_{0}^{M_1} e^{iuz} h_1(z) dz \qquad (2.105)$$

$$= \int_{0}^{M_1} e^{-iuz} h_1(-z) dz + \int_{0}^{M_1} e^{iuz} h_1(z) dz \qquad (2.106)$$

where $h_1(z) = \frac{1}{B} e^{iuB(k(z)-\frac{z}{B})+iuA} f_Z(z)$.

Now, we build a grid of N points: $(z_k = \beta k)_{k=0,1,\dots,N-1}$, and the corresponding grid in the space of the characteristic function $\varphi_Y(u)$ given by $(u_j)_{j=0,1,\dots,N-1}$, $u_j = -\frac{N}{2}\delta + \delta j$, where β and δ are the sizes of the subintervals in the corresponding grids.

43

Next using a trapezoidal rule we approximate both integrals at any point u_j as follows:

$$
\begin{aligned}
\int_0^{M_1} e^{iu_j z} h_1(z) dz &= \int_0^{N\beta} e^{iu_j z} h_1(z) dz \\
&\simeq \beta \sum_{k=0}^{N-1} e^{iu_j z_k} h_1(z_k) w_k = \beta \sum_{k=0}^{N-1} e^{i(-\frac{N}{2}\delta + \delta j)\beta k} h_1(z_k) w_k \\
&= \beta \sum_{k=0}^{N-1} e^{i(-\frac{N}{2}\delta + \delta j)\beta k} h_1(z_k) w_k = \beta \sum_{k=0}^{N-1} e^{i\delta\beta jk} e^{-i\frac{N}{2}\delta\beta k} h_1(z_k) w_k \\
&= \beta \sum_{k=0}^{N-1} e^{i\delta\beta jk} h_2(z_k)
\end{aligned}
$$

where $w_k = 2$ for $k = 1, \ldots N - 2$ and $w_0 = w_{N-1} = 1$, $h_2(z_k) = h_2(k) = e^{-i\frac{N}{2}\delta\beta k} h_1(z_k) w_k$

On the other hand, the inverse discrete FT of the vector $h_2 = (h_k)$ is given by:

$$
IDFT(h_2) = \frac{1}{N} \sum_{k=0}^{N-1} e^{i\frac{2\pi}{N} jk} h_2(k)
$$

Then, by identifying the terms we find the relationship between $\delta = \frac{2\pi}{N\beta}$.

In summary we have the approximations:

$$
\begin{aligned}
\int_0^{M_1} e^{iu_j z} h_1(z) dz &= \beta N IDFT(h_2) \\
\int_0^{M_1} e^{iu_j z} h_1(z) dz &= \beta DFT(h_3)
\end{aligned}
$$

where $h_3(k) = h_2(-z_k)$ and $DFT(h_3)$ is the discrete FT of the vector h_3.

As explained above discrete fractionary FT(DFRT) allows for a more efficient computation of the characteristic function. It generalizes DFT and it is defined for a vector h and $\alpha \in \mathbb{R}$ as:

$$
DFRT_k(h, \alpha) = \sum_{j=0}^{N-1} e^{-i2\pi jk\alpha} h(k)
$$

and its inverse:

$$IDFRT_j(h, \alpha) = \frac{1}{N} \sum_{k=0}^{N-1} e^{i2\pi jk\alpha} h(j)$$

The choice of $\alpha = \frac{1}{N}$ leads to the standard FT approach. Actual choice of α will depend on the interval on which the values of the characteristic function is calculated and the precision required.

Moreover, as in Bailey and Swarztrauler (1991,1994) the fractionary FT can be calculated by applying three times a standard FT with $2N$ points. Indeed:

$$DFRT_k(h, \alpha) = e^{-i\pi k^2 \alpha} \circ IDFT(DFT(y) \circ DFT(z))$$

where DFT and IDFT and the discrete direct and inverse FT respectively, and the vectors y and z are defined as:

$$y = ((h_j e^{-i\pi j^2 \alpha})_{j=0,\dots N-1}, 0_{j=0,\dots N-1}) \tag{2.107}$$

$$z = ((e^{i\pi j^2 \alpha})_{j=0,\dots N-1}, (e^{i\pi(N-j)^2\alpha})_{j=0,\dots N-1}) \tag{2.108}$$

Here $v \circ u$ represents the Hadammard or componentwise product.

Now, we proceed to compute the c.d.f. of the loss L_t given by equation (2.101). Notice that it is continuous everywhere except at $x = 0$. Moreover, for large values of the parameter λt for the intensity of frequencies, as the ones found in Operational Risk context, the jump size at zero is associate with the probability of zero losses is $exp(-\lambda t)$ which becomes a negligible quantity. Therefore, for practical purposes the c.d.f. is considered continuous.

For a large value M_2, using again a trapezoidal approximation on grids $x = \{x_k = \beta_2 k, k = 0, \dots, N-1\}$ and

$(u) = \{u_j = \delta_2 j, j = 0, \ldots, N-1\}$. We denote by $\hat{\varphi}_{L_t}(u)$ the estimated characteristic function using the previous procedure. Then:

$$
\begin{aligned}
F_{L_t}(x_k) - F_{L_t}(x_{k-1}) \quad &\simeq \quad \frac{1}{2\pi} \int_0^{M_2} \frac{e^{-ix_{k-1}u} - e^{-ix_k u}}{iu} \hat{\varphi}_{L_t}(u)\, du \\
&\simeq \quad \frac{\delta_2}{2\pi} \sum_{j=0}^{N-1} e^{-ix_k u_j \alpha} \frac{e^{-i(x_{k-1}-x_k)u_j} - 1}{iu_j} \hat{\varphi}_{L_t}(u_j) w_j \\
&= \quad \frac{\delta_2}{2\pi} \sum_{j=0}^{N-1} e^{-ix_k u_j \alpha} \frac{e^{-i\beta_2 u_j} - 1}{iu_j} \hat{\varphi}_{L_t}(u_j) w_j \\
&= \quad \frac{\delta_2}{2\pi} \sum_{j=0}^{N-1} e^{-i\delta_2 \beta_2 kj\alpha} \frac{e^{-i\beta_2 u_j} - 1}{iu_j} \hat{\varphi}_{L_t}(u_j) w_j \\
&= \quad \frac{\delta_2}{2\pi} \sum_{j=0}^{N-1} e^{-i\delta_2 \beta_2 kj\alpha} h_4(u_j) \\
&= \quad \frac{\delta_2}{2\pi} \sum_{j=0}^{N-1} e^{-i2\pi \delta_2 \beta_2 kj\alpha_1} h_4(u_j) = \frac{\delta_2}{2\pi} DFRT(h_4, \alpha_1) \\
&= \quad \frac{\delta_2}{2\pi} e^{-i\pi k^2 \alpha_1} \circ IDFT(DFT(y_1) \circ DFT(z_1))
\end{aligned}
$$

where $h_4(u_j) = \frac{e^{-i\beta_2 u_j} - 1}{iu_j} \hat{\varphi}_{L_t}(u_j) w_j$, $\alpha_1 = \frac{\beta_2 \delta_2}{2\pi} \alpha$ and

$$
y_1 \quad = \quad \left((h_4(j) e^{-i\pi j^2 \alpha_1})_{j=0,\ldots N-1}, 0_{j=0,\ldots N-1} \right)
$$

$$
z_1 \quad = \quad \left((e^{i\pi j^2 \alpha_1})_{j=0,\ldots N-1}, (e^{i\pi (N-j)^2 \alpha_1})_{j=0,\ldots N-1} \right)
$$

Finally, we recover the c.d.f. from the increments by:

$$
F_{L_t}(x_j) = \sum_{k=1}^{N-1} \left(F_{L_t}(x_k) - F_{L_t}(x_{k-1}) \right)
$$

with $F_{L_t}(x_0) = F_{L_t}(0) = 0$.

We invoke the procedure above several times to compute the c.d.f. of losses $F_{L_t}(x_j)$ in solving the equation $F_{L_t}(x) = \alpha$ by Newton-Rhapson to obtain the VaR_α and the $CVaR_\alpha$. The algorithm can be summarized as follows:

Algorithm 2.13. (1) *Initialization*

    (a) *Select FRFT parameters: α, β_1, β_2 δ_1 and δ_2.*

    (b) *Select loss distribution parameters: λ, A, B, g, h*

(2) *Compute $\varphi_Y(u)$ in a grid $[0, (N-1)\delta_1)$ to produce values $\hat{\varphi}_Y(u)$.*

(3) *Compute $\hat{\varphi}_{L_t}(u) = e^{\lambda(\hat{\varphi}_Y(u)-1)}$*

(4) *Compute $\Delta F_k = F(x_k) - F(x_{k-1})$ in a grid $[0, (N-1)\beta_2)$.*

(5) *Compute $F(x_k) = \sum_{l=1}^{k} \Delta F_l$.*

(6) *Solve $F_{L_t}(x) = \alpha$ by Newton-Rhapson to compute VaR_α.*

(7) *Compute*

$$CVaR_\alpha = \frac{1}{\alpha} \int_\alpha^1 VaR_\beta d\beta$$

by the trapezoidal rule.

Alternatively we compute the characteristic function of φ_Y as above then noting that:

$$\varphi_{\sum_{j=1}^{k} Y_k}(u) \quad = \quad \prod_{j=1}^{k} \varphi_Y(u) = \frac{e^{ikuA}}{B^k} \varphi_X(uB))^k = \frac{e^{ikuA}}{B^k} \left(\int_{\mathbb{R}} e^{iuBk(z)} f_Z(z) dz \right)^k$$

and an estimated of the characteristic function can be obtained as:

$$\hat{\varphi}_{L_t}(u) = e^{-\lambda t} \left(1 + \sum_{k=1}^{N_1} \frac{(\lambda t)^k \hat{\varphi}_Y}{k!} \right)$$

47

2.8   Computing VaR/CVaR by Panjer recurrence

Panjer's recursion has been introduced by Harry J. Panjer in 1981. By using this recursion, we would like to calculate the aggregate loss distribution C of $L_t = \sum_{i=0}^{N_t} Y_i$ assuming a discretized severity F. In this section, we will simply review the method to compute the VaR and CVaR.

