Optimal reinsurance strategy with bivariate Pareto risks

by

Evelyn Susanne Gaus

A Thesis Submitted in
Partial Fulfillment of the
Requirements for the Degree of

Master of Science
in
Mathematics

at

The University of Wisconsin-Milwaukee
May 2014

ABSTRACT

Optimal reinsurance strategy with bivariate Pareto risks

by

Evelyn Susanne Gaus

The University of Wisconsin-Milwaukee, 2014
Under the Supervision of Professor Wei Wei

In an insurance, one is often concerned with risks and extreme events which can cause large losses. The Pareto distribution is often used in actuarial sciences for modeling large losses. This thesis extends the study of Cai and Wei (2011) by considering a two-line business model with positive dependence through stochastic ordering (PDS) risks, where the risks are bivariate Pareto distributed. Cai and Wei (2011) showed that in individual reinsurance treaties the excess-of-loss treaty is the optimal reinsurance form for an insurer with PDS risks. We derive explicit expressions for the optimal retention levels in the excess-of-loss treaty by considering several risk functions including the criteria of minimizing the variance, minimizing moments of higher order and minimizing moments of fractional order of the total retained loss of the insurer. This will be followed by a comparison of retentions for different choices of the parameters of the bivariate Pareto distribution.

Table of Contents

List of Figures

List of Tables

ACKNOWLEDGEMENTS

I would like to thank Prof. Wei Wei who helped me developing the results of this thesis. Moreover, I would like to thank Prof. Eric Key and Prof. Vytaras Brazauskas for their feedback and for being part of the committee.

Chapter 1

Introduction

Reinsurance is a special sector in the insurance business. An insurer transfers risks to a reinsurance company to protect itself against large losses. Therefore, the insurer has a greater security for its solvency and equity. Furthermore, the insurance company can decrease its required risk capital or can increase its underwriting capacity in both number and size of risks. Reinsurance is divided into proportional reinsurance and non-proportional reinsurance. In proportional reinsurance the reinsurer receive a predetermined share or portion of the premium the insurer charges in its insurance contracts. In compensation the reinsurer indemnifies the insurer against losses in the same portion or share in the covered insurance contracts. Non-proportional reinsurance consists of the excess-of-loss treaty and the stop-loss treaty. In the excess-of-loss treaty the reinsurer indemnifies the insurer against losses that exceed a specified amount, known as the insurer's retention or deductible. In the stop-loss treaty the reinsurer indemnifies the insurer against losses up to a treaty limit.

1.1 Literature review

This thesis is based on the paper "Optimal reinsurance with positive dependent risks" by Jun Cai and Wei Wei. In this paper Cai and Wei considered positively dependent through stochastic ordering (PDS) risks in the individual risk model.

They proved that in individualized reinsurance treaties the excess-of-loss treaty is the optimal reinsurance form for an insurer with PDS dependent risks. Previously Denuit and Vermandele (1998) had considered the problem of the optimal reinsurance strategy in the individual risk model. They showed that the excess-of-loss reinsurance with equal retentions for each line of business is the optimal reinsurance strategy if the risks are exchangeable random variables. Cai and Wei (2011) extended the study of Denuit and Vermandele (2011) on individualized reinsurance treaties to dependent risks. After defining the notion of several positive dependent risks, Cai and Wei (2011) proved that convolution preservation of the convex order for PDS random variables holds. They showed that the optimal reinsurance in the individual risk model with PDS risks is the excess-of-loss treaty. Moreover, Cai and Wei (2011) derived explicit expressions for the optimal retention in the excess-of-loss treaty in an insurance business model with two lines of business under the criteria of minimizing the variance of the total retained loss of the insurer and maximizing the expected exponential utility for insurer.

1.2 Model introduction

Assume the insurer has $n = 2$ lines of business. Let the random variable X_i denote the loss in line i, with $i = 1, 2$. Furthermore, we consider the individual risk model, where the total loss of the insurer is given by $S_2 = X_1 + X_2$, that is the total loss modeled as the sum of the individual losses. The individual risk model is especially used in life and health insurance. The insurer applies a reinsurance strategy $I(x)$ to transfer risks of big losses to the reinsurer. Let $I_i(x)$ be increasing for $x \geq 0$ and satisfying $0 \leq I_i(x) \leq x$ for $i = 1, 2$. $I_i(x)$ is called the reinsurance strategy of line i. Therefore, the insurer retains the total loss $S_2^I = I_1(X_1) + I_2(X_2)$ and the reinsurer covers the remaining loss $S_2 - S_2^I$ with reinsurance strategy $I = (I_1, I_2)$.

The premium the insurer has to pay the reinsurer for taking on parts of the risks is calculated by the expected value principle, as in Denuit and Vermandele (1998).

The expected value principle is given by $(1+\theta_R)\mathbb{E}(S_2 - S_2^I)$ and equal to a constant P, where $\theta_R > 0$ denotes the security loading of the reinsurer.

$$(1+\theta_R)\mathbb{E}(S_2 - S_2^I) = P$$

This is equivalent to assuming that the expected total retained loss $\mathbb{E}(S_2^I)$ is fixed and equal to $p > 0$. Solving this equation for p, we get

$$\mathbb{E}(S_2^I) = p = \mathbb{E}(S_2) - \frac{P}{1+\theta_R}.$$

We will define in Chapter 2.2 that the range of p is given by $\left[\theta_1 + \theta_2, \frac{\alpha(\theta_1+\theta_2)}{\alpha-1}\right)$.

As in Cai and Wei (2011) for each $p > 0$ we define the class

$$\begin{aligned} D_2^p = \big\{ I = (I_1, I_2) \big| I_i(x) &\text{ is increasing in } x \geq 0 \text{ with } 0 \leq I_i(x) \leq x \\ &\text{for } i = 1,2 \text{ and } \mathbb{E}(S_2^I) = p \big\} \end{aligned}$$

of admissible reinsurance strategies. For a given positive convex function u on $(0, \infty)$ we wish to minimize $\mathbb{E}(u(S_2^I))$ over D_2^p. The function u is called risk function. The reinsurance strategy in the excess-of-loss treaty with two lines of business is given by $I_i(x) = x \wedge d_i$ for $i = 1, 2$ with the retention vector (d_1, d_2).

Assume that (X_1, X_2) are positive dependent risks, particularly PDS risks. In the individual risk model with individualized reinsurance treaties one is often concerned with positively dependent risks. This means that if we consider an insurance business with two lines of business it is very probable that if a loss in one line occurs, then a loss in the other line occurs as well. For example, consider an insurance with two lines of business, Health/Life and Property/Casualty. An event such as Hurricane Katrina can cause extremely large losses in both lines. It's reasonable to assume that the property losses and the number of dead or injured people are positively dependent.

Based on Cai and Wei (2011) we derive explicit expressions for the optimal retention vector (d_1^*, d_2^*) in the bivariate Pareto case such that

$$\mathbb{E}[u(X_1 \wedge d_1^* + X_2 \wedge d_2^*)] = \inf_{(d_1,d_2)\in L} \mathbb{E}[u(X_1 \wedge d_1 + X_2 \wedge d_2)] \tag{1.1}$$

where for $p > 0$

$$L = \left\{ (d_1, d_2) \in [0,\infty)^2 \middle| \int_0^{d_1} \bar{F}_1(x)dx + \int_0^{d_2} \bar{F}_2(x)dx = p \right\}.$$

and u is convex.

In this thesis we will focus on risks that are bivariate Pareto distributed and work with three different risk functions. In Chapter 2 we will recall the definitions of several Pareto distributions and examine some characteristics of the Pareto distribution. Furthermore, we will recall notions of positive dependence. In Chapter 3, we derive optimal retention levels for three different risk functions. In the first case, where we minimize the variance, we will use Theorem 4.4 of Cai and Wei (2011). In the other two cases, where we minimize moments of higher order and of fractional order, we derive explicit expressions for the optimal retentions (d_1^*, d_2^*) such that (1.1) holds. Finally, in Chapter 4 we analyze the optimal retentions for different choices of the parameter of the bivariate Pareto distribution.

Chapter 2

Preliminaries

In this chapter we will review the properties of several Pareto distributions and notions of positive dependence. Furthermore, we will derive important results for further calculations in Chapter 3 such as the linear transformation between $ParetoII$ and the $Lomax$ distribution and the conditional distribution of bivariate Pareto distributed random variables. Afterwards, we derive an explicit expression of d_2 in terms of d_1 such that (1.1) holds.

2.1 Pareto Distribution

The Pareto distribution belongs to the class of heavy-tailed distributions and is often used in insurance for modeling large losses that exceed a specific threshold. A heavy-tailed distribution means, roughly speaking, that there is a relatively large probability for the occurrence of events that cause large losses. Furthermore, the future worst case is expected to be much worse than the current worst case. To continue the previous example, Hurricane Katrina is a good example for heavy-tailed losses. Hurricane Andrew (1992) was the hurricane that caused the largest losses so far in the American history until Hurricane Katrina occurred. Hurricane Katrina caused insured losses in the amount of $62,200 in comparison to $17,000, the insured loss Hurricane Andrew caused[1].

