


TEXTS AND READINGS IN MATHEMATICS 54

# Lectures on Insurance Models

S. Ramasubramanian

HINDUSTAN BOOK AGENCY

TEXTS AND READINGS
IN MATHEMATICS **54**

# Lectures on
# Insurance Models

**Texts and Readings in Mathematics**

**Advisory Editor**
C. S. Seshadri, Chennai Mathematical Institute, Chennai.

**Managing Editor**
Rajendra Bhatia, Indian Statistical Institute, New Delhi.

**Editors**
R. B. Bapat, Indian Statistical Institute, New Delhi.
V. S. Borkar, Tata Inst. of Fundamental Research, Mumbai.
Probal Chaudhuri, Indian Statistical Institute, Kolkata.
V. S. Sunder, Inst. of Mathematical Sciences, Chennai.
M. Vanninathan, TIFR Centre, Bangalore.

# Lectures on
# Insurance Models

**S. Ramasubramanian**
Indian Statistical Institute
Bangalore

HINDUSTAN
BOOK AGENCY

Published in India by

Hindustan Book Agency (India)
P 19 Green Park Extension
New Delhi 110 016
India

email: info@hindbook.com
http://www.hindbook.com

Copyright © 2009, Hindustan Book Agency (India)

No part of the material protected by this copyright notice may be reproduced or utilized in any form or by any means, electronic or mechanical, including photocopying, recording or by any information storage and retrieval system, without written permission from the copyright owner, who has also the sole right to grant licences for translation into other languages and publication thereof.

All export rights for this edition vest exclusively with Hindustan Book Agency (India). Unauthorized export is a violation of Copyright Law and is subject to legal action.

ISBN 978-81-85931-93-7 ISBN 978-93-86279-44-6 (eBook)
DOI 10.1007/978-93-86279-44-6

# Contents

**Preface** **vii**

**1 Introduction** **1**
1.1 Main assumptions . . . . . 2
1.2 An overview . . . . . 3

**2 Poisson model** **6**
2.1 Preliminaries . . . . . 6
2.2 Interarrival and waiting time distributions . . . . . 10
2.3 Order statistics property . . . . . 17
2.4 Cramér-Lundberg Model . . . . . 20
2.5 Exercises . . . . . 23

**3 Renewal model** **29**
3.1 Preliminaries . . . . . 29
3.2 Renewal function . . . . . 36
3.3 Renewal theorems . . . . . 43
3.4 Sparre Andersen model . . . . . 54
3.4.1 Some basic premium principles . . . . . 57
3.5 Exercises . . . . . 61

**4 Claim size distributions** **69**
4.1 Light-tailed distributions . . . . . 69
4.2 Regularly varying functions and distributions . . . . . 72
4.3 Subexponential distributions . . . . . 84
4.4 Exercises . . . . . 101

**5 Ruin Problems** **107**
5.1 Risk process and an associated random walk . . . . . 107
5.2 Lundberg bound . . . . . 112

5.3 An integral equation in Cramér-Lundberg model . . . . . 115
5.4 Exact asymptotics in Cramér-Lundberg model . . . . . 119
5.4.1 Light-tailed claims . . . . . 119
5.4.2 Subexponential claims . . . . . 123
5.5 Exercises . . . . . 125

**6 Lundberg risk process with investment** **130**
6.1 Lundberg risk process and Markov property . . . . . 131
6.2 Risk process with investment . . . . . 141
6.3 Assorted comments . . . . . 166
6.4 Exercises . . . . . 168

**Appendix** **173**
A.1 Basic notions . . . . . 173
A.2 On the central limit problem . . . . . 180
A.3 Martingales . . . . . 185
A.4 Brownian motion and Ito integrals . . . . . 189

**Bibliography** **196**

**Index** **200**

# Preface

This introductory book is an expanded version of a series of lectures given during the academic year 2006-2007 at the Department of Mathematics, Indian Institute of Science (I.I.Sc.), Bangalore, organized by *IISc Mathematics Initiative*. The audience consisted of students/ faculty with background in mathematics, statistics or engineering.

Our objective is to introduce the basic mathematical concepts of insurance models. To maintain a reasonable length, we consider only models from collective risk theory (also called non-life insurance). The language of continuous time stochastic processes is the right medium to mathematically formulate these models. So Poisson process, more general renewal processes and later stochastic differential equations will be encountered in the sequel. As the so called heavy-tailed distributions have become increasingly important in insurance, tail behaviour of probability distributions on the positive real line is also studied at some length. Nothing is new in the book, except perhaps parts of Chapter 6 and some details of presentation.

Good background in basic probability theory and real analysis should suffice to follow most of the book. Some familiarity with measure theoretic probability will be more helpful. It is hoped that those preparing to take up a career in actuarial science would find this book a useful introductory text. Practitioners in the field who are interested in the underlying mathematical aspects of insurance models may find it informative. In addition, this can serve as a collateral text for courses on stochastic processes to graduate students in mathematics / statistics / operations research / engineering.

It is a pleasure to thank Mrinal Ghosh for inviting me to lecture at his department, and for encouraging me to write up the notes. Thanks are due to Vivek Borkar for his interest in the project, to B.V. Rao and two painstaking referees for their constructive suggestions on an earlier draft, and to E. Asha Lata for her efficient typing. Also I thank my wife Chandra for her constant support.

This book is dedicated to my parents.

S. Ramasubramanian

# Chapter 1

# Introduction

Catastrophes, natural as well as man-made, reinforce the fact that almost everyone would like to be assured of some (non-supernatural) agency to bank upon in times of grave need. If the affected parties are too poor, then it is the responsibility of governments and the "haves" to come to the rescue. It is not uncommon that places of worship, palatial buildings and schools serve as refuges for the affected. However, there are also significant sections of the population who are willing to pay regular premium to suitable agencies during normal times to have insurance cover to tide over crises.

Insurance policies have been prevalent in Europe since middle ages. Scientists like J. Graunt, E. Halley, A. De Moivre and C. Neumann are believed to have played major roles in formulating some of the procedures in connection with life insurance in 17th century. In India, as advocated by educationist and social reformer I. Vidyasagar, insurance schemes aimed at mitigating the plight of widows were introduced in late 19th century.

By 18th century the extreme uncertainties encountered in maritime trade gave rise to insurance covers against property loss; these included also fairly sophisticated schemes like reinsurance policies, that is, insuring of insurance companies. Complexities of 20th and 21st centuries have, of course, spawned a whole gamut of insurance schemes/products (like life, health/medical, accident, disability, fire, theft, car, pension funds, travel, housing, business related insurance, etc.). Thus insurance has become a part of modern life, whether we like it or not.

In an insurance set up, a considerable proportion of the financial risk is shifted to the insurance company. The implicit trust between the insured and the insurance company (perhaps formalized by a legal

contract) is at the core of the interaction. A reasonable mathematical theory of insurance can possibly provide a scientific basis for this trust. There were sporadic attempts at this, notably by T. Thiele who derived a differential equation in the context of life insurance in 19th century. The first systematic step in this direction seems to have been taken by the actuary F. Lundberg in 1903. It was further strengthened in the 1930's by the pioneering work of the actuary F. Esscher and the probabilist H. Cramér. Not surprisingly, these developments took place around the same time when probability theory acquired its identity. Mathematics of insurance is now considered a part of applied probability, and a major portion of it is described in terms of continuous time stochastic processes. For more details concerning history, and also for quick overview of many theoretical concepts and practical aspects of insurance, an excellent source is the encyclopedic work [TS].

To keep this introductory text to a reasonable length, we shall consider only some simple models of *risk theory* (also called non life insurance or collective risk theory). A typical model will involve an insurance company underwriting a large number of policies for some type of risk, like insurance against fire or theft in a particular but fairly large locality, or insurance against floods, or insurance against crop damage in a region. In such a homogeneous portfolio, the insured goods may have comparable value, and chances of mishap may not be very different from one insured to another. Such portfolios are used at least as first approximation to real situations.

## 1.1 Main assumptions

One of the objectives is modelling of claims that arrive in an insurance business. In this context the following assumptions are generally made.

1. The number of claims occurring in a given period is random. Claims happen at times $\{T_i\}$ satisfying $0 \leq T_1 \leq T_2 \leq \ldots$. We call them *claim arrival times* (or just *arrival times*).

2. The $i^{\text{th}}$ claim arriving at time $T_i$ causes a payment $X_i$. The sequence $\{X_i\}$ is assumed to be an i.i.d. (independent and identically distributed) sequence of nonnegative random variables. These are called *claim sizes*.

3. The claim sizes $\{X_i\}$ and the claim arrival times $\{T_j\}$ are assumed to be independent.

The second assumption indicates a certain homogeneity of policies which are nevertheless independent. The first two assumptions are fairly natural, whereas the third one is more of a mathematical convenience.

We also mention two stochastic processes central to our discussions. Take $T_0 = 0$. Define the *claim number process* by

$$\begin{aligned} N(t) &= \max\{i \geq 0 : T_i \leq t\}, \\ &= \text{number of claims occurring by time } t, \quad t \geq 0 \end{aligned} \tag{1.1}$$

Also define the *total claim amount process* by

$$S(t) = \begin{cases} 0, & \text{if } N(t) = 0 \\ X_1 + \ldots + X_k, & \text{if } N(t) = k, \quad k = 1, 2, \ldots \end{cases} \tag{1.2}$$

for $t \geq 0$. Clearly $S(t)$ = cumulative claim over $[0, t]$. We shall write $S(t) = \sum_{i=1}^{N(t)} X_i$, $t \geq 0$ which shall stand for (1.2). An implicit assumption here is that payment is immediate on receipt of the claim. Note that a sample path of $N$ and the corresponding sample path of $S$ have jumps at the same times $\{T_i\}$, by 1 for $N$ and by $X_i$ for $S$.

**Notation:** A function $f(\cdot)$ is said to be $o(h)$ if $\lim_{h \to 0} \frac{f(h)}{h} = 0$; we denote this by $f(\cdot) = o(h)$. That is, $f(\cdot)$ decays at a faster rate than $h$.

## 1.2 An overview

We now give an overview of the book, indicating along the way relation to the main objectives of risk theory. References to more comprehensive works are also given.

A primary task is to model times at which claims arrive, and to model payments involved when claims are made. In the notation of Section 1.1, this means modelling $\{T_i\}$ (or equivalently $N(\cdot)$) and $\{X_j\}$; consequently the total claim amount process $S(\cdot)$ can be modelled. A reasonable model should be realistic as well as amenable to analysis.

In Chapter 2 we discuss the versatile Poisson process and introduce the classical Cramér-Lundberg model, along with their basic properties. As indicated earlier, this is the first mathematical model in the subject and even after 100 years continues to hold a central place.

The renewal model, which is a genealization of the Poisson model, is considered in Chapter 3. An important role is played by the renewal function $t \mapsto E(N(t)+1)$, where $N(\cdot)$ is the counting process associated

with the renewal model. Renewal theorems, including Blackwell's renewal theorem, are proved. This gives the asymptotic behaviour of the solution to the renewal equation, which in turn, is used in the study of ruin problems later. Sparre Andersen model, the analogue of Cramér-Lundberg model in the renewal set up, is introduced in Section 3.4. Thus Chapters 2 and 3 give models for claim number process $N(\cdot)$. Both these models are useful in various disciplines besides insurance; see [KT], [Ro1], [Ro2] in addition to [EKM], [M], [RSST].

As the insurance company is open to enormous risk, it is important to indicate how premiums are to be charged so that the company does not go bankrupt. A safety loading factor is more or less imperative. Some basic premium principles are indicated at the end of Chapter 3.

Over the years actuaries have observed that claim sizes can sometimes be so huge that any finite variance assumption would be unrealistic. Such distributions/distribution functions were already familiar to analysts and probabilists, thanks to the work of Karamata on Tauberian theorems, and that of Feller on the central limit problem; see Section 4.2 and Appendix 2. Risk theory has been among the first to recognize that these heavy tailed distributions are not mere mathematical curiosities. A larger class, called subexponential distributions, is more or less synonymous with heavy tailed (claim size) distributions in insurance modelling. An account is given in Section 4.3. For more information see [BGT], [Fe], [EKM] and [RSST].

Ruin problems form an essential part of risk theory. Probability of ruin in finite time (or ruin probability, for short) is generally considered an objective indicator of the health of an insurance company. Hence obtaining estimates for ruin probability, asymptotic behaviour, etc. are considered very important. In Sections 5.1, 5.2 we look at the ruin problem for the renewal risk model (that is, Sparre Andersen model). Sections 5.3, 5.4 specialise to the Cramér-Lundberg model, with possibly heavy tailed claim size distribution. For additional information, besides [EKM], [RSST], one may look up the monograph [As] which is devoted to ruin problems.

In addition to income by way of premiums, an insurance company may have income from investments. Chapter 6 deals with the situation when, under the Cramér-Lundberg framework, the insurance company invests its surplus in a riskless bond and in a risky asset. The Markovian nature of the set up is used extensively. Surprisingly such analysis has been attempted only in the last decade. See [H], [HP1], [HP2] and

the references therein. Our approach is somewhat different. For the convenience of the reader, prerequisites are reviewed in an appendix.

# Chapter 2

# Poisson model

## 2.1 Preliminaries

We first consider the claim number process $\{N(t) : t \geq 0\}$. We assume that there is a basic probability space $(\Omega, \mathcal{F}, P)$ on which $N(t, \cdot)$ is a random variable for each $t \geq 0$. We list below some of the obvious/desired properties of $N$ (rather postulates for $N$), which may be taken into account in formulating a model for the claim number process.

• (N1): $N(0) \equiv 0$. For each $t \geq 0; N(t)$ is a nonnegative integer valued random variable.

• (N2): If $0 \leq s < t$ then $N(s) \leq N(t)$. Note that $N(t) - N(s)$ denotes the number of claims in the time interval $(s, t]$. So $N$ is a nondecreasing process.

• (N3): The process $\{N(t) : t \geq 0\}$ has *independent increments*; that is, if $0 < t_1 < t_2 < \ldots < t_n < \infty$ then $N(t_1), N(t_2) - N(t_1), \ldots, N(t_n) - N(t_{n-1})$ are independent random variables, for any $n = 2, 3, \ldots$. In other words, number of claims arriving in disjoint time intervals are independent.

• (N4): The process $\{N(t)\}$ has *stationary increments*; that is, if $0 \leq s < t, h > 0$ then the random variables $N(t) - N(s)$ and $N(t + h) - N(s + h)$ have the same distribution. This means that the probability law of the number of claim arrivals in any interval of time depends only on the length of the interval

• (N5): Probability of two or more claim arrivals in a very short span of time is negligible; that is,

$$P(N(h) \geq 2) = o(h), \quad \text{as} \quad h \downarrow 0. \tag{2.1}$$

- (N6): There exists $\lambda > 0$ such that

$$P(N(h) = 1) = \lambda h + o(h), \quad \text{as} \quad h \downarrow 0. \tag{2.2}$$

The number $\lambda$ is called the *claim arrival rate*. That is, in a very short time interval the probability of exactly one claim arrival is roughly proportional to the length of the interval.

The first two postulates are self evident. The hypothesis (N3) is quite intuitive; it is very reasonable at least as a first step in modelling. (N5), (N6) are indicative of the fact that between two arrivals there will be a gap, but may be very small; (note that bulk arrivals are not considered here). (N4) is a time homogeneity assumption; it is not very crucial.

To see what our postulates (N1) - (N6) lead to, put

$$P_n(t) = P(N(t)) = n), \ t \geq 0, \ n = 0, 1, 2, \ldots. \tag{2.3}$$

Observe that

$$\begin{aligned} P_0(t+h) &= P(N(t) = 0, \ N(t+h) - N(t) = 0) \text{ (by (N1),(N2))} \\ &= P(N(t) = 0) \ P(N(t+h) - N(t) = 0) \quad \text{(by (N3))} \\ &= P_0(t) \ P(N(h) = 0) \quad \text{(by (N4))} \\ &= P_0(t) \ [1 - \lambda h + o(h)], \quad \text{(by (N5), (N6))} \end{aligned}$$

whence it follows that $P_0(\cdot)$ is differentiable and

$$\frac{d}{dt} P_0(t) = -\lambda \ P_0(t), \quad t > 0. \tag{2.4}$$

By (N1), $P_0(0) = P(N(0) = 0) = 1$. So the differential equation (2.4) and this initial value give

$$\begin{aligned} P_0(t) &= P(N(t) = 0) = P(N(t+s) - N(s) = 0) \\ &= \exp(-\lambda t), \quad t \geq 0, \quad s \geq 0. \end{aligned} \tag{2.5}$$

Similarly for $n \geq 1$, using (N3) - (N6) we get

$$P_n(t+h) = P(N(t+h) = n) = I_1 + I_2 + I_3,$$

where

$$\begin{aligned} I_1 &= P(N(t) = n, \quad N(t+h) - N(t) = 0), \\ I_2 &= P(N(t) = n-1, \quad N(t+h) - N(t) = 1), \\ I_3 &= P(N(t+h) = n, \quad N(t+h) - N(t) \geq 2), \end{aligned}$$

and hence

$$P_n(t+h) = P_n(t)[1 - \lambda h + o(h)] + P_{n-1}(t)\ [\lambda h + o(h)] + o(h).$$

Right continuity and differentiability from the right of $P_n(\cdot)$ now follow. Similarly differentiability from the left can be proved; see Exercise 2.1.6 below. Consequently $P_n(\cdot)$ is differentiable and

$$\frac{d}{dt}\ P_n(t) = -\lambda P_n(t) + \lambda P_{n-1}(t), \quad t > 0 \tag{2.6}$$

for $n = 1, 2, \ldots$. To solve (2.6) inductively, set $Q_n(t) = e^{\lambda t} P_n(t)$, $t \geq 0$, $n = 0, 1, 2, \ldots$ By (2.6)

$$\frac{d}{dt}\ Q_n(t) = \lambda Q_{n-1}(t), \quad t > 0, \quad n = 1, 2, \ldots. \tag{2.7}$$

Clearly $Q_0(\cdot) \equiv 1$ by (2.5). So from (2.7) it follows that $Q_1(t) = \lambda t + c$ where $c$ is a constant. As $Q_1(0) = P_1(0) = P(N(0) = 1) = 0$, we have $c = 0$; hence $Q_1(t) = \lambda t$, $t \geq 0$. Using (2.7) for $Q_n(\cdot)$ and the initial value $Q_n(0) = 0$, $n \geq 2$ it is now simple to obtain inductively $Q(t) = \frac{1}{n!}\ \lambda^n t^n$. Hence we have proved

**Theorem 2.1** *Let the stochastic process $\{N(t) : t \geq 0\}$ satisfy the postulates (N1) - (N6). Then for any $t \geq 0$, $s \geq 0$, $k = 0, 1, 2, \ldots$*

$$\begin{aligned} P(N(t+s) - N(s) = k) &= P(N(t) = k) \\ &= \frac{(\lambda t)^k}{k!}\ e^{-\lambda t}. \end{aligned} \tag{2.8}$$

*Thus the random variable $N(t)$ has the Poisson distribution with parameter $\lambda t$. The stochastic process $\{N(t) : t \geq 0\}$ is called a time homogeneous Poisson process with arrival rate $\lambda > 0$.* □

The assumptions (N1) - (N6) are qualitative, whereas the conclusion above is quantitative. This is the reason why Poisson distribution and Poisson process come up in diverse fields like physics, chemistry, biology, engineering, economics, etc.; see [Fe], [KT], [Ro1], etc.

In view of the preceding theorem we have the following definition.

**Definition 2.2** Let $\lambda(\cdot)$ be a nonnegative measurable function on $[0, \infty)$ having finite integral over every bounded interval. A stochastic process $\{N(t) : t \geq 0\}$ is said to be a *Poisson process with intensity (or rate) function* $\lambda(\cdot)$ if the following hold.

(i) (N1), (N2), (N3) are satisfied.

(ii) For $P$ - a.e. $\omega \in \Omega$, the sample path $N(\cdot, \omega)$ is right continuous at every $t \geq 0$, and has left limit at every $t > 0$. That is, the process has r.c.l.l. sample paths.

