Capital Allocation

Kristof D'haeseleer

Acknowledgements

Two years ago, I already wrote a thesis for my studies in math. Since I didn't want to quit my life as a student, I started the study of "Master in Financial and Actuarial Engineering" in Leuven. Moreover I moved to Aarschot to begin a new part of my life with my girlfriend Sofie. She was my greatest support during this period. Furthermore I would like to thank my parents for giving me the opportunity to study for 6 years at the universities of Gent and Leuven.

This work has not been possible without the cooperation of many people of Secura-re. This company offered me a trainee ship to work on the subjects treated in this thesis. I am glad to say that I work now for this reinsurance company. Especially, I thank Jean-François Walhin and Stijn Desmedt. Stijn was always helpful to read my findings and to introduce new insights.

Finally I would like to thank my promotor, Prof. Dr. Jan Dhaene, for his ideas and interest in my work. I appreciate his clear and fast communication and the fact that I always could pass by.

Contents

Preface

The work I may present here, was made within the framework of a traineeship at Secura-re (a Belgian reinsurance company). The issues that are treated, were proposed by the R&D department. They are closely related to general problems and needs which a (re)insurance company faces. Therefore this paper can also be seen as a productive collaboration between the university and the world of insurance. Moreover, from the first part of this thesis, a scientific paper is submitted for publication in "Journal of Risk".
Both theoretical, practical as well as programming aspects were given a chance in this work. Therefore it was a pleasure to work with such varied subjects.

Because the title of this thesis, "Capital Allocation", is very general, I will precise a little more what you can expect when you are reading this piece of work. This paper consists of two main parts. In the first part, "EVA Optimization for Insurers Using a Multivariate Student-t Model", we study the class of multivariate elliptical distributions. With the achieved knowledge of that study, a model (based on the multivariate student-t distribution) is build to calculate the value creation of different underwriting policies of an insurance company.
Whereas in the first part a nice application was the result of the theoretical study of several papers, the opposite is reached in part II, "Stochastic Capital Model for Large MTPL Claims". Here we start from real data of motor insurance and build a stochastic model. Although this part is more practically orientated, several theoretical tools (statistical tests, the use of several distributions) are used. The main aim of building this model is to make simulations later on. From these simulations, we can determine the needs for the capital allocated to this business. In the next paragraphs we explain the way of working and the content of each part more in detail.

In the first chapter of "EVA Optimization for Insurers Using a Multivariate Student-t Model" some basics and definitions about risk measures and value creation are presented. The class of elliptical distributions is studied in chapter 2. At the end of this chapter we introduce the concept of tail dependency. In chapter 3 we build a model to calculate the Economic Value Added (EVA) for a given portfolio of insurance contracts. The model presented is a generalization of the model used in Walhin (2005). Here we use a multivariate elliptical distribution for the claims. We give numerical applications for the normal and student-t distribution. At the end we analyze the effect of stop loss reinsurance on the value creation of an insurance company.

The introduction of part II contains a short overview of the theoretical tools which are used to build the model. In the second chapter the different components of the model are described. We analyze the number of claims, the size of a claim and its evolution in time. In the last chapter, we put all these elements together and make simulations. In this way we obtain a distribution for the total amount of claims.

For me it was an enriching experience working on these topics. I hope you enjoy the reading!

PART I

EVA OPTIMIZATION FOR INSURERS USING A MULTIVARIATE STUDENT-T MODEL

General Introduction

In this part we analyze the underwriting policy of an insurance company. When the insurer sells an insurance contract he wants to know the probability of having a claim on that contract and to have an idea of the size of that claim. This will depend on the kind of contract that is underwritten: fire, health care, life, third party liability,... but also on the profile of the client (age, smoker, job,...). All risk, linked to the frequency and severity of a claim, is called underwriting risk. Of course this is not the only risk an insurer faces. Another important risk is the return on investment. First premiums are received while the claims are payed later on. The collected money has to be invested in several assets (bonds, stocks,...). When the financial markets are crashing, an insurance company can lose a lot of money. We call this market risk.
When the commercial department does not function well, this may lead to a decreasing premium income or excessive expenses. A severe failure of the computer system can paralyze the whole company. These risks are referred to as operational risks.
All risks (underwriting risk, market risk, operational risk,...) have an influence on the value creation. In this work we focus on the underwriting risk.

The goal is to determine a portfolio of policies for which the value creation is maximal, given a set of constraints. In order to be sure that the insurer can pay his claims, the regulator will impose him to hold a required capital. This capital cannot be funded with the premium income alone. An important part is provided by the shareholders, but they want to be rewarded for their investments. We assume the capital has a cost k.

Creating a positive economic value added (EVA) is not always possible. A lot of factors can influence the EVA. One of them is the distribution of the risks. To model the risks or claims, represented by a random vector $\boldsymbol{X}$, we assume $\boldsymbol{X}$ is multivariate elliptically distributed. This allows us to work with the student-t distribution which has fatter tails than those of the normal distribution. The thickness of the tails depends on the number of degrees of freedom, which can be chosen as a parameter. Moreover, the multivariate student-t distribution allows to model tail dependencies, whereas the tail dependency for the normal distribution equals zero.
A possible following stage in the research of this topic is the extension to logelliptical distributions because these are non symmetric distributions and also have fat tails. Here we treat the extension to the elliptical distributions.

Chapter 1

Introduction

1.1 Risk Measures and Value Creation

We denote the total amount of the claims on an insurance portfolio by a random variable S. The individual risks X_i are modelled by a random vector $\boldsymbol{X}$. So we have that $S = X_1 + \ldots + X_n$. To pay the claims the insurer receives a total premium income P from its clients. To ensure that the insurance company can pay its claims with a sufficiently high probability, the regulator demands the insurer to hold a certain amount of money, called the required solvency level (RSL). The RSL consists of the premium P and a certain amount of capital provided by the shareholders. We call this the risk adjusted capital (RAC). So we have

$$RSL = P + RAC. \tag{1.1}$$

There are several possibilities to determine the RSL. If we take the RSL equal to the Value at Risk with a given probability q:

$$VaR_q(S) = \inf\{s \in \mathbb{R} | F_S(s) \geq q\}\,,\ 0 < q < 1,$$

we have a probability of $1 - q$ that $S > VaR_q(S)$. A drawback of the VaR is that it does not take into account the situation once $S > VaR_q(S)$.
Another popular risk measure is the Tail Value at Risk at level q defined by

$$TVaR_q(S) = \frac{1}{1-q}\int_q^1 VaR_p(S)dp\,,\ 0 < q < 1.$$

Here an average of all the quantiles exceeding the VaR_q is taken, so more information about the upper tail is taken into account. Moreover the TVaR is a subadditive risk measure whereas the VaR is not. This means that for two lines of business the following property holds

$$TVaR_q(S_1 + S_2) \leq TVaR_q(S_1) + TVR_q(S_2).$$

In section 1.1.1 we give some remarks about the subadditivity of a risk measure (see also Dhaene (2004-2)).
Finally, we mention the Conditional Tail Expectation defined as

$$CTE_q(S) = E[S | S > VaR_q(S)]\,,\ 0 < q < 1. \tag{1.2}$$

We remark that the TVaR equals the CTE in case of continuous random variable (see Dhaene (2004-1)).
Henceforth we will use the TVaR. So equation (1.1) becomes

$$TVaR_q(S) = P + RAC.$$

As mentioned above the RAC is the capital provided by the shareholders who want to be rewarded for their investments. We define the economic value added (EVA) as

$$EVA = Margin - kRAC,$$

where the Margin is a loading of the pure premium $E[S]$ and k the cost of capital. The return on the invested capital can be measured by

$$RORAC = \frac{Margin}{RAC} = \frac{EVA}{RAC} + k,$$

where RORAC stands for "return on RAC".

1.1.1 Subadditivity of a risk measure

In Dhaene (2004-2) is explained why many people want a risk measure to be subadditive but they also warn that a risk measure can be too subadditive. We explain it briefly.
We already saw the example of the TVaR. In general a risk measure ρ is called subadditive if for two risks X_1, X_2

$$\rho(X_1 + X_2) \leq \rho(X_1) + \rho(X_2).$$

If we assume that the solvency capital requirement imposed by the regulator is given by ρ, the equation above can be explained as follows: for a risk measure ρ, the capital requirement $\rho(X_1 + X_2)$ for the merged risks is smaller than for the added capitals $\rho(X_1)$ and $\rho(X_2)$. There are some arguments to defend this statement. We therefore first have a look at the shortfall of a risk X defined as

$$(X - \rho(X))_+.$$

This is the difference between the claim X and the available capital for that claim if $X > \rho(X)$. In case the capital $\rho(X)$ exceeds the claim X, the shortfall is zero. If two risks X_1, X_2 are merged as one business, the shortfall decreases because

$$(X_1 + X_2 - \rho(X_1) - \rho(X_2))_+ \leq \sum_{i=1}^{2}(X_i - \rho(X_i))_+. \tag{1.3}$$

This is due to the diversification effect, the loss on one risk can be compensated by a gain on the other risk. Because of the decrease in shortfall, it is justified to require a less severe capital requirement for the merged risks. From the viewpoint of the regulator the smaller the shortfall or the higher the capital, the better. On the other hand the capital has a cost. An equilibrium has to be found between these two conflicting criteria. If the risk measure ρ is subadditive, the capital $\rho(X_1 + X_2)$ can be used. But we must take care that the shortfall $(X_1 + X_2 - \rho(X_1 + X_2))_+$ does not become greater than the right side of equation 1.3. In that case the risk measure is too subadditive. In Dhaene (2004-2) an alternative "regularity condition" is proposed to replace the subadditivity condition. This condition takes into account that the shortfall should not increase after a merge.

1.2 Capital Allocation

An insurance company can write several branches $S_1, \ldots, S_n$. The capital provided by the investors is given by

$$RAC = TVaR_q(S) - P,$$

with $S = S_1 + \ldots + S_n$. Let P_i denote the premium for line of business i (LOB$_i$). To distribute the capital over the different lines of business (LOB's) of the company, we use the following allocation rule (see also Denault (2001) and Kalkbrener (2005)):

$$RAC_i = E[S_i | S > VaR_q(S)] - P_i, \tag{1.4}$$

for a each line of business (LOB).

In Dhaene (2005) it is shown that this and other allocation principles can be seen as the minimum of the following optimization problem:

$$\min_{K_1,\ldots,K_n} \sum_{i=1}^{n} E\left[\frac{\epsilon(K_i; S_i)}{\nu_i}\zeta_i\right], \qquad \sum_{i=1}^{n} K_i = K \tag{1.5}$$

where the total capital K denotes the sum of RAC and P. The function ϵ is a measure of the "closeness" between the risk S_i and the capital for that risk K_i. The parameter ζ_i is used as a correction for parameter uncertainty and for the choice of the distribution of the risks. Via ζ_i one can also adapt the risk aversion toward the risk S_i. The other parameter ν_i is a scalar weight which expresses the level of exposure to risk S_i. If we set

$$\epsilon(K_i; S_i) = (S_i - K_i)^2$$

and all ζ_i equal to

$$h(S) = \frac{I(S > VaR_q(S))}{Pr(S > VaR_q(S))},$$

then the optimal solution of (1.5) is given by

$$K_i = E[S_i | S > VaR_q(S)]$$

which is the allocation rule (1.4) we will use.

The total premium is given by the sum

$$P = P_1 + \ldots + P_n.$$

It is clear that the allocation principle (1.4) is additive, i.e.:

$$RAC = RAC_1 + \ldots + RAC_n.$$

We define the value creation per LOB as:

$$EVA_i = Margin_i - kRAC_i. \tag{1.6}$$

We will investigate how many contracts an insurer can sell for the different LOB's such that the total EVA is maximal for a given RAC. If we want to calculate the expectation in equation (1.4) we have to know the distribution of S. A first possibility is to assume a normal distribution for S. In order to allow for fatter tails, we will also consider a student-t distribution, which is, just as the normal distribution, a member of the class of elliptical distributions. In the next chapter we give some notions about this class of distributions. For more details we refer to Landsman and Valdez (2003).

Chapter 2

Elliptical Distributions

In this chapter we shortly investigate the class of multivariate elliptical distributions. Two well known distributions which belong to this class are the normal and student-t distribution, but we also have a look at other elliptical distributions such as the logistic and exponential power distribution.

2.1 Definition

Let $\boldsymbol{X} = (X_1, \ldots, X_n)^T$ be a n-dimensional random vector. Let $\boldsymbol{\mu}$ be a column vector and $\boldsymbol{\Sigma}$ a $n \times n$ positive-definite matrix. We say that $\boldsymbol{X}$ has a multivariate elliptical distribution denoted by

$$\boldsymbol{X} \sim E_n(\boldsymbol{\mu}, \boldsymbol{\Sigma}, g_n),$$

if the function g_n satisfies the condition

$$\int_0^\infty x^{\frac{n}{2}-1} g_n(x)dx < \infty. \tag{2.1}$$

If the density $f_{\boldsymbol{X}}(\boldsymbol{x})$ exists, it can be written as

$$f_{\boldsymbol{X}}(\boldsymbol{x}) = \frac{c_n}{\sqrt{|\boldsymbol{\Sigma}|}} g_n \left[\frac{1}{2} (\boldsymbol{x} - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1} (\boldsymbol{x} - \boldsymbol{\mu}) \right], \tag{2.2}$$

where c_n is a normalizing constant such that

$$\int_{-\infty}^\infty f_{\boldsymbol{X}}(\boldsymbol{x}) d\boldsymbol{x} = 1.$$

The function g_n is called the density generator. The mean and variance-covariance of an elliptical distribution are given by

$$E[X] = \boldsymbol{\mu} \qquad \text{if} \quad \int_0^\infty g_1(x)dx < \infty \tag{2.3}$$

and

$$Cov[X] = -\psi'(0)\boldsymbol{\Sigma} \qquad \text{if} \quad |\psi'(0)| < \infty, \tag{2.4}$$

where ψ is the characteristic generator which is defined in the proof of property 2.1.1.

The following properties will be useful.

Property 2.1.1 *If* $\boldsymbol{X} \sim E_n(\boldsymbol{\mu}, \boldsymbol{\Sigma}, g_n)$ *and* A *is some* $m \times n$ *matrix of rank* $m \leq n$ *and* $\boldsymbol{b}$ *is some m-dimensional column vector, then* $A\boldsymbol{X} + \boldsymbol{b}$ *is distributed as*

$$A\boldsymbol{X} + \boldsymbol{b} \sim E_m(A\boldsymbol{\mu} + \boldsymbol{b}, A\boldsymbol{\Sigma}A^T, g_m)$$

Proof

In Landsman and Valdez (2003) the following alternative definition for an elliptical distribution is given:

The random vector $\boldsymbol{X}$ has a multivariate elliptical distribution, $\boldsymbol{X} \sim E_n(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \psi)$, if its characteristic function $\varphi_{\boldsymbol{X}}(\boldsymbol{t}) = E[e^{i\boldsymbol{t}^T\boldsymbol{X}}]$ can be written as

$$\varphi_{\boldsymbol{X}}(\boldsymbol{t}) = e^{i\boldsymbol{t}^T\boldsymbol{\mu}}\psi(\frac{1}{2}\boldsymbol{t}^T\boldsymbol{\Sigma}\boldsymbol{t}) \tag{2.5}$$

where $\psi(t) \in \Psi_n$ is a function called the characteristic generator and where Ψ_n is the class of functions $\psi(t) : [0, \infty) \to \mathbb{R}$ such that the function $\psi(\sum_{i=1}^n t_i^2)$ is an n-dimensional characteristic function.

From (2.5) it follows that

$$\begin{aligned} \varphi_{A\boldsymbol{X}+\boldsymbol{b}}(\boldsymbol{t}) &= E[e^{i\boldsymbol{t}^T(A\boldsymbol{X}+\boldsymbol{b})}] \\ &= E[e^{i\boldsymbol{t}^T A\boldsymbol{X}}]e^{i\boldsymbol{t}^T\boldsymbol{b}} \\ &= \varphi_X(A^T\boldsymbol{t})e^{i\boldsymbol{t}^T\boldsymbol{b}} \\ &= e^{i\boldsymbol{t}^T(A\boldsymbol{\mu}+\boldsymbol{b})}\psi(\frac{1}{2}\boldsymbol{t}^T A\boldsymbol{\Sigma}A^T\boldsymbol{t}) \end{aligned}$$

Moreover we can conclude that

$$A\boldsymbol{X} + \boldsymbol{b} \sim E_m(A\boldsymbol{\mu} + \boldsymbol{b}, A\boldsymbol{\Sigma}A^T, \psi)$$

which is equivalent to the definition we used

$$A\boldsymbol{X} + \boldsymbol{b} \sim E_m(A\boldsymbol{\mu} + \boldsymbol{b}, A\boldsymbol{\Sigma}A^T, g_m)$$

$\square$

Property 2.1.2 *Suppose* $(X_1, \ldots, X_n)^T = \boldsymbol{X} \sim E_n(\boldsymbol{\mu}, \boldsymbol{\Sigma}, g_n)$, *then the sum* $S = X_1 + \ldots + X_n$ *is also elliptically distributed:*

$$S \sim E_1(\boldsymbol{e}^T\boldsymbol{\mu}, \boldsymbol{e}^T\boldsymbol{\Sigma}\boldsymbol{e}, g_1).$$

Proof

If we take $A = \boldsymbol{e}^T = (1 \cdots 1)$ n-dimensional and $\boldsymbol{b} = 0$ in the previous property we immediately find that $\boldsymbol{e}^T\boldsymbol{X} = X_1 + \ldots + X_n = S$ and

$$S \sim E_1(\boldsymbol{e}^T\boldsymbol{\mu}, \boldsymbol{e}^T\boldsymbol{\Sigma}\boldsymbol{e}, g_1)$$

$\square$

Analogously the following property can be proven.

Property 2.1.3 *Suppose* $(X_1, \ldots, X_n)^T = \boldsymbol{X} \sim E_n(\boldsymbol{\mu}, \boldsymbol{\Sigma}, g_n)$, *then the marginal random vectors* X_i $i = 1, \ldots, n$ *are also elliptically distributed:*

$$X_i \sim E_1(\mu_i, \sigma_i^2, g_1).$$

We now take a look at some specific elliptical distributions.

2.2 Examples

2.2.1 Normal Distribution

If we choose the density generator

$$g_n(u) = e^{-u},$$

we say that $\boldsymbol{X} \sim E_n(\boldsymbol{\mu}, \boldsymbol{\Sigma}, g_n)$ is normally distributed and denote $\boldsymbol{X} \sim N_n(\boldsymbol{\mu}, \boldsymbol{\Sigma})$. The density function (2.2) can be written as

$$f_{\boldsymbol{X}}(\boldsymbol{x}) = \frac{c_n}{\sqrt{|\boldsymbol{\Sigma}|}} e^{-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1}(\boldsymbol{x}-\boldsymbol{\mu})}.$$

Note that,

$$Cov(\boldsymbol{X}) = \boldsymbol{\Sigma},$$

because the characteristic function of the normal distribution is given by $e^{i\boldsymbol{t}^T\boldsymbol{\mu}+\frac{1}{2}i^2\boldsymbol{t}^T\boldsymbol{\Sigma}\boldsymbol{t}}$. This expression can be rewritten as $e^{i\boldsymbol{t}^T\boldsymbol{\mu}}e^{-\frac{1}{2}\boldsymbol{t}^T\boldsymbol{\Sigma}\boldsymbol{t}}$. From (2.5) we now find that $\psi(t) = e^{-t}$ and thus $\psi'(0) = -1$. Substitution in (2.4) ends the proof.

