Chapter 7


The purpose of this chapter is to present and explore some global characteristics of the surplus (or reserve) process. The characteristics we consider are connected, in some sense or another, either with the profitability of insurance operations or with their viability, i.e., the degree of protection against adversity.

1 INTRODUCTION

Given a particular portfolio, we define the surplus process \( R_t \) as the fund “on hand” at time \( t \). Merely for illustrative purposes, we may view \( R_t \) as the capital at time \( t \) if we keep in mind that it is, of course, not the whole capital of the company but rather the aggregate amount of the cash flow corresponding to the portfolio under consideration.

In some results below, we consider the surplus process \( R_t \) as a whole, without specifying its interior structure. However, in most cases, we assume

\[
R_t = u + c_t - S(t),
\]

where \( u \) is a fixed initial surplus, \( c_t \) is the aggregate amount of the positive cash flow by time \( t \), and \( S(t) \) is the corresponding loss process. In this section, we mostly view \( c_t \) as the total premium collected by time \( t \), and \( S(t) \) as the aggregate claim paid by the same time \( t \). In a more general model, \( c_t \) may include results of investment, and \( S(t) \) may include other expenses different from claim coverage.

As in previous chapters, when adopting model (1.1) in the continuous time case, we usually set

\[
S(t) = \sum_{i=1}^{N_t} X_i,
\]

where \( N_t \) is the process that counts consecutive claims, and \( X_i \) is the amount of the \( i \)th claim. Then, in a typical model, \( c_t = (1 + \theta)E\{S(t)\} \), where \( \theta \) is a relative loading coefficient. In Chapter ??, we consider other principles of premium determination.

If \( N_t \) is a Poisson process with rate \( \lambda \), and \( m = E\{X_i\} \), then as we saw in Chapter ??, \( E\{S(t)\} = m\lambda t \), and hence \( c_t = (1 + \theta)m\lambda t \).
In this case, a typical realization of the surplus process looks as in Fig.1. The process grows linearly, and at the random moments of claim arrivals, the process drops by the amounts of claims.

Subject to certain regulations and conditions, the company can choose a premium and an initial surplus that it considers reasonable. In doing so, the company proceeds from its goals, which in turn are determined by quality criteria that the company establishes for itself.

If $R_t$ is viewed as a profit, one of possible criteria is the sum

$$\sum_{k=1}^{T} E\{g(R_k - R_{k-1})\},$$

where $T$ is the time horizon from which the company proceeds, and $g$ is a utility function. (Since $u$ stands for the initial surplus, we use a symbol different from that from Chapter 7.) Time may be discrete or continuous, and $E\{g(R_k - R_{k-1})\}$ is the expected utility of the profit during the $k$th period. Note that the profit $R_k - R_{k-1}$ may be negative.

Another criterion is the expected utility of the profit at the “final” time $T$, that is, $E\{g(R_T)\}$. Instead of the expected utility criterion, one may consider more flexible criteria from Section 7.

A different approach is connected with paying dividends. We use here the term “dividend” for brevity and understand it in a broad sense: it may concern real dividends paid to stockholders or an amount taken from the reserve for other purposes, say, for investment.

Consider the discrete time case, and denote by $d_t$ the dividend paid at time $t$. (Certainly, we do not exclude the case $d_t = 0$.) Then, instead of (1.1), for the surplus at time $t$ we write

$$R_t = u + c_t - S(t) - D_t,$$  \hspace{1cm} (1.3)

where $D_t = d_1 + \ldots + d_t$, the aggregate amount of dividends paid by time $t$.

The choice of a dividend to be paid at time $t$ may (and should) depend on the current situation. So, in general, $d_t$ is a random variable depending on the realization of the process.
until the time \( t \) and the strategy of paying dividends the company follows. For example, if the time horizon is large and the initial surplus is low, it may prove to be reasonable to pay fewer dividends in the beginning in order to let the cash process grow and to be able to pay more in the future. A natural criterion here is the expected discounted total payment, namely,

\[
E \left\{ \sum_{t=1}^{T} v^t d_t \right\},
\]

where \( v \) is a discount factor (see Section ??). The problem consists in finding a strategy maximizing (1.4). We consider this problem in Section 3.

Criteria of another type appeal to the viability of insurance operations, which amounts to keeping the surplus at a proper level. One example is

\[
P(R_t \geq k_t \text{ for all } t \leq T),
\]

where \( T \) is a time horizon, and \( k_t \) is a given level for the surplus at time \( t \). The formula above concerns both discrete and continuous time cases.

The goal of the company in this case is either to maximize (under some natural constraints) probability (1.5), or to keep it higher than a given security level.

The simplest and most frequently considered case is that of \( k_t = 0 \). In this case, a traditional notation for the probability (1.5) is \( \phi_T(u) \), so

\[
\phi_T(u) = P(R_t \geq 0 \text{ for all } t \leq T).
\]

The initial surplus \( u \) is presented explicitly in the notation to emphasize that the probability under consideration depends on \( u \). We call (1.6) the survival probability since it is the probability that the portfolio will be solvent during the period \([0, T]\). As in previous chapters, the quantity

\[
\psi_T(u) = 1 - \phi_T(u)
\]

is called the ruin probability regarding the finite time horizon \( T \). This is the probability that the surplus process \( R_t \) will assume a negative value during the period \([0, T]\); see also Fig.2.

The term “survival” is traditional, but it is important to emphasize that in the actuarial literature and in this book, it is used in two situations. In the first, it concerns the solvency of a portfolio as above. The second situation concerns life insurance, and a survival probability in this case is the probability that an individual will attain a certain age. We will consider the corresponding theory in Chapter ??.

To avoid confusion, we will also call the survival probability in the former context a non-ruin probability.

For \( T = \infty \), we set

\[
\phi(u) = P(R_t \geq 0 \text{ for all } t < \infty) \text{ and } \psi(u) = 1 - \phi(u),
\]

and call these two quantities infinite-horizon non-ruin (survival) and ruin probabilities, respectively. We will often omit the adjective “infinite-horizon” when it does not cause misunderstanding.

In the actuarial literature, non-ruin and ruin probabilities are often denoted by \( \tilde{\phi}(u) \) and \( \tilde{\psi}(u) \) in the discrete time case, and by \( \phi(u) \) and \( \psi(u) \) if time is continuous. We will use the
7. GLOBAL CHARACTERISTICS OF THE SURPLUS PROCESS

FIGURE 2. A realization of the surplus process in the case of ruin; \( \tau \) is the time of ruin. For the particular realization above, ruin has occurred at the moment of the fourth claim’s arrival.

symbols \( \phi \) and \( \psi \) below in both cases since, as a rule, we treat them simultaneously. The probability \( \psi(u) \) is one of the main objects of study in the next section.

The models we consider below are, certainly, idealized and do not reflect all main features of real surplus processes. For example, we do not touch on such important issue as investment income. So, the corresponding results cannot be viewed as direct instructions for decision making. These results provide rather some useful information about the behavior of the insurance process, and may (and should) be taken into account together with other factors of the insurance business. In particular, the ruin probability should be viewed as only one possible characteristic of the riskiness of the insurance process.

Regarding ruin models, one may encounter the following reasoning. Ruin models do not take into account that the company invests its collected premiums and, as a result, the capital grows. On the other hand, these models do not take inflation into account. Since these two factors acting in the opposite directions may compensate each other, ruin models may occur to be more adequate to reality than they seem at first glance.

2 RUIN MODELS

In this section, except Section 2.5, we assume \( R_t \) to be a process with independent increments.

Computing ruin probabilities, especially for a finite time horizon, is a difficult problem. There are skilful direct computational methods, though nowadays their significance is decreasing. Simulation of insurance processes, especially if it is carried out with the use of powerful computers, may lead to better accuracy than direct calculations do. But in this case, qualitative analysis that helps to see a general picture becomes even more important.

We start in the next section with a general theory and estimation of ruin probabilities. In Section 2.7, we will touch briefly on some computational aspects and consider some simple computational examples.
2. Ruin Models

2.1 Adjustment coefficients and ruin probabilities

2.1.1 Lundberg’s inequality

As above, let $R_t$ be a surplus process, and $u = R_0$, the initial surplus. Assume, as usual, that $R_t$ is a process with independent increments.

To simplify some formulations below, it is convenient to define also the claim surplus process $W_t = u - R_t$. In particular, if (1.1) is true, then $W_t = S(t) - c_t$, the total claim minus the total premium.

As in previous chapters, for a time interval $\Delta = (t, t+s]$, we set $W_\Delta = W_{t+s} - W_t$, the increment of the process $W_t$ over $\Delta$. Denote by $M(\Delta)(z)$ the m.g.f. of $W_\Delta$.

We call a number $\gamma > 0$ an adjustment coefficient if for any $\Delta$, $M(\Delta)(\gamma) = 1$. (2.1.1)

Some remarks.

1. Certainly, (2.1.1) is true for $\gamma = 0$. The above definition presupposes that under some condition there exists a positive $\gamma$ for which (2.1.1) also holds. Detailed analysis will be provided later, and for now we just clarify the significance of the definition of $\gamma$.

Let us consider, first, an arbitrary r.v. $\xi$ and denote by $\mu$ and $M(z)$ its mean and m.g.f., respectively. For the sake of simplicity, suppose for now that $M(z)$ is defined for all $z \geq 0$, and assume also that $P(\xi = 0) \neq 1$. (Otherwise, $M(z) \equiv 1$, and the situation is trivial.) Let $p$ be a solution to the equation $M(z) = 1$.

The m.g.f. $M(z)$ is convex, and $M'(0) = \mu$. Hence, if $\mu \geq 0$, then the function $M(z)$ is non-decreasing, and a typical graph of $M(z)$ looks as in Fig.3a. In this case, a positive solution $p$ does not exist.

If $M'(0) = \mu < 0$, then starting from the point $(0,1)$, the graph of the function $M(z)$ should “go down at least for a while”; see Fig.3bc. Since $M(z)$ is convex, two situations may be called typical. Either the graph looks as in Fig.3b, and a solution $p$
exists and unique; or the graph looks as in Fig.3c, and a finite positive solution does not exist. In this case, by convention, we set $p = \infty$. We will see below that it is possible only if $P(\xi \leq 0) = 1$.

In our particular case, where $\xi = W_\Delta$, among three possibilities considered above, only the second may be viewed as realistic. Indeed, if the mean claim surplus $E\{W_\Delta\} \geq 0$, it is too bad, and there is no reason for the company to function. As we will see, in this case the ruin probability equals one. On the other hand, if $P(W_\Delta(z) \leq 0) = 1$, for the company it is “too good” to be true: in this case, clients always pay more than the company pays them. Later, we will consider all of this in more detail.

2. It may seem surprising that (2.1.1) may be true with the same $\gamma$ for all $\Delta$. As we will see in examples below, the point here is that, as a rule,

$$M_\Delta(z) = \exp\{q_1(\Delta)q_2(z)\}, \quad \text{(2.1.2)}$$

where $q_1$ and $q_2$ are separate functions of $\Delta$ and $z$, respectively. So, we can try to find $\gamma$ for which $q_2(\gamma) = 0$, and then (2.1.1) will be true for all $\Delta$.

3. The reader who did not omit Section ?? on martingales, or who is familiar with this notion, may notice that the definition (2.1.1) implies that the process $Y_t = \exp\{\gamma W_t\}$ is a martingale. We show this in detail in Section 2.5. As a matter of fact, for subsequent results to be true, we need only the aforementioned property of $Y_t$, and the independence of increments is not necessary. Corresponding generalizations will be also considered in Section 2.5.

We proceed to results. Let $\psi(u)$ be a ruin probability as it was defined in Section 1.

**Proposition 1** (Lundberg’s inequality). If the adjustment coefficient $\gamma$ exists, then the ruin probability

$$\psi(u) \leq \exp\{-\gamma u\}. \quad \text{(2.1.3)}$$

**More remarks.**

4. This famous inequality gives an estimate for the ruin probability with some leeway: we will see in examples below that the real ruin probability is, as a rule, less than the r.-h.s. of (2.1.3). But the estimate is simple and tractable, and has the advantage that the total information about the process is accumulated in one parameter $\gamma$. In next sections we consider many examples.

5. If we face the situation illustrated in Fig.3a, we can set $\gamma = 0$ in (2.1.3). The inequality will become trivial (the r.-h.s. equals one), but will be still true.

6. If for any $\Delta$, the claim surplus $W_\Delta(z) \leq 0$ with probability one, then ruin is impossible. It is reflected in (2.1.3) either: in this case, we face the situation illustrated in Fig.3c, so we can set $\gamma = \infty$, and the the r.-h.s. of (2.1.3) equals zero for any $u > 0$.

7. This remark is **very important** and concerns the initial surplus $u$. It is reasonable to view it not as the initial surplus at the time when the insurance process had started, but
as the current surplus at the present time. At each time moment, we can recalculate the ruin probability, depending on the amount of the surplus at the current time.

EXAMPLE 1. Assume that the loss process $S_t$ is a homogeneous compound Poisson process with unit rate and claims having the standard exponential distribution. As we will compute it in Section 2.4.1, in this case, the adjustment coefficient $\gamma = \frac{\theta}{1+\theta}$ and the ruin probability itself is

$$\psi(u) = \frac{1}{1 + \theta} \exp\left(\frac{-\theta u}{1 + \theta}\right),$$

where $\theta$ is the security loading coefficient. Suppose that $\theta = 0.1$ and we have started the corresponding insurance business with an initial surplus of 30 units of money. Then the ruin probability $\psi(30) = \frac{1}{1 + \theta} \exp\{-0.1 \cdot 30/1.1\} \approx 0.0595$. Suppose that during some period we were lucky, and the total premium collected occurred to be 5 units larger than the total payment. So, the current surplus became equal to 35 units. In this case, we may forget about the past result. Now we are in a better position, and the new ruin probability equals $\psi(35) = \frac{1}{1 + \theta} \exp\{-0.1 \cdot 35/1.1\} \approx 0.038$. \hfill $\square$

### 2.1.2 Proof of Lundberg’s inequality

First, note that, by definition, $W_0 = 0$, and hence, $W_{[0,t]} = W_t$. Since (2.1.1) is true for any $\Delta$, we can write that for any $T > 0$,

$$1 = M\{\gamma_T = 1\} = E\{e^{\gamma_T W_T}\} = E\{e^{\gamma_T W_T} \mid T \leq \tau\} P(\tau \leq T) + E\{e^{\gamma_T W_T} \mid T > \tau\} P(\tau > T) \geq E\{e^{\gamma_T W_T} \mid T \leq \tau\} P(\tau \leq T) = E\{e^{\gamma_T (W_T - W_s)} \mid T \leq \tau\} P(\tau \leq T).$$

Since by definition $R_\tau < 0$, we have $W_t = u - R_\tau > u$. Consequently, if we replace $e^{\gamma_T W_T}$ by $e^{\gamma_T}$, then the resulting expression will not get larger, and we may write that

$$1 \geq E\{e^{\gamma_T (W_T - W_s)} \mid T \leq \tau\} P(\tau \leq T) = e^{\gamma_T} E\{e^{\gamma_T (W_T - W_s)} \mid T \leq \tau\} P(\tau \leq T).$$

Since $R_t$ is a process with independent increments, $W_t$ has the same property. So, given that $\tau$ equals some $s \leq T$, the distribution of the r.v. $W_T - W_s$ is equal to the distribution of the r.v. $W_T - W_s$, and does not depend on the values of the process $W_t$ for $t \leq s$. Then

$$E\{e^{\gamma_T (W_T - W_s)} \mid \tau = s \leq T\} = E\{\exp\{\gamma W_s(s, T)\} \mid \tau = s \leq T\} = E\{\exp\{\gamma W_s(s, T)\}\} = 1$$

by the same property (2.1.1). Let $F(s) = P(\tau \leq s \mid \tau \leq T)$, the conditional d.f. of $\tau$. Then, by the formula for total expectation,

$$E\{e^{\gamma_T (W_T - W_s)} \mid \tau \leq T\} = \int_0^T E\{e^{\gamma_T (W_T - W_s)} \mid \tau = s \leq T\} dF(s) = \int_0^T 1 \cdot dF(s) = 1.$$

From this and (2.1.4) it follows that $1 \geq e^{\gamma_T} P(\tau \leq T)$, and hence $P(\tau \leq T) \leq e^{-\gamma_T}$. The r.h.s. of the last inequality does not depend on $T$. So, we can write that $\psi(u) = P(\tau < \infty) = \lim_{T \to \infty} P(\tau \leq T) \leq e^{-\gamma_T}$. \hfill $\blacksquare$
7. GLOBAL CHARACTERISTICS OF THE SURPLUS PROCESS

2.1.3 A preliminary condition: for large time horizons the aggregate claim should be large

Next, we proceed to the statement of the main theorem. To this end, we need one more formal condition. Namely, since we consider an infinite horizon, we need an assumption concerning the behavior of the process at infinity. We assume that for “large” time \( t \) the surplus \( R_t \) is “large”. Formally, we require

\[
R_t \xrightarrow{P} +\infty \quad \text{as} \quad t \to \infty
\]  

(2.1.5)

(for this type of convergence see Section ???).

It is worthwhile to emphasize two things. First, assumption (2.1.5) does not concern the ruin issue. When imposing this assumption, we do not exclude that the process may take on negative values “on the way to infinity”.

Secondly, this formal assumption holds in all reasonable models of insurance processes without paying dividends, including all particular models we consider in this book (excepting models with paying dividends in Section 3). So, in the first reading, the examples in this subsection below may be even omitted. Note also, that the condition (2.1.5) is strongly connected with the LLN.