(iii) For $0 \leq s < t$ the random variable $N(t) - N(s)$ has a Poisson distribution with parameter $\int_s^t \lambda(y)\, dy$.

If $\lambda(\cdot) \equiv \lambda > 0$, a constant, then $N$ is said to be a *homogeneous Poisson process with rate* $\lambda > 0$. If $\lambda(\cdot) \equiv 1$ then $N$ is called a *standard homogeneous Poisson process.*

The above definition is not the most economical one. For example, existence of left limits follows from the paths being nondecreasing. This may be a small price for the sake of clarity.

As $N(\cdot)$ is nondecreasing, $N(0) \equiv 0, N(t, \omega) < \infty$ for $P$ - a.e. $\omega$, note that the sample paths have left and right limits at every $t > 0$. So condition (ii) above means that we have chosen a right continuous version.

**Proposition 2.3** *Let $\lambda(\cdot)$ be a nonnegative measurable function on $[0, \infty)$. Assume that $\mu(t) \triangleq \int_0^t \lambda(s)\, ds < \infty$, for all $t \geq 0$. Let $N(\cdot)$ be a Poisson process with rate function $\lambda(\cdot)$ and $\tilde{N}(\cdot)$ be a standard homogeneous Poisson process. Then the following hold.*

*(i) $\{\tilde{N}(\mu(t)) : t \geq 0\}$ is a Poisson process with rate function $\lambda(\cdot)$.*

*(ii) If $\lambda(\cdot) > 0$ and $\lim_{t \to \infty} \mu(t) = \infty$, then $\{N(\mu^{-1}(t)) : t \geq 0\}$ is a standard homogeneous Poisson process.*

**Proof:** Left as an exercise. □

Using Kolmogorov's consistency theorem and the theory of weak convergence of probability measures on $D[0, \infty)$, it can be established that, for a given rate function $\lambda(\cdot)$, a Poisson process as in Definition 2.2 exists and that the probability law induced on $D[0, \infty)$ is unique. Here $D[0, \infty)$ denotes the space of all real valued functions on $[0, \infty)$ which are right continuous at every $t \geq 0$ and have left limits at every $t > 0$; $D[0, \infty)$ is endowed with the Skorokhod topology. See [Bl1], [P1],

[EK]. We shall assume that a Poisson process exists for a given rate function. A constructive version of the homogeneous Poisson process will, however, be indicated later; see Theorem 2.5.

For a Poisson process, let $\mu(t) = \int_0^t \lambda(s)\, ds$, $t \geq 0$. The function $\mu(\cdot)$ can be considered as an intrinsic clock for $N$. If $N$ is homogeneous the intrinsic clock evolves linearly; that is, claims arrive roughly uniformly over time. If $N$ has nonconstant rate then the clock "slows down" or "accelerates" according to $\lambda(\cdot)$; in insurance context, nonconstant rate may refer to seasonal trends.

The Poisson process owes its versatility to the fact that many natural functionals connected with the model can be explicitly determined. The next two sections discuss a few examples which are relevant in the actuarial context as well.

In Section 2.2 it is shown that the interarrival times (that is, time between successive claim arrivals) of a homogeneous Poisson process form a sequence of independent identically distributed random variables having an exponential distribution. In fact, the homogeneous Poisson process is characterized by this property. Section 2.3 discusses the joint distribution of the first $n$ claim arrival times, given that exactly $n$ arrivals have taken place during $[0, t]$. This has connection with the order statistics of $n$ independent identically distributed random variables having uniform distribution. Both these properties play major role in the Cramér-Lundberg risk model as will be seen in the sequel.

## 2.2 Interarrival and waiting time distributions

Let $\{N(t) : t \geq 0\}$ be a time homogeneous Poisson process with arrival rate $\lambda > 0$. Set $T_0 \equiv 0$. For $n = 1, 2, \ldots$ define

$$\begin{aligned} T_n &= \inf\{t \geq 0 : N(t) = n\} \\ &= \text{time of arrival of } n\text{-th claim} \qquad (2.9) \\ A_n &= T_n - T_{n-1} \qquad (2.10) \\ &= \text{time between } (n-1)^{\text{th}} \text{ and } n^{\text{th}} \text{ claim arrivals.} \end{aligned}$$

The random variables $T_0, T_1, T_2, \ldots$ are called *claim arrival times* (or *waiting times*); the sequence $\{A_n : n = 1, 2, \ldots\}$ is called the sequence of *interarrival times*.

For any $s > 0$ note that $\{T_1 > s\} = \{N(s) = 0\}$. Hence by (2.8)

$$P(A_1 > s) = P(T_1 > s) = P(N(s) = 0) = e^{-\lambda s}.$$

Therefore the random variable $A_1$ or $T_1$ has *exponential distribution* with parameter $\lambda$; that is

$$\begin{aligned} & P(A_1 \in (a,b)) = P(T_1 \in (a,b)) \\ = & \int_a^b \lambda\, e^{-\lambda s}\, ds,\ 0 \le a < b < \infty. \end{aligned} \tag{2.11}$$

With more effort one can prove the following

**Theorem 2.4** *Let $\{N(t) : t \ge 0\}$ be a time homogeneous Poisson process with rate $\lambda > 0$. Let $\{T_i\}$ denote the claim arrival times, and $\{A_j\}$ the interarrival times defined by (2.9), (2.10). Then the following hold.*

*(a) Fix $n \ge 1$. Then the joint probability density function of $T_1, T_2, \ldots, T_n$ is given by*

$$\begin{aligned} & f_{(T_1,T_2,\ldots,T_n)}(t_1,\ldots,t_n) \\ = & \begin{cases} \lambda^n e^{-\lambda t_n}, & \text{if } 0 < t_1 < t_2 < \ldots < t_n < \infty \\ 0, & \text{otherwise.} \end{cases} \end{aligned} \tag{2.12}$$

*(b) $\{A_n : n = 1, 2, \ldots\}$ is a sequence of independent, identically distributed (i.i.d.) random variables having Exp($\lambda$) distribution.*

**Proof:** (a) In view of the discussion preceding the theorem (see (2.11)), note that we need to consider only $n \ge 2$. Let

$$G_n = \{(t_1, t_2, \ldots, t_n) : 0 < t_1 < t_2 < \ldots < t_n < \infty\}. \tag{2.13}$$

By postulate (N5) of Theorem 2.1 and Exercise 2.1.4 observe that $0 < T_1 < T_2 < \ldots < T_n < \infty$ a.s., that is, $P((T_1, T_2, \ldots, T_n) \in G_n) = 1$.

If a "rectangle" $I \triangleq (a_1, b_1] \times (a_2, b_2] \times \ldots \times (a_n, b_n] \subset G_n$ then $0 \le a_1 < b_1 \le a_2 < b_2 \le \ldots < b_{n-1} \le a_n < b_n < \infty$. Now by the independent increment property of the Poisson process we get, taking $b_0 = 0$,

$$
\begin{aligned}
& P((T_1,\dots,T_n)\in I) = P\left(\bigcap_{i=1}^{n}\{a_i < T_i \le b_i\}\right)\\
= \;& P(N(a_i)-N(b_{i-1})=0,\ N(b_i)-N(a_i)=1,\ 1\le i\le n-1,\\
& N(a_n)-N(b_{n-1})=0,\ N(b_n)-N(a_n)\ge 1)\\
= \;& \left[\prod_{i=1}^{n-1} P(N(a_i)-N(b_{i-1})=0)\cdot P(N(b_i)-N(a_i)=1)\right]\\
& \cdot P(N(a_n)-N(b_{n-1})=0)\cdot P(N(b_n)-N(a_n)\ge 1)\\
= \;& \left[\prod_{i=1}^{n-1} e^{-\lambda(a_i-b_{i-1})}\lambda(b_i-a_i)e^{-\lambda(b_i-a_i)}\right]\cdot\ e^{-\lambda(a_n-b_{n-1})}(1-e^{-\lambda(b_n-a_n)})\\
= \;& \left[\prod_{i=1}^{n-1}(b_i-a_i)\right]\lambda^n(e^{-\lambda a_n}-e^{-\lambda b_n})\\
= \;& \int_{a_1}^{b_1}\int_{a_2}^{b_2}\cdots\int_{a_{n-1}}^{b_{n-1}}\int_{a_n}^{b_n}\lambda^n e^{-\lambda t_n}dt_n\ dt_{n-1}\dots dt_2\ dt_1\\
= \;& \int_I\cdots\int\lambda^n e^{-\lambda t_n}dt_1\ dt_2\dots dt_n.
\end{aligned}
$$

Since rectangles of the form $I$ generate the Borel $\sigma$-algebra of $G_n$, (2.12) now follows.

(b) Observe that

$$
\begin{pmatrix} A_1\\ A_2\\ A_3\\ \cdot\\ \cdot\\ \cdot\\ A_n \end{pmatrix} = \begin{pmatrix} T_1\\ T_2-T_1\\ T_3-T_2\\ \cdot\\ \cdot\\ \cdot\\ T_n-T_{n-1} \end{pmatrix}
$$

$$
= \begin{pmatrix} 1 & 0 & 0 & 0 & \cdot & \cdot & \cdot & 0 & 0\\ -1 & 1 & 0 & 0 & \cdot & \cdot & \cdot & 0 & 0\\ 0 & -1 & 1 & 0 & \cdot & \cdot & \cdot & 0 & 0\\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot\\ 0 & 0 & 0 & 0 & \cdot & \cdot & \cdot & -1 & 1 \end{pmatrix} \begin{pmatrix} T_1\\ T_2\\ T_3\\ \cdot\\ \cdot\\ \cdot\\ T_n \end{pmatrix}. \qquad (2.14)
$$

The linear transformation given by the $(n \times n)$ matrix in (2.14) has determinant 1, and transforms the region $G_n$ in one-one fashion onto the orthant $\{(a_1, a_2, \ldots, a_n) : a_i > 0, 1 \leq i \leq n\}$. So by (2.12) and the change of variables formula the joint probability density function $f_{(A_1,\ldots,A_n)}$ of $A_1, \ldots, A_n$ is given by

$$\begin{aligned} & f_{(A_1,\ldots,A_n)}(a_1, a_2, \ldots, a_n) \\ = \; & f_{(T_1,\ldots,T_n)}(a_1, a_1 + a_2, \ldots, a_1 + a_2 + \ldots + a_n) \\ = \; & \begin{cases} \lambda\, e^{-\lambda a_1} \lambda\, e^{-\lambda a_2} \ldots \lambda\, e^{-\lambda a_n}, & \text{if } a_i > 0, 1 \leq i \leq n \\ 0, & \text{otherwise.} \end{cases} \end{aligned} \tag{2.15}$$

The required conclusion now follows from (2.15). □

As $A_i$ has Exp($\lambda$) distribution, its expectation is $E(A_i) = \frac{1}{\lambda}$; so $\frac{1}{\lambda}$ is the mean arrival time. Thus the arrival rate being $\lambda$ is consistent with this conclusion.

One can also go in the other direction as the next result indicates.

**Theorem 2.5** *Let $\{A_i : i \geq 1\}$ be a sequence of i.i.d. Exp($\lambda$) random variables, where $\lambda > 0$. Set $T_0 \equiv 0$, and $T_n = A_1 + \ldots + A_n$, for $n \geq 1$. Define the claim number process $N(\cdot)$ by*

$$N(t) = \max\{i \geq 0 : T_i \leq t\}, \; t \geq 0. \tag{2.16}$$

*Then $\{N(t) : t \geq 0\}$ is a time homogeneous Poisson process with rate $\lambda$.*

**Proof:** As $A_1$ has an exponential distribution, $A_1 > 0$ a.s. and hence $T_1 > 0$ a.s. So $N(0) = 0$ a.s. We next prove that $N(t)$ has a Poisson distribution with parameter $\lambda t$, for any $t > 0$. Indeed, for $k \geq 1$, $T_k = A_1 + A_2 + \ldots + A_k$ has $\Gamma(k, \lambda)$ distribution, being sum of $k$ independent Exp($\lambda$) random variables. So it can be verified that

$$\begin{aligned} P(T_k \leq t) &= \int_0^t \frac{\lambda^k}{(k-1)!}\, z^{k-1} e^{-\lambda z}\, dz \\ &= 1 - \sum_{j=0}^{k-1} e^{-\lambda t}\, \frac{(\lambda t)^j}{j!}. \end{aligned} \tag{2.17}$$

Consequently for $n = 0, 1, 2, \ldots$

$$\begin{aligned} P(N(t) = n) &= P(T_n \leq t) - P(T_{n+1} \leq t) \\ &= e^{-\lambda t}\, \frac{(\lambda t)^n}{n!} \end{aligned} \tag{2.18}$$

proving the claim.

To prove that $\{N(t)\}$ has independent stationary increments, we shall consider only the case of $N(t)$, $N(t+h)-N(t)$ being independent and $N(t+h)-N(t) \stackrel{d}{=} N(h)$, for $t,\ h>0$. Writing

$$q_{k,k+\ell}\ (t,t+h) \triangleq P(N(t)=k,\ N(t+h)-N(t)=\ell)$$

note that we need to establish (why ?)

$$q_{k,k+\ell}\ (t,t+h) = e^{-\lambda(t+h)}\ \frac{(\lambda t)^k}{k!}\ \frac{(\lambda h)^\ell}{\ell!} \tag{2.19}$$

for $t,\ h>0,\ k,\ \ell=0,1,2,\ldots$.

We shall consider the case $k \geq 1,\ \ell \geq 2$. (The other cases are simpler to handle.) Observe that

$$\begin{aligned} & \{N(t)=k, N(t+h)-N(t)=\ell\} \\ = \ & \{N(t)=k, N(t+h)=k+\ell\} \\ = \ & \{T_k \leq t < T_{k+1},\ T_{k+\ell} \leq t+h < T_{k+\ell+1}\}. \end{aligned} \tag{2.20}$$

We know that $T_k, A_{k+1}, \hat{T} \triangleq \sum_{i=k+2}^{k+\ell} A_i, A_{k+\ell+1}$ are independent random variables having respectively $\Gamma(k,\lambda)$, $\mathrm{Exp}(\lambda)$, $\Gamma(\ell-1,\lambda)$, $\mathrm{Exp}(\lambda)$ distributions. Therefore the joint probability density function of $(T_k, A_{k+1}, \hat{T}, A_{k+\ell+1})$ is given by

$$\begin{aligned} & f(z_1,z_2,z_3,z_4) \\ = \ & \begin{cases} \left(\frac{\lambda^k}{(k-1)!} z_1^{k-1} e^{-\lambda z_1}\right)(\lambda e^{-\lambda z_2})\left(\frac{\lambda^{\ell-1}}{(\ell-2)!}\ z_3^{\ell-2} e^{-\lambda z_3}\right)(\lambda e^{-\lambda z_4}), \\ \qquad\qquad \text{if } z_1,z_2,z_3,z_4>0 \\ 0, \text{ otherwise.} \end{cases} \end{aligned} \tag{2.21}$$

Using (2.20) we now get

$$\begin{aligned} & q_{k,k+\ell}(t,t+h) = P(T_k \le t < T_{k+1}, T_{k+\ell} \le t+h < T_{k+\ell+1}) \\ = \; & P(T_k \le t < T_k + A_{k+1} < T_k + A_{k+1} + \hat{T} \\ & \le t+h < T_k + A_{k+1} + \hat{T} + A_{k+\ell+1}) \\ = \; & \int\int_F \int\int f(z_1,z_2,z_3,z_4)dz_4dz_3dz_2dz_1 \\ = \; & \int_0^t \int_{t-z_1}^{t+h-z_1} \int_0^{t+h-(z_1+z_2)} \int_{t+h-(z_1+z_2+z_3)}^{\infty} f(z_1,z_2,z_3,z_4)dz_4dz_3dz_2dz_1 \end{aligned} \tag{2.22}$$

where $F = \{(z_1,z_2,z_3,z_4) : 0 \le z_1 \le t < (z_1+z_2) < (z_1+z_2+z_3) \le t+h < (z_1+z_2+z_3+z_4)\}$. Plugging (2.21) into (2.22), an elementary integration now leads to (2.19).

To show that $N(t_1), N(t_2)-N(t_1), \ldots, N(t_n)-N(t_{n-1})$ are independent random variables, for $0 < t_1 < t_2 < \ldots < t_n < \infty, n = 3,4,\ldots$, it is clear that the above approach will work though the book-keeping will be more tedious. Also the above argument shows that $N(t+s) - N(s)$ and $N(t)$ have the same distribution (viz. Poisson distribution with parameter $\lambda t$) for any $t \ge 0, s \ge 0$.

As $0 < T_1 < T_2 < \ldots < T_n < \infty$ a.s. it is seen from our construction (that is, (2.16)) that the sample paths of $N$ are right continuous at every $t \ge 0$ and have left limits at every $t > 0$. This completes the proof. □

In the jargon of the theory of stochastic processes, time homogeneous Poisson process is the renewal process with i.i.d. exponential arrival rates. Proposition 2.3 and Theorem 2.5 imply the existence of Poisson process for a large class of intensity functions, once a sequence of i.i.d. exponential random variables exist. The latter requirement is, of course, guaranteed by the consistency theorem.

The next result concerns age and residual life. These terms have come from reliability theory.

**Proposition 2.6** *Let $N(\cdot)$ be a time homogeneous Poisson process with rate $\lambda > 0$. Let $\{T_i\}$ denote the corresponding arrival times. Fix $t \ge 0$. Define*

$$B(t) = t - T_{N(t)}, \quad F(t) = T_{N(t)+1} - t. \tag{2.23}$$

*$B(t)$ is called the backward recurrence time or age, while $F(t)$ is called the forward recurrence time or excess life or residual life. Let $G_{B(t),F(t)}$ denote the distribution function of the two dimensional random variable $(B(t), F(t))$. Then*

$$\begin{aligned} & G_{B(t),F(t)}(x_1, x_2) \\ = & \begin{cases} (1 - e^{-\lambda x_1})(1 - e^{-\lambda x_2}), & \text{if } 0 \le x_1 < t, x_2 \ge 0 \\ (1 - e^{-\lambda x_2}), & \text{if } x_1 \ge t, x_2 \ge 0 \\ 0, & \text{otherwise.} \end{cases} \end{aligned} \tag{2.24}$$

*In particular (i) $B(t)$ and $F(t)$ are independent random variables, (ii) $F(t)$ has Exp($\lambda$) distribution, and (iii) $B(t)$ has a truncated exponential distribution (with parameter $\lambda$) with a jump at $t$, that is*

$$\begin{aligned} P(B(t) \le x_1) &= (1 - e^{-\lambda x_1}), \text{ if } 0 \le x_1 < t \\ P(B(t) = t) &= e^{-\lambda t}. \end{aligned}$$

*So $P(B(t) \le x_1) = 1$ if $x_1 \ge t$. Of course $P(B(t) \le x_1) = 0$ if $x_1 < 0$.*

**Proof:** Once (2.24) is established, the other assertions are easily obtained. As $t$ is fixed, note that the last arrival took place at time $T_{N(t)}$, and that the next arrival will happen at time $T_{N(t)+1}$. Clearly $B(t) \le t$ a.s. For $0 \le x_1 < t$ note that

$$\{B(t) < x_1\} = \{t - x_1 < T_{N(t)} \le t\} = \{N(t) - N(t - x_1) \ge 1\}.$$

Similarly for $x_2 > 0$

$$\{F(t) \le x_2\} = \{t < T_{N(t)+1} \le t + x_2\} = \{N(t + x_2) - N(t) \ge 1\}.$$

Therefore, as $N$ has independent increments, for $0 \le x_1 < t, x_2 \ge 0$ we get

$$\begin{aligned} & P(B(t) < x_1, F(t) \le x_2) \\ = & P(N(t) - N(t - x_1) \ge 1,\ N(t + x_2) - N(t) \ge 1) \\ = & P(N(t) - N(t - x_1) \ge 1) \cdot P(N(t + x_2) - N(t) \ge 1) \\ = & (1 - e^{-\lambda x_1})(1 - e^{-\lambda x_2}). \end{aligned}$$

Continuity of r.h.s of the above in $x_1$ over $[0, t)$ implies

$$P(B(t) \le x_1, F(t) \le x_2) = (1 - e^{-\lambda x_1})(1 - e^{-\lambda x_2})$$

for $0 \leq x_1 < t,\ x_2 \geq 0$. For $x_1 \geq t,\ x_2 \geq 0$ it is easily seen that

$$\begin{aligned} & P(B(t) \leq x_1,\ F(t) \leq x_2) = P(F(t) \leq x_2) \\ = \ & P(N(t+x_2) - N(t) \geq 1) = (1 - e^{-\lambda x_2}). \end{aligned}$$

Thus (2.24) is proved. □

It may look a bit paradoxical that the residual life and the interarrival time have the same distribution. But this is a manifestation of the memoryless property of the exponential distribution.

