UiO : Department of Mathematics
University of Oslo

Optimising Multivariate Reinsurance Contracts Under Uncertainty

Sigurd Fladby

Master's Thesis, Spring 2023

This master's thesis is submitted under the master's programme *Stochastic Modelling, Statistics and Risk Analysis*, with programme option *Finance, Insurance and Risk*, at the Department of Mathematics, University of Oslo. The scope of the thesis is 60 credits.

The front page depicts a section of the root system of the exceptional Lie group E_8, projected into the plane. Lie groups were invented by the Norwegian mathematician Sophus Lie (1842–1899) to express symmetries in differential equations and today they play a central role in various parts of mathematics.

Abstract

Abstract

Reinsurance is a contract between two parties in which financial risk is transferred from a cedent to a reinsurer for a fee, as a risk mitigation measure it is recognized by the solvency ii directive as a method for reducing SCR. The insurance layer contract, in which the reinsurer covers risk within an interval, has been shown to be the optimal form of reinsurance contract under the risk measure Value at risk. In previous research methods for finding the optimal solutions under various conditions has been characterized, as well as auxiliary results which has simplified the optimization problem. In this thesis the problem has been extended by introducing uncertainty to the risks. The optimization problem is largely the same, but the method for introducing uncertainty is new, and has required the use of an importance sampling scheme. We have seen that the optimal solution to the optimization problem can change when uncertainty is introduced and that simple summary statistics can indicate which kinds of solutions are optimal.

Acknowledgements

Thanks to everyone, and in particular to my supervisor.

Contents

Abstract **i**

Acknowledgements **ii**

Contents **iii**

List of Figures **iv**

List of Tables **iv**

1 Introduction **1**

2 Theory about relevant concepts **2**
2.1 Introduction 2
2.2 Useful definitions 2
2.3 Basic concepts in Non-life insurance 3
2.4 Reinsurance contracts 3
2.5 Risk measures 4
2.6 Optimization of portfolios 7
2.7 Methods used in our study 15
2.8 The algorithm 19

3 Numerical results **22**
3.1 Introduction 22
3.2 Introducing Uncertainty 22
3.3 Quantifying Uncertainty in Reinsurance Contract Optimization 24
3.4 Evaluating uncertainty 39
3.5 Further research 40

Appendices **41**

A The First Appendix - Probability distributions **42**

List of Figures

2.1 Each set represents different outcomes of the insurance layer contract for a bivariate problem. . . . 10

3.1 1e8 samples drawn from the lognormal distribution, with empiric mean 1.000 and empiric variance 2.251. We see realizations greater than or equal 200. . . . 23
3.2 Y_a . . . 25
3.3 Y_b . . . 26
3.4 Y_c . . . 26
3.5 Y_d . . . 27
3.6 Y_a . . . 27
3.7 Y_b . . . 28
3.8 Y_c . . . 28
3.9 Y_d . . . 29
3.10 Y_e . . . 30
3.11 Y_f . . . 31
3.12 Y_g . . . 31
3.13 Y_e . . . 32
3.14 Y_f . . . 32
3.15 Y_g . . . 33
3.16 Y_h . . . 34
3.17 Y_i . . . 35
3.18 Y_j . . . 35
3.19 Y_h . . . 36
3.20 Y_i . . . 36
3.21 Y_j . . . 37
3.22 Y_k. Note that the solution to this example is not symmetric, so the scale is not the same on the x- and y-axis. . . . 38

List of Tables

3.1 For each example a-j, the steps in the simulation study has been followed. X_{1-3} are auxiliary functions used to introduce uncertainty as discussed in the beginning of the chapter 24
3.2 For each distribution there has been generated 1e7 samples, the empiric mean and empiric variance has been calculated for these samples. 25
3.3 For each distribution there has been generated 1e7 samples, the empiric mean, empiric variance and empiric kurtosis has been calculated for these samples. 30
3.4 For each distribution there has been generated 1e8 samples, the empiric mean, empiric standard deviation and empiric kurtosis has been calculated for these samples. 34

CHAPTER 1

Introduction

Reinsurance is a method of risk diversification for insurance companies where one company cedes risk to another. There are many different types of contracts for reinsurance, some of which will be introduced later, but in all of them a *cedent* pays a reinsurer to take on some or all of the risk the cedent owns to begin with.
Bundling risks is often benficient to the cedent, but it isn't always feasible. In this thesis the goal is to investigate multivariate reinsurance contracts under uncertainty. This means that there are multiple risks that cannot be bundled together and needs to be reinsured separately.
We will first review existing methodology for reinsurance contracts without uncertainty in the univariate and multivariate cases, then develop methods for optimising reinsurance contracts under uncertainty in the multivariate case. At last we will develop methods for quantifying the effect of uncertainty on reinsurance optimisation.
In order to develop the aforementioned methods we will perform a simulation study, this will entail solving multiple problems with and without uncertainty and then finding the quantities that describes the changes in optimal solutions best.
The objective when optimizing reinsurance contracts is to minimize some *risk measure*. We will use value at risk, which is the preferred risk measure under solvency ii for capital requirements.

CHAPTER 2

Theory about relevant concepts

2.1 Introduction

In our problem we use numerical methods and theory of random variables to find how uncertainty influences optimal solutions of the reinsurance problem discussed in earlier articles. We start by clearly defining some very basic definitions and a short introduction to non-life insurance before we introduce reinsurance contracts and risk measures and see how they work together. This chapter will culminate in four sections on how the problem without uncertainty has been solved before. This will allow us to introduce the new element of uncertainty in the next chapter. Some of the results presented can be found in articles cited in the text and some of it is based on [Hus23].

2.2 Useful definitions

We start of with a couple of simple definitions. We will define random variables, cumulative distribution functions and probability density functions.

Definition 2.2.1. (Random variable) A random variable is a function $Y : \Omega \to \mathbf{R}$ for a probability space $(\Omega, \mathcal{F}, \mathbf{P})$ such that for all $x \in \mathbf{R}$, $\{\omega : x(\omega) \leq x\} \in \mathcal{F}$.

We assume that all our random variables are absolutely continuously distributed.

Definition 2.2.2. (Cumulative distribution function) The cumulative distribution function of a random variable X is the function F defined by

$$F(x) = P\{\omega \in \Omega : X(\omega) \leq x\}, x \in \mathbf{R}.$$

Definition 2.2.3. (Probability density function) The probability density function is the function f such that

$$F(x) = \int_{-\infty}^{x} f(y)dy, \quad -\infty < x < \infty$$

When the underlying random variable is absolutely continuously distributed this function exists.

Definition 2.2.4. (Survival function) We define the survival function $S_X(\alpha)$ as

$$S_X(\alpha) = P(X > \alpha) = 1 - F_X(\alpha)$$

2.3 Basic concepts in Non-life insurance

We continue with definitions of insurance contracts and related concepts which we need in order to understand what risk we are reinsuring.

Definition 2.3.1 (Insurance policy)**.** An insurance policy is a contract between a customer and an insurance company where the insurance company agrees to hold the financial risk of some adverse event happening to the customer in a defined time window.

Such a contract represents a risk to the insurance company, we let risks such as these be the random variables Z_i. The collection of these risks is also a random variable, which we will call X_j. We enumerate this random variable as well as because it is in the insurance company's interest to group these risks into different portfolios as they often have different characteristics and hence should be modeled differently. It is important for the insurance companies to know what kind of risk they hold so that they know what they need to charge for it, how much they need to hold in reserve and how much they need to reinsure. For clarity we define the relation between claim, premium and cost of the portfolio for the insurance company:

Definition 2.3.2 (Claim)**.** A claim occurs when the random variable $(Z_i > 0)$ we can also use Z_i to denote the payout on the policy (i).

Definition 2.3.3 (Premium)**.** A premium is what the customer pays the insurance company for the insurance contract. Generally this should be given by

$$\pi_i = (1 + \gamma)E[Z_i]$$

Where γ is the risk loading the insurance company charges to hold the risk of the type Z.

Definition 2.3.4 (Cost of portfolio)**.** We let

$$X_j = \sum_i^n Z_i$$

be the total payout of portfolio j and n be the number of policies in portfolio j.

The sum of the costs of portfolios, $\sum X_j$ is how much the insurance company has to pay out. The situation we are looking at in the next chapter is the case where we have two risks $Y_k = \sum X_j,\ k = 1, 2$, we are uncertain of the distribution of Y_k, and we want to reinsure each of the two risks separately. In this chapter we will be cover theory for an unspecified number of risks, but without the uncertainty.

2.4 Reinsurance contracts

An insurance contract is an agreement between a cedent and an insurer, where the insurer agrees to cover some risk the cedent is exposed to for a fee called a premium. These contracts can be between individuals and companies but insurance companies also need to disperse their own risk. This is done by reinsurance contracts, these often have a different structure but the concept of

ceding risk is the same.
In the article [Che+14] it is shown that insurance layer contracts are optimal when using Value at risk as risk measure. We will introduce Value at risk and discuss other risk measures later, but we will not introduce any other reinsurance contracts than insurance layer contracts and stop-loss contracts, where the latter is simply an insurance layer contract with the upper limit (b_i in this text) set to infinity. Following is a definition of an insurance layer contract. For risks X_i, contract parameters a_i, b_i we have that the retained risk for a cedent is given by $I_i(X_i)$ and the cost for the reinsurance company is given by $R_i(X_i)$:

$$R_i(X_i) = \begin{cases} 0, & \text{for } X_i < a_i \\ X_i - a_i, & \text{for } a_i \leq X_i \leq b_i \\ b_i - a_i, & \text{for } X_i > b_i \end{cases}$$

$$I_i(X_i) = \begin{cases} X_i, & \text{for } X_i < a_i \\ a_i, & \text{for } a_i \leq X_i \leq b_i \\ X_i - (b_i - a_i), & \text{for } X_i > b_i \end{cases}$$

The risk covered by the cedent is then I, but the cedent also have to pay a premium, which we assume is a function of the expectation of the cost for the reinsurer. We let $\pi_i = (1+\theta)E[R_i(X_i)]$ be the the price paid to the reinsurer as premium per risk i. Note that we have used a common loading factor θ, even if it may be natural to have different loading for different risks. We can now consider the total risk covered by the cedent:

$$\sum_{i=1}^{m} I_i(X_i) + (1+\theta)\sum_{i=1}^{m} E[R_i(X_i)] \tag{2.1}$$

This is a random variable, and we can use different risk measures to evaluate it.

2.5 Risk measures

Introduction to Risk Measures

We can use risk measures to assign a value to reinsurance contracts, generally risk measures are mappings from a set of random variables to the real numbers, this is a way to quantify risk. There are several different risk measures but we will start by defining a coherent risk measure and argue that Value at risk is the best option for our purpose, even if it isn't coherent. Coherency is a concept introduced in [Art+99], the article argues that risk measures should be constructed as follows:

Definition 2.5.1. For a set of random variables $\mathcal{L}$, a coherent risk measure $\rho : \mathcal{L} \rightarrow \mathbf{R} \bigcup \{\infty\}$ have the properties:

- Normalization: $\rho(0) = 0$. This means that an empty portfolio, i.e. one without any assets has a risk of zero.

- Monotonicity: If $Y_1 \leq Y_2$ a.s., then $\rho(Y_1) \leq \rho(Y_2)$. This means that if the portfolio Y_1 costs less than Y_2 almost surely then the risk given by the risk measure should reflect this.
- Translational invariance: For any $m \in \mathbf{R}$, $\rho(Y_1 + m) = \rho(Y_1) + m$
- Sub-additivity: $\rho(Z_1 + Z_2) \leq \rho(Z_1) + \rho(Z_2)$
- Positive homogeneity: If $\alpha \geq 0$, then $\rho(\alpha Z) = \alpha\rho(Z)$

This has later been refined to convex risk measures which has replaced the sub-additivity and positive homogeneity properties with the convexity property.

Definition 2.5.2. A risk measure ρ is a convex risk measure if

- Monotonicity: If $Y_1 \leq Y_2$, then $\rho(Y_1) \leq \rho(Y_2)$.
- Translational invariance: For any $m \in \mathrm{R}$, $\rho(Y_1 + m) = \rho(Y_1) + m$
- Normalization: $\rho(0) = 0$
- Law invariance: If $Y_1 \sim Y_2$ then $\rho(Y_1) = \rho(Y_2)$
- Convexity: $\rho(\lambda Y_1 + (1-\lambda)Y_2) \leq \lambda\rho(Y_1) + (1-\lambda)\rho(Y_2)$

There are a number of risk measures which are convex or coherent such as Conditional Tail expectation, Conditional Value at risk, Average Value at risk. For these risk measures it is shown in [Che+14] that stop-loss contracts are optimal reinsurance contracts in the optimization of univariate reinsurance contracts. We will, however, use the less constrained Value at Risk.

Value at risk

Value at risk is not a coherent risk measure as it lacks the sub-additivity property. This property ensures that the risk of multiple risks held together cannot be any worse than holding each risk separately. The sub-addivity problem and a concern that Value at risk does not appropriately consider the tails of the distribution has been points of contention for whether Value at risk is a good risk measure to use for capital requirements or not, and has been extensively discussed in particular after the financial crisis in 2008. The concensus now seems to be, however, that Value at risk is an invaluable tool and that the tail risk should be handled by risk managers. Either way, the solvency ii directive requiring the use of Value at risk as a risk measure makes it necessary to develop methods using Value at risk.

