DISCRETE TIME RISK MODELS WITH RANDOM PREMIUMS

by

LLEWELLYN HILLYER SMITH, JR.

DISSERTATION

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy at
The University of Texas at Arlington
May 2018

Arlington, Texas

ABSTRACT

DISCRETE TIME RISK MODELS

WITH RANDOM PREMIUMS

Llewellyn Hillyer Smith, Jr., Ph.D.

The University of Texas at Arlington, 2018

Supervising Professor: Andrzej Korzeniowski

Over the past century insurance companies relied to a large extent on the continuous time Mathematical Risk Model proposed by Lundberg, known for its ability to estimate the probability of ruin (capital reserve falling below zero), given the initial capital, linear premium rate and cumulative random size claims occurring at random times. In this Dissertation we introduce a discrete time risk model that allows random premiums, and derive the estimates of the ruin probabilities on both finite and infinite time horizons. Tools applied are drawn from modern probability and include, Martingales, Invariance Principle for Brownian motions, and Large Deviation Principle for the asymptotics of rare events. Our considerations can be dubbed “end of the day model”, as ruin is neither declared nor acted upon when it falls between successive discrete times.

ACKNOWLEDGEMENTS

I would like to thank, first and foremost, my advisor Professor Andrzej Korzeniowski. He has been remarkably patient and encouraging with me. His breadth and depth of mathematical knowledge is a constant source of amazement. His exceptional generosity with time and availability is immeasurably helpful. It is a true joy to work with him.

My committee has been wonderful. I would like to thank Professor Guojun Liao for his helpful support and kindness. I would like to thank Professor Benito Chen-Charpentier for inspiring me to think about the asymptotic behavior of systems. I would like to thank Professor Chien-Pai Han for asking such probing questions. I would like thank Professor Suvra Pal for being incredibly helpful.

I am deeply grateful for the support of my family and friends. Mom, Dad, and Denise have all had a hand in helping me achieve this goal. Matt has always been a phone call away. Lan Anh has been especially supportive, helpful, and faithful.

Dedicated to

My Lord and Savior Jesus Christ

Contents

Chapter 1

Introduction

Mathematical Risk Theory is concerned with the probability that the following stochastic process:

$$U(t) = u + ct - \sum_{i=1}^{N_t} X_i,$$

which models the risk reserve of an insurance company will ever have a negative value. The analysis of this process is usually carried out by numerical inversion of an associated Laplace Transform used to solve a renewal equation involving the probability of ruin in infinite time. Since the joint work of Hans Gerber and Elias Shiu [23] [24] in the late 1990's it has been customary to analyze the process in terms of an expected discounted penalty function. Various attempts have also been made to add a Lévy Process component to the model in [16], [1], [3], and [29] among many others. Over the years, discrete time versions of the model have been studied, see for example [27] for recent work along these lines. Lately, a stochastic income component has been added, see for example [5]. Much of the current theory and applications are elucidated in a number of reference works; [2], [6], and [25].

The process $U(t)$ represents the money available to pay the claims X_i which show up at random times N_t. The income stream ct is assumed to be a deterministic process. While this may be reasonable for an insurance company since their customers are contractually obligated to pay premiums in order to receive any coverage of their claims, it is certainly not a reasonable assumption for most models of business income. Also, while it may be true "on average" that an insurance company receives premiums as a continuous stream, even insurance companies receive discrete payments at discrete times. For this reason, we have replaced ct by a stochastic component. This has been done before, and leads to a partial differential equation while solving the resulting renewal equation in [5]. Additionally, we discretized the time variable t. Finally, we were able to find a corresponding representation of the process $U(t)$ as a discrete time zero-mean random walk.

This allows us to analyze the model with a wide variety of tools from modern probability theory.

Our model has the rare benefit of being both more practical since it encompasses a much larger class of possible business processes, and at the same time more amenable to mathematical analysis due to the wide variety of tools that can be applied to it.

This dissertation is organized as follows. In Chapter 2, we introduce a few selected results from the classical theory. The statement of sufficiency of $ct > E(N_t)E(X_i)$ is well-known, but our proof is new, and quite a bit shorter than any we have seen. We introduce the basic definitions used in the field.

In Chapter 3, we show how the continuous model can be approximated by the discrete risk process. Additionally, we introduce an important technical tool that will be used to replace the stochastic index variable N_t by a deterministic index variable n.

In Chapter 4, we derive our model, and show how a zero-mean random walk can be used to analyze risk processes.

In Chapter 5, we develop and elucidate various properties of our model.

In Chapter 6, we start applying our first tool, the Kolmogorov Maximal Inequality. to the question of getting an upper bound on the probability of ruin.

In Chapter 7, we point out a curious problem with relying on ultimate ruin as metric of solvency without also considering finite ruin probabilities.

In Chapter 8, we develop and apply martingale methods to the key question about the probability of ruin.

In Chapter 9, we solve for a general distribution of the time of ruin of a corresponding model that allows us to estimate both finite and ultimate survival probabilities at once. We do this using hitting time distributions of Brownian motion.

In Chapter 10, we apply large deviations principles to gain further insight about the decay rate of the probability of ruin away from the expected value of the process.

Chapter 2

Classical Continuous Time Risk Theory

Modern Insurance Risk Theory started in 1903 with Filip Lundberg's PhD Thesis. In the 1930's, Harold Crámer [7] put the material on a firmer mathematical foundation. In the 1960's, Feller [19] used renewal theory to express the probability of ruin as the unknown solution of a diffeo-integral equation.

We present some of the essential results of what is variously known as *mathematical risk theory* or *insurance mathematics*. We present the classical continuous time theory. The basic problem for an insurance business is that they sign contracts with their customers which - on paper - obligate them to potentially pay out much more money in total claims than they currently have available in cash reserves or ever will. The key word here is "potentially". That is, if every car insurance customer had a $50,000$ dollar car wreck tomorrow, then every car insurance company would simply go into bankruptcy and be unable to pay all of those claims. But such an event is extremely unlikely to occur. And the insurance businesses, themselves, are heavily regulated and monitored to prevent exactly this kind of thing from happening. Never-the-less, there is a great need for the insurance companies to study the probability distribution of the anticipated claims, and carefully track and predict the impact of the actual total claims on their current cash reserves. The following stochastic process tracks the current level of that reserve, which is, broadly speaking, initial capital plus total current income minus total current costs.

Definition 2.1 (Risk Reserve)**.** The stochastic process $U(t,\omega)$ for $t \in [0,\infty)$ and $\omega \in \Omega$ is called the *risk reserve* of an insurance company if and only if

$$U(t,\omega) = u + ct - \sum_{i=1}^{N_t(\omega)} X_i(\omega), \tag{RR}$$

where the real-valued random variable $X_i(\omega) \geq 0$ denotes the amount of the i^{th} claim. The non-negative integer valued random variable $N_t(\omega)$ is a counting process that reveals the number of claims up to, and including time $t \geq 0$. The gross premium rate, $c > 0$, is a real-valued constant, and $u \geq 0$ is the real-valued amount of the initial capital.

Notation 2.2. We often write $U(t)$ or U_t instead of $U(t,\omega)$, and likewise X_i for $X_i(\omega)$, and Y_i for $Y_i(\omega)$ especially when the argument or formula does not depend explicitly on a particular subset or value of Ω.

Remark 2.3. The X_i are assumed to be independent, and, identically distributed (i.i.d.), and each is independent of N_t.

Remark 2.4. The gross premium rate c is the total premium across all contracts in the risk pool collected per unit time (often measured in dollars per unit time). The number of contracts in the risk pool is assumed to be a fixed large positive integer K. Some individual contracts may pay less than $\frac{c}{K}$ per unit time, and some individual contracts may pay more than $\frac{c}{K}$ per unit time. But the insurance company will collect c per unit time from the collection of K contracts, as a whole, with the intention of covering all claims generated by the collection. This is known as a *collective insurance model*.

Remark 2.5. The two parameters c, the rate per unit time charged to cover a collection of K contracts, and u, the initial capital set aside from the start to cover the claims generated by this collection, are the two parameters which the insurance company can choose before offering to cover the losses generated by K contracts.

Remark 2.6. In any given finite time interval $(t_1, t_2]$ there can be at most finitely many claims. Additionally, no two claims occur at exactly the same time. At time 0 there is almost surely no claim.

Remark 2.7. We allow claims of size 0, subject to the same restrictions; no two claims of size 0 occur at the same time, and any finite interval contains at most finitely many of them.

Remark 2.8. For a fixed $\omega_0 \in \Omega$ the mapping $t \mapsto U(t, \omega_0)$ is assumed to be a function of t on $[0, \infty)$ which is right-continuous with left-limits for a.e. $\omega_0 \in \Omega$. The set $\{U(t, \omega_0) \mid t \in [0, \infty)\}$ is then called a particular *trajectory* for the stochastic process $U(t, \omega)$. And the mapping $\omega_0 \mapsto \{U(t, \omega_0) \mid t \in [0, \infty)\}$ associates with a single point, ω_0 in the sample space, Ω, the entire path of a particular trajectory. With this mapping in mind, it makes sense to refer to Ω as the set of all possible trajectories.

Now, we are ready to discuss how N_t and X_i might be distributed. Once we know the probability distributions of N_t and X_i, these two, together, induce a probability measure, $P(\cdot)$, on the set of all possible trajectories, Ω. In the classical theory, N_t is usually a Poisson Process.

Definition 2.9. A stochastic process N_t is said to be a *Poisson Process* if and only if

1. N_t has independent increments for $t \geq 0$;

2. $N_t \sim \text{Poisson}(\lambda t)$ for each fixed $t \geq 0$.

Here Poisson(λt) means a Poisson distribution with parameter λt. A random variable N with $N(\omega) \in \{0, a, 2a, 3a, \ldots\}$, where $a > 0$ and $\omega \in \Omega$, is said to have a Poisson distribution with parameter $\lambda > 0$ and jump size a if and only if

$$P(N = ak) = \frac{\lambda^k}{k!} e^{-\lambda} \qquad k = 0, 1, 2, 3, \ldots$$

Notation 2.10. Unless otherwise indicated, we assume $a \equiv 1$.

If N_t is a *Poisson process* then $\sum_{i=1}^{N_t} X_i$ is called a *compound Poisson process*. In the classical theory, X_i is often exponentially distributed. This brings us to the classical Crámer-Lundberg model. The typical path of a single trajectory starting with initial capital x would look like the line $x + ct$ interrupted by downward "jumps" at random times T_i. Ruin simply means that the current level of the *risk reserve* at some point along the trajectory has become negative.

Definition 2.11 (Crámer-Lundberg Model)**.** A *risk reserve* model (RR)

$$U(t,\omega) = u + ct - \sum_{i=1}^{N_t(\omega)} X_i(\omega) \tag{CL}$$

is called the *Crámer-Lundberg* risk model if and only if $\sum_{i=1}^{N_t} X_i$ is a *compound Poisson process*.

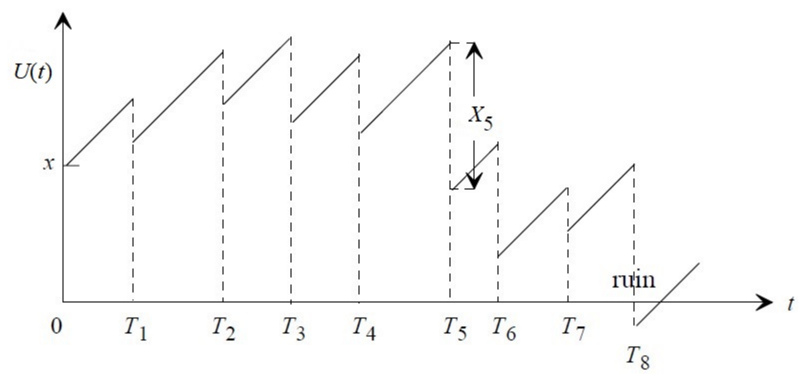

Figure 2.1
A single ruinous trajectory

The probability space Ω is said to be a *filtered probability space* and is denoted as $(\Omega, \mathscr{F}, \mathscr{F}_t, P)$. We assume, unless otherwise stated, that the sigma-algebra $\mathscr{F}$ is complete. A *filtration* $\mathscr{F}_t$ is by definition a nested sequence of increasing subsets of the sigma-algebra $\mathscr{F}$ such that if $s \leq t$ then $\mathscr{F}_s \subset \mathscr{F}_t$, and each $\mathscr{F}_t$ is a sub sigma-algebra of $\mathscr{F}$. We always assume that U_t is *adapted* to the filtration $\mathscr{F}_t$, that is we assume that U_t is measurable with respect to the sigma-algebra $\mathscr{F}_t$. Additionally, we require that the filtration $\mathscr{F}_t$ is right-continuous so that stopping times are well-defined, and we can employ martingale methods. Unless otherwise stated, we assume that $\mathscr{F}_t$ is the smallest filtration with these properties. And, finally, P is a probability measure on Ω.

We return, now, to the chief concern of an insurance company: the probability that they will be able to pay all of their claims. First, though, we ask, “Should any company ever be in the insurance business?” That is, can a company choose c, the collective premium rate, and provide

u, the initial capital, in response to knowing the distribution of N_t, the rate of claims, and the distribution of X_i, the size of each claim, and expect to stay in business in the short run, and make a profit in the long run? That is, does there exist some condition on c, and/or a condition on u from which we can conclude that some portion, S, of the set of all trajectories, Ω, never experience a negative cash reserve, $U(t,\omega) < 0$? And, if there is such a set S, can it be made sufficiently large? We now turn to address these questions.

2.1 Net Profit Condition

Choosing a correct premium rate at the start has profound theoretical and practical consequences. The premium rate c is chosen according to the net profit condition (NPC), which requires that

$$ct > E(N_t)E(X_i),$$

where $E(\cdot)$ represents the expectation of a random variable.

The following theorem shows that such a choice is sufficient to guarantee the existence of an arbitrarily large survival set, even if the variance is infinite.

Theorem 2.12. *Given $U(t) = u + ct - \sum_{i=1}^{N_t} X_i$, suppose $\sum_{i=1}^{N_t} X_i$ is a compound Poisson process. If $c > E(N_t)E(X_i)$ then for every $\eta \in (0,1)$ there exists an initial capital $u \geq 0$, and a survival set $S \subset \Omega$ with $P(S) \geq 1 - \eta$ such that for all $t \geq 0$, and for all $\omega \in S$,*

$$U(t,\omega) = u + ct - \sum_{i=1}^{N_t(\omega)} X_i(\omega) \geq 0.$$

Proof. Let $\eta \in (0,1)$ be given, we will show the existence of an initial capital u, and a set S satisfying the conclusion of the theorem.

Applying the expected value to both sides of the equation, $U(t) = u + ct - \sum_{i=1}^{N_t} X_i$ we get:

$$
\begin{aligned}
E(U_t) &= E(u+ct-\sum_{i=1}^{N_t} X_i) \\
&= E(u+ct)-E(\sum_{i=1}^{N_t} X_i) \\
&= u+ct-E(\sum_{i=1}^{N_t} X_i).
\end{aligned}
$$

Here we have utilized the fact that $\omega \mapsto (u+ct)$ for all $\omega \in \Omega$, hence $E(u+ct) = u+ct$. That is each trajectory associated with a particular $\omega_0 \in \Omega$ can be broken down into its "deterministic part", and its "stochastic part." Namely, every point $\omega \in \Omega$ maps to a trajectory which can be decomposed into a straight line, "the deterministic part", $u+ct$ which is the same for all $\omega \in \Omega$, minus a "stochastic part", $E(\sum_{i=1}^{N_t} X_i)$ which varies for each $\omega \in \Omega$.

Now, since $\sum_{i=1}^{N_t} X_i$ is a *compound Poisson process* we know that N_t is distributed as poisson(λt) by definition. Hence $E(N_t) = \lambda t$ with $\lambda > 0$, and, also, $c > 0$ is, by definition, a real number. And by our hypothesis, $c > E(N_t)E(X_i)$ for $t \in [0,\infty)$. In particular, for $t \in (0,\infty)$ which implies λt is positive since $\lambda > 0$, hence $c > \lambda t E(X_i)$ implies $E(X_i) = \mu < \infty$. Now, by Wald's equation, we have

$$
\begin{aligned}
E(\sum_{i=1}^{N_t} X_i) &= E(N_t)E(X_i) \\
&= \lambda t \mu.
\end{aligned}
$$

So we can write

$$
\begin{aligned}
E(U_t) &= u+ct-\lambda t \mu \\
&= u+(c-\lambda \mu)t.
\end{aligned}
$$

Now, as a consequence of the Strong Law of Large Numbers (SLLN) for compound Poisson processes, we have that

$$\begin{aligned}
\lim_{t\to\infty} \frac{U_t}{t} &= \frac{u}{t} + (c - \lambda\mu) \text{ a.s.} \\
&\geq 0 + (c - \lambda\mu) \text{ since } u \geq 0 \\
&> 0 \text{ since, by hypothesis, } (c - \lambda\mu) > 0.
\end{aligned}$$

So there exists $t_0 \geq 0$ such that for $t > t_0$ we have

$$\begin{aligned}
P\{\omega \in \Omega \mid \frac{U_t(\omega)}{t} > 0\} &= P\{\omega \in \Omega \mid U_t(\omega) > 0\} \\
&\geq 1 - \frac{\eta}{2}.
\end{aligned}$$

The set $B = \{\omega \in \Omega \mid U_t(\omega) \geq 0 \text{ for } t > t_0\}$ has probability greater than or equal to $1 - \frac{\eta}{2}$, and all trajectories included in B are greater than or equal to 0 for all $t > t_0$.