## 2.3 Order statistics property

This is another important property of the Poisson process. We begin with an elementary result.

**Proposition 2.7** *Let $N(\cdot)$ be a homogeneous Poisson process with rate $\lambda > 0$, and let $T_1$ be the first arrival time. Then for $0 \leq s \leq t$,*

$$P(T_1 < s | N(t) = 1) = \frac{s}{t};$$

*that is, given that exactly one arrival has taken place in $[0,t]$, the time of arrival is uniformly distributed over $[0,t]$.*

**Proof:** By the stationary independent increment property we obtain

$$\begin{aligned} P(T_1 < s \mid N(t) = 1) &= \frac{P(T_1 < s, N(t) = 1)}{P(N(t) = 1)} \\ &= \frac{P(N(s) = 1, N(t) - N(s) = 0)}{P(N(t) = 1)} \\ &= \frac{P(N(s) = 1) \cdot P(N(t) - N(s) = 0)}{P(N(t) = 1)} \\ &= \frac{\lambda s\ e^{-\lambda s}\ e^{-\lambda(t-s)}}{\lambda t\ e^{-\lambda t}} = \frac{s}{t}. \end{aligned}$$

□

A natural question is: If $N(t) = n$, what can one say about the conditional distribution of $(T_1, T_2, \ldots, T_n)$? For fixed $t > 0$, $n \geq 1$ note that $P(N(t) = n) > 0$. Put $\hat{\Omega} = \{N(t) = n\}, \hat{\mathcal{F}} = \mathcal{F} \cap \hat{\Omega} \triangleq \{E \cap \hat{\Omega} : E \in \mathcal{F}\}$; define the probability measure $\hat{P}$ on $(\hat{\Omega}, \hat{\mathcal{F}})$ by $\hat{P}(\hat{E}) = \frac{1}{P(\hat{\Omega})} P(E \cap \hat{\Omega}) = \frac{P(E \cap \hat{\Omega})}{P(N(t)=n)}$, for $\hat{E} = E \cap \hat{\Omega} \in \hat{\mathcal{F}}$. On the probability space

$(\hat{\Omega}, \hat{\mathcal{F}}, \hat{P})$ note that $(T_1, T_2, \ldots, T_n)$ is an $\mathbb{R}^n$-valued random variable. So the question is : What is the probability measure $\hat{P}(T_1, T_2, \ldots, T_n)^{-1}$ on $\mathbb{R}^n$? To appreciate the answer one needs the following.

**Theorem 2.8** *Let $Y_1, Y_2, \ldots, Y_n$ be i.i.d. random variables with common probability density function $f(\cdot)$. Let $Y_{(1)} \leq Y_{(2)} \ldots \leq Y_{(n)}$ denote the order statistics of $Y_1, Y_2, \ldots, Y_n$; that is, $Y_{(k)}(\omega)$ is the k-th smallest value among $\{Y_1(\omega), Y_2(\omega), \ldots, Y_n(\omega)\}$, $k = 1, 2, \ldots, n$ for any $\omega \in \Omega$. Then the joint probability density function of $Y_{(1)}, Y_{(2)}, \ldots, Y_{(n)}$ is given by*

$$\begin{aligned} & g(y_1, y_2, \ldots, y_n) \\ = & \begin{cases} (n!) f(y_1) f(y_2) \ldots f(y_n), & \text{if } y_1 < y_2 < \ldots y_n < \infty \\ 0, & \text{otherwise.} \end{cases} \end{aligned} \tag{2.25}$$

*In particular, if $V_1, V_2, \ldots, V_n$ are i.i.d. $U(0,t)$ random variables (that is, having uniform distribution over $[0,t]$), then the joint probability density function of the order statistics $V_{(1)}, V_{(2)}, \ldots, V_{(n)}$ is given by*

$$\begin{aligned} & f_{V_{(1)}, V_{(2)}, \ldots, V_{(n)}}(s_1, s_2, \ldots, s_n) \\ = & \begin{cases} (n!)/(t^n), & \text{if } 0 < s_1 < s_2 < \ldots < s_n < t \\ 0, & \text{otherwise.} \end{cases} \end{aligned} \tag{2.26}$$

**Proof:** We need to prove only the first assertion. Since $(Y_1, Y_2, \ldots, Y_n)$ is an $\mathbb{R}^n$-valued absolutely continuous random variable, with probability one there are no ties; that is, $Y_{(1)} < Y_{(2)} < \ldots < Y_{(n)}$ a.s. Clearly the joint probability density function of $Y_1, Y_2, \ldots, Y_n$ is

$$f_{Y_1, Y_2, \ldots, Y_n}(y_1, y_2, \ldots, y_n) = f(y_1)\ f(y_2) \ldots f(y_n).$$

Note that the $n$-dimensional random variable $(Y_{(1)}, Y_{(2)}, \ldots, Y_{(n)})$ takes values in the set $H = \{(y_1, y_2, \ldots, y_n) \in \mathbb{R}^n : y_1 < y_2 < \ldots < y_n\}$. Let $B \subseteq H$ be a Borel set. Let $B_i$, $i = 1, 2, \ldots, (n!)$ correspond to disjoint sets obtained by permutation of coordinates of elements of $B$, with $B_1 = B$. Observe that

$$
\begin{aligned}
& P((Y_{(1)}, Y_{(2)}, \ldots, Y_{(n)}) \in B) = P\left(Y_1, Y_2, \ldots, Y_n) \in \bigcup_{i=1}^{n!} B_i\right) \\
= & \sum_{i=1}^{n!} P((Y_1, Y_2, \ldots, Y_n) \in B_i) \\
= & \sum_{i=1}^{n!} \int_{B_i} f(z_1) f(z_2) \ldots f(z_n)\, dz_1\, dz_2 \ldots dz_n \\
= & \; n! \int_{B} f(z_1) f(z_2) \ldots f(z_n) dz_1\, dz_2 \ldots dz_n
\end{aligned}
$$

whence (2.25) follows. □

Here is the main result of this section

**Theorem 2.9** *Let $N(\cdot)$ be a homogeneous Poisson process with rate $\lambda > 0$, and $\{T_i\}$ the associated arrival times. Fix $t > 0$, $n = 1, 2, \ldots$. Then the conditional probability density function of $(T_1, T_2, \ldots, T_n)$ given $N(t) = n$ is*

$$
\begin{aligned}
& f_{T_1, T_2, \ldots, T_n}\left((s_1, s_2, \ldots . s_n) | N(t) = n\right) \\
= & \begin{cases} (n!)/(t^n), & \text{if } 0 < s_1 < s_2 < \ldots < s_n < t \\ 0, & \text{otherwise.} \end{cases}
\end{aligned} \tag{2.27}
$$

**Proof:** Denote for $t > 0$ , $n = 1, 2, \ldots$

$$
G_{n,t} = \{(s_1, s_2, \ldots, s_n) : 0 < s_1 < s_2 < \ldots < s_n < t\}. \tag{2.28}
$$

If a "rectangle" $I = (a_1, b_1] \times (a_2, b_2] \times \ldots \times (a_n, b_n] \subset G_{n,t}$ then $0 \leq a_1 < b_1 \leq a_2 < \ldots < b_{n-1} \leq a_n < b_n < t$. Now by (2.12) in Theorem

2.4

$$
\begin{aligned}
& P((T_1, T_2, \ldots, T_n) \in I | N(t) = n) \\
= & \frac{1}{P(N(t) = n)} P(a_1 < T_1 \leq b_1, \ldots, a_n < T_n \leq b_n, N(t) = n) \\
= & \frac{n!}{(\lambda t)^n e^{-\lambda t}} P(a_1 < T_1 \leq b_1, \ldots, a_n < T_n \leq b_n, T_{n+1} > t) \\
= & \frac{n!}{(\lambda t)^n e^{-\lambda t}} \int_{a_1}^{b_1} \int_{a_2}^{b_2} \cdots \int_{a_n}^{b_n} \int_{t}^{\infty} \lambda^{n+1} e^{-\lambda s_{n+1}} \, ds_{n+1} ds_n \ldots ds_2 ds_1 \\
= & \int_I \cdots \int \frac{n!}{t^n} \, ds_1 \, ds_2 \ldots ds_n
\end{aligned}
$$

Since rectangles of the form $I$ generate the Borel $\sigma$-algebra of $G_{n,t}$, (2.27) is now immediate. □

**Remark 2.10** By (2.26), (2.27) it is now clear that the conditional distribution of $(T_1, T_2, \ldots, T_n)$ given $N(t) = n$ is the same as the order statistics of $n$ i.i.d. $U(0, t)$ random variables. It is convenient to denote this fact by

$$
((T_1, T_2, \ldots, T_n) | N(t) = n) \stackrel{d}{=} (tU_{(1)}, tU_{(2)}, \ldots, tU_{(n)}) \tag{2.29}
$$

where $U_{(1)} < U_{(2)} < \ldots < U_{(n)}$ is the order statistics corresponding to i.i.d. $U(0, 1)$ random variables $U_1, U_2, \ldots, U_n$, and $\stackrel{d}{=}$ denotes that the two sides have the same distribution. Moreover, by Kolmogorov consistency theorem, the $U_i$'s can be chosen independently of $\{N(t) : t \geq 0\}$ or $\{T_j\}$. □

## 2.4 Cramér-Lundberg Model

This is the classical and a very important model in insurance; it has been extensively studied. As we will be seeing it often in the sequel, in this section we shall just introduce the model. The basic assumptions of the model are as follows.

1. Claims happen at the arrival times $0 < T_1 < T_2 < \ldots < T_n < \ldots$ of a homogeneous Poisson process $\{N(t) : t \geq 0\}$ with rate, say, $\lambda > 0$.

2. The $i^{\text{th}}$ claim arrival at time $T_i$ causes claim size $X_i$. The sequence $\{X_i\}$ is an i.i.d. sequence of nonnegative random variables.

3. The sequences $\{T_i\}$ and $\{X_j\}$ are independent; equivalently the process $\{N(t)\}$ and the sequence $\{X_j\}$ are independent.

The total claim amount upto time $t$ in this model is given by

$$S(t) = \sum_{i=1}^{N(t)} X_i, \quad t \geq 0 \tag{2.30}$$

which is the same as (1.2).

**Theorem 2.11** *Let (1) - (3) above hold. Then the process* $\{S(t) : t \geq 0\}$ *has independent increments and is time homogeneous.*

**Proof:** Let $k = 2, 3, \ldots, 0 < t_1 < t_2 < \ldots < t_k < \infty$. By our assumptions note that

$$\begin{aligned}
&\{X_1, \ldots, X_{N(t_1)}, N(t_1)\}, \\
&\{X_{N(t_1)+1}, \ldots, X_{N(t_2)}, N(t_2) - N(t_1)\}, \ldots, \\
&\{X_{N(t_{k-1})+1}, \ldots, X_{N(t_k)}, N(t_k) - N(t_{k-1})\}
\end{aligned}$$

are independent. Hence the increments

$$\begin{aligned}
S(t_1) - S(0) &= X_1 + \ldots + X_{N(t_1)}, \\
S(t_2) - S(t_1) &= X_{N(t_1)+1} + \ldots + X_{N(t_2)}, \ldots, \\
S(t_k) - S(t_{k-1}) &= X_{N(t_{k-1})+1} + \ldots + X_{N(t_k)}
\end{aligned}$$

are independent. Thus $\{S(t) : t \geq 0\}$ has independent increments.

Next let $t > 0$, $h > 0$. It is not difficult to show that $(X_1, \ldots, X_{N(h)}, N(h))$ has the same distribution as $(X_{N(t)+1}, \ldots, X_{N(t+h)}, N(t+h) - N(t))$. Hence it follows that $S(h) = X_1 + \ldots + X_{N(h)}$ has the same distribution as $S(t+h) - S(t) = X_{N(t)+1} + \ldots + X_{N(t+h)}$. Thus $S$ is time homogeneous. □

A process $S(\cdot)$ given by (2.30), where $\{N(t)\}, \{X_i\}$ satisfy (1) - (3) above, is sometimes called a *compound Poisson process*; the preceding result says that the process $S(\cdot)$ is a *Levy process*.

Next we look at the *discounted sum* corresponding to the Cramér-Lundberg model. Let $r > 0$ denote the interest rate. Define for $t \geq 0$

$$S_0(t) = \begin{cases} 0, \text{ if } N(t) = 0 \\ \sum_{i=1}^{k} e^{-rT_i} X_i, \text{ if } N(t) = k, \ k = 1, 2, \ldots \end{cases} \tag{2.31}$$

This is the "present value" (at time 0) of the cumulative claim amount over the time horizon $[0,t]$. Observe that by Theorem 2.9 and (2.29),

$$\begin{aligned}
E(S_0(t)) &= E\left(\sum_{i=1}^{N(t)} e^{-rT_i} X_i\right) \\
&= \sum_{n=1}^{\infty} E\left(\sum_{i=1}^{n} e^{-rT_i} X_i \mid N(t)=n\right) \cdot P(N(t)=n) \\
&= \sum_{n=1}^{\infty} \sum_{i=1}^{n} E[e^{-rT_i} X_i \mid N(t)=n] \cdot P(N(t)=n) \\
&= E(X_1) \sum_{n=1}^{\infty} \sum_{i=1}^{n} E(e^{-rT_i} \mid N(t)=n) \cdot P(N(t)=n) \\
&= E(X_1) \sum_{n=1}^{\infty} E\left(\sum_{i=1}^{n} e^{-rT_i} \mid N(t)=n\right) \cdot P(N(t)=n) \\
&= E(X_1) \sum_{n=1}^{\infty} E\left(\sum_{i=1}^{n} e^{-rtU_{(i)}}\right) \cdot P(N(t)=n) \\
&= E(X_1) \sum_{n=1}^{\infty} E\left(\sum_{i=1}^{n} e^{-rtU_i}\right) \cdot P(N(t)=n) \\
&= E(X_1) E(e^{-rtU_1}) E(N(t)) = \lambda t \; E(X_1) \left(\int_0^1 e^{-rty} dy\right) \\
&= \lambda \; E(X_1) \; \frac{1}{r}(1-e^{-rt})
\end{aligned}$$

where we have used independence of $\{X_i\}$ and $\{N(t)\}$, and $\sum_{i=1}^{n} e^{-rtU_{(i)}} = \sum_{i=1}^{n} e^{-rtU_i}$. Thus we have

**Theorem 2.12** *Let $S_0(\cdot)$ be the discounted sum process defined by (2.31) corresponding to the Cramér-Lundberg model. Let $\lambda > 0$, $r > 0$ denote respectively the claim arrival rate and the interest rate. Assume that the claim size has finite expectation. Then*

$$E(S_0(t)) = \lambda \; E(X_1) \frac{1}{r} \; (1-e^{-rt}), \quad t > 0. \tag{2.32}$$

□

In the Cramér-Lundberg model, the average/expected amount (discounted at 0) needed to take care of claims over $[0,t]$ is given by (2.32). A reasonable explanation can be thus. In insurance business, considerable part of the financial risk is on the company. A pilot study before starting the business might indicate average claim size, claim arrival rate, interest rate, etc. Also it may take some time for the premium income to stabilize, perhaps a year or so. Therefore, during the initial period any claim settlement will have to be met from the initial capital mobilized by the company. Hence (2.32) can be regarded as a rough estimate of the average initial capital needed to start the business.

A related process is

$$S_1(t) = \begin{cases} 0, \text{ if } N(t) = 0 \\ \sum_{i=1}^{k} e^{r(t-T_i)} X_i, \text{ if } N(t) = k, \ k = 1, 2, \ldots \end{cases} \tag{2.33}$$

for $t \geq 0$, where $r > 0$ again denotes interest rate. This process can be interpreted as follows. Let $t > 0$ be fixed. Suppose claims which arrive during $[0,t]$ are settled only at time $t$. So there is a delay in claim settlement, and hence the claim amount accrues an interest compounded continuously at rate $r > 0$. So $S_1(t)$ is the total liability of the insurance company at time $t$ due to claims occurring over $[0,t]$. Hence $E(S_1(t))$ is the expected amount that the company should have at time $t$ to honour its commitments.

**Notes:** The material considered in this Chapter is standard. One may look up [Bl2], [KT], [Re], [Ro1], [Ro2], [RSST], [M], etc.

## 2.5 Exercises

### Exercises on Section 2.1

**2.1.1.** Let $N$ be a homogeneous Poisson process with rate $\lambda > 0$.

(a) For any fixed $t > 0$, show that $P(N(t) = 0) > 0$, and $P(N(t) \geq k) > 0$ for $k = 1, 2, \ldots$.

(b) For any fixed $t \geq 0$, show that $P$ - a.e. $\omega$ the sample path $N(\cdot, \omega)$ does not have a jump at $t$.

(c) Show that (b) above does not mean that sample paths of $N(\cdot)$ are continuous. In fact, with probability one, sample paths have discontinuities.

(d) For $P$ - a.e. $\omega \in \Omega$, for any $t > 0$, show that the number of jumps of $N(\cdot, \omega)$ over $[0, t]$ is finite. That is, almost every sample path has only finitely many jumps over any time interval of finite length.

**2.1.2.** Let $N(\cdot)$ be a Poisson process with rate function $\lambda(\cdot)$. For $n = 2, 3, \ldots$ and $0 < t_1 < t_2 < \ldots < t_n < \infty$, find the joint distribution of $N(t_1), N(t_2), \ldots, N(t_n)$.

**2.1.3.** Let $N(\cdot)$ be as in 2.1.2. For $n = 1, 2, \ldots, 0 < t_1 < t_2 < \ldots < t_n < t_{n+1} < \infty$, nonnegative integers $k_i$ with $0 \leq k_1 \leq k_2 \leq \ldots \leq k_n \leq k_{n+1}$ show that

$$\begin{aligned} P(N(t_{n+1}) = k_{n+1} \quad | \quad & N(t_1) = k_1, \ldots, N(t_n) = k_n) \\ = \quad & P(N(t_{n+1}) = k_{n+1} \mid N(t_n) = k_n). \end{aligned}$$

That is, the *Markov property* holds for the Poisson process.

**2.1.4.** Let $\{N(t) : t \geq 0$ be a Poisson process with a continuous rate function $\lambda(\cdot)$. For any $t \geq 0$ show that

(a) $P(N(t+h) - N(t) \geq 2) = o(h)$, as $h \downarrow 0$

(b) $P(N(t+h) - N(t) = 1) = \lambda(t)h + o(h) = \lambda(t)\, h(1 + o(1))$, as $h \downarrow 0$.

**2.1.5.** Let $\{N(t)\}$ be a homogeneous Poisson process with rate $\lambda > 0$. Show that

$$\frac{N(t) - \lambda t}{\sqrt{\lambda t}} \Rightarrow N(0, 1), \text{ as } t \to \infty$$

where $\Rightarrow$ denotes weak convergence of probability measures (or convergence in distribution) and $N(0, 1)$ denotes the one dimensional standard normal distribution. (Hint: Use Fourier transforms / characteristic functions.)

**2.1.6.** Let $\{N(t) : t \geq 0\}$ satisfy (N1)-N(6). Fix $n$. Show that $t \mapsto P_n(t)$ is differentiable from the left for $t > 0$. (Hint: Derive an analogue of the display before (2.6) connecting $P_n(t)$ and $P_n(t - h)$).

