UiO : Department of Mathematics
University of Oslo

Optimal Reinsurance in a Multivariate Case

SMR5960

Andreas Christensen

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

This Master's Thesis aims to optimize reinsurance contracts in both univariate and multivariate cases, contributing to the advancement of reinsurance optimization techniques. First, we review the optimization methodology and identify the parts that can be solved analytically. We then develop Monte Carlo simulation methods to optimize a set of reinsurance contracts, using value-at-risk as the risk measure and exploring importance sampling to obtain more stable results and illustrate the methods with symmetrical and asymmetrical examples. Our findings provide insights for practitioners and researchers in the field and demonstrate the potential of Monte Carlo simulation and the importance sampling in optimizing multivariate reinsurance contracts.

Acknowledgements

I would like to express my gratitude to everyone who has supported me during my studies. Firstly, I would like to thank the Norwegian Student Actuarial Association for motivating me by providing insights into the applications of actuarial science throughout my master's degree. Secondly, I would like to thank my fellow students in study hall B1001 for creating a quiet and conducive environment that allowed me to concentrate on my studies. I would also like to express my appreciation to my supervisor for introducing me to the field of theoretical reinsurance and its intriguing applications. Lastly, I am grateful to my parents and siblings for their guidance and valuable advice, which helped me establish a strong foundation for my education.

Contents

Abstract **i**

Acknowledgements **ii**

Contents **iii**

List of Figures **v**

List of Tables **vii**

1 Introduction **2**

2 Preliminaries **5**
- 2.1 Insurance concepts . . . 5
- 2.2 Reinsurance contracts . . . 6
  - 2.2.1 Insurance layer contract . . . 6
  - 2.2.2 Premiums . . . 7

3 Theoretical background **10**
- 3.1 Risk measure . . . 10
  - 3.1.1 Value-at-risk . . . 10
- 3.2 Univariate reinsurance contracts . . . 13
  - 3.2.1 Partial derivatives total reinsurance premium . . . 13
  - 3.2.2 Value-at-risk . . . 13
- 3.3 Multivariate reinsurance contracts . . . 14
  - 3.3.1 Partial derivatives total reinsurance premium . . . 15
  - 3.3.2 Value-at-risk as the objective function . . . 16
  - 3.3.3 Minimizing value-at-risk . . . 16
- 3.4 Optimization problem . . . 19
  - 3.4.1 Optimal contract parameters . . . 19
  - 3.4.2 Expected reinsurance expenses . . . 19
  - 3.4.3 Constraint . . . 19
  - 3.4.4 Convexity and concavity . . . 23
  - 3.4.5 Hazard rate . . . 25
- 3.5 Generalized solutions of the optimization problem . . . 25
  - 3.5.1 Balanced solutions . . . 25
  - 3.5.2 Unbalanced solutions . . . 28

3.6 Distributions . . . 29
3.6.1 Copulas . . . 29

4 Optimal reinsurance **32**
4.1 Univariate reinsurance . . . 32
4.1.1 Partial derivatives of expected gain . . . 32
4.1.2 Objective function . . . 33
4.2 Multivariate reinsurance . . . 35
4.2.1 Partial derivatives of expected gain . . . 35
4.2.2 Optimizing a given b . . . 36
4.2.3 Optimizing b given a . . . 37
4.2.4 Optimal solution in the exponential case . . . 39
4.2.5 Lagrange approach . . . 42
4.2.6 The objective function . . . 43

5 Numerical results **46**
5.1 Optimization in the univariate case . . . 46
5.1.1 Simulation . . . 46
5.1.2 Results . . . 47
5.2 Optimization in the bivariate case . . . 48
5.2.1 Simulation . . . 48
5.2.2 Symmetrical . . . 49
5.2.3 Asymmetrical . . . 56
5.2.4 Change in dependency . . . 68
5.3 Discussion . . . 71
5.4 Conclusion . . . 73

Appendices **75**

A The First Appendix **76**

B The Second Appendix **78**

Bibliography **79**

List of Figures

2.1 Reinsurance concept 6
2.2 Retained risk 7

3.1 CDF and survival function visualizations 12
3.2 The sets $\mathcal{A}$, $\mathcal{B}$ and $\mathcal{C}$ for $a_i, b_i,\ i = 1, 2$ 17
3.3 $\mathcal{D}' = \{\mathbf{u} : 1 - \Delta < u_i < 1,\ \ i = 1, 2\}$ 21
3.4 Verification of importance sampling 22

4.1 Granulated sets $\mathcal{A}$, $\mathcal{B}$ and $\mathcal{C}$ for $i = 1, 2$ 39
4.2 Objective function $C_1(A)$ 44

5.1 Univariate LS and RS versus A 47
5.2 Univariate LS and RS versus A 47
5.3 Univariate objective function C_1 versus A 48
5.4 Univariate objective function C_1 versus A 48
5.5 Constraint and Iso-contours for example 1 in the log-normal distribution 50
5.6 Φ versus B_1 in example 1 51
5.7 C_1 versus A in example 1 51
5.8 Constraint and Iso-contours for example 6 in the Pareto distribution 52
5.9 Φ versus B_1 in example 4 52
5.10 C_1 versus A in example 4 52
5.11 Constraint and Iso-contours for example 8 in the truncated normal distribution 54
5.12 Φ versus B_1 in example 8 54
5.13 C_1 versus A in example 8 54
5.14 Constraint and Iso-contours for example 11 in the exponential distribution 56
5.15 Φ versus B_1 in example 11 56
5.16 C_1 versus A in example 11 56
5.17 Constraint and Iso-contours for example 13 and 15 in the log-normal distribution 58
5.18 Objective function C_1 versus A for example 15 58
5.19 Φ versus B_1 asymmetrical cases for lognormally distributed risks . 59
5.20 Constraint and Iso-contours for example 17 and 18 in the Pareto distribution 60
5.21 Φ versus B_1 asymmetrical cases for Pareto distributed risks 61

5.22 Objective function C_1 versus A for example 18 62
5.23 Constraint and Iso-contours for example 19 and 21 in the truncated normal distribution 63
5.24 Objective function C_1 versus A for example 21 64
5.25 Φ versus B_1 asymmetrical cases for truncated normally distributed risks 65
5.26 Constraint and Iso-contours for example 22 and 23 in the exponential distribution 66
5.27 Φ versus B_1 asymmetrical cases for exponentially distributed risks 67
5.28 Objective function C_1 versus A for example 23 68
5.29 Φ versus B_1 for change in dependency 70

List of Tables

5.1 Optimal contract parameters in the univariate case 47
5.2 Optimal values for symmetrical log-normally distributed risks . . . 49
5.3 Optimal contract parameters for symmetrical log-normally distributed risks 50
5.4 Optimal values for symmetrical Pareto distributed risks 51
5.5 Optimal contract parameters for symmetrical Pareto distributed risks 51
5.6 Optimal values for symmetrical truncated normally distributed risks 53
5.7 Optimal contract parameters for symmetrical truncated normally distributed risks 53
5.8 Optimal values for symmetrical exponentially distributed risks . . 55
5.9 Optimal contract parameters for symmetrical exponentially distributed risks 55
5.10 Optimal values for asymmetrical log-normally distributed risks . . 57
5.11 Optimal contract parameters for log-normally symmetrical distributed risks 57
5.12 Optimal values for asymmetrical Pareto distributed risks 59
5.13 Optimal contract parameters for asymmetrical Pareto distributed risks 59
5.14 Optimal values for asymmetrical truncated normally distributed risks 62
5.15 Optimal contract parameters for asymmetrical truncated normally distributed risks 62
5.16 Optimal values for asymmetrical exponentially distributed risks . . 65
5.17 Optimal contract parameters for asymmetrical exponentially distributed risks 66
5.18 Optimal values for truncated normally distributed risks with change in dependency 69
5.19 Optimal parameters for change in dependency ρ 69

CHAPTER 1

Introduction

The fundamental principle of the insurance industry involves transferring risk from the insured party to the insurer, who charges a premium in return. The insurer then assumes economic responsibility for mitigating the risk associated with the insured party.

In cases where the insured party's coverage amount is significant, the possibility of a large claim arises. This can result in irreversible harm to the insurer's portfolio or even lead to the failure to fulfill their financial obligation to compensate for the damages incurred on the insured object. As a precaution for these scenarios, insurance companies transfer risks between themselves to diversify risk and reduce the total exposure to potential claims. The concept of reinsurance has arisen, and the primary insurer that further transfers the risk to a reinsurer is known as the *cedent*. The cedent may transfer the risk for a large insurance contract or portfolio. This diversification strategy will, in return, optimize the cedent's financial position.

Usually, a single policyholder may incur a single risk factor, but since we are dealing with reinsurance, it is beneficial to mitigate risks. Insurance companies do this by pooling several risks to reduce the possibilities for financial losses and increase their ability to fulfill financial obligations to policyholders. The sole purpose is to limit the insurer's exposure to complex and correlated risks. However, note that pooling these risks does not necessarily mean that they are bundled, since they may contain different distributions. Therefore, we have a case of multivariate reinsurance contracts.

Despite the benefits of multivariate contracts, they are complex in both price and structure, which can lead to sub-optimal results for both parties involved. Necessary considerations include the reinsurance contract type and the risk measure to use. The main focus of this thesis is to review the methodology for optimizing reinsurance in univariate and multivariate cases, in particular by exploring the applications of optimization that can be done analytically, in our case, by using Lagrange multipliers. In addition, we examine the optimization of a collection of reinsurance contracts using Monte Carlo simulation methods. The primary risk measure will be value-at-risk, while the objective function will weigh value-at-risk for retained risk against the expected gain. We wish to minimize the value-at-risk and maximize the expected gain. However, maximizing expected gain can result in no reinsurance contracts being implemented and the same goes for a significant decrease in value-at-risk, as these goals contradict each other. Therefore, the objective function balances both outcomes to find the optimal solution.

The methods developed will rely heavily on Monte Carlo simulation and importance sampling techniques to perform the sampling of risks in the multivariate case. To showcase the efficacy of these methods, a range of examples will be presented in the bivariate case, encompassing both risks with identical distributions and risks with different distributions.

This Master's Thesis aims to provide fresh perspectives on optimizing multivariate reinsurance contracts. It also explores the advantages of incorporating value-at-risk as a risk measure and the potential use of importance sampling to enhance the reliability of Monte Carlo simulation outcomes.

2. Chapter 2 introduces the nature of insurance contracts and their applications to reinsurance. Definitions such as *stop-loss* and *insurance layer contracts* are introduced, in addition to how risks and premiums behave given a contract.

3. Chapter 3 deals with the theoretical background for our problem, with a special focus on risk measure, objective function, insurance layer contracts, optimization problems, and numerical simulation.

4. Chapter 4 further applies the introduced topics up until this point to our problem. A new objective function is introduced, in addition to how the optimization will be approached by optimizing each parameter. We will also investigate an analytical approach using Lagrange multipliers in the exponential case.

5. Lastly, Chapter 5 will analyze the tables and visualizations of the numerical results for our optimization problem in a univariate and bivariate case. Changes in parameters and behavior will be analyzed for different scenarios, including symmetrical, asymmetrical, balanced, unbalanced, and changes in dependencies.

CHAPTER 2

Preliminaries

2.1 Insurance concepts

This section will introduce some relevant definitions and terminology regarding the insurance applications of this thesis.

Firstly, an insurance contract is a contractual obligation where an insurer covers the insured's risk, referred to as the *policyholder*, in exchange for a fee, i.e., a *premium*. If the insured experiences a loss, they may file a claim to the insurer for reimbursement if the insurance policy validates the incident/loss. The insurance policy contains the terms and conditions evidenced in an insurance contract agreed upon by the parties (the insurer and the insured)[1].

Risk is the main subject of the insurance policy and states what type of incidents the insurance company covers. Risk is the exposure of an incident that may result in damage or loss of an asset, referred to as an *object*. The insurer determines the value of covering a specific risk X and calculates the corresponding premium exchanged for the risk.

Assume that the risk X is non-negative, in addition to being absolutely continuously distributed. Also, assume that the CDF of X, $F_X(x) = P(X \leq x)$, is strictly increasing whilst the survival function $S_X(x) = 1 - F_X(x) = P(X > x)$ is strictly decreasing.

A portfolio is a collection of insurance contracts that the insurance company classifies. It may be the entirety of insurance contracts, the number of business lines, or the number of product lines offered by the insurance company. We will assume that $X_1, \ldots, X_m$ are the independent non-negative random variables represented from m portfolios or business lines.

Reinsurance is insurance for the insurer, referred to as the *cedent*. The practice consists of insurers protecting their portfolio to reduce the likelihood of reimbursing a high-value claim, and they do this by transferring risk to a secondary insurer, referred to as the *reinsurer*, in exchange for a premium. The cedent reduces the portfolio's value-at-risk by this practice, which may contribute to the company's solvency, especially when dealing with enhanced exposure under larger quantities of assets under their portfolios.

[1] The usage of the terms *insurance contract* and *insurance policy* are commonly interchangeable, despite the definitions.

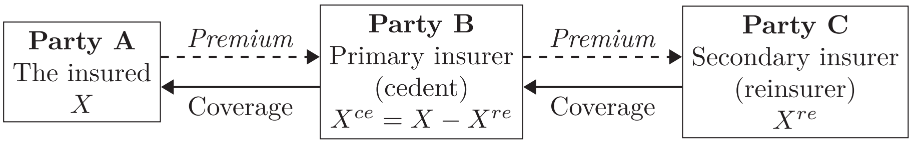


Figure 2.1: Reinsurance concept

2.2 Reinsurance contracts

We will now define the risk covered by the reinsurer as $X^{re} = R(X)$ and the risk covered by the cedent as the retained risk $X^{ce} = I(X) = X - R(X)$. It is preferable to introduce what types of reinsurance contracts there are to choose from, of which there are three:

(i) Stop-loss contract

(ii) Insurance layer contract

(iii) Proportional contract

The *stop-loss contract* is a reinsurance contract in which the cedent covers the risk up to a particular value a before the reinsurer handles the remaining risk $R(X) = \max(X - a, 0)$, where a is known as the *retention limit* of the cedent.

A slightly more complicated but favored approach (in the univariate case) when using the *value-at-risk* measure is the *insurance layer contract*, as shown in the work of Cheung et al. (2014)[Che14]. The risk covered by the reinsurer in an insurance layer contract is presented as an interval, where the reinsurer covers the risk between a and b, where $a < b$, denoted by $R(X)$. The cedent then covers the risk outside the interval $[a, b]$, denoted by $I(X) = X - R(X)$.

Remark 2.2.1 In some instances, the retention limit $b = \infty$ for the reinsurer may arise. Observe that the insurance layer contract $a \times b$ is modified to a stop-loss contract, also known as *Excess of Loss*.

In addition, we have the *proportional contract* (otherwise known as Pro Rata). Although this thesis will not cover it, it is worth mentioning. A proportional contract introduces a fixed percentage $0 \leq c \leq 1$, where the reinsurer and cedent respectively cover the proportion of risk c, denoted by $R(X) = cX$ and $I(X) = (1 - c)X$.

2.2.1 Insurance layer contract

Utilizing an insurance layer contract in the multivariate case, we are presented with the layer $[a_i, b_i]$ and are interested in optimizing the insurance contract using these parameters.

To avoid an issue of deceit, we assume that the loss functions $R_i(x_i)$ and $I_i(x_i)$ satisfy the following conditions in the bivariate case:

(i) $0 \leq R_i(x_i) \leq x_i \;\; \forall \;\; x_i \geq 0,$ both $R_i(x_i)$ and $I_i(x_i)$ are non-decreasing functions.

$$\textit{(ii)} \;\; 0 \leq \;\; R_i(y) - R_i(x) \leq y - x \;\; \forall \;\; 0 \leq x \leq y.$$

There are three possible scenarios when we are presented with this interval. Firstly, if $X_i < a_i$, the reinsurer covers nothing, and the cedent covers all the risk. Secondly, if $a_i < X_i < b_i$, the reinsurer covers up to $X_i - a_i$, and the cedent covers a_i. Lastly, if $X_i > b_i$, then the reinsurer covers the whole interval $b_i - a_i$, and the cedent covers the rest, which would be all the risk up to X_i, not including the interval $[a_i, b_i]$, which would then be $X_i - (b_i - a_i)$. These notations can be derived as follows:

$$R_i(X_i) = \begin{cases} 0 & \text{for } X_i < a_i \\ X_i - a_i & \text{for } a_i \leq X_i \leq b_i \\ b_i - a_i & \text{for } X_i > b_i \end{cases} \tag{2.1}$$

Algorithm 1 Insurance layer contract

1: **input** X_i, a_i, b_i:
2: **return** $R_i(X_i) \leftarrow \min(\max(X_i - a_i, 0), b_i - a_i)$

For $i = 1, \ldots, m$ where $a_i < b_i$. The retained risks for the cedent are denoted as follows: $I_i(X_i) = X_i - R_i(X_i)$, which yields:

$$I_i(X_i) = \begin{cases} X_i & \text{for } X_i < a_i \\ a_i & \text{for } a_i \leq X_i \leq b_i \\ X_i - (b_i - a_i) & \text{for } X_i > b_i \end{cases} \tag{2.2}$$

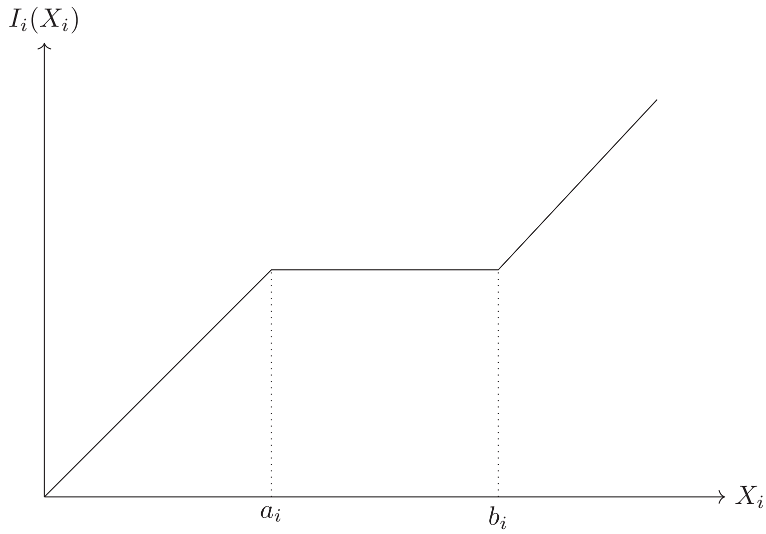


Figure 2.2: Retained risk

2.2.2 Premiums

We have presented the basic idea of retained and reinsured risk and will now introduce the corresponding premium for these transactions as compensation for the reinsurer. When an insurer covers an insured's risk, they must be compensated for the exposure. This compensation is given by the pure premium, which is the expected value of the risk $E[X]$.

Definition 2.2.2 (*Loading*) To avoid a break-even financial result, a loading γ is introduced:

$$\pi = (1+\gamma)E[X_i] \tag{2.3}$$

where $\gamma E[X_i]$ may be referred to as the cost of risk. Reinsurers utilize this pricing strategy too, where θ is used as loading. The price paid by the cedent for the ith contract is then denoted by

$$\pi_{X_i} = (1+\theta)E[R_i(X_i)], \quad i = 1, \ldots, m \tag{2.4}$$

We have obtained the premium from clients to the cedent, which is $(1+\gamma)\sum_{i=1}^{m} E[X_i]$. The premium from the cedent to the reinsurer is $(1+\theta)\sum_{i=1}^{m} E[R_i(X_i)]$. Additionally, we have the total risk $\sum_{i=1}^{m} X_i$ and the reinsured risk $\sum_{i=1}^{m} R_i(X_i)$. These components will be referred to as the *gain* for the cedent for now:

$$G = (1+\gamma)\sum_{i=1}^{m} E[X_i] - (1+\theta)\sum_{i=1}^{m} E[R_i(X_i)] - \sum_{i=1}^{m} X_i + \sum_{i=1}^{m} R_i(X_i) \tag{2.5}$$

Taking the expected gain into account, this can be further deduced:

$$\begin{aligned} E[G] = E\left[(1+\gamma)\sum_{i=1}^{m} E[X_i]\right] - E\left[(1+\theta)\sum_{i=1}^{m} E[R_i(X_i)]\right] \\ - E\left[\sum_{i=1}^{m} X_i\right] + E\left[\sum_{i=1}^{m} R_i(X_i)\right] \end{aligned}$$

$$\Rightarrow E[G] = \gamma \sum_{i=1}^{m} E[X_i] - \theta \sum_{i=1}^{m} E[R_i(X_i)] \tag{2.6}$$

We are given that we have fixed values for γ and θ, which do not weigh on abundant X_i-values. In other words, the price policies are calculated linearly.

In chapter 4, we will discuss the optimization problem that forms the core of this Master's Thesis and introduce the final objective function, where the gain for the cedent will serve as the numerator.

CHAPTER 3

Theoretical background

In order to comprehend the optimal reinsurance for both univariate and multivariate cases and derive the main outcomes of this thesis, this chapter further elaborates on the initial concepts. To prepare for our findings, we will delve into value-at-risk, optimization, Monte Carlo simulation, generalized solutions, and partial derivatives of the total reinsurance premium. Moreover, to interpret the results in chapter 5, we will briefly touch upon the topics of convexity, concavity, and hazard rate.

We need to assess the insurance layer contract which involves intervals, and this involves two parameters in the multivariate case: a_i and b_i. As the number of parameters n increases, the optimization problem becomes increasingly complex, as we will have $2n$ parameters to optimize. Given the high complexity of this calculation, we will focus our analysis on the bivariate case for our multivariate optimization problem.

3.1 Risk measure

When it comes to insurance contracts, insurance companies benefit from the law of large numbers. The vast compilation of insurance contracts provides valuable information to insurance companies, which in turn may be used to produce representative statistics. With this information, insurance companies can employ risk measures, which are mathematical tools that help to assess risk and determine the appropriate amount of assets to keep in reserve. Risk measures are a useful tool for assessing the level of risk associated with an insurance policy or portfolio. They allow you to determine the appropriate amount of assets to keep in reserve by quantifying the level of risk. Essentially, this method calculates the probability of a loss occurring.

Introduce some conditions for the risk X.

$$X_i, \quad i = 1, \ldots, m \text{ is non-negative and absolutely continuously distributed.}$$

For insurance companies, risk measures are, in short, used to manage and mitigate risk to ensure that a suitable premium associated with the policy or portfolio is charged.

3.1.1 Value-at-risk

A commonly used risk measure is value-at-risk. Although value-at-risk lacks the property of subadditivity, which makes it an incoherent risk measure, it benefits from its simplicity and straightforward implementation. In addition, it satisfies regulatory requirements and is an effective tool for measuring capital requirements.

Value-at-risk determines the worst loss over a target horizon within a given confidence interval. In other words, value-at-risk is a probability-based measure of the potential loss. We consider a risk, X, with a given probability distribution. Value-at-Risk is a specific percentile in the probability distribution of the risk. For example, if we choose the upper α percentile in the distribution, it means that there is a probability of α that the risk will be higher than this percentile.

Definition 3.1.1 (*Percentiles*) Percentile q_α defines the amount of which is used for *solvency capital* or *reserve.* [Bøl14, p. 6]

$$P\left(X > q_\alpha\right) = \alpha \tag{3.1}$$

and in our case, we use the value of α as a cut-off percentage for our value-at-risk calculation. The cumulative distribution function of a given data set can be represented by $F_X(q_\alpha) = 1 - \alpha$, with the corresponding survival function $S_X(q_\alpha) = \alpha$, which corresponds to the upper and lower percentiles, respectively.

Example 3.1.2 For instance, a 4% value-at-risk of \$500 over 1 month means that our losses should not exceed \$500 in the period of 1 month with a 96% probability.

This measure of risk is defined relative to some random variable X. The cumulative distribution of X is denoted by $F_X(x) = P(X \leq x)$. The survival function $S_X(x) = 1 - F_X(x) = P(X > x)$ is also introduced. The α-level value-at-risk associated with the risk X is given by $S_X^{-1}(\alpha)$ defined as:

$$V_\alpha[X] = S_X^{-1}(\alpha) = \inf\{x : P(X > x) \leq \alpha\} \tag{3.2}$$

If S_X is strictly decreasing, we have that $S_X^{-1}(\alpha) = r$ if and only if

$$P(X > r) \leq \alpha \leq P(X \geq r)$$

When S_X is strictly decreasing, the following holds:

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

Remark 3.1.3 If S_X is strictly decreasing for all $X \geq 0$, the cumulative distribution function F_X is strictly increasing for all $X \geq 0$. A continuous function is said to be strictly increasing if $F_X(x_1) < F_X(x_2) \ \forall \ x_1, x_2 \in \mathbb{R}$ where $x_1 < x_2$.

