BONUS-MALUS IN INSURANCE PORTFOLIOS

Weihong Ni

Supervisor: Dr. Corina Constantinescu

Department of Mathematical Sciences

University of Liverpool

A thesis submitted for the degree of

Doctor of Philosophy

Nov 1^{st} 2015

I would like to dedicate this thesis to my loving parents, kind supervisor, dear boyfriend and a group of lovely friends.

Acknowledgements

The completion of this work could not have been possible without the sponsorship by the Institute for Financial and Actuarial Mathematics (IFAM), University of Liverpool. Besides, strong support from the RARE -318984 project, a Marie Curie IRSES Fellowship within the 7th European Community Framework Programme has facilitated successful research exchange visits between several partner universities. Some work within this thesis was conducted via this programme. I would also like to acknowledge support from the knowledge exchange PRACTICE scheme of the University of Liverpool, a Higher Education Innovation Funding (HEIF) 2011-12 to 2014-15. In addition, I am in-debited to many people for their guidance and assistance for this research.

I would particularly like to express my deepest sense of gratitude to my supervisor **Corina Constantinescu** for her overwhelming attitude towards firm support and encouragement on her students. Her vision and foresight within and beyond the area of research, patient and clear explanations as well as timely advice has always been a great inspiration on me accomplishing this work.

I owe a debt of gratefulness to **Zbigniew Palmowski**, **Enkelejd Hashorva**, **Dominik Kortschak** for their professional view and important suggestions. I am also extremely thankful to **Junyi Guo**, **Kais Hamza**, **Andrea Collevecchio**, **Jiro Akahori** and **Yuri Imamura** for their hospitality and friendly support during my visits to partner universities. Special thanks are sincerely passed onto **Bo Li**, **Suhang Dai** and **Wei Zhu** for helpful remarks and meticulous attention to details on the work in this thesis. I would also like to thank

Joseph Lo for many fruitful discussions and for the practical perspective he brought to the paper, especially the suggestion regarding latent claims.

Last but not least, appreciation needs to be recorded for the anonymous reviewers for their helpful suggestions, and interesting ideas to improve and extend the work.

Weihong Ni

Abstract

This thesis constitutes a research work on Bonus-Malus (BM) systems in insurance portfolios, featuring designing pricing strategies and examining associated solvency risks. The first piece of work proposed two different pricing models via the Bayesian approach. Results imply adverse attitudes towards policyholders having a history of many small claims, when the modelling for claim severities takes different forms. On the other hand, the rest of the work dedicated to embedding a BM structure under a risk analysis framework, where the focus lies in measuring the underlying ruin probabilities. It was necessary to initially investigate a discrete model where such probability could be obtained through recursions. As for a continuous model, BM feature was reflected by a Bayesian estimator for premium adjustment. Such construction normally brings in a dependence structure to the risk model thus violating classical assumptions. One way was to inspect how different it is from a classical risk model. Then through some conditional arguments one could find accordingly a solution based on results in literature. From another perspective, it has been found that for a No Claim Discount (NCD) or a Bonus system, an alteration in premium rates could be transformed equivalently to an interchange of distribution between inter-claim times. Then some Markov properties were able to be diagnosed under higher dimensions, which leads to a further possibility of computations. Results can be found in the form of simulations.

Contents

List of Figures

Chapter 1

Preliminaries

This chapter serves as a foundation of the work to be presented in this thesis. As this research is to employ mathematical models in an insurance context, concepts from this background will be introduced. Technically, probability distributions and some statistical approaches will be demonstrated since it is strongly related to the contents in Chapter 3. Most of the definitions in the first three subsections follow Rolski et al. [2009] and Klugman et al. [1998]. Furthermore, a classical risk model will also be addressed here and further details can be found in Asmussen and Albrecher [2010].

1.1 Actuarial Concepts

Here are some basic definitions in the insurance context which will be seen through out the thesis:

- *An insurance premium*, or simply referred to as 'premium' in the sequel, is the amount of money that policyholders pay to an insurance company for the coverage of associated risks. There are many ways to calculate or estimate the value of this payment. (see premium calculation principles below)

- *Claims* are the amount of losses an insurer is entitled to pay for an insured product. The monetary value of a claim is also referred to as *claim size/cost* *or claim severity* and is considered as a non-negative random variable. *The number of claims* or sometimes called *claim counts/frequency* in a certain period is also a non-negative random variable. *The claim counting process* is often denoted by $\{N(t), t \geq 0\}$ where $N(t)$ is the number of claims up to time t.

- *An epoch of a claim*, or sometimes called *a claim arrival time*, is literally the time at which a claim happens (assuming its cost to be cleared instantaneously). If we denote the epochs by $\tau_1, \tau_2, \ldots$, then $T_n = \tau_n - \tau_{n-1}, n \geq 1$ are called the *inter-arrival times* in-between successive claims.

- *The counting process* denoted by $\{N(t)\}_{t \geq 0}$ counts the number of claims up to time t. It is a random process which will be discussed further in the sequel.

- *A risk surplus/reserve* is the amount of funds in operation in an insurance system which accumulates through premium incomes and drops by claims, whose process is normally described by the following equation.

$$U(t) = u + P(t) - S(t), \quad t \geq 0,$$

where $u = U(0)$ is the initial level of reserve, $P(t)$ is the *premium income* collected up to time t and $S(t) = \sum_{i=1}^{N(t)} Y_i$ is the *aggregate claim amount* with individual claim sizes $Y_i, i = 1, 2, \ldots N(t)$ and the number of claims up to time t as $N(t)$.

- *Premium calculation principles* are a list of different rules which could be followed when conducting premium calculations. Some commonly seen ones are displayed below while others could be found in Asmussen and Albrecher [2010]. Notice that it does not related to stochastic processes, but for a single risk X.

The net premium principle $p(X) = \mathbb{E}[X]$. This is the basic principle in the sense that premiums should be the expected value of losses. In our work Chapter 3.2, we followed exactly this principle and used the product of the expected individual loss and the expected frequency as a proposed premium.

The expected value principle $p(X) = \mathbb{E}[X] = (1+\eta)\mathbb{E}[X]$. Here, η is referred to as the safety loading. Normally, it is assumed in a classical risk model that $\eta > 0$, which is the so-called Net Profit Condition (NPC).

The variance principle $p(X) = \mathbb{E}[X] + \eta Var[X]$, which adds a variation of X.

The above concepts are relatively general but will be illustrated further in the following chapters. We will see how these elements can be modelled under different scenarios.

1.2 Probability Distributions

Discrete Distributions

Poisson distribution with the notation $Poi(\lambda)$ is the most commonly used discrete distribution in this topic defined as

$$\mathbb{P}(N = n) = \frac{\lambda^n}{n!} e^{-\lambda},$$

for $n = 0, 1, \ldots$ and $\lambda > 0$ with mean and variance both equal to λ. The moment generating function is

$$M_N(t) = \mathbb{E}[e^{tN}] = e^{\lambda(e^t - 1)}.$$

Negative Binomial distribution denoted by $NB(\alpha, p)$ is a discrete probability distribution defined by the following probability mass function.

$$\mathbb{P}(N = n) = \binom{\alpha + n - 1}{n} (1-p)^\alpha p^n,$$

for $n = 0, 1, \ldots$, $\alpha > 0, 0 < p < 1$. The mean is then $\frac{\alpha p}{1-p}$ while the variance equals to $\frac{\alpha p}{(1-p)^2}$. Its moment generating function is given by

$$M_N(t) = \left(\frac{1-p}{1-pe^t}\right)^\alpha.$$

Geometric distribution is another well-known discrete probability distribution. There are two ways to define such a random variable. Type 1 says that X is the number of trials until the first success of an experiment and $X \sim Geo(p)$ with $0 < p \leq 1$ as the success probability of an individual trial.

$$\mathbb{P}(X = k) = (1-p)^{k-1}p, k = 1, 2, \ldots.$$

In contrast, Type 2 defines a random variable Y as the number of failures until the first success of an experiment and $Y \sim Geo(p)$ with a successful rate of $0 < p \leq 1$ for each trial.

$$\mathbb{P}(Y = k) = (1-p)^k p, k = 0, 1, \ldots.$$

The expectations are given by $\mathbb{E}[X] = \frac{1}{p}$ and $\mathbb{E}[Y] = \frac{1-p}{p}$ respectively while the variances take the same form $Var[X] = Var[Y] = \frac{1-p}{p^2}$.

<u>Continuous Distributions</u>

Exponential distribution is one of the most widely considered continuous distributions when analysing an actuarial model. With the notation $X \sim Exp(\lambda)$, its probability density function is shown as

$$f(x) = \lambda e^{-\lambda x}, \quad x > 0.$$

The mean is $\frac{1}{\lambda}$ and the variance is $\frac{1}{\lambda^2}$. It has many nice properties such as the memory-less nature which will be explained further in later chapters when concrete models are taken into account and its moment generating function has a simple form

$$M_X(t) = \frac{\lambda}{\lambda - t}, \quad \text{for } t < \lambda.$$

Gamma distribution denoted as $\Gamma(\alpha, \lambda)$ has a probability density function

$$f(x) = \frac{\lambda^\alpha x^{\alpha-1} e^{-\lambda x}}{\Gamma(\alpha)}, \quad x \geq 0,$$

where $\alpha, \lambda > 0$ are real numbers and $\Gamma(\alpha)$ is a Gamma function

$$\Gamma(\alpha) = \int_0^\infty x^{\alpha-1} e^{-x} dx, \quad x > 0.$$

When $\alpha \in \mathbb{N}$, it is called an **Erlang distribution** $Erl(\alpha, \lambda)$ and $\Gamma(\alpha)$ simply becomes $(\alpha-1)!$. Erlang distribution could be considered as a sum of α independent exponential distributions with a common parameter λ.
Beta distribution $Beta(\alpha, \beta)$ has a probability density function shown below

$$f(x) = \frac{x^\alpha (1-x)^{\beta-1}}{B(\alpha, \beta)}, 0 < x < 1,$$

where $B(\alpha, \beta)$ is the Beta function

$$B(\alpha, \beta) = \int_0^1 t^{\alpha-1}(1-t)^{\beta-1} dt = \int_0^\infty \frac{t^{\beta-1}}{(1+t)^{\alpha+\beta}} dt, \ \alpha, \beta > 0. \tag{1.1}$$

Notice that there is a relation between the Gamma function and the Beta function.

$$B(\alpha, \beta) = B(\beta, \alpha) = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}. \tag{1.2}$$

Remark 1.1. (1.1) *is derived through a change of variable* $t = \frac{1}{x+1}$ *where* $x \in (0, \infty)$. *Thus, the integral becomes*

$$\int_0^\infty (1+x)^{1-\alpha} \left(\frac{x}{x+1}\right)^{\beta-1} x^{-2} dx = \int_0^\infty \frac{x^{\beta-1}}{(1+x)^{\alpha+\beta}} dx.$$

The first equality in (1.2) *results from the fact that Beta function is a convolution.*

The relation with Gamma function is illustrated below.

$$\begin{aligned}
B(\alpha,\beta)\Gamma(\alpha+\beta) &= \int_0^\infty \int_0^1 t^{\alpha-1}(1-t)^{\beta-1} y^{\alpha+\beta-1} e^{-y} dt\, dy \\
&= \int_0^\infty \int_0^1 (yt)^{\alpha-1}(y-yt)^{\beta-1} y\, dt\, e^{-y} dy \\
&= \int_0^\infty \int_0^y x^{\alpha-1}(y-x)^{\beta-1} dx\, e^{-y} dy \\
&= \int_0^\infty x^{\alpha-1} \int_x^\infty e^{-y}(y-x)^{\beta-1} dy\, dx \\
&= \int_0^\infty x^{\alpha-1} e^{-x} \int_0^\infty e^{-z} z^{\beta-1} dz\, dx \\
&= \Gamma(\beta) \int_0^\infty x^{\alpha-1} e^{-x} dx \\
&= \Gamma(\alpha)\Gamma(\beta)
\end{aligned}$$

which gives the desired result.

Lévy distribution is one of the few stable (1/2 stable) distributions that has an analytical probability density function. A standard Lévy$(0, c)$ is represented as

$$f(x) = \sqrt{\frac{c}{2\pi x^3}} exp\left(-\frac{c}{2x}\right), x > 0.$$

where c is the scale parameter. Notice that in Chapter 3.2, we are considering such a distribution with c replaced by $c^2/2$.

Weibull distribution. The density of X ∼Weibull(λ, k) could be written as

$$f(x) = \frac{k}{\lambda} x^{k-1} exp\left(-\frac{x^k}{\lambda}\right),\ x \geq 0.$$

When $0 < k < 1$, it is a *heavy-tailed distribution* which means the tail $\bar{F}(x) = 1 - F(x)$ is not bounded by an exponential tail $e^{-\sigma x}$ for all $\sigma > 0$, i.e.,

$$\lim_{x\to\infty} \frac{\bar{F}(x)}{e^{-\sigma x}} = \infty.$$

We used this distribution in Chapter 3.2 resulting from a mixing of Lévy and exponential distribution having a shape parameter $k = \frac{1}{2}$. Its r^{th} moment is calculated by

$$\mathbb{E}[X^r] = \int_0^\infty \frac{k}{\lambda} x^{k-1} exp\left(-\frac{x^k}{\lambda}\right) x^r dx.$$

To proceed with this, let $y = \frac{x^k}{\lambda} \in (0, \infty)$, then $x = (\lambda y)^{1-\frac{1}{k}}$, $dy = \frac{k}{\lambda} x^{k-1} dx$ substituting these back in yields

$$\begin{aligned}
\mathbb{E}[X^r] &= k \int_0^\infty y(\lambda y)^{\frac{r-1}{k}} \mathrm{e}^{-y} \frac{\lambda}{k} (\lambda y)^{\frac{1}{k}-1} dy \\
&= \lambda^{\frac{r}{k}} \int_0^\infty y^{\frac{r}{k}+1-1} \mathrm{e}^{-y} dy \\
&= \lambda^{\frac{r}{k}} \Gamma\left(\frac{r}{k} + 1\right).
\end{aligned}$$

When $k = \frac{1}{2}$, $\mathbb{E}[X] = 2\lambda^2$ and $Var[X] = \lambda^{\frac{2}{k}} \left[\Gamma(\frac{2}{k} + 1) - (\Gamma(\frac{1}{k} + 1))^2\right] = 20\lambda^4$.
Pareto distribution denoted by $X \sim$Par(s, m) is another heavy-tailed distribution with p.d.f

$$f(x) = \frac{sm^s}{(x + m)^{s+1}},\ x \geq 0,$$

Its tail has a power decay which is obviously not bounded by an exponential one. Its mean $\mathbb{E}[X]$ could be derived through

$$\begin{aligned}
\mathbb{E}[X] &= \int_0^\infty \frac{sm^s}{(x+m)^{s+1}} x dx \\
&= \frac{sm^s}{-s} \int_0^\infty x d(x+m)^{-s} \\
&= m^s \int_0^\infty (x+m)^{-s} dx \\
&= -\frac{m^s}{1-s} (x+m)^{-s+1}|_0^\infty \\
&= \frac{m}{s-1}, \quad s > 1.
\end{aligned}$$

1.3 Statistical Methods-Parameter Estimation

1.3.1 Method of Moments

This is a basic approach to estimate parameters in a statistical model. Suppose that there are k parameters $\theta_1, \theta_2, \ldots, \theta_k$ to be estimated in a distribution function $F_X(x; \boldsymbol{\theta})$ of the random variable X. The idea is to use the sample moments to represent those of the population. Then the $1^{st} - k^{th}$ moment can be expressed in terms of $\boldsymbol{\theta} = (\theta_1, \theta_2, \ldots, \theta_k)$.

$$\begin{aligned} \mathbb{E}[X] &= m_1 = g_1(\theta_1, \theta_2, \ldots, \theta_k); \\ \mathbb{E}[X^2] &= m_2 = g_2(\theta_1, \theta_2, \ldots, \theta_k); \\ &\vdots \\ \mathbb{E}[X^k] &= m_k = g_k(\theta_1, \theta_2, \ldots, \theta_k). \end{aligned}$$

Solving k parameters from a system of k equations will result in parameter estimators. On a given data set $x_1, \ldots, x_n$ of size n, we could use the sample moments calculated as

$$\hat{m}_i = \frac{1}{n} \sum_{j=1}^{n} x_j^i, \; for\, i = 1, \ldots, k$$

to estimate the population moments so that the estimates for parameters could be obtained by

$$\begin{aligned} \hat{m}_1 &= g_1(\hat{\theta}_1, \hat{\theta}_2, \ldots, \hat{\theta}_k); \\ \hat{m}_2 &= g_2(\hat{\theta}_1, \hat{\theta}_2, \ldots, \hat{\theta}_k); \\ &\vdots \\ \hat{m}_k &= g_k(\hat{\theta}_1, \hat{\theta}_2, \ldots, \hat{\theta}_k). \end{aligned}$$

This method is simple to implement but the estimation may not always be unbiased partly because it only considers several features rather than the whole behaviour of the data set. An unbiased estimator is defined to be an estimator whose expected value is the real value of the unknown parameter to be estimated. An alternative way is to use the Maximum Likelihood Estimation.

1.3.2 Maximum Likelihood Estimation (MLE)

The key reasoning behind this approach is to figure out a set of values for the parameters under which the possibilities of obtaining the observed data are maximised. We call the joint probability of observing the data $x_1, x_2, \ldots, x_n$ given the vector of parameters $\boldsymbol{\theta}$ the **Likelihood Function**.

$$L(\boldsymbol{\theta}) = \prod_{j=1}^{n} \mathbb{P}(X_j \in \mathrm{d}A_j|\boldsymbol{\theta}) = \prod_{j=1}^{n} f_X(x_j|\boldsymbol{\theta}),$$

where $f_X(x|\boldsymbol{\theta})$ denotes the probability density function for a random variable X given the parameter vector $\boldsymbol{\theta}$ and $\mathrm{d}A_j$ is the corresponding X_j's infinitesimal set. Sometimes, it is easier to consider the **Log-likelihood Function**

$$\ln L(\boldsymbol{\theta}) = \sum_{j=1}^{n} \ln f_X(x_j|\boldsymbol{\theta}).$$

MLE is to define an estimator $\hat{\boldsymbol{\theta}}$ for $\boldsymbol{\theta}$ so that the likelihood function or log-likelihood function is maximised.

1.3.3 Bayesian Estimation

While the previous methods make the assumption that the probability distribution for each sample is fixed and it is the difference among samples that causes the variation in data, Bayesian estimation takes the population distribution to be variable and relies on observed data to estimate the probability of a parameter taking a certain value. So instead of showing a deterministic value for a parameter, Bayesian estimation generates its distribution.

Several concepts need to be clarified first.

- **Prior Distribution** is the distribution of a parameter assumed before any observation, normally denoted by $\pi(\theta)$.

- **Posterior Distribution** is the estimated distribution for the parameter based on the observed data $x = (x_1, x_2, \ldots, k)$. Technically speaking, it is a conditional probability of θ given $\boldsymbol{x}$, i.e., $\pi_{\Theta|\boldsymbol{X}}(\theta|\boldsymbol{x})$.

- **Model Distribution** or sometimes called **Sampling Distribution** is the distribution of the underlying random variable X conditioning on a particular value for the parameter. We write it as $f_{\boldsymbol{X}|\Theta}(\boldsymbol{x}|\theta)$, which coincides with the likelihood function as defined in MLE.

- **Marginal Distribution** is the mixing distribution for X when the parameter is assumed to have a prior distribution and its pdf is

$$f_X(x) = \int f_{\boldsymbol{X}|\Theta}(\boldsymbol{x}|\theta)\pi(\theta)d\theta.$$

- **Predictive Distribution** is a distribution that predicts a new observation y when all previous data $\boldsymbol{x}$ is taken into account, $f_{Y|\boldsymbol{X}}(y|\boldsymbol{x})$.

The core of this approach lies with the Bayes' Theorem.

$$\pi_{\Theta|\boldsymbol{X}}(\theta|\boldsymbol{x}) = \frac{f_{\boldsymbol{X}|\Theta}(\boldsymbol{x}|\theta)\pi(\theta)}{\int f_{\boldsymbol{X}|\Theta}(\boldsymbol{x}|\theta)\pi(\theta)d\theta}.$$

In addition, the predictive distribution is computed by

$$f_{Y|\boldsymbol{X}}(y|\boldsymbol{x}) = \int f_{Y|\Theta}(y|\theta)\pi_{\Theta|\boldsymbol{X}}(\theta|\boldsymbol{x})d\theta.$$

In practice, it is normally required to present an estimate for the parameter. To find such an estimate, we often try to minimise the difference from the real value. A quadratic loss function $l(\hat{\theta}, \theta) = (\hat{\theta} - \theta)^2$ is commonly used in this context. As a result, the **Bayes estimate** is the mean of the parameter, whose reasoning is very similar to the least squared method applied in linear regressions.

Compared to other approaches, Bayesian analysis is more flexible because it allows the model to dynamically adjust to the observed data. However, as there are integrals or sums involved, computational complexity has increased and sometimes an explicit form is not easy to achieve. For instance, we had to introduce Bessel functions in our work (Chapter 3.2) to conduct the analysis.

1.4 Continuous risk models

The establishment of risk theory originated from Filip Lundberg whose work was then explained and further developed by Harald Cramér. Their work pioneered probabilistic modelling under an insurance context and boost the emerging theory of stochastic processes. One of the main goals of this thesis is to study this theory and make further extensions from a classical risk model. This section will briefly discuss the construction of a risk model under the continuous time horizon and how the probability of ruin are derived with demonstration on current existing results. We do not include a section describing a discrete risk model for now because it is to some extent similar to a continuous one and would be better explained under a specific model, e.g., (4.1).

Usually, **a risk surplus process** is written as.

$$U(t) = u + ct - \sum_{k=0}^{N(t)} Y_k,\ t \geq 0, \tag{1.3}$$

where $u = U(0)$. This appears very often in the classical insurance risk theory and describes the amount of surplus $U(t)$ of an insurance portfolio at time t, where c represents a constant rate of premium inflow, $N(t)$ is a claim counting process that counts the number of claims incurred during the time interval $(0, t]$ and $\{Y_k\}_{k\geq 0}$ is a sequence of independent and identically distributed (i.i.d.) claim sizes, independent of the claim arrival process $N(t)$. Normally it is assumed that $U(t) \to +\infty$ a.s. as $t \to +\infty$, which is equivalently to a net profit condition (NPC). One of the crucial quantities to investigate in this context is the probability that at some point in time the reserves in the portfolio will not be sufficient to cover the claims, i.e., $U(t) < 0$, which is called ruin. More formally,

Definition 1.2. *Let $\tau(u)$ denote the time that the surplus process drops below zero for the first time when the initial capital is u, i.e.,*

$$\tau(u) := \inf\{t \geq 0 : U(t) < 0 | U(0) = u\}, \tag{1.4}$$

then the event of ultimate ruin is $\{\tau(u) < \infty\}$. Thus, the **ultimate ruin prob-**

ability *denoted by* $\psi(u)$ *is defined by*

$$\psi(u) := \mathbb{P}(\tau(u) < \infty). \tag{1.5}$$

If on the other hand, only a finite time horizon is considered, i.e., $\{\tau(u) < T\}$, *then a* **finite ruin probability** *is defined as*

$$\psi(u, T) := \mathbb{P}(\tau(u) < T),$$

where $\tau(u) := \inf\{0 \leq t \leq T : U(t) < 0|U(0) = u\}$.

For simplicity, if it is not otherwise stated, we write τ instead of $\tau(u)$ and by default we assume $U(0) = u$ in the sequel. This thesis is only interested in the ultimate ruin probability (under an infinite time horizon).

We further call

$$S(t) = \sum_{k=0}^{N(t)} Y_k - ct,\ t \geq 0$$

the **claim surplus process** which has a supremum $M = \sup_{t\geq 0} S(t)$. Then equivalently,

$$\psi(u) = \mathbb{P}(M > u|U(0) = u).$$

Sometimes is is easier to work with the claim surplus process as a connection with a random walk could be built up. Intuitively, it could be understood that an increment per unit time is given by $\rho - c$ where

$$\frac{1}{t}\sum_{k=1}^{N(t)} Y_k \xrightarrow{a.s} \rho, \text{as } t \to \infty,$$

due to a strong law of large numbers. Then if $\rho - c \geq 0$, the process $\{S(t)\}$ will drift to infinity as $t \to \infty$. That indicates $\psi(u) = 1, \forall u$. On the contrary, if $\rho - c < 0$, $M < \infty$ $a.s.$ and $\psi(u) < 1$, which is the case worth studying Asmussen and Albrecher [2010]. Therefore, it again verifies our NPC assumption and we need

$$\eta = \frac{c - \rho}{\rho} > 0.$$

Obviously, the stochasticity of the process originates from two components - the claim jumps and the arrival process. For the former one, we only consider light-tailed case when calculating ruin probability in this thesis and often work with exponential distributed claims for simplicity, whereas the latter one draws the main attention here. The simplest example is the **Poisson process** which is defined in various ways. A summary can be found in Theorem 5.2.1 in Rolski et al. [2009].

Definition 1.3. *If $\{N(t)\}$ has stationary and independent increments, and for each fixed $t \geq 0$, the random variable $N(t)$ has a Poisson distribution, i.e., $X \sim Poi(\lambda t)$, then it is defined that $\{N(t)\}$ is a Poisson process with intensity λ.*

Proposition 1.4. *A well-known result is that when $\{N(t)\}$ is a Poisson process with intensity $\lambda > 0$, the sequence of inter-arrival times $\{\tau_n\}_{n\in\mathbb{N}}$ is that of i.i.d exponential random variables, i.e., $\tau_i \sim Exp(\lambda)$ for $i \in \mathbb{N}$.*

Proof. In general, we have $\mathbb{P}(N(t) = n) = \frac{(\lambda t)^n \mathrm{e}^{-\lambda t}}{n!}$. Let T_n denote the n^{th} claim arrival time, $T_n = \sum_{k=1}^{n} \tau_k$. Then, for the first claim arrival time $\sigma_1 = \tau_1$, $\mathbb{P}(\tau_1 > t) = \mathbb{P}(N(t) = 0) = \mathrm{e}^{-\lambda}$, which indicates that τ_1 has an exponential density $f_{\tau_1}(t) = \lambda \mathrm{e}^{-\lambda t}$. Then for any $0 < s < t$, $m \geq n \geq 0$, $\mathbb{P}(N(s) = n, N(t) = m) = \mathbb{P}(N(s) = n, N(t) - N(s) = m - n)$. By independent increment and stationary property, it further equals

$$\begin{aligned}
& \mathbb{P}(X(s) = n)\mathbb{P}(N(t) - N(s) = m - n) \\
= \; & \mathbb{P}(X(s) = n)\mathbb{P}(N(t - s) = m - n) \\
= \; & \frac{(\lambda s)^n \mathrm{e}^{-\lambda s}}{n!} \frac{(\lambda(t - s))^{(m-n)} \mathrm{e}^{-\lambda(t-s)}}{(m - n)!} \\
= \; & \frac{(\lambda t)^m \mathrm{e}^{-\lambda t}}{m!} \binom{m}{n} s^n (t - s)^{m-n}.
\end{aligned}$$

Then the conditional probability

$$\mathbb{P}(N(t) = m | N(s) = n) = \frac{(\lambda)^{m-n} \mathrm{e}^{-\lambda(t-s)}}{(m - n)!} (t - s)^{m-n}.$$

If $m = n$, that means $T_n \leq s < t < T_{n+1}$, i.e., $T_{n+1} - T_n > t - s$. We can then write

$$\mathbb{P}(N(t) = n | N(s) = n) = \mathbb{P}(T_{n+1} - T_n > t - s) = \mathbb{P}(\tau_n > t - s) = \mathrm{e}^{-\lambda(t-s)}.$$

Since τ_n is arbitrary, we have proved the assertion. ■

To be more general, $\{N(t)\}$ can be extended to a **renewal process**, which means given a sequence of arrival epochs $\{T_n\}_{n\geq 0}$, the inter-arrival times $\tau_n = T_n - T_{n-1}, n \geq 1$ with $T_0 = \tau_0 = 0$ are i.i.d. It is worth knowing that Poisson is the most commonly used renewal process and probably the easiest. The other extension is to adopt a **mixed Poisson process** whose definition is given by

Definition 1.5. *Rolski et al. [2009] The counting process $\{N(t), t \geq 0\}$ is called a mixed Poisson process if there exists a positive random variable, the mixing random variable Λ with distribution function $F(\lambda) = \mathbb{P}(\Lambda \leq \lambda)$ such that for each $n = 1, 2, \ldots$, for each sequence $\{k_r; r = 1, 2, \ldots, n\}$ of non-negative integers, and for $0 \leq a_1 \leq b_1 \leq a_2 \leq b_2 \leq \ldots \leq a_n \leq b_n$,*

$$\mathbb{P}\left(\bigcap_{r=1}^{n}\{N(b_r) - N(a_r) = k_r\}\right) = \int_0^\infty \prod_{r=1}^{n} \frac{(\lambda(b_r - a_r))^{k_r}}{k_r!} e^{-\lambda(b_r - a_r)} dF(\lambda).$$

The difference from a Poisson process is that a randomness is introduced in the intensity parameter λ, whose use will be addressed further in Chapter 5.1.

Furthermore, Chapter 5.2 considers a **regenerative process** where the renewal does not happen at each claim epoch. Rather, it renews after a few inter-claim times. A more formal definition is given by Asmussen and Albrecher [2010]

Definition 1.6. *Let $\{T_n\}$ be a renewal process. A stochastic process $\{X_t\}_{t\geq 0}$ with a general state space E is called* **regenerative** *with respect to $\{T_n\}$ if for any k, the post-T_k process $\{X_{T_k+t}\}_{t\geq 0}$ is independent of $T_0, T_1, \ldots, T_k$ (or equivalently of $\tau_0, \tau_1, \ldots, \tau_k$), and its distribution does not depend on k.*

1.5 Ruin probabilities

There are three main approaches in literature to tackle the problem of finding ruin probabilities in a **classical risk model**. To be precise, a classical risk model is the one where $\{N(t)\}_{t\geq 0}$ is a Poisson process with intensity λ, $\{Y_k\}_{k\in\mathbb{N}}$ is a sequence of i.i.d light-tailed claim distributions with a common distribution function $F(x)$ and is also independent from $\{N(t)\}_{t\geq 0}$.

The first one is the use of its renewal property from which an integral equation could be established sometimes also for the survival probability $\Phi(u) = 1 - \psi(u)$ by conditioning upon the first time a claim arrives, i.e., $\Phi(u) = \mathbb{E}[\Phi(u+cT_1-X_1)]$

$$\Phi(u) = \int_0^\infty \lambda e^{-\lambda t} dt \int_0^{u+ct} \Phi(u+ct-x)dF(x),$$

where u is the initial value of $U(t)$. In fact this has a connection with another integral equation which could be used directly to seek for a solution. First using the change of variables $y = u + ct$ and taking the derivative w.r.t u finally gives

$$\Phi'(u) = \frac{\lambda}{c}\Phi(u) - \frac{\lambda}{c}\int_0^u \Phi(u-x)dF(x).$$

Then integrating over $(0, t)$ yields,

$$\Phi(u) = 1 - \frac{\lambda\mu}{c} + \frac{\lambda}{c}\int_0^u \Phi(u-y)\bar{F}(y)dy.$$

Next replace survival probability by ruin probability (not necessary but more intuitive to work with) from where Laplace transform serves as the main tool as there is a convolution involved. Consequently, ruin probability could be defined by the **Pollaczek-Khinchin** formula.

$$\psi(u) = \left(1 - \frac{\lambda\mu}{c}\right)\sum_{n=1}^{\infty}\left(\frac{\lambda\mu}{c}\right)^n (\bar{F}_Y^s)^{*n}(u), \tag{1.6}$$

where $\mathbb{E}[Y] = \mu$, c is the premium rate and $\bar{F}_Y^s(y) = 1 - F_Y^s(y) = 1 - \mu\int_0^y \bar{F}_Y(x)dx, y \geq 0$ is the tail of the integrated tail distribution. The term $(\bar{F}_Y^s)^{*n}(u)$ is an n-fold convolution of F_Y. Since F_Y is the distribution function of Y, the n^{th} convolution power of F_Y gives the distribution function of the sum of n independent random variables with identical distribution F_Y. It is worth mentioning here that when claims are exponentially distributed with parameter $1/\mu$, the ruin probability is

$$\psi(u) = \frac{\lambda\mu}{c} \mathrm{e}^{-(\frac{1}{\mu} - \frac{\lambda}{c})u}, \quad u \geq 0$$

The second common method properly adopts the martingale techniques. In Subsection 5.2.4 there will be a detailed explanation of how ruin probability could be estimated by using a change of measure based on a construction of a martingale. Another popular way of analysis is through simulations for which Subsection 5.2.5 gives a more concrete explanation via identifying a Markov additive process. Of course there is a lot of work focusing on approximations as well as bounds and asymptotics about which Section 5.4 in Rolski et al. [2009] presented a good range of literature.

Chapter 2

Introduction

2.1 Bonus-Malus systems

A Bonus-Malus (BM) System is referring to a merit rating system where policyholders get discounts for no claims and they are punished for making claims. That is to say, the premiums for a policyholder to pay depend very much on his past claim records. This system has now been widely used in many European insurance companies for auto-mobile insurance. There are a few reasons for the necessity of the system. Initially, it has been claimed that BM systems could to some extent reduce the risk an insurer is faced with. Dionne and Ghali [2005] studied the influence of BM system on road safety in Tunisia. As a consequence, they found that BM systems did help to cut down the number of reported accidents from the insured who do not switch companies. On the other hand, Moreno et al. [2006] suggested BM systems as the only mechanism to cope with alleviating insurance fraud under the condition that policyholders are loyal to the insurance company. Insurance fraud is interpreted as misreporting the true loss of a claim in their paper. Because of the existence of asymmetry in information, adverse selection is possible. Dionne et al. [1999] found evidence of adverse selection in insurance market by conducting some empirical tests. Moreno et al. [2006] claimed that insurance fraud was traditionally resolved by auditing process which incurred costs to the company. They then demonstrated that, however, BM systems could assist in weakening this kind of phenomenon by adjusting premiums without any risk-bearing costs. Furthermore, BM systems play a significant role in eliminating the problem of moral hazard. By moral hazard, we mean the situation where policyholders are likely to maximise their own benefits when they are protected from the risk by sacrificing the insurer. For instance, a policyholder can be less careful when driving because he is assured for the risk. Thus, BM systems have been introduced to deal with this effect by imposing a financial punishment on policyholders who intend to behave like this. A structured list of the studies related to this system can be found in Lemaire [1995]. This thesis will first work on a pricing model for such system and then move onto a more risk analysis orientation.

However, due to the particular feature of a BM system, one of the biggest issues related to introducing the system is the bonus hunger problem. This means that in order to achieve a premium discount in the following year, the insured may not report some small claims and pay the costs themselves. So the insurance company is missing some useful information for calculating individual premiums. To our knowledge so far, this phenomenon has initially been addressed by Grenander [1957] and Straub [1968] from a game-theoretical perspective. Then Norberg [1975] found optimal premium strategies under two situations according to policyholders' behaviours. By using a least-squared-like approach, he found an explicit linear credibility formulae when policyholders choose a fixed barrier strategy. For the case when the barrier is determined through a comparison with the present value of all future increase in premiums caused by a particular claim, he gave some numerical results. In Chapter 3, we represent the premiums using the same structure as in Frangos and Vrontos [2001] where the analysis is implemented when the total claim size is kept fixed. As a consequence, our model generates a premium function that reflects a discouragement on the bonus hunger reaction.

The design of BM systems can date back to the middle of the 20th century since when only the number of claims reported in the past was considered in the calculation of the future premium a policyholder is to pay. In this case, Picard [1976] claimed that problems might arise if one had a claim worth much more money while the other had lots of small claims. The original objective to develop a fair system is then violated and the sharing of the premiums is not fair among these two policyholders. Thus, introducing the claim severity component is a natural consideration.

Following this thought, Lemaire [1995] first applied Picard [1976] method to Belgian data which distinguishes between small and large claims. However, Lemaire [1995] found that this classification would still lead to serious practical problems, since it is time consuming to assess the exact amount and many policyholders who have claims just over the limit argue a lot. Thus, instead of categorising the claim types, Frangos and Vrontos [2001] proposed a specific modelling of claim severities. Since the claim sizes in motor insurance seem to exhibit long-tails, distributions with this property are thought-out. For instances, Valdez and Frees used a distribution called the Burr XII long-tailed distribution to fit the claim sizes data from Singapore Motor Insurance. Another obvious choice to obtain, e.g. Pareto (Frangos and Vrontos [2001]), is mixing the exponential parameter with an Inverse Gamma. In fact, they used Negative Binomial distribution to model the claim frequency component and Pareto distribution to describe the claim severity component. Based on the Bayes' theorem, a posterior mean was adopted to represent the expections for both components. Additionally, they also incorporated several a priori information and used regression analysis to obtain the estimations of parameters. We have combined the result of mixing distribution presented in Albrecher et al. [2011] and the idea of applying the Bayes' theorem as proposed by Frangos and Vrontos [2001], for Weibull severities instead of Pareto, which will be shown in Section 3.2. We will discuss in detail the motivation and the consequences of choosing a Weibull distribution for the severities.