## Exercises on Section 2.2

**2.2.1.** (a) Let $N(\cdot)$ be a Poisson process. Show that $T_n$ given by (2.9) is a random variable.

(b) For $t \geq 0$ let $\mathcal{F}_t^N \triangleq \sigma\{N(r) : 0 \leq r \leq t\}$ be the smallest $\sigma$-algebra making the family $\{N(r) : 0 \leq r \leq t\}$ measurable; $\{\mathcal{F}_t^N : t \geq 0\}$ is said to be the *filtration generated by the process* $N$. For $t \geq 0$, $n = 1, 2, \ldots$

show that $\{T_n \leq t\} \in \mathcal{F}_t^N, \{A_n \leq t\} \in \mathcal{F}_t^N$. That is, each $T_n$ and hence each $A_n$ is a *stopping time w.r.t. the filtration* $\{\mathcal{F}_t^N\}$.

**2.2.2.** Let $T_1$ be the first arrival time corresponding to a time homogeneous Poisson process. For $t, s \geq 0$ show that

$$P(T_1 > t + s | T_s > s) = P(T_1 > t).$$

This is the so called *memoryless property* of the exponential distribution.

Conversely, if the above equation holds for all $s,\ t \geq 0$ and $T_1 > 0$ a.s., show that $T_1$ has an exponential distribution.

**2.2.3.** Let $N$ be a time homogeneous Poisson process with rate $\lambda > 0$. Let $T_1,\ T_2$ be the arrival time of the first two claims. For $0 \leq t_1 \leq t_2$ show that

$$\{T_1 \leq t_1, T_2 \leq t_2\} = \{N(t_1) = 1, N(t_2) - N(t_1) \geq 1\} \cup \{N(t_1) \geq 2\}$$

and hence that the joint distribution function of $T_1,\ T_2$ is given by

$$F_{(T_1,T_2)}\ (t_1, t_2) = -\lambda t_1 e^{-\lambda t_2} + H(t_1),\ \text{if } 0 \leq t_1 \leq t_2,$$

where $H$ is a function depending only on $t_1$. Find the joint probability density function of $T_1,\ T_2$ from the above.

**2.2.4.** Let $\{N(t) : t \geq 0\}$ be a homogeneous Poisson process with rate $\lambda > 0$. Let $n \geq 1$ be an integer and $G_n$ be defined by (2.13). Let $1 \leq m_1 \leq m_2 \leq \ldots \leq m_n < \infty$ each $m_i$ being an integer. For $(t_1, t_2, \ldots, t_n) \in G_n$ show that

$$\begin{aligned} & P(N(t_1) \geq m_1, N(t_2) \geq m_2, \ldots, N(t_n) \geq m_n) \\ = & \sum_{\text{finite}} \left\{ \begin{array}{l} [\text{Polynomial in } t_1, t_2, \ldots, t_n] \\ \times [\exp\{-\lambda(\text{linear function of } t_1, \ldots, t_n)\}] \end{array} \right\} \end{aligned}$$

and hence is a $C^\infty$-function on the region $G_n$.
(Hint: $\{N(t_1) \geq m_1,\ N(t_2) \geq m_2, \ldots, N(t_n) \geq m_n\} =$

$$\bigcup_{k=1}^{n} \bigcup_{\substack{1 \leq j_1 \leq j_2 \leq \cdots \leq j_{k-1} \leq m_n - 1, \\ j_\ell \geq m_\ell, \ell = 1, \ldots, k-1}} \left\{ \begin{array}{l} N(t_k) \geq m_n,\ N(t_{k-1}) = j_{k-1}, \\ \ldots, N(t_1) = j_1\} \end{array} \right\}.$$

Also, $\sum_{\text{finite}}$ denotes sum of a finite number of terms.)

**2.2.5.** Let $N$ be as in Exercise 2.2.4. Let $\{T_i\}$ be the arrival times. Fix an integer $n > 1$. Let $G_n$ be given by (2.13). Let $F_{T_1,\ldots,T_n}$ denote the

joint distribution function of $T_1, T_2, \ldots, T_n$. For $\underline{t} = (t_1, t_2, \ldots, t_n) \in G_n$ show that

$$F_{T_1,T_2,\ldots,T_n}(t_1, t_2, \ldots, t_n) = -\lambda^{n-1} t_1 t_2 \ldots t_{n-1} e^{-\lambda t_n} + H(\underline{t})$$

where $\frac{\partial^n}{\partial t_1 \partial t_2 \ldots \partial t_n} H(\underline{t}) = 0$.

(Hint: (i) Need to consider only $k = n$ in the hint to Exercise 2.2.4 with $m_i = i$.

(ii) Show that we need to look at only $\{\frac{-\lambda^{n-1}}{j_1!(j_2-j_1)!\ldots(n-1-j_{n-2})!}\} \times \{t_1^{j_1}(t_2-t_1)^{j_2-j_1} \ldots (t_n - t_{n-1})^{n-1-j_{n-2}} \cdot e^{-\lambda t_n}\}$ where $j_{n-1} = n-1$, $j_\ell \geq \ell$, $\ell = 1, 2, \ldots, n-1$, $j_1 \leq j_2 \leq \ldots \leq j_{n-1}$. That is, terms not of this form will be part of $H(\underline{t})$.

(iii) Unless $j_\ell = \ell$ for all $\ell = 1, 2, \ldots, n-1$ show that the coefficient of $t_1 t_2 \ldots t_{n-1}$ in the expression in (ii) above is zero.)

**2.2.6.** Let $\{T_n\}$ be as in Exercise 2.2.5.

(a) Show that $T_n$ has the gamma distribution $\Gamma(n, \lambda)$. This distribution is also called *Erlang distribution* with parameters $n, \lambda$ in queueing theory.

(b) Show that $0 < T_1 < T_2 < \ldots < T_n < \infty$ a.s.

**2.2.7.** Derive (2.19) in cases other than $k \geq 1$, $\ell \geq 2$.

**2.2.8.** Let $\{N(t) : t \geq 0\}$ be a Poisson process with rate function $\lambda(\cdot)$. Set $\mu(t) = \int_0^t \lambda(s)\, ds$, $t \geq 0$. Assume that $\lambda(\cdot)$ is continuous and $\lambda(r) > 0$ a.e.$r$ (w.r.t. the Lebesgue measure on $[0, \infty)$). Let $\{T_i\}$ be the claim arrival times, and $\{A_j\}$ the interarrival times for the process $N(\cdot)$ defined by analogues of (2.9), (2.10). Then the following hold.

(a) Fix $n \geq 1$. Then the probability density function of $(T_1, T_2, \ldots, T_n)$ is

$$f_{T_1,T_2,\ldots,T_n}(t_1, t_2, \ldots, t_n) = \begin{cases} e^{-\mu(t_n)} \lambda(t_1)\lambda(t_2)\ldots\lambda(t_n), & \text{if } 0 < t_1 < t_2 < \ldots < t_n < \infty \\ 0, & \text{otherwise.} \end{cases}$$

(b) The probability density function of $(A_1, A_2, \ldots, A_n)$ is given by

$$f_{A_1,A_2,\ldots,A_n}(a_1, a_2, \ldots, a_n) = \begin{cases} e^{-\mu(a_1+a_2+\ldots+a_n)} \left[\prod_{i=1}^n \lambda(a_1 + \ldots + a_i)\right], & \text{if } a_j > 0, 1 \leq j \leq n \\ 0, & \text{otherwise.} \end{cases}$$

Thus the interarrival times are i.i.d.'s only when $N(\cdot)$ is time homogeneous.

**2.2.9.** Show that $T_{N(t)}$ in (2.23) is a random variable.

**2.2.10.** Let $\{A_i\}$ be the sequence of interarrival times of a homogeneous Poisson process with rate $\lambda$. Fix $n \geq 1$. Find the distribution of $\min\{A_1, A_2, \ldots, A_n\}$.

**Exercises on Section 2.3**

**2.3.1.** Let $N(\cdot)$, $\lambda(\cdot)$, $\mu(\cdot)$, $\{T_i\}$ be as in Exercise 2.2.8. For any $t > 0$, $n = 1, 2, \ldots$ show that

$$((T_1, T_2, \ldots, T_n) | N(t) = n) \stackrel{d}{=} (Y_{(1)}, Y_{(2)}, \ldots, Y_{(n)})$$

where $0 < Y_{(1)} < Y_{(2)} < \ldots < Y_{(n)}$ is the order statistics corresponding to an i.i.d. sample $Y_1, Y_2, \ldots, Y_n$ with common probability density function $g(x) = \frac{\lambda(x)}{\mu(t)} I_{[0,t]}(x)$.

(Hint: Use part (a) of Exercise 2.2.8 and imitate the proof of Theorem 2.9.)

**2.3.2.** Let $N(\cdot), \{T_i\}, \{U_j\}$ be as in Theorem 2.9 and Remark 2.10. Let $n = 1, 2, \ldots$ and $t > 0$ be fixed. For any measurable symmetric function $g$ on $\mathbb{R}^n$ show that

$$(g(T_1, \ldots, T_n) \mid N(t) = n) \stackrel{d}{=} g(tU_1, tU_2, \ldots, tU_n).$$

**2.3.3.** Suppose that customers arrive at a train station at arrival times $0 < T_1 < T_2 < \ldots < T_n < \ldots$ of a homogeneous Poisson process $N(\cdot)$ with rate $\lambda > 0$. Suppose the train departs at time $t$. Note that $\sum_{i=1}^{N(t)} (t - T_i)$ denotes the sum of waiting times of customers arriving in $(0, t)$, provided $N(t) \geq 1$; (if $N(t) = 0$ then the quantity is taken as 0). Find $E\left[\sum_{i=1}^{N(t)} (t - T_i)\right]$.

**Exercises on Section 2.4**

**2.4.1.** (i) In the Cramér-Lundberg model assume that the claim size has finite expectation. Find $E(S(t))$.

(ii) Assuming that the claim size has finite second moment show that $\mathrm{Var}(S(t)) = \lambda t \, E(X_1^2)$.

**2.4.2.** If $F_1, F_2$ are distribution functions on $\mathbb{R}$, recall that the *convolution* of $F_1$ and $F_2$ is given by

$$F_1 * F_2(x) = \int_{\mathbb{R}} F_1(x-y)\, dF_2(y) = F_2 * F_1(x), \quad x \in \mathbb{R}.$$

Let $F$ denote the common distribution function of claim size $X_i$. Let $F^{*(0)}(x) = I_{[0,\infty)}(x)$, $x \in \mathbb{R}$ and $F^{*(1)} = F$. For $n = 2, 3, \ldots$ define $F^{*(n)}(x) = (F^{*(n-1)} * F)(x) = (F * \ldots * F)(x)$, $x \in \mathbb{R}$, the $n$-fold convolution of $F$. Let $S(\cdot)$ be given by (2.30). Show that for $x \in \mathbb{R}$

$$P(S(t) \leq x) = \sum_{n=0}^{\infty} \frac{(\lambda t)^n e^{-\lambda t}}{n!} F^{*(n)}(x).$$

**2.4.3.** Assume that the claim size $X_i$ has moment generating function $\phi_X$ in a nieghbourhood of 0. Show that the moment generating function of $S(t)$ is given by

$$\phi_{S(t)}(u) \triangleq E(e^{uS(t)}) = \exp\{\lambda t[\phi_X(u) - 1]\}$$

for $u$ in an appropriate neighbourhood of 0.

**2.4.4.** Let $S_1(\cdot)$ be defined by (2.33). Find $E(S_1(t))$.

**2.4.5.** Let $\{N(t)\}$, $\{T_i\}$, $\{X_j\}$ satisfy the assumptions (1) - (3) of Section 2.4. Show that there is a sequence $\{U_k\}$ of i.i.d. $U(0,1)$ random variables, which is independent of $\{N(t)\}$, $\{X_j\}$, such that

$$E\left[\sum_{i=1}^{N(t)} g(T_i, X_i)\right] = E\left[\sum_{i=1}^{N(t)} g(tU_i, X_i)\right]$$

holds for any bounded measurable function $g : \mathbb{R}^2 \to \mathbb{R}$.

(Hint: Enough to consider $g(t,x) = I_A(t)\, I_B(x)$ where $A$, $B$ are Borel sets in $\mathbb{R}$.)

# Chapter 3

# Renewal model

## 3.1 Preliminaries

Renewal process is a generalization of the homogeneous Poisson process obtained by dropping the assumption of exponential interarrival times. The corresponding counting process will serve as a model for claim number process. Our approach in this chapter relies heavily on [Fe], [RSST]. All the random variables are defined on a probability space $(\Omega, \mathcal{F}, P)$.

**Definition 3.1** Let $\{A_i : i = 1, 2, \ldots\}$ be a sequence of i.i.d. nonnegative random variables. Set $T_0 \equiv 0$, $T_n = A_1 + A_2 + \ldots + A_n$, $n \geq 1$. The sequence $\{T_n\}$ is said to be a *renewal sequence*. The associated counting process

$$\begin{aligned} N(t) &= \max\{i \geq 0 : T_i \leq t\} \\ &= \#\{i \geq 1 : T_i \leq t\}, \quad t \geq 0 \end{aligned} \tag{3.1}$$

is called a *renewal process*. The sequences $\{T_n\}$ and $\{A_n\}$ are referred to respectively as *arrival times* and *interarrival times of the renewal process* $N(\cdot)$. $T_n$ is also called $n^{th}$ *renewal epoch*.

To avoid trivialities we shall assume that

$$P(A_i = 0) < 1, \quad i \geq 1. \tag{3.2}$$

It is clear that existence of i.i.d. $\{A_i\}$ implies that of $N$; moreover $N(\cdot)$ has r.c.l.l. sample paths. Also note that $\{T_n : n = 0, 1, 2, \ldots\}$ is *random walk*.

We have already seen in Theorem 2.5 that if $\{A_i\}$ are i.i.d. $\text{Exp}(\lambda)$ random variables then $N(\cdot)$ is a homogeneous Poisson process with rate $\lambda > 0$.

If there are large gaps between arrivals then the Poisson process model for claim number process is not adequate. This is the case if interarrival times have thicker tails than the exponential distribution. Of course, many nice properties of the Poisson process are lost in the renewal model. However, many asymptotic properties are similar to the earlier model.

In this section we establish the strong law of large numbers for $N(\cdot)$ and the elementary renewal theorem.

To begin with, we want to establish that $N(t)$ has finite expectation. For this, and other purposes, it is convenient to treat $T_0 = 0$ also as one renewal. Indeed define for $t \geq 0$

$$N_0(t) \triangleq \sum_{i=0}^{\infty} I_{[0,t]}(T_i) = N(t) + 1. \tag{3.3}$$

We need the following lemma

**Lemma 3.2** *Let $\{\hat{A}_i\}$, $\{\tilde{A}_i\}$ be i.i.d. nonnegative random variables, and $\{\hat{T}_i\}$, $\{\tilde{T}_i\}$ be the respective renewal sequences. Let the corresponding counting processes $\hat{N}_0(\cdot)$, $\tilde{N}_0(\cdot)$ be defined by the analogues of (3.3). Suppose $\hat{A}_i \leq \tilde{A}_i$ a.s. for each $i$. Then $\hat{N}_0(t) \geq \tilde{N}_0(t)$ a.s. for $t \geq 0$. Hence $E(\tilde{N}_0(t)) \leq E(\hat{N}_0(t))$.*

**Proof:** Easy. □

**Theorem 3.3** *Let the interarrival time $\{A_i\}$ satisfy (3.2). Then $E(N(t)) < \infty$, $t \geq 0$.*

**Proof:** By assumption there exists $a > 0$ such that $P(A_i > a) > 0$. By rescaling, if necessary, assume $a = 1$. Define $\hat{A}_i = I_{(1,\infty)}(A_i)$, $i = 1, 2, \ldots$. Clearly $\{\hat{A}_i\}$ is a sequence of i.i.d. Bernoulli random variables with parameter $p = P(A_i > 1)$. Also $\hat{A}_i \leq A_i$ for all $i$. So by Lemma 3.2 it is enough to prove that $E(\hat{N}_0(t)) < \infty$.

Denoting by $\{\hat{T}_i\}$ the corresponding renewal sequence, for $k \geq 1$ note that $\hat{T}_k$ has a binomial distribution with parameters $k$ and $p$. By monotone convergence theorem, denoting $[t]$ = integral part of $t$

$$
\begin{aligned}
E(\hat{N}_0(t)) &= E\left(\sum_{k=0}^{\infty} I_{[0,t]}(\hat{T}_k)\right) = \sum_{k=0}^{\infty} E(I_{[0,t]}(\hat{T}_k)) \\
&= 1 + \sum_{k=1}^{\infty} P(\hat{T}_k \le [t]) \ (\because \hat{T}_0 \equiv 0, \hat{T}_k \text{ integer valued}) \\
&= 1 + \sum_{k=1}^{\infty} (1-p)^k + \sum_{k=1}^{\infty} \sum_{j=1}^{[t]} P(\hat{T}_k = j) \\
&= \frac{1}{p} + \sum_{j=1}^{[t]} \sum_{k=1}^{\infty} P(\hat{T}_k = j) \\
&= \frac{1}{p} + \sum_{j=1}^{[t]} \sum_{k=j}^{\infty} \binom{k}{j} p^j (1-p)^{k-j} \\
&\qquad \text{(because } P(\hat{T}_k = j) = 0 \text{ if } j < k\text{)} \\
&= \frac{1}{p} + \sum_{j=1}^{[t]} \frac{p^j}{j!} \sum_{\ell=0}^{\infty} (\ell+1)\dots(\ell+j)(1-p)^{\ell} \\
&\le \frac{1}{p} + \sum_{j=1}^{[t]} \frac{p^j}{j!} 2^j j^j \left[\sum_{\ell=0}^{\infty} \ell^j (1-p)^{\ell}\right] < \infty
\end{aligned}
$$

by ratio test, as $j$ varies only over a finite set. □

Here is the first asymptotic result.

**Theorem 3.4** *(SLLN for renewal process). In addition to the hypothesis of Theorem 3.3, let* $E(A_1) = \frac{1}{\lambda} < \infty$. *Then*

$$\lim_{t\to\infty} \frac{1}{t} N(t) = \lambda, \ a.s. \tag{3.4}$$

We need a lemma.