2.2.2 Student-t Distribution

If the density generator is equal to

$$g_n(u) = \left(1 + \frac{u}{k_p}\right)^{-p},$$

we say that $\boldsymbol{X} \sim E_n(\boldsymbol{\mu}, \boldsymbol{\Sigma}, g_n)$ is student-t distributed with parameters $p > n/2$ and k_p. The density function can be written as

$$f_{\boldsymbol{X}}(\boldsymbol{x}) = \frac{c_n}{\sqrt{|\boldsymbol{\Sigma}|}} \left(1 + \frac{(\boldsymbol{x}-\boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1}(\boldsymbol{x}-\boldsymbol{\mu})}{2k_p}\right)^{-p}.$$

Here, $Cov(\boldsymbol{X}) = \boldsymbol{\Sigma}$ does not always hold. If we take the parameters $k_p = \frac{m}{2}$ and $p = \frac{n+m}{2}$ (m, n integer), we have the classical student-t distribution with m degrees of freedom. The density function then becomes

$$f_{\boldsymbol{X}}(\boldsymbol{x}) = \frac{c_n}{\sqrt{|\boldsymbol{\Sigma}|}} \left(1 + \frac{(\boldsymbol{x}-\boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1} (\boldsymbol{x}-\boldsymbol{\mu})}{m}\right)^{-\frac{m+n}{2}}$$

and the covariance of $\boldsymbol{X}$ is equal to $Cov(\boldsymbol{X}) = \frac{m}{m-2}\boldsymbol{\Sigma}$ (see Demarta and McNeil (2005)). Remark that for m going to infinity, we obtain again the normal distribution.

2.2.3 Logistic Distribution

If we choose the density generator equal to

$$g_n(u) = \frac{e^{-u}}{(1+e^{-u})^2},$$

we say that $\boldsymbol{X} \sim E_n(\boldsymbol{\mu}, \boldsymbol{\Sigma}, g_n)$ has a logistic distribution. The density function can be written as

$$f_{\boldsymbol{X}}(\boldsymbol{x}) = \frac{c_n}{\sqrt{|\boldsymbol{\Sigma}|}} \frac{e^{-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1} (\boldsymbol{x}-\boldsymbol{\mu})}}{\left(1 + e^{-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1} (\boldsymbol{x}-\boldsymbol{\mu})}\right)^2}.$$

2.2.4 Exponential Power Distribution

If the density generator is equal to

$$g_n(u) = e^{-ru^s},$$

we say that $\boldsymbol{X} \sim E_n(\boldsymbol{\mu}, \boldsymbol{\Sigma}, g_n)$ is exponential power distributed with parameters $r, s > 0$. The density function can be written as

$$f_{\boldsymbol{X}}(\boldsymbol{x}) = \frac{c_n}{\sqrt{|\boldsymbol{\Sigma}|}} e^{-\frac{r}{2}\left((\boldsymbol{x}-\boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1} (\boldsymbol{x}-\boldsymbol{\mu})\right)^s}.$$

2.3 Tail Dependency

Before going to the next chapter where our model is presented, we have a look at another interesting theoretical concept for insurance business: Tail dependency. Consider two risks X_i, X_j. If the outcome of X_i is large, this means that we face a considerable loss. It might be of importance to know the probability that also the outcome of X_j will be large. In other words, we want to measure the dependency between the tails of the marginal distributions of X_i and X_j. In Embrechts et al. (2002) this probability is measured by tail dependency coefficients. The upper and lower tail dependency coefficient are defined by

$$\begin{aligned}\lambda_u &= \lim_{q\to 1} P(X_2 > F_2^{-1}(q) | X_1 > F_1^{-1}(q)) \\ \lambda_l &= \lim_{q\to 0} P(X_2 \leq F_2^{-1}(q) | X_1 \leq F_1^{-1}(q))\end{aligned}$$

if these limits $\lambda_u, \lambda_l \in [0,1]$ exist and where F_i is the marginal distribution function of X_i. When both coefficients are equal, we use one notation λ. For the student-t distributions there is a nice formula for λ (see Embrechts et al. (2002)):

Property 2.3.1 *For a student-t distributed random variable with m degrees of freedom, the tail dependency coefficient λ can be written as*

$$\lambda = 2F_t\left(-\frac{\sqrt{m+1}\sqrt{1-\rho}}{\sqrt{1+\rho}}\right) \tag{2.6}$$

where ρ is the correlation between two random variables and F_t the standard student-t distribution function.

In table 1 the tail dependency coefficients for different values of m and ρ are given.

	ρ			
m	-0.02	-0.01	0	0.01
3	0.134	0.137	0.139	0.142
4	0.085	0.087	0.089	0.091
5	0.055	0.056	0.058	0.058
6	0.036	0.037	0.038	0.040
7	0.023	0.024	0.025	0.027
8	0.016	0.016	0.017	0.018
9	0.010	0.011	0.012	0.012
10	0.007	0.007	0.008	0.008
15	0.001	0.001	0.001	0.001

Table 1: Tail dependency coefficients.

We see that λ increases for an increasing ρ and decreasing m. Later on we will see that the effect of tail dependency on our results is difficult to analyze.
Note that for the normal distribution $\lambda = 0$. There is no tail dependency in that case.

Chapter 3

The Model

In this chapter we present a model to calculate the Economic Value Added(EVA) for a given portfolio. The model presented is a generalization of the model used in Walhin (2005). Here we use a multivariate elliptical distribution for the claims. At the end some numerical applications are given.

3.1 Tail Conditional Expectation

In section 1.2 we already gave some formulas to calculate the Economic Value Added for each Line Of Business i (EVA_i). However in equation (1.4), i.e.

$$RAC_i = E[S_i|S > VaR_q(S)] - P_i,$$

we face this problem: how to calculate the conditional expectation? In theorem 3.1.3 an answer is given to this question, but we first prove two lemmas. (See also Landsman and Valdez (2003)). An alternative proof, proposed by Vanduffel et al. (2006), is given in theorem 3.1.4.

Lemma 3.1.1 *If* $\boldsymbol{X} = (X_1, \ldots, X_n)^T \sim E_n(\boldsymbol{\mu}, \boldsymbol{\Sigma}, g_n)$ *and* $S = X_1 + \ldots + X_n$*, then also*

$$\boldsymbol{X}_{i,S} = (X_i, S)^T \sim E_2(\boldsymbol{\mu}_{i,S}, \boldsymbol{\Sigma}_{i,S}, g_2) \qquad\qquad i = 1, \ldots, n.$$

If we denote $E[X_i] = \mu_i$*;* $Cov(X_i, X_j) = \sigma_{ij}$*;* $Var[X_i] = \sigma_i^2 = \sigma_{ii}$*;* $Cov(X_i, S) = \sigma_{iS}$ *and* $Var[S] = \sigma_S^2$*, we have*

$$\boldsymbol{\mu}_{i,S} = \begin{pmatrix} \mu_j \\ \sum_{j=1}^n \mu_j \end{pmatrix} \quad \textit{and} \quad \boldsymbol{\Sigma}_{i,S} = \begin{pmatrix} \sigma_i^2 & \sigma_{iS} \\ \sigma_{iS} & \sigma_S^2 \end{pmatrix}$$

Proof
If we apply property 2.1.1 by setting the matrix A equal to

$$\begin{array}{c} \begin{array}{ccccccc} & & & i & & & \\ & & & \downarrow & & & \end{array} \\ \begin{pmatrix} 0 & \cdots & 0 & 1 & 0 & \cdots & 0 \\ 1 & \cdots & 1 & 1 & 1 & \cdots & 1 \end{pmatrix}, \end{array}$$

we find that

$$A\boldsymbol{X} = (X_i, S)^T \sim E_2(A\boldsymbol{\mu}, A\boldsymbol{\Sigma}A^T, g_2).$$

Because

$$A\boldsymbol{\mu} = \begin{pmatrix} \mu_i \\ \sum_{j=1}^n \mu_j \end{pmatrix} = \boldsymbol{\mu}_{i,S}$$

and

$$\begin{aligned} A\boldsymbol{\Sigma}A^T &= \begin{pmatrix} \sigma_{i1} & \cdots & \sigma_{in} \\ \sum_l \sigma_{l1} & \cdots & \sum_l \sigma_{ln} \end{pmatrix} A^T \\ &= \begin{pmatrix} \sigma_i^2 & \sum_l \sigma_{il} \\ \sum_l \sigma_{li} & \sum_{i,j} \sigma_{ij}^2 \end{pmatrix} \\ &= \boldsymbol{\Sigma}_{i,S} \end{aligned}$$

we have that

$$\boldsymbol{X}_{i,S} \sim E_2(\boldsymbol{\mu}_{i,S}, \boldsymbol{\Sigma}_{i,S}, g_2).$$

□

Lemma 3.1.2 *Let $\boldsymbol{Y} = (Y_1, Y_2)^T \sim E_2(\boldsymbol{\mu}, \boldsymbol{\Sigma}, g_2)$ and suppose also that the condition in (2.3) holds. We have then*

$$E[Y_1|Y_2 > y_q] = \mu_1 + \lambda_2\sigma_1\sigma_2\rho_{12}$$

where

$$\begin{aligned} \rho_{12} &= corr(Y_1, Y_2) \\ \lambda_2 &= \frac{\frac{1}{\sigma_2}\overline{G}(\frac{1}{2}z_{2,q}^2)}{\overline{F}_Z(z_{2,q})}, \\ z_{2,q} &= \frac{y_q - \mu_2}{\sigma_2}, \\ G(x) &= c_1 \int_0^x g_1(u)du, \\ \overline{G}(x) &= G(\infty) - G(x), \\ F_Z(z) &= c_1 \int_{-\infty}^z g_1(\frac{1}{2}x^2)dx, \\ \overline{F}_Z(z) &= 1 - F_Z(z). \end{aligned} \tag{3.1}$$

Proof
We start with the left hand side of the equation

$$
\begin{aligned}
E[Y_1|Y_2 > y_q] &\overset{def}{=} \frac{1}{\overline{F}_{Y_2}(y_q)} \int_{-\infty}^{\infty}\int_{y_q}^{\infty} y_1 f_{\boldsymbol{Y}}(y_1, y_2) dy_2 dy_1 \\
&\overset{(2.2)}{=} \frac{1}{\overline{F}_{Y_2}(y_q)} \int_{-\infty}^{\infty}\int_{y_q}^{\infty} y_1 \frac{c_2}{\sqrt{|\boldsymbol{\Sigma}|}} g_2[\frac{1}{2}(\boldsymbol{y}-\boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1}(\boldsymbol{y}-\boldsymbol{\mu})] dy_2 dy_1 \\
&\overset{not}{=} \frac{1}{\overline{F}_Z(_q)} I_1
\end{aligned}
$$

We now work out some components of the integral I_1 separately:

$$
|\boldsymbol{\Sigma}| = \begin{vmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 \end{vmatrix} = \sigma_1^2\sigma_2^2 - \sigma_{12}^2 = \sigma_1^2\sigma_2^2(1-\rho_{12}^2)
$$

and

$$
\begin{aligned}
& (\boldsymbol{y}-\boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1}(\boldsymbol{y}-\boldsymbol{\mu}) \\
= & \frac{1}{\sigma_1^2\sigma_2^2(1-\rho_{12}^2)} \begin{pmatrix} y_1-\mu_1 & y_2-\mu_2 \end{pmatrix} \begin{pmatrix} \sigma_2^2 & -\sigma_{12} \\ -\sigma_{12} & \sigma_1^2 \end{pmatrix} \begin{pmatrix} y_1-\mu_1 \\ y_2-\mu_2 \end{pmatrix} \\
= & \frac{1}{\sigma_1^2\sigma_2^2(1-\rho_{12}^2)} \begin{pmatrix} (y_1-\mu_1)\sigma_2^2 - \sigma_{12}(y_2-\mu_2) & -\sigma_{12}(y_1-\mu_1) + (y_2-\mu_2)\sigma_1^2 \end{pmatrix} \begin{pmatrix} y_1-\mu_1 \\ y_2-\mu_2 \end{pmatrix} \\
= & \frac{1}{\sigma_1^2\sigma_2^2(1-\rho_{12}^2)} \left[(y_1-\mu_1)^2\sigma_2^2 - 2\sigma_{12}(y_2-\mu_2)(y_1-\mu_1) + (y_2-\mu_2)^2\sigma_1^2\right] \\
\overset{\frac{y_i-\mu_i}{\sigma_i}=z_i}{=} & \frac{1}{(1-\rho_{12}^2)}(z_1^2 - 2\rho_{12}z_1z_2 + z_2^2) \\
= & \frac{1}{(1-\rho_{12}^2)}\left((z_1-\rho_{12}z_2)^2 + (1-\rho_{12}^2)z_2^2\right) \\
= & \frac{(z_1-\rho_{12}z_2)^2}{(1-\rho_{12}^2)} + z_2^2
\end{aligned}
$$

By substituting these components, changing the order of integration and using the transformation $\frac{y_i-\mu_i}{\sigma_i} = z_i$, we can rewrite I_1 as

$$
\begin{aligned}
I_1 &= \frac{c_2}{\sqrt{1-\rho_{12}^2}} \int_{\frac{y_q-\mu_2}{\sigma_2}}^{\infty} \int_{-\infty}^{\infty} (\mu_1 + \sigma_1 z_1) g_2\left[\frac{1}{2}\frac{(z_1-\rho_{12}z_2)^2}{(1-\rho_{12}^2)} + \frac{1}{2}z_2^2\right] dz_1 dz_2 \\
&= I_{1a} + I_{1b}
\end{aligned}
$$

where (use the transformation $z_1' = \frac{(z_1-\rho_{12}z_2)}{\sqrt{1-\rho_{12}^2}}$ and property 2.1.3)

$$\begin{aligned}
I_{1a} &= \mu_1 \frac{c_2}{\sqrt{1-\rho_{12}^2}} \int_{z_{2,q}}^{\infty}\int_{-\infty}^{\infty} g_2\left[\frac{1}{2}\frac{(z_1-\rho_{12}z_2)^2}{(1-\rho_{12}^2)}+\frac{1}{2}z_2^2\right]dz_1dz_2 \\
&= \mu_1 \int_{z_{2,q}}^{\infty}\int_{-\infty}^{\infty} c_2g_2\left[\frac{1}{2}{z_1'}^2+\frac{1}{2}z_2^2\right]dz_1'dz_2 \\
&= \mu_1 \int_{z_{2,q}}^{\infty}\int_{-\infty}^{\infty} f_{\boldsymbol{Z}}(z_1',z_2)dz_1'dz_2 \\
&= \mu_1 \overline{F}_Z(z_{2,q})
\end{aligned}$$

and

$$\begin{aligned}
I_{1b} &= \sigma_1 \frac{c_2}{\sqrt{1-\rho_{12}^2}} \int_{z_{2,q}}^{\infty}\int_{-\infty}^{\infty} z_1 g_2\left[\frac{1}{2}\frac{(z_1-\rho_{12}z_2)^2}{(1-\rho_{12}^2)}+\frac{1}{2}z_2^2\right]dz_1dz_2 \\
&= \sigma_1 \int_{z_{2,q}}^{\infty}\int_{-\infty}^{\infty} (\sqrt{1-\rho_{12}^2}z_1'+\rho_{12}z_2)c_2g_2\left[\frac{1}{2}{z_1'}^2+\frac{1}{2}z_2^2\right]dz_1'dz_2 \\
&= I_{1b_1}+I_{1b_2}.
\end{aligned}$$

The first term I_{1b_1} disappears because the integrated function is odd i.e.

$$\sigma_1\sqrt{1-\rho_{12}^2}\int_{z_{2,q}}^{\infty}\underbrace{\int_{-\infty}^{\infty} z_1'c_2g_2\left[\frac{1}{2}{z_1'}^2+\frac{1}{2}z_2^2\right]dz_1'}_{=0}dz_2 = 0$$

We thus have

$$\begin{aligned}
I_{1b} &= I_{1b_2} \\
&= \rho_{12}\sigma_1 \int_{z_{2,q}}^{\infty} z_2 \int_{-\infty}^{\infty} c_2g_2\left[\frac{1}{2}{z_1'}^2+\frac{1}{2}z_2^2\right]dz_1'dz_2 \\
&= \rho_{12}\sigma_1 \int_{z_{2,q}}^{\infty} z_2 \int_{-\infty}^{\infty} f_{\boldsymbol{Z}}(z_1',z_2)dz_1'dz_2 \\
&= \rho_{12}\sigma_1 \int_{z_{2,q}}^{\infty} z_2 f_Z(z_2)dz_2 \\
&= \rho_{12}\sigma_1 \int_{z_{2,q}}^{\infty} z_2 c_1 g_1(\frac{1}{2}z_2^2)dz_2 \\
&\overset{\frac{1}{2}z_2^2=u}{=} \rho_{12}\sigma_1 \int_{\frac{1}{2}z_{2,q}^2}^{\infty} c_1 g_1(u)du \\
&\overset{def}{=} \rho_{12}\sigma_1 \overline{G}\left(\frac{1}{2}z_{2,q}^2\right)
\end{aligned}$$

By substituting all the integrals we finally find

$$\begin{aligned}E[Y_1|Y_2 > y_q] &= \frac{1}{\overline{F}_{Y_2}(y_q)}\left(\mu_1\overline{F}_Z(z_{2,q}) + \rho_{12}\sigma_1\overline{G}\left(\frac{1}{2}z_{2,q}^2\right)\right)\\ &= \mu_1 + \rho_{12}\sigma_1\sigma_2\frac{\frac{1}{\sigma_2}\overline{G}\left(\frac{1}{2}z_{2,q}^2\right)}{\overline{F}_Z(z_{2,q})}\\ &\overset{def}{=} \mu_1 + \rho_{12}\sigma_1\sigma_2\lambda_2\end{aligned}$$

because $\overline{F}_{Y_2}(y_q) = Pr(Y_2 > y_q) = Pr(\frac{Y_2-\mu_2}{\sigma_2} > \frac{y_q-\mu_2}{\sigma_2}) = Pr(Z > z_{2,q}) = \overline{F}_Z(z_{2,q})$.