Proposition 3.1.4 *Any strictly increasing continuous function ϕ has the following property (monotonicity) [Hus22]:*

$$V_\alpha[\phi(X)] = S_{\phi(X)}^{-1}(\alpha) = \phi(S_X^{-1}(\alpha)) \tag{3.3}$$

Proof. Since ϕ is a strictly increasing continuous function, it follows by (3.2) that (3.3) arises from:

$$\begin{aligned} V_\alpha[\phi(X)] &= \inf\{y : P(\phi(X) > y) \leq \alpha\} \\ &= \inf\{y : P(X > \phi^{-1}(y)) \leq \alpha\} \end{aligned}$$

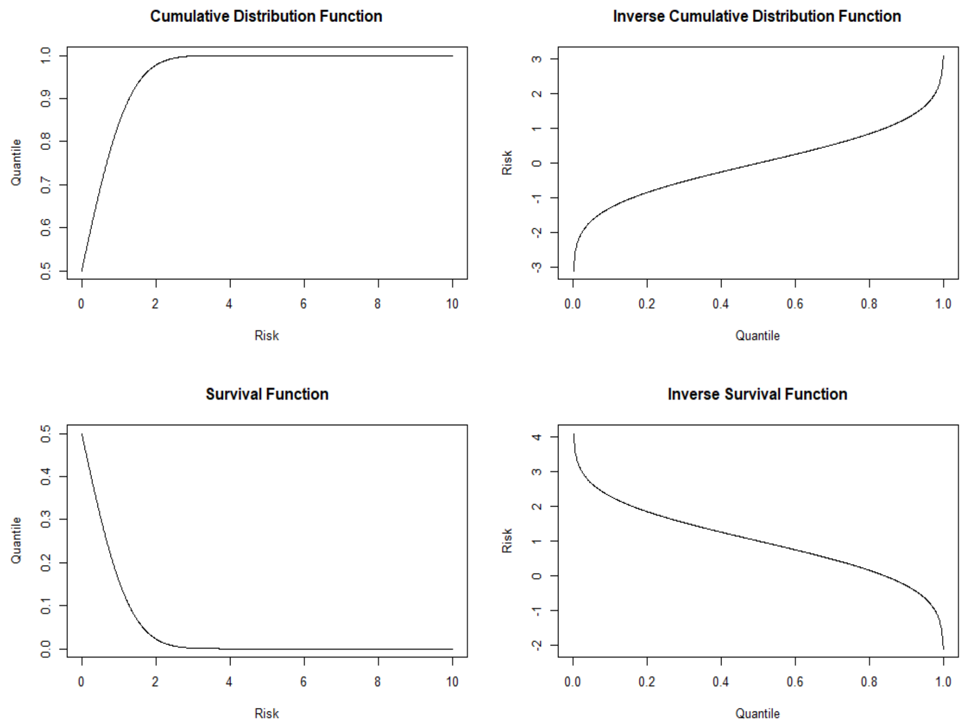


Figure 3.1: CDF and survival function visualizations

Substitute $y = \phi(x)$ and $\phi^{-1}(y) = x$:

$$\begin{aligned} V_\alpha[\phi(X)] &= \inf\{\phi(x) : P(X > x) \leq \alpha\} \\ &= \phi(\inf\{x : P(X > x) \leq \alpha\}) \\ &= \phi(S_X^{-1}(\alpha)) \end{aligned}$$

■

Proposition 3.1.5 *Value-at-risk has the following properties [Ury00, p. 276]: VaR, introduced as V_α, is equivariant to translation, meaning that the translation of input risks results in an equivalent translation of outputs for the risk measure, i.e., linearity:*

$$(i) \qquad V_\alpha[X + c] = V_\alpha[X] + c \tag{3.4}$$

V_α is positively homogeneous, i.e., multiplying a risk with a positive constant does not change the risk measure.

$$(ii) \qquad V_\alpha[cX] = cV_\alpha[X] \tag{3.5}$$

where c is a positive scalar. These properties are special cases of proposition 3.3.

Example 3.1.6 Let $a > 0$ and b be constants, and let X be a random variable, and define $\phi(X) = aX + b$. Then we have that:

$$V_\alpha[aX + b] = V_\alpha[\phi(X)] = \phi(V_\alpha[X]) = aV_\alpha[X] + b$$

3.2 Univariate reinsurance contracts

We will now continue to assume the univariate case and recall the risk covered in an insurance layer insurance contract (2.1), which in the univariate case is:

$$R(X) = \begin{cases} 0 & \text{for } X < a \\ X - a & \text{for } a \leq X \leq b \\ b - a & \text{for } X > b \end{cases} \tag{3.6}$$

for $m = 1$ where $a < b$. Also, recall the retained risk (2.2) for the cedent, denoted as follows in the univariate case $I(X) = X - R(X)$, which yields:

$$I(X) = \begin{cases} X & \text{for } X < a \\ a & \text{for } a \leq X \leq b \\ X - (b - a) & \text{for } X > b \end{cases} \tag{3.7}$$

3.2.1 Partial derivatives total reinsurance premium

Consider that for the total reinsurance premium denoted $E[R(X)]$, we have:

$$\begin{aligned} E[R(X)] &= \int_a^b (x-a) f_X(x) dx + \int_b^\infty (b-a) f_X(x) dx \\ &= \int_a^b x f_X(x) dx - aP(a < X \leq b) + (b-a)P(X > b) \\ &= \int_a^b x f_X(x) dx - aP(X > a) + bP(X > b) \end{aligned}$$

Hence, the derivatives of $E[R(X)]$ yields:

$$\begin{aligned} \frac{\partial E[R(X)]}{\partial a} &= -af_X(a) - P(X > a) + af_X(a) = -P(X > a) \\ \frac{\partial E[R(X)]}{\partial b} &= bf_X(b) + P(X > b) - bf_X(b) = P(X > b) \end{aligned}$$

We can use these results to study the impact of a change in one contract parameter on the total reinsurance premium in the univariate case, while keeping the other parameter constant

3.2.2 Value-at-risk

Recall the survival function $S_X(x) = 1 - F_X(x) = P(X > x)$ introduced in 3.1.1, it follows from the nature of our insurance layer contract that

$$S_{I(X)}(x) = P(I(X) > x) = \begin{cases} P(X > x) & \text{for } \ x < a \\ P(X > x + (b-a)) & \text{for } \ x \geq a \end{cases}$$

$$P(I(X) > x) = \begin{cases} S_X(x) & \text{for } \ x < a \\ S_X(x + (b-a)) & \text{for } \ x \geq a \end{cases}$$

Assume that $\alpha < S_X(a)$ or equivalently $a < S_X^{-1}(\alpha)$, and consider the following case:

$$\alpha \leq S_X(b) \text{ or } b \leq S_X^{-1}(\alpha)$$

The value-at-risk then becomes

$$V_\alpha[I(X)] = S_{I(X)}^{-1}(\alpha) = S_X^{-1}(\alpha) - (b - a)$$

also, consider the following case:

$$\alpha > S_X(b) \text{ or } b > S_X^{-1}(\alpha)$$

where the value-at-risk becomes

$$V_\alpha[I(X)] = S_{I(X)}^{-1}(\alpha) = a$$

Taking both these scenarios into account and combining them, we get the resulting value-at-risk in the univariate case:

$$V_\alpha[I(X)] = S_{I(X)}^{-1}(\alpha) = \max\{a, S_X^{-1}(\alpha) - (b - a)\}$$

It follows that the partial derivatives from this expression that

$$\frac{\partial}{\partial a} V_\alpha[I(X)] = \frac{\partial}{\partial a} a = 1$$

and

$$\frac{\partial}{\partial b} V_\alpha[I(X)] = \frac{\partial}{\partial b} \left(S_X^{-1}(\alpha) - (b - a)\right) = \begin{cases} -1 & \text{for } b < S_X^{-1}(\alpha) \\ 0 & \text{for } b > S_X^{-1}(\alpha) \end{cases}$$

Remark 3.2.1 In the case of $b = S_X^{-1}(\alpha)$, we get that

$$V_\alpha[I(X)] = \max\{a, S_X^{-1}(\alpha) - (b - a)\} = \max\{a, a\} = a$$

showing that $V_\alpha[I(X)]$ is continuous in b.

3.3 Multivariate reinsurance contracts

Proceeding with the thesis in the multivariate case. This involves considering m non-negative random variables, denoted as $X_1, \ldots, X_m$, which represent risks from m different business lines. We will use the bivariate case for illustrations and applied instances to simplify the implementations.

In this section, we will consider the problem of optimizing multivariate reinsurance contracts with respect to value-at-risk, expected costs, and gain for the cedent in preparation for a more advanced objective function that will be introduced in chapter 4. Recall the insurance layer contract for the reinsured (2.1) and the retained risk (2.2):

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

For $i = 1, \dots, m$ where $a_i < b_i$. The retained risks for the cedent are denoted as follows: $I_i(X_i) = X_i - R_i(X_i)$, which yields:

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

Since we cannot bundle our reinsurance contracts, we need to verify why our problem opens up to pooling single risk factors when optimizing multivariate reinsurance contracts. This approach opens up to mitigating an insurer's exposure to risks, improving their ability to fulfill financial obligations in addition to diversifying their portfolio. This will in turn provide coverage for losses that have differently distributed risks that may occur simultaneously or sequentially.

3.3.1 Partial derivatives total reinsurance premium

Consider that for the total reinsurance premium denoted as $\Phi = \sum_{i=1}^{m} \Phi_i = \sum_{i=1}^{m} E[R_i(X_i)]$, we have:

$$\begin{aligned} E[R_i(X_i)] &= \int_{a_i}^{b_i} (x - a_i) f_{X_i}(x)dx + (b_i - a_i)P(X_i > b_i) \\ &= \int_{a_i}^{b_i} x f_{X_i}(x)dx - a_i F_X(b_i) + a_i F_X(a_i) + (b_i - a_i)P(X_i > b_i) \\ &= \int_{a_i}^{b_i} x f_{X_i}(x)dx - a_i(1 - P(X_i > b_i)) \\ &\qquad\qquad + a_i(1 - P(X_i > a_i)) + (b_i - a_i)P(X_i > b_i) \\ &= \int_{a_i}^{b_i} x f_{X_i}(x)dx - a_i P(X_i > a_i) + b_i P(X_i > b_i), \quad i = 1, \dots, m \end{aligned}$$

Continuing, we will inspect the partial derivatives of the premium term. This is useful for analyzing the dependencies and determining the maximum/minimum points for a varying insurance layer contract parameter while holding the other constant to understand how our premium term behaves when either one changes.

$$\begin{aligned} \frac{\partial E[R_i(X_i)]}{\partial a_i} &= \frac{\partial}{\partial a_i} \int_{a_i}^{b_i} x f_{X_i}(x)dx - \frac{\partial}{\partial a_i} a_i \int_{a_i}^{\infty} f_{X_i}(x)dx \\ &= -a_i f_{X_i}(a_i) - P(X_i > a_i) + a_i f_{X_i}(a_i) = -P(X_i > a_i) \\ \frac{\partial E[R_i(X_i)]}{\partial b_i} &= \frac{\partial}{\partial b_i} \int_{a_i}^{b_i} x f_{X_i}(x)dx + \frac{\partial}{\partial b_i} b_i \int_{b_i}^{\infty} f_{X_i}(x)dx \\ &= b_i f_{X_i}(b_i) + P(X_i > b_i) - b_i f_{X_i}(b_i) = P(X_i > b_i) \end{aligned}$$

Since our parameters a_i and b_i are of the first order, the parameter not being partially differentiated is treated as a constant and therefore cancels out. Hence, the partial derivative of the premium term with respect to either a_i or b_i only depends on the respective parameter.

3.3.2 Value-at-risk as the objective function

The total risk covered by the cedent is given by:

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

where the first term is the retained risk term, and the latter is referred to as the premium term, i.e., the price paid by the cedent for the ith contract.
Recall the α-level value-at-risk (3.3)

$$V_\alpha[\phi(\mathbf{X})] = S^{-1}_{\phi(\mathbf{X})}(\alpha) = \phi(S^{-1}_{\mathbf{X}}(\alpha))$$

By (3.3), we have that the resulting α-level value-at-risk is given by:

$$\begin{aligned} C_0 = V_\alpha[\phi(\mathbf{X})] &= V_\alpha\left[\sum_{i=1}^{m} I_i(X_i) + (1+\theta)\sum_{i=1}^{m} E[R_i(X_i)]\right] \\ &= S^{-1}_{\sum_{i=1}^{m} I_i(X_i)+(1+\theta)\sum_{i=1}^{m} E[R_i(X_i)]}(\alpha) \end{aligned}$$

Utilizing the property of value-at-risk, we get the following:

$$C_0 = V_\alpha[\phi(\mathbf{X})] = S^{-1}_{\sum_{i=1}^{m} I_i(X_i)}(\alpha) + (1+\theta)\sum_{i=1}^{m} E[R_i(X_i)] \tag{3.8}$$

where the premium term is a constant.

3.3.3 Minimizing value-at-risk

We wish to minimize $C_0 = V_\alpha[\phi(\mathbf{X})]$. We will assume a multivariate case for the generalized concept with an expected risk premium $E[R_i(X_i)]$.

We denote three sets for the retained risk, according to the lower bound a_i:

$$\mathcal{A} = \{\mathbf{x} : \sum_{i=1}^{m} I_i(x_i) < \sum_{i=1}^{m} a_i\} \tag{3.9}$$

$$\mathcal{B} = \{\mathbf{x} : \sum_{i=1}^{m} I_i(x_i) = \sum_{i=1}^{m} a_i\} \tag{3.10}$$

$$\mathcal{C} = \{\mathbf{x} : \sum_{i=1}^{m} I_i(x_i) > \sum_{i=1}^{m} a_i\} \tag{3.11}$$

Figure 3.2 illustrates the bivariate case, i.e., $m = 2$. Observe that subset $\mathcal{B}$ includes the boundaries of the rectangle in addition to the borderline between $\mathcal{A}$ and $\mathcal{C}$. Since we assume that S_{X_i} are strictly decreasing for all i, it follows that $P(\mathbf{X} \in \mathcal{B} \cup \mathcal{C})$ and $P(\mathbf{X} \in \mathcal{C})$ are also strictly decreasing in a_i for all i.

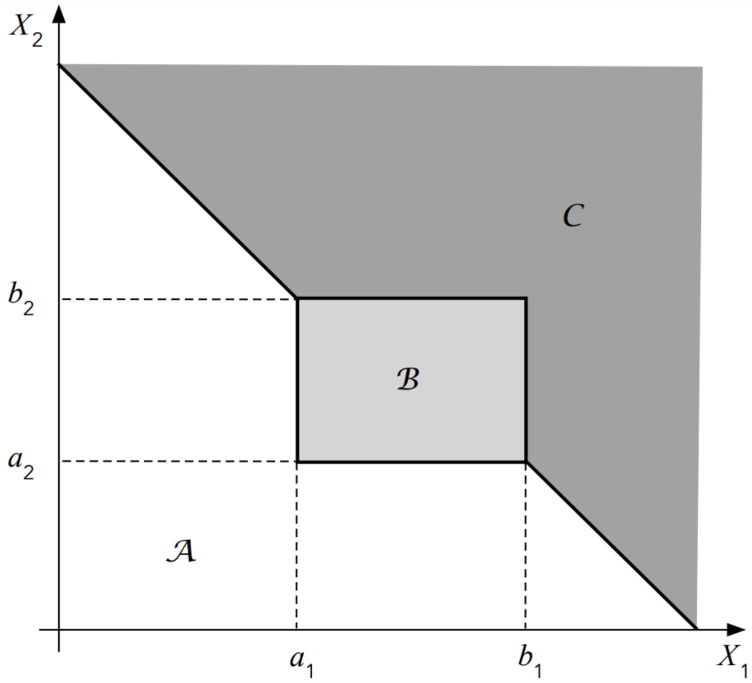


Figure 3.2: The sets $\mathcal{A}$, $\mathcal{B}$ and $\mathcal{C}$ for a_i, b_i, $i = 1, 2$

Proposition 3.3.1 *Assume that the contract parameters are chosen such that*

$$P(\mathbf{X} \in \mathcal{B} \cup \mathcal{C}) \geq \alpha$$

$$P(\mathbf{X} \in \mathcal{C}) \leq \alpha$$

Then

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

We know that if S_X is strictly decreasing, we have that $S_X^{-1}(\alpha) = r$ if and only if

$$P(X > r) \leq \alpha \leq P(X \geq r)$$

which results in the following for our retained risk:

$$P\left(\sum_{i=1}^{m} I_i(X_i) \geq \sum_{i=1}^{m} a_i\right) = P(\mathbf{X} \in \mathcal{B} \cup \mathcal{C}) \geq \alpha$$

Since the set $\mathcal{B}$ includes the boundaries of set $\mathcal{C}$ for our parameters. Meanwhile, we also have that for the set $\mathcal{C}$ that

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

In short, this will simplify the α-level to:

$$P(\mathbf{X} \in \mathcal{C}) \leq \alpha \leq P(\mathbf{X} \in \mathcal{B} \cup \mathcal{C})$$

which can equivalently be written as:

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

where we again utilize the properties of the introduced survival function, more specifically:

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

Similarly, one can also denote that $P(\mathbf{X} \in \mathcal{C})$ insinuates

$$S^{-1}_{\sum_{i=1}^{m} I_i(X_i)}(\alpha) = \sum_{i=1}^{m} a_i \tag{3.12}$$

by the indistinguishable property $S_X^{-1}(\alpha) = r$. Which is equivalent to the results produced in [HC20].

Theorem 3.3.2 *Assume that $a_1^*, b_1^*, \ldots, a_m^*, b_m^*$ are optimal contract parameters, and that*

$$P\left(\bigcap_{i=1}^{m} X_i > a_i^*\right) \geq \alpha \tag{3.13}$$

Then the following conditions must hold:

$$a_i^* = S_{X_i}^{-1}\left(\frac{1}{1+\theta}\right), \quad i = 1, \ldots, m \tag{3.14}$$

in addition to:

$$P(\mathbf{X} \in \mathcal{C}) = \alpha \tag{3.15}$$

Proof. The proof is given in [HC20]. ■

Observe that $S_{X_i}^{-1}\left(\frac{1}{1+\theta}\right) = a_i$ satisfies

$$S_{X_i}(a_i) = \frac{1}{1+\theta}, \quad i = 1, \ldots, m$$

where we denote $A_0 = (1+\theta)^{-1}$. Hence, we have that all observed values of a_i belong in the same percentile with respect to A_0. Thus, the optimal value A_0 may be utilized to find the corresponding optimal values for $a_1, \ldots, a_m$:

$$a_i^* = S_{X_i}^{-1}(A_0), \quad i = 1, \ldots, m$$

Introduce the same application for the other parameter $b_i, \ \ i = 1, \ldots, m$:

$$B_i = S_{X_i}(b_i) = P(X_i > b_i), \quad i = 1, \ldots, m$$

which is also subject to the solution of the optimization problem with respect to $B_1, \ldots, B_m$ instead of $b_1, \ldots, b_m$. The optimal values $B_1^*, \ldots, B_m^*$ may also be utilized to find the corresponding optimal values for $b_1, \ldots, b_m$:

$$b_i^* = S_{X_i}^{-1}(B_i^*), \quad i = 1, \ldots, m$$

3.4 Optimization problem

We must optimize our univariate and multivariate reinsurance contracts, and do this with a combination of maximizing our gain in addition to minimizing our value at risk. This is done by using a desired objective function with constraint to an α-level risk that is acceptable for our observations to be sampled from.

3.4.1 Optimal contract parameters

Theorem 3.4.1 *Recall (3.14) from the last chapter:*

$$a_i^* = S_{X_i}^{-1}\left(\frac{1}{1+\theta}\right) = S_{X_i}^{-1}(A_0), \quad i = 1, \dots, m$$

Assume that the values $a_1^*, \dots, a_m^*$ satisfies (3.15). The remaining optimal contract parameters $b_1^*, \dots, b_m^*$ may be found, as we know they already exist, as shown in the proof to the theorem (3.3.2). The optimal contract parameters may be found by solving the following optimization problem:

$$\begin{aligned} &\text{minimize} && \sum_{i=1}^{m} E[R_i(X_i)] \\ &\text{subject to} && P(\mathbf{X} \in \mathcal{C}) = \alpha \end{aligned} \tag{3.16}$$

with respect to $b_1, \dots, b_m$

Proof. The proof is given in [Hus22]. ■

3.4.2 Expected reinsurance expenses

Consider using the expected risk given in (3.16) as an example for an objective function, $\sum_{i=1}^{m} E[R_i(X_i)]$. Denote f_{X_i} as the density of X_i, $\ i = 1, \dots, m$. Let $\Phi = \sum_{i=1}^{m} E[R_i(X_i)]$. For an insurance layer contract $i = 1, \dots, m$ we know that:

$$\Phi_i = \int_{a_i}^{b_i} x f_{X_i}(x)dx - a_i P(X_i > a_i) + b_i P(X_i > b_i), \quad i = 1, \dots, m$$

where $\Phi_i = E[R_i(X_i)]$. Numerical integration will be used to compute Φ_i, $\ i = 1, \dots, m$ as functions of B_i, $\ i = 1, \dots, m$ [Hus22].

3.4.3 Constraint

Consider the constraint $P(\mathbf{X} \in \mathcal{C}) = \alpha$ that will assist in determining the set of B_i-values. Due to the irregular shape of the subset $\mathcal{C}$, an analytical solution to this problem is difficult. To bypass this issue, a simulation on the distribution of X_i will iterate the values of B_i until the subset $\mathcal{C}$ obtains a specific fraction of the simulations.

Recall the figure 3.2, of which it is easy to see that

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

Hence, we have that

$$P(\mathbf{X} \in \mathcal{C}) \leq P\left(\bigcup_{i=1}^{m} X_i > b_i\right) \leq \sum_{i=1}^{m} B_i$$

Indicating that the upper bound $B_1, \ldots, B_m$ is quite good due to the typically small numbers. Assume that we have generated N samples $\mathbf{X}_1, \ldots, \mathbf{X}_N$ from a given distribution of $\mathbf{X}$, using Monte Carlo simulation. We can estimate $p_{\mathcal{C}} = P(\mathbf{X} \in \mathcal{C})$ for given values of $B_1, \ldots, B_m$ by computing the resulting desired fraction of samples in the set $\mathcal{C}$ equal to a given α-level.

To obtain a stable estimate of the constraint set, we need a very large N, given that many of the observed risk $\mathbf{X}$ will fall outside the set $\mathcal{C}$, given the environment we are working in. A sufficient N will counter this effect by sampling an adequate amount of B_i-values affected in the desired set.

Assume that we can obtain a set $\mathcal{D}$ such that $\mathcal{C} \subset \mathcal{D}$ for all B_i that fall within the event $\{\mathbf{X} \in \mathcal{C}\}$, and such that $p_{\mathcal{D}} = P(\mathbf{X} \in \mathcal{D})$. Let $p_{\mathcal{C}|\mathcal{D}} = P(\mathbf{X} \in \mathcal{C} | \mathbf{X} \in \mathcal{D})$. Since $\mathcal{C} \subset \mathcal{D}$, we get:

$$p_{\mathcal{C}} = p_{\mathcal{C}|\mathcal{D}} \cdot p_{\mathcal{D}}$$

Continue by generating N samples $\mathbf{X}_1, \ldots, \mathbf{X}_N$ from the conditional distribution of $\mathbf{X}$ given $\mathbf{X} \in \mathcal{D}$. Estimate $p_{\mathcal{C}|\mathcal{D}}$ by

$$\hat{p}_{\mathcal{C}|\mathcal{D}} = \frac{1}{N} \sum_{i=1}^{N} I(\mathbf{X}_k \in \mathcal{C})$$

The unconditional probability $p_{\mathcal{C}}$ is estimated by:

$$\hat{p}_{\mathcal{C}} = \hat{p}_{\mathcal{C}|\mathcal{D}} \cdot p_{\mathcal{D}}$$

Further assume that $\mathbf{X}_1, \ldots, \mathbf{X}_N$ are generated by transforming independent uniformly distributed vectors $\mathbf{U}_1, \ldots, \mathbf{U}_N$:

$$\mathbf{X}_k = \psi(\mathbf{U}_k) = \left(F_1^{-1}(U_{1k}), \ldots, F_m^{-1}(U_{mk})\right), \quad k = 1, \ldots, N$$

where ψ is strictly increasing for each k, and the vectors $\mathbf{U}_1, \ldots, \mathbf{U}_N$ are sampled uniformly from the set $\mathcal{D}'$ given by

$$\mathcal{D}' = \{\mathbf{u} : 1 - \Delta < u_i < 1, \;\; i = 1, \ldots, m\}$$

Finally, let:

$$\mathcal{D} = \psi(\mathcal{D}') = \{\mathbf{x} = \psi(\mathbf{u}) : \mathbf{u} \in \mathcal{D}'\}$$

Δ must be chosen to be as small as possible, depending on the joint distribution of the risks $\mathbf{X}$, although $\Delta = 2\alpha$ is usually sufficient. In addition, a large enough Δ is preferable to observe enough samples $\mathcal{C} \subset \mathcal{D}$. The transformation ψ can be compiled from the inverse distribution functions of $X_1, \ldots, X_m$. In order to improve the stability of the probability estimates $\hat{p}_{\mathcal{C}}$ and $\hat{p}_{\mathcal{C}|\mathcal{D}}$, it is necessary to verify that $\mathbf{X_1}, \ldots, \mathbf{X}_N$ becomes distributed according to the conditional distribution of $\mathbf{X}$ given $\mathbf{X} \in \mathcal{D}$.