DISCRETE VERSION

We define the sequences:

$$f_k \quad := \quad \mathbb{P}(Y_i = kh); k \in \mathbb{N}_0 \tag{2.109}$$

$$p_n \quad := \quad \mathbb{P}(N = n); n \in \mathbb{N}_0 \tag{2.110}$$

$$g_n \quad := \quad \mathbb{P}(L = nh); n \in \mathbb{N}_0; \tag{2.111}$$

$$\tag{2.112}$$

where $\mathbb{N}_0 = \{0, 1, 2,...\}$

For $h > 0$, the size discretized interval. Y_i for $i \geq 1$, the sequence of i.i.d. severity random variables. N is a discrete random variable, independent of the sequence and L, the compound loss e.g a Poisson distribution.

Theorem 2.14. *Let $a,b \in \mathbb{R}$, with $a,b \geq 0$. If for $n = 1,2,....$ and $k = 1,....n$ a distribution follows the recursion*

$$p_n = \left(a + \frac{b}{n}\right)p_{n-1} \tag{2.113}$$

then

$$g_n = \frac{1}{1 - af_0} \sum_{k=1}^{n} (a + b\frac{k}{n}) g_{n-k} f_k \qquad (2.114)$$

and

$$g_0 = w_N(f_0) = \sum_{j=0}^{\infty} p_j f_0^j \qquad (2.115)$$

where w_N is the probability generating function of N.

The proof of theorem 2.14 is based on the book of Rolski(1999) , we need three lemmas.

Lemma 2.15. *For any $j \in \mathbb{N}_0$ and $n = 1,2,...$*

$$\mathbb{E}\left[Y_1 \middle| \sum_{i=1}^{n} Y_i = j\right] = \frac{j}{n} \qquad (2.116)$$

Proof: We set $L_n = \sum_{i=1}^{n} Y_i$. We have:

$$
\begin{aligned}
n\mathbb{E}[X_1|L_n = j] &= \sum_{i=1}^{n} \mathbb{E}[X_i|L_n = j] \\
&= \mathbb{E}\left[\sum_{i=1}^{n} Y_i \middle| L_n = j\right] \\
&= \mathbb{E}\left[L_n \middle| L_n = j\right] \\
&= j
\end{aligned}
$$

Lemma 2.16. *For any $j, k \in \mathbb{N}_0$; $n = 1,2,...$ and f_k^{*n} denoting the n-fold convolution of f_k.*

$$\mathbb{P}\left(Y_1 = k \,\middle|\, \sum_{i=1}^{n} Y_i = j\right) = \frac{f_k f_{j-k}^{*(n-1)}}{f_j^{*n}} \tag{2.117}$$

Proof:

$$\begin{aligned}
\mathbb{P}(X_1 = k | Y_1 + Y_2 + \ldots + Y_n = j) &= \frac{\mathbb{P}(Y_1 = k, Y_2 + \ldots + Y_n = j - k)}{\mathbb{P}(Y_1 + Y_2 + \ldots + Y_n = j)} \\
&= \frac{\mathbb{P}(Y_1 = k)\mathbb{P}(Y_2 + \ldots + X_n = j - k)}{\mathbb{P}(Y_1 + Y_2 + \ldots + Y_n = j)} \\
&= \frac{f_k f_{j-k}^{*(n-1)}}{f_j^{*n}}
\end{aligned}$$

Lemma 2.17. *For any $j \in \mathbb{N}_0$;*

$$\sum_{k=0}^{j} f_k f_{j-k}^{*(n-1)} = f_j^{*n} \tag{2.118}$$

Proof:

$$\begin{aligned}
1 &= \sum_{k=0}^{n} \mathbb{P}\left(Y_1 = k \,\middle|\, \sum_{i=1}^{n} Y_i = j\right) \\
&= \sum_{k=0}^{j} \mathbb{P}\left(Y_1 = k \,\middle|\, \sum_{i=1}^{n} Y_i = j\right) \\
&= \frac{\sum_{k=0}^{j} f_k f_{j-k}^{*(n-1)}}{f_j^{*n}}
\end{aligned}$$

Proof of the theorem.

For n=0

$$
\begin{aligned}
f_0^{*n}(y) &= \mathbb{P}(Y_1 + Y_2 + \dots + Y_n = 0) \\[2ex]
&= \mathbb{P}(Y_1 = 0, Y_2 = 0, \dots, Y_n = 0) \\[2ex]
&= \prod_{i=1}^{n} \mathbb{P}(Y_i = 0) \\[2ex]
&= \mathbb{P}(Y_1 = 0)^n \\[2ex]
&= f_0^n(Y)
\end{aligned}
$$

and for $n \geq 1$; $f_0^{*0} = 0$ and

$$
\begin{aligned}
g_0 &= \mathbb{P}(L = 0) \\[2ex]
&= \mathbb{P}\!\left(\sum_{i=1}^{N} Y_i = 0\right) \\[2ex]
&= \mathbb{P}(N = 0) + \sum_{n=1}^{\infty} \mathbb{P}\!\left(\sum_{i=1}^{n} Y_i = 0 \,\middle|\, N = n\right)\mathbb{P}(N = n) \\[2ex]
&= p_0 + \sum_{n=1}^{\infty} \mathbb{P}(Y_i = 0)^n p_n \\[2ex]
&= w_N(f_0)
\end{aligned}
$$

Also, for $n \geq 1$; $f_0^{*n} = 1$. We have:

$$
\begin{aligned}
g_j &= \sum_{n=0}^{\infty} p_n f_j^{*n} \\
&= p_0 f_j^{*0} + \sum_{n=1}^{\infty} p_n f_j^{*n} \\
&= 0 + \sum_{n=1}^{\infty} \left(a + \frac{b}{n} \right) p_{n-1} f_j^{*n} \\
&= \sum_{n=1}^{\infty} \left(a + b\mathbb{E}\left[\frac{Y_1}{j} \,\Big|\, \sum_{i=1}^{n} Y_i = j \right] \right) p_{n-1} f_j^{*n} \\
&= \sum_{n=1}^{\infty} \left(a + b \sum_{k=0}^{j} \frac{k}{j}.\mathbb{P}\left(Y_1 = k \,\Big|\, \sum_{i=1}^{n} Y_i = j \right) \right) p_{n-1} f_j^{*n} \\
&= \sum_{n=1}^{\infty} \left(a + b \sum_{k=0}^{j} \frac{k}{j}.\frac{f_k f_{j-k}^{*(n-1)}}{f_j^{*n}} \right) p_{n-1} f_j^{*n} \\
&= \sum_{n=1}^{\infty} \left(a f_j^{*n} + b \sum_{k=0}^{j} \frac{k}{j}.f_k f_{j-k}^{*(n-1)} \right) p_{n-1} \\
&= \sum_{n=1}^{\infty} \left(a \sum_{k=0}^{j} f_k f_{j-k}^{*(n-1)} + b \sum_{k=0}^{j} \frac{k}{j}.f_k f_{j-k}^{*(n-1)} \right) p_{n-1} \\
&= \sum_{n=1}^{\infty} \sum_{k=0}^{j} \left(a + b.\frac{k}{j} \right) f_k f_{j-k}^{*(n-1)} p_{n-1} \\
&= \sum_{k=0}^{j} \left(a + b.\frac{k}{j} \right) f_k \sum_{n=1}^{\infty} f_{j-k}^{*(n-1)} p_{n-1} \\
&= \sum_{k=0}^{j} \left(a + b.\frac{k}{j} \right) f_k \sum_{n=0}^{\infty} f_{j-k}^{*(n)} p_n \\
&= \sum_{k=0}^{j} \left(a + b.\frac{k}{j} \right) f_k g_{j-k} \\
&= a f_0 g_j + \sum_{k=1}^{j} \left(a + b.\frac{k}{j} \right) f_k g_{j-k}
\end{aligned}
$$

Rearranging, we get:

$$
(g_j - a f_0 g_j) = \sum_{k=1}^{j} \left(a + b\frac{k}{j} \right) f_k g_{j-k} \leftrightarrow g_j = \frac{1}{1 - a f_0} \sum_{k=1}^{j} (a + b\frac{k}{j}) g_{j-k} f_k
$$

The frequency can only have four discrete distributions: Poisson, binomial, negative

binomial and geometric. In our case as we decided, we will use the Poisson distribution which means that a $= 0$ and $b = \lambda$.We have then

$$p_n = \left(0 + \frac{\lambda}{n}\right)p_{n-1} = \frac{\lambda}{n} \cdot \frac{\lambda^{n-1}}{(n-1)!}e^{-\lambda} = \frac{\lambda^n}{n!}e^{-\lambda} \tag{2.119}$$

Example 3.6. 1. *Suppose $N \sim Poi(\lambda)$ and $f_k = \mathbb{P}(Y_i = k) = \frac{1}{4}$ for $k = 1, 2, 3, 4$ and $f_0 = 0$. $\mathbb{P}(L = nh)$ for $n = 0, 1, 2, 3, 4$. Assume $h = 1$. This example is very basic to be able to do the calculation by hand using Panjer's recursion.*

$f_0 = 0; \ f_1 = f_2 = f_3 = f_4 = \frac{1}{4}. \ a = 0; \ b = \lambda \ ; \ p_0 = e^{-\lambda}; \ w_N(t) = e^{\lambda(t-1)}$

