LISBON SCHOOL OF ECONOMICS & MANAGEMENT
UNIVERSIDADE DE LISBOA

MASTER

ACTUARIAL SCIENCE

MASTER´S FINAL WORK

DISSERTATION

RUIN PROBABILITY AND COPULAS: APPLICATIONS IN INSURANCE PRICING

RAGNAR LEVI GUDMUNDARSON

OCTOBER - 2020

MASTER

ACTUARIAL SCIENCE

MASTER´S FINAL WORK

DISSERTATION

RUIN PROBABILITY AND COPULAS: APPLICATIONS IN INSURANCE PRICING

RAGNAR LEVI GUDMUNDARSON

SUPERVISION:

ALEXANDRA BUGALHO DE MOURA
MANUEL CIDRAES CASTRO GUERRA

OCTOBER - 2020

Acknowledgements

I would like to thank my supervisors Professor Alexandra Bugalho de Moura and Professor Manuel Guerra for their invaluable advice, guidance, tireless support and availability. This work was only possible as a result of many enlightening discussions. The many meetings we had expanded my knowledge about stochastic processes, and emanated some exciting new ideas.

I would like to thank my family and in particular my girlfriend Berglind, for providing happy distractions to rest my mind outside of my research.

Contents

List of Figures

List of Symbols

X_t	Lundberg risk process
N_t	Jump process, usually Poisson
θ	The security loading parameter
Y_i	Claim severity random variable, usually sub-scripted to indicate claim number i
$F(x, y)$	Multivariate distribution function
$F(x)$	distribution function
$\bar{F}(x, y)$	Multivariate survival function
$\bar{F}(x)$	Survival function
S_t	Compound process
$V(x), V_T(t, x)$	Infinite/finite time ruin probability
$\bar{V}^(x), \bar{V}_T(t, x)$	Infinite/finite time survival probability
$V^*(x), V_T^*(t, x)$	The value function for infinite/finite time ruin probability
$\bar{V}^*(x), \bar{V}_T^*(t, x)$	The value function for the infinite/finite time survival probability
$\mathbb{P}, \mathbb{E}$	Probability and expected value
$\mathbb{P}_x, \mathbb{E}_x$	Conditional probability and conditional expected value (starting from x)
u	Initial surplus
r	Operational cost
μ	Claim severity mean
$\beta_i, \beta_{i,j}$	Logit demand parameter i for claim process j
C	Ordinary copula
ω	The (Lévy) copula dependence parameter
ν	The Lévy measure
$\mathcal{C}$	Lévy Copula
$\perp$	Independence
$\parallel$	Dependence
p	Probability of renewal/new customers
$\mathbb{S}$	State space
Ω	Sample space
$\mathcal{F}$	Event space, σ-algebra
$\{\mathcal{F}_t\}_{t\geq 0}$	The natural filtration
$\mathcal{B}(\mathbb{R})$	The Borel σ-algebra
C^1	The class of all differentiable functions whose derivative is continuous
$\overline{\mathbb{R}}$	The extended real number line

Abstract

In this thesis ruin probability of the Lundberg risk process is used as a criterion for determining the security loading of premium. Both single and aggregated claim processes are considered. The aggregated claim process is composed of two different homogeneous claim processes. Both independent and dependent cases are considered. Lévy copulas are used to model the dependence. Lévy copulas provide an elegant and flexible manner to model dependencies and can be a useful tool when modelling dependence of jump processes with applications in insurance and risk management. The optimal value function for minimum ruin probability is analysed and stochastic control theory is used to obtain the optimal loading of the expected premium principle minimizing the probability of ruin. Numerical simulations of different case studies are presented.

1 Introduction

Insurance is based on the idea that society asks for a protection against unforeseeable events which may cause serious (financial) damage. Insurance companies offer a financial protection against these events. The general idea is to build a community where everybody contributes a certain amount and those who are exposed to the damage receive financial reimbursement, as Wüthrich [1] mentioned.

The general (non-life) insurance premium pricing strategy usually begins by finding the so-called pure premium, which is the expected value of the total claims that will occur in one time unit. However, when pricing insurance policies, insurers must take into account the risk associated with the policy as well as additional costs (e.g. operational cost, capital costs, etc.). Therefore, a so-called security loading is added to cover the risk and additional costs. The security loading is often calculated using some premium calculation principle, which is one of the main research objectives in actuarial science. The term insurance premium is obtained once the security loading has been determined and added to the pure premium. The main concerns are usually whether the loading is an appropriate measurement of the risk and which premium principle to choose. Most premium principles have some parameter that needs to be determined and affects the security loading. The higher the loading the higher the premium and consequently, the underwriting risk will be lower. However, if the premium price is too high then the exposure will be too low, and the operational cost of the insurer will engulf the premium income resulting in financial instability. Insurers must use competitive prices that may require sophisticated premium calculations in order to secure stability. One possible way to determine the loading is to use some appropriate risk measure.

In this work, the dynamics of an insurance firm is modelled with the Lundberg risk process. The term surplus process is often used as a synonym for the Lundberg risk process. The ruin probability of the Lundberg risk process is used as a risk measure which is then used to determine the loading. The ruin probability is defined as the probability that the surplus/capital of an insurance company hits zero. Although profit can be used as a criterion for the security loading, it fails to look at the risk involved. The loadings based on two strategies will be compared, one that maximizes the profit and one which minimizes the ruin probability. A major part of price optimization is understanding the demand curve and risk behaviour of collective risk groups. This thesis, however, is not concerned with finding these curves but rather examines how it is possible to determine the security loading when the risk curve is known.

When insurers create a tariff they usually divide policyholders into independent homogeneous risk groups and the collective risk model is used to price the independent risk groups. These risk groups can have different premium loadings that depend on each group's risk behaviour. The ruin probability can be used to determine the premium for the independent groups. Furthermore, the independence of risk groups might not be very realistic as claims in an insurer's risk portfolio are subject to the same event cause and hence correlated. This thesis explains how Lévy copulas can be built and used to model claim process dependencies within the context of insurance and how dependencies can affect the riskiness of the insurance portfolio. The optimal security loading will be studied both under the independence assumption and the dependence assumption.

The competitiveness of insurance companies plays a major role in premium pricing. In order to stay relevant insurance companies must quote competitive prices. If two insurance products are considered, one can ask what the optimal strategy is if the exposure of one product is more sensitive to the added loading than the other product. The strategy will depend on the optimality criterion. When the two processes are considered together, is it better to increase the loading of the less competitive product while decreasing the loading of the more competitive product, or will it be the other way around? This question will be answered numerically, where the optimality criterion is to minimize the probability of ruin. The claim processes used can either be thought of as two completely different insurance products or they can be two different risk groups within the same product, for example, two different age groups. The context of the claim processes is not very important in this thesis, but rather shows the possible applications of ruin probabilities and copulas.

The thesis is organized as follows: Section 2 gives a brief overview of related works. Section 3 presents an introduction to risk theory and then defines the probability of ruin and further develops equations for the ruin probability. Section 4 provides a brief introduction to stochastic control theory and how it can be used to calculate the optimal loading parameter. Section 5 begins by showing how independent claim processes can be combined into a single claim process. Copulas will then be used to model dependent claim processes and to combine dependent claim processes into a single claim process. Section 6 gives examples and discussion on how these concepts can be used within the insurance pricing context.

2 Literature Review

Collective risk models are fundamental in actuarial science to model the aggregate claim amount of a line of business in an insurance company. The collective risk model has two random components, the number of claims and the severity of claims and is usually modelled with a compound process [2, Chapter 3]. A compound process is a jump process where jumps represent claims and the jump sizes represent the severity of the claims [3]. The classical Lundberg risk process has been studied extensively and there exist many variations, for example with reinsurance or investments [4]. It assumes that premia come in a continuous stream while claims happen at discrete times according to a Poission distribution.

Premium pricing strategies are mainly based on the assumption that the severity of claims and the number of claims are independent and thus the corresponding means can be estimated separately. The pure premium is defined as the expected total loss and represents the risk associated with the insurance product. Another assumption is that the risk can be divided into groups of homogeneous risks. The pure premia of these individual groups are usually modelled with generalized linear models (GLM). GLM's have been applied extensively in actuarial work and a good overview is provided in [5]. After the pure premium has been determined a loading is added to cover the risk involved and additional costs, resulting in the final insurance premium [6]. The difference between the insurance premium and the pure premium is called security loading. There are three types of security loadings representing different purposes: a) a loading to cover commissions, administrative costs and claim-settlements expenses; b) a loading to cover some profit; and c) a loading for the risk taken by the insurer when underwriting the policy [7].

A premium calculation principle is a mathematical rule that assigns to each risk a real nonnegative number [2]. Bühlmann [8] described a top-down approach where the total of all risks from a given portfolio should be considered and the insurance premium should be spread over the policies in a fair way. A lot of premium principles can be used to determine the loading, for example, the expected value principle, standard deviation principle, zero utility principle, and the risk adjusted principle [2]. Here the expected value principle will be considered.

Ruin is declared when the Lundberg process hits zero. This event is called the time of ruin. This definition is a little different than the usual definition where the probability of ruin is defined as the time when the process goes below zero [4, 9, 2]. The difference will not be consequential. Ruin probability is a classical measure of risk and has been studied extensively [4, 9, 10]. Ruin probability can be defined for both infinite time horizon (usually called the probability of ultimate ruin) and finite time horizon. Although there is no absolute meaning to the probability of ruin, it still measures the stability of insurance companies. A high ruin probability indicates instability, and risk mitigation techniques should be used like reinsurance or raising premia. In this study the security loading will be selected such that the probability of ruin will be minimized. One major recent trend in risk management is the replacement of (partial) risk by sophisticated risk control. Stochastic control provides a powerful way of calculating the optimal security loading with ruin probability as a criterion. Furthermore, Stochastic control theory has been widely used to control investment, reinsurance, exposure, product design, and optimal premium and the objective of control is often the ruin probability [11, 9]. The equations derived from stochastic control theory can become very complex and numerical methods are usually needed [12, 4].

Traditional risk theory has usually assumed independence due to its convenience, but it is generally not very realistic. Claims in an insurer's risk portfolio are correlated as they are subject to the same event cause [13]. Completely homogeneous risk groups are extremely rare. Dependence among risks has become a flourishing topic in actuarial literature [14]. Dependence has mostly been measured through linear correlation coefficients [15]. The popularity of linear coefficient is mainly due to the ease with which they can be parameterized, in terms of correlation matrices. Most random variables, however, are not jointly elliptically distributed and it could be very misleading to use linear coefficients [16]. This motivated the use of concordance measures. Two random variables are concordant when large values of one go with large values of the other [17]. The Lundberg risk model is a Lévy jump process [18]. Since the claim process is a Lévy process, the dependency of two claim processes is best explained through their Lévy measure [19]. This study will not go into details about Lévy processes but both Cont and Tankov [18] and Papapantoleon [20] provide a very good introduction. Sato [21] shows that bivariate claim processes can be decomposed into independent and dependent claim processes that are statistically independent. This decomposition theorem will be exploited and frequently used in this thesis.

Most non-life insurance products have a term of one year and therefore it can be argued that the one year ruin probability should be used. The one year ruin probability is simply the probability that the surplus/capital of an insurance company will hit zero within one year. However, the appropriateness of risk measures defined over fixed time horizons can be questioned, since ruin in a given time span can be minimized by increasing the probability of ruin in the aftermath of that period. Lundberg concluded that the actual assumptions behind the classical collective risk model are in fact less restrictive when time-invariant quantities like the infinite time ruin probability are considered [10]. Therefore, although the finite time ruin probability will be explained, the infinite time ruin probability will be used in the practical applications part of this thesis.

3 Risk Theory and the Probability of Ruin

Risk theory is introduced in this section along with the tools needed to calculate the ruin probability.

3.1 Basic Risk Theory

The Collective Risk Model considers the insurance pool as a collective. The amount of claims during a particular measurement period is given by:

$$S_t = Y_0 + Y_1 + ... + Y_{N_t} = \sum_{i=0}^{N_t} Y_i, \qquad Y_0 := 0. \tag{3.1}$$

S_t is the aggregate loss random variable, N_t is the number of claims after t time units have passed, and Y_i is the amount of the i-th claim (severity). The severity probability distribution that generates Y_i represents the collective severity distribution of the pool. In this thesis it is assumed that Y_i are positive absolutely continuous (with density). The severity distribution will be denoted as $F(x)$ and the severity survival distribution will be denoted as $\bar{F}(x)$. A collective distribution can even be used for large pools consisting of many different heterogeneous types of risks. In this case, the severity distribution and/or the claims count distribution would be a mixture of the individual severity distributions/claims count distribution applicable to the underlying classes. S_t is a compound process and if N_t is a Poisson process and Y_i are i.i.d. then S_t is a compound Poisson process.

In the insurance context the Poisson process is usually parameterized the following way, $N_t \sim Poi(\lambda mt)$. The volume $m > 0$ often measures the exposure in yearly units and λ measures the frequency per exposure. In the following sections m is frequently dropped (or set $m = 1$).

If the severity random variables Y_i are i.i.d and independent of N_t then the expected value of S_t can be written as:

$$\mathbb{E}[S_t] = \mathbb{E}\Big[N_t\, \mathbb{E}[Y \mid N_t]\Big] = \mathbb{E}[N_t]\, \mathbb{E}[Y] \tag{3.2}$$

which is a very useful result as now $\mathbb{E}[Y]$ and $\mathbb{E}[N]$ can be estimated separately.

The compound Poisson distribution has the so-called aggregation property and the disjoint decomposition property. These properties are extremely useful and partly explain the popularity of the compound Poisson model. The aggregation property, which is stated in proposition 3.1, will be used later on. Both properties can be found in Wütrich [1].

Proposition 3.1 (Aggregation Property of Compound Poisson). *Suppose $S_t^{(j)}$ are independent compound Poisson processes with Poisson parameters λ_j and severity distributions $F_j(x)$, j = 1,...,n. Then $S_t = S_t^{(1)} + S_t^{(2)} + ... + S_t^{(n)}$ is a compound Poisson with parameter $\lambda = \lambda_1 + ... + \lambda_n$ and severity distribution*

$$F(x) = \sum_{j=1}^{n} \frac{\lambda_j}{\lambda} F_j(x).$$

Proof. Follows easily from the moment generating function of S_t and the severity random variable Y and the probability generating function of the Poisson process and the independence assumption. □

Insurance pricing is usually done at discrete intervals. Here it is assumed that premiums are set at fixed periods. The period length will be fixed as t = 1. Meaning that premiums are set at integer times. The expected value of S_1 (one time unit) is called the collective pure premium of S, or simply pure premium. A premium calculation principle is a function H that assigns to each risk a real nonnegative number. The pure premium is simply calculated from equation (3.2):

$$P = \mathbb{E}[N_1]\,\mathbb{E}[Y] = \lambda m\,\mathbb{E}[Y].$$

The final insurance premium is the sum of the pure premium and the security loading and depends some unknown loading parameter θ, representing the security loading.

$$P_\theta = H(\theta, S_1)$$

From the insurer point of view the final premium should always be greater than the pure premium, otherwise the insurance company will be insolvent in the long term. Usually, the pure premium is calculated a priori and can be modified later on to include the loading. In this thesis the expected value premium principle is used and is given by the following form:

$$P_\theta = (1+\theta)\,\mathbb{E}[N_1]\,\mathbb{E}[Y].$$

It is one of the most used premium principles, mainly because of its simple form. The insurance premium will be denoted as c in the following sections.

The Lundberg risk model assumes that exposure is constant in time, that losses follow a compound Poisson process and that premiums arrive at a fixed continuous rate:

$$X_t = u + ct - \sum_{i=0}^{N_t} Y_i.$$

where u is the initial surplus, c is the risk premium rate, N_t is a time homogeneous Poisson process with intensity parameter λt (or $\lambda m t$), and Y_i are i.i.d. random variables.

3.2 Probability of Ruin

Consider the development of the capital of an insurer. This is a stochastic process (sometimes called the surplus process) that increases because of earned premia and decreases when claims occur. When the capital of an insurance company hits zero, the insurance company is said to be ruined. Ruin probability can be viewed as a risk measure of a company's portfolio and can be a useful risk management tool since it expresses the solvency of an insurer in relation to the available initial capital. Mathematically, the definition is the following:

Definition 3.1 (Probability of Ruin)**.** *Let $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, \mathbb{P})$ be a filtered probability space and $X = (X_t)_{t\in[0,\infty[}$ an surplus process which is adapted and Markov with respect to the filtration. The state space is $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$. Then the probability of ruin before time T of X is the function $V_T : [0,T]\times\mathbb{R} \mapsto [0,1]$ such that*

$$V_T(t,x) = \mathbb{P}\left(\exists s \in [t,T] : X_s \leq 0 \,\middle|\, X_t = x\right), \quad x \in \mathbb{R}.$$

Furthermore, if X is time homogeneous, the infinite time ruin probability is the function $V : \mathbb{R} \mapsto [0,1]$ such that

$$V(x) = \mathbb{P}\left(\exists s \in [0,+\infty[: X_s \leq 0 \,\middle|\, X_0 = x\right), \quad x \in \mathbb{R}.$$

The infinite time horizon ruin probability $V(x)$ of a time homogeneous process is indifferent of the current time t and only depends on the current surplus, x. Therefore, it is an easier probability to calculate as it is only a single parameter function. Moreover, from definition 3.1 it should be clear that $V_T(t,x) \leq V(x)$. Sometimes it is useful to use the survival probability which is defined in the following way:

Definition 3.2 (The Survival Probability)**.** *The survival probability or the non-ruin probability is*

$$\bar{V}_T(t,x) = 1 - V_T(t,x), \quad \textit{and} \quad \bar{V}(x) = 1 - V(x)$$

for the finite and infinite time ruin probability, respectively.

The following proposition is a powerful tool when it comes to the calculation of the ruin probability.

Proposition 3.2. *Let $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, \mathbb{P})$ be a filtered probability space and $X = (X_t)_{t\in[0,\infty[}$ be a right continuous strong Markov process with state space $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$. The probability of ruin $V_T(t,x)$ satisfies the following equation:*

$$\begin{aligned} V_T(t,x) &= \mathbb{E}\left(V_T(\theta, X_\theta)\mathbb{1}_{\theta<\tau} + \mathbb{1}_{\tau\leq\theta} \,\middle|\, X_t = x\right) \\ &= \mathbb{E}_{t,x}\left(V_T(\theta, X_\theta)\mathbb{1}_{\theta<\tau} + \mathbb{1}_{\tau\leq\theta}\right). \end{aligned}$$

where $\tau = \inf\{s \in [0,\infty[: X_s \leq 0\}$, $x > 0$, θ is a stopping time and $\mathbb{1}_A$ is the indicator function of an event A.

Furthermore, if X is time homogeneous, the infinite time horizon ruin probability, $V(x)$ satisfies the following equation:

$$\begin{aligned} V(x) &= \mathbb{E}\left(V(X_\theta)\mathbb{1}_{\theta<\tau} + \mathbb{1}_{\tau\leq\theta} \,\middle|\, X_0 = x\right) \\ &= \mathbb{E}_x\left(V(X_\theta)\mathbb{1}_{\theta<\tau} + \mathbb{1}_{\tau\leq\theta}\right). \end{aligned}$$

Proposition 3.2 is a version of the dynamical programming principle and the intuition is the following: Either the ruin occurs before or at some time θ ($\mathbb{E}[\mathbb{1}_{\tau\leq\theta}] = \mathbb{P}(\tau \leq \theta)$) or the ruin does not occur within the time θ but can happen afterwards ($\mathbb{E}[V_T(\theta, X_\theta)\mathbb{1}_{\theta<\tau}]$).

The following lemma is crucial for the proof of Proposition 3.2.

Lemma 3.3. *Let $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, \mathbb{P})$ be a filtered probability space, $X = (X_t)_{t\in[0,\infty[}$ be a right continuous stochastic process with state space $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ and $\{\mathcal{F}_t\}_{t\geq 0}$ be the natural filtration. Let T be the ruin time horizon and θ be a $\mathcal{F}_t$-stopping time such that $\mathbb{P}(t \leq \theta \leq T) = 1$. Then the set $\{\forall s \in [t,\theta] \quad X_s > 0\}$ is measurable with respect to the sigma algebra $\sigma(X_{s\wedge\theta}, s \geq 0)$.*

Proof of lemma 3.3. We calculate

$$\begin{aligned} \{\forall s \in [t,\theta] : \quad X_s > 0\} &= \bigcup_{n\in\mathbb{N}} \{\forall s \in [t,\theta] : \quad X_s \geq \frac{1}{n}\} \\ &= \bigcup_{n\in\mathbb{N}} \{\forall s \in [t,\theta] \cap \mathbb{Q} : \quad X_s \geq \frac{1}{n} \wedge X_\theta \geq \frac{1}{n}\}. \end{aligned}$$

The facts that $\mathbb{Q}$ is dense in $\mathbb{R}$ and that X is right continuous were used. Therefore, $\{X_s > 0 : \forall s \in [t,\theta]\}$ is measurable with respect to the sigma algebra $\sigma(X_{s\wedge\theta}, s \geq 0)$.