For example, let \( R_t = u + c_t - S_t \), where \( S_t \) is the aggregate claim, the premium \( c_t = (1 + \theta)E\{S_t\} \), and \( \theta \) is a positive relative loading coefficient. Then we can represent \( R_t \) by

\[
R_t = u + \theta E\{S_t\} - [S_t - E\{S_t\}],
\]

and (2.1.5) will be true if \( E\{S_t\} \to \infty \) (the aggregate claim is large on the average for large \( t \)), and the third term in (2.1.6) is small with respect to the second, that is, the deviation of the payment \( S_t \) from its expected value is smaller than the expected value itself. More precisely, we may require that \( S_t - E\{S_t\} = o(E\{S_t\}) \) in probability, which is equivalent to the relation

\[
P(\{|S_t - E\{S_t\}| > \epsilon E\{S_t\}\}) \to 0, \quad \text{as} \quad t \to \infty, \quad \text{for any} \quad \epsilon > 0.
\]  

(2.1.7)

This is just another form of the LLN. By Chebyshev’s inequality (see (??)),

\[
P(\{|S_t - E\{S_t\}| > \epsilon E\{S_t\}\}) \leq \frac{1}{\epsilon^2 (E\{S_t\})^2} \var{S_t}.
\]

We see that for (2.1.7) to be true, it suffices that

\[
\var{S_t} = o((E\{S_t\})^2).
\]  

(2.1.8)

This is a very mild condition.

EXAMPLE 1. Let time be discrete, and \( S_t = S_t = X_1 + \ldots + X_t \), where the \( X \)’s (claims at separate time moments) are i.i.d. Then \( E\{S_t\} = mt \) and \( \var{S_t} = \sigma^2 t \), where \( m = E\{X_t\} \) and \( \sigma^2 = \var{X_t} \). If \( m > 0 \), then \( E\{S_t\} \to \infty \), and (2.1.8) is also true.

EXAMPLE 2. Let \( S_t \) be a homogeneous compound Poisson process. In the notation of Section ??, \( E\{S_t\} = m\lambda t \), \( \var{S_t} = (\sigma^2 + m^2)\lambda t \).
2. Ruin Models

So, again both requirements, \( E\{S(t)\} \to \infty \) and (2.1.8), are true. \( \square \)

EXAMPLE 3. If \( S(t) \) is a non-homogeneous compound Poisson process, using the notation of Section ??, we write

\[
E\{S(t)\} = m\chi(t), \quad \text{Var}\{S(t)\} = (\sigma^2 + m^2)\chi(t),
\]

and (2.1.5) is true if \( \chi(t) \), the mean number of claims by time \( t \), converges to infinity. \( \square \)

For the reader who did not skip Chapter ??, or who is familiar with the notion of Brownian motion, consider

EXAMPLE 4. Let \( S(t) = \mu t + \sigma w_t \), a Brownian motion with drift. It is natural to assume \( \mu > 0 \). Then \( E\{S(t)\} = \mu t \to \infty \), \( \text{Var}\{S(t)\} = \sigma^2 t \), and (2.1.8) again holds. \( \square \)

The reader can suggest other examples. In particular, it is not necessary for the \( X_i \)'s above to be identically distributed.

Thus, we adopt the assumption (2.1.5) and turn to main result.

2.1.4 The main theorem

Let \( \tau = \tau_u \) be the moment of ruin, provided that it occurs; see also Fig.2. Formally, \( \tau = \min\{t : R_t < 0\} \). If the process \( R_t \) never assumes a negative value (no ruin occurs), we indicate this writing \( \tau = \infty \).

As was repeatedly noted in Chapters ?? and ??, the r.v. \( \tau \) may be defective; that is, it may happen that \( P(\tau_u < \infty) < 1 \). Moreover, this should be the case when we are modeling real surplus processes, since \( P(\tau_u < \infty) \) equals the ruin probability \( \psi(u) \), and we want this probability to be small.

Below we omit the index \( u \) in \( \tau_u \), and start with a corollary of the main theorem that will be stated later.

**Theorem 2** Let (2.1.5) be true, and \( \gamma > 0 \) be the adjustment coefficient defined above. Then

\[
\psi(u) = \frac{\exp\{-\gamma u\}}{E\{\exp\{-\gamma R_\tau\} | \tau < \infty\}}. \tag{2.1.9}
\]

So, the theorem gives a precise expression for the ruin probability. The denominator in (2.1.9) looks somewhat complicated, though, as we will see in Section 2.4, there are cases when it may be easily computed.

Lundberg’s inequality follows from Theorem 2 almost immediately. Indeed, by definition, \( R_\tau \leq 0 \) and \( \gamma > 0 \). Hence, \( \exp\{-\gamma R_\tau\} \geq 1 \). So, the denominator in (2.1.9) is not less than one, and we have come to (2.1.3).

We prove Theorem 2 in Section 2.5 by making use of a martingale technique. For the reader who does not plan yet to learn martingales, we gave above a direct proof of Lundberg’s inequality.

2.2 Computing adjustment coefficients

In the first subsection below, we consider conditions under which equation (2.1.1) has a positive solution. This has a rather mathematical significance: the point is that, as we will
see, if in a particular problem we manage to find a solution \( \gamma > 0 \), then we can be sure that it is unique and use it applying either Lundberg’s inequality or Theorem 2. So, the reader who is interested merely in applications may skip Subsection 2.2.1.

2.2.1 A general proposition

Consider a r.v. \( \xi \) whose m.g.f. \( M(z) \) exists and is finite for all \( z \in [0, z_0) \), where \( 0 < z_0 \leq \infty \).

We assume that \( z_0 \) is the largest number with this property, that is, \( M(z) = \infty \) for \( z > z_0 \).

To clarify this, let us recall that if, for example, \( \xi \) is uniformly distributed on some interval, \( M(z) \) exists for all \( z \) and hence \( z_0 = \infty \). On the other hand, if \( \xi \) is exponential, then \( M(z) \) exists only for \( z < 1/\mu \), where \( \mu = E\{\xi\} \). So, \( z_0 = 1/\mu \). (See Section ???.)

As for the point \( z_0 \) itself, \( M(z_0) \) may or may not exist. In Example 1 below, we consider the case when \( M(z_0) \) is finite. However, for the exponential distribution, \( M(z) \) does not exist at \( z_0 = 1/\mu \).

We do not consider \( M(z) \) at negative \( z \)'s.

Note also that, in general, we do not have to assume \( \mu \) to be finite. Since there exists a positive \( z \) for which \( M(z) < \infty \), we can write \( \mu = \frac{1}{z}E\{z\xi\} \leq \frac{1}{z}E\{e^{z\xi}\} = \frac{1}{z}M(z) < \infty \). However, we do not exclude the case where \( \mu = -\infty \), that is, the negative part of \( X \) has an infinite expectation. In this case, the reasoning below will remain correct.

Consider the equation

\[
M(z) = 1, \ z \geq 0. \tag{2.2.1}
\]

We assume that

\[
P(\xi = 0) \neq 1, \tag{2.2.2}
\]

since otherwise \( M(z) \equiv 1 \), and equation (2.2.1) is trivial.

**Proposition 3** If (2.2.2) holds, then the following is true.

(a) A positive solution to (2.2.1) exists if and only if \( \mu < 0 \) and

\[
M(z_0) \geq 1. \tag{2.2.3}
\]

(b) If \( \mu < 0 \) but (2.2.3) does not hold, then \( M(z) < 1 \) for all positive \( z \) for which \( M(z) \) exists.

(c) If the equation (2.2.1) has a positive solution, then this is the only positive solution, and (2.2.3) is true.

Before proving this proposition, we clarify the sense of its conditions. In Section 2.1.1, in particular in Fig.3, we have already shown the role played by \( \mu \). Consider now condition (2.2.3).

First of all, if \( M(z_0) = \infty \), whatever \( z_0 \) is, finite or infinite, then (2.2.3) holds automatically. In particular, this is the case when \( z_0 = \infty \) and

\[
P(\xi > 0) > 0. \tag{2.2.4}
\]

Indeed, (2.2.4) implies that there exists \( b > 0 \) such that

\[
P(\xi \geq b) > 0. \tag{2.2.5}
\]
2. Ruin Models

(If \( P(\xi > b) = 0 \) for all \( b > 0 \), we can write that \( P(\xi > b) = \lim_{b \to 0} P(\xi > b) = 0 \),
which would contradict (2.2.4).) Hence, denoting by \( F(x) \) the d.f. of \( \xi \), we can write

\[
M(z) = \int_{-\infty}^{\xi} e^{\xi z} F(x) e^{\xi} \int_{b}^{\xi} dF(x) = e^{\xi b} P(\xi > b) \to \infty.
\]

On the other hand, if \( P(\xi > 0) = 0 \), then in view of (2.2.2), \( M(z) = E \{ e^{\xi z} \} < 1 \) for all \( z > 0 \).
(Indeed, \( e^{\xi} < 1 \) with a positive probability, and \( P(e^{\xi} > 1) = 0 \).) So, in this case, there is no solution to (2.2.1), and condition (2.2.3) also does not hold. This situation has been illustrated in Fig.3c.

Now, let \( z_0 < \infty \). Again, if \( M(z_0) = \infty \), the condition (2.2.3) holds automatically. For example, let \( \xi = X - c \) where \( X \) is interpreted as a claim and is exponential, and \( c \) is viewed as a premium. Then \( M(z) = e^{-cz} \frac{1}{1-E(X)z} \to \infty \) as \( z \to 1/E(X) \). The situation is illustrated in Fig.4a.

However, it may happen that \( M(z_0) < \infty \), while \( M(z) = \infty \) for all \( z > z_0 \).

EXAMPLE 1. Let \( \xi = X - c \) where \( c > 0 \) and a positive r.v. \( X \) has the density

\[
f(x) = \frac{K}{1+x^2} e^{-x}, \ x \geq 0,
\]

and \( K \) is a constant for which \( \int_{0}^{\infty} f(x) dx = 1 \). Then

\[
M(z) = e^{-cz} \int_{0}^{\infty} e^{xz} \frac{K}{1+x^2} e^{-x} dx = Ke^{-cz} \int_{0}^{\infty} e^{(z-1)x} \frac{1}{1+x^2} dx.
\]

The last integral diverges (or equals infinity) for \( z > 1 \), it is finite for \( z < 1 \), and

\[
M(1) = Ke^{-c} \int_{0}^{\infty} \frac{1}{1+x^2} dx < \infty.
\]

Thus, \( M(z) \) is defined at \( z \leq 1 \). The last integral equals \( \arctan(\infty) = \pi/2 \), and setting \( K_1 = K\pi/2 \), we have

\[
M(1) = K_1 e^{-c}.
\]
Let $i$ plus during the $i$th period. As a rule, $Y_i = X_i - c$, where $X_i$ and $c$ are the aggregate claim and premium, respectively, corresponding to the $i$th period. Below, we omit the adjective "aggregate" if it cannot cause misunderstanding.

Suppose $Y_i$'s are i.i.d. and omit the index $i$ when this cannot cause misunderstanding. As a counterpart of condition (2.2.2) in our case, we assume that

$$P(Y = 0) < 1.$$  

(2.2.7)
2. Ruin Models

Formally, we should solve the equation (2.1.1) for all intervals \( \Delta = (k, k + t] \), where \( k, t \) are integers. However, since \( Y \)'s are identically distributed, it suffices to consider \( \Delta = (0, t] \), and accordingly the r.v. \( W_t \).

Because \( Y \)'s are independent, the m.g.f. \( M_\Delta(z) = M_{(0,t)}(z) = E\{\exp(zW_t)\} = (M_Y(z))^t \), where \( M_Y(z) \) is the common m.g.f. of the r.v.'s \( Y_t \). Thus, \( M_\Delta(z) = 1 \) if and only if

\[
M_Y(z) = 1. \tag{2.2.8}
\]

This is an equation for the adjustment coefficient \( \gamma \).

One may apply now Proposition 3, but as a matter of fact, we should not much worry about conditions of this Proposition. As we know from this proposition, if we manage to find a positive solution to (2.2.8), this solution will be unique and all conditions of Proposition 3 will be true automatically.

Let now \( Y_i = X_i - c \). Then \( E\{Y_i\} = m - c \), where \( m = E\{X_i\} \). The condition \( E\{Y\} < 0 \) is equivalent to the condition

\[
c > m,
\]

and condition (2.2.7) to the condition

\[
P(X_i = c) < 1,
\]

which is more than natural: nobody will pay a premium that is larger than the future payment with probability one, and the payment may be equal to the premium but, certainly, not always.

Note now that equation (2.2.8) may be rewritten as

\[
e^{-cz}M_X(z) = 1, \tag{2.2.9}
\]

where \( M_X(z) \) is the common m.g.f. of \( X \)'s. (See (??).)

In examples below, when considering a separate claim \( X_i \), we omit the index \( i \).

EXAMPLE 1. To deal with a distribution with some skewness (which is more realistic than to take, say, a uniform distribution), assume that the claim \( X \) has a \( \Gamma \)-density, say, \( f(x) = \frac{x}{\alpha} e^{-\alpha x} \). Then \( E\{X\} = 2 \), the m.g.f. \( M_X(z) = 1/(1 - z)^2 \) and exists for \( z < 1 \). We can rewrite (2.2.9) as

\[
e^{-cz} = (1 - z)^2, \tag{2.2.10}
\]

but we should accept only solutions \( z < 1 \). In Exercise 5a, the reader is suggested to graph the r.-h.s. and the l.-h.s. of (2.2.10).

It is impossible to write a solution to (2.2.10) explicitly, but it is easy to solve it numerically, even using a graphing calculator. For example, for \( c = 2.15 \), we will readily find that \( 0.136 < \gamma < 0.137 \). Thus,

\[
\psi(u) \leq \exp\{-\gamma u\} \leq \exp\{-0.136u\}.
\]

(To have a correct inequality, we should take the lower (!) bound for \( \gamma \).) \( \square \)

EXAMPLE 2. Let \( X \) be well approximated by a \( (m, \sigma^2) \)-normal distribution. Then \( M_X(z) = \exp\{mz + \sigma^2 z^2/2\} \), and equation (2.2.9) may be rewritten as

\[
-cz + mz + (\sigma^2 z^2/2) = 0.
\]
7. **GLOBAL CHARACTERISTICS OF THE SURPLUS PROCESS**

The positive root is

\[ \gamma = \frac{2(c - m)}{\sigma^2}. \]  

(2.2.11)

The particular expression (2.2.11) gives an idea of the following approximation in the general case. Taking the logarithm of both sides of (2.2.9), we can rewrite it as

\[ -cz + \ln M_X(z) = 0. \]  

(2.2.12)

Assume that the root of the equation is “small”. In Section ??,??, we derived the approximation formula

\[ \ln M_X(z) = mz + \frac{\sigma^2 z^2}{2} + o(z^2), \]

where \( m = E\{X\} \) and \( \sigma^2 = Var\{X\} \). So, we can write

\[ -cz + mz + (\sigma^2 z^2 / 2) + o(z^2) = 0. \]  

(2.2.13)

If we neglect the term \( o(z^2) \), the positive solution to (2.2.13) will lead us to the approximation

\[ \gamma \approx \frac{2(c - m)}{\sigma^2}. \]  

(2.2.14)

The accuracy of such an approximation can be estimated (with the use of bounds for the remainder in Taylor’s expansions), but we restrict ourselves to an example.

**EXAMPLE 3.** For the data from Example 1, \( \sigma^2 = 2 \), and (2.2.14) gives

\[ \gamma \approx 2 \cdot 0.15^2 = 0.15, \]

which is not so bad in comparison with the answer from Example 1 (\( \approx 0.136 \)).

Let now \( c = 2.05 \). Then, as is easy to compute, \( 0.05 < \gamma < 0.051 \), while (2.2.14) gives 0.05. \( \square \)

### 2.2.2.2 The adjustment coefficient for a group of insured units.

It is natural to view the loss (claim) r.v. \( X_i \) above as the aggregate claim coming from a portfolio of insured units. Here we present it in an explicit way. Let

\[ X_i = X_{i1} + ... + X_{in}, \]

(2.2.15)

where \( n_i \) is a certain number of units composing the portfolio in the \( i \)th period, and \( X_{ij} \) is the payment to the \( j \)th unit. We assume \( \{X_{ij}\} \) to be i.i.d. r.v.’s. Note that in this case, \( X_i \)'s are independent, but perhaps are not identically distributed because the number of terms in the sum (2.2.15) depends on \( i \).

Denote by \( \tilde{c} \) the premium for a particular unit, and set \( Y_{ij} = X_{ij} - \tilde{c} \), the claim surplus of the \( j \)th unit in the \( i \)th period. Since time is discrete, for any time interval \( \Delta = (t, t + k] \), where \( t, k \) are integers,

\[ W_{\Delta} = \sum_{i=t+1}^{t+k} Y_i = \sum_{i=t+1}^{t+k} \sum_{j=1}^{n_i} Y_{ij}, \]

(2.2.16)
and the m.g.f. $M_{\Delta}(z) = (M_Y(z))^s$ where $M_Y(z)$ is the common m.g.f. of r.v.'s $Y_{i,j}$, and $s = n_{t+1} + ... + n_{t+k}$, the total number of all terms in (2.2.16).

Thus, $M_{\Delta}(z) = 1$ iff $M_Y(z) = 1$, and we have come to a nice conclusion:

The adjustment coefficient for a homogeneous portfolio is equal to the adjustment coefficient for one separate insured unit. \hfill (2.2.17)

**EXAMPLE 4.** Considering the above model, it is natural to assume that the loss for a particular unit is equal to zero with a positive and “substantial” probability. Let, say, $X_{ij} = 0, 10, 20$ with probabilities 0.8, 0.1, 0.1, respectively. Then $E\{X_{ij}\} = 3$. Suppose $\hat{c} = 3.5$.