Now we construct another set A such that for all $\omega \in A$, we have $U_t(\omega) \geq 0$ for $0 \leq t \leq t_0$, and the probability of A is within $\frac{\eta}{2}$ of 1 as well, namely $P(A) \geq 1 - \frac{\eta}{2}$.

For a Poisson process the probability of getting k claims by time t_0 can be computed exactly as

$$p_k = \frac{\lambda^k}{k!} e^{-\lambda t_0} \qquad k = 0, 1, 2, 3, \ldots.$$

Observe that we can partition the sample space Ω into a countable partition by defining

$$A_k = \{\omega \in \Omega \mid N_{t_0}(\omega) = k\} \text{ for } k = 0, 1, 2, 3, \ldots,$$

and that each A_k has probability p_k of occurring. Now we have

$$\sum_{k=0}^{\infty} p_k = 1,$$

hence, there exists an N such that

$$\sum_{k=0}^{N} p_k \geq \sqrt{1 - \frac{\eta}{2}} \qquad \text{for } t \leq t_0.$$

Now, we turn our attention to the distribution of the X_i's. Define the kth convolution of F_X the distribution of X with itself as

$$F_X^{*k}(u) = P(X_1 + X_2 + \cdots + X_k \leq u),$$

where $F_X^0(u) \equiv 1$. Observe that $F_X^{*k}(u)$ is a true distribution function, in particular, for each fixed k, we have $\lim_{u \to \infty} F_X^{*k}(u) = 1$. Now, for $k = 0, 1, 2, \ldots, N$ define $u_k = inf\{u \geq 0 \mid \sqrt{1 - \frac{\eta}{2}} \leq F_X^{*k}(u)\}$. Let $u_{max} = max\{u_k\}$, hence we have that

$$F_X^{*k}(u_{max}) \geq \sqrt{1 - \frac{\eta}{2}} \text{ for all } k \in \{0, 1, 2, \ldots N\}.$$

Now, define

$$A_0(u_{max}) = \{\omega \in \Omega \mid N_{t_0}(\omega) = 0\}$$
$$A_k(u_{max}) = \{\omega \in \Omega \mid N_{t_0}(\omega) = k \;\&\; X_1(\omega) + \cdots + X_k(\omega) \leq u_{max}\}.$$

Set $A = \cup_{k=0}^{N} A_k(u_{max})$ and compute a lower bound on the probability of A.

$$\begin{aligned}
P(A) &= \sum_{k=0}^{N} P(A_k) \\
&= \sum_{k=0}^{N} p_k F_X^{*k}(u_{max}) \quad \text{by the indep. of } N_t \text{ and } Y_i \\
&\geq \sum_{k=0}^{N} p_k \sqrt{1 - \frac{\eta}{2}} \\
&= \sqrt{1 - \frac{\eta}{2}} \sum_{k=0}^{N} p_k \\
&\geq \sqrt{1 - \frac{\eta}{2}} \sqrt{1 - \frac{\eta}{2}} \\
&= 1 - \frac{\eta}{2}.
\end{aligned}$$

Let $S = A \cap B$, and observe that

$$
\begin{aligned}
P(S) &= P(A \cap B) \\
&= P(A) + P(B) - P(A \cup B) \\
&\geq 1 - \frac{\eta}{2} + 1 - \frac{\eta}{2} - 1 \\
&= 1 - \eta.
\end{aligned}
$$

Choosing $u = u_{max}$ gives

$$U(t,\omega) = u_{max} + ct - \sum_{i=1}^{N_t(\omega)} X_i(\omega).$$

and

$$
\begin{aligned}
U(t,\omega) &= u_{max} + ct - \sum_{i=1}^{N_t(\omega)} X_i(\omega) \\
&\geq u_{max} - \sum_{i=1}^{N_t(\omega)} X_i(\omega) \\
&\geq 0 \text{ for all } \omega \in S \text{ and } t \in (0, t_0],
\end{aligned}
$$

whereas

$$
\begin{aligned}
U(t,\omega) &= u_{max} + ct - \sum_{i=1}^{N_t(\omega)} X_i(\omega) \\
&\geq ct - \sum_{i=1}^{N_t(\omega)} X_i(\omega) \\
&\geq 0 \text{ for all } \omega \in S \text{ and } t \in (t_0, \infty).
\end{aligned}
$$

Hence for all $\omega \in S$, and $u = u_{max}$, we have $U(t,\omega) \geq 0$ with $P(S) \geq 1 - \eta$, as claimed. □

The set S can naturally be thought of as the "survival set," on which all of the trajectories in the set never go below the $x - axis$. In practice, one would want η to be no more than 0.005. That is, you would want to have at least a 99.5% probability of survival.

2.2 Probability of Ruin

The preceding theorem naturally leads to a number of new definitions. For each trajectory or equivalently for each $\omega_0 \in \Omega$, we can look at the set of all possible times, if any, at which it has a negative risk reserve (RR), and then define the "time of ruin" in the following way:

Definition 2.13 (Time of Ruin)**.** Given $U(t,\omega) = u + ct - \sum_{i=1}^{N(t,\omega)} X_i(\omega)$ then the value $T \in [0,\infty]$ of the extended-real valued function $(u,\omega_0) \mapsto T$ is said to be the *time of ruin* of a given trajectory if and only if

$$T(u,\omega) = inf\{t > 0 \mid U(t,\omega) < 0\}.$$

Observe that in the case that the set $\{t > 0 \mid U(t,\omega_0) < 0\}$ is empty then we have that $T = \infty$ because of the usual convention that $inf\ \emptyset = \infty$. Unfortunately, calling T the time of ruin is standard, even in the case that $T = \infty$. That is when $T = \infty$ we have no ruin at all, rather we have survival. Likewise, although we technically allow $T = 0$, we have as an assumption of our model that $U(0) = 0$ a.s., so the set of trajectories for which $T = 0$ has at most probability 0. Furthermore, observe that $T(\omega)$ is an extended real valued random variable defined on the sample space Ω. Now, we can define the probability of ultimate ruin in the following way:

Definition 2.14 (Probability of Ultimate Ruin)**.** The function $\Psi_{(0,\infty)}(u)$ for $u \geq 0$ is said to be the *probability of ultimate ruin* as a function of the initial capital u if and only if

$$\Psi_{(0,\infty)}(u) = P(\omega \in \Omega : T(u,\omega) < \infty).$$

That is, we simply mark a trajectory with the inf of the set of all times when it crosses below the line $y = 0$ and then we compute the probability of all trajectories which have been marked in such a way. The probability of *ultimate survival* or just *survival* if the meaning is clear from context can also be defined without using T in the following way:

Definition 2.15 (Ultimate Survival)**.** The function $\Phi(u)$ for $u \geq 0$ is said to be the probability of *ultimate survival* as a function of u if and only if

$$\Phi(u) = P(\omega \in \Omega \mid \forall t > 0 \text{ s.t. } u + ct - \sum_{i=1}^{N_t(\omega)} X_i(\omega) \geq 0).$$

Likewise, the *probability of ultimate ruin* can also be defined without using T in the following way:

Definition 2.16 (Probability of Ultimate Ruin)**.** The function $\Psi_{(0,\infty)}(u)$ for $u \geq 0$ is said to be the *probability of ultimate ruin* as a function of the initial capital u if and only if

$$\Psi_{(0,\infty)}(u) = P(\omega \in \Omega : \exists t > 0 \text{ s.t. } u + \theta\mu t - S_{N(t)}(\omega) < 0).$$

It should always be clear from context which definition is being used, but they are equivalent, so in general it won't matter. The probability of ultimate ruin is used as a metric of solvency.

2.3 Adjustment Coefficient and Lundberg Upper Bound

In the current practice, the main object of study in risk theory is the probability of ruin in infinite time rather than finite time. In the best case scenario, the probability of ruin in infinite time will decrease at an exponentially fast rate as the amount, u, of initial capital increases.

In general, except for certain starting values of u (e.g. $u = 0$), and/or certain claim size distributions (e.g. exponential, and finitely supported), there is no known closed form solution to $\Psi(u)$. This situation gave rise to the need to find an upper bound on $\Psi(u)$. This was done by Lundberg in 1909.

For a large class of claim size distributions, there exists a positive number $R > 0$, and a number $C \in (0,1)$ such that $\Psi(u) \leq Ce^{-Ru}$. This number R, if it exists, is called the *adjustment coefficient*. The reason for the name is that the number $R > 0$ reveals how much each additional dollar of initial capital will "adjust" the ruin probability downward. Observe that

$$Ce^{-R(u+1)} = Ce^{-Ru}e^{-R},$$

and with $R > 0$ we have $e^{-R} < 1$; hence each additional dollar lowers the ruin probability by a factor of e^{-R}. We can also "adjust" the ruin probability by increasing our premium rate c, in cases where R exists.

Before explaining the conditions under which the positive number R will exist, it is helpful to see an example where there is no such R. The Inverse Gaussian distribution is a standard example of a claim size distribution for which we may not be able to find an adjustment coefficient R for all possible models $U(t)$. For some values of c the line ct will not hit the incomplete moment generating function.

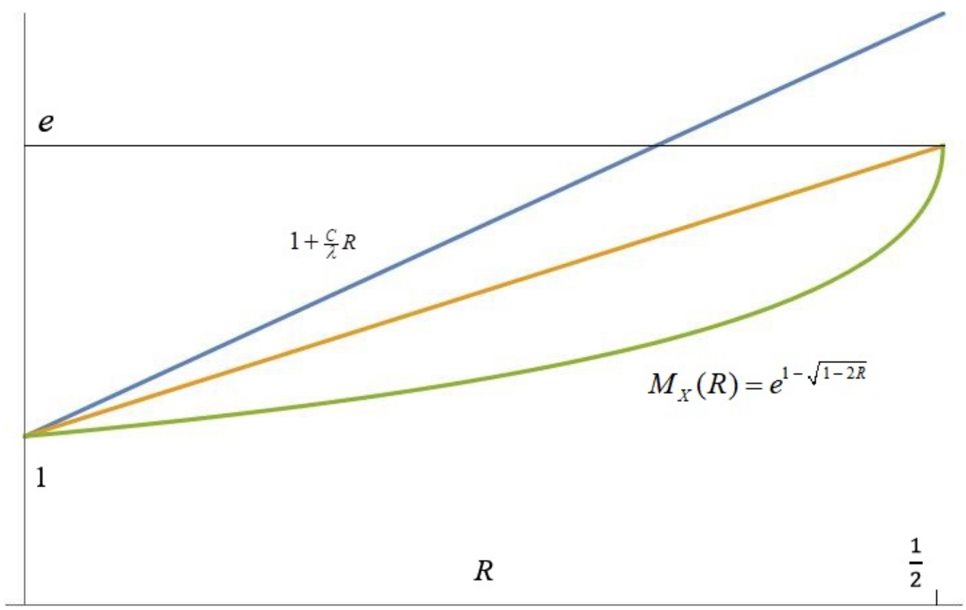

Figure 2.2
Incomplete Moment Generating Function of Inverse Gaussian

This next theorem is a standard argument. It establishes necessary and sufficient conditions to ensure the existence of an adjustment coefficient coefficient R.

Theorem 2.17. *The adjustment coefficient, R, exists if and only if the following equation can be solved for a unique positive R:*

$$1+(1+\theta)\mu R = m_X(R).$$

Here $m_X(s) = E_X(\exp(sX))$ *is the moment generating function of X,* $\mu = E(X)$*, and* θ *is the* safety loading *in the premium rate c. That is, assuming we have chosen c such that* $c - \lambda\mu > 0$*, then* θ *can be determined as*

$$\theta = \left(\frac{c}{\lambda\mu}\right) - 1.$$

Proof. Our equation, $1+(1+\theta)\mu R = m_X(R)$ can be thought of as equating a line on the left-hand side with a concave up curve on the right-hand side. Consider the equation with the variable s, where $s \geq 0$,

$$1+(1+\theta)\mu s = m_X(s).$$

Let

$$h(s) = m(s) - 1 - (1+\theta)\mu s$$

so

$$\begin{aligned} h(0) &= m(0) - 1 - (1+\theta)\mu \cdot 0 \\ &= m(0) - 1 \\ &= E_X(e^{0 \cdot X}) - 1 \\ &= \int_{-\infty}^{\infty} 1 \, \mathrm{dF}_X(x) - 1 \\ &= 1 - 1 = 0. \end{aligned}$$

Therefore, $h(0) = 0$, and graphically both the line, and the moment generating function go through the point $(0,1)$. Now, we consider the derivative of $h(s)$ at 0. We are assuming that, at the very least, $m_X(s)$ exists and is differentiable in an ε-neighborhood of 0. So, we can take the derivative of $h(s)$, and evaluate it at 0. Which leads to

$$\left[h'(s)\right]_{s=0} = \left[m'(s)\right]_{s=0} - (1+\theta)\mu.$$

Therefore,

$$\begin{aligned}
h'(0) &= \left[\frac{\mathrm{d}}{\mathrm{d}s}\int_{-\infty}^{\infty} e^{sx}\,\mathrm{dF}_X(x)\right]_{s=0} - (1+\theta)\mu \\
&= \left[\int_{-\infty}^{\infty} \frac{\mathrm{d}}{\mathrm{d}s} e^{sx}\,\mathrm{dF}_X(x)\right]_{s=0} - (1+\theta)\mu \\
&= \left[\int_{-\infty}^{\infty} xe^{sx}\,\mathrm{dF}_X(x)\right]_{s=0} - (1+\theta)\mu \\
&= \int_{-\infty}^{\infty} xe^{0x}\,\mathrm{dF}_X(x) - (1+\theta)\mu \\
&= \int_{-\infty}^{\infty} x\,\mathrm{dF}_X(x) - (1+\theta)\mu \\
&= \mu - (1+\theta)\mu = \mu - \mu - \theta\mu \\
&= \mu - \mu - \theta\mu \\
&= -\theta\mu < 0.
\end{aligned}$$

Hence, we see that $h(s)$ is decreasing in a neighborhood of 0. Likewise, looking at $h''(s)$, we see that

$$\begin{aligned}
h''(s) &= \frac{\mathrm{d}}{\mathrm{d}s}h'(s) = \frac{\mathrm{d}}{\mathrm{d}s}\left[\int_{-\infty}^{\infty} xe^{sx}\,\mathrm{dF}_X(x) - (1+\theta)\mu\right] \\
&= \frac{\mathrm{d}}{\mathrm{d}s}\left[\int_{-\infty}^{\infty} xe^{sx}\,\mathrm{dF}_X(x)\right] - 0 \\
&= \int_{-\infty}^{\infty} x^2 e^{sx}\,\mathrm{dF}_X(x) > 0.
\end{aligned}$$

Hence, $h(s)$ is concave up, but observe that $m''(s) = h''(s)$, so $m(s)$ is also concave up.

Now, if we assume that there exist $a > 0$, $b > 0$ such that $m(s)$ exists on $(-a, b)$, and that $\lim_{s\to b} m(s) = \infty$. Here b might possibly be ∞, and, observing that $m(s)$ is C^∞ on $(-a, b)$. Then, under these assumptions, we can see that the moment generating function, $m_X(s)$, intersects the line $1+(1+\theta)\mu s$ once, and only once on $(-a,b)$. Observe that both the line, and the moment generating function go through the point $(0,1)$. The line has positive slope (since $1+\theta > 0$, $\mu > 0$). The moment generating function has negative slope since $\frac{\mathrm{d}}{\mathrm{d}s}m(s)|_{s=0} < 0$. Recall that $m''(s) > 0$ on $(-a,b)$, so we need to show that $\lim_{s\to b}$ of $m_X(s)-(1+(1+\theta)\mu s) = \infty$. If $b < \infty$ then $\lim_{s\to b}(1+(1+\theta)\mu s) = (1+(1+\theta)\mu b) < \infty$. But, $\lim_{s\to b} m(s) = \infty$ so $\lim_{s\to b} m(s)-(1+(1+\theta)\mu s) = \infty$. Hence, $m(s)$ is now above the line as the line approaches b. Recall that $m(s)$ was below the line when $s=0$. Also, it had a negative derivative when $s=0$. Therefore, it must now have a positive derivative. It changed signs, and it changed signs only once since $m''(s) > 0$. It is C^∞ on $(-a,b)$. So, by the Intermediate Value Theorem (IVT), it assumed the value 0 on the interval (a,b). Any, by the previous derivative arguments, it assumed that value only once. Hence, there is one-and-only-one positive value R solving

$$1+(1+\theta)\mu R = m_X(R).$$

If $b=\infty$, then we will show that $\lim_{s\to\infty} m(s)-(1+(1+\theta)\mu s) = \infty$. Remember that $E(X)=\mu > 0$ and that $X \geq 0$. So, there exists a $\delta > 0$ and an $m > 0$ so that $P(X > \delta) \geq m > 0$. Observe that we are simply saying that a nonnegative random variable with positive expectation can be bounded away from 0 on a set of positive measure. So we have

$$\begin{aligned} m_X(s) &= \int_0^\infty e^{sx}\,\mathrm{dF}_X(x) \\ &\geq \int_\delta^\infty e^{sx}\,\mathrm{dF}_X(x) \\ &\geq e^{s\delta}m \end{aligned}$$

since $e^{s\delta} \leq e^{sx} \quad \forall s \geq \delta$ and $P_X[\delta,\infty) \geq m$. So we get

$$\lim_{s\to\infty}(m_X(s)-(1+(1+\theta)\mu s))$$

$$\geq (e^{s\delta}m - 1 - (1+\theta)\mu s) \geq \infty,$$

since $me^{s\delta}$ dominates a linear function in s. So, this shows that $m_X(s)$ goes to ∞ faster than $(1+(1+\theta)\mu s)$. Hence, it intersects it exactly once. And, again, there exists one-and-only-one R for which the equation is satisfied. $\square$

We have shown that there exists a unique positive constant R - known as the *adjustment coefficient* - in the case where the moment generating function is defined on an interval $(-a,b)$ about 0. And, on that interval, the moment generating function approaches ∞ as s approaches the right hand boundary b of our interval around 0. Furthermore, this number R, can be used to get an upper bound on the probability of ruin $\Psi(u)$. Namely,

$$\Psi(u) \leq e^{-Ru}.$$

This is a famous standard result in risk theory first proved by Lundberg [18]. We will often refer to it as a standard of comparison. Its primary value is that it guarantees a certain rate of decay of the probability of ultimate ruin as the initial capital u increases.

