A REVIEW OF RUIN PROBABILITY MODELS

by

Brian James Fullilove

(Under the direction of William P. McCormick)

Abstract

The probability that an insurance company can go bankrupt is a crucial quantity to be able to calculate. There are many ways to calculate such a probability. For example, we could model the arrival of the claims with a Poisson process. Alternatively, we could use a random walk in order to model the effects that claims have on an insurance company's surplus. The distribution of the claim sizes also could have an effect on the model. An additional model can use random walks with dependent steps in the form of a time series. This paper seeks to introduce several of the available models and contains the results of a simulation of one of Veraverbeke's (1977) results.

Index words: Ruin Probability, Collective Risk Theory, Random Walk, Compound Poisson Model

A REVIEW OF RUIN PROBABILITY MODELS

by

Brian James Fullilove

B.S., Morehouse College, 2006

A Thesis Submitted to the Graduate Faculty

of The University of Georgia in Partial Fulfillment

of the

Requirements for the Degree

Master of Science

Athens, Georgia

2009

A REVIEW OF RUIN PROBABILITY MODELS

by

Brian James Fullilove

Approved:

Major Professor: William P. McCormick

Committee: Lynne Seymour
T.N. Sriram

Electronic Version Approved:

Maureen Grasso
Dean of the Graduate School
The University of Georgia
December 2009

Acknowledgments

I would like to first and foremost thank my major professor, Dr. William P. McCormick, whose guidance and patience has proved invaluable.

Also, I would like to thank Dr. Lynne Seymour and Dr. T. N. Sriram for serving on my committee and for being excellent mentors. A special thanks goes to Ana Moura Bargo for her help with LaTeX. Moreover, I would like to thank the University of Georgia Department of Statistics as a whole, which has provided me with the tools and the opportunity to further my education. Lastly, I would like to thank my friends and family for their constant support, love, and understanding, even in the darkest of hours.

Table of Contents

Chapter 1

Introduction

Ruin Probability, also known as collective risk theory, is a branch of actuarial science that studies an insurer's vulnerability to insolvency based on mathematical modeling of the insurer's surplus.(Asmussen (2000)) There are several ways to model this. The classical model is the Compound Poisson model, which is named this way because the arrivals of insurance claims are assumed to follow a Poisson process with rate λ. In this model the claim sizes are $X_1, X_2, ...$ and are i.i.d. with some common distribution F. We will also, without loss of generality, be able to assume that the insurance premium, the amount that the insured pays for insurance, is $p = 1$.(Asmussen (2000))

A random walk based model is based on the number of claims that arrive, not the arrival rates of the claims. This model is named after the random walk S_n, which tracks the movement of an insurance company's surplus (claims vs. premiums). Since insurance companies will set rates in a manner that will generate profits, one could model the flow of insurance claims and premiums as a random walk with positive drift. However, for mathematical purposes we will view the flow of claims vs. premiums as a random walk with negative drift. This will be done in spite of a more intuitive model which would be "the other way around"-i.e. a model with positive drift, where premiums are positive and claims are negative. We will concentrate on the maximum of such a walk, which is in turn the minimum of the intuitive model. We will also present a simulation of the asymptotic results of this model.

The third ruin probability model that will be covered here is actually a generalization of the above random walk model. This model will remove the assumption that the steps of the random walk S_n are independent. Instead, we will assume the steps are autocorrelated and the basic structure of the model will be time-series based. Specifically, a two-sided linear process will be used to model the reserve.

Chapter 2

The Risk Reserve Process

Here we will introduce the risk reserve process, and provide some of the symbols that will be used throughout the paper, particularly in the Compound Poisson Model. Most of the material below comes from Asmussen (2000).

2.1 Risk reserve basics

A risk reserve process, R_t where $t \geq 0$, in general is a model for the progress of the reserves of an insurance company. The initial reserve will be denoted R_0. We will use $\psi(R_0)$ to denote the probability of ultimate ruin, which is the probability that an insurance company's reserve drops below zero. In other words, $\psi(R_0) = \mathbb{P}(\inf_{t\geq 0} R_t < 0)$. We can also discuss $\psi(R_0, T) = \mathbb{P}(\inf\limits_{0\leq t<T} R_t < 0)$, which is the probability of ultimate ruin before time T. Commonly, $\psi(R_0)$ and $\psi(R_0, T)$ are referred to as infinite and finite horizon ruin probabilities, respectively. While counter intuitive, it is mathematically easier to work with the Claim Surplus Process, denoted $S_t = R_0 - R_t$. With this notation, ruin occurs whenever $S_t > R_0$. Thus, ultimate ruin may now be written as $\psi(R_0) = \mathbb{P}(\sup\limits_{t\geq 0} S_t > R_0)$. Notice that we have switched our focus from an infimum of the R_t process to a supremum of the S_t process; this change will simplify some of the analysis. Now define, $\tau(R_0)$, which is the time until ultimate ruin, which can be expressed in a couple of different ways:

$$\tau(R_0) = \inf_{t\geq 0}(t : S_t > R_0) = \inf_{t\geq 0}(t : R_t < 0).$$

Also, we define $M = \sup\limits_{0\leq t<\infty} S_t$ and $M_T = \sup\limits_{0\leq t<T} S_t$, which serve as "maxima" for the Claim Surplus Process. The addition of M and $\tau(R_0)$ allows us to redefine $\psi(R_0)$ and $\psi(R_0, T)$ in the following, more succinct, manner.

$\psi(R_0) = \mathbb{P}[\tau(R_0) < \infty] = \mathbb{P}(M > R_0),$

$\psi(R_0, T) = \mathbb{P}(M_T > R_0) = \mathbb{P}[\tau(R_0) \leq T].$

We will also impose some restrictions on the risk reserve process in the next section. (Assmussen(2000))

2.2 Some assumptions on the risk reserve model

With probability one, there are only finitely many claims during finite time intervals: Let $N_t < \infty$ denote the number of arrivals from time 0 to time t.

We denote the inter-arrival times of claims by $\{T_1, T_2, ...\}$. The arrival of the nth claim will be denoted as $\sigma_n = \sum_{i=1}^{n} T_i$, and $N_t = \max\{n : \sigma_n \leq t\}$.

The size of the ith claim is denoted by X_i.

Premiums flow in at rate p, per unit time.

Therefore: $R_t = R_0 + pt - \sum_{i=1}^{N_t} X_i$ and $S_t = \sum_{i=1}^{N_t} X_i - pt$.

2.3 A couple of constants

Risk process models typically have the property that there exists a constant ρ such that $\frac{1}{t}\sum_{i=1}^{N_t} X_i \stackrel{a.s.}{\to} \rho$, as $t \to \infty$. The interpretation of ρ is as the average amount of claim per unit time. Another important quantity is the safety loading constant, denoted η and defined by the relative amount by which the premium rate p exceeds ρ. In other words, $\eta = \frac{p-\rho}{\rho}$. Any insurance company will try to ensure that $\eta > 0$; that is, an insurance company will set premiums so that the expected premium rate exceeds the expected claim rate. Typically, according to the theoretical literature, η is fairly small, falling within 10%-20%.

2.4 Some quick results

Below are some results that relate the above constants and other quantities of interest to each other.

Theorem 2.4.1. *Assume* $\frac{1}{t}\sum_{k=1}^{N_t} X_i \to \rho$ *a.s. , as* $t \to \infty$ *. If* $\eta < 0$ *, then* $M = \infty$ *a.s. and hence* $\psi(R_0) = 1$ *for all* R_0*. If* $\eta > 0$ *, then* $M < \infty$ *a.s. and hence* $\psi(R_0) < 1$ *for all sufficiently large* R_0*.*

Proof. Recall $S_t = \sum_{i=1}^{N_t} X_i - pt$.

Therefore, $\frac{S_t}{t} = \frac{\sum_{i=1}^{N_t} X_i - pt}{t}$.

Alternatively we can write, $\frac{S_t}{t} = \frac{1}{t}\sum_{i=1}^{N_t} X_i - p$.

Using $\frac{1}{t}\sum_{i=1}^{N_t} X_i \stackrel{a.s.}{\to} \rho$, as $t \to \infty$, it follows that $\frac{S_t}{t} \to \rho - p$, as $t \to \infty$.

Recall $\eta = \frac{p-\rho}{\rho}$.

If $\eta < 0$, then it must be true that $\rho > p$.

Therefore, $\frac{S_t}{t}$ approaches a positive number as $t \to \infty$.

Thus $S_t \to \infty$, and $M = \sup\limits_{0 \le t < \infty} S_t = \infty$.

Similarly, if $\eta > 0$, it must be true that $p > \rho$.

Thus, $\frac{S_t}{t}$ approaches a negative number as $t \to \infty$.

Therefore $S_t \to -\infty$, and $M < \infty$.

□

In order to simplify things, we would like to be able to relate ruin probability models where $p = 1$ to models where $p \neq 1$. The following theorem does just that.

Theorem 2.4.2. *Assume* $p \neq 1$ *and define* $R_t^{'} = R_{t/p}$ *. Then the connection between ruin probabilities for the given risk process* R_t*and those* $\psi^{'}(R_0)$, $\psi^{'}(R_0, T)$ *for* $R_t^{'}$ *is given by* $\psi(R_0) = \psi^{'}(R_0)$ *and* $\psi(R_0, T) = \psi^{'}(R_0, Tp)$*.*

Proof. We have that $\psi(R_0) = \mathbb{P}(\inf_{t \ge 0} R_t < 0)$

$= \mathbb{P}(\inf_{\frac{t}{p} \ge 0} R_{\frac{t}{p}} < 0)$, by simply replacing t with $\frac{t}{p}$.