□

Theorem 3.1.3
Let $\boldsymbol{X} = (S_1, \ldots, S_n)^T \sim E_n(\boldsymbol{\mu}, \boldsymbol{\Sigma}, g_n)$ such that g_1 satisfies the condition

$$\int_0^\infty g_1(x)dx < \infty$$

and let $S = S_1 + \ldots + S_n$. Then the tail conditional expectation can be written as

$$E[S_i|S > VaR_q(S)] = \mu_{S_i} + \lambda_S\sigma_{S_i}\sigma_S\rho_{S_i,S}, \qquad i = 1, \ldots, n$$

where

$$\begin{aligned}\lambda_S &= \frac{\frac{1}{\sigma_S}\overline{G}(\frac{1}{2}z_{S,q}^2)}{\overline{F}_Z(z_{S,q})},\\ z_{S,q} &= \frac{s_q - \mu_S}{\sigma_S},\\ s_q &= VaR_q(S).\end{aligned}$$

and $G(x), \overline{G}(x), F_Z(z)$ and $\overline{F}_Z(z)$ as in (3.1).

Proof
If we set $\boldsymbol{Y} = (S_i, S)^T$, we know from lemma 3.1.1 that $\boldsymbol{Y}$ is elliptically distributed. By using lemma 3.1.2, we have

$$\begin{aligned}E[S_i|S > VaR_q(S)] &\overset{not}{=} E[S_i|S > s_q]\\ &= \mu_{S_i} + \lambda_S\sigma_{S_i}\sigma_S\rho_{S_i,S}\end{aligned}$$

□

In the following theorem the alternative proof of Vanduffel (2006) is given.

Theorem 3.1.4
Let $\boldsymbol{X} = (S_1, \ldots, S_n)^T \sim E_n(\boldsymbol{\mu}, \boldsymbol{\Sigma}, g_n)$ and $S = S_1 + \cdots + S_n$. Then the tail conditional expectation can be written as

$$E[S_i|S > VaR_q(S)] = \mu_{S_i} + \frac{\sigma_{S_i,S}}{\sigma_S^2}(CTE_q(S) - \mu_S), \qquad i = 1, \ldots, n$$

Proof
From the Law of Total Probability, we can write

$$E[S_i|S > VaR_q(S)] = \int_{VaR_q(S)}^{\infty} E[S_i|S = s]dF_S(s|S > VaR_q(S)). \tag{3.2}$$

For the random vector $\boldsymbol{S} \sim E_n(\boldsymbol{\mu}, \boldsymbol{\Sigma}, g_n)$, one can prove that the following result holds:

$$E[S_i|S = s] = \mu_i + \frac{\sigma_{S_i,S}}{\sigma_S^2}(s - \mu_S).$$

Substituting this expression into equation (3.2) yields

$$\begin{aligned} E[S_i|S > VaR_q(S)] &= \mu_i + \frac{\sigma_{S_i,S}}{\sigma_S^2}\int_{VaR_q(S)}^{\infty} sdF_S(s|S > VaR_q(S)) - \frac{\sigma_{S_i,S}}{\sigma_S^2}\mu_S \\ &= \mu_{S_i} + \frac{\sigma_{S_i,S}}{\sigma_S^2}(CTE_q(S) - \mu_S), \end{aligned}$$

because of equation (1.2).

□

3.2 Construction

3.2.1 General Settings

Assume that an insurance company sells policies concerning three kinds of risks:

$$\begin{array}{lll} n_1 & \text{risks of type 1:} & X_1, \ldots, X_{n_1} \\ n_2 & \text{risks of type 2:} & X_{n_1+1}, \ldots, X_{n_1+n_2} \\ n_3 & \text{risks of type 3:} & X_{n_1+n_2+1}, \ldots, X_{n_1+n_2+n_3}. \end{array} \tag{3.3}$$

Let

$$\boldsymbol{X} = (X_1, ..., X_{n_1}, X_{n_1+1}, ..., X_{n_1+n_2}, X_{n_1+n_2+1}, ..., X_{n_1+n_2+n_3})^T,$$

$$S = \underbrace{X_1 + ... + X_{n_1}}_{S_1} + \underbrace{X_{n_1+1} + ... + X_{n_1+n_2}}_{S_2} + \underbrace{X_{n_1+n_2+1} + ...X_{n_1+n_2+n_3}}_{S_3}, \tag{3.4}$$

and

$$\boldsymbol{S} = (S_1, S_2, S_3)^T,$$

Suppose the $n_1 + n_2 + n_3$ risks are elliptically distributed as follows:

$$
\begin{array}{rcl}
\mu_{X_j} & = & \mu_1\,, \quad j = 1, \ldots, n_1 \\
\mu_{X_j} & = & \mu_2\,, \quad j = n_1 + 1, \ldots, n_1 + n_2 \\
\mu_{X_j} & = & \mu_3\,, \quad j = n_1 + n_2 + 1, \ldots, n_1 + n_2 + n_3 \\
\sigma_{X_j} & = & \sigma_1\,, \quad j = 1, \ldots, n_1 \\
\sigma_{X_j} & = & \sigma_2\,, \quad j = n_1 + 1, \ldots, n_1 + n_2 \\
\sigma_{X_j} & = & \sigma_3\,, \quad j = n_1 + n_2 + 1, \ldots, n_1 + n_2 + n_3 \\
\rho_{X_i,X_j} & = & \rho_{11}\,, \quad i = 1, \ldots, n_1, j = i + 1, \ldots, n_1 \\
\rho_{X_i,X_j} & = & \rho_{22}\,, \quad i = n_1 + 1, \ldots, n_1 + n_2, j = i + 1, \ldots, n_1 + n_2 \\
\rho_{X_i,X_j} & = & \rho_{33}\,, \quad i = n_1 + n_2 + 1, \ldots, n_1 + n_2 + n_3, j = i + 1, \ldots, n_1 + n_2 + n_3 \\
\rho_{X_i,X_j} & = & \rho_{12}\,, \quad i = 1, \ldots, n_1, j = n_1 + 1, \ldots, n_1 + n_2 \\
\rho_{X_i,X_j} & = & \rho_{13}\,, \quad i = 1, \ldots, n_1, j = n_1 + n_2 + 1, \ldots, n_1 + n_2 + n_3 \\
\rho_{X_i,X_j} & = & \rho_{23}\,, \quad i = n_1 + 1, \ldots, n_1 + n_2, j = n_1 + n_2 + 1, \ldots, n_1 + n_2 + n_3
\end{array}
$$

We define $\mathbf{\Sigma}$ as the $(n_1 + n_2 + n_3) \times (n_1 + n_2 + n_3)$ matrix with elements

$$
\begin{array}{rcl}
\mathbf{\Sigma}(j,j) & = & \sigma^2_{X_j}\,,\ 1 \leq j \leq n_1 + n_2 + n_3, \\
\mathbf{\Sigma}(i,j) & = & \rho_{X_i,X_j}\sigma_{X_i}\sigma_{X_j}\,,\ 1 \leq i \neq j \leq n_1 + n_2 + n_3.
\end{array}
$$

By definition we then have that $Cov(\boldsymbol{X}) = \mathbf{\Sigma}$ if $\boldsymbol{X}$ has a multivariate normal distribution. If we take for example $n_1 = n_2 = n_3 = 3$, we have

$$
\mathbf{\Sigma} = \begin{pmatrix}
\sigma_1^2 & \rho_{11}\sigma_1^2 & \rho_{11}\sigma_1^2 & \rho_{12}\sigma_1\sigma_2 & \rho_{12}\sigma_1\sigma_2 & \rho_{12}\sigma_1\sigma_2 & \rho_{13}\sigma_1\sigma_3 & \rho_{13}\sigma_1\sigma_3 & \rho_{13}\sigma_1\sigma_3 \\
\rho_{11}\sigma_1^2 & \sigma_1^2 & \rho_{11}\sigma_1^2 & \rho_{12}\sigma_1\sigma_2 & \rho_{12}\sigma_1\sigma_2 & \rho_{12}\sigma_1\sigma_2 & \rho_{13}\sigma_1\sigma_3 & \rho_{13}\sigma_1\sigma_3 & \rho_{13}\sigma_1\sigma_3 \\
\rho_{11}\sigma_1^2 & \rho_{11}\sigma_1^2 & \sigma_1^2 & \rho_{12}\sigma_1\sigma_2 & \rho_{12}\sigma_1\sigma_2 & \rho_{12}\sigma_1\sigma_2 & \rho_{13}\sigma_1\sigma_3 & \rho_{13}\sigma_1\sigma_3 & \rho_{13}\sigma_1\sigma_3 \\
\rho_{12}\sigma_1\sigma_2 & \rho_{12}\sigma_1\sigma_2 & \rho_{12}\sigma_1\sigma_2 & \sigma_2^2 & \rho_{22}\sigma_2^2 & \rho_{22}\sigma_2^2 & \rho_{23}\sigma_2\sigma_3 & \rho_{23}\sigma_2\sigma_3 & \rho_{23}\sigma_2\sigma_3 \\
\rho_{12}\sigma_1\sigma_2 & \rho_{12}\sigma_1\sigma_2 & \rho_{12}\sigma_1\sigma_2 & \rho_{22}\sigma_2^2 & \sigma_2^2 & \rho_{22}\sigma_2^2 & \rho_{23}\sigma_2\sigma_3 & \rho_{23}\sigma_2\sigma_3 & \rho_{23}\sigma_2\sigma_3 \\
\rho_{12}\sigma_1\sigma_2 & \rho_{12}\sigma_1\sigma_2 & \rho_{12}\sigma_1\sigma_2 & \rho_{22}\sigma_2^2 & \rho_{22}\sigma_2^2 & \sigma_2^2 & \rho_{23}\sigma_2\sigma_3 & \rho_{23}\sigma_2\sigma_3 & \rho_{23}\sigma_2\sigma_3 \\
\rho_{13}\sigma_1\sigma_3 & \rho_{13}\sigma_1\sigma_3 & \rho_{13}\sigma_1\sigma_3 & \rho_{23}\sigma_2\sigma_3 & \rho_{23}\sigma_2\sigma_3 & \rho_{23}\sigma_2\sigma_3 & \sigma_3^2 & \rho_{33}\sigma_3^2 & \rho_{33}\sigma_3^2 \\
\rho_{13}\sigma_1\sigma_3 & \rho_{13}\sigma_1\sigma_3 & \rho_{13}\sigma_1\sigma_3 & \rho_{23}\sigma_2\sigma_3 & \rho_{23}\sigma_2\sigma_3 & \rho_{23}\sigma_2\sigma_3 & \rho_{33}\sigma_3^2 & \sigma_3^2 & \rho_{33}\sigma_3^2 \\
\rho_{13}\sigma_1\sigma_3 & \rho_{13}\sigma_1\sigma_3 & \rho_{13}\sigma_1\sigma_3 & \rho_{23}\sigma_2\sigma_3 & \rho_{23}\sigma_2\sigma_3 & \rho_{23}\sigma_2\sigma_3 & \rho_{33}\sigma_3^2 & \rho_{33}\sigma_3^2 & \sigma_3^2
\end{pmatrix}. \tag{3.5}
$$

The matrix $\mathbf{\Sigma}$ has to be positive semi-definite. It is shown in Walhin (2005) that the following conditions are sufficient for positive semi-definiteness:

$$
\begin{array}{rcl}
\rho_{11}\rho_{22}\rho_{33} + 2\rho_{12}\rho_{13}\rho_{23} & \geq & \rho_{11}\rho_{23}^2 + \rho_{22}\rho_{13}^2 + \rho_{33}\rho_{12}^2 \\
\rho_{11}\rho_{22} & \geq & \rho_{12}^2 \\
\rho_{11}\rho_{33} & \geq & \rho_{13}^2 \\
\rho_{22}\rho_{33} & \geq & \rho_{23}^2 \\
\rho_{11} & \geq & 0 \\
\rho_{22} & \geq & 0 \\
\rho_{33} & \geq & 0.
\end{array} \tag{3.6}
$$

3.2.2 EVA Calculation

Using Property 2.1.2 and (3.4) we have that $S_k \sim E_1(\mu_{S_k}, \sigma^2_{S_k}, g_1)$ for all $k \in \{1,2,3\}$ where

$$\begin{aligned} \mu_{S_k} &= n_k \mu_k \\ \sigma^2_{S_k} &= n_k \sigma_k^2 + n_k(n_k - 1)\rho_{kk}\sigma_k^2 \end{aligned}$$

(this can easily been seen from (3.5)), but also that $S = S_1 + S_2 + S_3 \sim E_1(\mu_S, \sigma^2_S, g_1)$ where

$$\begin{aligned} \mu_S &= \sum_{k=1}^{3} \mu_{S_k} = \sum_{k=1}^{3} n_k \mu_k \\ \sigma_S^2 &= \sum_{k=1}^{3} \left(n_k \sigma_k^2 + n_k(n_k-1)\rho_{kk}\sigma_k^2 \right) + 2n_1 n_2 \rho_{1,2}\sigma_1\sigma_2 + 2n_1 n_3 \rho_{1,3}\sigma_1\sigma_3 + 2n_2 n_3 \rho_{2,3}\sigma_2\sigma_3. \end{aligned}$$

The last equation is found by taking the sum of all elements of the matrix $\mathbf{\Sigma}$. We can now define the premium as

$$P = \sum_{k=1}^{3} P_k = \sum_{k=1}^{3} (1+\alpha_k) n_k \mu_k,$$

so that

$$Margin = \sum_{k=1}^{3} \alpha_k n_k \mu_k = \sum_{k=1}^{3} Margin_k.$$

By using theorem 3.1.3 we can rewrite the capital allocation rule (1.4) as:

$$\begin{aligned} RAC_i &= E[S_i | S > VaR_q(S)] - P_i \\ &= \mu_{S_i} + \lambda_S \sigma_{S_i} \sigma_S \rho_{S_i,S} - P_i \\ &= \lambda_S \sigma_{S_i} \sigma_S \rho_{S_i,S} - Margin_i. \end{aligned} \tag{3.7}$$

We have now all the information to calculate the EVA$_i$ defined in (1.6):

$$EVA_i = Margin_i - kRAC_i.$$

3.3 Numerical Applications

Our goal was maximizing the EVA by changing the number of contracts n_1, n_2, n_3 with a capital constraint. We will investigate several cases with different parameter sets and distributions. We start with the following basic parameters (these are used in Walhin

(2005)):

$$\begin{aligned}
\mu_1 = \mu_2 = \mu_3 &= 1 \\
\sigma_1 = \sigma_2 = \sigma_3 &= 1 \\
\rho_{1,1} = \rho_{2,2} = \rho_{3,3} &= 0.1 \\
\rho_{1,2} &= -0.01 \\
\rho_{1,3} &= -0.01 \\
\rho_{2,3} &= 0.01 \\
\alpha_1 = \alpha_2 = \alpha_3 &= 0.1 \\
q &= 0.99 \\
\text{cost of capital} &= 0.15 \\
RAC &= 100
\end{aligned} \tag{3.8}$$

3.3.1 Normal Distribution

Here we assume that $\boldsymbol{X} \sim N_n(\boldsymbol{\mu}, \boldsymbol{\Sigma})$. For the basic parameters (3.8) we find:

LOB	n	RAC	Margin	EVA	RORAC
1	94	37.09	9.4	3.84	25.34%
2	79	30.95	7.9	3.26	25.53%
3	80	31.75	8	3.24	25.20%
Total	253	99.79	25.3	10.33	25.35%

Table 2: Optimal portfolio for basic parameters (3.8).

This means that to create the highest EVA on a given capital of 100, you should sell 94 policies of type 1, 79 policies of type 2 and 80 policies of type 3. How the capital is distributed over the 3 lines of business can be seen in the third column. Because LOB_1 has the most contracts, it absorbs the most capital.

We now create a portfolio where the margin of LOB_1 is negative and for which the risks in LOB_1 are negatively correlated with the risks in LOB_2 and LOB_3 by changing α_1 to -0.01 and $\rho_{1,2}$ and $\rho_{1,3}$ to -0.02 in (3.8). The results are given in table 3.

LOB	n	RAC	Margin	EVA	RORAC
1	24	-0.37	-0.24	-0.19	66.72%
2	93	50.16	9.3	1.78	18.54%
3	93	50.16	9.3	1.78	18.54%
Total	210	99.96	18.36	3.37	18.37%

Table 3: Change in (3.8) $\alpha_1 = -0.01$ and $\rho_{1,2} = \rho_{1,3} = -0.02$.

We see that in this situation the EVA has decreased. For LOB_1 the EVA is even negative. The manager of a company with such results could think about dropping this branch. We then find

LOB	n	RAC	Margin	EVA	RORAC
1	0	0	0	0	.
2	90	49.76	9	1.54	18.09%
3	90	49.76	9	1.54	18.09%
Total	180	99.51	18	3.07	18.09%

Table 4: Abandoning LOB_1.

It seems that it is better to keep LOB_1, even when the business itself creates a negative EVA_1 due to the negative margin. Because of the negative correlation to LOB_2 and LOB_3 the total EVA is higher than without LOB_1. This is an excellent example of diversification of the portfolio. Because RAC_1 is negative, the other branches can consume more capital than when LOB_1 is dropped.

If we start again from the basic parameters given in (3.8) but double the RAC to 200, we find an optimal portfolio as given in table 5.

LOB	n	RAC	Margin	EVA	RORAC
1	194	73.54	19.4	8.37	26.38%
2	165	63.06	16.5	7.04	26.17%
3	165	63.06	16.5	7.04	26.17%
Total	524	199.66	52.4	22.45	26.24%

Table 5: Doubled initial capital.

The EVA has more than doubled. So due to the diversification, there is a relative gain. The more risks you write, the less capital they individually consume.

Finally we change the parameters of the first n_1 risks (starting from (3.8)) into $\mu_1 = 2$ and $\sigma_1 = 2$. The risks of type 1 are now "more dangerous". This yields

LOB	n	RAC	Margin	EVA	RORAC
1	45	38.35	9	3.25	23.47%
2	78	30.39	7.8	3.24	25.67%
3	79	31.18	7.9	3.22	25.33%
Total	202	99.92	24.7	9.71	24.72%

Table 6: Enlarging the first type of risks.

We find a decrease in the EVA because the risks of type 1 are larger than before. These risks consume a lot of capital and there is less diversification.
All these results are in line with the calculations in Walhin (2005).