Chapter 3

Discrete Time Risk Theory

We introduce the discrete time risk model. This model eliminates the continuous time arrival process for individual claims. The aggregate claim process and the aggregate premium process arrive at the end of the current period, n, simultaneously.

One particular version of a discrete time risk model, building on the work of De Vylder [8], was used by Dickson and Waters [14], and studied extensively by Dickson in later works, [11], [12], and [13]. Suppose for example that the premium rate $c = 1$ is given, and that the mean of the claims $E(Z_i) < 1$ is given, then we could write

$$\text{(D)} \qquad U(n) = u + n - \sum_{i=1}^{n} Z_i,$$

where the initial capital $u \in \{0, 1, 2, \ldots\}$ is given, and $n \in \{1, 2, 3, \ldots\}$ is the total accumulated premium by the end of period, n. Furthermore, suppose that $Z_i \in \{0, 1, 2, 3, \ldots\}$ then observe that we would have $U(n) \in \mathbb{Z}$. Suppose further that $Z_i \sim Z$ i.i.d., with $P(Z = k) = p_k$, $k = 0, 1, \ldots$.

It is important to note that Dickson also showed that this model (D) is equivalent to the binomial model,

$$\text{(BM)} \qquad U(n) = u + n - \sum_{i=1}^{N(n)} Y_i, \quad n = 1, 2 \ldots$$

studied by Gerber [22], where $Y_i \sim Y$ i.i.d. claim sizes, $P(Y = k) = p_k$, $k = 1, 2, \ldots$, with $qE(Y) < 1$, and $N(n) = \varepsilon_1 + \cdots + \varepsilon_n$ is a Binomial Process where ε_i i.i.d. $P(\varepsilon_i = 1) = q$, $P(\varepsilon_i = 0) = 1 - q = p$. The significance of (D) and (BM) equivalence lies in (BM) involving a random sum of random variables, but (D) is a non-random sum of random variables, and thus a much simpler object to study.

Furthermore, Dickson showed that for (D) with

$$P(Z=0)=p,\;\; P(Z=k)=q(1-\alpha)\alpha^{(k-1)}, k=1,2,\dots \text{ and letting}$$
$$P(Y=k)=(1-\alpha)\alpha^{(k-1)}, \nu=E(Y)=\frac{1}{(1-\alpha)} \text{ and}$$
$$P(X=k)=q(1-q)^{k-1}, \mu=E(X)=\frac{1}{q},$$

the ruin probability is

$$\Psi^D(u)=\frac{q}{1-\alpha}\left(\frac{\alpha}{p}\right)^u=\frac{E(Y)}{E(X)}\left(\frac{P(Y>1)}{P(X>1)}\right)^u. \tag{3.1}$$

Notice that X and Y each have a geometric distribution.

3.1 Estimating the Ruin Probability of a Continuous Model

For the continuous Crámer-Lundberg model (CL), $U(t)=u+ct-\sum_{i=1}^{N(t)}Y_i$ there is a known probability of ruin for all $u\geq 0$ (see for example [2]). Provided that the Poisson counting process $N(t)$ has intensity $\frac{1}{\lambda}$ and that $X_i\sim X_1$ are i.i.d. inter-arrival times, with p.d.f. $\frac{1}{\mu}e^{-\frac{1}{\mu}x}$, $E(X_1)=\mu$, and $Y_i\sim Y_1$ are i.i.d. claim sizes, with p.d.f. $\frac{1}{\nu}e^{-\frac{1}{\nu}x}$, $E(Y_1)=\nu$. And given that $\nu<\mu$ then the probability of ultimate ruin is:

$$\Psi^{CL}(u)=\frac{\nu}{\mu}e^{-(\frac{1}{\nu}-\frac{1}{\mu})^u}=\frac{E(Y_1)}{E(X_1)}\left(\frac{P(Y_1>1)}{P(X_1>1)}\right)^u. \tag{3.2}$$

Observing the form of equation (3.1) and equation (3.2) suggests that there may be a connection between $\Psi^D(u)$, and $\Psi^{CL}(u)$ when q, $p\to 0$.

Theorem 3.3. *Given the discrete model, $U(t)$* (D)*, with the notation, and assumptions of equation* (3.1)*, and given the continuous model, $U(t)$* (CL)*, with the notation, and assumptions of equation* (3.2)*, suppose that $q_n=\frac{1}{\mu n}$, and $1-\alpha_n=\frac{1}{\nu n}$; then $\Psi_n^D(u)\to\Psi^{CL}(u)$ as $n\to\infty$.*

Proof. Observe that the rescaled geometric distribution converges to an exponential distribution. Namely, let $P(Z_\theta = k) = \theta(1-\theta)^{(k-1)} k = 1,2,\ldots$ Fix n and let $\theta = \frac{\lambda}{n}$, $x = \frac{k}{n}$ then the random variable Z_θ^n with values in $\{\frac{1}{n}, \frac{2}{n}, \ldots\}$ satisfies

$$\begin{aligned}
1 &= \sum_{k=1}^{\infty} P\left(Z^n_{\frac{\theta}{n}} = \frac{k}{n}\right) = \sum_{k=1}^{\infty} P\left(Z_{\frac{\theta}{n}} = k\right) \\
&= \sum_{k=1}^{\infty} \lambda \left(1 - \frac{\lambda}{n}\right)^{n \cdot \frac{k}{n}} \frac{1}{n} \overset{n\to\infty}{\longrightarrow} \int_0^\infty \lambda e^{-\lambda x} dx.
\end{aligned}$$

This means $Z^n_{\frac{\lambda}{n}}$ converges in distribution to W, $W \sim$ exponential with density $\lambda e^{-\lambda x}$

Therefore $P(Z^n_{\frac{\lambda}{n}} > 1) \to P(W > 1) = e^{-\lambda}$

Now

$$\Psi^D_n(u) = \frac{q_n}{1-\alpha_n} \left(\frac{P(Y^n_{\frac{1}{\nu n}} > 1)}{P(X^n_{\frac{1}{\mu n}} > 1)} \right)^u \to \frac{\nu}{\mu} \left(\frac{e^{-\frac{1}{\nu}}}{e^{-\frac{1}{\mu}}} \right)^u = \Psi^C(u).$$

□

We have shown the pointwise convergence of the probability of ruin associated with a sequence of discrete models to the probability of ruin for a continuous model. This establishes that we can use a discrete risk model and its associated probability of ruin to approximate the probability of ruin for a continuous risk model.

Chapter 4

A New Model Utilizing a Zero-Mean Random Walk

In the last chapter, we utilized a sequence of discrete models (D) of the form

$$U(n) = u + n - \sum_{i=1}^{n} Z_i.$$

To each model in the sequence we calculated its ruin probability and formed another sequence. This sequence of ruin probabilities approached the ruin probability of a continuous time model (CL) of the form

$$U(t) = u + ct - \sum_{i=1}^{N_t} X_i.$$

Observe, though, that our model (D) is actually a deterministic sum since the upper index on the sum is n whereas the model whose ruin probability we have approximated is a random sum since the upper index on the sum is N_t. Also, notice that the discrete model (D) is a random walk with an offset, and a linear drift.

Random walks and their close cousins Brownian motion have a long history in the analysis of financial time series. In fact, they were the predominant method of analysis before the advent of Geometric Brownian motion. Stocks and Bonds can never take on a negative value, and so Geometric Brownian which is never negative is a better model. On the other hand, the risk reserve (RR) of an insurance company can certainly take on negative values. So, it is natural to seek to analyze the risk reserve process with a random walk, and with Brownian motion.

The key questions in this chapter are 1) what is the best form of the random walk we will use, that is, what is "our model," and 2) is there a better form of the Crámer-Lundberg model whose ruin probability we may be able to approximate?

For the sake of comparing to the current standard, we outline a heuristic to suggest how one might approximate the ruin probability of the classical compound Poisson process used in insurance mathematics. But we believe that a random walk rather than a compound Poisson process is simply better suited to the task of modeling the risk reserve of an insurance company.

4.1 Derivation of the Model

Boikov [5] in 2002 introduced a generalized continuous time model, which we call the Continuous Boikov model:

$$\text{(CB)} \qquad U(t) = u + \sum_{j=1}^{N_2(t)} y_j - \sum_{i=1}^{N_1(t)} x_i,$$

where $U(t)$, u, $\sum_{i=1}^{N_1(t)} x_i$, with $\sum_{i=1}^{N_1(t)} x_i \equiv \sum_{i=1}^{N(t)} x_i$ are as before. On the other hand, $\sum_{j=1}^{N_2(t)} y_j$ represents the total premium arrivals y_j by time t, which arrive according to a Poisson Process $N_2(t)$ assumed to be independent of $N_1(t)$.

Our objective is to introduce and study a new model which can be considered a discrete time counterpart to (CB). Furthermore, our model allows for a considerable reduction in random complexity, by proposing an equivalent $U(t)$ representation that replaces random sums $\sum_{j=1}^{N_2(t)} y_j$, $\sum_{i=1}^{N_1(t)} x_i$, by $\sum_{j=1}^{t} \tilde{y}_j$, $\sum_{i=1}^{t} \tilde{x}_i$ for appropriately defined $\tilde{y}_j$, and $\tilde{x}_i$. The significance of our model is that it applies to actual real-world practice. "Ruin" is naturally defined as having a negative balance at the end of the day. Likewise, "ruin" has not occurred if the balance at the end of the day is not negative. This is irrespective of whether or not the balance may have been in the "red" at some point before the end of the day.

To this end, we introduce the well-known discrete binomial model, which can be written as

$$U_t = u + t - \sum_{i=1}^{N_t} X_i \quad t \in \{1, 2, 3, \dots\},$$

where $N_t = I_1 + I_2 + \cdots + I_t$ are i.i.d. Bernoulli random variables representing the presence or absence of a claim during period t. Dickson [14] was able to simplify this model to

$$\text{(DD)} \qquad U_t = u + t - \sum_{i=1}^{t} X_i \quad t \in \{1, 2, 3, \dots\},$$

which we call the Discrete Dickson (DD) model. Here the only change was to replace the random upper bound N_t with a deterministic upper bound t - namely, just the number of periods that have elapsed since time equals 0. He did this by allowing for a claim size of 0, and then defining the probability of a "claim" of size 0 by the exact probability of not getting any claim in period t.

In our new model, we propose to discretize time in the Boikov model (CB), and use the deterministic upper bound technique from the Dickson model (DD) applied to both summands. The model would then become

$$U_t = u + \sum_{j=1}^{t} Y_i - \sum_{i=1}^{t} X_i, \quad t \in \{1, 2, 3, \dots\}$$

which we often call "our model," henceforth, where $Y_j \in [0, \infty)$, and $X_j \in [0, \infty)$. Notice that Y_j and X_i are now allowed to assume the value 0. Finally, since we have the same upper bound t for both summands, we can look at

$$\begin{aligned} U_t &= u + \sum_{i=1}^{t} Y_i - \sum_{i=1}^{t} X_i \\ &= u + \sum_{i=1}^{t} (Y_i - X_i), \end{aligned}$$

where $(Y_i - X_i) \in \mathbb{R}$. In this way, we can model the surplus as a discrete random walk.

Recall that the safety loading requirement applied to the traditional model would yield, for some $\theta > 0$,

$$\begin{aligned} E(ct) &= (1+\theta)E(\sum_{i=1}^{N_t} X_i) \\ &= (1+\theta)\lambda t E(X_i). \end{aligned}$$

So, since $E(ct) = ct$, dividing both sides of the equation by t we get $c = (1+\theta)\lambda E(X_i)$. Now, applying this same logic to the new model we get, for some $\theta > 0$,

$$\begin{aligned} E(\sum_{i=1}^{t} Y_i) &= (1+\theta)E(\sum_{i=1}^{t} X_i) \\ &= (1+\theta)tE(X_i). \end{aligned}$$

So, since $E(\sum_{i=1}^{t} Y_i) = tE(Y_i)$, dividing both sides of the equation by t we see that

$$\begin{aligned} E(Y_i) &= (1+\theta)E(X_i) \\ &= \theta E(X_i) + E(X_i) \end{aligned}$$

is required to achieve a safety loading percent of θ.

We are now in a position to represent the model

$$U_t = u + \sum_{i=1}^{t} Y_i - \sum_{i=1}^{t} X_i$$

in a much more informative manner.

We will write the process U_t as the sum of a fixed offset u plus a deterministic linear drift, and finally plus a mean-zero random process. We will rewrite the process in the following way:

$$\begin{aligned} U_t &= u + \sum_{i=1}^{t} Y_i - \sum_{i=1}^{t} X_i \\ &= u + \sum_{i=1}^{t} (\theta E(X_i) + Y_i - \theta E(X_i)) - \sum_{i=1}^{t} X_i \\ &= u + \sum_{i=1}^{t} \theta\mu + \sum_{i=1}^{t} (Y_i - \theta\mu) - \sum_{i=1}^{t} X_i \\ &= u + \theta\mu t - \sum_{i=1}^{t} [X_i - (Y_i - \theta\mu)] \\ &= u + \theta\mu t - \sum_{i=1}^{t} [X_i - Y_i + \theta\mu] \\ &= u + \theta\mu t - \sum_{i=1}^{t} Z_i, \end{aligned}$$

where $Z_i = X_i - Y_i + \theta\mu$. Notice that

$$\begin{aligned} E(Z_i) &= E(X_i) - E(Y_i) + E(\theta E(X_i)) \\ &= \mu - (1+\theta)\mu + \theta\mu \\ &= \mu - \mu - \theta\mu + \theta\mu \\ &= 0. \end{aligned}$$

So, Z_i is a zero-mean process, as expected. Hence, we have represented our process as

$$\begin{aligned} U_n &= u + \theta\mu t - \sum_{i=1}^{t} X_i - Y_i + \theta\mu \\ &= u + \theta\mu t - \sum_{i=1}^{t} Z_i. \end{aligned}$$

That is, we have represented it as an initial value u, plus a deterministic linear drift $\theta\mu t$, and finally, minus a mean-zero random walk $\sum_{i=1}^{t} Z_i$. The first two terms, $u + \theta\mu t$, remain the same for every trajectory, hence they are best described as "deterministic." The last term, $-\sum_{i=1}^{t} Z_i$, is the only one that represents the random variation of the process.

4.2 Zero Net Profit Leads to Ruin

Consider the following example. Suppose Z_i has the following distribution:

$$\begin{aligned} P(Z_i = -1) &= \frac{1}{4} \\ P(Z_i = 0) &= \frac{1}{2} \\ P(Z_i = 1) &= \frac{1}{4}. \end{aligned}$$

We have indicated that adding an "atom" at 0 may be necessary to approximate the ruin probabilities of the classical continuous process. In some cases, an "atom" at 0 naturally appears in successive iterates of the random variable.

Observe that the distribution is symmetrically distributed about 0, but not equally distributed among its 3 possible values. Surprisingly, though, this distribution can be thought of as sampling from the following symmetric random walk at its even time points.