To compute the adjustment coefficient $\gamma$, we do not need any information about the numbers of units, and the equation for $\gamma$ is the same equation (2.2.9) where $X$ corresponds to $X_{ij}$, and $c$ should be replaced by $\hat{c}$. Using version (2.2.12), we have

$$-3.5z + \ln(0.8 + 0.1e^{10z} + 0.1e^{20z}) = 0.$$ 

An analytical solution is again impossible, but it is easy to solve the equation numerically. The reader can verify that the positive solution $\gamma \approx 0.022$. $\Box$

However, the situation is different if we consider the portfolio as a whole, and set each

$$X_i = \xi_{i1} + ... + \xi_{iK_i}, \quad (2.2.18)$$

where now $\xi_{ij}$ is the amount of the $j$th claim in the $i$th period, and the r.v. $K_i$ is the number of claims in this period. It makes sense to emphasize that, while $X_{ij}$’s above were payments to separate units (and could take on zero values), $\xi_{ij}$ are claims arriving at the system (and may be assumed to be positive). Suppose that all r.v.’s are independent, $\xi$’s are identically distributed, and the same is true for $K_i$’s.

If we know the distributions of $\xi$’s and $K_i$, we can find, at least theoretically, the distribution of $X_i$ and hence the adjustment coefficient. For example, if we accept the approximation (2.2.14), we can set there $m = \tilde{m}E\{K_i\}$ and $\sigma^2 = \tilde{\sigma}^2E\{K_i\} + \tilde{m}^2\text{Var}\{K_i\}$, where $\tilde{m}$ and $\tilde{\sigma}^2$ are the mean and variance of $\xi$’s, respectively. A particular example is given in Exercise 9. The case where $K_i$ are Poisson r.v.’s is the most interesting, but it is convenient for us to consider it later, in Section 2.2.4, after we explore the case of the Poisson process in continuous time.

**2.2.3 The case of a homogeneous compound Poisson process**

We turn to continuous time and consider the claim surplus process $W_t = S_{(t)} - c_t$, where

$$S_{(t)} = \sum_{i=1}^{N_t} X_i, \quad (2.2.19)$$

$N_t$ is a Poisson process with intensity $\lambda$, and $X_i$ are i.i.d. and do not depend on $N_t$. (See also Section ???.?) We assume $X$’s are positive.
The reader remembers that $E\{N_t\} = \lambda t$, and $E\{S_t\} = m\lambda t$, where $m = E\{X_i\}$. Denote by $N_\Delta$, $S_\Delta$ the increments of the corresponding processes over an interval $\Delta = (t, t + \delta]$. Then $E\{N_\Delta\} = \lambda \delta$, and $E\{S_\Delta\} = m\lambda \delta$.

The point is that the r.v. $S_\Delta$ is also a compound Poisson r.v. Indeed, we can write that

$$S_\Delta = S_{(t+\delta)} - S_t = \sum_{i=N_t+1}^{N_{t+\delta}} X_i.$$ 

The process $N_t$ is a process with independent increments, and at any point $t$ “everything starts over as if from the beginning”. So, the r.v. $S_\Delta$ has the same distribution as the r.v.

$$S_\Delta = \sum_{i=1}^{N_\Delta} X_i.$$ 

In particular, the m.g.f.

$$M_{S_\Delta}(z) = \exp\{\lambda \delta [M_X(z) - 1]\},$$

where $M_X(z)$ is the m.g.f. of $X$’s (see (??,??) in Proposition ??,??).

Set the premium

$$c_\Delta = (1 + \theta)E\{S_t\} = (1 + \theta)m\lambda t,$$

where again $\theta > 0$ is a security loading coefficient. Then the increment of the premium over the interval $\Delta$ is $c_\Delta = (1 + \theta)m\lambda \delta$, and the m.g.f. of $W_\Delta$ is

$$M_{W_\Delta}(z) = \exp\{-c_\Delta z\}M_{S_\Delta}(z) = \exp\{-(1 + \theta)m\lambda \delta z\} \exp\{\lambda \delta [M_X(z) - 1]\} = \exp\{\lambda \delta [M_X(z) - 1 - (1 + \theta)mz]\}.$$ 

Thus, $M_{W_\Delta}(z) = 1$ iff $M_X(z) - 1 - (1 + \theta)mz = 0$, which we write as

$$M_X(z) = 1 + (1 + \theta)mz,$$  \hspace{1cm} (2.2.20)

or

$$M_X(z) = 1 + cz,$$  \hspace{1cm} (2.2.21)

where $c = (1 + \theta)m$, the premium per one claim.
This is an equation for the adjustment coefficient.

It is noteworthy that this equation does not involve $\lambda$ and $\delta$, so the adjustment coefficient is specified only by the distribution of $X$ (compare with (2.2.17)).

Consider (2.2.21) and/or (2.2.20) setting for simplicity $M(z) = M_X(z)$. Since $X$ is positive, and $M(z)$ is convex, $M'(z) > M'(0) = m > 0$. Hence, $M(z)$ is strictly increasing, and if $M(z)$ is defined on $[0, \infty)$, then $M(z) \to \infty$ as $z \to \infty$. Note also that while $M'(0) = m$, the slope of the line specified by the r.-h.s. of (2.2.21) is $c > m$. See also Fig.5a. Hence, in this case a solution to (2.2.21) or (2.2.20) exists and unique.

If $M(z)$ is defined on a finite interval $[0, z_0]$, $z_0 < \infty$, but $M(z) \to \infty$ as $z \to z_0$, we have the same; see also Fig.5b.

Consider the last case where $M(z)$ is defined on $[0, z_0]$, $z_0 < \infty$, and $M(z_0) < \infty$. Then a solution exists for all sufficiently small $\theta \geq 0$. Indeed, the closer $\theta$ is to zero, the closer the line $1 + (1 + \theta)mz$ is to the line tangent to $M(z)$ at the origin; see Fig.6a. So, for small $\theta$, the line $1 + (1 + \theta)mz$ intersects the graph of $M(z)$.

The formal proof may run as follows. Since $X$ is positive and $e^x \geq 1 + x + \frac{1}{2}x^2$ for all $x \geq 0$, the m.g.f. $M(z) = E\{e^{zX}\} \geq E\{1 + ZX + \frac{1}{2}z^2X^2\} = 1 + mz + \frac{1}{2}z^2E\{X^2\}$. Because $E\{X^2\} > 0$ and $1 + (1 + \theta)mz \to 1 + mz$ as $\theta \to 0$, for sufficiently small $\theta$, we have $M(z_0) \geq 1 + mz_0 + \frac{1}{2}z_0^2E\{X^2\} > 1 + (1 + \theta)mz_0$.

For large $\theta$, a positive solution may not exist. In this case, we may set $\gamma = z_0$.

Indeed, let us first choose $\theta$ for which $\gamma$ exists and equals $z_0$. It would be the limiting case when $M(z_0) = 1 + (1 + \theta)mz_0$. For such a $\theta$, Lundberg’s bound for the ruin probability is $e^{-\nu} = e^{-\gamma z_0}$. But for a larger $\theta$, the ruin probability will be even smaller (the positive component of the surplus process gets larger). Hence, the same bound is true for such $\theta$’s also.

EXAMPLE 1. Let $X$ have a $\Gamma$-distribution with parameters $a, \nu$. Then $m = \nu/a$, and (2.2.20) amounts to $(1 - z/a)^{-\nu} = 1 + z(1 + \theta)^{\nu}/a$, or

$$
\left(1 + z(1 + \theta)^{\nu}/a\right) \left(1 - \frac{z}{a}\right)^{\nu} = 1, \tag{2.2.22}
$$

where we should consider only $z < a$, since for $z \geq a$ the m.g.f. does not exist.

Solving (2.2.22) numerically does not present any difficulty; for $\nu = 1, 2$ one can write an explicit solution. If $\nu = 1$, i.e., $X$ is exponentially distributed, after simple algebra, (2.2.22)
may be rewritten as
\[ z \left( \theta - \frac{z(1 + \theta)}{a} \right) = 0. \]

The positive root is
\[ \gamma = \frac{a \theta}{1 + \theta}. \quad (2.2.23) \]

For \( \nu = 2 \), since one root of (2.2.22) is \( z = 0 \), the equation may be reduced to a quadratic equation, and we should choose the positive root which is less than \( a \). See also Exercise 11. \( \square \)

Using expansion (??), we can get a counterpart of approximation (2.2.14), writing (2.2.21) as \( 1 + mz + \frac{1}{2}mz^2 + o(z^2) = 1 + cz \), where \( m_2 = E\{X^2\} \), the second moment of \( X \). Neglecting the term \( o(z) \), we come to the approximation
\[ \gamma \approx 2\left( c - m \right) m_2. \quad (2.2.24) \]

EXAMPLE 2. In the situation of Example 1, \( m = \frac{\nu}{a} \), \( m_2 = \frac{\nu}{a^2} + \left( \frac{\nu}{a} \right)^2 \), and since \( c = (1 + \theta)m \), approximation (2.2.24) gives
\[ \gamma \approx \frac{2\theta a}{1 + \nu}. \quad (2.2.25) \]

For \( \nu = 1 \), the precise \( \gamma \) in (2.2.23) differs from approximation (2.2.25) by the multiplier \( \frac{1}{1 + \theta} \), which for small \( \theta \) is close to one. See another example in Exercise 14. \( \square \)

2.2.4 The discrete time case revisited

We continue to consider the compound Poisson model of the previous section, but now we will refer to ruin only as the event when the current surplus becomes negative at the end of a unit interval; say, at the end of a year. Formally, it means that the non-ruin (survival) probability is defined as \( \tilde{\phi}(u) = P(R_t \geq 0, t = 1, 2, \ldots) \). Time is still continuous, but we check for ruin only at integer moments of time. Here, we follow a tradition and mark the non-ruin probability by a tilde to distinguish this probability from \( \phi(u) = P(R_t \geq 0 \text{ for all } t > 0) \). Let \( \tilde{\psi}(u) = 1 - \tilde{\phi}(u) \), and as usual, \( \psi(u) = 1 - \phi(u) \).

Clearly,
\[ \psi(u) \leq \psi(u). \quad (2.2.26) \]

(If ruin occurs at an integer time moment, then ruin has occurred at some moment.)

To find \( \tilde{\psi}(u) \), we apply the results of Section 2.2.2. Let us mark \( X \)'s from this section by a tilde to distinguish them from the individual claims \( X \) in this section. More precisely, let for \( k = 1, 2, \ldots \)
\[ \tilde{X}_k = \sum_{i=N_{k-1}+1}^{N_k} X_i, \]
the total claim for the unit time period \( (k - 1, k] \). Since \( N_t \) is a homogeneous Poisson process, at the beginning of each period, the claim process starts to run as from the very beginning, all r.v. \( \tilde{X}_k \) are i.i.d., and
\[ E\{\tilde{X}_k\} = mE\{N_k - N_{k-1}\} = m\lambda. \]
Then the premium per unit time is \( c = (1 + \theta)E\{\tilde{X}_k\} = (1 + \theta)m\lambda \). The r.v. \( \tilde{X}_k \) has a compound Poisson distribution, and
\[
M_{\tilde{X}_k}(z) = \exp\{\lambda(M_X(z) - 1)\}.
\]

Hence in our case, the equation (2.2.9) may be written as
\[
\exp\{- (1+\theta)m\lambda z\} \exp\{\lambda(M_X(z) - 1)\} = 1,
\]
which is equivalent to \( \lambda(M_X(z) - 1) - (1+\theta)m\lambda z = 0 \), or
\[
M_X(z) = 1 + (1+\theta)mz.
\]

Thus, we have arrived at the same equation (2.2.20). This means that the adjustment coefficient in the discrete time case (for the scheme we consider) is the same as in the case of continuous time. So, the bounds \( e^{-\gamma u} \) for \( \psi(u) \) and \( \tilde{\psi}(u) \) will be the same.

Does this contradict (2.2.26)? No, since we deal with upper bounds. To compare \( \psi(u) \) and \( \tilde{\psi}(u) \) we should also take into account the denominator \( E\{\exp\{-\gamma\tilde{R}_\tau|\tau < \infty\}\} \) in (2.1.9) which is different for these two cases. The r.v. \( -\tilde{R}_\tau = |R_\tau| \) is the deficit at the moment of ruin. If we consider only integer moments of time, \( R_\tau \) may become negative before the end of the period of ruin, and \( |R_\tau| \) corresponds to the deficit accumulated during this period, whereas in the general case, \( |R_\tau| \) is the deficit corresponding to claim at the moment of ruin. So, we should expect that, on the average, \( |R_\tau| \) is larger in the discrete time case.

Certainly, this is not a proof; the proof itself follows from (2.2.26) since if the numerator in (2.1.9) is the same for both cases, the denominator must take on different values in the continuous and discrete cases.

2.2.5 The case of the non-homogeneous compound Poisson process

This section presupposes that the reader is familiar with the non-homogeneous Poisson process considered in Chapter ??.

Since the equation for \( \gamma \) in the previous section does not depend on \( \lambda \), one may hope that, as a matter of fact, the homogeneity of the process \( N_t \) is not necessary, and we can get the same result in the general case. This is indeed true.

Let \( \lambda(t) \) be the intensity function for \( N_t \). As in Section ??.??, we set
\[
\chi(t) = \int_0^t \lambda(s)ds, \quad \chi_\Delta = \int_\Delta \lambda(s)ds
\]
for any interval \( \Delta \). As was shown in Chapter ??., \( E\{N_t\} = \chi(t) \), \( E\{N_\Delta\} = \chi_\Delta \), and hence \( E\{S_\Delta\} = m\chi_\Delta \). The r.v. \( S_\Delta \) is a compound Poisson r.v., and its m.g.f.
\[
M_{S_\Delta}(z) = \exp\{\chi_\Delta[M_X(z) - 1]\}.
\]

We set again the premium
\[
c_t = (1 + \theta)E\{S_{(t)}\} = (1 + \theta)m\chi(t).
\]
Then the increment of the premium over an interval $\Delta$ is 

$$c_\Delta = (1 + \theta)m\chi_\Delta,$$

and the m.g.f. of $W_\Delta$ is

$$M_{W_\Delta}(z) = \exp\left\{-c_\Delta z\right\}M_{S_\Delta}(z) = \exp\left\{- (1 + \theta)m\chi_\Delta z\right\}\exp\{\chi_\Delta [M_X(z) - 1]\}$$

$$= \exp\{\chi_\Delta [M_X(z) - 1 - (1 + \theta)mz]\}.$$

Thus, $M_{W_\Delta}(z) = 1$ iff $M_X(z) - 1 - (1 + \theta)mz = 0$, which again leads to (2.2.20). So, the equation does not involve $\chi(t)$, and the adjustment coefficient is again specified only by the distribution of $X$.

2.3 Trade-off between the premium and the initial surplus

Having the bound

$$\psi(u) \leq \exp\left\{-\gamma u\right\}, \quad (2.3.1)$$

we can estimate the initial surplus $u$ and/or the premium $c$, needed for the ruin probability to be less than a given desirable level $\beta$. Setting $\exp\{-\gamma u\} = \beta$, we get $u = -\frac{1}{\gamma}\ln\beta = s/\gamma$ where $s = \ln(1/\beta)$. If $\beta < 1$, then $s = \ln(1/\beta) > 0$. The estimate $s/\gamma$ has some leeway because we proceed not from the real ruin probability, but from its upper bound. In any case,

For the ruin probability to be less than $\beta$, it suffices

that the initial surplus $u \geq \frac{1}{\gamma}\ln(1/\beta) = \frac{s}{\gamma}. \quad (2.3.2)$

EXAMPLE 1. Let us revisit Example 2.2.2-1, where we found that $0.136 < \gamma < 0.137$. If we estimate $u$ proceeding from the inequality $u \geq s/\gamma$, then we should take the lower bound for $\gamma$, that is, 0.136. Indeed, if $u \geq \frac{s}{0.136}$, then we have a right to write that

$$u \geq \frac{s}{0.136} \geq \frac{s}{\gamma},$$

which implies that the ruin probability is less than $\beta$. In turn, $\frac{1}{0.136} < 7.36$. So, we take $u \geq 7.36s$, and it will be an estimate with some leeway.

For example, for $\beta = 0.05$, we have $s = \ln(1/0.05) = \ln 20 \leq 2.996 < 3$, and we come to $u \geq 3 \cdot 7.36 = 22.08$. $\Box$

However, there is another way to keep the ruin probability lower than a given level, namely, to increase the premium. To a certain degree, both characteristics – the initial surplus and the premium – are under the control of the insurer, and the determination of them consists in a trade-off between these characteristics.
2. Ruin Models

For the discrete time case, by (2.2.12),

\[ c = \frac{1}{\gamma} \ln M_X(\gamma). \]  

(2.3.3)

In the compound Poisson case, (2.2.21) implies

\[ c = \frac{1}{\gamma} [M_X(\gamma) - 1]. \]  

(2.3.4)

(Note that these representations do not differ much for small \( \gamma \). Indeed, \( M_X(\gamma) - 1 \) is small for \( \gamma \) close to 0 (say why). Since \( \ln(1 + x) = x + o(x) \) for small \( x \)'s, we can write \( \ln M_X(\gamma) = \ln[1 + (M_X(\gamma) - 1)] = M_X(\gamma) - 1 + o(M_X(\gamma) - 1) \), where the second term is small.)