Our motivation could be interpreted from both academic and practical perspectives. Technically speaking, since Pareto in Frangos and Vrontos [2001] turned out to fit well on the data especially in the tails, it is preferable to have some similar-shaped distributions, and Weibull distribution is a reasonable candidate. Even though Weibull distribution does not have tails as long as Pareto, in reality, it can rely on reinsurance which usually alleviates the burden of extremely large claims. Thus, practically speaking, if Weibull distribution fits well the moderate claims of an insurance company, this combined with reinsurance would be the best choice. The additional value of choosing a Weibull fit in some instances is that this could address the bonus hunger problem, as it will be illustrated later in Section 3.2. This is an advantage of such choice, since by carrying out this model, the tendency of policyholders for not reporting small claims could be discouraged.

We fitted an exponential, a Weibull and a Pareto on a given data set, but first let us look in theory what they are like. All the upper tails of the three distributions are written as follows with one common parameter θ.

$$\begin{aligned} Exponential : P(X > x) &= exp(-\theta x); \\ Weibull : P(X > x) &= exp(-\theta x^{\gamma}); \\ Pareto : P(X > x) &= \left(\frac{\theta}{\theta + x}\right)^{s}. \end{aligned}$$

It can be seen that if $\gamma < 1$, the tail of a Weibull distribution is fatter than that of the Exponential but thinner than that of the Pareto distribution. However, when $\gamma > 1$, the Weibull tail is lighter than that of the Exponential and a special case appears when $\gamma = 1$ where the Weibull distribution becomes an Exponential distribution. Thus, we aim to find a Weibull distribution with its shape parameter less than 1 so that a heavy-tail property can be retained. Fortunately enough, we have found that the heterogeneity of the claim severity could be described by a Lévy (1/2 Stable) distribution. Consequently, when the mixing of this distribution is carried out on an Exponential distribution, a Weibull distribution is obtained in the end (Albrecher et al. [2011]). What is more motivating is that the shape parameter γ is known and equal to $1/2$, which is less than 1, fitting our aim.

Section 3.1-3.3 discusses the procedure of deriving the premium levels for the proposed model. The core of this chapter lies in Section 3.2. It involves explanation of how the mixing distribution is achieved as well as the Bayesian approach. The premium formula is obtained in the end with several analysis described in the subsequent subsection. Section 3.3 is dedicated to applying both our Weibull model and the one using Pareto claim severities on some given data. Results suggest that the bonus-hunger problem in some sense could be alleviated because the system punishes less on people with many small claims when the total cost is fixed.

According to the findings in the first half of the chapter, it has been suggested to use a hybrid model, where claim severities are assumed to be distinguished by 'small' and 'big' taking Weibull and Pareto distributions respectively. The claim frequency component is altered accordingly. Bayesian approach was employed again just under a more complicated setting. As a consequence, net premiums for small claims behave similar to the previous model, whereas those for large ones see a monotone increase with the frequency. Both models tend to suggest a milder strategy towards policyholders with many small claims while total expenses are kept fixed. Therefore, they serve as an encouragement for reporting each additional small claims so that insurers are aware of these potential risks.

2.2 BM embeded in a discrete risk model

With the ever growing popularity of BM systems, one interesting question to study would be whether it really reduces the associated risk and how much it does. A common measure to assess risks that an insurer is exposed to is the ruin probability. Motivated by such kind of problems, we try to compute the probability of ruin for models incorporating BM structures in this thesis, starting from a discrete model.

From Chapter 4, we will step into this risk analysis world. In general, under a discrete time horizon, the risk surplus process can be written

$$U_n = u + cn - \sum_{i=1}^{n} Y_i,$$

where u is the starting reserve, c is the amount of premium income in a single year and $Y_i, i = 1 \dots n$ are assumed to be i.i.d representing an aggregate claim in each year. When considering a BM system, c becomes random, e.g. (4.1). Initially, Wagner [2002] worked on a similar risk model having a two-state Markov Chain and introduced a recursive relation for ruin probabilities. Then Wu et al. applied the same recursion approach in a model under a two-class BM setting, or more precisely a No Claim Discount system. However, the 'ruin probabilities' under these settings are not exactly as how it is defined in a continuous model. The resulting probability is actually for ruin starting from a specific state and is for an individual rather than a collective risk. But it is still worth studying for the sake of understanding the process and the dynamics of the system. One pioneering work under a discrete time framework was Dufresne [1988]. By computing first the stationary distribution of a BM system iteratively, he then showed an inherent relation between such distribution and the ruin probability. He also gave an example using a 22-class Swiss BM system with a specified rule and calculated associated ruin probabilities. This paper built up a strong connection of a BM system and ruin theory.

It could be seen that the most crucial step in discrete risk models is the recursions. One premise though is to identify the states first. In Chapter 4, we are able to figure out a five-state Markov chain for our three-class BM system. We did not simply add a third state from Wu et al.. The idea originates from a practical problem in a reinsurance company where concerns lie in catastrophic risks. Hence, the model set-up relies on the construction of this Markov chain. Then by recursion, what is left is only computational complexities and solving boundary conditions. Results are shown in the form of probability generating functions.

2.3 BM embedded in a continuous risk model

There is an extensive research literature under this topic. A brief introduction was given in Preliminaries 1.4 and 1.5. However, we would like to emphasise here some recent work on the ruin probabilities associated with BM systems. It started from the idea of randomising premiums income in a classical risk process. Temnov

[2004] assumed another Poisson process for premium incomes independent from the claim process and found a Pollaczek-Khinchin-formula-like ruin functions, whereas Wang et al. [2007] extended it by adding a stochastic investment return according to a Lévy process and obtained bounds for the underlying probability. While these work performed only under a quasi-BM structure, Afonso et al. [2009, 2015] conducted calculations for ruin probabilities under a realistic BM framework and even worked with real data to find out the effects BM systems have on ruin probabilities. They have a novel setting with the risk surplus denoted by

$$U(t) = u + \sum_{j=1}^{i-1} P_j + (t - i + 1)P_i - S(t),$$

where i is the integer representing the i^{th} policy year and $t \in [i-1, i)$, P_i is the premium in the corresponding year and $S(t)$ still describes the aggregate claims. The interesting idea here is that they first analysed ruin probabilities for a single year by conditioning on the reserve level at the beginning and the end of the year. Since the premium rate is constant within a year, a classical technique could be borrowed. Rather than moving forward by recursion, they used approximations and worked with some data. In this thesis, we will introduce two other ways to identify ruin probabilities in risk models with different architectures.

2.3.1 Premium adjusted via a Bayesian estimator

Comparing to the classical collective risk models, one of the main assumptions is that premiums are arriving at a constant rate c and thus the surplus of the company evolves over time as (1.3), where u is the initial capital and Y_k are the claim sizes (i.i.d. random variables) arriving according to a Poisson process $N(t)$ with intensity λ. Ruin is defined as the first time the surplus process crosses zero. The time of ruin is denoted by τ and the probability of ruin

$$\psi(u) = \mathbb{P}\left(\inf_{t\geq 0} U(t) < 0 | U(0) = u\right) = \mathbb{P}(\tau < \infty | U(0) = u), \tag{2.1}$$

is a function of the initial reserve u, as defined in Definition 1.2.

In an attempt to provide more realistic models, non-constant premium rates have been proposed in collective risk literature. One such approach considers the premium to be a function of the current level of the risk reserve $U(t)$, see e.g. Chapter VIII of Asmussen and Albrecher [2010]. Another approach explores adjusting the premium rate according to the claims history - main feature of BM merit systems, see Bühlmann [2007] for a contextual history of the models. One way to achieve this is via a randomisation of the Poisson parameter, either at the beginning of the process, Lundberg [1948], or iteratively during the whole time of the process, Ammeter [1948].

Furthermore, in Bühlmann [1972] it is assumed that the Poisson parameter has a Gamma distribution and additionally introduces a model where premiums are adjusted based on the claims experience to date - a first presence of a BM premium system within risk theory framework. Furthermore, Dubey [1977] builds upon Bühlmann [1972] and employs the Bayesian estimation of the premium adjustment. This method permits a general distribution of the Poisson parameter. In Constantinescu et al. [2012] the analysis is extended from a Poisson process to more general counting processes and Jasiulewicz [2001] obtains the ruin probability of a surplus Cox process with the premium rate being a function of the claim arrivals.

We will focus on the premium rates adjusted according to the claim history in Section 5.1 as introduced by Bühlmann [1972] and refined in Dubey [1977]. Specifically, the risk reserve process is defined as

$$U(t) = u + c\int_0^t \hat{\lambda}(s)\mathrm{d}s - \sum_{k=0}^{N(t)} Y_k, \quad t \geq 0, \tag{2.2}$$

where $\hat{\lambda}(t) = \mathbb{E}[\Lambda|N(t)]$ is the Bayesian esimator of Λ conditioning upon the counting process $\{N(t)\}_{t\geq 0}$, which is illustrated further by (2.3). Hence, instead of a constant premium rate as shown in the classical collective risk process (1.3), the premium rates are dynamically adjusted, by randomising the expected number of claims Λ over time. In fact, the underlying counting process $\{N(t), t \geq 0\}$

is a mixed Poisson process whose formal definition is given by Definition 1.5 and could also be found in Rolski et al. [2009]. This is an inhomogeneous Poisson process, but conditioning on the random variable Λ, $\{N(t), t \geq 0\}$ becomes a homogeneous Poisson process. The randomness of Λ reflects a heterogeneous environment in an insurance portfolio.

More precisely, in (2.2), the intensity is a random variable Λ which is estimated based on the history of claims as

$$\hat{\lambda}(t) = \mathbb{E}[\Lambda | N(t)]. \tag{2.3}$$

This is a Bayesian estimator, frequently used in the BM literature for the calculation of the premium in terms of past claim frequencies, see e.g. Ni et al. [2014a]. As a side note, Λ following a Gamma distribution produces a credibility estimator that constitutes the basis for pricing BM systems in the Swiss liability car insurance Dubey [1977].

We will look at the defectiveness of Λ and thus introduce and analyse two streams of risks, the 'historical' stream and the 'unforeseeable' stream. In general, one can write $\mathbb{P}(\Lambda = 0) = p$, where $0 \leq p < 1$[1]. The case where $p = 0$ is that Λ is non-defective whereas when $0 < p < 1$ it means that Λ is defective at $\{0\}$. We refer to the 'historical' stream as the former case where the probability of no claims is zero. In the 'unforeseeable' stream, we will have a positive probability of non-occurrence of a claim. Intuitively, claims will happen for sure in the 'historical' stream as $\Lambda > 0$ holds almost surely, which means Λ is non-defective and this is normally an assumption in a risk model. We can regard claims in this stream as those coming from policies on which we have historical data and a certain amount of knowledge. On the contrary, policies associated with the 'unforeseeable' stream contain less information at the beginning and bear a potential to either cause no claims at all or to incur a large amount of

[1] We omit it when $p = 1$ because it is not worth considering in this model as this is a situation where λ is deterministic and equals to 0. That simply means no claims occur and thus the ruin probability is 0. This could be verified by substituting $p = 1$ in (2.7), which gives the desired result.

claims. This stream could be understood as a collection of risks that are relatively new and innovative. For car insurance specifically, a recent example would be an autonomous vehicle technology introduced by Google Thrun [2010]. They have already done a lot of road tests of their self-driving cars for a couple of years and had very few accidents, and these occurred either due to other drivers in traffic, or when the car was operated by a human (Pritchard), meaning none of the accidents were caused by the cars themselves. If this good news carries on, then insurance companies would receive no claims from Google for the launch of these autonomous cars. Unlike the 'historical' one, that means the probability of claims occurring in the 'unforeseeable' stream is less than 1 because Λ is defective at $\{0\}$.

If a claim is in the 'unforeseeable' stream, it is usually followed by a series of claims, which might be a big concern for an insurance company. These claims are called 'latent'. One example of 'latent' claims is asbestos, which have led to a burst of related diseases and thus an increase in claims to be paid Brooks et al. [2013]. In recent years, we have started concerns about our health and safety as a result of new technology, pollution and so on. These risks which nowadays have been included in the internal models of insurance companies can be referred to as 'latent' or 'emerging' because their effects are unknown when signing a policy but might become more obvious as time moves on. 'Latent' claims normally come in as a bunch or a cluster due to the same cause of origin. So the positive part of Λ in the 'unforeseeable' stream would possibly be large. Therefore, two extremes could be seen in such a risk stream and thus premiums are adjusted in a very sensitive way to the number of claims observed. Like in the previous example of self-driving cars, the experimental safety seems quite satisfying but people still have very little information on it and we have no knowledge about its performance when it really enters the market. It is 'so far so good'. So based on the current information, the premium rate could still be kept low before a big jump at the first witness of an accident due to the functioning of the car itself. Hence, it would be justifiable to look further into a model incorporating the 'unforeseeable' stream. We will conduct analysis on the 'unforeseeable' stream alone as well as a model with a combination of the two streams in Section 5.1.

Since Section 5.1 is mainly based on Dubey [1977], a brief introduction will be presented here first. This is a paper of André Dubey, written in French, published in Mitteilungen der Vereinigung schweiz Versicherungsmathematiker, in 1977. It discusses the ruin probability for an adjusted risk surplus process shown as (2.2). It presents an analysis of the probability of ruin under three scenarios based on different ways of estimating $\hat{\lambda}(t)$.

1. $\hat{\lambda}(t) = \mathbb{E}[\Lambda|N(t)]$ (Bayes/Credibility estimation);
2. $\hat{\lambda}(t) = N(t)/t$;
3. $\hat{\lambda}(t) = \frac{a+N(t)}{b+t}$, where a, b are parameters (special case of 1.).

Note first that Scenario 3 is actually a special case of Scenario 1 when Λ is non-defective, i.e., $\mathbb{P}(\Lambda = 0) = 0$, with $\Lambda \sim Gamma(a, b)$. This is in fact a common assumption in a BM system context, see e.g. Ni et al. [2014a].

For Scenario 2, regardless of whether Λ is defective or not, a general relation between the ruin probability in the adjusted model which we denote here by ψ^A (A stands for 'Average' per unit time) and a classical one has been established in Dubey [1977].

$$\psi^A(u) = [1 - \mathbb{P}(\Lambda = 0)] \left(1 - F(u) + \int_0^u \psi^C(u-y)dF(y)\right).$$

The focus of our work is Scenario 1. As in Dubey [1977], we first distinguish two cases according to the defectiveness of Λ for Scenario 1:

(a) Non-Defective

$$\mathbb{P}(\Lambda = 0) = 0; \tag{2.4}$$

(b) Defective

$$\mathbb{P}(\Lambda = 0) = p > 0. \tag{2.5}$$

Dubey [1977] shows that for Scenario 1 under the condition (a), i.e. a risk model with premiums adjusted to the history of claims arriving according to a mixed Poisson process with parameter Λ continuous non-defective random variable, the ruin probability coincides with that of a classical risk model with constant premium rate. Translated in our language, when all claims are from the 'historical' stream, the following relation holds:

$$\psi^H(u) = \psi^C(u), \tag{2.6}$$

where $\psi^C(u)$ refers to the ruin probability in a classical risk model (1.3) with deterministic intensity $\lambda = 1$ and ψ^H describes the probability of ruin in a mixed Poisson model (2.2), with a non-defective parameter Λ estimated via the Bayesian estimator (2.3). Here H stands for the 'historical' stream of risks involved in the model. The most important step in the proof is that the premium collected in one period $P_{n+1}, n \geq 1$, conditioning on the previous claim arrival time $T_n, n \geq 1$ is simply exponentially distributed, i.e., $\mathbb{P}(P_{n+1} \geq x | T_n = y) = e^{-x}$.

On the other hand, for Scenario 1 under the condition (b) which we call the 'unforeseeable' stream here, he simply states results of ruin probabilities for specific claim sizes (i.e. exponentially distributed, or equal claims of size one). However, since no explicit relation between the ruin probability in such model versus a classical one has been presented in Dubey [1977], our work focuses first on the derivation and analysis on (b) under Scenario 1, meaning when only 'unforeseeable' risks are considered.

Obviously, a (2.6)-type relationship will no longer hold, but we can still establish that (see also Theorem 5.1)

$$\psi^L(u) = \psi^C(u) - p\psi^C\left(u + c\ln\frac{1}{p}\right), \tag{2.7}$$

whenever ψ^L denotes the ruin probability of a risk model (2.2) with only the 'unforeseeable' stream of risks and $\psi^C(u)$ again refers to the ruin probability in a classical risk model (1.3) with deterministic intensity $\lambda = 1$. We chose 'L' in the ψ^L to stand for 'Latent' as there will probably be such claims in this stream. Consequently, in a risk model with known claim distribution the ruin probability can be expressed explicitly. See Example 5.2 for exponential claims, as in Dubey [1977].

Our second contribution introduces a more realistic scenario of the model (2.2), still under Scenario 1, featuring a mixture of both known and unknown risks, whose ruin probability will be denoted by ψ^M, with M for 'Mixture'. In order to combine the above two cases we need a well chosen estimator for the intensity of this number of claims process. The estimator of choice is

$$\hat{\lambda}(t) = \mathbb{E}[\Lambda^{(1)} + \Lambda^{(2)} | N(t)], \tag{2.8}$$

where $\mathbb{P}(\Lambda^{(1)} = 0) = 0$ and $\mathbb{P}(\Lambda^{(2)} = 0) = p > 0$.

In this case, the relation between the ruin probability of the adjusted premium model ψ^M versus the classical one ψ^C is more elaborated, and can be derived only when $\Lambda^{(1)} \sim \Gamma(\alpha, \lambda_0)$ and $\Lambda^{(2)}|_{\Lambda^{(2)}>0} \sim \Gamma(\beta, \lambda_0)$ for some $\alpha, \beta, \lambda_0 > 0$,

$$\psi^M(u) = \frac{1-p}{B(\alpha,\beta)} \int_{(0,1)} \psi^C_\theta(u) \theta^{\alpha-1}(1-\theta)^{\beta-1}\, d\theta + p \cdot \psi^C_1(u).$$

where $B(\alpha, \beta)$ is a Beta function and $\psi^C_\theta(u)$ is the classical ruin probability conditioning on θ, whose claim sizes have common distribution function $H_\theta(y) = F(y) + (1-\theta)G(y)$, with F, G denoting claim distributions in the 'historical' stream and the 'unforeseeable' stream respectively.

$$\psi^C_\theta(u) = P\left(\tau < \infty \middle| U(0) = u, \frac{\Lambda^{(1)}}{\Lambda^{(1)} + \Lambda^{(2)}} = \theta\right).$$

For detailed proof of results and explanations, please refer to Section 5.1.

2.3.2 Premium varying according to number of claims in a past fixed window

Section 5.2 will explain another approach to reflect a BM feature into the risk models. It could be considered as equivalent to changing the distribution of the

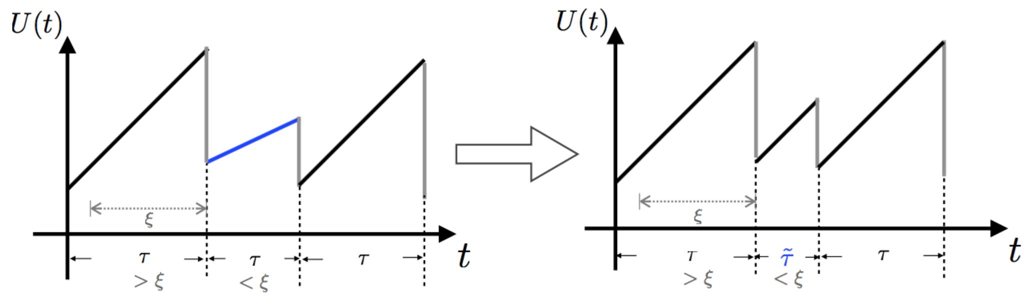


Figure 2.1: Model transformation

subsequent inter-claim time if the current one satisfies a certain condition thus introducing a dependence structure into the process. Without loss of generality, Figure 2.1 plots an example of such risk processes and demonstrates how we transfer a particular model to a form that is easier to be implemented. The graph on the left shows a two-level no claim discount system (Bonus system) where the premium rate decreases after a relatively long wait which exceeds a fixed number ξ. In reality, this fixed window could be understood as a calendar year for instance, because many insurances companies charge different premiums based on only last year's claim histories. After that, since the second waiting interval is less than ξ, the premium rate returns to its original value and so on and so forth. Equivalently, this could be transferred to a model where the adjustment on premium rates is reflected in inter arrival times switching between two different random variables, as long as the increment of $U(t)$ in this time interval is kept the same. That is to say, whenever a large inter arrival time, i.e., above ξ, is witnessed, the next one will switch to a different distribution. As we work only with the ruin probabilities under an infinite-time horizon, such transformation would not affect the results. For a model with realistic sense, $\tilde{\tau}$ is assumed to have a smaller mean than τ. Additionally, for computational reasons, we made an assumption that the inter-exchange of the randomness of inter arrival times only happens after a jump rather than a precise end of the fixed window.

It is mainly based on a regenerative structure which has been studied by Palmowski and Zwart [2007], Palmowski and Zwart [2010]. Such processes are more general than a renewal process and often involve a dependence architecture. Palmowski and Zwart [2007] derived asymptotic results for ruin probabilities when three forms of claim distributions were taken into account, heavy-tailed, intermediate and light-tailed case (Cramér assumption), for a general regenerative process. Section 5.2 will focus on the Cramér case only. On the other hand, a Markovian environment could still be found in such system even though there involves dependence. Rather than the usual Markov processes, a simple two-dimensional Markov chain can be identified which is the so-called discrete Markov additive process. Furthermore, results can be simulated through a change of measure technique. There is extensive literature on such approach to be found in Asmussen and Albrecher [2010]. But due to the simplicity of our model, a classical change of measure via exponential families is enough. There will also be some calculations based on integral equations in the end, yet not helping seeking for analytical solutions.

The rest of the thesis will be organised as follows. Chapter 3 describes two pricing models for BM system via a Bayesian estimation which is based on the papers Ni et al. [2014a] and Ni et al. [2014b]. Chapter 4 shows how to deal with ruin probabilities in a real-world example of discrete BM system. Inspired by Chapter 3, a Bayesian estimator is lodged in a continuous risk model discussed in the first half of Chapter 5. By introducing an innovative stream of risks, ruin probabilities are derived through a comparison with a classical one for two cases. The results of this section was recently published in the journal Insurance: Mathematics and Economics Li et al. [2015]. The second half in this chapter paved the way into a more risk analysis study by constructing a dependence structure mimicking a no claim discount system. It is still a working progress with some results presented at several conferences Constantinescu et al. [2015b]. Appendix A provides proofs as well as the data frame used in Chapter 3. Appendix B presents proofs and also some remarks for Chapter 5. A list of bibliography can be found at the end of the thesis.

Chapter 3

Pricing a BM system by addressing claim severity distributions

One of the pricing strategies for Bonus-Malus (BM) systems relies on the decomposition of the claims' randomness into one part accounting for claims' frequency and the other part for claims' severity. This chapter serves as a kick-off study through statistical analyses, aiming at providing an introduction and explanation of a BM system and addressing the issue of modelling claim costs. Two papers (Ni et al. [2014a], Ni et al. [2014b]) were published based on this chapter. Firstly, by mixing an Exponential with a Lévy distribution, we treated the claim severity component as a Weibull distribution. For a Negative Binomial number of claims, we employ the Bayesian approach to derive the BM premiums for Weibull severities. We then compared our closed form formulas for calculating premiums and numerical results with those for Pareto severities that were studied by Frangos and Vrontos [2001]. Based on our findings, we suggest a hybrid model for claim severities using the same approach, which will be discussed in the second section. Despite gaining deep understanding of a BM system from conducting this initial study, readers could also get prepared with Bayesian statistics/estimation which will be further incorporated in an insurance risk model in Chapter 5.1.

3.1 Modelling Claim Frequencies

The modelling of the claim counts is borrowed from Frangos and Vrontos [2001], so that a comparison of the results can be clearer. A very brief explanation will be given here in order to avoid duplications. More details are available in Frangos and Vrontos [2001].

Mixing the Poisson intensity Λ with a Gamma(α, τ) yields a Negative Binomial probability mass function (p.m.f).

$$P(N=n) = \int_0^\infty \frac{\mathrm{e}^{-\lambda}\lambda^n}{n!} \cdot \frac{\lambda^{\alpha-1}\tau^\alpha \mathrm{e}^{-\tau\lambda}}{\Gamma(\alpha)} \mathrm{d}\lambda = \binom{n+\alpha-1}{n} \left(\frac{\tau}{1+\tau}\right)^\alpha \left(\frac{1}{1+\tau}\right)^n.$$

Furthermore, by applying the Bayesian approach, the posterior distribution is given by,

$$\mu(\lambda|n_1, n_2, \ldots, n_t) = \frac{(\tau+t)^{K+\alpha}\lambda^{K+\alpha-1}\mathrm{e}^{-(t+\tau)\lambda}}{\Gamma(\alpha+K)},$$

where $K = \sum_{i=1}^t n_i$ represents the total claim frequency over t years with n_i denoting the claim numbers in each year respectively. It is easily seen that this posterior distribution is still a gamma but with new parameters $K+\alpha$ and $t+\tau$. When a quadratic loss function is considered, the posterior mean is the best estimate, which is given as follows.

$$\lambda_{t+1}(n_1, n_2, \ldots, n_t) = \frac{\alpha+K}{t+\tau}. \tag{3.1}$$

This also represents the expected claim frequency for the coming period, since the mean of a Poisson is λ itself.

3.2 Modelling Claim Severities

In this chapter, our focus lies in the claim severity distribution. We use the Weibull distribution to model the claim severities whose applications appear not only in reliability engineering and failure analysis, but also in insurance survival analysis and sometimes in reinsurance (Boland [2007]). Its probability density function (p.d.f) is

$$f(x) = c\gamma x^{\gamma-1} exp(-cx^{\gamma}), x \geq 0, c > 0, \gamma > 0,$$

and its cumulative density function (c.d.f) is

$$F(x) = 1 - exp(-cx^{\gamma}), x \geq 0.$$

It was found by Albrecher et al. [2011] that a mixing of a Lévy distribution on the exponential distribution would result in a Weibull distribution with its shape parameter known as 1/2. Suppose the exponential distribution is denoted as

$$f(X = x|\theta) = \theta \mathrm{e}^{-\theta x},$$

with a distribution function (c.d.f)

$$F(X \leq x|\theta) = 1 - \mathrm{e}^{-\theta x}.$$

And the parameter θ is assumed to be Lévy distributed which is also referred to as the stable (1/2) distribution. Then we have a prior distribution described below.

$$\pi(\Theta = \theta) = \frac{c}{2\sqrt{\pi\theta^3}} exp\left(-\frac{c^2}{4\theta}\right), \theta > 0.$$

Hence, we obtained the distribution function as follows. (Proof in Appendix A)

$$F(x) = 1 - exp(-c\sqrt{x}), x \geq 0.$$

It is the Weibull distribution with shape parameter equal to 1/2.

Furthermore, with a similar approach to Frangos and Vrontos [2001], we need to find the posterior distribution using the Bayes' Theorem. Suppose the insurance company receives a sequence of claim costs $\{x_1, x_2, \ldots, x_K\}$ from a policyholder with total of K claims over the time horizon considered. If we let $M = \sum_{i=1}^{K} x_i \geq 0$ to describe the total amount of all these claims, the posterior structure function is written in the following form according to Bayes' theorem.

$$\pi(\theta|x_1,x_2,\ldots,x_K)=\frac{[\prod_{i=1}^{K}f(x_i|\theta)]\pi(\theta)}{\int_0^\infty[\prod_{i=1}^{K}f(x_i|\theta)]\pi(\theta)\mathrm{d}\theta}=\frac{\theta^{K-\frac{2}{3}}exp\left(-\left(M\theta+\frac{c^2}{4\theta}\right)\right)}{\int_0^\infty\theta^{K-\frac{2}{3}}exp\left(-\left(M\theta+\frac{c^2}{4\theta}\right)\right)\mathrm{d}\theta}. \tag{3.2}$$

We know that after the integration, the denominator will become independent from θ. By omitting all terms which are not related to θ, we can obtain the kernel of this distribution.

$$\begin{aligned}\pi(\theta|x_1,x_2,\ldots,x_K) &\propto \theta^{K-\frac{2}{3}}exp\left(-\left(M\theta+\frac{c^2}{4\theta}\right)\right),\\ \pi(\theta|x_1,x_2,\ldots,x_K) &\propto \theta^{p}exp\left(-\left(\frac{\theta}{q}+\frac{r}{\theta}\right)\right).\end{aligned}$$

where, $p=K-\frac{3}{2}, q=\frac{1}{M}, r=\frac{c^2}{4}$. This is a form of a Generalized Inverse Gaussian distribution (Tremblay [1992]). Going back to (3.2), by slightly modifying the variables, it can be rewritten as,

$$\pi(\theta|x_1,x_2,\ldots,x_K)=\frac{\left(\frac{c}{2\sqrt{M}}\right)^{-\left(K-\frac{1}{2}\right)}\theta^{K-\frac{2}{3}}exp\left(-\left(M\theta+\frac{c^2}{4\theta}\right)\right)}{\int_0^\infty\left(\frac{2\sqrt{M}\theta}{c}\right)^{K-\frac{3}{2}}exp\left(-\frac{c\sqrt{M}}{2}\left(\frac{2\sqrt{M}\theta}{c}+\frac{c}{2\sqrt{M}\theta}\right)\right)\mathrm{d}\left(\frac{2\sqrt{M}\theta}{c}\right)}. \tag{3.3}$$

The integral on the denominator can be transformed to a modified Bessel function, whose integral representation is normally given as follows (Abramowitz and Stegun [1964]).

$$B_v(x)=\int_0^\infty \mathrm{e}^{-x\cosh t}\cosh(vt)\mathrm{d}t.$$

However, we cannot make a direct connection between this expression and what we have.

Proposition 3.1. *An alternative integral representation of the modified Bessel function is given as follows.*

$$B_v(x)=\frac{1}{2}\int_0^\infty exp\left(-\frac{1}{2}x\left(y+\frac{1}{y}\right)\right)y^{v-1}\mathrm{d}y, x>0.$$

Proof. Appendix A. ■

As compared to the integral in (3.3), it is not difficult to rewrite the posterior distribution in the form below.

$$\pi(\theta) = \frac{\left(\frac{c}{2\sqrt{M}}\right)^{-\left(K-\frac{1}{2}\right)} \theta^{K-\frac{3}{2}} exp\left(-\left(M\theta + \frac{c^2}{4\theta}\right)\right)}{2B_{K-\frac{1}{2}}(c\sqrt{M})}.$$

Or alternatively as,

$$\pi(\theta) = \frac{\left(\frac{\alpha'}{\beta'}\right)^{\frac{v}{2}} \theta^{v-1} exp\left(-\frac{1}{2}\left(\alpha'\theta + \frac{\beta'}{\theta}\right)\right)}{2B_v\left(\sqrt{\alpha'\beta'}\right)}.$$

where $\alpha' = 2M, \beta' = \frac{c^2}{2}, v = K - \frac{1}{2}$. From the properties of a Generalised Inverse Gaussian distribution, the expectation is shown below (Embrechts [1983]).

$$E[GIG] = \sqrt{\frac{\beta'}{\alpha'}} \frac{B_{v+1}\left(\sqrt{\alpha'\beta'}\right)}{B_v\left(\sqrt{\alpha'\beta'}\right)}.$$

Since our model distribution was assumed to be exponential whose conditional mean is given by $E(X|\theta) = \frac{1}{\theta}$. By integrating $1/\theta$ with respect to the posterior distribution $\pi(\theta)$, one gets

$$E[ClaimSeverity] = \frac{2\sqrt{M}}{c} \frac{B_{K-\frac{3}{2}}(c\sqrt{M})}{B_{K-\frac{1}{2}}(c\sqrt{M})}.$$

With the claim frequency (3.1) considered, this expression contributes to the closed form formula of the estimated net premium for the following period.

$$Premium = \frac{\alpha + K}{t + \tau} \cdot \left(\frac{2\sqrt{M}}{c} \frac{B_{K-\frac{3}{2}}(c\sqrt{M})}{B_{K-\frac{1}{2}}(c\sqrt{M})}\right).$$

Now the problem has reduced to calculate the ratio of the above two Bessel functions. As described by Lemaire [1995], two properties of the modified Bessel function could be considered here, i.e., for any $x > 0$, $B_v(x)$ satisfies the following two conditions.

$$\begin{aligned} B_{-v}(x) &= B_v(x), \\ B_{v+1}(x) &= \frac{2v}{x} B_v(x) + B_{v-1}(x). \end{aligned}$$

If we let,

$$\frac{B_{v-1}(c\sqrt{M})}{B_v(c\sqrt{M})} = \frac{B_{K-\frac{3}{2}}(c\sqrt{M})}{B_{K-\frac{1}{2}}(c\sqrt{M})} = Q_K(c\sqrt{M}).$$

Then it can be easily seen that $Q_1 = 1$ from the first condition. Additionally, we can write a recursive function for Q_K based on the second condition.

$$\frac{1}{Q_{K+1}(c\sqrt{M})} = \frac{2K-1}{c\sqrt{M}} + Q_K(c\sqrt{M}),\ K \geqslant 1.$$

This will finally contribute to the calculation of the premium.

On the other hand, however, it is not difficult to see that our premium model is not defined when $M = 0$. This denotes the scenario where there are no claims. So at the same time $K = 0$. Hence, we will redefine the premium for this case. Since we assumed the claim severity is Weibull distributed in a single year for each policyholder, it would be convenient to assume that our initial premium is equal to the product of the mean claim frequency (Negative Binomial) and severity (Weibull). Therefore, the base premium for any new entrant is set as follows.

$$P_0 = \frac{\alpha}{\tau} \cdot \frac{2}{c^2}.$$

Then after, premiums for the following years if no claims are filed will be given by,

$$Premium|_{M=0} = \left(\frac{\alpha}{t+\tau}\right)\left(\frac{2}{c^2}\right). \tag{3.4}$$

This means that when there are no claims reported the premium would be discounted with the time a policyholder is within the company.

In the following part, we would like to address further on the premium functions from Frangos and Vrontos [2001] as well as ours. Their premium function is concave. Nevertheless, our model presents a more complex shape, as one can see by analysing the difference equations of the premium functions with respect to the accumulated number of claims in t years.

Initially, let us see the premium expression given by Frangos and Vrontos [2001].

$$P_{FV} = \frac{\alpha + K}{t + \tau} \cdot \frac{m + M}{s + K - 1}$$

where $m > 0, s > 1$ are the parameters in the Pareto distribution originally coming from the Inverse Gamma distribution, and other notations are the same as above. The difference equation with respect to K is obtained as follows.

$$P_{FV}(K+1) - P_{FV}(K) = \frac{M + m}{t + \tau} \cdot \frac{s - \alpha - 1}{(s + K - 1)(s + K)}. \tag{3.5}$$

Since $K \geqslant 0, s > 1$, if $s - \alpha - 1 > 0$, which is normally the case, we can conclude that the premium is strictly increasing with K. This will be further illustrated with our following numerical example.

Subsequently, we look at the monotonicity of our Weibull model regarding the variable K when we keep M fixed. By analysing the ratio of two successive K values for the premium function we obtained above, we have,

$$\frac{Premium(K+1)}{Premium(K)} = \frac{\alpha + K + 1}{\alpha + K} \cdot \frac{Q_{K+1}(c\sqrt{M})}{Q_K(c\sqrt{M})}.$$

Clearly, the left half of the above formula is larger than 1. However, the right half is less than 1, which is explained as follows.

$$\frac{Q_{K+1}(c\sqrt{M})}{Q_K(c\sqrt{M})} = \frac{B^2_{K-\frac{1}{2}}(c\sqrt{M})}{B_{K-\frac{3}{2}}(c\sqrt{M})B_{K+\frac{1}{2}}(c\sqrt{M})} < 1.$$

The '$<$' comes from the Turán-type inequalities whose proof can be found in Ismail and Muldoon [1978] and Lorch [1994]:

$$B^2_v(x) < B_{v-1}(x)B_{v+1}(x).$$

Thus, the monotonicity is not seen analytically and may depend much on chosen parameters. This gives evidence that our premium function distinguishes from the Pareto model. In our numerical example, we have identified the specific pattern of our premium.

3.3 Numerical Illustration I

3.3.1 Parameter Estimation

In this small section, application of the model proposed in Section 3.2 will be illustrated in more details. Initially, a brief description of the data will be given. The data structure is originally obtained from Table 2.27 of Klugman et al. [1998]. But we skewed and scaled the data in the British currency. By keeping the group structure unchanged, the sample size was also shrunken to 250. The grouped data could be found in Appendix A. However, for some computational reasons, we have randomly generated a data set based on this grouped data. A summary of the data is shown below with its histogram underneath (Figure 3.1).

Table 3.1: Description of the Data

Min.	1^{st} Qu.	Median	Mean	3^{rd} Qu.	Max.
10	240	1395	4538	4103	102722

As can be seen from this histogram, most of the claims lie below £20,000. The situation where claim severities exceed £45,000 is very rare. In our sample, there are a total of 3 policyholders claiming more than this amount. In this work, we are treating these as outliers for illustration purposes.