Note that

$$\mathcal{D}' = \{\mathbf{u} : 1 - \Delta < u_i < 1, \;\; i = 1, \ldots, m\} = [0,1]^m \backslash \mathcal{E}'$$

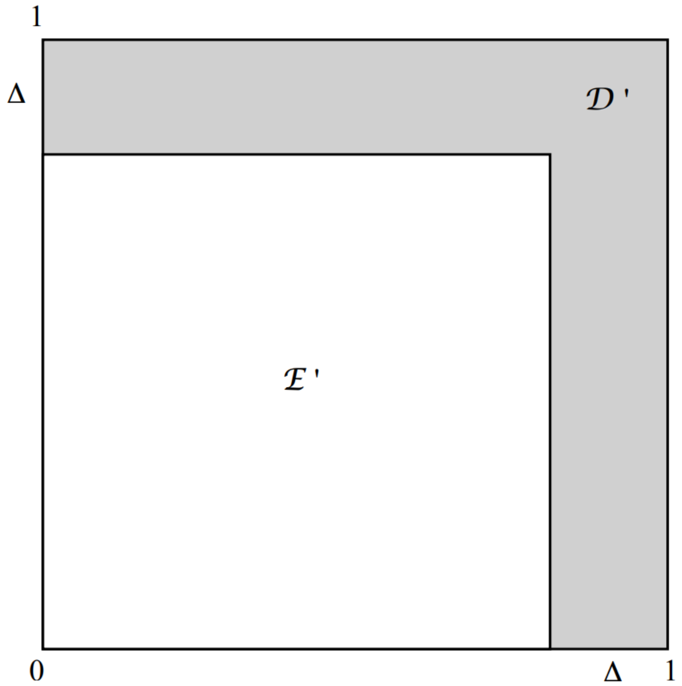


Figure 3.3: $\mathcal{D}' = \{\mathbf{u} : 1 - \Delta < u_i < 1, \quad i = 1, 2\}$

where

$$\mathcal{E} = \{\mathbf{u} : 0 < u_i < 1 - \Delta, \quad i = 1, \ldots, m\}$$

Hence, we know that the probability $p_{\mathcal{D}}$ is known because we have that

$$p_{\mathcal{D}} = P(\mathbf{X}_k \in \mathcal{D}) = P(\mathbf{U}_k \in \mathcal{D}') = 1 - (1 - \Delta)^m$$

Figure 3.4 shows a scatter plot where the blue observations represent simulations located in the $\mathcal{A}$- or $\mathcal{B}$-sets, while the red observations represent simulations located in the $\mathcal{C}$-set. The two scatterplots correspond to two extreme cases regarding the b-values. In the first case, b_1 is set to its minimum value, meaning that B_1 obtains its maximum value. At the same time, b_2 becomes maximal (typically infinite), and B_2 becomes minimal (typically 0). In the second scatterplot, the opposite is true. The purpose of the scatterplots is merely to verify whether importance sampling works, i.e., whether the blank area resulting from importance sampling overlaps with the $\mathcal{C}$-set with a corresponding Δ-value. If importance sampling works for these cases, it will also work for the less extreme cases. For example, in cases with equal risk distributions, one plot will be an approximate reflection of the other around the line $B_1 = B_2$, as illustrated in Figure 3.4, where the risks are sampled from $X_i \sim \text{Lognormal}(50, 50)$.

Although Monte Carlo serves beneficial results when using random sampling to yield numerical results, a disadvantage is that if the sample size m is insufficient, some values of X_i may not be sampled. For instance, these values may be very large or small, i.e., unlikely, at the tail of our distribution. If Monte Carlo fails to sample these values, the approximation of these numerical results

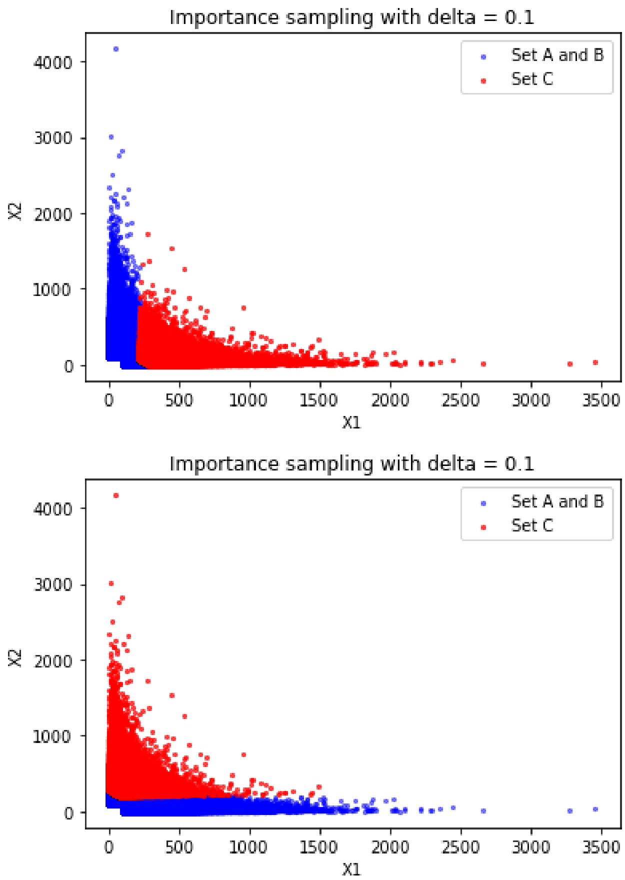


Figure 3.4: Verification of importance sampling

may be poor. To counteract this effect, we introduce *importance sampling*. Importance sampling is a Monte Carlo method used to predict the probability of rare events. The usual Monte Carlo method requires large sample sizes to be representative and may be computationally intensive for an algorithm [She19].

Importance sampling draws $X_i^*, \ \ i = 1, \ldots, m$ from a different distribution that assigns a higher probability to rare events. These observations are then down-weighted for the final numerical results to reduce bias while still representing these values. This approach uses *variance reduction* as a procedure to increase the precision of our estimates.

Example 3.4.2 Assume the approach of importance sampling when $m = 2$, then we have that the set $\mathcal{E}$, is defined by

$$\mathcal{E}' = \{\mathbf{u} : 0 < u_i < 1 - \Delta, \ \ i = 1, 2\}$$

whilst $\mathcal{D}' = [0, 1]^2 \backslash \mathcal{E}'$, as seen in the figure 3.3.

Consider the condition $m = 2$ from example 3.4.2, and assume the problem of sampling the vector $\mathbf{U} = (U_1, U_2)$ which is bivariate and uniformly distributed

$R[0,1]^2$ conditioned on that $\mathbf{U} \in \mathcal{D}'$, where

$$\mathcal{D}' = \{\mathbf{u} : 1-\Delta < u_i < 1, \;\; i = 1 \wedge 2\}, \;\; \text{where} \;\; 0 < \Delta < 1$$

Partition the set $\mathcal{D}'$ into two disjoint subsets:

$$\mathcal{D}'_1 = \{\mathbf{u} : 0 < u_1 < 1, \;\; 1-\Delta < u_2 < 1\}$$
$$\mathcal{D}'_2 = \{\mathbf{u} : 1-\Delta < u_1 < 1, \;\; 0 < u_2 < 1-\Delta\}$$

As a result of $\mathbf{U}$ being uniformly distributed on $[0,1]^2$, we have that the probabilities of the disjoint sets $\mathcal{D}'_1$ and $\mathcal{D}'_2$ are equal to their respective areas:

$$P(\mathbf{U} \in \mathcal{D}'_1) = \Delta$$
$$P(\mathbf{U} \in \mathcal{D}'_2) = \Delta(1-\Delta)$$

of which it follows that

$$\begin{aligned} P(\mathbf{U} \in \mathcal{D}') &= P(\mathbf{U} \in \mathcal{D}'_1) + P(\mathbf{U} \in \mathcal{D}'_2) \\ &= \Delta + \Delta(1-\Delta) = 2\Delta - \Delta^2 = 1 - (1-\Delta)^2 = p_{\mathcal{D}} \end{aligned}$$

Continuing, we can simplify for our conditional distribution of $\mathbf{U}$ given $\mathbf{U} \in \mathcal{D}'$ is uniform on $\mathcal{D}'$. Draw a random number U_0 sampled uniformly from the interval $[0, 2-\Delta]$. If $U_0 \leq 1$, $\mathbf{U}$ is sampled uniformly from the set $\mathcal{D}'_1$, implying that:

$$P(\mathbf{U} \in \mathcal{D}'_1 | \mathbf{U} \in \mathcal{D}') = \frac{P(\mathbf{U} \in \mathcal{D}'_1)}{P(\mathbf{U} \in \mathcal{D}')} = \frac{1}{2-\Delta}$$

and similarly for $U_0 > 1$, where we sample $\mathbf{U}$ uniformly on the set $\mathcal{D}'_2$, implying that:

$$P(\mathbf{U} \in \mathcal{D}'_2 | \mathbf{U} \in \mathcal{D}') = \frac{P(\mathbf{U} \in \mathcal{D}'_2)}{P(\mathbf{U} \in \mathcal{D}')} = \frac{1-\Delta}{2-\Delta}$$

Algorithm 2 Uniform sampling in the bivariate case algorithm

```
def bi_uniform_D(Δ):
    u_0 = uniform(0, 2 − Δ)
    if u_0 ≤ 1:
        u_1 = uniform(0, 1)
        u_2 = uniform(1 − Δ, 1)
    if u_0 > 1:
        u_1 = uniform(1 − Δ, 1)
        u_2 = uniform(0, 1 − Δ)
    return u_1, u_2
```

3.4.4 Convexity and concavity

Hazard rate is a method of predicting the time until a specific event occurs. In this case, it measures similarities between the propensity properties convexity and concavity with respect to the density function's relation to the survival function.

Definition 3.4.3 (*Convexity*) Let $S \subset X$ be a convex set. A function $f : S \longrightarrow \mathbb{R}^+$ is *convex* if

$$f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda) f(y), \;\; \forall \;\; x, y \in S \tag{3.17}$$

Definition 3.4.4 (*Concavity*) Let $S \subset X$ be a convex set, a function $f : S \longrightarrow \mathbb{R}^+$ is *concave* $\forall\ x, y \in S$ if

$$f(\lambda x + (1-\lambda)y) \geq \lambda f(x) + (1-\lambda) f(y), \quad \forall\ x, y \in S \tag{3.18}$$

Introduce the super- and sublevel sets of Φ 3.4.2 expressed in terms of $\mathbf{B} = B_i,\ \ i = 1, \dots, m$.

Definition 3.4.5 *(Level set)* A set of a real multivariate function where the function takes on a value c is known as a *level set*:

$$L_c(f) = \{\mathbf{x} \mid f(\mathbf{x}) = c\}, \text{ where } \mathbf{x} = (x_1, \dots, x_m)$$

or in our case:

$$L_c(\Phi) = \{\mathbf{B} \in [0,1]^m : \Phi(\mathbf{B}) = c\}$$

In simple terms, it means where the plane of $\mathbf{B}$ and the graph $\Phi(\mathbf{B})$ intersects.

Let

$$L_c^+(\Phi) = \{\mathbf{B} \in [0,1]^m : \Phi(\mathbf{B}) \geq c\} \tag{3.19}$$

$$L_c^-(\Phi) = \{\mathbf{B} \in [0,1]^m : \Phi(\mathbf{B}) \leq c\} \tag{3.20}$$

be the superlevel and sublevel sets of the function Φ relative to c, respectively.

The sets denoted by $L_c^+(\Phi)$ and $L_c^-(\Phi)$ are known as the superlevel and sublevel sets of a function Φ with respect to the level c. A function is considered quasiconvex when all of its sublevel sets are convex, while a function is considered quasiconcave when all of its superlevel sets are convex.

Proposition 3.4.6 *If $\Phi_1(B_1), \dots, \Phi_m(B_m)$ are convex functions, then Φ is a quasiconvex function of $\mathbf{B}$. If $\Phi_1(B_1), \dots, \Phi_m(B_m)$ are concave functions, then Φ is a quasiconcave function of B.*

Proof. We will prove the latter statement of proposition (3.4.6). Assume that $\Phi_1(B_1), \dots, \Phi_m(B_m)$ are concave functions, and let $\mathbf{B}^{(j)} = (B_1^{(j)}, \dots, B_m^{(j)}) \in L_c^+(\Phi)$, where $j = 1, 2$. We have that for all concave functions that any $\lambda \in [0,1]$, we get $\mathbf{B} = \lambda \mathbf{B}^{(1)} + (1-\lambda)\mathbf{B}^{(2)} \in L_c^+(\Phi)$. We know that

$$\Phi_i(\lambda B_i^{(1)} + (1-\lambda) B_i^{(2)}) \geq \lambda \Phi_i(B_i^{(1)}) + (1-\lambda)\Phi_i B_i^{(2)}, \ \ i = 1, \dots, m$$

Hence, for $i = 1, \dots, m$

$$\begin{aligned}
\Phi(\mathbf{B}) &= \Phi(\lambda \mathbf{B}^{(1)} + (1-\lambda)\mathbf{B}^{(2)}) \\
&= \sum_{i=1}^{m} \Phi_i(\lambda B_i^{(1)} + (1-\lambda) B_i^{(2)}) \geq \sum_{i=1}^{m} \lambda \Phi_i(B_i^{(1)}) + (1-\lambda)\Phi_i(B_i^{(2)}) \\
&= \lambda \Phi(\mathbf{B}^{(1)}) + (1-\lambda)\Phi(\mathbf{B}^{(2)}) \geq \lambda c + (1-\lambda)c = \lambda c + c - \lambda c = c
\end{aligned}$$

We then have that $\mathbf{B} \in L_c^+(\Phi)$ is concave. The claim that $L_c^+(\Phi)$ is convex is proved in a completely similar way. [Hus22] ∎

3.4.5 Hazard rate

Now for the relationship between hazard rate and convexity/concavity. Utilizing the expressions of $\Phi_1, \dots, \Phi_m$, we have that:

$$\frac{\partial \Phi_i}{\partial b_i} = P(X_i > b_i), \quad i = 1, \dots, m$$

Deriving the partial derivatives with respect to $B_1, \dots, B_m$, we get that:

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

where the partial derivative is increasing in B_i, and hence Φ_i convex. Equivalently, the following relation is also increasing in B_i:

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

which also yields a convex function Φ_i. Further, substitute $B_i = S_{X_i}(x)$. Since B_i is a decreasing function of x, we get that if the function Φ_i

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

is decreasing in x, then it is convex. Similarly, if the function Φ_i increases in x, it is concave. This is the *hazard rate* of the distribution of the risks X_i. [Hus22]

Theorem 3.4.7 *If the risks $X_1, \dots, X_m$ have decreasing hazard rates, then Φ is a quasiconvex function of $\mathbf{B}$. If $X_1, \dots, X_m$ have increasing hazard rates, then Φ is a quasiconcave function of $\mathbf{B}$. The ratio $\frac{f_{X_i}(x)}{S_{X_i}(x)}$ is the hazard rate of the distribution of X_i.*

3.5 Generalized solutions of the optimization problem

To determine the best amount of risk to transfer from an insurer to a reinsurer, optimization can be used in a general manner. However, since there may not be a single solution, a more comprehensive approach is needed. We will use a developed optimization method that considers a range of solutions, providing a nuanced and flexible answer to the question of risk. Our approach takes into account simulated observations, distribution types, solution types, and methods.

3.5.1 Balanced solutions

Recall A_0 to be the common value of $P(X_i > a_i)$ by subsection (3.3.3), being the common probability of all risks:

$$A_0 = S_{X_i}(a_i), \quad i = 1, \dots, m$$

and correspondingly, for the solution of b_i's, to reduce the set of possible combinations which satisfy $P\left(\sum_{i=1}^m I_i(X_i) > \sum_{i=1}^m a_i\right) = \alpha$. This is chosen such that

$$B = S_{X_i}(b_i), \quad i = 1, \dots, m$$

with the resulting values of the contract parameters:

$$b_i = S_{X_i}^{-1}(B), \;\; i = 1, \dots, m$$

It is preferable to produce a Monte Carlo simulation on the joint distribution of $X_i, \;\; i = 1, \dots, m$ to find a value B, rather than solving this problem analytically due to the varying space of the subset $\mathcal{C}$. Due to the nature of our survival function, we can see that an increasing value of B yields a decreasing value of $b_i, \;\; i = 1, \dots, m$. Also, since the risks $X_i, \;\; i = 1, \dots, m$ are assumed to be absolutely continuously distributed, the following 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 the value B. Estimating a value B_L such that

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

provides a lower bound on the correct value of B. Correspondingly, estimating a value B_U such that

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

provides an upper bound on the correct value of B.

Continuing, we will estimate the bounds of B_L and B_U. We have that for all the risks greater than b_i that:

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

Hence,

$$\begin{aligned} P\left(\sum_{i=1}^{m} I_i(X_i) > \sum_{i=1}^{m} a_i\right) &= P(\mathbf{X} \in \mathcal{C}) \leq P\left[\bigcup_{i=1}^{m} (X_i > b_i)\right] \\ &= 1 - P\left[\bigcup_{i=1}^{m} (X_i \leq b_i)\right] \\ &= 1 - (1 - B)^m \end{aligned}$$

We also have that if $B \in [0, 1]$ is such that $1 - (1 - B)^m = \alpha$, it follows by $1 - (1 - B)^m$ that

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

Let B_L denote this B-value since it is smaller than the correct B-value, denoted by

$$1 - (1 - B_L)^m = \alpha \Rightarrow B_L = 1 - \sqrt[m]{1 - \alpha}$$

Further, we must estimate an upper bound B_U for B. Similarly to the lower bound value B_L, we will use the definitions of our sets, of which we will focus on the a_i values lower than the risk X_i:

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

Hence,

$$\begin{aligned}
P\left(\sum_{i=1}^{m} I_i(X_i) > \sum_{i=1}^{m} a_i\right) &= P(\mathbf{X} \in \mathcal{C}) \geq P\left(\left[\bigcap_{i=1}^{m}(X_i > a_i)\right] \setminus \mathcal{B}\right) \\
&= P\left(\bigcap_{i=1}^{m}(X_i > a_i)\right) - P\left(\bigcap_{i=1}^{m}(a_i \leq X_i \leq b_i)\right) \\
&= A^m - (A - B)^m
\end{aligned}$$

Correspondingly,we have that if $B \in [0,1]$ is such that $A_0^m - (A_0 - B)^m = \alpha$, it follows by $A_0^m - (A_0 - B)^m$ that:

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

Let B_U denote this B-value since it is larger than the correct B-value, denoted by:

$$A_0^m - (A_0 - B)^m = \alpha \Rightarrow B_U = A_0 - \sqrt[m]{A_0^m - \alpha}$$

by using the *bisection method*, determining B is simplified such that the fraction of the simulated observations that belong to $\mathcal{C}$ is approximately α.

Remark 3.5.1 The difference between the upper bound B_U and the lower bound B_L is typically small.

Example 3.5.2 Assume that we have the values $m = 2$, $\alpha = 0.05$ and $\theta = \frac{1}{9}$. Then we have that

$$A_0 = \frac{1}{1 + \frac{1}{9}} = \frac{9}{10} = 0.9$$

and

$$\begin{aligned}
B_L &= 1 - \sqrt{1 - \alpha} = 1 - \sqrt{0.95} = 0.02532 \\
B_U &= A_0 - \sqrt{A_0^2 - \alpha} = 0.9 - \sqrt{0.81 - \alpha} = 0.02822
\end{aligned}$$

The bisection method will converge rapidly due to the small difference between our lower and upper bounds.

Definition 3.5.3 (*Bisection method*) The *bisection method* is a root-finding method for consistently bisecting an interval and extracting the subinterval where the function changes sign, implying it must contain a root. [Mør17, p. 242]

Algorithm 3 Bisection method

$a_0 = a$
$b_0 = b$
for $i = 1, \dots, m$
  $n_{i-1} = (a_i + b_i)/2$
  if $f(n_{i-1}) == 0$
    $a_i = b_i = n_{i-1}$
  if $f(a_{i-1})f(a_{n-1}) < 0$
    $a_i = a_{i-1}$
    $b_i = n_{i-1}$
  else
    $a_i = n_{i-1}$
    $b_i = b_{i-1}$
$n_m = (a_m + b_m)/2$

In short, balanced solutions allows us to reduce the large number of combinations that are produced by the equation

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

by assigning all possible B_i values to be equal. This does, however, not guarantee a globally optimal solution, which is why unbalanced, although more computationally heavy, may be more beneficial to use in some extreme cases.

3.5.2 Unbalanced solutions

Since an optimal balanced solution cannot be guaranteed to be optimal globally, an unbalanced approach may be proficient. Enable this approach by introducing $B_i = P(X_i > b_i) = S_{X_i}(b_i),\ \ i = 1, 2,$ for the case $m = 2$. A smaller m is preferable since a large m may be too computationally heavy and time-consuming.

Let B_1 be some sufficient number. Determine B_2 such that $P\left(\sum_{i=1}^{m} I_i(X_i) > \sum_{i=1}^{m} a_i\right) = \alpha$ holds. For Monte Carlo simulation on the distribution of X_1 and X_2, B_2 will be iterated until the subset $\mathcal{C}$ obtains the preferable fraction of the simulations. Utilizing the approach we did for balanced solutions, it is observable that for a given B_1. Continue by calculating the bounds for B_2.

We know that when $\mathcal{C} \subseteq (X_1 > b_1) \cup (X_2 > b_2)$ we have that

$$\begin{aligned} P((X_1, X_2) \in \mathcal{C}) &\leq P[(X_1 > b_1) \cup (X_2 > b_2)] \\ &= 1 - P[(X_1 > b_1) \cap (X_2 > b_2)] \\ &= 1 - (1 - B_1)(1 - B_2) \end{aligned}$$

Solve for the lower bound B_L for B_2. We know that $B_2 \in [0, 1]$ such that $1 - (1 - B_1)(1 - B_2) = \alpha$, whereas the probability for the retained risks $X_i,\ \ i = 1, 2$ being larger than the respective a_i's, we are left with a B_2 value less than the actual B_2 value:

$$1 - (1 - B_1)(1 - B_L) = \alpha \Rightarrow -(1 - B_L) = \frac{\alpha - 1}{1 - B_1}$$

$$\Rightarrow B_L = 1 - \frac{1-\alpha}{1-B_1} = \frac{\alpha - B_1}{1-B_1}$$

For the corresponding upper bound value, we have that

$$[(X_1 > b_1) \cup (X_2 > b_2)] \cap [(X_1 > a_1) \cap (X_2 > a_2)] \subseteq \mathcal{C}$$

and similarly:

$$\begin{aligned}
&P((X_1, X_2) \in \mathcal{C}) \geq P\left([(X_1 > b_1) \cup (X_2 > b_2)] \cap [(X_1 > a_1) \cap (X_2 > a_2)]\right) \\
&= P((X_1 > b_1) \cup (X_2 > b_2) | (X_1 > a_1) \cap (X_2 > a_2)) \cdot P((X_1 > a_1) \cap (X_2 > a_2)) \\
&= (1 - P((X_1 > b_1) \cup (X_2 > b_2) | (X_1 > a_1) \cap (X_2 > a_2))) \cdot P((X_1 > a_1) \cap (X_2 > a_2)) \\
&= \left(1 - \frac{(A - B_1)(A - B_U)}{A^2}\right) \cdot A^2 = A^2 - (A - B_1)(A - B_U)
\end{aligned}$$

Similarly to the lower bound of B_2, we have that $B_2 \in [0, 1]$ such that $A_0^2 - (A_0 - B_1)(A_0 - B_U) = \alpha$, and solve for B_U:

$$A_0^2 - (A_0 - B_1)(A_0 - B_U) = \alpha \Rightarrow A_0 - B_U = \frac{A_0^2 - \alpha}{A_0 - B_1}$$

$$B_U = A_0 - \frac{A_0^2 - \alpha}{A_0 - B_1} = \frac{\alpha - A_0 B_1}{A_0 - B_1}$$

Example 3.5.4 Assume that we have the values $m = 2$, $\alpha = 0.05$, $\theta = \frac{1}{9}$, $A_0 = 0.9$ and $B_1 = 0.03$, then we have that

$$\begin{aligned}
B_L &= \frac{\alpha - B_1}{1 - B_1} = \frac{0.05 - 0.03}{1 - 0.03} = \frac{0.02}{0.97} = 0.00262 \\
B_U &= \frac{\alpha - AB_1}{A - B_1} = \frac{0.05 - 0.9 \cdot 0.03}{0.9 - 0.03} = \frac{0.027}{0.87} = 0.03103
\end{aligned}$$

3.6 Distributions

This section will denote some pertinent distributions relevant to the multivariate optimization problem. A critical condition for the distributions we are about to explore is that $X < 0$ should not occur, given the nature of how we have defined risk respective to the parties involved. The bibliography used for said distributions is given in *Modern Mathematical Statistics with Applications* [DB12]. The listed distributions used in this thesis are *Log-normal distribution*, *Truncated normal distribution*, *Exponential distribution*, *Pareto distribution*, *Normal distribution* and *Gamma distribution* and can be found in Appendix A.

3.6.1 Copulas

A copula is a statistical tool that is used to model multivariate distributions by describing the dependencies between two or more random variables. It is expressed in terms of marginal distribution functions and a copula, which is based on a correlation parameter, denoted as $\rho \in [-1, 1]$.

A question that arises is why copula approach is preferred over regular multivariate dependency in modeling combined risks. The answer lies in the flexibility it offers in modeling the dependency between risks. By outputting a joint distribution, the copula approach can effectively capture and model the dependency between risks, making it a reliable choice for risk modeling.

Example 3.6.1 (*Gaussian copula*) The Gaussian copula is a tool for modeling the relationship between random variables that have a joint distribution following a multivariate normal distribution. It finds its application in finance and risk management to model the dependency between various financial variables like bonds and stocks. The Gaussian copula can model both positive and negative dependencies between variables, making it a useful tool to represent different economic scenarios, including recessions.