Our goal is to find \( c \) and \( u \) for which the ruin probability is not greater than a given security level \( \beta \). Since we have only an upper bound for the ruin probability, we have to proceed from this upper bound. For the upper bound in (2.3.1) to equal \( \beta \), we should have \( e^{-\gamma u} = \beta \), or

\[ \gamma = \frac{1}{u} \ln \left( \frac{1}{\beta} \right) = \frac{s}{u}. \]  

(2.3.5)

Now, if we replace \( \gamma \) in (2.3.3) and (2.3.4) by the r.-h.s. of (2.3.5), we will establish a relation between \( u \) and \( c \), which ensures the ruin probability to be not larger than \( \beta \). (As we understand, with leeway, since we proceeded from an upper bound for the ruin probability.)

Thus, in the discrete time case, we have

\[ c = \frac{u}{s} \ln M_X \left( \frac{s}{u} \right), \]  

(2.3.6)

and in the compound Poisson process case,

\[ c = \frac{u}{s} \left[ M_X \left( \frac{s}{u} \right) - 1 \right]. \]  

(2.3.7)

If, say, in the compound Poisson case, \( c \) is not equal to but larger than the r.-h.s. of (2.3.7), then the bound \( e^{-\gamma u} \) will be smaller than \( \beta \). From an economic point of view, it is quite understandable. If the premium is larger than required, then the ruin probability will be even smaller than we wish. Formally, it follows from the fact that the r.-h.s. of (2.3.4) is increasing in \( \gamma \), which it turn is true because \( M_X(\gamma) \) is convex. (Graph \( M(z) \) and consider the slope of the line connecting the point \((0, 1)\) and \((z, M(z))\). We skip formalities here.)

The same is true for the discrete time case. A detailed advice on how to show the monotoncity of the r.-h.s. of (2.3.3) is given in Exercise 12.

EXAMPLE 2. Suppose time is discrete, and \( X \) is \((m, \sigma^2)\)-normal. In this case, \( \ln M_X(z) = mz + \sigma^2 z^2/2 \), and the reader can readily get from (2.3.6) that

\[ c = m + \frac{s \sigma^2}{2} \cdot \frac{1}{u}. \]  

(2.3.8)
The same may also be obtained by substituting $\gamma$ in (2.3.5) by the explicit expression for $\gamma$ in (2.2.11).

The curve (2.3.8) is a hyperbola. Its graph is the boundary of the area depicted in Fig. 7. For all points $(u, c)$ above the curve (2.3.8) in Fig. 7, the ruin probability is less than $\beta$.

We see also that $c \to m$ as $u \to \infty$. This has a natural interpretation. If the initial surplus is large, then the security loading may be small, i.e., close to zero. This means that the premium per claim may be close to the expected loss per claim. The last quantity is $E\{X\} = m$.

Let, say, $m = 10$ and $\sigma^2 = 4$. To make our formulas nicer, we choose as $\beta$ not 0.05 or 0.01, but, say, $e^{-4} \approx 0.01832$. Then $s = 4$, which will make calculations simpler.

(The choice of $\beta$ is rather subjective anyhow. If, for example, $\beta = 0.05$ seems proper, one may choose $\beta = e^{-3} \approx 0.049787$, which is very close to 0.05.)

For $u = 8$, the premium $c$ should be equal to 11, that is, the relative loading is 10%. If it is too much, but we should keep the ruin probability at the same level, then we should increase the initial surplus. For example, for the 5% loading, $c = 10.5$ and (2.3.8) leads to $u = 16$.

Two more things are noteworthy. First, the fact that $c \to \infty$ as $u \to 0$ in (2.3.8) should not mislead us. This does not reflect the real situation, but rather the circumstance that we are dealing with an estimate which is not accurate for small $u$. Certainly, even if the initial surplus equals zero, for any $c > m$, there will be no ruin with some positive probability. If the premium $c$ is large (but not infinitely large), the ruin probability should be small. So, if $u = 0$, we do not need $c$ to be infinitely large for the ruin probability to be smaller than the level $\beta$.

On the other hand, for large $u$, (2.3.8) gives a good approximation not only for the normal case but for practically arbitrary $X$’s.

Indeed, a large $u$ leads to a small $\gamma$; see (2.3.5). As we saw in Section 2.2.2, in this case, one can use the approximation (2.2.14). This leads to (2.3.8) as an approximation. The larger $u$ is, the smaller $\gamma$ is, and hence the better the accuracy of this approximation. Certainly, this argument is heuristic, and for rigorous estimation we should evaluate the accuracy of the approximation quantitatively.

**EXAMPLE 3.** Consider the compound Poisson process with $X$’s uniformly distributed on $[0, 1]$. To make the example illustrative, set $\beta = e^{-4} \approx 0.018$ (see reasoning on this point in Example 2). Then $s = 4$. As we know, to evaluate the adjustment coefficient, we do not need any information about the intensity of the Poisson process.

In our case $M_X(z) = (e^z - 1)/z$, and (2.3.7) amounts to

$$c = \frac{u^2}{16} \left( \exp \left( \frac{4}{u} \right) - 1 - \frac{4}{u} \right).$$
The graph of this function is the border of the area depicted in Fig.8. All points above the border correspond to the ruin probabilities that are less than $\beta$. Note that $c \to 1/2$ as $u \to \infty$. (Set $x = 4/u$. Then $c = \frac{1}{2}(e^x - 1 - x) \to 1/2$ as $x \to 0$. The last fact may be proved by L'Hôpital's rule.)

As we already noted in Example 2, the convergence mentioned is not surprising. If the initial surplus is large, the premium per claim may be close to the expected loss per claim. In our case, this is $E\{X\} = 1/2$.

As to the fact that $c \to \infty$ as $u \to 0$, the corresponding remark from Example 2 applies to this case too. □

In conclusion, we mention an approximation concerning the case of the homogeneous compound Poisson process and a small security loading. Let $c = (1 + \theta)m$ as in Section 2.2.3, and $m_2 = E\{X^2\}$. Then

$$
\psi(u) - \frac{1}{1 + \theta} \exp \left\{ \frac{-2\theta mn}{(1 + \theta)m_2} \right\} \to 0 \quad \text{as } \theta \to 0, \text{ uniformly in } u > 0. \tag{2.3.9}
$$

This approximation is obtained in [?], and the accuracy of the approximation – in [?] and [?, Section 5.3]. Some refinements of (2.3.9) may be found in [?, Section 5.3].

The main point in (2.3.9) is that the approximation is true for all $u$, including $u$ depending on $\theta$. As above, let $s = \ln(1/\beta)$, and

$$
u = \frac{s}{m_2} \cdot \frac{1 + \theta}{\theta} \sim \frac{s}{m_2} \cdot \frac{1}{\theta} , \tag{2.3.10}
$$

for small $\theta$. Then, by (2.3.9), for small $\theta$,

$$
\psi(u) \approx \frac{1}{1 + \theta} \beta \approx \beta.
$$

Hence, (2.3.10) represents the trade-off between $\nu$ and $\theta$ for small $\theta$ and the security level $\beta$. Certainly, since we consider small $\theta$, the initial surplus $u$ is large.

Note also that (2.3.10) does not contradict what we got before. Since $M_X(z) = 1 + mz + m_2(z^2/2) + o(z^2)$ for small $z$ [see (?)], from (2.3.7) we get that for large $u$

$$
c = (1 + \theta)m = m + s \cdot \frac{m_2}{2u} + o \left( \frac{1}{u} \right).
$$

From the last relation it follows that

$$
\theta = s \frac{m_2}{2um} + o \left( \frac{1}{u} \right),
$$

which is consistent with (2.3.10) for small $\theta$ or (which is equivalent) for large $\nu$. ▶

### 2.4 Three cases when the ruin probability may be computed precisely

As was noted repeatedly, so far we have dealt only with estimates of the ruin probability. Now we will consider cases when the denominator in the main result (2.1.9) may be computed explicitly.
2.4.1 The case when the size of a separate claim is exponentially distributed

Let us consider the model \( R_t = u + ct - S(t) \); time may be discrete or continuous. Suppose claims \( X \)'s are exponentially distributed, \( E\{X\} = m \). Denote by \( \tau \) the ruin time. The value of the process \( R_t \) at time \( \tau \) is the r.v. \( R_\tau \).

At the moment \( \tau \) (if it occurs), the process \( R_t \) makes a jump down and crosses zero level. So, \( R_\tau < 0 \). Certainly, the jump may occur only if at this moment a sufficiently large claim arrived. Denote by \( R_{\tau-0} \) the value of the process before the jump, and by \( \tilde{X} \) the size of the jump, which is equal to the claim at the moment of ruin. Then \( R_\tau = R_{\tau-0} - \tilde{X} = -(\tilde{X} - R_{\tau-0}) \), where \( \tilde{X} - R_{\tau-0} \) is the overshoot. See also Fig.9. Denote the overshoot \( \tilde{X} - R_{\tau-0} \) by \( D \).

Clearly, \( D = |R_\tau| \); see again Fig.9.

Suppose the r.v. \( R_{\tau-0} \) assumed a value \( r \). Of course, the distribution of the claim \( \tilde{X} \) at the moment \( \tau \) depends on \( r \), since for ruin to occur the claim \( \tilde{X} \) must be larger then \( r \). So, the distribution of \( \tilde{X} \) is equal to the conditional distribution of the exponential r.v. \( X \), given that \( X > r \). However, in view of the memoryless property, the overshoot \( D = \tilde{X} - r \) does not depend on \( r \), and has the same distribution as each claim \( X \). That is, \( D \) has the exponential distribution with the same parameter \( a = 1/m \), where as usual \( m = E\{X\} \). (See also Section ???)

On the other hand, since \( D = |R_\tau| \), the denominator in (2.1.9) is

\[
E\{\exp\{-\gamma R_\tau\} | \tau < \infty\} = E\{\exp\{\gamma|R_\tau|\} | \tau < \infty\} = E\{\exp\{\gamma D\} | \tau < \infty\}.
\]

As we saw, the r.v. \( D \) does not depend on \( \tau \), and hence

\[
E\{\exp\{\gamma D\} | \tau < \infty\} = E\{\exp\{\gamma D\}\} = M_D(\gamma) = \frac{1}{1-m\gamma}.
\]

Thus, \( E\{\exp\{-\gamma R_\tau\} | \tau < \infty\} = 1/(1-m\gamma) \), and by (2.1.9),

\[
\psi(u) = (1-m\gamma) \exp\{-\gamma u\}.
\]

In the discrete time case, in accordance with (2.2.9), \( \gamma \) is the solution to the equation

\[
e^{-cz} = 1 - mz.
\]

\[\text{FIGURE 9.}\]
An explicit formula for the solution does not exist, but it is easy to solve such an equation numerically (see also Exercise 18).

For the compound Poisson process, (2.2.23) implies

$$\gamma = \frac{\theta}{m(1 + \theta)},$$

which together with (2.4.1) gives an explicit formula:

$$\psi(u) = \frac{1}{1 + \theta} \exp\left\{-\frac{\theta u}{m(1 + \theta)}\right\}.$$  (2.4.4)

Note also that in the discrete time case, if $X_i$ is an aggregate claim in the $i$th period, that is, $X_i$ is the sum of r.v.'s, then it will not be realistic to assume that the distribution of $X_i$ is exponential.

### 2.4.2 The case of the simple random walk

It is useful to check that Theorem 2 in this case leads to the result of Section ???. In the model of Section ???, we deal not with claims but with increments of the surplus process. Accordingly, the claim surplus process in our case is $W_t = Y_1 + \ldots + Y_t$, where the increment of the total claim surplus at the moment $i$ is $Y_i = -1$ or 1, with probabilities $p$ and $q = 1 - p$, respectively. ($Y_i$ indicates a loss, so when $Y_i = -1$, the surplus process moves up.) We assume $p > 1/2$.

Let $u$ be an integer. Since the process $R_t$ each time jumps up or down exactly by one unit, at the moment of ruin (if any), $R_\tau = -1$, and hence $E\{\exp\{-\gamma R_\tau\}\mid \tau < \infty\} = \exp\{\gamma\}$. Then, by (2.1.9),

$$\psi(u) = \frac{e^{-\gamma u}}{e^\gamma} = e^{-\gamma(u+1)}. $$  (2.4.5)

On the other hand, in this case, equation (2.2.8) amounts to

$$pe^{-z} + qe^z = 1.$$  (2.4.6)

Setting $e^z = x$, we rewrite (2.4.6) as $p + qx^2 = x$. There are two solutions to this equation: $x = 1$ and $x = p/q$. Since we are looking for a positive $z$, we should choose the latter solution. Hence $\gamma = \ln(p/q)$. Substituting it into (2.4.5) we have

$$\psi(u) = \frac{(q/p)^{u+1}}{r^{u+1}}, $$  (2.4.7)

where, as in Section ???, $r = (1 - p)/p$.

The difference between (2.4.7) and the formula $\psi(u) = r^u$ in Section ?? is explained by the fact that in Section ?? we defined the ruin time as the moment when the process first reaches zero level, while here the ruin time is the moment when the process takes on a negative value. In the framework of Section ??, the latter definition corresponds to the replacement of the initial capital $u$ by $u + 1$, which would lead to (2.4.5).
2.4.3 The case of Brownian motion

As in Section ??, let \( R_t = u + \mu t + \sigma w_t \), where \( w_t \) is a standard Brownian motion, and \( \mu > 0 \). Since \( R_t \) is now a continuous process, it will not overshoot zero level, but first will hit (touch) it. In this case, it is reasonable to redefine the notion of ruin setting

\[
\tau = \min\{t > 0 : R_t = 0\}. \tag{2.4.8}
\]

All results above continue to be true in this case.

As follows from the remark at the end of Section ??, once the process reaches zero level at time \( \tau \), in any arbitrary small interval \((\tau, \tau + \delta]\) the process will take negative values infinitely many times, rapidly oscillating around zero for a while. So, the previous definition \( \min\{t > 0 : R_t < 0\} \) is not proper: the minimum does not exist, and we should write \( \inf\{t > 0 : R_t < 0\} \). In view of the continuity of the process, the latter definition coincides with (2.4.8).

Thus, in our case, \( R_\tau = 0 \). Hence, the denominator \( E\{\exp\{-\gamma R_\tau\}|\tau < \infty\} = 1 \), and

\[
\psi(u) = \exp\{-\gamma u\}. \tag{2.1.1}
\]

Consequently,

\[
\psi(u) = \exp\{-2\mu u/\sigma^2\},
\]

which coincides with (??).

The precise formula for \( \psi_T(u) = P(\tau \leq T) \) was obtained in Section ??.

2.5 The martingale approach. A generalization of Theorem 2

We consider now the ruin problem in the martingale framework, presupposing that the reader is familiar with the basic notions of Section ???. To the author’s knowledge, the first use of martingales in Actuarial Modeling is due to H. Gerber (see, e.g., [?], [?]) and F. DeVylder [?].

The main goal of this section is not to prove Theorem 2, although we will do that. Rather, it is to show that this theorem is not a tricky analytical fact, but a direct and almost obvious corollary from the martingale stopping property.

Assume that all processes under consideration are functions of an original process \( \xi_t \) as it was defined in Section ???. As such a process, we can take the surplus process \( R_t \) itself, but it is more convenient to define the original process separately.

As in Section ??, we denote by \( \xi_t \) the collection \( \{\xi_u; 0 \leq u \leq t\} \), i.e., the whole history of the process until time \( t \).

In the framework of this section, the surplus process \( R_t \) is a rather general process. In particular, we will not assume that it is a process with independent increments. However, let us first look at what will happen when this condition holds.
It is convenient to define the independence of increments in terms of $\xi'$. More specifically, assume for a while that

for any $t$ and any interval $\Delta = (t, t + s]$, the r.v. $R_{\Delta}$ does not depend on $\xi'$. \hfill (2.5.1)

(Here, as in Section ??, $R_\Delta = R_{t+s} - R_t$.)

Since $R_t$ is completely determined by $\xi'$, from (2.5.1) it follows that $R_\Delta$ does not depend on $R_t$. Vice versa, if we take $R_t$ as the original process (and we can do that), then property (2.5.1) will follow from the independence of increments of $R_t$.

As in Section 2.1.4, let $W_t = u - R_t$, the claim surplus process. Note that condition (2.1.5) is equivalent to

$$W_t \xrightarrow{P} -\infty \text{ as } t \rightarrow \infty. \hfill (2.5.2)$$

If property (2.5.1) holds for $R_t$, it holds for $W_t$ too. Then, for any $z, t, s$ and interval $\Delta = (t, t + s]$,

$$E\left\{\exp^{W_{t+s}} | \xi'\right\} = E\left\{\exp^{W_t} \exp^{W_s} | \xi'\right\} = \exp^{W_t} E\left\{\exp^{W_s} | \xi'\right\} = \exp^{W_t} E\left\{\exp^{W_s}\right\} = \exp^{W_t} M_\Delta(z), \hfill (2.5.3)$$

where we denote by $M_\Delta(z)$ the m.g.f. of $W_\Delta$.

Thus, if $M_\Delta(z) = 1$, then $E\left\{\exp\{zW_{t+s}\} | \xi'\right\} = \exp\{zW_t\}$. As we know, there may be only one positive solution $\gamma$ (if any) to the equation $M_\Delta(z) = 1$. For $\gamma$ so defined, let $Y_t = \exp^{W_t}$. Then $E\{Y_{t+s} | \xi'\} = Y_t$, and hence

the process $Y_t = \exp^{W_t}$ is a martingale. \hfill (2.5.4)

As a matter of fact, (2.5.4) together with (2.5.2) is the only thing we need. So, we may refuse condition (2.5.1), adopting (2.5.4) itself as the original condition. As we saw, if (2.5.1) is true, then (2.5.4) is also true, but certainly the process $Y_t$ may be a martingale while increments of $R_t$ are dependent.

Thus, regarding the claim surplus process $W_t$, we eventually assume that

(a) $W_0 = 0$, which is more than natural;

(b) the condition (2.5.2) holds;

(c) there exists a number $\gamma$ for which (2.5.4) is true.