[1]Original Values of Munich Re 2012 in Mio. US$, Geo Risks Research, NatCatSERVICE, August 2012

In the literature many different types of the Pareto distribution are discussed. In this thesis we consider the Pareto distribution of Type 1 ($ParetoI$), Pareto distribution of Type 2 ($ParetoII$), Lomax distribution ($Lomax$) and the bivariate Pareto distribution of Type 1.

The distribution function of $X \sim ParetoI(\alpha, \sigma)$ is defined as

$$F(x) = 1 - \left(\frac{\sigma}{x}\right)^{\alpha},$$

with density function

$$f(x) = \frac{\alpha\sigma^{\alpha}}{x^{\alpha+1}},$$

where $x > \theta > 0$ and $\alpha > 0$.

For $X \sim ParetoII(\alpha, \sigma, \mu)$ the distribution function is defined as

$$F(x) = \left(1 + \frac{x-\mu}{\sigma}\right)^{-\alpha}, x > \mu, \mu \in \mathbb{R}, \sigma > 0, \alpha > 0.$$

If $\mu = 0$, the $ParetoII$ distribution is known as the $Lomax$ distribution. Mardia (1962) defined the density function of the bivariate Pareto distribution of Type 1 for a vector of random variables (X_1, X_2) as

$$f(x_1, x_2) = \begin{cases} \alpha(\alpha+1)(\theta_1\theta_2)^{\alpha+1}(\theta_2 x_1 + \theta_1 x_2 - \theta_1\theta_2)^{-(\alpha+2)} & , x_i > \theta_i > 0, i = 1, 2 \\ 0 & , \text{otherwise} \end{cases}$$

where $\alpha > 0$. The random variables X_i have the marginal distribution $ParetoI(\alpha, \theta_i)$ with parameters $\alpha > 0$ and $\theta_i > 0$ for $i = 1, 2$.

The Pareto distribution can be used to model losses whose values exceed a specific threshold. The parameter θ is the minimal possible value of the risk X and therefore a lower bound for X. The shape parameter α is the tail index. The tail index measures how fast the survival function decays to 0. For a heavy-tailed distribution, the survival function decays slowly to 0.

A helpful result for further calculations is the linear transformation between a *Lomax* and a *ParetoII* distributed random variable. Let $X \sim Lomax(\alpha, \sigma)$ with $\alpha > 0$ and $\sigma > 0$. Then it holds for the linear transformation $X + \mu$, that $X + \mu \sim ParetoII(\alpha, \sigma, \mu)$ with parameters $\alpha > 0, \sigma > 0, \mu \in \mathbb{R}$ and $x > \mu$. This is true, since $\mathbb{P}(X + \mu > x) = \mathbb{P}(X > x - \mu) = \left(1 + \frac{x-\mu}{\sigma}\right)^{-\alpha}$.

Furthermore, an important result for further calculations is the following conditional distribution of bivariate Pareto distributed random variables.

Lemma 2.1.1. *Let (X_1, X_2) be bivariate Pareto distributed with parameters $\theta_1 > 0, \theta_2 > 0$ and $\alpha > 0$. The conditional distribution of X_2 given $X_1 > d_1$ is $ParetoII(\alpha, \sigma_2, \mu_2)$ with $\mu_2 = \theta_2$ and $\sigma_2 = \frac{d_1\theta_2}{\theta_1}$ and the conditional distribution of X_1 given $X_2 > d_2$ is $ParetoII(\alpha, \sigma_1, \mu_1)$ with $\mu_1 = \theta_1$ and $\sigma_1 = \frac{d_2\theta_1}{\theta_2}$.*

Proof. It holds that

$$\mathbb{P}(X_2 > s | X_1 > d_1) = \left(1 + \frac{s - \theta_2}{\frac{d_1\theta_2}{\theta_1}}\right)^{-\alpha}.$$

Thus, X_2 given $X_1 > d_1$ has $ParetoII(\alpha, \sigma_2, \mu_2)$, with $\mu_2 = \theta_2$ and $\sigma_2 = \frac{d_1\theta_2}{\theta_1}$. Analogeous, it follows X_1 given $X_2 > d_2$ has $ParetoII(\alpha, \sigma_1, \mu_1)$ with $\mu_1 = \theta_1$ and $\sigma_1 = \frac{d_2\theta_1}{\theta_2}$. For the detailed calculations see Apendix. □

The distribution of a linear shifted *ParetoII* distributed random variable is still *ParetoII*. Let $X \sim ParetoII(\alpha, \sigma, \mu)$, then $X + d \sim ParetoII(\alpha, \sigma, \mu')$ with $\mu' = \mu + d$ for a constant $d \in \mathbb{R}$. This is true since $\mathbb{P}(X + d > x) = \mathbb{P}(X > x - d) = \left(1 + \frac{x-(\mu+d)}{\sigma}\right)^{-\alpha}$.

The notions of stochastically increasing (SI) and positively dependent through stochastic ordering (PDS) are given by the following definition. Recall, the support of a random variable Y, denoted by $S(Y)$, is a Borel set of $\mathbb{R}$ such that $\mathbb{P}\{Y \in S(Y)\} = 1$.

Definition 2.1.2. *A random vector* $(X_1, \ldots, X_n)$ *is said to be stochastically increasing (SI) in the random variable* Y*, denoted as* $(X_1, \ldots, X_n) \uparrow_{SI} Y$*, if* $\mathbb{E}[u(X_1, \ldots, X_n)|Y = y]$ *is increasing in* $y \in S(Y)$ *for any increasing function* $u : \mathbb{R}^n \to \mathbb{R}$ *such that the conditional expectation exists.*
Furthermore, the random vector $(X_1, \ldots, X_n)$ *is said to be positively dependent through the stochastic ordering (PDS) if* $(X_1, \ldots, X_{i-1}, X_{i+1}, \ldots, X_n) \uparrow_{SI} X_i$ *for any* $i = 1, \ldots, n$*.*

Using this definition, we can show that a bivariate Pareto distributed random vector is PDS.

Lemma 2.1.3. *If* (X_1, X_2) *is bivariate Pareto distributed with parameters* θ_1, θ_2 *and* $\alpha > 0$*, where* $x_1 > \theta_1 > 0$ *and* $x_2 > \theta_1 > 0$*, then* (X_1, X_2) *is PDS.*

Proof. It holds that (X_1, X_2) is PDS if $X_1 \uparrow_{SI} X_2$ and $X_2 \uparrow_{SI} X_1$.
If for any $x_1 \in \mathbb{R} : \mathbb{P}(X_1 > x_1|X_2 = x_2)$ is increasing in $x_2 \in S(X_2)$, then $X_1 \uparrow_{SI} X_2$.
Analogous, $X_2 \uparrow_{SI} X_1$ if for any $x_2 \in \mathbb{R} : \mathbb{P}(X_2 > x_2|X_1 = x_1)$ is increasing in $x_1 \in S(X_1)$.
Therefore, to prove $X_2 \uparrow_{SI} X_1$ we have to show that $\mathbb{P}(X_2 > x_2|X_1 = x_1)$ is increasing in $x_1 \in S(x_1)$ for all $x_2 \in \mathbb{R}$. In the following we will just go over the steps, for the detailed calculations see the appendix.
The conditional probability of X_2 given $X_1 = x_1$ is given by

$$\mathbb{P}(X_2 > x_2|X_1 = x_1) = (\theta_2 x_1)^{\alpha+1}(\theta_2 x_1 + \theta_1 x_2 - \theta_1\theta_2)^{-(\alpha+1)}.$$

The increasing property holds, since for $s, t \in S(x_1)$ with $s \leq t, x_2 \in \mathbb{R} : \mathbb{P}(X_2 > x_2|X_1 = s) \leq \mathbb{P}(X_2 > x_2|X_1 = t)$ is increasing in $x_1 \in S(x_1)$. With analogous calculations, it follows that $X_1 \uparrow_{SI} X_2$. Therefore, (X_1, X_2) is PDS. □

Notice, that throughout this thesis the survival function is $\bar{F}(x) = 1 - F(x) > 0$.

2.2 Explicit expression for d_2

In this section we want to derive a mapping from d_1 to d_2 to have an expression for d_2 in terms of d_1. Recall, the set L for $p > 0$ is given by

$$L = \left\{ (d_1, d_2) \in [0, \infty)^2 \middle| \int_0^{d_1} \bar{F}_1(x)dx + \int_0^{d_2} \bar{F}_2(x)dx = p > 0 \right\}.$$

For $p > 0$, we consider the equation

$$\int_0^{d_1} \bar{F}_1(x)dx + \int_0^{d_2} \bar{F}_2(x)dx = p$$

Note that we have a symmetric structure here and can solve this equation for d_1 or d_2. This symmetric structure is very interesting for further considerations of three lines of business or more. In this thesis we solve this equation for d_2. Since $\int_0^{d_1} \bar{F}_1(x)dx + \int_0^{d_2} \bar{F}_2(x)dx < \mathbb{E}(X_1) + \mathbb{E}(X_2)$, the range of p is bounded from above by $p < \mathbb{E}(X_1)+\mathbb{E}(X_2)$. Otherwise, in terms of the set L, if $p = \mathbb{E}(X_1)+\mathbb{E}(X_2)$ then $L = \{(\infty, -\infty)\}$. If $p > \mathbb{E}(X_1) + \mathbb{E}(X_2)$ then there exists no solution for (d_1, d_2), thus $L = \emptyset$. Moreover, we assume that $p \geq \theta_1+\theta_2$ because θ_1 and θ_2 are the minimal values for the losses X_1 and X_2. Thus, p has the range $p \in \left[\theta_1 + \theta_2, \frac{\alpha(\theta_1+\theta 2)}{\alpha-1}\right)$. We also need following Lemma of Cai and Wei (2011).