$\mathbb{P}(L = n) = g_n = \frac{\lambda}{n} \sum\limits_{k=1}^{n} k \cdot g_{n-k} f_k$ gives us by the recursion:

$$
\begin{aligned}
g_0 &= w_N(f_0) = w_N(0) = e^{-\lambda} \\[2mm]
g_1 &= \frac{\lambda}{1}g_0 f_1 = \frac{\lambda}{1}g_0\frac{1}{4} \\[2mm]
g_2 &= \frac{\lambda}{2}g_1 f_1 + \frac{2\lambda}{2}g_0 f_2 = \frac{\lambda}{2}g_1\frac{1}{4} + g_1 \\[2mm]
g_3 &= \frac{\lambda}{3}g_2 f_1 + \frac{2\lambda}{3}g_1 f_2 + \frac{3\lambda}{3}g_0 f_3 = \frac{\lambda}{3}g_2\frac{1}{4} + \frac{2\lambda}{3}g_1\frac{1}{4} + g_1 \\[2mm]
g_4 &= \frac{\lambda}{4}g_3 f_1 + \frac{2\lambda}{4}g_2 f_2 + \frac{3\lambda}{4}g_1 f_3 + \frac{4\lambda}{4}g_0 f_4 \\[2mm]
&= \frac{\lambda}{4}g_3\frac{1}{4} + \frac{2\lambda}{4}g_2\frac{1}{4} + \frac{3\lambda}{4}g_1\frac{1}{4} + g_1
\end{aligned}
$$

With the results obtained, the probability for the loss distribution can be calculated by:

$$
g_n = \begin{cases}
e^{-\lambda} & \text{for } n = 0 \\[2mm]
\dfrac{1}{4}\lambda g_0 & \text{for } n = 1 \\[2mm]
\dfrac{1}{4}\dfrac{\lambda}{n}\sum_{k=1}^{n-1} k \cdot g_{n-k} + g_1 & \text{for } n = 2,\, 3,\, 4
\end{cases}
$$

The algorithm to compute Var/CVaR by Panjer recursion is the following:

Algorithm 2.18.   (1) *Initialization: calculate f_0 and g_0 and set $G_0 = g_0$*

(2) *For $n = 1, 2, \ldots$*

    (a) *Calculate f_n.*

    (b) *Calculate $g_n = \dfrac{1}{1-af_0} \sum_{k=1}^{n} \left(a + b\dfrac{k}{n}\right) g_{n-k} f_k$.*

    (c) *Calculate $G_n = G_{n-1} + g_n$.*

    (d) *Interrupt the procedure if G_n is larger than the require quantile level α, e.g.*
        $\alpha = 0.999$. Then the estimate of the quantile q_α is $n \times h$

(3) *Then do an increment $(n+1)$ and return to step 2*

COMPUTING VAR/CVAR IN AGGREGATE LOSS

In Banks and Financial Institutions, Losses occurring in a certain Business line may have effects on the others and lead to losses on the others. Thus, we have a dependence existing between the losses of these business lines. To model this dependence, we are introducing in our thesis the concept of copula. Embrechts et al. (2003) introduced an application of copula to Risk Management.

3.1   Copulas

A d dimensional copula is a d variate probability distribution function on $[0,1] \times ... \times [0,1] = [0,1]^d$ with uniform marginal distributions. A copula $C(u)$ as a function may be defined as follow (see Cherubini(2004)).

$$C(u) = C(u_1, u_2, ... u_d) : [0,1]^d \to [0,1], \quad u = (u_1, u_2, ..., u_d)$$

The above definition implies that for any given copula $C(u)$ there exists a random vector $U = (U_1, U_2, ... U_d)$ with $U_i, i = 1, ..., d$ being a uniform random variable on $[0,1]$ and

$$C(u_1, ..., u_d) = P(U_1 \leq u_1, ..., U_d \leq u_d), \quad u_i \in [0,1], i = 1, ..., d$$

3.1.1   Properties of Copulas

This section presents some useful properties of copulas.

(1) Characterization of Copula: A function $C(u_1, ... u_d)$ defined on $[0,1]^d$ is a copula if and only of the following conditions are satisfied.

    a. $C(u_1, ..., u_d)$ is increasing in each $u_i \in [0,1]$

    b. $C(1, ..., 1, u_i, 1, ..., 1) = u_i$ for all $i = 1, ..., d$

    c. For all $(a_1, ..., a_d)$, $(b_1, ..., b_d) \in [0,1]^d$ with $a_i \le b_i$, $i = 1, ..., d$,

$$\sum_{i_1=1}^{2} \cdots \sum_{i_d=1}^{2} (-1)^{i_1 + ... + i_d} C(u_{1 i_1}, ..., u_{d i_d}) \ge 0$$

    where $u_{j1} = a_j$ and $u_{j2} = b_j$ for all $j = 1, ..., d$

(2) Sklar Theorem: Let F be a d dimensional PDF with marginal distributions $F_1, ..., F_d$. Then, there exists a copula $C(u_1, ..., u_d) : [0,1]^d \to [0,1]$ such that for all $x_1, x_2, ..., x_d \in [-\infty, \infty]$,

$$F(x_1, x_2, ..., x_d) = C(F_1(x_1), ..., F_d(x_d))$$

(3) Invariance Property: Let $(X_1, X_2, ..., X_d)$ be a random vector with continuous marginal distributions and copula C. Let $T_1, ... T_d$ be strictly increasing functions. Then $(T_1(X_1), ..., T_d(X_d))$ also has the copula C.

3.1.2 Fundamental Copulas

This presents some popular copulas and their properties.

(1) The Gaussian Copula: Let $\Sigma \in \mathbb{R}^{d \times d}$ be a correlation matrix. The Gaussian copula with parameter matrix Σ is defined by the following equation.

$$C_\Sigma(u) = \Phi_\Sigma \left(\Phi^{-1}(u_1), ..., \Phi^{-1}(u_d) \right)$$

where Φ^{-1} is the inverse cumulative distribution function of a standard normal distribution Φ_Σ is is the joint cumulative distribution function of a multivariate normal distribution with mean vector zero and correlation matrix Σ.

(2) The T Student Copula: Let $\Sigma \in \mathbb{R}^{d \times d}$ be a correlation matrix. The T student copula with parameter matrix Σ and ν degrees of freedom is defined by the following equation.

$$C_{\Sigma,\nu}(u) = T_{\Sigma,\nu}\left(T_\nu^{-1}(u_1), ..., T_\nu^{-1}(u_d)\right)$$

where T^{-1} is the inverse cumulative distribution function of univariate central Student T distribution with ν degrees of freedom and $T_{\Sigma,\nu}$ is is the joint cumulative distribution function of multivariate Student T distribution with mean vector zero and correlation matrix Σ.

(3) The independence Copula: The independence copula can be defined as follows.

$$C(u) = u_1 u_2 ... u_d$$

3.2 Aggregated losses in two cells

We study now the aggregated losses in two cells of the OR matrix corresponding to the loss event of type j. This loss event affects all business lines to a different extend. For the seek of concreteness we fix two cells (j, k) and (j, m), where (j, k) reflects the j-th loss event at the k-th business line, $j = 1, 2, \ldots, 8$, $k, m = 1, 2, \ldots, 7$. Loss events and business lines are explained in chapter 2.

The losses in the respective cells are given by:

$$L_t(j, m) \quad = \quad \sum_{l=1}^{N_t(j,m)} X_l^{(j,m)} \tag{3.1}$$

$$L_t(j, k) \quad = \quad \sum_{l=1}^{N_t(j,k)} X_l^{(j,k)} \tag{3.2}$$

The aggregated loss in both cells is:

$$L_t^A = L_t(j, m) + L_t(j, k) \tag{3.3}$$

The analysis of the total aggregated loss follows in a similar way, of course at much more computational effort.

We first consider the case of severities given by no-truncated g-and-h distributions coming from internal data. The case of a combination of internal and external truncated data follows in a similar way.

In addition to common loss events, in the seven business lines, it is convenient to take into account loss events that affect exclusively a particular business line. Equations (3.1) and (3.2) are rewritten as:

$$L_t(j, m) \quad = \quad \sum_{l=1}^{N_t^0} X_{0,l}^{(j,m)} + \sum_{l=1}^{N_t(j,m)} X_l^{(j,m)} \tag{3.4}$$

$$L_t(j, k) \quad = \quad \sum_{l=1}^{N_t^0} X_{0,l}^{(j,k)} + \sum_{l=1}^{N_t(j,k)} X_l^{(j,k)} \tag{3.5}$$

where $(N_t^0)_{t\geq0}$, $(N_t(j,m))_{t\geq0}$ and $(N_t(j,k))_{t\geq0}$ are independent Poisson processes with respective intensities λ_0, λ_{jm} and λ_{jk}.

The process $(N_t^0)_{t\geq0}$ accounts for the number of common loss events while $(N_t(j,m))_{t\geq0}$ and $(N_t(j,k))_{t\geq0}$ account for the number of loss events exclusively affecting cells (j,k) and

(j, m) respectively.

Severities coming from exclusive events in two different cells are independent. Also severities coming from a single cell are independent and equally distributed.

On the other hand severities coming from the same loss event, denoted by $X_l^{(j,m)}$ and $X_l^{(j,k)}$ have marginals coming from a g-and-h distribution with respective parameter set $\theta_k = (A_k, B_k, g_k, h_k)$ and $\theta_m = (A_m, B_m, g_m, h_m)$ and dependence driven by a copula C_j.

Next, we follow a Monte Carlo approach to compute the VaR and CVaR of the aggregated loss. The main difficulty is in generating the common dependent severities.

to this end we consider a conditional copula simulation approach, consisting in generating pairs of random variables (U_1, U_2) with dependence driven by the copula C_j , by conditioning on the first component and generating the second one using the inverse transform of the conditional distribution , i.e. $U_2 = C^{-1}(V_1/U_1)$, there U_1 and V_1 are uniform independent random variables in $(0, 1)$.

For copulas within the Archimedian class the inverse of the conditional copula is explicitly available.

Once the pair (U_1, U_2) is generated, successive monotone transforms lead to pair of random variables (Z_1, Z_2) and $(X_{0,l}^{(j,m)}, X_{0,l}^{(j,k)})$ with the same copula dependence than the original uniform random variables and the desired marginal distribution.

We summarize the results in the following proposition:

Proposition 3.1. *Let the loss processes in cells* (j, m) *and* (j, k), $j = 1, 2, \ldots, 8$, $k, m = 1, 2, \ldots, 7$ *be given by equations (3.4) and (3.5) and aggregated loss process given by equation*

(3.3).