We will consider independent, non-negative, absolutely continuously distributed risks $X_1, X_2, ..., X_m$. Each of these risks will represent business lines or some other kind of bundled risk, and our goal is to minimize this risk measure by altering the contract parameters of the reinsurance contract.

Definition 2.5.3. For the survival function $S_{X_i} = P(X_i > x) = 1 - F_{X_i}(x)$. The α-level Value at risk V_α for a random variable X_i is defined as

$$V_\alpha(X_i) = S_{X_i}^{-1}(\alpha) = \inf\{x : P(X_i > x) \leq \alpha\}$$

It has the following properties:

- Monotonicity: If $Y_1 \leq Y_2$, then $\rho(Y_1) \leq \rho(Y_2)$.
- Translational invariance: For any $m \in \mathbf{R}$, $\rho(Y_1 + m) = \rho(Y_1) + m$
- Normalization: $\rho(0) = 0$
- Law invariant: If Y_1 and Y_2 have the same distribution under $\mathbf{P}$, then $\rho(Y_1) = \rho(Y_2)$.

Value at risk is not a coherent risk measure as it does not have the sub-additive property, this property as described in the definition of coherent risk measures ensures that the risk of two portfolios put together is not worse than adding the risk of the two separate portfolios. It has also been critized for not taking into account the tails of distribution, which it obviously does not, this can lead to

We include a theorem which says that the inverse image of the survival function has a unique solution, which is extensively used in the proof for the optimal solution.

Theorem 2.5.4. *If the survival function S is strictly decreasing, then $V_\alpha(X_i) = x_0$ if and only if:*

$$P(X_i > x_0) \leq \alpha \leq P(X_i \geq x_0)$$

and in particular if

$$P(X > r) = \alpha, \text{ then } S_{X_i}^{-1}(\alpha) = r$$

Proof. Assume first that:

$$V_\alpha(X_i) = inf\{x : P(X_i > x) \leq \alpha\} = x_0$$

If $P(X_i > x_0) = S_{X_i} > \alpha$ then it follows, since S_{X_i} is right-continuous, that there exists $\epsilon > 0$ such that $S_{X_i} > \alpha$ for all $x \in [x_0, x_0 + \epsilon]$. However, this implies that:

$$inf\{x : P(X_i > x) \leq \alpha\} \geq x_0 + \epsilon$$

which contradicts our assumption that $V_\alpha(X_i) = X_0$. If $P(X_i \geq x_0)\alpha$, then it follows, since $P(X_i \geq x_0)$ is left-continuous, that there exists $\epsilon > 0$ such that $P(X \geq x) < \alpha$ for all $x \in [x_0 - \epsilon, x_0]$. However, this implies that $P(X_i > x_0 - \epsilon \leq P(X_i \geq x_0 - \epsilon) < \alpha$, and hence:

$$inf\{x : P(X_i > x) \leq \alpha\} \leq x_0 - \epsilon$$

which contradicts the assumption that $V_\alpha = X_0$. Thus, we have shown that:

$$V_\alpha(X_i) = x_0 \implies P(X_i > x_0) \leq \alpha \leq P(X_i \geq x_0).$$

In order to prove the converse implication, we assume that:

$$P(X_i > x_0) \leq \alpha \leq P(X_i \geq x_0)$$

Since $P(X_i > x_0) \leq \alpha$, it follows that $x_0 \in \{x : P(X_i > x) \leq \alpha\}$, and hence:

$$S_{X_i}^{-1}(\alpha) = inf\{x : P(X_i > x) \leq \alpha\} \leq x_0$$

Assume then that $S_{X_i}^{-1}(\alpha) = x_1 < x_0$. This implies that:

$$S_{X_i}(x) = P(X_i > x) \leq \alpha, \text{ for all } x \in (x_1, x_0).$$

Since we have assumed that S_{X_i} is strictly decreasing, this implies that there exists $x_2\ in(x_1, x_0)$ such that $P(X_i > x_2) < \alpha$, and hence:

$$P(X_i \geq x_0) \leq P(X_i > x_2) < \alpha,$$

which contradicts the assumption that $P(X_i \geq x_0) \geq \alpha$. We thus conclude that

$$P(X_i > x_0) \leq \alpha \leq P(X_i \geq x_0) \implies V_\alpha(X) = x_0$$

we still have to prove the last part:
Assume that $P(X_i > x_0) = \alpha$, then obviously $P(X_i > x_0) \leq \alpha$. We see that

$$P(X_i \geq x_0) \leq P(X_i > x_0) = \alpha$$

And again, since we assumed that S_{X_i} is strictly decreasing, it follows that

$$V_\alpha(X_i) = x_0.$$

■

2.6 Optimization of portfolios

In this section we will properly introduce the optimisation problem. We will first review how the univariate optimisation problem can be solved with an example, and and then turn to the multivariate problem. The multivariate problem has some extra challenges and we will review the necessary theory to solve it in more detail than for the univariate problem. In order to solve the multivariate problem it is also necessary to use numerical methods for most distributions and we will go through some of the strategies used in the method used.

Univariate optimization problem

The simplest example is the univariate example, we use it to ease into the theory. Our objective in this subsection will be to introduce a univariate optimization problem and solve it. We put ourselves into the position of the cedent with a portfolio, X, which is exponentially distributed. Underneath is a list of assumed properties:

- $X \sim Exponential(\frac{1}{50})$
- $\theta = 0.2$
- $\alpha = 0.01$

The problem is formulated as

$$min_{a,b} V_\alpha(I(X) + (1+\theta)E[R(X)])$$

In [Che+14] it is shown that the optimal value for a in this problem, when using Value at risk as a risk measure, is given by

$$\frac{\partial V_\alpha}{\partial a} = 1 - (1+\theta)S_X(a),$$

this implies that the optimal value for a is

$$a = S_X^{-1}(\frac{1}{1+\theta}) = -50 \cdot ln\left(\frac{1}{1.2}\right) = 9.12$$

Next we need to find b so that $P(X > b) = \alpha$. This is given by

$$\begin{aligned} \alpha &= 1 - (1 - e^{-\lambda b}) \\ ln(\alpha) &= -b\lambda \\ -50ln(0.01) &= b = 230.26 \end{aligned}$$

The resulting value at risk is then

$$V_\alpha(I(X) + (1+\theta)E[R(X)] = a + (1+\theta)E[R(X)] = 9.12 + 1.2 \cdot (50 - 9.12) = 58.18$$

whereas the α-level value at risk for the uninsured portfolio is

$$V_{0.01}(X) = inf\{x : P(X > x) \leq 0.01\} = 230.26$$

This shows us that reinsurance contracts can effectively be used to reduce the risk a company holds. This example is made extra simple by the closed form of the survival function of the exponential distribution. In the next chapter where we create new distributions with uncertainty, this is not an option, and we will have to use numerical methods to estimate the survival functions.

The multivariate problem

The total risk covered 2.1 consists of two terms, a premium term and a retained risk term and we want to minimize each of them. If we apply the α-level Value at risk we get

$$V_\alpha\left(\sum_{i=1}^m I_i(X_i) + (1+\theta)\sum_{i=1}^m E[R_i(X_i)]\right) = S^{-1}_{\sum_{i=1}^m I_i(X_i)}(\alpha) + (1+\theta)\sum_{i=1}^m E[R_i(X_i)] \tag{2.2}$$

We let $f_{X_i}(x)$ be the probability density function of the risk X_i. The premium term for each risk i is then given by

$$\begin{aligned} E[R_i(X_i)] &= \int_{a_i}^{b_i} (x_i - a_i) f_{X_i} dx_i + (b_i - a_i)P(X_i > b_i) \\ &= \int_{a_i}^{b_i} x f_{X_i} dx - \int_{a_i}^{b_i} a_i f_{X_i} dx + (b_i - a_i)P(X_i > b_i) \\ &= \int_{a_i}^{b_i} x f_{X_i} dx - a_i P(a_i < X_i \leq b_i) + (b_i - a_i)P(X_i > b_i) \end{aligned}$$

$$= \int_{a_i}^{b_i} x f_{X_i} dx + b_i P(X_i > b_i) - a_i P(X_i > a_i)$$

we are interested in the partial derivatives of the premium term with respect to a_i and b_i. We also introduce the notation ϕ_i to denote the expected reinsurance cost of risk i.

$$\frac{\partial}{\partial a_i}(1+\theta)\sum_{i=1}^{m}\phi_i = \frac{\partial}{\partial a_i}(1+\theta)\left(\sum_{i=1}^{m}\int_{a_i}^{b_i} x f_{X_i} dx + b_i P(X_i > b_i) - a_i P(X_i > a_i)\right)$$
$$= -(1+\theta)P(X_i > a_i)$$

and similarly for b_i:

$$\frac{\partial}{\partial b_i}(1+\theta)\sum_{i=1}^{m}\phi_i = \frac{\partial}{\partial b_i}(1+\theta)P(X_i > b_i)$$

We see that partial derivatives of the premium term only depends on the contract parameter we are differentiating with respect to, which is very convenient.

The retained risk term is less convenient to optimize. We start by defining the following sets given contract parameters $a_1, b_1, ..., a_m, b_m$:

$$\mathcal{A} = \{(X_1, ..., X_m) : \sum_{i=1}^{m} I_i(X_i) < \sum_{i=1}^{m} a_i\},$$
$$\mathcal{B} = \{(X_1, ..., X_m) : \sum_{i=1}^{m} I_i(X_i) = \sum_{i=1}^{m} a_i\},$$
$$\mathcal{C} = \{(X_1, ..., X_m) : \sum_{i=1}^{m} I_i(X_i) > \sum_{i=1}^{m} a_i\}.$$

Now we will include some results from [HC20], these allow us to find a_i with relative ease. The fourth theorem, from [Hus22], also simplifies the process of finding the b_i-values.

Theorem 2.6.1. *Assume that the contract parameters $a_1, b_1, ..., a_m, b_m$ are chosen so that:*

$$P((X_1, X_2, ..., X_m) \in \mathcal{B} \cup \mathcal{C}) \geq \alpha$$
$$P((X_1, X_2, ..., X_m) \in \mathcal{C}) \leq \alpha$$

Then we have:

$$V_\alpha\left(S^{-1}_{\sum_{i=1}^{m} I_i(X_i)}(\alpha)\right) = \sum_{i=1}^{m} a_i$$

Proof. The result is immediate from the definition of Value at risk and by the monotonicity of the survival function.

$$P(\sum_{i=1}^{m} I_i(X_i) \geq \sum_{i=1}^{m} a_i) = P((X_1, X_2, ..., X_m) \in \mathcal{B} \cup \mathcal{C}) \geq \alpha$$

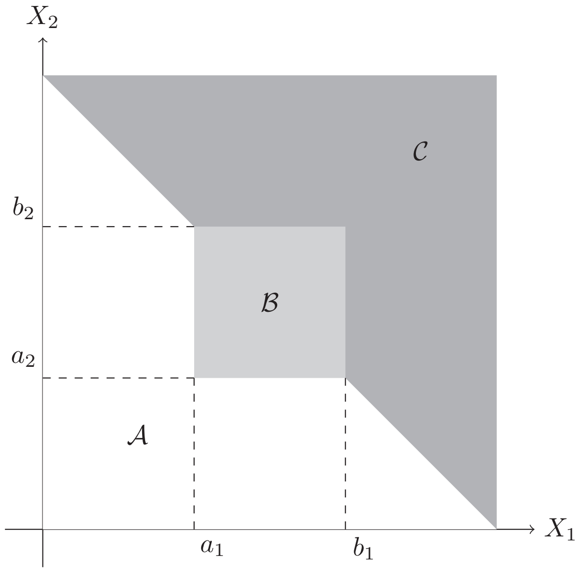


Figure 2.1: Each set represents different outcomes of the insurance layer contract for a bivariate problem.

and

$$P(\sum_{i=1}^{m} I_i(X_i) > \sum_{i=1}^{m} a_i) = P((X_1, X_2, ..., X_m) \in \mathcal{C}) \leq \alpha$$

■

Theorem 2.6.2. *If we assume that the contract parameters* $a_1, b_1, ..., a_m, b_m$ *are chosen so that*

$$P((X_1, X_2, ..., X_m) \in \mathcal{C} = \alpha$$

Then

$$S^{-1}_{\sum_{i=1}^{m} I_i(X_i(\alpha) = \sum_{i=1}^{m} a_i}$$

Proof.

$$P(\sum_{i=1}^{m} I_i(X_i) > \sum_{i=1}^{m} a_i) = P(X_1, X_2, ..., X_m) \in \mathcal{C}) = \alpha$$

■

Theorem 2.6.3. *Assume that* $a_1, b_1, a_2, b_2, ..., a_m, b_m$ *are optimal contract parameter values and that*

$$(1+\theta)^{-m} \geq \alpha \tag{2.3}$$

Then the following must hold true:

$$a_i = S^{-1}_{X_i}\left(\frac{1}{1+\theta}\right), \quad i = 1, ..., m \tag{2.4}$$

and:

$$P(\sum_{i=1}^{m} I_i(X_i) > \sum_{i=1}^{m} a_i) = P((X_1, X_2, ..., X_m) \in \mathcal{C}) = \alpha \tag{2.5}$$

Proof. Let $a_1, a_2, ..., a_m$ be chosen so that

$$P\left(\bigcap_{i=1}^{m} X_i > a_i\right) \geq \alpha,$$

then $b_1, b_2, ..., b_m$ should be chosen so that 2.5 holds. We can see that such values exist and are relevant because if $b_i = a_i$ for all i then

$$\sum_{i=1}^{m} I_i(X_i) = \sum_{i_i}^{m} X_i,$$

and

$$P\left(\sum_{i_i}^{m} I_i(X_i) > \sum_{i_i}^{m} a_i)\right) = P\left(\sum_{i_i}^{m} X_i > \sum_{i_i}^{m} a_i\right) > P\left(\bigcap_{i=1}^{m} X_i > a_i\right) \geq \alpha,$$

On the other hand, if $b_i = \infty$ for all i then $I_i(X_i)$ becomes a stop-loss contract instead and

$$P\left(\sum_{i_i}^{m} I_i(X_i) > \sum_{i_i}^{m} a_i)\right) = 0$$

Since $X_1, X_2, ..., X_m$ are assumed to be absolutely continuously distributed we can conclude that there must exist values $b_i > a_i$ such that (2.5) holds. Thus Theorem 2.6.2 implies that

$$S^{-1}_{\sum_{i=1}^{m} I_i(X_i)}(\alpha) = \sum_{i=1}^{m} a_i$$

Now, if we increase the b_i's, we still have :

$$P(\sum_{i=1}^{m} I_i(X_i) \geq \sum_{i=1}^{m} a_i) > P\left(\bigcap_{i=1}^{m} X_i > a_i\right) \geq \alpha$$

While at the same time:

$$P\left(\sum_{i=1}^{m} I_i(X_i) > \sum_{i=1}^{m} a_i)\right) \leq \alpha$$

hence by Theorem 2.6.2 the retained risk term remains the same while the premium term increases. This makes the value of V_α increase. On the other hand, if we decrease the values of the b_i's the retained risk term increases and the premium decreases. This means that for any $a_1, a_2, ..., a_m$ we can find $b_1, b_2, ..., b_m$ such that 2.5 holds, and it will be optimal with respect to the a_i's.