$= \mathbb{P}(\inf_{\frac{t}{p} \ge 0} R_t^{'} < 0)$ using the definition of $R_t^{'}$.

$= \mathbb{P}(\inf_{t \ge 0} R_t^{'} < 0) = \psi^{'}(R_0)$.

Note: $R_0 = R_0^{'}$.

For the second case observe,

$$\psi(R_0, T) = \mathbb{P}(\inf_{0\leq t\leq T} R_t < 0)$$
$$= \mathbb{P}(\inf_{0\leq \frac{t}{p}\leq T} R_{\frac{t}{p}} < 0)$$
$$= \mathbb{P}(\inf_{0\leq t\leq Tp} R_{\frac{t}{p}} < 0)$$
$$= \mathbb{P}(\inf_{0\leq t\leq Tp} R_t^{'} < 0) = \psi^{'}(R_0, Tp).$$

□

For the reserve process $R_t^{'}$, the premium collected in the interval $[0,t]$ equals the premium collected for the process $R_{\frac{t}{p}}$ in the interval $[0,t]$, but since R has premium rate p , this amount is $(\frac{t}{p})p = t$. Thus, we have that $R^{'}$ has premium rate equal to 1. This result shows that without loss of generality, we may take the premium $p = 1$. We will make use of this in Chapter 4.

Chapter 3

Claim size distributions

Let F be the claim size distribution. There are essentially two classes of distributions that are possible: light tailed and heavy-tailed distributions (the differences between them will be explained below). The subexponential class of distributions will be introduced. Also, the Dominated Convergence Theorem will be presented, as it is needed later in the paper.

3.1 Light-Tailed Distributions

A light-tailed distribution is a distribution, $F(x)$, such that its tail, $\bar{F}(x) = 1 - F(x)$, satisfies $\bar{F}(x) = O(e^{-sx})$, for some $s > 0$. The asymptotic notation $f(x) = O(g(x))$ means $\limsup_{x\to\infty} |\frac{f(x)}{g(x)}| < \infty$. Some examples of light-tailed distributions are described below.

3.1.1 The Exponential Distribution

Consider the exponential density function

$f(x) = \lambda e^{-\lambda x}$.

The parameter λ is referred to as the rate of the function. The exponential distribution is a key part of the compound Poisson model, which will be highlighted in Chapter 4. One of the exponential functions most crucial features is the "memoryless" property:

A random variable X is said to be *memoryless*, if $\mathbb{P}(X > s + t | X > t) = \mathbb{P}(X > s)$, $\forall s, t \geq 0$. (Ross(2007))

This is equivalently defined as

$\frac{\mathbb{P}(X>s+t, X>t)}{\mathbb{P}(X>t)} = \mathbb{P}(X > s)$

or $\mathbb{P}(X > s + t) = \mathbb{P}(X > s)\mathbb{P}(X > t)$.

For the exponential density function the memoryless property is verified below:

$\mathbb{P}(X > s+t) = e^{-\lambda(s+t)}$

$= e^{-\lambda s}e^{-\lambda t} = \mathbb{P}(X > s)\mathbb{P}(X > t)$.

Also it should be noted that the exponential function has mean $\mathbb{E}X = \frac{1}{\lambda}$ and variance $\mathbb{V}X = \frac{1}{\lambda^2}$.

3.1.2 The Gamma Distribution

Consider the gamma density function

$f(x) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x}$.

It has moment generating function $\hat{F}(s) = (\frac{\beta}{\beta - s})^\alpha$, $s < \beta$ (note that $\hat{F}(s) = \mathbb{E}(e^{sX})$, where $X \sim f$, i.e. $\hat{F}(s)$ is the m.g.f of f).

Which has mean, $\mathbb{E}X = \frac{\alpha}{\beta}$. The variance , $\mathbb{V}X = \frac{\alpha}{\beta^2}$.

The tail is $\bar{F}(x) = \frac{\Gamma(\beta x;\alpha)}{\Gamma(\alpha)}$, where $\Gamma(x;\alpha) = \int_x^\infty t^{\alpha-1} e^{-t} dt$.

Note when $\alpha = 1$ and $\beta = \lambda$ the gamma density function is the exponential density function.

3.2 Heavy-tailed Distributions

A function F is heavy-tailed if and only if its m.g.f $\hat{F}(s) = \infty$ for all $s > 0$. There are also different (and more strict) definitions, some of which will be presented in this section.

3.2.1 The Weibull distribution

$f(x) = crx^{r-1}e^{-cx^r}$ is the density function. $\bar{F} = e^{-cx^r}$ is the distribution's tail. This function is heavy tailed as long as $0 < r < 1$.

3.2.2 The lognormal distribution

This is the distribution of e^Y, where $Y \sim N(\mu, \sigma^2)$. Therefore its density is as follows:

$f(x) = \frac{1}{x\sigma\sqrt{2\pi}} \exp\{-\frac{1}{2}[\frac{log(x)-\mu}{\sigma}]^2\}$.

Asymptotically the tail is:

$\bar{F}(x) = \frac{\sigma}{log(x)\sqrt{2\pi}} \exp\{-\frac{1}{2}[\frac{log(x)-\mu}{\sigma}]^2\}$.

The mean of the lognormal distribution is $\mathbb{E}X = e^{\mu + \frac{\sigma}{2}^2}$.

3.2.3 The Pareto Distribution

The Pareto distribution will be of great use to us in Chapter 5. Its tail posseses properties that will be useful in illustrating Veraverbeke's theorem.

We will make use of one of its simpler forms, which is $f(x) = \frac{\alpha\beta^\alpha}{x^{\alpha+1}}$, $x \geq \beta$. Thus, it's tail is $\bar{F} = (\frac{\beta}{x})^\alpha$.

It should be noted that the Pareto distribution has a regularly varying tail (see Chapter 5.2) and is also subexponential (see Chapter 3.4).

3.2.4 The Loggamma distribution

The loggamma distribution is the distribution of e^Y, where $Y \sim Gamma(\alpha, \beta)$. Its density is as follows:

$f(x) = \frac{\beta^\alpha [log(x)]^{\alpha-1}}{x^{\beta+1}\Gamma(\alpha)}$. For $\alpha = 1$, the loggamma distribution is a Pareto distribution.

3.3 Regularly Varying Tails

The tail of a distribution F, is said to be regularly varying with exponent α under the following conditions:

$\bar{F}(x) \sim \frac{L(x)}{x^\alpha}$, $x \to \infty$, where $L(x)$ is slowly varying, i.e. $L(x)$ satisfies $\frac{L(xt)}{L(x)} \to 1$, for any $t > 0$, as $x \to \infty$.

3.4 Subexponential Distributions

A distribution F is subexponential if:

$\lim\limits_{x\to\infty} \frac{\bar{F}^{*2}(x)}{\bar{F}(x)} = 2$, where $\bar{F}^{*2}(x)$, is the convolution square of $\bar{F}$. The convolution square is the distribution of the sum of independent random variables $X_1, X_2 \sim F$. In other words,

$F^{*2}(x) = \int_{-\infty}^{\infty} F(x-y)\,dF(y)$. Thus $\bar{F}^{*2}(x) = 1 - F^{*2}(x) = \int_{-\infty}^{\infty} \bar{F}(x-y)dF(y)$. If F is the distribution function of a nonnegative random variable such that $F(0^-) = 0$, then $F^{*2}(x) = \int_0^x F(x-y)\,dF(y)$ and $\bar{F}^{*2}(x) = \int_0^\infty \bar{F}(x-y)\,dF(y)$. In terms of the above definition, this means $\mathbb{P}(X_1 + X_2 > x) \sim 2\mathbb{P}(X_1 > x)$.

Equivalently, a distribution is subexponential if:

$\bar{F}^{*n}(x) \sim n\bar{F}(x)$, for every $n \geq 1$. $\bar{F}^{*n}$ is the tail of the nth convolution power of F. F^{*n} is the distribution of the sum of n i.i.d. random variables with distribution F.

Examples of subexponential distributions include F such as:

$\overline{F}(x) \sim x^{-\alpha}$, $\alpha > 0$;

$\overline{F}(x) \sim \exp(-x^{\beta}), 0 < \beta < 1$;

$\overline{F}(x) \sim \exp(-\frac{x}{(\log x)^2})$.

Chapter 4

The Compound Poisson Model

Here we will discuss the compound Poisson model.

4.1 Introduction

The compound Poisson model keeps the assumptions from Chapter 2 with the following additions:

N_t, the number of arrivals in the interval $[0, t]$, is now much less general. N_t is now a Poisson Process with rate λ. This specification of N_t implies (Ross(2007)):

1. $N_0 = 0$.

2. The inter-arrival times T_i are now distributed exponentially with rate λ (which is the same as exponential with mean $\frac{1}{\lambda}$).

3. The claim sizes X_i , $\forall i \in \mathbb{N}$, are i.i.d. with common distribution function F, which is independent of N_t.