Let the pair (U_1, U_2) of random variables on $[0,1]$ with dependence driven by the copula C_j and uniform marginals and define the transforms:

$$X_1 \quad = \quad A_1 + B_1 k_1(Z_1) = A_1 + B_1 k_1(\Phi^{-1}(U_1)) \tag{3.6}$$

$$X_2 \quad = \quad A_2 + B_2 k_2(Z_2) = A_2 + B_2 k_2(\Phi^{-1}(U_2)) \tag{3.7}$$

where:

$$k_l(x) = \frac{e^{g_l x} - 1}{g_l} e^{\frac{h_l x^2}{2}}, x \in \mathbb{R}, g_l > 0, l = 1, 2 \tag{3.8}$$

Then the pair (X_1, X_2) has a copula dependence C_j and marginal g-and-h p.d.f. with parameters $\theta_l = (A_l, B_l, g_l, h_l), l = 1, 2..$

Proof. First, notice that $k_l \circ \Phi^{-1}$ is a strictly increasing function as the composition of two strictly increasing functions. By the inverse transform method $Z_l = \Phi^{-1}(U_l)$ has standard Gaussian p.d.f.

On the other hand by construction X_l has a g-and-h distribution with parameter θ_l.

Finally, strictly increasing transforms preserve the copula dependence. More precisely:

$$F_{X_1, X_2}(x_1, x_2) \quad = \quad P(X_1 \le x_1, X_2 \le x_2) = P(A_1 + B_1 k_1(\Phi^{-1}(U_1)) \le x_1, A_2 + B_2 k_2(\Phi^{-1}(U_2)) \le x_2) \tag{3.9}$$

$$= \quad P(U_1 \le \Phi \circ k_1^{-1}\left(\frac{x_1 - A_1}{B_1}\right), U_2 \le \Phi \circ k_2^{-1}\left(\frac{x_2 - A_2}{B_2}\right)) \tag{3.10}$$

$$= \quad F_{U_1, U_2}(\Phi \circ k_1^{-1}\left(\frac{x_1 - A_1}{B_1}\right), \Phi \circ k_2^{-1}\left(\frac{x_2 - A_2}{B_2}\right)) \tag{3.11}$$

$$= \quad C(\Phi \circ k_1^{-1}\left(\frac{x_1 - A_1}{B_1}\right), \Phi \circ k_2^{-1}\left(\frac{x_2 - A_2}{B_2}\right)) = C(F_{X_1}(x_1), F_{X_2}(x_2)) \tag{3.12}$$

where the last two equalities follow from the relationship between copulas and joint c.d.f. and the expression for the marginal c.d.f. in equation (2.7). □

Remark 3.2. *Once common severities are generated the procedure to simulate losses and to compute VaR and CVaR follow the lines of algorithm 3.9.*

Remark 3.3. *Pairs of random numbers coming from a random vector (U_1, U_2) with copula C and uniform marginals are obtained using a conditional generation approach. See for example Embrechts, Lindskog and McNeil (2001). The conditional probability distribution is obtained by differentiating the copula with respect to the corresponding variable. In the case of bivariate copulas for $0 < x, y < 1$:*

$$P(U_2 \leq y/U_1 = x) \quad = \quad \lim_{h \to 0} P(U_2 \leq y/x \leq U_1 < x + h) \tag{3.13}$$

$$= \quad \lim_{h \to 0} \frac{P(U_2 \leq y, x \leq U_1 < x + h)}{P(x \leq U_1 < x + h)}) \tag{3.14}$$

$$= \quad \lim_{h \to 0} \frac{P(U_2 \leq y, U_1 \leq x + h) - P(U_2 \leq y, U_1 \leq x)}{P(U_1 \leq x + h) - P(U_1 \leq x)}) \tag{3.15}$$

$$= \quad \lim_{h \to 0} \frac{F_{U_1, U_2}(x + h, y) - F_{U_1, U_2}(x, y)}{F_{U_1}(x + h) - F_{U_1}(x)} \tag{3.16}$$

$$= \quad \frac{\frac{\partial F_{U_1, U_2}(x, y)}{\partial x}}{f_{U_1}} = \frac{\partial C(x, y)}{\partial x} \tag{3.17}$$

The algorithm to compute the aggregated VaR and CVaR can be written as:

Algorithm 3.4. *1.- Initialization: Set N, number of simulated losses, $\theta_{jm}^0 = (A_{jm}^0, B_{jm}^0, g_{jm}^0, h_{jm}^0)$, $\theta_{jk}^0 = (A_{jk}^0, B_{jk}^0, g_{jk}^0, h_{jk}^0)$, parameters of the g-and-h distribution for common severities coming from an event type j at business lines m and k respectively, $\theta_{jm} = (A_{jm}, B_{jm}, g_{jm}, h_{jm})$, $\theta_{jk} = (A_{jk}, B_{jk}, g_{jk}, h_{jk})$, parameters of the g-and-h distribution for idiosyncratic severities coming from an event type j at business lines m and k respectively, $\lambda_0, \lambda_{jm}, \lambda_{jk}$ parameters of the Poisson processes, α, level of the Value-at-Risk, T, simulation time.*

2.- *Simulate Poisson processes N_T^0, $N_T(j,m)$ and $N_T(j,k)$.*

3.- *Generate (U_1, U_2) with copula C_j and uniform marginals, following:*

    3a) Generate independent uniform numbers U_1 and V_1 in $(0,1)$.

    3b) Set $U_2 = (D^{(1)}C)^{-1}(V_1/U_1)$

4.- *Compute $Z_1 = \Phi^{-1}(U_1)$ and $Z_2 = \Phi^{-1}(U_2)$.*

5.- *Compute g-and-h severities $X_0^{jm} = k_{jm}(Z_1)$ and $X_0^{jk} = k_{jk}(Z_2)$.*

6.- *Repeat steps 3-5 N_T^0 times to obtain two $1 \times N_T^0$ vectors of common severity losses in cells (j,m) and $j,k)$.*

7.- *Generate $N_T(j,m)$ severities with a g-and-h p.d.f. having parameter θ_{jm}.*

8.- *Generate $N_T(j,k)$ severities with a g-and-h p.d.f. having parameter θ_{jk}.*

9.- *Compute the loss $L_t(j,m)$ at cell (j,m) by adding common and idiosyncratic losses in steps 7 and 8.*

10.- *Compute the loss $L_t(j,k)$ at cell (j,k) by adding common and idiosyncratic losses in steps 7 and 9.*

11.- *Compute the aggregate loss $L_t^A = L_t(j,m) + L_t(j,k)$.*

12.- *Repeat steps 6-11 until complete a vector of $1 \times N$ aggregated losses.*

13.- *Compute the percentile of vector L_t^A to obtain VaR_α.*

14.- Integrate VaR_β on the interval $[\alpha, 1]$ to obtain $CVaR_\alpha$.

For implementation we specifically take a Clayton copula given by equation:

$$C(u, v, \theta) = (u^{-\theta} + v^{-\theta} - 1)^{-\frac{1}{\theta}}, \theta > 0$$

The Clayton copula has low dependence tail, with coefficient on lower dependence tail $\lambda_L = 2^{-\frac{1}{\theta}}$. Moreover, it has a simple structure, depending on a single parameter, with an explicit expression for the inverse of its first order partial derivative.

Thus:

$$D^{(1)}C_\theta(v/u) = (1 + u^\theta(v^{-\theta} - 1))^{-\frac{\theta+1}{\theta}}, \theta > 0$$

Inverting, we have:

$$(D^{(1)}C)_\theta^{-1}(v/u) = (1 + u^\theta(v^{-\frac{\theta}{\theta+1}} - 1) + 1)^{-\frac{1}{\theta}}, \theta > 0$$

Therefore step 3b) in the algorithm is done by setting:

$$U_2 = (1 + U_1^\theta(V_1^{-\frac{\theta}{\theta+1}} - 1) + 1)^{-\frac{1}{\theta}}$$

which guarantees that the pair (U_1, U_2) has a copula dependence C.

λ / measure(L)	Mean(L)	STD(L)	Skewness(L)	Excess of Kurtosis(L)	Median(L)
$\lambda = [200\ 100\ 100]$	100,841,939	13,671,902	31.4262	1,437	100,277,096
$\lambda = [100\ 200\ 100]$	70,389,713	10,334,457	37.2724	1,785	70,110,517
$\lambda = [100\ 100\ 200]$	70,698,855	25,627,416	59.0799	3,775	70,061,178

Table 3.1: Statistical Measures of the Aggregate Losses in two cells by Monte Carlo

measure/α	95%	97.5 %	99 %	99.5%
VaR	110,571,658	113,174,873	116,654,190	120,651,194
CVaR	121,799,690	130,599,525	169,114,639	259,787,044

Table 3.2: Dependence of VaR and CVaR with $\lambda = [200, 100, 100]$ on Quantile α

In table 3.1 we present a summary with relevant statistical information about empirical losses in two cells of the Operational risk Matrix, obtained by Monte Carlo simulation and assuming a Poisson probability distribution for the frequency

In table 3.2 results for the Var and CVaR for two cells are shown, using a model with Poisson frequencies of intensity $\lambda = [200, 100, 100]$ and a range of values for the VaR/CVaR levels. VaR is computed by obtaining the correct percentile from the sequence of generated losses.

In figure (3.1), perlossint, perlossint2, perlossint3 and perlossint4 respectively represent the variation of the VaR at quantiles 95%, 97.5 %, 99 % and 99.5%.

In figure (3.2), corvar1, corvar2, corvar3 and corvar4 respectively represent the variation

of the CVaR at quantile 95%, 97.5 %, 99 % and 99.5% and different types of losses.