In order to compare with Frangos and Vrontos [2001], both of the two distributions used to model claim severity will be fitted using our data set. Initially, the Pareto distribution is applied. Our estimates are obtained using R. The results of our maximum likelihood estimation for Pareto distribution is

$$\begin{cases} m = 1999.985031, \\ s = 1.343437. \end{cases}$$

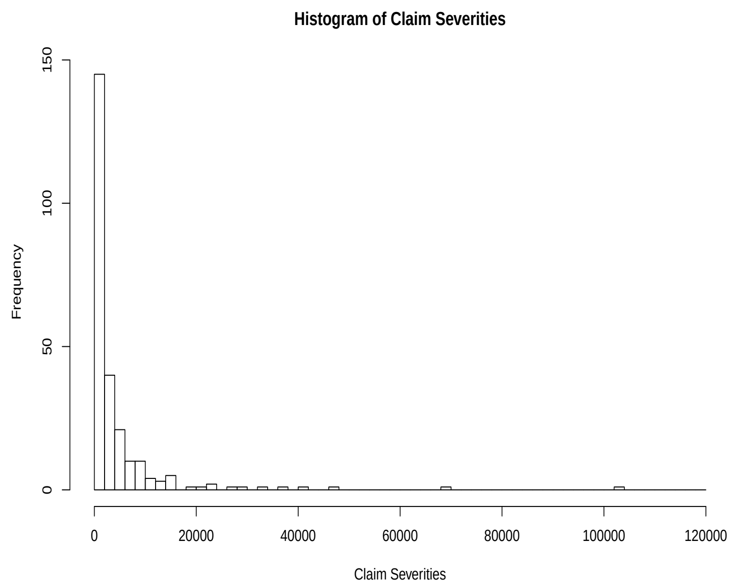


Figure 3.1: Histogram of the Data

In Figure 3.2 the dot-dashed curve is the fitted Pareto distribution for the data without the three outliers. On the other hand, the parameter in our Weibull distribution was also estimated through maximum likelihood estimation (We have a known shape parameter equal to 1/2). The estimated scale parameter is equal to 2227.752. However, in our case, the p.d.f of the Weibull distribution was written in the form,

$$f(x) = \frac{c}{2}x^{-\frac{1}{2}}\mathrm{e}^{-c\sqrt{x}}, x > 0$$

Hence, the estimate for our parameter c is obtained by modifying the scale parameter, where $c = 2227.752^{-\frac{1}{2}} = 0.02118686$. The fitted curve is shown by the solid line in the figure below.

There is another fitted curve in this graph, which is drawn by the dashed line, and it is an exponential distribution fitting to the data. The two mixing distributions appear to fit much better than the exponential distribution. On the other hand, in order to see clearly how the curves are fitted to the data, corresponding QQ plots are presented as follows. The tail behaviours of the three distributions can be seen clearly and are consistent with what is discussed in Chapter 2.1.

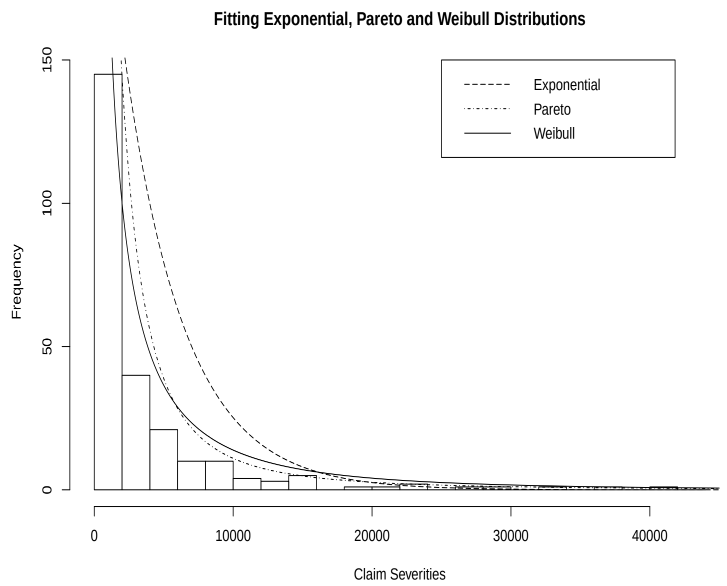


Figure 3.2: The Fitted Curves on the Data without the Outliers

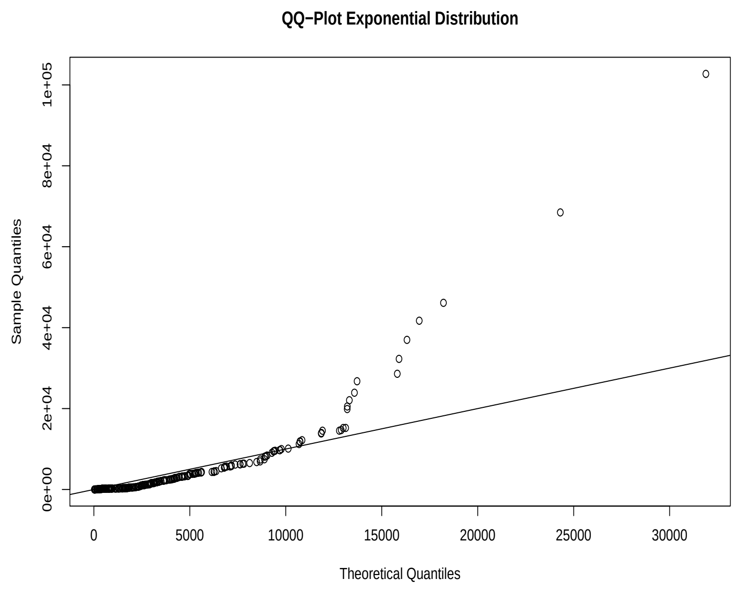


Figure 3.3: QQ-Plot of the Exponential Distribution versus the Sample Data

It is shown on this sketch that the exponential distribution fits the data relatively better up to around £12,500. At the tail where there are extremely large claim sizes, it fails to estimate the probability accurately.

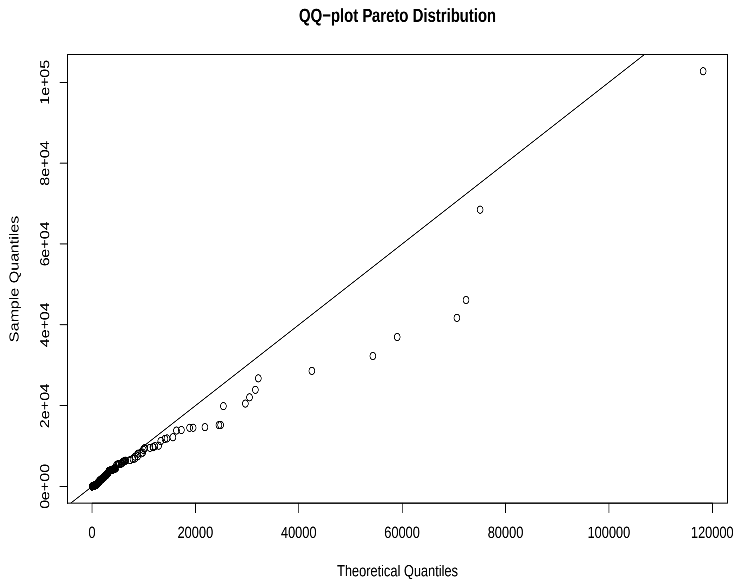


Figure 3.4: QQ-Plot of the Pareto Distribution versus the Sample Data

The first QQ plot (Figure 3.4) shows the goodness-of-fit of the Pareto distribution to our data. Several over-estimations for some small and medium sized claims are present. Nevertheless, it fits well especially for very expensive claims. This has emphasised its heavy-tail feature.

Figure 3.5 suggests that Weibull distribution fits very well up to £40,000, although there is slight perturbation. In the tail, it fits better than the Exponential distribution but worse than the Pareto distribution, as what is expected.

Overall, the exponential distribution does not perform well compared to the other two. While Weibull fits better for smaller claims, Pareto yields the best performance for very large claim sizes. From these plots, it is likely to suggest a mixture of strategies. When the claim sizes are moderate, the insurer is advised to adopt the Weibull claim severities. Particularly when reinsurance is in place, Weibull distribution can be the better choice. On the contrary, for very large claim sizes, Pareto distribution plays the key role due to its long-tail property.

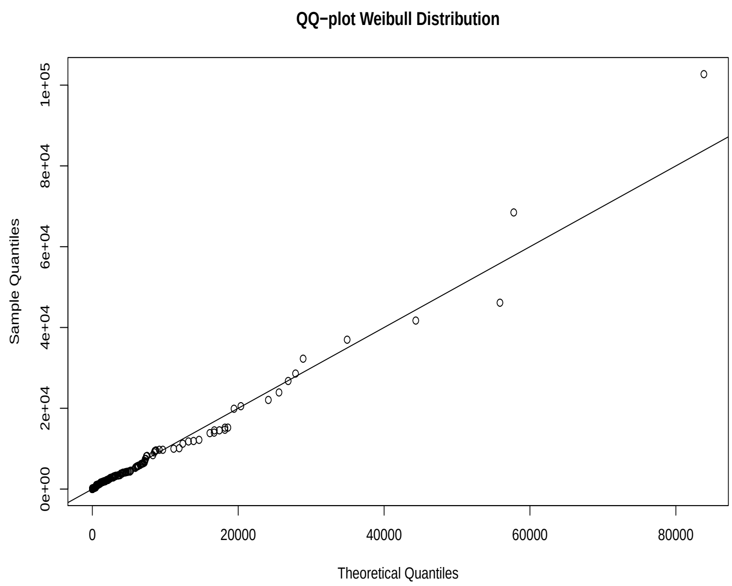


Figure 3.5: QQ-Plot of the Weibull Distribution versus the Sample Data

3.3.2 Calculations for Net Premiums

As mentioned before, the net premiums are calculated via the product of the expected claim frequency and the expected claim severity with independence between the two components assumed. Regarding the claim frequency component, we proceed as Frangos and Vrontos [2001]. Their estimates for the parameters α and τ are

$$\begin{cases} \alpha = 0.228, \\ \tau = 2.825. \end{cases}$$

In terms of the claim severity component, our first task is to estimate Q_K, as mentioned in the final part of last section. Since we have a recursive function for Q_K and its initial value is known as 1, this can be easily calculated. We used MATLAB to generate the solutions for $K = 1, 2, 3, 4, 5$ and $t = 1, 2, 3, 4, 5$. The first two tables underneath demonstrate the resulting premium rates for our Weibull model, and the other two show those for the Pareto model. The first column in each table denotes the scenario where no claims are reported to the company ($M = 0$). While we derive these values from (3.4), Frangos and Vrontos [2001] used the following formula to calculate the first column premium rates. Notice that tables with different total claim severities have the same first column.

$$P_{FV}|(M=0) = \frac{\alpha}{\tau + t} \cdot \frac{m}{s-1}.$$

Year	**Number of Claims**					
t	**0**	**1**	**2**	**3**	**4**	**5**
0	359.6	N/A				
1	265.6	2624.6	3082.1	3022.9	2856.7	2704.7
2	210.5	2080.6	2443.3	2396.4	2264.7	2144.2
3	174.4	1723.4	2023.9	1985.0	1875.9	1776.1
4	148.8	1470.9	1727.3	1694.2	1601.0	1515.8
5	129.8	1282.9	1506.6	1477.7	1396.4	1322.1

Table 3.2: Optimal Net Premiums with Weibull Severities and Total Claim Cost $M = 7,500$

Year	**Number of Claims**					
t	**0**	**1**	**2**	**3**	**4**	**5**
0	359.6	N/A				
1	265.6	3030.6	3735.4	3802.0	3677.7	3528.7
2	210.5	2402.5	2961.3	3014.0	2915.5	2797.4
3	174.4	1990.1	2452.9	2496.6	2415.0	2317.1
4	148.8	1698.5	2093.5	2130.8	2061.1	1977.6
5	129.8	1481.4	1826.0	1858.5	1797.7	1724.9

Table 3.3: Optimal Net Premiums with Weibull Severities and Total Claim Cost $M = 10,000$

First, let us look into details on our premiums. The upper table describes the premium levels for the situation where the accumulative claim costs are £7,500. The lower table gives the rates where the total claim amounts are £10,000. Overall, it follows the pattern that the premium decreases over time if claim frequencies are kept constant. How the BM system works will be illustrated in the next example. For instance, if a policyholder has a claim which costs £7,500 in his first policy year, his premium will be raised to £2624.6 (Table 3.2). If in the subsequent year, he has another claim whose severity is £2,500. The total accumulated number of claims is now 2 and the total size amounts to £10,000

(Table 3.3). He is then subject to pay £2961.3 in the next year. And if no more claims are filed in the following 3 years, his payment will reduce to £1826.0 from the beginning of year 6. Now it is essential to see how the BM system using the Pareto model works. Again the following two tables represent the total claim cost of £7,500 and £10,000 respectively.

Year	**Number of Claims**					
t	**0**	**1**	**2**	**3**	**4**	**5**
0	470.0	N/A				
1	347.1	2270.2	2361.3	2397.9	2417.6	2430.0
2	275.2	1799.7	1871.9	1900.9	1916.6	1926.4
3	227.9	1490.8	1550.6	1574.6	1587.6	1595.7
4	194.5	1272.3	1323.4	1343.9	1354.9	1361.9
5	169.7	1109.7	1154.3	1172.1	1181.8	1187.8

Table 3.4: Optimal Net Premiums with Pareto Severities and Total Claim Cost $M = 7,500$

Year	**Number of Claims**					
t	**0**	**1**	**2**	**3**	**4**	**5**
0	470.0	N/A				
1	347.1	2867.7	2982.7	3028.9	3053.9	3069.5
2	275.2	2273.3	2364.5	2401.2	2420.9	2433.3
3	227.9	1883.1	1958.6	1989.0	2005.3	2015.6
4	194.5	1607.2	1671.6	1697.5	1711.5	1720.3
5	169.7	1401.8	1458.0	1480.6	1492.8	1500.4

Table 3.5: Optimal Net Premiums with Pareto Severities and Total Claim Cost $M = 10,000$

For the same insured we described before, this system will in general punish less severely than the previous one. Specifically, this customer will pay £470 at the start. Due to one claim reported in the first year, his premium rate is raised to £2270.2 (Table 3.4). Then after a second claim in the subsequent year, an increase to £2364.5 occurs (Table 3.5). If he has no more claims till the end of the 5^{th} year, his premium payment could go down to £1458.0. Up to now, the flow of the two systems appears to be similar except that the punishment is less severe in the latter one.

However, for this dataset, there is an unexpected finding of our results. Unlike the Pareto fitting, when the total claim size is fixed, our premium is not strictly increasing with the number of claims, but starts to drop slightly for more than 2 or 3 claims in our example. The last few columns in Table 3.2 and Table 3.3 have demonstrated this phenomenon. In order to see how our premium rates behaves when compared to the Pareto model. We have plotted the premiums for large quantities of claims ($K = 100$) when the total claim severity is kept unchanged (Figure 3.6). Again, two cases where the total cost of claims is £$7,500$ and £$10,000$ respectively are analysed. This irregular behaviour of the severity component of this premium formula may be different depending on the nature of the dataset to which the Weibull model is applied because the monotonicity of our premium function is affected by the parameters as mentioned in section 3.2.

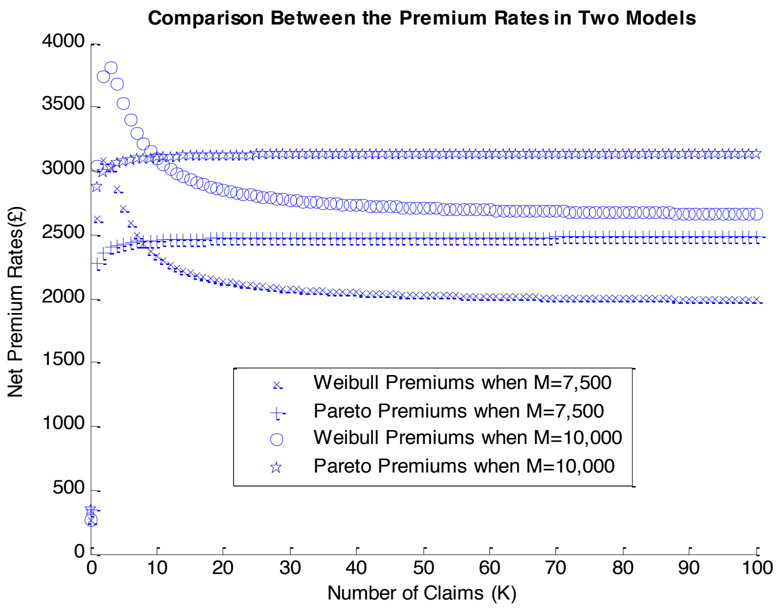


Figure 3.6: Behaviours of Premium Rates with respect to the Number of Claims when the Total Claim Sizes is £$7,500$ and £$10,000$

As presented in Figure 3.6, under a fixed total claim cost, our premium function is not strictly increasing with the number of claims. It reaches a peak at some point and then decreases and finally asymptotically converges. To our knowledge so far, this behaviour has not appeared in classical models. However, with this effect, the bonus hunger problems could in some way be alleviated. As seen in our previous tables, if one client has already claimed twice with a total cost slightly less than £7,500 in one particular year. In the system using Pareto claim severities, it is very likely that he will bear the cost himself if he has one more small accident and does not disclose this information to the insurer because of otherwise the growth in premium payment. However, in our system, there will be a reward if he reveals the true information to the company for a further small claim. It is where the premiums start to decrease. Notice that our premiums only drop a little when the additional claim size is not very large. If one has one more severe claim, he will be paying more than the current amount. Hence, our model is helpful to encourage the insured to be truthful to insurance companies. In this way, it is a flexible decision for the insurer to either fix the premium after a certain level of claim frequency or to reward the policyholders for disclosing true information about cheap claims.

On the other hand, this model actually distinguishes the premium payments more according to various claim costs. We punish more on those who make few large claims than the ones who report many small claims. An extreme example could be seen in Figure 3.6. When comparing an insured person who makes one claim worth £7,500 and the other who has 10 claims summing up to £7,500, which means he has an average of £750 for each claim, we could see that the former one is paying a similar amount to the latter one or even slightly more. Obviously this is not the case in the Pareto model. This observation implies that our model emphasises more on the claim severity component while the Pareto model addresses more on the frequency level.

It is also noticeable that our initial premium payments are lower than the Pareto model which might be more preferable to starting policyholders thus creating more competitiveness to the insurer. Hence, we proposed a slightly different pricing strategy for insurance companies here. This additional option for insurers leads to a diversity of premium strategies where a suitable one could be chosen from to adapt to certain needs.

As we have seen from our results, the BM system adopting the Weibull severities provides a lower entrance premium level and punish more on people who have large sized claims and less on those who make plenty of small claims. This could be considered as one of the options when insurers are deciding on pricing strategies. We would suggest our model to be adopted by those who prefer to offer a mild treatment to policyholders with many small claims. In practice, it is reasonable and this kind of strategy helps the alleviation of hunger for bonus phenomenon. Therefore, it is suggested that an insurance company could always consider all these factors when choosing among models. Sometimes a mixture of the models would be the most rational choice, which will be introduced in the next section.

3.4 A hybrid Model

Inspired by previous results, we assume that the claims which cost less than a threshold z are distributed according to a Weibull and those whose sizes are over z conform to a Pareto distribution. For convenience, we will refer the former kind as 'small claims' and the latter as 'big claims' in the sequel.

Based on the thought of implementing hybrid distributions on the claim size modelling, we can simply write our premiums in the following form which is a linear combination of the two expected aggregate claim costs.

$$Premium = \Pi[X]\Pi[N] = \Pi[X_w]\Pi[N_w](1-\rho) + \Pi[X_p]\Pi[N_p]\rho$$

where $\Pi[\cdot]$ denotes the posterior mean because we also applied Bayesian analysis under this circumstance. Here X denotes the claim severity, N the number of claims and ρ is the probability of being above the threshold.

Corollary 3.2. *Suppose the number of claims N is a negative binomial distributed random variable with p.m.f*

$$P(N=n) = \binom{n+\alpha-1}{n}\left(\frac{\tau}{1+\tau}\right)^{\alpha}\left(\frac{1}{1+\tau}\right)^{n}.$$

Distinguishing the claims by a limiting amount z results in two random variables counting the corresponding frequency in each category. Then one of the decomposed random variables follows a $NB(\alpha, \tau/\rho)$ and the other a $NB(\alpha, \tau/(1-\rho))$, where ρ relates to the probability of the categorisation.

Proof. We know that the number of claims is assumed to be Negative Binomial distributed, i.e., $N \sim NB(\alpha, \tau)$. Firstly, we denote the frequency of large claims by N_p and it can be written that,

$$N_p = I_1 + I_2 + \ldots + I_N, where\ I \sim Bernoulli(\rho).$$

That means,

$$\begin{aligned} Pr(I_j = 1) &= Pr(X_j > z) = 1 - F_X(z) = \rho; \\ Pr(I_j = 0) &= Pr(X_j \leq z) = F_X(z) = 1 - \rho. \end{aligned}$$

Its moment generating function (m.g.f) is then given by,

$$m_I(s) = E[\mathrm{e}^{sI}] = \rho\mathrm{e}^s + (1-\rho)\mathrm{e}^0 = 1 - \rho + \rho\mathrm{e}^s.$$

Hence, the m.g.f of N_p is computed as

$$\begin{aligned} m_{N_p}(s) &= P_N(m_I(s)) \\ &= \left(\frac{\tau}{\tau + \rho - \rho\mathrm{e}^s}\right)^\alpha \\ &= \left(\frac{\frac{\tau}{\tau+\rho}}{1 - \left(1 - \frac{\tau}{\tau+\rho}\right)\mathrm{e}^s}\right)^\alpha, \end{aligned}$$

where P_N represents the probability generating function (p.g.f) of N. It is thus clear that N_p is still Negative Binomial distributed with a new parameter, i.e., $N_p \sim NB(\alpha, \tau/\rho)$. Therefore, the mean claim frequency of large sized claims is, $E[N_p] = \frac{\alpha\rho}{\tau}$. Similarly, we can obtain that the distribution of N_w is also a Negative Binomial, i.e., $N_w \sim NB(\alpha, \tau/(1-\rho))$ and its mean is $E[N_w] = \frac{\alpha(1-\rho)}{\tau}$.
■

Therefore,

Corollary 3.3. *The posterior means of the random variables N_w and N_p are* *given by*

$$\begin{aligned}\Pi[N_p] &= \frac{\alpha+K_1}{\tau/\rho+t};\\ \Pi[N_w] &= \frac{\alpha+K_2}{\tau/(1-\rho)+t}.\end{aligned}$$

where K_1 and K_2 represent the number of small and large claims respectively.

Proof. Replacing all the τ with τ/ρ and $\tau/(1-\rho)$ respectively which does not affect the integration will lead to the results as claimed above. ■

Corollary 3.4. *The posterior expectation of the size of large claims is*

$$\Pi[X_p] = \frac{1}{\rho}\left(\frac{M_2+m}{M_2+m+z}\right)^{K_2+s}\left(z+\frac{M_2+m+z}{K_2+s}\right);$$

And the posterior expectation of that for small claims is given by

$$\begin{aligned}\Pi[X_w] &= \frac{1}{1-\rho}\frac{2\sqrt{M_1}}{c}\frac{B_{v-1}(c\sqrt{M_1})}{B_v(c\sqrt{M_1})}\\ &- \frac{1}{1-\rho}\left(\frac{M_1}{M_1+z}\right)^{\frac{v}{2}}\left[z\frac{B_v(c\sqrt{M_1+z})}{B_v(c\sqrt{M_1})}+\frac{2\sqrt{M_1+z}}{c}\cdot\frac{B_{v-1}(c\sqrt{M_1+z})}{B_v(c\sqrt{M_1})}\right].\end{aligned}$$

where $v = K_1 - \frac{1}{2}$ and $K_1, M_1 > 0, K_2, M_2 \geq 0$.

Proof. Since both distributions are the results of a mixing over exponential. We initially compute the conditional expectation of the exponential for each segment.

$$E[X|X\leq z] = \frac{1}{1-\rho}\left[\frac{1}{\theta}-\left(z+\frac{1}{\theta}\right)e^{-\theta z}\right]; \tag{3.6}$$

$$E[X|X> z] = \frac{1}{\rho}\left(z+\frac{1}{\theta}\right)e^{-\theta z}. \tag{3.7}$$

Since when $X \leq z$, we have the posterior distribution for θ as follows Ni et al. [2014a],

$$\pi_1(\theta) = \frac{\left(\frac{c}{2\sqrt{M_1}}\right)^{-\left(K_1-\frac{1}{2}\right)}\theta^{K_1-\frac{3}{2}}exp\left(-\left(M_1\theta+\frac{c^2}{4\theta}\right)\right)}{2B_{K_1-\frac{1}{2}}(c\sqrt{M_1})}. \tag{3.8}$$

Integrating (3.7) and (3.8) yields the desired result. Only one part of the integration is illustrated as shown below since other parts are computed in a very similar way.

$$\begin{aligned} I_1 &= \int_0^\infty \frac{1}{\theta} e^{-\theta z} \pi_1(\theta) d\theta = \left(\sqrt{\frac{M_1}{M_1+z}}\right)^{K_1-\frac{1}{2}} \frac{2\sqrt{M_1+z}}{c} \cdot \frac{1}{B_{K_1-\frac{1}{2}}(c\sqrt{M_1})} \\ & \int_0^\infty \left(\frac{2\sqrt{M_1+z}}{c}\theta\right)^{K-\frac{5}{2}} exp\left(-\frac{1}{2}c\sqrt{M_1+z}\left(\frac{2\sqrt{M_1+z}}{c}\theta + \frac{c}{2\sqrt{M_1+z}\theta}\right)\right) d\left(\frac{2\sqrt{M_1+z}}{c}\theta\right) \\ &= \left(\sqrt{\frac{M_1}{M_1+z}}\right)^{K_1-\frac{1}{2}} \frac{2\sqrt{M_1+z}}{c} \cdot \frac{B_{K_1-\frac{3}{2}}(c\sqrt{M_1+z})}{B_{K_1-\frac{1}{2}}(c\sqrt{M_1})}. \end{aligned}$$

The posterior distribution of θ for the second component, i.e., when $X > z$, is given by,

$$\pi_2(\theta) = \frac{\theta^{K_2+s-1} e^{-(M_2+m)\theta} (M_2+m)^{K_2+s}}{\Gamma(K_2+s)}. \tag{3.9}$$

Integrating (3.7) and (3.9) will lead to the value as claimed in Corollary 3.4. ■

Henceforth, our premium function is thus modified to

$$Premium = \frac{\alpha+K_1}{\tau/(1-\rho)+t} \cdot \Pi[X_w](1-\rho) + \frac{\alpha+K_2}{\tau/\rho+t} \cdot \Pi[X_p]\rho. \tag{3.10}$$

where $\Pi[X_w]$ and $\Pi[X_p]$ are as discussed in Corollary 3.4. In order to find the premium values, unlike in 3.2, we need not only the sum of total claim severity but also the respective sums of the severities of each kind of claims. Similarly, it is sufficient to know the number of small claims K_1 and the number of large claims K_2. Clearly we have $M = M_1 + M_2$ and $K = K_1 + K_2$ where M and K denote the total size and total number of claims for an individual over a period of t years.

3.5 Parameter Computations

In this section, we do not use conventional methods to do the parameter estimations. Instead, by using some properties of our model and under the following reasonable assumptions, we estimate the parameters in a special way.

Assumptions

1. The probability density function $f(x)$ for the claim sizes can be represented as follows

$$f(x) = \begin{cases} \frac{c}{2}x^{-\frac{1}{2}}\mathrm{e}^{-c\sqrt{x}} & , \quad 0 \leqslant x \leqslant z, \\ \frac{sm^s}{(x+m)^{s+1}} & , \quad x > z, \end{cases}$$

   with $f(z^-) = f(z^+)$. We assume that it is continuous but may not be differentiable at z. Hence, the probability of a given claims size which is less or more than the threshold value z can be written respectively by

$$\begin{aligned} Pr(X \leqslant z) &= F(z^-) = \int_0^z \frac{c}{2}x^{-\frac{1}{2}}\mathrm{e}^{-c\sqrt{x}}; \\ Pr(X > z) &= 1 - F(z^+) = \int_z^\infty \frac{sm^s}{(x+m)^{s+1}}. \end{aligned}$$

2. Both z and ρ are observations from the sample data. z is computed from the intersection of the Weibull and Pareto distributions as in Section 3.2. ρ is the proportion of claims in the portfolio that is over the value z. They could be obtained by running simulations, but we will illustrate our model by using one dataset.

3. The estimation of the parameters for the claim frequency component is based on the maximum likelihood method as often seen in literature. We use the same results as in Frangos and Vrontos [2001], namely

$$\begin{cases} \alpha = 0.228, \\ \tau = 2.825. \end{cases}$$

The Scale Parameter in Weibull Distribution Based on the fact that the proportion of claims which are below the threshold z is $1-\rho$, we can write

$$\int_0^z \frac{c}{2}x^{-\frac{1}{2}}\mathrm{e}^{-c\sqrt{x}}\mathrm{d}x = 1 - exp(-c\sqrt{z}) = 1 - \rho. \tag{3.11}$$

The Parameters in Pareto Distribution Similarly, the rest of the portfolio is ρ and thus,

$$\int_z^\infty \frac{sm^s}{(x+m)^{s+1}}\mathrm{d}x = \left(\frac{m}{z+m}\right)^s = \rho.$$

$$\begin{cases} \bar{F}(z) = \left(\frac{m}{z+m}\right)^s, \\ f(z^-) = f(z^+). \end{cases}$$

$\Leftrightarrow$

$$\begin{cases} \left(\frac{m}{m+z}\right)^s = \rho, \\ \frac{c}{2} z^{-\frac{1}{2}} e^{-c\sqrt{z}} = \frac{sm^s}{(z+m)^{s+1}}. \end{cases} \tag{3.12}$$

(3.11) and (3.12) will then be able to form a system of equations where all the parameters c, m and s can be calculated with known values of z and ρ.

3.6 Numerical Illustration II

Again the data was sourced from Klugman et al. [1998] and details could be seen in Appendix A. This serves as an example of illustration how our hybrid model works.

Applying the model on the same dataset as we have used before yields the following results. Initially, by observation, the value of z and ρ are observed as explained above and they are $z = 5784.47, \rho = 0.184$. Substitute these into (3.11), c can be easily obtained which is $c = 0.02225763$.

Subsequently, following the same steps to estimate the parameters in the Pareto distribution and combining (3.11) and (3.12), the estimates of m and s can be calculated as

$$\begin{cases} m = 1475.0447, \\ s = 1.0622451. \end{cases}$$

with the help of the values of z and ρ.

Hence, figure 3.7 shows a hybrid distribution fitting to the dataset. Compared to our former results, this yields a better fit.

Premium Calculations Again we analysed several scenarios similar to Section 3.2. But now we need to specify the number of small and large claims and fix the corresponding total costs accordingly. Note that since we have a threshold $z = 5784.47$, a reasonable setting of the scenarios should satisfy several constraints.

$$M_1 \leqslant min\{zK_1\}, \ M_2 \geqslant min\{zK_2\}, \ K_1, \ K_2 \in \mathbb{N}.$$

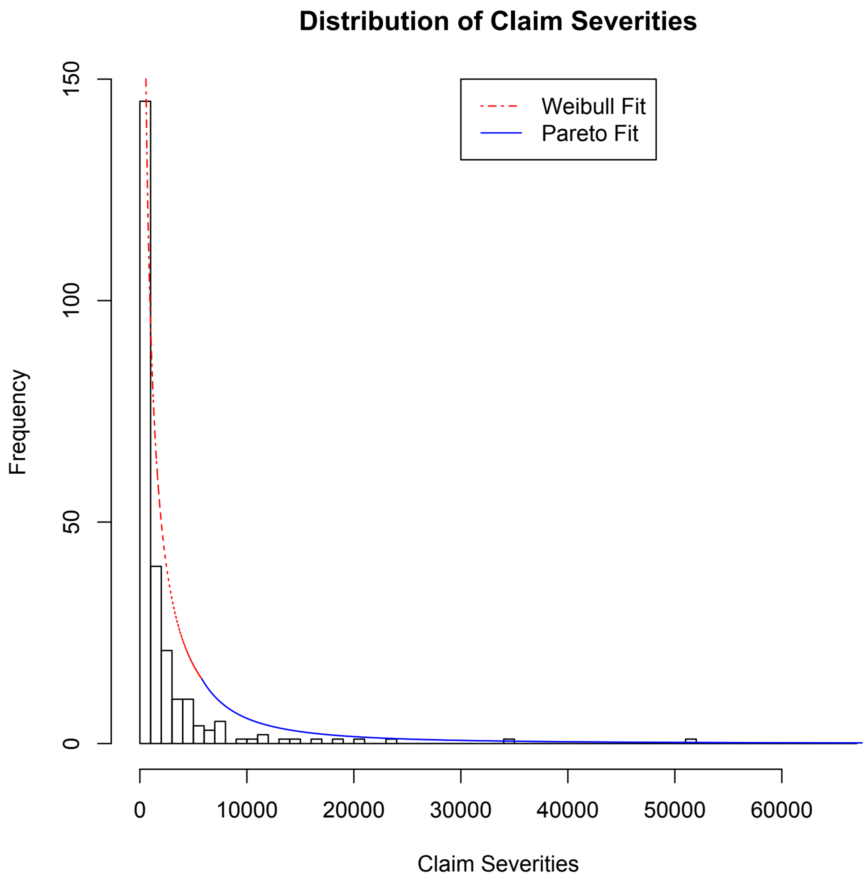


Figure 3.7: Distribution of Claim Severities.

For instance, if K_1 ranges from 1 to 5, $M_1 \leqslant z$ and similarly if $K_1 \geqslant 2$, $M_1 \leqslant 2z$. Thus, we will be looking at several special scenarios as described below.

Scenario 1 One simple scenario is that policyholders only make small claims and no big ones, i.e., $K_2 = M_2 = 0$. Fixing the total cost $M_1 = 5000$ and employing (5.4) would yield the following results as shown in Table 3.6.

Scenario 2 On the contrary, we also consider a case where a policyholder has only big claims and no small ones, i.e., $K_1 = M_1 = 0$. Keeping the total large claim size fixed at $M_2 = 30000$, we obtained Table 3.7. It is essential to mention it here that since (5.4) is not defined when $M_1 = 0$, we redefine the expected

Table 3.6: Premiums for Scenario 1

$M_1 = 5000, M_2 = K_2 = 0, t = 1, 2$

	$K_1 = 1$	2	3	4	5
$t = 1$	707.907	1061.181	1346.342	1530.187	1628.134
$t = 2$	617.699	906.294	1139.247	1289.434	1369.448

Table 3.7: Premiums for Scenario 2

$M_1 = K_1 = 0, M_2 = 30000, t = 1$

	$K_2 = 1$	2	3	4	5
$t = 1$	1471.1	1665.3	1694.0	1653.0	1577.8
$t = 2$	1360.4	1543.4	1570.5	1531.8	1461.0

claim size for small claims using,

$$\Pi[X_w|M_1 = 0] = \frac{1}{1-\rho} \cdot \frac{2}{c^2}$$

Scenario 3 This is when a policyholder has £5000 worth small claims and £30000 worth big claims in a single year. We analysed 25 different scenarios where the number of both the small and large claims vary from 1 to 5. And for simplicity, we only look at the first two years' premium levels. Results are demonstrated in Table 3.8 and Table 3.9 respectively. A clearer comparison could be seen in Figure 3.8 and Figure 3.9.

We have actually looked at premium levels for claimers only. Generally speaking, premiums decrease overtime. According to Table 3.6 and Table 3.7, it is also obvious that a policyholder with both small and large claims would definitely pay more than those who have small claims only knowing that their costs of small claims are the same. Furthermore, by comparing Scenario 2 and 3, each additional small claim adds quickly on the premium levels.

Moving into details, clearly the premium levels jump upwards column-wise. In a practical language, when the aggregate costs of claims are fixed, the higher

Table 3.8: Premiums for Scenario 3 $t = 1$

$M_1 = 5000, M_2 = 30000, t = 1$

	$K_2 = 1$	2	3	4	5
$K_1 = 1$	1658.7	1852.8	1881.6	1840.5	1765.4
2	2011.9	2206.1	2234.9	2193.8	2118.7
3	2297.1	2491.3	2520.0	2479.0	2403.9
4	2480.9	2675.1	2703.9	2662.8	2587.7
5	2578.9	2773.1	2801.8	2760.8	2685.6

Table 3.9: Premiums for Scenario 3 $t = 2$

$M_1 = 5000, M_2 = 30000, t = 2$

	$K_2 = 1$	2	3	4	5
$K_1 = 1$	1513.7	1696.7	1723.7	1685.1	1614.3
2	1802.3	1985.3	2012.3	1973.7	1902.9
3	2035.2	2218.2	2245.3	2206.6	2135.8
4	2185.4	2368.4	2395.5	2356.8	2286.0
5	2265.4	2448.4	2475.5	2436.8	2366.0

their frequency is the cheaper is each claim. The increase in premiums suggests that this system punishes severely on people who frequently make small claims. It is also noticeable that the increase of premiums with respect to K_1, i.e., the number of small claims, is faster than that with regard to K_2, i.e., the number of large claims. In fact, premiums almost stay quite stable with the change in K_2. There is even a decreasing trend starting from the 4th column in Table 3.7-3.9. However, this does not mean the rise in claim frequency would lead to lower premium level. Notice that the total claim size is fixed. So when the counts of large claims move towards right, it only implies that each individual claim actually costs less. Such reduction in premiums could be understood as when it comes to large claims, the system punishes less if the frequency is high. Figure 3.8 and 3.9 reinforces this statement and notice that the premium does not drop much.