4. The number of events in any interval of length t is Poisson distributed with mean λt . That is, for all $s, t \geq 0$, $\mathbb{P}[N_{t+s} - N_s = n] = e^{-\lambda t}\frac{(\lambda t)^n}{n!}$, $n \in \{0, 1, ...\}$.

For the sake of simplicity, we will assume that the premium rate is $p = 1$. We are allowed to do this based on a result from Chapter 2. Therefore, $R_t = R_0 + t - \sum_{i=1}^{N_t} X_i$ and $S_t = R_0 - R_t = \sum_{i=1}^{N_t} X_i - t$.

We are looking to find the basic formulas for the moments of $S_t = R_0 - R_t$. We first need the nth moment of x, which we will denote $\mu^{(n)} = \mathbb{E}X^n$. For example, $\mu = \mu^{(1)} = \mathbb{E}X$. Also, recall that $\eta = \frac{p-\rho}{\rho}$. Additionally, in the Compound Poisson Model, $p = 1$. Therefore, by solving for ρ, we find $\rho = \frac{1}{1+\eta}$.

It should also be noted here that $\rho = \lambda\mu$. The reasoning behind this is relatively simple. The quantity, ρ represents the average amount of claim per unit of time. The average claim per unit of time is obtained by multiplying the arrival rate by the expected claim size. The arrival rate is λ and the expected claim size is μ.

This equation, of course, can be obtained by more theoretical means as well. It will be presented as a theorem in the next section.

4.2 Some basic Compound Poisson results

Here we will highlight some basic results in regards to S_t, the claim surplus. The mean, variance and the moments of S_t are below.

Theorem 4.2.1. $\rho = \lambda\mu$.

Proof. By its definition let $\rho = \lim_{t\to\infty} \frac{1}{t} \sum_{i=1}^{N_t} X_i$.

$= \mathbb{E}(X_1) \lim_{t\to\infty} \mathbb{E}(\frac{N_t}{t})$, which is equivalent to...

$= \mathbb{E}(X_1)\lambda$, since λ is the arrival rate.

$= \mu\lambda$

□

Theorem 4.2.2. $\mathbb{E}S_t = t(\rho - 1)$

Proof. $\mathbb{E}S_t = \mathbb{E}(\sum_{i=1}^{N_t} X_i - t)$.

$= \mathbb{E}[\mathbb{E}(\sum_{i=1}^{N_t} X_i - t \mid N_t)]$.

$= \mathbb{E}(N_t\mu) - t$, since t is constant, $\mu = \mathbb{E}X$, and N_t is the number of X_i's.

$= \lambda t\mu - t$, since $\mathbb{E}N_t = \lambda t$.

Also, N_t and the X_i's are independent of each other.

$= t(\rho - 1)$, since $\rho = \lambda\mu$.

□

Theorem 4.2.3. *The variance of* S_t *is* $\mathbb{V}S_t = t\lambda\mu^{(2)}$

Proof. $\mathbb{V}S_t = \mathbb{V}(\sum_{i=1}^{N_t} X_i - t)$.

$= \mathbb{V}(\sum_{i=1}^{N_t} X_i)$, since $\mathbb{V}t = 0$.

$= \mathbb{V}[\mathbb{E}(\sum_{i=1}^{N_t} X_i \mid N_t)] + \mathbb{E}[\mathbb{V}(\sum_{i=1}^{N_t} X_i \mid N_t)]$, a common result from any calculus based statistics course.

$= \mathbb{V}(N_t\mu) + \mathbb{E}[N_t\mathbb{V}(X)]$.

$= \lambda t\mu^2 + \lambda t\mathbb{V}(X)$.

$= \lambda t\mu^{(2)}$.

□

Theorem 4.2.4. *The moment generating function of* S_t*,* $\mathbb{E}e^{sS_t} = e^{t\kappa(s)}$*,where* $\kappa(s) = \lambda(\hat{F}[s] - 1) - s$.

Proof. $\mathbb{E}e^{sS_t} = \mathbb{E}e^{s(\sum_{i=1}^{N_t} X_i - t)}$

$= e^{-st}\sum_{k=0}^{\infty}\mathbb{E}e^{s(X_1+\ldots+X_k)}\mathbb{P}(N_t = k)$

$= e^{-st}\sum_{k=0}^{\infty}\hat{F}[s]^k e^{-\lambda t}\frac{(\lambda t)^k}{k!}$

$= e^{-st-\lambda t+\hat{F}[s]\lambda t}$

$= e^{t\kappa(s)}$.

□

4.3 Ladder Heights

In order to discuss the Pollaczeck-Khinchine formula, which gives an expression of the ruin probability for the Compound Poisson model, we must first discuss the ladder height distribution.

Consider the claim process S_t of a general risk process and the time $\tau(R_0) = \inf_{t\geq 0}(t : S_t > R_0)$. We will assume $R_0 = 0$. Let $\tau_+ = \tau(0)$ and the associated ladder height be S_{τ_+} . Then the ladder height distribution is defined as

$G_+(x) = \mathbb{P}(S_{\tau_+} \leq x) = \mathbb{P}(S_{\tau_+} \leq x,\, \tau_+ < \infty)$.

Note that G_+ has no mass on $(-\infty, 0]$, and is typically defective; that is,

$||G_+|| = G_+(\infty) = \mathbb{P}(\tau_+ < \infty) = \psi(0) < 1$ when $\eta > 0$. Note that when $\eta > 0$, there is positive probability that S_t will never rise above 0.

Let $M_T = \sup_{0 \le t < T} S_t = S_{\tau_+(K_T)}$, where K_T denotes the last ladder height to be defined on the interval $[0, T)$: $K_T = \max\{k : \tau_+(k) < T\}$. Now, define M_t as the process of relative maxima. The term ladder height is motivated by the shape of this process. The first ladder step is $S_{\tau_+} = S_{\tau_+(1)}$, the second step is $S_{\tau_+(2)} - S_{\tau_+(1)}$, where $\tau_+(1)$ and $\tau_+(2)$ are the times of the first and second relative maxima, respectively. If $\eta > 0$, then there are only finitely many ladder steps. Thus $M = \sup_{0 \le t < \infty} S_{\tau_+(K)}$, where K is the last ladder height to be defined: $K = \max\{k : \tau_+(k) < \infty\}$. Therefore, M is the total height of the ladder, i.e. the sum of all of the ladder steps. Note that since the S_t process has negative drift ($\eta > 0$), there will be, with probability 1, a finite value of K. In regards to the compound Poisson model, the following result from Asmussen (2000) holds. A proof will not be provided, as it would require tools that are beyond the scope of this text.

For the compound Poisson model with $\rho = \lambda\mu < 1$, G_+ is given by the defective density $g_+(x) = \lambda\bar{F}(x) = \rho b_0(x)$, where $b_0(x) = \frac{\bar{F}(x)}{\mu}$, for $x > 0$.

4.4 The Pollaczeck-Khinchine formula

In order to present this formula, we will exploit the fact that M is the sum of ladder heights. Assume that $\eta > 0$. Equivalently, we assume $\rho < 1$. In the compound Poisson model, the ladder heights are i.i.d. This follows from noting that the process repeats itself after reaching a relative maximum. Decomposing M as the sum of ladder heights yields the following formula:

Theorem 4.4.1. *The distribution of M is $(1 - ||G_+||)\sum_{n=0}^{\infty} G_+^{*n}$, where G_+ is given by the defective density $g_+(x) = \lambda\bar{F}(x) = \rho b_0(x)$ on $(0, \infty)$, where $b_0(x) = \frac{\bar{F}(x)}{\mu}$.*

Proof. The probability that M is attained in precisely n ladder steps and does not exceed x is $G_+^{*n}(x)(1 - ||G_+||)$ (the term in parenthesis gives the probability that there are no further ladder steps after the nth). Summing over n , the formula for the distribution of M follows. The expression for g_+ was provided at the end of Chapter 4.3.

□

By combining Theorem 4.4.1 with $\psi(R_0) = \mathbb{P}(M > R_0)$ (from Chapter 2.1), we obtain a representation for $\psi(R_0)$ which is known as the Pollaczeck-Khinchine formula:

$\psi(R_0) = \mathbb{P}(M > R_0) = (1-\rho)\sum_{n=0}^{\infty} \rho^n \bar{F}_0^{*n}(R_0)$,

where F_0 has density $b_0 = \frac{\bar{F}x}{\mu}$. Thus the above expression of $\psi(R_0)$ represents the distribution of M as a geometric compound. Unfortunately, this formula is not very useful in terms of computing ruin probabilities, since the formula contains an infinite sum of convolution powers. However, we shall discuss a special case of Pollaczeck-Khinchine that yields a simpler result below.

4.4.1 Ruin probability when the initial reserve is zero

The case $R_0 = 0$ is interesting since it gives a formula for $\psi(R_0)$ which depends only on the mean of the claim size distribution. In fact in this case,

Theorem 4.4.2. $\psi(0) = \rho = \lambda\mu = \frac{1}{1+\eta}$.

Proof. Recall, $\tau_+ = \tau(0)$, thus

$\psi(0) = \mathbb{P}(\tau_+ < \infty)$

$= ||G_+||$

$= \lambda \int_0^\infty \bar{F}(x)dx$

$= \lambda\mu$, which is also equivalent to ρ and $\frac{1}{1+\eta}$. □

The above formula is often referred to as an insensitivity property. This references the fact that $\psi(R_0)$ depends on F only through μ.