Theorem 3.6.2 *(Sklar's Theorem)*
Let $F_1(x_1)$ and $F_2(x_2)$ be marginal cumulative distribution functions of two random variables x_1 and x_2. Denote $F_1(x_1) = P(X_1 \leq x_1)$ and $F_2(x_2) = P(X_2 \leq x_2)$. Let $F(x_1, x_2)$ be a joint cumulative distribution function $F(x_1, x_2) = P(X_1 \leq x_1, X_2 \leq x_2)$. Then $F(x_1, x_2)$ is linked to $F_1(x_1)$ and $F_2(x_2)$ through a copula. Thus

$$F(x_1, x_2) = C(F_1(x_1), F_2(x_2)) \tag{3.21}$$

Furthermore, if $F_1(x_1)$ and $F_2(x_2)$ are continuous, then C is unique.

Introduce a joint distribution function $C(u_1, u_2)$. Define the distribution functions $u_1 = F_1^{-1}(x_1)$ and $u_2 = F_2^{-1}(x_2)$. Then $x_1 = F_1^{-1}(u_1)$ and $x_2 = F_2^{-1}(u_2)$ are the percentiles of the distribution functions.

Algorithm 4 Inversion sampling

1: **input** $F^{-1}(u)$
2: **draw** $U \sim$ uniform
3: **return** $X \leftarrow F^{-1}(U)$

The distribution functions of X_1 and X_2 becomes $F_1(x_1)$ and $F_2(x_2)$ by the inversion sampling, regardless of the dependencies of U_1 and U_2. The resulting joint distribution function $C(u_1, u_2)$ is known as a *copula*. [Bøl14, p. 202]

Example 3.6.3 The financial crisis of 2008 serves as a great example of why copulas are crucial. The crisis was triggered by the housing market, and copulas were utilized to model the dependence between different types of mortgage-backed securities, including subprime and prime mortgages. The models assumed that the probability of default for each type of mortgage was independent of the possibility of default for the other types.

However, when the defaults of subprime mortgages increased, it became apparent that there was a correlation, and hence the defaults of other types of mortgages also increased. As the copula models did not take this correlation into account, the risk of mortgage-backed securities was greatly underestimated. This led to a widespread purchase of these securities, exposing investors to significantly more risk than originally anticipated, resulting in massive losses.

CHAPTER 4

Optimal reinsurance

In this chapter, we will utilize the theory we have covered so far to support our Master's thesis. Our objective is to introduce an objective function called C_1, which we will strive to minimize. By demonstrating how this optimization is carried out, we will provide examples and results in chapter 5. It is possible to pursue optimization by either minimizing the denominator of the objective function, maximizing the numerator, or both. Please note that if we maximize the numerator, which represents expected gain, no reinsurance contracts will be implemented. Similarly, if we minimize the denominator, our value-at-risk will be reduced to zero. Therefore, we will optimize both of these terms simultaneously to provide an optimal solution for realistic scenarios.

Our policy for pricing is based on a linear model, where the premium for coverage increases proportionally with the amount of coverage purchased. In contrast, a non-linear pricing policy be a convex function, i.e. a case where the price increases more if the risk is moved along tail. This risk aversion is put in place because an insurance company weigh more on higher losses rather than small. Many of the following results are given in the lecture notes for the course STK4400 "Risk and Reliability Analysis" at the University of Oslo as of spring 2023. [Hus23]

4.1 Univariate reinsurance

We start by recalling the univariate case. As this chapter will introduce a new objective function denoted C_1, we will do so for one dimension before resuming the multivariate reinsurance optimization problem.

4.1.1 Partial derivatives of expected gain

We will now introduce the *gain*[1] in the univariate case for the cedent by G:

$$G = (1+\gamma)E[X] - (1+\theta)E[R(X)] - X + R(X)$$

where γ is the risk loading for the client and θ is the risk loading for the cedent. It follows that

$$E[G] = (1+\gamma)E[X] - (1+\theta)E[R(X)] - E[X] + E[R(X)] = \gamma E[X] - \theta E[R(X)]$$

[1]This thesis does not account for the cost of capital

Furthermore, we want to investigate the derivatives of the expected gain, to observe how it behaves when one contract parameter changes while the other remains constant:

$$\frac{\partial E[G]}{\partial a} = -\theta \frac{\partial E[R(X)]}{\partial a}$$
$$\frac{\partial E[G]}{\partial b} = -\theta \frac{\partial E[R(X)]}{\partial b}$$

Notice that the value of $E[X]$ remains constant regardless of the values of a and b. Determining these derivatives, we can recall that from 3.2.1 that:

$$E[R(X)] = \int_a^b x f_X(x)dx - aP(X > a) + bP(X > b)$$

with the corresponding partial derivatives:

$$\frac{\partial E[R(X)]}{\partial a} = -P(X > a)$$
$$\frac{\partial E[R(X)]}{\partial b} = P(X > b)$$

This gives us the following results for the partial derivatives of expected gain:

$$\frac{\partial E[G]}{\partial a} = -\theta \frac{\partial E[R(X)]}{\partial a} = \theta P(X > a) > 0$$
$$\frac{\partial E[G]}{\partial b} = -\theta \frac{\partial E[R(X)]}{\partial b} = -\theta P(X > b) < 0$$

4.1.2 Objective function

Introduce the objective function C_1, where the goal is to find a and b such that

$$C_1(a, b) = \frac{V_\alpha[I(X)]}{E[G]}$$

is minimized. Given the considerations of our value-at-risk assumptions (3.2.2), if $\alpha < S_X(a)$ or equivalently, if $a < S_X^{-1}(\alpha)$, the derivatives of $C_1(a, b)$ with respect to a are such that:

$$\frac{\partial}{\partial a} C_1(a, b) = (E[G])^{-2} \cdot (E[G] - V_\alpha[I(X)] \cdot \theta P(X > a))$$

and similarly if $b < S_X^{-1}(\alpha)$ we have:

$$\frac{\partial}{\partial b} C_1(a, b) = (E[G])^{-2} \cdot (V_\alpha[I(X)] \cdot \theta P(X > b) - E[G])$$

while if $b > S_X^{-1}(\alpha)$ we have:

$$\frac{\partial}{\partial b} C_1(a, b) = (E[G])^{-2} \cdot (V_\alpha[I(X)] \cdot \theta P(X > b))$$

It follows that $b > S_X^{-1}(\alpha)$, we have

$$\frac{\partial}{\partial b} C_1(a, b) > 0$$

Thus, if we can show that for $b < S_X^{-1}$ we have:

$$\frac{\partial}{\partial b} C_1(a, b) < 0$$

then this would imply that the optimal b-value is $S_X^{-1}(\alpha)$. In pursuit of this, we need to first find the corresponding optimal value for a, since the sign of $\frac{\partial}{\partial b} C_1(a, b)$ depends on the value of a. Assume an optimal value for a such that $E[G] > 0$; this condition can be expressed as $a > a_{\min}$, where the lower bound $a_{\min}$ is the solution to $E[G] = 0$. Assume that there exists a unique $a \in (a_{\min}, b)$ such that $\frac{\partial}{\partial a} C_1(a, b) = 0$, given that $a_{\min} < b$. It follows that for this a-value, we have:

$$E[G] = V_\alpha[I(X)] \cdot \theta P(X > a)$$

Insert the expression for $E[G]$ into $\frac{\partial}{\partial b} C_1(a, b)$:

$$\begin{aligned}(E[G])^2 \cdot \frac{\partial}{\partial b} C_1(a, b) &= V_\alpha[I(X)] \cdot \theta P(X > b) - V_\alpha[I(X)] \cdot \theta P(X > a) \\ &= V_\alpha[I(X)] \cdot (\theta P(X > b) - \theta P(X > a))\end{aligned}$$

Since $a < b$, we have that $P(X > b) < P(X > a)$, hence $\frac{\partial}{\partial b} C_1(a, b) < 0$. These results conclude that the optimal value for b is $b^* = S_X^{-1}(\alpha)$. When $b = b^*$, it follows by $V_\alpha[I(X)] = \max\{a, S_X^{-1}(\alpha) - (b - a)\}$ that:

$$V_\alpha[I(X)] = \max\{a, S_X^{-1}(\alpha) - (b^* - a)\} = \max\{a, a\} = a$$

If $b > S_X^{-1}(\alpha)$, we have that:

$$P(I(X) > a) = P(X > b) < \alpha$$

$$P(I(X) \geq a) = P(X > a) \geq \alpha$$

So, by $P(X > x) \leq \alpha \leq P(X \geq x)$, we also get that $V_\alpha[I(X)] = a$. Simultaneously, if $b > b^*$, the expected gain is reduced. Thus, $C_1(a, b) > C_1(a, b^*)$. On the other hand, if $b < S_X^{-1}(\alpha)$, we have:

$$P(I(X) > a) = P(X > b) > \alpha$$

Hence, $V_\alpha[I(X)] > a$. Simultaneously, if $b < b^*$, the expected gain is increased. However, the increase in expected gain is not enough to compensate for the increase in $V_\alpha[I(X)]$. Thus, we still have $C_1(a, b) > C_1(a, b^*)$. We can conclude with that in order to minimize $C_1(a, b)$, the value of b should be chosen such that:

$$P(I(X) > a) = \alpha$$

Since the optimal value for b is determined with b^*, we can continue by optimizing a. This corresponding value, denoted a^*, is found by solving $\frac{\partial}{\partial a} C_1(a, b) = 0$ with respect to $a \in (a_{\min}, b^*)$ which is equivalent to:

$$E[G] = V_\alpha[I(X)] \cdot \theta P(X > a) = a \cdot \theta P(X > a)$$

Recall the expected gain $E[G] = \gamma E[X] - \theta E[R(X)]$. When $E[G] = 0$, we can simplify this to:

$$\gamma E[X] = \theta E[R(X)] \Rightarrow E[R(X)] = \frac{\gamma}{\theta} E[X]$$

where

$$E[R(X)] = \int_a^b x f_X(x)dx - aP(X > a) + bP(X > b)$$

The expected gain can consequentially be written as:

$$E[G] = \gamma E[X] - \theta \int_a^b x f_X(x)dx + a\theta P(X > a) - b\theta P(X > b)$$

So, when $E[G] = a\theta P(X > a)$, the remaining terms of the expression above must equal 0:

$$\gamma E[X] - \theta \int_a^b x f_X(x)dx - b\theta P(X > b) = 0$$

which is simplified to:

$$\int_a^b x f_X(x)dx + bP(X > b) = \frac{\gamma}{\theta} E[X]$$

$$\Rightarrow \int_a^b x f_X(x)dx = \frac{\gamma}{\theta} E[X] - bP(X > b)$$

We have that the right-hand side of the above expression is known. We can use this information to iterate until the expression is true and return the optimal a^*. We can also return its corresponding $A = S_X^{-1}(a^*)$ and the minimized objective function $C_1(a^*, b)$. The trapezoidal rule gives the estimation approach for our integral.

4.2 Multivariate reinsurance

Applying the theory until this point to the multivariate case, we can dive into how the new objective function and its expected gain behave when $i = 1, \ldots, m$ insurance contracts are involved.

4.2.1 Partial derivatives of expected gain

Recall the gain (2.5) in the multivariate case

$$G = (1 + \gamma) \sum_{i=1}^{m} E[X_i] - (1 + \theta) \sum_{i=1}^{m} E[R_i(X_i)] - \sum_{i=1}^{m} X_i + \sum_{i=1}^{m} R_i(X_i)$$

and its respective expected gain (2.6):

$$E[G] = \gamma \sum_{i=1}^{m} E[X_i] - \theta \sum_{i=1}^{m} E[R_i(X_i)]$$

where

$$E[R_i(X_i)] = \int_{a_i}^{b_i} x f_{X_i}(x)dx - a_i P(X_i > a_i) + b_i P(X_i > b_i), \quad i = 1, \ldots, m$$

Since our parameters a_i and b_i are first-order, we treat the parameter not taken as a partial derivative as a constant and cancel it out. Hence, the partial derivative of the premium term of either a_i or b_i only depends on the respective parameter. So, similarly to the total reinsurance premium 3.3.1 in the multivariate case and the expected gain in the univariate case, we have:

$$\begin{aligned}\frac{\partial E[G]}{\partial a_i} &= \frac{\partial}{\partial a_i}\left[\gamma\sum_{i=1}^{m}E[X_i] - \theta\sum_{i=1}^{m}E[R_i(X_i)]\right] - \theta\frac{\partial}{\partial a_i}E[R_i(X_i)] \\ &= \theta P(X_i > a_i) > 0 \\ \frac{\partial E[G]}{\partial b_i} &= \frac{\partial}{\partial b_i}\left[\gamma\sum_{i=1}^{m}E[X_i] - \theta\sum_{i=1}^{m}E[R_i(X_i)]\right] - \theta\frac{\partial}{\partial b_i}E[R_i(X_i)] \\ &= -\theta P(X_i > b_i) > 0\end{aligned}$$

Further, we have the m-dimensional case where our goal is to find $\mathbf{a} = (a_1, \ldots, a_m)$ and $\mathbf{b} = (b_1, \ldots, b_m)$ in order to minimize the objective function $C_1(\mathbf{a}, \mathbf{b})$:

$$C_1(\mathbf{a}, \mathbf{b}) = \frac{V_\alpha[\sum_{i=1}^{m} I_i(X_i)]}{E[G]}$$

where $V_\alpha\left[\sum_{i=1}^{m} I_i(X_i)\right]$ denotes the $(1-\alpha)$ percentile of the distribution of $\sum_{i=1}^{m} I_i(X_i)$. We have that

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

4.2.2 Optimizing a given b

Given that we already have a $\mathbf{a}$ value, we will now consider an optimization problem for $\mathbf{b}$. Assume a fixed value for $\mathbf{b}$ has been chosen, denoted $\mathbf{b}_0$ such that

$$V_\alpha\left[\sum_{j=1}^{m} I_j(X_j)\right] \approx \sum_{j=1}^{m} a_j$$

The optimal value for $\mathbf{a}$ is then found by solving the partial derivative of the objective function with respect to the desired variable, $\mathbf{a}$:

$$\frac{\partial}{\partial a_i}C_1 = 0, \quad i = 1, \ldots, m$$

$$\frac{\partial}{\partial a_i}\frac{V_\alpha[\sum_{j=1}^{m} I_j(X_j)]}{E[G]} = 0$$

$$\frac{\partial}{\partial a_i}C_1 = \frac{\frac{\partial}{\partial a_i}V_\alpha[\sum_{j=1}^{m} I_j(X_j)]\cdot E[G] - \frac{\partial}{\partial a_i}E[G]\cdot V_\alpha[\sum_{j=1}^{m} I_j(X_j)]}{(E[G])^2}$$

where previous results in (3.3.3) has shown us that

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

such that

$$\frac{\partial}{\partial a_i} C_1 = \frac{E[G] - \theta P(X_i > a_i) \cdot V_\alpha[\sum_{j=1}^m I_j(X_j)]}{(E[G])^2} = 0$$

We see that finding the optimal value for **a** holds if and only if

$$E[G] = \theta P(X_i > a_i)) \cdot V_\alpha \left[\sum_{j=1}^m I_j(X_j) \right] = \theta P(X_i > a_i) \cdot \sum_{j=1}^m a_j, \quad i = 1, \ldots, m$$

which may also be denoted as

$$P(X_i > a_i) = \frac{E[G]}{\theta \sum_{j=1}^m a_j}, \quad i = 1, \ldots, m$$

It follows that $P(X_1 > a_1) = \ldots = P(X_m > a_m) = A$ for some probability A. Express **a** in terms of A, as introduced in Theorem 3.3.2:

$$a_i = S_{X_i}^{-1}(A), \quad i = 1, \ldots, m$$

Continue by inserting in the expression for $E[G]$, and we get that:

$$E[G] = \theta A \cdot \sum_{j=1}^m S_{X_j}^{-1}(A)$$

where A can be easily determined, with $a_1, \ldots, a_m$ being iterated. Denote the resulting **a** as $\mathbf{a}_1 = (a_{1,1}, \ldots, a_{1,m})$.

4.2.3 Optimizing b given a

Assume that an optimal value $\mathbf{a} = \mathbf{a}_1$ has been found for an initial value $\mathbf{b} = \mathbf{b}_0$, as shown in (4.2.2). We now want to find an optimal **b**, denoted correspondingly as $\mathbf{b}_1 = (b_{1,1}, \ldots, b_{1,m})$.

The optimal value of **b** should be chosen such that the objective function $C_1(\mathbf{a}_1, \mathbf{b})$ is minimized with respect to **b**:

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

We have that

$$P\left(\sum_{i=1}^m I_i(X_i) > \sum_{i=1}^m a_{1,i} \right) < \alpha, \quad \mathbf{b}' > \mathbf{b}$$

$$P\left(\sum_{i=1}^m I_i(X_i) \geq \sum_{i=1}^m a_{1,i} \right) \geq \alpha, \quad \mathbf{b}' > \mathbf{b}$$

and respectively $V_\alpha[\sum_{i=1}^m I_i(X_i)] = \sum_{i=1}^m a_{1,i}$. Intuitively, the expected gain $E[G]$ will be reduced if a chosen **b** is larger than the initial value of $\mathbf{b} = \mathbf{b}_0$. This is because the layer of our multivariate reinsurance contract is compressed. Hence, $C_1(\mathbf{a}_1, \mathbf{b}') > C_1(\mathbf{a}_1, \mathbf{b})$.

Correspondingly, we have that

$$P\left(\sum_{i=1}^{m} I_i(X_i) > \sum_{i=1}^{m} a_{1,i}\right) > \alpha, \quad \mathbf{b}' > \mathbf{b}$$

where $V_\alpha[\sum_{i=1}^{m} I_i(X_i)] > \sum_{i=1}^{m} a_{1,i}$. In this case, the expected gain $E[G]$ is *not* increased, as the layer of our multivariate reinsurance contract is extended. Moreover, the objective function is not affected as a result of this since the expected gain does not compensate for the increase in $V_\alpha[\sum_{i=1}^{m} I_i(X_i)]$. Hence, $C_1(\mathbf{a}_1, \mathbf{b}') > C_1(\mathbf{a}_1, \mathbf{b})$. Continuing, we have that for a given $\mathbf{a} = \mathbf{a}_1$, the following optimization problem arises:

$$\begin{aligned} &\text{minimize} \quad C_1(\mathbf{a}_1, \mathbf{b}) = \frac{V_\alpha[\sum_{i=1}^{m} I_i(X_i)]}{E[G]} \\ &\text{subject to} \quad P\left(\sum_{i=1}^{m} I_i(X_i) > \sum_{i=1}^{m} a_{1,i}\right) = \alpha \end{aligned} \tag{4.1}$$

Recall that the numerator $V_\alpha[\sum_{i=1}^{m} I_i(X_i)] = \sum_{i=1}^{m} a_{1,i}$, meaning that the objective function is minimized when the denominator is maximized, as the numerator is a constant. Hence, the optimization problem can be denoted as follows:

$$\begin{aligned} &\text{maximize} \quad E[G] = \gamma \sum_{i=1}^{n} E[X_i] - \theta \sum_{i=1}^{n} E[R_i(X_i)] \\ &\text{subject to} \quad P\left(\sum_{i=1}^{m} I_i(X_i) > \sum_{i=1}^{m} a_{1,i}\right) = \alpha \end{aligned} \tag{4.2}$$

Given that the retained risk is not dependent on the reinsurance contract, as defined by $\mathbf{b}$, the optimization problem may be further simplified to:

$$\begin{aligned} &\text{minimize} \quad \sum_{i=1}^{n} E[R_i(X_i)] \\ &\text{subject to} \quad P\left(\sum_{i=1}^{m} I_i(X_i) > \sum_{i=1}^{m} a_{1,i}\right) = \alpha \end{aligned} \tag{4.3}$$

where

$$\begin{aligned} E[R_i(X_i)] &= \int_{a_i}^{b_i} x f_{X_i}(x)dx - a_i P(X_i > a_i) + b_i P(X_i > b_i) \\ &= \int_{a_i}^{b_i} x f_{X_i}(x)dx - a_i A + b_i B_i, \quad i = 1, \dots, m \end{aligned} \tag{4.4}$$

Once $\mathbf{a}_1$ is determined, the next procedure is to determine the resulting $\mathbf{a}$, but by $\mathbf{b}_1$ instead of $\mathbf{b}_0$, and the resulting $\mathbf{a}$ is denoted $\mathbf{a}_2 = (a_{2,1} \dots, a_{2,m})$. The next step for finding $\mathbf{b}$ denoted $\mathbf{b}_2 = (b_{2,1}, \dots, b_{2,m})$ is to find the optimal $\mathbf{b}$ by using the resulting $\mathbf{a}_2$ rather than $\mathbf{a}_1$. This process is iterated until the solutions converge.

4.2.4 Optimal solution in the exponential case

We will now apply the techniques introduced above to the bivariate exponential case. Recall that the exponential distribution is a special case of the gamma distribution. An essential property of the exponential distribution is that it is memoryless and non-negative, which is necessary given the state of the transactions.

Introduce the granulated sets for the set $\mathcal{C}$:

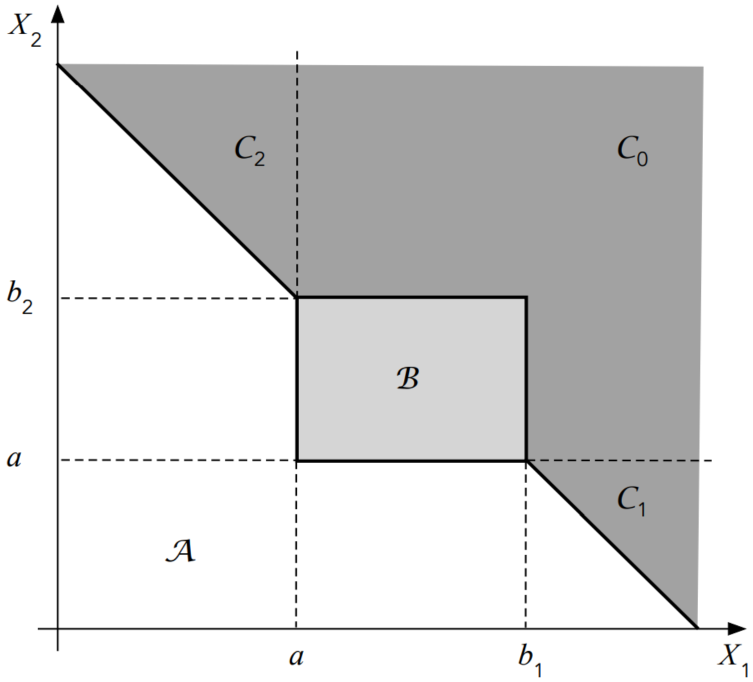


Figure 4.1: Granulated sets $\mathcal{A}$, $\mathcal{B}$ and $\mathcal{C}$ for $i = 1, 2$

We have that $X_i \sim \exp(\lambda)$, $i = 1, 2$.

$$P(X_i > x) = e^{-\lambda x}, \quad i = 1, 2$$

by $a_i = S_{X_i}^{-1}(A)$ it follows that the optimal values for a_i, $\; i = 1, 2$ must satisfy:

$$P(X_i > a_i) = e^{-\lambda a_i} = A, \quad i = 1, 2$$

solving for the optimal values a_i, we get that

$$a_i = -\frac{\ln(A)}{\lambda}, \quad i = 1, 2$$

utilizing the condition

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

We get that the optimal values for b_i, $\; i = 1, 2$ is simplified to

$$P(I_1(X_1) + I_2(X_2) > 2a) = \alpha$$

where the event $\{I_1(X_1)+I_2(X_2)>2\alpha\}$ corresponds to the set $\mathcal{C}$, also denoted as the granulated subsets $\mathcal{C}_0\cup\mathcal{C}_1\cup\mathcal{C}_2$, as visualized in figure 4.1. Further we will solve for the area involving the subsets $\mathcal{C}_0$, $\mathcal{C}_1$ and $\mathcal{C}_2$:

$$\begin{aligned} P((X_1,X_2)\in\mathcal{C}_0) &= P((X_1,X_2)\in\mathcal{C}_0\cup\mathcal{B}) - P((X_1,X_2)\in\mathcal{B}) \\ &= e^{-2\lambda a} - (e^{-\lambda a}-e^{-\lambda b_1})\cdot(e^{-\lambda a}-e^{-\lambda b_2}) \end{aligned} \tag{4.5}$$

where the common value of $a_i,\ \ i=1,2$ is denoted as a.
Continuing for the subset $\mathcal{C}_1$:

$$\begin{aligned} P((X_1,X_2)\in\mathcal{C}_1) &= P(X_1+X_2>b_1+a\cap X_2\le a) \\ &= \int_0^a P(X_1+x_2>b_1+a)\lambda e^{-\lambda x_2}dx_2 \\ &= \int_0^a e^{-\lambda(b_1+a-x_2)}\lambda e^{-\lambda x_2}dx_2 \\ &= \int_0^a \lambda e^{-\lambda(b_1+a)}dx_2 = \lambda a e^{-\lambda(b_1+a)} \end{aligned} \tag{4.6}$$

and similarly for the subset $\mathcal{C}_2$:

$$P((X_1,X_2)\in\mathcal{C}_2) = \lambda a e^{-\lambda(b_2+a)}$$

Introduce

$$P(X_i>b_i)=S_{X_i}(b_i)=B_i=e^{-\lambda b_i},\ \ i=1,2$$

We want to estimate α given the condition $P(I_1(X_1)+I_2(X_2)>2a)=\alpha$. Simplify by also using the area for each subset as found above:

$$\begin{aligned} P(I_1(X_1)+&I_2(X_2)>2a) \\ &= P((X_1,X_2)\in\mathcal{C}_0)+P((X_1,X_2)\in\mathcal{C}_1)+P((X_1,X_2)\in\mathcal{C}_2) \\ &= e^{-2\lambda a}-(e^{-\lambda a}-e^{-\lambda b_1})\cdot(e^{-\lambda a}-e^{-\lambda b_2})+\lambda a\cdot e^{-\lambda a}\cdot(e^{-\lambda b_1}+e^{-\lambda b_2}) \\ &= A^2-(A-B_1)\cdot(A-B_2)+\lambda a\cdot A(B_1+B_2) \\ &= A(1+\lambda a)(B_1+B_2)-B_1B_2=\alpha \end{aligned} \tag{4.7}$$

insert $a=-\frac{\ln(A)}{\lambda}$:

$$\Rightarrow A(1-\ln(A))(B_1+B_2)-B_1B_2=\alpha$$

$$\Rightarrow A(1-\ln(A))(B_1+B_2)=\alpha+B_1B_2$$

$$B_1+B_2=\frac{\alpha+B_1B_2}{A(1-\ln(A))}$$

Turning our focus to the objective function, which is used to optimize the parameters defined by our insurance layer contract, we have that

$$C_1=\frac{V_\alpha[\sum_{i=1}^m I_i(X_i)]}{E[G]}=\frac{V_\alpha[I_1(X_1)+I_2(X_2)]}{E[G]}=\frac{2a}{E[G]}$$

Minimizing the objective function C_1 with respect to the optimal values B_1 and B_2 for a given A is equivalent with maximizing $E[G]$ subject to $B_1+B_2=\frac{\alpha+B_1B_2}{A(1-\ln(A))}$.