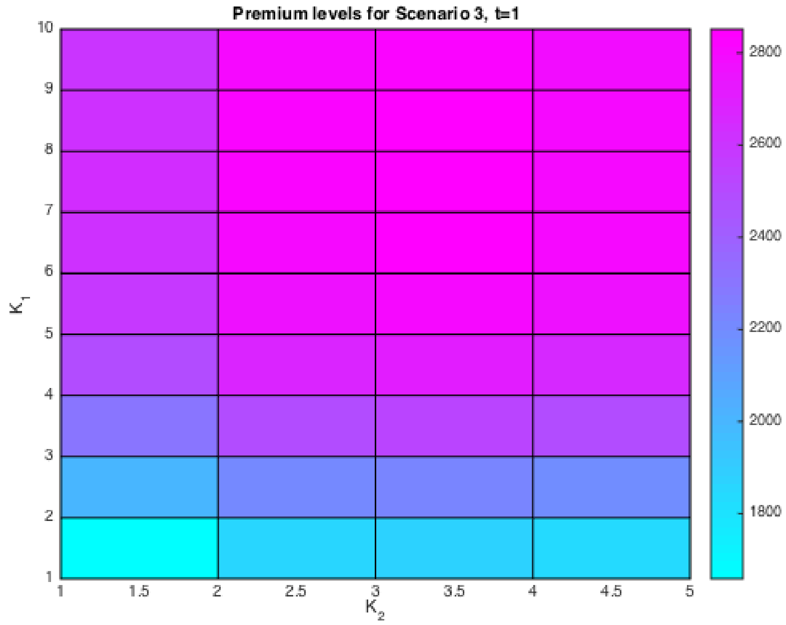


Figure 3.8: Premiums for Scenario 3 $t = 1$.

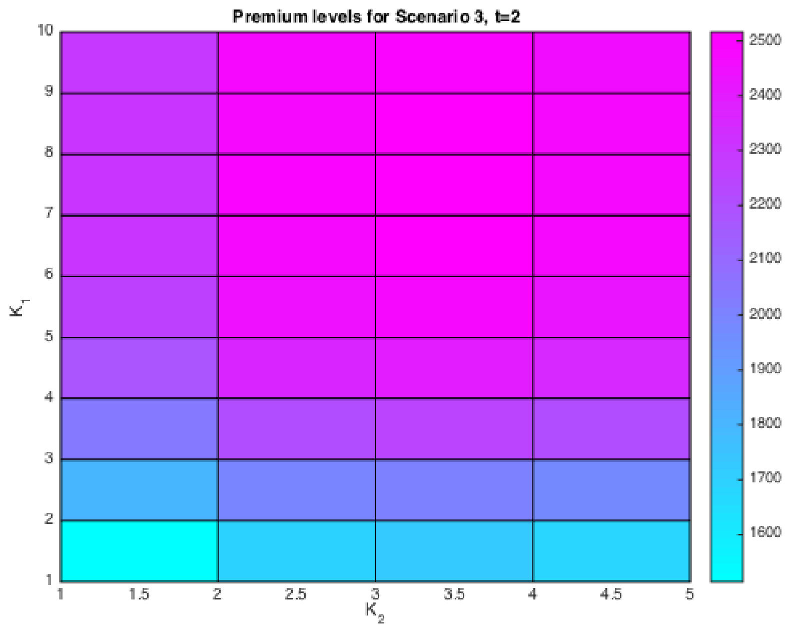


Figure 3.9: Premiums for Scenario 3 $t = 2$.

This means that our model makes more emphasise on the total claim severity rather than the claim frequency component for large claims and vice versa for small ones. In other words, the proposed model punishes more on someone making a lot of small claims while such punishment is not obvious when large claims are made. On the other hand, when the frequency of large sized claims is raised, the per-claim cost actually is smaller which might fall into our smaller size claims category. For instance, the K_2 cannot increase further after a sum of 5 claims worth £$30,000$ in total (Table 3.7, 3.8, 3.9), because otherwise the average size of the claims would fall below z and will be reconsidered as small claims. We penalise less because the drivers having claims valued near the threshold are actually not affecting the income of a company too much and in addition they have informed the insurer about their claims. Such information is valuable in estimation and forecast. Technically speaking, this is due to the fact that $\frac{M_2+m}{M_2+m+z}$ is between $(0,1)$ as can be seen in $\Pi[X_p]$ in Corollary 3.4.

However, for small claims, the premium still rises as the frequency increases when the total costs are kept constant. These behaviours can also be seen from Figure 3.8 and Figure 3.9, the colour gets warmer when moving upwards on the K_1 axis. Practically speaking, one reason would be that frequently dealing with small claims would probably induce more administrative costs which should be offset by forcing higher premiums. In addition, it is very likely that people with many small claims would create a big loss in the future. They are potential risks and likely to cause a sudden loss to the insurer.

That is to say, this model assigns more attention on potential risks and is relatively milder in penalising those who already reported a larger claim. In fact, it is often the case that these people would be more careful in the future, while those constantly filing small claimers possibly have a potential to create an unexpected attack to the insurance company.

Another interesting question to ask is how this model compares to the previous one in Section 3.2. So let us look at the same example where a policyholder reports one claim worth £7500 and then £2500 in the subsequent year. That means a 'big' claim in the first and a 'small' in the second year using the hybrid model and the resulting premium for the second year should be £554.7 and £646.3 for the third year. These figures are much smaller than those of both the Weibull and

Pareto models. It could be the reason that the fitting of this hybrid distribution performs better for small claims and also the data does not contain many large claims.

3.7 Summary

In conclusion, this chapter first extended Frangos and Vrontos [2001] by choosing a different severity distribution. They worked on both the a priori and the a posteriori information, whereas we analysed the BM system only considering the a posteriori information. For the latter case, while Frangos and Vrontos [2001] adopted Negative Binomial distribution for to describe the claim frequencies and Parteo distribution to model the claim severity component, this work maintains other modelling factors and only alters the claim severity distribution to Weibull. By comparing the two models when applied to the same dataset, although we provide worse estimation in the tail, we offer cheaper initial premiums and more reasonable especially to those who claim many times but with small severities. Furthermore, on this dataset our model seems to discourage the hunger for bonus phenomenon, which is empirically significant.

Then based on these results, the idea of a mixed strategy for claim severity distributions was adopted. Under the given dataset, a similar trend can be witnessed in the large claims zone, i.e., the increase in frequencies not necessarily gives rise to that in premiums when the total expenses are fixed. It seems that as long as the total costs for large claims are kept constant, the system is kind with the increase in frequency for these claims. However, it holds the opposite attitude within the small claims zone. Notice that penalties on the increasing number of large claims have an upper limit due to the fact that a rising frequency for these claims would possibly mean claims fall into the 'small' categorisation again.

Future extensions are possible with directions including study on the a priori information (regression analysis), sensitivity analysis as well as improved parameter estimation techniques. Moreover, one consideration on modelling would be to incorporate the deductibles as a more realistic case. This piece of work has laid a foundation on BM systems and facilitated studies in Bayesian inference, enough to provide a reasonable introduction to the following researches.

Chapter 4

Risk Analysis of a BM system in discrete risk models

This chapter serves as the first step into the risk theory analysis world. For sure, it is a very popular topic not only among the academics but also practitioners. Normally, in insurance industry, one good measure for solvency/insolvency is the ruin probability, which has already been introduced in the preliminaries. Just to ring a bell, it describes how likely it is that an insurer does not have adequate funds to cover claims. Taking an initial move, this chapter looks at a discrete model for a BM system with three classes. For simplicity, we assume all monetary terms here including premium income, claims sizes as well as the initial capital to be integers only. Although it is an unrealistic assumption, this piece of work helped to establish insights into industry and comprehension in risk models. The construction of the BM system is inspired by an industrial partner. It suggests the use of a BM system for reinsurance companies, resembling a merit rating to customers who are insurance companies under this concern. This also interprets the BM idea into a collective risk model where the ruin probability is well-defined. This work was done in collaboration with Dr. Bo Li[1] from Nankai University.

[1]School of Mathematical Sciences, Nankai University, Tianjin 300071, China

4.1 Model Set-up

The motivation of this piece of work originates from a reinsurance company who would like to employ the BM idea in order to deal with catastrophic risks. For a peaceful period, ceding companies pay the base reinsurance premium. When a catastrophic event occurs, the re-insurer raises the premiums by 50% for the subsequent year. Such events are revealed by the aggregate claims observed in the previous year. If the aggregate amount of claims in a year is below the value of concern, then the reinsurance premium returns to the base level. It also could happen that no claims are reported for a consecutive 3 years from all the cedents being insured. Then a 50% discount is offered. In this way, the re-insurer could gain its competitiveness in the market during 'good' years whereas the same time acquires the security when extreme events happen. Notice that rather than an adjustment according to individual histories as in the BM systems used in car insurance, this model relies more on the whole external environment and provides mutual bonuses and maluses on the premium levels for all ceding companies. In the following contents, we use a BM system to model this idea. So when it comes to moving to a class, it implies the environment changes and the reinsurance premium levels for all ceding companies vary at the same time.

The BM system we are trying to model consists of 3 classes, $S1$, $S2$ and $S3$. Correspondingly, three premium levels assigned to each of these classes are 50%, 100% and 150% respectively. The transition rule is stated as follows. Under normal circumstances, ceding companies pay the base reinsurance premium which is 100%. With 3 consecutive years of no claims, ceding companies are able to jump to $S1$, paying 50% of the base premium. However, if someone has claims in a single year with the total severity exceeding a predetermined threshold R, this customer will move to $S3$ and pay 150% of the base premium. This system is relatively mild in the sense that it only punishes customers who report catastrophic claims. Of course, such system is not limited to auto-mobile insurance.

According to the description, by assuming that claims made in each year are independent from each other, we are able to establish the following Markov chain $\{J_n\}_{n\in\mathbb{N}}$, where $J_0, J_1 \ldots$ are states occupied in each policy year for ceding companies. In this Markov chain, we recognised 5 states and they are:

0. One or more claims made in the preceding year but the aggregate amount is not big enough to be considered as a catastrophe;
1. No claims in the preceding year, but claims made in the second last year;
2. No claims in the last 2 consecutive years period, but claims made in the third last year;
3. No claims made in the last 3 consecutive years period;
4. Aggregate claims exceeding the threshold in the preceding year.

Translating into the mathematical language, we write $\{J_n\}_{n\in\mathbb{N}}$ with state space $E = \{0, 1, 2, 3, 4\}$. Let q denote the probability of cedents making no claims and p the probability of making claims with aggregate amount less than R, whereas r represents the probability of having huge claims in a single policy year. Then, the transition probability matrix of $\{J_n\}_{n\in\mathbb{N}}$ is given by

$$\mathbf{P} = \begin{array}{c} \\ \boxed{0} \\ \boxed{1} \\ \boxed{2} \\ \boxed{3} \\ \boxed{4} \end{array} \begin{array}{c} \begin{array}{ccccc} \boxed{0} & \boxed{1} & \boxed{2} & \boxed{3} & \boxed{4} \end{array} \\ \begin{pmatrix} p & q & 0 & 0 & r \\ p & 0 & q & 0 & r \\ p & 0 & 0 & q & r \\ p & 0 & 0 & q & r \\ p & q & 0 & 0 & r \end{pmatrix} \end{array}.$$

Obviously, we have $p + q + r = 1$. Since $\mathbf{P}$ is irreducible and aperiodic, there exists a unique stationary distribution $\boldsymbol{\pi} = (p, q(p+r), q^2(p+r), q^3, r)$. We make the simplest assumption that premium levels take values 1, 2, 3 respectively with regard to the three different percentages as described at the beginning of this section. Specifically, the current premium level at each state is displayed below.

States	$\boxed{0}$	$\boxed{1}$	$\boxed{2}$	$\boxed{3}$	$\boxed{4}$
Premiums	2	2	3	1	3

For example, if the current state is 0, i.e., ceding companies had claims last year which cost less than the threshold in total, then the current premium level is 2 which is the base premium.

Let us now consider a discrete risk surplus process for an individual from the company's perspective.

$$U(n) = u + \sum_{i=1}^{n} C_i - \sum_{i=1}^{N(n)} Z_i, \; n = 0, 1, \dots, \tag{4.1}$$

where $u \in \mathbb{N}$ is the initial capital reserved for this individual and C_i here is a random variable denoting the premium payment received in the i^{th} year. $N(n) = \sum_{i=1}^{n} \mathbb{1}\{Z_i > 0\}$ represents the total number of years with claims (either normal or extreme ones) with the i.i.d random variable Z_i describing the aggregate claim size in the i^{th} year. For convenience, we define Z in the following way.

$$Z \overset{d}{=} \begin{cases} X, \; \mathbb{P}(0 < Z \leq R) = p; \\ Y, \; \mathbb{P}(Z > R) = r; \\ 0, \; \mathbb{P}(Z = 0) = q, \end{cases}$$

recalling that R is the threshold we set to distinguish normal from extreme claims. X and Y have distribution functions $F(x)$ with mean μ_F and $G(x)$ with mean μ_G, respectively. Furthermore, we define the ruin time by

$$T(u) = \inf\{n : U(n) < 0 | U(0) = u\},$$

and the ruin probability in this case by

$$\psi_i(u) = \mathbb{P}(T(u) < \infty | J_0 = i), \;\; i \in E, \tag{4.2}$$

where J_0 stands for the starting state as mentioned before. Thus, corresponding survival probabilities are

$$\Phi_i(u) = 1 - \psi_i(u), \;\; i \in E. \tag{4.3}$$

Since the claims are assumed to be i.i.d, a safety loading condition needs to be satisfied.

$$p(2-\mu_F)+q(p+r)2+q^2(p+r)2+q^3+r(3-\mu_G) = 2+r-q^3-(p+r)\mu > 0, \quad (4.4)$$

where $\mu = \frac{p}{p+r}\mu_F + \frac{r}{p+r}\mu_G$.

4.2 Recursive relations

As this discrete process also renews at the end of each policy year, it is reasonable to establish the following relations similar to a classical model, except that at each step the increment has several possibilities. Further implementing the law of total probability, the survival probabilities $\Phi_i, i \in E$ for $u \in \mathbb{N}$ can be written as,

$$\Phi_0(u) = pE[\Phi_0(u+2-X)] + q\Phi_1(u+2) + rE[\Phi_4(u+2-Y)] \quad (4.5)$$
$$\Phi_1(u) = pE[\Phi_0(u+2-X)] + q\Phi_2(u+2) + rE[\Phi_4(u+2-Y)] \quad (4.6)$$
$$\Phi_2(u) = pE[\Phi_0(u+2-X)] + q\Phi_3(u+2) + rE[\Phi_4(u+2-Y)] \quad (4.7)$$
$$\Phi_3(u) = pE[\Phi_0(u+1-X)] + q\Phi_3(u+1) + rE[\Phi_4(u+1-Y)] \quad (4.8)$$
$$\Phi_4(u) = pE[\Phi_0(u+3-X)] + q\Phi_1(u+3) + rE[\Phi_4(u+3-Y)] \quad (4.9)$$

Pairing equations (4.5) with (4.9) and (4.7) with (4.8) suggests that

$$\Phi_4(u) = \Phi_0(u+1) \quad \text{and} \quad \Phi_2(u) = \Phi_3(u+1).$$

Additionally, it can be observed from (4.5)-(4.7) that

$$\begin{aligned} & pE[\Phi_0(u+2-X)] + rE[\Phi_4(u+2-Y)] \\ = \; & \Phi_0(u) - q\Phi_1(u+2) \\ = \; & \Phi_1(u) - q\Phi_2(u+2) \\ = \; & \Phi_3(u+1) - q\Phi_3(u+2) \quad \text{for } u \geq 0 \end{aligned}$$

Then we have, in terms of $\Phi_3(\cdot)$, for $u \geq 0$

$$\begin{cases} \Phi_0(u)=\Phi_3(u+1) - q\Phi_3(u+2) + q\Phi_3(u+3) - q^2\Phi_3(u+4) + q^2\Phi_3(u+5) \\ \Phi_1(u)=\Phi_3(u+1) - q\Phi_3(u+2) + q\Phi_3(u+3) \\ \Phi_2(u)=\Phi_3(u+1) \\ \Phi_4(u)=\Phi_3(u+2) - q\Phi_3(u+3) + q\Phi_3(u+4) - q^2\Phi_3(u+5) + q^2\Phi_3(u+6) \end{cases} \tag{4.10}$$

Note that $\Phi(\cdot)$ is only defined on $[0, \infty)$, (4.8) has limited X by $X \leq u+1$ and Y by $Y \leq u+1$, for $u \geq 0$. Before we proceed, let us define a probability generating transform to be used in the sequel.

Definition 4.1. *A probability generating transform on a function $\Phi(n), n \in \mathbb{N}$ with respect to n is*

$$\hat{\Phi}(s) = \sum_{n=0}^{\infty} s^n \Phi(n).$$

Denote this also in the operator form by $\mathcal{P}_s\Phi$.

It resembles the Laplace transform except that it is in a discrete time horizon based on probability generating functions. Then we could derive a similar property for its operation on convolutions.

Lemma 4.2. *Given two functions $l_1(\cdot)$ and $l_2(\cdot)$ defined on $\mathbb{N}$ and their convolution $l_1 * l_2(x)$, the probability generating transform of this convolution is the product*

$$\mathcal{P}_s\{l_1 * l_2(x)\} = \hat{l_1}(s)\hat{l_2}(s). \tag{4.11}$$

Proof.

$$\begin{aligned} \mathcal{P}_s\{l_1 * l_2(x)\} &= \sum_{x=0}^{\infty} s^x \sum_{n=0}^{x} l_1(x-n) l_2(n) \\ &= \sum_{n=0}^{\infty} s^n l_2(n) \sum_{x=n}^{\infty} s^{x-n} l_1(x-n) \\ &= \hat{l_1}(s)\hat{l_2}(s) \end{aligned}$$

■

Then we can start our calculations. First we give the following statement.

Theorem 4.3. *A shifted probability generating transform on the survival function initialising from state 3, $\mathcal{P}_s\Phi_3(u+6)$, will satisfy the following relation.*

$$\begin{aligned} -\mathcal{P}_s\Phi_3(u+6)w_0(s) &= \Phi_3(0)+\Phi_3(1)w_1(s)+\Phi_3(2)w_2(s) \\ &\quad +\Phi_3(3)w_3(s)+\Phi_3(4)w_4(s)+\Phi_3(5)w_5(s), \end{aligned} \tag{4.12}$$

where

$$\begin{aligned} w_0(s) &= s^6 - qs^5 - p\frac{\hat{F}(s)}{s}\left(s^5 - qs^4 + qs^3 - q^2s^2 + q^2s\right) - r\frac{\hat{G}(s)}{s}\left(s^4 - qs^3 + qs^2 - q^2s + q^2\right); \\ w_1(s) &= s - q - p\frac{\hat{F}(s)}{s}; \\ w_2(s) &= s^2 - qs - p\frac{\hat{F}(s)}{s}(s-q) - r\frac{\hat{G}(s)}{s}; \\ w_3(s) &= s^3 - qs^2 - p\frac{\hat{F}(s)}{s}(s^2 - qs + q) - r\frac{\hat{G}(s)}{s}(s-q); \\ w_4(s) &= s^4 - qs^3 - p\frac{\hat{F}(s)}{s}(s^3 - qs^2 + qs - q^2) - r\frac{\hat{G}(s)}{s}(s^2 - qs + q); \\ w_5(s) &= s^5 - qs^4 - p\frac{\hat{F}(s)}{s}(s^4 - qs^3 + qs^2 - q^2s + q^2) - r\frac{\hat{G}(s)}{s}(s^3 - qs^2 + qs - q^2). \end{aligned}$$

Proof. First taking the probability generating transform on both side of the equation (4.8) with respect to u gives us

$$\hat{\Phi}_3(s) = \frac{p}{s}\left[\hat{\Phi}_0(s)\hat{F}(s) - \Phi_0 * F(0)\right] + \frac{q}{s}\left[\hat{\Phi}_3(s) - \Phi_3(0)\right] + \frac{r}{s}\left[\hat{\Phi}_4(s)\hat{G}(s) - \Phi_4 * G(0)\right].$$

Those minus terms are due to shifted transform. One example could be

$$\begin{aligned} \sum_{u=0}^{\infty} s^u E[\Phi_0(u+1-X)] &= \sum_{u=0}^{\infty} s^u \sum_{x=0}^{u+1} \Phi_0(u+1-x)F(x) \\ &= \frac{1}{s}\sum_{u=1}^{\infty} s^k \sum_{x=0}^{k} \Phi_0(k-x)F(x) \\ &= \frac{1}{s}\left[\sum_{u=0}^{\infty} s^k \sum_{x=0}^{k} \Phi_0(k-x)F(x) - \Phi_0 * F(0)\right] \\ &= \frac{1}{s}\left[\hat{\Phi}_0(s)\hat{F}(s) - \Phi_0 * F(0)\right]. \end{aligned}$$

Then plugging (4.10) in yields terms with shifted transform again. For instance,

$$\begin{aligned}\hat{\Phi}_0(s) &= \sum_{u=0}^{\infty} s^u \Phi_0(u) = \sum_{u=0}^{\infty} s^u \Phi_3(u+1) - q \sum_{u=0}^{\infty} s^u \Phi_3(u+2) \\ &\quad + q \sum_{u=0}^{\infty} s^u \Phi_3(u+3) - q^3 \sum_{u=0}^{\infty} s^u \Phi_3(u+4) + q^2 \sum_{u=0}^{\infty} s^u \Phi_3(u+5),\end{aligned}$$

with

$$\sum_{u=0}^{\infty} s^u \Phi_3(u+1) = \frac{1}{s} \sum_{k=1}^{\infty} s^k \Phi_3(k) = \frac{1}{s}(\hat{\Phi}_3(s) - \Phi_3(0)),$$

and so on. Rearranging these terms will lead to

$$\begin{aligned}& (s-q)\hat{\Phi}_3(s) + q\Phi_3(0) \\ = \; & \frac{1}{s^6}\hat{\Phi}_3(s)\left(sp\hat{F}(s)(s^4 - qs^3 + qs^2 - q^2 s + q^2) + r\hat{G}(s)(s^4 - qs^3 + qs^2 - q^2 s + q^2)\right) \\ & + \frac{1}{s^6}\Phi_3(0)\left(-sp\hat{F}(s)(s^4 - qs^3 + qs^2 - q^2 s + q^2) - r\hat{G}(s)(s^4 - qs^3 + qs^2 - q^2 s + q^2)\right) \\ & + \frac{1}{s^5}\Phi_3(1)\left(sp\hat{F}(s)(qs^3 - qs^2 + q^2 s - q^2) + r\hat{G}(s)(-s^4 + qs^3 - qs^2 + q^2 s - q^2)\right) \\ & + \frac{1}{s^4}\Phi_3(2)\left(sp\hat{F}(s)(-qs^2 + q^2 s - q^2) + r\hat{G}(s)(qs^3 - qs^2 + q^2 s - q^2)\right) \\ & + \frac{1}{s^3}\Phi_3(3)\left(sp\hat{F}(s)(q^2 s - q^2) + r\hat{G}(s)(-qs^2 + q^2 s - q^2)\right) \\ & + \frac{1}{s^2}\Phi_3(4)\left(sp\hat{F}(s)(-q^2) + r\hat{G}(s)(q^2 s - q^2)\right) \\ & + \frac{1}{s}\Phi_3(5)\left(r\hat{G}(s)(-q^2)\right).\end{aligned}$$

We know that $\mathcal{P}_s\Phi_3(u+6) = \sum_{u=0}^{\infty} s^u \Phi_3(u+6) = \frac{1}{s^6}(\hat{\Phi}_3(s) - \Phi_3(0) - \sum_{k=1}^{5} s^k \Phi_3(k))$.

Substituting this into the right hand side of the above equation will give us

$$
\begin{aligned}
&(s-q)\hat{\Phi}_3(s) + q\Phi_3(0) \\
=&\mathcal{P}_s\Phi_3(u+6)\left(p\hat{F}(s)(s^5 - qs^4 + qs^3 - q^2s^2 + q^2s) + r\hat{G}(s)(s^4 - qs^3 + qs^2 - q^2s + q^2)\right) \\
&+ \Phi_3(5)\left(p\hat{F}(s)(s^4 - qs^3 + qs^2 - q^2s + q^2) + r\hat{G}(s)(s^3 - qs^2 + qs - q^2)\right) \\
&+ \Phi_3(4)\left(p\hat{F}(s)(s^3 - qs^2 + qs - q^2) + r\hat{G}(s)(s^2 - qs + q)\right) \\
&+ \Phi_3(3)\left(p\hat{F}(s)(s^2 - qs + q) + r\hat{G}(s)(s - q)\right) \\
&+ \Phi_3(2)\left(p\hat{F}(s)(s - q) + r\hat{G}(s)\right) \\
&+ \Phi_3(1)\left(p\hat{F}(s)\right).
\end{aligned}
$$

Further replacing $\hat{\Phi}_3(s)$ by $\mathcal{P}_s\Phi_3(u+6)$ eventually shows the result as stated in the theorem. ■

4.3 Boundary conditions

It is clear from Theorem 4.3 that once we know the boundary values for $\Phi_3(0) - \Phi_3(5)$, we could have an explicit form of the such transform on survival probability. Hence, this section explains how to find these conditions.

Remark 4.4. *Under the positive safety loading condition (4.4), it could be known that $\Phi_3(u) \to 1$ as u goes to ∞, thus $(1-s)\hat{\Phi}_3(s) = \Phi_3(0) + \sum_{u\geq 1} s^n(\Phi_3(u) - \Phi_3(u-1)) \to 1$, as $s \to 1^-$. Actually the conclusion can also be derived from Tauber theorem, let $s \to 1$, (4.12) gives*

$$
\Phi_3(0) + r\Phi_3(1) + pq\Phi_3(2) + rq\Phi_3(3) + pq^2\Phi_3(4) + rq^2\Phi_3(5) = 2 + r - q^3 - (1-q)\mu \tag{4.13}
$$

Remark 4.5. *If $G(1) = 0$, which is equivalent to $\frac{\hat{G}(s)}{s}|_{s=0} = 0$, then it follows from (4.12) that*

$$
\Phi_3(0) - q\Phi_3(1) = pf_1\left(\Phi_3(1) - q\Phi_3(2) + q\Phi_3(3) - q^2\Phi_3(4) + q^2\Phi_3(5)\right) \tag{4.14}
$$

In fact, the above identity can also be derived from (4.8) *by setting* $u = 0$.

In the later of this section, we may assume that $G(1) = 0$ (this assumption also makes sense in practice as Y represents huge claim costs) and introduce a new probability measure $\{h_u = (1-q)^{-1}(pf_{u+1} + rg_{u+2})\}_{u\geq 0}$, theu $\sum_{u\geq 0} h_u = 1$, aud

$$
\begin{aligned}
&(1-q)\hat{H}(s) = \sum_{u\geq 0} s^u h_u = p\frac{\hat{F}(s)}{s} + r\frac{\hat{G}(s)}{s^2},\\
&(1-q)\mu_H = p(\mu_F - 1) + r(\mu_G - 2) < 1 + q - q^3;\\
&\frac{\hat{H}(s)-1}{s-1} = \sum_{u\geq 1} \frac{s^u - 1}{s-1} h_u = \sum_{u\geq 1}\sum_{k=0}^{u-1} s^k h_u = \sum_{k\geq 0} s^k (1 - H(k)),
\end{aligned}
$$

where the inequality comes from the safety loading condition (4.4), and the coefficient of $\mathcal{P}_s\Phi_3(u+6)$ can be rewritten as

$$
\begin{aligned}
&s^6 - qs^5 - s(s^4 - qs^3 + qs^2 - q^2 s + q^2)(1-q)\hat{H}(s)\\
=&s\left(s^5 - qs^4 - (1-q)\hat{H}(s)\left(s^4 - q(s-1)(s^2+q)\right)\right)\\
=&s(s-1)\left(s^4 + (1-q)q(s^2+q) - (1-q)\frac{\hat{H}(s)-1}{s-1}\left(s^4 - q(s-1)(s^2+q)\right)\right)
\end{aligned}
$$

To find the probability generating transform $\mathcal{P}_s\Phi_3(u+6)$ from equation (4.12) explicitly, 4 more boundary conditions have to be solved. Here, we use a similar argument as in Albrecher and Boxma [2005], by proving a sufficient condition that

$$
\left(s^4 + (1-q)q(s^2+q) - (1-q)\frac{\hat{H}(s)-1}{s-1}\left(s^4 - q(s-1)(s^2+q)\right)\right)
$$

has exactly 4 roots in $\{s \in \mathcal{C} : |s| < 1\}$ under the safety loading condition (4.4).

Let $M(x)$ be a matrix in the form

$$M(s) \stackrel{\text{def}}{=} \begin{pmatrix} (1-q)\hat{H}(s)-s & q & 0 & 0 \\ (1-q)\hat{H}(s) & -s^2 & q & 0 \\ (1-q)\hat{H}(s) & 0 & -s^2 & q \\ (1-q)\hat{H}(s) & 0 & 0 & q-s \end{pmatrix} \tag{4.15}$$

then $\text{Det}[M(s)] = s\left(s^5 - qs^4 - (1-q)\hat{H}(s)\left(s^4 - q(s-1)(s^2+q)\right)\right)$, $\left.\dfrac{\text{Det}\,[M(s)]}{s-1}\right|_{s=1} = 1 + q - q^3 - (1-q)\mu$. The existence of roots is proved by the Rouche's theorem for the case of a matrix and the following lemma, see De Smit [1983] for example.

Lemma 4.6. *If $A = (a_{ij})$ is a complex $n \times n$-matrix, then $Det(A) \neq 0$ if one of following condition holds*

1. $|a_{ii}| > \sum_{j=1, j\neq i}^{n} |a_{ij}|$ *for all* $i = 1, 2, \cdots, n$, *or*
2. *A is indecomposable and* $|a_{ii}| \geq \sum_{j=1, j\neq i}^{n} |a_{ij}|$ *for all* $i = 1, 2, \cdots, n$, *with strictly inequality for at least one* i

It is known that, for $s \in \{s \in \mathbb{C} : |s| \leq 1\}$, $|\hat{H}(s)| \leq 1$, $|q - s| = (1-q)$ only if $s = 1$. Lemma 4.6 shows $\text{Det}\,[M(s)] \neq 0$ for $s \neq 1$, hence $\left(\dfrac{1}{1-s}\text{Det}\,[M(s)]\right) \neq 0$ for $s \neq 1$, and $\left.\left(\dfrac{1}{1-s}\text{Det}\,[M(s)]\right)\right|_{s=1} = -(1 + q - q^3 - (1-q)\mu_H) \neq 0$.

Theorem 4.7. $\dfrac{Det\,[M(s)]}{1-s} = 0$ *has 5 roots in the complex plane* $\{s \in \mathbb{C} : |s| < 1\}$.

Proof of Theorem 4.7. Based on the observation that

$$\begin{pmatrix} \frac{1}{1-s} & \frac{q}{1-s} & \frac{q^2}{1-s} & \frac{q^3}{(1-s)(1-q)} \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} M(s) = \begin{pmatrix} \frac{\hat{H}(s)-s}{1-s} & q(1+s) & q^2(1+s) & \frac{q^3}{1-q} \\ (1-q)\hat{H}(s) & -s^2 & q & 0 \\ (1-q)\hat{H}(s) & 0 & -s^2 & q \\ (1-q)\hat{H}(s) & 0 & 0 & q-s \end{pmatrix},$$

we introduce an $L(\lambda, s)$ on $[0,1] \times \{s \in \mathcal{C} : |s| \leq 1\}$

$$L(\lambda, s) = \begin{pmatrix} \lambda\frac{\hat{H}(s)-s}{1-s} + (1-\lambda) & \lambda q(1+s) & \lambda q^2(1+s) & \lambda\frac{q^3}{1-q} \\ (1-q)\hat{H}(s) & -s^2 & q & 0 \\ (1-q)\hat{H}(s) & 0 & -s^2 & q \\ (1-q)\hat{H}(s) & 0 & 0 & q-s \end{pmatrix}.$$

Then $\mathrm{Det}\,[L(\lambda, s)]$ is analytical in $\{s \in \mathcal{C} : |s| < 1\}$ and continuous with respect to (λ, s), $\mathrm{Det}\,[L(\lambda, s)] = \lambda \mathrm{Det}\,[L(1, s)] + (1-\lambda)s^4(q-s)$, $\mathrm{Det}\,[L(1, s)] = \dfrac{1}{(1-s)}\mathrm{Det}\,[M(s)]$ for $s \neq 1$.

Following the idea of De Smit [1983], if for every $\lambda \in [0,1]$, $\mathrm{Det}\,[L(\lambda, s)] \neq 0$ on the boundary $\{s \in \mathcal{C} : |s| = 1\}$, then the number of roots of $\mathrm{Det}\,[L(\lambda, s)] = 0$ remains constant. Taking their multiplicities into consideration, since $\mathrm{Det}\,[L(0, s)] = -\xi s^4(s-q)$ has exactly 5 roots within the bounded regime $\{s \in \mathcal{C} : |s| < 1\}$, our conclusion can be proved.

Case 1 For $s = 1$, $\mathrm{Det}\,[L(\lambda, s)] = -\lambda(1+q-q^3-(1-q)\mu_H) - (1-\lambda)(1-q) < 0$ for all $\lambda \in [0,1]$.

Case 2 For $s \neq 1$, $|q-s| > (1-q) \geq |(1-q)\hat{H}(s)|$

$$\begin{pmatrix} (1-s) & -q(1-s) & -q^2(1-s) & -\frac{q^3(1-s)}{1-q} \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} L(\lambda, s)$$
$$= \begin{pmatrix} (1-\lambda)(1-s) + \lambda\left((1-q)\hat{H}(s) - s\right) & \lambda q & 0 & 0 \\ (1-q)\hat{H}(s) & -s^2 & q & 0 \\ (1-q)\hat{H}(s) & 0 & -s^2 & q \\ (1-q)\hat{H}(s) & 0 & 0 & q-s \end{pmatrix}$$

and $|(1-\lambda)(1-s) + \lambda\left((1-q)\hat{H}(s) - s\right)| = |(1-\lambda) + \lambda\left((1-q)\hat{H}(s)\right) - s| \geq 1 - (1-\lambda) - \lambda(1-q) = \lambda q$, once again, Lemma 4.6 says the matrix is invertible and $\mathrm{Det}\,[L(\lambda, s)] \neq 0$. ■

4.4 Discussions

As $\mathcal{P}_s\Phi_3(u+6)$ has an explicit form and we proved the existence of boundary conditions, next step is to take an inverse transform and obtain a formulae for $\Phi_3(u)$, which still needs some tedious calculations in the future. Once $\Phi_3(u)$ is achieved, survival probabilities originating from other states are not difficult to be derived.

This section serves as an attempt to seek ruin probabilities of a simple discrete BM system, where ruin probabilities refer to the probability of not being able to cover claims using fund reserved from a reinsurance company's perspective. It is true that even for a BM system with only three classes, although the model itself does not seem difficult, the calculations are complex and tedious. There are also several unrealistic assumptions, such as the premium levels taking exact values 1, 2, 3. It has been found that changing these values would further complicate all the calculations. Above all, it is a model with everything being discrete, which is not consistent with real life scenarios.

However, the main idea here is to understand the ruin theory and how a renewal process is established. Also, the construction of a proper Markov chain is crucial in the first place. Such thought will be held and passed on to the next stage of research, and this time a continuous time framework will be considered, which is slightly different when setting up the model.

Chapter 5

Risk Analysis of a BM system in continuous risk models

Having looked at pricing BM systems and the renewal aspect of ruin theory in previous chapters, we are now ready to move on to more realistic models with these two concepts combined. This chapter will demonstrate two ways of incorporating a BM feature into a classical risk model under a continuous time framework. First, a Bayesian estimation would be adopted to reflect an overtime premium adjustment according to historical claim frequencies. Dubey [1977] pioneered this idea, whereas we further extended his model and interpret it in a practical manner. Ruin probabilities were obtained in terms of a link with classical results. This was a joint work with Dr. Bo Li and my supervisor Dr. Corina Constantinescu, presented at several conferences and published at Insurance: Mathematics and Economics in 2015. The second section in this chapter demonstrates an approach via constructing a dependence structure between consecutive inter-arrival times thus forming a regenerative process which was extensively supported by Prof. Zbigniew Palmowski (University of Wroclaw, Poland). As a joint work with Prof. Palmowski, Suhang Dai (PhD at University of Liverpool) and my supervisor, the work has been presented at both local and international seminars and conferences and the paper is expected to be submitted soon. In this piece of work we presented some analytical calculations and also simulated ruin probabilities using two different methods.