Chapter 5

The Random Walk Model

5.1 Veraverbeke's Theorem

Veraverbeke's Theorem (1977) relates the tail behavior of the maximum of certain random walks with the tail behavior of their increments. Let X_i, $i \geq 1$, be a sequence of independent and identically distributed random variables, each having mean μ, $\mu < 0$. Define a random walk S_n by $S_0 = 0$, $S_n = S_{n-1} + X_n$, $n \in \{0, 1, ...\}$. We are interested in $M = \max_{n \geq 0} S_n$. However, let us further discuss the distribution of X_i, the increments of the random walk. Let each $X_i \sim f$, where f is a probability density function with cumulative distribution function F. Let $\bar{F} = 1 - F$, be its tail. If the tail is regularly varying (this will be explained in the next section) and the X_i's have negative mean (which we have assumed already) and $-\alpha < -1$, where $-\alpha$ is the index of regular variation, then $\mathbb{P}(M > t) \sim \frac{1}{-\mu} \int_t^\infty \bar{F}(u)\, du$, as t tends to infinity; this is equivalent to $\mathbb{P}(M > t) \sim \frac{t\bar{F}(t)}{-\mu(\alpha-1)}$, by Karamata's Theorem. (Barbe(2008)) We seek to verify Veraverbeke's Theorem through simulation. The easiest function to use for f is a Pareto distribution.

5.2 Why Pareto?

If X has a Pareto distribution with parameters α and β then

$f(x) = \frac{\alpha\beta^\alpha}{x^{\alpha+1}}$, $x \geq \beta$.

It's tail is $\bar{F} = (\frac{\beta}{x})^\alpha$. We need for its tail to be regularly varying:

$\bar{F}(y) \sim \frac{L(y)}{y^\alpha}$, $y \to \infty$, where $L(y)$ is slowly varying, i.e. $L(y)$ satisfies $\frac{L(yt)}{L(y)} \to 1$, as $y \to \infty$.(Asmussen (2000)) Of course, we will be shifting the Pareto so that it has a negative mean, however this does not effect the properties of the tail.

5.3 The Simulation

The classical function to use to simulate Veraverbeke's results is in fact the Pareto distribution. It has a regularly varying tail and from a programming perspective, is relatively easy to create. We need to ensure that μ, the mean of each step of the random walk, is negative. This further promotes our choice of using Pareto as the distribution of the X_i's , since its mean has a simple, and closed, form. Note that this is not the case for the log-gamma distribution, which also has regularly varying tails. Depending on the value of its parameters, the mean of the log-gamma distribution does not have a closed form mean. Thus it is not a good distribution to base our simulation on.

To generate random variables from a Pareto distribution, we used an inverse uniform transformation. To be more specific, we randomly generate a random uniform value $Z \sim Uni(0,1)$. We then set α and β to some fixed values. Next, we transform each Z generated to a $Y \sim Pareto(\alpha, \beta)$ by using the formula $Y = \frac{\beta}{Z^{1/\alpha}}$. We then create a variable $X = Y - \frac{\alpha\beta}{\alpha-1} + \mu$, which is a Pareto random variable minus its mean with a negative number μ added to it . This ensures that X has mean μ and has Pareto distribution.

Keep in mind that $Y = \frac{\beta}{Z^{1/\alpha}} > \beta$. Thus the tail distribution function is $\bar{F} = \mathbb{P}(Y \geq x) = (\frac{x}{\beta})^{-\alpha}$, for $x \geq \beta$. We calculate the mean, which is finite when $\alpha > 1$, as follows:

$\mathbb{E}X = \int_0^\infty \bar{F}(x)\,dx$

$= \int_0^\beta 1\,dx + \int_\beta^\infty (\frac{x}{\beta})^{-\alpha}\,dx$

$= \beta + \beta \int_1^\infty x^{-\alpha}\,dx$

$= \beta + \beta\frac{1}{\alpha-1} = \beta(\frac{\alpha}{\alpha-1})$, which is the expected result.

These Pareto distributed random variables will be our X_i's, i.e. the steps of our random walk S_n.

The simulation algorithm is as follows.

1. Generate N_1 uniform random variables, $Z \sim Uni(0,1)$.

2. Choose values for two variables α and β, the parameters of the Pareto distribution.

3. Transform each Z generated to a $Y \sim Pareto(\alpha, \beta)$ by using the formula $Y = \frac{\beta}{Z^{1/\alpha}}$.

4. Create a new variable, $X = Y - \frac{\alpha\beta}{\alpha-1} + \mu$, where $\mu < 0$, to ensure that the random walk will have negative drift.

5. Obtain $S_n = X_1 + \ldots + X_n$, for each $n \in \{1, \ldots, N_1\}$.

6. Find $M = \max_{0 \leq n \leq N_1} S_n$.

7. Repeat steps 1-6, N_2 times, creating N_2 different M values, each denoted M_k, where $k \in \{1, \ldots, N_2\}$.

8. Set a value for t, which is the same t from Chapter 5.1.

9. Create a variable c, to count the number of M_k's that exceed t.

10. Calculate simulated $\mathbb{P}(M > t) = \frac{c}{N_2}$.

11. Compare this value to $\frac{t\bar{F}(t)}{-\mu(\alpha-1)}$, which is the theoretical result.

To clarify, we generated $M_k = \max_{0 \leq n \leq N_1} S_{n,k}$, where $S_{n,k}$ is the position of the nth step of the kth random walk. In the initial version the program generated 1000 random walks of 1000 steps in length, i.e. $N_1 = N_2 = 1000$. This number was chosen for two reasons. The first reason is solely practical. The simulation was run using SAS, and generating 1 million random variables takes about 5 minutes of processing time. Significantly larger numbers seem to demand too much processing power. Secondly, the probability of not obtaining the true max of the walk is very low for a large enough $-\mu$. The set of the maxima of the random walks that were used to generate a $\mathbb{P}(M > t)$, which would simply be the proportion of the maxima that were above a set value for t. We then compare this probability to $\frac{t\bar{F}(t)}{-\mu(\alpha-1)}$, from Karamata's Theorem, where $\bar{F}(t) = (\frac{\beta}{t})^{\alpha}$.

For example, we ran this simulation setting the parameters of the Pareto distribution as $\alpha = 2$ and $\beta = 1$. We also set $\mu = -1$.

For $t = 10$, simulated $\mathbb{P}(M > t) = .102$. The expression$\frac{t\bar{F}(t)}{-u(\alpha-1)} = \frac{1}{10} = .1$.

For $t = 100$, simulated $\mathbb{P}(M > t) = .011$. The expression $\frac{t\bar{F}(t)}{-u(\alpha-1)} = \frac{1}{100} = .01$.

For $t = 1000$, simulated $\mathbb{P}(M > t) = .002$. The expression $\frac{t\bar{F}(t)}{-u(\alpha-1)} = \frac{1}{1000} = .001$.

For $t = 10000$, simulated $\mathbb{P}(M > t) = .000$. The expression $\frac{t\bar{F}(t)}{-\mu(\alpha-1)} = \frac{1}{10000} = .0001$.

Notice that the simulated results (the simulated $\mathbb{P}(M > t)$'s) seem quite close to their theoretical ($\frac{t\bar{F}(t)}{-\mu(\alpha-1)}$) counterparts. This is somewhat surprising given the asymptotic nature of Veraverbeke's Theorem.

5.4 Global maximum of a negative drift random walk as a model for stationary waiting time in a G/G/1 queue.

Consider a queueing model where arrivals occur according to a renewal process, the service times are iid from a general distribution and there is one server. The queue has infinite capacity as queue length is unrestricted. Such a queueing model is referred to as G/G/1. Such a process is generally not Markovian. For the process to be Markovian, the arrivals follow a Poisson process and the service distribution is exponential. This model is denoted M/M/1.

Denote the interarrival time by σ_n. Set σ_n to be the length of time between the arrival of customer $n-1$ and n. Then the time of arrival of customer n is given by $t_n = \sigma_1 + \cdots + \sigma_n$ where the time of customer 0 is taken to be 0.

Let τ_n denote the service time of customer n. It is assumed that $\sigma_n, n \geq 1$ and $\tau_n, n \geq 0$ are iid sequences and are independent of each other. Traffic intensity, denoted p is defined as

$p = \frac{\mathbb{E}\tau_0}{\mathbb{E}\sigma_1}$.

Note when $p < 1$, then the average service time is less then the average interarrival time so that it is unlikely that queue lengths will grow excessively large. We assume the stability condition that $p < 1$.

The waiting time process W_n is defined as the length of time from customer n's arrival until commencement of his service. Note that the W_n satisfies a recursion. We have $W_0 = 0$ and

$W_{n+1} = (W_n + \tau_n - \sigma_{n+1})_+$, $n \geq 0$.

To see why this equation is true note that if customer n arrived at time t to the queue, then he departs the system at time $t + W_n + \tau_n$. This is clear because he waits W_n units of time for service to begin and then τ_n time units later his service finishes. Therefore customer $n+1$ who arrives at time $t + \sigma_{n+1}$ will have the wait for his service given by

$$W_{n+1} = \begin{cases} 0 & \text{if } t + \sigma_{n+1} > t + W_n + \tau_n, \\ W_n + \tau_n - \sigma_{n+1} & \text{if } t + \sigma_{n+1} \leq t + W_n + \tau_n. \end{cases}$$

This is equivalent to $(W_n + \tau_n - \sigma_{n+1})_+$.