We have already remarked that the first term $\gamma \sum_{i=1}^{m} E[X_i]$ in $E[G]$ may be disregarded in the maximization problem, given that the term is not dependent on the values of B_1 and B_2. Maximizing $E[G]$ is then equivalent to minimizing the latter term $-\theta \sum_{i=1}^{m} E[R_i(X_i)]$ with respect to the values B_1 and B_2 subject to $B_1 + B_2 = \frac{\alpha + B_1 B_2}{A(1-\ln(A))}$.
Recall the solution for the expected reinsurance premiums in subsection 3.4.2:

$$\Phi_i = E[R_i(X_i)] = \int_{a_i}^{b_i} x f_{X_i}(x)dx - a_i P(X_i > a_i) + b_i P(X_i > b_i), \quad i = 1, \dots, m$$

Applying the exponential case to this solution, we get that the latter term of the expected gain $E[G]$ may be denoted as

$$\begin{aligned} E[R_i(X_i)] &= \int_{a_i}^{b_i} x\lambda e^{-\lambda x} dx - ae^{-\lambda a} + b_i e^{-\lambda b_i} \\ &= \frac{(a\lambda + 1)e^{-\lambda a} - (b_i\lambda + 1)e^{-\lambda b_i}}{\lambda} - ae^{-\lambda a} + b_i e^{-\lambda b_i} \\ &= \frac{a\lambda e^{-\lambda a} + e^{-\lambda a} - b_i \lambda e^{-\lambda b_i} - e^{-\lambda b_i} + \lambda a e^{-\lambda a} + \lambda b_i e^{-\lambda b_i}}{\lambda} \\ &= \frac{e^{-\lambda a} - e^{-\lambda b_i}}{\lambda} = \frac{A - B_i}{\lambda}, \quad i = 1, 2 \end{aligned} \tag{4.8}$$

Maximizing $E[G]$ subject to $B_1 + B_2 = \frac{\alpha + B_1 B_2}{A(1-\ln(A))}$ is equivalent to minimizing $\lambda^{-1}(2A - (B_1 + B_2))$ with respect to B_1 and B_2 subject to $B_1 + B_2$. Assuming $\lambda > 0$ is a known constant and that A is determined, this is equivalent to maximizing $B_1 + B_2$ subject to $B_1 + B_2 = \frac{\alpha + B_1 B_2}{K}$ where $K = A(1 - \ln(A))$.

Remark 4.2.1 Given that $(1 - \ln(A)) > 1$, it follows that $K > A$ since $\frac{K}{A} = 1 - \ln(A)$.

We have that $B_1 + B_2$ may be simplified in order to express B_1 in terms of B_2:

$$B_1 + B_2 = \frac{\alpha + B_1 B_2}{K} \Rightarrow B_1 = \frac{\alpha - K B_2}{K - B_2}$$

and the goal of optimizing the objective function C_1 may now be expressed as a function of the value B_2 alone. This function is denoted by ϕ:

$$\phi(B_2) = \frac{\alpha - K B_2}{K - B_2} + B_2$$

with the derivative being calculated as follows:

$$\phi'(B_2) = 1 - \frac{K^2 - \alpha}{(K - B_2)^2} = 1 - \frac{K^2 - \alpha}{(B_2 - K)^2}$$

Given the remark (4.2.1), we need to assume $K^2 > \alpha$, to solve for the optimal value of the equation $\phi'(B_2) = 0$:

$$1 - \frac{K^2 - \alpha}{(B_2 - K)^2} = 0$$

$$K^2 - \alpha = (B_2 - K)^2$$

$$B_2 = K \pm \sqrt{K^2 - \alpha}$$

Proposition 4.2.2 *The optimal value of B_2 is given as $B_2 = K - \sqrt{K^2 - \alpha}$ with the corresponding value of $B_1 = K - \sqrt{K^2 - \alpha}$.*

Proof. To have a reasonable solution for the insurance layer contract, we must have $b_2 > a$ given the boundaries of our interval, hence $B_2 < A$. The proposed solution $B_2 = K \pm \sqrt{K^2 - \alpha}$ involves $B_2 = K + \sqrt{K^2 - \alpha}$, of which the property $B_2 > K > A$ is not valid for the nature of our reinsurance contract. Therefore, we let $B_2 = K - \sqrt{K^2 - \alpha}$, yielding the property $0 < B_2 < A < K$. The solution is valid given that:

$$\alpha < A^2(1 - \ln(A^2))$$

Furthermore, the derivative of our function ϕ is decreasing in B_2 around the solution, meaning that this solution corresponds to a maximum value of the objective function C_1. Correspondingly, we can solve for B_1:

$$\begin{aligned} B_1 &= \frac{\alpha - KB_2}{K - B_2} = \frac{\alpha - K(K - \sqrt{K^2 - \alpha})}{K - (K - \sqrt{K^2 - \alpha})} \\ &= \frac{\alpha - K^2 + K\sqrt{K^2 - \alpha}}{\sqrt{K^2 - \alpha}} \\ &= K - \sqrt{K^2 - \alpha} \end{aligned} \tag{4.9}$$

Hence, $B_i = K - \sqrt{K^2 - \alpha}, \;\; i = 1, 2$. Our solution is then a *balanced solution*, as introduced in 3.5.1. Let B denote the common value of B_1 and B_2. ∎

4.2.5 Lagrange approach

Continuing, we will expand our approach for optimal values of B_1 and B_2 by applying it to the analytical approach using Lagrange multipliers. Although the generalized form may be exponentially complicated and extensive given an increase in i parameters, it is beneficial for finding local minima and maxima of the objective function.

Introduce

$$\Lambda(B_1, B_2) = \Phi(B_1, B_2) - L \cdot \Psi(B_1, B_2)$$

where

$$L \text{ is denoted as the Lagrange multiplier}$$

$$\Phi(B_1, B_2) = B_1 + B_2$$

$$\Psi(B_1, B_2) = B_1 + B_2 - \frac{\alpha - B_1 B_2}{K}$$

A potential extrema point of $\Phi(B_1, B_2)$ is found by solving

$$\nabla\Lambda(B_1, B_2) = 0$$

where L is determined by $\Psi(B_1, B_2) = 0$. We have that

$$\nabla\Lambda(B_1, B_2) = \nabla\Phi(B_1, B_2) - L \cdot \nabla\Psi(B_1, B_2) = 0$$

$$\Rightarrow \nabla\Phi(B_1, B_2) = L \cdot \nabla\Psi(B_1, B_2)$$

Moreover,

$$\nabla\Phi(B_1, B_2) = (1, 1)$$

$$\nabla\Psi(B_1, B_2) = \left(1 - \frac{B_2}{K}, 1 - \frac{B_1}{K}\right)$$

and

$$L\left(1 - \frac{B_i}{K}\right) = 1, \quad i = 1, 2$$

with the resulting solution when solving for B_i:

$$B_i = K(1 - L^{-1}), \quad i = 1, 2$$

Let the common value of B_i, $i = 1, 2$ be B, as our extreme point yields a balanced solution for the optimal values. Continuing, we need to determine L such that $\Psi(B_1, B_2) = 0$ holds:

$$B_1 + B_2 = \frac{\alpha + B_1 B_2}{K} \Rightarrow 2B = \frac{\alpha + B^2}{K}$$

with the following solution for B, given that this is the desired quantity rather than L:

$$B = K \pm \sqrt{K^2 - \alpha}$$

Recall that $0 < B < A < K$, and the positive coefficient is not a valid solution, hence

$$B_i = B = K - \sqrt{K^2 - \alpha}, \quad i = 1, 2$$

4.2.6 The objective function

Turning our focus back to the objective function C_1, we can continue by deriving the optimal value for A as well, given that the optimal value for B is expressed in terms of A:

$$C_1 = \frac{V_\alpha[I_1(X_1) + I_2(X_2)]}{E[G]} = \frac{2a}{E[G]}$$

with the solution for a being $a = -\frac{\ln(A)}{\lambda}$, yielding the expected gain:

$$E[G] = \gamma(E[X_1] + E[X_2]) - \theta(E[R_1(X_1)] + E[R_2(X_2)]) = \frac{2\gamma}{\lambda} - \frac{2\theta(A - B)}{\lambda}$$

Hence

$$C_1(A) = \frac{\frac{-2\ln(A)}{\lambda}}{\frac{2\gamma}{\lambda} - \frac{2\theta(A-B)}{\lambda}} = \frac{-\ln(A)}{\gamma - \theta(A - B)}$$

where

$$B = K - \sqrt{K^2 - \alpha} = A(1 - \ln(A)) - \sqrt{A^2(1 - \ln(A))^2 - \alpha}$$

We can see that these results do not depend on the constant λ and that the optimal value of A can be found by minimizing the introduced objective function $C_1(A)$ by either $C_1'(A) = 0$ or by a numerical algorithm.

In order to minimize $C_1(A)$ we need to allocate a suitable interval $[A_{\min}, A_{\max}]$. Note that the denominator $\gamma - \theta(A - B)$ which is equal to

$\frac{\lambda}{2} \cdot E[G]$ after the simplification in (4.2.6), can be negative for large values of A. The upper bound A_{max} needs to be chosen so that the denominator is positive regardless of whatever value B holds.
Let

$$A_{\text{max}} < \frac{\gamma}{\theta}$$

Furthermore, a large value of C_1 is not optimal either, and to counter this effect, a proportional property can be introduced for good measure:

$$A_{\text{max}} = \frac{3}{4} \cdot \frac{\gamma}{\theta}$$

In addition, for the lower bound A_{min}, we have that the square root in the formula for B is valid if

$$A_{\text{min}} = \min\{A : A^2(1 - \ln(A))^2 \geq \alpha\}$$

Solving the following equation, the value of A_{min} can be found:

$$A_{\text{min}}^2(1 - \ln(A_{\text{min}}))^2 = \alpha$$

by using the bisection method on the interval $[\alpha, A_{\text{min}}]$, as introduced in (3.5.3).

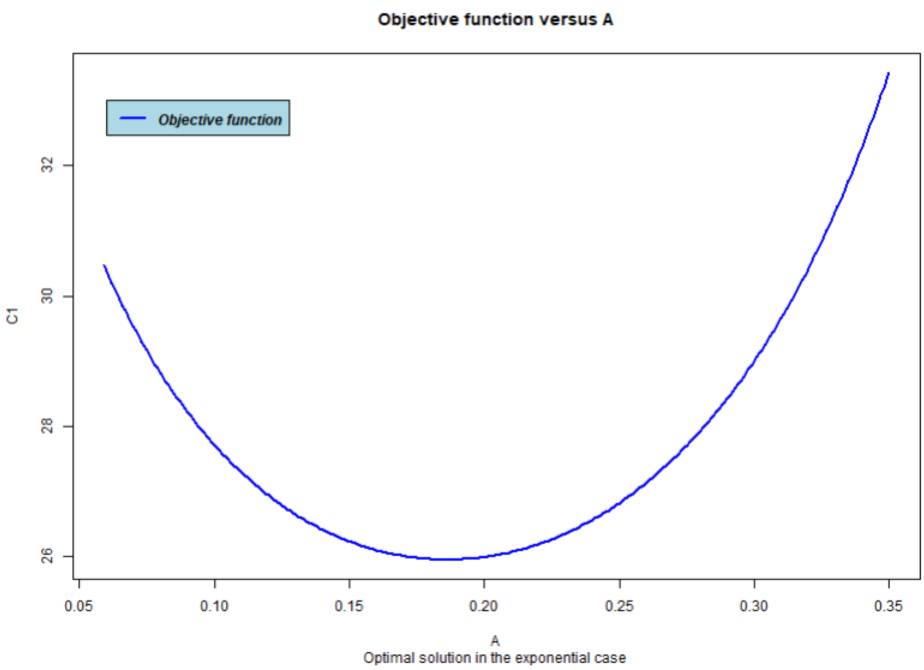


Figure 4.2: Objective function $C_1(A)$

Example 4.2.3 Figure 4.2 demonstrates the objective function C_1 as a function of A over a given interval $[A'_{\text{min}}, A_{\text{max}}]$. The optimal solution is given by $A_{\text{opt}} = 0.186$ and $B_{\text{opt}} = 0.010$, while the resulting value of the objective function is $C_1 = 25.947$. This example has $\gamma = 0.1$, $\theta = 0.2$ and $\alpha = 0.01$.

CHAPTER 5

Numerical results

This chapter will denote the numerical results outputted by the simulation. We will review various distributions from appendix A, in addition to the effects of the optimal solution when there is a change in dependency and multiple combinations of the characteristics of the optimization problem in the univariate and multivariate case, such as symmetry, balance, and dependency.

All the following examples consider either the univariate or bivariate case, with tables and visualizations respective to each distribution. We let $\alpha = 0.01$, $\theta = 0.20$, $\gamma = 0.10$, $\rho = 0.00$ and $\Delta = 0.10$. The iso-curves are graphed by Φ, which is returned from the objective function C_1. The iso-curves are computed using numerical integration, while the constraint curves are calculated using the importance sampling method with $N = 5\ 000\ 000$.

5.1 Optimization in the univariate case

We start with the univariate case, i.e., $m = 1$. We expect an insurance layer contract for the following results, as there are no additional distributions to offset our optimization.

5.1.1 Simulation

This section will apply the theoretical background in the univariate case to how the simulation works and how the optimized parameters behave depending on distributions. We know that the optimal A is yielded when $E[G] = a\theta P(X > a)$ is found. This was simplified in 4.1.2, where we saw that this is the same as solving

$$\gamma E[X] - \theta \int_a^b x f_X(x)dx - b\theta P(X > b) = 0$$

or equivalently, as utilized by the simulation:

$$\gamma E[X] - \theta \int_a^b x f_X(x)dx + a\theta P(X > b) - b\theta P(X > b) = a\theta P(X > b)$$

where the point of intersection represents when the lefthand and righthand sides are equal.

5.1.2 Results

The following figures and parameters are yielded for each distribution, compared to an increase in standard deviation:

	$E[X]$	sd$[X]$	Opt C_1	Opt A	Opt B	Opt a	Opt b
Truncated normal	30	30	24.86	0.201	0.01	49.29	125.72
Log-normal	30	30	23.17	0.216	0.01	40.82	147.20
Pareto	30	30	14.88	0.336	0.01	27.61	118.38
Exponential	30	30	25.95	0.193	0.01	49.40	138.15
Truncated normal	30	50	28.69	0.174	0.01	57.71	141.30
Log-normal	30	50	33.51	0.149	0.01	51.18	225.68
Pareto	30	50	15.78	0.317	0.01	27.45	135.35
Exponential	30	30	25.95	0.193	0.01	49.40	138.15

Table 5.1: Optimal contract parameters in the univariate case

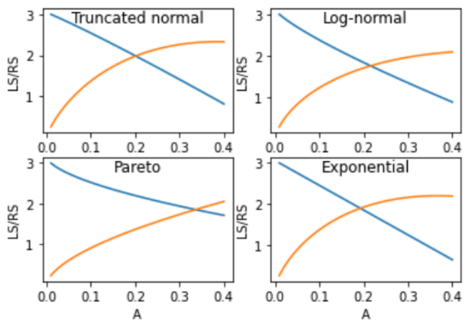


Figure 5.1: Univariate LS and RS versus A

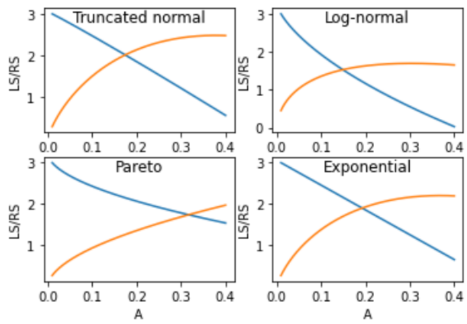


Figure 5.2: Univariate LS and RS versus A

Comparing these two figures and their respective tables, we can naturally see that when the standard deviation increases, so do our objective function and insurance layer contract. Furthermore, we can see that this affects some distributions more than others. For instance, the log-normal distribution experiences a higher jump in the objective function in addition to insurance layer contract parameters a and b. Note that the left side, LS, is the blue graph, while the right side, RS, is the orange graph.

Moreover, both the truncated normal and Pareto distributions are influenced differently. Truncated normal distribution experiences a more considerable jump in contract parameter a, while Pareto slightly decreases. Both distributions have an increased b-value, although Pareto has a somewhat more significant proportion increase.

We can see from figures 5.1 and 5.2 that the left-hand sign and right-hand sign of both the truncated normal distribution and log-normal distribution intersect at a lower value of A. At the same time, Pareto and exponential seem relatively unchanged. This is because the exponential distribution does not get affected by these distinctions.

Investigating further, we can see the effects as standard deviation increases in figures 5.3 and 5.4. Log-normal provides a more apparent minimized C_1 value

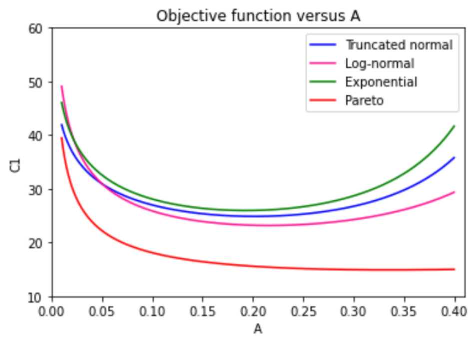


Figure 5.3: Univariate objective function C_1 versus A

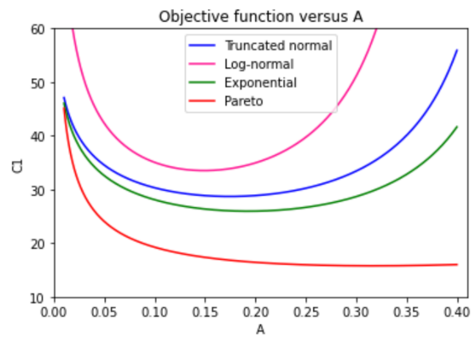


Figure 5.4: Univariate objective function C_1 versus A

but increases the objective function. The objective function for the truncated normal distribution also has a jump, but not as drastic as log-normal. Both the Pareto and exponential distributions remain relatively unchanged. Note that $E[X] < sd[X]$ is not always sufficient since we may omit some risks when truncation occurs. As for the Pareto distribution, the opposite is true since the distribution is prone to losses with low frequency but high severity.

5.2 Optimization in the bivariate case

Now let us focus on a scenario involving two variables, known as a bivariate scenario ($m = 2$). We aim to determine whether the outcomes will fall under a stop-loss or an insurance layer contract. In a symmetrical scenario, we can expect most of the distributions to yield insurance layer contracts, except for the truncated normal distribution, which will influence the mean, standard deviation, and truncation points. However, one distribution may result in a stop-loss contract in asymmetrical scenarios. At the same time, the other may lead to an insurance layer contract, depending on the level of disparity between the distributions.

5.2.1 Simulation

This section will denote the application of the theoretical background from chapters 3 and 4 to the numeric simulations through simulated distributions and methods in the bivariate case.

The simulation consists of 4 stages:

- Stage 1: Calculate the optimal A-value by solving

$$E[G] = A \cdot \theta \cdot (a_1 + a_2)$$

of which the solution will minimize the objective function

$$C_1 = \frac{(a_1 + a_2)}{E[G]}$$

based on initial guesses of B_1 and B_2

- Stage 2: Use the current optimal value of A found in the previous stage in order to optimize B_1 and B_2. Firstly, calculate expected reinsurance costs $R_1(X_1)$ and $R_2(X_2)$ as functions of B_i, and store the results. Secondly, calculate the isoquant curves for $R_1(X_1)+R_2(X_2)$ as functions of (B_1, B_2). Thirdly, generate risks sampled from the set $\mathcal{D}$ in order to determine $\max(B_1)$ and $\min(b_1)$, followed by determining the corresponding $\max(B_2)$ and $\min(b_2)$, respectively. Lastly, determine the constraint set, i.e., the set of values of B_1 and B_2 such that $P(\mathcal{C}) = \alpha$. For each (B_1, B_2) in the constraint set, calculate $\Phi = \sum_{i=1}^{m} E[R_i(X_i)]$. The optimal value of (B_1, B_2) is the value that minimizes Φ.

- Stage 3: Again, we find the optimal value of A based on the optimized values of B_1 and B_2 by solving

$$E[G] = \theta \cdot A \cdot (a_1 + a_2)$$

of which the solution will minimize the objective function

$$C_1 = \frac{(a_1 + a_2)}{E[G]}$$

based on initial guesses of B_1 and B_2

- Stage 4: Stage 4 allows for calculating the balanced and unbalanced solutions.

We will examine different characteristics, such as symmetrical, asymmetrical, balanced, unbalanced, and changes in dependencies, to analyze methods for optimizing a set of reinsurance contracts using value-at-risk as a risk measure and C_1 as the objective function. In addition, we will attempt to replicate various combinations of these characteristics to locate significant differences between cases. We will utilize the log-normal, truncated normal, Pareto, and exponential distribution.

5.2.2 Symmetrical

We start by looking at symmetrical cases. Then, we will work with optimal reinsurance in the bivariate case where the risks X_1 and X_2 are sampled from the same distribution.

Log-normal distribution

Lognormal distribution can model positive values with a skewed distribution and is useful for modeling financial data with extreme values. However, this distribution can also be sensitive to outliers in our data.

	$E[X_1]$	$E[X_2]$	sd$[X_1]$	sd$[X_2]$	$\min(\Phi)$	Opt C_1	Opt A
Example 1	50	50	50	50	20.62	23.55	0.212
Example 2	100	100	50	50	28.30	15.55	0.322
Example 3	50	50	30	30	15.88	17.00	0.294

Table 5.2: Optimal values for symmetrical log-normally distributed risks

With corresponding optimal contract parameters

	a_1	a_2	b_1	b_2	B_1	B_2
Example 1	68.72	68.72	276.87	271.58	0.00673	0.00717
Example 2	98.79	98.79	264.69	290.15	0.00717	0.00807
Example 3	57.88	57.88	167.25	163.715	0.00706	0.00785

Table 5.3: Optimal contract parameters for symmetrical log-normally distributed risks

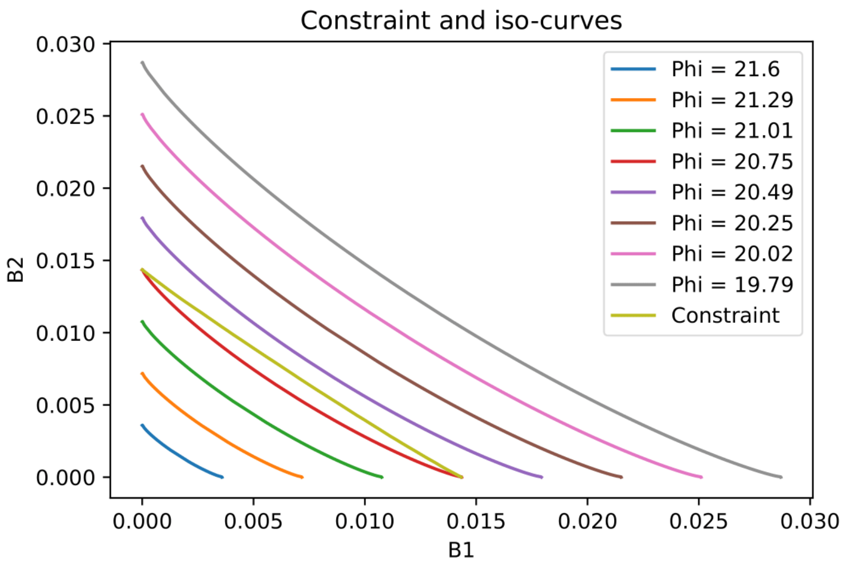


Figure 5.5: Constraint and Iso-contours for example 1 in the log-normal distribution

The optimal a and optimal b seem to be affected more if there is a change in mean rather than a change in standard deviation. Naturally, the parameters are increased when the mean increases, and otherwise, if the opposite is true. The same goes for variations in the standard deviation. We can also notice that when the mean is closer to being equal to the standard deviation, the objective function C_1 increases, thus yielding a higher value for the optimal A.