Suppose $V_k \quad k = 0,1,2,3,\ldots$ has the following distribution. For $k = 0$, $P(V_0 = 0) = 1$, and for $k \geq 1$

$$P(V_k = -\frac{1}{2}) = \frac{1}{2}$$
$$P(V_k = \frac{1}{2}) = \frac{1}{2}.$$

Now, for $i = 0$, $P(Z_0 = 0) = 1$, and for $i \geq 1$ we use the following equivalence $Z_i \equiv V_{2i-1} + V_{2i}$ for $i = 1,2,3,\ldots$. Now, the distribution of Z_i can be determined by

$$\begin{aligned}
P(Z_i = -1) &= P(V_{2i-1} = -\frac{1}{2} \text{ and } V_{2i} = -\frac{1}{2}) \\
&= \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4} \\
P(Z_i = 0) &= P(V_{2i-1} = -\frac{1}{2} \text{ and } V_{2i} = \frac{1}{2}) \\
&\quad + P(V_{2i-1} = \frac{1}{2} \text{ and } V_{2i} = -\frac{1}{2}) \\
&= \frac{1}{2} \cdot \frac{1}{2} + \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4} + \frac{1}{4} = \frac{1}{2} \\
P(Z_i = 1) &= P(V_{2i-1} = \frac{1}{2} \text{ and } V_{2i} = \frac{1}{2}) \\
&= \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4},
\end{aligned}$$

which agrees with our original distribution of Z_i. At the moment, we don't need to use this observation, but it is important to note that this Z_i process is equivalent to the sum of two symmetric simple random walks.

Now, we calculate the probability of ruin for this example. Consider the case where $u \geq 0$, $\theta = 0$, and Z_i has the distribution from above.

Proposition 4.1. *Given the model $U_n = u + \theta\mu n - \sum_{i=1}^{n} Z_i$ with no safety loading, that is $\theta = 0$ and any arbitrary amount of initial capital, $u \geq 0$, suppose Z_i has the distribution above. Then the probability of ultimate ruin is* 1.

Proof. Due to independent and stationary increments, the probability of survival only depends on the level u, and not the time n. That is, $P($ Survival with $U_0 = 0$ for $Z_n \quad n = 1,2,3,\dots) = P($ Survival with $U_k = 0$ for $Z_n \quad n = k+1, k+2, k+3, \dots k \geq 0)$. Of course, $U_k = u + \theta\mu k + \sum_{i=1}^{k} Z_i$ is the current level of the risk reserve.

Now, continuing with our proof, let $u = 0$, $\theta = 0$, $U_0 \equiv 0$, and for $i > 0$ $P(Z_i = -1) = \frac{1}{4}$, $P(Z_i = 0) = \frac{1}{2}$, and $P(Z_i = 1) = \frac{1}{4}$.

By using the total probability formula, we get

$$\begin{aligned} P(\text{ Survival from } U_0 = 0) &= \frac{1}{4} P(\text{ Survival from } U_1 = 1) \\ &+ \frac{1}{2} P(\text{ Survival from } U_1 = 0) \\ &\leq \frac{1}{4} + \frac{1}{2} = \frac{3}{4}. \end{aligned}$$

This makes perfect sense, since there was a $\frac{1}{4}$ chance of getting $U_1 = -1$ which is “ruin”. Recalling that

$$P(\text{ Survival from } U_0 = 0) = P(\text{ Survival from } U_1 = 0),$$

we get

$$\begin{aligned} P(\text{ Survival from } U_0 = 0) &= \frac{1}{4} P(\text{ Survival from } U_1 = 1) \\ &+ \frac{1}{2} P(\text{ Survival from } U_1 = 0). \end{aligned}$$

Now, skipping the phrase "survival from," and skipping the index, we can write

$$P(U=0)=\frac{1}{4}P(U=1)+\frac{1}{2}P(U=0)$$

And then, subtracting $\frac{1}{2}P(U=0)$ from both sides we get

$$\frac{1}{2}P(U=0)=\frac{1}{4}P(U=1).$$

Finally, multiplying both sides by 2, we see that

$$P(U=0)=\frac{1}{2}P(U=1)\leq\frac{1}{2}.$$

But

$$P(U=1)=\frac{1}{4}P(U=2)+\frac{1}{2}P(U=1)+\frac{1}{4}P(U=0),$$

and using the substitution $P(U=0)=\frac{1}{2}P(U=1)$ we get

$$\begin{aligned}P(U=1)&=\frac{1}{4}P(U=2)+\frac{1}{2}P(U=1)+\frac{1}{4}[\frac{1}{2}P(U=1)]\\&=\frac{1}{4}P(U=2)+\frac{1}{2}P(U=1)+\frac{1}{8}P(U=1)\\&=\frac{1}{4}P(U=2)+\frac{5}{8}P(U=1).\end{aligned}$$

So, subtracting $\frac{5}{8}P(U=1)$ from both sides, we get

$$\frac{3}{8}P(U=1)=\frac{1}{4}P(U=2).$$

And, now, using the substitution $2P(U=0)=P(U=1)$ we have

$$\begin{aligned}\frac{3}{8}[2P(U=0)]&=\frac{1}{4}P(U=2)\text{ so}\\\frac{3}{4}P(U=0)&=\frac{1}{4}P(U=2)\text{ or}\\P(U=0)&=\frac{1}{3}P(U=2)\leq\frac{1}{3}.\end{aligned}$$

So now we can use our substitutions, and our known probabilities, to get

$$\begin{aligned}
P(U=0) &= \frac{1}{3}P(U=2) \\
&= \frac{1}{3}[\frac{1}{4}P(U=3)+\frac{1}{2}P(U=2)+\frac{1}{4}P(U=1)] \\
&= \frac{1}{3}\{\frac{1}{4}P(U=3)+\frac{1}{2}[3P(U=0)]+\frac{1}{4}[2P(U=0)]\} \\
&= \frac{1}{12}P(U=3)+\frac{1}{2}P(U=0)+\frac{1}{6}P(U=0) \\
&= \frac{1}{12}P(U=3)+\frac{2}{3}P(U=0) \\
&= \frac{1}{4}P(U=3) \leq \frac{1}{4}.
\end{aligned}$$

Now, we will make an induction hypothesis. Suppose the following equation holds for $U=1, U=2, \ldots, U=k$

$$P(U=0) = \frac{1}{k+1}P(U=k).$$

We need to show that $P(U=0) = \frac{1}{k+2}P(U=k+1)$ in order to complete the proof. Assuming, by our induction hypothesis, $P(U=0) = \frac{1}{k+1}P(U=k)$, is true, we can use it to compute

$$\begin{aligned}
P(U=0) &= \frac{1}{k+1}\{\frac{1}{4}P(U=k+1)+\frac{1}{2}P(U=k)+\frac{1}{4}P(U=k-1)\} \\
&= \frac{1}{k+1}\{\frac{1}{4}P(U=k+1)+\frac{1}{2}[(k+1)P(U=0)]+\frac{1}{4}[kP(U=0)]\} \\
&= \frac{1}{k+1}\frac{1}{4}P(U=k+1)+\frac{1}{k+1}\frac{k+1}{2}P(U=0)+\frac{1}{k+1}\frac{k}{4}P(U=0).
\end{aligned}$$

Now, gathering all the terms that involve $P(U=0)$ on one side of the equation, and solving for

$P(U=k+1)$ on the other side, we get

$$\begin{aligned}
P(U=0)-\frac{1}{2}P(U=0)-\frac{k}{k+1}\frac{1}{4}P(U=0) &= \frac{1}{k+1}\frac{1}{4}P(U=k+1)\\
\frac{1}{2}P(U=0)-\frac{k}{k+1}\frac{1}{4}P(U=0) &= \frac{1}{k+1}\frac{1}{4}P(U=k+1)\\
2P(U=0)-\frac{k}{k+1}P(U=0) &= \frac{1}{k+1}P(U=k+1)\\
2\frac{k+1}{k+1}P(U=0)-\frac{k}{k+1}P(U=0) &= \frac{1}{k+1}P(U=k+1)\\
\frac{2k+2-k}{k+1}P(U=0) &= \frac{1}{k+1}P(U=k+1)\\
\frac{k+2}{k+1}P(U=0) &= \frac{1}{k+1}P(U=k+1)\\
P(U=0) &= \frac{1}{k+2}P(U=k+1).
\end{aligned}$$

Now, we have proven our inductive hypothesis is true. In other words,

$$P(U=0)=\frac{1}{k+1}P(U=k)$$

is true for all $k>1$. So now, we can take the limit as $k\to\infty$

$$\begin{aligned}
\lim_{k\to\infty}P(U=0) &= \lim_{k\to\infty}\frac{1}{k+1}P(U=k)\\
&\le \lim_{k\to\infty}\frac{1}{k+1}\to 0,
\end{aligned}$$

so $P(U=0)=0$ as claimed. As a consequence, observe the we have shown that $P(U=0)=\frac{1}{k+1}P(U=k)$, and hence that $P(U=k)=(k+1)P(U=0)$ for all k. But we have also proven that $P(U=0)=0$, so

$$\begin{aligned}
P(U=k) &= (k+1)P(U=0)\\
&= (k+1)0\\
&= 0 \text{ for all } k.
\end{aligned}$$

□

Chapter 5

Properties of the Model

We now introduce some of the basic facts, definitions, lemmas, and theorems associated with the analysis of our model. Usually these facts have some clear connection with the classical continuous risk process ($U_t = u_0 + (1+\theta)\mu t - \sum_{i=1}^{N_t} X_i$). Often they are simply routine facts that would need to be established for any risk theory model. Sometimes they are the particular peculiarities encountered when using a zero-mean ($Z_i(\omega)$) process, and its sum ($S_n(\omega)$) in the context of mathematical risk theory. Occasionally, they are the specific results, derivable from our model, that provide a foundation for future chapters.

Given our model,

$$U_n(\omega) = u_0 + \theta\mu n - \sum_{i=1}^{n} Z_i(\omega) \text{ for all } \omega \in \Omega,$$

we introduce the following definitions.

The probability of ultimate ruin is used as a metric to compare different models, and to determine if an insurance business is deemed solvent with probability η. In the setting of a classical continuous risk process, this number is difficult to calculate explicitly. There are very few known examples where it can be calculated in a closed form. Also, this ignores the obvious shortcoming that we are concerning ourselves with the probability of ruin on an infinite time scale $(0,\infty)$, when it might be more reasonable to study the probability of ruin on a finite time scale, $(0,N)$. None-the-less, here is the definition of the probability of ultimate ruin adapted to the setting of our model.

Definition 5.1 (Probability of Ultimate Ruin)**.** The function $\Psi_{(0,\infty)}(u)$ for initial capital $u \geq 0$ is said to be the *probability of ultimate ruin* as a function of u if and only if

$$\Psi_{(0,\infty)}(u) = P(\omega \in \Omega : \exists n > 0 \text{ s.t. } u + \theta\mu n - S_n(\omega) < 0).$$

The probability of ruin by time T is also used as a metric of solvency in finite time. Naturally, it leads to the question of how to find specific values of T, and how to get a bound on the probability of ruin after time T.

Definition 5.2 (Probability of Ruin by Time T)**.** The function $\Psi_{(0,T]}(u)$ for $u \geq 0$ is said to be the *probability of ruin by time T* as a function of the initial capital u for $T \in \{1,2,\dots\}$ if and only if

$$\Psi_{(0,T]}(u) = P(\omega \in \Omega : \exists n \in (0,T] \text{ s.t. } u + \theta\mu n - S_n(\omega) < 0).$$

The probability of ruin at time T provides a natural way to construct a countable partition of the sample space Ω. We will construct that partition and use it often.

Definition 5.3 (Probability of Ruin at Time T)**.** The function $\Psi_{[T]}(u)$ for $u \geq 0$ is said to be the *probability of ruin at time T* as a function of the initial capital u for $T \in \{1,2,\dots\}$ if and only if

$$\Psi_{[T]}(u) = P(\omega \in \Omega : \forall m \in (0,T-1] \text{ s.t. } u + \theta\mu m - S_m(\omega) \geq 0 \;\&\; u + \theta\mu T - S_T(\omega) < 0).$$

The distribution of ruin probabilities is the principal object in our model. We will establish some results about its structure that will allow us to compare ruin probabilities at different time periods, and different initial capital levels.

Definition 5.4 (Distribution of Ruin Probabilities)**.** The infinite sequence $(\Psi_{[1]}(u), \Psi_{[2]}(u), \dots)$ is said to be the *Distribution of Ruin Probabilities* as a function of the initial capital u and n.

For our model we have $P(\omega \in \Omega : u_0 - S_0(\omega) \geq 0) = 1$ a.s., since, by assumption, there is almost surely no claim at time 0, ($S_0(\omega) = 0$ a.s.), and $u_0 \geq 0$. Let $\Psi_{[0]}(u) = 0$ to denote this fact. Furthermore, in our model we only allow the initial capital to be in whole units, i.e. $u_0 \in \{0,1,2,3,\dots\}$. Although after time $n = 0$, we certainly allow $U_n(\omega)$ to be any real number.

For a given fixed u_0 we might naturally wonder if the *Distribution of Ruin Probabilities* $(\Psi_{[1]}(u_0), \Psi_{[2]}(u_0), \dots)$ is a true probability distribution rather than a defective probability distribution. In particular, do we have $\sum_{n=1}^{\infty}\Psi_{[n]}(u_0)=1$, rather than only $\sum_{n=1}^{\infty}\Psi_{[n]}(u_0)\leq 1$? That question is answered in the next lemma.

5.1 Defective Probability of Ruin

Lemma 5.5. *Given $U_n(\omega)=u_0+\theta\mu n-S_n(\omega)$, and $u_0\geq 0$ fixed, then $\sum_{n=1}^{\infty}\Psi_{[n]}(u_0)\leq 1$.*

Proof. Let $A_0=\{\omega\in\Omega:\forall n>0 \text{ s.t. } u_0+\theta\mu n-S_n(\omega)\geq 0\}$.

Let $A_1=\{\omega\in\Omega:u_0+\theta\mu 1-S_1(\omega)<0\}$.

Let $A_2=\{\omega\in\Omega:\forall m\in(0,1] \text{ s.t. } u_0+\theta\mu m-S_m(\omega)\geq 0 \;\&\; u_0+\theta\mu 2-S_2(\omega)<0\}$.

Let $A_k=\{\omega\in\Omega:\forall m\in(0,k-1] \text{ s.t. } u_0+\theta\mu m-S_m(\omega)\geq 0 \;\&\; u_0+\theta\mu k-S_k(\omega)<0\}$.

Now we show that the A_k's form a partition of the sample space Ω. Suppose $i\geq 0$ and $j>i$ are given. Consider the case of $i=0$, and observe that $\omega_0\in A_0$ means $\forall n>0\; u_0+\theta\mu n-S_n(\omega_0)\geq 0$. Hence, $\omega_0\notin A_j\;\forall j\geq 1$ by definition of the sets A_j for $j\geq 1$. Consider the case $i\geq 1$ and $j>i$, then $\omega_0\in A_i$ would imply that $u_0+\theta\mu i-S_i(\omega_0)<0$, but then $\omega_0\notin A_j$ since for all A_j with $j>i$ we have $u_0+\theta\mu i-S_i(\omega_0)\geq 0$ Therefore, we have shown that $A_i\cap A_j=\emptyset$ for all $i\geq 0$, $j>i$. Now, let $\omega_1\in\Omega$ be arbitrary, and suppose $\forall n>0\; u_0+\theta\mu n-S_n(\omega_1)\geq 0$; then $\omega_1\in A_0$, as otherwise, there exists k for which $u_0+\theta\mu k-S_k(\omega_1)<0$ so $\omega_1\in A_k$. Hence $\Omega\subset\cup_{k=0}^{\infty}A_k$. Now we compute the probability of our partition of Ω

$$
\begin{aligned}
\Omega &= \cup_{k=0}^{\infty}A_k \text{ taking the probability of both sides}\\
P(\Omega) &= P(\cup_{k=0}^{\infty}A_k) \text{ giving}\\
1 &= \sum_{k=0}^{\infty}P(A_k) \text{ since the } A_k \text{ are disjoint}\\
1 &= P(A_0)+\sum_{k=1}^{\infty}P(A_k) \text{ now move } P(A_0) \text{ to the other side}\\
1-P(A_0) &= \sum_{k=1}^{\infty}P(A_k)\leq 1.
\end{aligned}
$$

□

So we only have $\sum_{n=1}^{\infty} \Psi_{[n]}(u_0) \leq 1$, and if $P(A_0) > 0$, then we have $\sum_{n=1}^{\infty} \Psi_{[n]}(u_0) < 1$. From the definition of A_0 we see that it defines the subset of the sample space Ω whose trajectories never go into ruin. So, in general, we usually have $P(A_0) > 0$. Also, from the definition of A_k we see that the *probability of ultimate ruin* $\Psi_{(0,\infty)}(u) = P(\cup_{k=1}^{\infty} A_k)$.

Notation 5.6. In the event that we are considering two different models A, and B, which differ only in their θ values, and/or their μ values, we may designate this fact by writing $u + \theta_A \mu_A n - S_n$, or $u + \theta_B n - S_n$, and sometimes append the letter A or B in the upper right of the appropriate symbol.

5.2 Ordering Relations

Now we define the infinite array of ruin probabilities. This allows us to establish ordering relations on the set of all $\Psi^A_{[k]}(j)$.