Next we can find optimal a_i's conditioned on the b_i's by adding the derivatives with respect to a_i.

$$\frac{\partial}{\partial a_i} S^{-1}_{\sum_{i=1}^m I_i(X_i)}(\alpha) = 1$$

and hence the derivative V_α is

$$\frac{\partial}{\partial a_i} V_\alpha = 1 - (1+\theta)P(X_i > a_i), \quad i = 1, 2, ..., m.$$

and if we solve for a_i we get

$$\frac{\partial}{\partial a_i} V_\alpha = 1 - (1+\theta)P(X_i > a_i) = 0, \quad i = 1, 2, ..., m.$$
$$P(X_i > a_i) = (1+\theta)^{-1}, \quad i = 1, 2, ..., m.$$

Thus, by the assumption $(1+\theta)^{-m} \geq \alpha$ and that the X_i's are independent it follows that

$$P(\bigcap_{i=1}^{m} X_i > a_i) = (1+\theta)^{-m} \geq \alpha$$

■

From this we can conclude that $b_1, b_2, ..., b_m$ exists and that for any set of a_i's we can find optimal b_i's. Lastly we include the theorem which allows us to find these optimal b_i's in an efficient manner.

Theorem 2.6.4. *Assume that $a_1^*, ..., a_m^*$ given by* (2.4) *satisifies* (2.3)*. Then the remaining optimal contract parameters $b_1^*, ..., b_m^*$ can be found by solving the following optimization problem:*

$$\textit{Minimize: } \sum_{i=1}^{m} E[R_i(X_i)]$$
$$\textit{subject to: } P((X_1, ..., X_m) \in \mathcal{C}) = \alpha$$

with respect to $b_1, ..., b_m$

Proof. We start out by noting that the constraint in the optimization problem is the same as (2.5), which is used in 2.6.3, so the constraint is well-justified. It also justfies writing the constraint as

$$P\left(\sum_{i=1}^{m} I_i(X_i) > \sum_{i=1}^{m} a_i^*\right) = \alpha$$

Then, by 2.5.4 it follows that under the constraint the retained risk term is given by:

$$S_{\sum_{i=1}^m I(X_i)}(\alpha) = \sum_{i=1}^{m} a_i^*.$$

Hence, the resulting total α-level Value at risk becomes:

$$V_\alpha = \sum_{i=1}^{m} a_i^* + (1+\theta)\sum_{i=1}^{m} E[R_i(X_i)]$$

From this it follows that minimizing V_α is equivalent to minimizing $\sum_{i=1}^{m} E[R_i(X_i)]$ subject to the constraint in this theorem with respect to $b_1, ..., b_m$. ■

Finding the optimal solution

We see that determining the a_i's is quite straight forward, but that the lack of constraints on the b_i's yields infinitely many possible solutions when $m > 1$, so we need an algorithm to find the optimal set of b_i's. We will mainly deal with the situation when $m = 2$, as our focus is not increasing m, but rather adding uncertainty to the parameters of the distributions. However, we start by considering the balanced solution for a general m.

Balanced case for general m

Let A be the common value of $P(X_i > a_i)$ such that $A = S_{X_i}(a_i), i = 1, ..., m$. We let $B = S_{X_i}(b_i), i = 1, ..., m$ such that $P(X_i > b_i)$ has the same value for all i. The resulting values of the contract parameters are then given by:

$$b_i = S_{X_i}^{-1}(B),$$

It is inconvenient to solve this problem analytically, instead we use Monte Carlo simulation on the joint distribution of $X_1, ..., X_m$. Based on this sample we can iterate the value of B until the subset C contatins the desired fraction of simulations. The iteration is done as follows: Since $X_1, ..., X_m$ is assumed to be absolutely continuously distributed it follows that the probability

$$P\left(\sum_{i=1}^{m} I_i(X_i) > \sum_{i=1}^{m} a_i\right)$$

is a continuous and increasing function of B. Thus we can find a value $B = B_l$ such that this value is less than α, as an lower bound and equivalently for an upper bound $B = B_u$. In order to find these bounds we consider the figure 2.1 with the sets and note that

$$C \subseteq \bigcup_{i=1}^{m} (X_i > b_i)$$

hence, we have:

$$P\left(\sum_{i=1}^{m} I_i(X_i) > \sum_{i=1}^{m} a_i\right) = P\left[(X_1, ..., X_m) \in \mathcal{C}\right]$$
$$\leq P\left[\bigcup_{i=1,...,m} (X_i > b_i)\right] = 1 - P\left[\bigcap_{i=1,...,m} (X_i \leq b_i)\right] = 1 - (1-B)^m$$

If $B \in [0,1]$ is such that $1-(1-B)^m = \alpha$, it follows from the previous equation that

$$P\left(\sum_{i=1}^{m} I_i(X_i) > \sum_{i=1}^{m} a_i\right) \leq \alpha$$

Thus this value is smaller than the correct B-value. We let $B_l = 1 - \sqrt[m]{1-\alpha}$. For the upper bound we consider the set

$$\left[\bigcap_{i=1,...,m} (X_i > a_i)\right] \backslash \mathcal{B} \subseteq \mathcal{C}$$

hence, we have

$$P\left(\sum_{i=1}^{m} I_i(X_i) > \sum_{i=1}^{m} a_i\right) = P[(X_1,...,X_m) \in \mathcal{C}] \geq P\left(\left[\bigcap_{i=1,...,m} (X_i > a_i)\right] \backslash \mathcal{B}\right)$$

$$= P\left(\bigcap_{i=1,...,m} (X_i > a_i)\right) - P\left(\bigcap_{i=1,...,m} (a_i \leq X_i \leq b_i)\right) = A^m - (A-B)^m$$

Similarly to the lower bound, if $B \in [0,1]$ is such that $A^m - (A-B)^m = \alpha$ then

$$P\left(\sum_{i=1}^{m} I_i(X_i) > \sum_{i=1}^{m} a_i\right) \geq \alpha$$

and we denote this value by $B_u = A - \sqrt[m]{A^m - \alpha}$. Consequently we can deduce that the true value of B is in between B_u and B_l, we can thus use the bisection method to find the true value.

Unbalanced case for m = 2

For the unbalanced case we restrict ourselves to m = 2 which is the case we will concern ourselves with in the next chapter.Our first goal is to find the relevant interval for the value of the b_i's. We let A be the common value of $P(X_i > a_i)$ such that $A = S_{X_i}(a_i), i = 1,...,m$, we denote $B_i = S_{X_i}(b_i), i = 1,2$. Due to the nature of the reinsurance contracts B_i must be in the interval $[0, 1-A]$, which is a fine starting point for finding the best value, but we can improve the upper border.

If we again consult figure 2.1, we see that if $B_1 = 0$, then B_2 must have the corresponding highest possible value. This value is given by the following equation:

$$P(I(X_2), I(X_1) \in C | b_1 = S_{X_1}^{-1}(0)) = \alpha,$$

where a_1, a_2, b_1 is given. The algorithm for this process is described in the algorithm section. After this we create an array of evenly spaced values and find B_1- and B_2-values which yield $P(C) = \alpha$ in the same manner. For each of these values we can find the corresponding b_i-values, and then the corresponding total risk, and we designate the lowest total risk as the optimal solution.

2.7 Methods used in our study

In this section we define methods which we have used in our study to characterize the optimal solutions.

Border and inner point solutions

In the previous section we saw that there were different types of solutions, and we are interested in characterizing them more formally. In order to do this we introduce the concept of balanced solutions, unbalanced solutions, border solutions and inner point solutions. We start by introducing B_i, which is given by $B_i = F_{X_i}^{-1}(b_i)$. It is easy to see that the possible values for B_i are $[0, A_i]$, where $A_i = F_{X_i}^{-1}$. In no examples do we see $B_i = A_i$, this is because it is irreconcilable with our distributions and choice of α-values. $B_i = 0$ is a more realistic solution, it is the case when we get a stop-loss contract as the optimal solution. This is what we will refer to as a border solution.

Definition 2.7.1. (Border solution). A solution to the optimization problem where $B_i = 0$, for at least one i.

An inner point solution is the opposite of a border solution

Definition 2.7.2. (Inner point solution). A solution where $B_i \neq 0$, for all i.

We have not seen any solutions where $B_1 = 0, B_2 = 0$, this is not feasible when a_i is chosen according to the theory, and the cost of reinsurance as it is in our examples. The types of solution we see in our simulation study are instead on one out of two forms:

1. one stop loss contract and one insurance layer contract with $B_i > 0$.
2. both contracts with $B_i > 0$.

In the next section on hazard rates there is a discussion on which kinds of problems leads to which kinds of solution.
The concept of unbalanced and balanced solutions is another approach to classifying the solutions, which we use less in this study as all our solutions are unbalanced.

Definition 2.7.3. (Balanced solution) We say that a solution is balanced if $b_1 = b_2$.

The balanced solution is much easier to find computationally, as it reduces to number of potential solutions. In the bivariate case this is not necessary, and the concept of balanced solutions is not that interesting. Our examples are often close to being balanced, but not quite balanced.

Hazard rate and the tails of distributions

While the speed of the main algorithm is reasonable it is convenient to quickly check properties of distributions we consider using, and additionally the hazard rate plots are also informative in their own right. In this section we look at a method for determining what kind of solution a distribution will have. We start by defining the hazard rate.

Definition 2.7.4. (Hazard rate) Let $y(x)$ be the probability density function of a random variable Y, and $S(x)$ the survival function of the same random variable. Then the hazard function is defined as

$$h(x) = \frac{y(x)}{S(x)},$$

for $S(x) > 0$.

The hazard rate yields, for each $x \in [0, \infty)$ the probability mass function of the discretized approximation of y(x) divided by the probability of incurring a claim that is greater than or equal to x.

In order to explain why we can use the hazard rate to say something about the solution of a problem, we need the notion of super- and sublevel sets and quasiconvexity and quasiconcavity. First, we remind the reader of the updated objective function found in 2.6.4.

$$\phi = \sum_{i=1}^{m} \phi_i = \sum_{i=1}^{m} E[R_i(X_i)] = \sum_{i}^{m} \int_{a_i}^{b_i} x \cdot f_{X_i}(x)dx + b_i P(X_i > b_i) - a_i P(X_i > a_i)$$

We will look at the objective function in terms of $B = (B_1, B_2, ..., B_m)$ and define the super- and sublevel sets from this.

Definition 2.7.5. The superlevel sets are the sets

$$L_c^+(\boldsymbol{\phi}) = \{B \in [0, 1]^m : \phi(B) \geq c\}$$

and the sublevel sets are the sets

$$L_c^-(\boldsymbol{\phi}) = \{B \in [0, 1]^m : \phi(B) \leq c\}$$

meaning that the superlevel set of ϕ is a set of values of $B \in [0, 1]^m$ which yields the same or higher value of expected reinsurance cost as $\boldsymbol{\phi}(c)$. We say that ϕ is quasiconvex if all its sublevel sets are convex and $\boldsymbol{\phi}$ is quasiconcave if all its superlevel sets are convex. These sets can be plotted and form iso-contours. Next, we see what kind of conclusions we can draw from these contours.

Theorem 2.7.6.