Let $X_{n+1} = \tau_n - \sigma_{n+1}$, $n \geq 0$ so that $W_{n+1} = (W_n + X_{n+1})_+$.

Let us observe that we can solve the recursion.

Theorem 5.4.1. $W_n = \max\{0, X_n, X_{n+1} + X_n, \cdots, \sum_{i=1}^n X_i\}$.

Proof. We establish the result by induction. For $n = 0$, we have $max\{0\} = 0 = W_0$. Assume the proposition is true for n. For $n+1$ we have

$W_{n+1} = (W_n + X_{n+1})_+$

$= (\max\{0, X_n, X_{n+1} + X_n, \cdots, \sum_{i=1}^n X_i\} + X_{n+1})_+$

$= (\max\{X_{n+1}, X_n + X_{n+1}, \cdots, \sum_{i=1}^{n+1} X_i\})_+$

$= \max\{0, X_{n+1}, X_n + X_{n+1}, \cdots, \sum_{i=1}^{n+1} X_i\}$

□

This establishes the induction step. So the recursion holds for all nonnegative integers n.

Theorem 5.4.2. *Let* $S_n = \sum_{i=1}^n X_i$, *then* $\mathbb{P}(W_n \leq w) = \mathbb{P}(\max_{0 \leq k \leq n} S_k \leq w)$

Proof. Observe that $(X_1, X_2, \cdots, X_n) \stackrel{d}{=} (X_n, X_{n-1}, \cdots, X_1)$

Therefore $(S_1, S_2, \cdots, S_n) \stackrel{d}{=} (X_n, X_{n-1} + X_n, \cdots, \sum_{i=1}^n X_i)$.

Hence $W_n = max\{0, X_n, X_{n-1} + X_n, \cdots, \sum_{i=1}^n X_i\} \stackrel{d}{=} max\{0, S_1, S_2, \cdots, S_n\}$.

Which establishes the above theorem.

□

The stationary waiting time distribution is the limiting distribution of W_n. Under the stability condition $p < 1$ we have

$\mathbb{E}X_{n+1} = \mathbb{E}(\tau_n - \sigma_{n+1}) < 0 \ , \ n \geq 0.$

Therefore the random walk S_n has negative drift implying that $S_n \to -\infty$, a.s.. Hence $W_n \stackrel{a.s.}{\to} W_\infty$, where W_∞ is an a.s. finite, non-negative random variable.

Since, $\max\limits_{0\leq k\leq n} S_k \uparrow \max\limits_{k\geq 0} S_k$,

we have $\mathbb{P}(W_\infty \leq w) = \lim\limits_{n\to\infty} \mathbb{P}(W_n \leq w)$

$= \lim\limits_{n\to\infty} \mathbb{P}(\max\limits_{0\leq k\leq n} S_k \leq w)$

$= \mathbb{P}(\max\limits_{k\geq 0} S_k \leq w).$

Thus we see that the distribution of the global maximum of a random walk with negative drift gives the distribution of the stationary waiting time.

5.5 Ladder Height Variables in the Random Walk Model

As was noted, the stationary waiting distribution in a G/G/1 queue is given by that of the global maximum, i.e. the maximum over all time, of a random walk with negative drift. The value of such a representation would only lie in the possibility that if one may be able to derive a closed form for the distribution or an approximation to it. Ladder heights variables, introduced in Chapter 4.3 will provide us with a tool in this direction.

Let S_n be a random walk with $S_0 = 0$. Similar to Chapter 4.3 let the time of first passage to $(0, \infty)$ be denoted τ_1. In other words,

$\tau_1 = \min\{n \geq 1 : S_n > 0\}.$

We take the minimum of the empty set to be infinity, thus $\tau_1 = \infty$ if $S_n \leq 0$, for all $n \geq 1$. On the set $\{\tau_1 < \infty\}$, define $H_1 = S_{\tau_1}$. Of course $H_1 > 0$ whenever it is defined. On the set $\{\tau_1 = \infty\}$, the variable H_1 is left undefined. Thus τ_1 and H_1 are possibly defective random variables. This means that $\mathbb{P}(\tau_1 < \infty) \leq 1$ and $\mathbb{P}(H_1 < \infty) \leq 1$, with the strict inequality possible. The defect, $\mathbb{P}(\tau_1 = \infty)$, is the same for both. That is

$1 - F_{\tau_1}(\infty) = \mathbb{P}(\tau_1 = \infty) = 1 - F_{H_1}(\infty)$

where F_{τ_1} and F_{H_1} denote the respective distributions of τ_1 and H_1. In the case of interest to us, a random walk with negative drift,

$\mathbb{P}(\tau_1 = \infty) = \mathbb{P}(S_1 \leq 0, S_2 \leq 0, \ldots) > 0$

so that the variables are strictly defective. The variable H_1 is referred to as the first ladder height. The variable τ_1 will be referred to as the ladder epoch. Successive such variables, such as second ladder height and epoch, may be defined. For example, on the set $\{\tau_1 < \infty\}$ define the following

$\tau_2 = \min\{n > \tau_1 : S_n > S_{\tau_1}\} - \tau_1$

where as before $\tau_2 = \infty$ if the value S_{τ_1} is never exceeded. This says that S_{τ_1} is the global maximum of the random walk. Note τ_2 is defined as the *additional* number of steps required for the random walk to exceed the value S_{τ_1}. The value $\tau_1 + \tau_2$ is referred to as the second ladder epoch.

If $\tau_2 < \infty$, define $H_2 = S_{\tau_1+\tau_2} - S_{\tau_1} = S_{\tau_1+\tau_2} - H_1$.

The variable $H_1 + H_2$ is referred to as the second ladder height. The reason for providing definitions of second ladder variable in terms of increments is that the random walk

$S_{\tau_1+1} - S_{\tau_1}, S_{\tau_1+2} - S_{\tau_1}, \ldots, S_{\tau_1+n} - S_{\tau_1}, \ldots$

say $Y_1, Y_1 + Y_2, \ldots, \sum_{i=1}^{n} Y_i, \ldots$ is a probabilistic replica of the original random walk, S_n. Therefore τ_2 has the same distribution as τ_1 and by independent increments is independent of τ_1. Similarly, H_2 is an independent copy of H_1, in terms of its distribution. In this way i.i.d pairs are defined (τ_i, H_i), $i \in \{1, \ldots, k\}$ where k denotes the last finite ladder height variable, i.e. $\tau_k < \infty$ but $\tau_{k+1} = \infty$.

Note that the number of variables defined is itself a geometric random variable with probability mass function $p(1-p)^k$, $k \in \{0, 1, \ldots\}$ with $p = \mathbb{P}(\tau_1 = \infty)$. For a random walk with negative drift $p > 0$. Therefore, with probability one there is a last ladder height epoch, say $\tau_1 + \ldots + \tau_k$. Note that

$W_\infty = \max_{n \geq 0} S_n = H_1 + \ldots + H_k$.

Also observe that if $H_i \sim H$, then define $\widetilde{H} = (1-p)^{-1}H$ which is a proper distribution, $\widetilde{H}(\infty) = 1$. Let $\widetilde{H}_i$ be an i.i.d. sequence of random variables having distribution $\widetilde{H}$. From our representation of W_∞, we have

$$\mathbb{P}(W_\infty \le w) = \mathbb{P}(W_\infty \le w|\tau_1 = \infty)p + \textstyle\sum_{k=2}^{\infty} \mathbb{P}(W_\infty \le w|\tau_{k-1} < \infty, \tau_k = \infty)p(1-p)^{k-1}$$
$$= I_{[0,\infty)}(w)p + \textstyle\sum_{k=2}^{\infty} \mathbb{P}(\widetilde{H}_1 + \ldots + \widetilde{H}_{k-1} \le w)p(1-p)^{k-1}$$
$$= I_{[0,\infty)}(w)p + \textstyle\sum_{k=2}^{\infty} \widetilde{H}^{*(k-1)}(w)p(1-p)^{k-1}$$
$$= \textstyle\sum_{k=1}^{\infty} \widetilde{H}^{*(k-1)}(w)p(1-p)^{k-1}$$
$$= \textstyle\sum_{k=0}^{\infty} \widetilde{H}^{*(k)}(w)p(1-p)^{k}$$

where $\widetilde{H}^{*(0)}$ is the degenerate distribution at zero. Thus the ladder height variables provide a representation of the distribution of W_∞ as that of a compound geometric distribution. We record this in the following theorem.

Theorem 5.5.1. *Let $p = \mathbb{P}(\tau_1 = \infty)$ be the positive defect of the ladder height distribution in the case of a negative drift random walk.*

Define $\widetilde{H}(x) = \mathbb{P}(H_1 \le x|\tau_1 < \infty) = (1-p)^{-1}H(x)$

where H denotes the defective ladder height distribution. Then

$$W_\infty \sim \textstyle\sum_{k=0}^{\infty} p(1-p)^k \widetilde{H}^{*(k)}.$$

5.6 Dominated Convergence Theorem

We will need to make use of the following theorem in Chapter 5.7, so it is stated below. (Bartle (2001))

Theorem 5.6.1. *Let (f_k) be a sequence in $\mathfrak{R}^*(I)$. In other words, let (f_k) be Riemann Integrable. Also let $f(x) = \lim f_k(x)$ for all $x \in I$, where I is defined as the interval $[a, b]$. Suppose that there exist functions $\alpha, \omega \in \mathfrak{R}^*(I)$ such that*

$\alpha(x) \le f_k(x) \le \omega(x)$ for $x \in I$, $k \in \mathbb{N}$.