Definition 5.7 (Infinite Array of Ruin Probabilities)**.** For a given model A,
the following is called the *Infinite Array of Ruin Probabilities*

$$\begin{array}{ccccc}
\vdots & \vdots & \vdots & \vdots & \\
\Psi^A_{[1]}(3) & \Psi^A_{[2]}(3) & \Psi^A_{[3]}(3) & \Psi^A_{[4]}(3) & \dots \\
\Psi^A_{[1]}(2) & \Psi^A_{[2]}(2) & \Psi^A_{[3]}(2) & \Psi^A_{[4]}(2) & \dots \\
\Psi^A_{[1]}(1) & \Psi^A_{[2]}(1) & \Psi^A_{[3]}(1) & \Psi^A_{[4]}(1) & \dots \\
\Psi^A_{[1]}(0) & \Psi^A_{[2]}(0) & \Psi^A_{[3]}(0) & \Psi^A_{[4]}(0) & \dots
\end{array}$$

Notation 5.8. In the event that A is clear from context, then we omit it.

We prove the following fact about the row sum of the *Infinite Array of Ruin Probabilities*.

Lemma 5.9. *Given $U_n(\omega) = u_0 + \theta \mu n - S_n(\omega)$, with $u_0 = j$,
then $\sum_{n=1}^{\infty} \Psi_{[n]}(j) \leq 1$.*

Proof. Apply previous lemma with $u_0 = j$. □

We show the inequality relations for a given fixed column.

Lemma 5.10. *Given $U_n(\omega) = u_0 + \theta\mu n - S_n(\omega)$, suppose n is fixed,*

Then the following inequalities hold: if $k > j$ then $\Psi_{[n]}(k) \leq \Psi_{[n]}(j)$.

Proof.

Let $A_k = \{\omega \in \Omega : \forall m \in (0, n-1]$ s.t. $k + \theta\mu m - S_m(\omega) \geq 0$ & $k + \theta\mu n - S_n(\omega) < 0\}$.

Let $A_j = \{\omega \in \Omega : \forall m \in (0, n-1]$ s.t. $j + \theta\mu m - S_m(\omega) \geq 0$ & $j + \theta\mu n - S_n(\omega) < 0\}$.

Suppose $\omega_0 \in A_k$ is given; then $S_n(\omega_0) \geq k + \theta\mu n$ by the definition of A_k.

But, $k + \theta\mu n > j + \theta\mu n$ since $k > j$ by hypothesis, which means $S_n(\omega_0) > j + \theta\mu n$ so $\omega_0 \in A_j$, giving $A_k \subset A_j$. Therefore, $P(A_k) \leq P(A_j)$ by the monotonicity of measure.

So, applying the definition of *probability of ruin at time n* we see

$\Psi_{[n]}(k) = P(A_k) \leq P(A_j) = \Psi_{[n]}(j)$. □

Corollary 5.11. *Given $U_n(\omega) = u_0 + \theta\mu n - S_n(\omega)$, suppose n is fixed.*

Then the following inequality holds: for all $k > 0$ then $\Psi_{[n]}(k) \leq \Psi_{[n]}(0)$.

Proof. Apply previous lemma with $j = 0$. □

We establish the following results for the limit of a sequence, and the limit of the tail of a series.

Lemma 5.12. *Given $U_n(\omega) = u_0 + \theta\mu n - S_n(\omega)$, suppose $u_0 = j$ is fixed.*

Then $\lim_{n\to\infty} \Psi_{[n]}(j) = 0$, and $\lim_{n\to\infty} \sum_{k=n}^{\infty} \Psi_{[k]}(j) = 0$.

Proof. By previous lemma we have $\sum_{n=1}^{\infty} \Psi_{[n]}(j) \leq 1$.

So, by necessity, $\lim_{n\to\infty} \Psi_{[n]}(j) = 0$, and $\lim_{n\to\infty} \sum_{k=n}^{\infty} \Psi_{[k]}(j) = 0$. □

5.3 Uniform Convergence

Now, we establish some uniform convergence results.

Lemma 5.13. *Given $U_n(\omega) = u_0 + \theta\mu n - S_n(\omega)$, suppose $u_0 = u$ is allowed to vary. Then $\lim\limits_{n\to\infty} \Psi_{[n]}(u) = 0$, uniformly in u and $\lim\limits_{n\to\infty} \sum_{k=n}^{\infty} \Psi_{[k]}(u) = 0$ uniformly in u.*

Proof. Let $\varepsilon > 0$ be given, then $\exists N_1$ such that $\Psi_{[N_1]}(0) < \varepsilon$ since $\lim\limits_{n\to\infty} \Psi_{[n]}(0) = 0$, but by the previous lemma $\Psi_{[n]}(0) \geq \Psi_{[n]}(u)$ for all $u \geq 0$ so $\Psi_{[N_1]}(u) < \varepsilon$ for all $u \geq 0$ hence $\Psi_{[n]}(u)$ converges uniformly to 0. Likewise, $\exists N_2$ such that $\sum_{n=N_2}^{\infty} \Psi_{[n]}(0) < \varepsilon$, since $\lim\limits_{n\to\infty} \sum_{k=n}^{\infty} \Psi_{[k]}(0) = 0$. But, again, $\Psi_{[n]}(0) \geq \Psi_{[n]}(u)$ for all $u \geq 0$ so $\sum_{n=N_2}^{\infty} \Psi_{[n]}(u) < \varepsilon$ for all $u \geq 0$ hence $\sum_{k=n}^{\infty} \Psi_{[k]}(u)$ converges uniformly to 0. □

Using the uniform convergence result, we get a uniform bound on the convergence of

$$\lim_{n\to\infty} \sum_{k=n}^{\infty} \Psi_{[k]}(u).$$

Lemma 5.14. *Given $U_n(\omega) = u_0 + \theta\mu n - S_n(\omega)$ then, for every $\varepsilon > 0$ there exists an N such that $\sum_{k=N}^{\infty} \Psi_{[k]}(u) < \varepsilon$ for all $u \geq 0$.*

Proof. Let $\varepsilon > 0$ be given. Since $\sum_{k=n}^{\infty} \Psi_{[k]}(0)$ converges as $n \to \infty$, we know there exists an N such that $\sum_{k=N}^{\infty} \Psi_{[k]}(0) < \varepsilon$. But, by the previous uniform convergence result we also have

$$\sum_{k=N}^{\infty} \Psi_{[k]}(u) \leq \sum_{k=N}^{\infty} \Psi_{[k]}(0) < \varepsilon \text{ for all } u.$$

□

With this result in hand, it now makes sense to introduce the following augmented *Infinite Array of Ruin Probabilities* where we have reduced the number of columns from ∞ to N, and replaced the appropriate entry by the sum which is greater than or equal to the sum of the elements we have eliminated. The following is a representation of that effect.

$$
\begin{array}{ccccc}
\vdots & \vdots & \vdots & & \vdots \\
\Psi^A_{[1]}(3) & \Psi^A_{[2]}(3) & \Psi^A_{[3]}(3) & \dots & \sum_{k=N}^{\infty}\Psi_{[k]}(0) \\
\Psi^A_{[1]}(2) & \Psi^A_{[2]}(2) & \Psi^A_{[3]}(2) & \dots & \sum_{k=N}^{\infty}\Psi_{[k]}(0) \\
\Psi^A_{[1]}(1) & \Psi^A_{[2]}(1) & \Psi^A_{[3]}(1) & \dots & \sum_{k=N}^{\infty}\Psi_{[k]}(0) \\
\Psi^A_{[1]}(0) & \Psi^A_{[2]}(0) & \Psi^A_{[3]}(0) & \dots & \sum_{k=N}^{\infty}\Psi_{[k]}(0)
\end{array}
$$

For the sake of completeness, and to set the stage for further analysis, we verify that $\frac{U_n}{n}$ almost surely converges to $\theta\mu$.

Lemma 5.15. *Suppose $U_n(\omega) = u_0 + \theta\mu n - S_n(\omega)$, where $S_n(\omega) = Z_1(\omega) + Z_2(\omega) + \cdots + Z_n(\omega)$ and $Z_i(\omega)$ are i.i.d. with $E(Z_i(\omega)) = 0$. Then*

$$\lim_{n\to\infty} \frac{U_n(\omega)}{n} = \theta\mu \text{ a.s.}$$

Proof.

$$
\begin{aligned}
\lim_{n\to\infty} \frac{U_n(\omega)}{n} &= \lim_{n\to\infty} \left(\frac{u_0}{n} + \frac{\theta\mu n}{n} - \frac{S_n(\omega)}{n}\right) \\
&= \lim_{n\to\infty} \left(\frac{u_0}{n}\right) + \lim_{n\to\infty} (\theta\mu) - \lim_{n\to\infty} \left(\frac{S_n(\omega)}{n}\right) \\
&= 0 + \theta\mu - \lim_{n\to\infty} \left(\frac{S_n(\omega)}{n}\right) \\
&= 0 + \theta\mu - 0 \text{ a.s. by the SLLN} \\
&= \theta\mu.
\end{aligned}
$$

□

Notice that, in general, we would not expect to find a specific algorithm, $N(\varepsilon,\eta)$ such that, for a given ε, and a given η, this algorithm would finish in a finite number of steps, and produce an N that would satisfy the following:

$$n \geq N(\varepsilon,\eta) \Rightarrow |S_n(\omega)| < \varepsilon \text{ for all } \omega \in \Omega_{N(\varepsilon,\eta)} \subset \Omega \text{ with } P(\Omega_{N(\varepsilon,\eta)}) \geq 1-\eta.$$

And, even if we could construct such an algorithm $N(\varepsilon,\eta)$, we would not expect it to involve all (or even some) of the parameters u_0,θ,μ, or $Var(Z_i)$. The construction of such an algorithm, $N(\varepsilon,\eta)$, that involves the parameters, u_0,θ,μ, and $Var(Z_i)$, is a goal.

There exist various recursive algorithms (see for example Dickson) to compute estimates of the probability of ruin for a discrete risk process. But, as Asmussen notes in his book Ruin Probabilities, while much work has been done on these algorithms, they invariably lose track of the model parameters. That is, while they produce good numerical results, it is difficult to get qualitative information out of them. For example, one might like to know how the solution would change if some model parameter changed. We are not attempting to construct another algorithm that computes the probability of finite or ultimate ruin in a discrete risk setting, *and*, in addition, also preserves the model parameters. Rather, we let the idea of numerical precision become secondary, and accept (as Asmussen remarks) that it is easier to get good results for large n, or small n, but difficult to get good results for both small and large simultaneously.

We let the goal of obtaining qualitative information become primary; we are attempting to get formulas for upper bound estimates on the probability of ruin (or certain portions of that probability) that involve the model parameters explicitly. That is, we would like to understand dependence of this upper bound estimate on certain model parameters. This is a fairly natural thing to attempt, since the Lundberg upper bound, e^{-Ru} in the continuous setting does something similar to this.

Chapter 6

Probability of Ruin by the Kolmogorov Maximal Inequality

We introduce the following Lemma, to demonstrate that our model has the same limiting behavior as the classical Crámer-Lundberg model. That is, we show that as the initial capital goes to infinity then the probability of ruin goes to zero. Additionally, we actually get more than that. We get an expression for the probability of ruin in the "tail" that involves the model parameters themselves.

Lemma 6.1. *Suppose $U_n = u_0 + \theta\mu n - S_n$, where $S_n = Z_1 + Z_2 + \cdots + Z_n$, and Z_i are i.i.d. with $E(Z_i) = 0$. Then for every integer q there exists an initial capital u_0 s.t.*

$$P(Ultimate\ Ruin) = P(\exists n > 0 \ s.t\ U_n < 0) < \frac{1}{q}.$$

Proof. Given a subsequence $n_k \nearrow \infty$, $n_0 = 0$, and an integer l,

$$\begin{aligned}
P(\text{Ultimate Ruin}) &= P(\text{ ruin occurs in } (0, n_{l-1}]) + P(\cup_{k=l}^{\infty}(\text{ ruin occurs in } (n_{k-1}, n_k]) \\
&= P(\exists m \in [1, n_{l-1}] \text{ s.t } u_0 + \theta\mu m < S_m) + \sum_{k=l}^{\infty} P(\text{ ruin occurs in } (n_{k-1}, n_k]) \\
&\leq P(\max_{1 \leq m \leq n_l} |S_m| > u_0) + \sum_{k=l}^{\infty} \max_{n_k \leq m \leq n_{k+1}} |S_m| > \theta\mu n_k^{2/3} \\
&\leq P(\max_{1 \leq m \leq n_l} |S_m| > u_0) + \sum_{k=l}^{\infty} \max_{1 \leq m \leq n_{k+1}} |S_m| > \theta\mu n_k^{2/3}.
\end{aligned}$$

Based on the figure below, choose $u_0 = \theta\mu n_l^{(2/3)}$ and denote by $b_k = P(\max_{1 \leq m \leq n_k} |S_m| > \theta\mu n_k^{2/3})$ Then

$$P(\text{Ultimate Ruin}) \leq b_l + \sum_{k=l}^{\infty} b_k.$$

To complete the proof it suffices to show that for a properly chosen subsequence $\{n_k\}_{k=1}^{\infty}$, that the $\sum_k b_k$ converges, as then $\exists l = l(q)$ such that $b_l + \sum_{k=1}^{\infty} b_k < \frac{1}{q}$. To this end, by Kolmogorov's

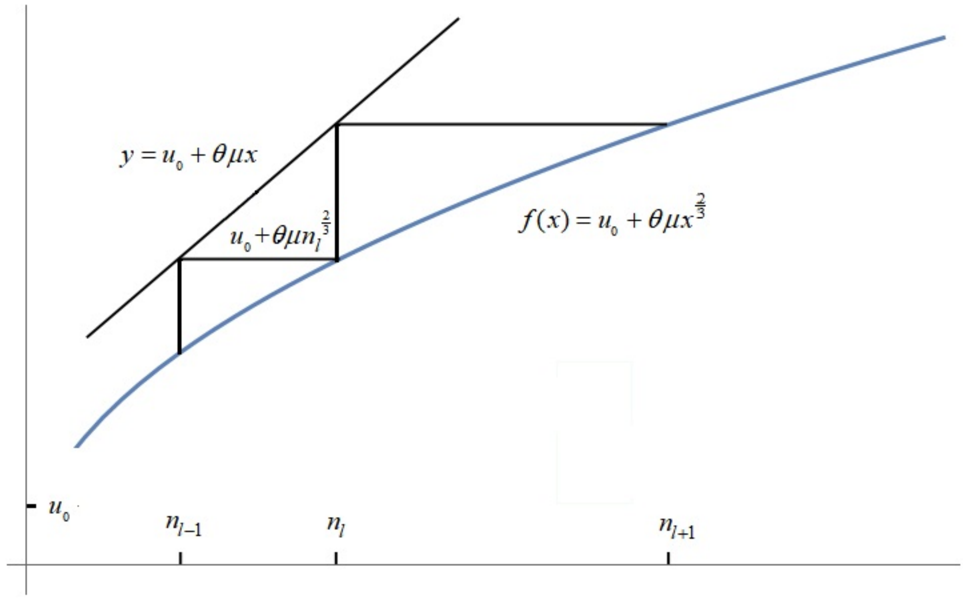

Figure 6.1
Approximating via $f(x) \sim x^{2/3}$

Maximal Inequality with $n_k = k^6$, $k = 1, 2, 3, \ldots,$

$$\begin{aligned} P(\max_{(1 \le m \le n_k)} |S_m| \ge n_k^{2/3}\theta\mu) &\le \frac{var(S_{n_k})}{[n_k^{2/3}\theta\mu]^2} \\ &\le \frac{n_k var(Z_i)}{n_k^{4/3}\theta^2\mu^2} \\ &\le \frac{var(Z_1)}{\theta^2\mu^2}\frac{1}{n_k^{1/3}} \text{ but } n_k = k^6 \\ &\le \frac{var(Z_1)}{\theta^2\mu^2}\frac{1}{(k^6)^{1/3}} \\ &\le \frac{var(Z_1)}{\theta^2\mu^2}\frac{1}{k^2}. \end{aligned}$$

Hence,

$$\sum_{k=1}^{\infty} b_k \le \frac{var(Z_1)}{\theta^2\mu^2}\sum_{k=1}^{\infty}\frac{1}{k^2} = \frac{var(Z_1)}{\theta^2\mu^2}\frac{\pi^2}{6} < \infty.$$

□

Corollary 6.2. *For every integer q there exists an initial capital u_0, and time horizon T such that*

$$P(\text{ Ruin by Time } T\ ) \leq \frac{1}{q}.$$

Proof. Choose $u_0 = T\sqrt{var(Z_i)}$ and $T = q$. Then by Kolmogorov's Maximal Inequality,

$$\begin{aligned}
P(\text{ Ruin by time T }) &= P(\exists\ 1 \leq m \leq T \text{ s.t. } u_0 + \theta\mu m < S_m) \\
&\leq P(\exists\ 1 \leq m \leq T \text{ s.t. } u_0 < S_m) = P(\max_{1\leq m\leq T} |S_m| > u_0) \\
&\leq \frac{var(S_T)}{u_0^2} = \frac{T var(Z_1)}{T^2 var(Z_1)} = \frac{1}{T} \leq \frac{1}{q}
\end{aligned}$$

□

Remark 6.3. The above shows that starting with the capital $T\sqrt{var(Z_1)}$ the probability of survival during the time interval $(0,T]$ is at least $1-\frac{1}{T}$.