We have that the hazard rate of the lognormal distribution depends on its properties. For instance, a log-normally distributed risk with a standard deviation higher than the mean may have a decreasing hazard rate. However, suppose the risk is log-normally distributed with a mean higher than or equal to the standard deviation. In that case, there is an increasing hazard rate (regarding our resulting iso curves and their convexity/concavity). In reality, the hazard rate will first increase and then decrease as it moves further out the tail. However, this also depends on our expected value and standard deviation parameters.

We can see from figure 5.6 that we indeed have a balanced, optimal, and unique solution in example 1. Although we have slight differences from our

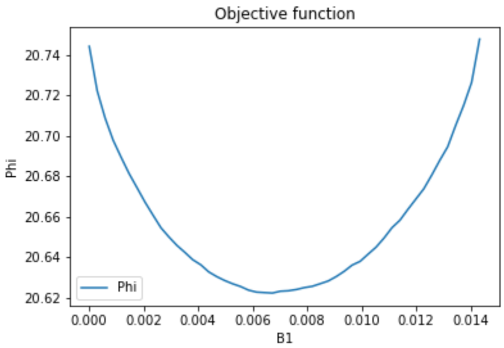


Figure 5.6: Φ versus B_1 in example 1

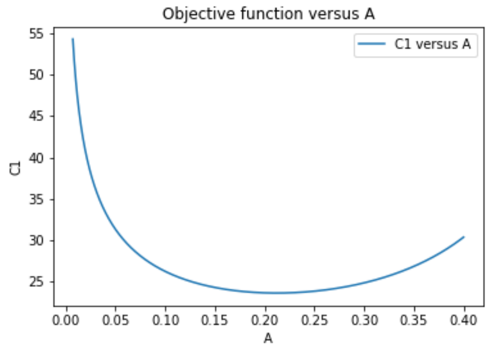


Figure 5.7: C_1 versus A in example 1

numerical outputs in table 5.3, we can argue that the optimal balanced solution is calculated as follows $(B_1^*, B_2^*) = (0.007, 0.007)$, with insurance layer contract for risks X_1 and X_2. From figure 5.7, we also have a clear minimum for our objective function C_1 in the interval $A \in [0, 0.4]$, where $A = 0.212$.

Pareto distribution

Pareto distribution can model heavy-tailed distributions and capture extreme values that occur with a low probability. Although a good application for our data, it may not be appropriate for simulating data with a finite upper bound and requiring a large sample size to accurately determine parameters, which may be limited by processing power.

	$E[X_1]$	$E[X_2]$	sd$[X_1]$	sd$[X_2]$	min(Φ)	Opt C_1	Opt A
Example 4	50	50	30	30	16.25	13.83	0.362
Example 5	110	110	30	30	21.43	11.90	0.400
Example 6	50	50	70	70	21.01	16.08	0.311

Table 5.4: Optimal values for symmetrical Pareto distributed risks

With corresponding optimal contract parameters

	a_1	a_2	b_1	b_2	B_1	B_2
Example 4	46.64	46.64	192.04	196.45	0.00525	0.00512
Example 5	105.40	105.40	254.65	257.89	0.00579	0.00545
Example 6	46.56	46.56	283.61	292.70	0.00554	0.00517

Table 5.5: Optimal contract parameters for symmetrical Pareto distributed risks

In the case of the Pareto distribution, we can see that an increase in mean has a much more significant impact on the a_i parameters, which slightly decreases when the standard deviation is increased. The resulting b_i parameters are also increased more when there is a higher standard deviation, resulting in a much larger premium for the cedent. The objective function C_1 sees a lower minimum for a higher mean than a higher standard deviation. The expected gain is then

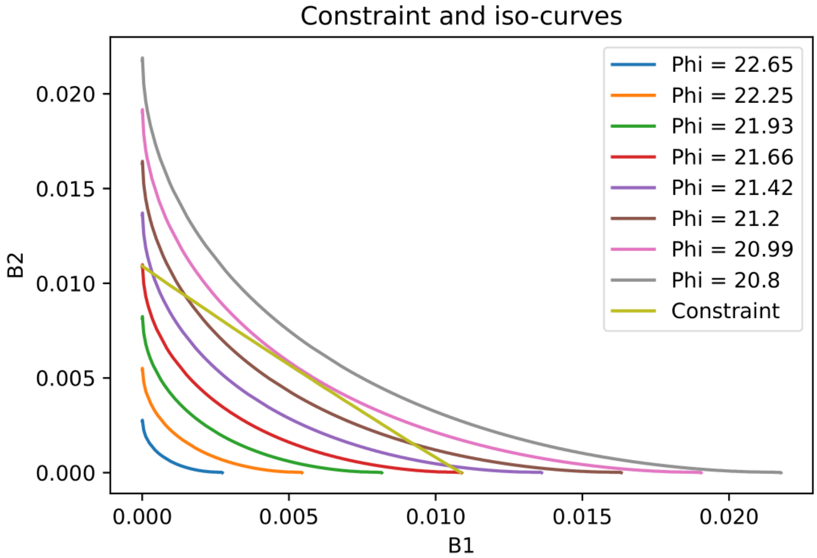


Figure 5.8: Constraint and Iso-contours for example 6 in the Pareto distribution

increased by this change of mean since $a_1 + a_2$ is increased. This is evident from the optimal A being almost 25% larger for the example 5 than example 6.

In figure 5.8, we can have that X_1 and X_2 is Pareto distributed with mean 50 and standard deviation 70. Since the Pareto distribution has a decreasing hazard rate, we expect by Theorem 3.4.7 that Φ to be quasiconvex. This is evident by the illustration, which shows that the sublevels are convex. The constraint curve is bending away from the origin. We can see that the solution is slightly unbalanced, with $(B_1^*, B_2^*) = (0.00554, 0.00517)$ and the corresponding $(b_1^*, b_2^*) = (283, 293)$. Although, due to numerical error, this example is indeed balanced, i.e., $B_1 = B_2 \approx 0.005$ due to symmetry.

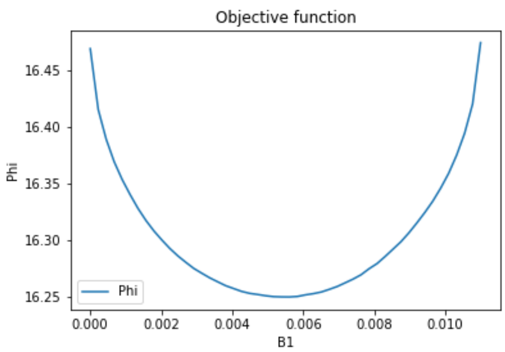


Figure 5.9: Φ versus B_1 in example 4

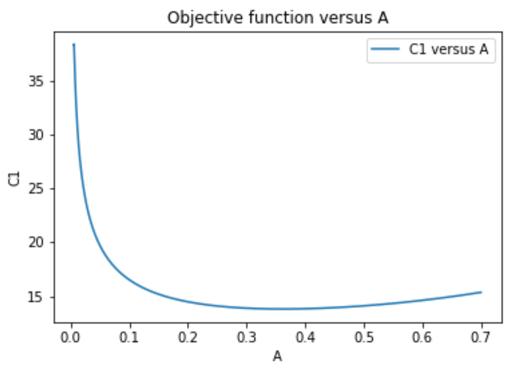


Figure 5.10: C_1 versus A in example 4

We can see from figure 5.9 that we indeed have a balanced, optimal, and unique solution in example 4. Although we have slight differences from our numerical outputs in table 5.5, we can argue that the optimal balanced solution is calculated as follows $(B_1^*, B_2^*) = (0.005, 0.005)$, with insurance layer contract for risks X_1 and X_2. From figure 5.10, we also have a clear minimum for our objective function C_1 in the interval $A \in [0, 0.7]$, where $A = 0.362$. Note that we have expanded our x-axis for the Pareto distribution when examining A since a halt at 0.4 would be unclear whether or not there was a minimum in our interval.

Truncated normal distribution

Truncated normal distribution is able to model continuous data with a bounded range and can also capture both positive and negative values (although we assume $X_i \geq 0, \quad i = 1, \ldots, m$). Moreover, it may be computationally limited to simulate in addition to sensitivity depending on the truncation points.

	$E[X_1]$	$E[X_2]$	sd$[X_1]$	sd$[X_2]$	min(Φ)	Opt C_1	Opt A
Example 7	50	50	30	30	12.19	17.48	0.286
Example 8	100	100	30	30	16.09	12.96	0.386
Example 9	50	50	60	60	14.76	24.68	0.203

Table 5.6: Optimal values for symmetrical truncated normally distributed risks

With corresponding optimal contract parameters

	a_1	a_2	b_1	b_2	B_1	B_2
Example 7	65.92	65.92	116.69	205.66	0.00229	$4.586 \cdot 10^{-6}$
Example 8	108.67	108.67	∞	162.32	0	0.01908
Example 9	83.69	83.69	∞	185.42	0	0.0222

Table 5.7: Optimal contract parameters for symmetrical truncated normally distributed risks

In this example, we can see that an increase in mean will affect the results more than an increase in standard deviation. This can be seen by the resulting expected reinsurance cost Φ. In addition, we have that $a_1 + a_2$ is higher than in the previous example, yet the optimal C_1 is decreased. So this must mean that the expected gain is increased when the mean increases even though the expected reinsurance cost increases. This is to be expected since changes to the mean do not affect the premium for the reinsurance contract, even though it increases the retained risk.

In figure 5.11, we let X_1 and X_2 be truncated normally distributed with mean 100 and standard deviation 30. The truncated normal distribution has an increasing hazard rate. Thus, by Theorem 3.4.7 Φ is quasiconcave. Hence the superlevel sets are convex. We can see from figure 5.11 that the iso-curves are indeed convex since they are bending away from the origin. The optimal solution is the one where $B_1 = 0$ and $B_2 = 0.019$. The corresponding solution for the b_i-values are $(b_1^*, b_2^*) = (\infty, 162)$. Note that there will indeed be two (unbalanced) optimal solutions by symmetry, where the b_i and B_i values in the bivariate case switch values.

Since the constraint curve is approximately linear, a strongly unbalanced solution like the one illustrated will always be optimal when X_1 and X_2 share

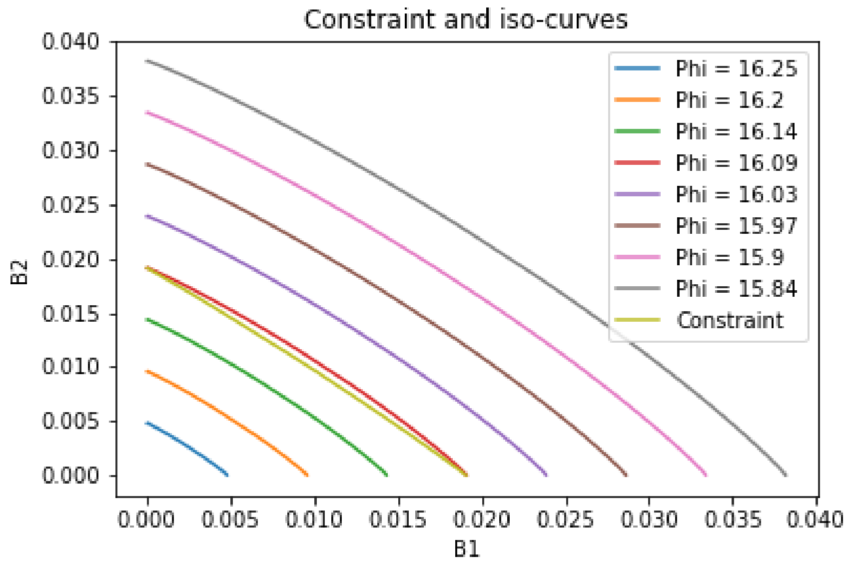


Figure 5.11: Constraint and Iso-contours for example 8 in the truncated normal distribution

the characteristics for the distribution, and this distribution has an increasing hazard rate.

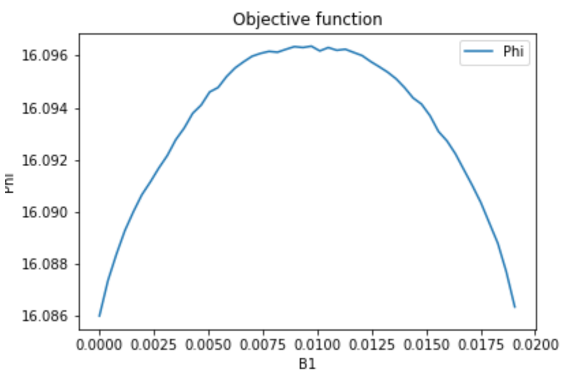


Figure 5.12: Φ versus B_1 in example 8

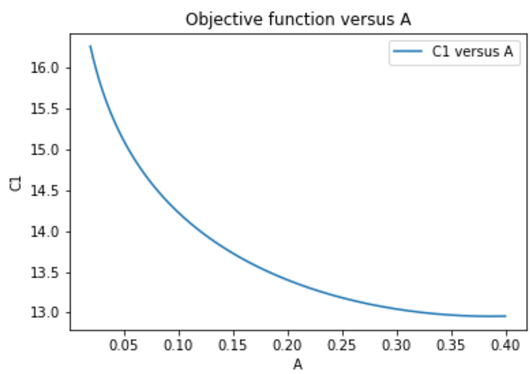


Figure 5.13: C_1 versus A in example 8

From figure 5.12, we indeed have two optimal solutions due to symmetry. The calculated optimal values are $(B_1^*, B_2^*) = (0, 0.019)$ and $(B_1^*, B_2^*) = (0.019, 0)$ in example 8, as verified from table 5.7. Although difficult to determine visually, we also have from figure 5.13 that there is a minimum for our objective function C_1 in the interval $A \in [0, 0.4]$, where $A = 0.386$, as verified by table 5.7.

Exponential distribution

The exponential distribution is able to model continuous data with non-negative ranges and models concerning waiting times between events. It may not be useful for modeling data containing heavy tails and assuming a constant failure rate, which may not always be true.

	$E[X_1]$	$E[X_2]$	sd$[X_1]$	sd$[X_2]$	min(Φ)	Opt C_1	Opt A
Example 10	50	50	50	50	17.96	25.94	0.193
Example 11	60	60	60	60	21.56	25.94	0.193
Example 12	70	70	70	70	25.15	25.94	0.193

Table 5.8: Optimal values for symmetrical exponentially distributed risks

With corresponding optimal contract parameters

	a_1	a_2	b_1	b_2	B_1	B_2
Example 10	82.33	82.33	220.58	241.79	0.0121	0.00794
Example 11	98.79	98.79	264.69	290.15	0.0121	0.00794
Example 12	115.26	115.26	308.81	338.51	0.0121	0.00794

Table 5.9: Optimal contract parameters for symmetrical exponentially distributed risks

We can see that our expected reinsurance cost increases with the increased mean. However, the optimal C_1 remains the same for all our examples, and the same goes for our optimal A. In addition, the parameters for each example increase proportionally with the mean, and all are chosen to be insurance layer contracts.

Our constraint overlaps slightly with an iso-curve at the edges in figure 5.14, bending slightly away from the origin. Therefore, We will have a balanced solution by symmetry since the constraint curve will touch the isocurve for the minimal value at a unique point. Note that the hazard rate of the exponential distribution is constant, meaning that it decreases over time, even though the iso-curves are linear.

We can see from figure 5.15 that we indeed have an unbalanced, optimal, and unique solution in example 11. Although we have slight differences from our numerical outputs in table 5.9, we have that the optimal unbalanced solution is calculated as follows $(B_1^*, B_2^*) = (0.001, 0.008)$, with insurance layer contract for risks X_1 and X_2, as seen in table 5.9. From figure 5.16, we also have a clear minimum for our objective function C_1 in the interval $A \in [0, 0.4]$, where $A = 0.193$.

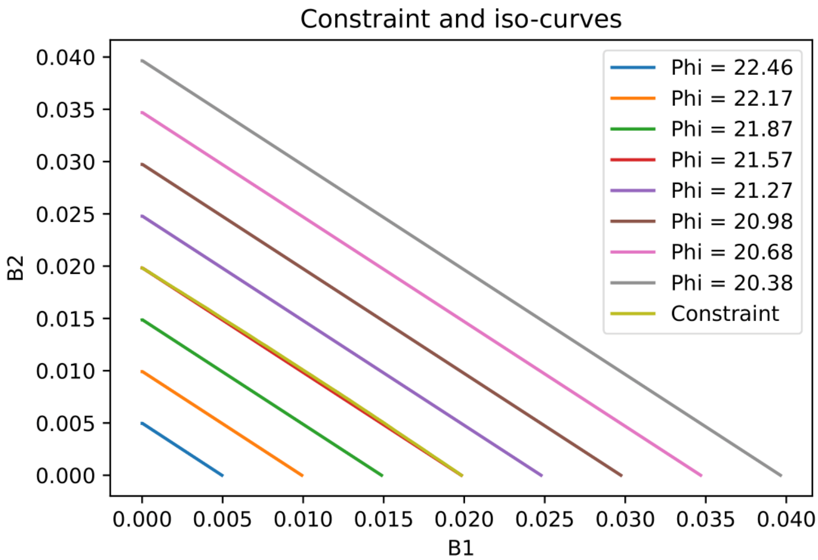


Figure 5.14: Constraint and Iso-contours for example 11 in the exponential distribution

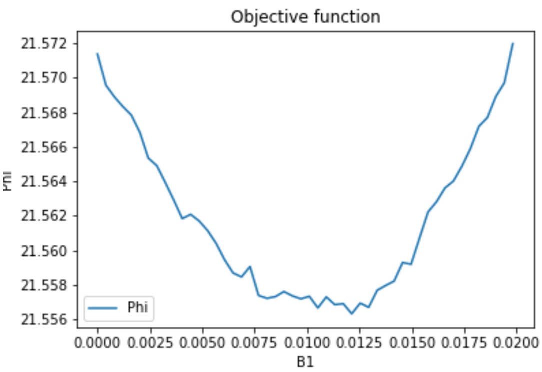


Figure 5.15: Φ versus B_1 in example 11

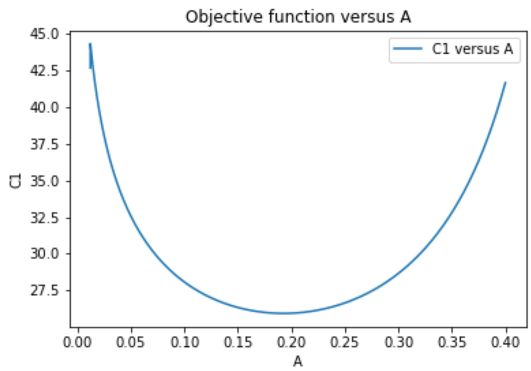


Figure 5.16: C_1 versus A in example 11

5.2.3 Asymmetrical

Continuing, we will look at how the optimal contracts and the corresponding parameters will behave when we introduce asymmetrical conditions, i.e., where the distributions of risks X_1 and X_2 are different.

Log-normal distribution

	$E[X_1]$	$E[X_2]$	$\text{sd}[X_1]$	$\text{sd}[X_2]$	$\min(\Phi)$	Opt C_1	Opt A
Example 3	50	50	30	30	15.87	17.00	0.294
Example 13	50	50	30	50	18.34	19.77	0.253
Example 14	50	100	30	30	18.35	14.09	0.355
Example 15	50	100	30	50	22.11	16.01	0.312

Table 5.10: Optimal values for asymmetrical log-normally distributed risks

With corresponding optimal contract parameters

	a_1	a_2	b_1	b_2	B_1	B_2
Example 3	57.87	57.87	167.24	163.71	0.00706	0.00785
Example 13	61.99	61.50	∞	227.46	0	0.01269
Example 14	52.69	106.83	143.91	365.44	0.01451	$2.539 \cdot 10^{-6}$
Example 15	56.22	112.67	202.97	262.91	0.00253	0.01124

Table 5.11: Optimal contract parameters for log-normally symmetrical distributed risks

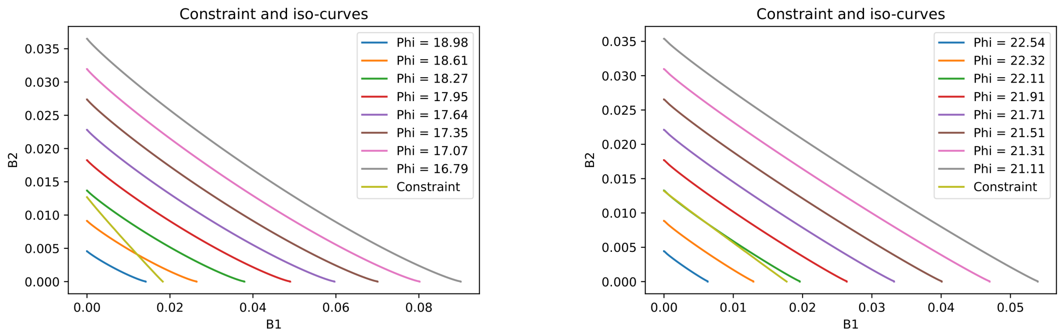
We will use example 3 as a benchmark when comparing asymmetrical log-normal distributed risks. We can see in example 13 that an increase in standard deviation for risk X_2 yields an increased optimal Φ and C_1, in addition to the insurance layer contract for risk X_1 yielding a stop-loss contract. For example 14, we can see that just an increase in the mean for risk X_2 yields an insurance layer contract for both risk X_1 and X_2, yet the risk X_2 is on the verge of being a stop-loss contract when considering the low value of B_2. In reality, this is due to a numerical error by the simulation, and example 14 is deemed a stop-loss contract.

The changes in the mean do not affect the tails as much as in the standard deviation for log-normal distributed risks, which is why we see the effects noted above. Continuing, by increasing both the mean and standard deviation for the risk X_2, we can see that there is no stop-loss contract, but the insurance layer contract for risk X_2 has shifted to define higher values of a_2 and b_2. The insurance layer contract for risk X_1 has increased, both with decreased a_1 and increased b_1.

Figure 5.17 shows iso-curves for the objective function Φ along with the constraint curve. The superlevel sets are convex, and the constraint curve is approximately linear.

Since the two risks have different distributions for both visualized examples, the optimal combination of B_1 and B_2 is not balanced. Detailed calculations for example 13 and 15 yield $(B_1^*, B_2^*) = (0, 0.0013)$ and $(B_1^*, B_2^*) = (0.0025, 0.0112)$, respectively. Corresponding contract b_i^* parameters yield $(b_1^*, b_2^*) = (\infty, 227)$ and $(b_1^*, b_2^*) = (202, 263)$, respectively.

By figure 5.18, we can see a clear minimum when C_1 is plotted as a function of A in example 15, where $A = 0.312$. Note that we have selected $A \in [0, 0.4]$ for our visualization.



Example 13 Example 15

Figure 5.17: Constraint and Iso-contours for example 13 and 15 in the log-normal distribution

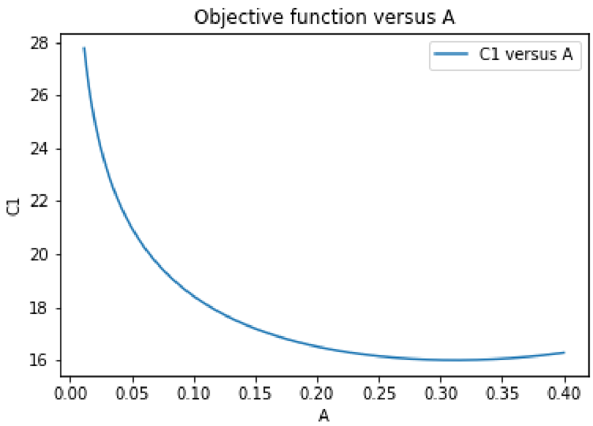
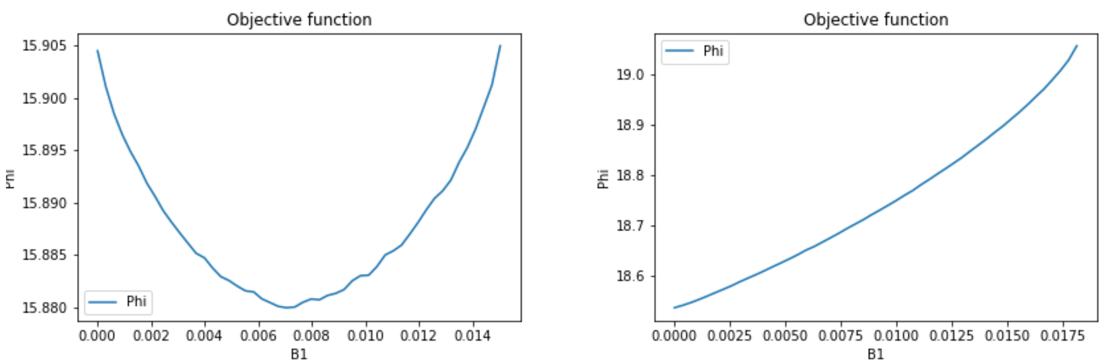
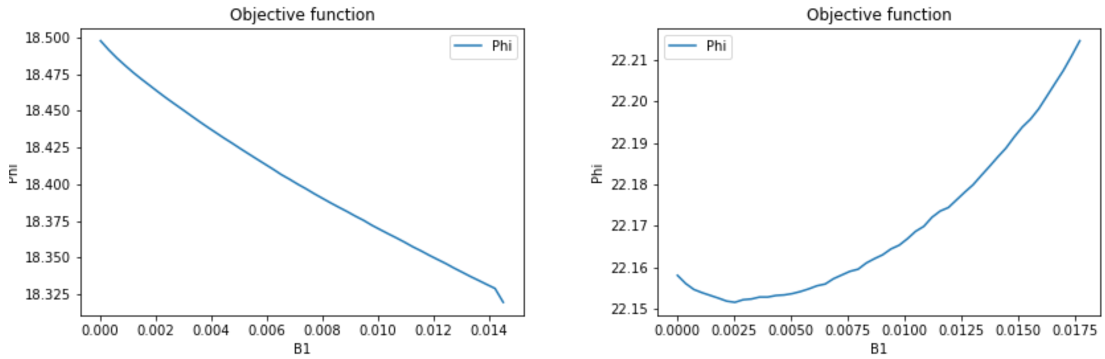


Figure 5.18: Objective function C_1 versus A for example 15



Example 3 Example 13

Looking at Φ compared to B_1 in figure 5.19, we can see in the symmetrical example that a clear minimum is balanced. Hence it is also an optimal and



Example 14 Example 15

Figure 5.19: Φ versus B_1 asymmetrical cases for lognormally distributed risks

unique solution, where $(B_1^*, B_2^*) = (0.007, 0.007)$. Furthermore, we can see that as our standard deviation for risk X_2 is increased, we have an increasing function Φ of B_1, and oppositely when only the mean is increased for risk X_2, where expected reinsurance costs are decreasing. We can, therefore, quickly determine that example 13 has a stop loss contract where $B_1 = 0$, and example 14 has a stop loss where $B_2 = 0$, as verified by table 5.11.