□

Proof of Proposition 3.2. The probability of ruin with time horizon T can be written as:

$$\begin{aligned} V_T(t,x) &= \mathbb{P}_{t,x}\left(\exists s \in [t,\theta] : X_s \leq 0\right) + \mathbb{P}_{t,x}\left(\forall s \in [t,\theta] : X_s > 0 \wedge \exists s' \in]\theta, T] : X_{s'} \leq 0\right) \\ &= \mathbb{E}_{t,x}\left(\mathbb{1}_{\tau\leq\theta}\right) + \mathbb{E}_{t,x}\left(\mathbb{P}\left(\forall s \in [t,\theta] : X_s > 0 \wedge \exists s' \in]\theta, T] : X_{s'} \leq 0 \,\middle|\, \sigma(X_{s\wedge\theta, s\geq 0})\right)\right) \\ &= \mathbb{E}_{t,x}\left(\mathbb{1}_{\tau\leq\theta}\right) + \mathbb{E}_{t,x}\left(\mathbb{E}\left(\mathbb{1}_{\{\forall s\in[t,\theta]:X_s>0\}\wedge\{\exists s'\in]\theta,T]:X_{s'}\leq 0\}} \,\middle|\, \sigma(X_{s\wedge\theta, s\geq 0})\right)\right) \\ &= \mathbb{E}_{t,x}\left(\mathbb{1}_{\tau\leq\theta}\right) + \mathbb{E}_{t,x}\left(\mathbb{E}\left(\mathbb{1}_{\{\forall s\in[t,\theta]:X_s>0\}}\mathbb{1}_{\{\exists s'\in]\theta,T]:X_{s'}\leq 0\}} \,\middle|\, \sigma(X_{s\wedge\theta, s\geq 0})\right)\right) \\ &= \mathbb{E}_{t,x}\left(\mathbb{1}_{\tau\leq\theta}\right) + \mathbb{E}_{t,x}\left(\mathbb{1}_{\{\forall s\in[t,\theta]:X_s>0\}}\,\mathbb{E}\left(\mathbb{1}_{\{\exists s'\in]\theta,T]:X_{s'}\leq 0\}} \,\middle|\, \sigma(X_{s\wedge\theta, s\geq 0})\right)\right) \\ &= \mathbb{E}_{t,x}\left(\mathbb{1}_{\tau\leq\theta}\right) + \mathbb{E}_{t,x}\left(\mathbb{1}_{\{\forall s\in[t,\theta]:X_s>0\}}\,\mathbb{E}\left(\mathbb{1}_{\{\exists s'\in]\theta,T]:X_{s'}\leq 0\}} \,\middle|\, X_\theta = x_\theta\right)\right) \\ &= \mathbb{E}_{t,x}\left(\mathbb{1}_{\tau\leq\theta}\right) + \mathbb{E}_{t,x}\left(\mathbb{1}_{\tau>\theta}\,\mathbb{E}\left(\mathbb{1}_{\{\exists s'\in]\theta,T]:X_\theta\leq 0\}} \,\middle|\, X_\theta = x_\theta\right)\right) \\ &= \mathbb{E}_{t,x}\left(\mathbb{1}_{\tau\leq\theta} + \mathbb{1}_{\theta<\tau}V_T(\theta, X_\theta)\right) \end{aligned}$$

where Lemma 3.3 and the strong Markov property of $(X_t)_{t\in[0,\infty[}$ were used. The probability of ruin with infinite time horizon can be written as:

$$
\begin{aligned}
V(x) &= \mathbb{P}_x\left(\exists s \geq [t,\theta] : X_s \leq 0\right) + \mathbb{P}_x\left(\forall s \in [t,\theta] : X_s > 0 \wedge \exists s' > \theta : X_{s'} \leq 0\right) \\
&= \mathbb{E}_x\left(\mathbb{1}_{\tau\leq\theta}\right) + \mathbb{E}_x\left(\mathbb{E}\left(\mathbb{1}_{\{\forall s\in[t,\theta]:X_s>0\}\wedge\{\exists s'>\theta:X_{s'}\leq 0\}} \,\middle|\, \sigma(X_{s\wedge\theta,s\geq 0})\right)\right) \\
&= \mathbb{E}_x\left(\mathbb{1}_{\tau\leq\theta}\right) + \mathbb{E}_x\left(\mathbb{1}_{\{\forall s\in[t,\theta]:X_s>0\}}\,\mathbb{E}\left(\mathbb{1}_{\{\exists s'>\theta:X_{s'}\leq 0\}} \,\middle|\, X_{s'} = x_\theta\right)\right) \\
&= \mathbb{E}_x\left(\mathbb{1}_{\tau\leq\theta} + \mathbb{1}_{\theta<\tau}V(X_\theta)\right).
\end{aligned}
$$

□

3.3 Integro-differential Equations for the Ruin Probability

The following proposition, which is based on Grandell [22], shows how Proposition 3.2 can be used in practice.

Proposition 3.4. *Let*

$$X_t = u + ct - \sum_{i=0}^{N_t} Y_i$$

where Y_i are iid positive absolutely continuous random variables with $E[Y] < \infty$, N_t is a Poisson(λt) and the premium rate $c > \lambda\mathbb{E}[Y]$. If $V(X), \bar{V}(x) \in C^1(]0,\infty[)$ then the probability of ruin with infinite time horizon satisfies the following equation:

$$0 = c\frac{d}{dx}V(x) + \lambda\Big(\int_0^x V(x-y)dF(y) - V(x) + 1 - F(x)\Big) \quad x > 0 \tag{3.3}$$

with the following boundary condition:

$$\begin{cases} V(x) = 1 & x \leq 0 \\ \lim_{x\to 0^+} V(x) = \frac{\lambda}{c}\mathbb{E}[Y]. \end{cases}$$

Furthermore, the probability of non-ruin satisfies the following equation:

$$\bar{V}(x) - \bar{V}(\epsilon) = \frac{\lambda}{c}\int_\epsilon^x \bar{V}(x-y)\bar{F}(y)dy \tag{3.4}$$

for some $\epsilon > 0$ with the following boundary condition:

$$\begin{cases} \bar{V}(x) = 0 & x \leq 1 \\ \lim_{x\to 0^+} \bar{V}(x) = 1 - \frac{\lambda}{c}\mathbb{E}[Y]. \end{cases}$$

Proof of Proposition 3.4. Let $\theta = h$ and note that if $N_t = 0$ then $X_t = u + ct$. From Proposition 3.2 the infinite time ruin probability can be written as:

$$
\begin{aligned}
0 &= \mathbb{E}_x\left(V(X_h)\mathbb{1}_{h<\tau} + \mathbb{1}_{\tau\leq h} - V(x)\right) \\
&= \mathbb{E}_x\left((V(X_h) - V(x))\mathbb{1}_{h<\tau} + \mathbb{1}_{\tau\leq h}(1 - V(x))\right) \\
&= \sum_{n=0}^{\infty}\mathbb{E}_x\left((V(X_h) - V(x))\mathbb{1}_{h<\tau} + \mathbb{1}_{\tau\leq h}(1 - V(x)) \mid N_h = n\right)\mathbb{P}\left(N_h = n\right) \\
&= \mathbb{E}_x\left((V(x+ch) - V(x)) \mid N_h = 0\right)\mathbb{P}\left(N_h = 0\right) \quad + \\
&\qquad \mathbb{E}_x\left((V(X_h) - V(x))\mathbb{1}_{h<\tau} + \mathbb{1}_{\tau\leq h}(1 - V(x)) \mid N_h = 1\right)\mathbb{P}\left(N_h = 1\right) + o(h) \\
&= (V(x+ch) - V(x))e^{-\lambda h} \quad + \\
&\qquad \mathbb{E}_x\left((V(X_h) - V(x))\mathbb{1}_{h<\tau} + \mathbb{1}_{\tau\leq h}(1 - V(x)) \mid N_h = 1\right)e^{-\lambda h}\lambda h + o(h)
\end{aligned}
$$

$$\Rightarrow$$

$$
\begin{aligned}
0 = \lim_{h\to 0^+} \Bigg(& \frac{(V(x+ch)-V(x))}{h} e^{-\lambda h} \quad + \\
& \mathbb{E}_x\Big((V(X_h)-V(x))\mathbb{1}_{x+ch-y\geq 0} + \mathbb{1}_{x+ch-y<0}(1-V(x)) \mid N_h = 1\Big) e^{-\lambda h}\lambda + \frac{o(h)}{h}\Bigg) \\
= c\frac{d}{dx}V(x) & + \lambda\Big(\int_0^x V(x-y)dF(y) - V(x)\,\mathbb{P}(Y\leq x) + (1-V(x))\,\mathbb{P}(Y>x)\Big) \\
= c\frac{d}{dx}V(x) & + \lambda\Big(\int_0^x V(x-y)dF(y) - V(x)F(x) + (1-V(x))(1-F(x))\Big) \\
= c\frac{d}{dx}V(x) & + \lambda\Big(\int_0^x V(x-y)dF(y) - V(x) + 1 - F(x)\Big)
\end{aligned}
$$

$$\Leftrightarrow$$

$$0 = c\frac{d}{dx}V(x) + \lambda\Big(\int_0^x V(x-y)dF(y) - V(x) + 1 - F(x)\Big). \tag{3.5}$$

This is an integro-differential equation and by definition:

$$V(x) = 1 \quad x \leq 0$$

From the law of large numbers $\lim_{t\to\infty} X_t/t = c - \lambda\,\mathbb{E}[Y]$ with probability one. Note that if $c < \lambda\,\mathbb{E}[Y]$ then $\lim_{t\to\infty} X_t/t < 0$ with probability one. Since trajectories of X have finite variation on bounded intervals, it follows that trajectories are bounded below with probability one. Set X_0, and for any $b > 0$ let A_b be the set of trajectories of X which are bounded below by $-b$.. Then, for any monotonically increasing sequence $\{b_j\}_{j\in\mathbb{N}}$ such that $\lim_{j\to\infty} b_j = \infty$ we have $A_{b_1} \subset A_{b_2} \subset \ldots$ and $\lim_{j\to\infty}\mathbb{P}(A_{b_j}) = \mathbb{P}\left(\cup_{j=1}^{\infty} A_{b_j}\right) = 1$. Therefore,

$$
\begin{aligned}
\lim_{x\to\infty} V(x) &\leq \lim_{x\to\infty} \mathbb{P}\left(\inf_{t\geq 0} X_s < 1 \,\middle|\, X_0 = x\right) \\
&= \lim_{x\to\infty} \big(1 - \mathbb{P}(A_{x-1})\big) = 0.
\end{aligned}
$$

Now V is monotonically decreasing, hence $\lim_{x\to 0^+} V(x)$ exists. To prove the boundary condition take the non-ruin probability and $\bar{V}(x) = 1 - V(x)$, equation 3.5 becomes.

$$
\begin{aligned}
-\frac{d}{dx}V(x) &= \frac{\lambda}{c}\int_0^x \big(V(x-y)-V(x)\big)f(y)dy + \big(1-V(x)\big)\big(1-\int_0^x f(y)dy\big) \\
-\frac{d}{dx}V(x) &= \frac{\lambda}{c}(1-V(x)) + \frac{\lambda}{c}\int_0^x (V(x-y)-1)f(y)dy \\
\frac{d}{dx}\bar{V}(x) &= \frac{\lambda}{c}\bar{V}(x) + \frac{\lambda}{c}\int_0^x (1-\bar{V}(x-y)-1)f(y)dy \\
\frac{d}{dx}\bar{V}(x) &= \frac{\lambda}{c}\bar{V}(x) - \frac{\lambda}{c}\int_0^x \bar{V}(x-y)f(y)dy.
\end{aligned} \tag{3.6}
$$

Integrating with respect to x, using $d\bar{F}(y) = -f(y)dy$, integrating by parts and for some $\epsilon > 0$ the equation becomes:

$$
\begin{aligned}
\bar{V}(u) - \bar{V}(\epsilon) &= \frac{\lambda}{c}\int_\epsilon^u \bar{V}(x)dx + \frac{\lambda}{c}\int_\epsilon^u \Big([\bar{V}(x-y)\bar{F}(y)]_{y=0}^{y=x} + \int_0^x \bar{V}'(x-y)\bar{F}(y)dy\Big)dx \\
&= \frac{\lambda}{c}\Big(\int_\epsilon^u \bar{V}(x)dx + \int_\epsilon^u \bar{V}(0^+)\bar{F}(x)dx - \int_\epsilon^u \bar{V}(x)dx + \int_\epsilon^u\int_0^x \bar{V}'(x-y)\bar{F}(y)dydx\Big) \\
&= \frac{\lambda}{c}\Big(\int_\epsilon^u \bar{V}(0^+)\bar{F}(x)dx + \int_\epsilon^u\int_0^x \bar{V}'(x-y)\bar{F}(y)dydx\Big).
\end{aligned} \tag{3.7}
$$

Using Fubini's theorem on the double integral gives

$$
\begin{aligned}
\int_\epsilon^u\int_0^x \bar{V}'(x-y)\bar{F}(y)dydx &= \int_\epsilon^u\int_y^u \bar{V}'(x-y)dx\bar{F}(y)dy + \int_0^\epsilon\int_\epsilon^u \bar{V}'(x-y)dx\bar{F}(y)dy \\
&= \int_\epsilon^u \big(\bar{V}(u-y) - \bar{V}(0^+)\big)\bar{F}(y)dy + \int_0^\epsilon\int_\epsilon^u \bar{V}'(x-y)dx\bar{F}(y)dy.
\end{aligned} \tag{3.8}
$$

Plugging the results into equation (3.7) gives

$$\begin{aligned}
&= \frac{\lambda}{c}\Big(\int_{\epsilon}^{u}\bar{V}(0^+)\bar{F}(x)dx + \int_{\epsilon}^{u}\big(\bar{V}(u-y)-\bar{V}(0^+)\big)\bar{F}(y)dy + \int_{0}^{\epsilon}\int_{\epsilon}^{u}\bar{V}'(x-y)dx\bar{F}(y)dy\Big)\\
&= \frac{\lambda}{c}\Big(\int_{\epsilon}^{u}\bar{V}(u-y)\bar{F}(y)dy + \int_{0}^{\epsilon}\int_{\epsilon}^{u}\bar{V}'(x-y)dx\bar{F}(y)dy\Big).
\end{aligned} \tag{3.9}$$

Letting $\epsilon \to 0^+$ the equation becomes:

$$\bar{V}(u) - \bar{V}(0^+) = \frac{\lambda}{c}\int_{0^+}^{u}\bar{V}(u-y)\bar{F}(y)dy. \tag{3.10}$$

Letting $u \to \infty$ and noting that $\bar{V}$ is bounded the monotone convergence theorem can be used:

$$\bar{V}(\infty) = \bar{V}(0^+) + \bar{V}(\infty)\frac{\lambda}{c}\,\mathbb{E}[Y]$$

which finally results in the boundary condition:

$$1 = \bar{V}(0^+) + \frac{\lambda}{c}\,\mathbb{E}[Y] \Leftrightarrow \bar{V}(0^+) = 1 - \frac{\lambda}{c}\,\mathbb{E}[Y] \Leftrightarrow V(0^+) = \frac{\lambda}{c}\,\mathbb{E}[Y].$$

□

Remark: From proposition 3.4 it can be seen that the ruin probability of the general Lundberg risk process is independent of λ under the expected value premium principle $c = (1+\theta)\,\mathbb{E}[X_1] = (1+\theta)\lambda\,\mathbb{E}[Y]$ and the variance premium principle $c = \mathbb{E}[X_1] + \alpha Var[X] = \lambda\,\mathbb{E}[Y] + \alpha\lambda\,\mathbb{E}[Y^2]$. This is because the ruin probability itself does not depend on λ and the operational time scale can take any arbitrary value (look at the ratio $\frac{\lambda}{c}$). This means that the risk is not influenced by the volume of the portfolio reflected in λ (or λm).

Explicit formula for the infinite time ruin probability can be found when the claim distribution is exponentially distributed, using equation 3.4. Following the proof of Grandell [22] closely:

Proposition 3.5 (The probability of ruin when the claims are exponentially distributed)**.** *Let $X_t = u + ct - \sum_{i=0}^{N_t} Y_i$ where Y_i are i.i.d. exponential random variables with rate parameter μ and N_t is a Poisson(λt). Then the Probability of ruin satisfies the following equation:*

$$V(x) = \frac{\lambda}{c\mu}e^{-(\mu-\frac{\lambda}{c})x}. \tag{3.11}$$

Proof of proposition 3.5. Equation 3.6 is used, differentiate with respect to x to get:

$$\begin{aligned}
\bar{V}'(x) &= \frac{\lambda}{c}\bar{V}(x) - \frac{\lambda}{c}\int_{0}^{x}\bar{V}(x-y)f(y)dy\\
\bar{V}''(x) &= \frac{\lambda}{c}\bar{V}'(x) - \frac{\lambda\mu}{c}\frac{d}{dx}\int_{0}^{x}\bar{V}(x-y)e^{-\mu y}dy\\
\bar{V}''(x) &= \frac{\lambda}{c}\bar{V}'(x) - \frac{\lambda\mu}{c}\frac{d}{dx}\int_{0}^{x}\bar{V}(y)e^{-\mu(x-y)}dy\\
\bar{V}''(x) &= \frac{\lambda}{c}\bar{V}'(x) - \frac{\lambda\mu}{c}\Big(\bar{V}(x) - \mu\int_{0}^{x}\bar{V}(y)e^{-\mu(x-y)}dy\Big)\\
\bar{V}''(x) &= -(\mu - \frac{\lambda}{c})\bar{V}'(x).
\end{aligned}$$

Which gives:

$$\bar{V}(x) = C_1 + C_2e^{-(\mu-\frac{\lambda}{c})x}.$$

Because $\bar{V}(\infty) = 1$ and $\bar{V}(0^+) = 1 - \frac{\lambda}{c}\,\mathbb{E}[Y]$ the solution is:

$$\bar{V}(x) = 1 - \frac{\lambda\,\mathbb{E}[Y]}{c}e^{-(\mu-\frac{\lambda}{c})x}.$$

□

Proposition 3.6 (Equation for Finite Time Ruin Probability)**.** *Let*

$$X_t = u + ct - \sum_{i=0}^{N_t} Y_i$$

where Y_i are i.i.d. absolutely continuous random variables and N_t is a Poisson(λt). Then the finite time ruin probability satisfies the following equation:

$$0 = \frac{\partial}{\partial t} V_T(t,x) + c\frac{\partial}{\partial x} V_T(t,x) + \lambda\Big(\int_0^x V_T(t,x-y)dF(y) - V_T(t,x) + 1 - F(x)\Big). \tag{3.12}$$

Proof. Let $\theta = t + h$ and note that if $N_t = 0$ then $X_t = u + ct$. Proposition 3.2 is used to calculate the finite time ruin probability:

$$\begin{aligned}
0 &= \mathbb{E}_{x,t}\left(V_T(t+h, X_{t+h})\mathbb{1}_{t+h<\tau} + \mathbb{1}_{\tau\le t+h} - V_T(t,x)\right)\\
&= \mathbb{E}_{x,t}\left((V_T(t+h, X_{t+h}) - V_T(t,x))\mathbb{1}_{t+h<\tau} + \mathbb{1}_{\tau\le t+h}(1 - V_T(t,x))\right)\\
&= \sum_{n=0}^{\infty} \mathbb{E}_{x,t}\left((V_T(t+h, X_{t+h}) - V_T(t,x))\mathbb{1}_{t+h<\tau} + \mathbb{1}_{\tau\le t+h}(1 - V_T(t,x)) \;\middle|\; N_{t+h} - N_t = n\right)\mathbb{P}\left(N_{t+h} - N_t = n\right)\\
&= \mathbb{E}_{x,t}\Big((V_T(t+h, x+ch) - V_T(t,x)) \mid N_h = 0\Big)\,\mathbb{P}\left(N_h = 0\right) \quad +\\
&\quad\; \mathbb{E}_{x,t}\Big((V_T(t+h, X_{t+h}) - V_T(t,x))\mathbb{1}_{t+h<\tau} + \mathbb{1}_{\tau\le t+h}(1 - V_T(t,x)) \mid N_h = 1\Big)\,\mathbb{P}\left(N_h = 1\right) + o(h)\\
&= (V_T(t+h, x+ch) - V_T(t,x))e^{-\lambda h} \quad +\\
&\quad\; \mathbb{E}_{x,t}\Big((V_T(t+h, X_{t+h}) - V_T(t,x))\mathbb{1}_{t+h<\tau} + \mathbb{1}_{\tau\le t+h}(1 - V_T(t,x)) \mid N_h = 1\Big)e^{-\lambda h}\lambda h + o(h)
\end{aligned}$$

$$\Leftrightarrow$$

$$\begin{aligned}
0 &= \lim_{h\to 0}\Bigg(\frac{(V_T(t+h, x+ch) - V_T(t,x))}{h}e^{-\lambda h} \quad +\\
&\qquad \mathbb{E}_x\Big((V_T(t+h, X_{t+h}) - V_T(t,x))\mathbb{1}_{x+ch-y\ge 0} + \mathbb{1}_{x+ch-y<0}(1 - V_T(t,x)) \mid N_h = 1\Big)e^{-\lambda h}\lambda + o(h)\Bigg)\\
&= \frac{\partial}{\partial t}V_T(t,x) + c\frac{\partial}{\partial x}V_T(t,x) + \lambda\Big(\int_0^x V_T(t,x-y)dF(y) - V_T(t,x)\,\mathbb{P}(Y \le x) + (1 - V_T(t,x))\,\mathbb{P}(Y > x)\Big)\\
&= \frac{\partial}{\partial t}V_T(t,x) + c\frac{\partial}{\partial x}V_T(t,x) + \lambda\Big(\int_0^x V_T(t,x-y)dF(y) - V_T(t,x)F(x) + (1 - V_T(t,x))(1 - F(x))\Big)\\
&= \frac{\partial}{\partial t}V_T(t,x) + c\frac{\partial}{\partial x}V_T(t,x) + \lambda\Big(\int_0^x V_T(t,x-y)dF(y) - V_T(t,x) + 1 - F(x)\Big)
\end{aligned}$$

$$\Leftrightarrow$$

$$0 = \frac{\partial}{\partial t} V_T(t,x) + c\frac{\partial}{\partial x} V_T(t,x) + \lambda\Big(\int_0^x V_T(t,x-y)dF(y) - V_T(t,x) + 1 - F(x)\Big). \tag{3.13}$$

□

Proposition 3.6 lacks a boundary condition. It is more complicated to find one because now the probability of ruin is a function of two variables. Although, the following can be stated:

$$V_T(T,x) = \begin{cases} 1 & \text{if } x \le 0 \\ 0 & \text{if } x > 0 \end{cases}, \quad V_T(t,x) = 1 \quad x \le 0$$

but there might be a discontinuity so the partial derivative might not exist around $x = 0$ and $t = T$. One possible numerical approach consists of approximating the Lundberg process by a discrete process for which the probabilities of ruin can be explicitly computed backwards from $t = T$, as explained in Appendix A.2

3.4 Numerical Examples

It is impossible to derive $V(x)$ and $V_T(t,x)$ analytically for most cases, therefore one has to build numerical schemes to solve them. The numerical schemes used in this thesis can be seen in the appendix. To showcase the schemes let $\mathbb{E}[Y] = \lambda = 1$ and let $c = (1+\theta)\lambda\mathbb{E}[Y]$ in the following examples so that the time unit is the mean time between claims and the cash unit is the mean claim severity.

Example 3.1 (Probability of Ruin in Infinite Time)**.** *Consider the Lundberg risk process where the claims are exponentially distributed with scale parameter μ. To solve equation* (3.3) *the numerical scheme in appendix A.1 can be used or even the explicit formula in the case of exponential claims,* (3.11) *can be used. The result is the following.*

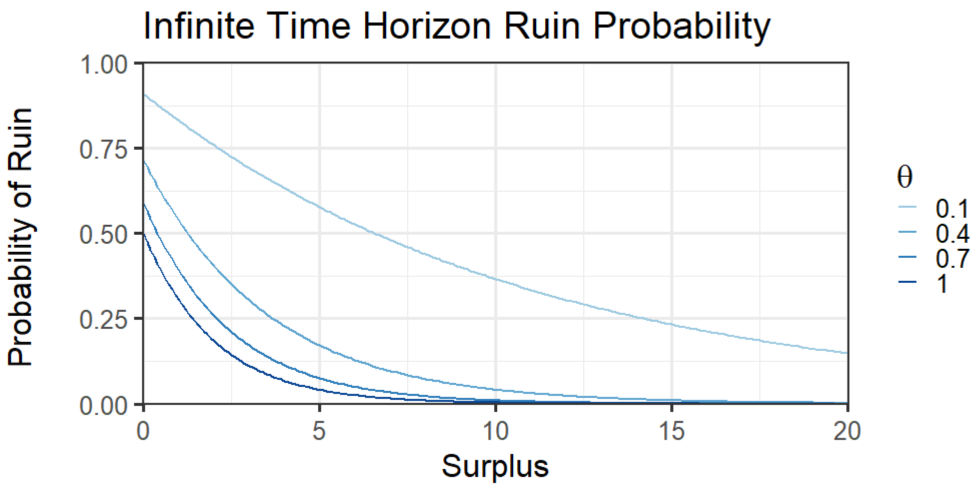

Figure 1: Infinite time horizon ruin probability of a Poisson claim process with exponential claims, $\mu = 1$, $\lambda = 1$ and different values of θ. The higher the θ the lower the probability.