**Lemma 3.5** *Let* $\{Z_n : n = 1, 2, \dots\}, Z, \{M(t) : t \ge 0\}$ *be random variables on the same probability space such that (i)* $Z_n \to Z$ *a.s., (ii)* $M(t)$ *integer valued, (iii)* $M(t) \to \infty$ *a.s. Then* $Z_{M(t)} \to Z$ *a.s.*

**Proof:** Put $\Omega_1 = \{\omega \in \Omega : M(t,\omega) \to \infty\}$, $\Omega_2 = \{\omega \in \Omega : Z_n(\omega) \to Z(\omega)\}$. Note that $P(\Omega_1) = P(\Omega_2) = 1$ and hence $P(\Omega_1 \cap \Omega_2) = 1$. Clearly $Z_{M(t,\omega)}(\omega) \to Z(\omega)$ for any $\omega \in \Omega_1 \cap \Omega_2$. □

**Proof of Theorem:** As the sample paths of $N$ are nondecreasing and $P(A_i = 0) < 1$, note that $\lim_{t\to\infty} N(t)$ exists and is strictly positive with probability one. Now $E(A_i) < \infty$ for all $i$ implies $P(T_n < \infty) = 1$ for every $n$. It then follows that $P\left(\lim_{t\to\infty} N(t) = \infty\right) = 1$. Observe that

$$\{N(t) = n\} = \{T_n \leq t < T_{n+1}\}, \quad n = 0, 1, 2, \ldots \tag{3.5}$$

and hence $T_{N(t)} \leq t < T_{N(t)+1}$, $t \geq 0$. Therefore

$$\frac{T_{N(t)}}{N(t)} \leq \frac{t}{N(t)} \leq \frac{T_{N(t)+1}}{N(t)+1} \cdot \frac{N(t)+1}{N(t)} \tag{3.6}$$

for all sufficiently large $t$, with probability one. By Lemma 3.5 note that $P\left(\lim_{t\to\infty} \frac{T_{N(t)}}{N(t)} = \lambda^{-1}\right) = 1$. The required conclusion (3.4) is now clear from (3.6). □

Theorem 3.4 suggests that $E(N(t))$ will be asymptotically linear. Indeed we have

**Theorem 3.6** *(Elementary renewal theorem): Let the hypotheses be as in Theorem 3.4. Then*

$$\lim_{t\to\infty} \frac{1}{t} E(N(t)) = \lambda. \tag{3.7}$$

**Proof:** By Theorem 3.3 we know $E(N(t)) < \infty$, $t \geq 0$. As $\frac{1}{t} N(t)$ is nonnegative, by Theorem 3.4 and Fatou's lemma

$$\lambda = E\left[\liminf_{t\to\infty} \frac{N(t)}{t}\right] \leq \liminf_{t\to\infty} \frac{1}{t} E(N(t)). \tag{3.8}$$

So it remains to establish

$$\limsup_{t\to\infty} \frac{1}{t} E(N(t)) \leq \lambda. \tag{3.9}$$

For this we consider truncated interarrival times. For $b > 0$ set

$$\left.\begin{array}{lcl} A_i^{(b)} & = & \min\{A_i, b\},\ i = 1, 2, \ldots \\ T_0^{(b)} & \equiv & 0,\ T_n^{(b)} = A_1^{(b)} + \ldots + A_n^{(b)},\ n = 1, 2, \ldots \\ N^{(b)}(t) & = & \max\{i \geq 0 : T_i^{(b)} \leq t\},\ t \geq 0. \end{array}\right\} \tag{3.10}$$

Clearly $\{A_i^{(b)}\}$ is an i.i.d. sequence of nonnegative random variables, $\{T_n^{(b)}\}$ is the corresponding renewal sequence, and $\{N^{(b)}(t)\}$ is the associated renewal process. Observe that $T_n \geq T_n^{(b)}$, $n \geq 0$ and $N(t) \leq N^{(b)}(t)$, $t \geq 0$.

Assuming *Wald's identity*

$$E\left[T^{(b)}_{N^{(b)}(t)+1}\right] = E[N^{(b)}(t)+1] \cdot E[A_1^{(b)}], \quad t > 0 \tag{3.11}$$

we shall proceed with the proof. (3.11) will be proved later. By (3.11) we get

$$\begin{aligned} & \frac{1}{t}\, E(N(t)) \leq \frac{1}{t}\, E(N^{(b)}(t)) = \frac{1}{t}\, E[N^{(b)}(t)+1] - \frac{1}{t} \\ = & \frac{1}{t\, E(A_1^{(b)})}\, E\,[T^{(b)}_{N^{(b)}(t)+1}] - \frac{1}{t} \\ = & -\frac{1}{t\, E(A_1^{(b)})}\, E\,[T^{(b)}_{N^{(b)}(t)}] + \frac{1}{t\, E[A_1^{(b)}]}\, E\,[A^{(b)}_{N^{(b)}(t)+1}] - \frac{1}{t}. \end{aligned}$$

Therefore, since $T^{(b)}_{N^{(b)}(t)} \leq t$, $A^{(b)}_{N^{(b)}(t)+1} \leq b$, we have

$$\limsup_{t\to\infty} \frac{1}{t}\, E(N(t)) \leq \frac{1}{E[A_1^{(b)}]}, \text{ for any } b > 0.$$

Letting $b \uparrow \infty$ in the above we get (3.9). □

It remains to prove Wald's identity (3.11). A proof, which exploits the fact that $\{T_n\}$ is a random walk, is indicated in Exercises 3.1.3, 3.1.4. We give below another approach using the versatile martingale framework, especially the optional sampling theorem. This method will also be useful elsewhere.

Set $\hat{A}_i^{(b)} = A_i^{(b)} - E(A_i^{(b)}) = A_i^{(b)} - E(A_1^{(b)})$, $i \geq 1$, $\hat{T}_0^{(b)} \equiv 0$, $\hat{T}_n^{(b)} = \hat{A}_1^{(b)} + \ldots + \hat{A}_n^{(b)} = T_n^{(b)} - nE(A_1^{(b)})$, $n \geq 1$. Let $\mathcal{F}_0^{(b)} = \{\phi, \Omega\}$, $\mathcal{F}_n^{(b)} = \sigma\{A_i^{(b)} : i \leq n\}$, $n \geq 1$. It is easily verified that $\{\hat{T}_n^{(b)} : n = 0, 1, 2, \ldots\}$ is a *martingale* w.r.t. $\{\mathcal{F}_n^{(b)}\}$. (See Exercise 3.1.5 or Appendix A.3 for a definition.) It is also easily seen that $(N^{(b)}(t)+1)$ is a stopping time w.r.t. $\{\mathcal{F}_n^{(b)}\}$; see Exercise 3.1.3.

Put $\tau = N^{(b)}(t) + 1$, remembering that $b > 0$, $t > 0$ are fixed. We want to apply the optional sampling theorem to the stopping times $\tau$ and 0, for the martingale $\{\hat{T}_n^{(b)}\}$. See [V1, V2], [BW] for more information

on martingales, stopping times, optional sampling theorem, etc. A brief account is given in Appendix A.3. For this we need to verify

$$E\ |\hat{T}_\tau^{(b)}| < \infty \tag{3.12}$$

and

$$\lim_{k\to\infty} E[|\hat{T}_k^{(b)}| : \{\tau > k\}] = 0. \tag{3.13}$$

(In (3.13) and elsewhere for a random variable $Y$ and a measurable set $B$, $E[Y : B] \triangleq \int_B Y\ dP = E[I_B \cdot Y]$). Note that $P(\tau < \infty) = 1$ and so $\hat{T}_\tau^{(b)}$ is well defined.

By Theorem 3.3, $E(\tau) < \infty$ and hence

$$\begin{aligned} E|\hat{T}_\tau^{(b)}| &= E[|T_\tau^{(b)} - \tau E(A_1^{(b)})|] \le E|T_\tau^{(b)}| + E(A_1^{(b)})E(\tau) \\ &= E(T^{(b)}_{N^{(b)}(t)+1}) + E(A_1^{(b)}) \cdot E(\tau) \\ &= E(T^{(b)}_{N^{(b)}(t)} + A^{(b)}_{N^{(b)}(t)+1}) + E(A_1^{(b)})E(\tau) \\ &\le t + b + E(A_1^{(b)}) \cdot E(\tau) < \infty, \end{aligned}$$

proving (3.12). Next note that $\{\tau > k\} = \{N^{(b)}(t) > k-1\} = \{N^{(b)}(t) \ge k\} = \{T_k^{(b)} \le t\}$. Therefore

$$\begin{aligned} & E\ [|\hat{T}_k^{(b)}| : \{\tau > k\}] = E\ [|T_k^{(b)} - kE(A_1^{(b)})| : \{\tau > k\}] \\ \le\ & E\ [T_k^{(b)} : \{\tau > k\}] + E\ (A_1^{(b)})\ E[k : \{\tau > k\}] \\ \le\ & E\ [T_k^{(b)} : \{T_k^{(b)} \le t\}] + E\ (A_1^{(b)})\ E[\tau : \{\tau > k\}] \\ \le\ & t \cdot P(T_k^{(b)} \le t) + E(A_1^{(b)})\ E[\tau : \{\tau > k\}]. \end{aligned}$$

As $t$ is fixed, by the law of large numbers, applied to $\{A_i^{(b)}\}$, note that $P(T_k^{(b)} \le t) \to 0$ as $k \to \infty$. Since $E(\tau) < \infty$, the second term on r.h.s. also goes to 0. Thus (3.13) follows.

So by the optional sampling theorem $E[\hat{T}_\tau^{(b)}] = E[\hat{T}_0^{(b)}] = 0$, whence (3.11) follows. This completes the proof of the elementary renewal theorem. □

Unlike the Poisson process, it is in general not possible to determine the exact distribution of the claim number process $N(\cdot)$ in the renewal model. We give below a *central limit theorem* which may provide an approximation for large $t$. We follow [RSST].

**Theorem 3.7** *Assume that* $0 < Var(A_1) < \infty$. *Denote* $E(A_1) = \frac{1}{\lambda}$, $c = \lambda^3 Var(A_1)$. *Then for every* $x \in \mathbb{R}$

$$\lim_{t\to\infty} P\left(\frac{N(t) - \lambda t}{\sqrt{ct}} \le x\right) = \Phi(x) \tag{3.14}$$

*where* $\Phi(\cdot)$ *is the distribution function of the standard normal distribution. Moreover the convergence in (3.14) is uniform over* $x \in \mathbb{R}$.

**Proof:** By Lindeberg-Levy central limit theorem

$$\lim_{n\to\infty} \sup_{x\in\mathbb{R}} \left| P\left(\frac{T_n - n\lambda^{-1}}{\sqrt{n\mathrm{Var}(A_1)}} \le x\right) - \Phi(x)\right| = 0. \tag{3.15}$$

In (3.15) we have also used the fact: If $G_n$, $n \ge 1$, $G$ are distribution functions on $\mathbb{R}$ such that $\{G_n\}$ converges weakly to $G$, and $G$ is continuous, then $G_n(x) \to G(x)$ uniformly over $x \in \mathbb{R}$. See [Bl2]. Also this fact will give us uniform convergence in (3.14) once pointwise convergence is established.

Clearly $\{N(t) < \ell\} = \{T_\ell > t\}$. Denoting $m(t) = \lfloor x\sqrt{ct} + \lambda t \rfloor \triangleq$ largest integer not exceeding $(x\sqrt{ct} + \lambda t)$, we get

$$\begin{aligned} & P\left(\frac{N(t) - \lambda t}{\sqrt{ct}} \le x\right) = P(N(t) \le x\sqrt{ct} + \lambda t) \\ = \; & P(N(t) < m(t) + 1) = P(T_{m(t)+1} > t) \\ = \; & P\left(\frac{T_{m(t)+1} - \lambda^{-1}(m(t)+1)}{\sqrt{(m(t)+1)\mathrm{Var}(A_1)}} > \frac{t - \lambda^{-1}(m(t)+1)}{\sqrt{(m(t)+1)\mathrm{Var}(A_1)}}\right). \end{aligned}$$

Note that $m(t) \to \infty$ as $t \to \infty$. It is easily seen that

$$\begin{aligned} \lim_{t\to\infty} \frac{t - (m(t)+1)\lambda^{-1}}{\sqrt{(m(t)+1)\mathrm{Var}(A_1)}} &= \lim_{t\to\infty} \frac{t - m(t)\lambda^{-1}}{\sqrt{m(t)\mathrm{Var}(A_1)}} \\ = \lim_{t\to\infty} \frac{t - (x\sqrt{ct} + \lambda t)\lambda^{-1}}{[(x\sqrt{ct} + \lambda t)\mathrm{Var}(A_1)]^{1/2}} &= -x \end{aligned} \tag{3.16}$$

holds for each $x \in \mathbb{R}$. Since the convergence in (3.15) is uniform, and $1 - \Phi(-x) = \Phi(x)$, note that (3.16) now implies (3.14) pointwise. □

## 3.2 Renewal function

As already seen in Section 3.1, $E(N(t))$ plays an important role. Since it may be the only quantity amenable to analysis in a fairly general set up, it is important to understand it well.

For any right continuous nondecreasing function $G$ on $\mathbb{R}$ with $\lim_{x\to-\infty} G(x) = 0$, the associated measure is denoted by $\mu_G$. We know that $\mu_G((a,b]) = G(b) - G(a)$, $a, b \in \mathbb{R}$, $a \leq b$. By slight abuse of terminology, $G$ will be referred to as the "distribution function" of $\mu_G$.

**Definition 3.8** The function

$$H(t) \triangleq E(N_0(t)) = E(N(t)+1), \quad t \geq 0 \tag{3.17}$$

is called the *renewal function* corresponding to the renewal process $N(\cdot)$.

We have already seen in Exercise 3.1.3 that for fixed $t$, $N_0(t)$ is a stopping time w.r.t. the natural filtration, while $N(t)$ is not.

To appreciate the role of the renewal function we first show that it contains all the information about the interarrival distribution.

**Definition 3.9** Let $G$ be a right continuous nondecreasing function on $\mathbb{R}$ such that $G(0-) = 0$, $0 \leq G(x) < \infty$, $x \geq 0$. (It is possible that $G(+\infty) \triangleq \lim_{x\to\infty} G(x) = +\infty$.) Note that $\mu_G$ is a $\sigma$-finite measure supported on $[0,\infty)$. Define

$$\hat{\ell}_G(s) = \int_0^\infty e^{-sx} dG(x) = \int_{[0,\infty)} e^{-sx} d\mu_G(x), \; s \geq 0 \tag{3.18}$$

whenever the integral exists. The function $\hat{\ell}_G$ is called the *Laplace-Stieltjes transform of* $G$.

If $\mu_G$ is a finite measure then $\hat{\ell}_G(s)$ exists for all $s \geq 0$.

**Theorem 3.10** *(a) Let $F$ be the distribution function of a probability measure supported on $[0,\infty)$. Then $\hat{\ell}_F$ has derivatives $\hat{\ell}_F^{(n)}$ of all orders on $(0,\infty)$ and*

$$(-1)^n \hat{\ell}_F^{(n)}(s) \geq 0, \quad s > 0. \tag{3.19}$$

*Moreover at any point of continuity of $F$*

$$F(x) = \lim_{s\to\infty} \sum_{n\le sx} \frac{(-s)^n}{n!}\, \hat{\ell}_F^{(n)}(s). \tag{3.20}$$

*(b) If $F_1$, $F_2$ are distribution functions of two probability measures supported on $[0,\infty)$ and $\hat{\ell}_{F_1} = \hat{\ell}_{F_2}$, then $F_1 = F_2$.*

We need a lemma whose proof may be amusing.

**Lemma 3.11** *Let $x > 0$, $z > 0$. Then*

$$\lim_{s\to\infty} e^{-sx} \sum_{k\le sz} \frac{(sx)^k}{k!} = \begin{cases} 0, & \text{if } x > z \\ 1/2, & \text{if } x = z \\ 1, & \text{if } x < z. \end{cases} \tag{3.21}$$

**Proof:** The l.h.s. of (3.21) suggests bringing in a homogeneous Poisson process $Y$ with parameter $x > 0$. Note that

$$\begin{aligned} \text{l.h.s of (3.21)} &= \lim_{s\to\infty} P(Y(s) \le sz) \\ &= \lim_{s\to\infty} P\left(\frac{Y(s) - sx}{\sqrt{sx}} \le \frac{sz - sx}{\sqrt{sx}}\right). \end{aligned}$$

Let $\Phi$ denote the distribution function of $N(0,1)$ distribution. In Exercise 2.1.5 (CLT for Poisson process), observe that the convergence is uniform over $\mathbb{R}$ as $\Phi$ is continuous. So the above limit

$$= \lim_{s\to\infty} \Phi\left(\frac{(z-x)}{\sqrt{x}}\sqrt{s}\right) = \text{r.h.s of (3.21)}.$$

□

**Proof of theorem:** As $F$ is the distribution function of a probability measure, by repeated application of the dominated convergence theorem

$$(-1)^n\, \hat{\ell}_F^{(n)}(s) = \int_0^\infty e^{-sx}\, x^n\, dF(x) \ge 0, \quad s > 0. \tag{3.22}$$

By (3.22), Lemma 3.11 and again the dominated convergence theorem we get

$$\sum_{n\le sz}(-1)^n\,\frac{s^n}{n!}\,\hat{\ell}_F^{(n)}(s) = \int_0^\infty \sum_{n\le sz} e^{-sx}\,\frac{(sx)^n}{n!}\,dF(x)$$

$$\to \int_0^\infty \left\{ I_{[0,z)}(x) + \frac{1}{2} I_{\{z\}}(x) \right\} dF(x)$$

$$= F(z-) + \frac{1}{2}(F(z) - F(z-)), \text{ as } s \to \infty. \tag{3.23}$$

If $z$ is a point of continuity of $F$, the conclusion (3.20) is obvious. Clearly part (b) is a corollary of (3.20). □

Let $F$ denote the common distribution function of the interarrival times; let $\mu_F$ denote the corresponding probability measure supported on $[0,\infty)$. By (3.2) note that $\mu_F \neq \delta_0$. Denote $\mu_F^{*(0)} = \delta_0$, $\mu_F^{*(1)} = \mu_F$, $\mu_F^{*(n)} = \mu_F^{*(n-1)} * \mu_F = \mu_F * \mu_F^{*(n-1)}$, $n \ge 1$. Similarly $F^{*(0)} = I_{[0,\infty)}$, $F^{*(1)} = F$, $F^{*(n)} = F^{*(n-1)} * F$, $n \ge 1$. (See Exercise 2.4.2 for definition of convolution.) Clearly $\mu_F^{*(n)}$ is a probability measure supported on $[0,\infty)$ and $F^{*(n)}$ is the corresponding distribution function.

**Theorem 3.12** *Let the hypothesis be as in Theorem 3.3. Then the renewal function $H$ is a right continuous nondecreasing function on $\mathbb{R}$ such that $H(0-) = 0$, $0 \le H(t) < \infty$, $t \ge 0$, and hence $\mu_H$ is a $\sigma$-finite measure supported on $[0,\infty)$. Moreover*

$$H = \sum_{n=0}^{\infty} F^{*(n)}, \text{ that is }, \ \mu_H = \sum_{n=0}^{\infty} \mu_F^{*(n)}. \tag{3.24}$$

**Proof:** By definition of $N(\cdot)$ and Theorem 3.3 the first assertion is clear. As $T_k \sim \mu_F^{*(k)}$ for all $k \ge 0$, by (3.17), (3.3) and the monotone convergence theorem

$$H(t) = \sum_{n=0}^{\infty} P(T_n \le t) = \sum_{n=0}^{\infty} F^{*(n)}(t),$$

completing the proof. □

**Theorem 3.13** *Let the hypothesis be as in Theorem 3.3. Then the Laplace-Stieltjes transform $\hat{\ell}_H$ of $H$ is well defined on $[0,\infty)$. Moreover $\hat{\ell}_H(s) > 1$, $s > 0$ and*

$$\hat{\ell}_H(s) = \frac{1}{1-\hat{\ell}_F(s)}, \quad s > 0. \tag{3.25}$$

*Consequently*

$$\hat{\ell}_F(s) = 1 - \frac{1}{\hat{\ell}_H(s)}, \quad s > 0. \tag{3.26}$$

**Proof:** Note that $\hat{\ell}_F(s) = E(e^{-sA_1})$. As $\{A_i\}$ is an i.i.d. sequence, clearly $\hat{\ell}_{F^{*(n)}}(s) = E[e^{-sT_n}] = [\hat{\ell}_F(s)]^n$ for all $n$. Hence for $s > 0$, by Theorem 3.12

$$\hat{\ell}_H(s) = \sum_{n=0}^{\infty} \int_0^{\infty} e^{-sx} dF^{*(n)}(x) = \frac{1}{1-\hat{\ell}_F(s)}$$

where the geometric series converges as $\hat{\ell}_F(s) < 1$, $s > 0$, completing the proof. □

From (3.26) and the inversion formula (3.20) note that $F$ can be obtained from the renewal function $H$ in principle. The measure $\mu_H$ is also called the *renewal measure* of the process $N$.