Keeping the above results in mind, we can see that a combined increase for the mean and standard deviation in example 15 yielded an unbalanced, optimal, and unique solution, with a layer insurance contract for both risks.

Pareto distribution

	$E[X_1]$	$E[X_2]$	sd$[X_1]$	sd$[X_2]$	min(Φ)	Opt C_1	Opt A
Example 4	50	50	30	30	16.25	13.83	0.362
Example 16	50	50	30	50	17.90	14.49	0.345
Example 17	50	150	30	30	19.80	11.90	0.400
Example 18	50	150	30	50	25.73	12.64	0.396

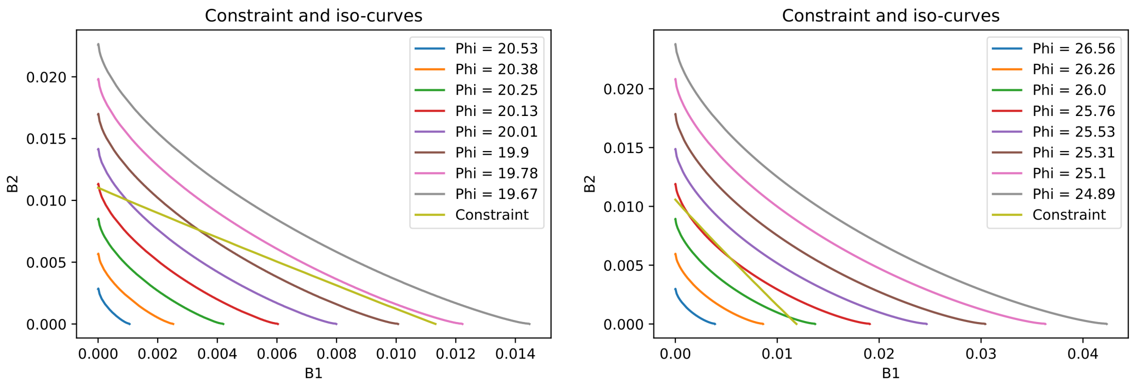
Table 5.12: Optimal values for asymmetrical Pareto distributed risks

With corresponding optimal contract parameters

	a_1	a_2	b_1	b_2	B_1	B_2
Example 4	46.64	46.64	192.04	196.45	0.00525	0.00512
Example 16	47.39	45.51	260.15	211.54	0.00229	0.00845
Example 17	45.07	145.74	162.10	349.37	0.00924	0.00193
Example 18	45.24	142.40	233.57	367.82	0.00315	0.00762

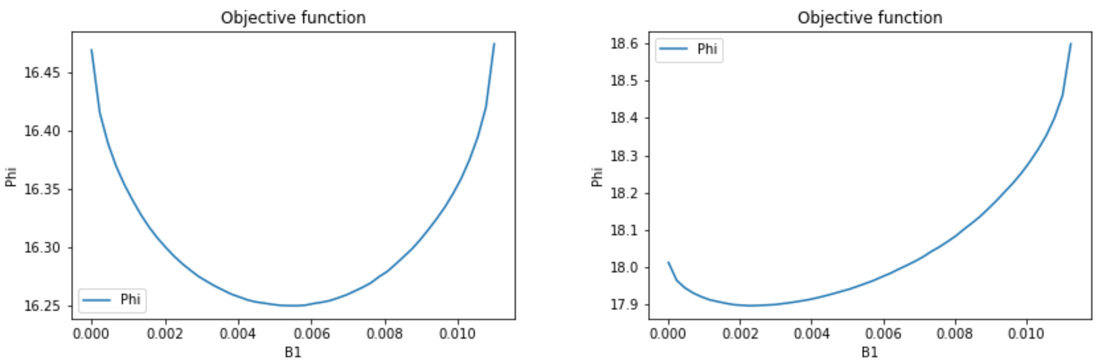
Table 5.13: Optimal contract parameters for asymmetrical Pareto distributed risks

We compare the example 4 to the asymmetrical cases for more depth. We can see that when compared to example 16, the optimal value of Φ and C_1 is naturally increased since one of the risks has an increased standard deviation. However, when the standard deviation of risk X_2 is increased, it seems to affect the contract yielded for the risk X_1. As a result, the insurance layer contract is increased, resulting in a higher total reinsurance cost for the cedent. However, this increase is not too large, so although there is a significant increase in layer $a_1 \times b_1$, this is due to the heavy tail of Pareto distribution.

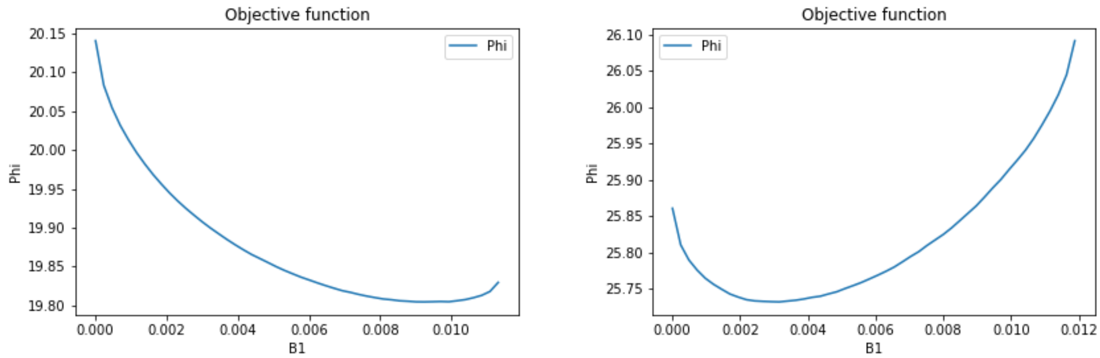


Example 17 Example 18

Figure 5.20: Constraint and Iso-contours for example 17 and 18 in the Pareto distribution



Example 4 Example 16



Example 17 Example 18

Figure 5.21: Φ versus B_1 asymmetrical cases for Pareto distributed risks

As for the mean, when it is increased for the mean in example 17, we can see that the insurance layer contract for risk X_1 has a decreased interval, and the risk X_2 is more affected. Increasing the mean and standard deviation for risk X_2 observes the combined effects, affecting the layer contracts of both X_1 and X_2.

Since the Pareto distribution has a decreasing hazard rate, we expect by Theorem 3.4.7 that Φ to be quasiconvex. This is evident by figure 5.20, which illustrates that the sublevels are convex. Provided that the constraint curve is roughly linear or slightly curved away from the origin, an unbalanced solution similar to the examples above will remain the most advantageous, assuming X_1 and X_2 follow the same distribution with a decreasing hazard rate.

In addition, we have that two risks have different distributions for both visualized examples. The optimal combination of B_1 and B_2 is not balanced. Detailed calculations for example 17 and 18 yield $(B_1^*, B_2^*) = (0.009, 0.002)$ and $(B_1^*, B_2^*) = (0.003, 0.007)$, respectively. Corresponding contract b_i^* parameters yield $(b_1^*, b_2^*) = (162, 349)$ and $(b_1^*, b_2^*) = (234, 368)$, respectively.

Looking more closely at expected reinsurance cost, we can see that in example 16, the objective function will first decrease and then increase for higher values of B_1, yielding an optimal value at $(B_1^*, B_2^*) = (0.002, 0.008)$, verifying that we have an inner point solution which is optimal and unique. The same results are verified for example 17 and 18, as mentioned previously. However, as seen from figure 5.21, Φ increases when there is a higher standard deviation and decreases when there is a higher mean. The combination of these gives us a similar graph to example 16, suggesting that standard deviation has more effect on Φ than the mean.

For example 18, we have expanded our search interval for our optimal value of A when C_1 is minimal. The reason for this is that at $A \in [0, 0.4]$, it is hard to visually determine whether the Pareto distribution yields a clear minimum since numerical results yielded $A = 0.362$.

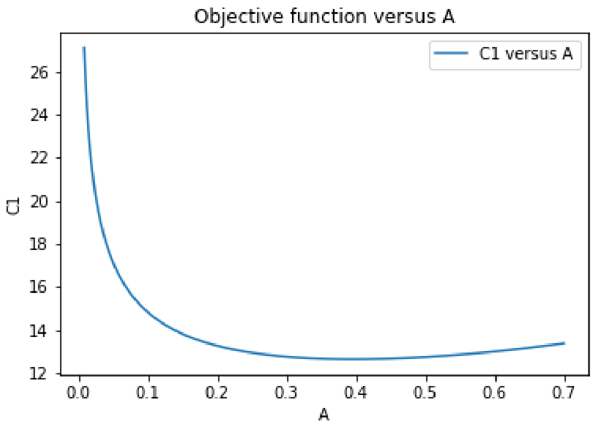


Figure 5.22: Objective function C_1 versus A for example 18

Truncated normal distribution

	$E[X_1]$	$E[X_2]$	sd$[X_1]$	sd$[X_2]$	min(Φ)	Opt C_1	Opt A
Example 7	50	50	30	30	12.19	17.48	0.286
Example 19	50	50	30	40	13.79	19.38	0.258
Example 20	50	100	30	30	15.03	14.30	0.350
Example 21	50	100	30	40	16.04	15.24	0.328

Table 5.14: Optimal values for asymmetrical truncated normally distributed risks

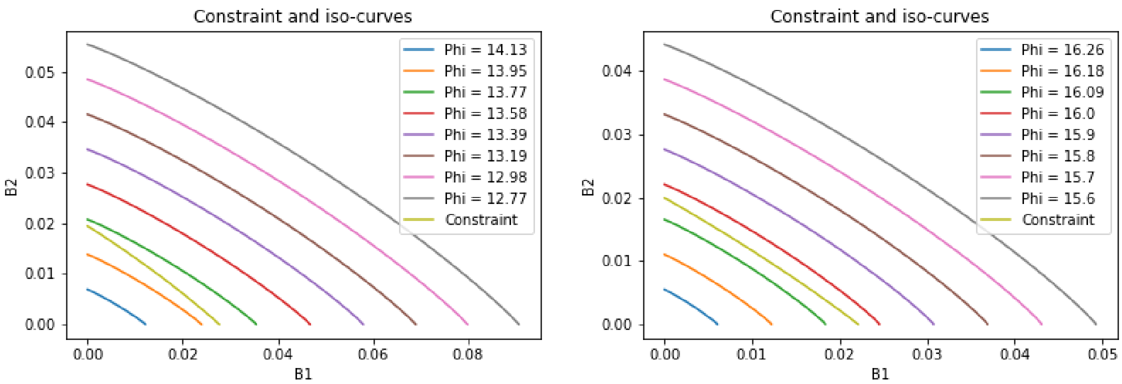
With corresponding optimal contract parameters

	a_1	a_2	b_1	b_2	B_1	B_2
Example 7	65.92	65.92	116.69	205.66	0.00229	$4.586 \cdot 10^{-6}$
Example 19	68.84	70.86	∞	155.27	0	0.01945
Example 20	59.82	111.57	119.18	234.30	0.01950	$3.976 \cdot 10^{-6}$
Example 21	61.82	117.68	∞	183.71	0	0.01998

Table 5.15: Optimal contract parameters for asymmetrical truncated normally distributed risks

As for our truncated normal distribution, when we compare it to the symmetrical example 7, we have some interesting results. We can see for example 7 and 20 that we nearly get stop-loss contracts, and example 19 and 21 yield stop-loss contract contracts. Following the same procedure, we can see that when the standard deviation is increased for X_2, the contract parameters for the risk X_1 are affected. Our total reinsurance cost seems proportional to the mean and standard deviation changes. Still, the risk is spread differently depending on whether the mean or the standard deviation changes. Note that similarly to previous calculations, example 7 and 21 are in reality stop-loss contracts. These small values of B_2 are due to numerical errors in the simulation.

We can see in example 20 that with a change in the mean for risk X_2, the distribution is moved along the x-axis for the distribution, and thus the risk has shifted, yielding higher a_2 and b_2 values for insurance layer contract. However, a change in both the mean and the standard deviation for the risk X_2 experiences a shift in the insurance layer contract for risk X_2, in addition to the increase in b_1 for risk X_1. Observe that when comparing example 19 to 21, the contract parameter a_1 is decreased rather than increased, so this risk is heavily affected by the changes in risk X_2. Similarly to the previous asymmetrical examples, this distribution is also affected more by a difference in the standard deviation rather than the expected value.



Example 19 Example 21

Figure 5.23: Constraint and Iso-contours for example 19 and 21 in the truncated normal distribution

The truncated normal distribution has an increasing hazard rate. Thus, by Theorem 3.4.7 Φ is quasiconcave. Hence the superlevel sets are convex. Moreover, we can see from figure 5.23 that the iso-curves are indeed convex since they bend away from the origin.

Since the two risks have different distributions for both visualized examples, the optimal combination of B_1 and B_2 is not balanced. Detailed calculations for example 19 and 21 yield $(B_1^*, B_2^*) = (0, 0.0195)$ and $(B_1^*, B_2^*) = (0, 0.02)$, respectively. Corresponding contract b_i^* parameters yield $(b_1^*, b_2^*) = (\infty, 155.27)$ and $(b_1^*, b_2^*) = (\infty, 183.71)$, respectively. In both examples, we have $b_1^* > b_2^*$, which means that risk X_1 gets better reinsurance coverage for both examples than risk X_2. This is due to risk X_2 having a higher standard deviation and thus exposing to higher severity claims.

Provided that the constraint curve is roughly linear or slightly curved away from the origin, an unbalanced solution like the above examples will remain the most advantageous, assuming X_1 and X_2 follow the same distribution with an increasing hazard rate.

For our objective function C_1 by figure 5.24, we can see that there is a clear minimum in the interval $A \in [0, 0.4]$, where $A = 0.328$. Note that these values for A are only valid for values $B \leq A$ due to B being calculated as a function of A; $B = K - \sqrt{K^2 - A}$. This avoids negative values for our expected gain $E[G]$.

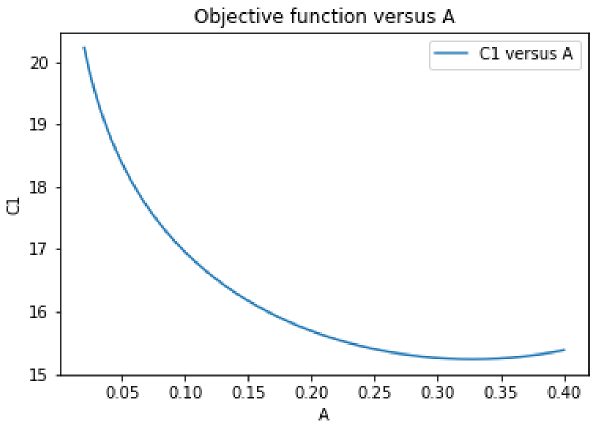
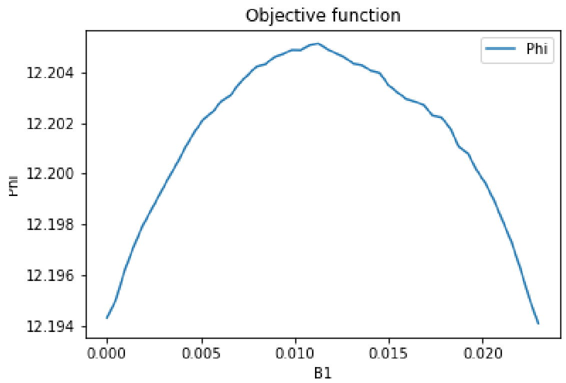
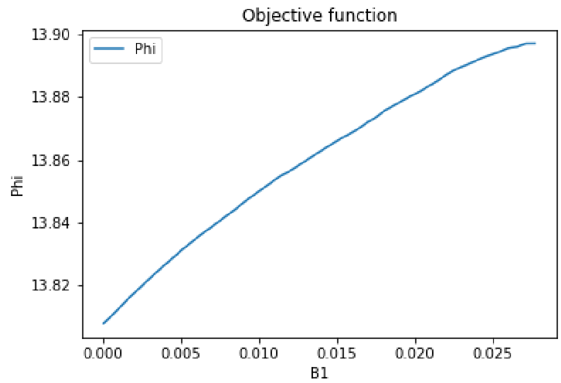


Figure 5.24: Objective function C_1 versus A for example 21

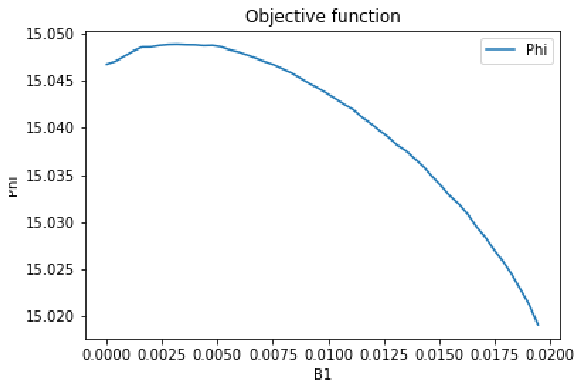
We can see from figure 5.25 the three different asymmetrical examples for truncated normally distributed risks. The first plot, example 19, has an increased standard deviation for risk 2. The expected reinsurance risk then increases as B_1 increases. Thus, the minimum value is $B_1 = 0$, as verified from table 5.15. We can also see a clear minimum when the mean is increased rather than the standard deviation, as Φ will now decrease for higher values of B_1. In the last example, we have both an increase in standard deviation and mean. This will yield two optimal unbalanced solutions, where $(B_1^*, B_2^*) = (0, 0.02)$ and $(B_1^*, B_2^*) = (0.02, 0)$. All of our optimal solutions for the asymmetrical examples are border solutions. Note that the small irregularities in our graph are caused by the small difference in our y-axis, Φ, causing the bisection method to show the small jumps when looking at similar values.



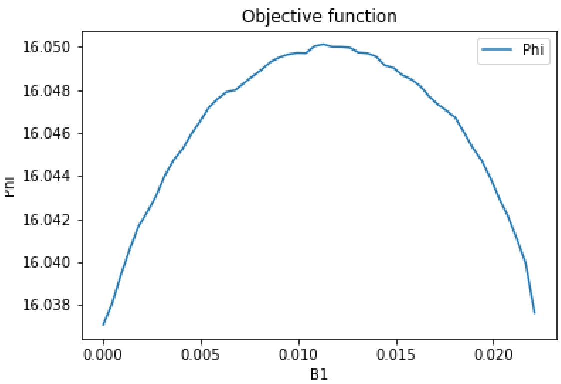
Example 7



Example 19



Example 20



Example 21

Figure 5.25: Φ versus B_1 asymmetrical cases for truncated normally distributed risks

Exponential distribution

	$E[X_1]$	$E[X_2]$	$\text{sd}[X_1]$	$\text{sd}[X_2]$	$\min(\Phi)$	Opt C_1	Opt A
Example 10	50	50	50	50	17.96	25.94	0.193
Example 22	40	60	40	60	17.97	25.95	0.193
Example 23	50	400	50	400	80.90	25.96	0.193

Table 5.16: Optimal values for asymmetrical exponentially distributed risks

With corresponding optimal contract parameters

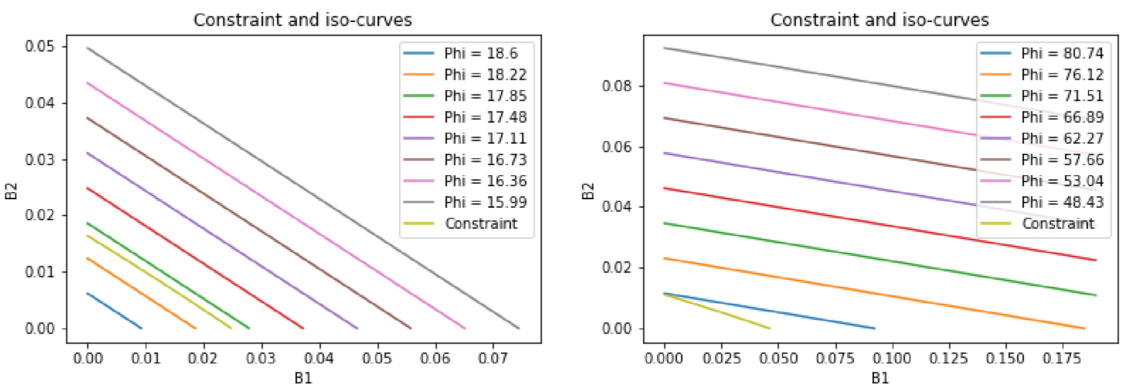
	a_1	a_2	b_1	b_2	B_1	B_2
Example 10	82.32	82.32	220.58	241.79	0.01213	0.00794
Example 22	65.86	98.80	164.90	308.22	0.01620	0.00587
Example 23	82.35	658.78	∞	1798.99	0	0.01114

Table 5.17: Optimal contract parameters for asymmetrical exponentially distributed risks

Lastly, we have the asymmetrical exponential case. We will use the symmetrical example 10 to compare our results. We can see that examples 10 and 22 have different distributions. Still, the optimal value of Φ is almost identical (this could, in reality, be the same, but we have a slight deviation due to numerical error). Furthermore, the optimal values of C_1 and A are almost identical, which is due to the hazard rate of the exponential distribution as a result of the mean being equal to the standard deviation:

$$h(x) = \frac{f(x)}{1 - F(x)} = \frac{\lambda e^{-\lambda x}}{1 - (1 - e^{\lambda x})} = \lambda$$

In the symmetrical case, we have $a_1 = a_2$ and a slight deviation for $b_1 = b_2$. This difference grows if some of the expected value is shifted, i.e., $E[X_1] < E[X_2]$ in example 22. We then get that the mean of risk X_2 is shifted to the right on the x-axis of our distribution and, consequently, will have higher values for our layer-contract parameters a_2 and b_2. In example 23, we can see a more extreme version, where $8 \cdot E[X_1] = E[X_2]$. This results in a stop-loss contract for risk X_1, although a_1 remains the same as in example 10. Moreover, the contract parameters of risk X_2 are significantly increased. Note that these changes in expected value will also involve equal modifications to the standard deviation since the expected value is equal to the standard deviation in our exponential distribution.

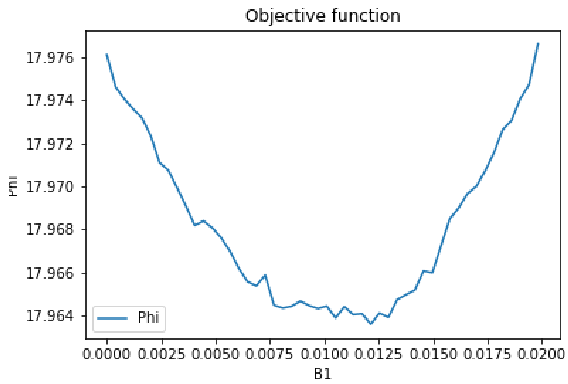


Example 22 Example 23

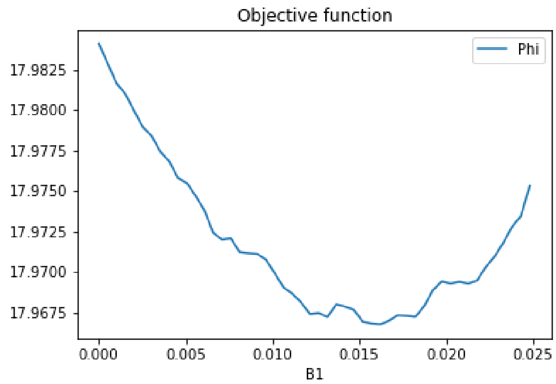
Figure 5.26: Constraint and Iso-contours for example 22 and 23 in the exponential distribution

Since the two risks have different distributions for both visualized examples, the optimal combination of B_1 and B_2 is not balanced. Detailed calculations for example 22 and 23 yield $(B_1^*, B_2^*) = (0.0162, 0.0058)$ and $(B_1^*, B_2^*) = (0, 0.0111)$, respectively. Corresponding contract b_i^* parameters yield $(b_1^*, b_2^*) = (165, 308)$ and $(b_1^*, b_2^*) = (\infty, 1799)$, respectively.