3.3.2 Student-t Distribution

As already mentioned above, the normal distribution does not have large tails. It is therefore interesting to use the student-t distribution with m degrees of freedom to model S. The smaller m, the larger the tails. For m going to infinity we find the normal distribution. In section 2.2.2 we saw that $Cov(\boldsymbol{X}) = \frac{m}{m-2}\boldsymbol{\Sigma}$. We will consider only student-t distributions where $m \geq 3$. To make the calculations comparable to those of the previous section 3.3.1, we want that $Cov(\boldsymbol{X}) = \boldsymbol{\Sigma}$ holds. Therefore, we will multiply the basic parameters σ_i with a factor $\sqrt{\frac{m-2}{m}}$. We have a new basic parameter set as given by (3.9).

$$\begin{aligned} \mu_1 = \mu_2 = \mu_3 &= 1 \\ \sigma_1' = \sigma_2' = \sigma_3' &= \sqrt{\frac{m-2}{m}} \\ \rho_{1,1} = \rho_{2,2} = \rho_{3,3} &= 0.1 \\ \rho_{1,2} &= -0.01 \\ \rho_{1,3} &= -0.01 \\ \rho_{2,3} &= 0.01 \\ \alpha_1 = \alpha_2 = \alpha_3 &= 0.1 \\ q &= 0.99 \\ \text{cost of capital} &= 0.15 \\ RAC &= 100 \end{aligned} \tag{3.9}$$

From (3.5) it is now clear that

$$Cov(\boldsymbol{X}) = \boldsymbol{\Sigma}$$

because we have $Cov(\boldsymbol{X}) = \frac{m}{m-2}\boldsymbol{\Sigma}'$ and $\boldsymbol{\Sigma}' = \frac{m-2}{m}\boldsymbol{\Sigma}$.
We will calculate the optimal number of contracts to maximize the EVA, just as we did for the the normal distribution, but this time for different values of m of the student-t distribution.

m=3

With the adapted basic parameters (3.9) it is better to sell no contracts at all. If you underwrite some policies anyway, the EVA will be negative. Denoting the tail dependency coefficient between LOB_i and LOB_j with λ_{ij}, we find from (2.6): $\lambda_{11} = \lambda_{22} = \lambda_{33} = 0.168$, $\lambda_{12} = \lambda_{13} = 0.137$ and $\lambda_{23} = 0.142$.
Changing the parameters α_1 to -0.01 and $\rho_{1,2}$ and $\rho_{1,3}$ to -0.02 also leads to an economic value added of 0. Since n_1 is then zero, LOB_1 is already dropped. The value for λ_{12} and λ_{13} has slightly decreased to 0.134.
Of course, enlarging the first n_1 risks by putting $\mu_1 = 2$ and $\sigma_1' = 2\sqrt{\frac{m-2}{m}}$, makes it only worse. The EVA remains zero.
Taking twice the capital RAC yields

LOB	n	RAC	Margin	EVA	RORAC
1	115	73.47	11.5	0.48	15.65%
2	98	63.23	9.8	0.32	15.50%
3	98	63.23	9.8	0.32	15.50%
Total	311	199.92	31.1	1.11	15.56%

Table 7: Student-t (m=3) for the basic parameters (3.9).

Now we have a positive EVA. Since there is more capital available, we can diversify more and the underwriting policy becomes profitable with a maximal EVA for n_1, n_2, n_3 resp. $115, 98, 98$. This means that great insurers have an advantage to the smaller ones. They can underwrite these kind of risks whereas an insurer with less capital is not able to do this, although the risks are the same.

m=4

In table 8 the results are given for the basic parameter set (3.9).

LOB	n	RAC	Margin	EVA	RORAC
1	61	36.59	6.1	0.61	16.67%
2	52	31.49	5.2	0.48	16.51%
3	52	31.49	5.2	0.48	16.51%
Total	165	99.57	16.5	1.56	16.57%

Table 8: Student-t (m=4) for the basic parameters (3.9).

The tail dependency coefficients are: $\lambda_{11} = \lambda_{22} = \lambda_{33} = 0.113$, $\lambda_{12} = \lambda_{13} = 0.087$ and $\lambda_{23} = 0.091$.
When we change the parameter α_1 to -0.01 and $\rho_{1,2}$ and $\rho_{1,3}$ to -0.02, it is not possible to create a positive EVA. We have $\lambda_{12} = \lambda_{13} = 0.085$. Dropping LOB_1 has no sense of course.

If we double the RAC, we get again more than twice the EVA:

LOB	n	RAC	Margin	EVA	RORAC
1	129	74.29	12.9	1.76	17.36%
2	109	62.65	10.9	1.50	17.40%
3	109	62.65	10.9	1.50	17.40%
Total	347	199.58	34.7	4.76	17.39%

Table 9: Student-t (m=4) with double capital.

Starting from (3.9) and enlarging the risks of type 1, leads to the results as given in table 10.

LOB	n	RAC	Margin	EVA	RORAC
1	28	36.26	5.6	0.16	15.44%
2	52	31.84	5.2	0.42	16.33%
3	52	31.84	5.2	0.42	16.33%
Total	132	99.94	16	1.01	16.01%

Table 10: Student-t (m=4) with larger risks in LOB_1.

Higher Degrees of Freedom

As we have seen, changing α_1 to -0.01 and $\rho_{1,2}$ and $\rho_{1,3}$ to -0.02 in (3.9) results in an EVA of zero for $m \in \{3, 4\}$. Only from $m = 8$ on the EVA becomes positive for this parameter set, we then have

LOB	n	RAC	Margin	EVA	RORAC
1	17	-0.49	-0.17	-0.10	34.83%
2	76	49.52	7.6	0.17	15.35%
3	77	50.75	7.7	0.09	15.17%
Total	170	99.78	15.13	0.16	15.16%

Table 11: Student-t (m=8) with $\alpha_1 = -0.01$ and $\rho_{1,2} = \rho_{1,3} = -0.02$.

Let us now compare this result with table 12, where we assume the basic parameters (3.9) and $m = 8$.

LOB	n	RAC	Margin	EVA	RORAC
1	76	36.41	7.6	2.14	20.87%
2	65	31.61	6.5	1.76	20.56%
3	65	31.61	6.5	1.76	20.56%
Total	206	99.64	20.6	5.65	20.67%

Table 12: Student-t (m=8) with basic parameters (3.9).

We conclude that the EVA has decreased. This was expected because the margin of LOB_1 was made negative. In spite of this negative margin, it remains possible to underwrite these risks and to create a positive EVA due to the negative correlation to the other branches. Moreover, if we drop LOB_1 the EVA becomes zero again.

Setting m equal to 7, α_1 to -0.01 and $\rho_{1,2}$ and $\rho_{1,3}$ to -0.02 results in an EVA of zero. As before, omitting LOB_1 is not an interesting option (although it seems attractive because α_1 became negative). But if we raise the capital to 500 we can get a positive economic value added.

LOB	n	RAC	Margin	EVA	RORAC
1	110	-6.44	-1.1	-0.13	17.08%
2	395	253.10	39.5	1.54	15.61%
3	395	253.10	39.5	1.54	15.61%
Total	900	499.75	77.9	2.94	15.59%

Table 13: Student-t (m=7) with $RAC = 500$.

We can conclude that it is possible to underwrite risks with a negative margin because the negative dependency with the other risks and/or because a higher capital offers possibilities of diversification.

Summary

In table 14 we present the values for the EVA for the different cases we discussed:

- case 1: basic parameters (3.8) or (3.9).
- case 2: let $\alpha_1 = -0.01$ and $\rho_{1,2} = \rho_{1,3} = -0.02$ in case 1.
- case 3: omit LOB_1 in case 2.
- case 4: let $RAC = 500$ in case 2.
- case 5: let $RAC = 200$ in case 1.
- case 6: double μ_1 and σ_1 in case 1.

We have put the EVA in bold and the optimal numbers of contracts (n_1 n_2 n_3) between brackets.

Distribution	**Cases**					
STUD-T	1	2	3	4	5	6
$m = 3$	**0** (0 0 0)	**0** (0 0 0)	**0** (0 0 0)	**0** (0 0 0)	**1.11** (115 98 98)	**0** (0 0 0)
$m = 4$	**1.56** (61 52 52)	**0** (0 0 0)	**0** (0 0 0)	**0** (0 0 0)	**4.76** (129 109 109)	**1.01** (28 52 52)
$m = 5$	**3.07** (67 56 57)	**0** (0 0 0)	**0** (0 0 0)	**0** (0 0 0)	**7.82** (140 119 119)	**2.51** (31 56 57)
$m = 6$	**4.20** (71 60 61)	**0** (0 0 0)	**0** (0 0 0)	**0.007** (107 308 308)	**10.05** (148 126 126)	**3.61** (33 60 60)
$m = 7$	**5.03** (74 63 63)	**0** (0 0 0)	**0** (0 0 0)	**2.94** (111 395 395)	**11.74** (155 131 131)	**4.43** (35 62 62)
$m = 8$	**5.65** (76 65 65)	**0.16** (17 76 77)	**0** (0 0 0)	**5.19** (115 406 407)	**13.04** (160 135 135)	**5.05** (36 64 64)
$m = 9$	**6.16** (78 66 67)	**0.52** (19 78 79)	**0.31** (0 76 77)	**6.96** (117 415 416)	**14.06** (164 138 138)	**5.56** (37 65 66)
$m = 10$	**6.56** (80 67 68)	**0.80** (21 80 80)	**0.58** (0 77 78)	**8.39** (120 422 423)	**14.91** (167 141 141)	**5.96** (38 66 67)
$m = 15$	**7.83** (85 71 72)	**1.65** (18 84 84)	**1.42** (0 82 82)	**12.71** (126 444 445)	**17.42** (176 149 149)	**7.22** (40 71 71)
NORMAL						
	10.33 (94 79 80)	**3.37** (24 93 93)	**3.07** (0 90 90)	**21.37** (143 489 489)	**22.45** (194 165 165)	**9.71** (45 78 79)

Table 14: Optimal portfolio and maximal EVA for different distributions and cases.

The lower m, the fatter the tails. It is therefore logical that the EVA decreases as m decreases. Compared to the normal distribution, the student-t distribution results in a lower value creation. The conclusions for the different cases are the same for both distributions. Furthermore we remark that the larger the tails are, the stronger the relative effect of changing a parameter is. From case 1 to case 2 for example, the EVA decreases to 30% of the initial value for the normal distribution. Considering the same change but for the student-t with $m = 9$, results in a decrease to 8%. The same reasoning can be made for other changes. We conclude that the distribution of the underwriting risk plays an important role.

If we have a look at table 1, we see that λ depends on m and ρ. The EVA also changes for different values of m and ρ. However, the effect of the tail dependency on the EVA is difficult to analyze. In general an increasing λ leads to a decrease of the EVA.

In table 15 we calculate the optimal portfolio with the following non symmetric parameters: $\mu_1 = 2$, $\mu_2 = 3$, $\mu_3 = 4$; $\sigma'_1 = \sqrt{\frac{m-2}{m}}$, $\sigma'_2 = 2\sqrt{\frac{m-2}{m}}$, $\sigma'_3 = 3\sqrt{\frac{m-2}{m}}$; $\rho_{1,1} = 0.05$, $\rho_{2,2} = 0.1$, $\rho_{3,3} = 0.15$, $\rho_{1,2} = \rho_{1,3} = \rho_{2,3} = 0.01$; $\alpha_1 = \alpha_2 = \alpha_3 = 0.1$; $q = 0.99$, $k = 0.15$ and RAC= 100.

LOB	n	RAC	Margin	EVA	RORAC
1	207	75.00	41.4	30.15	0.55%
2	28	16.88	8.4	5.87	0.50%
3	9	7.95	3.6	2.41	0.45%
Total	244	99.83	53.4	38.42	0.53%

Table 15: Optimal portfolio with non symmetric parameters.

By taking these parameters, the risks of LOB 3 are larger than those of LOB 2 which are "less attractive" than those of LOB 1. However, it remains interesting to underwrite risks of LOB 3. Abandoning LOB 3 results in a lower EVA. We can again conclude that local decisions are not appropriate. One always has to consider the total EVA for all possible scenarios.

To end this section, we may discuss equation (3.7)

$$RAC_i = \lambda_S \sigma_{S_i} \sigma_S \rho_{S_i,S} - Margin_i.$$

In table 11 we see that $RAC_1 < 0$ because ρ_{12} and ρ_{13} were set to -0.02. Having $\rho_{S_i,S} < 0$ is not necessary to make RAC$_i$ negative. We can also take a high margin. An example is shown in table 16 with all ρ_{ij} equal to zero. For the other parameters we set μ_i and σ_i equal to 1, α_i equal to 0.1 and m equal to 4. The results are given for a portfolio with 500 contracts for each LOB. An optimal portfolio does not exist in this case: the more contracts you underwrite, the more value you create. All the capital is provided by the margin.

LOB	n	RAC	Margin	EVA	RORAC
1	500	-2.34	50	50.35	-21.35%
2	500	-2.34	50	50.35	-21.35%
3	500	-2.34	50	50.35	-21.35%
Total	1500	-7.03	150	151.05	-21.35%

Table 16: Portfolio with negative RAC$_i$.

3.4 Reinsurance

In this section we analyze the effect of an unlimited stop-loss reinsurance with a threshold d. The retention for the cedent is then $min(S, d)$, so less capital is required. Taking the same RAC allows for more risks to be underwritten. On the other hand a reinsurance premium has to be paid. We investigate the effects on the economic value added. We assume that there is no credit risk, i.e. that the reinsurer never defaults.

The pure reinsurance premium is defined as

$$\begin{aligned} PP^{Re} &= E\left[(S-d)_+\right] \\ &= \int_d^\infty (1 - F_S(s))ds \\ &\overset{\frac{s-\mu_S}{\sigma_S}=y}{=} \sigma_S \int_{\frac{d-\mu_S}{\sigma_S}}^\infty (1 - F_Y(y))dy \end{aligned} \tag{3.10}$$

where we set $\frac{S-\mu_S}{\sigma_S}$ equal to Y. The second equality can be proven both analytically as in a graphical way (for more details we refer to Kaas et al. (2004)). The commercial premium P^{Re} contains a loading β, such that $P^{Re} = (1+\beta)PP^{Re}$. The retention is denoted as S^{Ret}.

In the future we will use the threshold $d = VaR_{q_{Ret}}(S) = VaR_{99\%}(S)$. The formulas for S^{Ret} are a little different from those for S. We define them below

$$\begin{aligned}
VaR_{99\%}(S^{Ret}) &= d \\
TVaR_{99\%}(S^{Ret}) &= d \\
RAC(S^{Ret}) &= d - (P - P^{Re}) \\
Margin(S^{Ret}) &= Margin(S) - \beta PP^{Re} \\
EVA(S^{Ret}) &= Margin(S^{Ret}) - kRAC(S^{Ret}) \\
RORAC(S^{Ret}) &= \frac{Margin(S^{Ret})}{RAC(S^{Ret})}
\end{aligned}$$

We recalculate the optimal portfolio (n_1, n_2, n_3), but maximize the EVA for S^{Ret} instead of the EVA for S. We will not analyze the value creation and the RORAC for each line of business because we consider a stop-loss reinsurance on the whole portfolio. As before we give the results for the different distributions.

3.4.1 Normal Distribution

Remark that the pure reinsurance premium PP^{Re}, as defined in (3.10), for this distribution can be rewritten as

$$PP^{Re} = \sigma_S \phi(\frac{d - \mu_S}{\sigma_S}) - (d - \mu_S)(1 - \Phi(\frac{d - \mu_S}{\sigma_S}))$$

A proof of this expression can be found in Kaas et al. (2004). In table 17 we find the optimal underwriting policy for the basic parameters (3.8) and different values of the loading β.

β	n_1	n_2	n_3	d	RAC(S^{Ret})	EVA(S^{Ret})	RORAC(S^{Ret})
no	94	79	80	362.19	99.79	10.33	25.35%
1	113	95	95	432.64	99.72	15.15	30.20%
4	112	95	95	431.23	99.98	14.45	29.45%
9	111	94	94	427.01	99.97	13.23	28.23%
14	110	93	93	422.78	99.95	12.02	27.03%
19	109	92	92	418.55	99.91	10.84	25.85%

Table 17: Optimal portfolio with reinsurance for normal distribution.

We see that buying reinsurance allows to increase the optimal number of contracts for all lines of business, even when the commercial premium is twenty times higher than the pure premium. In the next section we consider the student-t based model.

3.4.2 Student-t Distribution

m=3

As already mentioned above, the lower m, the larger the tails. The table with the results for reinsurance is given below for $m = 3$. We have taken the basic parameter set (3.9).

β	n_1	n_2	n_3	d	RAC(S^{Ret})	EVA(S^{Ret})	RORAC(S^{Ret})
no	0	0	0	0	0	0	0
1	95	80	80	378.96	99.81	9.86	24.88%
4	93	78	79	371.66	99.96	7.37	22.37%
9	89	76	76	358.51	99.78	3.40	18.41%
14	0	0	0	0	0	0	0
19	0	0	0	0	0	0	0

Table 18: Optimal portfolio with reinsurance for student-t distribution (m=3).

This is an interesting result. In table 18, we see that without buying a reinsurance cover, no value can be created so it is better not to sell any contracts. But now, reinsurance allows to underwrite risks and to create value. However the stop-loss premium cannot be too high, otherwise the EVA again becomes zero again.

m=4

β	n_1	n_2	n_3	d	RAC(S^{Ret})	EVA(S^{Ret})	RORAC(S^{Ret})
no	61	52	52	248.31	99.57	1.56	16.57%
1	94	79	79	375.89	99.66	9.76	24.80%
4	93	78	78	371.49	100	7.97	22.97%
9	90	76	76	361.22	99.71	5.02	20.04%
14	88	74	74	352.43	99.70	2.23	17.24%
19	0	0	0	0	0	0	0

Table 19: Optimal portfolio with reinsurance for student-t distribution (m=4).

There can be a positive EVA created without buying reinsurance, but the EVA with reinsurance is much higher as long as the reinsurance premium is not unreasonably high.

The same conclusions can be made for higher degrees of freedom. The EVA will become higher and higher because the tails shrink as m increases. We present the tables once more for $m \in \{5, 10, 15\}$.

m=5

β	n_1	n_2	n_3	d	RAC(S^{Ret})	EVA(S^{Ret})	RORAC(S^{Ret})
no	67	56	57	268.84	99.56	3.07	18.08%
1	96	81	82	384.09	99.99	10.50	25.50%
4	95	80	80	378.25	99.74	8.95	23.97%
9	93	78	79	370.96	99.87	6.50	21.51%
14	91	77	77	363.66	99.92	4.14	19.15%
19	89	75	76	356.37	99.89	1.87	16.87%

Table 20: Optimal portfolio with reinsurance for student-t distribution (m=5).

m=10

β	n_1	n_2	n_3	d	RAC(S^{Ret})	EVA(S^{Ret})	RORAC(S^{Ret})
no	80	67	68	314.49	99.57	6.56	21.59%
1	103	88	88	406.33	99.98	12.63	27.63%
4	102	87	87	402.02	99.79	11.54	26.56%
9	101	85	86	396.28	99.78	9.81	24.83%
14	100	84	84	390.54	99.73	8.12	23.14%
19	98	83	83	384.80	99.65	6.47	21.50%

Table 21: Optimal portfolio with reinsurance for student-t distribution (m=10).

m=15

β	n_1	n_2	n_3	d	RAC(S^{Ret})	EVA(S^{Ret})	RORAC(S^{Ret})
no	85	71	72	331.06	99.82	7.83	22.84%
1	107	90	90	415.20	99.99	13.46	28.46%
4	106	89	89	410.92	99.73	12.48	27.51%
9	105	88	88	406.65	99.93	10.96	25.97%
14	103	87	87	400.94	99.77	9.44	24.46%
19	102	86	86	396.66	99.92	7.99	23.00%

Table 22: Optimal portfolio with reinsurance for student-t distribution (m=15).