- *If $\phi_1(B_1), ..., \phi_m(B_m)$ are convex functions, then ϕ is a quasiconvex function of B.*
- *If $\phi_1(B_1), ..., \phi_m(B_m)$ are concave functions, then ϕ is a quasiconcave function of B.*

We prove the first part of the theorem, the second is proved in the same manner:

Proof. We assume that $\phi_1(B_1), ..., \phi_m(B_m)$ are convex function, and we let $\mathbf{B}^{(j)=(B_1^{(j)},...,B_m^{(j)}}) \in L_c^-(\boldsymbol{\phi}), j = 1, 2$. In order to show that $L_c^-(\boldsymbol{\phi})$ is convex we must show that for any $\lambda \in [0, 1]$, we also have that $\mathbf{B} = \lambda \mathbf{B}^{(1)} + (1 + \lambda)\mathbf{B}^{(2)} \in L_c^-(\boldsymbol{\phi})$.
Since $\phi_1, ..., \phi_m$ are convex, we know that for $i = 1, ..., m$, we have

$$\phi_i(\lambda B_i^{(1)} + (1 + \lambda)B_i^{(2)}) \leq \lambda \phi_i(B_i^{(1)}) + (1 - \lambda)B_i^{(2)}$$

Hence we get that:

$$\begin{aligned}\boldsymbol{\phi}(\mathbf{B}) &= \boldsymbol{\phi}(\lambda\mathbf{B}^{(1)} + (1-\lambda)B_i^{(2)}) \\ &= \sum_{i=1}^{m} \phi_i(\lambda B_i^{(1)} + (1-\lambda)B_i^{(2)}) \\ &\leq \sum_{i=1}^{m} \lambda\phi_i(B_i^{(1)}) + (1-\lambda)\phi_i(B_i^{(2)}) \\ &= \lambda\boldsymbol{\phi}(\mathbf{B}^{(1)}) + (1-\lambda)\boldsymbol{\phi}(\mathbf{B}_i^{(2)}) \\ &\leq \lambda c + (1-\lambda)c = c.\end{aligned}$$

■

If we now look at the derivatives of the components of the objective function $\phi_1, \phi_2, ..., \phi_m$ we will end up with the definition of the hazard rate. We first find the derivative with respect to b_i:

$$\begin{aligned}\frac{\partial\phi_i}{\partial b_i} &= \frac{\int_{a_i}^{b_i} x \cdot f_{X_i}(x)dx + b_i P(X_i > b_i) - a_i P(X_i > a_i)}{\partial b_i} \\ &= P(X_i > b_i)\end{aligned}$$

The derivative with respect to B_i is then given by:

$$\frac{\partial\phi_i}{\partial B_i} = \frac{\partial\phi_i}{\partial b_i}\frac{\partial b_i}{\partial B_i} = -\frac{B_i}{f_{X_i}(S_{X_i}^{-1}(B_i))}, \quad i = 1, ..., m$$

we thus see that ϕ_i is convex if

$$-\frac{B_i}{f_{X_i}(S_{X_i}^{-1}(B_i))}$$

is increasing in B_i. This is equivalent to

$$\frac{f_{X_i}(S_{X_i}^{-1}(B_i))}{B_i}$$

increasing in B_i. We may now substitute $B_i = S_{X_i}(x)$, and we are left with the expression

$$\frac{f_{X_i}(x)}{S_{X_i}(x)},$$

which is the hazard rate presented at the beginning of the section. Since B_i is a decreasing function of x, ϕ_i is now instead *convex* if it is *decreasing*. Similarly we have that ϕ_i is *concave* if the hazard rate is *increasing*. We can conclude with the following theorem:

Theorem 2.7.7.

- *If the hazard rate of the distribuion X is decreasing in B_i, then phi_i is convex.*
- *If the hazard rate of X is increasing in B_i, then ϕ_i is concave.*

Iso-curves

An Iso-curve plot is a tool used to describe how changes in distribution of resources changes the potential output. The idea being that there is a limited amount of resources and it can be distributed between two or more alternatives. We use Iso-curves for our bivariate problem to illustrate what the optimal solution looks like.

We are, for every iso-curve, letting the expected reinsurance cost, $\sum_{i=1}^{m} E[R_i(X_i)]$, be the same and plotting the different B_1, B_2-values that yields this value. The constraint is plotted in the same picture and provides the border for feasible solutions. The iso-curve with the lowest $\sum_{i=1}^{m} E[R_i(X_i)]$ that also intersects with the constraint is the lowest feasible expected reinsurance cost, and the optimal solution is characterized by the B_i-values which gives this solution.

Given that the constraint will look the same for all our problems the iso-curves are more interesting, as they can change from distribution to distribution, or if we introduce uncertainty. The shape of the iso-curve is one way of characterizing whether a solution will be on the border of the solution space or an inner point.

Kurtosis

In chapter 3 we will consider methods of quantifying the uncertainty added with our mixture distributions. In this thesis this has been done by looking at summary statistics of samples drawn from the distributions and comparing them. We have primarily used standard deviation and kurtosis, and kurtosis requires a short introduction. Kurtosis is the fourth standardized moment, and is defined as

$$Kurt[X] = E\left[\left(\frac{X-\mu}{\sigma}\right)^4\right]$$

We have used the Fisher definition of kurtosis which is the standard in the python pandas package, the only difference being that it subtracts 3 from the standard kurtosis estimate above.

$$Kurt[X] = E\left[\left(\frac{X-\mu}{\sigma}\right)^4\right] - 3$$

The purpose of this estimate compared to the other is to give the normal distribution a kurtosis of 0. One of our examples has approximately the same tail as a normal distribution, and this is reflected in a slightly negative value which we will see in the next section, however this does not in any interesting way influence the way we use the kurtosis.

The kurtosis estimates used in the next chapter have been implemented by creating large samples from the distributions we are interested in and then using the mentioned python pandas function. The estimation of kurtosis on samples is, as we will see, somewhat unstable as one outlier in the fourth power will have a large influence on the estimate. We will, however, see some systematic changes in the results which seem reasonable.

2.8 The algorithm

Now that we know why the algorithm works we will go through how the algorithm is implemented in practice. This section will consist of three algorithms and short explanations of them.

Algorithm 1 Main algorithm

```
1: Set variables: seed, number of simulations, number of integration points,
   number of iterations, epsilon, alpha, gamma, theta, Del, Distribution 1 and
   Distribution 2
2: Calculate the integrals of the distributions at the integration points.
3: Find optimal a_1, a_2
4: Calculate expected reinsurance cost for b_i-values from a_i to the highest
   percentile of the distribution.
5: Determine viable combinations of b_1 and b_2 given α
6: Calculate iso-curves
7: Calculate constraint
8: Determine optimal b_1, b_2 given objective function
```

where epsilon is the slack in the bisection method and Del is the size of the importance sampling region. Step 2 is done by numerical integration at each integration interval, for each interval the middle value as been chosen. Expected reinsurance cost is used in the value at risk calculation, and is calculated for all relevant values of a_i and b_i. In Step 5 and our importance sampling scheme is used, this is discussed in more detail in its own section. Step 6, 7 and 8 is done according to the theory in the previous section.

Determine viable combinations b_1, b_2

As mentioned in the unbalanced section, there is an algorithm for obtaining the other b_i-value given that the first is known.

Algorithm 2 viable combinations of b_1, b_2

```
Determine the max value of B1 and B2
Define vector BB1. BB1 should be evenly spaced values for B1 from 0 to
the max determined in previous step.
Define BB2 as an empty vector of the same size as BB1.
for i in range(BB2) do
    BB2[i] is given by the B2 value which yields P(C) = α, we use our
    importance sampling scheme as well as a bisection method
end for
```

Importance sampling

Importance sampling is a Monte Carlo technique that estimates attributes from one distribution by drawing samples from another distribution and adjusting the results by a specific factor. In our study, we employ importance sampling to reduce the number of simulations required for estimating the probability of landing in set C (as described in 2.1) given different contract parameters a_i, b_i.

To achieve this, we generate two independent bivariate uniform variables, constrained to the importance sample region $D = ([1-\Delta, 1] \times [0,1]) \cup ([0,1] \times [1-\Delta, 1])$, using the following algorithm:

Algorithm 3 Importance sampling

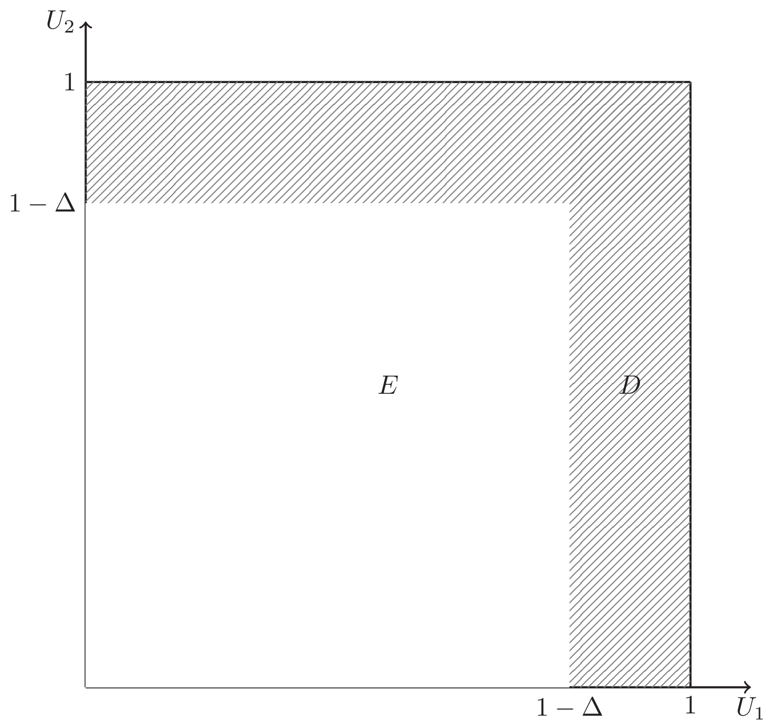
set $u0 = uniform(0, 2-\Delta$
if $(u0 \leq 1)$ **then**
    $u1 = uniform(0,1)$
    $u2 = uniform(1-\Delta, 1)$
else
    $u1 = uniform(1-\Delta, 1)$
    $u2 = uniform(0, 1-\Delta)$
end if
return $u1, u2$

We let $u_1 = u_{1,1}, u_{1,2}, ..., u_{1,n}$, $u_2 = u_{2,1}, u_{2,2}, ..., u_{2,n}$, and $(x_1, x_2) = (F_1^{-1}(u_1), F_2^{-1}(u_2))$, where F_1^{-1} and F_2^{-1} are the inverse cumulative distribution functions. This way we ensure that X_1, X_2 have the desired conditional distribution. These variables are then used to estimate the appropriate values of b_i.

At this stage of the optimization process, we have already determined the optimal values for a_i, and our objective is to find the corresponding b_i values such that $P(X_1, X_2 \in C) = \alpha$. With this importance sampling scheme, we can compute $P(X_1, X_2 \in C | X_1, X_2 \in D)$ for vectors x_1, x_2, and contract parameters a_1, a_2, b_1, b_2. Then, $P(X_1, X_2 \in C)$ can be calculated as:

$$P(X_1, X_2 \in C) = P(X_1, X_2 \in C | X_1, X_2 \in D) \cdot P(D)$$

where $P(D) = 1 - (1-\Delta)^2$, as can be seen in the figure below 2.8.



Find optimal solution

In step 8 of 1 we have two arrays B_1, B_2, for each index of these arrays we have viable combinations of b_1, b_2 that yield a probability of exceeding the insurance layer contract equal to α. We calculate the expected reinsurance cost for each combination and chose the one with the lowest associated cost. According to 2.6.4 these b_i characterizes the optimal solution to the optimization problem.

CHAPTER 3

Numerical results

3.1 Introduction

In this chapter we will introduce uncertainty to our distributions, we will see that adding this uncertainty will affect the optimal solution and we will attempt to explain how this works.

The script we are using creates distributions with uncertainty by introducing uncertainty to one of the parameters of the original distribution. This has to be done numerically in order to create as many different distributions as we are interested in, there are some choices of distributions and parameters which have nice analytical distributions, but they will not be considered in this text.

Since we are doing this numerically we can use any combination of two distributions we can think of, which lets us create an endless number of distributions with different characteristics. By choosing different distributions we can simulate uncertainty and use this to investigate how the uncertainty influences the properties of the resulting solution. The key factor which we expect will have an influence on the result is the heaviness of the tails of distributions, so we are particularly interested in combinations of distributions which leads to significant change in the tail of the distributions.

In order to see how the uncertainty affects the optimal solution we will develop a method for quantifying the uncertainty we add to a distribution and see if we can create a framework where we can see how much uncertainty affects reinsurance optimization.

3.2 Introducing Uncertainty

We want to create distributions $Y_{1,2}$ that have one parameter which has a distribution of its own, $\sigma \cdot X$. An example we will use is a construction where

$$\begin{aligned}
X_1 &= Lognormal(1.0, 1.5) \\
Y_1 &= T_N(50, 30 \cdot X_1) \\
Y_2 &= T_N(50, 30 \cdot X_1),
\end{aligned}$$

where $T_N(\mu, \sigma)$ is the Truncated normal distributed random variable with mean μ and standard deviation σ. Note also that in this chapter the random variables X_{1-3} only specify the distribution we are imposing on the parameters, none of these are correlated in between distributions Y_{a-j} or Y_{1-2}.