Then, $f \in \mathfrak{R}^*(I)$ and

$$\textstyle\int_I \lim_{k\to\infty} f_k(x) = \lim_{k\to\infty} \int_I f_k.$$

5.7 Veraverbeke's Approach

In Theorem 5.5.1 we saw that the stationary waiting time distribution for a stable queue, i.e. its traffic intensity is less than 1, is given as a compound geometric distribution. Veraverbeke (1977) exploited this fact by obtaining an approximation to the distribution when the underlying random walk has increments with right tail distribution being heavy-tailed in the sense of belonging to the class of subexponential distributions (see Chapter 3.4). We remark that distribution functions F with a regularly varying tail belong to the class of subexponential functions. Let the class of subexponential functions be denoted $\mathbb{S}$. Note that if $F \in \mathbb{S}$, then for every $n \geq 1$,

$$\lim_{x\to\infty} \frac{\overline{F^{*n}}(x)}{\overline{F}(x)} = n.$$

The term subexponential derives from the property that if $F \in \mathbb{S}$

$$\lim_{x\to\infty} e^{\epsilon x}\overline{F}(x) = \infty, \forall \epsilon > 0.$$

A basic result on subexponential distributions is the following result which is Lemma 7 of section 4 in Chapter 4 of Athreya and Ney, Branching Processes (1972).

Theorem 5.7.1. (Athreya and Ney) *Let $G \in \mathbb{S}$. Then for any $\epsilon > 0$ there is a constant $D < \infty$ such that $\frac{\overline{G^{*n}}(x)}{\overline{G}(x)} \leq D(1+\epsilon)^n, \forall n$ and $x > 0$.*

By combining the above and Theorem 5.5.1 we see that if $\widetilde{H} \in \mathbb{S}$ and $x > 0$

$\mathbb{P}(W_\infty > x) = \sum_{k=1}^{\infty} p(1-p)^k \overline{\widetilde{H^{*k}}(x)}$ because $\overline{\widetilde{H^{*0}}(x)} = 1 - I_{[0,\infty)}(x) = 0$, for $x > 0$. Thus

$$\lim_{x\to\infty} \frac{\mathbb{P}(W_\infty > x)}{1-\widetilde{H}(x)}$$
$$= \lim_{x\to\infty} \sum_{k=1}^{\infty} p(1-p)^k \left(\frac{\overline{\widetilde{H^{*k}}(x)}}{1-\widetilde{H}(x)}\right)$$
$$= \sum_{k=1}^{\infty} p(1-p)^k k$$
$$= \frac{1-p}{p}$$

where the limit may be interchanged with the summation by the Dominated Convergence Theorem (presented as Theorem 5.6) using Theorem 5.7.1. Thus we have the following proposition.

Theorem 5.7.2. *If the ladder height distribution H is subexponential, then* $\mathbb{P}(W_\infty > x) \sim \frac{1-p}{p}\overline{\widetilde{H}}(x) = \frac{1}{p}\overline{H}(x)$ *where* $\overline{H}(x) = H(\infty) - H(x) = 1 - p - H(x)$.

Proof. From Theorem 5.7.2 one sees that a first order asymptotic expansion for the tail distribution of W_∞ is attainable once such an asymptotic equivalent to $\overline{H}(x)$ is obtained. Towards that end, Veraverbeke (1977) shows that if the underlying distribution F of the random walk has finite mean and

$F_1(x) = \int_0^x \overline{F}(t)dt / \int_0^\infty \overline{F}(t)dt$ then $H \in \mathbb{S}$ and furthermore $\overline{H}(x) \sim \frac{p}{(-\mu)} \int_x^\infty \overline{F}(t)dt$ as $x \to \infty$ where p is the defect of H, i.e. $H(\infty) = 1 - p$. In light of the above asymptotic equivalence, one obtains by Theorem 5.7.2 the result.

□

Theorem 5.7.3. *Assuming $F_1 \in \mathbb{S}$ we have* $\mathbb{P}(W_\infty > x) \sim \frac{1}{(-\mu)} \int_x^\infty \overline{F}(t)dt$ *as* $x \to \infty$. *In particular the above holds if $\overline{F}$ is regularly varying.*

Chapter 6

Dependent steps random walk model

Before we discuss the next ruin probability model we need to introduce a few concepts.

6.1 Time series and other notations

A time series is an ordered sequence of observations. The ordering is usually though time, but the ordering can be taken though other dimensions, such as space. The key feature of a time series that we will use in this paper is the dependence among observations. Note that before this point of the paper, we have always assumed that the step sizes were independent from one another.

There are two basic types of stationary time series, Autoregressive (AR) and Moving Average (MA).

Let Z_t be a stochastic process. An Autoregressive model of order p, denoted AR(p), is given by

$$Z_t = \phi_1 Z_{t-1} + \cdots + \phi_p Z_{t-p} + \varepsilon_t$$

where ε_t, known as a "white noise series", is a process with mean zero and finite variance σ^2 , i.e. $\mathbb{E}\varepsilon_t = 0$ and $\mathbb{E}\varepsilon_t^2 = \sigma^2 < \infty$.

Let Z_t be a stochastic process. A Moving Average model of order q, denoted MA(q), is given by

$$Z_t = \varepsilon_t - \theta_1 \varepsilon_{t-1} - \cdots - \theta_q \varepsilon_{t-q}$$

An AR(p) and an MA(q) model can be mixed in order to form an ARMA(p,q) model. This is given by

$$Z_t = \phi_1 Z_{t-1} + \cdots + \phi_p Z_{t-p} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \cdots - \theta_q \varepsilon_{t-q}$$

Before the next model type is introduced, we need to introduce the "backshift" operator. The "backshift" operator denoted $B(Z_t) = Z_{t-1}$.

We define powers of B iteratively. For example,

$B^2 = B(B(Z_t)) = Z_{t-2}$. Thus, $B^k = B(B^{k-1}) = Z_{t-k}$.

An ARIMA model is a "differenced" version of an ARMA model. We usually will fit a differenced model if there is some evidence of a change in the mean value of a series over time. Using the backshift operator, we can now define the ARIMA(p,d,q) model. An ARIMA model may be represented as

$(1 - \phi_1 B - \cdots - \phi_p B^p)(1 - B)^d Z_t = (1 - \theta_1 B - \cdots - \theta_q B^q)\varepsilon_t$.

6.2 The tail of an infinite series of independent random variables

We will now introduce some of the terminology that will be present in Mikosh and Samorodnitsky's ruin probability model. Let us begin by considering the right tail of an infinite series say,

$X = \sum_{j=-\infty}^{\infty} \varphi_j \varepsilon_j$

where ε_n for all $n \in \mathbb{Z}$ is a sequence of i.i.d. random variables that satisfies regularly varying and tail balance conditions for some $\alpha > 0$. The regularly varying and tail balance conditions are below.

$\mathbb{P}(|\varepsilon| > \lambda) = \frac{L(\lambda)}{\lambda^\alpha}$,

$\lim\limits_{\lambda\to\infty} \frac{\mathbb{P}(\varepsilon>\lambda)}{\mathbb{P}(|\varepsilon|>\lambda)} = p$,

and $\lim\limits_{\lambda\to\infty} \frac{\mathbb{P}(\varepsilon<-\lambda)}{\mathbb{P}(|\varepsilon|>\lambda)} = q$, as $\lambda \to \infty$, for $0 < p \leq 1$. L is a slowly varying (at infinity) function. (Mikosch (2000))

In addition to the above conditions, the φ_j's are such that the infinite series X converges. Under certain conditions the following holds true.

$\frac{\mathbb{P}(X>x)}{\mathbb{P}(|\varepsilon|>x)} \sim \sum_{j=-\infty}^{\infty} |\varphi_j|^\alpha [pI_{\varphi_j>0} + qI_{\varphi_j<0}]$ which will be defined as $||\varphi||_\alpha^\alpha$

There are some conditions on φ_j.

$$\begin{cases} \sum_{j=-\infty}^{\infty} \varphi_j^2 < \infty & \text{for } \alpha > 2, \\ \sum_{j=-\infty}^{\infty} |\varphi_j|^{\alpha-\varepsilon} & \text{for some } \varepsilon > 0, \text{ for } \alpha \le 2. \end{cases}$$

This leads us to a theorem, which we will not prove in this paper, see the Mikosch and Samorodnitsky(2000) for details.

Theorem 6.2.1. *Let the i.i.d sequence of ε_n satisfy the regular variation and tail balance conditions with an $\alpha > 0$. If $\alpha > 1$, assume that $\mathbb{E}(\varepsilon) = 0$. If the coefficients φ_n satisfy the above conditions, then the infinite series X exists and $\frac{\mathbb{P}(X>x)}{\mathbb{P}(|\varepsilon|>x)} \sim \sum_{j=-\infty}^{\infty} |\varphi_j|^{\alpha}[pI_{\varphi_j>0} + qI_{\varphi_j<0}] = ||\varphi||_{\alpha}^{\alpha}$ holds.*

The theorem above will help us when we are defining the model in the next section.