Lemma 2.2.1. *On the set L, the mapping from d_1 to d_2 is one-to-one. Denote the mapping as $d_2 = L(d_1)$. Then, $L(d_1)$ is continuous, differentiable and strictly decreasing in d_1, with $\frac{\partial d_2}{\partial d_1} = -\frac{\bar{F}_1(d_1)}{\bar{F}_2(d_2)}$.*

The proof of this Lemma can be found in Lemma 4.1 in Cai and Wei (2011).

Theorem 2.2.2. *If (X_1, X_2) is bivariate Pareto distributed, with parameters θ_1, θ_2 and α, where $x_1 > \theta_1 > 0$, $x_2 > \theta_2 > 0$ and $\alpha > 1$ then*

$$L(d_1) = d_2 = \left(\frac{\alpha(\theta_1 + \theta_2) - \theta_1^\alpha d_1^{1-\alpha} + p(1-\alpha)}{\theta_2^\alpha} \right)^{\frac{1}{1-\alpha}}$$

with domain $(\underline{d_1}, \infty)$, where $\underline{d_1} = \left(\frac{\alpha(\theta_1+\theta_2)+p(1-\alpha)}{\theta_1^\alpha}\right)^{\frac{1}{1-\alpha}}$.

Proof. The survival function of X_1 is given by $\bar{F}_1(x_1) = \left(\frac{\theta_1}{x_1}\right)^\alpha$ for $\theta_1 < x_1$. Therefore,

$$\int_0^{d_1} \bar{F}_1(x)dx = \theta_1 + \int_{\theta_1}^{d_1} \theta_1^\alpha x^{-\alpha} dx = \frac{\theta_1 d_1^{1-\alpha} - \alpha\theta_1}{1-\alpha}$$

Analogous, the survival function for X_2 is given through $\bar{F}_2(x_2) = \left(\frac{\theta_2}{x_2}\right)^\alpha$ for $\theta_2 < x_2$ and it follows:

$$\int_0^{d_2} \bar{F}_2(x)dx = \frac{\theta_2 d_2^{1-\alpha} - \alpha\theta_2}{1-\alpha},$$

Thus, we get the following equation

$$\int_0^{d_1} \bar{F}_1(x)dx + \int_0^{d_2} \bar{F}_2(x)dx = p$$

$$\frac{\theta_1 d_1^{1-\alpha} - \alpha\theta_1}{1-\alpha} + \frac{\theta_2 d_2^{1-\alpha} - \alpha\theta_2}{1-\alpha} = p$$

$$\theta_1^\alpha d_1^{1-\alpha} - \alpha\theta_1 + \theta_2^\alpha d_2^{1-\alpha} - \alpha\theta_2 = p(1-\alpha)$$

Solving this equation for d_2 results in

$$d_2 = \left(\frac{\alpha(\theta_1+\theta_2) - \theta_1^\alpha d_1^{1-\alpha} + p(1-\alpha)}{\theta_2^\alpha}\right)^{\frac{1}{1-\alpha}}$$

Therefore, the function $L(d_1)$ is defined through $L(d_1) := d_2$ with domain $(\underline{d_1}, \infty)$ and is continuous, differentiable and strictly decreasing. The lower limit of the domain is given through

$$\begin{aligned}
\underline{d_1} &= \lim_{d_2\to\infty} L^{-1}(d_2) \\
&= \lim_{d_2\to\infty} \left(\frac{\alpha(\theta_1+\theta_2) - \theta_2^\alpha d_2^{1-\alpha} + p(1-\alpha)}{\theta_1^\alpha}\right)^{\frac{1}{1-\alpha}} \\
&= \left(\frac{\alpha(\theta_1+\theta_2) + p(1-\alpha)}{\theta_1^\alpha}\right)^{\frac{1}{1-\alpha}}.
\end{aligned}$$

□

Example 2.2.3. *For the parameter choices* $\theta_1 = 10, \theta_2 = 20, \alpha = 3$ *and* $p = 38$*, the function* $L(d_1)$ *has following representation:*

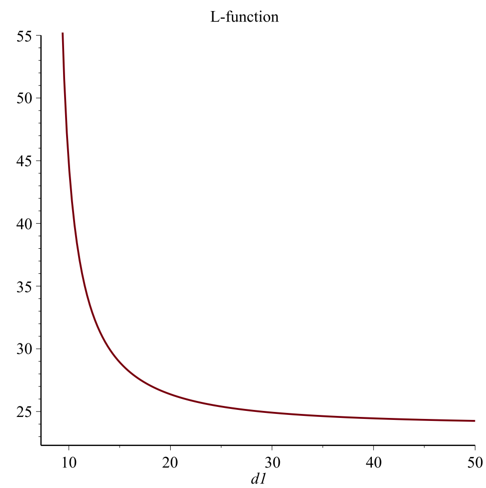

Figure 2.1: $L(d_1)$ for $\theta_1 = 10, \theta_2 = 20, \alpha = 3$ and $p = 38$

The function $L(d_1)$ *is continous, differentiable and strictly decreasing in* d_1*.*

Chapter 3

Optimal retention levels

In this chapter we want to derive the optimal retention levels in the excess-of-loss treaty for three different risk functions. Afterwards we compare the optimal retentions for different choices of the parameters of the bivariate Pareto distribution. Recall, that we want to derive explicit expressions for the optimal retention vector (d_1^*, d_2^*) in the bivariate Pareto case such that

$$\mathbb{E}[u(X_1 \wedge d_1^* + X_2 \wedge d_2^*)] = \inf_{(d_1,d_2)\in L} \mathbb{E}[u(X_1 \wedge d_1 + X_2 \wedge d_2)]$$

where for $p > 0$

$$L = \left\{ (d_1, d_2) \in [0, \infty)^2 \middle| \int_0^{d_1} \bar{F}_1(x)dx + \int_0^{d_2} \bar{F}_2(x)dx = p \right\}.$$

for different risk functions $u(x)$.

Consider the objective function

$$M(d_1, d_2) = \mathbb{E}[u(X_1 \wedge d_1 + X_2 \wedge d_2)].$$

Note that $d_2 = L(d_1)$, but we write in terms of d_2 to keep a symmetric structure. With Lemma 4.2 in Cai and Wei (2011), it follows that for u continuous and monotonic, such that $\mathbb{E}(|u(X_1 + X_2)|) < \infty$, $M(d_1, d_2)$ is continous in $d_1 \in (\underline{d_1}, \infty)$. Furthermore we need following Lemma in Cai and Wei (2011)

Lemma 3.0.4. *Assume $u(x) \in C^1(\mathbb{R})$, i.e. $u'(x)$ is continuous on $\mathbb{R}$. Then $\frac{\partial^+}{\partial d_1} M(d_1, d_2)$ is right continuous in $d_1 \in (\underline{d_1}, \infty)$ and*

$$\frac{\partial^+}{\partial d_1} M(d_1, d_2) = \bar{F}_1(d_1) \left(\mathbb{E}[u'(X_1 \wedge d_1 + X_2 \wedge d_2)|X_1 > d_1] \right.$$
$$\left. - \mathbb{E}[u'(X_1 \wedge d_1 + X_2 \wedge d_2)|X_2 > d_2]\right)$$

The proof of this Lemma can be found in Lemma 4.3 in Cai and Wei (2011).

So, for any risk function $u(x)$ with $\mathbb{E}(|u(X_1 + X_2)|) < \infty$ and $u(x) \in C^1(\mathbb{R})$, we define

$$C_u(d_1) = \mathbb{E}[u'(X_1 \wedge d_1 + X_2 \wedge d_2)|X_1 > d_1] - \mathbb{E}[u'(X_1 \wedge d_1 + X_2 \wedge d_2)|X_2 > d_2]$$
$$= \mathbb{E}[u'(d_1 + (X_2 \wedge d_2)|X_1 > d_1] - \mathbb{E}[u'(X_1 \wedge d_1) + d_2|X_2 > d_2].$$

We get the optimal retentions (d_1^*, d_2^*) if we solve $C_u(d_1) = 0$. In the following, we derive for several risk functions $u(x)$ the function $C_u(d_1)$. Notice, that the condition $C_u(d_1) = 0$ is only a necessary condition. We will check for every risk function that $M(d_1, d_2)$ indeed attains its minimum at (d_1^*, d_2^*).

3.1 Minimizing variance

Cai and Wei (2011) derived the following general explicit expression for the retentions in the optimal excess-of-loss-treaty in the bivariate case with risk function $u(x) = x^2$. Here we minimize the variance, since the expectation is fixed.