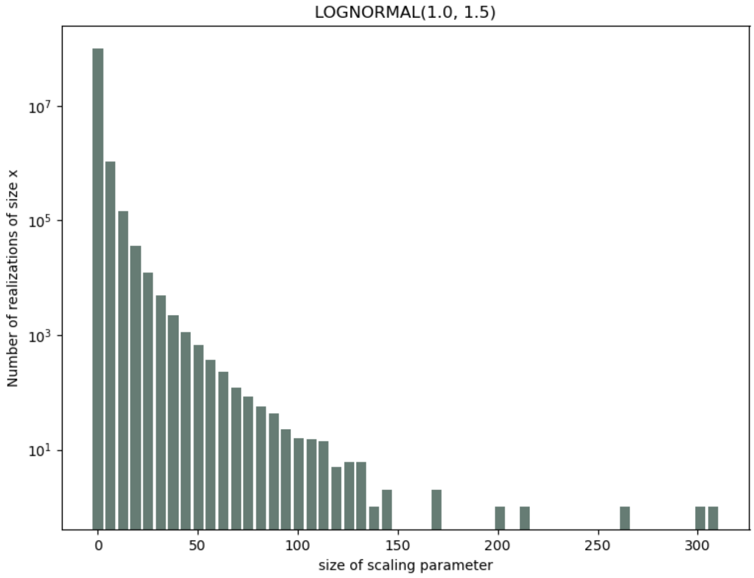


Figure 3.1: 1e8 samples drawn from the lognormal distribution, with empiric mean 1.000 and empiric variance 2.251. We see realizations greater than or equal 200.

Specifically we create the distributions Y_{a-j} by selecting a parameter which sets the number of different realisations we want from X, n. We then use this X to create new parameters for our main distribution. If we want to have uncertainty on the standard deviation of a truncated normal distribution we may define a vector with values from the lognormal distribution lnv with length n and define a vector of standard deviations for the mixed distributions as $\sigma[i] = lnv[i] \cdot y$ for some suitable y. We can then find the probability density function of this new distribution, Y, in x by finding the probability density function for each of the underlying distributions with standard deviation $\sigma[i]$ and summing them each with a weight $\frac{1}{n}$, we let the value of these be $pdf[i](x)$. Hence the probability density function of the distribution Y for a value x is given by

$$y(x) = \frac{1}{n} \sum_{i=1}^{n} pdf[i](x)$$

This new distribution is the distribution we are after. We may also manually set the distribution parameters (i.e. $\sigma[i]$'s) for the mixture distribution but in practice this results in simpler distributions. This way we end up with a new distribution, characterized by the mean, standard deviation and weights of the underlying distributions. While we don't have the analytic probability density function, cumulative distribution function or survival function we can approximate them with arbitrary accuracy and reasonable speed.

3.3 Quantifying Uncertainty in Reinsurance Contract Optimization

In this section, we explore the impact of uncertainty on reinsurance contract optimization through a simulation study. We introduce examples with and without added uncertainty and obtain optimal solutions to these examples. Our goal is to identify a suitable measure of dispersion to represent the degree of uncertainty in the resulting distributions, quantify the uncertainty's influence, and interpret its effect on the increased objective function value observed in the optimization process. We will also see how the uncertainty influences the types of solution. The simulations has been set up as follows:

1. Simulate 1e7 samples for each distribution, plot the distributions and find measures of dispersion.
2. Plot the hazard rate.
3. Find and plot the optimal solutions to the reinsurance contract optimization.
4. Compare potential measures of dispersion of the distributions to the final value of the objective function.

The simulations has been performed for the truncated normal distribution, exponential distribution and pareto distribution. For each such distribution there are two or three examples with added uncertainty, and an example without uncertainty is also simulated as a baseline. Underneath is a list of all examples:

Example	Distribution	μ	σ	Empiric mean	Empiric variance
a	$Y_a \sim TN_I(\mu, \sigma)$	50	30	50.00	30.00
b	$Y_b \sim TN_I(\mu, \sigma)$	50	$30 \cdot X_1$	55.29	52.87
c	$Y_c \sim TN_I(\mu, \sigma)$	$50 \cdot X_1$	30	51.89	44.49
d	$Y_d \sim TN_I(\mu, \sigma)$	50	X_3	67.54	66.22
e	$Y_e \sim EXP(\mu)$	50		49.99	49.98
f	$Y_f \sim EXP(\mu)$	$50 \cdot X_1$		49.12	103.59
g	$Y_g \sim EXP(\mu)$	X_3		50.01	64.03
h	$Y_h \sim Pareto(\mu, \sigma)$	50	30	50.00	30.56
i	$Y_i \sim Pareto(\mu, \sigma)$	X_2	30	50.00	39.23
j	$Y_j \sim Pareto(\mu, \sigma)$	50	X_3	50.00	31.33

Table 3.1: For each example a-j, the steps in the simulation study has been followed. X_{1-3} are auxiliary functions used to introduce uncertainty as discussed in the beginning of the chapter

- $X_1 : Lognormal(1.0, 1.5)$
- $X_2 : P(X_2 = 10) = 0.2, P(X_2 = 30) = 0.2, P(X_2 = 50) = 0.2, P(X_2 = 70) = 0.2, P(X_2 = 90) = 0.2$
- $X_3 : P(X_4 = 10) = 0.2, P(X_3 = 20) = 0.2, P(X_3 = 30) = 0.2, P(X_3 = 40) = 0.2, P(X_3 = 50) = 0.2$

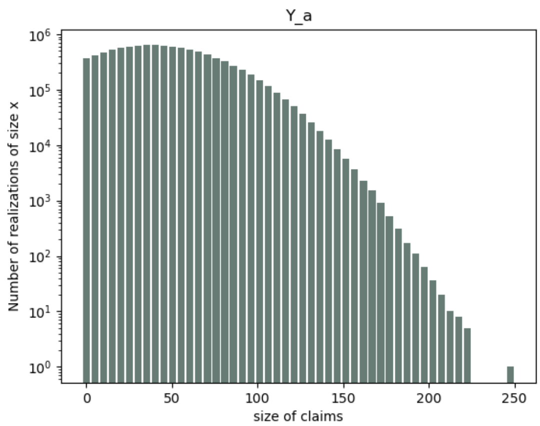
Truncated normal distribution

The truncated normal distribution is the first example because it has one of the more fascinating responses to the added uncertainty because the optimal solution changes depending on what the resulting distribution looks like. We will introduce four examples based on the truncated normal distribution shown in the table 3.2 below.

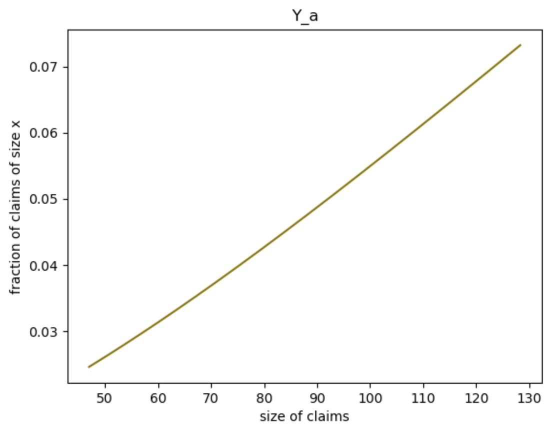
Example	μ	σ	Empiric mean	Empiric stdev	Kurtosis
a	50	30	50.00	30.00	-0.05
b	50	$30 \cdot X_1$	55.29	52.87	49.34
c	$50 \cdot X_1$	30	51.89	44.49	5.03
d	50	X_2	50.14	32.96	6.71

Table 3.2: For each distribution there has been generated 1e7 samples, the empiric mean and empiric variance has been calculated for these samples.

We first look at samples drawn from the distributions. In figure 3.2-3.5 we see how the distributions are skewed to the right by the added uncertainty in the claims sample plots to the left. We note that Y_b (the distribution with the mixed standard deviation) has much higher extreme values than the other distributions. There is a significant increase in empiric variance in Y_b and Y_c. We can also see a dramatic increase in kurtosis in Y_b, as discussed in chapter 2, this estimate is somewhat prone to being influenced by outliers in the numeric estimation, however the estimates in this example seem significant and may indicate that a higher kurtosis leads to a balanced solution, while a lower kurtosis leads to an unbalanced solution.



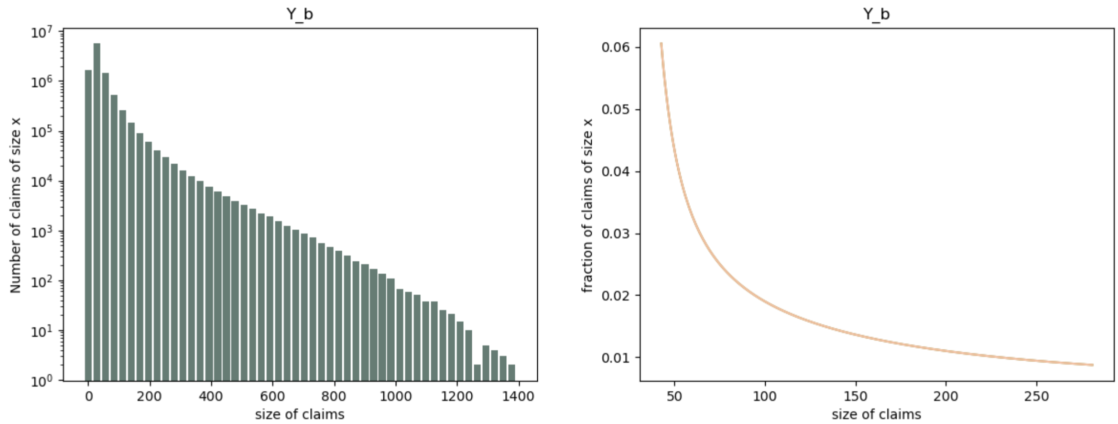
claims sample Y_a



hazard rate Y_a

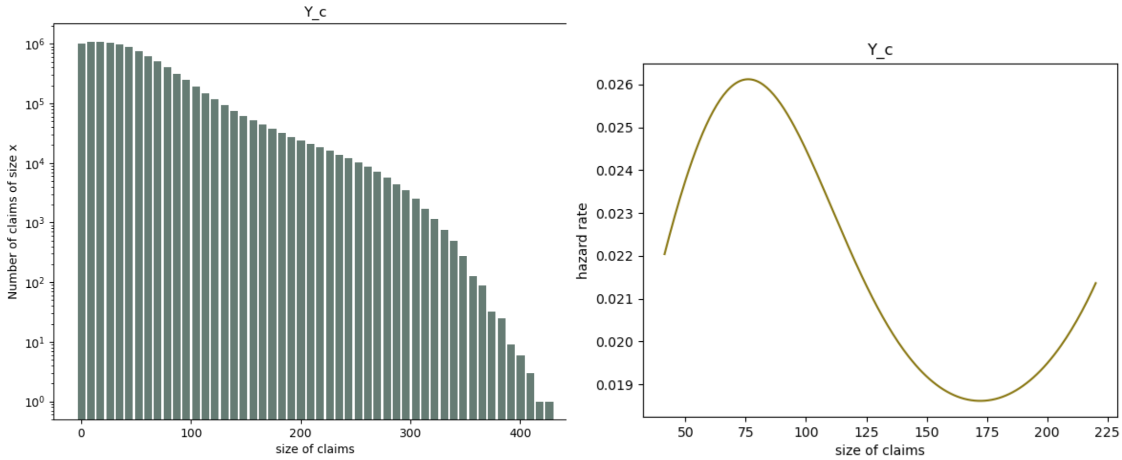
Figure 3.2: Y_a

Next, we look at the hazard rate of the distributions, from chapter 2 we know that a monotone increase in hazard rate will lead to a solution which can be characterized as a border point solution, and opposite for a monotone decreases in hazard rates. Distribution Y_a has a strictly increasing hazard rate



claims sample Y_b hazard rate Y_b

Figure 3.3: Y_b

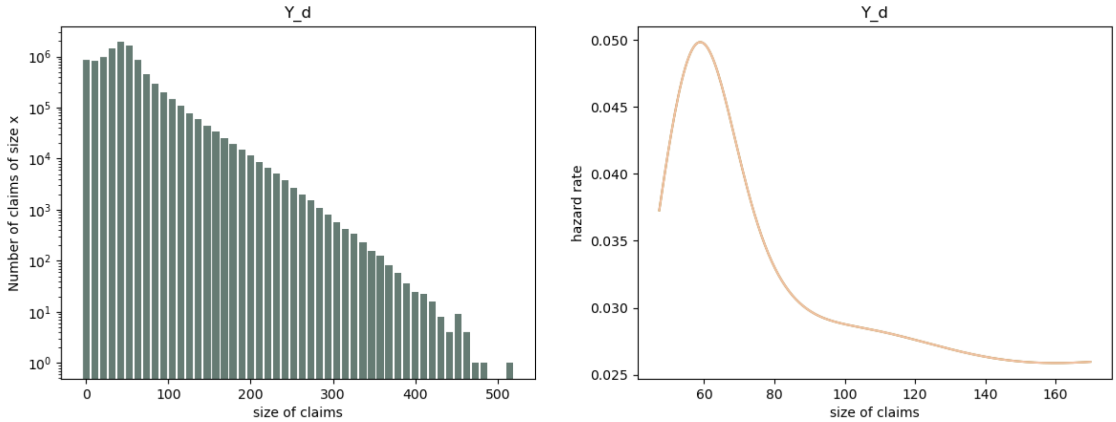


claims sample Y_c hazard rate Y_c

Figure 3.4: Y_c

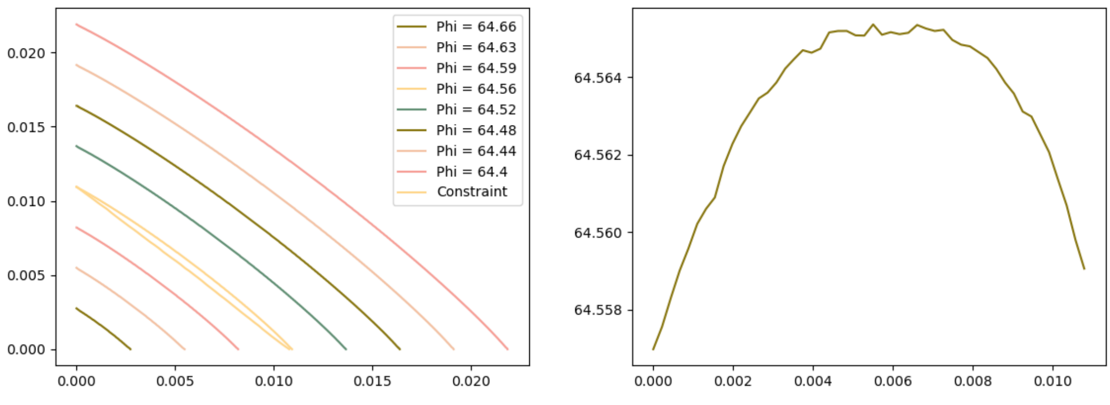
while it is strictly decreasing for Y_b. Again, Y_c, Y_d has unexpected shapes, as they are not monotone, this means that the theory for hazard rates does not apply. We include them because they underpin the main result in the end of the section.