Our next step is to apply the martingale stopping property (??). If we had had a right to do that, we would have written

$$1 = E\{e^0\} = E\{e^{W_0}\} = E\{Y_0\} = E\{Y_\tau\} = E\{\exp\{\gamma(u - R_\tau)\}\} = \exp^{W_t} E\{\exp\{-\gamma R_\tau\}\}. \hfill (2.5.5)$$

Then it would have remained to recall that $\tau$, and hence $R_\tau$, is an improper (or defective) r.v., that is, $P(\tau < \infty) < 1$. If ruin does not occur, then we can say that $\tau = \infty$. Thus, we can write – at a somewhat heuristic level – that

$$E\{\exp\{-\gamma R_\tau\}\} = E\{\exp\{-\gamma R_\tau\} | \tau < \infty\} P(\tau < \infty) + E\{\exp\{-\gamma R_\tau\} | \tau = \infty\} P(\tau = \infty). \hfill (2.5.6)$$
7. GLOBAL CHARACTERISTICS OF THE SURPLUS PROCESS

In view of (2.5.2), \( R_t = u - W_t \overset{P}{\rightarrow} +\infty \) as \( t \rightarrow \infty \). Hence, if \( \tau = \infty \), we can set (again reasoning a bit heuristically) that \( R\tau = \infty \), and hence \( E\{\exp(-\gamma R\tau) \mid \tau = \infty \} = 0 \).

Thus, from (2.5.5) and (2.5.9) it follows that
\[
1 = e^{\gamma u}E\{\exp(-\gamma R\tau) \mid \tau < \infty \}P(\tau < \infty),
\]
and we come to the basic result (2.1.9):
\[
P(\tau < \infty) = \frac{e^{-\gamma u}}{E\{\exp(-\gamma R\tau) \mid \tau < \infty \}}. \tag{2.5.7}
\]

The problem is that, since \( \tau \) is improper and is, consequently, not a stopping time, we cannot apply the martingale stopping property directly. This obstacle, however, may be easily overcome if we apply a sort of truncation, which we do in the proof below.

So, we state and prove the following theorem.

**Theorem 4** Let the above conditions (a)-(c) hold. Then (2.5.7) is true.

**Proof.** Let a fixed \( T > 0 \), and \( \tau_T = \min\{T, \tau\} \). Since \( T \) is fixed, \( \tau_T \) is bounded, and Condition 1 of Theorem ?? holds. Applying this theorem and using the notion of expectation over a set (see Section ??), we have
\[
1 = E\{Y_0\} = E\{Y_{\tau_T}\} = E\{\exp(\gamma W_{\tau_T})\} = E\{\exp(\gamma W_T); \tau \leq T\} + E\{\exp(\gamma W_T); \tau > T\}
= E\{\exp(\gamma W_T); \tau \leq T\} + E\{\exp(\gamma W_T); \tau > T\}
= E\{\exp(\gamma(u - R_T)); \tau \leq T\} + E\{\exp(\gamma W_T); \tau > T\}. \tag{2.5.8}
\]

Let \( T \rightarrow \infty \). Then the first term in (2.5.8)
\[
E\{\exp(\gamma(u - R_T)); \tau \leq T\} \rightarrow e^{\gamma u}E\{\exp(-\gamma R_T); \tau < \infty\}
= e^{\gamma u}E\{\exp(-\gamma R_T) \mid \tau < \infty\}P(\tau < \infty). \tag{2.5.9}
\]

It remains to prove that the second term in (2.5.8) vanishes as \( T \rightarrow \infty \). Note that by the definition of \( \tau \), if \( T < \tau \), then \( R_T \geq 0 \), and \( W_T = u - R_T \leq u \). Fixing a number \( k > 0 \), we have
\[
E\{\exp(\gamma W_T); \tau > T\} = E\{\exp(\gamma W_T); \tau > T, W_T > -k\} + E\{\exp(\gamma W_T); \tau > T, W_T \leq -k\}
\leq E\{\exp(\gamma u); \tau > T, W_T > -k\} + E\{\exp(-\gamma k); \tau > T, W_T \leq -k\}
\leq e^{\gamma u}E\{1; W_T > -k\} + e^{-\gamma k}E\{1; \tau > T, W_T \leq -k\}
\leq e^{\gamma u}E\{1; W_T > -k\} + e^{-\gamma k}E\{1\} = e^{\gamma u}P(W_T > -k) + e^{-\gamma k}. \tag{2.5.10}
\]

Let \( T \rightarrow \infty \). Condition (2.5.2), by definition, means that \( P(W_T > -k) \rightarrow 0 \) as \( T \rightarrow \infty \), for any fixed \( k \). Hence, from (2.5.10) it follows that
\[
\lim_{T \to \infty} E\{\exp(\gamma W_T); \tau > T\} \leq e^{-\gamma k}
\]
for any \( k \). The term on the left does not depend on \( k \), so we can let \( k \to \infty \), which for \( \gamma > 0 \) implies that the limit on the left is zero. ■
2.6 The renewal approach

2.6.1 The first surplus below the initial level

In this subsection, we consider the compound Poisson process case, assuming that $R_t = u + ct - S(t)$, the process $S(t)$ is defined in (2.2.19), the process $N_t$ is a homogeneous Poisson process with a constant intensity $\lambda$, and $c = (1 + \theta)m\lambda$, where $m = E\{X_i\} > 0$. When considering a separate $X_i$, we will omit the index $i$.

Since it does not make sense to consider claims equal to zero, we assume also that $P(X > 0) = 1$.

It will be convenient for us to indicate the dependence of the ruin time on $u$ explicitly, so we set $\tau_u = \min\{t : R_t < 0 | R_0 = u\}$.

Let $q$ be the probability that the process will ever fall below the initial level $u$. It may happen if and only if the process $ct - S(t)$ falls below zero level. Hence, $q$ does not depend on $u$, and equals $\psi(0)$, the ruin probability in the case when the initial surplus equals zero.

For the same reason, the size of the drop below the level $u$ at the moment when the process first crosses this level does not depend on $u$ either. The distribution of the size of the drop mentioned coincides with the distribution of $|R_{\tau_0}|$, the absolute value of the deficit of the surplus at the ruin time, if the process starts from zero.

It proves that $q$ and the distribution of $|R_{\tau_0}|$ may be represented in a simple form. Let $F(x)$ be the d.f. of $X$’s.

**Theorem 5** For any $x \geq 0$,

$$P(\tau_0 < \infty, |R_{\tau_0}| \leq x) = \frac{1}{(1 + \theta)m} \int_0^x (1 - F(y))dy. \quad (2.6.2)$$

We prove (2.6.2) in Section 2.6.4, and now we discuss several interesting corollaries from this theorem.

First, setting $x = \infty$ in (2.6.2), we get that

$$P(\tau_0 < \infty) = \psi(0) = \frac{1}{(1 + \theta)m} \int_0^\infty (1 - F(y))dy = \frac{1}{1 + \theta}. \quad (2.6.3)$$

by virtue of the formula (2.2.19). Recall also that $q = \psi(0)$.

The formula (2.6.3) is very interesting – the ruin probability for $u = 0$ depends only on the security loading, and does not depend on the distribution of $X$’s at all.

Now note that, since $\tau_0$ is an improper r.v., i.e., $P(\tau_0 < \infty) = \psi(0) < 1$, the overshoot (or deficit at the moment of ruin) $|R_{\tau_0}|$ is also improper: it is defined only in the case $\tau_0 < \infty$. Let us consider the conditional distribution of $|R_{\tau_0}|$ given $\tau_0 < \infty$, more precisely the conditional d.f.

$$F_1(x) = P(|R_{\tau_0}| \leq x | \tau_0 < \infty).$$

From (2.6.2)-(2.6.3) it follows that

$$F_1(x) = \frac{P(\tau_0 < \infty, |R_{\tau_0}| \leq x)}{P(\tau_0 < \infty)} = \frac{1}{m} \int_0^x (1 - F(y))dy.$$
The conditional density equals
\[ f_1(x) = F_1'(x) = \frac{1}{m} (1 - F(x)). \] (2.6.4)

### 2.6.2 The renewal approximation

Let us return to \( R_t \). Starting from level \( u \), with probability \( q = 1/(1 + \theta) \), the process at some time will drop below the initial level \( u \). If it happens, the size of the drop below the level \( u \) will be a r.v. \( Y_1 \) having the above d.f. \( F_1(x) \).

The process is homogeneous, and the time between consecutive drops have the lack of memory property. Consequently, after the drop mentioned, the process will start to run as if it is at the beginning, with the exception that the starting position is now \( L_1 = u - Y_1 \). See Fig.10.

Since the next drop cannot occur immediately after the first drop, starting from \( L_1 \), the process will be moving up for a while. Hence, \( L_1 \) is a local minimum of the process.

The process will fall below the new level \( L_1 \) with the same probability \( q \). If it happens, the size of the new drop below the level \( L_1 \) will be a r.v. \( Y_2 \) which will not depend on \( Y_1 \), and will have the same distribution \( F_1 \). The value of the process at the moment when it falls below the level \( L_1 \), is \( L_2 = u - Y_1 - Y_2 \). See again Fig.10.

Continuing to reason in the same way, we define the r.v. \( Y_n \) as the size of the \( n \)th drop below the previous \((n-1)\)th level, and the r.v. \( L_n = u - Y_1 - ... - Y_n \). The r.v. \( L_n \) is the \( n \)th local minimum of \( R_t \).

The process \( L_n \) is called a renewal process, and values of \( L_n \) – record values. See Fig.10.

Since the probability of falling below the current value, that is, \( q \), is less than one, the sequence of record values, or drops, is not infinite, but will run up to the moment when the process leaves the lowest level, and will never fall below it. Denote by \( K \) the total number of record values, not counting \( u \). Then \( P(K = n) = q^n p \), where \( p = 1 - q \) so \( K \) has a geometric distribution. Then the lowest level of the process \( R_t \) is

\[ L = \min_t R_t = L_K = u - Z_K, \]

where \( Z_K = \sum_{k=1}^{K} Y_k \), and the r.v.’s \( Y_k \) are independent and have the common d.f. \( F_1 \).
If $K = 0$, that is, the process never falls below the initial level, then we set $Z_K = 0$, and $L = u$. It occurs with probability $p = 1 - q$.

It is easy to understand that the non-ruin probability

$$\phi(u) = P(L \geq 0) = P(Z_K \leq u), \quad (2.6.5)$$

and we have come to a familiar object, namely, the distribution of the sum of a random number of independent r.v.’s.

The ruin probability $\psi(u) = 1 - \phi(u)$.

In accordance with (2.6.5) and (2.6.3),

$$\phi(u) = p \sum_{n=0}^{\infty} q^n F_{1^n}(u), \text{ where } p = \frac{\theta}{1 + \theta} \text{ and } q = \frac{1}{1 + \theta}. \quad (2.6.6)$$

We can apply to (2.6.6) methods of Section ???.

EXAMPLE 1. Let $X$’s take on values 1 or 2 with probabilities 0.75 and 0.25, respectively. The reader is invited to verify that in this case, $m = 1.25$, and the density

$$f_1(x) = \frac{1}{m} (1 - F(x)) = \begin{cases} 0.8 & \text{if } x \in [0, 1] \\ 0.2 & \text{if } x \in (1, 2], \end{cases}$$

and equals 0 otherwise. This means that $f_1(x) = 0.8g_1(x) + 0.2g_2(x)$, where $g_1(x)$ and $g_2(x)$ are the densities of the uniform distributions on $[0, 1]$ and $(1, 2]$, respectively. In other words, $f_1$ is a mixture of uniform distributions; see also Exercise 19.

Let $\theta = 0.2$. Then $q = \frac{1}{1 + 8} = \frac{5}{6}$. For an integer $k$, the part of (2.6.6) corresponding to summation $\sum_{n=k+1}^{\infty}$ does not exceed $q^{k+1}$; see Section ???. for detail. For example, $q^{26} \leq 0.009$, and if we are satisfied with such an accuracy, we can restrict ourselves to $\sum_{n=0}^{25}$.

Numerical estimation of such a sum is not a very complicated problem. Denoting by $G_1, G_2$ the corresponding uniform d.f.’s, for the convolution $F_{1^n}^*$ we can write

$$F_{1^n}^* = (0.8G_1 + 0.2G_2)^n = \sum_{k=0}^{n} \binom{n}{k} (0.8)^k (0.2)^{n-k} G_1^*G_2^*;$$

see (???.). There exist explicit, though cumbersome, formulas for convolutions of uniform distributions, so with good software one should not have a problem in calculations. $\square$

In accordance with (2.6.5), $\phi(u)$ is the d.f. of $Z_K$. Next, we compute the m.g.f. of $Z_K$ or, equivalently, that of its d.f. $\phi(u)$. By (2.6.5) and/or (2.6.6),

$$M_\phi(z) = \int_0^\infty e^{zu} d\phi(u) = \frac{p}{1 - qM_Y(z)}, \quad (2.6.7)$$

where $M_Y(z)$ is the m.g.f. of the r.v.’s $Y_i$. Using (2.6.4) and integrating by parts, we get that for all $z > 0$,

$$M_Y(z) = \frac{1}{m} \int_0^\infty e^{zx} (1 - F(x)) dx = -\frac{1}{mz} (1 - F(0)) + \frac{1}{mz} \int_0^\infty e^{zx} dF(x) = \frac{1}{mz} [M_X(z) - 1],$$

2. Ruin Models
because, in view of (2.6.1), $F(0) = 0$. Inserting this into (2.6.7), and substituting values for $p$ and $q$, it is easy to calculate that

$$M_{\phi}(z) = \frac{\theta mz}{1 + (1 + \theta)mz - M_X(z)}, \quad (2.6.8)$$

If for a particular $X$, the m.g.f. (2.6.8) is familiar for us, we can determine $\phi(u)$. For some cases, it is convenient to rewrite (2.6.8) as

$$M_{\phi}(z) = \frac{\theta}{1 + \theta} + \frac{1}{1 + \theta} \cdot \frac{\theta (M_X(z) - 1)}{1 + (1 + \theta)mz - M_X(z)}. \quad (2.6.9)$$

The last formula reflects the following circumstance. With probability $p$ the r.v. $K = 0$, and since in this case $Z_K = 0$, the d.f. of $Z_K$ – that is, $\phi(u)$ – makes a jump of $p$ at zero. In view of (2.6.6), this may be represented in the following form:

$$\phi(u) = p + p \sum_{n=1}^{\infty} q^n F^*(n)(u). \quad (2.6.10)$$

The two terms in (2.6.9) correspond to the respective two terms in (2.6.10).

**EXAMPLE 2.** Let $F(x)$ be a mixture of exponential distributions, say, the tail $F(x) = 1 - F(x) = \frac{1}{2} e^{-x} + \frac{1}{2} e^{-x/3}$. Then $m = \frac{1}{2} \cdot 1 + \frac{1}{2} \cdot 3 = 2$, and for $z < 1/2$,

$$M_X(z) = \frac{1}{2} \cdot \frac{1}{1 - z} + \frac{1}{2} \cdot \frac{1}{1 - 3z}.$$  

To make calculations more illustrative, set $\theta = 0.5$ in our example, realizing that it is not very realistic. Substituting it into (2.6.9), the reader can verify that in this case,

$$M_{\phi}(z) = \frac{1}{3} + \frac{2}{3} \cdot \frac{2 - 3z}{2 - 18z + 18z^2}. \quad (2.6.11)$$

The equation $2 - 18z + 18z^2 = 0$ has two solutions: $z_1 = 0.5 + \sqrt{3}/6 \approx 0.87$, and $z_2 = 1 - z_1 \approx 0.13$.

Using the method of partial fractions we write

$$\frac{2 - 3z}{2 - 18z + 18z^2} = \frac{2 - 3z}{18(z_1 - z)(z_2 - z)} = \frac{1}{18} \left( \frac{c_1}{z_1 - z} + \frac{c_2}{z_2 - z} \right), \quad (2.6.12)$$

where $c_1$, $c_2$ are constants that we should find. Putting the r.-h.s. of (2.6.12) into the common denominator, we get that $c_1 + c_2 = 3$, $c_1 z_2 + c_2 z_1 = 2$.

We will write all solutions up to the second digit. Solving the equations for $c_1$ and $c_2$, we readily get that $c_1 = 0.83$, $c_2 = 2.17$. Then

$$\frac{2 - 3z}{2 - 18z + 18z^2} = \left( \frac{0.05}{1 - z/z_1} + \frac{0.95}{1 - z/z_2} \right), \quad (2.6.13)$$
where the denominators are precise but the coefficients in the numerators are computed up to the second digit. Together with (2.6.12), we have with the same accuracy that

\[
M_{\theta}(z) = \frac{1}{3} + \frac{2}{3} \left( \frac{0.05}{1 - z/z_1} + \frac{0.95}{1 - z/z_2} \right).
\]

The last term is a mixture of exponential m.g.f.'s. Consequently,

\[
\phi(u) = \frac{1}{3} + \frac{2}{3} (0.05 F_{z_1}(u) + 0.95 F_{z_2}(u)),
\]

where \( F_z \) stands for the exponential d.f. with parameter \( z \). Eventually,

\[
\psi(u) = 1 - \phi(u) = \frac{2}{3} (0.05 F_{z_1}(u) + 0.95 F_{z_2}(u))
\]

\[
= \frac{2}{3} (0.05 \exp\{-z_1 u\} + 0.95 \exp\{-z_2 u\}) \approx 0.03 \exp\{-0.87 u\} + 0.63 \exp\{-0.13 u\}.
\]

\[
\square
\]

2.6.3 The Cramér-Lundberg approximation

In conclusion, we present without a proof one more celebrated result of Risk Theory.