Theorem 3.1.1. *Assume (X_1, X_2) is PDS and $\mathbb{E}[(X_1 + X_2)^2] < \infty$.*
For $d_1 \in (d_1, \infty)$, define

$$C_1(d_1) = \mathbb{E}[(X_2 - L(d_1)) \wedge 0|X_1 > d_1] - \mathbb{E}[(X_1 - d_1) \wedge 0|X_2 > L(d_1)]. \qquad (3.1)$$

Denote $r_1 = \sup\{d_1|C_1(d_1) < 0\}$ and $r_2 = \inf\{d_1|C_1(d_1) > 0\}$.
Then $\underline{d_1} < r_1 \leq r_2 < \infty$ and for any $d_1^* \in [r_1, r_2]$, the retention vector $(d_1^*, L(d_1^*))$ is a solution to 1.1.

The proof of this Theorem can be found in Theorem 4.4 of Cai and Wei (2011). To apply Theorem 3.1.1, we have to check the assumptions of this Theorem in the case that the random variable vector (X_1, X_2) is bivariate Pareto distributed. Afterwards, we derive an explicit expression for the function $C_1(d_1)$. Recall, that with Lemma 2.1.3 follows that (X_1, X_2) is PDS. An important assumption of Theorem 3.1.1 is that $\mathbb{E}[(X_1 + X_2)^2]$ exists. However, this is just true for $\alpha > 2$.

Lemma 3.1.2. *If (X_1, X_2) is bivariate Pareto distributed,with parameters θ_1, θ_2 and $\alpha > 0$, where $x_1 > \theta_1 > 0$ and $x_2 > \theta_1 > 0$, then $\mathbb{E}((X_1 + X_2)^2) < \infty$ for $\alpha > 2$.*

Proof. With the Minkowski Inequality follows $\mathbb{E}((X_1 + X_2)^2) \leq \mathbb{E}(X_1^2) + \mathbb{E}(X_2^2)$. Since $X_i \sim ParetoI(\alpha, \theta_i)$ for $i = 1, 2$ it holds that

$$\mathbb{E}(X_i^2) = \frac{\alpha\theta_i^2}{\alpha - 2}, \text{ for } i = 1, 2$$

Therefore, it follows $\mathbb{E}((X_1 + X_2)^2) < \infty$ if $\alpha > 2$. □

Thus, throughout this section we will assume $\alpha > 2$. Based on Lemma 2.1.2 the conditional distribution of $X_2 - d_2$ given $X_1 > d_1$ is $ParetoII(\alpha, \sigma_2, \mu_2)$, where $\mu_2 = \theta_2 - d_2$ and $\sigma_2 = \frac{d_1\theta_2}{\theta_1}$. Moreover, the conditional distribution of $X_1 - d_1$ given $X_2 > d_2$ has $ParetoII(\alpha, \sigma_1, \mu_1)$ distribution, with $\mu_1 = \theta_1 - d_1$ and $\sigma_1 = \frac{d_2\theta_1}{\theta_2}$. Through applying the linear transformation between *Lomax* and *ParetoII* distribution and we can transform $X_2 - d_2$ given $X_1 > d_1$, respectively $X_1 - d_1$ given $X_2 > d_2$ in random variables with *Lomax* distribution. To get an explicit representation of $C_1(d_1)$ we can apply

$$\mathbb{E}(X \wedge x) = \frac{\sigma}{\alpha - 1}\left(1 - \left(\frac{\sigma}{x + \sigma}\right)^{\alpha - 1}\right), \tag{3.2}$$

with $x \geq 0, \alpha \neq 1$ for $X \sim Lomax(\alpha, \mu)$.

Consider the first expectation $\mathbb{E}[(X_2 - d_2) \wedge 0 | X_1 > d_1]$ of $C_1(d_1)$.
For $d_2 > \theta_2$, $\sigma_2 = \frac{d_1\theta_2}{\theta_1}$ and $\mu_2 = \theta_2 - d_2$ we get through applying the linear

transformation and formula (3.2):

$$
\begin{aligned}
&\mathbb{E}[(X_2-d_2)\wedge 0|X_1>d_1]\\
&=\mathbb{E}[((X_2-d_2)-(\theta_2-d_2))\wedge(d_2-\theta_2))|X_1>d_1]+\mathbb{E}[\theta_2-d_2|X_1>d_1]\\
&=\frac{\sigma_2}{\alpha-1}\left(1-\left(\frac{\sigma_2}{(d_2-\theta_2)+\sigma_2}\right)^{\alpha-1}\right)+\theta_2-d_2\\
&=\frac{d_1\theta_2}{\theta_1(\alpha-1)}-\frac{(d_1\theta_2)^\alpha(\theta_1 d_2+\theta_2 d_1-\theta_1\theta_2)^{1-\alpha}}{\theta_1(\alpha-1)}+\theta_2-d_2.
\end{aligned}
$$

For the second expectation of $C_1(d_1)$ we get through analogous calculations with $d_1>\theta_1$, $\sigma_1=\frac{d_2\theta_1}{\theta_2}$ and $\mu_1=\theta_1-d_1$

$$
\mathbb{E}[(X_1-d_1))\wedge 0|X_2>d_2]=\frac{d_2\theta_1}{\theta_2(\alpha-1)}-\frac{(d_2\theta_1)^\alpha(\theta_2 d_1+\theta_1 L(d_1)-\theta_1\theta_2)^{1-\alpha}}{\theta_2(\alpha-1)}+\theta_1-d_1.
$$

Altogether, the representation of $C_1(d_1)$ is

$$
\begin{aligned}
C_1(d_1)&=\mathbb{E}[(X_2-d_2)\wedge 0|X_1>d_1]-\mathbb{E}[(X_1-d_1)\wedge 0|X_2>d_2]\\
&=\frac{d_1\theta_2}{\theta_1(\alpha-1)}-\frac{(d_1\theta_2)^\alpha(\theta_1 d_2+\theta_2 d_1-\theta_1\theta_2)^{1-\alpha}}{\theta_1(\alpha-1)}+\theta_2-d_2\\
&-\frac{d_2\theta_1}{\theta_2(\alpha-1)}+\frac{(d_2\theta_1)^\alpha(\theta_2 d_1+\theta_1 d_2-\theta_1\theta_2)^{1-\alpha}}{\theta_2(\alpha-1)}-\theta_1+d_1
\end{aligned}
$$

with

$$
d_2=\left(\frac{\alpha(\theta_1+\theta_2)-\theta_1^\alpha d_1^{1-\alpha}+p(1-\alpha)}{\theta_2^\alpha}\right)^{\frac{1}{1-\alpha}}
$$

In following example we see the plot of the $C_1(d_1)$ for a specific parameter choice. The optimal retention is at $d_1 = d_1^*$, where $C_1(d_1) = 0$.

Example 3.1.3. *For the parameter choices* $\theta_1 = 10, \theta_2 = 20, \alpha = 3.5$ *and* $p = 38$, *the function* $C_1(d_1)$ *has following representation:*

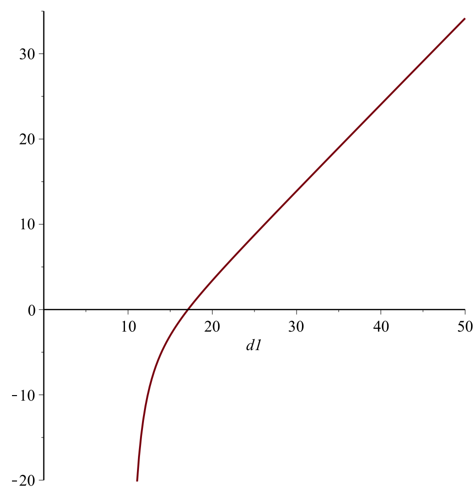

Figure 3.1: $C_1(d_1)$ for $\theta_1 = 10, \theta_2 = 20, \alpha = 3$ and $p = 38$

$M(d_1, d_2)$ attains indeed its minimum at $d_1^* = d_1$ since for any $d_1 < d_1^* : C_1(d_1) < 0$ and $d_1 > d_1^* : C_1(d_1) > 0$.

3.2 Minimizing moments of higher order

In this section we are interested in solving the optimization problem (1.1) under the aspect of minimizing moments of higher order. We consider the risk function $u(x) = x^3$ and derive a explicit expression for $C_u(d_1)$. Recall, that $C_u(d_1)$ is given by

$$C_u(d_1) = \mathbb{E}[u'(d_1 + (X_2 \wedge d_2)|X_1 > d_1] - \mathbb{E}[u'(X_1 \wedge d_1) + d_2|X_2 > d_2].$$

The risk function $u(x) = x^3$ is in $C^1(\mathbb{R})$ and convex. Moreover, for $\alpha > 3$, it follows directly with the Minkowski Inequality that $\mathbb{E}((X_1 + X_2)^3) < \infty$, since $\mathbb{E}(X_i^3) < \infty$ for $\alpha > 3$ and $i = 1, 2$.