5.1 Via a Bayesian estimation approach

Classical compound Poisson risk models consider the premium rate to be constant. By adjusting the premium rate to the claims history, one can emulate a Bonus-Malus system within the ruin theory context. One way to implement such adjustment is by considering the Poisson parameter to be a continuous random variable and use credibility theory arguments to adjust the premium rate a posteriori. Depending on the defectiveness of this random variable, respectively referred to as 'unforeseeable' (defective) versus 'historical' (non-defective) risks, one obtains different relations between the ruin probability with constant versus adjusted premium rate. A combination of these two kinds of risks also leads to a relation between the two ruin probabilities, when the a posteriori estimator of the number of claims is carefully chosen. This section will present main results from the published paper Li et al. [2015].

5.1.1 Ruin probability under the 'unforeseeable' stream

The 'Unforeseeable' stream of risks are defined in the introduction 2.3.1. Let us first consider the risk model

$$U(t) = u + c\int_0^t \hat{\lambda}(s)\mathrm{d}s - \sum_{k=0}^{N(t)} Y_k, \quad t \geq 0. \tag{5.1}$$

under such stream. Risks under this stream normally do not have clear information at present and we cannot expect claims to occur for sure. However, once they broke out in a negative way, it would possibly be too late for an insurance company to control the losses. Technically speaking, the expectation of Λ conditioning on $\{\Lambda > 0\}$ here is very large and practically it could be assumed to be much more than the average number claims in the 'historical' stream. Hence, in the 'unforeseeable' stream, there are two extremes. It could either be with no claims at all or a burst of claims in the future. The significance of this model is to control such kind of uncertainty based on observations as a single witness of a claim in the 'unforeseeable' stream would mean a dramatic increase in the premium rate thus helping the insurer to control the losses.

Moving into technical details, the main difference here from Dubey [1977] is that Λ has a mix distribution with mass at $\{0\}$, which led us to consider the sets $\{\Lambda = 0\}$ and $\{\Lambda > 0\}$ separately. And due to the fact that no claims are expected to be attributed to the risk process on the set $\{\Lambda = 0\} = \{T_1 = \infty\}$, we have

$$\begin{aligned} &\mathbb{P}(\tau < \infty | U(0) = u) = \mathbb{P}(\tau < \infty, \Lambda > 0 | U(0) = u) \\ &= \mathbb{P}(\tau < \infty | U(0) = u, \Lambda > 0)\mathbb{P}(\Lambda > 0). \end{aligned} \tag{5.2}$$

Thus, it is enough to perform the analysis under the measure $\mathbb{P}(\cdot|\Lambda > 0)$ in this case and derive the following expression for the underlying ruin probability.

Theorem 5.1. *The probability of ruin of the adjusted surplus process (5.1) with Poisson intensity Λ, a random variable with (2.5) is given by*

$$\psi^L(u) = \psi^C(u) - p\psi^C\left(u + c\ln\frac{1}{p}\right), \tag{5.3}$$

where ψ^C denotes the ruin probability in a classical risk model with jump intensity $\lambda = 1$.

Proof. Let $T_i, i = 1, \dots$ represent the arrival time of the i^{th} claim with the convention that $T_0 = 0$, and $\{cP_i, i = 1, 2, \cdots\}$ denote the premium collected in-between the $(i-1)^{th}$ and the i^{th} claim. Similar to Dubey Dubey [1977], we first analyse the conditional distribution of the sequence of premiums $\{P_n, n \geq 1\}$. From the definition of the Bayesian estimator, we have for $n \geq 0$

$$\mathbb{E}[\Lambda|N(t) = n] = \frac{\mathbb{E}[\Lambda; N(t) = n]}{\mathbb{P}[N(t) = n]} = \frac{\int_{[0,\infty)} \lambda^{n+1}e^{-\lambda t}\mathbb{P}(\Lambda \in \lambda)}{\int_{[0,\infty)} \lambda^{n}e^{-\lambda t}\mathbb{P}(\Lambda \in \lambda)} = \frac{-V^{(n+1)}(t)}{V^{(n)}(t)},$$

where $V(x) = \mathbb{E}(e^{-\Lambda x})$. It follows that $P_1 = \ln\left(\frac{V(0)}{V(T_1)}\right)$ and

$$P_{n+1} = \ln\{V^{(n)}(T_n)/V^{(n)}(T_{n+1})\}, \ n \geq 1. \tag{5.4}$$

On the one hand, adopting very similar steps in Dubey [1977] yields for $n \geq 1$,

$$\mathbb{P}(P_{n+1} \geq x | T_n = y, \Lambda > 0) = e^{-x},\ n \geq 1. \tag{5.5}$$

On the other hand, regarding the distribution of P_1, it could be first noticed that $V(x)$ is a continuous and decreasing function and $V(x) \in (p, 1]$, since (2.5). Hence, for $x \leq -\ln p$,

$$\begin{aligned} & \mathbb{P}(P_1 \geq x | \Lambda > 0) = \mathbb{P}(V(T_1) \leq e^{-x} | \Lambda > 0) \\ = \ & \mathbb{P}(T_1 \geq V^{-1}(e^{-x}) | \Lambda > 0) = \int_{(0,+\infty)} e^{-\lambda V^{-1}(e^{-x})} \cdot \mathbb{P}(\Lambda \in d\lambda | \Lambda > 0) \\ = \ & \frac{1}{1-p}(V(V^{-1}(e^{-x})) - p) = \frac{e^{-x} - p}{1-p}. \end{aligned}$$

That is to say, under the measure $\mathbb{P}(\cdot | \Lambda > 0)$, the sequence $\{P_n, n \geq 2\}$ again follows an exponential distribution with parameter 1 and P_1 conforms to a truncated exponential distribution.

Furthermore, the independence structure for the sequence of premiums still holds under the measure $\mathbb{P}(\cdot | \Lambda > 0)$. As presented in Dubey [1977], Recall (5.4) and (5.5) and let us consider the joint Laplace transform of the premiums,

$$\begin{aligned} & \mathbb{E}\left[e^{-\sum_{i=1}^{n+1} s_i P_i} \Big| \Lambda > 0\right] = \mathbb{E}\left[\mathbb{E}\left(e^{-\sum_{i=1}^{n+1} s_i P_i} \Big| T_1, T_2, \cdots, T_n, \Lambda > 0\right)\right] \\ = \ & \mathbb{E}\left\{\mathbb{E}\left[e^{-\sum_{i=1}^{n} s_i P_i} \cdot \mathbb{E}\left(e^{-s_{n+1} P_{n+1}} \Big| T_1, T_2, \cdots, T_n, \Lambda > 0\right) \Big| T_1, T_2, \cdots, T_n, \Lambda > 0\right]\right\} \\ = \ & \mathbb{E}\left\{\mathbb{E}\left[e^{-\sum_{i=1}^{n} s_i P_i} \cdot \mathbb{E}\left(e^{-s_{n+1} P_{n+1}} \Big| T_n, \Lambda > 0\right) \Big| T_1, T_2, \cdots, T_n, \Lambda > 0\right]\right\} \\ = \ & \frac{1}{1+s_{n+1}} \mathbb{E}\left[\mathbb{E}\left(e^{-\sum_{i=1}^{n} s_i P_i} \Big| T_1, T_2, \cdots, T_n, \Lambda > 0\right)\right] \\ = \ & \frac{1}{1+s_{n+1}} \mathbb{E}\left[e^{-\sum_{i=1}^{n} s_i P_i} \Big| \Lambda > 0\right], \end{aligned}$$

and this finishes the proof of our desired independence structure of premiums.

As a result, conditioning on premium collected before the first claim and associate claim size, we have

$$\begin{aligned}&\mathbb{P}(\tau < \infty | U(0) = u, \Lambda > 0)\\ =& \int_0^{ln(1/p)} \frac{1}{1-p} e^{-t} \left(\int_0^{u+ct} \psi^C(u+ct-y)dF(y) + \overline{F}(u+ct) \right) dt,\end{aligned}$$

where $\psi^C(u)$ is the ruin probability of classical risk model with jump intensity 1, which should satisfy

$$\psi^C(u) = \int_0^{\infty} e^{-t} \left(\psi^C * F(u+ct) + \overline{F}(u+ct) \right) dt.$$

Eventually, following the identity above we are able to rewrite the formula for the ruin probability from (5.2) as follows

$$\begin{aligned}\psi^L(u) &= \int_0^{\infty} e^{-t} \left(\psi^C * F(u+ct) + \overline{F}(u+ct) \right) dt\\ &\quad - \int_{ln(1/p)}^{\infty} e^{-t} \left(\psi^C * F(u+ct) + \overline{F}(u+ct) \right) dt\\ &= \psi^C(u) - \int_0^{\infty} e^{-(x+\ln\frac{1}{p})} \left(\psi^C * F(u + c(x + \ln\frac{1}{p})) + \overline{F}(u + c(x+\ln\frac{1}{p}) \right) dx\\ &= \psi^C(u) - p\psi^C(u + c\ln(1/p)).\end{aligned}$$

This completes the proof. ■

Now we apply this formulae to calculate ruin probabilities for specific claim distributions.

Example 5.2. *(As in Dubey [1977]) Y_i follows an exponential distribution with $\mathbb{E}(Y_k) = 1$. A classical risk model (with jump intensity 1) gives an explicit ruin function, $\psi^C(u) = \frac{1}{c} exp\left(-\frac{c-1}{c}u\right)$. Substituting this into (5.3) yields,*

$$\psi^L(u) = \frac{1}{c} e^{-\frac{c-1}{c}u} - \frac{p}{c} e^{-\frac{c-1}{c}\left(u + c\ln\frac{1}{p}\right)} = (1-p^c)\psi^C(u).$$

Example 5.3. *(As in Dubey [1977]) $Y_k = 1$. An approximation has been shown in classical models, $\psi^C(u) \sim \frac{c-1}{1+cr-c} e^{-ru}$, where r is the positive solution to $e^x =$*

$1+cx$. *Applying* (5.3) *with the above identity also verifies our result.*

$$\psi^L(u) \sim \frac{c-1}{1+cr-c}e^{-ru} - p\frac{c-1}{1+cr-c}e^{-r\left(u+c\ln\frac{1}{p}\right)} \sim (1-p^{cr+1})\psi^C(u),$$

when $u \to \infty$ with r the same as above.

Example 5.4. $Y_k \sim Gamma(\frac{m}{n}, \alpha)$ *with density function* $f_Y(x) = \frac{\alpha^{\frac{m}{n}}}{\Gamma(\frac{m}{n})}x^{-\frac{m}{n}}e^{-\alpha x}, x \geq 0$. *It is worth emphasising that $\frac{m}{n}$ taking integer values also covers the case of Erlang distributed claims. Employing recent results from Constantinescu et al. [2015a], where* $\psi^C(u) = 1 - e^{-\alpha u}u^{\frac{1}{n}-1}\sum_{k=0}^{m+n-1} m_k E_{\frac{1}{n},\frac{1}{n}}\left(s_k u^{\frac{1}{n}}\right)$, *the required ruin probability is demonstrated in the following equations.*

$$\begin{aligned}\psi^L(u) &= 1-p-e^{-\alpha u}u^{\frac{1}{n}-1}\sum_{k=0}^{m+n-1} m_k E_{\frac{1}{n},\frac{1}{n}}\left(s_k u^{\frac{1}{n}}\right) \\ &\quad + pe^{-\alpha\left(u+c\ln\frac{1}{p}\right)}\left(u+c\ln\frac{1}{p}\right)^{\frac{1}{n}-1}\sum_{k=0}^{m+n-1} m_k E_{\frac{1}{n},\frac{1}{n}}\left(s_k\left(u+c\ln\frac{1}{p}\right)^{\frac{1}{n}}\right),\end{aligned}$$

where $E_{\frac{1}{n},\frac{1}{n}}\left(s_k u^{\frac{1}{n}}\right) = \sum_{i=0}^{\infty}\frac{\left(s_k u^{\frac{1}{n}}\right)^i}{\Gamma\left(\frac{k+1}{n}\right)}$, *$m_k$ is a constant to be determined, and s_k solves the equation* $cx^{m+n} - (c\alpha+1)x^m + \alpha^{\frac{m}{n}} = 0$.

All these examples calculate ruin probabilities for an 'unforeseeable' risk stream through connection with classical results. We will derive a similar approach to obtain ruin probabilities while both risk streams are taken into account in the next subsection.

5.1.2 Ruin Probability with both the 'historical' stream and the 'unforeseeable' stream

We are further proposing a third estimator for Λ by combining the two scenarios indicating a consideration for both kinds of risks.

$$\hat{\lambda}(t) = \mathbb{E}[\Lambda^{(1)} + \Lambda^{(2)}|N(t)], \tag{5.6}$$

where $\mathbb{P}(\Lambda^{(1)} = 0) = 0$ and $\mathbb{P}(\Lambda^{(2)} = 0) = p > 0$.

This combined model actually describes the true situation in an insurance company. When signing a policy, the insurer decides which stream to assign a contract to. For the commonly known risks, this model simply adjust the premium rate according to observed claims using parameters that are estimated from historical data. For those unknown or innovative risks, since not enough information is known at the beginning, it is reasonable to set a lower base premium until the first claim appears. Then the premium would see a dramatic increase as a big amount of subsequent claims are expected to burst. Examples of such unknown risks could be the launch of autonomous vehicles, the consumption of genetically modified organism (GMO) food, or maybe the use of hydrogen vehicles in the future etc. Our model would deal with this kind of uncertainty and the use of dynamical adjustment of premiums would help to alleviate sudden attacks on the insurance company. The main reason for distinguishing between the two risk streams is that we could set different parameters for premium adjustment. For the 'unforeseeable' stream, premiums are expected to be very sensitive to the number of claims because once a claim occurs, many more are expected to follow. Hence, one reasonable assumption is that $\mathbb{E}[\Lambda^{(1)}] < \mathbb{E}[\Lambda^{(2)}|\Lambda^{(2)} > 0]$. Note this is a practical assumption that does not affect the results in this work (not needed in the sequel).

Correspondingly, denoting the respective claim sizes by $Y^{(1)}$ and $Y^{(2)}$ and claim counts by $N^{(1)}$ and $N^{(2)}$, then the risk surplus process considered in this section satisfies the following equation

$$\begin{aligned} dU(t) &= c\hat{\lambda}(t)dt - dS^{(1)}(t) - dS^{(2)}(t) \\ &= c\hat{\lambda}(t)dt - dS(t). \end{aligned} \tag{5.7}$$

where $S^{(i)}(t) = \sum_{j=1}^{N^{(i)}(t)} Y_j^{(i)}$, $i = 1, 2$, $S(t) = S^{(1)}(t) + S^{(2)}(t) = \sum_{k=1}^{N(t)} Y_k$; $N(t) = N^{(1)}(t) + N^{(2)}(t)$ and $\hat{\lambda}(t)$ as defined in (5.6). Let F and G denote the common distribution function of $Y^{(1)}$ and $Y^{(2)}$ respectively. We propose the following results for the underlying risk surplus process.

Lemma 5.5. *Conditioning on* $\{\Lambda^{(1)} + \Lambda^{(2)} = \lambda\}$, $N(\cdot)$ *is a Poisson process with* *intensity* λ. *Conditioning on* $\left\{\frac{\Lambda^{(1)}}{\Lambda^{(1)}+\Lambda^{(2)}} = \theta\right\}$, *the claim sizes are i.i.d with a common distribution function* $H_\theta(y) \stackrel{def}{=} \theta F(y) + (1-\theta)G(y)$, *where* F *and* G *are the distribution functions for* $Y^{(1)}$ *and* $Y^{(2)}$ *respectively.*

Lemma 5.6. *If* $(\Lambda^{(1)} + \Lambda^{(2)})$ *is independent from* $\left(\frac{\Lambda^{(1)}}{\Lambda^{(1)}+\Lambda^{(2)}}\right)$ *under* $\mathbb{P}(\cdot|\Lambda^{(2)} > 0)$, *then for a given* $\theta \in (0, 1]$, *under* $\mathbb{P}\left(\cdot\middle|\frac{\Lambda^{(1)}}{\Lambda^{(1)}+\Lambda^{(2)}} = \theta\right)$, *Model* (5.7) *could be reduced to Case 1 in Dubey [1977] model.*

Both proofs could be seen from Appendix B. Generally speaking, there is a strong dependence between the sequence of claim sizes and the sequence of claim times. However, under the assumption of Lemma 5.6, we claim that for any fixed $\theta \in (0, 1]$, under $\mathbb{P}\left(\cdot\middle|\frac{\Lambda^{(1)}}{\Lambda^{(1)}+\Lambda^{(2)}} = \theta\right)$, the risk surplus process of this extended model has the same law as

$$dU_\theta(t) = c\hat{\lambda}(t)dt - d\sum_{k=1}^{N(t)} Y_k^\theta,$$

which has a mixed Poisson process $N(\cdot)$ and i.i.d claim sizes having common distribution function $H_\theta(y)$. Since the random variable N is positive, the underlying conditioned surplus process is then reduced to the same one as Dubey [1977].

Analysing the assumption of Lemma 5.6, it has been found that this independence property is satisfied if and only if the two variables have Gamma distributions with the same scale parameter. (See Lukacs's proportion-sum independence theorem in Lukacs [1955].) More precisely, we propose the following lemma whose proof is presented in Appendix as well.

Lemma 5.7. *If* $\Lambda^{(1)} \sim \Gamma(\alpha, \lambda_0)$ *and* $\Lambda^{(2)}|_{\Lambda^{(2)}>0} \sim \Gamma(\beta, \lambda_0)$ *for some* $\alpha, \beta, \lambda_0 > 0$, *then we have*

$$\left(\Lambda^{(1)} + \Lambda^{(2)}\right)|_{\Lambda^{(2)}>0} \sim \Gamma(\alpha+\beta, \lambda_0), \qquad \left(\frac{\Lambda^{(1)}}{\Lambda^{(1)}+\Lambda^{(2)}}\right)\Bigg|_{\Lambda^{(2)}>0} \sim Beta(\alpha, \beta),$$

and they are independent.

In other words, we found particular distribution functions for $\Lambda^{(1)}$ and $\Lambda^{(2)}|_{\Lambda^{(2)}>0}$ in order to ensure the desired condition satisfied. Additionally, a specific distribution for $\left(\frac{\Lambda^{(1)}}{\Lambda^{(1)}+\Lambda^{(2)}}\right)\Big|_{\Lambda^{(2)}>0}$ could also be determined which is a Beta in this case.

In the following part, we explain two possible methods to calculate the ruin probability. To simplify notations, we denote $\Theta = \frac{\Lambda^{(1)}}{\Lambda^{(1)}+\Lambda^{(2)}}$ in the sequel.

First of all, according to previous discussions, we know that when both $\Lambda^{(1)}$ and $\Lambda^{(2)}|_{\Lambda^{(2)}>0}$ are $Gamma(\lambda_0, \alpha)$ and $Gamma(\lambda_0, \beta)$, respectively. $\Theta|_{\Theta\neq 1}$ is $Beta(\alpha, \beta)$ distributed (Lemma 5.7). In addition, for a fixed $\theta \in (0, 1]$, conditioning on $\{\Theta = \theta\}$, the surplus process can be reduced to the one in Dubey [1977] where the conditional ruin probability (5.2) coincides with that of the classical risk process with parameter $(c, 1, H_\theta(\cdot))$ (Lemma 5.5, 5.6). Hence, the ruin probability for the underlying risk surplus process depends on Θ and could be derived as,

$$\begin{aligned}
\psi^M(u) &= \mathbb{E}(\psi^H_\Theta(u)) \\
&= \frac{1-p}{B(\alpha,\beta)} \int_{(0,1)} \psi^H_\theta(u)\theta^{\alpha-1}(1-\theta)^{\beta-1}\, d\theta + p \cdot \psi^H_1(u) \\
&= \frac{1-p}{B(\alpha,\beta)} \int_{(0,1)} \psi^C_\theta(u)\theta^{\alpha-1}(1-\theta)^{\beta-1}\, d\theta + p \cdot \psi^C_1(u).
\end{aligned}$$

where $\psi^C_\theta(u)$ is the ruin probability in a classical risk model with $(c, 1, H_\theta(\cdot))$ conditioning on θ,

$$\psi^C_\theta(u) = P\left(\tau < \infty \middle| U(0) = u, \frac{\Lambda^{(1)}}{\Lambda^{(1)}+\Lambda^{(2)}} = \theta\right). \tag{5.8}$$

Firstly, notice here that $\psi^H_1(u)$ denotes the ruin probability when $\Theta = 1$, i.e., on the set $\{\Lambda^{(2)} = 0\}$. This clearly reduces the model to Case 1 in Dubey [1977] which means $\psi^H(u) = \psi^C(u)$ (Recall $\psi^C(u)$ from Theorem 5.1). Secondly, since $\psi^C_\theta(u)$ is dependent on the claim size distribution $H_\theta(\cdot)$, the ruin probability is only possible to be calculated for a specific distribution of claims.

However, even for a mixture of two exponential distributions where we could employ the result from Constantinescu and Lo [2013], due to computational complexity, it is not trivial to obtain an explicit formula for the probability of ruin.

Alternatively, we found that when the Pollaczek-Khinchin formula is employed, for $\theta \in (0,1]$, the survival probability can be expressed as

$$1-\psi_\theta^M(u)=\begin{cases}(1-\frac{\mu_\theta}{c})\sum_{n\geq 0}\left(\frac{\mu_\theta}{c}\right)^n H_{e,\theta}^{*n}(u) & \text{if} \quad \mu_\theta<c,\\ 0 & \text{if} \quad \mu_\theta\geq c,\end{cases} \tag{5.9}$$

where $\mu_\theta=\theta\mu_F+(1-\theta)\mu_G$, $H_{e,\theta}(dy)=\frac{1}{\mu_\theta}(1-H_\theta(y))\,dy=\mu_\theta^{-1}\overline{H}_\theta(y)\,dy$ is the integrated tail distribution of H_θ, and $H_{e,\theta}^{*n}(u)$ is the n-th convolution of $H_{e,\theta}$.

Theorem 5.8. *If $\max\{\mu_F,\mu_G\}<c$, then we obtain, for $u>0$,*

$$\begin{aligned}&\psi_\Theta^M(u)|_{\Theta\neq 1}\\ =\;& 1-(1-\eta)\sum_{l\geq 0,m\geq 0}\eta^l\rho^m\binom{m+l}{l}\frac{B(l+1+\alpha,m+\beta)}{B(\alpha,\beta)}F_e^{*l}*G_e^{*m}(u)\\ &-(1-\rho)\sum_{l\geq 0,m\geq 0}\eta^l\rho^m\binom{m+l}{l}\frac{B(l+\alpha,m+1+\beta)}{B(\alpha,\beta)}F_e^{*l}*G_e^{*m}(u)\end{aligned} \tag{5.10}$$

where $\eta=\mu_F/c$, $\rho=\mu_G/c$, and $F_e(y)=\frac{1}{\mu_F}\int_0^y(1-F(x))\,dx$, $G_e(y)=\frac{1}{\mu_G}\int_0^y(1-G(x))\,dx$.

This proof is to be seen in Appendix B. If we further introduce $F^\gamma(t,u)$ and $G^\gamma(t,u)$ as follows, for $t\in(0,1)$ and $\gamma>0$,

$$F^\gamma(t,u)=\sum_{l\geq 0}\binom{-\gamma}{l}(-t\eta)^l\left(F_e\right)^{*l}(u),\; G^\gamma(t,u)=\sum_{l\geq 0}\binom{-\gamma}{l}(-t\rho)^l\left(G_e\right)^{*l}(u) \tag{5.11}$$

Together with the notations introduced above, the ruin probability could be rewritten in the following way which is proved in Appendix B,

Corollary 5.9. *If $\max\{\mu_F,\mu_G\}<c$, then for $u>0$,*

$$\begin{aligned}\psi_\Theta^M(u)|_{\Theta\neq 1}\;=\;& 1-\alpha(1-\eta)\int_0^1(1-t)^{\alpha+\beta-1}\int_0^u F^{\alpha+1}(t,u-y)G^\beta(t,dy)\,dt\\ &-\beta(1-\rho)\int_0^1(1-t)^{\alpha+\beta-1}\int_0^u F^\alpha(t,u-y)G^{\beta+1}(t,dy)\,dt\end{aligned} \tag{5.12}$$

For a better illustration of the result, we consider exponential distributions with different parameters.

Corollary 5.10. *If $F \sim \exp(\zeta_1)$, $G \sim \exp(\zeta_2)$ and α, β are integers, the probability of ruin could be shown by the following formulae.*

$$\begin{aligned}
\psi_{\Theta}^{M}(u)|_{\Theta\neq 1} &= 1-\alpha(1-\eta)\left[\frac{1}{\alpha+\beta}+e^{-\zeta_1 u}\sum_{j=1}^{\beta}\binom{\beta}{j}(\rho\zeta_2)^j\sum_{i=1}^{\alpha+1}\binom{\alpha+1}{i}(\eta\zeta_1)^i\frac{u^{i+j-1}}{\Gamma(i+j)}\right.\\
&\quad\left.\times\int_0^1(1-t)^{\alpha+\beta-1}t^{i+j}e^{\zeta_1 t\eta u}M_{X(i,j)}(-[(\zeta_1\eta-\zeta_2\rho)t-\zeta_1+\zeta_2]u)\,dt\right]\\
&\quad-\beta(1-\rho)\left[\frac{1}{\alpha+\beta}+e^{-\zeta_1 u}\sum_{j=1}^{\beta+1}\binom{\beta+1}{j}(\rho\zeta_2)^j\sum_{i=1}^{\alpha}\binom{\alpha}{i}(\eta\zeta_1)^i\frac{u^{i+j-1}}{\Gamma(i+j)}\right.\\
&\quad\left.\times\int_0^1(1-t)^{\alpha+\beta-1}t^{i+j}e^{\zeta_1 t\eta u}M_{X(i,j)}(-[(\zeta_1\eta-\zeta_2\rho)t-\zeta_1+\zeta_2]u)\,dt\right] \qquad (5.13)
\end{aligned}$$

where ${}_1F_1(\cdot)$ is a hyper-geometric function with order 1,1 whose definition is given as follows.

$${}_1F_1(a;b;z)=\sum_{k=0}^{\infty}\frac{(a)_k}{(b)_k}\frac{z^k}{k!},$$

where $(c)_k = c(c+1)\ldots(c+k-1)$ with $(c)_0 = 1$.

Detailed proof is available in Appendix B.

Remark 5.11. *Alternatively, we could take the Laplace Transform of (5.12) and express the result as*

$$\begin{aligned}
\hat{\psi}_{\Theta}^{M}(s)|_{\Theta\neq 1} &= \frac{1}{s}-\alpha(1-\eta)\int_0^1(1-t)^{\alpha+\beta-1}\left(1-\frac{t\eta\zeta_1}{\zeta_1+s}\right)^{-(\alpha+1)}\left(1-\frac{t\rho\zeta_2}{\zeta_2+s}\right)^{-\beta}dt\\
&\quad-\beta(1-\rho)\int_0^1(1-t)^{\alpha+\beta-1}\left(1-\frac{t\eta\zeta_1}{\zeta_1+s}\right)^{-\alpha}\left(1-\frac{t\rho\zeta_2}{\zeta_2+s}\right)^{-(\beta+1)}dt.
\end{aligned}$$

5.2 Via a dependence structure

In this section, as a connection to a Bonus-only system (Figure 2.1), we analyse the ruin probability for the Cramér renewal risk process with consideration of an inter-arrival time depending on the number of claims that have come within a past fixed time-window. It is naturally believed that the use of BM system would possibly help reduce the solvency issue meaning decreasing ruin probabilities. This part of the thesis aims to test this belief based. This adjusted model could actually be explained through a regenerative structure. Asymptotic results of ruin probabilities for the Cramér case of the claim distributions will be discussed, while those for the heavy-tailed and intermediate case as defined by Palmowski and Zwart [2007] will be omitted in this thesis, but could be found in the paper Constantinescu et al. [2015b]. The focus here will lie in simulation methods as well as the construction of a Markov additive process. The former one will focus on overcoming the drawbacks of a crude Monte Carlo simulation and using importance sampling method to simulate infinite time ruin probabilities, whereas the latter one will be a further extension which helps to simulate ruin probabilities in an alternative way. Additionally, the use of integral equations will be demonstrated in Subsection 5.2.6, although no explicit solutions are obtained.

5.2.1 The model

Before diving into details, recall the collective renewal risk model from (1.3). In this section, for notation purposes we redefine the ruin probability as follows.

$$\psi(x) = \mathbb{P}(T(x) < \infty \mid U(0) = x),$$

where $U(0) = x \geq 0$ is the initial capital in the portfolio and

$$T(x) = \inf \{t \geq 0 : U(t) < 0 \mid U(0) = x\}$$

is the time of ruin for an initial surplus x.

To be more specific, let $\{\tau_k\}_{k\geq 0}$ be the sequence of inter-claim times. In this section we will analyse the model when the distribution F_{τ_k} of τ_k depends on the number of claims that appeared within a fixed time window ξ as follows,

$$\mathbb{P}(\tau_k \leq x) = F_{\tau_k}\left(x, N\left(\sum_{i=0}^{k-1}\tau_i\right) - N\left(\sum_{i=0}^{k-1}\tau_i - \xi\right)\right), \quad k = 1, 2, \ldots$$

For $k = 0$, τ_0 denotes the time waited until the first claim. Then

$$W_k = \sum_{i=0}^{k-1}\tau_i, k = 1, 2 \ldots$$

represents the k^{th} arrival time and $W_0 = 0$. The number of arrivals up to time t is given by

$$N_t = \sum_{k=1}^{\infty}\mathbb{1}(W_k \leq t).$$

Similar to the ordinary renewal process Asmussen and Albrecher [2010], no claims are assumed at $W_0 = 0$, but notice here that the inter-arrival time starts from τ_0. It is true that when the dependence structure is introduced, a direct use of renewal theory is no longer applicable as clearly $\{W_k\}_{k\geq 0}$ is not a renewal process. However, taking a second look, we found that even though it is not renewal at each jump time, the process in fact renews after several jumps and we call it a 'regeneration'. We define the regenerative epochs for our model here by

Definition 5.12. *Regeneration epochs T_{k+1}, for $k = 0, 1, \ldots, l = 1, 2, \ldots$ are defined as*

$$\begin{aligned} T_{k+1} &= \min\left\{W_{l+1} \geq T_k : N\left(W_l\right) - N\left(W_l - \xi\right) = 0\right\}, \\ &= \min\left\{\sum_{i=0}^{l}\tau_i \geq T_k : N\left(\sum_{i=0}^{l-1}\tau_i\right) - N\left(\sum_{i=0}^{l-1}\tau_i - \xi\right) = 0\right\} \end{aligned}$$

with $T_0 = 0$.

A rigorous definition of a regenerative process can be found in the preliminaries Definition 1.6.

Therefore, it is true that the risk process $U(t)$ is regenerative, with regeneration epochs T_k being the arrival times with zero number of arrivals within the last time window with length ξ. In this case, $\mathbb{P}(\tau_k \leq x) = F_{\tau_k}(x)$. Notice that we define the regenerative epochs in such a way that the concern only lies in whether there are claims or not in the past fixed window ξ rather than how many of them. The asymptotic results derived in Palmowski and Zwart [2010, 2007], where a general regenerative framework was studied, could then be applied to find the asymptotics of the ruin probability under Cramér assumptions.

Moving into details, let us consider the claim surplus process denoted by

$$S(t) = \sum_{k=1}^{N(t)} Y_k - c\,t,$$

and

$$X_1 = S(T_1), \qquad M_1 = \sup_{0 \le t \le T_1} S(t), \qquad M = \sup_{t \ge 0} S(t). \tag{5.14}$$

Equivalently, our purpose is to find $\psi(u) = \mathbb{P}(M > u)$.

The simplest case that we focus on is when we choose an inter arrival time from two distributions of random variables τ and $\tilde{\tau}$. τ corresponds to the situation where in a past time-window of length ξ there is at least one claim. Otherwise we assign $\tilde{\tau}$ as the inter-arrival time. Hence,

$$\mathbb{P}(\tau_k \le x) = \begin{cases} \mathbb{P}(\tau \le x), \text{ if } N(W_k) - N(W_k - \xi) \ge 1; \\ \mathbb{P}(\tilde{\tau} \le x), \text{ otherwise.} \end{cases}$$

$k = 1, 2, \ldots$. It is a natural choice since usually in and insurance company a long "silence" translates into a different behaviour of the arrival process just right after it. To rephrase it, our current model incorporates a dependence structure between two consecutive inter-arrival times. Whenever an inter-arrival time exceeds ξ, the next one would have a distribution as $\tilde{\tau}$. Otherwise, it conforms to τ. Without loss of generality, we assume $\mathbb{P}(\tau_0 \le x) = \mathbb{P}(\tilde{\tau} \le x)$.

More interestingly, such model set-up would fit into a basic Bonus system, i.e., a system where policyholders enjoy discounts when they do not file claims for a certain period (but with no penalties). A detailed illustration was given in the introduction 2.3.2. Just to recall, the key idea is that the change in premium rate

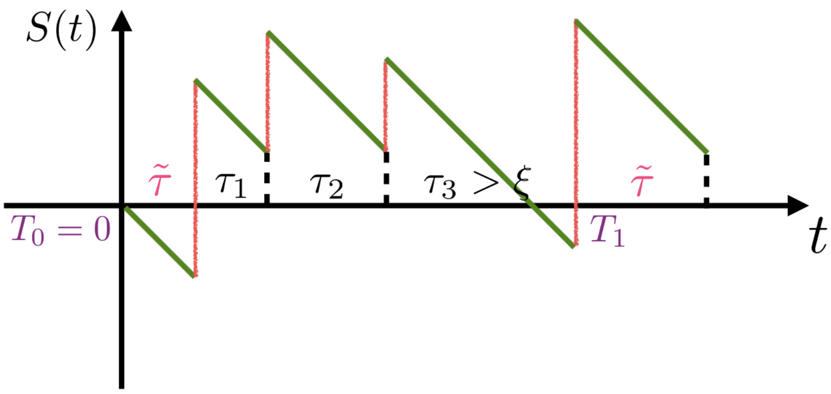


Figure 5.1: A sample path of the regenerative process

could be equivalently transformed to an alteration in inter-arrival times. Hence, a sample path of the claim surplus process we will be working with is visually described by Figure 5.1, assuming starting from $\tilde{\tau}$, up to the first regenerative epoch. Here T_1 is the first time when the last inter-claim time is larger than ξ and then the process regenerates so on and so forth. It is obvious that the process renews at each regenerative epoch.

Recall from (5.14) that X_1 is the end value at the first regenerative epoch. Then it is not difficult to observe that it has the same law as

$$X_1 \stackrel{d}{=} (Y_0 - \tilde{\tau}) + \mathbf{I}_{\{\tilde{\tau} \le \xi\}} \left(\sum_{k=1}^{N-1} (Y_k - \tau_k^{\xi}) + \left(Y_N - \overline{\tau}_N^{\xi} \right) \right), \tag{5.15}$$

where N is a geometric random variable with parameter $p = \mathbb{P}(\tau > \xi)$. Here $\mathbb{P}(N = k) = (1-p)^{k-1}p$, $k = 1, 2 \ldots$ and $\tau_k^{\le \xi} = \mathbb{E}[\tau_k | \tau_k \le \xi]$, $\tau_k^{>\xi} = \mathbb{E}[\tau_k | \tau_k > \xi]$.

5.2.2 Asymptotic results

A review of the extension of the classical Cramér case from random walks to perturbed random walks and regenerative processes can be found in Palmowski and Zwart [2007]. In this subsection, we directly apply their theorems as our model is described by a specific regenerative process.

Theorem 5.13 (as in Palmowski and Zwart [2007]). *Assume that there exists a solution $\kappa > 0$ to the equation*

$$\mathbb{E}[\mathrm{e}^{\kappa X_1}] = 1 \text{ such that } m = \mathbb{E}[X_1 \mathrm{e}^{\kappa X_1}] < \infty.$$

Assume furthermore that X_1 is non-lattice and that $\mathbb{E}[\mathrm{e}^{\kappa M_1}] < \infty$. Then

$$\psi(x) \sim K\mathrm{e}^{-\kappa x}$$

with $K = \frac{1}{\kappa m}\mathbb{E}[\mathrm{e}^{\kappa M_1} - \mathrm{e}^{\kappa(M+X_1)}; M_1 > M + X_1]$ for independent M of X_1 and M_1.