In the figures (3.6, 3.7,3.8, 3.9), we see that Y_a, Y_c and Y_d has a border solution while Y_b has an inner point solution. We note that even if the standard deviation of all the 'new' distributions is increased, only the mixed standard deviation with increased kurtosis gets a change in solution type.



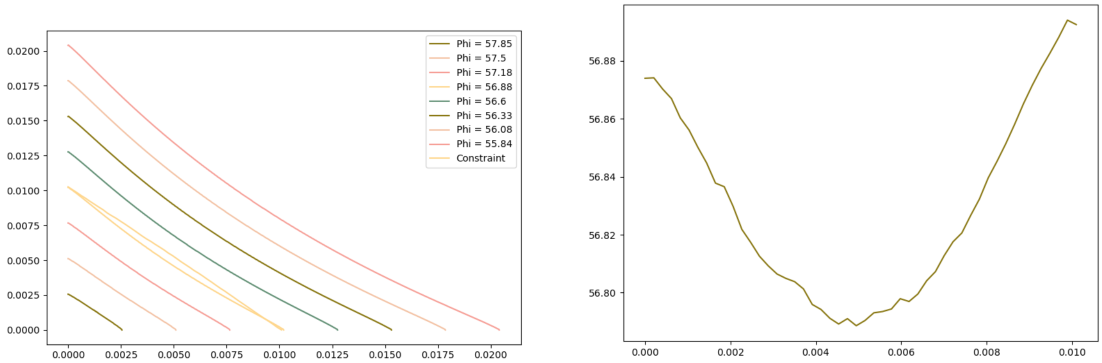
claims sample Y_d hazard rate Y_d

Figure 3.5: Y_d



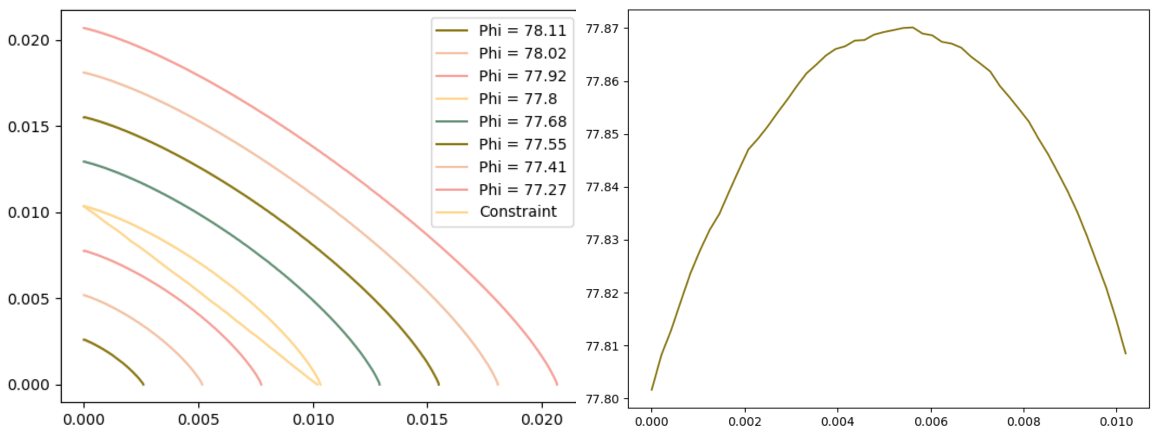
Iso-curves Y_a Optimal B_1 Y_a

Figure 3.6: Y_a



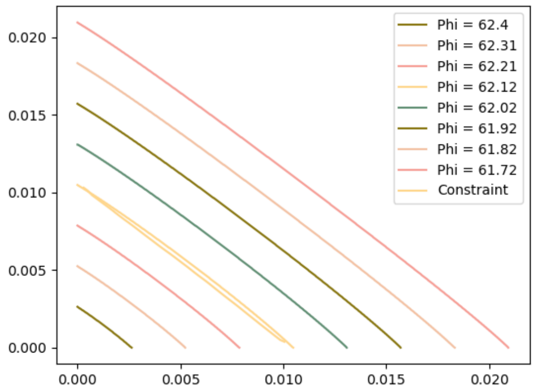
Iso-curves Y_b Optimal B_1 Y_b

Figure 3.7: Y_b

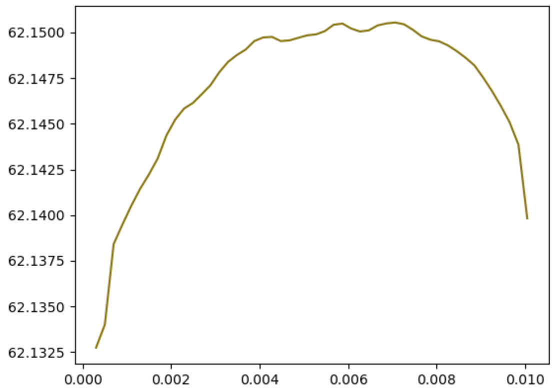


Iso-curves Y_c Optimal B_1 Y_c

Figure 3.8: Y_c



Iso-curves Y_d



Optimal B_1 Y_d

Figure 3.9: Y_d

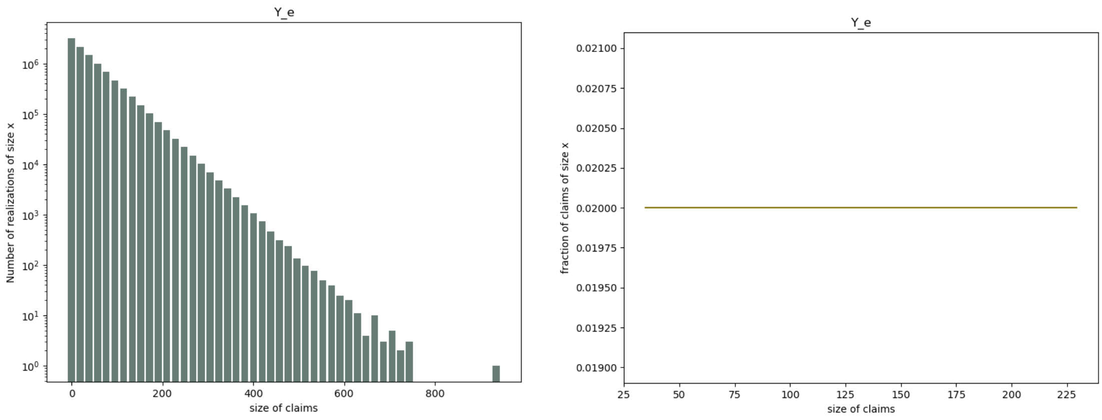
Exponential distribution

The next example we look at is the exponential distribution, we have already looked at the analytical solution in the previous section but now we introduce uncertainty to this example as well. The exponential distribution is often thought of as the boundary between light- and heavy-tailed distributions, so it is also interesting to see if the solutions change. We will only have three examples due to the exponential distribution only having one parameter.

Example	μ	Empiric mean	Empiric stdev	Empiric kurtosis
Y_e	50	49.99	49.98	6.02
Y_f	$50 \cdot X_1$	49.12	103.59	99.29
Y_g	X_3	50.01	64.03	11.82

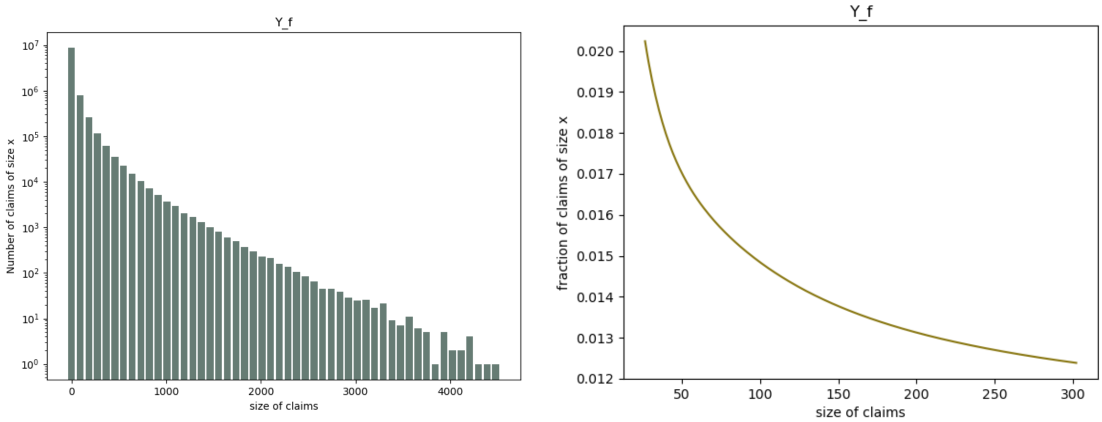
Table 3.3: For each distribution there has been generated 1e7 samples, the empiric mean, empiric variance and empiric kurtosis has been calculated for these samples.

In the table 3.3 we see that Y_f has a large increase in empiric standard deviation and kurtosis, while Y_g has a more moderate increase in both. This is reiterated in the claims sample plots, where Y_f is much more skewed. If we now compare this to the plots with the Iso-curves and Optimal B_1 we see that Y_f has a much more curved iso-curve and contstraint, indicating that the uncertainty added to this example has had the most effect. While both Y_e and Y_g has inner point solutions, we see that the iso-curve almost has the same shape as the constraint and that differences in the values of ϕ in the Optimal B_1-plots are very small.