Throughout this section we will assume $\alpha > 3$. We define,

$$\begin{aligned} C_2(d_1) &= \mathbb{E}[u'(d_1 + (X_2 \wedge d_2))|X_1 > d_1] - \mathbb{E}[u'((X_1 \wedge d_1) + d_2)|X_2 > d_2] \\ &= \mathbb{E}[(d_1 + (X_2 \wedge d_2))^2|X_1 > d_1] - \mathbb{E}[((X_1 \wedge d_1) + d_2)^2|X_2 > d_2] \\ &= \mathbb{E}[((X_2 + d_1) \wedge (d_1 + d_2))^2|X_1 > d_1] - \mathbb{E}[((X_1 + d_2) \wedge (d_1 + d_2))^2|X_2 > d_2]. \end{aligned}$$

In contrast to the previous section, we won't use the linear transformation between $ParetoII$ and $Lomax$ distribution here. In the following, we derive a general formula for $\mathbb{E}((X \wedge c)^2)$ for $X \sim ParetoII(\alpha, \mu, \sigma)$ and use this formula to get an explicit expression of $C_2(d_1)$.

Theorem 3.2.1. *If* $X \sim ParetoII(\alpha, \sigma, \mu)$, *with* $x > \mu, \mu \in \mathbb{R}, \sigma > 0$ *and* $\alpha > 3$ *then*

$$\mathbb{E}((X \wedge c)^2) = \alpha\sigma^\alpha \Bigg(\frac{(c + \sigma - \mu)^{-\alpha}(2\mu\alpha c + 4\mu\sigma - 2\alpha\sigma c - (\alpha - 1)\alpha c^2 - 2\mu^2 - 2\sigma^2)}{(\alpha - 2)(\alpha - 1)\alpha} - \frac{\sigma^{-\alpha}(3\mu^2\alpha + 2\mu^2 - 2\alpha\sigma\mu - \alpha^2\mu^2 - 2\sigma^2)}{(\alpha - 2)(\alpha - 1)\alpha} + \frac{c^2(c + \sigma - \mu)^{-\alpha}}{\alpha} \Bigg)$$

Proof.

$$\begin{aligned}\mathbb{E}((X \wedge c)^2) &= \int_{-\infty}^{\infty} (X \wedge c)^2 f(x) dx \\ &= \int_{\mu}^{c} x^2 \frac{\alpha\sigma^\alpha}{(x+\sigma-\mu)^{\alpha+1}} dx + \int_{c}^{\infty} c^2 \frac{\alpha\sigma^\alpha}{(x+\sigma-\mu)^{\alpha+1}} dx\end{aligned}$$

Consider the first integral. By integration by parts twice we see

$$\begin{aligned}&\int_{\mu}^{c} x^2 \frac{\alpha\sigma^\alpha}{(x+\sigma-\mu)^{\alpha+1}} \\ &= \alpha\sigma^\alpha \left(\frac{(x+\sigma-\mu)^{-\alpha}(2\mu\alpha x + 4\mu\sigma - 2\alpha\sigma x - (\alpha-1)\alpha x^2 - 2\mu^2 - 2\sigma^2)}{\alpha(\alpha-1)(\alpha-2)} \Bigg|_{x=\mu}^{c} \right).\end{aligned}$$

Thus, we get

$$\begin{aligned}\mathbb{E}((X \wedge c)^2) = \alpha\sigma^\alpha &\left(\frac{(x+\sigma-\mu)^{-\alpha}(2\mu\alpha x + 4\mu\sigma - 2\alpha\sigma x - (\alpha-1)\alpha x^2 - 2\mu^2 - 2\sigma^2)}{\alpha(\alpha-1)(\alpha-2)} \Bigg|_{x=\mu}^{c} \right. \\ &\left. + \frac{c^2(x+\sigma-\mu)^{-\alpha}}{-\alpha} \Bigg|_{x=c}^{\infty} \right) \\ &= \alpha\sigma^\alpha \left(\frac{(c+\sigma-\mu)^{-\alpha}(2\mu\alpha c + 4\mu\sigma - 2\alpha\sigma c - (\alpha-1)\alpha c^2 - 2\mu^2 - 2\sigma^2)}{(\alpha-2)(\alpha-1)\alpha} \right. \\ &\left. - \frac{\sigma^{-\alpha}(3\mu^2\alpha + 2\mu^2 - 2\alpha\sigma\mu - \alpha^2\mu^2 - 2\sigma^2)}{(\alpha-2)(\alpha-1)\alpha} + \frac{c^2(c+\sigma-\mu)^{-\alpha}}{\alpha} \right).\end{aligned}$$

□

Recall, with Lemma 2.1.2 follows, that the conditional distribution of $X_2 + d_1$ given $X_1 > d_1$ is $ParetoII(\alpha, \sigma_2, \mu_2)$, where $\mu_2 = \theta_2 + d_1$ and $\sigma_2 = \frac{d_1\theta_2}{\theta_1}$ and the conditional distribution $X_1 + d_2$ given $X_2 > d_2$ has $ParetoII(\alpha, \sigma_1, \mu_1)$ distribution, with $\mu_1 = \theta_1 + d_2$, $\sigma_1 = \frac{d_2\theta_1}{\theta_2}$ and $\alpha > 3$. By applying Theorem 3.2.1 we get for $C_2(d_1)$following representation:

$$
\begin{aligned}
C_2(d_1) &= \mathbb{E}[((X_2 + d_1) \wedge (d_1 + d_2))^2 | X_1 > d_1] - \mathbb{E}[((X_1 + d_2) \wedge (d_1 + d_2))^2 | X_2 > d_2] \\
&= \alpha\sigma_2^{\alpha} \left(\frac{(c + \sigma_2 - \mu_2)^{-\alpha}(2\mu_2\alpha c + 4\mu_2\sigma_2 - 2\alpha\sigma_2 c - (\alpha-1)\alpha c^2 - 2\mu_2^2 - 2\sigma_2^2)}{(\alpha-2)(\alpha-1)\alpha} \right. \\
&- \left. \frac{\sigma_2^{-\alpha}(3\mu_2^2\alpha + 2\mu_2^2 - 2\alpha\sigma_2\mu_2 - \alpha^2\mu_2^2 - 2\sigma_2^2)}{(\alpha-2)(\alpha-1)\alpha} + \frac{c^2(c + \sigma_2 - \mu_2)^{-\alpha}}{\alpha} \right) \\
&- \alpha\sigma_1^{\alpha} \left(\frac{(c + \sigma_1 - \mu_1)^{-\alpha}(2\mu_1\alpha c + 4\mu_1\sigma_1 - 2\alpha\sigma_1 c - (\alpha-1)\alpha c^2 - 2\mu_1^2 - 2\sigma_1^2)}{(\alpha-2)(\alpha-1)\alpha} \right. \\
&- \left. \frac{\sigma_1^{-\alpha}(3\mu_1^2\alpha + 2\mu_1^2 - 2\alpha\sigma_1\mu_1 - \alpha^2\mu_1^2 - 2\sigma_1^2)}{(\alpha-2)(\alpha-1)\alpha} + \frac{c^2(c + \sigma_1 - \mu_1)^{-\alpha}}{\alpha} \right).
\end{aligned}
$$

To get the optimal retention vector (d_1^*, d_2^*), we solve $C_2(d_1) = 0$.

Example 3.2.2. *For the parameter choices $\theta_1 = 10, \theta_2 = 20, \alpha = 3.5$ and $p = 38$, the function $C_2(d_1)$ has following plot:*

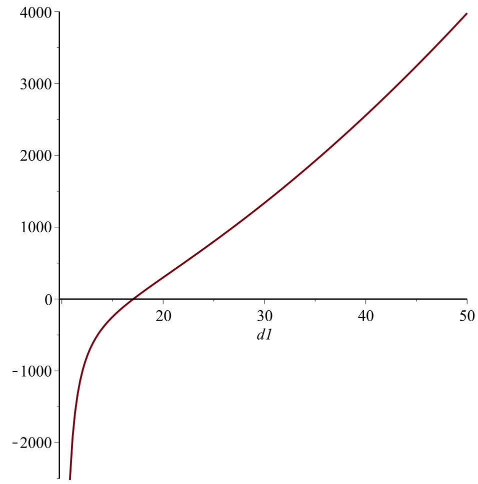

Figure 3.2: $C_2(d_1)$ for $\theta_1 = 10, \theta_2 = 20, \alpha = 3$ and $p = 38$

Since for any $d_1 < d_1^* : C_1(d_1) < 0$ and $d_1 > d_1^* : C_1(d_1) > 0$, the objective function attains its minimum at $d_1^* = d_1$.