Example 3.2 (Probability of Ruin in Finite Time)**.** *Consider the Lundberg risk process where the claims are exponentially distributed with scale parameter μ and a time horizon $T = 10$. To solve the ruin probability for a finite time horizon, the numerical scheme in Appendix A.2 was used.*

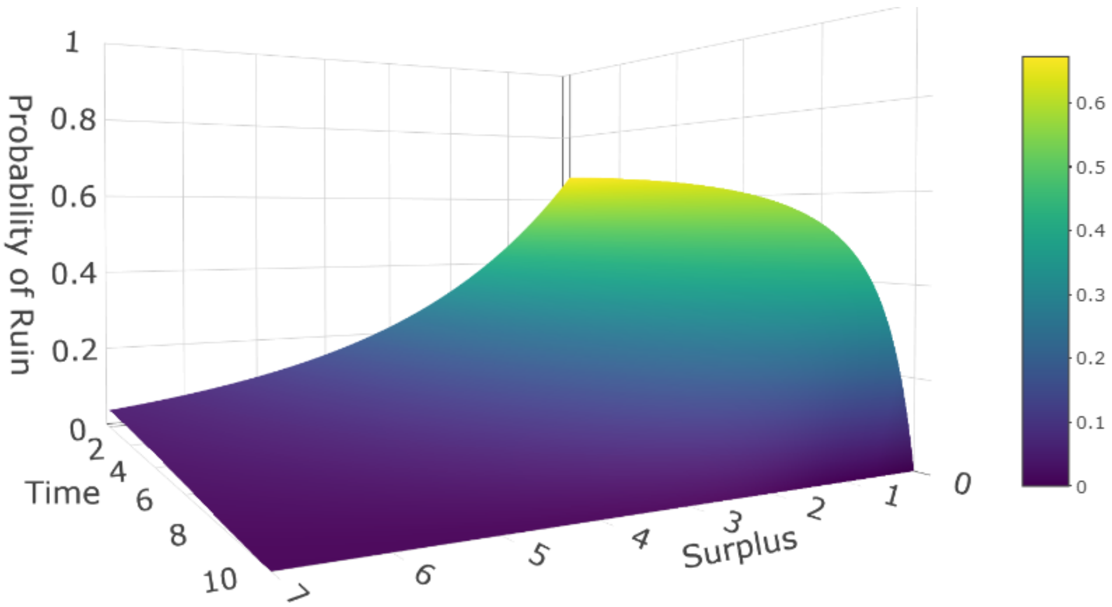

Figure 2: The figure shows the probability of ruin before time $T = 10$ as a function of the surplus, x and time, t. Note how the probability changes along the t axis showing the time inhomogeneity of the ruin probability.

The function for the ruin probability at finite time was sliced at times $t = 0, 2, 4, 6, 8$, to further compare the ruin probability in infinite time and finite time. The result can be seen in figure 3.

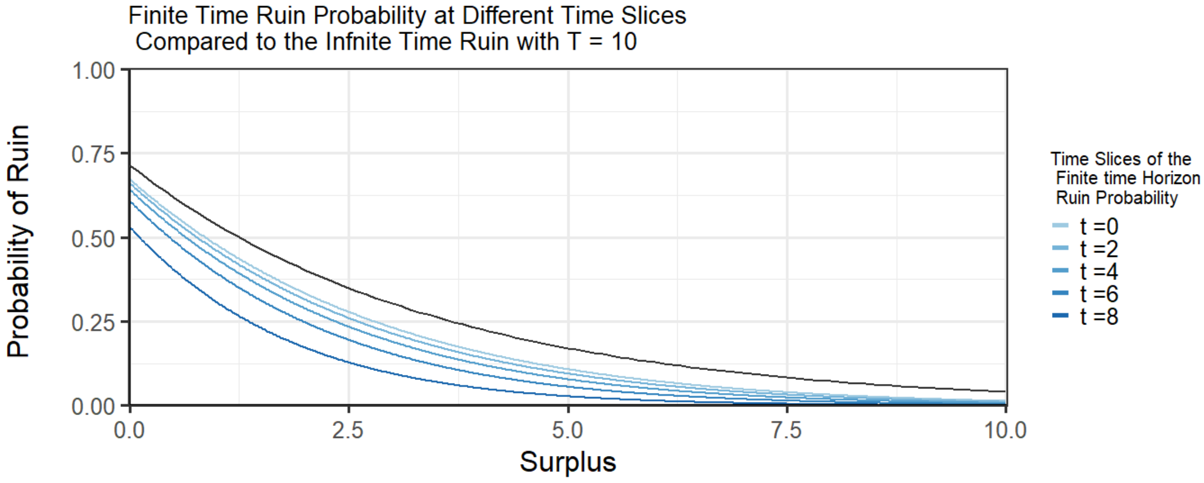

Figure 3: Finite and infinite time ruin probability comparison. The black curve shows the ruin probability in infinite time and the blue curves are the ruin probability for finite time at different times.

The ruin probability decreases with increasing time for each surplus value. Furthermore, the finite time ruin probability decreases more rapidly with increasing surplus the higher t is. Note that the ruin probability is smaller for smaller values of $T - t$ indicating that the finite time horizon is insensitive to what happens after the horizon time T. The highest ruin probability is given by the infinite time horizon case as expected. Note that the lower the t the closer it is to the infinite time horizon case. This indicates that the larger the value T, the closer the curve for $t = 0$ should be to the infinite time horizon curve.

4 Stochastic Control of Premium

Insurers are interested in controlling certain parameters to ensure smooth operation of the firm. There are many parameters that can be controlled, including investment of risky assets, reinsurance, new business, and setting premia [11]. This thesis is concerned with setting premia.

4.1 Probability of Ruin as a Function of the Premium Loading Parameter

An insurer can control the volume of his business by setting the premium c. A reasonable assumption is that the higher the premium rate c, the smaller the number of contracts in his portfolio, which means that the claim intensity (or business volume) will decrease. Therefore, $\mathbb{E}^{\theta}[N_1]$ will depend on θ as indicated by the superscript. It is reasonable to assume that $\mathbb{E}^{\infty}[N_1] = 0$ as abnormal premium rates will not attract customers. To capture these concepts let $\mathbb{E}^{\theta}[N_1] = \lambda p(\theta)$. Here λ is the average number of claims per unit of time for the whole market, and $p(\theta)$ is the probability that a potential claim is filed as an actual claim to the particular insurer under consideration. In other words, $p(\theta)$ reflects the demand or sensitivity to the loading parameter θ. $p(\theta)$ can be interpreted as probability because when an insurance firm gives insurance quotes, people will either purchase the contract or not. It is possible to study the demand of insurance contracts with a logit glm model, the demands (or renewals) are simply a Bernoulli random variable as described in Hardin and Tabari [23]. In simple terms the probability of renewal, $p(\theta)$, will be

$$p(\theta) = \frac{1}{1 + e^{\beta_0 + \beta_1 \theta}}$$

where β_0 and β_1 are determined from the glm and θ is the loading parameter. β_1 will be a positive number so $p \to 0$ when $\theta \to \infty$ and $p \to 1$ when $\theta \to -\infty$. Other methods can be used to determine the demand such as Gaussian processes see Zhang and Walton [24].

There is one issue when the minimization of the ruin probability is the goal. Namely, looking at figure 1, it is obvious that $\theta = \infty$ minimizes the ruin probability (if $x > 0$) because if $u > 0$ and $\theta \to \infty$ then $\mathbb{E}^{\theta}[N_1] \to 0$ which means ruin will not happen as there will be no incoming claims. Therefore, one needs to set some constraints or introduce new variables, such as an interest rate or cost. A cost (function) will be introduced to force the optimal premium loading to a bounded interval. This is not unrealistic as insurance products have some operation cost associated with them. The cost will be denoted as r and the expression for the net premium income will be:

$$c(\theta) = (1+\theta)\,\mathbb{E}^\theta[N_1]\,\mathbb{E}[Y] - r.$$

The surplus process is the same as in Chapter 3 but with slightly different notation to emphasize the dependency on θ:

$$X_t = u + c(\theta)t - \sum_{i=0}^{N_t} Y_i. \tag{4.1}$$

The equation for the infinite time horizon ruin probability, equation (3.3) becomes:

$$\begin{cases} \frac{d}{dx}V(x,\theta) = \frac{\mathbb{E}^\theta[N_1]}{c(\theta)}\Big(V(x,\theta) - \int_0^x V(x-y,\theta)dF(y) + F(x) - 1\Big), \\ V(0^+,\theta) = \frac{\mathbb{E}^\theta[N_1]}{c(\theta)}\,\mathbb{E}[Y]. \end{cases} \tag{4.2}$$

It is vital to understand how equation (4.2) behaves with respect to the loading parameter θ. To explore the relationship, a new parameter $\alpha = \frac{\mathbb{E}^\theta[N_1]}{c(\theta)}$ is defined. The next proposition proves that $V(x,\theta)$ is strictly increasing with respect to α.

Proposition 4.1. *If $V(x,\theta)$ satisfies equation (4.2) then $V(x,\theta)$ is strictly increasing with respect to the parameter $\alpha = \frac{\mathbb{E}^\theta[N_1]}{c(\theta)}$*

Proof of proposition 4.1. It is possible to integrate Equation (4.2) on the interval $]0,x]$ to obtain the following equation:

$$V(x,\theta) = \frac{\mathbb{E}^\theta[N_1]}{c(\theta)}\left(\mathbb{E}[Y] + \int_0^x \Big(V(z,\theta) - \int_0^z V(z-y)dF(y) + F(z) - 1\Big)dz\right). \tag{4.3}$$

To prove the proposition, we will study equations of the general form:

$$u(x) = \alpha\left(g(x) + \int_0^x \Big(u(z) - \int_0^z u(z-y)dF(y)\Big)dz\right). \tag{4.4}$$

For any function $h : [0,+\infty] \mapsto \mathbb{R}$ measurable and locally bounded, the following operator is defined:

$$\begin{aligned} (\Psi h)(x) &= \int_0^x \Big(h(z) - \int_0^z h(z-y)dF(y)\Big)dz \\ \Psi^0 h = h, &\qquad \Psi^n h = \Psi(\Psi^{n-1}h), \quad n \in \mathbb{N}. \end{aligned}$$

Notice that the transformation $h \mapsto \Psi h$ is linear.

Let, $\|h\|_{[0,x]} = \sup_{z\in[0,x]} |h(z)|$

Then:

$$\begin{aligned} |(\Psi h)(x)| &\le \int_0^x \Big(|h(z)| + \int_0^z |h(z-y)|dF(y)\Big)dz \\ &\le 2x\,\|h\|_{[0,x]}\,. \end{aligned}$$

Suppose that the following inequality holds, for some $n \in \mathbb{N}$:

$$\|(\Psi^n h)\|_{[0,x]} \le \frac{2^n x^n}{n!}\,\|h\|_{[0,x]}$$

then,

$$\begin{aligned} |(\Psi^{n+1}h)(x)| &\le \int_0^x \Big(|(\Psi^n h)(z)| + \int_0^z |(\Psi^n h)(z-y)|dF(y)\Big)dz \\ &\le \int_0^x 2\frac{2^n z^n}{n!}\,\|h\|_{[0,x]}\,dz = \frac{2^{n+1}x^{n+1}}{(n+1)!}\,\|h\|_{[0,x]}\,. \end{aligned}$$

Thus, by induction

$$\|(\Psi^n h)\|_{[0,x]} \leq \frac{2^n x^n}{n!} \|h\|_{[0,x]}.$$

Therefore, for every $x \in [0, \infty[$ fixed there is some $n \in \mathbb{N}$ such that Ψ^n is a contraction in the space of functions $h : [0, x] \mapsto \mathbb{R}$, measurable and bounded. It follows by the contraction principle that equation (4.3) has one unique solution. Further, $\lim_{n\to\infty}(\alpha^n \Psi^n)h = 0$, uniformly in $[0, x]$ for any given h and any fixed $x \in [0, +\infty[$. Let $u_{\alpha,g}$ be the solution of equation (4.3) for given g and α. Then,

$$\begin{aligned} u_{\alpha,g} &= \alpha(g + \Psi u_{\alpha,g}) = \alpha g + \alpha\Psi(\alpha(g + \Psi u_{\alpha,g})) \\ &= \alpha g + \alpha^2\Psi g + \alpha^2\Psi u_{\alpha,g} \\ &= \alpha g + \alpha^2\Psi g + \cdots + \alpha^{n+1}\Psi^n g + \alpha^{n+1}\Psi^{n+1} u_{\alpha,g}. \end{aligned}$$

Because the remainder $\lim_{n\to\infty} \alpha^n\Psi^n u_{\alpha,g}(x) = 0$ then the series can be written as:

$$u_{\alpha,g} = \sum_{n=0}^{\infty} \alpha^{n+1}\Psi^n g.$$

which converges uniformly with respect to α on compact intervals. Taking the derivative with respect to α the series becomes

$$\begin{aligned} \frac{d}{d\alpha}u_{\alpha,g}(x) &= \sum_{n=0}^{\infty}(n+1)\alpha^n(\Psi^n g)(x) \\ &= \sum_{n=0}^{\infty} \alpha^n(\Psi^n g)(x) + \sum_{n=1}^{\infty} n\alpha^n(\Psi^n g)(x) \\ &= \frac{1}{\alpha}u_{\alpha,g} + \sum_{n=1}^{\infty} \alpha^n(\Psi^n g)(x) + \sum_{n=2}^{\infty}(n-1)\alpha^n(\Psi^n g)(x) \\ &= \frac{1}{\alpha}u_{\alpha,g} + \sum_{n=0}^{\infty} \alpha^{n+1}(\Psi^{n+1} g)(x) + \sum_{n=1}^{\infty} n\alpha^{n+1}(\Psi^{n+1} g)(x) \\ &= \frac{1}{\alpha}u_{\alpha,g} + (\Psi u_{\alpha,g})(x) + \sum_{n=1}^{\infty} \alpha^{n+1}(\Psi^{n+1} g)(x) + \sum_{n=2}^{\infty}(n-1)\alpha^{n+1}(\Psi^{n+1} g)(x) \\ &= \frac{1}{\alpha}u\alpha, g + (\Psi u_{\alpha,g})(x) + \sum_{n=0}^{\infty} \alpha^{n+2}(\Psi^{n+2} g)(x) + \sum_{n=1}^{\infty} n\alpha^{n+2}(\Psi^{n+2} g)(x) \\ &= \frac{1}{\alpha}u_{\alpha,g} + (\Psi u_{\alpha,g})(x) + (\alpha\Psi^2 u_{\alpha,g})(x) + \cdots + (\alpha^{k-1}\Psi^k u_{\alpha,g})(x) + \sum_{n=1}^{\infty} n\alpha^{n+k}(\Psi^{n+k} g)(x) \\ &= \sum_{n=0}^{\infty} \alpha^{n-1}(\Psi^n u_{\alpha,g})(x) = \frac{1}{\alpha^2}u_{\alpha,u_{\alpha,g}}. \end{aligned}$$

For any $h : [0, x] \mapsto \mathbb{R}$ locally absolutely continuous function:

$$\begin{aligned} (\Psi h)(x) &= \int_0^x \Big(h(z) - \int_0^z h(z-y)dF(y)\Big)dz \\ &= \int_0^x \Big(h(z) - [h(z-y)F(y)]_{y=0}^{y=z} - \int_0^z h'(z-y)F(y)dy\Big)dz \\ &= \int_0^x (h(z) - h(0)F(z))dz - \int_0^x \int_y^x h'(z-y)F(y)dzdy \\ &= \int_0^x (h(z) - h(0)F(z))dz - \int_0^x \big(h(x-y) - h(0)\big)F(y)dy \\ &= \int_0^x h(z)dz - \int_0^x h(x-y)F(y)dy = \int_0^x h(z)(1 - F(x-z))dz. \end{aligned}$$

Thus, $h > 0$ implies $(\Psi h) > 0$, which implies $(\Psi^n h) > 0$, $\forall n \in \mathbb{N}$, and therefore $u_{\alpha,h} > 0$ for any $\alpha > 0$. This argument shows that $\frac{d}{d\alpha}V = \frac{1}{\alpha^2}u_{\alpha,V} > 0$ as $V > 0$. Therefore V is strictly increasing with α.

□

According to Proposition 4.1 if one wants to find θ minimizing the probability of ruin then it is sufficient to find θ minimizing $\frac{\mathbb{E}^\theta[N_1]}{c(\theta)}$. Another quantity of interest is the expected profit per unit time or, equivalently, the expected surplus at $t = 1$:

$$\begin{aligned}\mathbb{E}^\theta[X_1 \mid X_0 = x] &= (1+\theta)\,\mathbb{E}^\theta[N_1]\,\mathbb{E}[Y] - r - \mathbb{E}^\theta[N_1]\,\mathbb{E}[Y] \\ &= \theta\,\mathbb{E}^\theta[N_1]\,\mathbb{E}[Y] - r.\end{aligned}$$

For $p(\theta) = \frac{1}{1+e^{\beta_0+\beta_1\theta}}$, the value of θ minimizing the probability of ultimate ruin can be found with direct differentiation of α and is

$$\theta_{ruin} = \frac{1}{\beta_1}\Big(\ln\big(\frac{\lambda\,\mathbb{E}[Y]}{r\beta_1}\big) - \beta_0\Big) \tag{4.5}$$

while the value of theta maximizing the expected profit is θ_{profit}, the unique solution of the equation

$$1 + e^{\beta_0+\beta_1\theta} - \beta_1\theta e^{\beta_0+\beta_1\theta} = 0. \tag{4.6}$$

In general, θ_{ruin} does not coincide with θ_{profit}, as can be seen in the following example.

Example 4.1. *Consider process 4.1. Let Y_i be i.i.d Gamma distributed random variables with shape parameter $\alpha = 2$ and scale parameter $k = 5$. Let N_t be a Poisson process with mean $\mathbb{E}[N_t] = \lambda p(\theta)t$ with $\lambda = 80$, and $p(\theta) = \frac{1}{1+e^{-0.6+0.4\theta}}$ where θ is the loading parameter. Let the cost be $r = 0.08 * \lambda * k * \alpha$.*

Using equation (4.5) gives $\theta_{ruin} \approx 0.4349$

While equation (4.6) gives (numerically), $\theta_{profit} \approx 0.3586$. A numerical solver in the R programming language was used

With the new notation, the equation for the finite time horizon ruin probability, equation (3.12) becomes:

$$\frac{\partial}{\partial t}V_T(t,x,\theta) + c(\theta)\frac{\partial}{\partial x}V_T(t,x,\theta) + \mathbb{E}^\theta[N_1]\Big(\int_0^x V_T(t,x-y,\theta)dF(y) - V_T(t,x,\theta) + 1 - F(x)\Big) = 0. \tag{4.7}$$

It is not clear that there is one θ^* that minimizes $V_T(t,x,\theta)$ for all $x > 0$. The following example should give some hints about the problem.

Example 4.2. *Consider process 4.1. Let Y_i be i.i.d Gamma distributed random variables with shape parameter $\alpha = 2$ and scale parameter $k = 5$. Let N_t be a Poisson process with mean $\mathbb{E}[N_t] = \lambda p(\theta)t$ with $\lambda = 80$, and $p(\theta) = \frac{1}{1+e^{-0.6+0.4\theta}}$ where θ is the loading parameter. Let the cost be $r = 0.08 * \lambda * k * \alpha$.*

Consider two cases when $\theta_1 = 0.7$ and $\theta_2 = 1.2$. The following graph shows the ruin probability at $t = 0$ for $T = 1$ or $V_1(0, x, \theta_1)$.

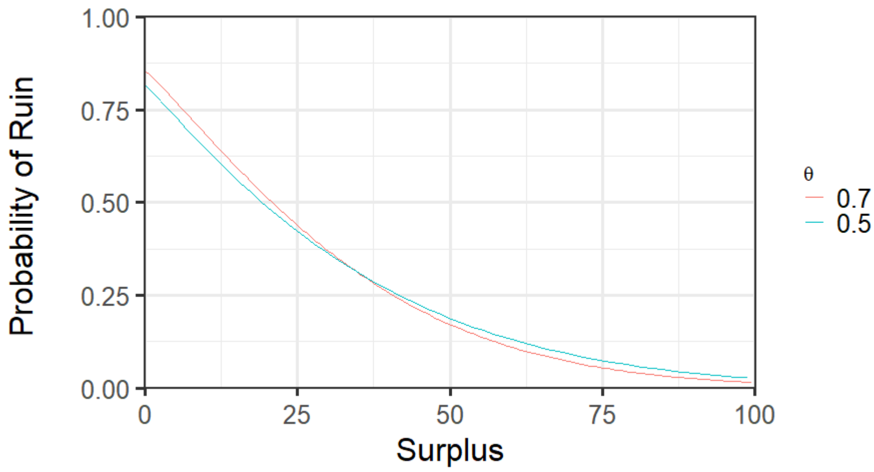

Figure 4: Finite horizon ruin probability with $T = 1$ for two different values of θ.

From Example 4.2 it can be seen that it is possible to find two values θ_1 and θ_2 such that the the ruin probability curves

$$x \mapsto V_T(0, x, \theta_i), i = 1, 2$$

intersect. This strongly indicates that $V_T(t, x, \theta)$ is not monotonic with respect to any single parameter depending on θ. In other words, θ that obtains the minimum ruin probability in finite time depends on the current surplus x and time t.

4.2 Optimal Loading with Periodic Updates

Stochastic control has been studied widely and there are multiple methods one can use depending on the nature of the problem. The dynamic programming principle, namely proposition 3.2 will be used to find the optimal loading parameter which is the parameter that minimizes the probability of ruin. The definition of a value function is needed before the model and assumptions are discussed

Definition 4.1. *Let $V_T(t, x, \theta)$ be an objective function where $t \in [0, T]$, $x \in \mathbb{S}$ and θ is the loading parameter The value function for the finite horizon ruin probability is defined as:*

$$V_T^*(t, x) = \inf_{\theta} V_T(t, x, \theta).$$

The value function for the finite horizon survival probability is defined as:

$$\bar{V}_T^*(t, x) = \sup_{\theta} \bar{V}_T(t, x, \theta).$$

The optimal loading parameter is denoted as θ^*. Furthermore, if the process is time homogeneous and the time horizon is infinite, then the value function is defined as:

$$V^*(x) = \inf_{\theta} V(x, \theta)$$
$$\bar{V}^*(x) = \sup_{\theta} \bar{V}(x, \theta)$$

for the infinite time ruin and survival probability, respectively.

Normally, premium rates are set at discrete times and the goal is to set premium rates (loading parameter) such that the ruin probability is minimized. The exposure (i.e., the number of policies) is determined according to the new loading and remains unchanged for one year (i.e., no one can enter or leave the policy for one year). For simplicity, it assumed that the process is Markov. With normalized time interval of length 1, the game is the following: Starting at t = 0, the insurer selects a loading such that the probability of ruin is minimized for a given current surplus x_0. This procedure is repeated at the beginning of each time interval.

$$\theta_0^* = \arg\min_{\theta} V_T(0, x_0, \theta).$$

The process evolves until the next time period is reached (given the insurer has not ruined). At the next decision point, t = 1, the insurer can pick a new loading parameter θ_1 given the current surplus x_1 such that the ruin probability of the surplus process is minimized.

$$\theta_1^* = \arg\min_{\theta} V_T(1, x_1, \theta).$$

This premium loading selection problem continues until $t = T - 1$ or continues infinitely often if $T = \infty$.