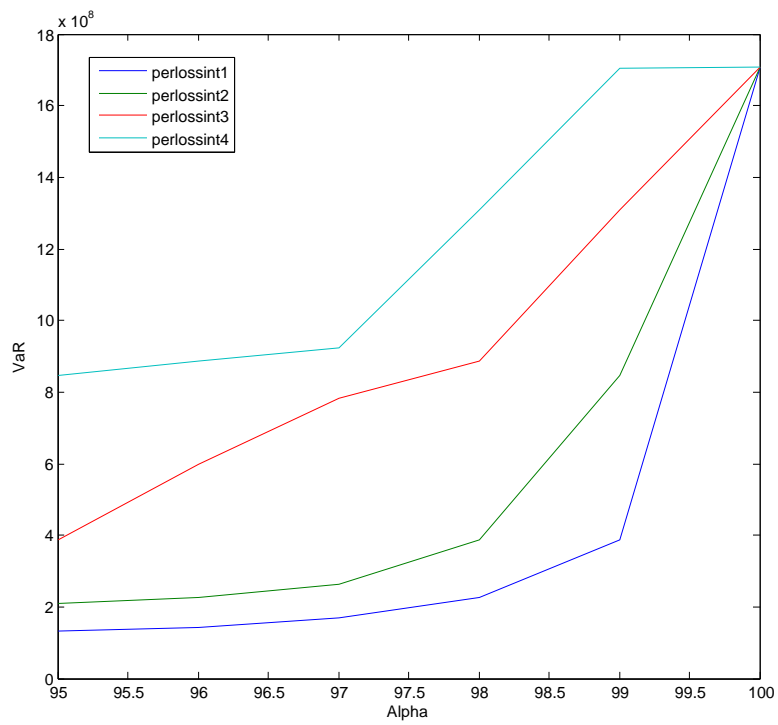Results are obtained using losstwocell.m in Appendix 1.

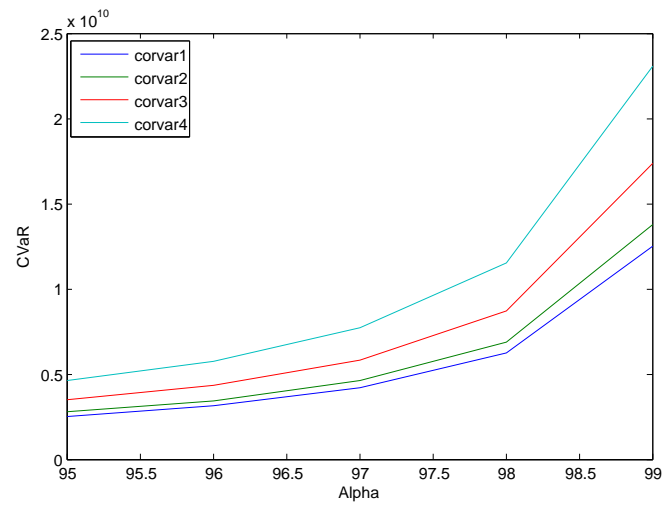Figure 3.1: Dependence of VaR on quantile α

Figure 3.2: Dependence of CVaR on quantile α

3.3 Aggregated losses with mixed truncated g-and-h probability distribution

Consider now the case where severities in both cells (j, m) and (j, k) come from a mixing of truncated internal and external data.

Now, severities on each cell can be written as a combination of internal and external g-and-h random variables truncated at respective thresholds T_I and T_e. Internal severities are related by a copula C_I, while the dependence of external severities is driven by the copula C_e. Internal are external data are independent among them.

The following proposition provides the joint c.d.f. and p.d.f. of severities on both cells, suggesting a method to simulate them.

Proposition 3.5. *Let U_1 and U_2 two independent random variables on $(0, 1)$, $0 < p < 1$ and*

$$Y_1 \quad = \quad 1_{[U_1 < p]} X_1^I + 1_{[U_1 > p]} X_1^e \tag{3.18}$$

$$Y_2 \quad = \quad 1_{[U_2 < p]} X_2^I + 1_{[U_2 > p]} X_2^e \tag{3.19}$$

where (X_1^I, X_2^I) and (X_1^e, X_2^e) are couples of independent random variables indicating internal and external severities losses in cells 1 and 2. The pair (X_1^I, X_2^I) has a dependence given by the copula C_I and the pair (X_1^e, X_2^e) has a dependence given by the copula C_e. The marginal c.d.f. are truncated g-and-h truncated at levels T_l^I and T_l^e, i.e. $X_l^I \sim g - and - h(A_l^I, B_l^I, g_l^I, h_l^I, T_I)$ and $X_l^e \sim g - and - h(A_l^e, B_l^e, g_l^e, h_l^e, T_e)$.

Then, the joint c.d.f. F_{Y_1, Y_2} is given by:

$$F_{Y_1, Y_2}(y_1, y_2) \quad = \quad C_I(F_{X_1^I}(x_1), F_{X_2^I}(x_2))p^2 + [F_{X_1^I}(x_1)F_{X_1^e}(x_2) + F_{X_1^e}(x_1)F_{X_2^I}(x_2)]p(1 - p)$$

$$+ \quad C_e(F_{X_1^e}(x_1), F_{X_2^e}(x_2))(1 - p)^2$$

Proof. From the Law of total Probability:

$$
\begin{aligned}
F_{Y_1,Y_2}(y_1,y_2) &= \mathbb{P}(Y_1 \le y_1, Y_2 \le y_2) \\
&= \mathbb{P}(X_1^I \le y_1, X_2^I \le y_2/U_1 < p, U_2 < p)p^2 + \mathbb{P}(X_1^I \le y_1, X_2^e \le y_2/U_1 < p, U_2 \ge p)p(1-p) \\
&+ \mathbb{P}(X_1^e \le y_1, X_2^I \le y_2/U_1 \ge p, U_2 < p)p(1-p) + P(X_1^e \le y_1, X_2^e \le y_2/U_1 \ge p, U_2 \ge p)(1-p)^2 \\
&= C_I(F_{X_1^I}(x_1), F_{X_2^I}(x_2))p^2 + [F_{X_1^I}(x_1)F_{X_1^e}(x_2)]p(1-p) + [F_{X_1^e}(x_1)F_{X_2^I}(x_2)]p(1-p) \\
&+ C_e(F_{X_1^e}(x_1), F_{X_2^e}(x_2))(1-p)^2 \\
&= C_I(F_{X_1^I}(x_1), F_{X_2^I}(x_2))p^2 + [F_{X_1^I}(x_1)F_{X_1^e}(x_2) + F_{X_1^e}(x_1)F_{X_2^I}(x_2)]p(1-p) \\
&+ C_e(F_{X_1^e}(x_1), F_{X_2^e}(x_2))(1-p)^2
\end{aligned}
$$

Therefore equation for the joint c.d.f. immediately follows from the relationship between copulas and joint c.d.f. □

Chapter 4

METHODS FOR PARAMETER ESTIMATION

In the chapter we review some methods for the estimation of parameters in the model.

We first consider data coming from a well delimited mixing of internal and external data.

The data set consists in observations of the number of internal and external data over the

same period, respectively n_I and n_e respectively. As in previous chapters we assume both

frequencies are independent Poisson distributed with parameters λ_I and λ_e.

Severities coming form internal data have a common distribution $g - and - h(A_I, B_I, g_I, h_I)$

with known truncation level T_I, while external severities have a common distribution $g -$

$and - h(A_e, B_e, g_e, h_e)$ with unknown truncation level T_e.

The set of internal and external observed severities are $x_1, x_2, \ldots x_{n_e}$ and $y_1, y_2, \ldots y_{n_I}$. The

parametric space is:

$$\Theta = \{\theta = (\theta_I, \theta_e, T_e, \lambda_I, \lambda_e),\ \theta_I = (A_I, B_I, g_I, h_I),\ \theta_e = (A_e, B_e, g_e, h_e)\}$$

4.1   Method of Moments

The method of moments is susceptible of application as expression for the moments of a

g-and-h are available for moderated values of the skewness parameter h, see for example

Iglewicz and Martinez (1984).

In the simplest case of non-truncated severity data, well delimited between external and

internal observations the problem requires solving four non-linear equation that equal the-

oretical truncated moments given by equation (2.23) with the empirical four first moments.

Hence, assuming the condition $h < \frac{1}{4}$ to guarantee the existence of the first four moments:

$$E(X/X > T) \;=\; \frac{1}{B_I^2(1 - \Phi(u(T_I, \theta_I)))}\left[A I_0(T_I, \theta_I) + \frac{B_I}{g_I}I_1(T_I, \theta_I)\right] = \bar{x}_I := \frac{1}{n_I}\sum_{j=1}^{n_I} x_j$$

$$E(X^2/X > T) \;=\; \frac{1}{B_I^2(1 - \Phi(u(T_I, \theta_I)))}\left[A_I^2 I_0(T_I, \theta_I) + 2\frac{B_I}{g_I}I_1(T_I, \theta_I) + \frac{B_I^2}{g_I^2}I_2(T_I, \theta_I)\right] = \frac{1}{n_I}\sum_{j=1}^{n_I} x_j^2$$

$$E(X^3/X > T) \;=\; \frac{1}{B_I^2(1 - \Phi(u(T_I, \theta_I)))}\left[A_I^3 I_0(T_I, \theta_I) + 3A_I^2\frac{B_I}{g_I}I_1(T_I, \theta_I)\right.$$

$$+ \;\; \left. 3A_I\frac{B_I^2}{g_I^2}I_2(T_I, \theta_I) + \frac{B_I^3}{g_I^3}I_3(T_I, \theta_I)\right] = \frac{1}{n_I}\sum_{j=1}^{n_I} x_j^3$$

$$E(X^4/X > T) \;=\; \frac{1}{B_I^2(1 - \Phi(u(T_I, \theta_I)))}\left[A_I^4 I_0(T_I, \theta_I) + 4A_I^3\frac{B_I}{g_I}I_1(T_I, \theta_I) + 6A_I^2\frac{B_I^2}{g_I^2}I_2(T_I, \theta_I)\right.$$