3.3 Minimizing moments of fractional order

Another interesting aspect is solving the optimization problem under the criteria of minimizing moments of fractional order. We consider the risk function $u(x) = x^{1.5}$ for $\alpha = 2.5$ and $\alpha = 3.5$. The general form of the C-function is given by

$$C_u(d_1) = \mathbb{E}[u'(d_1 + (X_2 \wedge d_2)|X_1 > d_1] - \mathbb{E}[u'(X_1 \wedge d_1) + d_2|X_2 > d_2].$$

The risk function $u(x) = x^{1.5}$ is convex and in $C^1(\mathbb{R})$. We define

$$\begin{aligned} C_3(d_1) &= \mathbb{E}[u'(d_1 + (X_2 \wedge d_2))|X_1 > d_1] - \mathbb{E}[u'((X_1 \wedge d_1) + d_2)|X_2 > d_2] \\ &= \mathbb{E}[(d_1 + (X_2 \wedge d_2))^{0.5}|X_1 > d_1] - \mathbb{E}[((X_1 \wedge d_1) + d_2)^{0.5}|X_2 > d_2] \end{aligned}$$

$$= \mathbb{E}[((X_2+d_1)\wedge(d_1+d_2))^{0.5}|X_1>d_1] - \mathbb{E}[((X_1+d_2)\wedge(d_1+d_2))^{0.5}|X_2>d_2].$$

Since (X_1, X_2) is bivariate Pareto distributed with parameters $\theta_1 > 0, \theta_2 > 0$ and α, where $x_1 > \theta_1 > 0$ and $x_2 > \theta_2 > 0$ it holds with Minkowski Inequality that $\mathbb{E}(|(X_1 + X_2)^{1.5}|) < \infty$ for $\alpha > 1.5$.

First we are interested in the case with $\alpha = 2.5$. Similar to Section 3.2 we derive a general formula for $\mathbb{E}((X \wedge c)^{0.5})$ if $X \sim ParetoII(\alpha, \mu, \sigma)$ and $\alpha = 2.5$ to calculate $C_3(d_1)$.

Theorem 3.3.1. *If $X \sim ParetoII(\alpha, \mu, \sigma)$, with $x > \mu, \mu \in \mathbb{R}, \sigma > 0$ and $\alpha = 2.5$, then*

$$\mathbb{E}((X\wedge c)^{0.5}) = 2.5\sigma^{2.5}\left[\frac{\frac{4}{15}c^{1.5}(c+\frac{5}{2}(\sigma-\mu))}{(\sigma-\mu)^2(\sigma-\mu+c)^{2.5}} - \frac{\frac{4}{15}\mu^{1.5}(\sigma-\frac{3}{2}\mu)}{(\sigma-\mu)^2\sigma^{2.5}} + \frac{\frac{2}{5}c^{0.5}}{(c+\sigma-\mu)^{2.5}}\right]$$

Proof.

$$\mathbb{E}((X\wedge c)^{0.5}) = \int_{-\infty}^{\infty}(X\wedge c)^{0.5}f(x)dx$$

$$= \int_{\mu}^{c} x^{0.5}\frac{2.5\sigma^{2.5}}{(x+\sigma-\mu)^{3.5}}dx + \int_{c}^{\infty} c^{0.5}\frac{2.5\sigma^{2.5}}{(x+\sigma-\mu)^{3.5}}dx$$

$$= 2.5\sigma^{2.5}\left(\left[\frac{\frac{4}{15}x^{1.5}(x+\frac{5}{2}(\sigma-\mu))}{(\sigma-\mu)^2(\sigma-\mu+x)^{2.5}}\Bigg|_{x=\mu}^{c}\right] - \left[\frac{\frac{2}{5}c^{0.5}}{(x+\sigma-\mu)^{2.5}}\Bigg|_{x=c}^{\infty}\right]\right)$$

$$= 2.5\sigma^{2.5}\left[\frac{\frac{4}{15}c^{1.5}(c+\frac{5}{2}(\sigma-\mu))}{(\sigma-\mu)^2(\sigma-\mu+c)^{2.5}} - \frac{\frac{4}{15}\mu^{1.5}(\sigma-\frac{3}{2}\mu)}{(\sigma-\mu)^2\sigma^{2.5}} + \frac{\frac{2}{5}c^{0.5}}{(c+\sigma-\mu)^{2.5}}\right]$$

□

Recall that the conditional distribution of X_1+d_2 given $X_2 > d_2$ is $ParetoII(\alpha, \sigma_1, \mu_1)$ with $\mu_1 = d_2+\theta_1, \sigma_1 = \frac{d_2\theta_1}{\theta_2}$ and the conditional distribution of X_2+d_1 given $X_1 > d_1$

is $ParetoII(\alpha, \sigma_2, \mu_2)$ with $\mu_2 = d_1 + \theta_2, \sigma_2 = \frac{d_1\theta_2}{\theta_1}$. Thus, through applying Theorem 3.3.1 we get for $\alpha = 2.5$ following formula for $C_3(d_1)$:

$$C_3(d_1) = \mathbb{E}[((X_2 + d_1) \wedge (d_1 + d_2))^{0.5} | X_1 > d_1] - \mathbb{E}[((X_1 + d_2) \wedge (d_1 + d_2))^{0.5} | X_2 > d_2]$$

$$= 2.5\sigma_2^{2.5} \left[\frac{\frac{4}{15}(d_1 + d_2)^{1.5}(d_1 + d_2 + \frac{5}{2}(\sigma_2 - \mu_2))}{(\sigma_2 - \mu_2)^2(\sigma_2 - \mu_2 + d_1 + d_2)^{2.5}} - \frac{\frac{4}{15}\mu_2^{1.5}(\sigma_2 - \frac{3}{2}\mu_2)}{(\sigma_2 - \mu_2)^2\sigma^{2.5}} \right.$$

$$\left. + \frac{\frac{2}{5}(d_1 + d_2)^{0.5}}{(d_1 + d_2 + \sigma_2 - \mu_2)^{2.5}} \right] - 2.5\sigma_1^{2.5} \left[\frac{\frac{4}{15}(d_1 + d_2)^{1.5}(d_1 + d_2 + \frac{5}{2}(\sigma_1 - \mu_1))}{(\sigma_1 - \mu_1)^2(\sigma_1 - \mu_1 + d_1 + d_2)^{2.5}} \right.$$

$$\left. - \frac{\frac{4}{15}\mu_1^{1.5}(\sigma_1 - \frac{3}{2}\mu_1)}{(\sigma_1 - \mu_1)^2\sigma^{2.5}} + \frac{\frac{2}{5}(d_1 + d_2)^{0.5}}{(d_1 + d_2 + \sigma_1 - \mu_1)^{2.5}} \right].$$

Now we consider the case $\alpha = 3.5$. A general formula for $\mathbb{E}((X \wedge c)^{0.5})$ if $X \sim ParetoII(\alpha, \mu, \sigma)$, with $\alpha = 3.5$ is given by following Theorem.

Theorem 3.3.2. *If $X \sim ParetoII(\alpha, \mu, \sigma)$, with $x > \mu, \mu \in \mathbb{R}, \sigma > 0$ and $\alpha = 3.5$, then*

$$\mathbb{E}((X \wedge c)^{0.5}) = 3.5\sigma^{3.5} \left[\frac{\frac{2}{3}c^{1.5}((\sigma - \mu)^2 + 0.8(\sigma - \mu)c + \frac{8}{35}c^2)}{(\sigma - \mu)^3(\sigma - \mu + c)^{3.5}} \right.$$

$$\left. - \frac{\frac{2}{3}\mu^{1.5}((\sigma - \mu)^2 + 0.8\mu(\sigma - \mu) + \frac{8}{35}\mu^2))}{(\sigma - \mu)^3\sigma^{3.5}} + \frac{\frac{2}{7}c^{0.5}}{(c + \sigma - \mu)^{3.5}} \right]$$

Proof.

$$\mathbb{E}((X \wedge c)^{0.5}) = \int_{-\infty}^{\infty} (X \wedge c)^{0.5} f(x) dx$$

$$= \int_{\mu}^{c} x^{0.5} \frac{3.5\sigma^{3.5}}{(x + \sigma - \mu)^{4.5}} dx + \int_{c}^{\infty} c^{0.5} \frac{3.5\sigma^{3.5}}{(x + \sigma - \mu)^{4.5}} dx$$

$$= 3.5 \cdot \sigma^{3.5} \left(\left[\frac{\frac{2}{3}x^{1.5}((\sigma - \mu)^2 + 0.8(\sigma - \mu)x + \frac{8}{35}x^2)}{(\sigma - \mu)^3(\sigma - \mu + x)^{3.5}} \Bigg|_{x=\mu}^{c} \right] \right.$$

$$- \left[\frac{\frac{2}{7}c^{0.5}}{(x+\sigma-\mu)^{3.5}} \Bigg|_{x=c}^{\infty} \right] \Bigg)$$

$$= 3.5\sigma^{3.5} \left[\frac{\frac{2}{3}c^{1.5}((\sigma-\mu)^2 + 0.8(\sigma-\mu)c + \frac{8}{35}c^2)}{(\sigma-\mu)^3(\sigma-\mu+c)^{3.5}} \right.$$

$$\left. - \frac{\frac{2}{3}\mu^{1.5}((\sigma-\mu)^2 + 0.8\mu(\sigma-\mu) + \frac{8}{35}\mu^2))}{(\sigma-\mu)^3\sigma^{3.5}} + \frac{\frac{2}{7}c^{0.5}}{(c+\sigma-\mu)^{3.5}} \right]$$