**Theorem 6** Let \( \gamma > 0 \) be the adjustment coefficient satisfying (2.2.20). Then,

\[
\psi(u) \sim C e^{-\gamma u} \quad \text{as} \quad u \to \infty,
\]

(2.6.14)

where

\[
C = \frac{m\theta}{M_X'(\gamma) - m(1 + \theta)}.
\]

(2.6.15)

Proofs may be found, e.g., in [?] or [?].

To clarify the significance of the last formula, assume that \( X \) is exponential. Then \( M_X(z) = 1/(1 - mz) \), and \( M_X'(z) = m/(1 - mz)^2 \). By (2.4.3), \( \gamma = \theta/[m(1 + \theta)] \). Substituting it into (2.6.15), we readily get \( C = 1/(1 + \theta) \), which is consistent with the precise formula (2.4.4).

Another example is given in Exercise 23.

The reader may find further interesting approximations for the ruin probability and ruin time as well as further references, e.g., in [?] and [?].

2.6.4 Proof of Theorem 5 from Section 2.6.1

Usually, this theorem is proved with the use of differential equations. Below, we mainly follow a different proof from [?] by S. Asmussen. In part, we do it for diversity, but also because the latter proof is direct and illustrative.
7. GLOBAL CHARACTERISTICS OF THE SURPLUS PROCESS

So, we consider the case when $R_0 = 0$. Set $\tau = \tau_0$, and for a set $A$ from the real line denote by $I_A(x)$ the indicator $A$ — that is, $I_A(x) = 1$ if $x \in A$ and $I_A(x) = 0$ otherwise. Let

$$\eta_A = \int_0^\tau I_A(R_t)dt.$$  \hfill (2.6.16)

Since $I_A(R_t) = 1$ or 0, depending whether $R_t$ got into $A$ or not, the r.v. $\eta_A$ is equal to the amount of time the process $R_t$ spent in $A$ before the time moment $\tau$, that is, before ruin. Our proof is based on

**Lemma 7** For any bounded set $A$,

$$E\{\eta_A\} = \frac{1}{c}|A|,$$

where $|A|$ is the length of $A$ (and $c = (1 + \theta)m\lambda$, the premium rate).

We will prove this in the end of the section. From Lemma 7 and (2.6.16) it follows that

$$|A| = cE\{\eta_A\} = cE\left\{\int_0^\tau I_A(R_t)dt\right\}. \hfill (2.6.17)$$

Next, we show that (2.6.17) implies that for any bounded function $g(y)$ defined and integrable on $[0, \infty)$,

$$\int_0^\infty g(v)dv = cE\left\{\int_0^\tau g(R_t)dt\right\}. \hfill (2.6.18)$$

If $g(v) = I_A(v)$, then (2.6.18) coincides with (2.6.17). Consider a piecewise constant function

$$g(v) = \sum_k g_k I_{A_k}(v), \hfill (2.6.19)$$

where $g_1, g_2, \ldots$ are numbers, and $A_1, A_2, \ldots$ are disjoint sets. The function $g(v)$ takes on the constant value $g_k$ if $v \in A_k$. Since sets $A_k$ are disjoint,

$$\int_0^\infty g(v)dv = \sum_k g_k|A_k|. \hfill (2.6.20)$$

By (2.6.17),

$$\int_0^\infty g(v)dv = \sum_k g_k|A_k| = \sum_k g_k cE\left\{\int_0^\tau I_{A_k}(R_t)dt\right\} = cE\left\{\int_0^\tau \sum_k g_k I_{A_k}(R_t)dt\right\} = cE\left\{\int_0^\tau g(R_t)dt\right\},$$

which proves (2.6.18) for any function of the type (2.6.19). Since any bounded function may be approximated with any desired accuracy by a piecewise constant function, (2.6.18) is true for any bounded $g(y)$.

Having (2.6.18), we can turn to the direct proof of Theorem 5. Our reasoning is close to what we already did in Section 2.4.1.
The process $R_t$ jumps down during an infinitesimally small interval $[t, t + dt]$ only if $N_t$ jumps up (a claim arrives). In accordance with (2.6.21), the probability that this happens equals $\lambda dt$. Denote by $R_{t-0}$ the value of the process before this jump, and by $\tilde{X}$ the size of the jump (that is, the claim). We omit an index in $\tilde{X}$. Since $dt$ is infinitesimally small, we may identify the value of the process after the jump with $R_t$, so $R_t = R_{t-0} - \tilde{X}$; see also Fig.11.

Consider the event $\mathcal{E}_{tx}(dt)$ consisting in the following:

(i) During an interval $(t, t + dt]$ ruin occurred.

(ii) It occurred at the first time, and hence $t < \tau$ and $R_{t-0} > 0$.

(iii) The overshoot $|R_t|$ has exceeded some level $x > 0$.

Then $P(\mathcal{E}_{tx}(dt)) = P(t < \tau, \tilde{X} > R_{t-0} + x)\lambda dt$.

Let $I(\mathcal{E})$ stands for the indicator of an event $\mathcal{E}$, that is, $I(\mathcal{E}) = 1$ if $\mathcal{E}$ occurs, and $= 0$ otherwise. (In the function $I_A(x)$ above, $A$ is a set from the real line, while $\mathcal{E}$ is an event in the original space $\Omega$ of elementary outcomes.)

Using the formula for total expectation (2.6.21), we can write that $P(\mathcal{E}_{tx}(dt)) = E\{I(t \leq \tau)I(\tilde{X} > R_{t-0} + x)\} \lambda dt = E\{I(t \leq \tau)E\{I(\tilde{X} > R_{t-0} + x)\mid R_{t-0}, I(t \leq \tau)\}\} \lambda dt = E\{I(t \leq \tau)E\{\tilde{F}(R_{t-0} + x)\mid R_{t-0}, I(t \leq \tau)\}\} \lambda dt$.

Given $R_{t-0}$ and $t < \tau$, the conditional probability $P(\tilde{X} > R_{t-0} + x \mid R_{t-0}, I(t \leq \tau))$ is the probability that the amount of a claim will be larger than $R_{t-0} + x$. Then, setting $\tilde{F}(x) = P(X_t > x)$, where $X_t$ is a claim, we have

$$P(\mathcal{E}_{tx}(dt)) = E\{I(t \leq \tau)\tilde{F}(R_{t-0} + x)\} \lambda dt. \quad (2.6.21)$$

For a fixed $t$, the probability that a jump occurs at time $t$ is zero. Consequently, for a fixed $t$, the distributions of the r.v.'s $R_{t-0}$ and $R_t$ are the same. Then we can replace $R_{t-0}$ by $R_t$ in the right member of (2.6.21). Thus,

$$P(\mathcal{E}_{tx}(dt)) = E\{\tilde{F}(R_t + x)I(t \leq \tau)\} \lambda dt.$$ 

Summing up the probabilities $P(\mathcal{E}_{tx}(dt))$, or more precisely, integrating in $t$, we have

$$P(\tau < \infty, |R_t| > x) = \int_0^{\infty} P(\mathcal{E}_{tx}(dt)) = \lambda \int_0^{\infty} E\{\tilde{F}(R_t + x)I(t \leq \tau)\} dt$$

$$= \lambda E\left\{\int_0^{\infty} \tilde{F}(R_t + x)I(t \leq \tau) dt\right\} = \lambda E\left\{\int_{\tau}^{\infty} \tilde{F}(R_t + x) dt\right\}.$$
Consecutively applying (2.6.18), the fact that $c = (1 + \theta)m\lambda$, and the variable change $y = x + \nu$, we get that

$$P(\tau < \infty, |R_t| > x) = \frac{\lambda}{c} \int_0^{\infty} F(x + \nu) d\nu = \frac{1}{(1 + \theta)m} \int_0^{\infty} F(x + \nu) d\nu$$

(2.6.22)

Setting $x = 0$, and recalling that $m = \int_0^{\infty} F(y) dy$ by virtue of (???)?, we write

$$P(\tau < \infty) = P(\tau < \infty, |R_t| > 0) = \frac{1}{(1 + \theta)m} \int_0^{\infty} F(\nu) d\nu = \frac{1}{1 + \theta}.$$  

(2.6.23)

To get (2.6.2), it remains to subtract (2.6.22) from (2.6.23).

To complete the proof, we should provide

**Proof of Lemma 7.** Let us fix, for a while, $t > 0$ and consider for $s \in [0, t]$ the process $\tilde{R}_s = R_t - R_{t-s}$. The process $\tilde{R}_s$ may be interpreted as $R_t$ in reversed time. Note that $\tilde{R}_0 = 0$, and $\tilde{R}_t = R_t$ because $R_0 = 0$. The process $\tilde{R}_s$ moves up linearly with the same slope $c$, and drops down with the same intensity $\lambda$ as $R_s$. The distribution of jumps is also the same as for $R_s$, and the only difference is that, if $R_s$ has a jump at a point $s_1$, the corresponding jump of $\tilde{R}_s$ occurs at time $t - s_1$; see also Fig.12.

However, the last fact has no effect on the distribution of trajectories of $\tilde{R}_s$, since the intensity of jumps does not depend on time, and jumps are equally likely to occur at any time. Thus, the distribution of the process $\tilde{R}_s$ is the same as for $R_s$, that is, the probability of any collection of possible trajectories of $\tilde{R}_s$ equals the same probability for $R_s$.

On the other hand, $\tilde{R}_t = R_t = \tilde{R}_s + R_{t-s}$. So, $\tilde{R}_t \geq \tilde{R}_s$ for all $s \leq t$ if and only if $R_{t-s} \geq 0$ for all $s \leq t$. This is the same as $R_s \geq 0$ for all $s \leq t$. Thus, for any set $A$, the event

$$\{R_t \in A, \tau < \tau\} = \{R_t \in A, R_s \geq 0 \text{ for all } s \leq t\} = \{\tilde{R}_t \in A, \tilde{R}_s \geq \tilde{R}_s \text{ for all } s \leq t\}.$$  

(2.6.24)

(The last step is true because $\tilde{R}_t = R_t$.) Since the processes $R_s$ and $\tilde{R}_s$ have the same distribution, this implies, in turn, that

$$P(R_t \in A, \tau < \tau) = P(\tilde{R}_t \in A, \tilde{R}_s \geq \tilde{R}_s \text{ for all } s \leq t) = P(R_t \in A, R_s \geq R_t \text{ for all } s \leq t).$$
2. Ruin Models

The time spent in $A$ at the moments of leadership.

![Diagram](image)

**FIGURE 13.** “Thick” segments indicate moments of leadership.

Note also that $I_A(R_t) = I(R_t \in A)$, by definition of $I(E)$. Then, by (2.6.24), for a bounded set $A$,

$$
E \{ \eta_A \} = E \left\{ \int_0^T I_A(R_t) dt \right\} = E \left\{ \int_0^\infty I(R_t \in A) I(t \leq \tau) dt \right\} = E \left\{ \int_0^\infty I(R_t \in A, t \leq \tau) dt \right\}
$$

$$
= \int_0^\infty E \{ I(R_t \in A, t \leq \tau) \} dt = \int_0^\infty P(R_t \in A, t \leq \tau) dt
$$

$$
= \int_0^\infty P(R_t \in A, R_t \geq R_s \forall s \leq t) dt = E \left\{ \int_0^\infty I(R_t \in A, R_t \geq R_s \forall s \leq t) dt \right\} (2.6.25)
$$

Since $A$ is bounded, there exists $M$ such that $A \subset [0, M]$. Denote the last integral in (2.6.25) by $J$. This is the total time when $R_t$ is in $A$, being at the same time the largest value with respect to all previous moments $s \leq t$. We can also call such $t$’s moments of leadership.

If $R_T > M$ at some moment $T$, then “in the future”, for $t > T$, at a leadership moment the value of the process will be larger than $M$, and hence will not be in $A$.

Consequently, $J$ is exactly (!) equal to $|A|/c$, the length of $A$ divided by the slope of $R_t$ at points of growth – see Fig.13. Note also that $J \leq (|A|/c) \leq M/c$ in any case. It remains to use the condition $R_t \xrightarrow{T} \infty$. Let $T > 0$. The last expected value in (2.6.25) equals

$$
E \{ J \} = E \{ J | R_T > M \} P(R_T > M) + E \{ J | R_T \leq M \} P(R_T \leq M)
$$

$$
= (|A|/c) P(R_T > M) + E \{ J | R_T \leq M \} P(R_T \leq M). \quad (2.6.26)
$$

Let $T \to \infty$. The first term in (2.6.26) converges to $|A|/c$, since $P(R_T > M) \to 1$ for any $M$. The second term does not exceed $(|A|/c) P(R_T \leq M) \to 0$ as $T \to \infty$. ■

2.7 Some recurrent relations and computational aspects

Here, we briefly discuss how to compute ruin probabilities for finite time horizons using recursive methods. The relations we consider are based on the first step analysis. We restrict
7. GLOBAL CHARACTERISTICS OF THE SURPLUS PROCESS

ourselves to the case $R_t = u + c_t - S(t)$, where $u$ is the initial surplus, $S(t)$ is the loss process, and $c_t$ is the aggregate amount of (positive) cash collected by time $t$. It will be convenient for us to consider the non-ruin probability $\phi_T(u)$. The ruin probability $\psi_T(u) = 1 - \phi_T(u)$.

We start with a particular problem which requires only common sense.

EXAMPLE 1 ([?, N2]). BIB is a new insurer writing homeowners policies. You are given: (a) Initial surplus $= \$15$. (b) Number of insured homes $= 3$. (c) Premium per home $= \$10$. (d) Premiums are paid at the start of each year. (e) Size of each claim $= \$40$. (e) Claims are paid immediately. (f) There are no expenses. (g) There is no investment income.

Each homeowner files at most one claim per year. The probability that a given homeowner files a claim in year 1 is $20\%$, and in year 2, it is $10\%$. Claims are independent.

Calculate the probability that BIB has positive surplus at the end of year 2.

The insurer will not have a positive income at the end if there is ruin in the middle of the period, so we are computing the non-ruin probability. At the beginning, the insurer has $\$45$, and in order to not be ruined in the first stage, there should not be more than one claim of $\$40$.

If there is no claim in the first period, then the insurer will have $\$75$ at the beginning of the second period, and to have a positive cash, the insurer must not have more than one claim. If there is one claim in the beginning, the insurer will have just $\$35$ at the beginning of the second period. In this case, there will be no ruin only if there is no claim in this period.

The number of claims during each period has a binomial distribution, so the non-ruin probability

$$\phi_2(15) = (0.8)^3 \left[ (0.9)^3 + \binom{3}{1}(0.9)^2(0.1) \right] + \binom{3}{1}(0.8)^2(0.2)(0.9)^3 = 0.7776. \square$$

Now, we present the same logic in a more formal way. First, let time be discrete, and $S(t) = S_t = X_1 + ... + X_t$, where $X_j$ is the size of the $j$th claim, and $X$'s are i.i.d. r.v.'s. For ruin not to happen during time interval $[0, T]$, the first claim $X_1$ should not exceed $u + c_1$, and starting from the new level $u + c_1 - X_1$, the process should not take on negative values during time $T - 1$. We can unify both cases: $X_1 \leq u + c_1$ and $X_1 > u + c_1$, setting by definition the non-ruin probability $\phi_T(u) = 0$, if $u < 0$. Then, given $X_1$, the conditional non-ruin probability during the last $T - 1$ periods after the first period is $\phi_{T-1}(u + c_1 - X_1)$.

In view of the independence of $X$'s, from this it follows that the non-ruin probability

$$\phi_T(u) = E\{\phi_{T-1}(u + c_1 - X_1)\}. \hfill (2.7.1)$$

For $T = \infty$, setting $\phi(u) = \phi_\infty(u)$, we can rewrite (2.7.1) as

$$\phi(u) = E\{\phi(u + c_1 - X_1)\}, \hfill (2.7.2)$$

which is an equation for $\phi(u)$.

\footnote{Reprinted with permission of the Casualty Actuarial Society.}
Consider, for example, the discrete case when $X_i$ take on values $x_1, x_2, \ldots$ with probabilities $f_1, f_2, \ldots$, respectively. Then (2.7.1) may be written as

$$\phi_T(u) = \sum_{j} \phi_{T-1}(u + c_i - x_j) f_j.$$  

(2.7.3)

It is worth emphasizing that in the last sum, as a matter of fact, terms for which $x_j > u + c_i$, vanish.

For $T = \infty$, we may write (2.7.3) as

$$\phi(u) = \sum_{j} \phi(u + c_i - x_j) f_j.$$  

The reader is invited to make sure on her/his own that when $c_i = 1$, and $X_i$ takes on values 0 or 1, the last equation leads to the classical equation for the ruin probability for the simple random walk; see (?? ??).

For $X_i$'s taking many values and for a finite $T$, calculations are not so nice as they were in Section ?? ??, and one should use numerical procedures. Here, we consider only simple examples in order to demonstrate the logic of calculations.