Then K is bounded from above by

$$\bar{K} = \mathbb{E}[\mathrm{e}^{\kappa M_1}]/(\kappa m), \tag{5.16}$$

and even further its upper limited is

$$\tilde{K} = \mathbb{E}[\mathrm{e}^{\kappa(X_1+T_1)}]/(\kappa m). \tag{5.17}$$

Note that by (5.15) the Cramér adjustment coefficient $\kappa > 0$ solves

$$\begin{aligned}
\mathbb{E}[\mathrm{e}^{\kappa X_1}] &= \tilde{p}\mathbb{E}[\mathrm{e}^{\kappa Y}\mathbb{E}[\mathrm{e}^{-\kappa\tilde{\tau}}|\tilde{\tau} > \xi] + \tilde{q}\mathbb{E}[\mathrm{e}^{\kappa Y}]\mathbb{E}[\mathrm{e}^{-\kappa\tilde{\tau}}|\tilde{\tau} \le \xi] \cdot \mathbb{E}[\mathrm{e}^{\kappa Y}]\mathbb{E}[\mathrm{e}^{-\kappa\tau}|\tau > \xi] \\
&\quad \cdot \sum_{k=1}^{\infty} p(1-p)^{k-1}\left[\mathbb{E}[\mathrm{e}^{\kappa Y}]\mathbb{E}[\mathrm{e}^{-\kappa\tau}|\tau \le \xi]\right]^{k-1} \\
&= p\tilde{q}\frac{(\mathbb{E}[\mathrm{e}^{\kappa Y}])^2\mathbb{E}[\mathrm{e}^{-\kappa\tau}|\tau > \xi]\mathbb{E}[\mathrm{e}^{-\kappa\tilde{\tau}}|\tilde{\tau} \le \xi]}{1-(1-p)\mathbb{E}[\mathrm{e}^{\kappa Y}]\mathbb{E}[\mathrm{e}^{-\kappa\tau}|\tau \le \xi]} + \tilde{p}\mathbb{E}[\mathrm{e}^{\kappa Y}]\mathbb{E}[\mathrm{e}^{-\kappa\tilde{\tau}}|\tilde{\tau} > \xi] \\
&= 1,
\end{aligned} \tag{5.18}$$

where $p = \mathbb{P}(\tau > \xi)$, $q = 1-p = \mathbb{P}(\tau \le \xi)$ and $\tilde{p} = \mathbb{P}(\tilde{\tau} > \xi)$, $\tilde{q} = 1-\tilde{p} = \mathbb{P}(\tilde{\tau} \le \xi)$. Let us define the [1]m.g.f $\mathbb{E}[e^{\theta X_1}]$ as

$$\varphi(\theta) = p\tilde{q}\frac{(\mathbb{E}[\mathrm{e}^{\theta Y}])^2\mathbb{E}[\mathrm{e}^{-\theta\tau}|\tau > \xi]\mathbb{E}[\mathrm{e}^{-\theta\tilde{\tau}}|\tilde{\tau} \le \xi]}{1-(1-p)\mathbb{E}[\mathrm{e}^{\theta Y}]\mathbb{E}[\mathrm{e}^{-\theta\tau}|\tau \le \xi]} + \tilde{p}\mathbb{E}[\mathrm{e}^{\theta Y}]\mathbb{E}[\mathrm{e}^{-\theta\tilde{\tau}}|\tilde{\tau} > \xi] \tag{5.19}$$

[1]Here the m.g.f exists since it was assumed that $\mathbb{E}[\mathrm{e}^{\kappa M_1}] < \infty$ and $\mathbb{E}[\mathrm{e}^{\kappa X_1}] \le \mathbb{E}[\mathrm{e}^{\kappa M_1}] < \infty$

Assume there exists a κ such that

$$\varphi(\kappa) = 1. \tag{5.20}$$

At the same time, we can also identify the constant $\tilde{K}$:

$$\begin{aligned}
\tilde{K} &= \mathbb{E}\left[\mathrm{e}^{\kappa(X_1+T_1)}\right]/\kappa m = \mathbb{E}\left[\mathrm{e}^{\kappa\sum_{i=1}^{N(T_1)} Y_i}\right]/\kappa m \\
&= \frac{1}{\kappa m}\sum_{n=1}^{\infty}\left(\mathbb{E}\left[\mathrm{e}^{\kappa Y}\right]\right)^n \mathbb{P}(N=n) \\
&= \left(\mathbb{P}(\tilde{\tau} > \xi)\mathbb{E}[\mathrm{e}^{\kappa Y}] + \mathbb{P}(\tilde{\tau} \le \xi)\sum_{n=2}^{\infty}\left(\mathbb{E}[\mathrm{e}^{\kappa Y}]\right)^n \mathbb{P}(N=n)\right)\frac{1}{\kappa m} \\
&= \left(\mathbb{P}(\tilde{\tau} > \xi)\mathbb{E}[\mathrm{e}^{\kappa Y}] + \frac{\mathbb{P}(\tilde{\tau} \le \xi)\mathbb{P}(\tau > \xi)(\mathbb{E}[\mathrm{e}^{\kappa Y}])^2}{1-\mathbb{P}(\tau \le \xi)\mathbb{E}[\mathrm{e}^{\kappa Y}]}\right)\frac{1}{\kappa m}. \qquad (5.21)
\end{aligned}$$

under assumption that

$$m = \varphi_k'(\kappa) < \infty.$$

Remark 5.14. *In addition, according to this, one could obtain a net profit condition (NPC) via (5.15). Like the usual NPC, we need the increment of such 'random walk' to be negative. One obvious reason is that if it were positive, the process would drift to infinity thus resulting in a ruin probability equal to 1. Here, since the underlying process renews at each regenerative epoch, we must have*

$$\mathbb{E}[X_1] = \mathbb{P}(\tilde{\tau} \le \xi)\left[\mathbb{E}[Y] - \mathbb{E}[\tau^{>\xi}] + \mathbb{E}[N-1](\mathbb{E}[Y] - \mathbb{E}[\tau^{\le\xi}])\right] + (\mathbb{E}[Y] - \mathbb{E}[\tilde{\tau}]) < 0.$$

Example 5.15. *A special example of exponentially distributed $\tau \sim Exp(\lambda_1)$, $\tilde{\tau} \sim Exp(\lambda_2)$ and $Y \sim Exp(\beta)$ would lead to*

$$\begin{aligned}
\varphi(\theta) &= \frac{\lambda_1\lambda_2\left(\mathrm{e}^{-\lambda_1\xi} - \mathrm{e}^{-\lambda_2\xi}\right)\frac{\hat{B}^2(\theta)\mathrm{e}^{-\theta\xi}}{(\lambda_1+\theta)(\lambda_2+\theta)} + \frac{\lambda_2}{\lambda_2+\theta}\hat{B}(\theta)\mathrm{e}^{-(\lambda_2+\theta)\xi}}{1-\frac{\lambda_1}{\lambda_1+\theta}\left(1-\mathrm{e}^{-(\lambda_1+\theta)\xi}\right)\hat{B}(\theta)} \qquad (5.22) \\
&= \left[\lambda_1\lambda_2\left(\mathrm{e}^{-\lambda_1\xi} - \mathrm{e}^{-\lambda_2\xi}\right)\frac{\beta^2\mathrm{e}^{-\xi\theta}}{(\beta-\theta)^2(\lambda_1+\theta)(\lambda_2+\theta)} + \frac{\lambda_2}{\lambda_2+\theta}\frac{\beta}{\beta-\theta}\mathrm{e}^{-(\lambda_2+\theta)\xi}\right] \div \\
&\quad \left[1 - \frac{\lambda_1}{\lambda_1+\theta}\left(1-\mathrm{e}^{-(\lambda_1+\theta)\xi}\right)\frac{\beta}{\beta-\theta}\right], \qquad (5.23)
\end{aligned}$$

and

$$\tilde{K} = \frac{\beta}{\beta - \kappa} \cdot \frac{\beta e^{-\lambda_1 \xi} - \kappa e^{-\lambda_2 \xi}}{\beta e^{-\lambda_1 \xi} - \kappa}.$$

That gives

$$\psi(x) \sim \frac{\beta}{\beta - \kappa} \cdot \frac{\beta e^{-\lambda_1 \xi} - \kappa e^{-\lambda_2 \xi}}{\beta e^{-\lambda_1 \xi} - \kappa} e^{-\kappa x}.$$

On the other hand, since

$$\begin{aligned} \mathbb{E}[\tau^{\leq \xi}] &= \mathbb{E}[\tau | \tau \leq \xi] = \frac{\frac{1}{\lambda_1} - \left(\xi + \frac{1}{\lambda_1}\right) e^{-\lambda_1 \xi}}{1 - e^{-\lambda_1 \xi}}; \\ \mathbb{E}[\tau^{> \xi}] &= \mathbb{E}[\tau | \tau > \xi] = \xi + \frac{1}{\lambda_1}, \end{aligned}$$

it will eventually lead to

$$\left(\frac{1}{\beta} - \frac{1}{\lambda_1}\right)(1 - e^{-\lambda_2 \xi}) + \left(\frac{1}{\beta} - \frac{1}{\lambda_2}\right) e^{\lambda_1 \xi} < 0, \tag{5.24}$$

where $\lambda_1, \lambda_2, \beta, \xi \geq 0$ *and* $\mathbb{E}[Y] = \frac{1}{\beta}$.

Furthermore, as a connection with Subsection 5.2.3 and 5.2.5, it is worth mentioning here that the above identity should coincide with

$$\left(\frac{1}{\beta} - \frac{1}{\lambda_1}\right) \pi_1 + \left(\frac{1}{\beta} - \frac{1}{\lambda_2}\right) \pi_2 < 0, \tag{5.25}$$

where

$$\pi_1 = \frac{1 - e^{-\lambda_2 \xi}}{1 - e^{-\lambda_2 \xi} + e^{-\lambda_1 \xi}}; \tag{5.26}$$

$$\pi_2 = \frac{e^{-\lambda_1 \xi}}{1 - e^{-\lambda_1 \xi} - e^{-\lambda_2 \xi}}, \tag{5.27}$$

denoting the steady state distribution in the Markovian environment of τ *and* $\tilde{\tau}$*, which is clearly defined in Section 5.2.5. That is to say, when the process becomes stationary, the probability to have an inter-arrival time less or equal to* ξ *(State 1) would be* π_1 *while that for it being larger than* ξ *(State 2) is represented by* $\pi_2 = 1 - \pi_1$*. The graph (Figure 5.2) below shows an example of this distribution. It could be seen that the probability for State 1 in our case is monotonically increasing with*

ξ. The blue line represents the ratio of probabilities between State 1 and State 2 thus having the same monotonicity as the green line. This will be analysed further via simulation.

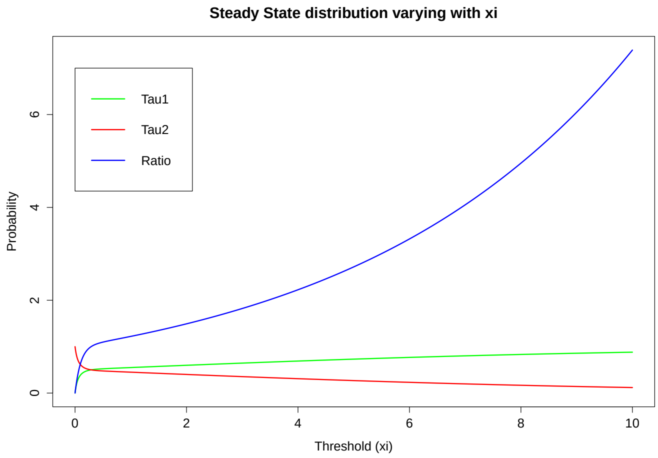


Figure 5.2: Steady State distribution when $\lambda_1 = 0.2, \lambda_2 = 10$

5.2.3 Monte Carlo Simulation

In this subsection, we show some results via a crude Monte Carlo simulation method. The key idea is to simulate the process according to the model setting and simply count the number of times it gets to ruin. Due to the nature of this approach, a 'maximum' time should be set beforehand, which means we are in fact simulating a finite time ruin probability. However, the drawback of it may be ignored for now as long as we are not getting a lot of zeros.

Our first task is to compare the simulated results with a classical analytical ruin probability. Hence, for the simplest case of exponentially distributed claim costs, we plotted both the classic ruin probabilities and our simulated ones on the same graph as shown below (Figure 5.3).

Solid lines show classical ruin probabilities (infinite-time) as a function of initial capital u, and each of them denotes an individual choice of Poisson parameters ($\lambda_1 = 0.15$, $\lambda_2 = 0.45$) with the middle one being the average of the other

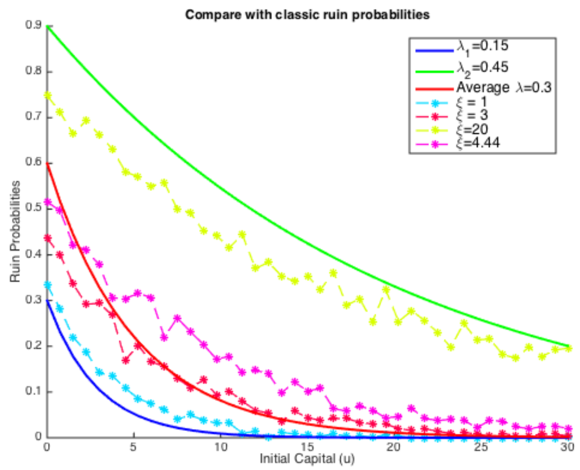


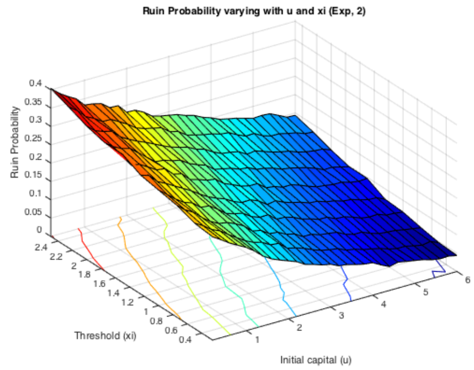
Figure 5.3: Classic ruin probabilities vs our model (Exponential claims)

two ($\lambda = 0.3$). It is clear that the larger the Poisson parameter, the higher is the ruin probability. On the other hand, those dotted lines are simulated results from our risk model with dependence for the same given pair of Poisson parameters $\lambda_1 = 0.15$ and $\lambda_2 = 0.45$. The four layers here correspond to four different choices of values for ξ, i.e., $\xi = 1, \xi = 3, \xi = 4.44, \xi = 20$. If $\xi \to 0$, the simulated ruin probability (in fact finite-time) tends to a classical case with the lower claim arrival intensities (λ_1 here), which explains the blue dotted line lying around the dark blue solid line. On the contrary, if $\xi \to \infty$, simulated ruin probabilities approach the other end. This phenomenon is also theoretically supported by (5.45) and (5.46) if either these limits ($\xi \to 0$ and $\xi \to \infty$) is taken. This then triggered us to search for a ξ such that the simulated ruin probability coincides with a classical one. Let us see an example here, if $\xi = \frac{\frac{1}{\lambda_1}+\frac{1}{\lambda_2}}{2} = 4.44$ based on the parameters we chose in Figure 5.3. That implies the choice of our fixed window is the average length of the two kinds of inter-arrival times. However, as can be seen from Figure 5.3, the dotted line with $\xi = 3$ lies closer than the one with $\xi = 4.44$ to the red solid line. This suggests that the choice of ξ will influence the simulated ruin probabilities and thus the comparison with a classical one. It is also very likely that there exists a ξ such that our simulated ruin probabilities concur with the classic one.

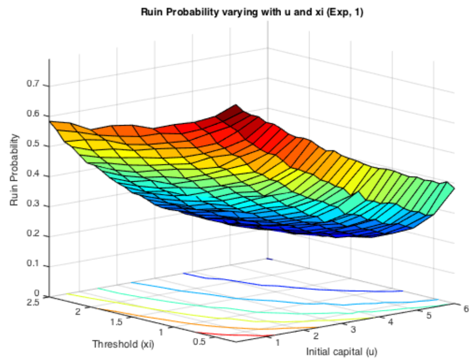
It could be concluded that under two given parameters for Poisson intensity, simulated finite ruin probabilities in our model lie between two extreme but have many possibilities in-between. The comparison depends extensively on the value of ξ. These results also confirmed Theorem 5.13 that the tail of the ruin function in our case still has an exponential decay and ξ is strongly related to the solution for κ. In other words, when the dependence is introduced, it is not for sure that ruin probabilities would see an improvement.

While the first half of the Monte Carlo simulation looked at the influence of ξ on simulated ruin probabilities, the second step is to see the effects of claim sizes. Typical representation of light-tailed and heavy-tailed distributions - Exponential and Pareto - were assumed for claim severities and inter arrival times were switching between two different exponentially distributed random variables with parameters λ_1 and λ_2. Two cases were simulated - either $\lambda_1 > \lambda_2$ or $\lambda_1 < \lambda_2$. It is expected that the effects from claim severity distributions on infinite time ruin probabilities would be tiny as they normally affects more severely in the deficit at ruin. Here, since we simulate finite-time ruin probabilities, we are curious whether the same conclusion can be drawn.

Figure 5.4 displays the two cases for Exponential claims while Figure 5.5 does that for Pareto claims. All of these four graphs demonstrate a decreasing trend for simulated finite-time ruin probabilities over the amount of initial capital, which is as expected. In general, the differences between ruin probabilities for Exponentially distributed claim costs and those for Pareto ones are not significant. To be more precise, the exact values of these disparities are plotted in Figure 5.6. The color bar shows the scale of the graph, and yellow represents values around 0. Indeed, the differences are very small. Furthermore, it can be seen that the disparities behave differently when $\lambda_1 < \lambda_2$ and when $\lambda_1 > \lambda_2$. For the former case, ruin probabilities for Pareto claims tend to be smaller than those for Exponential claims when the initial capital is not little, whereas there seems to be no distinction between the two claim distributions in the latter case. One way to explain this is that claim distributions would have more impact on the deficit at ruin because the claim frequency is not affected, the same as in an infinite-time



(a) Ruin probabilities when $\lambda_1 = 0.45, \lambda_2 = 0.15, \beta = 0.5$

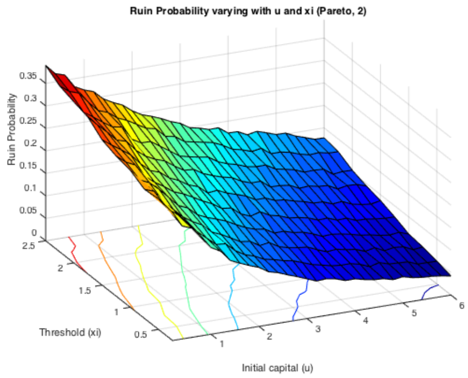


(b) Ruin probabilities when $\lambda_1 = 0.15, \lambda_2 = 0.45, \beta = 0.5$

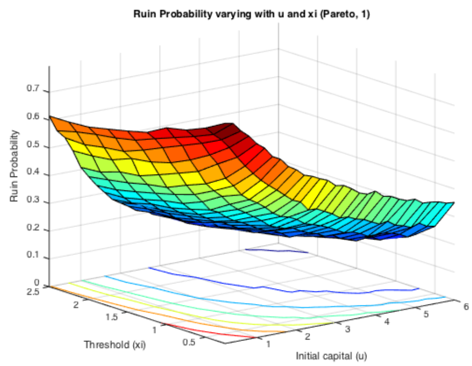
Figure 5.4: Examples: Ruin probabilities for Exponential Claims

ruin case. However, this is just a sample simulated result from which we cannot draw a general conclusion.

On the other hand, it could be seen from the projections on the $y-z$ plane that the magnitude of λ_1 and λ_2 causes different monotonicity of ruin probabilities with respect to the fixed window ξ. If $\lambda_1 > \lambda_2$, the probability of ruin is monotonically increasing with the increase of ξ. If $\lambda_1 < \lambda_2$, it appears to be the opposite monotonicity. This conclusion for monotonicity is true for both models with heavy-tailed claims and those with light-tailed ones. Such behaviour could also



(a) Ruin probabilities when $\lambda_1 = 0.45, \lambda_2 = 0.15, \alpha = 2$



(b) Ruin probabilities when $\lambda_1 = 0.15, \lambda_2 = 0.45, \alpha = 2$

Figure 5.5: Examples: Ruin probabilities for Pareto claims

be theoretically verified if we look at the stationary distribution of the Markov Chain created by the exchange of inter claim times given by (5.26) and (5.27). The increase of ξ will raise the probability of getting an inter-claim time smaller than ξ at steady state, i.e.,

$$\xi \uparrow \;\Rightarrow\; \pi_1 \uparrow,\; \pi_2 \downarrow .$$

And that directly leads to an increasing number of τ. The ruin probability is

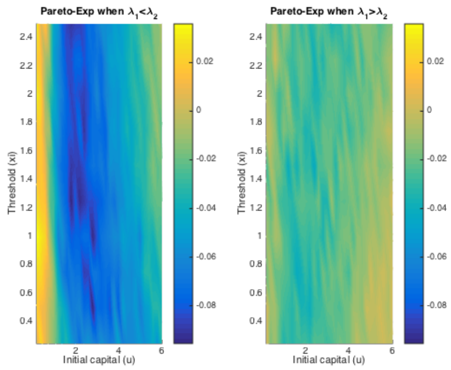


Figure 5.6: Differences in ruin probabilities using two claim distributions

associated with

$$S_T = \sum_{k=1}^{N_1(T)+N_2(T)} Y_k - \sum_{i=1}^{N_1(T)} \tau - \sum_{j=1}^{N_2(T)} \tilde{\tau}$$

for any fixed time T, where $N_1(T)$ and $N_2(T)$ denote the number of times τ and $\tilde{\tau}$ appearing in the process. Notice that $\sum_{i=1}^{N_1(T)} \tau + \sum_{j=1}^{N_2(T)} \tilde{\tau} = T$ stays the same even though the value of ξ alters. So now the magnitude of S_T depends only on $N_1(T) + N_2(T)$ and the distribution of i.i.d Y_k. The change of ξ alters only the former value. Intuitively, a rise in π_1 indicates an increase in $N_1(T)$ and a decrease in $N_2(T)$ whose amount is denoted by ΔN_1 and ΔN_2 respectively. Since the sum of τs and $\tilde{\tau}$s is kept constant, we have

$$\begin{aligned} |\Delta N_1|\mathbb{E}[\tau] &= |\Delta N_2|\mathbb{E}[\tilde{\tau}] \\ \left|\frac{\Delta N_1}{\Delta N_2}\right| &= \frac{\mathbb{E}[\tilde{\tau}]}{\mathbb{E}[\tau]} \end{aligned}$$

If $\lambda_1 > \lambda_2$, then $\mathbb{E}[\tau] < \mathbb{E}[\tilde{\tau}]$, which implies $\left|\frac{\Delta N_1}{\Delta N_2}\right| > 1$. That is to say, the increase of $N_1(T)$ is more than the drop in $N_2(T)$ so that $N_1(T)+N_2(T)$ sees a rise in the end. Thus, it leads to a higher ruin probability. On the contrary, when $\lambda_1 < \lambda_2$, i.e., $\mathbb{E}[\tau] > \mathbb{E}[\tilde{\tau}]$, as ξ goes up, ruin probabilities would experience a monotone decay. This reasoning is visually reflected in Figure 5.4-5.5 shown above and it could also be noticed that the distribution of claims does not affect such monotonicity.

Therefore, these results suggest that when $\lambda_1 < \lambda_2$, the larger choice of the fixed window ξ, the smaller the ruin probability will be, and vice versa. On the contrary, when $\lambda_1 > \lambda_2$, the larger choice of the fixed window ξ, the larger the ruin probability will be, and vice versa. In fact $\lambda_1 < \lambda_2$ was mentioned in the introduction (Figure 2.1) to be an assumption for a Bonus system. Such observation suggests that if the insurer opts to investigate claims histories less frequently, i.e., choosing a larger ξ, the ruin probability tends to be smaller. This potentially implies a smaller ruin probability if no premium discount is offered to policyholders. It seems that to minimise an insurer's probability of ruin probably relies more on premium incomes. The use of Bonus systems may not help in decreasing such probabilities. The case of $\lambda_1 > \lambda_2$ could be referred to as a Malus system which is unusual in the real world which leads to an opposite conclusion to the other case. This again addresses the significance of premium income to an insurer. In a system with purely maluses, the ruin probability could be reduced if the insurer reviews the policyholders' behaviours more frequently indicating more premium incomes.

5.2.4 Importance sampling and change of measure

One cause of the drawback of using the crude Monte Carlo simulation is that ruin probability tends to zero very quickly, when the initial capital u is large. This has been explained by the Cramér theorem that asymptotically ruin probability has an exponentially decay with respect to u. The other reason of not simply adopting a crude Monte Carlo simulation is that we are anyway trying to simulate an infinite time ruin probability under a finite time horizon. In order to overcome this effect, the importance sampling technique has been brought in. The key idea behind is to find an equivalent probability measure under which the process has a probability of ruin equal to 1.

Let us start from something trivial. For the moment, we only consider the "ruin probability" when the time between regenerative epochs is ignored. In other words, we now look at our process from a macro perspective and it is renewed at each regenerative time epoch, so we omit the situations where ruin happens within these intervals. We refer to it as the "macro" process which coincides with a classical risk process and its corresponding ruin probability as the "macro" ruin probability in the sequel. We can then define the macro ruin time as

$$T^*(u) = \inf\{T_i \geq 0 : U(T_i) < 0,\, i = 1, \ldots \mid U(0) = u\}. \tag{5.28}$$

Consequently, the macro ruin probability denoted by $\psi^*(u) = \mathbb{P}(T^*(u) < \infty \mid U(0) = u)$ should be smaller or equal than the ruin probability associated with our actual risk process $\psi(u)$. But for illustration purposes, it is worth covering the nature of change of measure under the framework of this macro process first before we dig into more complex scenarios.

Theorem 5.16. *If we do the change of measure as shown below,*

$$\begin{aligned} \mathbb{Q}(Y \in dy) &= \frac{\mathbb{P}(Y \in dy)\mathrm{e}^{\kappa Y}}{\mathbb{E}[\mathrm{e}^{\kappa Y}]}; \\ \mathbb{Q}(\tau^{\leq \xi} \in dx) &= \frac{\mathbb{P}(\tau \in dx)\mathrm{e}^{-\kappa x}}{\int_0^\xi \mathrm{e}^{-\kappa x}\mathbb{P}(\tau \in dx)},\ x \in (0, \xi]; \\ \mathbb{Q}(\tau^{> \xi} \in dx) &= \frac{\mathbb{P}(\tau \in dx)\mathrm{e}^{-\kappa x}}{\int_\xi^\infty \mathrm{e}^{-\kappa x}\mathbb{P}(\tau \in dx)},\ x \in (\xi, \infty), \end{aligned}$$

with $\tilde{\tau}^{\leq \xi}$ and $\tilde{\tau}^{> \xi}$ defined in a similar way, then we could establish the same relation for m.g.f as in the classical case,

$$\varphi_{\mathbb{Q}}(\theta) = \varphi(\theta + \kappa)/\varphi(\kappa) = \varphi(\theta + \kappa), \tag{5.29}$$

where we assume there exists a κ s.t. $\varphi(\kappa) = 1$.

Proof. Rewrite equation (5.22),

$$\begin{aligned} \varphi(\theta + \kappa) &= \mathbb{E}[\mathrm{e}^{(\theta+\kappa)Y}]\mathbb{E}[\mathrm{e}^{-(\theta+\kappa)\tilde{\tau}}, \tilde{\tau} > \xi] + \mathbb{E}[\mathrm{e}^{(\theta+\kappa)Y}]\mathbb{E}[\mathrm{e}^{-(\theta+\kappa)\tilde{\tau}}, \tilde{\tau} \leq \xi] \\ &\quad \cdot \mathbb{E}[\mathrm{e}^{(\theta+\kappa)Y}]\mathbb{E}[\mathrm{e}^{-(\theta+\kappa)\tau}, \tau > \xi] \\ &\quad \cdot \sum_{k=1}^{\infty} \left((1-p)\mathbb{E}[\mathrm{e}^{(\theta+\kappa)Y}]\mathbb{E}[\mathrm{e}^{-(\theta+\kappa)\tau}, \tau \leq \xi]\right)^{k-1} \end{aligned} \tag{5.30}$$

First, we notice that

$$\mathbb{E}[\mathrm{e}^{(\theta+\kappa)Y}] = \int \mathrm{e}^{(\theta+\kappa)Y}\mathbb{P}(Y \in dy) = \mathbb{E}[\mathrm{e}^{\kappa Y}] \int \mathrm{e}^{\theta Y}\mathbb{Q}(Y \in dy) = \mathbb{E}[\mathrm{e}^{\kappa Y}]\mathbb{E}_{\mathbb{Q}}[\mathrm{e}^{\theta Y}]. \tag{5.31}$$

So similarly, for $\tau^{\leq\xi}, \tau^{>\xi}$,

$$\mathbb{E}[\mathrm{e}^{-(\theta+\kappa)\tau}, \tau > \xi] = \mathbb{E}[\mathrm{e}^{-\kappa\tau}, \tau > \xi]\mathbb{E}_{\mathbb{Q}}[\mathrm{e}^{-\theta\tau^{>\xi}}], \tag{5.32}$$

$$\mathbb{E}[\mathrm{e}^{-(\theta+\kappa)\tau}, \tau \leq \xi] = \mathbb{E}[\mathrm{e}^{-\kappa\tau}, \tau \leq \xi]\mathbb{E}_{\mathbb{Q}}[\mathrm{e}^{-\theta\tau^{\leq\xi}}]. \tag{5.33}$$

Also, $\tilde{\tau}^{\leq\xi}, \tilde{\tau}^{>\xi}$ have the same form. Then equation (5.30) could be modified to

$$\begin{aligned}
\varphi(\theta+\kappa) &= \mathbb{E}[\mathrm{e}^{\kappa Y}]\mathbb{E}[\mathrm{e}^{-\kappa\tilde{\tau}}, \tilde{\tau} > \xi] \cdot \left[\mathbb{E}_{\mathbb{Q}}[\mathrm{e}^{\theta Y}]\mathbb{E}_{\mathbb{Q}}[\mathrm{e}^{-\theta\tilde{\tau}^{>\xi}}]\right] \\
&+ \mathbb{E}[\mathrm{e}^{\kappa Y}]\mathbb{E}[\mathrm{e}^{-\kappa\tilde{\tau}}, \tilde{\tau} \leq \xi] \cdot \left[\mathbb{E}_{\mathbb{Q}}[\mathrm{e}^{\theta Y}]\mathbb{E}_{\mathbb{Q}}[\mathrm{e}^{-\theta\tilde{\tau}^{\leq\xi}}]\right] \\
&\cdot \mathbb{E}[\mathrm{e}^{\kappa Y}]\mathbb{E}[\mathrm{e}^{-\kappa\tau}, \tau > \xi] \cdot \left[\mathbb{E}_{\mathbb{Q}}[\mathrm{e}^{\theta Y}]\mathbb{E}_{\mathbb{Q}}[\mathrm{e}^{-\theta\tau^{>\xi}}]\right] \\
&\cdot \sum_{k=1}^{\infty} \left(\mathbb{E}[\mathrm{e}^{\kappa Y}]\mathbb{E}[\mathrm{e}^{-\kappa\tau}, \tau \leq \xi]\right)^{k-1} \cdot \left[\mathbb{E}_{\mathbb{Q}}[\mathrm{e}^{\theta Y}]\mathbb{E}_{\mathbb{Q}}[\mathrm{e}^{-\theta\tau^{\leq\xi}}]\right]^{k-1}.
\end{aligned}$$

Now let,

$$\begin{aligned}
\tilde{p}_\kappa &= \mathbb{E}[\mathrm{e}^{\kappa Y}]\mathbb{E}[\mathrm{e}^{-\kappa\tilde{\tau}}, \tilde{\tau} > \xi], \quad \tilde{q}_\kappa = 1 - \tilde{p}_\kappa, \\
p_\kappa &= (\mathbb{E}[\mathrm{e}^{\kappa Y}])^2\mathbb{E}[\mathrm{e}^{-\kappa\tilde{\tau}}, \tilde{\tau} \leq \xi]\mathbb{E}[\mathrm{e}^{-\kappa\tau}, \tau > \xi], \quad q_\kappa = 1 - p_\kappa,
\end{aligned}$$

Thus,

$$\begin{aligned}
\varphi(\theta+\kappa) &= \tilde{p}_\kappa \cdot \left[\mathbb{E}_{\mathbb{Q}}[\mathrm{e}^{\theta Y}]\mathbb{E}_{\mathbb{Q}}[\mathrm{e}^{-\theta\tilde{\tau}^{>\xi}}]\right] + p_\kappa\tilde{q}_\kappa \left[(\mathbb{E}_{\mathbb{Q}}[\mathrm{e}^{\theta Y}])^2\mathbb{E}_{\mathbb{Q}}[\mathrm{e}^{-\theta\tilde{\tau}^{\leq\xi}}]\mathbb{E}_{\mathbb{Q}}[\mathrm{e}^{-\theta\tau^{>\xi}}]\right] \\
&\quad \cdot \sum_{k=1}^{\infty} (1-p_\kappa)^{k-1} \cdot \left[\mathbb{E}_{\mathbb{Q}}[\mathrm{e}^{\theta Y}]\mathbb{E}_{\mathbb{Q}}[\mathrm{e}^{-\theta\tau^{\leq\xi}}]\right]^{k-1} \\
&= \varphi_{\mathbb{Q}}(\theta).
\end{aligned}$$

■

To analyse (5.29) further, $\varphi_{\mathbb{Q}}(\theta)$ can be considered as if the function $\varphi(\theta)$ shifted to the left by κ. We know that the net profit condition for the macro process requires $\mathbb{E}[X_1] < 0$, i.e., $\varphi'(0) < 0$. Additionally, (5.19) should have a positive root κ if the tail of the claim cost distribution is Exponentially bounded. That is to say, $\varphi'(0) > 0$ would result in a positive drift of the macro claim surplus process and then cause a macro ruin to happen for sure. $\varphi(\theta+\kappa)$ makes this true. Hence, for any stopping time $T^*(u)$, we can write for a macro ruin probability as

$$\psi^*(u) = \mathbb{E}[\mathbf{1}_{T^*(u)<\infty}] = \mathbb{E}_{\mathbb{Q}}[\mathrm{e}^{-\kappa S(T^*(u))+T^*(u)\ln\varphi(\kappa)}\mathbf{1}_{T^*(u)<\infty}],$$

with $\mathbb{E}_{\mathbb{Q}}[\mathbf{1}_{T^*(u)<\infty}] = 1$. For a strict and detailed proof please refer to Asmussen and Albrecher [2010] (Chapter IV. Theorem 4.3). However, intuitively, since we have (5.29) to hold, it is equivalent to say

$$\mathbb{P}_{\mathbb{Q}}(X_1 \in dx) = \frac{\mathbb{P}(X_1 \in dx)\mathrm{e}^{\kappa x}}{\int \mathbb{P}(X_1 \in dx)\mathrm{e}^{\kappa x}dx}. \tag{5.34}$$

Then the likelihood ratio could be obtained as,

$$\begin{aligned} L_n &= \prod_{i=1}^{n} \frac{d\mathbb{P}}{d\mathbb{Q}}(X_i) \\ &= \prod_{i=1}^{n} \frac{\left(\mathbb{E}[e^{\kappa X_i}]\right)^n}{e^{\kappa X_i}} \\ &= \mathrm{e}^{n\ln\varphi(\kappa)-\kappa\sum_{i=1}^{n}X_i}, \end{aligned} \tag{5.35}$$

for a fixed integer $n \geq 0$. Obviously, $\{L_n, n \geq 0\}$ is a Wald martingale, i.e., $\mathbb{E}[\mathrm{e}^{\kappa S_n - n\ln\varphi(\kappa)}] = 1$, where $S_n = S(T_n)$. We could thus consider the macro process as a discrete classic risk model. Define a new stopping time $N^*(u) = \inf\{n \geq 0;\ S_n > u\}$ such that $\{N^*(u) < \infty\}$ is equivalent to $\{T^*(u) < \infty\}$. It is then true that for a stopping time $N^*(u)$ and $G \subseteq \{N^*(u) < \infty\}$,

$$\mathbb{P}\{G\} = \mathbb{E}_{\mathbb{Q}}\left[\frac{1}{L_{N^*(u)}};\ G\right],$$

according to Asmussen and Albrecher [2010] Chapter III. Theorem 1.3. Furthermore, since we have $\mathbb{E}_{\mathbb{Q}}[G] = 1$, i.e., $N^*(u) < \infty$, the optional stopping theorem could be applied thus achieving the desired result.

This means that we could simulate macro ruin probabilities under the new measure under which such ruin happens for sure. In general, the change of measure suggests that

$$\begin{aligned}
Y^{(\kappa)} &\sim \mathbf{Exp}(\beta - \kappa); \\
\tilde{\tau}^{>\xi} &\sim \mathbf{Exp}(\lambda_2 + \kappa) \ on \ (\xi, \infty); \\
\tau^{>\xi} &\sim \mathbf{Exp}(\lambda_1 + \kappa) \ on \ (\xi, \infty); \\
\tau^{\leq\xi} &\sim \mathbf{Exp}(\lambda_1 + \kappa) \ on \ (0, \xi]; \\
\tilde{\tau}^{\leq\xi} &\sim \mathbf{Exp}(\lambda_2 + \kappa) \ on \ (0, \xi].
\end{aligned}$$

Equivalently, the underlying process is given as

$$\mathbf{I}_{\{Z>\tilde{p}_\kappa\}}(Y_0 - \tilde{\tau}_0^{>\xi}) + \mathbf{I}_{\{Z\leq\tilde{p}_\kappa\}}\left(\sum_{i=1}^{N-1}(Y_i - \tilde{\tau}_i^{\leq\xi}) + \left(Y_N - \tilde{\tau}_N^{>\xi}\right) + \left(Y_0 - \tilde{\tau}_0^{\leq\xi}\right)\right), \tag{5.36}$$

where $N \sim Geo(p_\kappa)$ and $Z \sim U(0,1)$.