□

Through applying Theorem 3.3.2 we get for $\alpha = 3.5$

$$C_3(d_1) = \mathbb{E}[((X_2+d_1) \wedge (d_1+d_2))^{0.5} | X_1 > d_1] - \mathbb{E}[((X_1+d_2) \wedge (d_1+d_2))^{0.5} | X_2 > d_2]$$

$$= 3.5\sigma_2^{3.5} \left[\frac{\frac{2}{3}(d_1+L(d_1))^{1.5}((\sigma_2-\mu_2)^2 + 0.8(\sigma_2-\mu_2)(d_1+L(d_1)) + \frac{8}{35}(d_1+L(d_1))^2)}{(\sigma_2-\mu_2)^3(\sigma_2-\mu_2+(d_1+L(d_1)))^{3.5}} \right.$$

$$\left. - \frac{\frac{2}{3}\mu_2^{1.5}((\sigma_2-\mu_2)^2 + 0.8\mu_2(\sigma_2-\mu_2) + \frac{8}{35}\mu_2^2))}{(\sigma_2-\mu_2)^3\sigma_2^{3.5}} + \frac{\frac{2}{7}(d_1+L(d_1))^{0.5}}{((d_1+L(d_1))+\sigma_2-\mu_2)^{3.5}} \right]$$

$$- 3.5\sigma_1^{3.5} \left[\frac{\frac{2}{3}(d_1+L(d_1)^{1.5}((\sigma_1-\mu_1)^2 + 0.8(\sigma_1-\mu_1)(d_1+L(d_1) + \frac{8}{35}(d_1+L(d_1)^2)}{(\sigma_1-\mu_1)^3(\sigma_1-\mu_1+(d_1+L(d_1))^{3.5}} \right.$$

$$\left. - \frac{\frac{2}{3}\mu_1^{1.5}((\sigma_1-\mu_1)^2 + 0.8\mu_1(\sigma_1-\mu_1) + \frac{8}{35}\mu_1^2))}{(\sigma_1-\mu_1)^3\sigma_1^{3.5}} + \frac{\frac{2}{7}(d_1+L(d_1)^{0.5}}{((d_1+L(d_1)+\sigma_1-\mu_1)^{3.5}} \right],$$

Example 3.3.3. *For the parameter choices* $\theta_1 = 10, \theta_2 = 20, \alpha = 2.5$ *and* $p = 38$, *the function* $C_3(d_1)$ *has following representation:*

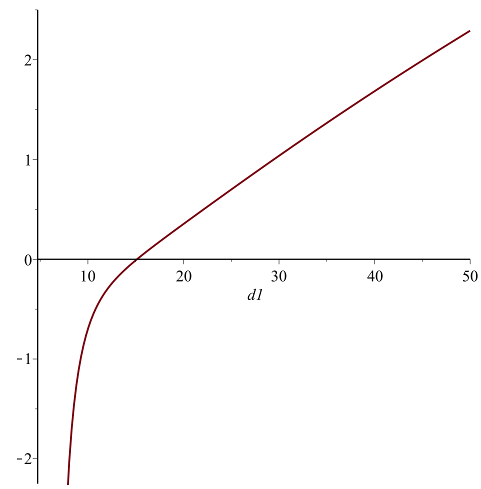

Figure 3.3: $C_3(d_1)$ for $\theta_1 = 10, \theta_2 = 20, \alpha = 2.5$ and $p = 38$

$M(d_1, d_2)$ attains its minimum at $d_1^* = d_1$ since for any $d_1 < d_1^* : C_3(d_1) < 0$ and $d_1 > d_1^* : C_3(d_1) > 0$ for $\alpha = 2.5$. Moreover, for any $d_1 < d_1^* : C_3(d_1) < 0$ and $d_1 > d_1^* : C_3(d_1) > 0$ for $\alpha = 3.5$.

3.4 Sensitivity Analysis

In this section we analyze the optimal retentions (d_1^*, d_2^*) for different choices of the parameters of the bivariate Pareto distribution. For this we use the previous derived formulas of $C_u(d_1)$ of the different risk functions and calculate the values $d_1 = d_1^*$ where $C_u(d_1) = 0$.

Let (X_1, X_2) be bivariate Pareto distributed with parameters $\theta_1 > 0, \theta_2 > 0$ and $\alpha > 0$. We fix $\theta_1 = 10$ and $\theta_2 = 20$. Thus, the minimal possible value for risk X_1 is

10 and for risk X_2 is 20. Since we are considering a *Pareto* distribution we expect for risk X_2 a larger expected claim if a claim occurs. Intuitively, the deductible for risk X_2 should be larger than for risk X_1. Furthermore, we consider two different fixed values of p. Recall that $p \in \left[\theta_1 + \theta_2, \frac{\alpha(\theta_1+\theta_2)}{\alpha-1}\right)$. In Table 3.1 we see the possible values for p for different choices of the parameter.

θ_1	θ_2	α	Interval of p
10	20	2.5	[30,50)
10	20	3	[30,45)
10	20	3.5	[30,42)
10	20	4	[30,40)
10	20	4.5	[30, 38.57)

Table 3.1: Interval of p

We will fix the value of p equal to 35 and 38.

First, we consider the case, where we minimize the variance with risk function $u(x) = x^2$. Recall, that $\alpha > 2$ in this case. We analyze the values for the optimal retention vector (d_1^*, d_2^*) if α increases in 0.5-steps from 2.5 on.

θ_1	θ_2	α	p	d_1	d_2
10	20	2.5	35	12.652	23.732
10	20	3	35	12.833	23.966
10	20	3.5	35	13.050	24.246
10	20	4	35	13.317	24.586
10	20	4.5	35	13.654	25.016

Table 3.2: Retentions (d_1, d_2) for $p = 35$, $u(x) = x^2$, shifting α

θ_1	θ_2	α	p	d_1	d_2
10	20	2.5	38	15.119	27.185
10	20	3	38	15.906	28.217
10	20	3.5	38	17.109	29.787
10	20	4	38	19.295	32.624
10	20	4.5	38	25.719	40.913

Table 3.3: Retentions (d_1, d_2) for $p = 38$, $u(x) = x^2$, shifting α

Consider the risk function $u(x) = x^3$. In Table 3.4 and 3.4 we see the values for the optimal retention vector (d_1^*, d_2^*) if α increases from 3.5 on in 0.5-steps till 4.5. Recall, that in this case $\alpha > 3$ and again we fix p equal to 35 and 38.

θ_1	θ_2	α	p	d_1	d_2
10	20	3.5	35	13.039	24.254
10	20	4	35	13.303	24.596
10	20	4.5	35	13.638	25.027

Table 3.4: Retentions (d_1, d_2) for $p = 35$, $u(x) = x^3$, shifting α

θ_1	θ_2	α	p	d_1	d_2
10	20	3.5	38	17.055	29.821
10	20	4	38	19.208	32.669
10	20	4.5	38	25.511	40.989

Table 3.5: Retentions (d_1, d_2) for $p = 38$, $u(x) = x^3$, shifting α

For the risk function $u(x) = x^{1.5}$ it holds that $\alpha > 2$. Again, we fix $\theta_1 = 10$, $\theta_2 = 20$ and p equal to 35 and 38. We derived the formulas of $C_2(d_1)$ for $\alpha = 2.5$ and $\alpha = 3.5$. In table 3.9 we see the explicit values for the optimal retentions (d_1, d_2).

θ_1	θ_2	α	p	d_1	d_2
10	20	2.5	35	12.656	23.728
10	20	3.5	35	13.056	24.241
10	20	2.5	38	15.134	27.174
10	20	3.5	38	17.136	29.770

Table 3.6: Retentions (d_1, d_2) for $p = 35$ and $p = 38$, $u(x) = x^{1.5}$, shifting α

We see that for increasing α, the deductibles (d_1, d_2) are increasing for all three risk functions. The shape parameter α indicates how fast the tail of the distribution goes to 0. If α is getting larger the survival function is steeper and decays faster to 0. The probability that a loss occurs is higher for larger values of α. It is more probable that a loss exceed a specific threshold for α larger. Therefore, the optimal retentions for the insurer are larger if α increases for fixed parameters.

For $p = 38$ the retentions have, in general, higher values. This makes sense because p is equal to $\mathbb{E}(S_2^I)$, the expected total retained loss of the insurer. For a larger p the insurer retains a larger value and the reinsurer covers less of the loss, therefore the retentions is higher for larger p. Furthermore, we can observe that values for the retentions are minimally smaller for the risk function $u(x) = x^3$ than for the risk function $u(x) = x^2$ and $u(x) = x^{1.5}$. For the risk function $u(x) = x^{1.5}$ the optimal deductibles have the largest values.

We also consider what happens if we change the parameters θ_1 and θ_2 for the different risk functions. We fix $p = 38$ for all risk functions. The parameter α is fixed equal to 2.5 for the risk function $u(x) = x^2$. For the risk function $u(x) = x^3$ we fix $\alpha = 3.5$.