Before continuing it is important to realize the shortcomings of the model. Stochastic control theory usually assumes that the stochastic process of interest is Markov, roughly speaking, given the present the future is independent of the past. The surplus process of an insurance product is, however, not Markov. To see this imagine that an insurer is changing the premium loading at time $t \in \mathbb{N}$, the insurance policies bought at, say, $t - 0.5$ still affect the Poisson rate parameter meaning that the exposure is not fixed between decision times Therefore, the surplus process of a real insurance policy is not Markov as the present is dependent on the past. It is assumed that the claims are paid immediately which is not very realistic as partial payments may be made to cover ongoing losses. Lastly, the premiums are assumed to come in a continuous stream. This is perhaps not so unrealistic as in insurance accounting, incomes are not realized immediately but rather divided in equally throughout the policy term (usually monthly) and then the income is realized once the month has passed so in reality premiums come in at discrete times but the income is not realized immediately.

4.3 Stochastic Control in Finite Time Horizon

Optimizing the premium loading parameter with respect to finite time horizon is relatively straight forward when the dynamic programming principle is used. Although the game involves selecting θ_t such that the ruin probability of the surplus process is minimized from the current time t to the horizon time T, the best strategy is to start from the horizon time T and calculate recursively backward in time until the current time t is reached. The reason is that the ruin probability at $t = T$ is known, namely $V_T^*(T, x) = V_T(T, x, \theta) = 0, x > 0, \theta$ because the game is over. This can be written the following way:

For $t = T - 1$, the equation is

$$\begin{aligned} V_T^*(T-1, x) &= \inf_\theta \mathbb{E}_{T-1,x} \left[V_T^*(T, X_T) \mathbb{1}_{T<\tau(\theta)} + \mathbb{1}_{\tau(\theta)\leq T} \right] \\ &= \inf_\theta \mathbb{E}_{T-1,x} \left[\mathbb{1}_{\tau(\theta)\leq T} \right]. \end{aligned}$$

This means that equation (3.12) can be used with $T = 1$, $t = 0$ and boundary condition $V_T(T, x) = 0$, for $x > 0$. To find the value function $V_T^*(T - 1, x)$ and the associated loading parameter $\theta_{T-1}(x)$, it is possible to calculate $V_1(0, x)$ using equation (3.12) multiple times for different values of θ and pick θ that minimizes the ruin probability for each surplus x considered.

Once $V_T^*(T - 1, x)$ is known proceed to $t = T - 2$ and the equation becomes:

$$V_T^*(T-2, x) = \inf_\theta \mathbb{E}_{T-2,x} \left[V_T^*(T-1, X_{T-1}) \mathbb{1}_{T-1<\tau(\theta)} + \mathbb{1}_{\tau(\theta)\leq T-1} \right].$$

Again, equation (3.12) can be used with $T = 1$, $t = 0$ but with modified boundary condition $V_T(T, x) = V_T^*(T - 1, x)$.

To find the value function $V_T^*(T-2, x)$ the same method can be used. To put this more rigorously, the algorithm is on the next page.

: Finite Time Value Function Recursion

Let the symbol $\vec{}$ denote a (computer) vector.

Import $H\Big(t, T, \vec{x}, B, \theta\Big)$, a computer function that computes equation (3.12) with boundary condition B at $t = T$ for each $x_i \in \vec{x}$ with loading parameter θ. Note that $H\Big(t, T, \vec{x}, B, \theta\Big)$ returns a vector, where the elements give the probability of ruin for surplus $x_i \in \vec{x}$ at time t

Initialize

$\vec{\theta}$ vector of loading parameters considered

$\vec{x}$ the surplus grid with elements $x_1, ..., x_n$

$\vec{t}$ the decision times with elements $T, T-1, .., 0$

ValueFunctionList = list() a list which has the same length as $\vec{t}$ and stores the value functions and the corresponding loading parameter for each decision time t

VList = list() a list which has the same length as $\vec{\theta}$. The list will store information of ruin probabilities at time t for each θ

BoundaryCondition $\leftarrow \vec{0}$ vector for the boundary condition Loop that finds the value function V^*.

for *currentTime in* $\vec{t}$ **do**

  This for loop calculates the ruin probability for each $\theta_i \in \vec{\theta}$ at each $x_i \in \vec{x}$

  for *thetaConsidered in* $\vec{\theta}$ **do**

    VList[thetaConsidered] $\leftarrow H\Big(t = 0, T = 1, \vec{x} = \vec{x}, B = BoundaryCondition, \theta = thetaConsidered\Big)$

  end

  $\vec{V}^*$ initialize arbitrary vector which has the same length as $\vec{x}$. The vector stores the optimal ruin probability value for each $x \in \vec{x}$

  $\vec{\theta}^*$ initialize arbitrary vector which has the same length as $\vec{x}$. The vector stores the optimal loading parameter value for each $x \in \vec{x}$

  Loop thorugh $\vec{x}$ and find the lowest value of V(t,x) for each $\theta_i \in \vec{\theta}$

  for *currentX in* $\vec{x}$ **do**

    $\vec{V}^*[currentX] \leftarrow$ Minimum value $V_T(currentTime, currentX)$ in VList

    $\vec{\theta}^*[currentX] \leftarrow$ corresponding θ

  end

  Assign the value function and optimal loading parameters to the ValueFunctionList

  ValueFunctionList[currentTime] $\leftarrow \vec{V}^*, \vec{\theta}^*$

  Finally update the boundary condition

  BoundaryCondition $\leftarrow \vec{V}^*$

end

Return a list with the value function and the loading parameter for each decision time

Return ValueFunctionList

To see the algorithm in action an example is provided.

Example 4.3. *Consider the process given in* (4.1). *Let Y_i be i.i.d Gamma distributed random variables with shape parameter $\alpha = 2$ and scale parameter $k = 5$. Let N_t be a Poisson process with mean $\mathbb{E}[N_t] = \lambda p(\theta) t$ with $\lambda = 80$, and $p(\theta) = \frac{1}{1+e^{-0.6+0.4\theta}}$ where θ is the loading parameter. Let the cost be $r = 0.08 * \lambda * k * \alpha$. The value function and optimal loading parameter are found by using the finite time value function recursion. The grid of the loading parameter was an unequal spacing of points lying in the interval* $[0.38, 1.4]$. *The results can be found on the following graphs*

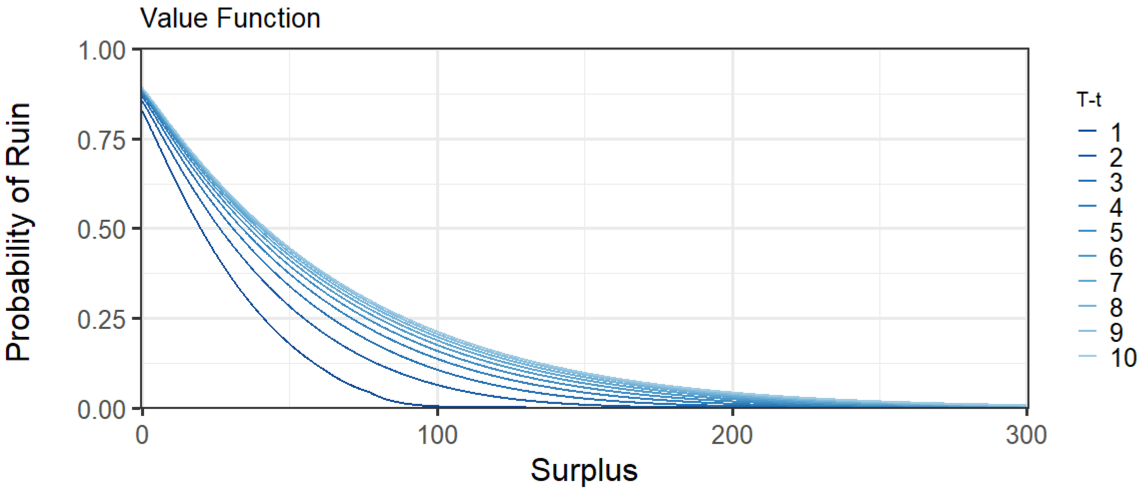

Figure 5: The contour plot of the value function. The blue curves indicate the distance from the current time t to the time horizon T.

Figure 5 shows the value function at different distances from the time horizon. It can be seen that the value functions increase with increasing distance from the current time t to the horizon time T, $|T-t|$. Indicating that the further away from the horizon the higher the probability of ruin for a given surplus value x. The same conclusion could be drawn from figure 3. The roughness for curve $|T-t|=1$ is likely because the grid considered for θ was very sparse for the larger values for θ

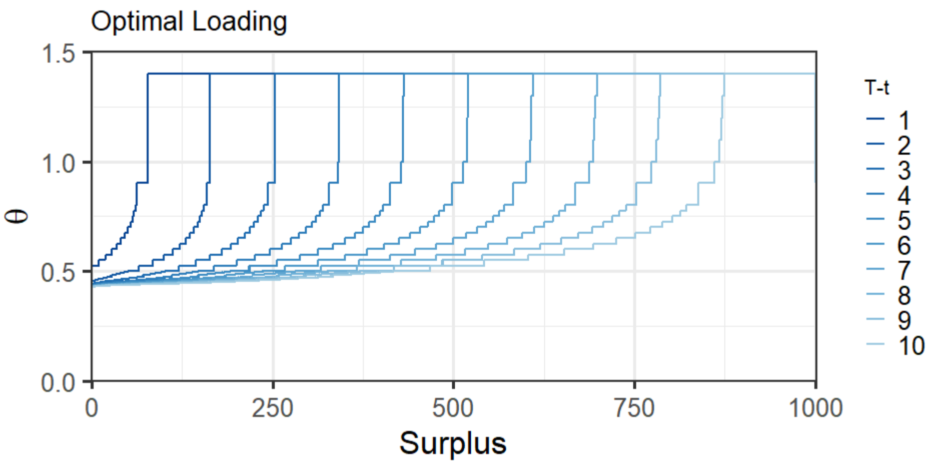

Figure 6: The contour plot of the optimal loading parameter. The blue curves indicate the distance from the current time t to the time horizon T. The considered range of the loading parameter was $[0.38, 1.4]$

Figure 6 shows which θ's are optimal at different distances from the time horizon for different values of the current surplus x. It can be seen that the smaller $|T-t|$ values favor larger values of θ. This means that when t is close to T the optimal strategy is to increase the premiums such that the exposure becomes almost zero. Therefore, it is best to take no more risk and just pay the operational cost when the game is almost finished, that way you will not get ruined, given that the surplus x is high enough. Conversely, the insurance firm has to take on some risk when the surplus is not high, otherwise the insurance firm will get ruined. Furthermore, the firm has to set competitive premium rates to ensure that the premium income is higher than the operational cost. However, if the insurance firm has very high surplus then the optimal strategy is always to set high premium rate, pay the operational cost and sit out the remaining time, that way the insurance firm always wins the game of ruin. This can always be done if $x > r(T-t)$. Note also that the values seem to be converging to one value of theta, they should (probably) be converging to the optimal loading given in the infinite time horizon problem because if T is big then the loading selection process behaves (probably) like the infinite time selection process. Here the convergence point is ≈ 0.425. This indicates that the finite time ruin probability might not be such a good criterion for the selection of the loading parameter.

The reason that $\theta(t,x) = 1.4$ for high values of x is because the maximum considered value in the simulation was $\theta_{max} = 1.4$. Also, the θ grid in the simulation was sparse resulting in a non smooth contour plot. The calculations took a long time but that is because the algorithm in appendix A.2 is very slow for $c \approx 0$. Additionally, algorithm

1 is not computationally efficient.

4.4 Stochastic Control with Infinite Time Horizon

The infinite ruin probability can also be used as a criterion. In this case $T = \infty$. Premium rates are set at discrete times but unlike the finite horizon case, there is no final time that can be used to calculate the value function recursively. The equation from Proposition 3.2 is:

$$V(x, \theta) = \mathbb{E}_x \Big(V(X_1, \theta) \mathbb{1}_{1<\tau} + \mathbb{1}_{\tau \leq 1} \Big). \tag{4.8}$$

Note that the expectation is really the finite time ruin probability with $T = 1$ and modified boundary condition $V_1(T, x) = V(x)$. Therefore, to calculate the ruin probability $V(x)$, it is possible to utilize the fact that the function is a contraction mapping and use an algorithm called the value iteration. The algorithm calculates $V(x, \theta)$ for multiple pairs (x, θ) and chooses the pair which minimizes V. This process of picking pairs is repeated until V does not change.

: Infinite Time Value Iteration

Let the symbol $\vec{\ }$ denote a (computer) vector.

Import $H\Big(t, T, \vec{x}, B, \theta\Big)$, a computer function that computes equation (3.12) with boundary condition B at $t = T$ for each $x_i \in \vec{x}$ with loading parameter θ. Note that $H\Big(t, T, \vec{x}, B, \theta\Big)$ returns vector. A vector where the elements give the probability of ruin for surplus $x_i \in \vec{x}$ at time t

Initialize
$\vec{\theta}$ vector of loading parameters considered
$\vec{x}$ the surplus grid with elements $x_1, ..., x_n$
V^* arbitrarily, e.g. $V^*(x_i) = 0$ for all $x_i \in \vec{x}$
$\vec{\theta^*}$ vector that stores the optimal loading, same size as $\vec{x}$

Loop that finds the value function V^*.
count Number of iterations
for *k in 1:count* **do**
    for x_i *in* $\vec{x}$ **do**
        $V^*(x_i) \leftarrow \max_{\theta \in \vec{\theta}} H\Big(0, 1, \vec{x}, V^*, \theta\Big)$
        $\theta^*(x_i) \leftarrow \arg\min_{\theta \in \vec{\theta}} H\Big(0, 1, \vec{x}, V^*, \theta\Big)$
    end
end

Return an approximation of the value function and the corresponding loading parameters
Return V^*, θ^*

Example 4.4. *Consider the process given in* (4.1)*. Let Y_i be i.i.d Gamma distributed random variables with shape parameter $\alpha = 2$ and scale parameter $k = 5$. Let N_t be a Poisson process with mean $\mathbb{E}[N_t] = \lambda p(\theta) t$ with $\lambda = 0.08$, and $p(\theta) = \frac{1}{1+e^{-0.6+0.4\theta}}$ where θ is the loading parameter. Let the cost be $r = 0.08 * \lambda * k * \alpha$. The value function iteration and the $V(x)$ obtained when α is minimized are compared in the following figure. The value function considered the grid $\theta_0 = 0.2$ to $\theta_n = 0.5$ and $\theta_i - \theta_{i-1} = 0.01$. The x grid considered was $x_0 = 0^+$ to $x_n = 600$ with $x_i - x_{i-1} = 2$ for $i = 1, .., n$.*

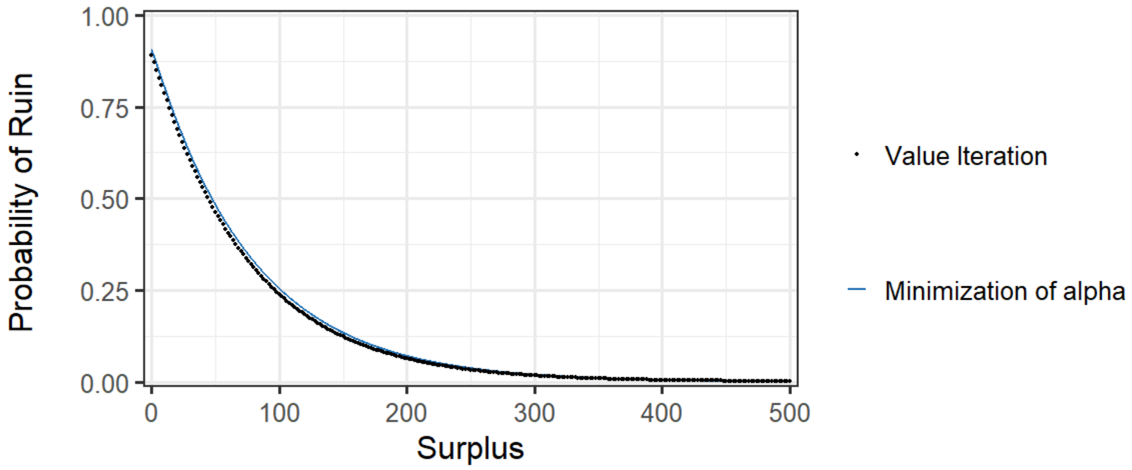

Figure 7: The blue curve gives $V(x)$ obtained when α is minimized while the black dots give the outcome from the value function iteration.

The blue curve gives $V(x)$ obtained when α is minimized while the black dots give the outcome from the value iteration. The value iteration gave the range of $\theta(x) \in [0.41, 0.44]$ while the θ that minimized $V(x)$ was $\theta^* = 0.4349$. More specifically, the mapping is: $\theta\big((0,10]\big) \mapsto 0.41$, $\theta\big((10,44]\big) \mapsto 0.42$, $\theta\big((44,420]\big) \mapsto 0.43$ and $\theta\big((420,\infty)\big) \mapsto 0.44$. This indicates that the static problem and the periodic problem for the infinite time ruin probability is simply the same problem.

It is hard to say whether the optimal loading parameter is independent for periodic updates is independent of the current surplus. The algorithm considered is computationally inefficient as $V(x,\theta)$ has to be fully calculated using the numerical scheme A.2 for each x and each θ. Additionally, algorithm 2 discretizes a continuous process which requires a fine grid of points in order to achieve a good approximation.

4.5 Random Exposure

In reality the average number of potential claims per unit time is a random variable. This means that the intensity parameter of the Poisson process is a random variable meaning that the claim process is a mixed Poisson process. The model discussed in this chapter assumes that no policyholder leaves the insurance company and that the insurance process is Markov. With these assumptions it is possible to incorporate a mixed Poisson distribution to the claim numbers. The process is the same as before

$$X_t = u + c(\theta)t - S_t \tag{4.9}$$

where

$$N_t \mid \epsilon \sim Poisson\big(\lambda p(\theta,\epsilon)\big), \quad S_t = \sum_{i=0}^{N_t} Y_i,$$
$$p(\theta,\epsilon) = \frac{1}{1+e^{\beta_0+\beta_1\theta+\epsilon}}.$$

Here ϵ is some absolutely continuous random variable with distribution $F_\epsilon(x)$. In simple terms, this process describes an insurance business where the premium is unknown at the beginning but stays the same once it is known. The objective is to find θ such that the ruin probability of the process is minimized. As discussed before, the process is not very realistic as customers do not buy insurance at the same time but are instead spread over the year. However, the smaller the decision interval the more realistic the problem becomes. For example, it could be assumed that customers buy insurance at the beginning of each month rather than year. Let $F_\epsilon(z)$ be the distribution of ϵ and assuming that $\mathbb{P}(N_h \geq 2) = o(h)$ then using Proposition 3.2 the integro-differential equation becomes:

$$\begin{aligned} 0 &= \sum_{n=0}^{\infty} \mathbb{E}_x \left((V(X_h) - V(x))\mathbb{1}_{h<\tau} + \mathbb{1}_{\tau\leq h}(1-V(x)) \;\middle|\; N_h = n \right) \mathbb{P}\left(N_h = n \mid z\right) \\ &= \sum_{n=0}^{\infty} \mathbb{E}_x \left((V(X_h) - V(x))\mathbb{1}_{h<\tau} + \mathbb{1}_{\tau\leq h}(1-V(x)) \;\middle|\; N_h = n \right) \int \mathbb{P}\left(N_h = n \mid z\right) dF_\epsilon(z) \end{aligned}$$

$$= \mathbb{E}_x \Big((V(x+c(\theta)h) - V(x)) \mid N_h = 0 \Big) \int e^{-\lambda p(\theta,z)h} f_\epsilon(z)dz \quad +$$

$$\mathbb{E}_x \Big((V(X_h) - V(x)) \mathbb{1}_{h<\tau} + \mathbb{1}_{\tau \le h}(1 - V(x)) \mid N_h = 1 \Big) \int e^{-\lambda p(\theta,z)h} \lambda p(\theta,z) h f_\epsilon(z)dz + o(h)$$

$$\Rightarrow$$

$$0 = \lim_{h \to 0^+} \Bigg(\frac{(V(x+ch) - V(x))}{h} \int e^{-\lambda p(\theta,z)h} f_\epsilon(z)dz \quad +$$

$$\mathbb{E}_x \Big((V(X_h) - V(x)) \mathbb{1}_{x+ch-y \ge 0} + \mathbb{1}_{x+ch-y<0}(1 - V(x)) \mid N_h = 1 \Big) \int e^{-\lambda p(\theta,z)h} \lambda p(\theta,z) f_\epsilon(z)dz + \frac{o(h)}{h} \Bigg)$$

$$\Leftrightarrow$$

$$0 = c\frac{d}{dx}V(x)A + B\big(\int_0^x V(x-y)dF(y) - V(x) + 1 - F(x) \big) \tag{4.10}$$

where

$$A = \lim_{h \to 0} \int e^{-\lambda p(\theta,z)h} f_\epsilon(z)dz, \quad B = \lim_{h \to 0} \int e^{-\lambda p(\theta,z)h} \lambda p(\theta,z) f_\epsilon(z)dz.$$

Given that the assumptions hold and that A and B are well defined then the integro-differential equation is very similar to the one in Proposition 3.4. Also, if $N \mid \lambda \sim Poisson(\lambda)$ and $\lambda \sim Gamma(\alpha, \beta)$ is a gamma random variable with shape parameter α and rate parameter β then equation (4.10) is the same as the infinitesimal generator for a compound negative binomial with $A = 1$ and $B = \alpha\frac{\beta}{1+\beta}$.

5 Multiclaim Processes

Up until now only claim processes of homogeneous groups have been considered, but most financial applications require multivariate modelling with dependence structure, for example portfolio optimization. However, it is difficult to construct multidimensional models with jumps. In this chapter, tools and framework will be provided for the modelling of multivariate models with jumps. The model will have the same flexibility as the Gaussian multivariate case. This section (definitions and propositions) is mostly based on Cont and Tankov [18, Chapter 5] and Tankov [19].

5.1 Independent vs. Dependent Compound Processes

A brief introduction on how one can calculate the ruin probability of the sum of independent surplus processes will be given. Consider the surplus process $\boldsymbol{X} = (X_t^{(1)}, ..., X_t^{(n)})$ where

$$X_t^{(1)} = u^{(1)} + c^{(1)}t - \sum_{i=0}^{N_t^{(1)}} Y_i^{(1)}$$

$$\vdots$$

$$X_t^{(n)} = u^{(n)} + c^{(n)}t - \sum_{i=0}^{N_t^{(n)}} Y_i^{(n)}$$

where $N_t^{(j)}$ are Poisson r.v. with intensity λ_j and $Y_i^{(j)}$ are i.i.d. absolutely continuous random variables which follow the severity distribution $F_j(x)$. If these processes are independent, it is relatively easy to combine them into a single process using the aggregation property in Proposition 3.1. The aggregation property allows the combination of multiple surplus processes into a single risk process as follows:

$$X_t = \sum_{j=1}^{n} u^{(j)} + \sum_{j=1}^{n} c^{(j)} t - \sum_{i=0}^{N_t} Y_i$$

where N_t is a Poisson r.v. with $\lambda = \lambda_1 + ... + \lambda_n$ and Y_i are i.i.d. random variables, which follows the severity distribution $F(x) = \sum_{j=1}^{n} \frac{\lambda_j}{\lambda} F_j(x)$. Finally, to calculate the infinite time ruin probability, one can use equation (3.3) or (3.4). This is, however, not true for dependent Poisson processes and the remaining part of this chapter will be used to develop tools that can be used to combine dependent compound Poisson processes.