**EXAMPLE 2.** Let the unit of time be a year, and the premium $c = 4$ be paid at the beginning of each year. Assume that the losses $X = 2, 4, 10$ with probabilities $f_1 = 0.5, f_2 = 0.4, f_3 = 0.1$, respectively, are paid at the end of each year. Let $T = 2$. By (2.7.3),

$$\phi_2(u) = \phi_1(u + 4 - 2) \frac{1}{2} + \phi_1(u + 4 - 4) \frac{2}{5} + \phi_1(u + 4 - 10) \frac{1}{10}$$

$$= \frac{1}{10} (5\phi_1(u + 2) + 4\phi_1(u) + \phi_1(u - 6)).$$  

(2.7.4)

Here, it makes sense to consider only integer $u$'s. We have $\phi_1(u) = 1$ for $u = 6, 7, \ldots$. If $u = 0, \ldots, 5$, a ruin may happen in one period only if the biggest claim occurs, so $\phi_1(u) = 0.9$. Thus,

$$\phi_2(u) = \frac{1}{10} (5 \cdot 0.9 + 4 \cdot 0.9 + 0) = 0.81 \quad \text{for } u = 0, \ldots, 3;$$

$$\phi_2(u) = \frac{1}{10} (5 \cdot 1 + 4 \cdot 0.9 + 0) = 0.86 \quad \text{for } u = 4, 5;$$

$$\phi_2(u) = \frac{1}{10} (5 \cdot 1 + 4 \cdot 1 + 0.9) = 0.99 \quad \text{for } u = 6, \ldots, 11;$$

$$\phi_2(u) = \frac{1}{10} (5 \cdot 1 + 4 \cdot 1 + 1) = 1 \quad \text{for } u = 12, 13, \ldots. \quad \Box$$

**EXAMPLE 3.** Consider the same problem but assume that the available surplus is invested with a risk free interest $r$. This means that the cash flow $c_1$ in (2.7.3) should include the growth of the capital, and $u + c$ above should be replaced by $(u + c)\alpha$, where $\alpha = 1 + r$.

Then instead of (2.7.4), we should write

$$\phi_2(u) = \frac{1}{10} (5\phi_1(\alpha(u + 4) - 2) + 4\phi_1(\alpha(u + 4) - 4) + \phi_1(\alpha(u + 4) - 10)),$$
and \( \phi_1(u) \) should also be recomputed. To make calculations illustrative set \( r = 1/9 \). Now we consider all \( u \)'s, not only integers. The function \( g(\alpha) = \alpha(u + 4) \) equals 10 for \( u = 5 \), so \( \phi_1(u) = 1 \) for \( u \geq 5 \). For \( u < 5 \) we have \( \phi_1(u) = 0.9 \). Note also that \( \alpha(u + 4) - 4 \geq 5 \) if \( u \geq 4.1 \), and \( \alpha(u + 4) - 2 \geq 5 \) if \( u \geq 2.3 \). Thus,

\[
\begin{align*}
\phi_2(u) &= \frac{1}{10} (5 \cdot 0.9 + 4 \cdot 0.9 + 0) = 0.81 \quad \text{for } 0 \leq u < 2.3; \\
\phi_2(u) &= \frac{1}{10} (5 \cdot 1 + 4 \cdot 0.9 + 0) = 0.86 \quad \text{for } 2.3 \leq u < 4.1; \\
\phi_2(u) &= \frac{1}{10} (5 \cdot 1 + 4 \cdot 1 + 0.9) = 0.99 \quad \text{for } 4.1 \leq u < 5; \\
\phi_2(u) &= \frac{1}{10} (5 \cdot 1 + 4 \cdot 1 + 1) = 1 \quad \text{for } u \geq 5. \quad \square
\end{align*}
\]

Nothing prevents us from continuing the recurrence procedure. Applying the same formula (2.7.3) to its interior terms we can write

\[
\phi_T(u) = \sum_j \left( \sum_i \phi_{T-i}(u+c_i-x_j-x_i) f_j \right) f_j \\
= \sum_j \sum_i \phi_{T-i}(u+c_i-x_j-x_i) f_j f_i,
\]

(2.7.5)

moving in the same way up to the moment when we will come to \( \phi_0(u) = 1 \) for \( u \geq 0 \), and = 0 for \( u < 0 \). Note that \( c_2 \) is the cumulative cash by time 2, and again terms inside the sum in (2.7.5) are equal to zero if \( u + c_2 - x_j - x_i < 0 \). Calculations may be tedious even if we write a corresponding program, but the program itself should not be too complicated.

Note also that equations (2.7.3)-(2.7.5) are the so called backward equations: we condition the behavior of the process with respect to what may happen in the first period. Another approach may concern the so called forward equations. In this case, we assume that there was no ruin during the first \( T - 1 \) periods, and consider the behavior of the process in the last period. Regarding the later methods, see, e.g., [?, Section 7.3].

The same logic may be applied to processes in continuous time. Assume that \( S(t) \) is a homogeneous compound Poisson process with intensity \( \lambda \). Let \( m \) and \( F(x) \) be the mean value and the d.f., respectively, of a separate claim. As usual, we set \( c_i = (1 + \theta) m \lambda \), where \( \theta \) is a security loading. For simplicity, consider the case \( T = \infty \).

Set \( \eta = \min \{ t : S(t) > 0 \} \), the moment of the first claim, and set \( Z = S(\eta) \), the value of the first claim. For the process under consideration, \( \eta \) is exponential with parameter \( \lambda \), the d.f. of \( Z \) is \( F \), and \( \eta \) and \( Z \) are independent.

Let again \( \phi(u) = \phi_\infty(u) \). As before, we set \( \phi(u) = 0 \) if \( u < 0 \).

At the moment \( \eta \), the conditional non-ruin probability is equal to \( \phi_{T-\eta}(u + c\eta - Z) \). It is equal to zero if \( Z > c\eta + u \).

In view of the memoryless property, at the moment \( \eta \), the process starts over from the new level. Since \( T = \infty \), the time horizon with respect to the new starting moment \( \eta \) is again infinite. Hence,

\[
\phi(u) = E \left\{ \phi(u + c\eta - Z) \right\}.
\]
This is an equation for $\phi(u)$. Since we know the distributions of $Z$ and $\eta$, we can rewrite it as

$$\phi(u) = \int_{0}^{\infty} \int_{0}^{\infty} \phi(u + ct - z) dF(z) \lambda e^{-\lambda t} dt.$$ 

Since $\phi(u) = 0$ for $u < 0$, it may be written as

$$\phi(u) = \lambda \int_{0}^{\infty} e^{-\lambda t} \left( \int_{0}^{u + ct} \phi(u + ct - z) dF(z) \right) dt. \quad (2.7.6)$$

The theory of solutions to equations of this type is well developed and uses various mathematical methods; see, e.g., [?], [?], [?], [?], [?]. All these methods are not very simple but give, in particular, an alternative way to obtain many results we got above by making use of the martingale or renewal approaches.

3 CRITERIA CONNECTED WITH PAYING DIVIDENDS

In the situation described in the previous section, two things may happen:

- either during some finite time period, ruin will occur (for a large initial surplus and/or large premiums, the probability of this event is small), or

- the company will avoid ruin, and the surplus $R_t$ will unlimitedly grow: $R_t \to \infty$ as $t \to \infty$. (See also condition (2.1.5).)

The last property is not realistic: no company will keep an excessive surplus of high liquidity while having an opportunity to invest a part of it or pay dividends. Moreover, a law and a general usual insurance practice requires paying some dividends if the surplus exceeds a certain level. On the other hand, as one may guess and as we will see below, the probability that the company will remain solvent forever would be zero unless the company allows the surplus to grow. In other words, condition (2.1.5) is essential for the ruin probability not to equal one.

To resolve these issues, we should consider, as an alternative to ruin probability, other quality criteria – for example, the expected discounted amount of dividends to be paid or and the expected life of the company. The idea to use these criteria was first aired by B. De Finetti in 1957 and was considered later by K. Borch and other scholars (see, e.g., [?], [?], [?], [?], [?]).

The goal of this section is to illustrate some ideas and results in this area. We restrict ourselves to the discrete time case and consider the model (1.3)-(1.4) from Section 1.
3.1 A general model

Denote by \( d_t \) the dividend paid at time \( t = 1, 2, \ldots \). The surplus process is governed by the relation

\[
R_t = u + ct - S_t - D_t,
\]

where \( S_t = X_1 + \ldots + X_t \), the claims \( X_t \) are i.i.d. r.v.'s, \( D_t = d_1 + \ldots + d_t \), and \( c \) is a premium per unit interval of time. As before, when it does not cause misunderstanding, we omit the index \( i \) in \( X_i \).

Assume that at the moment, if any, when \( R_t < 0 \), the company stops operating.

Let \( v < 1 \) be a discount factor. We consider an infinite time horizon, and the criterion

\[
E \left\{ \sum_{t=1}^{\infty} v^t d_t \right\},
\]

the expected total amount of discounted dividends to be paid.

The variables \( d_t \) represent a strategy of paying dividends. In general, since \( d_t \) may depend on the history of the process until time \( t \), it is a r.v.

Since \( v < 1 \), dividends to be paid in the future are less valuable than payments now. However, it does not mean that the company should pay large dividends in the beginning. If the company pays too much in earlier stages, this will reduce the current surplus and will make possible an earlier ruin. In this case, the total amount of dividends may be small because the time of functioning will be small.

Thus, an optimal strategy should reflect a trade-off between two issues: the desire to pay dividends in earlier stages, and the necessity to keep the company functioning during a sufficiently long period.

We will show in Section 3.3 that under some mild conditions the optimal strategy maximizing (3.1.2) has the following threshold structure.

- If at the end of an underwriting period the surplus \( R_t \) exceeds an optimal threshold level \( z^* \), then the amount \( R_t - z^* \) is paid out as the dividend payment during that period.

- If the surplus \( R_t \) is less than \( z^* \), then no dividends are paid, and the company keeps the surplus \( R_t \) until the next underwriting period.

In other words,

\[
d_t = \max\{R_t - z^*, 0\}.
\]

We will see in Section 3.3 that the optimal level \( z^* \) does not depend on the initial surplus \( u \).

To find the level \( z^* \), we consider the threshold strategy for all \( z \)'s and the function

\[
V(u, z) = E \left\{ \sum_{t=1}^{\infty} v^t d_t \right\},
\]

where \( d_t = \max\{R_t - z, 0\} \). Once we know \( V(u, z) \), we can try to find its maximizer in \( z \), that is, the optimal level \( z^* \). Since this level does not depend on \( u \), we can do that for \( u = 0 \).
Certainly, this does not mean that the amount of dividends itself does not depend on $u$. Let us consider this in more detail.

If the initial level $u > z$, then by definition of the strategy $d_t$, the company should immediately pay off the surplus $u - z$, that is,

$$V(u, z) = V(z, z) \quad \text{for } u > z.$$  

Let $u \leq z$. Since the goal of the company is to maximize the amount of dividends, the initial surplus $u$ that the company keeps for functioning may be viewed as an investment for getting dividends in the future. Then the variable $V(u, z) - u$ may be viewed as the profit of the company. We will prove in Section 3.3 that in the case of the optimal level $z^*$,

the function $V(u, z^*) - u$ is increasing in $u$ when $u < z^*$.  

This means that the optimal behavior is to start with the initial surplus $u = z^*$ and proceed following the optimal threshold strategy.

The next question concerns the ruin probability. We will see that, under the above threshold strategy, it is equal to one, provided that with a positive probability the claim may exceed the premium. Such a condition is natural since otherwise nobody will pay such a premium.

Let $P(X > c + a) \geq \delta > 0$ for some $a > 0$. Let $k = \lceil z^* / a \rceil + 1$, where as usual $[x]$ is the integer part of $x$. Then with the probability $\delta^k$, all claim surpluses $X_t - c$, $t = 1, ..., k$, will be larger than $a$, and $R_k$ will be negative. If it does not happen during the first $k$ steps, then it will happen with the same positive probability during the next $k$ steps, and so on. So, the probability that ruin will ever happen is one.

[More rigorously, let $A_i = \bigcap_{l=ik+1}^{(i+1)k} \{X_t > c + a\}$. For ruin to occur, it suffices that at least one of the events $A_i$ occurs. The probability of this is one, since $A_i$’s are independent and $P(A_i) > \delta^k > 0$.]

The fact that in the case of a threshold strategy the ruin probability equals one is not a reason to refuse the approach above: we nevertheless deal with the maximal amount of dividends. Moreover, if the time before the ruin is sufficiently large, say, larger than the time horizon for the company, the fact mentioned is not essential.

Nevertheless, we can apply a more cautious approach by introducing into consideration the expected time of ruin. Let $D(u, z)$ be the mentioned expected time for the initial surplus $u$ and the threshold level $z$. Having at its disposal both characteristics, $V(u, z)$ and $D(u, z)$, the insurer can establish a more flexible criterion. One example consists in maximizing $V(u, z)$ under the restriction

$$D(u, z) \geq D_0,$$

where $D_0$ is a given level determined by the preferences of the insurer.

Analytical solutions of the problems above are complicated even in simple cases (see, e.g., [?], [?]). So, we restrict ourselves to one example, namely, to the simple random walk model. Results for this model illustrate well what we can expect in more general cases. As to the general situation, it is worth emphasizing that numerical solutions based on simulation of the process $R_t$ are quite tractable, and with use of modern software do not present essential difficulties.
3.2 The case of the simple random walk

Let \( c = 1 \), and the size of the claim at each period is the r.v.

\[
X = \begin{cases} 
0 & \text{with probability } p, \\
2 & \text{with probability } q,
\end{cases}
\]

where \( q < p \). In this case, \( m = E\{X\} < 1 \), and hence \( c > m \).

Thus, for each period, the profit of the company is \( c - X = \pm 1 \) with probabilities \( p \) and \( q \), respectively.

Consider the threshold strategy with a level \( z \). Assume that \( u \) and \( z \) are integers, and let \( w_n(u, z) \) be the probability that the first dividend will be paid at the moment \( n \). By the definition of the strategy we use,

\[
w_0(u, z) = 0 \text{ for } u \leq z; \quad w_0(u, z) = 1 \text{ for } u > z;
\]

\[
w_n(u, z) = 0 \text{ for } u < 0, \text{ since in this case the insurer is ruined in the very beginning};
\]

\[
w_n(u, z) = 0 \text{ for } u > z \text{ and } n > 0, \text{ since in this case the first payment occurred}
\]

at the initial time. (3.2.1)

We apply the first step approach in a way similar to what we did in Section ???. With probability \( p \) the process moves up, the surplus becomes \( u + 1 \), the random walk starts over, and the probability that a dividend will be paid at time \( n \) “becomes” \( w_{n-1}(u + 1, z) \). The same concerns the case when the process in the first step moves down. Thus,

\[
w_n(u, z) = pw_{n-1}(u + 1, z) + qw_{n-1}(u - 1, z). \quad (3.2.2)
\]

Let

\[
\hat{w}(u, z) = \sum_{n=0}^{\infty} v^n w_n(u, z), \quad (3.2.3)
\]

the generating function of the sequence of probabilities \( \{w_n\} \). (See also Section ???.) We chose the same letter \( v \in (0, 1) \) as for discount on purpose.

Applying (3.2.2) for \( u \leq z \), and taking into account conditions (3.2.1), we have

\[
\hat{w}(u, z) = \sum_{n=1}^{\infty} v^n (pw_{n-1}(u, z) + qw_{n-1}(u - 1, z)) = p\nu \hat{w}(u + 1, z) + q\nu \hat{w}(u - 1, z).
\]

So, for the generating function we have the equation

\[
\hat{w}(u, z) = p\nu \hat{w}(u + 1, z) + q\nu \hat{w}(u - 1, z) \quad (3.2.4)
\]

for \( u \leq z \). Similarly one can get that

\[
\hat{w}(u, z) = 0 \text{ for } u < 0, \quad \hat{w}(u, z) = 1 \text{ for } u > z. \quad (3.2.5)
\]

Setting \( \hat{w}(u, z) = r^{u+1} \), where \( r \) is a number, and inserting it into (3.2.4), we see that such a function satisfies (3.2.4) if

\[
r = p\nu r^2 + q\nu. \quad (3.2.6)
\]
Thus, if \( r_1 \) and \( r_2 \) are the roots of the quadratic equation (3.2.6), the functions \( r_1^{u+1} \) and \( r_2^{u+1} \) are solutions to (3.2.4). Without going too deeply into the theory, note that then any solution \( \hat{w}(u, z) = c_1 r_1^{u+1} + c_2 r_2^{u+1} \), where \( c_1, c_2 \) are constants. To find constants \( c_1, c_2 \), we use (3.2.5), writing \( \hat{w}(z+1, z) = 1, \hat{w}(-1, z) = 0 \). Eventually it leads to the solution

\[
\hat{w}(u, z) = \frac{r_1^{u+1} - r_2^{u+1}}{r_1^{z+2} - r_2^{z+2}}.
\] (3.2.7)

Next, we consider a connection between \( V(u, z) \) and \( \hat{w}(u, z) \). Since \( u \) and \( z \) are integers, each dividend paid is equal to one. A dividend is paid if the surplus equals \( z + 1 \), and once a dividend is paid, the surplus becomes equal to \( z \). The probability that paying dividends starts from a moment \( n \) is \( w_n \), and after that “everything starts over” from the level \( z \). Hence, for \( u \leq z \),

\[
V(u, z) = \sum_{n=0}^{\infty} w_n(u, z) [v^n \cdot 1 + v^n V(z, z)] = [1 + V(z, z)] \hat{w}(u, z).
\] (3.2.8)

Setting \( u = z \), we have \( V(z, z) = [1 + V(z, z)] \hat{w}(z, z) \), from which it follows that

\[
V(z, z) = \frac{\hat{w}(z, z)}{1 - \hat{w}(z, z)}.
\] (3.2.9)

Combining (3.2.8) and (3.2.9), we have

\[
V(u, z) = \frac{\hat{w}(u, z)}{1 - \hat{w}(z, z)}.
\]

Substituting (3.2.7), we get eventually that for \( u \leq z \),

\[
V(u, z) = \frac{r_1^{u+1} - r_2^{u+1}}{(r_1^{z+2} - r_2^{z+2}) - (r_1^{z+1} - r_2^{z+1})}.
\] (3.2.10)

The denominator depends only on \( z \), while the numerator – only on \( u \). So, as was expected, the optimal level \( z^* \) does not depend on \( u \).