To sum up, this subsection used the importance sampling approach avoiding the common drawback from a crude Monte Carlo simulation and suggesting a way to simulate infinite-time macro ruin probabilities according to (5.2.4).

5.2.5 Embedded Markov additive process

In spite of the nice result we can get for ruin probabilities in the last subsection, it is not solving the underlying problem we proposed. Therefore, we study the nature of our process in more depth. Again, for simplicity, we assume everything to be exponential distributed with $\tau \sim Exp(\lambda_1)$, $\tilde{\tau} \sim Exp(\lambda_2)$ and $Y \sim Exp(\beta)$, respectively.

Recall our process described by (5.14). Note that ruin happens only at the moments of claim arrivals $\sigma_k = \sum_{i=1}^{k} \tau_i$ and $\sigma_0 = 0$. From time σ_k to σ_{k+1}, the distribution of the increment $S(\sigma_{k+1}) - S(\sigma_k)$ is only dependent on the relation between τ_k and ξ. Hence, we could transfer the original model into a new one by adding a Markov state process $\{J_n\}_{n\geq 0}$ defined on $E = \{1, 2\}$. $i \in E$ represents the occupying state of $\{J_k\}$ at time σ_k. For instance, state 1 describes a status where the current inter-arrival time is less or equal than ξ while state 2 refers to the opposite situation. For convenience, we construct τ_0 based on the choice of J_0: $J_0 = 1$ implies $\tau_0 < \xi$ and $\tau_0 \geq \xi$ otherwise. As we mentioned in Section 2 before, the two state Markov chain $\{J_n\}$ has a transition probability matrix as follows with the ij^{th} element being $p_{ij},\ i, j \in E$.

$$\mathbf{P} = \begin{bmatrix} q & p \\ \tilde{q} & \tilde{p} \end{bmatrix},$$

where $p = \mathbb{P}(\tau > \xi), q = 1-p = \mathbb{P}(\tau \leq \xi)$ and $\tilde{p} = \mathbb{P}(\tilde{\tau} > \xi),\ \tilde{q} = 1-\tilde{p} = \mathbb{P}(\tilde{\tau} \leq \xi)$. We also define a new process $\{S_n\}_{n\geq 0}$ whose increment $\Delta S_{n+1} = S_{n+1} - S_n$ is governed by $\{J_n\}$. More specifically, two scenarios could be analysed to explain this process. Given $n = 0, 1, \ldots$, scenario 1 is when $J_n = 1$, i.e., $\tau_n \leq \xi$ and $\tau_{n+1} \stackrel{d}{=} \tau$. Then comparing τ with ξ, there is a chance q of obtaining $J_{n+1} = 1$ given $\tau \leq \xi$, and p having $J_{n+1} = 2$ given $\tau > \xi$, with the corresponding increment being $\Delta S_{n+1} \stackrel{d}{=} Y - \tau^{\leq\xi}$ and $\Delta S_{n+1} \stackrel{d}{=} Y - \tau^{>\xi}$, respectively. On the contrary, scenario 2 represents the situation where the current state is $J_n = 2$, i.e., $\tau_n > \xi$ and $\tau_{n+1} \stackrel{d}{=} \tilde{\tau}$. Thus, all the variables above are presented in the same way only with a tilde sign added on τ, p and q.

$\{S_n, J_n\}$ is a discrete time bivariate Markov process or referred to as a Markov additive process (MAP). In fact, $Z(\sigma_n) := (S(\sigma_n), J_n)$ $(n = 0, 1, 2, \ldots)$ coincides with $\{S_n, J_n\}$, starting at $Z(0) = (0, J_0)$, where J_0 is the initial state taking value 1 or 2. The moment of ruin is the first passage time of S_n given a process Z_n over level $u > 0$, defined by

$$T^{(i)}(u) = \inf\{n \in \mathbb{N} :\ S_n > u | Z(0) = (0, i),\ \text{for } i = 1, 2\}.$$

Note that $\sigma_{T^{(2)}(u)} = T(u)$ such that the event $\{T^{(2)}(u) < \infty\}$ is equivalent to

$\{T(u) < \infty\}$. That implies

$$\psi(u) = \mathbb{P}(T^{(2)}(u) < \infty).$$

Moving into details, this MAP is specified by a kernel matrix with the ij^{th} entry given by a measure $F_{ij}(dx) = \mathbb{P}_i(J_1 = j,\, \Delta S_1 \in dx)$. Here $\mathbb{P}_i$ and $\mathbb{E}_i$ denotes the probability measure conditional on the set $\{J_0 = i\}$ and its corresponding expectation, respectively. Then for $\theta > 0$, a m.g.f on the measure $F_{ij}(dx)$ is $\hat{F}_{ij}[\theta] = \mathbb{E}_i[\mathrm{e}^{\theta \Delta S_1}; J_1 = j]$. These elements consist of a matrix $\hat{\mathbf{F}}[\theta]$ and in our case

$$\hat{\mathbf{F}}[\theta] = \left[\begin{array}{cc} \mathbb{E}(\mathrm{e}^{\theta Y}\mathrm{e}^{-\theta\tau}; \tau \leq \xi) & \mathbb{E}(\mathrm{e}^{\theta Y}\mathrm{e}^{-\theta\tau}; \tau > \xi) \\ \mathbb{E}(\mathrm{e}^{\theta Y}\mathrm{e}^{-\theta\tilde{\tau}}; \tilde{\tau} \leq \xi) & \mathbb{E}(\mathrm{e}^{\theta Y}\mathrm{e}^{-\theta\tilde{\tau}}; \tilde{\tau} > \xi) \end{array} \right].$$

Additionally, define $\hat{F}_{n,ij}[\theta] = \mathbb{E}_i[\mathrm{e}^{\theta(S_n - S_0)}; J_n = j]$, then the following equation can be proved to hold Asmussen and Albrecher [2010] (Chapter III. 4).

$$\hat{\mathbf{F}}_n[\theta] = (\hat{\mathbf{F}}[\theta])^n.$$

Rather than considering a continuous time MAP, we propose similar results for a discrete time one. Initially,

Lemma 5.17.

$$\mathbb{E}_{J_n}[\mathrm{e}^{\theta(S_{n+1}-S_n)} v^{(\theta)}_{J_{n+1}}] = \lambda(\theta)\mathrm{v}^{(\theta)}_{\mathrm{J_n}}, \tag{5.37}$$

where $\lambda(\theta)$ is the eigenvalue of $\hat{\mathbf{F}}[\theta]$ and $\mathbf{v} = (v_1, v_2)^T$ is the corresponding right eigenvector.

Proof.

$$\mathbb{E}_{J_n}[\mathrm{e}^{\theta(S_{n+1}-S_n)} v^{(\theta)}_{J_{n+1}}] = \mathbf{e}^T_{J_n}\hat{\mathbf{F}}_1[\theta]\mathbf{v} = \mathbf{e}^T_{J_n}\lambda(\theta)\mathbf{v} = \lambda(\theta) v^{(\theta)}_{J_n},$$

where $\mathbf{e}_{J_n}$ is a standard basis vector. ■

Therefore,

Lemma 5.18. *The following sequence*

$$\left\{ \mathrm{e}^{\theta S_n - n \ln \lambda(\theta)} v^{(\theta)}_{J_n} \right\}_{n\in\mathbb{N}} \tag{5.38}$$

is a martingale.

Proof. Let $M_n = \mathrm{e}^{\theta S_n - n \ln \lambda(\theta)} v^{(\theta)}_{J_n}$. Then,

$$\begin{aligned}
\mathbb{E}[M_{n+1}|\mathcal{F}_n] &= \mathbb{E}[\mathrm{e}^{\theta S_{n+1} - (n+1)\ln\lambda(\theta)} v^{(\theta)}_{J_{n+1}} | \mathcal{F}_n] \\
&= \mathbb{E}[\mathrm{e}^{\theta(S_{n+1}-S_n)} v^{(\theta)}_{J_{n+1}} | \mathcal{F}_n] \mathrm{e}^{\theta S_n - (n+1)\ln\lambda(\theta)} \\
&= \mathbb{E}_{J_n}[\mathrm{e}^{\theta(S_{n+1}-S_n)} v^{(\theta)}_{J_{n+1}}] \mathrm{e}^{\theta S_n - (n+1)\ln\lambda(\theta)} \\
&= \lambda(\theta) v^{(\theta)}_{J_n} \mathrm{e}^{\theta S_n - (n+1)\ln\lambda(\theta)} \\
&= M_n
\end{aligned}$$

■

Next, if we implement the exponential change of measure as in Theorem 5.16, there would be a similar result to what has been described in the last section. Firstly, in addition to Lemma 5.18, it is also true that

$$L_n = \mathrm{e}^{\theta S_n - n \ln \lambda(\theta)} \frac{v^{(\theta)}_{J_n}}{v^{(\theta)}_{J_0}}, \; n \in \mathbb{N} \tag{5.39}$$

is a martingale due to the Markov property. Then,

Lemma 5.19. *Define a new conditional probability measure $\mathbb{Q}^{(\theta)}_i(dx) = \mathbb{Q}^{(\theta)}(dx|J_0 = i)$ and the Randon-Nikodym derivative is*

$$L_n = \frac{\mathbb{Q}^{(\theta)}_i(dx)}{\mathbb{P}_i(dx)} \mid \mathcal{F}_n,$$

where L_n is defined by (5.39) given some $\theta > 0$. Then under the new measure, the MAP $\{Z^{(\theta)}_n\}_{n\in\mathbb{N}}$ is specified by

$$\hat{\mathbf{F}}^{(\theta)}[\gamma] = \mathrm{e}^{-\ln\lambda(\theta)} \left(\mathbf{v}^{(\theta)}_{diag}\right)^{-1} \hat{\mathbf{F}}[\theta+\gamma]\mathbf{v}^{(\theta)}_{diag} \tag{5.40}$$

Proof. Initially, $F_{ij}^{(\theta)}(dx)$ can be written as

$$\begin{aligned} F_{ij}^{(\theta)}(dx) &= \mathbb{Q}_i^{(\theta)}(S_1 \in dx, J_1 = j) = \mathbb{E}^{\mathbb{Q}^{(\theta)}}[\mathbf{1}_{\{S_1 \in dx, J_1 = j\}} | J_0 = i] = \mathbb{E}_i[L_1 \mathbf{1}_{\{S_1 \in dx, J_1 = j\}}] \\ &= \mathrm{e}^{\theta x - \ln \lambda(\theta)} \frac{v_j^{(\theta)}}{v_i^{(\theta)}} F_{ij}(dx). \end{aligned}$$

This shows that the new measure is exponentially proportional to the old one, which ensures that $F_{ij}^{(\theta)}$ is absolutely continuous with respect to F_{ij}. Further transferring it into the matrix form yields the desired result. ■

Corollary 5.20. *Under the new measure $\mathbb{Q}^{(\theta)}$, the MAP $\{Z_n^{(\theta)}\}_{n \in \mathbb{N}}$ consists of a Markov state process $\{J_n^{(\theta)}\}_{n \in \mathbb{N}}$ which has a transition probability matrix*

$$\mathbf{P}^{(\theta)} = \begin{bmatrix} q_\theta & p_\theta \\ \tilde{q}_\theta & \tilde{p}_\theta \end{bmatrix}, \tag{5.41}$$

where

$$\begin{aligned} \tilde{p}_\theta = \frac{\beta \lambda_2}{(\beta - \theta)(\lambda_2 + \theta)} e^{-(\lambda_2 + \theta)\xi} \quad &, \quad \tilde{q}_\theta = 1 - \tilde{p}_\theta; \\ q_\theta = \frac{\beta \lambda_1}{(\beta - \theta)(\lambda_1 + \theta)} (1 - e^{-(\lambda_1 + \theta)\xi}) \quad &, \quad p_\theta = 1 - q_\theta, \end{aligned}$$

and an additive component $\{S_n^{(\theta)}\}_{n \in \mathbb{N}}$ with random variables $Y, \tau^{>\xi}, \tau^{<\xi}, \tilde{\tau}^{>\xi}, \tilde{\tau}^{<\xi}$ under the new measure $\mathbb{Q}^{(\theta)}$ given by Theorem 5.16 in terms of θ rather than κ.

In fact, when $\theta = \kappa$, $\mathbb{Q}^{(\theta)}$ coincides with $\mathbb{Q}$ defined by Theorem 5.16. Recall $T^*(u)$ from (5.28) and $\psi(u) \geq \psi^*(u)$. Since $\sigma_{T^{(2)}(u)} \leq T^*(u)$, then $\mathbb{Q}(T^*(u) < \infty) = 1$ implies $\mathbb{Q}^{(\kappa)}(T^{(2)}(u) < \infty) = 1$. In addition,

Lemma 5.21. *The ruin probability for the underlying process (5.14) is*

$$\psi(u) = v_2^{(\kappa)} e^{-\kappa u} \mathbb{E}_2^{(\kappa)} \left[\frac{e^{-\kappa \varepsilon(T^{(2)}(u))}}{v_{J_{T^{(2)}(u)}}^{(\kappa)}} \right], \tag{5.42}$$

where $\varepsilon(T^{(2)}(u)) = S_{T^{(2)}(u))}^{(\kappa)} - u$ denotes the overshoot at the time of ruin $T^{(2)}(u))$.

In addition, it has been discovered that

Remark 5.22. $\hat{\mathbf{F}}[\kappa]$ *has an eigenvalue equal to 1 and*

$$\mathbf{v}^{(\kappa)} = \begin{bmatrix} \frac{\beta\lambda_1}{(\beta-\kappa)(\lambda_1+\kappa)} - q_\kappa \\ p_\kappa \end{bmatrix}$$

is the corresponding right eigenvector.

Proof. Let λ denote the eigenvalue of $\hat{\mathbf{F}}[\kappa]$. Thus, we can write,

$$(\mathbb{E}[\mathrm{e}^{\kappa Y}]\mathbb{E}[\mathrm{e}^{-\kappa\tau}, \tau \leq \xi]-\lambda)(\mathbb{E}[\mathrm{e}^{\kappa Y}]\mathbb{E}[\mathrm{e}^{-\kappa\tilde{\tau}}\tilde{\tau} > \xi]-\lambda) = (\mathbb{E}[\mathrm{e}^{\kappa Y}])^2\mathbb{E}[\mathrm{e}^{-\kappa\tau}, \tau > \xi]\mathbb{E}[\mathrm{e}^{-\kappa\tilde{\tau}}, \tilde{\tau} \leq \xi].$$

Recall (5.18), clearly $\lambda = 1$ is a solution to the above equation. That directly leads to $\hat{\mathbf{F}}\mathbf{v} = \mathbf{v}$ and one can obtain

$$\frac{v_1}{v_2} = \frac{\mathbb{E}[\mathrm{e}^{-\kappa\tau}, \tau > \xi]}{1 - \mathbb{E}[\mathrm{e}^{-\kappa\tau}, \tau \leq \xi]}.$$

Plugging in the parameters completes the proof. ■

Example 5.23. *Here is an example. Assume Y, τ and $\tilde{\tau}$ have exponential distribution with parameters $\beta = 3$, $\lambda_1 = 1$ and $\lambda_2 = 2$, respectively. The smallest positive real root of Equation (5.18) is calculated for $\kappa = 1.1439$ and its corresponding right eigenvector is $\mathbf{v}^{(k)} = [0.5790, 0.8153]'$. Then the ruin function is plotted as in Figure 5.7. Without surprise, it shows an exponential decay, which again confirms Theorem 5.13.*

5.2.6 Method using renewal equations

This subsection shows several intermediate results toward solving the ruin probability under the simplest case - everything being Exponentially distributed. Trials include differentiation and the adoption of Laplace Transform which naturally leads to the use of Dickson-Hipp operator Li and Garrido [2004] in this case. Unfortunately, neither of the methods solved for ruin probabilities analytically. The former approach stops with a system of second order (Negative) Delay-differential Equations which is to be solved. The latter one establishes the relations between ruin probabilities for two states which is not yet enough. However, I would still

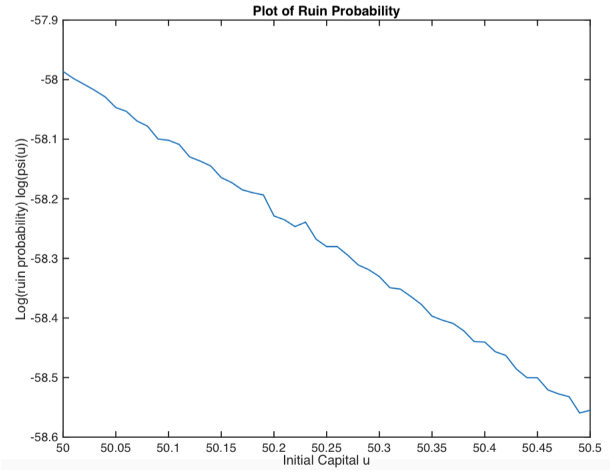


Figure 5.7: Example of ruin function generated by an MAP

like to leave these calculations in the thesis for future references.

Assume that τ and $\tilde{\tau}$ are exponentially distributed with density functions $f_1(t) = \lambda_1 e^{-\lambda_1 t}$ and $f_2(t) = \lambda_2 e^{-\lambda_2 t}$ respectively. Then the renewal integral equations can be written as,

$$\psi_1(u) = \int_0^{\xi} f_1(t) g_1(u+t)dt + \int_{\xi}^{\infty} f_1(t) g_2(u+t)dt; \qquad (5.43)$$

$$\psi_2(u) = \int_0^{\xi} f_2(t) g_1(u+t)dt + \int_{\xi}^{\infty} f_2(t) g_2(u+t)dt. \qquad (5.44)$$

where $\psi_1(u), \psi_2(u)$ correspond to the ruin probabilities with the first inter-arrival time being τ and $\tilde{\tau}$ respectively, and

$$g_i(u) = \int_0^u \psi_i(u-y) b(y)dy + \int_u^{\infty} b(y)dy, \; i = 1, 2$$

with $b(y)$ being the density function of the claim sizes.

Differentiation

Lemma 5.24. *[IDE] Assume $\psi_i(u), i = 1, 2$ is first order differentiable, then*

$$\begin{aligned}\lambda_1\psi_1(u) - \frac{d}{du}\psi_1(u) &= \lambda_1 g_1(u) + f_1(\xi)(g_2(u+\xi) - g_1(u+\xi)); \quad (5.45)\\ \lambda_2\psi_2(u) - \frac{d}{du}\psi_2(u) &= \lambda_2 g_1(u) + f_2(\xi)(g_2(u+\xi) - g_1(u+\xi)). \quad (5.46)\end{aligned}$$

Proof. Taking the derivative on both sides of (5.43) yields

$$\begin{aligned}\frac{d\psi_1(u)}{du} &= \int_0^\xi f_1(t)\frac{dg_1(u+t)}{du}dt + \int_\xi^\infty f_1(t)\frac{dg_2(u+t)}{du}dt\\ &= \int_0^\xi f_1(t)dg_1(u+t) + \int_\xi^\infty f_1(t)dg_2(u+t)\\ &= f_1(t)g_1(u+t)\big|_0^\xi + f_1(t)g_2(u+t)\big|_\xi^\infty\\ &\quad +\lambda_1\int_0^\xi f_1(t)g_1(u+t)dt + \lambda_1\int_\xi^\infty f_1(t)g_2(u+t)dt\\ &= f_1(\xi)(g_1(u+\xi) - g_2(u+\xi)) - \lambda_1 g_1(u) + \lambda_1\psi_1(u).\end{aligned}$$

Rearranging the equation gives the proposed result as stated in the lemma. Similarly, we obtain the equation with respect to $\psi_2(u)$. ■

We can further obtain a system of second-order (negative) delay-differential equations (NDDE). Normally, the definition of a delay-differential equation (DDE) can be found in Erneux [2008]. The ones we obtained have negative shifts, so they are simply referred to as NDDEs.

Lemma 5.25. *[NDDE] Assume $\psi_i(u), i = 1, 2$ is second order differentiable, then*

$$\begin{aligned}(\lambda_1 - \beta)\psi_1'(u) - \psi_1''(u) &= \beta f_1(\xi)[\psi_2(u+\xi) - \psi_1(u+\xi)]; \quad (5.47)\\ (\lambda_2 - \beta)\psi_2'(u) - \psi_2''(u) &= \beta f_2(\xi)[\psi_2(u+\xi) - \psi_1(u+\xi)]\\ &\quad -\lambda_2\beta(\psi_2(u) - \psi_1(u)). \quad (5.48)\end{aligned}$$

Proof. We differentiate on both sides of (5.45) and (5.46).

$$\begin{aligned}\lambda_1\psi_1'(u) - \psi_1''(u) &= \lambda_1 g_1'(u) + f_1(\xi)\left(\frac{d}{du}g_2(u+\xi) - \frac{d}{du}g_1(u+\xi)\right); \quad (5.49)\\ \lambda_2\psi_2'(u) - \psi_2''(u) &= \lambda_2 g_1'(u) + f_2(\xi)\left(\frac{d}{du}g_2(u+\xi) - \frac{d}{du}g_1(u+\xi)\right). \quad (5.50)\end{aligned}$$

First notice that we have, for $i = 1, 2$,

$$\frac{dg_i(u)}{du} = \psi_i(u)b(0) + \int_0^u \psi_i(y)\frac{db(u-y)}{du}dy - b(u) \tag{5.51}$$
$$= \psi_i(u)b(0) - \beta g_i(u). \tag{5.52}$$

And similarly,

$$\frac{dg_i(u+\xi)}{du} = \psi_i(u+\xi)b(0) + \int_0^{u+\xi} \psi_i(y)\frac{db(u+\xi-y)}{du}dy - b(u+\xi)$$
$$= \psi_i(u+\xi)b(0) - \beta g_i(u+\xi).$$

Hence,

$$\frac{dg_2(u+\xi)}{du} - \frac{dg_1(u+\xi)}{du}$$
$$= \beta[\psi_2(u+\xi) - \psi_1(u+\xi)] - \beta[g_2(u+\xi) - g_1(u+\xi)]. \tag{5.53}$$

Plugging (5.52) and (5.53) into (5.49) and (5.50) gives the assertion. ■

Laplace Transform

On the other hand, a linear combination of (5.45) and (5.46) simply gives,

$$f_2(\xi)\left[\lambda_1\psi_1(u) - \frac{d}{du}\psi_1(u)\right] - f_1(\xi)\left[\lambda_2\psi_2(u) - \frac{d}{du}\psi_2(u)\right] \tag{5.54}$$
$$= \lambda_1 f_2(\xi)g_1(u) - \lambda_2 f_1(\xi)g_1(u). \tag{5.55}$$

By taking the Laplace Transform on both sides we obtain,

$$\left[\rho(\lambda_1 - s) - (\rho\lambda_1 - \lambda_2)\hat{b}(s)\right]\hat{\psi}_1(s) - (\lambda_2 - s)\hat{\psi}_2(s)$$
$$= (\rho\lambda_1 - \lambda_2)\frac{1-\hat{b}(s)}{s} - \rho\psi_1(0) + \psi_2(0).$$

This leads to

$$\hat{\psi}_2(s) = \frac{\rho(\lambda_1 - s) - (\rho\lambda_1 - \lambda_2)\hat{b}(s)}{\lambda_2 - s}\cdot\hat{\psi}_1(s) + \frac{\rho\psi_1(0) - \psi_2(0)}{\lambda_2 - s} - \frac{(\rho\lambda_1 - \lambda_2)(1-\hat{b}(s))}{s(\lambda_2 - s)}.$$

As a result,

$$\hat{\psi}_2(s) = \left(\rho + w_1 \cdot \frac{-\lambda_2}{s-\lambda_2} - w_2 \cdot \frac{\beta}{s+\beta}\right)\hat{\psi}_1(s) + w_3 \cdot \frac{-\lambda_2}{s-\lambda_2} - w_4 \cdot \frac{\beta}{s+\beta}, \tag{5.56}$$

where

$$\begin{aligned}
w_1 &= \frac{\rho(\lambda_1-\lambda_2)+(1-\rho)\beta}{\beta+\lambda_2};\\
w_2 &= \frac{\rho\lambda_1-\lambda_2}{\beta+\lambda_2};\\
w_3 &= \frac{(\beta+\lambda_2)(\rho\psi_1(0)-\psi_2(0))-\rho\lambda_1+\lambda_2}{\lambda_2(\beta+\lambda_2)};\\
w_4 &= \frac{\rho\lambda_1-\lambda_2}{\beta(\beta+\lambda_2)}.
\end{aligned}$$

Besides, the Dickson-Hipp operator appeared in Li and Garrido [2004] is defined as follows and will be used in the sequel.

Definition 5.26. *For any integrable function $f : [0,\infty) \mapsto \mathbb{R}$ and a real constant $s > 0$, the Dickson-Hipp operator $\mathcal{T}_s$ is defined by*

$$\mathcal{T}_s f(x) = e^{sx}\int_x^\infty e^{-sy} f(y)dy.$$

It has a number of useful properties as follows

Lemma 5.27. *For any integrable functions $f, g : [0,\infty) \mapsto \mathbb{R}$ and real contants $s, t > 0$, we have*

1. $$\mathcal{T}_s\mathcal{T}_t f = \frac{\mathcal{T}_s f(x) - \mathcal{T}_t f(x)}{t-s}.$$

2. *Denote $*$ as the convolution operator,*

$$\mathcal{T}_s(f * g) = \mathcal{T}_s g(0)\cdot\mathcal{T}_s f + (\mathcal{T}_s g) * f,$$

3. *Denote $\mathcal{I}$ as the identity operator and $\mathcal{D}$ the differential operator,*

$$(s\mathcal{I} - \mathcal{D})\mathcal{T}_s f = f.$$

Proof. Proof can be seen in Feng [2008]. ■

Using these properties, we could show for $i = 1, 2$:

Lemma 5.28. *1.*

$$\mathcal{T}_s g_i(\xi) = \hat{b}(s) \cdot \mathcal{T}_s \psi_i(\xi) + \mathcal{T}_s b * \psi_i(\xi) + \frac{\bar{B}(\xi) - \mathcal{T}_s b(\xi)}{s}.$$

2.

$$\mathcal{T}_s(\mathcal{D}\psi_i)(\xi) = s\mathcal{T}_s\psi_i(\xi) - \psi_i(\xi).$$

Proof. The proof is similar to the properties of Laplace transform. ■

Remark 5.29. *The Dickson-Hipp operator on $g_i(u+\xi), i = 1, 2$ is equivalent to $\mathcal{T}_s g_i(2\xi), i = 1, 2$.*

Proof. Let $h_i(u) = g_i(u+\xi), u \geq 0, i = 1, 2$, then we can write

$$\mathcal{T}_s h(\xi) = e^{s\xi} \int_\xi^\infty e^{-sy} h_i(y) dy = e^{s\xi} \int_\xi^\infty e^{-sy} g_i(y+\xi) dy.$$

Let $z = y + \xi$, then $y = z - \xi$ and $z \in [2\xi, \infty)$, the above equation evolves to

$$\mathcal{T}_s h(\xi) = e^{2s\xi} \int_{2\xi}^\infty e^{-sz} g_i(z) dz = \mathcal{T}_s g_i(2\xi), \tag{5.57}$$

for $i = 1, 2$. ■

Lemma 5.30. *The relation between $\mathcal{T}_s\psi_1(\xi)$ and $\mathcal{T}_s\psi_2(\xi)$ is illustrated by*

$$\begin{aligned}
\mathcal{T}_s\psi_2(\xi) \quad = \quad & \frac{1}{f_1(\xi)(\lambda_2 - s)} \Big\{ [f_2(\xi)(\lambda_1 - s - \lambda_1\hat{b}(s)) + f_1(\xi)\lambda_2\hat{b}(s)]\mathcal{T}_s\psi_1(\xi) \\
& + [\lambda_2 f_1(\xi) - \lambda_1 f_2(\xi)] \left(\mathcal{T}_s b * \psi_1(\xi) + \frac{\bar{B}(\xi) - \mathcal{T}_s b(\xi)}{s} \right) \\
& + f_2(\xi)\psi_1(\xi) - f_1(\xi)\psi_2(\xi)] \Big\}.
\end{aligned}$$

Proof. That is done by applying the Dickson-Hipp transform $\mathcal{T}_s \cdot (\xi)$ on both sides of (5.45) and (5.46).

$$(\lambda_1 - s - \lambda_1\hat{b}(s))\mathcal{T}_s\psi_1(\xi) + \psi_1(\xi) - \lambda_1\mathcal{T}_s b * \psi_1(\xi) - \lambda_1\frac{\bar{B}(\xi) - \mathcal{T}_s b(\xi)}{s}$$
$$= f_1(\xi)[\mathcal{T}_s g_2(2\xi) - \mathcal{T}_s g_1(2\xi)];$$
$$(\lambda_2 - s)\mathcal{T}_s\psi_2(\xi) + \psi_2(\xi) - \lambda_2\hat{b}(s)\mathcal{T}_s\psi_1(\xi) + \lambda_2\mathcal{T}_s b * \psi_1(\xi) - \lambda_2\frac{\bar{B}(\xi) - \mathcal{T}_s b(\xi)}{s}$$
$$= f_2(\xi)[\mathcal{T}_s g_2(2\xi) - \mathcal{T}_s g_1(2\xi)].$$

Similar to equation (5.55), we have now

$$\begin{aligned} & f_2(\xi)\big[\lambda_1\mathcal{T}_s\psi_1(\xi) - s\mathcal{T}_s\psi_1(\xi) + \psi_1(\xi)\big] - f_1(\xi)\big[\lambda_2\mathcal{T}_s\psi_2(\xi) - s\mathcal{T}_s\psi_2(\xi) + \psi_2(\xi)\big] \\ = \;& \lambda_1 f_2(\xi)\Big[\hat{b}(s)\cdot\mathcal{T}_s\psi_1(\xi) + \mathcal{T}_s b * \psi_1(\xi) + \frac{\bar{B}(\xi) - \mathcal{T}_s b(\xi)}{s}\Big] \\ & -\lambda_2 f_1(\xi)\Big[\hat{b}(s)\cdot\mathcal{T}_s\psi_2(\xi) + \mathcal{T}_s b * \psi_2(\xi) + \frac{\bar{B}(\xi) - \mathcal{T}_s b(\xi)}{s}\Big]. \end{aligned}$$

Rearranging the above identity presents us the results as stated in the lemma.

■

Now, we expand $\mathcal{L}g_i(u+\xi)(s)$.

Lemma 5.31.

$$\mathcal{L}g_i(u+\xi)(s) = \mathcal{T}_s g_i(\xi) = \hat{b}(s)\cdot\mathcal{T}_s\psi_i(\xi) + \mathcal{T}_s b * \psi_i(\xi) + \frac{\bar{B}(\xi) - \mathcal{T}_s b(\xi)}{s}. \quad (5.58)$$

Proof.

$$
\begin{aligned}
\mathcal{L}g_i(u+\xi)(s) &= \int_0^\infty e^{-su}\int_0^{u+\xi}\psi_i(y)b(u+\xi-y)dydu + \int_0^\infty e^{-su}\int_{u+\xi}^\infty b(y)dydu \\
&= \int_0^\infty e^{-su}\int_\xi^{u+\xi}\psi_i(y)b(u+\xi-y)dydu \\
&\quad + \int_0^\infty e^{-su}\int_0^\xi \psi_i(y)b(u+\xi-y)dydu + \frac{\bar{B}(\xi)-\mathcal{T}_s b(\xi)}{s} \\
&= \int_\xi^\infty \int_{y-\xi}^\infty e^{-su}\psi_i(y)b(u+\xi-y)dudy + \mathcal{T}_s b * \psi_i(\xi) + \frac{\bar{B}(\xi)-\mathcal{T}_s b(\xi)}{s} \\
&= \int_\xi^\infty e^{-sy+s\xi}\psi_i(y)dy \cdot \int_{y-\xi}^\infty e^{-s(u+\xi-y)}b(u+\xi-y)du \\
&\quad + \mathcal{T}_s b * \psi_i(\xi) + \frac{\bar{B}(\xi)-\mathcal{T}_s b(\xi)}{s} \\
&= \hat{b}(s)\cdot \mathcal{T}_s\psi_i(\xi) + \mathcal{T}_s b * \psi_i(\xi) + \frac{\bar{B}(\xi)-\mathcal{T}_s b(\xi)}{s}.
\end{aligned}
$$

■

<u>Other remarks</u>

Remark 5.32. *If we write*

$$
\begin{aligned}
\psi_1(u) &\sim K_1 e^{-\kappa u}; && (5.59)\\
\psi_2(u) &\sim K_2 e^{-\kappa u}, && (5.60)
\end{aligned}
$$

then the following relation is shown to hold for the two constant components (K_1 and K_2) in the ruin functions.

$$
\begin{aligned}
\frac{K_1}{K_2} &= \frac{\beta f_1(\xi)e^{-\kappa\xi}}{\beta f_1(\xi)e^{-\kappa\xi}-\kappa(\lambda_1-\beta)-\kappa^2}; \\
\frac{\lambda_2\beta-\beta f_2(\xi)e^{-\kappa\xi}}{\lambda_2\beta-(\lambda_2-\beta)\kappa-\kappa^2-\beta f_2(\xi)e^{-\kappa\xi}} &= \frac{\beta f_1(\xi)e^{-\kappa\xi}-(\lambda_1-\beta)\kappa-\kappa^2}{\beta f_1(\xi)e^{-\kappa\xi}}, \qquad (5.61)
\end{aligned}
$$

where κ is the solution to (5.22).

Proof. As shown in Theorem 5.13, asymptotically we have (5.59) and (5.60). Plugging these back into (5.47) and (5.48) will lead to the proposed result. Via some further calculations, we could find that (5.61) coincides with (5.22). ■

This subsection shows how the integral equation method could be implemented on our model. Calculations suggest that a relation of $\psi_1(u)$ and $\psi_2(u)$ can be expressed in terms of Laplace transform and a system of second order NDDE was obtained. Although Dickson-Hipp operator turns on the light towards dealing with modified Laplace transforms, solutions are not yet explicit. One way of proceeding with this is to seek for some numerical methods solving the NDDEs. Figuring out the boundary conditions though remains the main challenge.

5.3 Discussions

In a BM system, premiums are adjusted according to claim histories on the purpose of providing a fair share of risks. We follow the idea in this work and first extend it to a broader concept where risks are distinguished not only among policyholders, but also among themselves. There are known risks for which historical data is available and expectations can be made based on past observations. On the other hand, in modern era, while we are enjoying the benefits brought by the ever improving technology, we might also encounter some hidden or unknown risks, which might possibly create significant losses to insurance companies. Therefore, it is worthwhile accounting for such 'unforeseeable' risks. When incorporating these risks together with the classical (referred 'historical' here) ones in a risk model, the number of claims can be described as a Poisson with a random parameter Λ, continuous random variable, that can be defective at $\{0\}$ or not.

Considering the ruin probabilities for the 'unforeseeable' stream alone (2.5) and then a combination of both the 'unforeseeable' and the 'historical' streams (5.6), we derive relationships between the probability of ruin in the classical case (1.3), versus the case where the premiums are adjusted to the history of claims (5.1). Unlike the case for the 'historical' stream only Dubey [1977], we found that the ruin probability for a risk model distinguishing the 'historical' stream and the 'unforeseeable' stream is different from that in a classical case. The differences are amenable and thus this theory should encourage insurance companies to use adjusted premium rates in an attempt to reward their good customers and at the same time to protect the insurer itself, as in a classical BM system. The separation of risks would allow an insurance company not only to fairly distribute the premiums among its customers, but to also correctly incorporate their exposure to 'historical' risks versus the 'unforeseeable' ones, brought on for instance by the progress in technology.

From another perspective, we found that a simple Bonus system could be reflected by a dependence structure embedded in a risk model. For the simplest case, we made inter-arrival times switch between two random variables by comparing them with a fixed window ξ. Such interchange was equivalently converted from the change of premium rates based on recent claims as shown by Figure 2.1 emulating a basic no claim discount (NCD) system where there are only two classes - either a base or discounted level. Theoretically speaking, it also works for a merely Malus system. Yet in practice, such system does not exist as it probably sounds more tempting if an insurance company offers awards rather than a penalty.

Several different approaches have been undertaken to study the ruin probability under the framework of a regenerative process. It is not surprising under the Cramér assumption, the ruin function still has an exponential tail. By Monte Carlo simulations, it has been discovered that the underlying probability has opposite monotonicity with respect to ξ when two random variables for the inter-claim times swap parameters. It has also been found that the use of BM systems may not reduce ruin probabilities when we made a comparison between our results and the classical ones. Furthermore, we explained how we could construct a discrete Markov additive process from the model under concern when everything is exponentially distributed. By a change of measure via exponential families, ruin probabilities were possible to be simulated through a more convenient form (5.42). Last but not least, calculations using integral equations have also been carried out and a system of a second order NDDEs could be achieved, yet it remains difficult to seek for analytical solutions.