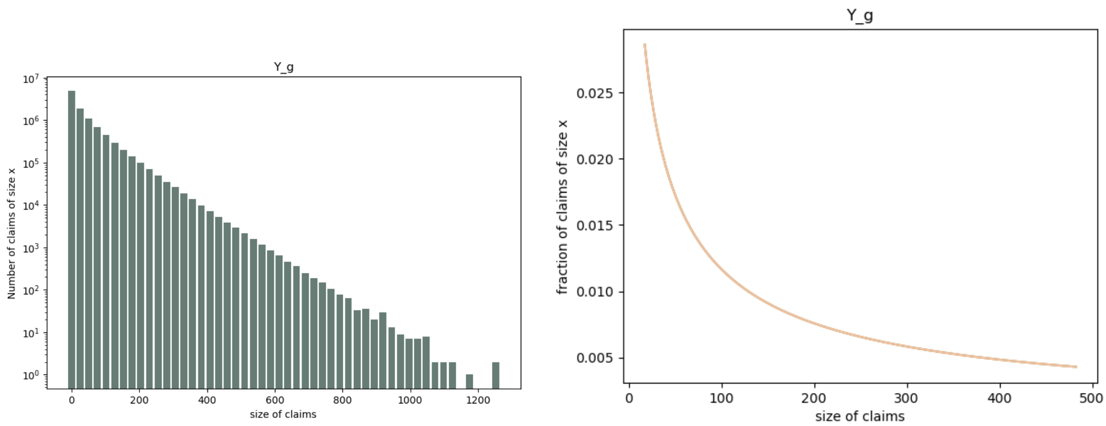
claims sample Y_e hazard rate Y_e

Figure 3.10: Y_e



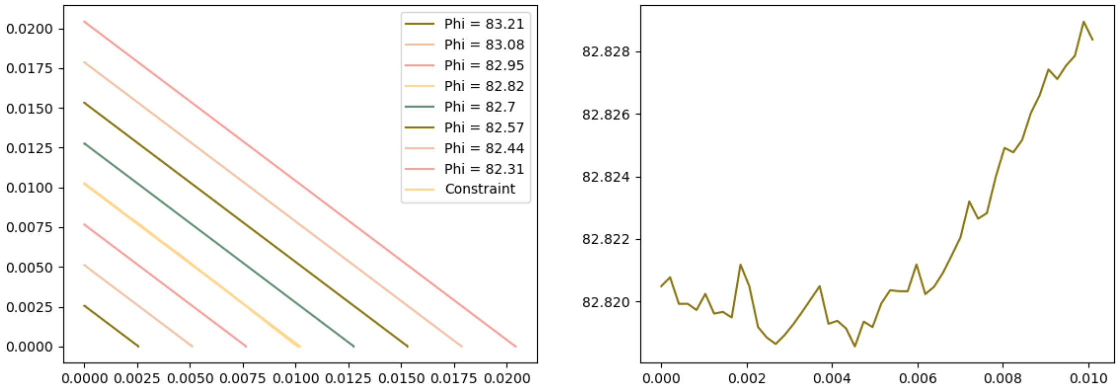
claims sample Y_f hazard rate Y_f

Figure 3.11: Y_f



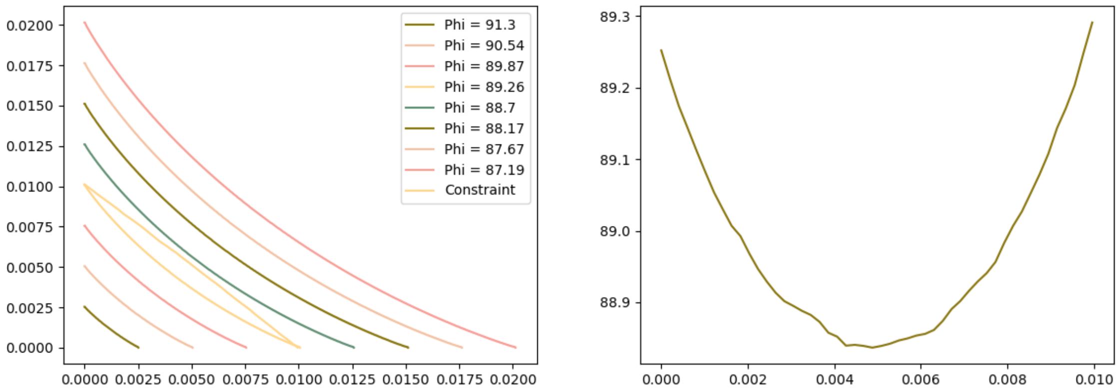
claims sample Y_g hazard rate Y_g

Figure 3.12: Y_g



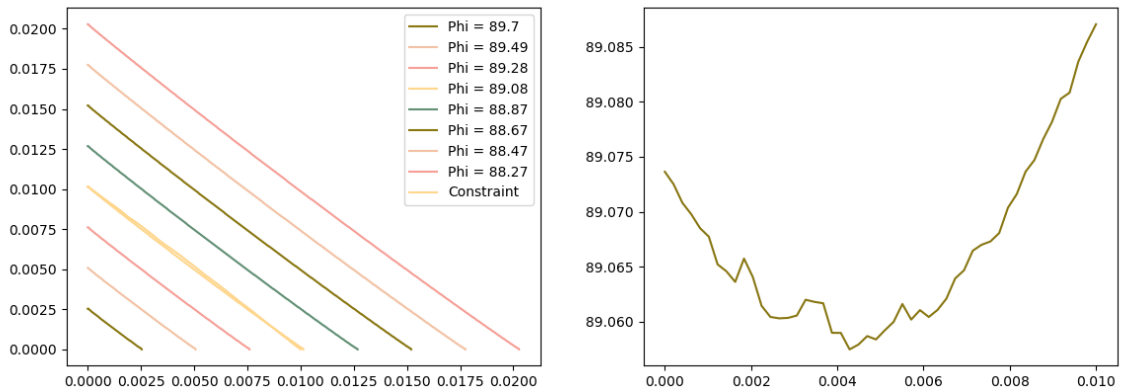
Iso-curves Y_e Optimal B_1 Y_e

Figure 3.13: Y_e



Iso-curves Y_f Optimal B_1 Y_f

Figure 3.14: Y_f



Iso-curves Y_g Optimal B_1 Y_g

Figure 3.15: Y_g

Pareto distribution

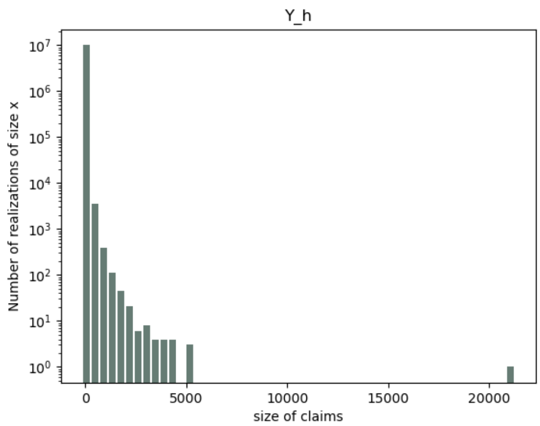
Example	μ	σ	Empiric mean	Empiric stdev	Kurtosis
Y_h	50	30	50.00	30.56	25318.27
Y_i	X_3	30	50.00	39.23	825.30
Y_j	50	X_4	50.00	31.33	4198.78

Table 3.4: For each distribution there has been generated 1e8 samples, the empiric mean, empiric standard deviation and empiric kurtosis has been calculated for these samples.

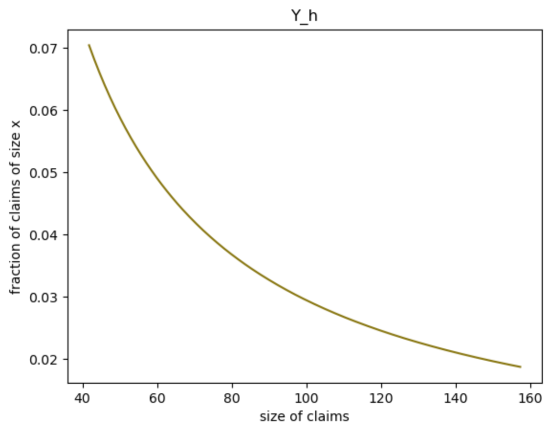
For the pareto distribution the empiric kurtosis is too unstable to draw any clear conclusions from, however, typically the kurtosis for Y_j is higher than for Y_i, which is in line with the results from the exponential distribution and truncated normal distribution. The standard deviation is, again, higher for the example with the mixed mean than for the mixed standard deviation.

In the claims sample plots we see that Y_h has one extreme outlier, and Y_i and Y_j has multiple less extreme outliers, but heavier tails overall. The hazard rates for examples Y_h and Y_j are quite as expected, and while Y_i might look a little odd it is only because of how the uncertainty on the mean is constructed in conjunction with the heavy tail.

The iso-curves and Optimal B_1's are quite similar for all the pareto examples, which is to be expected as all the distributions, irrespective of uncertainty, have very heavy tails.