6.3 Mikosch and Samorodnitsky's model

Veraverbeke's assumption of independent step sizes might be unrealistic. For example, in a queuing context a typical model has steps distributed as the difference between service times and inter-arrival times, and the independence assumption is agreed not to hold.

Note that, $x \in \mathbb{R}$, $x^+ = max(0, x)$, which is read “the positive part of x ”.

Similarly, $x^- = -min(0, x)$, which is read “the negative part of x “.

Now set $S_n = X_1 + \cdots + X_n$.

Recall, if the steps $X_i \sim F$, $i \in \mathbb{Z}$ were to be i.i.d. subexponential random variables (see Chapter 3.4 for a definition of subexponential) and they generated a random walk

$S_n = X_1 + \cdots + X_n$, for $n \ge 1$ and $S_0 = 0$, then

$\mathbb{P}(\sup_{n\ge 0} S_n > \lambda) \sim \frac{1}{\mu} \int_{\lambda}^{\infty} \bar{F}(u)\, du$, as $\lambda \to \infty$, where $\mathbb{E}X = -\mu$, $-\mu < 0$.

In fact, we will model the random walk steps X_n , $n \in \mathbb{Z}$, of the random walk $S_n = \sum_{i=1}^{n} X_i$ as a two-sided linear process. We will denote the process

$X_i = -\mu + \sum_{j=-\infty}^{\infty} \varphi_{i-j}\varepsilon_j$, $i \in \mathbb{Z}$, where ε_i is a sequence of i.i.d random variables with mean zero (i.e. $\mathbb{E}\varepsilon = 0$)and $\mu > 0$ is a constant. (Mikosch (2000)) [1] The ε_j are commonly referred to as "white noise". Also, note that the X_i's resemble the series X from Chapter 6.2.

While they do resemble a Moving Average (MA) process, notice that the steps X_i's of the random walk are not technically from an MA, ARMA or ARIMA process. ARIMA processes are always represented as one-sided (i.e. casual) linear processes. In other words, in an ARMA or ARIMA process, $\varphi_n = 0$ for all $n < 0$.

We will assume that $\varepsilon = \varepsilon_0$, will satisfy the regular variation and tail balance conditions. These are as follows:

$\mathbb{P}(|\varepsilon| > \lambda) = \frac{L(\lambda)}{\lambda^\alpha}$,

$\lim\limits_{\lambda\to\infty} \frac{\mathbb{P}(\varepsilon>\lambda)}{\mathbb{P}(|\varepsilon|>\lambda)} = p$,

and $\lim\limits_{\lambda\to\infty} \frac{\mathbb{P}(\varepsilon<-\lambda)}{\mathbb{P}(|\varepsilon|>\lambda)} = q$, as $\lambda \to \infty$, for some exponent of variation $\alpha > 1$ and $0 < p \leq 1$. L is a slowly varying (at infinity) function. (Mikosch(2000)) We also note that the coefficients φ_j satisfy the following two conditions:

All of the φ_j's cannot be equal to zero, some of them can equal zero, just not all of them.

Secondly, $\sum_{j=-\infty}^{\infty} |j\varphi_j| < \infty$.

Notice that these conditions bear a strong resemblance to the regular variation and tail balance conditions set forth in Chapter 5.

However, it is worth mentioning that the above condition excludes linear processes with long range dependence which can be described by the condition $\sum_{j=-\infty}^{\infty} |j\varphi_j| = \infty$. Thus, $\sum_{j=-\infty}^{\infty} |j\varphi_j| < \infty$ is referred to as weak dependence.

Combining the weak dependence with the regular variation and tail balance conditions along with the fact that $\mathbb{E}\varepsilon = 0$ imply that $X_i = -\mu + \sum_{j=-\infty}^{\infty} \varphi_{i-j}\varepsilon_j$, $i \in \mathbb{Z}$, converges absolutely with probability 1 and that $\mathbb{E}X_i = -\mu$. Also, by using Theorem 6.2.1, we obtain

$\frac{\mathbb{P}(X>x)}{\mathbb{P}(|\varepsilon|>x)} \sim \sum_{j=-\infty}^{\infty} |\varphi_j|^\alpha [pI_{\varphi_j>0} + qI_{\varphi_j<0}] = ||\varphi||_\alpha^\alpha$ as $\lambda \to \infty$.

[1] Note that this is a bit of an abuse of the terminology “random walk” since the step sizes are not i.i.d.

Let us not forget that we are interested in the ruin probability. This is defined as

$\psi(\lambda) = \mathbb{P}(\sup_{n\geq 0} S_n > \lambda)$ as $\lambda \to \infty$.

Assume that the S_n process had i.i.d steps we would obtain

$\psi_{iid}(\lambda) \sim \frac{1}{\mu(\alpha-1)}\lambda\mathbb{P}(X > \lambda)$

$\sim \frac{||\varphi||_\alpha^\alpha}{\alpha-1}\frac{1}{\mu}\lambda\mathbb{P}(|\epsilon| > \lambda)$ as $\lambda \to \infty$.

Where $\psi_{iid}(\lambda)$, is the ruin probability under the i.i.d. condition.

Now let's begin to build the dependent steps, time series based model. Because of the heavy tails, the event

$\{\sup_n S_n > \lambda\}$, is expected to occur because of a single large positive or a very small negative value of the noise, ε_n. The largest ever contribution of the "important" noise variables to the state of the random walk is as follows.

$S_n = -n\mu + \sum_{k=1}^{n}\sum_{j=-\infty}^{\infty}\varphi_{k-j}\varepsilon_j$

$= -n\mu + \sum_{j=-\infty}^{\infty}\varepsilon_j\sum_{k=1-j}^{n-j}\varphi_k.$

Think about large ε_n's first. A possibly large contribution of ε_j^+ to S_n is multiplied by $\sum_{k=1-j}^{n-j}\varphi_k$. We do not expect every ε_j^+ to make a sizable contribution to the tail of the process, because of the negative drift, the contribution of each ε_j^+ dissipates with time. Therefore, the important noise variables are the ones with high j's, in which case the $\sum_{k=1-j}^{n-j}\varphi_k$ becomes approximately $\sum_{k=-\infty}^{n-j}\varphi_k$. Represent the largest that this factor becomes with

$m_\varphi^+ = \underset{-\infty<n<\infty}{sup}\sum_{k=-\infty}^{n}\varphi_k$

Similarly, for the small negative values of ε_j we utilize

$m_\varphi^- = \underset{-\infty<n<\infty}{sup}\sum_{k=-\infty}^{n}(-\varphi_k).$

Mikosch and Samorodnitsky (2000) show that the following ruin probability results hold asymptotically:

$\psi(\lambda) \sim \sum_{j=1}^{\infty}[\mathbb{P}(m_\varphi^+\varepsilon_j^+ > \lambda + j\mu) + \mathbb{P}(m_\varphi^-\varepsilon_j^- > \lambda + j\mu)]$

$\sim \int_1^\infty \mathbb{P}(m_\varphi^+\varepsilon^+ > \lambda + y\mu)\,dy + \int_1^\infty \mathbb{P}(m_\varphi^-\varepsilon^- > \lambda + y\mu)\,dy$

$\sim \frac{m_\varphi^+}{\mu}\int_{\lambda/m_\varphi^+}^\infty \mathbb{P}(\varepsilon > y)\,dy + \frac{m_\varphi^-}{\mu}\int_{\lambda/m_\varphi^-}^\infty \mathbb{P}(\varepsilon < -y)\,dy$.

By applying Karamata's theorem, we expect to have

$$\psi(\lambda) \sim \frac{[p(m_\varphi^+)^\alpha + q(m_\varphi^-)^\alpha]}{\alpha - 1} \frac{1}{\mu} \lambda \mathbb{P}(|\varepsilon| > \lambda) \sim \frac{[p(m_\varphi^+)^\alpha + q(m_\varphi^-)^\alpha]}{||\varphi||_\alpha^\alpha} \frac{1}{\mu(\alpha - 1)} \lambda \mathbb{P}(X > \lambda), \text{ as } \lambda \to \infty\ .$$

Chapter 7

Recent Developments

Of course, this paper is not all inclusive of all possible ruin probability models. Below are a few models that have appeared in recent literature.

Hult and Lindskog(2008) study a model similar to the one in Chapter 5, but they further complicate matters by giving the insurance company a chance to deposit part of its capital into a bank account that yields interest. In addition the insurance company may invest in n stocks, denoted "risky assets" by Hult and Lindskog. The interest rate of the bank account at time t is denoted r_t, where the r_t's follow a $c\grave{a}dl\grave{a}g$ adapted process ($c\grave{a}dl\grave{a}g$ means that the process is right continuous and has left limits). The prices of the stocks at time t are a spot process S_t^k, where $k \in \{1, \cdots, n\}$. The spot prices are assumed to form strictly positive semimartingales. Now let π_t^0 denote the fraction of the capital deposited into the bank account at time t. Let π_t^k denote the fraction of capital invested in the kth stock at time t. By construction, $\pi_t^0 + \cdots + \pi_t^n = 1$. The cumulative premiums minus claims up to time t are modeled by a $L\acute{e}vy$ process, denoted εY_t, whose downward jumps are assumed to have a heavy-tailed, regularly varying, distribution. Thus, by letting $R_0 > 0$ denote initial capital, the evolution of the risk reserve R_t^ε is given by the stochastic integral equation.