Now observe that $F^{*(n)}(t) = \int_0^t F^{*(n-1)}(t-s)\, dF(s)$, $t \geq 0$ as $F$, $F^{*(n-1)}$ are supported on $[0,\infty)$. Therefore by (3.24) and monotone convergence theorem, for $t \geq 0$

$$\begin{aligned} H(t) &= 1 + \sum_{n=1}^{\infty} F^{*(n)}(t) = I_{[0,\infty)}(t) + \sum_{n=1}^{\infty} F^{*(n)}(t) \ (\because t \geq 0) \\ &= I_{[0,\infty)}(t) + \sum_{n=1}^{\infty} \int_0^t F^{*(n-1)}(t-s)\, dF(s) \\ &= I_{[0,\infty)}(t) + \int_0^t \sum_{k=0}^{\infty} F^{*(k)}(t-s)\, dF(s) \\ &= I_{[0,\infty)}(t) + \int_0^t H(t-s)\, dF(s). \end{aligned}$$

Thus we obtain

**Theorem 3.14** *With the hypothesis as in Theorem 3.3, the renewal function satisfies the renewal equation*

$$H(t) = I_{[0,\infty)}(t) + \int_0^t H(t-s)\, dF(s), \quad t \geq 0. \tag{3.27}$$

**Definition 3.15** A convolution equation of the form

$$U(t) = u(t) + \int_0^t U(t-y)\, dF(y), \quad t \geq 0 \tag{3.28}$$

is called a *renewal equation*. Here $F$ is the distribution function of a probability measure supported on $[0,\infty)$, $u(\cdot)$ is a known function and $U(\cdot)$ is the unknown function both defined on $[0,\infty)$.

Theorems 3.12 and 3.14 suggest the following result regarding solution to (3.28). We say a measurable function $u$ is *locally bounded* if $u$ is bounded on every bounded set.

**Theorem 3.16** *Let $F$ be the distribution function of a probability measure $\mu_F \neq \delta_0$ supported on $[0,\infty)$. Let $H = \sum_{n=0}^{\infty} F^{*(n)}$, $\mu_H = \sum_{n=0}^{\infty} \mu_F^{*(n)}$ denote respectively the corresponding "renewal function", "renewal measure". Consider the renewal equation (3.28) where $u(\cdot)$ is a locally bounded measurable function supported on $[0,\infty)$. Then*

$$U(t) = \int_0^t u(t-s)\, dH(s) = \int_{[0,t]} u(t-s)\, d\mu_H(s) \tag{3.29}$$

*is the unique solution to (3.28) in the class of locally bounded measurable functions supported on $[0,\infty)$.*

**Proof:** As $\mu_F \neq \delta_0$, by Theorems 3.3 and 3.12, $H$ and $\mu_H$ are well defined. As $\mu_H([0,s]) < \infty$ for any $0 < s < \infty$, for any locally bounded measurable function $\alpha$ supported on $[0,\infty)$ note that

$t \mapsto \int_0^t \alpha(t-s)\, d\mu_H(s)$ is a well defined locally bounded measurable function supported on $[0,\infty)$. Moreover

$$\int_0^t \alpha(t-s)\, d\mu_H(s) = \sum_{n=0}^{\infty} \int_0^t \alpha(t-s)\, d\mu_F^{*(n)}(s), \quad t \geq 0.$$

Thus $U(\cdot)$ given by (3.29) is a well defined locally bounded function supported on $[0,\infty)$. Also by Exercise 3.2.1 (see below), note that

$$\begin{aligned}
U(t) &= \int_0^t u(t-s)\, d\mu_H(s) = u(t) + \sum_{k=1}^{\infty} \int_0^t u(t-s)\, d\mu_F^{*(k)}(s) \\
&= u(t) + \sum_{j=0}^{\infty} \int_0^t u(t-s)\, d(\mu_F^{*(j)} * \mu_F)(s) \\
&= u(t) + \sum_{j=0}^{\infty} \int_0^t \left[ \int_0^{\infty} u(t-(s+r))\, d\mu_F^{*(j)}(s) \right] d\mu_F(r) \\
&= u(t) + \sum_{j=0}^{\infty} \int_0^t \int_0^{t-r} u(t-r-s)\, d\mu_F^{*(j)}(s)\, d\mu_F(r) \\
&= u(t) + \int_0^t \left[ \sum_{j=0}^{\infty} \int_0^{t-r} u(t-r-s)\, d\mu_F^{*(j)}(s) \right] d\mu_F(r) \\
&= u(t) + \int_0^t U(t-r)\, d\mu_F(r)
\end{aligned}$$

proving that $U$ satisfies (3.28).

To prove uniqueness, let $U_1$, $U_2$ be two locally bounded solutions supported on $[0,\infty)$. Using the renewal equation (3.28), iterating and by Exercise 3.2.1 we get for $k = 1, 2, \ldots$

$$\begin{aligned}
U_1(t) - U_2(t) &= \int_0^t [U_1(t-s) - U_2(t-s)]\, d\mu_F(s) \\
&= \int_0^t [U_1(t-s) - U_2(t-s)]\, d\mu_F^{*(k)}(s).
\end{aligned}$$

As $U_1$, $U_2$ are locally bounded it now follows

$$|U_1(t) - U_2(t)| \leq F^{*(k)}(t) \left[ \sup_{0 \leq s \leq t} |U_1(s) - U_2(s)| \right]$$

for any $k \geq 1$. Since $H(t) < \infty$ we have $F^{*(k)}(t) \to 0$ as $k \to \infty$. So $U_1 = U_2$. □

In the preceding discussions we have taken all the $A_i$'s having the same distribution. For some purposes it is convenient to consider the following situation.

**Definition 3.17** Let $A_i$, $i = 1, 2, \ldots$ be a sequence of nonnegative independent random variables. Let $A_1$ have distribution function $G$, and each $A_i$, $i \geq 2$ have distribution function $F$. Set $T_0^{(D)} \equiv 0$, $T_n^{(D)} = A_1 + A_2 + \ldots + A_n$, $n \geq 1$. The sequence $\{T_n^{(D)}\}$ is called a *delayed renewal sequence*. The associated counting process

$$N^{(D)}(t) \triangleq \max\{i \geq 0 : T_i^{(D)} \leq t\}, \quad t \geq 0 \tag{3.30}$$

is called a *delayed renewal process*. Also the function

$$H^{(D)}(t) \triangleq E(N^{(D)}(t)), \quad t \geq 0 \tag{3.31}$$

is called the *delayed renewal function*. Note that

$$N^{(D)}(t) = \sum_{i=1}^{\infty} I_{[0,t]}(T_i^{(D)}).$$

Then imitating the proofs of Theorems 3.12, 3.13 one can get the following

**Theorem 3.18** *Let the hypotheses of Theorem 3.3 and Definition 3.17 hold. Then*

$$H^{(D)}(t) = G(t) + \sum_{k=1}^{\infty} (G * F^{*(k)})(t), \quad t \geq 0. \tag{3.32}$$

*Moreover $H^D(\cdot)$ satisfies the renewal equation*

$$H^{(D)}(t) = G(t) + \int_0^t H^{(D)}(t-s)\, dF(s), \quad t \geq 0. \tag{3.33}$$

□

A special case of interest is when the distribution function of $A_1$ is given by

$$F_I(t) = \frac{1}{E(A_2)} \int_0^t (1 - F(y))\, dy, \quad t \geq 0. \tag{3.34}$$

$F_I$ is called the *integrated tail distribution of* $F$. Sometimes it is called the *stationary excess distribution* or *equilibrium distribution of* $F$. This plays an important role later. Let $H_I(\cdot)$ denote the corresponding renewal function.

**Proposition 3.19** *Let the hypotheses be as in Theorem 3.18, with the distribution of* $A_1$ *given by (3.34). Then* $H_I(t) = \lambda t$, $t \geq 0$ *where* $E(A_2) = \lambda^{-1}$.

**Proof:** By integration by parts

$$F_I(t) = \lambda t - \int_0^t \lambda(t - s)\, dF(s).$$

Consequently the renewal equation (3.33) for $H_I$ can be written as

$$\begin{aligned} H_I(t) &= F_I(t) + \int_0^t H_I(t - s)\, dF(s) \\ &= \lambda t + \int_0^t [H_I(t - s) - \lambda(t - s)]\, dF(s), \end{aligned}$$

whence the required assertion follows. □

## 3.3 Renewal theorems

Renewal theorems concern the asymptotic behaviour of the renewal function $H(\cdot)$, and of the solution $U(\cdot)$ to the renewal equation. By the elementary renewal theorem we know that $H(\cdot)$ is asymptotically linear. We also know that $H(\cdot)$ is the "distribution function" of a $\sigma$-finite measure supported on $[0, \infty)$, and that the expression for $U(\cdot)$ involves a Stieltjes integral with respect to $H(\cdot)$. So to study the asymptotic behaviour of $U(\cdot)$ it is helpful to know the asymptotic behaviour of $H(t + h) - H(t)$ as $t \to \infty$, for each fixed $h > 0$.

**Definition 3.20** A distribution function $G$ (or equivalently the corresponding measure $\mu_G$) is said to be a *lattice distribution* if there exists $a \geq 0$ such that $\mu_G(E^c) = 0$ where $E = \{ka : k \in \mathbb{Z}\}$. A distribution is called *nonlattice* if it is not lattice. (Note that a nonlattice distribution cannot be supported on a discrete subgroup of $\mathbb{R}$).

**Theorem 3.21** *(Blackwell's renewal theorem): Let the interarrival time distribution $F$ be supported on $[0, \infty)$, be nonlattice, and have finite expectation $\lambda^{-1}$. Let $H$ be the corresponding renewal function. Then for any $h \geq 0$*

$$\lim_{t \to \infty} H(t+h) - H(t) = \lim_{t \to \infty} E[N(t+h) - N(t)] = \lambda h. \tag{3.35}$$

□

We shall eventually take up the proof of the theorem. Before that we shall explore certain implications of the result.

**Definition 3.22** Let $G_n$, $n = 1, 2, \ldots$ and $G$ be nondecreasing right continuous functions on $\mathbb{R}$. We say that $G_n$ *converges weakly* to $G$, denoted $G_n \xrightarrow{w} G$ if $\lim_{n \to \infty} G_n(b) - G_n(a) = G(b) - G(a)$ for every $a < b$, $a, b \in \mathbb{R}$ being continuity points of $G$. This notion is also denoted $\mu_{G_n} \xrightarrow{w} \mu_G$.

In view of the above definition, Blackwell's renewal theorem can be interpreted as follows. For $t \geq 0$ define the $\sigma$-finite measure $\nu_t$ by setting $\nu_t(E) = \mu_H(E + t)$ for any Borel set $E$; (that is, $\nu_t((a, b]) = H(t+b) - H(t+a)$, $a, b \in \mathbb{R}$, $a < b$). Then under the hypotheses of the theorem $\nu_t \xrightarrow{w} \lambda \, ds$ as $t \to \infty$, where $ds$ denotes the Lebesgue measure.

Suppose $u(\cdot)$ is a continuous function on $[0, \infty)$ vanishing outside a finite interval. If $U(\cdot)$ denotes the solution to the corresponding renewal equation (3.28) then (see Exercise 3.3.1 (b))

$$\lim_{t \to \infty} U(t) = \lambda \int_0^\infty u(s) \, ds \tag{3.36}$$

using Blackwell's renewal theorem.

As another illustration, consider $g(t) = H(t) - H(t-h)$ as a function on $\mathbb{R}$, by setting $H(\cdot) = 0$ on $(-\infty, 0)$; here $h \geq 0$ is fixed. By Exercise

3.2.4 we know that $g(\cdot)$ is the solution to (3.28) with $u(t) = I_{[0,h)}(t)$. In this case (3.35) can be written as

$$\lim_{t\to\infty} g(t) = \lambda \int_0^\infty u(s)\, ds. \tag{3.37}$$

In the Poisson process model, and in the delayed renewal model given in Proposition 3.19 note that (3.35) holds exactly for all $t$.

Having guessed the asymptotic behaviour of the solution to the renewal equation, we need the following definition to carry it through for a large class of functions.

**Definition 3.23** Let $u(\cdot)$ be a function on $[0,\infty)$. For $h > 0$, $n = 1, 2, \ldots$ denote

$$\begin{aligned} \underline{m}_n &= \inf\{u(x) : (n-1)h \le x \le nh\} \\ \overline{m}_n &= \sup\{u(x) : (n-1)h \le x \le nh\}. \end{aligned}$$

The function $u(\cdot)$ is said to be *directly Riemann integrable* if
(i) the two series $\underline{u}(h) = h\sum_n \underline{m}_n$, $\overline{u}(h) = h\sum_n \overline{m}_n$ converge absolutely for any $h > 0$, and
(ii) $\lim_{h\downarrow 0} [\overline{u}(h) - \underline{u}(h)] = 0$.

The above notion is perhaps due to Wiener. The following lemma from [RSST] gives a useful sufficient condition.

**Lemma 3.24** *Let $u_1 : [0,\infty) \to (0,\infty)$ be nondecreasing, $u_2 : [0,\infty) \to [0,\infty)$ be nonincreasing such that $\int_0^\infty u_1(x)\, u_2(x)\, dx < \infty$ and $\lim_{h\to 0} c(h) = 1$, where*

$$c(h) = \sup\left\{ \frac{u_1(x+r)}{u_1(x)} : x \ge 0, \quad 0 \le r \le h \right\}.$$

*Then the product $u(x) = u_1(x)\, u_2(x)$ is directly Riemann integrable.*

**Proof:** Note that $c(h) > 0$. By definition of $c(\cdot)$

$$\begin{aligned} &\sup\{u(x) : (n-1)h \le x \le nh\} \le u_1(nh)\, u_2((n-1)h) \\ \le\ & c(2h)\, u_1((n-2)h)\, u_2((n-1)h),\ n = 2, 3, \ldots \end{aligned}$$

Note that $u_1((n-2)h)\ u_2((n-1)h) \le u_1(x)\ u_2(x)$ for $(n-2)h \le x \le (n-1)h$. Therefore

$$\begin{aligned} \overline{u}(h) &\le h \sup\{u(x) : 0 \le x \le h\} + c(2h) \int_0^\infty u_1(x)\ u_2(x)\ dx \\ &< \infty, \text{ for any } h > 0. \end{aligned}$$

Similarly $\underline{u}(h) \ge \frac{1}{c(2h)} \int_h^\infty u_1(x)\ u_2(x)\ dx$. Since $c(h) \to 1$ as $h \to 0$ the required conclusion now follows. □

**Lemma 3.25** *The following classes of functions on $[0,\infty)$ are directly Riemann integrable.*
*(i) Continuous function vanishing outside a finite interval.*
*(ii) Indicator function of a finite interval.*
*(iii) Riemann integrable function vanishing outside a finite interval.*
*(iv) $u(t) = F(t+\xi) - F(t)$, $t \ge 0$ where $F$ is the distribution function of a probability measure supported on $[0,\infty)$ and $\xi \ge 0$ is fixed.*
*(v) $u(t) = F(t) - F(t-\xi)$, $t \ge 0$ where $F$, $\xi$ are as in (iv).*
*(vi) $u(t) = 1 - F(at)$, $t \ge 0$ where $F$ is the distribution function of a probability measure supported on $[0,\infty)$ having finite expectation, and $a > 0$ is fixed.*

**Proof:** We shall indicate the proof only in case (iv). Let $a_n = \sup\{u(t) : n-1 \le t \le n\}$. Clearly $a_n \le F(n+\xi) - F(n-1) \le F(n+M) - F(n-1)$, where $M$ is the smallest integer $\ge \xi$. Therefore for all sufficiently large $k$, $\sum_{n=1}^{k} a_n \le \sum_{i=k}^{k+M} F(i) \le M+1$. Now apply Exercise 3.3.3 (b). □

**Theorem 3.26** *Let notation and the hypotheses be as in Theorem 3.21. Then the following are equivalent.*
*(a) Blackwell's renewal theorem holds, that is, (3.35) is satisfied for any $h \ge 0$.*
*(b) (Key renewal theorem) Let $u(\cdot)$ be a directly Riemann integrable function supported on $[0,\infty)$. Let $U(\cdot)$ be the locally bounded solution to the renewal equation (3.28). Then (3.36) holds, that is $\lim_{t\to\infty} U(t) = \lambda \int_0^\infty u(s)\ ds$.*

**Proof:** As any directly Riemann integrable function is locally bounded and integrable (see Exercise 3.3.4), the statement in part (b) is meaningful.

**(b) $\Rightarrow$ (a) :** Take $u(t) = I_{[0,h)}(t)$. We have already seen that (3.37) is the same as (3.35).

**(a) $\Rightarrow$ (b) :** In view of Exercise 3.3.6 (a) it is enough to consider $u(\cdot) \geq 0$. With $h > 0$ fixed but arbitrary, and notation as in Definition 3.23, set

$$\begin{aligned} u_*(t) &= \sum_{n=1}^{\infty} \underline{m}_n \, I_{[(n-1)h,nh)}(t), \\ u^*(t) &= \sum_{n=1}^{\infty} \overline{m}_n \, I_{[(n-1)h,nh)}(t). \end{aligned}$$

Clearly $0 \leq u_* \leq u \leq u^*$ and hence by (3.29)

$$\int_0^t u_*(t-r) \, dH(r) \leq U(t) \leq \int_0^t u^*(t-r) \, dH(r), \; t \geq 0. \tag{3.38}$$

By direct Riemann integrability, $\sum \overline{m}_n < \infty$. By (3.35)

$$\begin{aligned} & \sup\{H(t-(n-1)h) - H(t-nh) : t \geq 0, \; n = 1, 2, \ldots\} \\ \triangleq \; & C_h < \infty. \end{aligned} \tag{3.39}$$

Now (3.35), (3.39) and the dominated convergence theorem applied to the measure given by $\{\overline{m}_n\}$ imply

$$\begin{aligned} & \lim_{t\to\infty} \int_0^t u^*(t-r) dH(r) \\ = \; & \lim_{t\to\infty} \sum_{n=1}^{\infty} \overline{m}_n \{H(t-(n-1)h) - H(t-nh)\} = \lambda h \sum_{n=1}^{\infty} \overline{m}_n. \end{aligned}$$

In a similar fashion

$$\lim_{t\to\infty} \int_0^t u_*(t-r) \, dH(r) = \lambda h \sum_{n=1}^{\infty} \overline{m}_n.$$

Consequently by (3.38) we now get, for any $h > 0$

$$\begin{aligned} & \lambda h \left( \sum_{n=1}^{\infty} \underline{m}_n \right) = \lim_{t \to \infty} \int_0^t u_*(t-r)\, dH(r) \\ \leq & \liminf_{t \to \infty} U(t) \leq \limsup_{t \to \infty} U(t) \\ \leq & \lim_{t \to \infty} \int_0^t u^*(t-r)\, dH(r) = \lambda h \left( \sum_{n=1}^{\infty} \overline{m}_n \right). \end{aligned} \tag{3.40}$$

By direct Riemann integrability

$$\lim_{h \downarrow 0} h \sum_n \underline{m}_n = \lim_{h \downarrow 0} h \sum_n \overline{m}_n = \int_0^{\infty} u(s)\, ds. \tag{3.41}$$

The required conclusion is now clear from (3.40), (3.41). □

**Note:** We now make some remarks on a result to be used in the proof of Theorem 3.21. Let $F$ be a probability distribution on $\mathbb{R}$; consider the convolution equation $\zeta(x) = \int_{-\infty}^{\infty} \zeta(x-y)\, dF(y)$, $x \in \mathbb{R}$. If $\zeta(\cdot) \equiv$ constant, clearly $\zeta$ satisfies the equation. In case $F$ is nonlattice, the celebrated Choquet-Deny theorem asserts that constants are the only solutions in the class of bounded continuous functions. See [Fe] for an analytical proof.

If $F$ is not a probability distribution, then even non-zero constants do not satisfy the convolution equation. See Section 5.4.1 for implications to the asymptotics.

Now we take up

**Proof of Theorem 3.21:** We first claim that for any fixed finite interval $I = (a, b]$,

$$\begin{aligned} \sup\{\nu_t(I) & \triangleq \mu_H(I+t) = H(b+t) - H(a+t) : t \geq 0\} \\ & < \infty. \end{aligned} \tag{3.42}$$

Indeed if $u(t) = 1 - F(t)$, $t \geq 0$ then the corresponding solution to the renewal equation is $U(t) = 1$, $t \geq 0$, and consequently by the representation (3.29)

$$\begin{aligned} 1 &= \int_0^t u(t-r)\, dH(r) \geq \int_{t-h}^t (1 - F(t-r))\, dH(r) \\ &\geq (1 - F(h))\, [H(t) - H(t-h)] \end{aligned}$$

for any $t > 0$, $h > 0$. As $\mu_F \neq \delta_0$ we can choose $h > 0$ such that $(1 - F(h)) > 0$. Consequently the above implies

$$\sup_t \{H(t) - H(t-h) : t \geq 0\} < \infty, \text{ for some } h > 0.$$

Since any finite interval is a finite union of intervals of length $\leq h$, the claim (3.42) now follows.