To conclude, this chapter looked at BM systems from a risk assessment orientation under a continuous time horizon. Some results were obtained by comparing with a classical risk model while others, although not in explicit form, were simulated and could be presented in a nice form under a simple example. Future extensions are possible when a more complicated structure of risks are constructed, e.g., introducing a third or even more risk divisions. In the other model, it would be interesting to dig into more depth on solving the 'advanced' system of ODEs. Reconstructing the model by incorporating more classes would be rather realistic and also theoretically appealing.

Chapter 6

Concluding Remarks

This thesis has dedicated work into the modelling of BM systems as well as analyses of risks such systems are exposed to. The ever-growing popularity of BM systems in reality is probably due to its feature of customising risk distributions. It provides mutual benefits to insurers and policyholders. While the former gains more accurate estimation of risks thus being able to offer more competitive policies, the latter has access to discounts in payment with careful driving. The use of this system though has created the so-called bonus hunger issue. Although it would not give rise to direct losses to an insurer, it bares the potential of incurring higher damages in the future.

Section 3.2 undertook Bayesian approach to reflect the use of a personalised history to estimate expected individual claims. By adopting a Weibull distribution for the claim severities, an alleviation on bonus hunger concerns could be identified. The proposed model suggests a very active strategy to encourage drivers to report each additional small claims. In this way, insurance companies are likely to keep track on the real cookies of an individual thus being able to take preventive steps in controlling the risks. Furthermore, a hybrid model (Weibull, Pareto) was employed in modelling the severity component. Results suggest that a mild strategy is encouraged for people filing many large claims when the aggregate expenses are fixed, whereas it is more harsh when people report many small claims. Rather than conclusion drawn from the previous model, this hybrid model proposes a pricing strategy that is strict towards drivers constantly reporting many small claims and mild towards people with frequent large claims as such frequency actually indicates cheaper per-claim cost.

In practice, such a model might be argued for its complexity and inefficiency due to very competitive environment in car insurance markets. However, insurance companies are free to make their own choices to create a discrete scale of the premium levels. With exact premium values calculated based on claim histories, it would be flexible and simple to conduct this further step. On the other hand, it remains to be the insurer's decision on whether to penalise policyholders with many small claims. It completely depends on the risk preferences of an insurer. For insurers who are more risk adverse, the hybrid model would fit better for their choice. In addition, these insurers are advised to keep the premiums constant after the drop point if they opt to use the first model we proposed.

From another perspective, the rest of the thesis looked into measuring the associated risks with a BM system. We first start from a discrete model presented in Chapter 4. Motivated by improving a reinsurance company's market competitiveness, we suggest the use of a simple BM system for their portfolio. Then the collective risks can be assessed via ruin probabilities. Through identifying a Markov chain related to our particular example, we were able to establish ruin functions via recursive relations. Attaining them in transformed forms and analysing boundary conditions, we could not find a nicer way rather than tedious calculations to compute analytical solutions even for a simple and everything-discrete model. However, this work has provided us with an in depth understanding of the renewal feature of the system laying a good foundation for later work.

In addition, a lot of work has been done carrying out the on risk analysis under a continuous framework. Therefore, Chapter 5 is divided into two parts with each one based on a paper Li et al. [2015]; Constantinescu et al. [2015b]. The first paper has recently been published at *Insurance: Mathematics and Economics* and the second one is under progress to submit. The first one involves the idea mentioned in Chapter 3 and adopted a premium adjustment according to a Bayesian estimator. This work relies mostly on connecting the underlying risk model with a classical one. Thus, properly modifying the results would lead to a desired answer. The other novelty of this work is the suggestion of dividing risks into two streams so that those associated with modern technologies are taken into account. Insurers are advised to launch such 'innovative' policies under the guidance of ruin probabilities proposed here. On the other hand, Section 5.2 incorporates a dependence structure in the risk model in order to imitate the dynamics of a no claim discount system (Bonus system). Asymptotic results show that ruin function retains the exponential tail. Besides, simulations imply a strong connection of ruin probabilities with the chosen fixed window. Further using a Markovian structure and change of measure, we obtained a nicer form to be implemented through a more sophisticated simulation. In the end, integral equations have been presented which lead to a system of second order NDDEs as well as a relation in terms of Laplace transforms.

To wrap up, this work has provided insights into BM systems and risk theory. Aiming at connecting these two, results were obtained under several model settings, both discrete and continuous. However, the author still feels the need for developing a more general framework. One interesting direction is to consider deductibles in a BM system. Some work has already been done using MLE for other models Paulsen and Stubø [2011]. It possibly needs refinement when a Bayesian estimation is used. Another intriguing path could be through Afonso et al. [2009] which resembles a BM feature in the best way so far. Extensions are possible if we seek for the Markovian environment hidden behind or we could consider a dependence only on the number of claims.

Appdx A

Derivation of Weibull Distribution

Lemma .1. *As in Albrecher et al. [2011], if an Exponential distribution has a Lévy random parameter, then the mixing distribution is given by*

$$F(x) = 1 - exp(-c\sqrt{x}),$$

which represents a Weibull distribuion with shape parameter $1/2$.

Proof. Some techniques were used to calculate the distribution function of the mixing distribution.

$$F(x) = \int_0^\infty (1 - e^{-\theta x}) \frac{c}{2\sqrt{\pi\theta^3}} exp\left(-\frac{c^2}{4\theta}\right) d\theta.$$

By the change of variables, let

$$\delta = \frac{c}{2\sqrt{\theta}},\ \theta = \frac{c^2}{2\sigma^2},$$

with $\theta \in (0, \infty)$, ν will decrease from ∞ to 0. Then $d\delta = -\frac{1}{2}\frac{c}{2\sqrt{\theta^3}}d\theta$ and $-2d\delta =$

$\frac{c}{2}\theta^{-\frac{3}{2}}d\theta$. The integral is then modified to the following form.

$$\begin{aligned} F(x) &= \int_0^\infty \left(1 - e^{-\frac{c^2}{4\delta^2}x}\right)\frac{2}{\sqrt{\pi}}e^{-\delta^2}d\delta \\ &= 1 - \frac{2}{\sqrt{\pi}}\int_0^\infty exp\left\{-\left(\frac{c^2}{4\delta^2}x + \delta^2\right)\right\}d\delta. \end{aligned} \tag{1}$$

The equating of $\int_0^\infty \frac{2}{\sqrt{\pi}}e^{-\delta^2}d\delta = 1$ comes from the well-known fact that $\int_0^\infty e^{-\delta^2}d\delta = \frac{\sqrt{\pi}}{2}$. Before continuing the above equation, we let,

$$\frac{c^2x}{4} = a \geq 0,$$

and use the letter I to denote the integral in equation (16), then it becomes,

$$I = \frac{2}{\sqrt{\pi}}\int_0^\infty e^{-\left(\frac{a}{\delta^2}+\delta^2\right)}d\delta.$$

On the other hand, if we look at the integral,

$$I_0 = \int_0^\infty e^{-\left(\delta-\frac{\sqrt{a}}{\delta}\right)^2}d\left(\delta - \frac{\sqrt{a}}{\delta}\right) = \int_0^\infty e^{-\left(\delta-\frac{\sqrt{a}}{\delta}\right)^2}d\delta + \int_0^\infty e^{-\left(\delta-\frac{\sqrt{a}}{\delta}\right)^2}\frac{\sqrt{a}}{\delta^2}d\delta.$$

Now let $\epsilon = \frac{\sqrt{a}}{\delta}$, with δ increasing from 0 to infinity. ϵ is decreasing from infinity to 0. And again we have $d\epsilon = -\frac{\sqrt{a}}{\delta^2}d\delta$. Hence, the latter term in the above integral is altered. We have,

$$I_0 = 2\int_0^\infty e^{-\left(\delta-\frac{\sqrt{a}}{\delta}\right)^2}d\delta. \tag{2}$$

On the other hand, if we let,

$$g = \delta - \frac{\sqrt{a}}{\delta} \in (-\infty,\, \infty),$$

Then,

$$\frac{dg}{d\delta} = 1 + \frac{\sqrt{a}}{\delta^2} > 0.$$

This indicates that g is monotonically increasing from minus infinity to infinity. Thus, the integral I_0 is alternatively written as,

$$I_0 = \int_0^{\infty} e^{-\left(\delta - \frac{\sqrt{a}}{\delta}\right)^2} \left(1 + \frac{\sqrt{a}}{\delta^2}\right) d\delta = \int_{-\infty}^{\infty} e^{-g^2} dg = \sqrt{\pi} \tag{3}$$

Combining the results (6) and (7), we know that

$$\int_0^{\infty} e^{-\left(\delta - \frac{\sqrt{a}}{\delta}\right)^2} d\delta = \frac{\sqrt{\pi}}{2}.$$

Recall that the integral we would like to solve is actually

$$I = \frac{2}{\sqrt{\pi}} e^{-2\sqrt{a}} \int_0^{\infty} e^{-\left(\delta - \frac{\sqrt{a}}{\delta}\right)^2} d\delta = e^{-2\sqrt{a}} = exp(-c\sqrt{x}),$$

Therefore, we have proved that the resulting distribution function is

$$F(x) = 1 - exp(-c\sqrt{x}).$$

■

Proof of Proposition 3.1

Proof. The form we often see on various mathematical handbooks regarding the modified Bessel function is presented below.

$$B_v(x) = \int_0^{\infty} e^{-xcosht} cosh(vt) dt,$$

Expanding the cosh terms, we can write,

$$B_v(x) = \int_0^\infty exp\left(-\frac{x}{2}(e^t + e^{-t})\right)\frac{e^{vt} + e^{-vt}}{2}dt.$$

It can be transformed initially by changing of the variable. We let $y = e^t$ with $t \in (0, \infty)$, $y \in (1, \infty)$ and $t = lny$, $dt = \frac{1}{y}dy$. Substituting into the above formula yields,

$$B_v(x) = \frac{1}{2}\int_1^\infty exp\left(-\frac{x}{2}(y + \frac{1}{y})\right)y^{v-1}dy + \frac{1}{2}\int_1^\infty exp\left(-\frac{x}{2}(y + \frac{1}{y})\right)y^{-(v+1)}dy.$$

Before continuing, we set $z = \frac{1}{y}$, $y = \frac{1}{z}$ with $y \in (1, \infty)$, $z \in (0, 1)$ and $dy = -\frac{1}{z^2}dz$. The second integral is then modified. Notice here the sign of the integral will change due to the alteration of the domain of the variable. It follows that,

$$\begin{aligned} B_v(x) &= \frac{1}{2}\int_1^\infty exp\left(-\frac{x}{2}\left(y + \frac{1}{y}\right)\right)y^{v-1}dy - \frac{1}{2}\int_0^1 exp\left(-\frac{x}{2}\left(\frac{1}{z} + z\right)\right)z^{v+1}\left(-\frac{1}{z^2}\right)dz \\ &= \frac{1}{2}\int_0^\infty exp\left(-\frac{x}{2}\left(y + \frac{1}{y}\right)\right)y^{v-1}dy. \end{aligned}$$

This has completed the proof. ■

Data used in Chapter 3

Groups(Range of Claims Severities in GBP)	**Number of Claims**
0-250	52
250-500	43
500-1500	33
1500-2500	28
2500-3500	22
3500-4500	19
4500-6500	14
6500-9500	11
9500-13500	9
13500-17500	7
17500-25000	4
25000-35000	3
35000-45000	2
45000-65000	1
65000-95000	1
95000-135000	1

Table 1: Grouped Data for Claim Severities

Appdx B

Proof of Lemma 5.5

Proof. It is true that under the condition $\{\Lambda^{(1)} + \Lambda^{(2)} = \lambda, \frac{\Lambda^{(1)}}{\Lambda^{(1)}+\Lambda^{(2)}} = \theta\}$, $S(\cdot)$ is a compound Poisson process with parameter $(\lambda, H_\theta(y))$, and for any $n \in \mathbb{N}$ and $t_k, x_k \geq 0$, we have,

$$\mathbb{P}\left(\tau_k > t_k, Y_k \leq y_k, k = 1, \ldots, n \middle| \Lambda^{(1)} + \Lambda^{(2)} = \lambda, \frac{\Lambda^{(1)}}{\Lambda^{(1)} + \Lambda^{(2)}} = \theta\right) = \prod_{k=1}^{n} e^{-\lambda t_k} H_\theta(y_k), \tag{4}$$

where τ_k denotes the inter-arrival time between the $(k-1)^{th}$ and the k^{th} claim. Then,

$$\mathbb{P}\left(Y_k \leq y_k, k = 1, \ldots, n \middle| \frac{\Lambda^{(1)}}{\Lambda^{(1)} + \Lambda^{(2)}} = \theta\right) = \prod_{k=1}^{n} H_\theta(y_k);$$

$$\mathbb{P}\left(\tau_k > t_k, k = 1, \ldots, n \middle| \Lambda^{(1)} + \Lambda^{(2)} = \lambda\right) = \prod_{k=1}^{n} e^{-\lambda t_k}.$$

The conclusion of the first assertion is straight forward. ■

Proof of Lemma 5.6

Proof. If $(\Lambda^{(1)} + \Lambda^{(2)})$ and $\left(\frac{\Lambda^{(1)}}{\Lambda^{(1)}+\Lambda^{(2)}}\right)$ are conditionally independent under $\{\Lambda^{(2)} > 0\}$. Given any $\theta \in (0,1)$, the conditional independence implies

$$\mathbb{P}\left((\Lambda^{(1)} + \Lambda^{(2)}) \in d\lambda \middle| \frac{\Lambda^{(1)}}{\Lambda^{(1)} + \Lambda^{(2)}} = \theta\right) = \mathbb{P}\left((\Lambda^{(1)} + \Lambda^{(2)}) \in d\lambda \middle| \Lambda^{(2)} > 0\right),$$

since $\left\{\frac{\Lambda^{(1)}}{\Lambda^{(1)}+\Lambda^{(2)}} = 1\right\} = \{\Lambda^{(2)} = 0\}$ and $\left\{\frac{\Lambda^{(1)}}{\Lambda^{(1)}+\Lambda^{(2)}} \in B\right\} = \left\{\frac{\Lambda^{(1)}}{\Lambda^{(1)}+\Lambda^{(2)}} \in B\right\} \cap \{\Lambda^{(2)} > 0\}$, $\forall B \in \mathcal{B}(0,1)$. Therefore, it follows from identity (4) that $\forall A \in \mathcal{B}(\mathbb{R}^+)$,

$$\begin{aligned}
&\mathbb{P}\left(\tau_k > t_k, Y_k \leq y_k, k = 1, \ldots, n, \Lambda^{(1)} + \Lambda^{(2)} \in A \middle| \frac{\Lambda^{(1)}}{\Lambda^{(1)} + \Lambda^{(2)}} = \theta\right) \\
= &\int_{\lambda \in A} \prod_{k=1}^{n} e^{-\lambda t_k} H_\theta(y_k) \mathbb{P}\left(\Lambda^{(1)} + \Lambda^{(2)} \in d\lambda \middle| \frac{\Lambda^{(1)}}{\Lambda^{(1)} + \Lambda^{(2)}} = \theta\right) \\
= &\prod_{k=1}^{n} H_\theta(y_k) \left[\int_{\lambda \in A} \prod_{k=1}^{n} e^{-\lambda t_k} \mathbb{P}\left(\Lambda^{(1)} + \Lambda^{(2)} \in d\lambda \middle| \Lambda^{(2)} > 0\right)\right] \\
= &\prod_{k=1}^{n} H_\theta(y_k) \times \mathbb{P}(\tau_k > t_k, k = 1, \ldots, n, \Lambda^{(1)} + \Lambda^{(2)} \in A | \Lambda^{(2)} > 0).
\end{aligned}$$

The identity above implies that, for every $\theta \in (0,1)$, under measure $\mathbb{P}\left(\cdot \middle| \frac{\Lambda^{(1)}}{\Lambda^{(1)} + \Lambda^{(2)}} = \theta\right)$, the claim sizes $\{Y_k, k \geq 1\}$ are i.i.d with a common distribution function $H_\theta(\cdot)$, the counting process $N(\cdot)$ is a mixed Poisson process with intensity $(\Lambda^{(1)}+\Lambda^{(2)})|_{\Lambda^{(2)}>0}$. More importantly, they are mutually independent so is the case conditioning on $\{\Lambda^{(2)} = 0\}$ with $N(\cdot)$ as a mixed Poisson process with intensity $\Lambda^{(1)}$. ■

Proof of Lemma 5.7

Proof. Basically, denoting $\gamma_1 = \Lambda^{(1)}, \gamma_2 = \Lambda^{(2)}|_{\Lambda^{(2)}>0}$, we have

$$\begin{aligned}
&\mathbb{P}\left((\gamma_1+\gamma_2)\in du, \frac{\gamma_1}{\gamma_1+\gamma_2}\in dv\right) = f_{\gamma_1}(uv)f_{\gamma_2}(u(1-v))u\,du\,dv \\
= \;& \frac{\lambda_0^\alpha\lambda_0^\beta}{\Gamma(\alpha)\Gamma(\beta)}(uv)^{\alpha-1}(u(1-v))^{\beta-1}e^{-\lambda_0 u}u\,du\,dv \\
= \;& \left(\frac{\lambda_0^{\alpha+\beta}}{\Gamma(\alpha+\beta)}u^{\alpha+\beta-1}e^{-\lambda_0 u}\,du\right)\cdot\left(\frac{1}{B(\alpha,\beta)}v^{\alpha-1}(1-v)^{\beta-1}\,dv\right),
\end{aligned}$$

where $f_{\gamma_1}(\cdot), f_{\gamma_2}(\cdot)$ are the density functions for γ_1 and γ_2 respectively, and $B(\alpha,\beta) = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}$. And the lemma is proved. ■

Proof of Theorem 5.8

Proof. If we let $\mu_F/c = \eta, \mu_G/c = \rho$, then for any fixed $\theta \in (0,1)$ such that $\theta\mu_F + (1-\theta)\mu_G < c$,

$$\mu_\theta H_{e,\theta}(y) = \int_0^y (1-\theta F(x) - (1-\theta)G(x))\,dx = \theta\mu_F F_e(y) + (1-\theta)\mu_G G_e(y),$$

where $F_e(y) = \frac{1}{\mu_F}\int_0^y (1-F(x))\,dx$, $G_e(y) = \frac{1}{\mu_G}\int_0^y (1-G(x))\,dx$. Hence,

$$\begin{aligned}
1-\psi_\theta^M(u) &= (\theta(1-\eta)+(1-\theta)(1-\rho))\sum_{n\geq 0}\left(\frac{1}{c}\right)^n(\theta\mu_F F_e(\cdot)+(1-\theta)\mu_G G_e(\cdot))^{*n}(u) \\
&= (\theta(1-\eta)+(1-\theta)(1-\rho))\sum_{n\geq 0}\sum_{0\leq l\leq n}\binom{n}{l}\theta^l(1-\theta)^{n-l}\left(\frac{\mu_F^l\mu_G^{n-l}}{c^n}\right) \\
&\quad F_e^{*l} * G_e^{*(n-l)}(u) \\
&= (1-\eta)\sum_{l\geq 0, m\geq 0}\binom{m+l}{l}\eta^l\rho^m\theta^{l+1}(1-\theta)^m\left(F_e^{*l} * G_e^{*m}\right)(u) \\
&\quad +(1-\rho)\sum_{l\geq 0, m\geq 0}\binom{m+l}{l}\eta^l\rho^m\theta^l(1-\theta)^{m+1}\left(F_e^{*l} * G_e^{*m}\right)(u).
\end{aligned}$$

Then an integration over Θ on $\{\Theta \neq 1\}$ using the probability density function of $Beta(\alpha, \beta)$ on both sides will lead to the desired result as shown in the theorem.
■

Proof of Corollary 5.9

Proof. It can be seen that for $l, m \geq 0$, we have

$$\begin{aligned}
& \binom{m+l}{l} \frac{B(l+1+\alpha, m+\beta)}{B(\alpha,\beta)} \\
= \;& \frac{\Gamma(l+1+\alpha)}{\Gamma(l+1)\,\Gamma(\alpha)} \, \frac{\Gamma(m+\beta)}{\Gamma(m+1)\,\Gamma(\beta)} \, \frac{(m+l)!\Gamma(\alpha+\beta)}{\Gamma(\alpha+\beta+m+l+1)} \\
= \;& \alpha \frac{(\alpha+l)(\alpha+l-1)\cdots(\alpha+1)}{l(l-1)\cdots 1} \cdot \frac{(\beta+m-1)(\beta+m-2)\cdots(\beta+1)\beta}{m(m-1)\cdots 1} \\
& \cdot B(\alpha+\beta, m+l+1) \\
= \;& \alpha(-1)^{l+m} \binom{-\alpha-1}{l} \binom{-\beta}{m} \int_0^1 t^{m+l}(1-t)^{\alpha+\beta-1}\,dt,
\end{aligned}$$

by adopting the property of a negative binomial distribution function where it allows for positive α, β. We further introduced notations from (5.11) through which we could write,

$$\begin{aligned}
& \sum_{l\geq 0, m\geq 0} \eta^l \rho^m \binom{m+l}{l} \frac{B(l+1+\alpha, m+\beta)}{B(\alpha,\beta)} F_e^{*l} * G_e^{*m}(u) \\
= \;& \alpha \int_0^1 (1-t)^{\alpha+\beta-1} \int_0^u F^{\alpha+1}(t, u-y) G^{\beta}(t, dy)\,dt.
\end{aligned}$$

Clearly, $F^\gamma(t,u)\,(G^\gamma(t,u))$ increases on $[0,1)\times\mathbb{R}^+$ with respect to (t,u), $F_\gamma(t,0) = 1, F^\gamma(t,\infty) = (1-t\eta)^{-\gamma}$, and $G^\gamma(t,0) = 1, G_\gamma(t,\infty) = (1-t\rho)^{-\gamma}$. Actually, taking the Laplace transform of $F^\gamma(t_0,\cdot)$ yields,

$$\int_{[0,\infty)} e^{-su} F^\gamma(t_0, du) = \sum_{l\geq 0} \binom{-\gamma}{l} (-t_0\eta)^l (\hat{F}_e(s))^l = \left(1 - t_0\eta\hat{F}_e(s)\right)^{-\gamma},$$

which demonstrates that $F^{\gamma}(t,u)$ is proportional to a cumulative distribution function of a γ-convolution of compound geometry distribution.

Similarly, we have

$$\binom{m+l}{l}\frac{B(l+\alpha,m+1+\beta)}{B(\alpha,\beta)}=\beta(-1)^{l+m}\binom{-\alpha}{l}\binom{-\beta-1}{m}\int_0^1 t^{m+l}(1-t)^{\alpha+\beta-1}\,dt.$$

These directly lead to the equations shown in Corollary 5.9. ■

Proof of Corollary 5.10

Proof. In fact, $\hat{F}_e(s)=\frac{\zeta_1}{\zeta_1+s}$, $\eta=(\zeta_1 c)^{-1}$, for $t_0\in(0,1)$,

$$\left(1-\frac{t_0\eta\zeta_1}{\zeta_1+s}\right)^{-1}=\int_0^{\infty}e^{-sy}\left(\delta_0(dy)+t_0\eta\zeta_1 e^{-\zeta_1(1-t_0\eta)y}\right)\,dy,$$

then, for any $\gamma\in\mathbb{N}$, we have

$$F^{\gamma}(t_0,dy)=\delta_0(\,dy)+\left(\sum_{l=1}^{\gamma}\binom{\gamma}{l}(t_0\eta\zeta_1)^l\frac{y^{l-1}}{\Gamma(l)}e^{-\zeta_1(1-t_0\eta)y}\right)\,dy, \tag{5}$$

where δ_0 denotes the Dirac measure centered at 0. Similarly, we have $\hat{G}_e(s)=\frac{\zeta_2}{\zeta_2+s}$, $\rho=(\zeta_2 c)^{-1}$ and

$$\int_{[0,\infty)}e^{-sy}G^{\gamma}(t_0,dy)=\sum_{l\geq 0}\binom{-\gamma}{l}(-t_0\rho)^l(\hat{G}_e(s))^l=\left(1-t_0\rho\hat{G}_e(s)\right)^{-\gamma}.$$

Hence, for any $\gamma\in\mathbb{N}$,

$$G^{\gamma}(t_0,dy)=\delta_0(\,dy)+\left(\sum_{l=1}^{\gamma}\binom{\gamma}{l}(t_0\rho\zeta_2)^l\frac{u^{l-1}}{\Gamma(l)}e^{-\zeta_2(1-t_0\rho)y}\right)\,dy.$$

Before continuing (5.12), first the following convolution is calculated,

$$
\begin{aligned}
&\int_0^u F^{\alpha+1}(t,u-y)G^{\beta}(t,y)dy \\
= \quad & 1+e^{-\zeta_1 u+\zeta_1 t\eta u}\sum_{j=1}^{\beta}\binom{\beta}{j}\frac{(t\rho\zeta_2)^j}{\Gamma(j)}\sum_{i=1}^{\alpha+1}\binom{\alpha+1}{i}\frac{(t\eta\zeta_1)^i}{\Gamma(i)} \qquad (6)\\
&\int_0^u e^{-[(\zeta_1\eta-\zeta_2\rho)t-\zeta_1+\zeta_2]y}(u-y)^{i-1}y^{j-1}dy \qquad (7)\\
= \quad & 1+e^{-\zeta_1 u+\zeta_1 t\eta u}\sum_{j=1}^{\beta}\binom{\beta}{j}(t\rho\zeta_2)^j\sum_{i=1}^{\alpha+1}\binom{\alpha+1}{i}(t\eta\zeta_1)^i\frac{u^{i+j-1}}{\Gamma(i+j)} \qquad (8)\\
&{}_1F_1(i,i+j,-[(\zeta_1\eta-\zeta_2\rho)t-\zeta_1+\zeta_2]u),
\end{aligned}
$$

where 1 results from an integration of the product of two Dirac measures, $1 = \int_0^u \delta_0^2(dy)$, and ${}_1F_1(\cdot)$ is a hyper-geometric function with order 1,1 whose definition is given as follows.

$$
{}_1F_1(a;b;z)=\sum_{k=0}^{\infty}\frac{(a)_k}{(b)_k}\frac{z^k}{k!},
$$

where $(c)_k = c(c+1)\ldots(c+k-1)$ with $(c)_0 = 1$. In fact, it relates to a moment generating function of a Beta distributed random variable X with parameters i, j, i.e., $X \sim Beta(i,j)$.

$$
M_X(-[(\zeta_1\eta-\zeta_2\rho)t-\zeta_1+\zeta_2]u) = {}_1F_1(i,i+j,-[(\zeta_1\eta-\zeta_2\rho)t-\zeta_1+\zeta_2]u),
$$

which could be seen from the nature of the integral in (6). Thus, (5.12) could be written as,

$$
\begin{aligned}
\psi^M(u)|_{\Theta\neq 1} \;=\;& 1-\alpha(1-\eta)\left[\frac{1}{\alpha+\beta}+e^{-\zeta_1 u}\sum_{j=1}^{\beta}\binom{\beta}{j}(\rho\zeta_2)^j\sum_{i=1}^{\alpha+1}\binom{\alpha+1}{i}(\eta\zeta_1)^i\frac{u^{i+j-1}}{\Gamma(i+j)}\right.\\
&\left.\times\int_0^1(1-t)^{\alpha+\beta-1}t^{i+j}e^{\zeta_1 t\eta u}M_{X(i,j)}(-[(\zeta_1\eta-\zeta_2\rho)t-\zeta_1+\zeta_2]u)\,dt\right]\\
&-\beta(1-\rho)\left[\frac{1}{\alpha+\beta}+e^{-\zeta_1 u}\sum_{j=1}^{\beta+1}\binom{\beta+1}{j}(\rho\zeta_2)^j\sum_{i=1}^{\alpha}\binom{\alpha}{i}(\eta\zeta_1)^i\frac{u^{i+j-1}}{\Gamma(i+j)}\right.\\
&\left.\times\int_0^1(1-t)^{\alpha+\beta-1}t^{i+j}e^{\zeta_1 t\eta u}M_{X(i,j)}(-[(\zeta_1\eta-\zeta_2\rho)t-\zeta_1+\zeta_2]u)\,dt\right]. \quad (9)
\end{aligned}
$$

■

References

M. Abramowitz and I. A. Stegun. *Handbook of mathematical functions with formulas, graphs and mathematical tables.* New York: Dover., 1964. 35

Lourdes B Afonso, Alfredo D Egídio dos Reis, and Howard R Waters. Calculating continuous time ruin probabilities for a large portfolio with varying premiums. *ASTIN Bulletin*, 39(01):117–136, 2009. 23, 119

Lourdes B Afonso, Cardoso Rui, Alfredo D Egídio dos Reis, and Gracinda Guerreiro. Bonus malus systems and finite and continuous time ruin probabilities in motor insurance. *Preprint*, 2015. 23

H. Albrecher, C. Constantinescu, and S. Loisel. Explicit ruin formulas for models with dependence among risks. *Insurance: Mathematics and Economics*, 48(2): 265–270, 2011. ISSN 01676687. 19, 20, 34, 120

Hansjörg Albrecher and Onno J Boxma. On the discounted penalty function in a markov-dependent risk model. *Insurance: Mathematics and Economics*, 37 (3):650–672, 2005. 69

Hans Ammeter. A generalization of the collective theory of risk in regard to fluctuating basic-probabilities. *Scandinavian Actuarial Journal*, 1948(1-2):171–198, 1948. 24

Søren Asmussen and Hansjörg Albrecher. *Ruin probabilities / Søren Asmussen and Hansjörg Albrecher.* Advanced series on statistical science & applied probability: 14. Singapore ; World Scientific, 2010., 2010. 1, 2, 12, 14, 24, 31, 85, 100, 103

P. J. Boland. *Statistical and probabilistic methods in actuarial science.* Boca Raton, FL ; Chapman & Hall/CRC., 2007. ISBN 1584886951. 33

R. Brooks, B. Gravelsons, and G. Macra. Update from uk asbestos and deafness working parties. 2013. 26

H Bühlmann. Ruinwahrscheinlichkeit bei erfahrungstarifiertem portefeuille. *Mitteilungen der Vereinigung Schweizerischer Versicherungsmathematiker*, 72: 211–224, 1972. 24

H. Bühlmann. The history of ASTIN. *ASTIN Bulletin*, 37(2):191–202, 2007. 24

C. Constantinescu, G. Samorodnitsky, and W. Zhu. Risk models with gamma claims. *Preprint*, 2015a. 78

Corina Constantinescu and Joseph Lo. Ruin theory starter kit#. 2013. 81

Corina Constantinescu, Véronique Maume-Deschamps, and Ragnar Norberg. Risk processes with dependence and premium adjusted to solvency targets. *European Actuarial Journal*, 2(1):1–20, 2012. 24

Corina Constantinescu, Suhang Dai, Weihong Ni, and Zbigniew Palmowski. Ruin probabilities with dependence on the number of claims within fixed window time. *Working paper*, 2015b. 31, 84, 118

Jos HA De Smit. The queue gi/m/s with customers of different types or the queue gi/hm/s. *Advances in Applied Probability*, pages 392–419, 1983. 70, 71

Georges Dionne and Olfa Ghali. The (1992) bonus-malus system in tunisia: An empirical evaluation. *Journal of Risk and Insurance*, 72(4):609–633, 2005. 17

Georges Dionne, Christian Gouriéroux, and Charles Vanasse. *Evidence of adverse selection in automobile insurance markets.* Springer, 1999. 17

André Dubey. Probabilité de ruine lorsque le paramètre de poisson est ajusté a posteriori. *Mitteilungen der Vereinigung schweiz Versicherungsmathematiker*, 2:130–141, 1977. 24, 25, 27, 28, 29, 73, 75, 76, 77, 80, 81, 114

François Dufresne. Distributions stationnaires d'un système bonus-malus et probabilité de ruine. *ASTIN Bulletin*, 18(01):31–46, 1988. 22

P. Embrechts. A property of the generalized inverse gaussian distribution with some applications. *Journal of Applied Probability*, (3):537, 1983. ISSN 00219002. 36

Thomas Erneux. *Applied delay differential equations.* Surveys and tutorials in the applied mathematical sciences: 3. New York ; Springer, 2008., 2008. 108

Runhuan Feng. *A generalization of the discounted penalty function in ruin theory.* PhD thesis, University of Waterloo, 2008. 111

Nicholas E Frangos and Spyridon D Vrontos. Design of optimal bonus-malus systems with a frequency and a severity component on an individual basis in automobile insurance. *ASTIN Bulletin*, 31(01):1–22, 2001. 18, 19, 32, 33, 34, 37, 38, 39, 43, 52, 59

Ulf Grenander. Some remarks on bonus systems in automobile insurance. *Scandinavian Actuarial Journal*, 1957(3-4):180–197, 1957. 18

M. E. H. Ismail and M. E. Muldoon. Monotonicity of the zeros of a cross-product of Bessel functions. *SIAM Journal on Mathematical Analysis*, 9(4):759–767, 1978. 38

Helena Jasiulewicz. Probability of ruin with variable premium rate in a markovian environment. *Insurance: Mathematics and Economics*, 29(2):291–296, 2001. 24

S. A. Klugman, H. H. Panjer, and G. E. Willmot. *Loss models : from data to decisions.* John Wiley & Sons, 1998. 1, 39, 53

J. Lemaire. *Bonus-malus systems in automobile insurance.* Boston, MA Kluwer Academic Publishers., 1995. ISBN 079239545X. 18, 19, 36

Bo Li, Weihong Ni, and Corina Constantinescu. Risk models with premiums adjusted to claims number. *Insurance: Mathematics and Economics*, 65:94–102, 2015. 31, 74, 118

Shuanming Li and Jose Garrido. On ruin for the erlang (n) risk process. *Insurance: Mathematics and Economics*, 34(3):391–408, 2004. 106, 110

L. Lorch. Monotonicity of the zeros of a cross product of Bessel functions. *Methods and Applications of Analysis*, 1(1):75–80, 1994. 38

Eugene Lukacs. A characterization of the gamma distribution. *The Annals of Mathematical Statistics*, 26:319–324, 1955. 80

Ove Lundberg. *On Random Processes and Their Application to Sickness and Accident Statistics.* University of Stockholm Thesis, Uppsala., 1948. 24

Ignacio Moreno, Francisco J Vázquez, and Richard Watt. Can bonus-malus allieviate insurance fraud? *Journal of Risk and Insurance*, 73(1):123–151, 2006. 17

W. Ni, C. Constantinescu, and A. Pantelous. Bonus-malus systems with weibull distributed claim severities. *Annals of Actuarial Science*, 8(2):217–233, 2014a. ISSN 1748-4995. 25, 27, 31, 32, 50

Weihong Ni, Bo Li, Corina Constantinescu, and Athanasios A Pantelous. Bonus-malus systems with hybrid claim severity distributions. *Vulnerability, Uncertainty, and Risk: Quantification, Mitigation, and Management*, pages 1234–1244, 2014b. 31, 32

Ragnar Norberg. Credibility premium plans which make allowance for bonus hunger. *Scandinavian Actuarial Journal*, 1975(2):73–86, 1975. 18

Z. Palmowski and B. Zwart. On perturbed random walks. *J. Appl. Probab.*, 2010. 30, 86

Zbigniew Palmowski and Bert Zwart. Tail asymptotics of the supremum of a regenerative process. *Journal of applied probability*, pages 349–365, 2007. 30, 31, 84, 86, 87, 88

Jostein Paulsen and Knut Stubø. On maximum likelihood and pseudo-maximum likelihood estimation in compound insurance models with deductibles. *ASTIN Bulletin*, 41(01):1–28, 2011. 119

P. Picard. Généralisation de l'etude sur la survenance des sinistres en assurance automobile. *Bulletin Trimestriel de l'Institut des Actuaires Français*, 296:204–268, 1976. 18, 19

J. Pritchard. Google acknowledges 11 accidents with its self-driving cars. http://bigstory.ap.org/article/297ef1bfb75847de95d856fb08dc0687/ap-exclusive-self-driving-cars-getting-dinged-california. Accessed: 2015-10-29. 26

Tomasz Rolski, Hanspeter Schmidli, Volker Schmidt, and Jozef Teugels. *Stochastic processes for insurance and finance*, volume 505. John Wiley & Sons, 2009. 1, 13, 14, 16, 25

Erwin Straub. *Zur Theorie der Prämienstufen-Systeme*. PhD thesis, Diss. Math. ETH Zürich, Nr. 4250, 0000. Ref.: Bühlmann, H.; Korref.: Weinberg, F., 1968. 18

Gregory Temnov. Risk process with random income. *Journal of Mathematical Sciences*, 123(1):3780–3794, 2004. 22

Sebastian Thrun. What we're driving at. *The Official Google Blog*, 2010. 26

L. Tremblay. Using the poisson inverse gaussian in bonus-malus systems. *ASTIN Bulletin*, 22(1):97–106, 1992. 35

E. A. Valdez and E. W. Frees. Longitudinal modeling of singapore motor insurance. Working Paper 2005. 19

Christian Wagner. Time in the red in a two state markov model. *Insurance: Mathematics and economics*, 31(3):365–372, 2002. 22

Rongming Wang, Lin Xu, and Dingjun Yao. Ruin problems with stochastic premium stochastic return on investments. *Frontiers of Mathematics in China*, 2(3):467–490, 2007. 23

Xueyuan Wu, Mi Chen, and Junyi Guo. On a discrete-time risk model with claim correlated premiums. 22