Conclusion

We made a lot of EVA-calculations for different parameter sets within a multivariate normal and student-t framework. Similar calculations can be made for other elliptical distributions such as the logistic and exponential power distribution. Our goal was to maximize the EVA under a capital constraint by varying the number of contracts underwritten. To model the claim amount S we used a multivariate normal or a multivariate student-t distribution. When the distribution for the claims seems to have large tails, it is more appropriate to choose a student-t distribution. The thickness of the tails then depends on the number of degrees of freedom m.

We saw that the EVA is lower for a student-t distribution and decreases for a decreasing m. For student-t distributions we have tail dependencies whereas in the normal case the tail dependency is zero.

The following conclusions which can be made for the normal case still hold for the student-t distribution. First, when the capital is doubled, the number of contracts underwritten and the EVA are more than doubled. Second, when the risks are enlarged, the EVA decreases. Finally, negative correlated risks have a positive impact on the EVA, even when the margin is negative.

It is important to evaluate the portfolio as a whole because it is possible to write businesses with negative margins due to the correlation structure between the LOB's. In particular making decisions on the basis of local EVA's may lead to destroying value situations. Therefore only the global EVA makes sense.

In the last section it was shown that a stop-loss reinsurance has a positive effect on the EVA as long as the premium is not unreasonably high.

Finally we remark that our model only takes into account the underwriting risk and the cost of the capital. Other risks and expenses may have an impact on the EVA.

PART II

STOCHASTIC CAPITAL MODEL FOR LARGE MTPL CLAIMS

General Introduction

In this part we develop a stochastic model for large claims in the Motor Third Party Liability (MTPL) business. With the results of the simulations of this model, capital requirements for this long-tailed branch can be made. Important elements for this line of business are typically the number of new claims, the severity and the development of these claims. Also the inflation (containing social and super-inflation) can have a considerable impact but is not yet covered in the model we present.

The data on which the model is based, represent indexed claims of the French (MTPL) market which are retrieved by Secura. Since we work with reinsurance data, we have to be careful because we only have data of large claims. Small claims are paid by the insurer. Data of these claims are not representative to model large claims. Therefore only the claims above a certain threshold, called Amin, are considered. For the French market Amin is about 2.7 million. Note that the retention of the insurer and the priority of the data are different concepts.
The model we build consists of 3 main parts. First we analyze the number of new claims (above Amin) that appear in each development year (DY). Secondly, we try to determine the size of these new claims. In the last stage, the development of all claims is analyzed.

To build the model we make use of statistical tests such as the "VanDerWaerden" and "Wilcoxon"-test. With these tests we group data with the same distribution. For these grouped data we fit a known parameterized distribution such as for example a gamma- or a Pareto distribution. To choose the distribution and its parameters we use mean excess plots and Maximum Likelihood Estimation (MLE). All these tools are shortly introduced in the first chapter.

The aim of the model is to make simulations for the claims and analyze the impact on the capital allocation requirements for this line of business. The results of the simulations are presented in the last chapter.

Chapter 1

Introduction

1.1 AY, DY and CY

In insurance we can distinguish two kind of business: short tail and long tail business. As you already may have experienced, a claim on an insurance contract is not immediately paid. There are several reasons for the delay. If someone is injured due to a car accident, it is possible that his revalidation takes several years. All that time, the amount the insurer has to pay is not known for sure. The time before the final compensation can be paid also depends on the jurisdiction. Another example of long tail development is the pollution of a ground. The pollution itself can have taken place many years before the consequences are visible. To measure the total damage it can even take some years extra. The year in which a claim occurs is called the accident year (AY). The year a claim is reported can be different from the AY (this depends on the contract). From that year on the (re)insurer will make an estimation of the cost because the claim amount itself can still change during several development years (DY). To present the claims in a clear structure one often uses a triangle with row index the AY and column index the DY.

	DY					
AY	0	1	2	$\cdots$	9	10
1994	100	110	108	$\cdots$	115	115
1995	109	130	136	$\cdots$	148	?
$\vdots$	$\vdots$	$\vdots$				$\vdots$
2003	200	215	?	$\cdots$	?	?
2004	90	?	?	$\cdots$	?	?

Table 1.1: Triangle representation of claims.

In this form claims on the same diagonal are claims of the same calender year. Therefore the lower (resp. upper) triangle represents the future (resp. past). Business (such as MTPL, health care,...) with many DY's are called long tailed. Other lines of business where the size of the claims is paid rather shortly, are referred to as short tail.

1.2 Overview Distributions

In this section we give the main features of several distributions that are used in the model.

1.2.1 Pareto

In reinsurance the Pareto distribution is used a lot because it plays an important role in extreme value theory (see also Walhin (2001)). This distribution has a fat tail and a zero-probability for X smaller than the parameter A. The density function $f_X(x)$ and the probability function $F_X(x)$ are given by

$$\begin{aligned} f_X(x) &= \frac{\alpha A^\alpha}{x^{\alpha+1}} \\ F_X(x) &= 1 - \left(\frac{A}{x}\right)^\alpha, \end{aligned}$$

where $x \geq A$, $\alpha > 0$ and $A > 0$. The loglikelihood function for n observations $x_1, \ldots, x_n$ is given by

$$l(\alpha, A) = n \ln \alpha + n\alpha \ln A - (\alpha + 1) \sum_{i=1}^{n} \ln x_i,$$

It can be shown that the maximum likelihood estimators for the parameters A and α are

$$\begin{aligned} \hat{A} &= \min_i x_i \\ \hat{\alpha} &= \frac{n}{\sum_i \ln\left(\frac{x_i}{\hat{A}}\right)}. \end{aligned}$$

1.2.2 Gamma

The gamma distribution with parameters k and θ is characterized by the following density function

$$f_X(x) = x^{k-1} \frac{e^{-\frac{x}{\theta}}}{\theta^k \Gamma(k)}.$$

The loglikelihood function can be expressed as

$$(k-1) \sum_{i=1}^{n} \ln x_i - \frac{\sum_{i=1}^{n} x_i}{\theta} - nk \ln \theta - n \ln \Gamma(k).$$

The maximum likelihood estimators for the parameters can not be written in an explicit form. When necessary, we can find them by maximizing the log function in SAS. Note that this distribution is often defined with parameters $\alpha = k$ and $\beta = \frac{1}{\theta}$.

1.2.3 Exponential

If we set the parameter k of the gamma distribution equal to 1, we obtain the exponential distribution. The density and the distribution function are then given by

$$\begin{aligned} f_X(x) &= \lambda e^{-\lambda x} \\ F_X(x) &= 1 - e^{-\lambda x}. \end{aligned}$$

Because the gamma distribution is an extension with an extra parameter, the exponential fits for our data always perform worse than the gamma fits.

1.2.4 Gamma-Pareto

During the analysis of the data we stumble upon observed distributions for which neither the Pareto, nor the gamma distribution lead to a good fit. In some cases we have used a mixed distribution composed of a gamma and a Pareto. The density function of this mixed distribution is defined as

$$\begin{aligned} f_X(x) &= \begin{cases} 0 & \text{if } x < a, \\ f_X^{gamma}(x-a) & \text{if } a < x < b, \\ P(X > b-a) f_X^{Pareto}(x) & \text{if } b < x \end{cases} \\ \Leftrightarrow \quad f_X(x) &= \begin{cases} 0 & \text{if } x < a, \\ f_X(x) = (x-a)^{k-1} \frac{e^{-\frac{x-a}{\theta}}}{\theta^k \Gamma(k)} & \text{if } a < x < b, \\ P(X > b-a) \frac{\alpha A^\alpha}{x^{\alpha+1}} & \text{if } b < x \end{cases} \end{aligned}$$

where $P(X > b - a)$ is calculated as $1 - \int_a^b f_X^{gamma}(x-a)dx$.

1.3 Statistical Tests

To test whether a random variable X_1 differs significantly from another variable X_2, we can use a simple linear rank statistic defined as

$$S_j = \sum_{i=1}^{n} \delta_{ij} a(R_i), \quad j = 1, 2. \tag{1.1}$$

where R_i denotes the rank of an observation i, $a(R_i)$ is the score of the rank and δ_{ij} indicates whether observation i is an outcome of X_1 or X_2. For the Wilcoxon scores we have

$$a(R_i) = R_i.$$

The Van der Waerden scores are the normal quantiles given by

$$a(R_i) = \Phi^{-1}(\frac{R_i}{n+1}).$$

Both tests are predefined in SAS. If the test is significant, this means that we can not assume that both random variables have the same distribution. We explain it more intuitively for the Wilcoxon test. These scores equal the rank in this case. The statistic is the sum of all ranks of observations from the same random variable. If the distributions for both random variables X_1, X_2 do not differ, the ordered outcomes of X_1, X_2 should more or less be spread equally. In the most extreme case (i.e. when all outcomes of X_1 are smaller than those of X_2, we obtain the smallest test statistic for S_1 and the highest value for S_2. So the bigger the difference between S_1 and S_2, the higher the probability that the distributions differ. (To determine the significance of this difference, the p-value from the output in SAS is used.)
For the VDW test the scores are equal to the normal quantiles of the proportion of the ranks and the total number of observations. By this transformation of the ranks, extreme values have more weight compared to the Wilcoxon test. In this last test the difference between the scores of to succeeding observations was always 1 ($= R_{i+1} - R_i$). For the VDW test this difference depends on the quantiles and is given by $\Phi^{-1}(\frac{R_{i+1}}{n+1}) - \Phi^{-1}(\frac{R_i}{n+1})$. For extreme observations this difference is higher.

1.4 Mean Excess Plot

During our analysis of the data we often make use of a mean excess plot. The mean excess function of a random variable X is defined as

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

The mean excess $e_X(u)$ can be interpreted as the expected claim size for a layer above the threshold u, given that the claim size exceeds u. In Embrechts et al. (1997) the mean excess functions for some standard distributions are given. We resume shortly the functions for the distributions mentioned in section 1.2. The most simple mean excess function is a constant for the exponential distribution. For the Pareto distribution it is an increasing line and for the gamma distribution where $k > 1$ the function $e_X(u)$ is decreasing.
If we plot the mean excess function for the observed data, we can get an idea about the distribution of those data. For n ordered observations $x_1, ..., x_n$ the mean excess function is calculated as

$$e_X(x_i) = \frac{\sum_{k=i+1}^{n}(x_k - x_i)}{n - i + 1}$$

Since the mean excess for the last observations is based on few data, we often do not display it for the last 5 observations.

Chapter 2

Model Components

2.1 Number of New Claims

In reinsurance we are mainly interested in large claims. Small claim information is not relevant and therefore often only the data above a certain priority, called Amin, are considered. (For the data we use Amin is equal to 2718158.) As a consequence, it can happen that a claim size is lower than Amin during the first DY's but develops to an amount higher than Amin. The year that Amin is exceeded, the claim enters the data set. Note that all past claims are indexed to the same (quotation) year, so the same priority can be used for all data.
In this section we model the number of new claims appearing in the data set for each DY. A new claim for DY x is defined as a claim which is lower than Amin in DY $x-1$ and higher than Amin in DY x. For the special case DY 0, we consider the new claims as those above Amin. In table 2.1 you find the numbers (NB) of new claims for each DY. We will predict the number of new claims for a new quotation year (in our case it is AY 2005) by using regression. The numbers NB are based on claims of the past. Since the premium income is not the same each year, we must correct the numbers NB before starting the regression. The premium income is considered as a measure for the exposure of the risk. Note that just as the claims, all premiums are indexed. The claim index is based on the wages. The premium index is calculated as the total premium income over the number of contracts.
For the first DY (DY 0) we observe 500 new claims. For this DY we have data for all AY's (1994-2004). Thus, the first number (500) corresponds to the premium income between 1994 and 2004. The second number (295) for DY 1 is based on the AY's 1994-2003 (because the data for AY 2004 and DY 1 are still unknown, see also table 1.1). As a consequence this number corresponds to the premium income of the years between 1994 and 2003. For the same reason the third number (128) corresponds to the premium income of the period 1994-2002, ... and so on. The numbers (NBC_{2005}) we want, are those for the one year period 2005 (for each DY). The exposure of risk for this year is expressed by the expected premium income (EPI_{2005}). Hence,

$$NBC_{2005,j} = EPI_{2005} \cdot \frac{NB_j}{P_{period_j}} = EPI_{2005} \cdot \frac{\sum_{i=1994}^{2004-j} n_{ij}}{\sum_{i=1994}^{2004-j} P_i},$$

where P_i denotes the premium income for AY i and n_{ij} the number of claims in AY i and DY j. The corrected numbers NBC_{2005} can be found in the second column of table 2.1. In the next sections we fit a regression curve through the numbers NBC_{2005}.

DY	period$_j$	NB	NBC$_{2005}$	$\log NBC_{2005}$
0	1994-2004	500	50.81	3.928
1	1994-2003	295	33.11	3.500
2	1994-2002	128	16.52	2.805
3	1994-2001	74	11.14	2.410
4	1994-2000	59	10.50	2.349
5	1994-1999	38	8.10	2.092
6	1994-1998	19	4.97	1.604
7	1994-1997	20	6.68	1.899
8	1994-1996	5	2.26	0.817
9	1994-1995	7	4.86	1.581
10	1994	2	2.85	1.048

Table 2.1: Observed numbers of new claims.

2.1.1 Normal Regression

Since we observe a logarithmic evolution in NBC_{2005}, we fit a linear regression model for $\log(NBC_{2005})$ with an intercept and the covariate DY:

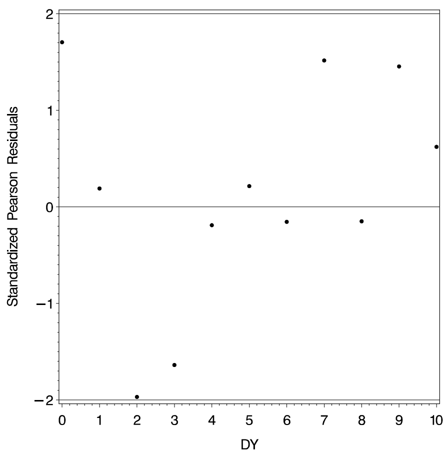
$$\log(NBC_{2005}) = \beta_0 + \beta_1 DY + \epsilon$$

where $\epsilon \sim N(0, \sigma^2)$. Both parameters β_0 and β_1 are significantly different from zero and are resp. estimated by 3.84421 and -0.36272 both with p-values > 0.0001. We find for this model an adjusted R^2 equal to 0.9189. Note that we made a weighted regression based on the original numbers NB. The motivation to do so is that we want to give more credibility to large numbers. (The first DY's are then considered to be more reliable.) In figure 2.1 we can see that the standardized residuals all lay in the confidence band $(-2, 2)$. The results of the fit are shown in the second column of table 2.2. The standard deviation of $\log(NBC_{2005})$ is printed as well. To find the predicted values for NBC_{2005} itself, we have to be careful. We can not just take the exponential of $\log(NBC_{2005})$ because of the following property of the well-known lognormal distribution:

Property 2.1.1 *If $Y \sim LN(\mu, \sigma^2)$, then $\log Y \sim N(\mu, \sigma^2)$ and*

$$\begin{aligned} E[Y] &= e^{\mu + \frac{1}{2}\sigma^2} \\ Var[Y] &= e^{2\mu + \sigma^2}(e^{\sigma^2} - 1) \end{aligned}$$

With this property we can transform our results from $\log(NBC_{2005})$ to results for NBC_{2005} (see table 2.2).



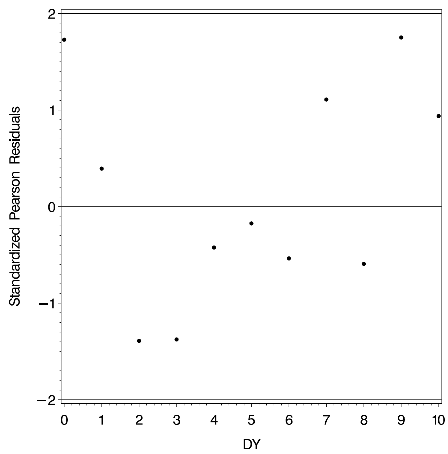


Figure 2.1: Residual plot normal regression.

Figure 2.2: Residual plot Poisson regression.

We face two main problems with this regression. We see that the confidence interval for the value of NBC_{2005} in DY 11 reaches to negative values because of the large value for $\sigma_{NBC_{2005}}$. Another disadvantage is the impossibility to extrapolate to higher DY's. We explain this briefly: due to the fact that the regression is weighted, $\sigma_{\log(NBC_{2005})}$ depends on these weights. For new values of DY (11, 12,...) we have no weights and thus no standard deviation. Although we know from property 2.1.1 that this standard deviation is needed to compute the expected value of NBC_{2005}. With the Poisson regression in the next section we try to solve these problems.

DY	$\log(\widehat{NBC_{2005}})$	$\sigma_{\log(\widehat{NBC_{2005}})}$	$\widehat{NBC_{2005}}$	$\sigma_{\widehat{NBC_{2005}}}$
0	3.844	0.123	47.08	5.84
1	3.482	0.140	32.83	4.61
2	3.119	0.200	23.08	4.66
3	2.756	0.262	16.29	4.34
4	2.393	0.298	11.45	3.49
5	2.031	0.373	8.17	3.15
6	1.668	0.518	6.062	3.36
7	1.305	0.518	4.22	2.34
8	0.943	0.982	4.16	5.30
9	0.580	0.849	2.56	2.63
10	0.217	1.538	4.05	12.59

Table 2.2: Fitted numbers of new claims (normal regression).

2.1.2 Poisson Regression

The numbers of new claims can be considered as counts. Therefore a Poisson regression is more appropriate to fit these numbers. With Poisson regression we mean that we use a gen-

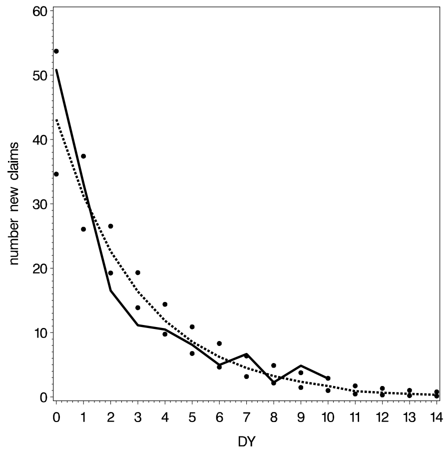


Figure 2.3: Observed (dotted line) and fitted (line) evolution with confidence interval for the mean (dots) of NBC_{2005} with Poisson regression.

eralized linear model where a Poisson distribution is assumed for the outcomes (NBC_{2005}) and a log transformation for the link function. The link function connects the linear predictor

$$\eta = \beta_0 + \beta_1 DY$$

to the outcomes NBC_{2005}. The predicted values for NBC_{2005} are then given by

$$NBC_{2005} = log^{-1}\eta = e^{\eta}.$$

The estimates for β_0 and β_1 are given by resp. 3.7641 and -0.3229. Both are strongly significant with p-values < 0.0001. The residual plot for this regression is given in figure 2.2. The predicted values and the 95% confidence bounds for this mean are printed in table 2.3. We are now able to extrapolate the results to higher DY's and the predicted values are always positive. Comparing the two residual plots we can exclude neither the normal model nor the Poisson. In figure 2.3 we plot the observed and estimated evolution of the number of new claims, as well as the confidence interval for the mean for Poisson. Note that this is not the same as the interval for the predictions NBC_{2005}. To check whether these predicted values differ not too much from the observed values one has to look at the residual plot 2.2.