5.2 Copulas for Random Variables

Before dependence concepts for compound Poisson processes are discussed, it is necessary to introduce copulas for random variables. Informally, copulas are functions that use the one-dimensional marginal distributions as building blocks for the joint multivariate distribution function. The idea is to glue or couple the marginals together in order to construct the dependence structure in an easy manner. Copulas have been of interest to statisticians for two main reasons: firstly, as a way of studying scale-free measures of dependence; and secondly, as a starting point for constructing families of bivariate distributions, sometimes with a view to simulation. In other words, copulas reveal the true nature of dependence between variables and lead to flexible multivariate models [14]. For more details on random variable copulas see Nelsen [17]. The chapter will describe bivariate processes, but the theory can be easily extended to n-dimensions.

The law of a two-dimensional random vector (X, Y) is typically described via its cumulative distribution function:

$$F(x, y) = \mathbb{P}(X \le x, Y \le y)$$

and the marginal laws can simply be obtained from the two-dimensional distribution function:

$$F_1(x) = F(x, \infty), \qquad F_2(y) = F(\infty, y).$$

However, the knowledge of the marginal laws alone is not enough to determine the two-dimensional random vector as the distribution function $F(x, y)$ contains information about dependence.

The copula of a two-dimensional random vector is a two argument function that defines the dependence structure. The distribution function can be computed from the copula $C(x, y)$ and the marginals: $F(x, y) = C(F_1(x), F_2(y))$. If $F_1(x)$ and $F_2(y)$ are absolutely continuous then the copulas are unique and can be computed from the distribution function and the margins: $C(x, y) = F(F_1^{-1}(x), F_2^{-1}(y))$. The following definitions from Cont and Tankov [18, Definition 5.4 and Definition 5.5] define F-volume, 2-increasing function and grounded function. They are essential for ordinary copulas and Lévy copulas.

Definition 5.1 (F-volume). *Let I_1, I_2 be nonempty subsets of $\overline{\mathbb{R}}$, let F be a real valued function of n variables such that $Dom(F) = I_1 \times I_2$ and for $\boldsymbol{a} \le \boldsymbol{b}$ $(a_k \le b_k)$ for k = 1,2, let $B = [\boldsymbol{a}, \boldsymbol{b}]$ be a 2-box whose vertices are in $Dom(F)$. Then the F-volume of B is defined by*

$$Vol_F(B) = F(b_1, b_2) - F(b_1, a_2) - F(a_1, b_2) + F(a_1, a_2).$$

Definition 5.2 (2-increasing function, grounded function, margins)**.**

- *A real function f of 2 variables is called 2-increasing if $Vol_F(B) \ge 0$ for all 2-boxes whose vertices lie in $Dom(F)$.*
- *Suppose that the domain of F is $I_1 \times I_2$ where each I_k has a smallest element a_k, k =1,2. F is said to be grounded if for every $x \in I_1$, $F(x, \min I_2) = 0$ and for every $y \in I_2$, $F(\min I_1, y) = 0$.*
- *If each I_k is nonempty and has a greatest element , then (one-dimensional) margins of F are functions F_k, with $Dom(F_k) = I_k$ defined by $F_1(x) = F(x, \max I_2)$ and $F_2(x) = F(\max I_1, x)$ for all x in I_k.*

In Cont and Tankov [18, Definition 5.3], copulas are defined the following way:

Definition 5.3 (Copula)**.** *A two-dimensional copula is a function C with domain $[0, 1]^2$ such that*

1. *C is grounded and 2-increasing.*
2. *C has margins C_k, k=1,2 which satisfy $C_k(u) = u$ for all u in [0,1].*

From the previous definition, copulas can be considered as a distribution function on $[0,1]^2$ with uniform margins. The next Theorem is central for the construction of ordinary copulas [18, Theorem 5.1].

Theorem 5.1 (Sklar's Theorem)**.** *Let F be a two-dimensional distribution function with margins F_1, F_2. Then there exists a two-dimensional copula C such that for all $x \in \bar{\mathbb{R}}^2$,*

$$F(x_1, x_2) = C(F_1(x), F_2(y)) \tag{5.1}$$

if F_1, F_2 are absolutely continuous then C is unique, otherwise C is uniquely determined on $RangeF_1 \times RangeF_2$. Conversely, if C is a copula and F_1 and F_2 are distribution functions, then the function F defined by (5.1) *is two-dimensional distribution function with margins F_1 and F_2.*

5.3 Dependence Concepts for Compound Poisson Processes

The compound Poisson process is a particular type of a Lévy process. This thesis will not go into the details of Lévy processes. For deeper analysis of Lévy processes, see Cont and Tankov [18], Papapantoleon [20] or Sato [21]. Unlike continuous random variables, dependence concepts for compound Poisson processes are best explained through their Lévy measure. Luckily, the formulation of the two copulas is very similar.

A Lévy process is a càdlàg process with stationary and independent increments. A Lévy process can be decomposed into a diffusion and a jump process where the jump prcess is characterized by its Lévy measure.

Definition 5.4 (Lévy measure)**.** *A Lévy measure in $\mathbb{R}^2$ is a Borel measure satisfying*

$$\nu(\{0\}) = 0 \qquad \textit{and} \qquad \int_{\mathbb{R}^2} (1 \wedge |x|^2)\nu(dx) < \infty.$$

Intuitively speaking, if ν is the Lévy measure for a 2-dimensional Lévy process X, and $A \subset \mathbb{R}^2$ is a Borel set, then $\nu(A)$ is the expected number of jumps such that $X_t - X_{t^-} \in A$, per unit of time. From the condition $\int_{\mathbb{R}^2}(1 \wedge |x|^2)\nu(dx) < \infty$, it follows that almost every trajectory has a finite number of jumps larger than some constant $\epsilon > 0$. However, the number of small jumps ($|X_t - X_{t^-}| < \epsilon$) may be infinite with positive probability.

For a compound Poisson process the Lévy measure is

$$\nu(dx) = \lambda dF(x)$$

where λ is the expected number of jumps and F is the jump size distribution. It is possible to construct a probability measure if ν is a finite measure, i.e. $dF(x) := \frac{\nu(dx)}{\lambda}$. The Lévy measure carries the information about the jump component of the process and is fundamental to the theory of Lévy processes. Recall that the support of a Borel measure μ is the complement of the largest open set A such that $\mu(A) = 0$.

Proposition 5.2. *Let (X_t, Y_t) be a compound Poisson Lévy process with Lévy measure ν. Its components are independent if and only if the support of ν is contained in the set $\{(x,y) : xy = 0\}$, that is, if and only if they never jump together. In this case*

$$\nu(A) = \nu_X(A_X) + \nu_Y(A_Y) \tag{5.2}$$

where $A_X = \{x : (x,0) \in A\}$ and $A_Y = \{y : (0,y) \in A\}$, and ν_x and ν_y are Lévy measures of X_t and Y_t.

The Lévy measure plays a similar role as the probability measure for random variables. An important difference is that for independent random variables we have $\mu([a,b]\times[c,d]) = \mu_X([a,b])\mu_Y([c,d])$, where μ is the distribution of the random variable, while for independent jump processes we have equation (5.2). The following two properties of increasing functions will be convenient later on for the construction of Lévy copulas [18, Lemma 5.1]
.

Lemma 5.3. *Let H be a 2-increasing function.*

1. *Increasing variable changes: if the functions f_1 and f_2 are increasing then $H(f_1(x), f_2(x))$ is 2-increasing.*
2. *Transformation by a function with positive derivatives: if H is grounded and f is increasing, convex and satisfies $f(0) = 0$ then the composition $f \circ H$ is a 2-increasing grounded function.*

Tail integrals play a significant role in the theory of Lévy copulas and are defined as follows [18, Definition 5.7]:

Definition 5.5. *A 2-dimensional tail integral is a function $U : [0,\infty]^2 \to [0,\infty]$ such that*

1. *U is a 2-increasing function.*

2. *U is equal to zero if one of its arguments is equal to* ∞.

3. *U is finite everywhere except at zero,* $U(0,0)=\infty$.

The role of distribution function is now played by the tail integral. The marginals of a tail integral are defined similarly to the marginals of a distribution function

$$U(x_1,0)=U_1(x_1),\quad U(0,x_2)=U_2(x_2).$$

Definition 5.6 (Lévy Copula for Processes with Positive Jumps)**.** *A two-dimensional Lévy copula for Lévy processes with positive jumps, or for short, a positive Lévy copula, is a 2-increasing grounded function* $\mathcal{C}:[0,\infty]^2\to[0,\infty]$ *with uniform margins, that is,* $\mathcal{C}(x,\infty)=\mathcal{C}(\infty,x)=x$.

The next theorem is similar to Sklar's theorem for random variable copulas [18, Theorem 5.3]:

Proposition 5.4. *Let U be a two-dimensional tail integral with margins* U_1 *and* U_2. *There exists a positive Lévy copula* $\mathcal{C}$ *such that*

$$U(x_1,x_2)=\mathcal{C}(U_1(x_1),U_2(x_2)) \tag{5.3}$$

if U_1 *and* U_2 *are absolutely continuous, this Lévy copula is unique, otherwise it is unique on Range*$(U_1)\times$ *Range*(U_2), *the product of ranges of one-dimensional tail integrals.*
Conversely, if $\mathcal{C}$ *is a positive Lévy copula and* U_1, U_2 *are one-dimensional tail integrals then equation* (5.3) *defines a two-dimensional tail integral.*

The proposition shows that Lévy copulas link multidimensional tail integrals to their marginals in the same way as the copulas link the distribution functions to their marginals. For every Lévy measure ν on $[0,\infty]^2$, one can define its tail integral as follows:

Definition 5.7. *Tail integral of a Lévy measure* ν *on* $[0,\infty]^2$ *is*

$$\begin{aligned}&U(x_1,x_2)=0\quad if\quad x_1=\infty\quad or\quad x_2=\infty.\\&U(x_1,x_2)=\nu\big([x_1,\infty]\times[x_2,\infty[\big)\quad for\quad (x_1,x_2)\in]0,\infty[^2.\\&U(0,0)=\infty.\end{aligned} \tag{5.4}$$

Tail integrals should define a positive Lévy measure. The following lemma shows that the integrability condition of Definition 5.4 holds [18, Lemma 5.2]:

Lemma 5.5. *Let U be a two-dimensional tail integral with margins* U_1 *and* U_2. *U defines a Lévy measure on* $[0,\infty[^2\backslash\{0\}$, *that is the condition of 5.4 is satisfied if and only if the margins of U correspond to Lévy measures on* $[0,\infty]$, *that is, for* $k=1,2$.

$$\int_0^1 x^2 dU_k(x)<\infty.$$

These results combined with the following theorem are fundamental to the theory of Lévy copulas [18, Theorem 5.4].

Theorem 5.6. *Let* (X_t,Y_t) *be a pure jump Lévy process with positive jumps having tail integral U and marginal tail integrals* U_1 *and* U_2. *There exists a two-dimensional positive Lévy copula* $\mathcal{C}$ *which characterizes the dependence structure of* (X_t,Y_t), *that is for all* $x_1,x_2\in[0,\infty]$,

$$U(x_1,x_2)=\mathcal{C}(U_1(x_1),U_2(x_2)) \tag{5.5}$$

if U_1 *and* U_2 *are absolutely continuous, this Lévy copula is unique, otherwise it is unique on Range*$(U_1)\times$ *Range*(U_2), *the product of ranges of one-dimensional tail integrals.*

Conversely, let X_t *and* Y_t *be two one-dimensional pure jump Lévy process with positive jump with positive jumps having tail integrals* U_1, U_2 *and let* $\mathcal{C}$ *be a two-dimensional positive Lévy copula. Then there exists a two-dimensional Lévy process with Lévy copula* $\mathcal{C}$ *and marginal tail integrals* U_1 *and* U_2. *Its tail integral is given by equation* (5.5).

The first part of Theorem 5.6 states that all types of dependence in the jump component of a Lévy process, including perfect dependence and independence, can be represented with Lévy copulas and the second part shows that one can construct multivariate jump Lévy process models by specifying one-dimensional laws for the components.

5.4 Independence and Perfect Dependence Copulas

In this subsection Lévy copulas for the two extreme cases of independence and perfect dependence will be presented. Independence means that the jumps of one process will not affect the jumps of the other process while perfect dependence means that if a process jumps then the other process will jump as well. To emphasize, this thesis is only concerned with positive jumps and therefore positive dependence. So if one process jumps then the other process will jump in the same direction (both processes have positive increments). From Proposition 5.2 the Lévy measure of a compound Poisson with independent components is $\nu(A) = \nu_X(A_X) + \nu_Y(A_Y)$. The tail integral is given by:

$$U(x_1, x_2) = U_1(x_1)\mathbb{1}_{x_2=0} + U_2(x_2)\mathbb{1}_{x_1=0}$$

and to calculate the Lévy copula one can "invert the arguments" of equation (5.5) to get

$$\mathcal{C}_\perp(x_1, x_2) = U(U_1^{-1}(x_1), U_2^{-1}(x_2)) = x_1\mathbb{1}_{x_2=\infty} + x_2\mathbb{1}_{x_1=\infty}.$$

The symbol $\perp$ is used to denote independence. To analyze perfectly dependent processes, the notion of an increasing set is needed [18, Definition 5.9 and Definition 5.10].

Definition 5.8 (Increasing Set)**.** *A subset I of $\mathbb{R}^2$ is called increasing if for every two distinct vectors $(v_1, v_2) \in I$ and $(u_1, u_2) \in I$ either $v_k < u_k\ \forall k$ or $v_k > u_k\ \forall k$.*

Definition 5.9. *Let $\boldsymbol{X} = (X_t^1, X_t^2)$ be a Lévy process with positive jumps. Its jumps are said to be perfectly dependent or comonotonic if there exists an increasing subset I of $]0,\infty[^2$ such that every jump $\Delta\boldsymbol{X}$ of $\boldsymbol{X}$ is in I.*

That is, if the two jump processes are perfectly dependent, one of them can be reconstructed from the trajectory of the other [18, Proposition 5.4].

Proposition 5.7. *Let $\boldsymbol{X} = (X_t^1, X_t^2)$ be a Lévy process with positive jumps. If its jumps are perfectly dependent, then the Lévy copula of $\boldsymbol{X}$ is the perfect dependence Lévy copula defined by*

$$\mathcal{C}_\parallel(x_1, x_2) = min(x_1, x_2).$$

Conversely, if the Lévy copula of $\boldsymbol{X}$ is given by $\mathcal{C}_\parallel$ and the tail integrals of components of $\boldsymbol{X}$ are continuous, then the jumps of $\boldsymbol{X}$ are completely dependent.

The symbol $\parallel$ is used to denote perfect dependence.

5.5 Dependence of Compound Poisson Processes

Take two surplus processes X_1 and X_2. Let $\mathcal{C}$ be the Lévy copula and let U_1 and U_2 be the tail integrals of X_1 and X_2. The Lévy measure of a compound Poisson is simply the intensity times the density of the severity or $\nu(x) = \lambda f(x)$. From equation (5.4) it can be seen that the tail integral is simply $U(x) = \lambda\bar{F}(x)$ where $\bar{F}$ is the survival function. Therefore, the intensities of λ_1 and λ_2 are

$$\lambda_i = \lambda_i\bar{F}_i(0) = \lim_{x\to 0+} U_i(x).$$

Note that $\lambda_i = U_i(0)$ can not be written because according to the definition of a tail integral $U(0) = \infty$. The intensity of the two dimensional process (X_1, X_2) is equal to:

$$\lambda = \lim_{x_1,x_2\to 0+} \big(U_1(x_1) + U_2(x_2) - U(x_1, x_2)\big) = \lambda_1 + \lambda_2 - \mathcal{C}(\lambda_1, \lambda_2).$$

Since every positive Lévy copula satisfies $0 \le \mathcal{C}(x, y) \le \min(x, y)$ it follows that $\lambda \in [\max(\lambda_1, \lambda_2), \lambda_1 + \lambda_2]$.

The tail integrals of the independent components are

$$\begin{aligned} U_1^\perp(x_1) &= \nu([x_1,\infty[, \{0\}) = U(x_1, 0) - \lim_{x_2\to 0+} U(x_1, x_2) = U_1(x) - C(U_1(x_1), \lambda_2), \\ U_2^\perp(x_2) &= U_2(x_2) - C(\lambda_1, U_2(x_2)). \end{aligned} \tag{5.6}$$

The tail integral of the common component is

$$U_{\|}(x_1,x_2) = \lim_{x_1'\to x_1+, x_2'\to x_2+} U(x_1',x_2') = \mathcal{C}(min(U_1(x_1),\lambda_1), min(U_2(x_2),\lambda_2)). \tag{5.7}$$

Finally, to find the intensities, the tail integrals can be used. The intensity of the common Poisson shocks is equal to:

$$\lambda^{\|} = \mathcal{C}(\lambda_1,\lambda_2). \tag{5.8}$$

The intensity of the independent Poisson processes is equal to:

$$\begin{aligned} \lambda_1^{\perp} &= \lambda_1 - \lambda^{\|}, \\ \lambda_2^{\perp} &= \lambda_2 - \lambda^{\|}. \end{aligned} \tag{5.9}$$

5.5.1 Construction of Lévy Copulas

Lévy copulas can be constructed by "inverting the arguments" of equation (5.5) to get

$$\mathcal{C}(x_1,x_2) = U(U_1^{-1}(x_1), U_2^{-1}(x_2)).$$

However, there are few multivariate Lévy models available. Therefore a more general approach is needed [18, Proposition 5.5].

Proposition 5.8. *Let C be an (ordinary) 2-copula and f(x) an increasing convex function from $[0,1]$ to $[0,\infty]$. Then*

$$\mathcal{C}(x,y) = f\big(C(f^{-1}(x), f^{-1}(y))\big)$$

defines a two-dimensional positive Lévy Copula.

Proof. $\mathcal{C}$ is a 2-increasing function from Lemma 5.3. For groundedness note:

$$\mathcal{C}(0,y) = f\big(C(0,f^{-1}(y))\big) = 0$$

and the marginal property follows because

$$\mathcal{C}(\infty,y) = f\big(C(1,f^{-1}(y))\big) = f\big(f^{-1}(y)\big) = y.$$

□

What the proposition says is that it is possible to take an ordinary copula for random variables and create a Lévy copula from an increasing convex function. Common choices of f are $f(x) = x/(1-x)$ and $f(x) = -log(1-x)$.

By analogy with (ordinary) Archimedean copulas, one can also construct Archimedean Lévy copulas [18, Proposition 5.6].

Proposition 5.9. *Let ϕ be a strictly decreasing convex function from $[0,\infty]$ to $[0,\infty]$ such that $\phi(0)=\infty$ and $\phi(\infty)=0$. Then*

$$\mathcal{C}(x,y) = \phi^{-1}\big(\phi(x)+\phi(y)\big)$$

defines a two-dimensional positive Lévy copula.

Proof. $\mathcal{C}$ has to be a 2-increasing grounded function with uniform marginals. Uniform marginals follow because:

$$\mathcal{C}(\infty,y) = \phi^{-1}(0+\phi(y)) = y.$$

The groundedness follows because:

$$\mathcal{C}(0,y) = \phi^{-1}(\infty+\phi(y)) = 0.$$

The fact that $\mathcal{C}$ is 2-increasing comes from Lemma 5.3. Let $H(x,y) = x+y$ and note that H is 2-increasing ($Vol_H(B) = 0$). Next take $f_1(x) = f_2(x) = -\phi(x)$ which are increasing functions so $H(-\phi(x), -\phi(y))$ is 2-increasing. Finally note that $\phi^{-1}(-u)$ is an increasing convex function so it follows that $\mathcal{C} = \phi^{-1} \circ H \circ -\phi$ is 2-increasing.

□

5.6 Examples of Lévy Copulas

There is an extensive list of copulas which can be created with propositions 5.8 or 5.9. Examples can be found in Benth et al. [25] or Kettler [26].

Example 5.1 (Clayton family of Lévy copulas)**.** *For $\phi(u) = u^{-\omega}$ with $\omega > 0$ the following copula is obtained:*

$$\mathcal{C}_\omega(x, y) = (x^{-\omega} + y^{-\omega})^{-\frac{1}{\omega}}$$

If $x \leq y$ then:

$$x^{-\omega} \leq x^{-\omega} + y^{-\omega} \leq 2x^{-\omega} \Leftrightarrow 2^{-\frac{1}{\omega}}x \leq C(x, y) \leq x.$$

Let $\omega \to \infty$ such that $\mathcal{C}(x, y) = x$. Therefore, if $\omega \to \infty$ the perfect dependence case is obtained, $\mathcal{C}(x, y) = min(x, y)$.

Now consider when $\omega \to 0$ then $\mathcal{C}(x, y) = 0$ unless $x = \infty$ then $\mathcal{C}(x, y) = y$ and likewise if $y = \infty$ then $\mathcal{C}(x, y) = x$. Therefore the case of independence is obtained, $\mathcal{C}(x, y) = x\mathbb{1}_{y=\infty} + y\mathbb{1}_{x=\infty}$.

The Clayton Lévy copula gives is a smooth transition from the independence copula to the perfect dependence copula.

Example 5.2 (Gumbel Lévy Copula)**.** *The Gumbel Lévy Copula is generated by $\phi(u) = [log(u + 1)]^{-\omega}$ with $\omega > 0$. The inverse is simply $\phi^{-1}(u) = e^{u^{-\frac{1}{\omega}}} - 1$ which gives the Gumbel Lévy copula:*

$$\mathcal{C}(x, y) = \exp\left\{\left[(log(x + 1))^{-\omega} + (log(y + 1))^{-\omega}\right]^{-\frac{1}{\omega}}\right\} - 1.$$

With the same arguments as the ones for the Clayton Lévy copula the independence copula is obtained when $\omega \to 0+$ and the complete dependence Lévy copula is obtained when $\omega \to \infty$.