Chapter 7

Finite versus Infinite Time Horizon for Ruin

We introduce the idea of "switching" from one model $U_n^A = u^A + \theta^A \mu^A - S_n$, called model "A," to another model $U_n^B = u^B + \theta^B \mu^B - S_n$, called model "B," with the same random zero-mean sum S_n, but perhaps different initial capital values, and different θ, and μ. We mean to compare two different alternatives, A, and B, and to conclude that the current standard of using the probability of ultimate ruin for a given starting capital, $\Psi_{(0,\infty)}(u)$, as a metric of solvency may have an unexpected side-effect in some cases. It would allow choosing a model B which requires a company to borrow less money u^B than the capital required for another model u^A. In particular, if the company can demonstrate that their choices $\theta^B \mu^B$ would lead to the same or smaller probability of ultimate ruin as the choice $\theta^A \mu^A$, then the various regulatory agencies might score the two models the same. But, surprisingly, model B may have a higher probability of ruin on some set of finite intervals, $(0,k]$ for some, potentially large, finite N, and all $k \in [1, N-1]$.

In keeping with the notation used in the classical Crámer-Lundberg continuous model, we usually write our model as

$$
\begin{aligned}
U_n(\omega) &= u_0 + \theta \mu n - \sum_{i=1}^{n} (X_i(\omega) - Y_i(\omega) + \theta \mu) \text{ for all } \omega \in \Omega \\
&= u_0 + \theta \mu n - \sum_{i=1}^{n} Z_i(\omega) \text{ for all } \omega \in \Omega,
\end{aligned}
$$

but actually, in our derivation we had

$$
\begin{aligned}
\theta \mu &= \frac{E(Y_i) - E(X_i)}{E(X_i)} E(X_i) \\
&= E(Y_i) - E(X_i) \\
&= |E(Y_i) - E(X_i)| \text{ since } E(Y_i) - E(X_i) > 0 \\
&= d(E(Y_i), E(X_i)) \text{ where } d \text{ is ordinary distance.}
\end{aligned}
$$

In our model, the distribution of the X_i's is relatively fixed and represents the expected distribution of the claims. The distribution, or at least the expectation of the distribution of the $E(Y_i)$ is relatively flexible, and represents the average premium amount our customers have signed a contract to pay.

Suppose that $E(Y_i)$ can assume any positive real number, then we can represent the distance between the different choices of $E(Y_i)$ and fixed choice of $\mu = E(X_i)$ as $d_1, d_2, d_3, \dots$ for some sequence of choices for the mean of the Y_i's where the only thing that changes is the mean of our premium distribution, and not its variance. If we had that situation then we could adopt the following notation and our model would read,

$$U_n(\omega) = u_0 + d_j n - \sum_{i=1}^{n} Z_i(\omega) \text{ for all } \omega \in \Omega.$$

We begin with a preliminary lemma. It provides a method for showing the existence another model B which we can choose to have less starting capital, u^B, than the starting capital, u^A of our original model A. But, the probability of ultimate ruin, $\Psi^B_{(0,\infty)}(u^B)$, for model B can be made to match or be lower than the probability of ultimate ruin, $\Psi^A_{(0,\infty)}(u^A)$, of our initial model A. This is done by choosing a larger distance, $d_B = d(E(Y_i^B), E(X_i))$, between the expected value of the premium size distribution for B, than the corresponding distance for model A.

Lemma 7.1. *Suppose $U_n^A(\omega) = u^A + d_A n - \sum_{i=1}^n Z_i(\omega)$, where $d_A = d(E(Y_i^A), E(X_i)) > 0$, with $\Psi^A_{(0,\infty)}(u^A) \in [0,1]$. Then for every initial capital $u^B < u^A$ there exists a model $U_n^B(\omega) = u^B + d_B n - \sum_{i=1}^n Z_i(\omega)$, where $d_B = d(E(Y_i^B), E(X_i)) > 0$, such that*

$$\Psi^B_{(0,\infty)}(u^B) \leq \Psi^A_{(0,\infty)}(u^A). \tag{7.2}$$

Proof. We are trying to show the existence of another model with less initial capital, $u^B < u^A$, but the same or lower probability of ultimate ruin.

Notation 7.3. The notation, $\bar{S}^{intercept}_{slope}$ designates subsets of Ω which contain all of the non-survival, or ruinous, trajectories with respect to a given *expected reserve line*.

We form the following sets:

$$\bar{S}_{d_A}^{u^A} = \{\omega \in \Omega : \exists n > 0 \text{ s.t. } u^A + d_A n - S_n(\omega) < 0\}.$$
$$\bar{S}_{d_A}^{u^B} = \{\omega \in \Omega : \exists n > 0 \text{ s.t. } u^B + d_A n - S_n(\omega) < 0\}.$$

Suppose $\omega_0 \in \bar{S}_{d_A}^{u^A}$ then $u^A + d_A N < S_N(\omega)$ for some N. But, by assumption, $u^B < u^A$ so $u^B + d_A N < u^A + d_A N < S_N(\omega)$ giving $\omega_0 \in \bar{S}_{d_A}^{u^B}$. This is true for all $\omega_0 \in \bar{S}_{d_A}^{u^A}$, which implies $\bar{S}_{d_A}^{u^A} \subset \bar{S}_{d_A}^{u^B}$, and therefore, $P(\bar{S}_{d_A}^{u^B}) \geq P(\bar{S}_{d_A}^{u^A})$ giving $\Psi_{(0,\infty)}^{A_1}(u^B) \geq \Psi_{(0,\infty)}^{A}(u^A)$. This inequality was anticipated, since the *expected reserve line* of our starting model, $u^A + d_A n$, is parallel to, and above, the *expected reserve line* $u^B + d_A n$.

Now, we seek to construct a model B_1 whose *expected reserve line* is above the *expected reserve line* of our starting model, for all n, but whose initial capital is, u^B, less than u^A. This simply means that we solve

$$u^B + d_1 \cdot 1 = u^A + d_A \cdot 1$$

for d_1. Which gives $d_1 = (u^A - u^B) + d_A$ which is non-zero by our model assumption that $d_A > 0$, meaning $E(Y_i) > E(X_i)$. Now, we form the following set,

$$\bar{S}_{d_1}^{u^B} = \{\omega \in \Omega : \exists n > 0 \text{ s.t. } u^B + d_1 n - S_n(\omega) < 0\},$$

and compare it to $\bar{S}_{d_A}^{u^A}$ the set from our original starting model.

Suppose $\omega_0 \in \bar{S}_{d_1}^{u^B}$ then $u^B + d_1 N < S_N(\omega)$ for some N. But, by construction, $d_1 = (u^A - u^B) + d_A \geq d_A$ so, supposing $N = 1$ then we would get $u^B + d_1 \cdot 1 = u^A + d_A$ hence $\omega_0 \in \bar{S}_{d_A}^{u^A}$. Suppose $N > 1$, then write N as $1 + (N-1)$, and observe that $u^B + d_1 N$ becomes $u^B + d_1(1 + (N-1)) = u^B + d_1 \cdot 1 + d_1(N-1) > u^A + d_A$, and gives $\omega_0 \in \bar{S}_{d_A}^{u^A}$ as well. Again, this is true for all $\omega_0 \in \bar{S}_{d_1}^{u^B}$, which implies $\bar{S}_{d_1}^{u^B} \subset \bar{S}_{d_A}^{u^A}$, and therefore, $P(\bar{S}_{d_A}^{u^A}) \geq P(\bar{S}_{d_1}^{u^B})$ giving $\Psi_{(0,\infty)}^{A}(u^A) \geq \Psi_{(0,\infty)}^{B_1}(u^B)$. But, we already had $\Psi_{(0,\infty)}^{A_1}(u^B) \geq \Psi_{(0,\infty)}^{A}(u^A)$. Putting all this together, and summarizing, we have shown the following:

$$\bar{S}^{u^B}_{d_1} \subset \bar{S}^{u^A}_{d_A} \subset \bar{S}^{u^B}_{d_A}$$
$$P(\bar{S}^{u^B}_{d_1}) \leq P(\bar{S}^{u^A}_{d_A}) \leq P(\bar{S}^{u^B}_{d_A})$$
$$\Psi^{B_1}_{(0,\infty)}(u^B) \leq \Psi^{A}_{(0,\infty)}(u^A) \leq \Psi^{A_1}_{(0,\infty)}(u^B).$$

The model, B_1 satisfies the conclusion of the lemma. □

We don't expect the model B_1 to be particularly sharp, in the sense that there might be another model B_∞ with a much shallower slope d_∞ that also satisfies the conclusion of the theorem. We now establish sufficient conditions to ensure that we can find the smallest slope d_∞. That is, we will be able to find a model B_∞ whose *expected reserve line* has the smallest slope d_∞, but whose probability of ruin is equal to the probability of ruin of model A. We will also be able to ensure that the *expected reserve line* of model B_∞ intersects the *expected reserve line* of model A. The basic idea of the theorem is to form a partial ordering (where the ordering is by set inclusion) of subsets of Ω, and then use the fact that a probability measure is continuous on a monotone class. In particular, if we have a nested sequence of increasing subsets of Ω,

$$\bar{S}_{d_1} \subset \bar{S}_{d_2} \subset \cdots \subset \bar{S}_{d_\infty} \text{ with } \cup_{k=1}^{\infty} \bar{S}_{d_k} = \bar{S}_{d_\infty} \text{ then } \lim_{k\to\infty} P(\bar{S}_{d_k}) = P(\bar{S}_{d_\infty}).$$

The following diagram gives an overview of the theorem. We have labeled the initial capital of model A as u_0^A. The other 3 models each have the same initial capital which we have labeled $u_0^{B_\infty}$. The model with the lowest ruin probability, and hence highest survival probability is model $U_n^{B_1}$. The model with the highest ruin probability, and hence lowest survival probability is model $U_n^{A_1}$. The original given model is U_n^A. The model with a lower initial capital than model U_n^A but none-the-less an equal probability of ruin is model $U_n^{B_\infty}$. All of the models with initial capital $u_0^{B_\infty}$ form a single monotone class.

Notation 7.4. We abbreviate u_0^A as u_A, u_0^B, as u_B, B_∞, as B, and d_∞ as d_B.

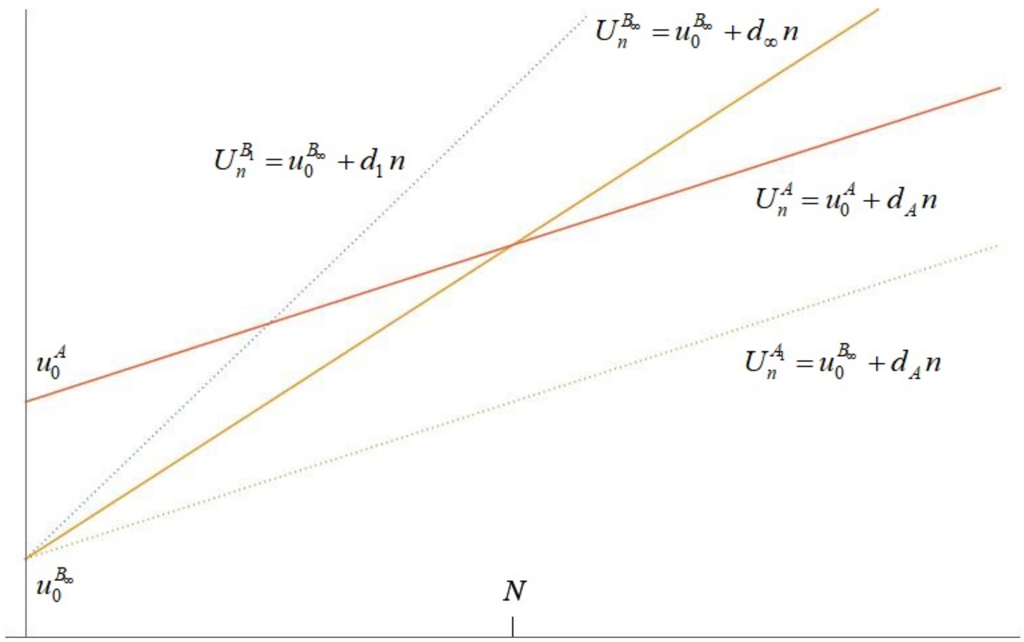

Figure 7.1
Expected Reserve Line for 4 Models

Theorem 7.5. *Given* $U_n^A(\omega) = u^A + d_A n - \sum_{i=1}^n Z_i(\omega)$ *where* $d_A = d(E(Y_i^A), E(X_i)) > 0$ *with* $\Psi^A_{(0,\infty)}(u^A) \in (0,1)$. *Suppose* X_i *is supported on* $(0,\infty)$, *then for every initial capital* $u^B < u^A$ *there exists a model* $U_n^B(\omega) = u^B + d_B n - \sum_{i=1}^n Z_i(\omega)$ *where* $d_B = d(E(Y_i^B), E(X_i)) > 0$ *such that*

$$\Psi^B_{(0,\infty)}(u^B) = \Psi^A_{(0,\infty)}(u^A). \tag{7.6}$$

Proof. Applying Lemma (7.2), and using the same notation as in the Lemma, and, in view of the fact that X_i is supported on $(0,\infty)$, we have the following sharp inequalities:

$$0 < P(\bar{S}^{u^B}_{d_1}) < P(\bar{S}^{u^A}_{d_A}) < P(\bar{S}^{u^B}_{d_A}) < 1$$
$$0 < \Psi^{B_1}_{(0,\infty)}(u^B) < \Psi^{A}_{(0,\infty)}(u^A) < \Psi^{A_1}_{(0,\infty)}(u^B) < 1.$$

Notice that we have a monotone class with

$$\bar{S}^{u^B}_{d_1} \subset \bar{S}^{u^B}_{d_\alpha} \subset \bar{S}^{u^B}_{d_A} \quad \text{for } d_A \le d_\alpha \le d_1,$$

where our set "in the middle" is $\bar{S}^{u^B}_{d_\alpha}$ and not $\bar{S}^{u^A}_{d_A}$. Note, with our notation, there are two clear ways to generate a monotone class: either keep the slope fixed, and allow the initial capital to vary, or keep the the initial capital fixed, and allow the slope to vary. We have chosen the later method, but done so in such a way that the "top" of the chain is a subset of the sample space with probability of ruin lower than the probability of ruin of model A, and the "bottom" of the chain is a subset of the sample space with probability of ruin greater than the probability of ruin of model A. Also, our assumption on the support of X_i ensures that $P(\bar{S}^{u^B}_{d_{\alpha_1}}) \neq P(\bar{S}^{u^B}_{d_{\alpha_2}})$ for $d_{\alpha_1} \neq d_{\alpha_2}$.

Now, let $g(d_\alpha) = P(\bar{S}^{u^B}_{d_\alpha}) - P(\bar{S}^{u^A}_{d_A})$, and restrict g to the monotone class just defined. Notice that $g(d_1) > 0$, and $g(d_A) < 0$. But, $P(\cdot)$ is continuous on monotone classes, and by the intermediate value theorem (IVT) g assumes all of its intermediate values. Hence there exists d_B with $P(\bar{S}^{u^B}_{d_B}) = P(\bar{S}^{u^A}_{d_A})$, which gives $\Psi^B_{(0,\infty)}(u^B) = \Psi^A_{(0,\infty)}(u^A)$ as claimed. □

The theorem can essentially be thought of as establishing conditions under which we can "charge ourselves" a "claim" of size $[(u_A + d_A n) - (u_B + d_\infty n))] > 0$ at time $0 \leq n < N$, and then completely mitigate its impact on the ultimate survival probability by immediately increasing the gross premium rate from d_A to d_∞. That is, we can "switch" to a different model with a lower initial capital, but the same ultimate probability of ruin. Unfortunately, though, we may have dramatically increased our probability of ruin in the short term. In the next few chapters we will see an example which exhibits this phenomenon. The main point is that relying entirely on the probability of ultimate ruin as a metric of solvency is missing crucial information about survival in finite time. Models with the same ultimate survival probability can have very different finite survival profiles. In Chapter 9, we will develop the machinery necessary to see this effect more clearly.

Chapter 8

Probability of Ruin by Stopping a Martingale

We show how martingale methods can be applied to our model.

Let $U_n = u + \sum_{i=1}^{n} Y_i - \sum_{i=1}^{n} X_i$, $E(Y_i) = (1+\theta)E(X_i)$, $\theta > 0$ and $X_i \sim X_1, Y_i \sim Y_1$ are independent and identically distributed nonnegative random variables with $0 < E(X) < \infty$.