2.2 New Claim Sizes

In the previous section we predicted the number of new claims. A new claim was defined as a claim that exceeds the threshold Amin from one DY to another. In this section we analyze the size of these new claims. By definition they have only values above Amin. We look for a distribution for these claim amounts to be able to simulate new claims later on. Before searching a distribution we check whether the distribution of the claims are significantly different. This is done by the Van Der Waerden and Wilcoxon score tests (see also section 1.3). In this process we want to reach a compromise between the accuracy and complexity of the model, taking into account the amount of data we dispose of.

DY	$\widehat{NBC_{2005}}$	lower	upper
0	43.13	34.62	53.72
1	31.22	26.06	37.41
2	22.61	19.26	26.54
3	16.37	13.87	19.32
4	11.85	9.75	14.40
5	8.58	6.76	10.89
6	6.21	4.64	8.31
7	4.50	3.18	6.37
8	3.26	2.17	4.90
9	2.36	1.47	3.77
10	1.71	1.00	2.91
11	1.24	0.68	2.24
12	0.89	0.46	1.73
13	0.65	0.31	1.33
14	0.47	0.21	1.03
15	0.34	0.14	0.80

Table 2.3: Fitted numbers of new claims (Poisson regression).

2.2.1 Distribution Testing

Because we expect that the sizes of new claims can be different for different DY's, we take in the tests the DY as class variable. The test will tell us whether the distributions of claim amounts in each DY class differ from each other or not. Depending on the results of these tests, we can group several DY's and treat them as one class. The decision of merging two or more DY's not only depends on the tests. For the later DY's fewer data are available and then it is hard to rely on the tests. Moreover it seems reasonable to assume that the distribution of claim amounts in the last DY's remains the same. We expect for example that a new claim in the last DY's has less probability to be high comparing to a new claim appearing in the first DY's. To keep some structure in the merging of DY classes, we only group successive DY's into one new class.
As mentioned before we consider in the first test all DY's as a different class variable and obtain the VDW scores as printed in table 2.4. We see that the score for DY 0 is strongly significant and differs a lot from the other scores. We can give an intuitive explanation about this difference. Development year 0 is not really a "development" year. It is the year in which the claims occur. The claims in the other DY's do not occur in that DY but appear above Amin in that year due to the evolution. Therefore, it is logical that there is a difference in distribution for the "occurrence of claims above Amin" and the "appearing above Amin of claims due to evolution". We therefore look for a seperate distribution for the claim amounts in DY 0.
We first continue to analyze the distribution of the sizes in the remaining DY's 1-10. The results of this test are shown in table 2.5. The p-value (equal to 0.0008) indicates that we can not assume the same distribution for the claim amounts in all these DY's. As

DY	N	Sum of Scores	Expected Under H0	Std Dev Under H0	Mean Score
0	500	115.47726	0.0	16.711622	0.230955
1	295	-13.85642	0.0	14.730323	-0.046971
2	128	-15.43895	0.0	10.611421	-0.120617
3	74	-22.84129	0.0	8.279368	-0.308666
4	59	-26.84037	0.0	7.444266	-0.454922
5	38	-23.68526	0.0	6.031690	-0.623296
6	19	-11.80306	0.0	4.301429	-0.621214
7	20	-5.63410	0.0	4.411217	-0.281705
8	5	2.04777	0.0	2.220238	0.409553
9	7	4.38838	0.0	2.624719	0.626911
10	2	-1.81396	0.0	1.406045	-0.906979

Table 2.4: Van der Waerden Scores for Variable Size, classified by Variable DY, ($p < .0001$).

DY	N	Sum of Scores	Expected Under H0	Std Dev Under H0	Mean Score
1	295	38.321165	0.0	12.568697	0.129902
2	128	7.517144	0.0	10.053021	0.058728
3	74	-9.729555	0.0	8.031582	-0.131480
4	59	-15.712536	0.0	7.264782	-0.266314
5	38	-16.549922	0.0	5.933465	-0.435524
6	19	-7.969618	0.0	4.260539	-0.419454
7	20	-2.239295	0.0	4.367739	-0.111965
8	5	3.002107	0.0	2.209838	0.600421
9	7	4.822947	0.0	2.610640	0.688992
10	2	-1.462438	0.0	1.400886	-0.731219

Table 2.5: Van der Waerden Scores for Variable Size, classified by Variable DY, ($p = .0008$).

remarked before, the last DY's contain fewer data. So there is a need to group the last DY's. Performing similar tests for DY 9-10, 8-10, 7-10,...,3-10 we find that all tests yield a non significant p-value (> 0.05). For example the last test for the DY's 3-10 yields a p-value of 0.1081. Actually this p-value is already fairly small, but we choose not to split this group further. The p-value is importantly influenced by the last DY's with few observations. Adding DY 2 leads to a p-value of 0.0143 which is significant with a 5% significance level. If we finally compare the remaining DY's 1 and 2 we find $p = 0.4361$. Summarized we have the following 3 classes of grouped DY's:

- class 0: the sizes of new claims of DY 0.
- class 1: the sizes of new claims of DY 1-2.
- class 3: the sizes of new claims of DY 3-10.

In the next section we fit for each of these classes a distribution on the claim sizes.

2.2.2 Distribution Fitting

To have an idea of the distribution, mean excess plots are calculated. We fit 3 kinds of parameterized distributions: a gamma distribution, a Pareto distribution and the combined gamma-Pareto distribution (where on the first part of the data a gamma is fitted and for the second part a Pareto is used). Note that the exponential distribution can be seen as a special case of the gamma distribution. For more details about these distributions we refer to section 1.2. Maximum likelihood estimation is used to determine the parameters. The values of the loglikelihood are presented in table 2.6. We see that the combined

class	Pareto	exponential	gamma	gamma-Pareto
0	-7650.68	-208.61	-208.21	-199.56
1	-6352.01	-86.25	-75.70	-70.70
3	-3274.80	26.67	50.81	53.11

Table 2.6: Loglikelihood values of fitted distributions.

class	Pareto		expon	gamma		gamma-Par			
	A	α	λ	k	θ	k	θ	A=b	α
0	2720040	2.50	1.79	1.05	0.53	1.01	0.57	6733187	6.52
1	2719024	3.08	2.22	0.77	0.59	0.76	0.61	5400000	3.09
3	2726542	4.21	3.06	0.60	0.54	0.66	0.43	4450000	1.57

Table 2.7: Optimal parameters.

gamma-Pareto distribution gives us the highest loglikelihood. Also in the graphs where the cumulative distribution function of the observed data is compared with the fitted distribution, we can see that the gamma-Pareto is most in line with the observed data. The point as from which observation in the data set we change the gamma to a Pareto distribution, is determined using the mean excess plots. We now analyze the distribution for each class (0, 1 and 3) more in detail.

Class 0

This class only contains the new claims in DY 0. These new claims are already from the beginning above Amin. We first have a look at the mean excess plot of these data in figure 2.4. We observe a decreasing line until 7 million. At the end we notice a crack and the line becomes increasing. According to section 1.4 the mean excess plot of a gamma distribution is a decreasing curve; it is an increasing line for the Pareto distribution and a straight line for an exponential distribution. Based on these findings we have plotted the gamma-Pareto distribution (see section 1.2.4) with break point b equal to 6733187 which corresponds to the crack in mean excess plot. We searched the size for wich the loglikelihood function was maximal and started around the crack. This optimization is done by hand. In figure 2.7 the fit with the gamma-Pareto distribution is given. In figure 2.8 the fits with other distributions are displayed. As the loglikelihood values in table 2.6 already predicted, we

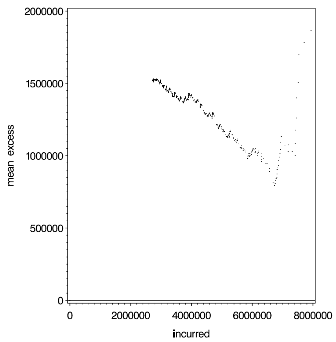


Figure 2.4: Mean excess plot for class 0.

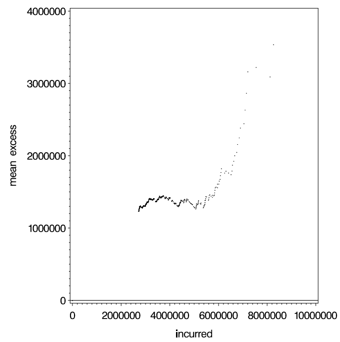


Figure 2.5: Mean excess plot for class 1.

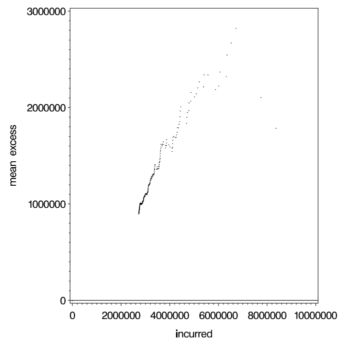


Figure 2.6: Mean excess plot for class 3.

see also from the graphs that the mixed gamma-Pareto distribution provides the best fit. The MLE parameters as denoted in section 1.2 are given in table 2.7.

Class 1

In this class we have grouped the new claim amounts with DY 1 and DY 2. Just as before we have a look at the mean excess plot in figure 2.5 to get a rough idea of the distribution. We observe a straight line on the left side of the graph and an increasing line at the end. We can then expect an exponential distribution followed by a Pareto. Since the gamma distribution is an extension of the exponential we can use again the mixed gamma-Pareto distribution to fit the data. The result can be seen in figure 2.9. The point b where the distribution changes corresponds to the crack in the mean excess plot and is 5.4 million. The values of all parameters are printed in table 2.7.

Class 3

The last DY's (DY 3-10) are put together in class 3. At first sight the mean excess plot (figure 2.6) shows an increasing line. We might think that these data are Pareto distributed. If we however have a look at table 2.6 this distribution seems not to fit the claims well. We therefore have also fitted other distributions such as the combined gamma-Pareto distribution. This distribution yields the best plot (figure 2.11) and also the best likelihood value. The explanation for this can be found in the number of parameters. With the mixed gamma-Pareto distribution we can play with 4 parameters. The Pareto fit has to do it with only 2 parameters. To compare with the gamma-Pareto fit, we display the other distribution fits in figure 2.12.

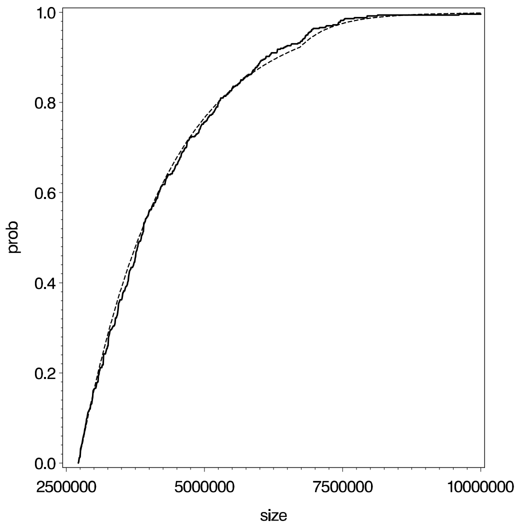


Figure 2.7: Observed and fitted gamma-Pareto (— — —) for class 0.

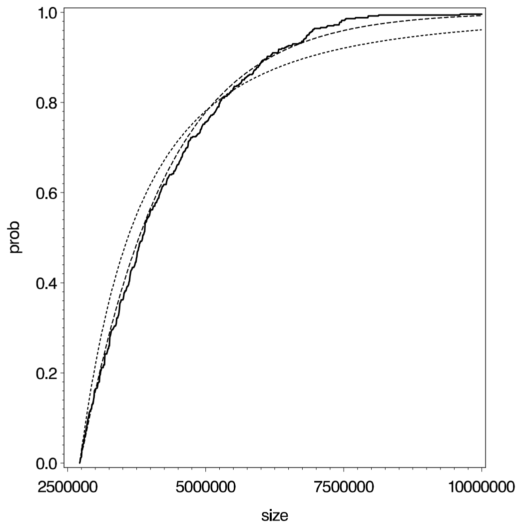


Figure 2.8: Observed and fitted Pareto (···) and gamma (— — —) for class 0.

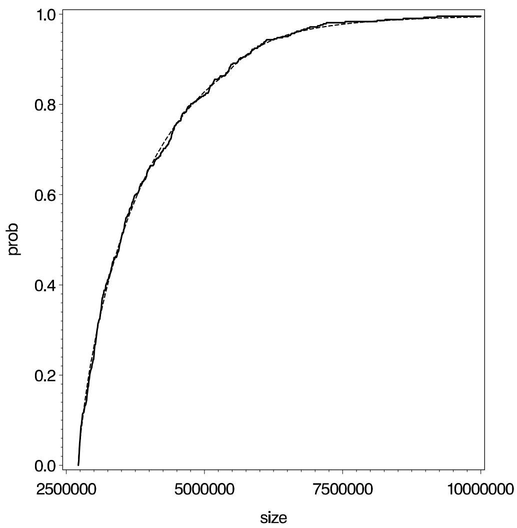


Figure 2.9: Observed and fitted gamma-Pareto (— — —) for class 1.

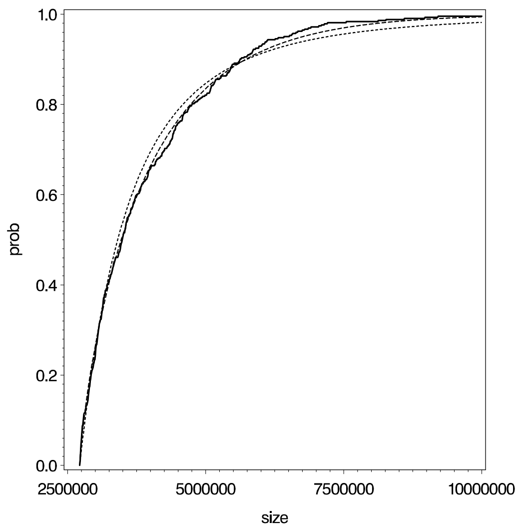


Figure 2.10: Observed and fitted Pareto (···) and gamma (— — —) for class 1.

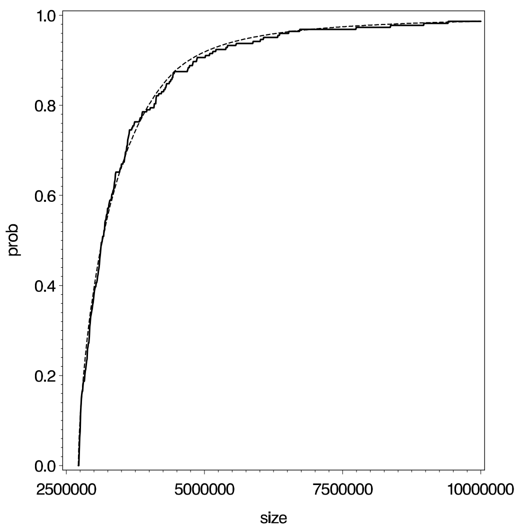


Figure 2.11: Observed and fitted gamma-Pareto (— — —) for class 3.

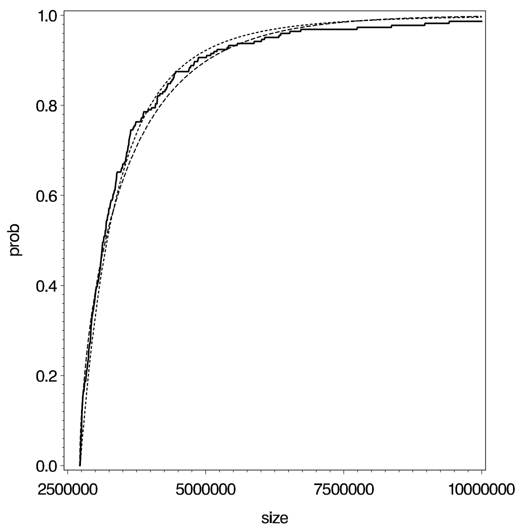


Figure 2.12: Observed and fitted Pareto $(\cdots)$ and gamma (— — —) for class 3.

2.3 Evolution of Claims

In the previous sections we showed how many new claims appear in a DY and how the size of these claims is distributed. We now investigate how the claims evolve over time. Therefore we analyze the ratio of a claim for DY x which is defined as the size of a claim in DY x divided by the size of the same claim in DY $x-1$. By definition we have no ratios for DY 0. Of course we only look at ratios of claims for which the size in DY $x-1$ is greater than Amin.

Before analyzing these ratios more in detail we transform them to the log scale. The decreasing claims, having a ratio between 0 and 1, have a negative logratio. On the other hand, positive logratios correspond to increasing claims. From figure 2.13 we see a concentration of logratios around 0. Therefore we divide the logratios into 3 main groups: positive, stable and negative logratios. A stable logratio is defined as a logratio for which the absolute value is smaller than 0.05. This group represents all the claims that have not changed more than ± 5%. A positive (negative) logratio has a logratio greater (smaller) than 0.05 (-0.05).

In the next sections we have a closer look at these groups. We look for a distribution of the logratios in each group. To determine whether all the logratios have the same distribution we again perform some tests. We check whether the DY or the size of the claim has an influence on the logratio. We have to make 2 tests: once with class variable DY and once with a class variable linked to the size of the claim. We start with the investigation of the negative logratios.

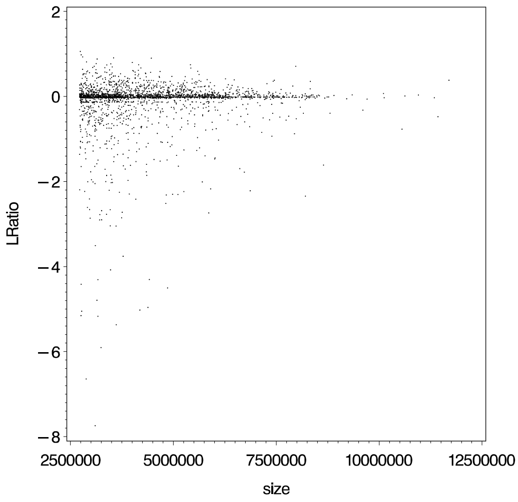


Figure 2.13: Logratios of all claims.