By (3.42) and the Helly selection principle outlined in Exercise 3.3.2 there is a sequence $t_k \to \infty$ and a $\sigma$-finite measure $\nu$ on $\mathbb{R}$ such that $\nu_{t_k} \xrightarrow{w} \nu$. Note that $\nu$ need not be supported on $[0, \infty)$. (As $\nu_t$ is supported on $[-t, \infty)$ we should expect $\nu$ to be supported on $\mathbb{R}$ and it turns out to be so.)

Now let $u$ be a continuous function supported on $[0, a]$ for some $a > 0$. Let $U$ be the corresponding solution to the renewal equation (3.28). By (3.29), for $t_k > a$ note that

$$\begin{aligned} U(x + t_k) &= \int_0^{x+t_k} u(x + t_k - r)\, dH(r) \\ &= \int_{(-t_k) \vee (x-a)}^{x} u(x - s)\, d\nu_{t_k}(s) = \int_{x-a}^{x} u(x - s)\, d\nu_{t_k}(s) \\ &= \int_{-\infty}^{\infty} u(x - s)\, d\nu_{t_k}(s), \text{ if } x \geq 0. \end{aligned}$$

If $x < 0$ then the above equality holds for $t_k > a + |x|$. As $\nu_{t_k} \xrightarrow{w} \nu$ it now follows by Exercise 3.3.1 (a) that

$$U(x + t_k) \to \int_{-\infty}^{\infty} u(x - s)\, d\nu(s) \triangleq \zeta(x) \tag{3.43}$$

as $t_k \to \infty$ for any $x \in \mathbb{R}$. By the dominated convergence theorem and (3.42) it is clear that $\zeta$ is a bounded continuous function on $\mathbb{R}$. Since $U$ solves the renewal equation (3.28), by (3.43) and the dominated convergence theorem again we get

$$\begin{aligned}
\zeta(x) &= \lim_{t_k\to\infty}\left[u(x+t_k)+\int_0^{x+t_k} U(x+t_k-y)\,dF(y)\right]\\
&= \lim_{t_k\to\infty}\int_0^{x+t_k} U(x+t_k-y)\,dF(y)\\
&\quad (\because u \text{ is supported on } [0,a])\\
&= \lim_{t_k\to\infty}\int_{-\infty}^{\infty} U(x+t_k-y)\,dF(y)\\
&\quad (\because U,\ F \text{ supported on } [0,\infty))\\
&= \int_{-\infty}^{\infty}\zeta(x-y)\,dF(y).
\end{aligned}$$

Thus

$$\zeta(x)=\int_{-\infty}^{\infty}\zeta(x-y)\,dF(y),\quad x\in\mathbb{R}. \tag{3.44}$$

As $F$ is nonlattice, the only solution to the convolution equation (3.44) is $\zeta(\cdot)\equiv$ constant by the *Choquet-Deny theorem* (see [Fe]). This means that $\int_{-\infty}^{\infty} u(x-s)\,d\nu(s)$ is independent of $x\in\mathbb{R}$ for any continuous function with compact support in $[0,\infty)$. So by standard measure theoretic arguments it follows that $\nu(I+x)=\nu(I)$ for any finite interval $I\subset\mathbb{R}$. Thus $\nu$ is proportional to the Lebesgue measure. That is, we have proved

$$\lim_{t_k\to\infty}\,[H(t_k+h)-H(t_k)]=\alpha h \tag{3.45}$$

where $\alpha$ is a constant independent of $h$.

To complete the proof we need to show that $\alpha$ is independent of the sequence $\{t_k\}$ and that $\alpha=\lambda$. By (3.45), and the proof of (a) $\Rightarrow$ (b) in Theorem 3.26 we get

$$\lim_{t_k\to\infty}\,U(t_k)=\alpha\int_0^{\infty}u(s)\,ds \tag{3.46}$$

for any directly Riemann integrable function $u(\cdot)$, where $U(\cdot)$ stands for the corresponding solution to the renewal equation (3.28). In particular take $u(t) = 1 - F(t)$, $t \geq 0$. By Lemma 3.25 (vi), as $F$ has finite expectation, we have r.h.s of (3.46) $= \alpha\lambda^{-1}$. Since l.h.s of (3.46) in this case is 1, we get $\alpha = \lambda$, completing the proof. □

As an application of the key renewal theorem we shall now prove a limit theorem for the residual waiting time. From the definition of $N(\cdot)$ we know that for every $t \geq 0$

$$T_{N(t)} \leq t < T_{N(t)+1}. \tag{3.47}$$

Recall (see Proposition 2.6) that $T_{N(t)+1} - t$ is called the *residual life* or *residual waiting time* at epoch $t$.

**Theorem 3.27** *Let the notation and hypotheses be as in Theorem 3.21. Write for* $t \geq 0$

$$Q(t,\xi) = P(T_{N(t)+1} - t \leq \xi), \quad \xi \in \mathbb{R}. \tag{3.48}$$

*So* $Q(t,\cdot)$ *is the distribution function of the residual waiting time at epoch* $t$. *Then*

$$\lim_{t\to\infty} Q(t,\xi) = \lambda \int_0^{\xi} (1 - F(s))\, ds \triangleq F_I(\xi). \tag{3.49}$$

*That is, the limit distribution of the residual waiting time is the same as the integrated tail distribution given by (3.34).*

**Proof:** By (3.47) observe that

$$\begin{aligned} Q(t,\xi) &= P(T_{N(t)} \leq t < T_{N(t)+1} \leq t+\xi) \\ &= \sum_{k=0}^{\infty} P(T_k \leq t < T_k + A_{k+1} \leq t+\xi) \\ &= \sum_{k=0}^{\infty} \int_0^t \int_{t-x}^{t-x+\xi} dF(y)\, dF^{*(k)}(x) \\ &= \int_0^t [F(t-x+\xi) - F(t-x)]\, dH(x) \end{aligned} \tag{3.50}$$

by Theorem 3.12. So by Theorem 3.16, for fixed $\xi \geq 0$ the function $Q(\cdot, \xi)$ solves the renewal equation corresponding to $u(t) = F(t+\xi) - F(t)$, $t \geq 0$. By Lemma 3.25 (iv), $u(\cdot)$ is directly Riemann integrable. So the key renewal theorem (Theorem 3.26 (b)) and Exercise 3.3.7 can now be used to obtain (3.49). □

The next result shows that the limiting distribution is independent of the initial distribution. Indeed we have

**Theorem 3.28** *Let the notation be as in Definition 3.17. Let the hypotheses of Theorem 3.21 hold. For $t \geq 0$, denote by*

$$Q^{(D)}(t,\xi) = P\left(T^{(D)}_{N^{(D)}(t)+1} - t \leq \xi\right), \quad \xi \in \mathbb{R} \tag{3.51}$$

*the distribution function of the residual waiting time at epoch $t$ corresponding to the delayed renewal process $N^{(D)}$. Then*

$$\lim_{t\to\infty} Q^{(D)}(t,\xi) = F_I(\xi). \tag{3.52}$$

*Thus the limiting distribution of the residual waiting time is independent of the initial distribution $G$.*

**Proof:** Proceeding as in the derivation of (3.50), and using (3.24)

$$\begin{aligned}
& Q^{(D)}(t,\xi) \\
= \; & P(t < A_1 \leq t+\xi) + \sum_{k=1}^{\infty} P\left(T^{(D)}_k \leq t < T^{(D)}_k + A_{k+1} \leq t+\xi\right) \\
= \; & G(t+\xi) - G(t) \\
& + \int_0^t \sum_{k=1}^{\infty} [F(t-x+\xi) - F(t-x)]\, dG * F^{*(k-1)}(x) \\
= \; & G(t+\xi) - G(t) \\
& + \int_0^t [F(t-x+\xi) - F(t-x)]\, d(G*H)(x).
\end{aligned} \tag{3.53}$$

It can be shown (see Exercise 3.3.9) that

$$\lim_{t\to\infty} \int_0^t [F(t-x+\xi) - F(t-x)]\, d(G*H)(x) = F_I(\xi). \tag{3.54}$$

As $\lim_{t\to\infty} G(t+\xi) - G(t) = 0$, the required conclusion (3.52) is now clear from (3.53), (3.54). □

The next result reinforces the steady state behaviour of the integrated tail distribution.

**Theorem 3.29** *Let the set up be as in Definition 3.17. Assume that* $E(A_i) = \lambda^{-1} < \infty$, $i \geq 2$, *and that the initial distribution* $G = F_I$ *given by (3.34). Let* $Q^{(I)}(t,\cdot)$ *denote the distribution function of the corresponding residual waiting time (at epoch* $t$*) defined by the analogue of (3.51). Then*

$$Q^{(I)}(t,\xi) = F_I(\xi), \ \xi \geq 0, \ \textit{for all } t \geq 0. \tag{3.55}$$

**Proof:** The analogue of (3.53) in this case is

$$Q^{(I)}(t,\xi) = F_I(t+\xi) - F_I(t) + \int_0^t [F(t-x+\xi) - F(t-x)]\, dH_I(x)$$

where $H_I$ denotes the corresponding delayed renewal function. By Proposition 3.19 we know that $dH_I(x) = \lambda\, dx$. Hence it follows by (3.34) that

$$\begin{aligned} & Q^{(I)}(t,\xi) \\ = \ & \lambda \left[ \xi - \int_0^{t+\xi} F(y)dy + \int_0^t F(y)dy + \int_\xi^{t+\xi} F(y)dy - \int_0^t F(y)dy \right] \\ = \ & F_I(\xi) \end{aligned}$$

completing the proof. □

**Remark 3.30** There is another approach to prove Blackwell's renewal theorem using the so called *coupling technique*. It goes as follows. Take two independent delayed renewal processes $N$ and $\tilde{N}$ with the same interarrival time distribution $F$, but the initial distribution corresponding to $N$ is $G$ while that of $\tilde{N}$ is the equilibrium distribution $F_I$. We know that $E(\tilde{N}(t)) = \lambda t$, $t \geq 0$; also the residual waiting time distribution for $\tilde{N}$ is $F_I$ for any epoch $t$. (Note also that Blackwell's theorem or key renewal theorem is not used in the proof of Theorem 3.29). The coupling method involves in showing that the processes $N$ and $\tilde{N}$ come arbitrarily close to each other in finite time resulting in coupling with

probability one. After coupling the two processes become (statistically) more and more indistinguishable. So one might as well "observe" the process $\tilde{N}$ whose asymptotics are known. Of course, all these have to be translated into mathematics; an interested reader can look up [BW] for details. There is yet another approach to renewal theorems using functional analysis; see [Ru].

## 3.4 Sparre Andersen model

This important model is a generalization of the Cramér-Lundberg model. Sparre Andersen proposed this set up as a way of studying $\mathbb{R}$-valued random walks in late 1940's. Soon connections with queueing theory and risk theory were discovered, leading to a vast literature. The approach also simplified many asymptotic results that had been obtained earlier using complicated techniques; see Section 5.1 below for a sample. An interested reader may look up Chapter XII of [Fe] to get an idea of the implications of Sparre Andersen's work.

The basic assumptions are as follows.

1. The claim number process $\{N(t) : t \geq 0\}$ is a renewal process.
2. The claim sizes $\{X_i\}$ form an i.i.d. sequence of nonnegative random variables.
3. $\{N(t)\}$ and $\{X_i\}$ are independent.

As in the earlier model, the *total claim amount* upto time $t$ is given by

$$S(t) = \sum_{i=1}^{N(t)} X_i, \quad t \geq 0 \tag{3.56}$$

which is the same as (1.2).

**Theorem 3.31** *Consider the Sparre Andersen model with the inter-arrival times $\{A_i\}$ and the claim sizes $\{X_j\}$ having finite expectation. Then*

*(i)*
$$\lim_{t\to\infty} \frac{1}{t} S(t) = \frac{E(X_1)}{E(A_1)} \ a.s.; \tag{3.57}$$

*(ii)*
$$\lim_{t\to\infty} \frac{1}{t} E(S(t)) = \frac{E(X_1)}{E(A_1)}; \tag{3.58}$$

*(iii) for any* $h > 0$

$$\lim_{t\to\infty} E(S(t+h)) - E(S(t)) = \frac{E(X_1)}{E(A_1)}\ h. \tag{3.59}$$

*Thus analogues of SLLN, elementary renewal theorem, and Blackwell's renewal theorem hold for* $\{S(t)\}$.

**Proof:** Note that $\frac{S(t)}{t} = \frac{S(t)}{N(t)}\ \frac{N(t)}{t}$. Denote $\Omega_1 = \left\{\omega \in \Omega : \frac{N(t)}{t} \to \frac{1}{E(A_1)}\right\}$, and $\Omega_2 = \{\omega \in \Omega : \frac{S(t)}{N(t)} \to E(X_1)\}$. By Theorem 3.4 we know that $P(\Omega_1) = 1$. By SLLN for the i.i.d. sequence $\{X_i\}$ we know that $\frac{1}{n}\sum_{i=1}^{n} X_i \to E(X_1)$ a.s. As $N(t) \to \infty$ a.s. apply Lemma 3.5 to get $P(\Omega_2) = 1$. Hence $P(\Omega_1 \cap \Omega_2) = 1$; so (3.57) is proved. Rest of the proof is left as exercise; see Exercise 3.4.1. □

Our next objective is to prove a central limit theorem for $\{S(t)\}$ following [EKM]. We first need a lemma.

**Lemma 3.32** *Let* $\{Z_n : n \geq 1\}$, $Z$, $\{\alpha_n : n \geq 1\}$, $\{\xi_n : n \geq 1\}$ *be random variables such that* $Z_n \xrightarrow{d} Z$, $\alpha_n \to a$ *in probability,* $\xi_n \to b$ *in probability, as* $n \to \infty$, *where* $a, b \in \mathbb{R}$ *are constants. Then* $\alpha_n Z_n + \xi_n \xrightarrow{d} aZ + b$.

**Proof:** This is basically Slutsky's theorem; see [CT, p.272] for a proof. □

**Theorem 3.33** *(Central limit theorem): Consider the Sparre Andersen model with the interarrival times* $\{A_i\}$ *and the claim sizes* $\{X_j\}$ *having finite second moments. Denote* $\lambda^{-1} = E(A_1)$, $\sigma_A^2 = Var\ (A_1)$, $\mu = E(X_1)$ *and* $\sigma_X^2 = Var\ (X_1)$. *Then*

$$\frac{1}{(\sigma_X^2 + \lambda^2\mu^2\sigma_A^2)^{1/2}}\ \frac{1}{\sqrt{\lambda t}}(S(t) - \lambda\mu t) \xrightarrow{d} N(0,1) \tag{3.60}$$

*as* $t \to \infty$.

**Proof:** Denote $\sigma = (\sigma_X^2 + \lambda^2\mu^2\sigma_A^2)^{1/2}$, $Y_i = (X_i - \lambda\mu A_i)$, $i \geq 1$, $Z_n = \frac{1}{\sigma}\sum_{i=1}^{n}(X_i - \lambda\mu A_i) = \frac{1}{\sigma}\sum_{i=1}^{n} Y_i$, $n \geq 1$. Note that

$$
\begin{aligned}
& \frac{1}{\sigma}\frac{1}{\sqrt{\lambda t}}\left[S(t)-\lambda\mu t\right] \\
= & \frac{1}{\sigma}\frac{1}{\sqrt{\lambda t}}\left[S(t)-\lambda\mu T_{N(t)}\right]+\frac{\lambda\mu}{\sigma\sqrt{\lambda t}}(T_{N(t)}-t) \\
= & \frac{1}{\sigma}\frac{1}{\sqrt{\lambda t}}\sum_{i=1}^{N(t)}(X_i-\lambda\mu A_i)+\frac{\lambda\mu}{\sigma\sqrt{\lambda t}}\left(T_{N(t)}-t\right) \\
= & \frac{1}{\sigma}\frac{1}{\sqrt{\lambda t}}\sum_{i=1}^{N(t)}Y_i+\frac{\lambda\mu}{\sigma\sqrt{\lambda t}}\left(T_{N(t)}-t\right) \\
= & \frac{1}{\sqrt{\lambda t}}Z_{[\lambda t]}+\frac{1}{\sqrt{\lambda t}}\left[Z_{N(t)}-Z_{[\lambda t]}\right]+\frac{\lambda\mu}{\sigma\sqrt{\lambda t}}\left(T_{N(t)}-t\right) \\
= & \left(\frac{[\lambda t]}{\lambda t}\right)^{1/2}\frac{1}{\sqrt{[\lambda t]}}Z_{[\lambda t]}+\frac{1}{\sqrt{\lambda t}}\left(Z_{N(t)}-Z_{[\lambda t]}\right) \\
& +\frac{\lambda\mu}{\sigma\sqrt{\lambda t}}\left(T_{N(t)}-t\right). \qquad (3.61)
\end{aligned}
$$

By independence of $\{X_\ell\}$ and $\{A_j\}$, note that $\{Y_i\}$ is an i.i.d. sequence with mean zero and variance $\sigma^2$. So by the classical central limit theorem $\frac{1}{\sqrt{n}}Z_n \xrightarrow{d} N(0,1)$, and hence $\frac{1}{\sqrt{[\lambda t]}}Z_{[\lambda t]} \xrightarrow{d} N(0,1)$ as $t\to\infty$. In view of the preceding lemma and (3.61), to complete the proof it is enough to show that

$$
\begin{aligned}
\frac{1}{\sqrt{\lambda t}}\left[Z_{N(t)}-Z_{[\lambda t]}\right] &\to 0 \text{ in probability}, \qquad (3.62)\\
\frac{1}{\sqrt{t}}\left(T_{N(t)}-t\right) &\to 0 \text{ in probability}. \qquad (3.63)
\end{aligned}
$$

Let $\epsilon>0$ be arbitrary; put $C_\epsilon=\left\{\frac{1}{\sqrt{\lambda t}}\,|Z_{N(t)}-Z_{[\lambda t]}\,|>\epsilon\right\}$. Let $\delta>0$, put

$$
\begin{aligned}
B_1 &= \left\{\left|\frac{N(t)}{t}-\lambda\right|>\delta\right\},\\
B_2 &= \left\{\left|\frac{N(t)}{t}-\lambda\right|\le\delta,\frac{1}{\sqrt{\lambda t}}\mid Z_{N(t)}-Z_{[\lambda t]}\mid>\epsilon\right\}.
\end{aligned}
$$

Clearly $C_\epsilon\subseteq B_1\cup B_2$; (dependence of these sets on $t$ is suppressed). By SLLN, $P(B_1)\to 0$ as $t\to\infty$ for any $\delta>0$. Next by Kolmogorov's

inequality

$$\begin{aligned} P(B_2) &\leq P\left(\max\left\{|\, Z_n - Z_{[\lambda t]}\,| : \left|\frac{n}{t} - \lambda\right| \leq \delta\right\} > \epsilon\sqrt{\lambda t}\right) \\ &\leq \frac{2}{\epsilon^2 \lambda t} \text{ Var } \left(Z_{[(\lambda+\delta)t]} - Z_{[\lambda t]}\right) \leq K\frac{\delta}{\epsilon^2} \end{aligned}$$

where $K$ is a positive constant independent of $t, \epsilon, \delta$. For fixed $\epsilon > 0$, first choose $\delta > 0$ such that $P(B_2)$ is very small for all $t$. Then choose $t$ large enough so that $P(B_1)$ is small. Thus $P(C_\epsilon) \to 0$ as $t \to \infty$ for any $\epsilon > 0$. This proves (3.62).