Theorem 8.1. *Suppose* $\exists r \neq 0$ *such that* $E(e^{r(X-Y)}) = 1$.
Then $r > 0$ *and*

$$P(ultimate\ ruin) \leq e^{-ru}. \tag{8.2}$$

Proof. Let $W_i = X_i - Y_i$ so $W \sim X - Y$ and $E(W) = -\theta E(X) < 0$. By Jensen's inequality, for any convex $\Phi(X)$, $\Phi(E(W)) \leq E(\Phi(W))$, so taking $\Phi(x) = e^{rx}$ we have $e^{rE(W)} \leq E(e^{rW}) = 1$ and thereby r must be positive. Now for any $a, b > 0$ consider $S_n = \sum_{i=1}^{n} W_i$ and

$$\begin{aligned} P_{a,b} &= P(\exists n\ S_n > a \text{ and } \min_{1\leq i\leq n} S_i > -b) \\ &= P(S_n \text{ crosses } a \text{ before crossing } -b) \end{aligned}$$

and

$$\begin{aligned} P_{b,a} &= P(\exists n\ S_n \leq -b \text{ and } \max_{1\leq i\leq n} S_i < a) \\ &= P(S_n \text{ crosses } -b \text{ before crossing } a). \end{aligned}$$

Define a stopping time N by

$$\begin{aligned} N &= min\{n \mid S_n \geq a \text{ or } S_n \leq -b\} \\ &= \text{smallest } n \text{ s.t. } S_n \text{ exits the interval } (a,b) \text{ and} \end{aligned}$$

$N = \infty$ in the case that no such n exists.

Then $M_n = e^{rS_n}$ is a martingale because

$$\begin{aligned} E\left[e^{rS_n} \mid W_1, W_2, \ldots, W_{n-1}\right] &= E\left[e^{rS_{n-1}+uW_n} \mid W_1, \ldots, W_{n-1}\right] \\ &= e^{rS_{n-1}} E\left[e^{uW_n} \mid W_1, \ldots, W_{n-1}\right] \\ &= M_{n-1}, \end{aligned}$$

as the conditional expectation becomes expectation due to the independence of $\{W_i\}$, and by assumption equals 1. It is standard to check that $P(N < \infty) = 1$ (Durrett [17]) and

$$EM_n = EM_N = 1.$$

Now,

$$\begin{aligned} 1 &= E\left[e^{rS_N} \mid S_N \geq a\right] P_{a,b} + E\left[e^{rS_N} \mid S_N \leq -b\right] P_{b,a} \\ &\geq e^{ra} P_{a,b}, \end{aligned}$$

giving

$$P_{a,b} \leq e^{-ra} \text{ for any } b > 0. \tag{8.3}$$

Taking $b = k \in \mathbb{N}$,

$$\begin{aligned} P_{a,k} &= P(S_n \text{ crosses } a \text{ before } -k) \\ &= P(\cup_{n=1}^{\infty} \{S_n \geq a \, \cap \min_{1 \leq i \leq n} S_i > -k\}), \end{aligned}$$

and setting

$$\begin{aligned} C_{n,k} &= \cup_{n=1}^{\infty} \{S_n \geq a\} \cap \{\min_{1 \leq i \leq n} S_i > -k\} \\ &= \cup_{n=1}^{\infty} A_n \cap B_{n_k},\ B_{n_k} \subset B_{n_{k+1}} \end{aligned}$$

we have $C_{n_k} \nearrow$ in k. By the continuity of $P(\cdot)$ for monotone sequences

$$
\begin{aligned}
\lim_{k\to\infty} P(C_{n_k}) &= P(\cup_{k=1}^{\infty} C_{n_k}) \\
&= P(\cup_{k=1}^{\infty} \cup_{n=1}^{\infty} A_n \cap B_{n_k}) \\
&= P(\cup_{n=1}^{\infty} A_n \cap (\cup_{k=1}^{\infty} B_{n_k})) \\
&= P(\cup_{n=1}^{\infty} A_n \cap \{\min_{1\le i\le n} S_i > -\infty\}) \\
&= P(\cup_{n=1}^{\infty} A_n) \text{ since } P(\min_{1\le i\le n} S_i > -\infty) = 1.
\end{aligned}
$$

Consequently

$$
\begin{aligned}
\lim_{k\to\infty} P_{a,k} = \lim_{k\to\infty} P(C_{n_k}) \\
&= P(\cup_{n=1}^{\infty} A_n) \\
&= P(\exists n \;\; S_n \ge a) \\
&= P(S_n \text{ ever crosses } \; a),
\end{aligned}
$$

and by (8.3),

$$
P(S_n \text{ ever crosses } \; a) \le e^{-ra}. \tag{8.4}
$$

Turning to the proof of (8.2),

$$
\begin{aligned}
P(ultimate\ ruin) &= P(\exists n\, U_n < 0) \\
&\le P(\exists n\, U_n \le 0) \\
&= P(\exists n\, u \le \sum_{i=1}^{n} (X_i = Y_i)) \\
&= P(\exists n\, u \le \sum_{i=1}^{n} W_i) \\
&= P(\exists n \le S_n)
\end{aligned}
$$

$$= P(S_n \text{ ever crosses } u)$$
$$\le e^{-ru} \text{ by (8.4) with } a = u.$$

$\square$

Example 8.5 (Exponential case).

Let the claim size $X \sim$ exponential with mean μ and premium size $Y \sim$ exponential with mean $\lambda = (1+\theta)\mu$

Then the condition

$$E(e^{r(X-Y)}) = 1 \tag{8.6}$$

reads

$$M_{X-Y}(r) = \frac{\frac{1}{\mu}}{\frac{1}{\mu} - r} \cdot \frac{\frac{1}{\lambda}}{\frac{1}{\lambda} + r} = 1.$$

Solving for r we obtain

$$r = \frac{\theta}{(1+\theta)\mu} \pm \frac{\sqrt{(2+\theta)^2 - (\frac{4}{\mu})^2}}{\mu(1+\theta)}.$$

Remark 8.7. For solutions r to be well-defined and positive we must have

1. $(2+\theta)^2 - (\frac{4}{\mu})^2 \ge 0$,

   which is always satisfied for $\mu \ge 1$ whereas for $0 < \mu < 1$, $\theta \ge 2(\frac{1}{\mu} - 1) > 0$.

2. r^+ is always a positive solution; however, r^- can also be a solution if $0 < r^-$, or equivalently $2(\frac{1}{\mu} - 1) \le \theta \le \frac{1}{\mu^2} - 1$ with $0 < \mu < 1$.

   Consequently $r = r^+$ if $\theta > \frac{1}{\mu^2} - 1$ and $r = r^-$ if $2(\frac{1}{\mu} - 1) \le \theta \le \frac{1}{\mu^2} - 1$.

3. The first term $\frac{\theta}{(1+\theta)\mu}$ in r is identical to the adjustment coefficient R in the continuous model $U(t) = u + ct - \sum_{i=1}^{N(t)} X_i$ with claim size exponential with mean μ.

Example 8.8 (Binomial case). Let $Y \sim Bin(p_Y, k)$, and $X \sim Bin(p_X, k)$ and $p_X < p_Y$, which gives $\theta = \frac{kp_Y - kp_X}{kp_X} = \frac{p_Y}{p_X} - 1$ Then the condition (8.6) reads,

$$M_{X-Y}(r) = [(p_X e^r + 1 - p_X)(p_Y e^{-r} + 1 - p_Y)]^k = 1.$$

Solving for r gives

$$r = \ln\left[\frac{(1-p_X)p_Y}{(1-p_Y)p_X}\right],$$

and (8.2) reads

$$P(ultimate\ ruin) \leq e^{-ru} = \left[\frac{(1-p_Y)p_X}{(1-p_X)p_Y}\right]^u$$

$$\alpha = \frac{(1-p_Y)p_X}{(1-p_X)p_Y} \text{ where } 0 < \alpha < 1 \text{ given } p_X < p_Y.$$

For example, for $p_X = .5$, $p_Y = .67$, $\theta = .34$ (safety loading of 34%) $\alpha = .492$ or

$$P(ultimate\ ruin\ with\ initial\ capital\ u) < \left(\frac{1}{2}\right)^u.$$

Namely, every extra dollar of initial capital halves the probability of ultimate ruin!

Chapter 9

Probability of Ruin via Donsker Invariance Principle

In this chapter we will apply the celebrated Donsker's Invariance Principle [15] for Brownian Motion to derive the probability of ruin for the risk reserve process for both finite and infinite time horizon under the mild natural assumption of finite second moments for the claims X, and the premiums Y.

Typically the Invariance Principle is stated on the Space of Continuous Functions $C[0,1]$ [4], however our considerations require $C[0,\infty)$ and therefore the formulation here follows [28] pp. 66-71. To this end consider $\{\xi_j\}_{j=1}^{\infty}$ i.i.d. with $E(\xi_j)=0$, $Var(\xi_j)<\infty$, and $S_k=\sum_{j=1}^{k}\xi_j$, $S_0=0$. Define a continuous process $\{Y_t, t\geq 0\}$ obtained by linear interpolation

(9.1)
$$Y_t = S_{\lfloor t\rfloor} + t - \lfloor t\rfloor \xi_{\lfloor t\rfloor+1},\; t\geq 0,$$

where $\lfloor t\rfloor$ is the greatest integer less than or equal to t. Scaling appropriately both time and space consider a sequence of processes, $n\in\mathbb{N}$

(9.2)
$$X_t^n = \frac{1}{\sigma\sqrt{n}} Y_{nt},\; t\geq 0.$$

Donsker's Invariance Principle (Theorem 9.20, [28]). *Let $(\Omega,\mathscr{F},P)$ be the probability space on which $\{\xi_j\}_{j=1}^{\infty}$ is defined. Let P_n be the probability induced by $\{X_t^n\}_{t\geq 0}$ given by* (9.2) *on $(C[0,\infty),\mathscr{B}(C[0,\infty))$.*
Then

P_n converges weakly to the Wiener measure.

The Wiener measure is the probability measure on $C[0,\infty)$ under which the coordinate mapping process $W_t(\circ)\equiv w(t)$ is the standard Brownian motion $B(t)$. Equivalently,

(9.3)
$$X_t^n \Longrightarrow B(t) \text{ as } n\to\infty,$$

where "$\Longrightarrow$" stands for convergence in distribution.

Remark 9.4. Since $(t-\lfloor t\rfloor)\xi_{\lfloor t\rfloor+1} \to 0$ in probability, (9.3) can be stated, due to (9.1) and (9.3), as

$$\frac{S_{\lfloor nt\rfloor}}{\sigma\sqrt{n}} \Longrightarrow B(t) \text{ as } n \to \infty, \tag{9.5}$$

where, again, $\Longrightarrow$ is convergence in distribution.

Turning to the risk reserve process

$$U_n = u + \theta\mu n - S_n,\ n \in \mathbb{N},$$

where $S_n = \sum_{i=1}^{n} Z_i$, $Z_i = X_i - Y_i + \theta\mu$, $E(Z_i) = 0$, $E(X_i) = \mu$, $E(Y_i) = (1+\theta)\mu$, X_i, Y_i are i.i.d. and X_i independent of Y_i. Denote the variance of Z_i by

$$\begin{aligned}\sigma^2 &= Var(X_i - Y_i + \theta\mu)\\ &= Var(X_i) + Var(Y_i)\\ &= \sigma_X^2 + \sigma_Y^2,\end{aligned}$$

and set $\sigma = \sqrt{\sigma_X^2 + \sigma_Y^2}$.

First we rename n as t and write

$$U(t) = u + \theta\mu t - S_t,\ t \in \mathbb{N}.$$

Then we extend the above to

$$U(t) = u + \theta\mu t - S_t,\ \ t \geq 0, \tag{9.6}$$

where S_t is obtained by linear interpolation

$$S_t = S_{\lfloor t\rfloor} + (t-\lfloor t\rfloor)Z_{\lfloor t\rfloor+1}.$$

Now we rescale the time by a factor $\frac{1}{n}$, and consider a family of processes for $n \in \mathbb{N}$ corresponding to (9.6).

$$U^n(t) = \sqrt{n}u + \frac{1}{\sqrt{n}}\theta\mu tn - S_{\lfloor nt \rfloor} \tag{9.7}$$

$$= \sqrt{n}\left[u + \theta\mu t - \frac{S_{\lfloor nt \rfloor}}{\sqrt{n}}\right]. \tag{9.8}$$

Lemma 9.9. *Given the above notation and assumptions,*

$$\frac{U^n(t)}{\sqrt{n}} \Longrightarrow u + \theta\mu t - \sigma B(t).$$

Proof. Apply Donsker's Invariance Principle to

$$\sigma\frac{S_{\lfloor nt \rfloor}}{\sigma\sqrt{n}} \Longrightarrow \sigma B(t).$$

□

Now, let

$$T_n = inf\{t > 0 \mid U^n(t) \leq 0\}$$

and observe that by (9.7) for all $n \in \mathbb{N}$

$$T_n = inf\{t > 0 \mid u + \theta\mu t - \frac{S_{\lfloor nt \rfloor}}{\sqrt{n}} \leq 0\}. \tag{9.10}$$

Define

$$T = inf\{t > 0 \mid u + \theta\mu t - \sigma B(t) \leq 0\}. \tag{9.11}$$

Before stating our main result we need the following fact concerning Brownian motion:

Fact 9.12 ([28], p. 196)**.** *Given Brownian motion with drift $W(t) = ct + B(t)$. Let τ be the hitting time of the barrier $a \neq 0$ given $c \neq 0$. Then the density*

$$f_\tau(t) = \frac{a}{\sqrt{2\pi t^3}}e^{-\frac{(a-ct)}{2t}}, \quad t \geq 0 \tag{9.13}$$

is Inverse Gaussian and

$$P(\tau \le t) = \int_0^t f_\tau(s)\,\mathrm{d}s$$
$$P(\tau < \infty) = e^{ca-|ca|},$$

where

$$e^{ca-|ca|} = \begin{cases} 1, & \textit{if a and c have the same sign}, \\ e^{2ca} < 1, & \textit{otherwise.} \end{cases}$$

The case $P(\tau < \infty) = 1$ corresponds to a drift c pointing toward the barrier, whereas $P(\tau < \infty) < 1$ corresponds to a drift c pointing in the direction opposite to the barrier, and signifies the fact that the density $f_\tau(t)$ is defective, i.e. $P(\tau = \infty) = 1 - e^{2ca}$.

Theorem 9.14.

$$\lim_{n\to\infty} P(T_n \le t) = P(T \le t) \tag{9.15}$$
$$= \int_0^t \frac{u}{\sqrt{2\pi s^3}} e^{-\frac{(u+\theta\mu s)^2}{2s\sigma^2}}\,\mathrm{d}s. \tag{9.16}$$

Proof. The main benefit of the functional Central Limit Theorem, such as Donsker's Invariance Principle, is that continuous functionals of processes converging in distribution (here $U^n \Longrightarrow u + \theta\mu t + \sigma B(t)$) also converge in distribution. The random variables T_n, T, defined by (9.10) and (9.11) respectively, both satisfy $T_n \Longrightarrow T$ because both are derived from a continuous map on the space of trajectories on $C[0,\infty)$. Namely,

$$\tau : C[0,\infty) \mapsto \mathbb{R}$$

defined by $\tau(x) = inf\{t > 0 \mid X(t) \le 0\}$ if non-empty, and $+\infty$ otherwise, is measurable, and almost surely continuous with respect to $u + \theta\mu t + \sigma B(t)$. This establishes the first equality of (9.16).

To show the second equality in (9.16) notice that

$$\begin{aligned} T &= inf\{t > 0 \mid u + \theta\mu t - \sigma B(t) \le 0\} \\ &= inf\{t > 0 \mid \frac{u}{\sigma} \le -\frac{\theta\mu}{\sigma} + B(t)\}, \end{aligned}$$

whence by Fact (9.13) with $a = \frac{\mu}{\sigma}$, $c = -\frac{\theta\mu}{\sigma}$ the proof is complete. □

Corollary 9.17.

$$P(ultimate\ ruin) = P(T < \infty) = e^{-\frac{2\theta\mu}{\sigma^2}u}.$$

Note, the limiting distribution T of T_n is considered as corresponding to T for the original process (9.6). Similar analysis for the classical Lundberg model $u + ct - \sum_{i=1}^{N(t)} X_i$, $N(t)$ a Poisson process was done by Iglehart [26] on the space $D[0,\infty)$ of right continuous left limit functions.

As we indicated at the end of chapter 7, we did not yet have the machinery to fully illustrate Theorem (7.5); now we do. By way of review, Theorem (7.5) established a disconcerting result: if we say that two different processes are equivalent when they have the same probability of ultimate ruin - which is the standard practice - then we may call two different processes equivalent, or interchangeable, when, in fact, they have significantly different finite ruin probabilities.

Example 9.18. Let the expected value of the claim size distribution be 1, that is, $\mu = 1$. And let the standard deviation of the claims plus premium be 1, that is, $\sigma = 1$. Suppose we have two different models A and B with the initial capital for A being $u = 24$ and the safety loading $\theta = 0.11$, whereas for B we have initial capital, 48, and safety loading 0.055. So by the above corollary we have

$$P(ultimate\ ruin\ A) = e^{-2(0.11)24} = 0.00509 = e^{-2(0.055)48} = P(ultimate\ ruin\ B).$$

That is, model A and model B both satisfy the same standard of having approximately a $\frac{1}{2}$ of 1% probability of ultimate ruin. But, the probability of ruin in finite time for model A and model B correspond to very different Inverse Gaussian distributions.