2.3.1 Negative Logratios

We want to check whether the distribution of these logratios changes for a different DY or a different amount of the claim. Therefore we again execute Vanderwaerden and Wilcoxon tests. First we have to create a class variable that is linked with the size of the claims. A naive way to order n claims is to give the $\frac{n}{10}$ smallest claims label 1, the second $\frac{n}{10}$ smallest claims label 2,... and so on until label 10 for the largest claims. But in this way all groups have an equal number of claims but the total claim amount is much higher for the last groups. To avoid this we make groups with an approximately equal total amount of claim sizes. Therefore, we first order the claims by size. The first claim then gets label 1, the second too,... until the total amount of claims with label 1 exceeds $\frac{S}{10}$, where S denotes the total claim size of all the claims (with a negative logratio). The following claim then receives the label 2 and the same procedure is followed. For each claim we now create the variable sizeG which takes the value equal to the label.
We are now able to analyze the impact of the claim size by making the VDW test with class variable sizeG. From table 2.8 we conclude that the claim size has no significant influence on the distribution of the negative logratios. In table 2.9, we give the results of the VDW test, but now with class variable DY. Since the p-value is smaller than 0.05 we conclude that the distribution of the logratios is not equal for all DY's. We see that the sum of score for DY 0 is the most different from 0. Removing this DY from the data results in a p-value of 0.3111. This result makes us believe that the we can use one distribution for the logratios for the DY's from 2 to 10.
We thus have two groups for the negative logratios: a group with DY 1 and a group with DY 2-10. In the following sections we search a distribution for these groups. To make the calculations easier we analyze the absolute values of the logratios. Therefore only positive values appear in the plots.

sizeG	N	Sum of Scores	Expected Under H0	Std Dev Under H0	Mean Score
1	86	1.707301	0.0	8.488778	0.019852
2	80	-10.902694	0.0	8.236577	-0.136284
3	73	0.449487	0.0	7.922537	0.006157
4	68	-7.874117	0.0	7.683796	-0.115796
5	62	3.489306	0.0	7.379594	0.056279
6	56	2.697291	0.0	7.053702	0.048166
7	49	-11.358874	0.0	6.641810	-0.231814
8	43	8.749050	0.0	6.256751	0.203466
9	38	3.857375	0.0	5.908916	0.101510
10	28	9.185875	0.0	5.118507	0.328067

Table 2.8: VDW Scores for NEG logratio, classified by Variable sizeG (p= 0.2923).

DY	N	Sum of Scores	Expected Under H0	Std Dev Under H0	Mean Score
1	143	-47.012956	0.0	10.299405	-0.328762
2	129	6.978081	0.0	9.936663	0.054094
3	101	22.069401	0.0	9.059453	0.218509
4	61	-9.525428	0.0	7.326861	-0.156155
5	57	4.839757	0.0	7.109647	0.084908
6	43	13.937684	0.0	6.256751	0.324132
7	22	4.475116	0.0	4.561530	0.203414
8	14	1.038820	0.0	3.664696	0.074201
9	11	1.624861	0.0	3.256958	0.147715
10	2	1.574665	0.0	1.399654	0.787332

Table 2.9: VDW Scores for NEG logratio, classified by Variable DY (p= 0.0004).

Negative, DY 1

The mean excess plot for this class (see figure 2.14) is not very clear. Because we notice a straight line in the first part of the graph and an increasing trend at the end, the mixed gamma-Pareto distribution can be used. We also tried the other distributions from section 1.2, but their values for the loglikelihood function were all lower than the value corresponding to the gamma-Pareto distribution. The optimal parameters are given by $k = 0.76, \theta = 1.38, A = 2.02, \alpha = 2.89$. The result of this fit is shown in figure 2.15.

Negative, DY 2-10

Before fitting a distribution we again have a look at the mean excess plot in figure 2.16. We there see an increasing line. This points to a Pareto distribution. Nevertheless the combined gamma-Pareto performs better. In the latter case we let the Pareto part start around 0.5 so we can almost say that it is a Pareto. The loglikelihood value equals -200.66 while for the pure Pareto it is equal to -288.90. In figure 2.17 the observed and fitted

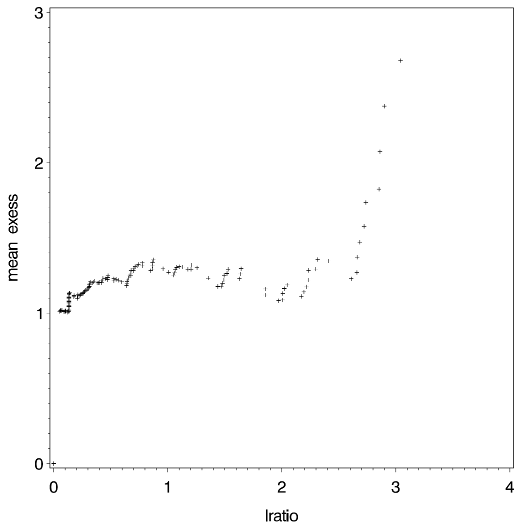


Figure 2.14: Mean excess plot of negative logratios (DY 1).

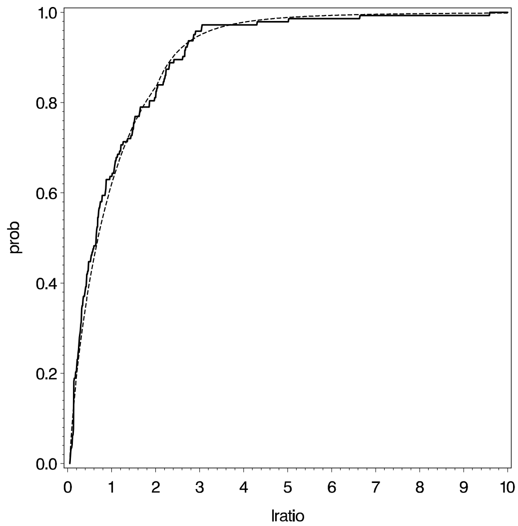


Figure 2.15: Observed and fitted (− − −) distribution of negative logratios (DY 1).

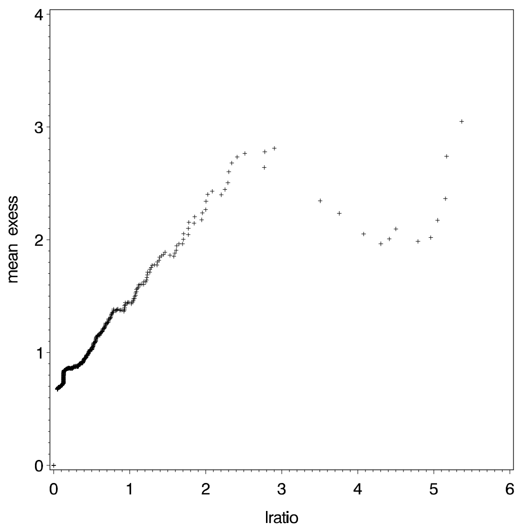


Figure 2.16: Mean excess plot of negative logratios (DY 2-10).

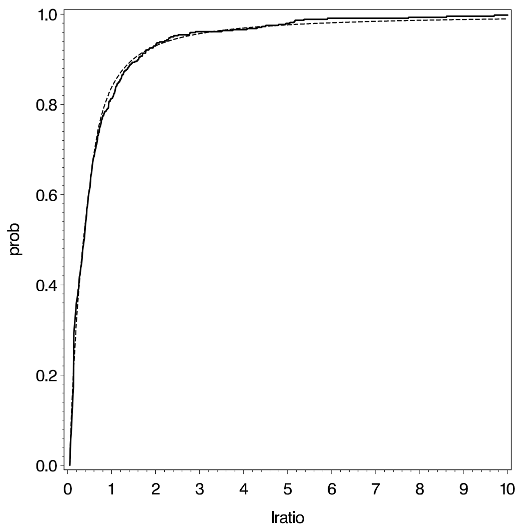


Figure 2.17: Observed and fitted (− − −) distribution of negative logratios (DY 2-10).

cumulative distribution are plotted. The MLE parameters are given by $k = 0.77, \theta = 0.65, A = 0.52, \alpha = 1.17$.

2.3.2 Positive Logratios

We follow the same procedure as for the negative logratios: first we perform tests to distinguish different groups of logratios for different sizes or DY's and then we search an appropriate distribution for each group.
First we check whether there is a difference according to the variable DY (this was the case for the negative logratios). The Vanderwaerden test gives us an insignificant global p-value of 0.2501. In table 2.10 we observe that all scores are laying close to zero.
In a second test for the positive logratios we analyze the effect of the claim size on the logratio. Just as we did for the negative logratios we divide the claims in 10 classes with an equal total amount of claims in each class. Again the created class variable is called sizeG. From the scores in table 2.11 and the associated p-value of 0.0002, we conclude that the distribution of logratios is not the same for all these sizes. We have to split up the logratios in several groups. Every group can only contain successive classes of sizeG and the p-value of a test on the group with class variable sizeG must be higher than 0.05. After several tests we have chosen the following 3 groups.

DY	N	Sum of Scores	Expected Under H0	Std Dev Under H0	Mean Score
1	87	4.736154	0.0	8.510944	0.054439
2	126	5.656416	0.0	9.820221	0.044892
3	113	-10.976351	0.0	9.434999	-0.097136
4	76	9.359394	0.0	8.044792	0.123150
5	70	-1.087122	0.0	7.767451	-0.015530
6	42	-9.147394	0.0	6.182813	-0.217795
7	29	8.882773	0.0	5.200465	0.306303
8	12	-0.346338	0.0	3.397444	-0.028862
9	10	-5.841802	0.0	3.106981	-0.584180
10	5	-1.235731	0.0	2.206753	-0.247146

Table 2.10: VDW Scores for positive logratio, classified by Variable DY (p= 0.2501).

- Small: the variable "sizeG" $\in \{1, 2, 3, 4, 5\}$
- Medium: the variable "sizeG" $\in \{6, 7, 8\}$
- Large: the variable "sizeG" $\in \{9, 10\}$

For these groups we have p-values of resp. 0.8580, 0.2294 and 0.8629. In the next subsections we try to find distributions for the positive logratios for the groups Small, Medium and Large. The procedures remain the same: we first have a look at the mean excess plot to have an idea of the distribution. We estimate the parameters for some distributions and compare the values of the maximum loglikelihood function. Finally we choose a distribution.

sizeG	N	Sum of Scores	Expected Under H0	Std Dev Under H0	Mean Score
1	86	9.261923	0.0	8.470644	0.107697
2	76	7.055532	0.0	8.044792	0.092836
3	70	9.594875	0.0	7.767451	0.137070
4	62	16.874933	0.0	7.368384	0.272176
5	58	12.531376	0.0	7.154734	0.216058
6	52	-12.931624	0.0	6.814140	-0.248685
7	48	-8.581122	0.0	6.572043	-0.178773
8	45	1.333144	0.0	6.381613	0.029625
9	39	-19.101619	0.0	5.974810	-0.489785
10	34	-16.037419	0.0	5.604881	-0.471689

Table 2.11: VDW Scores for positive logratio, classified by Variable DY (p= 0.0002).

Positive Small Claims

The mean excess plot in figure 2.18 shows a decreasing line. According to the findings in section 1.4 an Pareto or gamma-Pareto is not appropriate (In all the previous plots the gamma-Pareto was taken). This time we try a gamma distribution to fit the logratios best. Via maximum likelihood optimization we find the optimal parameters $k = 1.05$ and $\theta = 0.21$. These parameters lead to the fitted curve in figure 2.19.

Positive Medium Claims

Also for this group we see more or less a decreasing trend of the points in the mean excess plot (see figure 2.20). Although the trend is less clear than for the Small group, we fit a gamma distribution. The result is shown in figure 2.19. We can see that the curve lays a little below the observed cumulative distribution around logratio 0.5 and a little above around logratio 0.2. By assuming the fitted distribution, we attach more (less) probability to have an increasing claim with logratio 0.5 (0.2). This is a conservative assumption. The parameters of the fit are given by $k = 1.14$ and $\theta = 0.13$.

Positive Large Claims

The mean excess (figure 2.22) for this group is even more difficult to interpret than for the Medium group. An explanation can be found in the number of observations for the group Large. This group only contains the positive logratios for which sizeG is equal to 9 or 10. There are only 73 of these logratios. We tried all the distributions described in section 1.2 and the gamma distribution yields the highest value for the loglikelihood function (92.23). The values for the other distributions were 92.20 for the exponential, 91.52 for the gamma-Pareto and 87.04 for the Pareto. Moreover, also the other groups (Small and Medium) also have a gamma distribution. The parameters $k = 1.04$ and $\theta = 0.10$ lead to the gamma fit made in figure 2.23.

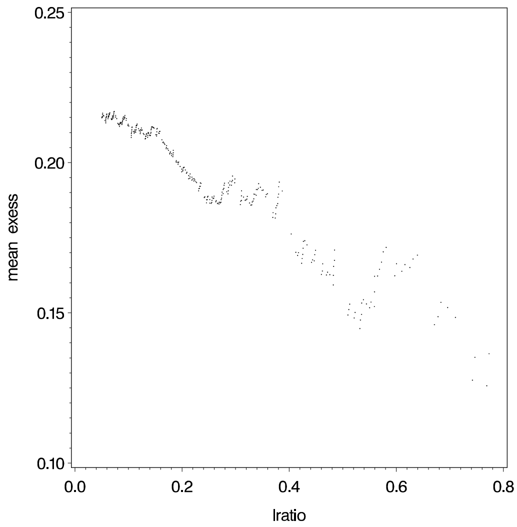


Figure 2.18: Mean excess plot of positive log-ratios (Small group).

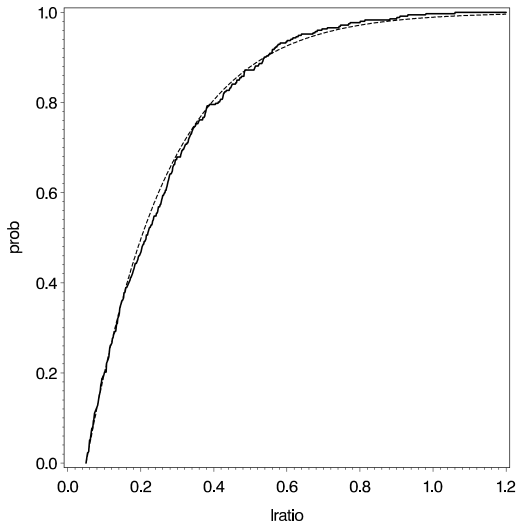


Figure 2.19: Observed and fitted $(- - -)$ distribution of positive logratios (Small).

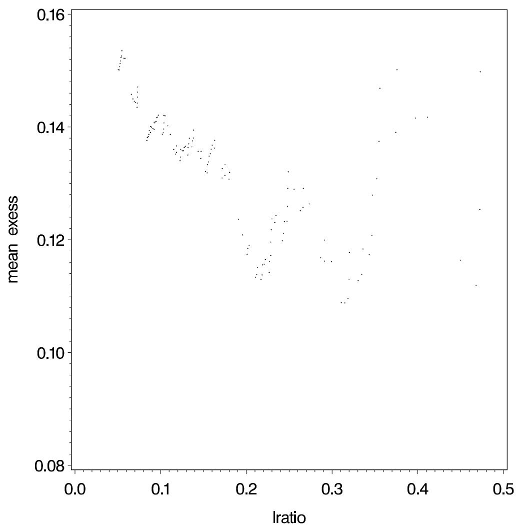


Figure 2.20: Mean excess plot of positive log-ratios (Medium group).

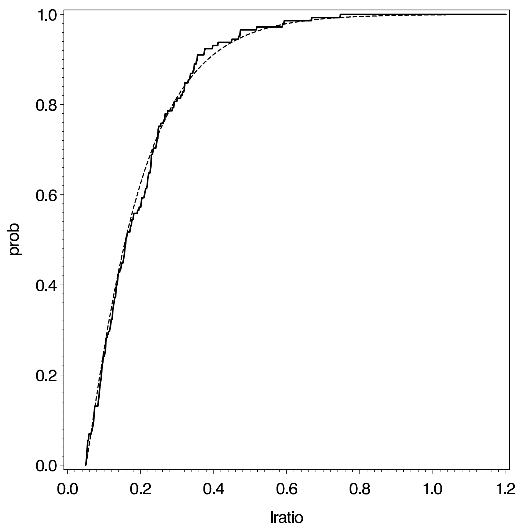
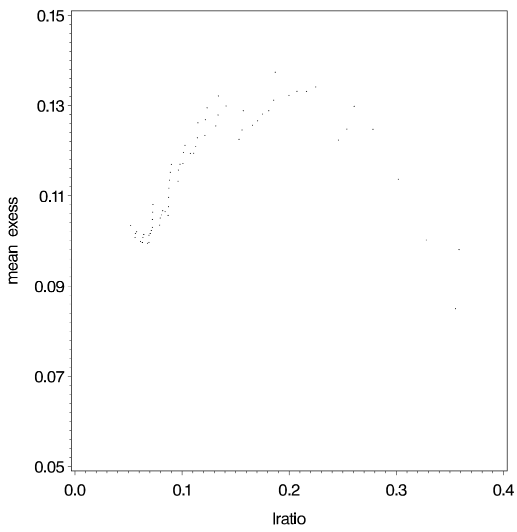


Figure 2.21: Observed and fitted $(- - -)$ distribution of positive logratios (Medium).



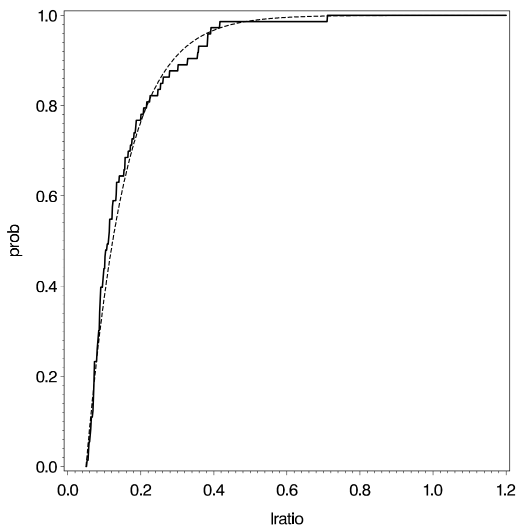


Figure 2.22: Mean excess plot of positive logratios (Large group).

Figure 2.23: Observed and fitted (− − −) distribution of positive logratios (Large).

Note that the smaller θ, the smaller the right tail of the gamma distribution. Comparing the values of θ for the different groups of positive logratios we can make the following conclusion: the larger the claims, the smaller θ, so the smaller the probability of having a highly positive logratio. This is a reassuring result for the reinsurer.

class		x	n	P(X=x)
[-0.050,	-0.045]	-0.05	14	0.007
]-0.045,	-0.035]	-0.04	46	0.022
]-0.035,	-0.025]	-0.03	115	0.056
]-0.025,	-0.015]	-0.02	155	0.075
]-0.015,	-0.005]	-0.01	894	0.433
]-0.005,	0.005]	0	278	0.135
] 0.005,	0.015]	0.01	171	0.083
] 0.015,	0.025]	0.02	201	0.099
] 0.025,	0.035]	0.03	44	0.021
] 0.035,	0.045]	0.04	71	0.034
] 0.045,	0.050]	0.05	72	0.035
			2061	1

Table 2.12: Discrete distribution for stable logratios.