$R_t^\varepsilon = R_0 + \int_{0+}^t \pi_s^0 R_{s-}^\varepsilon r_{s-} ds + \sum_{k=1}^n \int_{0+}^t \pi_s^k R_{s-}^\varepsilon \frac{dS_s^k}{S_{s-}^k} + \varepsilon Y_t,\ t \geq 0.$

Avram, Palmowski and Pistorius (2007) study the joint (two-dimensional) ruin problem. They assume that there are two insurance companies that split both the claims and the premiums in specified proportions. Specifically, say these proportions are δ_1 and δ_2, where $\delta_1 + \delta_2 = 1$. Let the premium rates be labeled c_1 and c_2. Also, let $S(t)$ represent the cumulative amount of claims up to time t. We will assume that $S(t)$ is a spectrally positive $L\acute{e}vy$ process, which is a Lévy process with only upward jumps. The risk process for the ith company, labeled U_i is as follows

$U_i(t) := -\delta_i S(t) + c_i t + u_i$, where $i \in \{1, 2\}$ and u_i are the initial cash reserves of the insurance companies. $S(t)$ will follow the classical model in a matter similar to Chapter 4. Thus $S(t)$ will be defined as the following

$S(t) = \sum_{k=1}^{N(t)} \sigma_k$, where $N(t)$ is a Poisson process with rate λ and the claims σ_k are i.i.d. random variables independent of $N(t)$. The σ_k's have a distribution function, say $F(x)$, as well as a mean of μ^{-1}. We will also assume that the second company (the company for which $i = 2$) is in some way, shape, or form, subordinate to the $i = 1$ company. Avram, Palmowski and Pistorius refer to the second company as the reinsurer. Thus, we assume that the reinsurer gets a smaller share of the profits per amount paid, denoted p_i. In other words we assume

$p_1 = \frac{c_1}{\delta_1} > \frac{c_2}{\delta_2} = p_2$.

Let $\rho = \frac{\lambda}{\mu}$. Just as in Chapter 2 we assume that both of the insurance companies will attempt to avoid ruin by setting the premiums in such a manner that the rate of incoming premiums is higher than the rate of incoming claims. In other words we assume $p_i > \rho$. This implies that in the absence of ruin, $U_i(t) \to \infty$ as $t \to \infty$, where $i \in \{1, 2\}$. In a manner very similar to Chapter 2, ruin occurs when at least one insurance company is ruined at time $\tau = \tau(u_1, u_2)$ where $\tau(u_1, u_2)$ is defined as follows

$\tau(u_1, u_2) = \inf\{t \geq 0 : U_1(t) < 0 \text{ or } U_2(t) < 0\}$.

We can also think of ruin geometrically as the first exit time of $(U_1(t), U_2(t))$ from the positive (first) quadrant. Also familiar to us is the following definition of ruin probability:

$\psi(u_1, u_2) = \mathbb{P}[\tau(u_1, u_2) < \infty]$.

Analytical solutions in multi-dimensional problems are rare, however Avram, Palmowski and Pistorius are able to present a closed form solution as long as the σ_i are exponentially distributed with intensity μ.

The solution to the two-dimensional ruin problem depends on the relative sizes of the proportions δ_1, δ_2 and premium rates c_1, c_2. To think about this geometrically, let $\boldsymbol{\delta} = (\delta_1, \delta_2)$ and $\mathbf{c} = (c_1, c_2)$ be two vectors with origin (u_1, u_2) in the "first quadrant" created by the lines "$y = u_1$" and "$x = u_2$" [1]. It is assumed throughout that the angle of the vector δ with the u_1 axis is larger than the angle of $\mathbf{c}$. In other words we assume $\delta_2 c_1 > \delta_1 c_2$. Starting with initial capital $(u_1, u_2) \in \mathcal{C}$, where $\mathcal{C}$ is the cone

$\mathcal{C} = \{(u_1, u_2) : u_2 \leq \frac{\delta_2}{\delta_1} u_1\}$,

the process (U_1, U_2) will hit the u_1 axis at time τ. In this case, τ is also equivalent to

$\tau(u_1, u_2) = \inf\{t \geq 0 : S(t) > b(t)\}$,

where $b(t) = \min\{\frac{u_1 + c_1 t}{\delta_1}, \frac{u_2 + c_2 t}{\delta_2}\}$.

In the case that $(u_1, u_2) \in \mathcal{C}$, i.e. $\frac{u_2}{\delta_2} \leq \frac{u_1}{\delta_1}$, $b(t)$ is linear, specifically $b(t) = \frac{u_2 + c_2 t}{\delta_2}$ and ruin will always happen to the second company. Thus the two-dimensional problem may be viewed as a one-dimensional problem with linear barrier b, premium rate c_2, and claim rate $\delta_2 \sigma$, i.e.

$\psi(u_1, u_2) = \psi_2(u_2) := \mathbb{P}(\tau_2(u_2) < \infty)$.

where $\tau_2(u_2) = \inf\{t \geq 0 : U_2(t) < 0\}$ and ψ_2 is the ruin probability of the re-insurer (the second insurance company). In the case of claims having structure $\mathbb{P}[\sigma > x] = \beta e^{\mathbf{B}x}\mathbf{1}$, the ruin probability may be written in matrix exponential form,

$\psi_2(u_2) = \boldsymbol{\eta} e^{\delta_2^{-1}(\mathbf{B} + \mathbf{b}\boldsymbol{\eta}) u_2} \mathbf{1}$

For the case $\frac{u_2}{\delta_2} > \frac{u_1}{\delta_1}$

The solution is the Laplace Transform

$\psi(u_1, u_2, s) := \mathbb{E}[e^{-s\tau(u_1, u_2)} \mathbf{1}_{\tau(u_1, u_2) < \infty}]$.

Other papers focus on the the distribution of the time to ruin, denoted

$\tau = \inf\{t > 0 : U(t) < 0\}$.

Borovkov and Dickson (2007) explore this subject within the Sparre Anderson model, where the cash surplus process is

---

[1] Note that I mean x and y in the traditional Cartesian plane sense

$U(t) = u + ct - \sum_{j \le N(t)} X_j$, where $u \ge 0$ is the initial cash reserve, $c > 0$ is the premium rate and $N(t)$ is a delayed renewal process generated by a sequence of inter-claim times, i.e.

$N(t) = \inf\{T_0 + \cdots + T_j \ge t\}$, where $j \ge 0$.

The X_j, where $j \ge 1$ are the sequence of claims that follow the exponential distribution. Note, the X_1 claim size occurs at time T_0, and so on. Also, we assume that the $\{T_j\}$ and $\{X_j\}$ sequences are both i.i.d. and jointly independent.

Again, let X_j be distributed in the following manner.

$\mathbb{P}(X_j > x) = e^{-\lambda x}$ where $x \ge 0$.

Let T_0 and T_1 (which is equal in distribution to $T_j, j > 1$) have densities $f_0(t)$ and $f(t)$, respectively. Borovkov and Dickson prove that the ruin time has the defective density $p_\tau(t)$ shown below.

$p_\tau(t) = e^{-\lambda(u+ct)}\{f_0(t) + \sum_{n=1}^{\infty} \frac{\lambda^n (u+ct)^{n-1}}{n!}[u(f^{*n} * f_0)(t) + c(f^{*n} * f_1)(t)]\}$,

where $f_1(t) = tf_0(t)$, $g * f$ is the convolution of functions g and f, and g^{*n}, $n \ge 2$ is the n-fold convolution of g with itself.

Bibliography

[1] ASMUSSEN, S. BINSWANGER, K. (1997). Simulation of ruin probabilities for subexponential claims. *Astin Bull.* **37** no. 2, 297–318.

[2] ASMUSSEN, S. (2000). Ruin probabilities. World Scientific Publishing Co, Singapore.

[3] AVRAM, F. PALMOWSKI, Z. PISTORIUS, M. (2008). A two-dimensional ruin problem on the positive quadrant. *Insurance Math. Econom.* **42**, no. 1, 227–234.

[4] BARBE, Ph. MCCORMICK, W.P. (2008). Veraverbeke's Theorem at Large. Preprint. Available at http://arXiv.org/abs/0802.3638.

[5] BARTLE, R. (2001). A Modern Theory of Integration. American Mathematical Society. Providence, Rhode Island.

[6] BOROVKOV, K. and DICKSON, D. (2008). On the ruin time distribution for a Sparre Andersen process with exponential claim sizes. *Insurance Math. Econom.* **42**, no. 3, 1104–1108.

[7] HULT, H. LINDSKOG, F. (2008). Ruin Probabilities Under General Investments and Heavy Tailed Claims. Preprint. Available at http://arxiv.org/abs/0809.4372.

[8] MIKOSCH, T. SAMORODNITSKY, G. (2000). The Supremum of a Negative Drift Random Walk with Dependent Heavy-Tailed Steps. *Ann. Appl. Probab.* **10**, no. 3, 1025–1064.

[9] ROSS, S. (2007). Introduction to Probability Models. Academic Press, Burlington.