5.7 Simulation of a Bivariate Process with a Lévy Copula

The compound Poisson processes S_1 and S_2 can be decomposed into common jumps and independent jumps in order to simulate the bivariate Lévy copula. That is $S_1^{\perp}$, $S_2^{\perp}$ and $(S_1^{\|}, S_2^{\|})$. Let $F_i^{\perp}(x)$ denote the marginal distribution of independent jumps and $F^{\|}(x)$ denote the marginal distribution of common jumps. $\bar{F}_i^{\perp}(x)$ and $\bar{F}^{\|}(x)$ are the corresponding survival distributions. The survival function of the common jumps can be found by using the tail integral of the common jumps, equation (5.7):

$$\bar{F}^{\|}(x_1, x_2) = \frac{1}{\lambda^{\|}}\mathcal{C}(\lambda_1\bar{F}_1(x_1), \lambda_2\bar{F}_2(x_2))$$

and the distribution function is related to the survival distribution by

$$F^{\|}(x_1, x_2) = 1 - \bar{F}^{\|}(x_1, 0) - \bar{F}^{\|}(0, x_2) + \bar{F}^{\|}(x_1, x_2)$$

the density $f^{\|}$ is given by

$$f^{\|}(x_1, x_2) = \frac{\partial^2 F^{\|}(x_1, x_2)}{\partial x_1 \partial x_2} = \frac{\lambda_1\lambda_2 f_1(x_1) f_2(x_2)}{\lambda^{\|}} \frac{\partial^2 \mathcal{C}(x_1, x_2)}{\partial x_1 \partial x_2}\Big|_{x_1=\lambda_1 S_1(x_1), x_2=\lambda_2 S_2(x_2)}.$$

The survival function of the independent jumps is found by using the tail integral of the independent jumps, equation (5.6)

$$\bar{F}_1^{\perp}(x) = \frac{\lambda_1}{\lambda_1^{\perp}}\bar{F}_1(x) - \frac{1}{\lambda_1^{\perp}}\mathcal{C}(\lambda_1\bar{F}_1(x), \lambda_2),$$
$$\bar{F}_2^{\perp}(x) = \frac{\lambda_2}{\lambda_2^{\perp}}\bar{F}_2(x) - \frac{1}{\lambda_2^{\perp}}\mathcal{C}(\lambda_1, \lambda_2\bar{F}_2(x)).$$

Let $Y^{1\|}$ and $Y^{2\|}$ denote the common jump sizes of process 1 and process 2 respectively, then we can write:

$$\begin{aligned}
\mathbb{P}\left(Y^{2\|} \leq y \,\middle|\, Y^{1\|} = x\right) &= \lim_{h\to 0} \mathbb{P}\left(Y^{2\|} \leq y \,\middle|\, x \leq Y^{1\|} \leq x+h\right) \\
&= \lim_{h\to 0} \frac{\mathbb{P}\left(Y^{2\|} \leq y, x \leq Y^{1\|} \leq x+h\right)}{\mathbb{P}\left(x \leq Y^{1\|} \leq x+h\right)} \\
&= \lim_{h\to 0} \frac{F^{\|}(x+h,y) - F^{\|}(x,y)}{F_1^{\|}(x+h) - F_1^{\|}(x)} \\
&= \frac{\frac{\partial F^{\|}(x,y)}{\partial x}}{f_1^{\|}(x)} = H_x(y).
\end{aligned}$$

To simulate the process, the steps of van Velsen [27] can be followed:

- Draw $N_1^{\perp}$ and $N_2^{\perp}$ from a Poisson distribution with frequency $\lambda_1^{\perp}$ and $\lambda_2^{\perp}$, respectively.
- Draw $N^{\|}$ from a Poisson distribution with frequency $\lambda^{\|}$.
- Draw $N_1^{\perp}$ times from a uniform $[0,1]$ distribution. The resulting draws are the jump times of $S_1^{\perp}$. Do the same for $N_2^{\perp}$.
- Draw $N_1^{\perp}$ times from a uniform $[0,1]$ distribution and apply the inverse of $F_1^{\perp}$ to each draw. The resulting numbers are the jump sizes of $S_1^{\perp}$. Do the same to determine $S_2^{\perp}$.
- Draw $N^{\|}$ times from a unifrom $[0,1]$ and apply the inverse of the marginal distribution function $F_1^{\|}$ defined as $F_1^{\|}(x) = F^{\|}(x,\infty)$. The resulting numbers x_i with $i = 1,..,N^{\|}$ are the jump sized of $S_1^{\|}$.
- Draw $N^{\|}$ times from a uniform $[0,1]$ distribution. The resulting draws are denoted by u_i with $i = 1,...,N^{\|}$. Apply the inverse of the distribution function $H_{x_i}(y)$ to u_i for all $i = 1,...,N^{\|}$. The resulting numbers y_i with $i = 1,...,N^{\|}$ are the jump sizes of $S_2^{\|}$

Example 5.3. *Let the marginal jump size distribution be exponential*

$$F_i(x) = 1 - e^{-\mu_j x}$$

and let the dependence be explained by the Clayton Lévy copula. The independent jump intensities are

$$\lambda_i^{\perp} = \lim_{x_i\to 0} U_i^{\perp}(x_i) = \lim_{x_i\to 0} \left(U_i(x_i) - \mathcal{C}(\lambda_j, U_i(x_i))\right) = \lambda_i - \lambda^{\|}, \quad j \neq i.$$

The common jump intensity is found by using equation (5.8)*. The survival function of the common jumps is*

$$\bar{F}^{\|}(x_1,x_2) = \frac{1}{\lambda^{\|}}\left((\lambda_1 \bar{F}_1(x_1))^{-\omega} + (\lambda_2 \bar{F}_2(x_2))^{-\omega}\right)^{-\frac{1}{\omega}}.$$

The distribution of the common jumps is

$$F^{\|}(x_1,x_2) = 1 - \frac{1}{\lambda^{\|}}\left((\lambda_1 \bar{F}_1(x_1))^{-\omega} + (\lambda_2)^{-\omega}\right)^{-\frac{1}{\omega}} - \frac{1}{\lambda^{\|}}\left((\lambda_1)^{-\omega} + (\lambda_2 \bar{F}_2(x_2))^{-\omega}\right)^{-\frac{1}{\omega}} + \bar{F}^{\|}(x_1,x_2)$$

and the following equation is also needed:

$$\begin{aligned}
\frac{\partial F^{\|}(x_1,x_2)}{\partial x_1} = &\frac{1}{\lambda^{\|}}\frac{1}{\omega}\left((\lambda_1 \bar{F}_1(x_1))^{-\omega} + \lambda_2^{-\omega}\right)^{-\frac{1}{\omega}-1}\left(\omega \lambda_1^{-\omega} \bar{F}_1(x_1)^{-\omega-1} f_1(x_1)\right) \quad - \\
&\frac{1}{\lambda^{\|}}\frac{1}{\omega}\left((\lambda_1 \bar{F}_1(x_1))^{-\omega} + (\lambda_2 \bar{F}_2(x_2))^{-\omega}\right)^{-\frac{1}{\omega}-1}\left(\omega \lambda_1^{-\omega} \bar{F}_1(x_1)^{-\omega-1} f_1(x_1)\right).
\end{aligned}$$

Finally, $f_1^{\|}(x)$ can be approximated via

$$f_1^{\|}(x) = \frac{F^{\|}(x+h,\infty) - F^{\|}(x-h,\infty)}{2h}$$

for small values of h.

Let the jump intensities be $\lambda_1 = 0.8$ and $\lambda_2 = 0.5$ then $\lambda^{\|} = 0.42$. Let the rate parameters of the claims be $\mu_1 = 0.9$ and $\mu_2 = 0.6$ and let the dependency parameter be $\omega = 2$. The following figure shows one trajectory of two dependent jump processes.

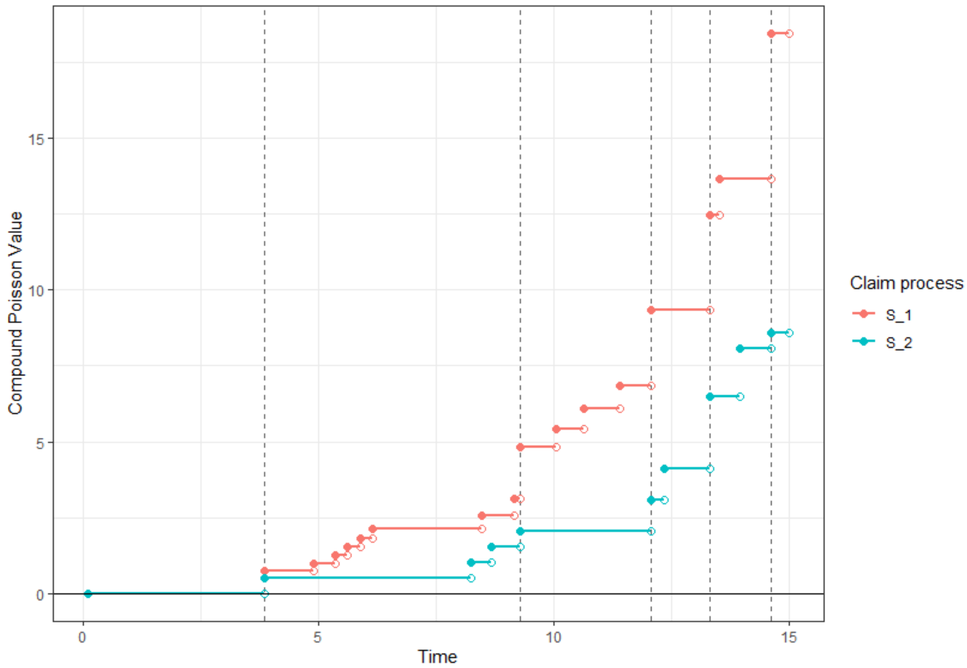

Figure 8: The figure shows one trajectory of two dependent jump processes, discussed in Example 5.3. The black dashed lines show where common jumps happen.

5.8 Ruin Probability of Multivariate Surplus Processes

It is possible to incorporate Lévy copulas in the ruin theory of the previous chapters to calculate the ruin probability of a surplus process composed of two dependent processes. Consider the following surplus processes:

$$X_t^{(1)} = u^{(1)} + c^{(1)}t - \sum_{i=0}^{N_t^{(1)}} Y_i^{(1)},$$
$$X_t^{(2)} = u^{(2)} + c^{(2)}t - \sum_{i=0}^{N_t^{(2)}} Y_i^{(2)}.$$

Aggregation of the these processes gives a new process:

$$X_t = X_t^{(1)} + X_t^{(2)} = u + ct - S_t^{(1)} - S_t^{(2)}$$

where $u = u^{(1)} + u^{(2)}$, $c = c^{(1)} + c^{(2)}$, and $S_t^{(j)} = \sum_{i=0}^{N_t^{(j)}} Y_i^{(j)}$. If $S_t^{(1)}$ and $S_t^{(2)}$ are independent then it is easy to combine them into a new compound Poisson process using Proposition 3.1 and compute the ruin probability using Proposition 3.4. If the two processes are dependent, a similar strategy can be used. First decide which copula structure best describes the dependency. Then decompose the jumps of the two surplus processes into independent jumps and common jumps. Let $S_t^{i\perp}$ denote independent jumps and $S_t^{\parallel}$ denote common jumps. The process becomes:

$$X_t = X_t^{(1)} + X_t^{(2)} = u + ct - S_t^{1\perp} - S_t^{2\perp} - S_t^{\parallel}.$$

Furthermore, note that $S_t^{1\perp}$, $S_t^{2\perp}$ and $S_t^{\parallel}$ are all mutually independent so it is still possible to use Proposition 3.1. First the distribution functions of the independent and common jump sizes need to be determined.

Let $Z = Y^{1\parallel} + Y^{2\parallel}$ be the random variable of common jump size. The distribution of Z can be written as

$$F^{\parallel}(z) = \int_{-\infty}^{\infty} \int_{0}^{z-x} f^{\parallel}(x, y) dy dx$$

and so the density becomes

$$f^{\parallel}(z) = \int_{0}^{z} f^{\parallel}(x, z - x) dx. \tag{5.10}$$

This is perhaps not easy to compute analytically, but it can be computed numerically. The procedure is to use equation (5.7), to find the survival function

$$\bar{F}^{\|}(x_1, x_2) = \frac{1}{\lambda^{\|}} \mathcal{C}(\lambda_1 \bar{F}_1(x_1), \lambda_2 \bar{F}_2(x_2))$$

and the distribution function is related to $\bar{F}^{\|}$ by

$$F^{\|}(x_1, x_2) = 1 - \bar{F}^{\|}(x_1, 0) - \bar{F}^{\|}(0, x_2) + \bar{F}^{\|}(x_1, x_2).$$

The form of the copula and the marginal distribution functions $\bar{F}_1$ and $\bar{F}_2$ are known, thus the joint density can be found using

$$f^{\|}(x_1, x_2) = \frac{\partial^2 F^{\|}(x_1, x_2)}{\partial x_1 \partial x_2} = \frac{\lambda_1 \lambda_2 f_1(x_1) f_2(x_2)}{\lambda^{\|}} \frac{\partial^2 \mathcal{C}(x_1, x_2)}{\partial x_1 \partial x_2}\Big|_{x_1 = \lambda_1 S_1(x_1), x_2 = \lambda_2 S_2(x_2)}$$

from this expression an explicit formula can be obtained and thus equation (5.10) can be used.

The distribution function of $Y^{1\|}$ and $Y^{2\|}$ can be found with:

$$F_{Y^{1\|}}(x) = F_{Y^{\|}}(x, \infty), \quad F_{Y^{2\|}}(x) = F_{Y^{\|}}(\infty, x).$$

The survival function of the independent jumps is found by using the tail integral of the independent jumps, equation (5.6)

$$\bar{F}_1^{\perp}(x) = \frac{\lambda_1}{\lambda_1^{\perp}} \bar{F}_1(x) - \frac{1}{\lambda_1^{\perp}} \mathcal{C}(\lambda_1 \bar{F}_1(x), \lambda_2),$$
$$\bar{F}_2^{\perp}(x) = \frac{\lambda_2}{\lambda_2^{\perp}} \bar{F}_2(x) - \frac{1}{\lambda_2^{\perp}} \mathcal{C}(\lambda_1, \lambda_2 \bar{F}_2(x)).$$

The corresponding densities can be found with differentiation. The intensities $\lambda_1^{\|}$, $\lambda_2^{\|}$ and $\lambda^{\perp}$ are found by using equations (5.9) and (5.8). Now using Proposition 3.1 the process becomes

$$X_t = u + ct - S_t^*$$

where S_t^* is a compound Poisson with intensity parameter $\lambda^* = \lambda_1^{\perp} + \lambda_1^{\perp} + \lambda^{\|}$ and claim severity density:

$$f^*(x) = \frac{\lambda_1^{\perp}}{\lambda^*} f_1^{\perp}(x) + \frac{\lambda_2^{\perp}}{\lambda^*} f_2^{\perp}(x) + \frac{\lambda^{\|}}{\lambda^*} f^{\|}(x).$$

Example 5.4 (Exponential claims)**.** *Let $Y^{(1)}$ and $Y^{(2)}$ be exponential random variables, both with rate parameter $\mu_1 = \mu_2 = 1$. Let $N_t^{(1)}$ and $N_t^{(2)}$ be Poisson processes both with intensity $\lambda_1 = \lambda_2 = 1$. Let*

$$S_t^{(j)} = \sum_{i=0}^{N_t^{(j)}} Y_i^{(j)}$$

be compound Poisson for j = 1,2 and let $S_t^{(1)}$ and $S_t^{(2)}$ be dependent according to a Clayton Lévy copula. Finally let $c = (1+\theta_1)\mu_1 \mathbb{E}[Y^{(1)}] + (1+\theta_2)\mu_2 \mathbb{E}[Y^{(2)}]$ where $\theta_1 = \theta_2 = 0.3$. Figure 9 shows the ruin probability for the following process:

$$X_t = u + ct - S_t^*$$

for different values of ω.

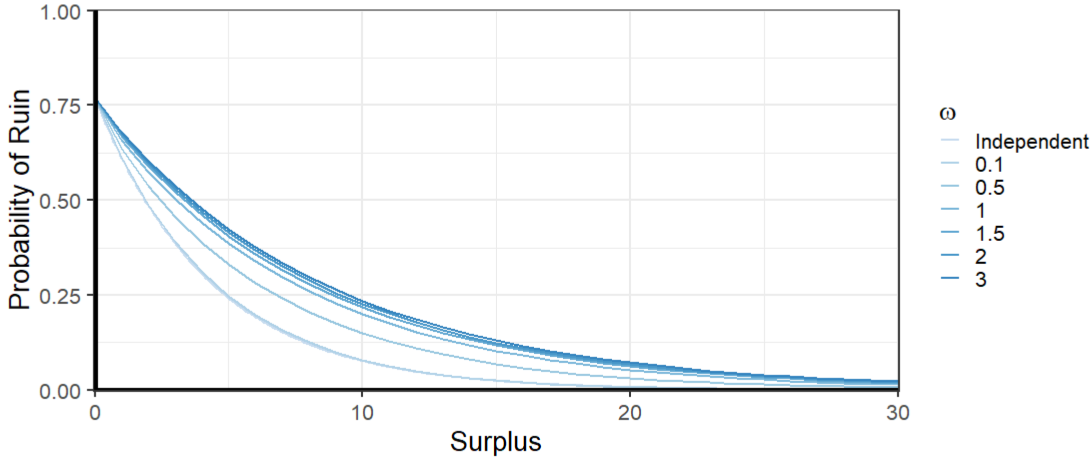

Figure 9: The figure shows ruin probabilities of the sum of two exponential claim processes for different levels of dependencies as a function of the current surplus. The higher the dependency (ω), the higher the probability of ruin.

Example 5.5 (Gamma claims)**.** *In this example we will consider the case when the claims are gamma distributed with shape parameter a and scale parameter k. The premium income c is defined as:*

$$c(\theta) = \sum_{j=1}^{2} \Big((1+\theta_j)\, \mathbb{E}[N_1^{(j)}] E[Y^{(j)}] - r_j \Big)$$

where the intensity of the two Poisson processes is a constant, $\mathbb{E}[N_1^{(1)}] = \lambda_1$ and $\mathbb{E}[N_1^{(2)}] = \lambda_2$. Using $a = (2,2)$, $k = (500,500)$, $r = (100,100)$, $\lambda = (800,800)$, and $\theta_1 = \theta_2 = 0.3$ figure 10 is produced.

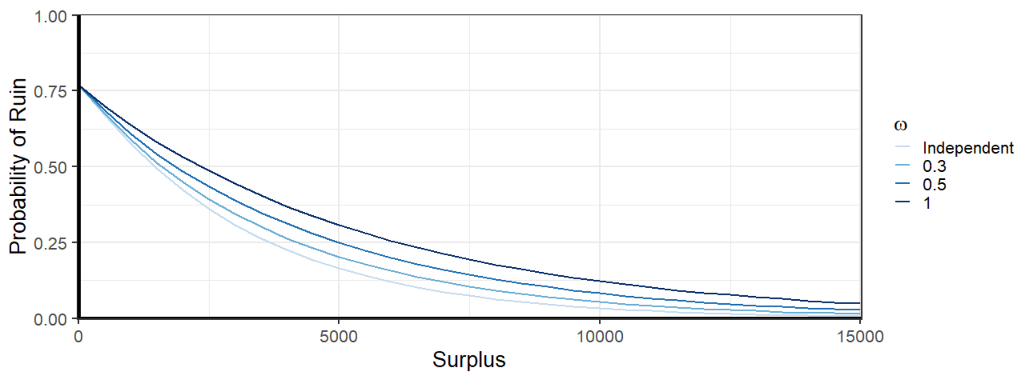

Figure 10: The figure shows ruin probabilities of the sum of two gamma claim processes for different levels of dependencies. The higher the dependency (ω), the higher the probability of ruin.

For both examples, Figures 9 and 10, show that the higher the dependency the higher the probability of ruin. Another interesting fact is that $V(0^+)$ looks to be the same for all values of ω. It is not exactly the same but very similar. The reason is that the formula for $V(0^+)$ is, as shown before:

$$V(0^+) = \frac{\mathbb{E}^{\theta}[N_1]\, \mathbb{E}[Y]}{c}.$$

If there would be no fixed cost the value would simply be $V(0^+) = \frac{1}{1+\theta}$, meaning that the value would be independent of all parameters and the severity distribution, but it would only depend on the loading parameter. When the fixed cost is introduced the value gets skewed and now depends on all the parameters. Nevertheless it is still very similar.

6 Applications to Insurance Pricing

In this chapter, the static loading parameter problem with infinite horizon will be considered in the case of two dependent surplus processes. It turns out that the results in Section 4.1 and Chapter 5 cannot be combined in a trivial way to obtain optimal loadings in the presence of dependencies because changes in exposure modify the dependence structure in the insurer's portfolio. In Section 6.1 we discuss this effect, and how to account for it under simplifying assumptions. The remaining sections in this chapter deal with a numerical example.

6.1 Models and Assumptions

Consider the following surplus processes for individuals risks:

$$X_t^{(j)} = u^{(j)} + c^{(j)}(\theta_j)t - \sum_{i=0}^{N_t^{(j)}} Y_i^{(j)} \tag{6.1}$$

where $j = 1, 2$. The intensity of $N_t^{(j)}$ depends on the loading parameter, θ_j. $c^{(j)}(\theta_j)$ is the premium income and is calculated according to the expected value principle $c^{(j)}(\theta) = (1 + \theta_j)\,\mathbb{E}[Y^{(j)}]\,\mathbb{E}^{\theta}[N_1^{(j)}] - r^j$ where $\mathbb{E}[Y^{(j)}]$ is the mean of the severity random variables $Y_i^{(j)}$, $\mathbb{E}^{\theta}[N_1^{(j)}]$ is the expected claim number and r^j is some cost.

As in previous sections let $\mathbb{E}^{\theta}[N_1^{(j)}] = \lambda_j p^{(j)}(\theta_j)$ where λ_j is the average number of potential claims per unit of time in the whole market for risk j, and $p^{(j)}(\theta)$ is the probability that a potential claim is filed as an actual claim to the particular insurer under consideration. Let $p^{(j)}(\theta_j)$ be a logit function of θ_j:

$$p^{(j)}(\theta_j) = \frac{1}{1 + e^{\beta_{1,j} + \beta_{2,j}\theta_j}}$$

where $\beta_{1,j} \in \mathbb{R}$ and $\beta_{2,j} > 0$.

The claim process for the insurance company and the whole market are not the same as the insurance company only has a subset of customers in the whole market. This fact has to be accounted for and some assumptions have to be made before the aggregation.

Assumption 1: Acquisition of insurance for risk 1 in our company is stochastically independent of acquisition of insurance for risk 2.

Assumption 2: All individuals in the market are identical, concerning the risks they carry.

Consider individuals who hold a policy for risk 1 but not for risk 2 and vice-versa, the proportion of such individuals in the insurer's portfolio is not independent of exposure, and this influences the distribution of the aggregate risk process, as we show bellow.

Let $p^{(1,0)}(\theta_1, \theta_2)$ denote the probability that a given individual holds a policy for risk 1 but not for risk 2, $p^{(0,2)}(\theta_1, \theta_2)$ denote the probability that a given individual holds a policy for risk 2 but not for risk 1 and $p^{(1,2)}(\theta_1, \theta_2)$ denote the probability that a given individual holds a policy for risk 1 and for risk 2.