2.3.3 Stable Logratios

We have already found distributions for the negative and positive logratios. For the stable logratios (i.e. the logratios between -0.05 and 0.05) we assume a discrete distribution based on the number of occurrences. The log ratios are divided in 11 classes. The probability P(X=x) that a logratio is equal to x is calculated as the number of logratios between $x - 0.005$ and $x + 0.005$ divided by the total number of stable logratios. The presentation in table 2.12 makes this more clear. The number of logratios in each class is denoted by n.

2.3.4 Probability on pos.,stable or neg. logratio

We now have distributions for all the logratios (negative, positive or stable) but we do not yet know the probability with which we need to use the different distributions. In other words we still have to determine the probability to have a negative, positive or stable logratio. In the figures 2.24 and 2.25 we have plotted these probabilities for the given data once against the DY and once against the size of the claims. Just as we have done for the claims with a negative or positive logratios, we now label all claims according to their size and so create 10 classes of claim sizes.
From figure 2.24 we see that the probability to have a stable claim increases with the DY. This is in line with our expectations. If these probabilities remain constant for all claim sizes, they can be used to make simulations. Unfortunately we see in figure 2.25 that the probabilities vary according to the size. To solve this problem we will make 3 groups of claims for which the probabilities are assumed to be constant. The first group contains the claim sizes 1, 2, 3 and 4 and is called "Little". The second group, called "Middle", contains the sizes 5, 6 and 7. The highest claim sizes 8, 9 and 10 form the group "Big". We have used the terms "Little, Middle and Big" to avoid confusion with the early used terms "Small, Medium and Large" for the positive logratio groups.
In the figures 2.26, 2.27 and 2.28 the probabilities from 2.24 are recalculated for the 3 groups separately. All values can be found in table 2.13.

		Observed probabilities										Regression	
		DY										parameters	
		1	2	3	4	5	6	7	8	9	10	β_0	β_1
Lit	+	0.37	0.27	0.22	0.17	0.17	0.17	0.12	0.05	0.14	0.1	-1.23	-0.103
	=	0.49	0.52	0.58	0.66	0.64	0.72	0.74	0.86	0.84	0.8	-0.29	0.207
	-	0.14	0.22	0.2	0.17	0.19	0.11	0.14	0.09	0.02	0.1	-0.50	-0.222
Nor	+	0.24	0.15	0.17	0.09	0.14	0.13	0.07	0.08	0.07	0	-0.66	-0.147
	=	0.48	0.56	0.56	0.7	0.64	0.69	0.76	0.9	0.76	0.8	-0.22	0.198
	-	0.28	0.29	0.27	0.21	0.21	0.18	0.16	0.03	0.17	0.2	-1.17	-0.175
Big	+	0.23	0.22	0.17	0.12	0.12	0.1	0.07	0.16	0.07	0	-1.42	-0.104
	=	0.58	0.66	0.64	0.76	0.75	0.76	0.86	0.75	0.78	1	0.22	0.172
	-	0.2	0.12	0.19	0.12	0.13	0.14	0.07	0.09	0.15	0	-1.01	-0.187

Table 2.13: Observed values and regression parameters for the probabilities.

For each of the 9 obtained curves we fit a regression. We use a generalized linear model

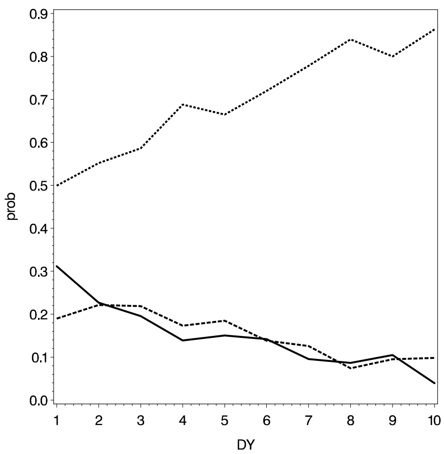


Figure 2.24: prob. on pos.(− − −), neg.(—) or stable (· · ·) logratio

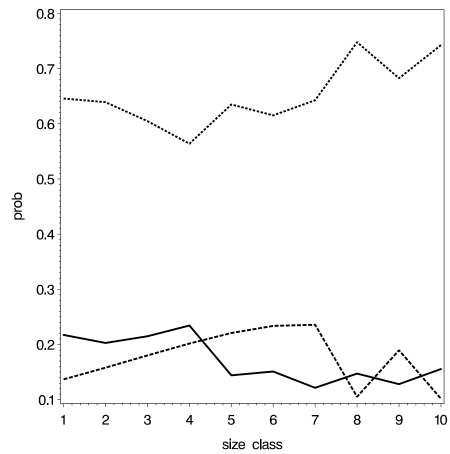


Figure 2.25: prob. on pos.(− − −), neg.(—) or stable (· · ·) logratio

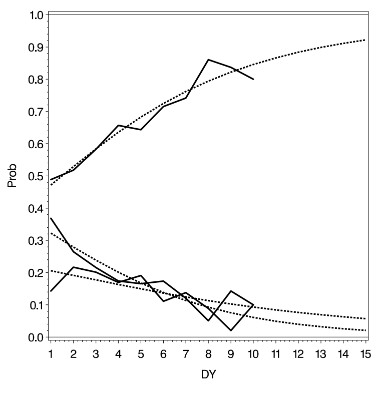


Figure 2.26: probabilities for "little" claims.

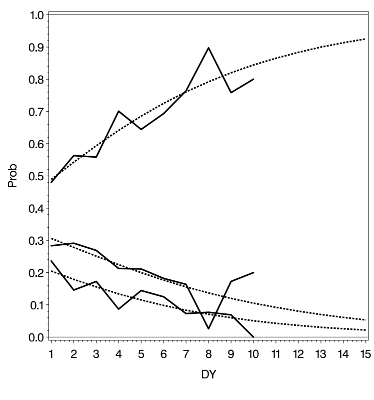


Figure 2.27: probabilities for "middle" claims.

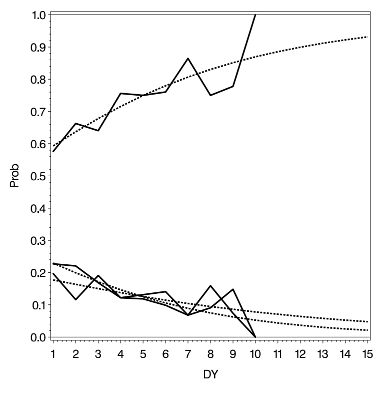


Figure 2.28: probabilities for "big" claims.

with normal errors and the logit function g as link. Thus,

$$\begin{aligned} g(p) &= \log(\frac{p}{1-p}) \\ &= \beta_0 + \beta_1 DY \end{aligned}$$

where p denotes the probability to be fitted. We choose this link function because we want that the predictions all lay between 0 and 1. We have corrected the regression by adding weights based on the number of observations in each DY. This gives a more correct view of the data because there are less observations in the last DY's. We see for all groups that the probability to have a stable logratio increases with the DY where as the other probabilities decrease with the DY. A claim has always the highest probability to remain stable.
For one DY, the sum of the probabilities to have a positive, negative and stable logratio must be 1. Since we fit the regressions for the positive, negative and stable logratios separately, this property is not fulfilled. Therefore we rescale the predictions in order to have sum 1. Since the deviation is never bigger than 2% the rescaled probabilities hardly differ from the predicted. The fits are shown in the figures 2.26, 2.27 and 2.28.

Chapter 3

Simulations

3.1 Method

In this section we explain how the simulation program is build up with the several components of chapter 2. The program can be summarized by the following pseudo code:

```
for i = 1 to "number of simulations" do
  number = [ Ps(l_0), Ps(l_1), ... ,Ps(l_15) ]
  for DY = 0 to 15
    class = 0 if DY = 0
            1 if DY = 1-2
            3 if DY = 3-15
    if number[DY] not 0 then do;
      for m = 1 to number[DY] do
        new claim size = a draw from the distrib. corresponding to class
        next = new claim size (initialization)
        if DY not 15 then do j=1 to 15-DY;
          size1 = "little" or "middle" or "big"
          state = "pos" or "stab" or "neg"
                  depending on size1 and the probabilities in table 2.13
          size2 = "small" or "medium" or "large"
          table = table with 6 distributions of logratios (determined in chapter 2)
          logratio = draw from table[1] if state = "neg" and DY = 1
                   = draw from table[2] if state = "neg" and DY not 1
                   = draw from table[3] if state = "stab"
                   = draw from table[4] if state = "pos" and size2 = "small"
                   = draw from table[5] if state = "pos" and size2 = "medium"
                   = draw from table[6] if state = "pos" and size2 = "large"
          next = next * exp(logratio);
        end;
      end;
    end;
  end;
end;
```

For each simulation we create a vector "number" where component i is an outcome from the Poisson distribution with parameter λ_i. The values of all λ_i's were estimated in section 2.1.2. We now know how many claims we have to generate for each DY. Depending on the DY, 3 kinds of distributions were distinguished for the new claim sizes. From the corresponding distribution a draw for a new claim size is taken. This claim needs to be developed. In the loop with parameter j, we search a logratio with which the claim will be multiplied. The 6 possible distributions of logratios are described in section 2.3. The choice of distribution depends on the probabilities we found in table 2.13. For example, if the new claim is appearing in DY 7, eight (j=15-7) evolutions will take place before the ultimate situation. This procedure is followed for all generated claims. In this way we have a stochastic model from which several statistics can be calculated such as the total amount of claims. Running this code for a huge number of simulations, leads to the distributions of these statistics. Interesting results are presented in the next section.

3.2 Results

Before we start the simulations, we have a look at several statistics of the distributions for the new claims and for the positive and negative logratios. In table 3.1, the mean μ, the standard deviation σ, the skewness s, the kurtosis κ, the 95% and 99% quantiles and finally the TVar at levels 95% and 99% for each distribution are presented. We had 3 distributions for the new claims, 2 for the negative logratios and 3 for the positive logratios. Note that we have Poisson distributions for the numbers and a discrete distribution presented in table 2.12 for the stable logratios. All these distributions are used in the simulations. We

	μ	σ	s	κ	Q_{95}	Q_{99}	TVaR_{95}	TVaR_{99}
class 0	4225222	1360659	1.35	5.39	7005740	8206869	7784578	9209392
class 1	3943276	1477108	5.84	136.54	6395710	8910106	8160536	11895195
class 3	3770471	14739916	277	82942	5833087	11415140	11895958	29606982
neg DY=1	-1.06	1.42	-44.12	6489	-3.02	-5.23	-4.62	-8.14
neg DY>1	-1.27	49.57	-267.19	77905	-2.65	-10.37	-16.20	-62.08
pos small	0.26	0.21	1.97	8.90	0.68	1.02	0.89	1.22
pos medium	0.20	0.14	1.88	8.42	0.48	0.70	0.61	0.83
pos large	0.15	0.10	1.97	8.96	0.36	0.52	0.46	0.62

Table 3.1: Statistics of used distributions.

make 100000 simulations. First of all we are interested in the total amount of claims for the quotation year 2005. The distribution of this total is shown in figure 3.1. In figure 3.2 we see the evolution of the claim distribution for different DY's. Because of the appearance of new claims and the evolution of old claims we expect that the distribution moves to the right for later DY's. This shift is clear between the DY's 0 and 1. For DY's 5 and 10, the difference is less big and between DY 10 and 15 we even observe a shift to the left. This can be explained by the fact that few claims appear in the last DY's and that the evolution of known claims is decreasing. In table 3.2 we see that the expected total (μ) of claims decreases after DY 8.
Note that we made the simulations for 15 years of development. Since we only have data

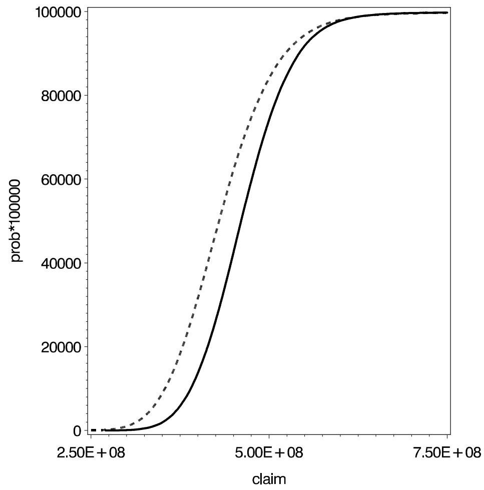


Figure 3.1: Total claim distribution for DY 10 (—) and DY 15 (– – –).

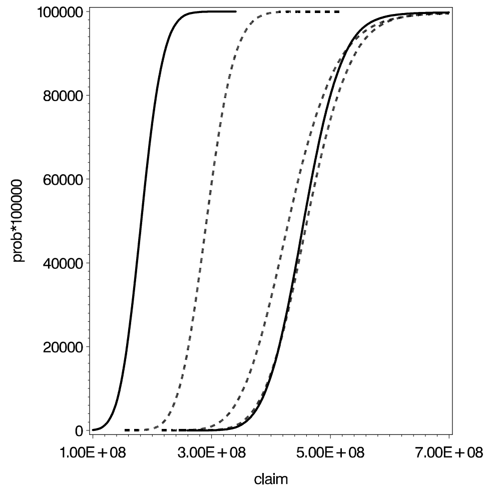


Figure 3.2: Total claim distribution for DY 0 (— left), DY 1 (– – – left), DY 5 (— right), DY 10 (– – – right) and DY 15 (– – – center).

for 10 years of development, we must be careful with the results of the last 5 years. In the simulations, we use for these years the same distributions (table 3.1 and table 2.12) as those for DY 10. Only the poisson generated number of new claims and the probabilities to have a positive, negative or stable logratio are different. Because the evolution of claims is decreasing for DY 10, this is implicitly assumed for DY 11-15. Because we had no data for these years, we must be careful with this assumption. It can be possible that the decrease should be less strong or even zero. In the last columns of table 3.2 we calculated the Value

DY	μ (million)	σ (million)	s	κ	Q_{95} (million)	Q_{99} (million)	TVaR$_{95}$ (million)	TVaR$_{99}$ (million)
0	182	29	0.18	3.05	232	254	245	266
1	292	37	0.16	3.05	356	384	373	398
2	367	43	0.16	3.07	438	470	458	488
3	412	49	1.36	17.85	491	532	523	589
4	441	54	1.61	20.00	525	574	564	648
5	458	56	1.60	18.61	547	600	590	683
6	468	58	1.57	17.67	560	617	605	702
7	471	60	1.60	17.66	567	626	614	716
8	472	62	1.61	18.02	571	630	618	723
9	469	64	1.78	22.99	571	633	620	728
10	465	65	1.68	19.61	569	634	620	730
11	460	66	1.67	18.90	567	636	619	732
12	454	68	1.63	17.45	564	636	618	733
13	448	69	1.64	17.22	561	635	617	736
14	442	71	1.64	16.63	558	637	616	739
15	436	72	1.67	16.89	556	637	615	739

Table 3.2: Characteristics of claim distribution after each DY.

at Risk (Q) and Tail Value at Risk (TVaR) at the levels 95% and 99%. From these risk measures the required capital can be determined. Note that super-inflation can have a considerable impact on the claims. This impact is not captured in this model.

Conclusion

Our goal was to build a stochastic model for the large MTPL claims of a given portfolio, based on claims from the past. The model consists of 3 main components:

- How many claims do appear each development year?
  We fitted a normal and a Poisson regression through these numbers. Finally the Poisson was selected because of the possibility to extrapolate the numbers to later development years. Moreover with the normal regression negative predictions were fitted which is impossible for counted numbers.
- What is the size of the new claims?
  During the analysis of claims, we distinguished 3 different distributions depending on the development year. For each of them a gamma-Pareto distribution was fitted.
- How do the claims evolve over time?
  We analyzed 3 kinds of evolutions: increasing, decreasing and stable claims. For the increasing and decreasing claims, subgroups according to the size or development year were made. For each group a distribution was found.

Finally, all pieces of this model were put together to make simulations. Note that the risk of super-inflation is not modelled in this work. We conclude that the total amount of claims increases a lot in the first years. From development year 8 this amount slightly decreases. From the distribution of the total claim amount risk measures were calculated from which the capital needed for this portfolio can be determined.

Bibliography

[1] M. Denault. (2001) *Coherent Allocation of Risk Capital.*, Journal of Risk, 4, 1-33.

[2] S. Demarta and A.J. McNeil. (2005) *The t copula and related copulas.* International Statistical Review, to appear 2005.

[3] J. Dhaene, et al. (2004) *Economic capital, risk measures and comonotonicity.* Belgian Actuarial Bulletin. Vol 4. No. 1.

[4] J. Dhaene, R.J.A. Laeven, S. Vanduel, G. Darkiewicz, M.J. Goovaerts (2004), *Can a coherent risk measure be too subadditive?*, to appear, 2004.

[5] J. Dhaene, A. Tsanakas, E.A. Valdez, S. Vanduffel (2005) *Optimal capital allocation principles.* Working paper, to appear.

[6] K. D'haeseleer, S. Desmedt and J.F. Walhin (2006) *Economic value added optimization for insurers using a multivariate student-t model.* Working paper 06-09, Institut des sciences actuarielles, Université Catholique de Louvain.

[7] P. Embrechts, C. Klüppelberg and T. Mikosch. (1997) Modelling Extremal Events. Springer, Berlin.

[8] P. Embrechts, et al. (2002) *Correlation and dependency in risk management: Properties and pitfalls.* In Risk Management: Value at Risk and Beyond, pp 176-223. Cambridge University Press, Cambridge.

[9] R. Kaas et al. (2004) *Modern Actuarial Risk Theory*, Kluwer Academic Publishers, Dordrecht.

[10] M. Kalkbrener (2005) *An axiomatic approach to capital allocation.* Mathematical Finance. Vol. 15, No. 3, pp 425-437.

[11] Z. Landsman and E.A. Valdez. (2003) *Tail conditional expectations for elliptical distributions.* North American Actuarial Journal, Vol. 7, oct. 01, pp 55-71

[12] Z. Landsman and E.A. Valdez. (2005) *Tail conditional variance for elliptically contoured distributions.* To appear in Belgian Actuarial Bulletin.

[13] H. Panjer. (2001) *Measurement of risk, solvency requirements and allocation of capital within financial conglomerates.* Research Report 01-15, Institute of Insurance and Pension Research, University of Waterloo.

[14] S. Vanduffel, J. Dhaene. (2006) *Some results on Denault's capital aloocation rule.* Working paper, to appear.

[15] J.F. Walhin, L. Herfurth and P. De Longueville. (2001) *The practical pricing of excess of loss treaties: actuarial, financial, economic and commercial aspects.* Belgian Actuarial Bulletin, 1:40-57.

[16] J.F. Walhin. (2005) *Value creation for insurers.* Working paper 05-07, Institut des sciences actuarielles, Université Catholique de Louvain.