We skip detailed calculations leading to an optimal \( z \). To find it, one should take the derivative of the denominator in (3.2.10), set it equal to zero, and divide the whole equation by \( r_2^z \). Then the unknown \( z \) will be contained only in the expression \((r_1/r_2)^z\). Solving the equation with respect to this expression, one can readily get that the optimal level

\[
z^* = \frac{1}{\ln(r_1) + |\ln(r_2)|} \ln \left( \frac{\ln(r_2)}{\ln(r_1)} \cdot \frac{r_2(1 - r_2)}{r_1(r_1 - 1)} \right)
\] (3.2.11)

where \( r_1 > 1 \) is the larger and \( r_2 < 1 \) is the smaller root of equation (3.2.6). (The values in (3.2.11) may be negative; in this case one should set \( z^* = 0 \).)

Table 1 shows the values of \( z^* \) for different values of \( p, q = 1 - p, \) and \( v \). These calculations, as well as simulation of the process with \( X \)’s having different distributions, and a proof of (3.1.4) were provided by Sarah Borg in her master’s thesis [?].
7. GLOBAL CHARACTERISTICS OF THE SURPLUS PROCESS

It can be seen that for each value of \( p \), the value of \( z^* \) increases as \( \nu \) increases. The higher \( \nu \) is, the more the company is concerned about the future payments. So, the company increases the level of the surplus in order to increase the time before ruin.

We see that \( z^* \) initially increases in \( p \), and then decreases as \( p \) gets closer to 1. It is also understandable. Consider the extreme case \( p = 1 \). Then with probability one there will be no claim, and therefore the whole surplus could be paid out as dividends. So, \( z^* \) in this case should be zero. Then, if \( p \) is close to one, we should expect \( z^* \) to be small.

<table>
<thead>
<tr>
<th>( p )</th>
<th>0.60</th>
<th>0.65</th>
<th>0.70</th>
<th>0.75</th>
<th>0.80</th>
<th>0.85</th>
<th>0.90</th>
<th>0.95</th>
<th>0.98</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \nu )</td>
<td>0.90</td>
<td>0.92</td>
<td>0.94</td>
<td>0.96</td>
<td>0.98</td>
<td>1.11</td>
<td>0.95</td>
<td>0.68</td>
<td>0.007</td>
</tr>
<tr>
<td></td>
<td>0.15</td>
<td>0.54</td>
<td>1.14</td>
<td>2.21</td>
<td>4.65</td>
<td>1.02</td>
<td>1.46</td>
<td>2.07</td>
<td>3.01</td>
</tr>
<tr>
<td></td>
<td>0.69</td>
<td>1.14</td>
<td>1.79</td>
<td>2.86</td>
<td>4.99</td>
<td>1.17</td>
<td>1.57</td>
<td>2.10</td>
<td>2.89</td>
</tr>
<tr>
<td></td>
<td>1.02</td>
<td>1.46</td>
<td>2.07</td>
<td>3.01</td>
<td>4.73</td>
<td>1.19</td>
<td>1.54</td>
<td>1.99</td>
<td>2.63</td>
</tr>
<tr>
<td></td>
<td>1.11</td>
<td>1.40</td>
<td>1.78</td>
<td>2.30</td>
<td>3.17</td>
<td>0.95</td>
<td>1.18</td>
<td>1.48</td>
<td>1.89</td>
</tr>
<tr>
<td></td>
<td>0.68</td>
<td>0.85</td>
<td>1.07</td>
<td>1.38</td>
<td>1.88</td>
<td>0.68</td>
<td>0.85</td>
<td>1.07</td>
<td>1.38</td>
</tr>
<tr>
<td></td>
<td>0.07</td>
<td>0.09</td>
<td>0.19</td>
<td>0.034</td>
<td>0.58</td>
<td>0.007</td>
<td>0.09</td>
<td>0.19</td>
<td>0.034</td>
</tr>
</tbody>
</table>

Next, we briefly consider the expected life \( D(u,z) \) for the random walk model. Assume, as before, that \( u, z \) are integers, and \( p > q \).

The same first step approach leads to the equation

\[
D(u,z) = 1 + pD(u+1,z) + qD(u-1,z), \quad 1 \leq u \leq z.
\]

As can be verified by direct substitution, the solution to this equation is

\[
D(u,z) = \frac{p}{(p-q)^2} \left[ \left( \frac{p}{q} \right)^{z+1} - \left( \frac{p}{q} \right)^{z-u} \right] - \frac{u+1}{p-q}.
\]

A general expression for \( D(u,z) \) in terms of some series and other examples may be found in [?].

3.3 Finding an optimal strategy

In this section, we prove that the optimal strategy has properties described in Section 3.1. Assume that the optimal strategy exists and set

\[
V(u) = \max E \left\{ \sum_{t=1}^{\infty} \delta^t d_t \right\}, \tag{3.3.1}
\]

where max is over all possible strategies \( \{d_1, d_2, \ldots\} \) of paying dividends, and \( u \) is an initial surplus. So, \( V(u) \) is the expected discounted amount of dividends under the optimal strategy. It is convenient to consider the function \( V(u) \) for all \( u \), setting \( V(u) = 0 \) for all \( u < 0 \). We assume also that \( V(u) \) is continuous at all points \( u \) except perhaps \( u = 0 \).
Consider a time moment $t$. If the company is still functioning, it has a surplus $R = R_t \geq 0$, and should specify its policy for the next period. The process we consider is a Markov process, which means, in particular, that the policy may depend on the current surplus but does not depend on what strategy the company chose before time $t$. The company again faces the infinite time period, and should solve the optimization problem as if it is at the very beginning.

The company receives the next premium $c$, and pays out the claim $X = X_{t+1}$ and a dividend $d$ (which perhaps equals zero). So, the surplus $R_{t+1} = R_t + c - d - X$. If, at the next time, the company applies the optimal strategy (which we do not know yet but assume that it exists), then given $X$, the expected discounted amount of dividends after time $t+1$ will be $V(R_t + c - d - X)$.

From the standpoint of the present time $t$, the total amount of dividends is $d + vE\{V(R + c - d - X)\}$, where $R = R_t$ and $d$ cannot exceed the current surplus $R$. To find the optimal behavior at the period $[t, t+1]$, we should maximize the last expression in $d$, which leads to the equation

$$V(R) = \max_{0 \leq d \leq R} \{d + vE\{V(R + c - d - X)\}\}. \tag{3.3.2}$$

The reader familiar with the optimization theory recognizes in the above reasoning the so called optimality principle, and realizes that we have derived the Bellman equation.

Let the function

$$w(y) = vE\{V(y - X)\} - y.$$ 

Then

$$V(R) = \max_{0 \leq d \leq R} \{R + c + w(R + c - d)\} = R + c + \max_{0 \leq d \leq R} w(R + c - d) = R + c + \max_{c \leq y \leq c + R} w(y), \tag{3.3.3}$$

where we changed variables, setting $y = R + c - d$.

Assume now that the function $w(y)$ has a unique maximum at a point $y_0$. This is an implicit condition we impose to find the optimal solution. Consider three cases.

(i) $y_0 \leq c$. In this case, $\max_{c \leq y \leq c + R} w(y)$ is attained at the point $y = c$ (graph $w(y)$ with a unique maximum at $y_0$, and place $c$ on the right of $y_0$). Then the optimal $d = R + c - y = R$.

(ii) $c < y_0 \leq c + R$. Then $\max_{c \leq y \leq c + R} w(y)$ is attained at the point $y = y_0$, and the optimal payment $d = R + c - y_0 = R - \tilde{z}$, where $\tilde{z} = y_0 - c > 0$.

(iii) $y_0 > c + R$. Then $\max_{c \leq y \leq c + R} w(y)$ is attained at the point $y = c + R$, and the optimal payment $d = 0$.

Setting $z^* = \max(0, y_0 - c) \geq 0$, we see that in all three cases above, the optimal payment

$$d = \begin{cases} R - z^* & \text{if } R > z^*, \\ 0 & \text{if } R \leq z^*. \end{cases} \tag{3.3.4}$$

So, the fact that the optimal strategy has the threshold structure is proved.
7. GLOBAL CHARACTERISTICS OF THE SURPLUS PROCESS

To prove (3.1.4), we write (3.3.3) as

\[ V(u) - u = c + \max_{c \leq y \leq c+u} w(y). \] (3.3.5)

If \( y_0 \leq c \), then \( V(u) - u = c + w(c) \) and, hence, does not depend on \( u \). It is natural – in this case the optimal dividend payment at the initial moment would be \( d = u \), and the company will start from zero level.

Let \( y_0 > c \). If \( u > z^* = y_0 - c \), then as was proved, the optimal behavior consists in immediate payment of the surplus \( u - z^* \) as a dividend. After that the process starts from the level \( z^* \).

If \( u < z^* \), there should be no dividend at the initial moment, and \( V(u) - u = c + w(c+u) \).
In this case, \( c + u < y_0 \), and hence \( w(c+u) \) is increasing in \( u \) up to the moment when \( c + u = y_0 \). This is equivalent to \( u = z^* \).

4 EXERCISES

Sections 1 and 2

1. Is \( \psi_T(u) \), as a function of \( T \), increasing?

2. Consider the process \( R_t \) in continuous time, and set \( \tilde{\phi}_n(u) = P(R_t \geq 0 \text{ for all } t = 1,2,\ldots,n) \), that is, we count only integer moments of time. Which is larger: \( \tilde{\phi}_n(u) \) or \( \phi_n(u) \)?

3. Show that in general, not assuming that \( R_t = u + c_t - S(t) \), for (2.1.5) to be true it suffices to require

\[ E\{R_t\} \to \infty \quad \text{and} \quad Var\{R_t\} = o\left([E\{R_t\}]^2\right). \]

4. Look up Exercises ??-?? from Chapter ??.

5. Problems below concern Example 2.2.2-1.

(a) Graph the r.-h. and l.-h. sides of (2.2.10). Show that for (2.2.10) to have a positive solution \( \gamma < 1 \), the premium \( c \) should be indeed greater than 2. Check the numerical answers in Example 2.2.2-1.

(b) Write a program (it suffices to provide a spreadsheet) which would allow to compute \( \gamma \) for different \( c \)'s. Compare the results with what the approximation (2.2.14) gives.

(c) Show that \( \gamma \to 1 \) as \( c \to \infty \). Explain why in this case, (2.1.3) is not a good estimate for large \( c \). (Advice: Show that the ruin probability should vanish when \( c \to \infty \).)
6. Assume that for some \( c > m \) there exists a positive solution \( \gamma \) to equation (2.2.9).

(a) Proceeding from the results of Section 2.2.1 and using Figures 4ab, show at a heuristic level that \( \gamma \to 0 \) as \( c \to m \) from the right (being greater than \( m \)), and in this case, the ruin probability converges to one. Explain why it is not surprising from an economic point of view.

(b) Prove that \( \gamma \to 0 \) as \( c \to m \) rigorously.

(c) Consider the case \( \gamma \to \infty \). Explain at a heuristic level that in this case, the ruin probability should converge to zero. Show that it follows from (2.1.3). (Hint: If \( c \) is very large, with probability close to one, the surplus at the next step will be very large.)

7. Assume that \( M_X(z) \) in (2.2.9) is defined for all \( z \).

(a) Assume also that a positive solution \( \gamma \) to equation (2.2.9) exists for all \( c > m \). Proceeding from the results of Section 2.2.1 and using Figures 3-5, show at a heuristic level that \( \gamma \to \infty \) as \( c \to \infty \), and the ruin probability converges to zero. Explain why it is not surprising from an economic point of view.

(b) Prove the result of Exercise 7a rigorously.

(c) Let the r.v. \( X \) above be bounded by a number \( b \). Show that in this case a positive solution to (2.2.9) does not exist for \( c > b \), and that \( \gamma \to \infty \) as \( c \to b \) if \( X \) is not degenerate.

8. If you should solve (2.2.22) for two different \( a \)’s and the same \( \nu \), would you solve (2.2.22) twice or just one time?

9. Using (2.2.14), estimate the ruin probability in the situation (2.2.18) for \( u = 100, \theta = 0.1, \xi’s \) having the standard exponential distribution, and \( K_i’s \) having the geometric distribution with parameter \( p = 0.1 \).

10. Estimate the adjustment coefficient in the situation of Example 2.2.3-1 for \( \nu = 3, \theta = 0.1, \) and \( a = 1 \) and 2. (Advice: First solve Exercise 8.)

11. Find the adjustment coefficient in the situation of Example 2.2.3-1 for \( \nu = 2 \). Show that for small \( \theta \) the answer does not contradict approximation (2.2.25). (Advice: First solve Exercise 8, and think for what \( a \) solutions to (2.2.22) are simpler.)

12. Making use of the result of Exercise ???, show that the r.-h.s. of (2.3.3) is non-decreasing in \( \gamma \).

13. A flow of claims is represented by a compound Poisson process \( S_N \) in continuous time. The mean time between adjacent claims is half an hour. The random value \( X \) of a particular claim is uniformly distributed on \( [0,10] \) (say, the unit of money is $1000). The initial surplus (capital) is 100, the relative loading coefficient \( \theta = 0.2 \).

(a) Estimate the ruin probability.

(b) For which the initial surplus is the ruin probability less than 0.05?
7. GLOBAL CHARACTERISTICS OF THE SURPLUS PROCESS

(c) Let \( u = 50 \). Find \( \theta \) for which the ruin probability is less than 0.05.

14. For a particular group of clients, a flow of claims arriving at an insurance company may be represented as a compound Poisson process. Let the amount of a particular claim be equal to either 2, 3, or 4, with probabilities 1/4, 1/2, and 1/4, respectively. Let the mean number of claims the company receives per day be 10. Assume that the company chooses for its activity a relative loading coefficient \( \theta = 0.1 \).

(a) Write an equation for the adjustment coefficient \( \gamma \).
(b) Does this equation involve \( \lambda \)?
(c) Find an approximate solution using software. Compare it with approximation (2.2.24).
(d) Find an approximate value of the initial capital for which the ruin probability for the company will be less than 0.03.
(e) Think how to answer all questions above, if a particular claim is, say, uniformly distributed on [2, 4]. Do you expect the ruin probability to be smaller?

15. Think how to answer all questions in Exercise 14 if the number of claims during a day is exactly equal to 10, that is, we consider the discrete time scheme, and day is a unit of time.

16. In the case of the compound Poisson process, for some data the ruin probability \( \psi(u) = 0.3e^{-2u} + 0.4e^{-u/2} \). Find \( \theta \) and \( \gamma \).

17. In the discrete time case, for the claim \( X \) in one period, we have \( E\{X\} = 3, \) \( Var\{X\} = 2.5 \). The required level \( \beta = 0.05 \). Estimate proper combinations of the initial surplus and the loading coefficient for “large” \( u \).

18. Provide a graph illustrating a solution to (2.4.2).

19. (a) Without calculating anything, show that if \( X \) is a discrete r.v., the density (2.6.4) is a mixture of uniform distributions.
(b) Find the density (2.6.4) for the case when \( X \) takes on only one value.
(c) Find the density (2.6.4) for the case when \( X \) takes on values 1,2,3 with probabilities 1/5,2/5,2/5, respectively.

20. Show that, if \( X \) is exponential with a parameter \( a \), then the density (2.6.4) is exponential with the same parameter. Explain that this fact is non-surprising in light of what was discussed in Section 2.4.1.

21. Clarify from a heuristic point of view the significance of the fact that the density (2.6.4) is decreasing.

22. In the framework of Section 2.6, let the r.v.’s \( X_i \) be continuous. Does it mean that the r.v. \( Z_K \) is continuous? (Advice: Think about \( \phi(0) \) and how it is connected with \( Z_K \).)
23. (a) Making use of (2.2.24), show that the constant $C$ in (2.6.15) is close to one for small $\theta$.
   
   (b) Making use of (2.6.14), estimate the ruin probability for large $u$ in the case when $X$ has the $\Gamma$-distribution with parameters $a = 1$, $\nu = 2$, and $\theta = 0.1$.

24. Realize why the values of $\psi_2(u)$ in Examples 2.7-2 and 3 are the same, while ranges for $u$ are different.

25. Assume that in the situations of Example 2.7-2, you are asked only to find $u$ for which $\phi(u) \geq 0.99$. Realize that in this case you can avoid most of the calculations, coming to the answers very quickly. Find this answer.

26. For $X = 2, 10$ with probabilities 0.6, 0.4, respectively, solve the problem of Example 2.7-2 and the problem of Example 2.7-3 for $r = 0.2$.

27. Show that for $T < \infty$, the counterpart of (2.7.6) will be
   
   $$\phi_T(u) = \lambda \int_0^T e^{-\lambda t} \left( \int_0^{u+ct} \phi_{T-t}(u+ct-z) dF(z) \right) dt + e^{-\lambda T}.$$
   
   *(Hint: The conditional non-ruin probability given $\eta \geq T$, is certainly one (no claim arrived within the period $[0, T]$).)*

Section 3

28. Regarding the model of Section 3.1, give a common sense explanation to why one should expect a higher optimal level $z^*$ for higher values of discount $v$.

29. Can the optimal level $z^*$ be zero? *(Advice: Think about the case of small $v$).*

30. In the model of Section 3.2, let $p = 0.7$, $v = 0.9$. Using Excel or other software, provide the graph of $V(0, z)$. Interpret it. Why are too small and too large values of $z$ not optimal? Estimate the optimal $z^*$. Explain why for any $u < z^*$, the maximizer of $V(u, z)$ will be the same.

31. Using Excel or other software, graph $z^*$ against $p$ for $v = 0.9$. 