θ_1	θ_2	α	p	d_1	d_2
10	20	2.5	38	15.119	27.185
15	20	2.5	38	16.562	21.802
15	15	2.5	38	21.086	21.086

Table 3.7: Retentions (d_1, d_2), $u(x) = x^2$, shifting θ_1, θ_2

θ_1	θ_2	α	p	d_1	d_2
10	20	3.5	38	17.057	29.821
15	20	3.5	38	16.648	21.899
15	15	3.5	38	23.277	23.277

Table 3.8: Retentions (d_1, d_2), $u(x) = x^3$, shifting θ_1, θ_2

θ_1	θ_2	α	p	d_1	d_2
10	20	2.5	38	15.134	27.174
15	20	2.5	38	16.563	21.802
15	15	2.5	38	21.086	21.086

Table 3.9: Retentions (d_1, d_2) for $\alpha = 2.5$, $u(x) = x^{1.5}$, shifting θ_1, θ_2

θ_1	θ_2	α	p	d_1	d_2
10	20	3.5	38	17.136	29.770
15	20	3.5	38	16.648	21.899
15	15	3.5	38	23.278	23.278

Table 3.10: Retentions (d_1, d_2) for $\alpha = 3.5$, $u(x) = x^{1.5}$

We see that for the same choice of the parameter θ_1 and θ_2, the retentions are equal in all risk functions u. For $\alpha = 2.5$ the optimal retentions for the risk functions $u(x) = x^2$ and $u(x) = x^{1.5}$ just differ minimally. The same observation can be made for the other two risk functions, where $\alpha = 3.5$. If the values of θ_1 and θ_2 are closer together, the optimal retentions are minimally higher than θ_1 and θ_2.

Chapter 4

Summary

Based on the study of Cai and Wei (2011) we developed the optimal retentions (d_1^*, d_2^*) for an insurance with two lines of business and where the risk variables are bivariate Pareto distributed and positive dependent through stochastic ordering. We minimized the objective function $M(d_1, d_2)$ under the aspects of minimizing the variance, minimizing moments of higher order and of fractional order. Especially the consideration of the risk function $u(x) = x^3$ and $u(x) = x^{1.5}$, where we minimized moments of higher order and of fractional order are new contributions to the existing literature. The shifting of the shape parameter α of the bivariate Pareto distribution has been analyzed under specific assumptions and for different risk functions. We observed that the optimal retentions just differ minimally for the different risk functions in this case. Furthermore, the changes in the values of the optimal retentions if we shift the parameter α has shown what we expected from a heavy-tailed distribution like the Pareto distribution.

An interesting topic for future researches is to consider this optimization problem in an business with three lines or more. It could be possible to find optimal retentions for adequate risk functions by using Lagrange multiplier. Another approach for this optimization problem would be to consider other distributions of regular variation. Since in an insurance one is often concerned about risks that exceed a specific threshold and also about positive dependence between risks, other possible distributions for this kind of research would, besides of the Pareto distribution, be the Burr distribution or the Log-Gamma distribution, for example.

Bibliography

[1] Jun Cai, Wei Wei, *Optimal reinsurance with positively dependent risks* , Insurance: Mathematics and Economics 50, 57-63,2011

[2] M. Denuit, C. Vermandele, *Optimal reinsurance and stop-loss order* , Insurance: Mathematics and Economics 22, 229-233, 1998

[3] K. V. Mardia, *Multivariate Paretor Distributions*, 1962

[4] A. Welshaus, *ASM Study Manual for Exam C/Exam 4 - Construction and Evaluation of Acturial Models*

[5] P. Hoeppe, T. Kuczinski, *Munich Re, GeoRisikoForschung, NatCatSERVICE*, August 2012

APPENDIX

Proof of Lemma 2.1.1

Proof.

$$
\begin{aligned}
\mathbb{P}(X_2 > s | X_1 > d_1) &= \mathbb{P}(X_2 > d_1 | X_1 > d_1) \\
&= \frac{1}{\mathbb{P}(X_1 > d_1)} \int_s^\infty \int_{d_1}^\infty \frac{\alpha(\alpha+1)(\theta_1\theta_2)^{\alpha+1}}{(\theta_2 x_1 + \theta_1 x_2 - \theta_1\theta_2)^{(\alpha+2)}} dx_1 dx_2 \\
&= \left(\frac{d_1}{\theta_1}\right)^\alpha \alpha(\alpha+1)(\theta_1\theta_2)^{\alpha+1} \int_s^\infty \frac{(\theta_2 d_1 + \theta_1 x_2 - \theta_1\theta_2)^{-(\alpha+1)}}{(\alpha+1)\theta_2} dx_2 \\
&= (d_1\theta_2)^\alpha \alpha\theta_1 \left(\frac{(\theta_2 d_1 + \theta_1 x_2 - \theta_1\theta_2)^{-\alpha}}{\alpha\theta_1} \Bigg|_{x_2=s}^\infty \right) \\
&= \left(\frac{\theta_2 d_1 + \theta_1 s - \theta_1\theta_2}{d_1\theta_2}\right)^{-\alpha} \\
&= \left(1 + \frac{s - \theta_2}{\frac{d_1\theta_2}{\theta_1}}\right)^{-\alpha}
\end{aligned}
$$

Thus, $X_2 | X_1 > d_1 \sim ParetoII(\alpha, \sigma_2, \mu_2)$, with $\mu_2 = \theta_2$ and $\sigma_2 = \frac{d_1\theta_2}{\theta_1}$.
Analogeous, it follows $X_1 | X_2 > d_2 \sim ParetoII(\alpha, \sigma_1, \mu_1)$ with
$\mu_1 = \theta_1$ and $\sigma_1 = \frac{d_2\theta_1}{\theta_2}$. □

Proof of Lemma 2.1.3

Proof. Thus, we first calculate the conditional probability $X_2 > x_2|X_1 = x_1$. Since $x_2 > \theta_2$ and $x_1 > \theta_1$,

$$\begin{aligned}
\mathbb{P}(X_2 > x_2|X_1 = x_1) &= \int_{x_2}^{\infty} f(s|x_1)ds \\
&= \int_{x_2}^{\infty} \frac{f(s, x_1)}{f(x_1)}ds \\
&= \int_{x_2}^{\infty} \frac{(\alpha+1)\alpha(\theta_1\theta_2)^{\alpha+1}(\theta_2 x_1 + \theta_1 s - \theta_1\theta_2)^{-(\alpha+2)}}{\alpha\theta_1^{\alpha} x_1^{-(\alpha+1)}}ds \\
&= \int_{x_2}^{\infty} (\alpha+1)\theta_1\theta_2^{\alpha+1}x_1^{\alpha+1}(\theta_2 x_1 + \theta_1 s - \theta_1\theta_2)^{-(\alpha+2)}ds.
\end{aligned}$$

With substitution $t = \theta_2 x_1 + \theta_1 s - \theta_1\theta_2$ it follows

$$\begin{aligned}
\mathbb{P}(X_2 > x_2|X_1 = x_1) &= \int_{\theta_2 x_1 + \theta_1 x_2 - \theta_1\theta_2}^{\infty} (\alpha+1)\theta_2^{\alpha+1}x_1^{\alpha+1}t^{-(\alpha+2)} \\
&= -\theta_2^{\alpha+1}x_1^{\alpha+1}t^{-(\alpha+1)}\Big|_{t=\theta_2 x_1 + \theta_1 x_2 - \theta_1\theta_2}^{\infty} \\
&= (\theta_2 x_1)^{\alpha+1}(\theta_2 x_1 + \theta_1 x_2 - \theta_1\theta_2)^{-(\alpha+1)}.
\end{aligned}$$

Now, we can prove the increasing property. It holds that $\mathbb{P}(X_2 > x_2|X_1 = x_1)$ is increasing in $x_1 \in S(x_1)$, if for $s, t \in S(x_1)$ with $s \leq t, x_2 \in \mathbb{R}$:
$\mathbb{P}(X_2 > x_2|X_1 = s) \leq \mathbb{P}(X_2 > x_2|X_1 = t)$.

$$\begin{aligned}
&\mathbb{P}(X_2 > x_2|X_1 = s) \leq \mathbb{P}(X_2 > x_2|X_1 = t) \\
\Leftrightarrow\quad &\theta_2^{\alpha+1}s^{\alpha+1}(\theta_2 s + \theta_1 x_2 - \theta_1\theta_2)^{-(\alpha+1)} \leq \theta_2^{\alpha+1}t^{\alpha+1}(\theta_2 t + \theta_1 x_2 - \theta_1\theta_2)^{-(\alpha+1)}
\end{aligned}$$

$$\Leftrightarrow \quad (\theta_2 st + \theta_1 x_2 s - \theta_1\theta_2 s)^{(\alpha+1)} \leq (\theta_2 st + \theta_1 x_2 t - \theta_1\theta_2 t)^{(\alpha+1)}$$

$$\Leftrightarrow \quad \theta_2 st + \theta_1 x_2 s - \theta_1\theta_2 s \leq \theta_2 ts + \theta_1 x_2 t - \theta_1\theta_2 t$$

$$\Leftrightarrow \quad s(\theta_1 x_2 - \theta_1\theta_2) \leq t(\theta_1 x_2 - \theta_1\theta_2)$$

$$\Leftrightarrow \quad s \leq t$$

With analogeous calculations, it follows that $X_1 \uparrow_{SI} X_2$. Therefore, (X_1, X_2) is PDS. $\square$