Under Assumption 1, the probability that a given individual holds a policy for both risks is

$$p^{(1,2)}(\theta_1, \theta_2) = p^{(1)}(\theta_1)p^{(2)}(\theta_2).$$

The probability that a given individual holds a policy for risk 1 but not for risk 2 and the probability that a given individual holds a policy for risk 2 but not for risk 1 are, respectively

$$\begin{aligned} p^{(1,0)}(\theta_1, \theta_2) &= p^{(1)}(\theta_1) - p^{(1,2)}(\theta_1, \theta_2) = p^{(1)}\big(\theta_1)(1 - p^{(2)}(\theta_2)\big), \\ p^{(0,2)}(\theta_1, \theta_2) &= p^{(2)}(\theta_1) - p^{(1,2)}(\theta_1, \theta_2) = p^{(2)}(\theta_2)\big(1 - p^{(1)}(\theta_1)\big). \end{aligned}$$

Let $Z = Z^{(1)} + Z^{(2)}$ be the process of cumulative claims in the **market**. Following Section 5, Z can be decomposed into independent and common components

$$Z_t^{(j)} = \sum_{i=0}^{N_t^{j\perp}} Y_i^{j\perp} + \sum_{i=0}^{N_t^{\|}} Y_i^{j\|}, \quad j = 1, 2$$

where $N_t^{1\perp}$, $N_t^{2\perp}$, $N_t^{\|}$, $Y_j^{1\perp}$, $Y_j^{2\perp}$, $Y_j^{1\|}$, and $Y_j^{2\|}$ are distributed as in Section 5. In particular $N_t^{1\perp}$, $N_t^{2\perp}$, and $N_t^{\|}$ are independent Poisson processes with intensities $\lambda_1^{\perp}$, $\lambda_2^{\perp}$, and $\lambda^{\|}$

Under Assumptions 1 and 2, the **company's** surplus process can be written as

$$X_t = u^{(1)} + u^{(2)} + \big(c^{(1)}(\theta_1) + c^{(2)}(\theta_2)\big)t - \sum_{i=0}^{\tilde{N}_t^{1\perp}} \tilde{Y}_i^{1\perp} - \sum_{i=0}^{\tilde{N}_t^{2\perp}} \tilde{Y}_i^{1\perp} - \sum_{i=0}^{\tilde{N}_t^{\parallel}} \Big(Y_i^{1\parallel} + Y_i^{2\parallel}\Big). \tag{6.2}$$

Here, $\tilde{N}_t^{1\perp}$, $\tilde{N}_t^{2\perp}$, and $\tilde{N}_t^{\parallel}$ count the number of claims **received by the company** concerning only risk 1, only risk 2, and both risks, respectively. Their intensities, are, respectively

$$\begin{aligned}
\tilde{\lambda}_1^{\perp} &= p^{(1,0)}(\theta_1,\theta_2)\big(\lambda_1^{\perp} + \lambda^{\parallel}\big) + p^{(1,2)}(\theta_1,\theta_2)\lambda_1^{\perp} = p^{(1)}(\theta_1)\lambda_1^{\perp} + p^{(1,0)}(\theta_1,\theta_2)\lambda^{\parallel},\\
\tilde{\lambda}_2^{\perp} &= p^{(2)}(\theta_2)\lambda_2^{\perp} + p^{(0,2)}(\theta_1,\theta_2)\lambda^{\parallel},\\
\tilde{\lambda}^{\parallel} &= p^{(1,2)}(\theta_1,\theta_2)\lambda^{\parallel}.
\end{aligned}$$

From this, it can be seen that the distribution of the single risk claim amounts $\tilde{Y}^{1\perp}$ (resp., $\tilde{Y}^{2\perp}$) is a mixture of the distributions $Y^{1\perp}$ and $Y^{1\parallel}$ (resp., $Y^{2\perp}$ and $Y^{2\parallel}$) with weights:

$$\frac{p^{(1)}\lambda_1^{\perp}}{p^{(1)}\lambda_1^{\perp} + p^{(1,0)}\lambda^{\parallel}}, \frac{p^{(1,0)}\lambda^{\parallel}}{p^{(1)}\lambda_1^{\perp} + p^{(1,0)}\lambda^{\parallel}}, \quad \Big(resp., \frac{p^{(2)}\lambda_2^{\perp}}{p^{(2)}\lambda_2^{\perp} + p^{(0,2)}\lambda^{\parallel}}, \frac{p^{(0,2)}\lambda^{\parallel}}{p^{(2)}\lambda_1^{\perp} + p^{(0,2)}\lambda^{\parallel}}\Big).$$

This is because, for example, that some customers only have risk 1 but risk 1 can jump together with risk 2. Therefore, the aggregate process (6.2) coincides in distribution with a process

$$X_t^{indp} = u^{(1)} + u^{(2)} + \big(c^{(1)}(\theta_1) + c^{(2)}(\theta_2)\big)t - \sum_{i=0}^{\tilde{N}_t} \tilde{Y}_i \tag{6.3}$$

where $\tilde{N}_t$ is a Poisson process with intensity

$$\tilde{\lambda} = p^{(1)}\lambda_1^{\perp} + p^{(2)}\lambda_2^{\perp} + \big(p^{(1,0)} + p^{(0,2)} + p^{(1,2)}\big)\lambda^{\parallel}$$

and $\tilde{Y}_i$, $i \in \mathbb{N}$ are i.i.d random variables with distribution

$$F_{\tilde{Y}} = \frac{p^{(1)}\lambda_1^{\perp}}{\tilde{\lambda}} F_{Y^{1\perp}} + \frac{p^{(2)}\lambda_2^{\perp}}{\tilde{\lambda}} F_{Y^{2\perp}} + \frac{p^{(1,0)}\lambda^{\parallel}}{\tilde{\lambda}} F_{Y^{1\parallel}} + \frac{p^{(0,2)}\lambda^{\parallel}}{\tilde{\lambda}} F_{Y^{2\parallel}} + \frac{p^{(1,2)}\lambda^{\parallel}}{\tilde{\lambda}} F_{Y^{1\parallel}+Y^{2\parallel}}.$$

Notice that the argument above includes the independent case, with $\lambda^{\parallel} = 0$. Also, looking at the term $p^{(1,2)}$ in the previous equation it can be seen that for small insurance companies where the acquisition of customers is more or less random $p^{(1,2)}$ is small. This means that the claims act as if they are independent even though the market is highly correlated. As the company grows the dependency grows nonlinearly.

From the above, it follows that equation (4.2) for the aggregate process takes the form

$$\frac{d}{dx}V(x,\theta_1,\theta_2) = \frac{\mathbb{E}^{\theta_1,\theta_2}[\tilde{N}_1]}{c^{(1)}(\theta_1) + c^{(2)}(\theta_2)}\Big(V(x,\theta_1,\theta_2) + \int_0^x V(x-y,\theta_1,\theta_2)dF_{\tilde{Y}}(y) + F_{\tilde{Y}}(x) - 1\Big).$$

Thus, the argument used in Section 4.1 cannot be used to deduce that the loadings minimizing the ruin probability can be found simply by minimizing $\alpha(\theta_1,\theta_2) = \frac{\mathbb{E}^{\theta_1,\theta_2}[\tilde{N}_1]}{c^1(\theta_1)+c^2(\theta_2)}$.

Similarly, the expected aggregate profit becomes

$$\mathbb{E}[X_1|X_0 = 0] = c^{(1)}(\theta_1) + c^{(2)}(\theta_2) - \mathbb{E}^{\theta_1,\theta_2}[\tilde{N}_1]\,\mathbb{E}^{\theta_1,\theta_2}[\tilde{Y}].$$

6.2 Numerical Results

Throughout this subsection, Y_i are assumed to be i.i.d gamma distributed random variables with shape parameter, a, and scale parameter, k. Therefore, the mean is, $\mathbb{E}[Y] = ak$. The constants used are $a = (2,2)$, $k = (500,500)$, $\lambda = (800,800)$, $\beta_1 = (-0.5,-0.5)$, $\beta_2 = (4,4.5)$ and r is taken to be 20% of the pure premium if the exposure was 40%, that is $r = 0.4 * 0.2kaN$. The operational cost is therefore 8% of of the expected total amount of claims in the market. The only difference in the two processes lies in β_2, that is, surplus process 2 is more sensitive to the security loading parameter. The dependence parameter of the Clayton Lévy copula is set to $\omega = 0.5$. The programming language R was used for every calculation.

6.2.1 Single Surplus Process

The surplus processes are first considered separately according to equation (6.1). The ruin probability and the expected profit is graphed as a function of θ for the two processes. θ^*_{ruin} was found by minimizing α.

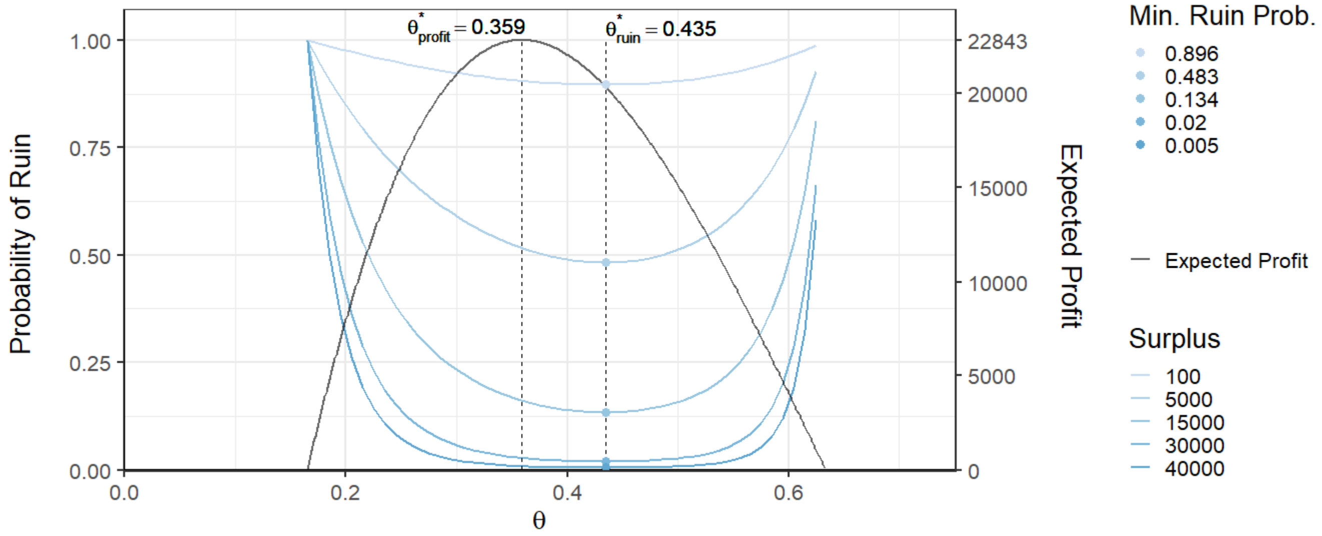

Figure 11: Surplus process 1. The blue lines show the ruin probability as a function of θ for a given surplus x. The black line shows the expected profit per time unit as a function of θ. The blue dots show the minimum ruin probability for each surplus. θ^*_{profit} and θ^*_{ruin} denote the optimal security loading parameter for the expected profit and for the probability of ruin, respectively.

From figure 11 it can be seen that the optimal security loading parameter for the ruin probability is, $\theta^*_{ruin} = 0.435$, while the θ that maximizes the expected profit is lower or $\theta^*_{profit} = 0.359$. Moreover, in this example, the maximum expected profit is 22.843 units and is given at θ^*_{profit} units. The expected profit taken at the point θ^*_{ruin} gives a lower expected profit.

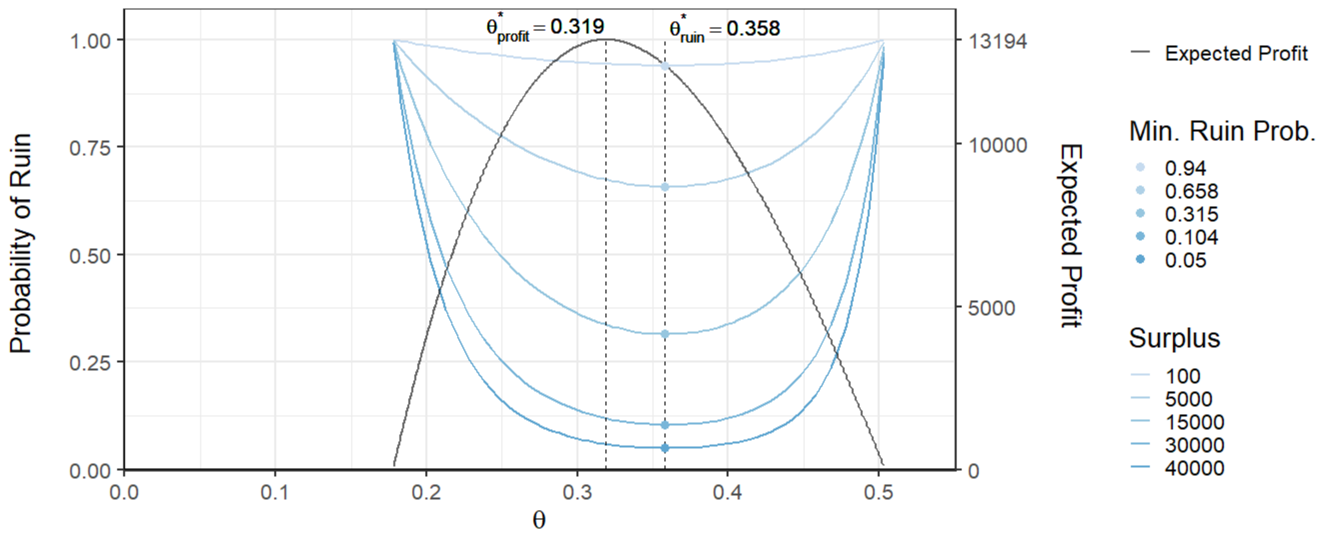

Figure 12: Surplus process 2. The blue lines show the ruin probability as a function of θ for a given surplus x. The black line shows the expected profit per time unit as a function of θ. The blue dots show the minimum ruin probability for each surplus. θ^*_{profit} and θ^*_{ruin} denote the optimal security loading parameter for the expected profit and for the probability of ruin, respectively.

From figure 12 it can be seen that the optimal security loading parameter for the ruin probability is $\theta^*_{ruin} = 0.358$, while the θ that maximizes the expected profit is again lower or $\theta^*_{profit} = 0.319$.

Obviously, for both processes, the ruin probability decreases with increasing surplus. Moreover, it can be seen that surplus process X_2 has higher probability of ruin than surplus process X_1 for the same amount of surplus. The sensitivity of the demand curve affects the ruin probability and θ^*_{ruin} greatly. The more sensitive to the exposure the demand curve is, the closer the θ^*_{profit} and θ^*_{ruin} are. A more sensitive curve also has higher probability of ruin for a given surplus, which indicates that more competitive insurance products are riskier. These effects can be seen if the two figures (11 and 12) are compared. Conversely, if the demand curve is not sensitive to the price, then the gap between θ^*_{profit} and θ^*_{ruin} can become quite large. Additionally, it can be seen from the curve at surplus = 100 that the ruin probability for θ^*_{profit} and θ^*_{ruin} are similar but as the surplus grows the values start to differ and once the surplus is great enough the two θ values of θ result in similar ruin probabilities again. Meaning that if the insurance firm has high surplus then they can choose arbitrary θ without risking the chance of ruin. If the surplus is great enough then the value of θ does not matter as much. However, having too much reserves can be bad for insurance companies as it can be seen as a negative leverage. The bowl shape of the blue curves is because of the interplay between the fixed cost and the demand curve.

θ^*_{ruin} should give the minimum ruin probability at all surplus values. This can be tested by graphing multiple ruin probability curves and compare it with the one obtained by θ^*_{ruin}. Figure 13 shows that θ^*_{ruin} gives the minimum ruin probability indeed.

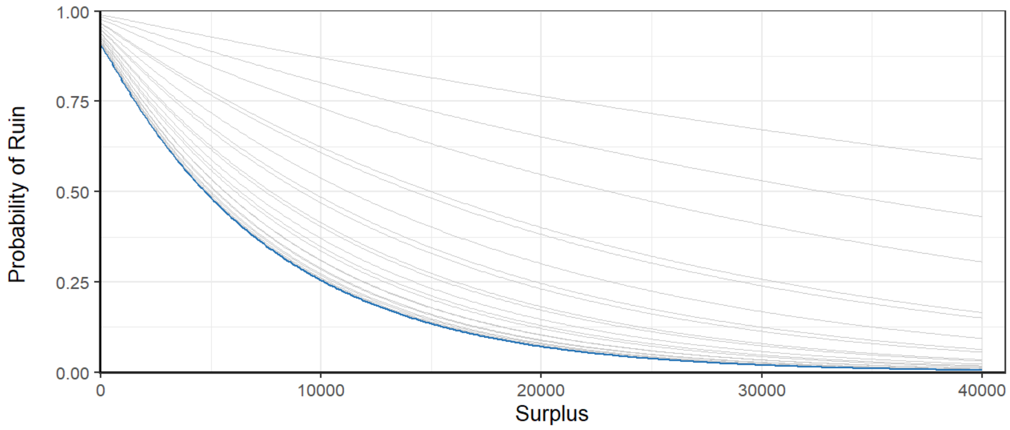

Figure 13: The figure compares the optimal value function of surplus process, X_1 (blue line) to other cost functions (grey lines). The blue line is achieved by setting $\theta = \theta^*_{ruin}$. All the cost functions lie above the optimal value function, as is expected.

6.2.2 Two Aggregated Surplus Processes with Common Loading

Next, the two surplus processes, X_1 and X_2 are aggregated, as explained in the beginning of this section. Both independent (Process (6.3)) and dependent cases are analyzed. Dependence is considered via a Clayton Lévy copula (Process (6.2)).

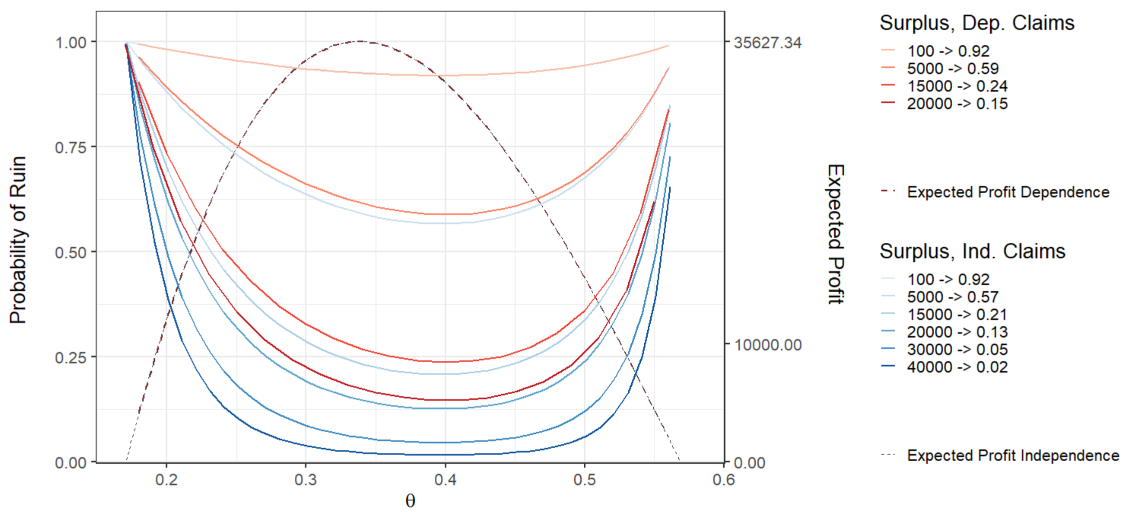

Figure 14: Ruin probability when X_1 and X_2 are aggregated as a function of the security loading parameter, θ, both when they are independent and dependent via Clayton Lévy copula with $\omega = 0.5$. The blue curves show the ruin probability when the two processes are independent for different values of the surplus and the red curves show the same for the dependent case. The curves have similar shapes, but the ruin probability is higher in the case of dependence, for the same surplus. The values in the legend show the minimum ruin probability for a given surplus (surplus $\rightarrow$ probability).

Figure 14 shows the ruin probability of the aggregated surplus process as a function of the security loading parameter, θ, both when they are independent and dependent via Clayton Lévy copula. The red curves represent dependence while the blue curves represent independence.

Firstly, it can be seen that the expected profit is the same for dependence and independence and from the figure, $\theta^*_{profit} \approx 0.34$. The reason is that the claim mean and the claim frequency is almost the same (numerically) for dependence and independence.

Secondly, the dependent case has a higher probability of ruin than the independent case for the same amount of surplus. That is not surprising as figures 9 and 10 gave the same result. However, the ruin probability is almost the same for small surplus values as can be seen from the figure. Interestingly, the optimal loading for dependence and independence seem to be the same and numerically the values are $\theta^*_{ruin,dep} = 0.4 = \theta^*_{ruin,indp}$. The surplus value does not change the optimal loading θ^*, as expected. When Figure 14 is compared to Figure 10 it can be noted that the ruin probability difference is greater in the latter figure, showing that the risk in the market is greater. The reason why the ruin probability difference is relatively small is because of the probability $p^{(1,2)}(\theta)$. The fact that the insurance company does not always have the both claims $Y^{1\|}$ and $Y^{1\|}$ when a common jump occurs reduces the risk.

Finally, the difference of the two curves (red and blue) for a given surplus seems to be increasing with increasing surplus, meaning that the ruin probability in the independent case decreases more rapidly with increasing surplus then for the dependent case. Therefore, it is clear that the dependent case is riskier. The same effects can be observed in Figure 10.

The optimal security loading parameter is around 0.4, which is very close to the weighted average of the optimal loading parameter of the isolated surplus processes where the weight is the exposure ratio of each surplus process, that is

$$\theta_{weighted} = \frac{0.435\frac{1}{1+\exp(-0.6+4*0.4)} + 0.358\frac{1}{1+\exp(-0.6+4.5*0.4)}}{\frac{1}{1+\exp(-0.6+4*0.4)} + \frac{1}{1+\exp(-0.6+4.5*0.4)}} \approx 0.4$$

which strongly indicates that the optimal value, θ^*_{ruin}, is simply the weighted average.

6.2.3 Two Aggregated Surplus Processes with Separate Loadings

It is more realistic to consider θ as a vector so that the loading parameter can be changed for each surplus process separately to spread the total premium over the policies in an optimal way. The two surplus processes, X_1 and X_2 are aggregated as before and the constants are the same but let $\theta = (\theta_1, \theta_2)$.