$$+ \;\; \left. 4A_I\frac{B_I^3}{g_I^3}I_3(T_I, \theta_I) + \frac{B_I^4}{g_I^4}I_4(T_I, \theta_I)\right] = \frac{1}{n_I}\sum_{j=1}^{n_I} x_j^4$$

where:

$$J_0(T_I, \theta_I) \;=\; 1; J_1(T_I, \theta_I) = 1 - \exp^{\frac{1}{2(1-h_I)}}$$

$$J_2(T_I, \theta_I) \;=\; 1 - J_1(T_I, \theta_I) + \exp^{\frac{2}{1-2h_I}}$$

$$J_3(T_I, \theta_I) \;=\; 1 - J_2(T_I, \theta_I) + \exp^{\frac{9}{2(1-3h_I)}}$$

$$J_4(T_I, \theta_I) \;=\; 1 - J_3(T_I, \theta_I) + \exp^{\frac{8}{1-4h_I}}$$

$$K_0(T_I, \theta_I) \;=\; \Phi(\frac{A_I - T}{B_I}); K_1(T_I, \theta_I) = K_0(T_I, \theta_I) - \exp^{\frac{1}{2(1-h_I)}}\Phi((\frac{A_I - T}{B_I} - \frac{1}{1 - h_I}))(1 - h_I)^{\frac{1}{2}})$$

$$K_2(T_I, \theta_I) \;=\; K_1(T_I, \theta_I) + \exp^{\frac{2}{1-2h_I}}\Phi((\frac{A_I - T_I}{B_I} - \frac{2}{1 - 2h_I}))(1 - 2h_I)^{\frac{1}{2}})$$

$$K_3(T_I, \theta_I) \;=\; K_2(T_I, \theta_I) - \exp^{\frac{9}{2(1-3h_I)}}\Phi((\frac{A_I - T_I}{B_I} - \frac{3}{1 - 3h_I}))(1 - 3h_I)^{\frac{1}{2}})$$

$$K_4(T_I, \theta_I) \;=\; K_3(T_I, \theta_I) + \exp^{\frac{8}{1-4h_I}}\Phi((\frac{A_I - T}{B_I} - \frac{4}{1 - 4h_I}))(1 - 4h_I)^{\frac{1}{2}})$$

$$I_l(T_I, \theta_I)) \;=\; J_l(T_I, \theta_I) - K_l(T_I, \theta_I), l = 1, 2, 3, 4$$

or equivalently the parameters solve the equations:

$$AI_0(T_I, \theta_I) + \frac{B_I}{g_I} I_1(T_I, \theta_I) - B_I^2(1 - \Phi(u(T_I, \theta_I)))\bar{x}_I = 0 \tag{4.1}$$

$$A_I^2 I_0(T_I, \theta_I) + 2\frac{B_I}{g_I} I_1(T_I, \theta_I) + \frac{B_I^2}{g_I^2} I_2(T_I, \theta_I) - B_I^2(1 - \Phi(u(T_I, \theta_I)))\frac{1}{n_I}\sum_{j=1}^{n_I} x_j^2 = 0 \tag{4.2}$$

$$A_I^3 I_0(T_I, \theta_I) + 3A_I^2 \frac{B_I}{g_I} I_1(T_I, \theta_I) \tag{4.3}$$

$$+ \quad 3A_I \frac{B_I^2}{g_I^2} I_2(T_I, \theta_I) + \frac{B_I^3}{g_I^3} I_3(T_I, \theta_I) - B_I^2(1 - \Phi(u(T_I, \theta_I)))\frac{1}{n_I}\sum_{j=1}^{n_I} x_j^3 = 0 \tag{4.4}$$

$$A_I^4 I_0(T_I, \theta_I) + 4A_I^3 \frac{B_I}{g_I} I_1(T_I, \theta_I) + 6A_I^2 \frac{B_I^2}{g_I^2} I_2(T_I, \theta_I) \tag{4.5}$$

$$+ \quad 4A_I \frac{B_I^3}{g_I^3} I_3(T_I, \theta_I) + \frac{B_I^4}{g_I^4} I_4(T_I, \theta_I) - B_I^2(1 - \Phi(u(T_I, \theta_I)))\frac{1}{n_I}\sum_{j=1}^{n_I} x_j^4 = 0 \tag{4.6}$$

with similar equations for external data.

4.2 Maximum Likelihood Estimation

The likelihood, after neglecting terms not depending on θ can be written as:

$$
\begin{aligned}
l(x_I, x_e, \theta) \quad &:= \quad \log L(x_I, x_e, \theta) = \log L(x_I, \theta_I) - \lambda_I + n_I \log(\lambda_I) \\
&+ \quad \log L(x_e, \theta_e, T_e) - \lambda_e + n_e \log(\lambda_e) \\
&= \quad l(x_I, \theta_I) - \lambda_I + n_I \log(\lambda_I) + l(x_e, \theta_e, T_e) - \lambda_e + n_e \log(\lambda_e)
\end{aligned}
$$

As the likelihood splits into four well delimited terms, each one depending on a different part of the parameter vector, they can be treated separately.

We first focus on the part corresponding to the internal data. Then from equation:

$$
\begin{aligned}
l(x_I, \theta_I) &= -\log(B_I)n_I - \log(1 - \Phi(k^{-1}(\frac{T - A_I}{B_I})))n_I - \sum_{m=1}^{n_I} \frac{(k^{-1}(\frac{x_m - A_I}{B_I}))^2}{2} \\
&\quad - \sum_{m=1}^{n_I} \log(k'(k^{-1}(\frac{x_m - A_I}{B_I})))^{-1} \\
&= -\log(B_I)n_I - \log(1 - \Phi(u(T_I, \theta_I)))n_I - \sum_{m=1}^{n_I} \frac{(u^2(x_m, \theta_I)}{2} \\
&\quad - \sum_{m=1}^{n_I} \log(k'(u(x_m, \theta_I)))^{-1}
\end{aligned}
$$

where $u(x, \theta_I)$ solves:

$$
B_I k(u_{(}x, \theta_I)) + A_I - T_I = B_I \frac{e^{gu(x,\theta_I)} - 1}{g} e^{\frac{1}{2}hu^2(x,\theta_I)} + A_I - T_I = 0
$$

Similarly:

$$
\begin{aligned}
l(x_e, \theta_e) &= -\log(B_e)n_e - \log(1 - \Phi(k^{-1}(\frac{T - A_e}{B_e})))n_e - \sum_{m=1}^{n_e} \frac{(k^{-1}(\frac{x_m - A_e}{B_e}))^2}{2} \\
&\quad - \sum_{m=1}^{n_e} \log(k'(k^{-1}(\frac{x_m - A_e}{B_e})))^{-1} \\
&= -\log(B_e)n_e - \log(1 - \Phi(u(T_e - \theta_e)))n_e - \sum_{m=1}^{n_e} \frac{(u^2(x_m, \theta_e)}{2} \\
&\quad - \sum_{m=1}^{n_e} \log(k'(u(x_m, \theta_e)))^{-1}
\end{aligned}
$$

where $v_k(\theta_e)$ solves:

$$
B_e k(v(x, \theta_e)) + A_e - T_e = B_e \frac{e^{gv(x,\theta_e)} - 1}{g} e^{\frac{1}{2}hv^2(x,\theta_e)} + A_e - T_e = 0
$$

The estimations of λ_I and λ_e are computed separately, as they do not depend on the remainder parameter it easily follows that $\hat{\lambda}_I = \frac{n_I}{P_I}$ and $\hat{\lambda}_e = \frac{n_e}{P_e}$, where P_I and P_e are the length of the periods in which the process is observed. To simplify we take $P_I = P_e = 1$.

By differentiating with respect to the parameters we obtain the following set of normal

equations, using chain rule and implicit differentiation:

$$
\begin{aligned}
\frac{\partial l(x_I, \theta_I)}{\partial A_I} &= -\frac{f_Z(u(T_I, \theta_I))}{(1 - \Phi(u_k(\theta_I)))} k''(u(T_I, \theta_I)) \frac{\partial u(T_I, \theta_I)}{\partial A_I} n_I - \sum_{m=1}^{n_I} \frac{(u^2(x_m, \theta_I))}{2} \\
&\quad - \sum_{m=1}^{n_I} \log(k'(u(x_m, \theta_I)))^{-1}
\end{aligned}
$$

and similar expressions for other parameters. Here

$$
\begin{aligned}
k''(x) &= \frac{1}{g} \left[g(g + hx)e^{gx} + he^{gx} - h \right] e^{\frac{1}{2}hx^2} + \frac{1}{g} \left[g(g + hx)e^{gx} + he^{gx} - h \right] e^{\frac{1}{2}hx^2} \\
&\quad + \frac{1}{g} ((g + hx)e^{gx} - hx) hx e^{\frac{1}{2}hx^2} \\
&= \frac{1}{g} e^{\frac{1}{2}hx^2} \left[g(g + hx) + h + (g + hx)e^{gx} - h^2 x^2 - h \right]
\end{aligned}
$$

Maximum likelihood estimators are defined as the vector $\hat{\theta} = argmin_{\theta \in \Theta} l(x, \theta)$, i.e. the

value that minimizes the log-likelihood for all values of θ in the parametric space. They

solve the following system of equations:

$$
\begin{aligned}
\frac{\partial l(x_I, \theta_I)}{\partial A_I} &= 0; \frac{\partial l(x_I, \theta_I)}{\partial B_I} = 0 \\
\frac{\partial l(x_I, \theta_I)}{\partial g_I} &= 0; \frac{\partial l(x_I, \theta_I)}{\partial h_I} = 0
\end{aligned}
$$

Chapter 5

CONCLUSION

In our thesis, under the framework of the Loss Distribution Approach, we explored various methods to compute the Value-at-Risk and Expected Shortfall. Due to the heavy-tailedness and the complexity of the g-and-h severity, we have faced some numerical challenges. Nevertheless, we implemented some of them and verified their accuracy in a way that was not done before. After exploring the one cell case, we analyzed the two cells case while considering and modeling the dependency among cells via copula. For the single cell case:

1) We explored and implemented the Monte Carlo algorithm for internal and mixed data,

2) we reviewed the properties of the recursive convolution and applied truncation to the algorithm for mixed data,

3) We explored the Fractionary Fourier Transform algorithm, and finally,

4) We reviewed the theorem and algorithm of Panjer's recurrence.