For model A the probability of ruin by time t corresponds to the following integral:

$$\int_0^t \frac{24}{\sqrt{2\pi s^3}} e^{-\frac{(24+0.11s)^2}{2s}} \, \mathrm{d}s.$$

We then see that after 4 years, or 1460 days, the probability of ruin in the tail for model A is negligible. That is, we have already equaled our ultimate ruin probability since

$$\int_0^{1460} \frac{24}{\sqrt{2\pi s^3}} e^{-\frac{(24+0.11s)^2}{2s}} \, \mathrm{d}s = 0.00509.$$

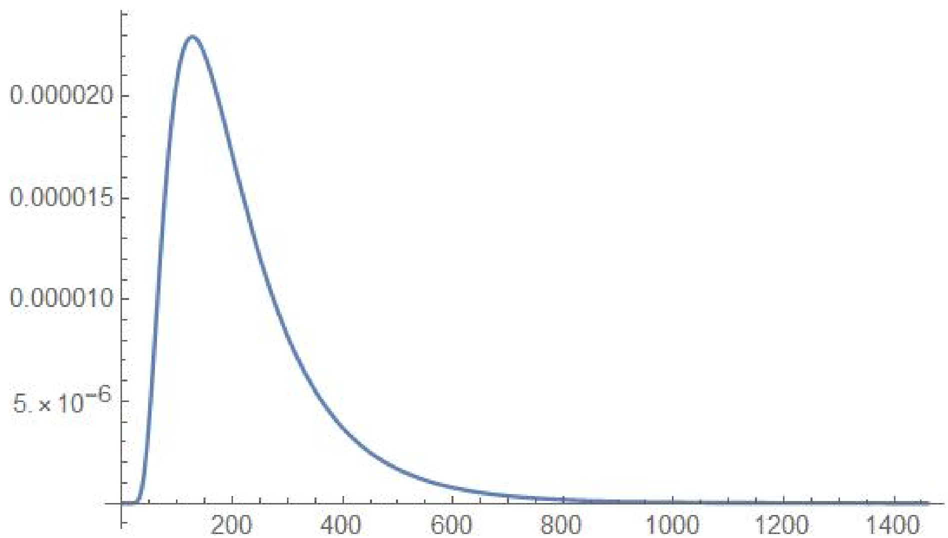

Figure 9.1
Distribution of finite ruin for A

But, for B it takes 4 times as long for the tail ruin probability to become negligible. For model B, see the figure below, the probability of ruin by time t corresponds to the following integral:

$$\int_0^t \frac{48}{\sqrt{2\pi s^3}} e^{-\frac{(48+0.055s)^2}{2s}} \, \mathrm{d}s.$$

And, after 4 years we still have

$$\int_0^{1460} \frac{48}{\sqrt{2\pi s^3}} e^{-\frac{(48+0.055s)^2}{2s}} \, \mathrm{d}s = 0.00447.$$

But, by year 20 we have

$$\int_0^{5840} \frac{48}{\sqrt{2\pi s^3}} e^{-\frac{(48+0.055s)^2}{2s}}\, \mathrm{d}s = 0.00509.$$

So, by year 20 the finite ruin probability for model B finally equals its ultimate ruin probability. So, model A is significantly more risky than model B in the first 4 years, but the current standard of regarding them as being equivalent is failing to highlight this distinction.

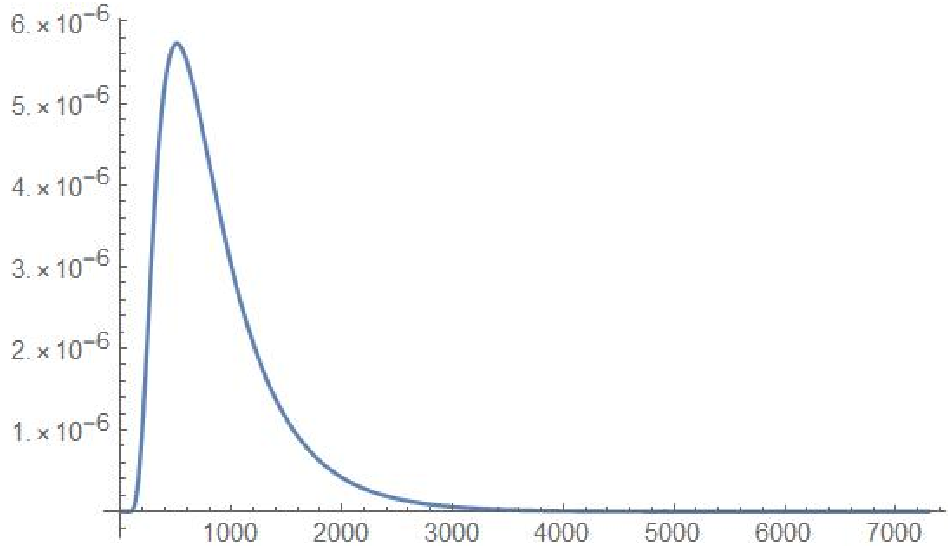

Figure 9.2
Distribution of finite ruin for B

Chapter 10

Probability of Ruin via Large Deviation Principle

This section is concerned with the derivation of the upper bound for the probability of ruin on the interval $[n,\infty)$, which we will refer to as *tail ruin probability*. Our arguments are based on the *rate function* - a key ingredient of the Large Deviation Principle (LDP), so for the sake of completeness we discuss the relevant basic facts.

Large deviations refer to a class of results which describe how probabilities of atypical events away from typical events decay to zero. It turns out that very often $P(A_n) \sim \exp^{-\alpha n}$ for large n where the constant $\alpha > 0$ is directly computable. One of the most important rates is concerned with the Strong Law of Large Numbers (SLNN), and reads $P(\frac{X_1+\cdots+X_n}{n} \in A) \sim e^{-\alpha n}$, whenever $E(X_i) = \mu \notin A$. The first large deviation result dates back to the late 1930's, and is attributed to Crámer [7]. The most significant developments came in the 1980's with the pioneering books of Varadhan [30] for a large class of Stochastic processes(including Markov diffusions), Freidlin and Wentzell [21] for dynamical systems(perturbed by Ito processes) and Freidlin [20] for PDEs. More recent monographs on LDP are Dembo [9] and Hollander [10]. Since our analysis is related to Crámer's LDP we state it without proof.

Crámer's Large Deviation (Th. I.4, [10])**.**

Let (X_i) be i.i.d with the moment generating function $\Phi(t) = E(e^{tX_1}) < \infty$, $t \in \mathbb{R}$, and $S_n = \sum_{i=1}^n X_i$. Then for any $a > E(X_1)$

$$\lim_{n\to\infty} \frac{1}{n} \log P(S_n \geq an) = -I(a)$$

where

$$I(a) = \sup_{t\in\mathbb{R}} [at - \log \Phi(t)]$$

is the rate function.

Remark 10.1. The above result has a straightforward extension to $\Phi(t) < \infty$ for t from some subset of $\mathbb{R}$. Key properties of the *rate function* $I(\cdot)$ are as follows: $0 \le I(\cdot)$, $I(\cdot)$ is convex, $I(E(X_i)) = I(\mu) = 0$, $I(x)$ may assume $+\infty$, $I(x) \searrow$ for $x < \mu$, $I(x) \nearrow$ for $x > \mu$. $I(x)$ is called a convex conjugate or the Legendre Transform of the convex function $\log \Phi(t)$.

Upper Bound Lemma. *Assume $E(e^{tX}) < \infty$ for $t \ge 0$. Then for all $x > E(X) = \mu$*

$$P(S_n \ge nx) \le e^{-nI(x)}. \tag{10.2}$$

Proof. By Markov's inequality,

$$\begin{aligned} P(S_n \ge nx) &= P(e^{tS_n - txn} \ge 1) \\ &\le E(e^{tS_n - txn}) \\ &= e^{-txn}(M_X(t))^n \\ &= e^{-n(xt - \log M_x(t))}. \end{aligned}$$

Now, since t is arbitrary, we can optimize this upper bound by maximizing $h(t) = xt - \log M_X(t)$ over t. We have

$$h'(t) = x - \frac{M_X'(t)}{M_X(t)}|_{t=0} = x - \mu > 0,$$

and therefore $h(t) > 0$ in some vicinity of $t = 0$, because $h(0) = 0$. This in turn, since $h(t)$ is concave down, shows that $h(t)$ has a unique strictly positive max that can be obtained by solving $x - \frac{M_X'(t)}{M_X(t)} = 0$ for some $t = t(x)$ and therefore $\max_t I(x) = x\,t(x) - \log(M_X(t(x))) = I(x)$. □

Theorem 10.3. *Consider $U_n = u + \theta\mu n - S_n$, where $S_n = \sum_{i=1}^n Z_i$, Z_i are i.i.d. $E(Z_i) = 0$. Then we have the following upper bounds for ruin probability:*

$$P(\textit{ruin in } [n, 2n]) \le \frac{e^{-nI(\theta\mu + \frac{u}{2n})}}{1 - e^{-I(\theta\mu + \frac{u}{2n})}} \tag{10.4}$$

$$P(\textit{ruin in } [n, \infty) \le \frac{e^{-nI(\theta\mu)}}{1 - e^{-I(\theta\mu)}}. \tag{10.5}$$

Proof. We have

$$
\begin{aligned}
P(ruin\ in\ [n,2n] &\leq P(\exists n \leq k \leq 2n \mid u+\theta\mu k - S_k < 0) \\
&= P(\cup_{k=n}^{2n} S_k > k\theta\mu + u) \\
&\leq \sum_{k=n}^{2n} P(S_k > k\theta\mu + u) \\
&\leq \sum_{k=n}^{2n} e^{-kI(\theta\mu+\frac{u}{k})} \text{ by (10.2)} \\
&\leq \sum_{k=n}^{2n} e^{-kI(\theta\mu+\frac{u}{2n})} \\
&\leq \frac{e^{-nI(\theta\mu+\frac{u}{2n})}}{1-e^{-I(\theta\mu+\frac{u}{2n})}}.
\end{aligned}
$$

The proof of (10.5) is straightforward and is omitted. □

In the example below we illustrate how the upper bound (10.4) can be used to estimate the ruin probability on a finite time block, given that the probability of ultimate ruin has a known upper bound.

Example 10.6. Consider our previous example 8.5 where the claim size $\sim$ exponential with mean μ and premium size $\sim$ exponential with mean $(1+\theta)\mu$. Then

$$P(ruin\ in\ [0,\infty)) \leq e^{-ru},$$

where

$$r = \frac{\theta + \sqrt{(2+\theta)^2 - (\frac{4}{\mu})^2}}{(1+\theta)\mu}.$$

Let us now choose $\mu = 25$, $\theta = .2$ and $u = 75$.

Then we have

$$P(Ultimate\ Ruin) \leq e^{-ru} = e^{-(.075805)(75)} = .002515,$$

or $\frac{1}{4}$ of 1%.

On the other hand,

$$\begin{aligned} P(ruin\ between\ 5th\ and\ 10th\ year) &= P(ruin\ \in [1825, 3650]) \\ &\leq \frac{e^{-(.00843)(1825)}}{1 - e^{-(.00843)}} \\ &= .0000248. \end{aligned}$$

i.e. 100 fold smaller or negligible.

Chapter 11

Future Work

There are two natural directions to extend the model introduced and studied in this dissertation. Namely, the first is to relax the independence of the increments of the random walk with positive drift, and allow Markov or other correlated random walks. The second direction is to let the initial capital u be invested in equities according to the Ito SDE of the form

$$\begin{cases} \mathrm{d}A_t = mA_t\,\mathrm{d}t + vA_t\,\mathrm{d}W_t, & W_t = \text{ Brownian Motion,} \\ A_0 = u, & u = \text{ initial capital, } m = \text{ mean return, } v = \text{ volatility.} \end{cases} \tag{11.1}$$

Since the solution to (11.1) is known explicitly as Geometric Brownian Motion $A_t = ue^{(m-\frac{1}{2}v^2)t+vW_t}$ one can consider its evolution at discrete time $t = 0,\ 1,\ 2,\ \ldots,$ and study the risk process of the form

$$U_n = A_n + \theta\mu n - \sum_{i=1}^{n} Z_i. \tag{11.2}$$

Preliminary analysis indicates that for $m \geq \theta\mu$ the upper bound for the probability of ruin for (11.2) is strictly smaller than for our model with $A_n \equiv u$.

Bibliography

[1] Albrecher, H., Renaud, J.-F., and Zhou, X. A Lévy insurance risk process with tax. *J. Appl. Probab. 45*, 2 (2008), 363–375.

[2] Asmussen, S., and Albrecher, H. *Ruin probabilities*, second ed., vol. 14 of *Advanced Series on Statistical Science & Applied Probability*. World Scientific Publishing Co. Pte. Ltd., Hackensack, NJ, 2010.

[3] Biffis, E., and Morales, M. On a generalization of the gerber–shiu function to path-dependent penalties. *Insurance: Mathematics and Economics 46* (2010), 92.

[4] Billingsley, P. *Convergence of probability measures*, second ed. Wiley Series in Probability and Statistics: Probability and Statistics. John Wiley & Sons, Inc., New York, 1999. A Wiley-Interscience Publication.

[5] Boĭkov, A. V. The Cramer-Lundberg model with stochastic premiums. *Teor. Veroyatnost. i Primenen. 47*, 3 (2002), 549–553.

[6] Bowers, N. L., Gerber, H. U., Hickman, J. C., Jones, D. A., and Nesbitt, C. J. *Actuarial mathematics*. Society of Actuaries, Schaumburg, IL, 1997.

[7] Cramér, H. Sur un nouveau théorèm-limite de la théorie des probabilités. *Actualités Sci. Indust. 736* (1938), 5–23.

[8] De Vylder, F., and Goovaerts, M. J. Recursive calculation of finite-time ruin probabilities. *Insurance Math. Econom. 7*, 1 (1988), 1–7.

[9] Dembo, A., and Zeitouni, O. *Large deviations techniques and applications*, second ed., vol. 38 of *Applications of Mathematics (New York)*. Springer-Verlag, New York, 1998.

[10] den Hollander, F. *Large deviations*, vol. 14 of *Fields Institute Monographs*. American Mathematical Society, Providence, RI, 2000.

[11] Dickson, D. C. M. On the distribution of the surplus prior to ruin. *Insurance Math. Econom. 11*, 3 (1992), 191–207.

[12] Dickson, D. C. M. Some comments on the compound binomial model. *ASTIN Bulletin: The Journal of the IAA 24*, 1 (1994), 33–45.

[13] Dickson, D. C. M. *Insurance risk and ruin*, second ed. International Series on Actuarial Science. Cambridge University Press, Cambridge, 2017.

[14] Dickson, D. C. M., and Waters, H. R. Recursive calculation of survival probabilities. *ASTIN Bulletin: The Journal of the IAA 21*, 2 (1991), 199–221.

[15] Donsker, M. D. An invariance principle for certain probability limit theorems. *Mem. Amer. Math. Soc., No. 6* (1951), 12.

[16] Dufresne, F., and Gerber, H. U. Risk theory for the compound poisson process that is perturbed by diffusion. *Insurance: Mathematics and Economics 10* (1991), 51.

[17] Durrett, R. *Probability: theory and examples*. Cambridge university press, 2013.

[18] F., L. Über die theorie der rückversicherung [on the theory of reinsurance]. *Transactions VI International Congressof Actuaries 1* (1909), 877.

[19] Feller, W. *An introduction to probability theory and its applications. Vol. II*. Second edition. John Wiley & Sons, Inc., New York-London-Sydney, 1971.

[20] Freidlin, M. *Functional integration and partial differential equations*, vol. 109 of *Annals of Mathematics Studies*. Princeton University Press, Princeton, NJ, 1985.

[21] Freidlin, M. I., and Wentzell, A. D. *Random perturbations of dynamical systems*, vol. 260 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, New York, 1984. Translated from the Russian by Joseph Szücs.

[22] Gerber, H. U. Mathematical fun with the compound binomial process. *ASTIN Bulletin: The Journal of the IAA 18* (1988), 161.

[23] Gerber, H. U., and Landry, B. On the discounted penalty at ruin in a jump-diffusion and the perpetual put option. *Insurance Math. Econom. 22*, 3 (1998), 263–276.

[24] Gerber, H. U., and Shiu, E. S. W. On the time value of ruin. *N. Am. Actuar. J. 2*, 1 (1998), 48–78. With discussion and a reply by the authors.

[25] Grandell, J. *Aspects of risk theory*. Springer Series in Statistics: Probability and its Applications. Springer-Verlag, New York, 1991.

[26] Iglehart, D. L. Diffusion approximations in collective risk theory. *J. Appl. Probability 6* (1969), 285–292.

[27] Jasiulewicz, H., and Kordecki, W. Ruin probability of a discrete-time risk process with proportional reinsurance and investment for the exponential and Pareto distributions. *Oper. Res. Decis. 25*, 3 (2015), 17–57.

[28] Karatzas, I., and Shreve, S. E. *Brownian motion and stochastic calculus*, second ed., vol. 113 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1991.

[29] Schmidli, H. *Stochastic control in insurance*. Probability and its Applications (New York). Springer-Verlag London, Ltd., London, 2008.

[30] Varadhan, S. R. S. *Large deviations and applications*, vol. 46 of *CBMS-NSF Regional Conference Series in Applied Mathematics*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1984.