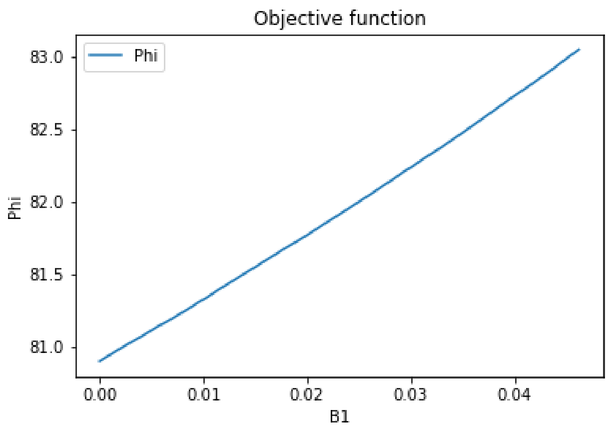
An unbalanced solution is optimal for the asymmetrical cases in the exponential distribution since the constraint curve is approximately linear or slightly bending away from the origin, touching a linear iso-curve for a minimal value at a unique point.



Example 10



Example 22



Example 23

Figure 5.27: Φ versus B_1 asymmetrical cases for exponentially distributed risks

Investigating the functions given in figure 5.27, we can see that examples 10 and 22 have a clear minimal point of Φ when plotted against B_1. Notice, however, that the functions are somewhat irregular. This is due to slight differences of Φ when it moves along B_1, and the bisection method causes the small jumps. As for example 23, we see a clear stop-loss contract where the minimum is $B_1 = 0$ since Φ increases linearly as B_1 increases, as verified by table 5.17.

For our objective function C_1 by figure 5.28, we can see a clear minimum in the interval $A \in [0, 0.4]$, where $A = 0.193$.

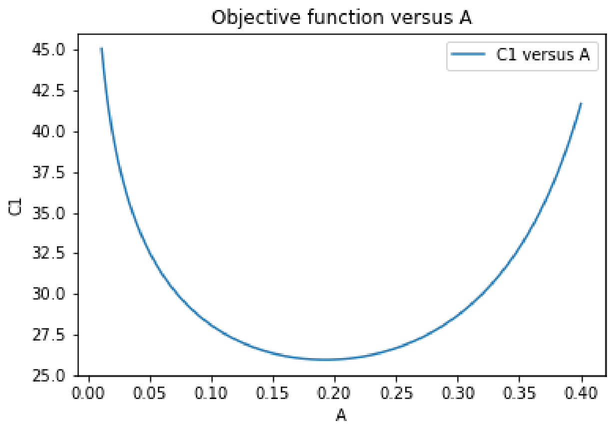


Figure 5.28: Objective function C_1 versus A for example 23

5.2.4 Change in dependency

Finally, let us explore the consequences of modifying the dependency by adjusting the correlation in the interval $\rho \in [-1, 1]$. When dealing with log-normal and truncated normal distributions, an increase in risk can lead to a higher skewness and a heavier tail, resulting in a greater likelihood of extreme losses. In the case of Pareto and exponential distributions, shifting to the left or right may occur depending on the correlation's sign, affecting the mean and variance and, as a result, the likelihood of more extreme losses.

The effect of dependency changes on each distribution depends on its shape, parameters, and risk correlation. A positive correlation between risks allows the risks to move together, which can result in a higher level of risk for the reinsurer. This is because losses in one particular risk are likely to be accompanied by losses in the other risk. On the other hand, if the risks are negatively correlated, the reinsurer's risk level may be reduced. This is because losses in one risk will offset the profits in another. Depending on their properties, some distributions may be more sensitive to correlation changes than others. Therefore, assessing each distribution's sensitivity to correlation changes is crucial when modeling reinsurance contracts.

Let us delve into the truncated normal distribution in this section. This distribution has interesting properties, such as its ability to manage both positive and negative skewness and its flexibility in modeling extreme risks. Moreover, the truncated normal distribution can be truncated on either end, making it ideal for modeling risks with upper or lower limits, such as insurance deductibles. Finally, it is worth noting that this distribution is sensitive to correlation changes, which can considerably impact pricing and risk allocation in multivariate reinsurance contracts.

Algorithm 5 Correlation

```
input ρ, Δ, dist1, dist2, N
c_1 = 1/2 · (√(1+ρ) + √(1−ρ))
c_2 = 1/2 · (√(1+ρ) − √(1−ρ))
if ρ = 0:
    for j = 0, . . . , N do
        u = bisection(Δ)
        drawX_1[j] ← dist1(u[0])
        drawX_2[j] ← dist2(u[1])
    end for
else
    for j = 0, . . . , N do
        u = bisection(Δ)
        g_1 ← gaussian_invcdf(u[0])
        g_2 ← gaussian_invcdf(u[1])
        h_1 ← c_1 · g_1 + c_2 · g_2
        h_2 ← c_1 · g_1 + c_2 · g_2
        v_1 ← gaussian_cdf(h_1)
        v_2 ← gaussian_cdf(h_2)
        drawX_1[j] ← dist1(v_1)
        drawX_2[j] ← dist2(v_2)
    end for
return X_1[j], X_2[j]
```

	$E[X_1]$	$E[X_2]$	sd$[X_1]$	sd$[X_2]$	min(Φ)	Opt C_1	Opt A
$\rho = -0.60$	100	100	30	30	15.00	12.79	0.391
$\rho = 0.00$	100	100	30	30	16.09	12.96	0.386
$\rho = 0.60$	100	100	30	30	16.19	12.97	0.385

Table 5.18: Optimal values for truncated normally distributed risks with change in dependency

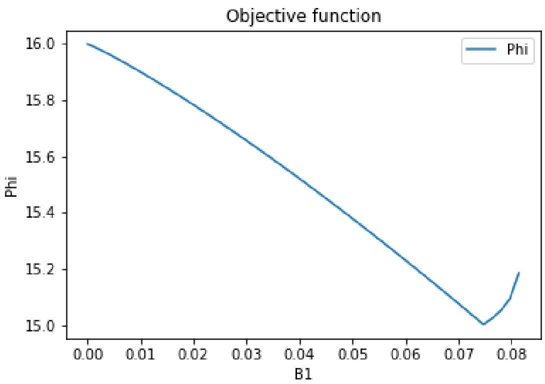
With corresponding optimal contract parameters

	a_1	a_2	b_1	b_2	B_1	B_2
$\rho = -0.60$	108.28	108.28	143.33	158.54	0.07472	0.02575
$\rho = 0.00$	108.67	108.67	∞	162.32	0	0.01909
$\rho = 0.60$	108.71	108.71	172.97	181.08	0.00762	0.00350

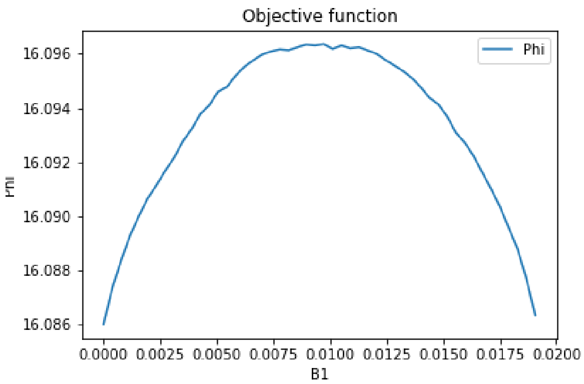
Table 5.19: Optimal parameters for change in dependency ρ

By these tables, we can see that a negative correlation yields a lower Φ, as the level of risk is reduced due to one risk being less accompanied by a loss from the other. So, we also have increased our gain $E[G]$ since optimal A is higher with a negative correlation. Thus the optimal value for our objective function C_1 finds a more optimal minimum value. On the other hand, if the correlation is positive, the risks are allowed to move together; this is because the level of risk is more exposed when the loss of a risk accompanies the loss of another risk.

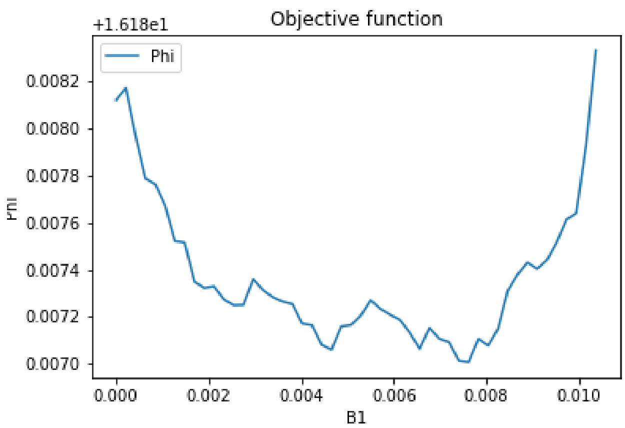
As for our contract parameters, we can see that the positive and negative correlation yields an insurance layer contract for both risks X_1 and X_2. This is due to the increased certainty when both risks are correlated. If there is no dependency between the risks, $\rho = 0.00$, we have that risk X_1 sees an optimal value when there is an insurance layer contract, and risk X_2 does the same for a stop-loss contract.



$\rho = -0.6$



$\rho = 0.0$



$\rho = 0.6$

Figure 5.29: Φ versus B_1 for change in dependency

Based on the information presented in figure 5.29, it is evident that in the case of a negative correlation, the stop-loss contract will be $(B_1^*, B_2^*) = (0.07, 0.03)$ with corresponding $(b_1^*, b_2^*) = (143, 159)$. Adjusting the correlation to $\rho = 0.0$, we get two optimal solutions by symmetry, $(B_1^*, B_2^*) = (0, 0.019)$ and $(b_1^*, b_2^*) = (\infty, 162)$, with corresponding symmetric results $(B_1^*, B_2^*) = (0.019, 0)$ and $(b_1^*, b_2^*) = (162, \infty)$. Lastly, for a positive correlation, we get a clear minimum $(B_1^*, B_2^*) = (0.007, 0.003)$ with $(b_1^*, b_2^*) = (173, 181)$. There are some irregularities for this plot due to small adjustments of Φ, so note that regulating the B_1 parameter has little effect on our objective function.

5.3 Discussion

In chapter 5, we have introduced 4 distributions; log-normal, truncated normal, Pareto, and exponential distribution. We have examined the symmetrical and asymmetrical cases for each distribution and observed various results depending on the distributions.

The log-normal distribution is characterized by a positively skewed shape, not dissimilar to *Weibull* and *Gamma* distribution. For the univariate case, we observed that the log-normal distribution experienced a higher jump for the objective function C_1 when the standard deviation is increased. Consequentially, this also applies to the insurance layer contracts, which also increased in order to capture these potential losses. This is also captured in figures 5.3 and 5.4.

As for the multivariate case, we have that log-normally distributed risks that are symmetrical also captured this change. All of the examples resulted in insurance layer contracts that were unique and balanced solutions. The expected reinsurance cost, Φ, was increased when the mean was increased, resulting from the distribution shifting to the right on the x-axis.

However, C_1 was decreased for an increase in mean, meaning that this insurance contract was preferable if one was to choose. A decrease in standard deviation decreased the expected reinsurance cost and further minimized the objective function C_1. Asymmetrical cases saw similar capabilities, yet an increase in either the mean or the standard deviation for X_2 yielded a stop-loss contract for risk X_1 and X_2, respectively. This is because when the uncertainty is increased for risk X_2, a stop-loss for X_1 is not deemed very expensive for the cedent. Yet, when the mean for X_2 is increased, the corresponding contract sees the lowest C_1 value of all the examples, and the added constant for this risk makes it feasible to have a stop-loss contract when the standard deviation is more minor in comparison. Note that there are some jumps in the visualizations for Φ against B_1, indicating that there is not much to gain from adjusting the B_1 values.

The Pareto distribution is characterized by a heavy tail and a high degree of skewness, which means that extreme events are relatively more likely to occur than other distributions. For example, in the univariate case, an increase in standard deviation increased the objective function C_1, yet the optimal a barely rose. Hence, the resulting value for C_1 because caused by a decrease in the numerator, i.e., the expected gain $E[G]$. The Pareto distribution handles losses with low frequency and high severity better, which is why we do not observe drastic changes in our optimal values and corresponding contract parameters.

As for the multivariate case, Pareto distributed symmetrical risks observed similar results. An increase in standard deviation was less influential on the parameters than an increase in mean. An increase in mean did, however, yield a lower C_1, which is caused by the fact that the mean is proportionally more significant, resulting in less impact from our standard deviation when held constant. The total reinsurance cost was fairly similar, yet the layer insurance contract spanned a larger interval for the increase in standard deviation.

As for the asymmetrical cases, we observed only layer-insurance contracts as well. Similarly to log-normal, an increase in standard deviation for risk X_2 yielded an increase in b_1, and an increase in mean yielded an increase in both a_2 and

b_2. All of these asymmetrical examples contained unbalanced and unique results.

The truncated normal distribution is often used to model the behavior of risks that are constrained to a specific range of values; specifically, it is a normal distribution with values that are boundaries. In the univariate case, truncated normally distributed risks had an increase in optimal a for an increase in standard deviation, similar to log-normally distributed risks. In contrast, the Pareto distribution observed a slight decrease. There was also an increase in optimal b. The increase in C_1 was less than for log-normally distributed risks but more than for Pareto distributed risks.

For the multivariate case, we observed that all examples yielded stop-loss contracts. An increase in mean and standard deviation produced stop-loss contracts for symmetrical distributions, in contrast to Pareto and log-normally distributed risks in the symmetrical case. Moreover, both stop-loss contracts were yielded for risk X_1, while our baseline example 7 resulted in a stop-loss for risk X_2. A mean with twice the value yielded a lower objective function C_1, arguably due to the exact reasons for log-normally and Pareto-distributed risks. The reasoning for the resulting stop-loss contracts may be due to that high-severity claims do not occur, but this does, however, depend on our truncation points. Note that by symmetry, these results yield two optimal and unbalanced solutions.

In the asymmetrical case, we also observed stop-loss contracts only. However, not dissimilar to the previous distributions, an increase in standard deviation for risk X_2 yielded a stop-loss contract for risk X_1, and oppositely when there was an increase for the mean in risk X_2. An increase in mean and standard deviation for risk X_2 resulted in a stop-loss contract for risk X_2, suggesting that the standard deviation has more influence on our model than the mean.

Finally, we have the exponential distribution. The exponential distribution is a probability distribution commonly applied to model waiting times or interarrival times between events. This distribution is beneficial for reinsurance policies that pertain to risks with a temporal component, such as weather-related or mortality risks. In the univariate case, we cannot observe a change in optimal C_1 and corresponding optimal A for an increase in standard deviation without increasing the mean due to the mean being equal to the standard deviation.

Continuing in the multivariate case, we can see that our optimal C_1 and corresponding optimal A remain constant for symmetrical exponentially distributed risks. However, there is a shift in the insurance layer contracts due to our distribution shifting to the right on the x-axis for the probability density function in the exponential distribution. Nevertheless, the optimal B_1 and B_2 values are also held constant, yielding one optimal, unique, and unbalanced solution for all examples. Note that for changing B_1 rigorously, jumps in Φ will occur due to the bisection method, which indicates that we do not see much of an improvement in adjusting the total reinsurance cost. Hence, the exponential distribution is relatively stable when optimizing various distribution parameters.

In the asymmetrical case, we can see that an increase in mean and standard deviation for risk X_2 while simultaneously decreasing these values for risk X_1 yields two insurance layer contracts, both unbalanced, unique, and optimal solutions. Risk X_1 has a decreased interval $a_1 \times b_1$ layer insurance contract, while risk X_2 has an increased interval. This indicates that risk X_1 is not as affected by the adjustments in X_2 like some previous distributions are. An extreme example where risk X_2 has significantly increased these parameters yields a stop-loss contract for risk X_1, which is because the potential losses in risk X_1 are not as bad as they could be in risk X_2, allowing the optimization to be able to cover all possible losses in risk X_1. Similarly to the symmetrical case, these examples also had jumps for the layer insurance contracts, meaning adjusting B_1 has little say in our expected reinsurance cost.

We also briefly investigated changes in dependency for truncated normally distributed risks. A negative correlation had smaller b_i values and very similar a_i values compared to no dependence due to the risks being less accompanied by one another, reducing the exposure to losses. Conversely, a positive correlation allows the risks to be more accompanied by one another, increasing the exposure to a level of risk.

5.4 Conclusion

This Thesis has provided a multivariate reinsurance optimization problem using the objective function C_1 to find the optimal value by reducing the risk measure value-at-risk and increasing expected gain $E[G]$. The main results would either result in an insurance layer contract or a stop-loss contract, depending on the nature of our distribution and the change in dependency for a bivariate insurance contract. We have reviewed univariate and multivariate examples, and the multivariate examples are either symmetrical or asymmetrical. The resulting visualizations and calculations emphasized whether a solution was balanced, unbalanced, had one or several optimal and unique solutions, resulted in stop-loss or insurance layer contract, and the convexity/concavity of the constraint given by our optimization problem.

We have reviewed the log-normal, Pareto, truncated normal and exponential distribution. In addition, this Thesis has also reviewed methodology for optimizing reinsurance contracts for importance sampling simulation and the utilization of analytical results in the exponential case by Lagrange multipliers.

Further work

Further research on optimizing multivariate reinsurance contracts could explore the potential to optimize other risk measures and adjust the objective function. Additionally, the dependency among risks could be explored in greater detail. As for numerical analysis, extending the scope from bivariate ($m = 2$) to trivariate ($m = 3$) cases may also be explored.

Appendices

APPENDIX A

The First Appendix

Definition A.0.1 (*Normal distribution*) If a continuous rv (random variable) X is defined as a *normal distribution* with the following mean and variance

$$E(X) = \mu, \quad \mu \in (-\infty, \infty)$$

$$\text{Var}(X) = \sigma^2, \quad 0 < \sigma$$

then the pdf (probability density function) is given by:

$$f(x; \mu, \sigma) = \frac{1}{\sqrt{2\pi}\sigma} e^{(x-y)^2/(2\sigma^2)}, \quad x \in (-\infty, \infty) \tag{A.1}$$

If a function is given by the normal distribution, it is commonly denoted $f(x) \sim N(\mu, \sigma^2)$. Some noteworthy properties of this distribution is that both the median and mean are the same. The distribution is symmetric, in addition to that both the mean μ and standard deviation σ are the only two characteristics necessary to define the pdf of a rv.

Definition A.0.2 (*Log-normal distribution*) If a non-negative rv X is defined as a *lognormal distribution* if the rv $Y = \ln(X)$ has a normal distribution with the following mean and variance

$$E(X) = e^{\mu + \frac{\sigma^2}{2}}, \quad \mu \in (-\infty, \infty)$$

$$\text{Var}(X) = e^{2\mu + \sigma^2} \cdot \left(e^{\sigma^2} - 1\right), \quad 0 < \sigma$$

then the pdf is given by:

$$f(x; \mu, \sigma) = \begin{cases} \frac{1}{\sqrt{2\pi}\sigma x} e^{-[\ln(x) - \mu]^2/(2\sigma^2)} & x \geq 0 \\ 0 & x < 0 \end{cases} \tag{A.2}$$

Definition A.0.3 (*Truncated normal distribution*) The *truncated normal distribution* is derived from a normal distributed rv X, with a lower and/or upper bound for the rv. $X \sim N(\mu, \sigma^2)$. The truncated normal distribution is defined by $f(x; \mu, \sigma, a, b)$, where $-\infty \leq a < b \leq \infty$, with the rv in the interval $a < x < b$:

$$f(x; \mu, \sigma, a, b) = \frac{1}{\sigma} \frac{\phi\left(\frac{x-\mu}{\sigma}\right)}{\Phi\left(\frac{b-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{\sigma}\right)} \tag{A.3}$$

$f = 0$ if $a < x < b$ is not valid.

Definition A.0.4 (*Pareto distribution*) The *Pareto distribution* is given by a continuous random variable X if its probability density function is given by:

$$f_X(x; x_m, \alpha) = \begin{cases} \frac{\alpha x_m^\alpha}{x^{\alpha+1}} & x \geq x_m \\ 0 & x < x_m \end{cases} \tag{A.4}$$

where the scale parameter x_m is the minimum possible value of X, and the shape parameter α is known as the *tail index*. The corresponding expected value and variance are given respectively as:

$$E(X) = \begin{cases} \infty & \alpha \leq 1 \\ \frac{\alpha x_m}{\alpha - 1} & \alpha < 1 \end{cases}$$

and

$$Var(X) = \begin{cases} \infty & \alpha \in (1, 2] \\ \left(\frac{x_m}{\alpha-1}\right)^2 \frac{\alpha}{\alpha-2} & \alpha > 2 \end{cases}$$

Note that the variance does not exist if $\alpha \leq 1$.

Definition A.0.5 (*Gamma distribution*) The *gamma distribution* is a continuous probability distribution that takes on two parameters. *Shape parameter* $\alpha \in \mathbb{R}^+$ and *rate parameter* $\beta = \frac{1}{\theta} \in \mathbb{R}^+$, where θ is denoted as the *scale parameter*. The rv is defined as $X \sim \Gamma(\alpha, \beta)$.

$$E(X) = \frac{\alpha}{\beta}, \quad \alpha, \beta > 0$$

$$\text{Var}(X) = \frac{\alpha}{\beta^2}, \quad \alpha, \beta > 0$$

The pdf is given by:

$$f(x; \alpha, \beta) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x}, \quad x > 0 \tag{A.5}$$

With $\Gamma(\alpha)$ being denoted as the *gamma function*.

Definition A.0.6 (*Exponential distribution*) X is said to have an *exponential distribution* with parameter λ if the pdf of X is

$$f(x; \lambda) = \begin{cases} \lambda e^{-\lambda x} & x \geq 0 \\ 0 & \text{otherwise} \end{cases} \tag{A.6}$$

The exponential pdf is a special case of the general gamma pdf (A.5), where $\alpha = 1$ and $\beta = \frac{1}{\lambda}$.

We then get:

$$E(X) = \alpha\beta = \frac{1}{\lambda}, \quad \lambda > 0$$

$$\text{Var}(X) = \sigma^2 = \alpha\beta^2 = \frac{1}{\lambda^2}, \quad \lambda > 0$$

APPENDIX B

The Second Appendix

Figure 3.1:

```
df_pnorm ← curve(pnorm, from = 0, to = 10, n = 1000, plot = FALSE)
df_qnorm ← curve(qnorm, from = 0, to = 1, n = 1000, plot = FALSE)
df_pnorm$x ← df_pnorm$x
df_qnorm$x ← df_qnorm$x
par(mfrow=c(2,2))
cdf ← plot(df_pnorm$x, df_pnorm$y, type = 'l', main = 'Cumulative Distribution Function', xlab = 'Risk', ylab = 'Quantile')
inverse_cdf ← plot(df_qnorm$x, df_qnorm$y, type = 'l', main = 'Inverse Cumulative Distribution Function', xlab = 'Quantile', ylab = 'Risk')
survival_function ← plot(df_pnorm$x,1-df_pnorm$y, type = 'l', main = 'Survival Function', xlab = 'Risk', ylab = 'Quantile')
inverse_survival_function ← plot(df_qnorm$x,1-df_qnorm$y, type = 'l', main = 'Inverse Survival Function', xlab = 'Quantile', ylab = 'Risk')
```

Figure 4.2:

```
gamma ← 0.1
theta ← 0.2
alpha ← 0.01
A ← seq(0.059, 0.35, by = 0.001)
C_1 ← function(A){(-log(A))/(gamma - theta*(A - B))}
B ← A*(1 - log(A)) - sqrt((A^2)*(1 - log(A))^2 - alpha)
C ← C_1(A)
plot(A, C, type = 'l', xlab = 'A', ylab = 'C1', col = 'blue', lwd = 2, main = 'Objective function versus A', sub = 'Optimal solution in the exponential case')
legend(0.06, 33, col = 'blue', legend = c('Objective function'), lty = 1, lwd = 2, text.font = 4, bg = 'lightblue')
```

Python codes for optimization of univariate and multivariate reinsurance contracts are given in the following repository:

https://github.com/drdrechr/SMR5960_simulation

Bibliography

[Bøl14] E. Bølviken. *Computation and Modelling in Insurance and Finance*. Cambridge University Press, 2014, pp. xxvi + 685.

[Che14] Cheung. *Optimal Reinsurance Under General Law-Invariant Risk Measures*. Scandinavian Actuarial Journal, 2014.

[DB12] J. L. Devore and K. N. Berk. *Modern Mathematical Statistics with Applications*. Springer, 2012, pp. xii + 845.

[HC20] A. B. Huseby and D. Christensen. *Optimal Reinsurance Contracts in the Multivariate Case*. Research Publishing, Singapore, 2020.

[Hus22] A. B. Huseby. *Optimizing Multiple Reinsurance Contracts*. Research Publishing, Singapore, 2022.

[Hus23] A. B. Huseby. *Lecture notes STK4400 "Risk and Reliability Analysis"*. University of Oslo, Spring 2023.

[Mør17] K. Mørken. *Numerical Algorithms and Digital Representation*. Department of Mathematics, University of Oslo, 2017, pp. vi + 472.

[She19] R. Sheldon M. *Introduction to Probability Models*. Academic Press, 2019, pp. xv + 826.

[Ury00] S. P. Uryasev. *Probabilistic Constrained Optimization*. Springer, 2000, pp. xii + 307.