For the two cells case, we used the Monte Carlo Simulation to Compute The VaR and CVaR. We used the Clayton Copula, which has a simple structure with an explicit expression for the inverse of its first order partial derivative, to model the dependence among cells. Previous studies have proven that Monte Carlo is very reliable. Fast Fourier Transform is

quick but is dependent on the parameters chosen. A wrong estimation of these parameters could lead to erroneous results. One could compare FFT and Monte Carlo to obtain accurate parameters.

We saw that the methods of computing the two risk measures of our analysis, namely VaR and ES/CVaR were actually proven and good methods for stable risk measures. One could think in our case (heavy-tail severity) that CVaR, given the fact that it is higher, is better since the end goal is for the banks to have enough capital reserve to avoid bankruptcy. Also, in the future perhaps, one could find more results on the g-and-h distribution and use the methods introduced in this analysis.

Appendices

Appendix A

ONE CELL CASE

(1) densitydependgh.m plots the density functions of g-and-h dependeding on parameters g and h

```matlab
1       %running different density g-and-h
2       A=0; B=1; gv=[1,2,3];h=0.25;
3       hv=[ .2 , .25 , .5 , .8 ];g=2;
4       fg1=densigandh(A,B,gv(1),h);
5       fg2=densigandh(A,B,gv(2),h);
6       fg3=densigandh(A,B,gv(3),h);
7       fh1=densigandh(A,B,g,hv(1));
8       fh2=densigandh(A,B,g,hv(2));
9       fh3=densigandh(A,B,g,hv(3));
10      x=linspace(-5,5,100);
11      plot(x,fg1,x,fg2,x,fg3);
12      plot(x,fh1,x,fh2,x,fh3)
13      fn=densigandh(A,B,g,h);
14      [y, fs] = gandhsim(A,B,g,h,100000 );
```

and below is the code gandhsim.m of the function gandhsim

```matlab
1    %generate n numbers coming from a g-and-h distribution(y)
2         % and compute the empirical density by ks filter(normal kernel)
3         %inputs: parameters: a,b,g,h; n:no. of values
4         function [y, f] = gandhsim(a,b,g,h,n )
5         rn=randn(1,n);
6         y=a+b*((exp(g*rn)-1).*exp(h*rn.^2/2))./g;
7         xi = linspace(-5,5,100);
8         f = ksdensity(y,xi,'function','pdf');
```

(2) losssincel.m returns the loss, the VaR and the CVaR for a single cell

```matlab
1     function [L, f, orvar, esh]=losssincel(A,B,g,h,lambda,alpha,n)
2     %monte carlo simualtion loss in a single cell,
3     %freq Poisson, sev g-and-h, n: no. of losses
4     %typical data in or: h=0.25; g=2; A=100000; B=1; lambda=100; alpha =
          95
5     N=poissrnd(lambda,1,n);
6     L=zeros(1,n);
7     for k=1:n
8     rn=randn(1,N(k));
9     y=A+B*((exp(g*rn)-1).*exp(h*rn.^2/2))./g;
10    L(k)=sum(y);
11    end
12    xx = linspace(0,max(L),300);
13    f = ksdensity(L,xx,'function','pdf');
```

```matlab
14      %fnor=pdf('Normal',xx,mean(L), std(L));

15      %plot(xx,f,xx,fnor)

16      LS=(L-mean(L))./std(L);

17      h=kstest(LS);

18      orvar=prctile(L,alpha);

19      normvar1=norminv(0.01*alpha,mean(L),std(L));

20      alf=linspace(95,99.5,100);

21   %comparing vars with poisson-gandh and normal distribution

22      pernorm=norminv(0.01*alf,mean(L),std(L));

23      perloss=prctile(L,alf);

24      iint=linspace((alpha./100),1,100);

25      perlossint=prctile(L,iint*100);%VaR on percentiles defined by iint

26      esh=trapz(iint,perlossint)./(1-alpha./100);%computing expected
            shortfall
```

(3) losssincelmix.m returns the loss, the VaR and the CVaR for a single cell mixing internal and external data

```matlab
1    function [L, f, orvar, esh]=losssincelmix(A1,B1,g1,h1,lambda,A2,B2,g2,
          h2,p,alpha,n)

2   %monte carlo simualtion loss in a single cell,

3   %freq mixed Poisson, sev two g-and-h, n: no. of losses

4   %p: probability of selecting the first of the severities

5   %typical data in or: h1=0.25; g1=2; A1=10000; B1=1; lambda=100; p=0.5;
          h2=0.3; g2=2; A2=20000; B2=1; lambda=100; p=0.5;alpha=0.95
```

```matlab
 6    L1=losssincel(A1,B1,g1,h1,p*lambda,alpha,n);

 7    L2=losssincel(A2,B2,g2,h2,(1-p)*lambda,alpha,n);

 8    L=L1+L2;

 9    xx = linspace(0,max(L),300);

10    f = ksdensity(L,xx,'function','pdf');

11    %fnor=pdf('Normal',xx,mean(L), std(L));

12    %plot(xx,f,xx,fnor)

13    %LS=(L-mean(L))./std(L);

14    %h=kstest(LS);

15    orvar=prctile(L,alpha*100);

16    normvar1=norminv(alpha,mean(L),std(L));

17    alf=linspace(95,99.5,100);

18    %comparing vars with poisson-gandh and normal distribution

19    pernorm=norminv(0.01*alf,mean(L),std(L));

20    perloss=prctile(L,alf);

21    iint=linspace(alpha,1,100); %Desire multiplied alpha by 100

22    perlossint=prctile(L,iint*100);%VaR on percentiles defined by iint %
          desire divided iint by 100

23    esh=trapz(iint,perlossint)./(1-alpha);%computing expected shortfall
```

Appendix B

TWO CELLS CASE

losstwocell.m computes the aggregate loss of two cell, then the VaR and CVaR

```matlab
1  function [LA, orvar, corvar]=losstwocell(lambda, AV,BV,gv,hv, theta,alpha,n)
2
3  %monte carlo simualtion aggregate loss in two cells, common losses are
4  %generated from a clayton copula
5  %inputs: freq Poisson 1x3 vector , sev g-and-h parameters AV,BV,gV
6  %hv, each a 1x4 vector, n: no. of simulated losses
7  % theta: param calyton copula
8  %theta: parameter clayton copula
9  %typical data in or: lambda=[200 100 100]; AV=[200000 200000 100000 100000];
        BV=[2 2 1 1]; gv=[ 3 3 2 2]; hv=[0.5 0.5 0.25 0.25]; theta=5; alpha=0.99;
10 theta0jm=[AV(1), BV(1), gv(1) hv(1)];
11 theta0jk=[AV(2), BV(2), gv(2) hv(2)];
12 thetajm=[AV(3), BV(3), gv(3) hv(3)];
13 thetajk=[AV(4), BV(4), gv(4) hv(4)];
14 N0=poissrnd(lambda(1),1,n);
15 N1=poissrnd(lambda(2),1,n);
16 N2=poissrnd(lambda(3),1,n);
17 LAjmp=zeros(1,n);
```

```matlab
18  LAjkp=zeros(1,n);
19  LA0jm=zeros(1,n);
20  LA0jk=zeros(1,n);
21  LAjk=zeros(1,n);
22  LAjm=zeros(1,n);
23  LA=zeros(1,n);
24  for k=1:n
25  [loss0jm, loss0jk]=losss0(theta0jm,theta0jk, theta,N0(k));
26  lossjm=losssgandh(thetajm,N1(k));
27  lossjk=losssgandh(thetajk,N2(k));
28  LAjmp(k)=sum(lossjm);
29  LAjkp(k)=sum(lossjk);
30  LA0jm(k)=sum(loss0jm);
31  LA0jk(k)=sum(loss0jk);
32  LAjm(k)=LAjmp(k)+LA0jm(k);
33  LAjk(k)=LAjkp(k)+LA0jk(k);
34  LA(k)=LAjm(k)+LAjk(k);
35  end
36  xx = linspace(0,max(LA),300);
37  f = ksdensity(LA,xx,'function','pdf');
38  %fnor=pdf('Normal',xx,mean(L), std(L));
39  %plot(xx,f,xx,fnor)
40  LS=(LA-mean(LA))./std(LA);
41  h=kstest(LS);
42  orvar=prctile(LA,100*alpha);
43  normvar1=norminv(alpha,mean(LA),std(LA));
```

83

```matlab
44  alf=linspace(95,99.5,100);

45  %comparing vars with poisson-gandh and normal distribution

46  pernorm=norminv(0.01*alf,mean(LA),std(LA));

47  perloss=prctile(LA,alf);

48  iint=linspace(alpha,1,100);

49  perlossint=prctile(LA,100*iint);%VaR on percentiles defined by iint

50  corvar=trapz(iint,perlossint)./(1-alpha);%computing expected shortfall
```

The functions loss0 and lossgandh

(1) loss0.m

```matlab
1   function [loss01, loss02]=losss0(theta01,theta02, theta,n)

2   %monte carlo simualtion  losses in two cells, common losses are

3   %generated from a clayton copula

4   %inputs: lambda, freq Poisson  , sev g-and-h parameters theta01 and

5   %theta02 in both cells

6   % n: no. of simulated losses

7   % theta: param calyton copula

8   %typical data in or:  AV=[10000  20000]; n = 500;  BV=[1 1]; gv=[3  2];
        hv=[0.5  0.25]; theta01=[AV(1),BV(1),gv(1),hv(1)];

9   %theta02=[AV(2),BV(2),gv(2),hv(2)]; theta=10;

10  u=copularnd('Clayton',theta,n);

11  z=norminv(u,0,1);

12  loss01p=theta01(1)+theta01(2)*((exp(theta01(3)*z(:,1))-1).*exp(theta01(4)
        *z(:,1).^2/2))./theta01(3);
```

84

```
13  loss02p=theta02(1)+theta02(2)*((exp(theta02(3)*z(:,2))-1).*exp(theta02(4)
        *z(:,2).^2/2))./theta02(3);
14  loss01=loss01p.*(loss01p>0);
15  loss02=loss02p.*(loss02p>0);
16  %plot(loss01,loss02,'.')
```

(2) lossgandh.m

```
1  function L=losssgandh(thetajm,n)
2  %monte carlo simualtion loss in a single cell,
3  %outpuit: n severity losses from a g-and-h distribution with param
        thetajm
4  %inputs: lambda,freq Poisson, thetajm sev g-and-h parameters, n: no. of
        losses
5  %typical data in or: h=0.25; g=2; A=100000; B=1; lambda=100
6  %thetajm=[A,B,g,h];
7  z=randn(1,n);
8  L=thetajm(1)+thetajm(2)*((exp(thetajm(3)*z)-1).*exp(thetajm(4)*z.^2/2))./
        thetajm(3);
```

REFERENCES

[1] Antoine Frachot, Pierre Georges, and Thierry Roncalli. Loss distribution approach for operational risk. March 2001. Working Paper, Crédit Lyonnais, Groupe de Recherche Opérationnelle.