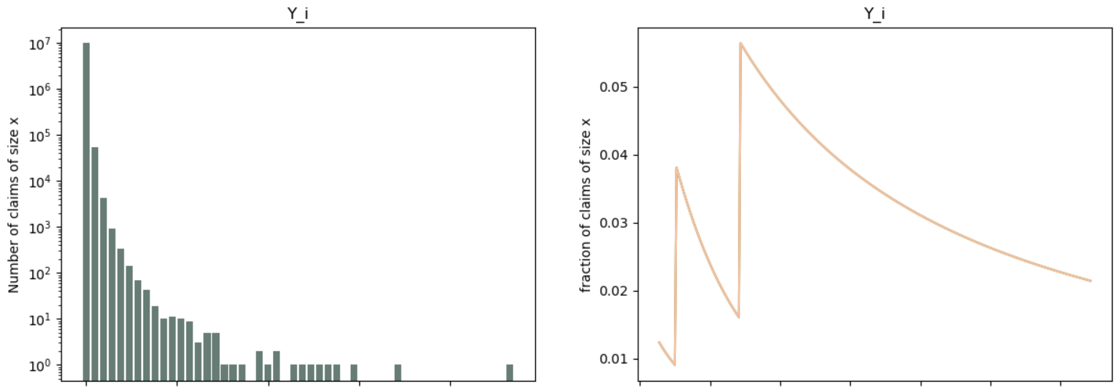


claims sample Y_h



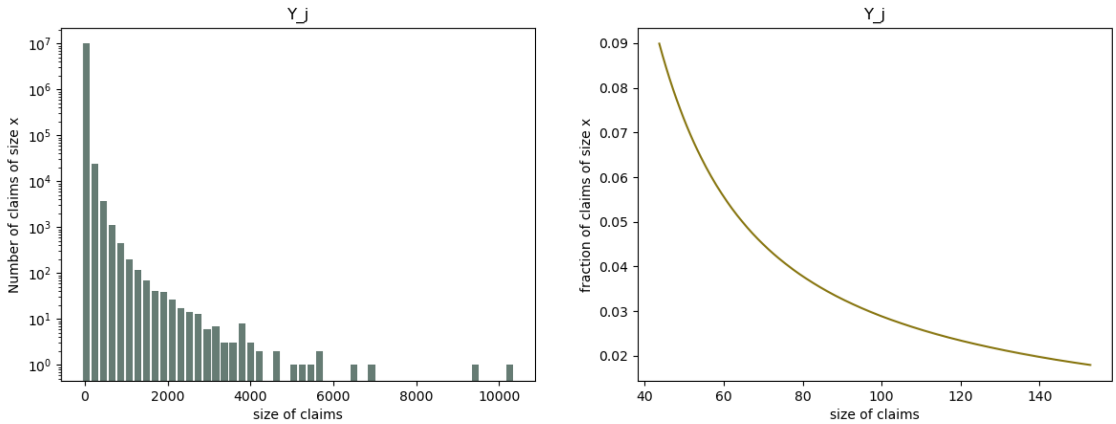
hazard rate Y_h

Figure 3.16: Y_h



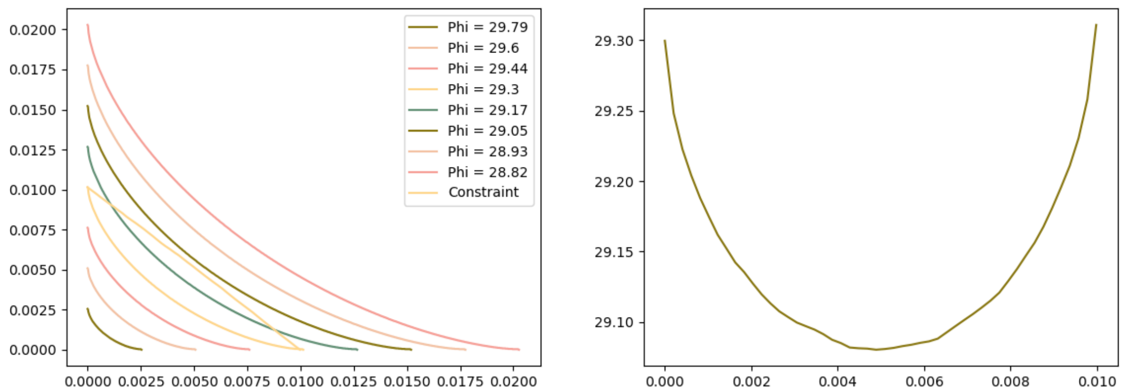
claims sample Y_i hazard rate Y_i

Figure 3.17: Y_i



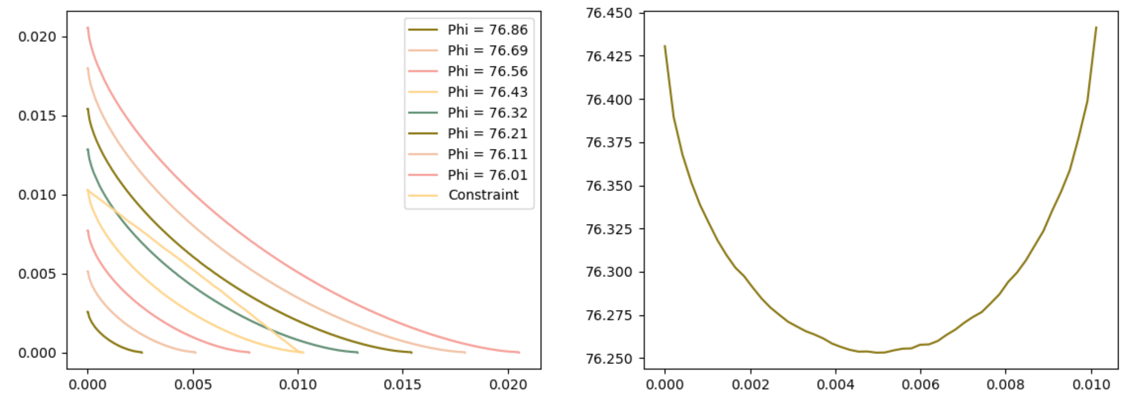
claims sample Y_j hazard rate Y_j

Figure 3.18: Y_j



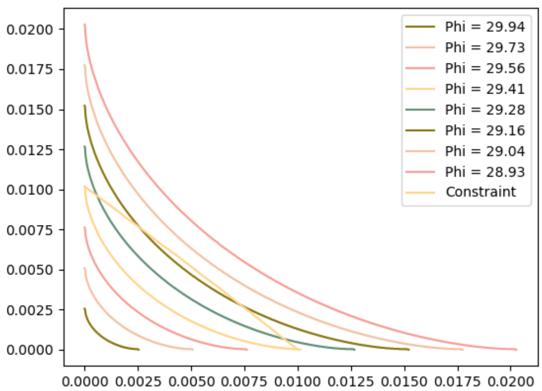
Iso-curves Y_h Optimal B_1 Y_h

Figure 3.19: Y_h

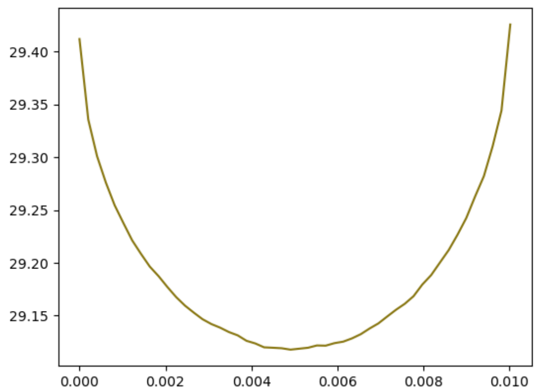


Iso-curves Y_i Optimal B_1 Y_i

Figure 3.20: Y_i



Iso-curves Y_j

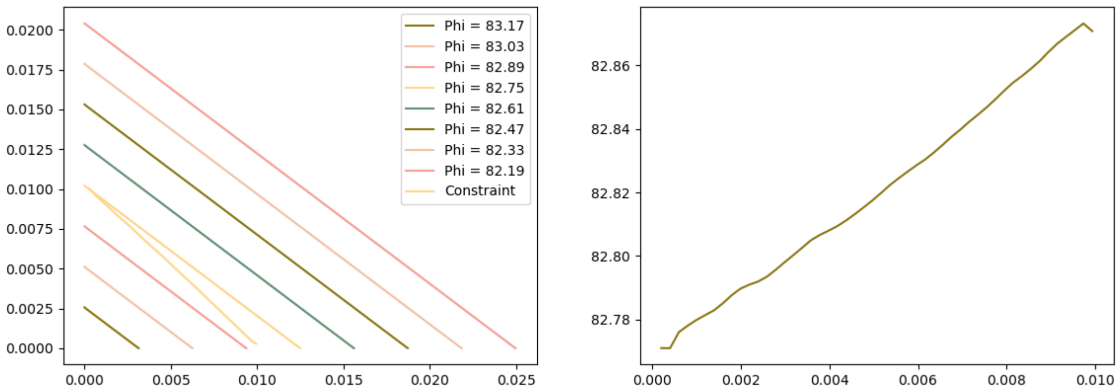


Optimal B_1 Y_j

Figure 3.21: Y_j

Asymmetric distributions

We include one asymmetric example of the exponential distribution, the idea being that for asymmetric distributions the iso-curves will be tilted, so that holding more of the light tailed risk is optimal. We are interested to see whether there are any other significant changes.



Iso-curves Y_k Optimal B_1 Y_k

Figure 3.22: Y_k. Note that the solution to this example is not symmetric, so the scale is not the same on the x- and y-axis.

We see that we get a very clear increase in our objective function when b_1 is increased, which is due to the tilted constraint. The optimal solution is thus a border solution as expected. From the iso-curve plot we can theorise that increasing the curve on the iso-curves, again could lead to an inner point solution.

Summary

We have seen that uncertainty has a real impact on the solutions of optimal reinsurance contracts, in this section we will try to draw some conclusions from the observations. We will focus on the change in standard deviation and kurtosis and how they impact

1. the type of solution.
2. the contract parameters

We need a couple of large tables to get an overview of the results

Example	a_1	a_2	b_1	b_2	Empiric stdev	Empiric Kurtosis
a	19.13	19.13	∞	127.40	29.99	-0.05
b	28.00	28.00	368.95	359.24	52.77	49.34
c	13.95	13.95	∞	219.28	44.49	5.03
d	20.60	20.60	306.44	169.00	32.96	6.71
e	9.12	9.12	269.77	257.86	49.98	6.02
f	3.31	3.31	669.22	647.77	103.59	99.29
g	5.53	5.53	372.40	345.84	64.04	11.82
h	35.12	35.12	201.21	196.19	30.56	25318
i	12.28	12.28	205.59	201.65	39.32	825.30
j	35.25	35.25	199.25	193.83	31.33	4198.78
k	8.20	10.03	352.23	253.84	-	-

Example	Uninsured V_α	Retained risk	Premium	V_α
a	257.27	39.50	76.24	115.73
b	563.53	56.01	68.15	124.15
c	441.69	29.46	91.80	121.27
d	340.68	42.49	73.27	115.76
e	460.52	18.23	99.38	117.61
f	968.48	6.63	106.60	113.23
g	606.21	11.06	106.87	117.93
h	315.62	70.25	34.90	105.14
i	337.92	24.55	91.50	116.06
j	306.55	70.50	34.94	105.44
k	460.52	18.23	99.33	117.56

We make the following observations

- The uninsured risk is varying a lot, the Value at Risk after reinsurance is quite stable.
- We see, however, that increases in the standard deviation is associated with increases in V_α, the exception is Y_f where a high empiric standard deviation and kurtosis leads to a balanced solution with low retained risk and the lowest V_α of the exponential examples.
- In the previous section we saw that the examples with uncertainty added to the standard deviation were associated with more curved iso-curves, here we see that the same examples has a higher increase in kurtosis. This indicates that an increase in kurtosis is associated with a an increased likelihood of an inner point solution being the optimal solution.
- The examples where standard deviation is increased significantly and kurtosis only has a mild increase, Y_c, Y_f and Y_i, also has a significant reduction in retained risk.

3.4 Evaluating uncertainty

In order to have a measure of how much adding uncertainty costs, we need a measure of how much uncertainty we are adding. We have seen that uncertainty added has two different consequences which is reflected in changes in two different summary statistics, standard deviation and kurtosis. We conclude that it is the dispersion parameter standard deviation that best reflects increases in V_α, while increases in the tailedness parameter kurtosis doesn't increase the V_α as much but can lead to different optimal solutions.

3.5 Further research

A natural next step is to see what happens if we increase the number of contracts and try to find optimal unbalanced solutions. This will require updating the algorithm, and will be significantly more costly computationally, and probably becomes unfeasible after m = 3 or m = 4. Furthermore it might be beneficial to look at more examples. In the examples presented in this thesis there is a quite clear connection changes in solution and increases in kurtosis, and between increases in standard deviation and increases in V_α. However, to improve the certainty of these results it might be beneficial to look at more distributions.

Appendices

APPENDIX A

The First Appendix - Probability distributions

Pareto distribution

The Pareto distribution is a heavy tailed distribution with a shape parameter α, a scale parameter σ and an optional location parameter μ. If the location parameter is used the distribution is called a type 2 distribution and if not a type 1 distribution. The two are equivalent if $\sigma = \mu$. This can be shown by comparing the inverse cumulative distributions.

- Pareto type 1 cdf $F_I(x) = 1 - \left(\frac{x}{\sigma}\right)^{-\alpha}$
- Pareto type 1 inverse cdf $F_I^{-1}(u) = \sigma(1-u)^{\frac{-1}{\alpha}}$
- Pareto type 2 cdf $F_{II}(x) = 1 - \left(1 + \frac{x-\mu}{\sigma}\right)^{-\alpha}$
- Pareto type 2 inverse cdf $F_{II}^{-1}(u) = \sigma(1-u)^{\frac{-1}{\alpha}} + \mu - \sigma$

Lomax distribution

If we let $\mu = 0$, we get the Lomax distribution, which is also a heavy tailed distribution.

- $F_L(x) = 1 - \left(1 + \frac{x}{\sigma}\right)^{-\alpha}$
- $F_L^{-1}(u) = \sigma(1-u)^{\frac{-1}{\alpha}} - \sigma$

Gamma Exponential distribution

The Gamma Exponential distribution is given by an exponential distribution where the rate parameter is modelled by a Gamma distribution $X|\lambda \sim Exp(\lambda)$. Depending on how the parameters are chosen we can end up with some very convenient distributions. The conditional density of this distribution is given by

$$f_{X|\lambda}(x|\lambda) = \lambda e^{-\lambda x}$$

If we integrate this for $t \in (x, \infty)$ we get the survival function

$$P(X > x|\lambda) = \int_x^\infty \lambda e^{-\lambda t} dt = e^{-\lambda x}$$

Next we can find the unconditional survival function by integrating $P(X > x|\lambda)g(\lambda)d\lambda$ where $g(\lambda)$ is the density of the gamma distribution.

$$\begin{aligned}
P(X > x) &= \int_0^\infty P(X > x|\lambda)g(\lambda)d\lambda = \int_0^\infty e^{-\lambda x}\frac{\beta^\alpha}{\Gamma(\alpha)}\lambda^{\alpha-1}e^{-\beta\lambda}d\lambda \\
&= \frac{\beta^\alpha}{\Gamma(\alpha)}\int_0^\infty \lambda^{\alpha-1}e^{-(\beta+x)\lambda}d\lambda = \frac{\beta^\alpha}{\Gamma(\alpha)}\frac{\Gamma(\alpha)}{(\beta+x)^\alpha} \\
&= \frac{\beta^\alpha}{(\beta+x)^\alpha}
\end{aligned}$$

This, in turn can be written as $\left[1+\frac{x}{\beta}\right]^{-\alpha}$, so $X \sim$ Lomax (α, β)

Truncated Normal distribution

The truncated normal distribution is a normal distribution with a limited interval. We have that if $Y \sim N(\mu, \sigma)$ and $X =^d (Y|Y \in I)$, where $I \subseteq \mathbb{R}$ is some interval of $\mathbb{R}$, then $X \sim TN_I(\mu, \sigma)$. As we are interested in reinsurance of positive claims we let $I = [0, \infty)$. We have that the inverse cumulative distribution function of the truncated normal distribution, truncated to the interval I, is given by

$$X = \mu + \sigma \cdot \Phi^{-1}\left(\Phi\left(\frac{-\mu}{\sigma}\right) + U \cdot \left(1 - \Phi\left(\frac{-\mu}{\sigma}\right)\right)\right)$$

where Φ is the cumulative distribution function of the standard normal distribution and U is the uniform distribution on $(0, 1)$. We need X to be non-negative with probability one. We see that X is increasing in U, and if we let U $= 0$, then

$$\begin{aligned}
X &= \mu + \sigma \cdot \Phi^{-1}\left(\Phi\left(\frac{-\mu}{\sigma}\right)\right) \\
&= \mu + \sigma \cdot \left(\frac{-\mu}{\sigma}\right) = \mu - \mu = 0
\end{aligned}$$

Next, we can verify that the distribution is as postulated. We let $x \geq 0$ and consider the probability $P(X > x)$:

$$\begin{aligned}
P(X > x) &= P\left(\frac{X-\mu}{\sigma} > \frac{x-\mu}{\sigma}\right) \\
&= P\left(\Phi^{-1}\left(\Phi\left(\frac{-\mu}{\sigma}\right) + U \cdot (1 - \Phi(\frac{-\mu}{\sigma}))\right) > \frac{x-\mu}{\sigma}\right) \\
&= P\left(\Phi(\frac{-\mu}{\sigma}) + U \cdot (1 - \Phi(\frac{-\mu}{\sigma}) > \Phi(\frac{x-\mu}{\sigma})\right) \\
&= P\left(U > \frac{\Phi(\frac{x-\mu}{\sigma}) - \Phi(\frac{-\mu}{\sigma})}{1 - \Phi(\frac{-\mu}{\sigma})}\right) = 1 - \frac{\Phi(\frac{x-\mu}{\sigma}) - \Phi(\frac{-\mu}{\sigma})}{1 - \Phi(\frac{-\mu}{\sigma})} \\
&= \frac{1 - \Phi(\frac{x-\mu}{\sigma})}{1 - \Phi(\frac{-\mu}{\sigma})} = \frac{P\left(Y > x\right)}{P\left(Y > 0\right)} \\
&= P\left(Y > x|Y > 0\right)
\end{aligned}$$

for a normally distributed Y with mean μ and standard deviation σ.

The inversion method

The inversion method is used to simulate distributions, as long as we can find the inverse distribution function it is possible to generate samples of the cumulative distribution function of the distribution in question by using the uniform distribution. Let $F(x)$ be a strictly increasing distribution function with inverse $x = F^{-1}(u)$ and let $X = F^{-1}(U)$, where $U \sim Uniform(0, 1)$. Then

$$F_X(x) = P(X \leq x) = P(F^{-1}(U) \leq F^{-1}(u)) = P((U) \leq u)$$

This enables us to sample from any of the distribution specified in the appendix as we have a formula for the inverse of each of them.

Bibliography

[Art+99] Artzner, P. et al. 'Coherent Measures of Risk'. In: *Mathematical Finance* vol. 9, no. 3 (1999), pp. 203–228. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/1467-9965.00068.

[Che+14] Cheung, K. et al. 'Optimal reinsurance under general law-invariant risk measures'. In: *Scandinavian Actuarial Journal* vol. 2014, no. 1 (2014), pp. 72–91.

[HC20] Huseby, A. B. and Christensen, D. 'Optimal Reinsurance Contracts in the Multivariate Case'. In: Jan. 2020, pp. 465–472.

[Hus22] Huseby, A. 'Optimizing Multiple Reinsurance Contracts'. In: *Proceedings of the 32th European Safety and Reliability Conference*. 2022.

[Hus23] Huseby, A. B. *Lecture notes STK4400 "Risk and reliability analysis". University of Oslo, spring 2023.* 2023.