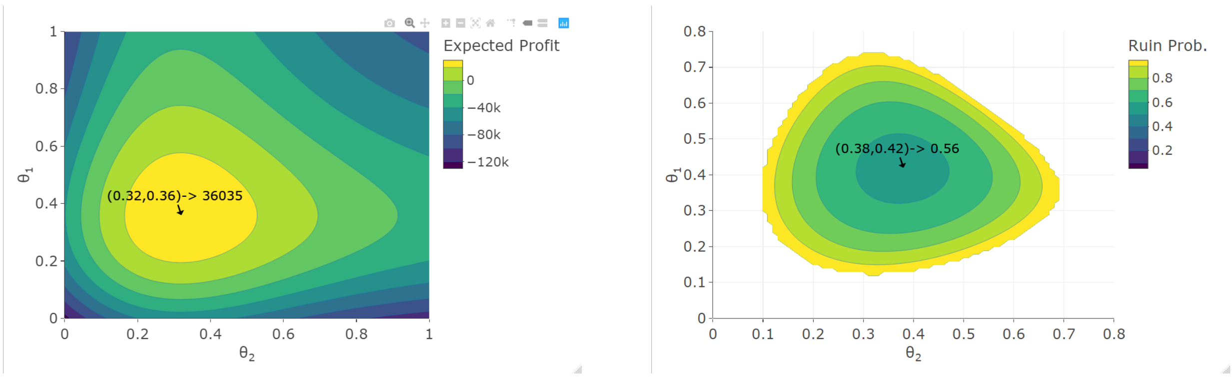

Figure 15: Expected profit (left) and the ruin probability (right) when X_1 and X_2 are aggregated, as a function of the security loading parameters, θ_1 and θ_2. The processes are assumed to be independent and the surplus is fixed at $x = 5000$. The parenthesis in the right figure shows the optimal values of θ_1 and θ_2 with the ruin probability as a criterion. The arrow shows which values θ_1 and θ_2 are mapped into, thus showing the minimum ruin probability. The parenthesis in the left figure shows the same for the expected profit. The shape of the contour plot is due to the fact that the θ grid considered is sparser for values that give high ruin probability

Figure 15 shows the expected profit (left) and the ruin probability (right), when X_1 and X_2 are assumed to be independent and aggregated, as a function of the security loading parameters, θ_1 and θ_2. The surplus is fixed at $x = 5000$ and the optimal values are shown in the figure. It should be noted that many surplus values were tested and they all gave the same value for θ_1 and θ_2, only the ruin probability level changed. Note that the optimal loading parameters for the expected profit are the same as those for the individual surplus processes. However, the optimal loading parameters for the ruin probability change when compared to the individual one (compare it with Figures 11 and 12). When compared to the optimal loading parameter for the individual surplus process, θ_1 decreases from 0.435 to 0.42 and θ_2 increases from 0.358 to 0.38. Therefore, the optimal security loading parameter decision is to decrease the loading parameter of the less sensitive surplus process while increasing the loading parameter of the more sensitive surplus process. Additionally, when compared to Figure 14, the minimum ruin probability for one shared loading is 0.57 while the ruin for two loadings is 0.56, showing only a marginal difference. When the same is done for other surplus values a similar difference is found. The expected profit is marginally higher.

Lastly, consider the case when the surplus processes are assumed to be dependent via Lévy Clayton copula and that the loadings can be changed for each surplus process separately.

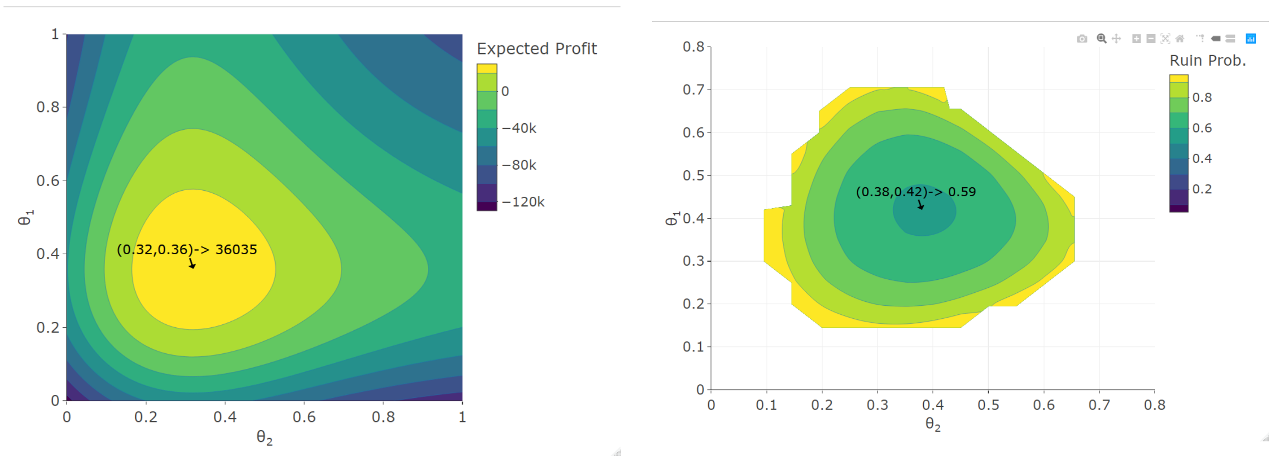

Figure 16: Expected profit (left) and the ruin probability (right) when X_1 and X_2 are aggregated, as a function of the security loading parameters, θ_1 and θ_2. The processes are assumed to be dependent via Clayton Lévy copula and the surplus is fixed at $x = 5000$. The parenthesis in the right figure shows the optimal values of θ_1 and θ_2 with the ruin probability as a criterion. The arrow shows which values θ_1 and θ_2 are mapped into, thus showing the minimum ruin probability. The parenthesis in the left figure shows the same except for the expected profit. The shape of the contour plot is due to the fact that the θ grid considered is sparser for values that give high ruin probability

Figure 16 shows the ruin probability when X_1 and X_2 are aggregated as a function of the security loading parameters, θ_1 and θ_2. The shape of the contour plots is due to the fact that the θ grid considered is sparser for values that give high ruin probability. The processes are assumed to be dependent via Clayton Lévy copula and the surplus is fixed at $x = 5000$.

It can be seen that the optimal loadings θ_1 and θ_2 are the same as the ones in the case of independence and the minimum ruin probability is higher (compared to Figure 15). The expected profit is the same as the independent case, both the value and the loadings. Again, the optimal security loading parameter decision is to decrease the loading parameter of the less sensitive surplus process while increasing the loading parameter of the more sensitive surplus process. The difference between the ruin probability in Figure 16 vs Figure 15 is only 0.03 still the surplus is low compared to the expected profit, if the surplus would be increased to ≈ 20.000 the difference would become greater. The difference would then decrease again if the surplus were increases to ≈ 40.000.

Additionally, when compared to Figure 14, the minimum ruin probability for one common loading is 0.59, which is the same as the ruin probability for separate loading selections, therefore the difference is only marginal.

7 Conclusion and Future Research

One of the objectives in this thesis was to determine the security loading parameter in an intelligent manner. The ruin probability was chosen as a criterion and the expected premium principle used. In this thesis, methods that can be used to calculate the infinite and finite time ruin probability of a claim process (numerically) were obtained. It was determined that the infinite time ruin probability is increasing with the ratio of intensity and premium income. The dynamic programming principle was then used in order to find the optimal loading parameter for periodic updates of the premium. Lévy Copulas proved to be a useful tool when modelling dependence of jump processes. Bivariate jump processes were only considered but the theory can easily be extended to higher dimensions. It was shown that if the loadings of the single claim process and the loadings of the bivariate claim process are compared, then the optimal decision is to increase the loading of the more sensitive process while decreasing the other. The optimal loadings for the dependent risk process are not the same as those for the independent risks. Finally, it could be seen that the dependency constant increases the ruin probability of the claim process, indicating that a portfolio of dependent claims is riskier than a portfolio of independent claims, for the same surplus.

The model considered in this thesis is very simple and a lot of modifications can be done. For example, claim severity and premium income discounting can be added. There is also the possibility to add lags to the claims as claims are usually not paid immediately. This is done in loss reserving and might be an interesting dynamic to add. The ruin probability would most likely decrease as the claims start to lag while the premium income stays the same. Premium income is discrete so a discrete premium income in the claim process could be considered. However, if the premium income is discrete, then the claim process is no longer stochastically continuous, which will make the analysis more difficult because the infinitesimal generator is only defined for stochastically dependent processes. One would have to break the ruin probability into parts where a finite ruin probability is considered between premium incomes. A more realistic model would consider investments and reinsurance. The problem would then be about assessing the impact of reinsurance and determine the role of investments in minimizing the ruin probability and to determine the optimal reinsurance and/or the security loading. Finally, it would be interesting to consider dependency in the acquisition of different policies for a customer. This dependency structure could be characterized with ordinary copulas. Considering another Lévy copula between the claim processes would of course be interesting as well. Another possible idea is to implement reinforcement learning (RL). RL iteratively interacts with a simulation of the insurance model and selects actions that minimizes the insurer's ruin according to the feedback from the data.

It can be concluded that ruin probabilities can be used as a risk management tool namely in determining the loading parameter of premia and that Lévy copulas provide a flexible and elegant modelling of dependent jump processes allowing insurance companies to take dependencies into account.

A Numerical Schemes

A.1 Numerical scheme for equation 3.4, using linear approximation

Consider the process

$$X_t = u + ct - \sum_{i=0}^{N_t} Y_i$$

where Y_i are iid absolutely continuous random variables with distribution $F(x)$ and N_t is a $Poisson(\lambda t)$. To approximate equation 3.4, take a grid of points $\epsilon = x_0 < x_1 < ...x_n$, $x_i \in \mathbb{R}, \forall i \in \mathbb{N}, \epsilon > 0$, with equal interval lengths, $h = x_i - x_{i-1}$. A linear approximation is used to approximate $\bar{V}(x)$

$$\bar{V}(z) = \bar{V}(x_{j-1}) + \frac{\bar{V}(x_j) - \bar{V}(x_{j-1})}{h}(z - x_{j-1}), \quad z \in [x_{j-1}, x_j], j \leq i$$

where $\frac{\bar{V}(x_j)-\bar{V}(x_{j-1})}{h}$ is an approximation of the derivative, $\bar{V}'(x_{j-1})$, using the so-called forward difference
Let $\bar{V}_i$ denote the approximation of $\bar{V}(x_i)$. Let $\bar{S}(x) = \int_0^x \bar{F}(y)dy$ and $\bar{\bar{S}}(x) = \int_0^x \bar{S}(y)dy$. For each $x_i, i > 0$ solve the following equation

$$\bar{V}(x_i) - \bar{V}(0^+) = \frac{\lambda}{c}\int_0^x \bar{V}(x_i - y)\bar{F}(y)dy$$

The goal is to develop a recursive method from x_0 as the value of $\bar{V}_0$ is known.

if $\boldsymbol{i = 0}$

Set $\bar{V}_0 = 1 - \frac{\lambda}{c}\,\mathbb{E}[Y]$

if $\boldsymbol{i = 1}$

Calculate

$$\begin{aligned}
\bar{V}_1 &= \bar{V}_0 + \frac{\lambda}{c}\int_{x_0}^{x_1}((\bar{V}_0 + \frac{\bar{V}_1 - \bar{V}_0}{h})(x_1 - y - x_0))\bar{F}(y)dy \\
&= \bar{V}_0 + \frac{\lambda}{c}\Big(\bar{V}_0(\bar{S}(x_1) - \bar{S}(x_0)) + \\
&\quad \frac{\bar{V}_1 - \bar{V}_0}{h}\big([x_1 - y - x_0]_{x_0}^{x_1} + \bar{\bar{S}}(x_1) - \bar{\bar{S}}(x_0)\big)\Big) \\
&= \bar{V}_0 + \frac{\lambda}{c}(a_{1,1} + \frac{\bar{V}_1 - \bar{V}_0}{h}a_{2,1})
\end{aligned}$$

$$\Leftrightarrow$$

$$(1 - \frac{\lambda a_{2,1}}{ch})\bar{V}_1 = \bar{V}_0 + \frac{\lambda}{c}\bar{V}_0(a_{1,1} - \frac{a_{2,1}}{h})$$

if $\boldsymbol{i > 1}$

Calculate

$$\begin{aligned}
\bar{V}(x_i) - \bar{V}(0) &= \frac{\lambda}{c}\Big(\sum_{j=2}^{i}\int_{x_{j-1}}^{x_j}\bar{V}(x_i - y)S(y)dy + \int_{x_0}^{x_1}\bar{V}(x_i - y)S(y)dy\Big) \\
&= \frac{\lambda}{c}\Bigg(\sum_{j=2}^{i}\Big(\bar{V}_{i-j}(\bar{S}(x_j) - \bar{S}(x_{j-1})) + \frac{\bar{V}_{i-j+1} - \bar{V}_{i-j}}{h}\big([x_i - y - x_{i-j}]_{x_{j-1}}^{x_j} + \bar{\bar{S}}(x_j) - \bar{\bar{S}}(x_{j-1})\big)\Big) + \\
&\quad \bar{V}_{i-1}(\bar{S}(x_1) - \bar{S}(x_0)) + \frac{\bar{V}_i - \bar{V}_{i-1}}{h}\big([x_i - y - x_{i-1}]_{x_0}^{x_1} + \bar{\bar{S}}(x_1) - \bar{\bar{S}}(x_0)\big)\Bigg) \\
&= \frac{\lambda}{c}\Bigg(\sum_{j=2}^{i}\Big(\bar{V}_{i-j}(\bar{S}(x_j) - \bar{S}(x_{j-1})) + \frac{\bar{V}_{i-j+1} - \bar{V}_{i-j}}{h}\big([x_i - y - x_{i-j}]_{x_{j-1}}^{x_j} + \bar{\bar{S}}(x_j) - \bar{\bar{S}}(x_{j-1})\big)\Big) + \\
&\quad \bar{V}_{i-1}a_{1,i} + \frac{\bar{V}_i - \bar{V}_{i-1}}{h}a_{2,i}\Bigg)
\end{aligned}$$

$$\Leftrightarrow$$

$$(1-\frac{\lambda a_{2,i}}{ch})\bar{V}_i = \bar{V}_0 + \frac{\lambda}{c}\Big(\sum_{j=2}^{i}\Big(\bar{V}_{i-j}(\bar{S}(x_j)-\bar{S}(x_{j-1})) + \frac{\bar{V}_{i-j+1}-\bar{V}_{i-j}}{h}\big([x_i - y - x_{i-j}]_{x_{j-1}}^{x_j} + \bar{\bar{S}}(x_j) - \bar{\bar{S}}(x_{j-1})\big)\Big) + \bar{V}_{i-1}a_{1,i} - \frac{\bar{V}_{i-1}}{h}a_{2,i}\Big)$$

: Estimation of $\bar{V}(x)$

Let the symbol $\vec{}$ denote a vector.

Initialize
$\vec{x}$ for some $x_0, ..., x_n$
$\vec{\bar{V}}$ with length equal to the length of $\vec{x}$
$N_x \leftarrow$ length of $\vec{x}$

 # loop to estimate each value in $\vec{\bar{V}}$
for i *in* $0, ..., (N_x - 1)$ **do**
  if $i = 0$ **then**
    $\bar{V}_0 \leftarrow 1 - \frac{\lambda}{c}\mathbb{E}[Y]$
  else if $i = 1$ **then**
    $\bar{V}_1 \leftarrow$ take case $i = 1$ from above and isolate $\bar{V}_1$
  else
    $\bar{V}_i \leftarrow$ take case $i > 1$ from above and isolate $\bar{V}_i$
  end
end
Return $\vec{\bar{V}}$, $\vec{x}$

A.2 Utilize the Probability Distributions and the Discretization of the Lundberg Process

Consider the process

$$X_t = u + ct - \sum_{i=0}^{N_t} Y_i$$

where Y_i are iid continuous random variables and N_t is a $Poisson(\lambda t)$. To approximate equation 3.12, take $t \in [0, T]$ and fix $K, M \in \mathbb{N}$ large enough. Next, create a grid of points such that $t_i = i\frac{T}{K}$, $i = 0, 1, ..., K$ and $x_j = j\frac{cT}{K}$, $j = 0, 1, ..., M$. W denotes the random variable of the arrival time of the next claim. Let $V_{i,j}$ denote the approximation of $V(t_i, x_j)$.

$i <= K - 1$ and $j < M$

$$\begin{aligned}
V_{i,j} &= \mathbb{P}\big(W > \frac{T}{K}\big)V_{i+1,j+1} + \\
&\quad \mathbb{P}\big(W \le \frac{T}{K}\big)\Big(\mathbb{P}\big(Y \le \frac{cT}{K}\big)V_{i+1,j} + \\
&\quad \mathbb{P}\big(\frac{cT}{K} < Y \le 2\frac{cT}{K}\big)V_{i+1,j-1} + \\
&\quad \mathbb{P}\big(2\frac{cT}{K} < Y \le 3\frac{cT}{K}\big)V_{i+1,j-2} + ... + \\
&\quad \mathbb{P}\big(j\frac{cT}{K} < Y \le (j+1)\frac{cT}{K}\big)V_{i+1,0} + \mathbb{P}\big(Y > (j+1)\frac{cT}{K}\big)\Big) \\
&= \mathbb{P}\big(W > \frac{T}{K}\big)V_{i+1,j+1} + \\
&\quad \mathbb{P}\big(W \le \frac{T}{K}\big)\Big(\sum_{r=0}^{j}\mathbb{P}\big(r\frac{cT}{K} < Y \le (r+1)\frac{cT}{K}\big)V_{i+1,j-r} + \mathbb{P}\big(Y > (j+1)\frac{cT}{K}\big)\Big)
\end{aligned}$$

$i \le K - 1$ and $j = M$

The algorithm uses future values with respect to space so we need to give it an end condition. Because the compound Poisson process has finite variation and if M is large enough then set $V_{i+1,M+1} = 0$ in the equation (as this index does not exist in the matrix). The equation becomes

$$V_{i,M} = \mathbb{P}\left(W \leq \frac{T}{K}\right)\Big(\sum_{r=0}^{j} \mathbb{P}\left(r\frac{cT}{K} < Y \leq (r+1)\frac{cT}{K}\right)V_{i+1,j-r} + \mathbb{P}\left(Y > (j+1)\frac{cT}{K}\right)\Big)$$

: Estimation of $\bar{V_T}(T, x)$

Let the symbol $\vec{}$ denote a vector.

Initialize
T as finite time horizon
K gives final time value t_K
$\vec{t} \leftarrow [0, ..., K]\frac{T}{K}$
M gives final surplus value x_M, high enough such that $V_T(t, x_M) \approx 0, \forall t \in [0, T[$ $\vec{\vec{V}}$ with length equal to the length of $\vec{x}$
$\vec{x} \leftarrow [0, ..., M]\frac{cT}{K}$ $\boldsymbol{V}$ a matrix of size $(K+1) \times (M+1)$
$V_{K,:} \leftarrow \vec{0}$, (all columns in row K)

 # loop to estimate each value in $\vec{\vec{V}}$
for i *in* $K-1, ..., 0$ **do**
  for j *in* $0, ..., M$ **do**
    if $j < M$ **then**
      $V_{i,j} \leftarrow \mathbb{P}\left(W > \frac{T}{K}\right)V_{i+1,j+1} + \mathbb{P}\left(W \leq \frac{T}{K}\right)\Big(\sum_{r=0}^{j} \mathbb{P}\left(r\frac{cT}{K} < Y \leq (r+1)\frac{cT}{K}\right)V_{i+1,j-r} + \mathbb{P}\left(Y > (j+1)\frac{cT}{K}\right)\Big)$
    else
      $V_{i,M} \leftarrow \mathbb{P}\left(W \leq \frac{T}{K}\right)\Big(\sum_{r=0}^{j} \mathbb{P}\left(r\frac{cT}{K} < Y \leq (r+1)\frac{cT}{K}\right)V_{i+1,j-r} + \mathbb{P}\left(Y > (j+1)\frac{cT}{K}\right)\Big)$
    end
  end
end
Return $\vec{\vec{V}}$, $\vec{x}$

References

[1] M. V. Wuthrich, "Non-life insurance: mathematics & statistics," *Available at SSRN 2319328*, 2019.

[2] R. Kaas, M. Goovaerts, J. Dhaene, and M. Denuit, *Modern actuarial risk theory: using R*, vol. 128. Springer Science & Business Media, 2008.

[3] S. A. Klugman, H. H. Panjer, and G. E. Willmot, *Loss models: from data to decisions*, vol. 715. John Wiley & Sons, 2012.

[4] C. Hipp and M. Plum, "Optimal investment for insurers," *Insurance: Mathematics and Economics*, vol. 27, no. 2, pp. 215–228, 2000.

[5] E. Ohlsson and B. Johansson, *Non-life insurance pricing with generalized linear models*, vol. 2. Springer, 2010.

[6] M. Denuit, X. Maréchal, S. Pitrebois, and J.-F. Walhin, *Actuarial modelling of claim counts: Risk classification, credibility and bonus-malus systems.* John Wiley & Sons, 2007.

[7] Y. Kahane, "The theory of insurance risk premiums—a re-examination in the light of recent developments in capital market theory," *ASTIN Bulletin: The Journal of the IAA*, vol. 10, no. 2, pp. 223–239, 1979.

[8] H. Bühlmann, "Premium calculation from top down," *Astin Bulletin - ASTIN BULL*, vol. 15, pp. 89–101, 11 1985.

[9] C. Kasumo, J. Kasozi, and D. Kuznetsov, "On minimizing the ultimate ruin probability of an insurer by reinsurance," *Journal of Applied Mathematics*, vol. 2018, 2018.

[10] J. Trufin, H. Albrecher, and M. M. Denuit, "Properties of a risk measure derived from ruin theory," *The Geneva Risk and Insurance Review*, vol. 36, no. 2, pp. 174–188, 2011.

[11] C. Hipp, "Stochastic control with application in insurance," in *Stochastic methods in finance*, pp. 127–164, Springer, 2004.

[12] H. J. Kappen, "An introduction to stochastic control theory, path integrals and reinforcement learning," in *AIP conference proceedings*, vol. 887, pp. 149–181, American Institute of Physics, 2007.

[13] A. Campana, P. Ferretti, *et al.*, "Distortion risk measures and discrete risks," tech. rep., 2005.

[14] F. Barning, "Counting processes and copulas: Applications in in-surance," Master's thesis, Université du Québec à Montréal, 2018.

[15] S. Britt and A. Napoli, "Linear correlation as a measure of dependency," in *XVth General Insurance Seminar, Institute of Actuaries of Australia*, 2005.

[16] D. Y. Straumann, "Correlation and dependency in risk management: properties and pitfalls," in *Risk Management: Value at Risk and Beyond*, Cambridge University Press, 2001.

[17] R. B. Nelsen, *An introduction to copulas.* Springer Science & Business Media, 2007.

[18] R. Cont and P. Tankov, "Financial modelling with jump processes, chapman & hall/crc financ," *Math. Ser*, 2004.

[19] P. Tankov, "Lévy copulas: review of recent results," in *The fascination of probability, statistics and their applications*, pp. 127–151, Springer, 2016.

[20] A. Papapantoleon, "An introduction to lévy processes with applications in finance," *arXiv preprint arXiv:0804.0482*, 2008.

[21] K.-i. Sato, S. Ken-Iti, and A. Katok, *Lévy processes and infinitely divisible distributions.* Cambridge university press, 1999.

[22] J. Grandell, *Aspects of risk theory.* Springer Science & Business Media, 2012.

[23] P. Hardin and S. Tabari, "Modelling non-life insurance policyholder price sensitivity: A statistical analysis performed with logistic regression," 2017.

[24] Y. Zhang and N. Walton, "Adaptive pricing in insurance: Generalized linear models and gaussian process regression approaches," *arXiv preprint arXiv:1907.05381*, 2019.

[25] F. E. Benth, V. A. Kholodnyi, and P. Laurence, “Quantitative energy finance,” in *Modelling, pricing, and hedging in energy and commodity markets*, Springer, 2014.

[26] P. C. Kettler, “Lévy-copula-driven financial processes,” *Preprint series. Pure mathematics http://urn. nb. no/URN: NBN: no-8076*, 2006.

[27] J. van Velsen, “Parameter estimation of a levy copula of a discretely observed bivariate compound poisson process with an application to operational risk modelling,” *arXiv preprint arXiv:1212.0092*, 2012.