[2] Antoine Frachot, Thierry Roncalli, and Eric Salomon. The Correlation Problem in Operational Risk. January 2004.

[3] Basel Committee on Banking Supervision, Enhancing Bank Transparency - Public disclosure and supervisory information that promote safety and soundness in banking systems; September 1998.

[4] Basel Committee on Banking Supervision. International Convergence of Capital Measurement and Capital Standards: A Revised Framework. Bank for International Settlements, Basel, June 2006.

[5] Basel Committee on Banking Supervision. Operational Risk Management. Bank for International Settlements, Basel, September 1998.

[6] Basel Committee on Banking Supervision. Operational Risk: Supporting Document to the New Basel Capital Accord. Bank for International Settlements, Basel,January 2001.

[7] Bjørn Sundt and William S. Jewell. Further results on recursive evaluation of compound distributions. Astin Bulletin, 12:27-39, 1981.

[8] Boris Iglewicz and Jorge Martinez. Some properties of the tukey g and h family of distributions. Communications in Statistics - Theory and Methods, 13(3):353-369, 1984.

[9] Bruno Dupire. Monte Carlo Methodologies and Applications for Pricing and Risk Management. Risk Books, London, 1998.

[10] Christian P. Robert and George Casella. Monte Carlo Statistical Methods. Springer Texts in Statistics. Springer-Verlag, Secaucus, NJ, USA, 2nd edition, August 2004.

[11] Christian P. Robert. Simulation of truncated normal variables. Statistics and Computing, 5(2):121-125, 1995.

[12] Daren B. H. Cline. Convolutions of distributions with exponential and subexponential tails. Journal of the Australian Mathematical Society. Series A, 43:347-365, 1987.

[13] Dirk P. Kroese, Thomas Taimre, and Zdravko I. Botev. Handbook of Monte Carlo methods. Wiley Series in Probability and Statistics. Wiley, Hoboken, NJ, USA, 2011.

[14] Dominik Kortschak and Hansjörg Albrecher. Asymptotic results for the sum of dependent non-identically distributed random variables. Methodology and Computing in Applied Probability, 11(3):279-306, 2008.

[15] Edward Omey and Eric Willekens. Second order behaviour of the tail of a subordinated

probability distribution. Stochastic Processes and their Applications, 21 (2):339-353, 1986.

[16] Eric Willekens. Asymptotic approximations of compound distributions and some applications. Bulletin de la Société Mathématique de Belgique. Sé rie B, 41(1): 55-61, 1989.

[17] Falko Aue and Michael Kalkbrener. LDA at work: Deutsche banks approach to quantifying operational risk. Journal of Operational Risk, 1(4):49-93, 2006.

[18] Gareth W. Peters, Adam M. Johansen, and Arnaud Doucet. Simulation of the annual loss distribution in operational risk via Panjer recursions and volterra integral equations for value-at-risk and expected shortfall estimation. Journal of Operational Risk, 2(3):29-58, 2007.

[19] Garold J. Borse. Numerical Methods with MATLAB: A Resource for Scientists and Engineers. International Thomson Publishing, 1st edition, 1996.

[20] George S. Fishman. Monte Carlo: Concepts, algorithms, and applications. Springer Series in Operations Research. Springer-Verlag, New York, NY, USA, 1996.

[21] Hansjørg Albrecher, Christian Hipp, and Dominik Kortschak. Higher-order expansions for compound distributions and ruin probabilities with subexponential claims. Scandinavian Actuarial Journal, 2010(2):105-135, 2010.

[22] Harry H. Panjer. Recursive evaluation of a family of compound distributions. Astin Bulletin, 12:22-26, 1981.

[23] Irving W. Burr. Cumulative frequency functions. The Annals of Mathematical Statistics, 13(2):215-232, 1942.

[24] Jean-Michel Marin, Kerrie Mengersen, and Christian P. Robert. Bayesian modeling and inference on mixtures of distributions. In Dipak K. Dey and C. R. Rao, editors, Handbook of Statistics, volume 25, chapter 16, pages 459-507. Springer, New York, NY, USA, 2005.

[25] Jim Pitman. Probability. Springer texts in statistics. Springer-Verlag, New York, NY, USA, 1993.

[26] John P. Robertson. The computation of aggregate loss distributions. Proceedings of the Casualty Actuarial Society, 79(1):57-133, 1992.

[27] Kabir Dutta and Jason Perry. A tale of tails: An empirical analysis of loss distribution models for estimating operational risk capital. Working Paper 06-13, Federal Reserve Bank of Boston, April 2007.

[28] Klaus Böcker and Claudia Klüppelberg. Operational VAR: a closed-form approximation. Risk, pages 90-93, December 2005. .

[29] Leonid Koralov and Yakov G. Sinai. Theory of Probability and Random Processes. Springer, 2nd edition, September 2007.

[30] M. C. Jones. The relationship between moments of truncated and original distributions plus some other simple structural properties of weighted distributions. Metrika, 37(3-4):233-243, 1990.

[31] Matthew Pritsker. Evaluating value at risk methodologies: Accuracy versus computational time. Journal of Financial Services Research, 12(2):201-242, 1997.

[32] Matthias Degen, Paul Embrechts, and Dominik D. Lambrigger. The quantitative modeling of operational risk: Between g-and-h and evt. Risk and Insurance, 37(2):1-32, 2007.

[33] Merran Evans, Nicholas Hastings, and Brian Peacock. Statistical Distributions. Wiley-Interscience, June 2000.

[34] Nicholas Metropolis and Stan Ulam. The Monte Carlo method. Journal of the American Statistical Association, 44(247):335-341, 1949.

[35] Luo, X.L., Shevchenko, P.V. and Donnelly J.B. (2007). Addressing the impact of data truncation and parameter uncertainty on operational risk estimates. The Journal of Operational Risk 2(4), 3-26.

[36] Paul Glasserman. Monte Carlo Methods in Financial Engineering, volume 53. Springer, 2004.

[37] Pavel V. Shevchenko. Implementing loss distribution approach for operational risk. Applied Stochastic Models in Business and Industry, 26(3):277-307, 2010.

[38] Pavel V. Shevchenko. Modelling Operational Risk Using Bayesian Inference, 2011.

[39] Philippe Barbe and William P. McCormick. Asymptotic expansions for infinite weighted

convolutions of heavy tail distributions and applications. Memoirs of the American Mathematical Society, 2004.

[40] Philippe Barbe and William P. McCormick. Asymptotic expansions for infinite weighted convolutions of rapidly varying subexponential distributions. Probability Theory and Related Fields, 141(1-2):155-180, 2008.

[41] R. Tyrrell Rockafellar and Stanislav Uryasev. Conditional value-at-risk for general loss distributions. Journal of Banking & Finance, 26(7):1443-1471, 2002.

[42] Rayner, G.D. and MacGillivray, H.L. Numerical maximum likelihood estimation for the g-and-k and generalized g-and-h distributions. Statistics and Computing, 2002, 12,1,57-75.

[43] Rhonda K. Kowalchuk and Todd C. Headrick. Simulating multivariate g-and-h distributions. The British journal of mathematical and statistical psychology, 63(1):63-74, 2010.

[44] Roger B. Nelsen. An Introduction to Copulas. Springer Series in Statistics. Springer, 2nd edition, January 2006.

[45] Rupert Brotherton-Ratcliffe. Monte carlo motoring. In Bruno Dupire, editor, Monte Carlo Methodologies and Applications for Pricing and Risk Management, pages 331-334. Risk Books, London, 1998.

[46] Søren Asmussen. Ruin probabilities. Advanced series on statistical science and applied probability, 2. World Scientific, Singapore, 2000.

[47] Sergey Foss, Dmitry Korshunov, and Stan Zachary. An Introduction to Heavy-Tailed and Subexponential Distributions, volume 38 of Springer Series in Operations Research and Financial Engineering. Springer, 2011.

[48] Sergey Foss, Dmitry Korshunov, and Stan Zachary. Convolutions of long-tailed and subexponential distributions. Journal of Applied Probability, 46(3):756-767, 2009.

[49] Sheldon M. Ross. Introduction to probability and statistics for engineers and scientists. Academic Press, 4th edition, 2009.

[50] Todd C. Headrick, Rhonda K. Kowalchuk, and Yanyan Sheng. Parametric probability densities and distribution functions for tukey g-and-h transformations and their use for fitting data. Applied Mathematical Sciences, 2(9):449-462, 2008.

[51] Umberto Cherubini, Elisa Luciano and Walter Vecchiato. Copula methods in finance.

[52] Wikipedia - Fractional Fourier transform.

[53] William Feller. An introduction to probability theory and its applications. Vol. II. Second edition. John Wiley & Sons Inc., New York, NY, USA, 1971.