Next for $\epsilon > 0$,

$$\begin{aligned} & P\left(\frac{1}{\sqrt{t}}\left|T_{N(t)} - t\right| > \epsilon\right) = P(t - T_{N(t)} > \epsilon\sqrt{t}) \\ = & \sum_{k=0}^{\infty} P(t - T_{N(t)} > \epsilon\sqrt{t} \mid N(t) = k) \cdot P(N(t) = k) \\ = & \sum_{k=0}^{\infty} P(A_{k+1} > \epsilon\sqrt{t} \mid N(t) = k) \cdot P(N(t) = k) \\ = & \sum_{k=0}^{\infty} P(A_1 > \epsilon\sqrt{t}) \cdot P(N(t) = k) \ (\because \{A_i\} \text{ i.i.d.'s}) \\ = & P(A_1 > \epsilon\sqrt{t}) \to 0 \text{ as } t \to \infty \end{aligned}$$

which implies (3.63), completing the proof. □

The rest of this section is devoted to an elementary discussion of premium principles.

### 3.4.1 Some basic premium principles

One of the main objectives of an insurance company is to determine the premium, or the premium rate to be more precise. Loss to the insurance company is described by the total claim amount $S(\cdot)$. Clearly the premium rate is to be chosen so that losses over time are covered. The premium income $p(t)$ is taken to be a deterministic function; here $p(t)$ = premium income upto time $t$.

Though $E(S(t))$ may be a rather crude estimate for $S(t)$, it is still a very useful approximation. Often it may not be possible or easy to determine even asymptotic behaviour of higher moments of $S(t)$. Also in view of the SLLN, if $p(t) < E(S(t))$ for large $t$ then one may say that

the company loses on the average; similarly if $p(t) > E(S(t))$ for large $t$, then the company may gain on the average. With this perspective, the following premium principles are in vogue; also these are perhaps elementary to implement.

1. *Net or equivalence principle.* According to this system, $p_{\text{net}}(t) = E(S(t))$, $t \geq 0$. In this situation, the insurance company does not intend to lose or gain on average; so it is sometimes called "fair market premium" or "bechmark premium". It can serve only as a theoretical value. Due to fluctuations the deviations from the mean can be positive as well as negative, and hence it would be unwise to charge premium according to this principle. In fact we shall prove later that this principle leads to "ruin" of the company in finite time with probability one.

2. *Expected value principle.* In view of our preliminary remarks, it seems reasonable to choose a premium by "loading" the expected total claim amount by suitable "safety factor". Accordingly

$$p_{EV}(t) \approx (1 + \rho)\ E(S(t)), \tag{3.64}$$

at least approximately, where $\rho > 0$ is the *safety loading factor.* By Theorem 3.31 (ii), in the Sparre Andersen model

$$E(S(t)) = \frac{E(X_1)}{E(A_1)}\ t + o(t), \text{ as } t \to \infty.$$

Hence

$$p_{EV}(t) = (1 + \rho)\ \frac{E(X_1)}{E(A_1)}\ t, \quad t \geq 0 \tag{3.65}$$

is a reasonable way to fix the premium. This is perhaps the most commonly used and easy to implement procedure. Of course, too high a loading may make the premium unattractive to the prospective buyer of insurance policy.

The expected value principle, however, does not take fluctuations into account. The next two approaches take cognisance of this when the second moments exist.

3. *Variance principle.* In this case, assuming that the second moments exist,

$$p_{\text{Var}}(t) \approx E(S(t)) + \alpha\ \text{Var}\ (S(t)) \tag{3.66}$$

where $\alpha > 0$ is a constant. In view of the central limit theorem for the Sparre Andersen model one can take

$$\begin{aligned} & p_{\mathrm{Var}}(t) \\ = \; & \frac{E(X_1)}{E(A_1)}t + \alpha\left(\mathrm{Var}(X_1) + \left(\frac{E(X_1)}{E(A_1)}\right)^2 \mathrm{Var}(A_1)\right)\frac{1}{E(A_1)}t \\ = \; & \lambda\mu t + \alpha(\sigma_X^2 + \lambda^2\mu^2\sigma_A^2)\,\lambda t,\; t \geq 0. \end{aligned} \tag{3.67}$$

4. *Standard deviation principle.* This is similar to the preceding one is spirit; it is desired that

$$p_{SD}(t) \approx E(S(t)) + \alpha\sqrt{\mathrm{Var}(S(t))}, \tag{3.68}$$

where $\alpha$ is a positive constant. So

$$p_{SD}(t) = \lambda\mu t + \alpha(\sigma_X^2 + \lambda^2\mu^2\sigma_A^2)^{1/2}\sqrt{\lambda}\sqrt{t}, \quad t \geq 0 \tag{3.69}$$

in the Sparre Andersen model when the second moments exist.

Note that the total claim amount exceeding the premium income is an event to be avoided; so it is desirable to fix the premium in order to make the probability of this eventuality small. When second moments exist, using the central limit theorem, justification for $p_{\mathrm{Var}}(\cdot)$ and $p_{SD}(\cdot)$ can be given; see Exercises 3.4.5, 3.4.6.

5. *Exponential principle.* According to this notion

$$p_{\mathrm{Exp}}(t) = \frac{1}{a}\,\log E[e^{aS(t)}], \quad t > 0 \tag{3.70}$$

where $a > 0$ is a constant called *risk aversion factor*. The rationale for the exponential principle can be given in terms of *utility theory* of economics. This theory was developed by von Neumann and Morgenstern. We give below a brief discussion in the context of insurance.

*Some remarks on utility functions vis-a-vis insurance.* We have seen that it is desirable to have $p(t) > E(S(t))$, certainly from the point of view of the insurer (i.e. insurance company). But this would mean that the insured (i.e. the end users who avail of the policy, considered collectively) pay more than what they expect to get on the average. The very fact that the insurance business exists indicates that a large

section of the population is willing to pay a price (for being insured) that is greater than the expected claim amount. In economics, an explanation for such apparent anomalous situations is given in terms of the so called utility functions. This postulates the existence of a utility function which determines decisions, in our case, collectively for the insured.

Let $I \subseteq \mathbb{R}$ be an interval; it need not be a finite interval. A function $v : I \to \mathbb{R}$ is called a *utility function* if it is nondecreasing and concave. For each wealth $x \in I$ a utility $v(x)$ is attached; $v(x)$ can be different from $x$. Since utility may be assumed to increase with wealth $v(\cdot)$ is assumed to be nondecreasing; this is also phrased as "marginal utility is nonnegative", that is, $v'(x) \geq 0$ if $v'(\cdot)$ makes sense. Next, note that giving Re.1 to a poor person makes more sense than giving it to a millionaire. So increments of $v(x)$ for small values of $x$ should be more than those for large values of $x$; that is, if $x_1 < x_2$, $h > 0$ then $v(x_1 + h) - v(x_1) \geq v(x_2 + h) - v(x_2)$; again, this may be termed as "decreasing marginal utility" or $v''(x) \leq 0$ whenever $v''(\cdot)$ makes sense. Thus $v(\cdot)$ should be concave.

The concavity assumption is also known as *risk aversion.* A decision maker is risk averse if a fixed loss is preferred over a random loss that has the same expected value. If we consider a random loss $Y$, then the expected utility is $E[v(-Y)]$. By Jensen's inequality $E[v(-Y)] \leq v(E(-Y))$ as $v$ is concave. So the fixed loss $E(-Y)$ is given greater utility than the expected utility of the random loss. Risk averse decision makers are generally considered 'reasonable' decision makers.

We assume that there is a utility function $v$ collectively for the potential buyers of insurance policy. For the insured the loss is $S(t)$ in the absence of insurance; (if the initial wealth is $x$ then it gets depleted to $x - S(t)$ if there is no insurance; for simplicity we take $x = 0$). So in the absence of insurance the expected utility is $E[v(-S(t))]$. If insurance cover is taken, similarly the utility at time is $v(-p(t))$, where $p(t)$ denotes the premium paid upto $t$. Therefore insurance will be taken by a potential buyer only if

$$E[v(-S(t))] \leq v(-p(t)), \quad t \geq 0; \tag{3.71}$$

that is, only if taking insurance provides greater utility than not having insurance cover. As $v$ is increasing, the "maximum premium" the insured may be willing to pay is determined by the solution $p^+(\cdot)$ to the equation

$$E[v(-S(t))] = v(-p^+(t)), \quad t \geq 0. \tag{3.72}$$

Equation (3.72) is called the *utility equilibrium equation*, and $p^+(\cdot)$ is called *zero utility premium*.

Our discussion regarding utility function and the insureds is applicable in an analogous manner to the insurance company as well. If $V$ is the utility function for the insurer, then the company will undertake insuring the risk only if

$$E[V(p(t) - S(t))] \geq V(0) \tag{3.73}$$

(or more generally only if $E[V(u + p(t) - S(t))] \geq V(u)$, where $u \geq 0$ is the initial capital). So $p^-(\cdot)$, the minimum premium the company will want, can be obtained as the solution the corresponding utility equilibrium equation

$$E[V(p^-(t) - S(t))] = V(0). \tag{3.74}$$

A deal improving the expected utility for both sides will be possible if $p^+(\cdot) \geq p^-(\cdot)$; this is a win-win situation.

Some useful utility functions are:
(i) linear utility: $v(x) = x$;
(ii) quadratic utility: $v(x) = \begin{cases} -(\alpha - x)^2, & x \leq \alpha \\ 0, & x \geq \alpha \end{cases}$ where $\alpha > 0$ is a constant;
(iii) exponential utility: $v(x) = -ae^{-ax}$, where $a > 0$.

The exponential premium principle (3.70) can be justified as the zero utility premium $p^+(\cdot)$ corresponding to the exponential utility function. Though the exponential principle is very useful and has nice theoretical features (see Exercise 3.4.8), as existence of moment generating function of $S(\cdot)$ is assumed, it is not applicable for heavy tailed cases.

For more information on premium principles, utility function, etc. see [RSST], [KGDD].

## 3.5 Exercises

### Exercises on Section 3.1

**3.1.1.** Suppose the i.i.d. interarrival times $\{A_i\}$ satisfy $P(A_i > 0) = 1$. For any $t > 0$ show that $P(N(t) - N(t-) > 1) = 0$.

**3.1.2.** Let the interarrival times $\{A_i\}$ satisfy (3.2). Assume that $P(A_i = \infty) = 0$; that is, the distribution of $A_i$ is nondefective.

(i) Show that $P(N(t) \to \infty) = 1$.

(ii) If $E(A_i) = \infty$ show that $\frac{1}{t}\, N(t) \to 0$ a.s.

In other words, SLLN holds in this case as well.

**3.1.3.** Let the interarrival times $\{A_i\}$ satisfy (3.2). Let $\{T_n\}$, $\{N(t)\}$ be the corresponding renewal sequence, renewal process respectively. Put $\mathcal{F}_0 = \{\phi, \Omega\}$, the trivial $\sigma$-algebra, $\mathcal{F}_n = \sigma\{A_i : i \leq n\}$ = smallest $\sigma$-algebra making $\{A_i : i \leq n\}$ measurable, $n = 1, 2, \ldots$; that is $\{\mathcal{F}_n : n \geq 0\}$ is the natural filtration associated with $\{A_i\}$.

(i) Fix $t \geq 0$. Show that $(N(t) + 1)$ is a *stopping time w.r.t.* $\{\mathcal{F}_n\}$; that is, $\{N(t) + 1 \leq k\} \in \mathcal{F}_k$ for any $k = 0, 1, 2, \ldots$. (Hint: Use (3.5).)

(ii) Show that $(N(t) + 1) = \inf\{k \geq 0 : T_k > t\}$.

(iii) Show that $N(t)$ is not a stopping time w.r.t. $\{\mathcal{F}_n\}$.

(iv) Show that the event $\{N(t)+1 \geq k\}$ is independent of the family $\{A_i : i \geq k\}$ of random variables, for fixed $t \geq 0$, $k = 0, 1, 2, \ldots$ .

**3.1.4.** (Continuation). Let $g$ be a bounded nonnegative measurable function on $[0, \infty)$. Fix $t \geq 0$. For $i = 0, 1, 2, \ldots$ define

$$Y_i = I_{[i,\infty)}(N(t) + 1) = \begin{cases} 0, & \text{if } i > N(t) + 1 \\ 1, & \text{if } i \leq N(t) + 1. \end{cases}$$

(i) Show that $\sum_{i=1}^{N(t)+1} g(A_i) = \sum_{i=1}^{\infty} g(A_i) Y_i$.

(ii) For fixed $i = 1, 2, \ldots$ show that $Y_i$ and $A_i$ are independent.

(iii) Show that

$$\begin{aligned} E\left[\sum_{i=1}^{N(t)+1} g(A_i)\right] &= E[g(A_1)] \sum_{i=1}^{\infty} P(N(t) + 1 \geq i) \\ &= E\,[g(A_1)] \cdot E\,[N(t) + 1] \end{aligned}$$

which is Wald's identity.

**3.1.5.** Let $\{A_i\}$, $\{T_n\}$, $\{\mathcal{F}_n\}$ be as in Exercise 3.1.3. Assume that $E(A_i) < \infty$. Set $\hat{A}_i = A_i - E(A_i)$, $i \geq 1$, and $\hat{T}_n = \hat{A}_1 + \ldots + \hat{A}_n = T_n - nE(A_1)$, $n \geq 1$; of course $\hat{T}_0 \equiv 0$. Show that $\{\hat{T}_n : n = 0, 1, 2, \ldots\}$ is a *martingale* w.r.t. $\{\mathcal{F}_n\}$; that is, $\hat{T}_n$ is integrable and $E\,[\hat{T}_{n+1} \mid \mathcal{F}_n] = T_n$, $n \geq 1$.

**3.1.6.** Let $c > 0$ be a constant such that the last equality in (3.16) holds. Show that $c = \lambda^3 \mathrm{Var}(A_1)$. This observation is due to Srikant Iyer and Adimurthy. (Note: If $E(N(t))$ is asymptotically $\lambda t$ and $0 < \mathrm{Var} A_1 < \infty$ one hopes for a CLT for $N(\cdot)$. Exercise 2.1.5 suggests

a scaling factor in terms of $\sqrt{t}$. Proof of Theorem 3.7 implies that if (3.16) holds then (3.14) follows. The above observation shows that the choice of $c$ in Theorem 3.7 is the appropriate one.)

**3.1.7.** This is a variation of Lemma 3.5. Let $\{Z_n : n \geq 1\}, Z, \{M(t) : t \geq 0\}$ be such that (i) $Z_n \to Z$ a.s., (ii) $M(t)$ integer valued, (iii) $M(t) \to \infty$ in probability. Then $Z_{M(t)} \to Z$ in probability. (Hint: For any sequence $\{t_k\}$ get a subsequence $\{t_{k_i}\}$ such that $M(t_{k_i}) \to \infty$ a.s. as $i \to \infty$, and use Lemma 3.5).

## Exercises on Section 3.2

**3.2.1.** Let $\mu$, $\nu$ be finite measures on $\mathbb{R}$. Let the convolution $\mu * \nu$ be defined by $(\mu * \nu)(E) = \int \mu(E - x)\, d\nu(x)$, for any Borel set $E$. For any Borel measurable function $g$ on $\mathbb{R}$ show that

$$\int g(z)\, d(\mu * \nu)(z) = \int \int g(z + x)\, d\mu(z)\, d\nu(x).$$

**3.2.2.** (a) Complete the proof of Theorem 3.18.
(b) Let the hypotheses of Definition 3.17 hold. Show that $G * H = H * G$ is well defined and $H^{(D)} = G * H$.
(c) Assuming $E(A_i) = \lambda^{-1} < \infty$ for $i \geq 2$ show that $\frac{1}{t} H^{(D)}(t) \to \lambda$ as $t \to \infty$.

**3.2.3.** Find the integrated tail distribution $F_I$ in the case of a homogeneous Poisson process, and interpret this as a residual life distribution.

**3.2.4.** Extend the renewal function $H$ to the real line by setting $H(t) = 0$, $t < 0$. Let $z \geq 0$ be fixed. Consider $g(t) = H(t) - H(t - z)$ as a function on $\mathbb{R}$. Note that $g(t)$ is the expected number of renewal epochs in the interval $(t - z, t]$. Show that $g(\cdot)$ is the unique locally bounded function supported on $[0, \infty)$ solving the renewal equation (3.28) with $u(t) = I_{[0,z)}(t)$.

**3.2.5.** Consider the renewal equation (3.28) with the following $u(\cdot)$ supported on $[0, \infty)$. In each case find the locally bounded solution supported on $[0, \infty)$.
(i) $u(t) = F(t)$, (ii) $u(t) = 1 - F(t)$; (iii) $u(t) = G(t) - G(t - h)$ where $h > 0$ is fixed and $G$ is the distribution function of a probability measure supported on $[0, \infty)$.

(*Ans.* (i) $U(t) = E(N(t))$; (ii) $U(t) = 1$; (iii) $U(t) = H^{(D)}(t) - H^{(D)}(t - h)$ where $H^{(D)}$ is the delayed renewal function corresponding to the initial distribution $G$.)

**3.2.6.** Let the hypotheses and notation be as in Definition 3.17. Let the interarrival time $A_i$, $i \geq 2$ have finite expectation $\lambda^{-1}$. If $H^{(D)}(t) = \lambda t$, $t \geq 0$ show that $G(t) = F_I(t)$ given by (3.34).

**Exercises on Section 3.3**

**3.3.1.** (a) Let $G_n$, $n \geq 1$, $G$ be as in Definition 3.22; that is $G_n \xrightarrow{w} G$. Let $f$ be a continuous function with compact support in $\mathbb{R}$. Show that $\lim_{n\to\infty} \int f(y)\, dG_n(y) = \int f(y)\, dG(y)$.

(b) Let the hypotheses of Blackwell's renewal theorem hold. Let $u(\cdot)$ be a continuous function with compact support in $[0, \infty)$, and $U(\cdot)$ the corresponding solution to the renewal equation (3.28). Using (a) above show that

$$\lim_{t\to\infty} U(t) = \lambda \int_0^\infty u(s)\, ds.$$

**3.3.2.** Suppose $\{\mu_n\}$ is a sequence of $\sigma$-finite measures on $\mathbb{R}$ such that $\{\mu_k([-x, x]) : k = 1, 2, \ldots\}$ is a bounded sequence of numbers for each $x \in \mathbb{R}$. Show that there exist a subsequence $\{\mu_{n_k}\}$ and a $\sigma$-finite measure $\mu$ such that $\mu_{n_k} \xrightarrow{w} \mu$ in the sense of Definition 3.22 (Hint: Imitate the proof of *Helly selection principle* involving Cantor's diagonalization argument for every $[-K, K]$. Take further subsequences. See [Fe].)

**3.3.3.** (a) Let $u(\cdot)$ be a nonnegative continuous function on $[0, \infty)$. Let $a_n = \sup\{u(x) : n - 1 \leq x \leq n\}$, $n = 1, 2, \ldots$. Show that $u(\cdot)$ is directly Riemann integrable if and only if $\sum_n a_n < \infty$.

(b) Show that the assertion in (a) holds even if $u(\cdot)$ is a nonnegative right continuous function having left limit at every $t$.

**3.3.4.** (a) Let $u(\cdot)$ be directly Riemann integrable. Show that $u(\cdot)$ is integrable and $\lim_{x\to+\infty} u(x) = 0$. What is $\lim_{h\downarrow 0} \overline{u}(h)$? (Hint: Take $h = 2^{-k}$, $k = 1, 2, \ldots$.)

(b) Show by an example that the converse of the above is not true in general.

**3.3.5.** Show that a nonincreasing, nonnegative function on $[0, \infty)$ is directly Riemann integrable if and only if it is integrable.

**3.3.6.** (a) Let $u(\cdot)$ be directly Riemann integrable. Show that $u^+(\cdot)$, $u^-(\cdot)$, $|u|$ are directly Riemann integrable. (Recall $u^+(x) = \max\{0, u(x)\}$, $u^-(x) = \max\{0, -u(x)\}$.)

(b) Let $u(\cdot)$ be continuous and $|u|$ be directly Riemann integrable. Show that $u(\cdot)$ is directly Riemann integrable. Can the continuity assumption be dropped?

**3.3.7.** (a) Complete the proof of Lemma 3.25.

(b) Find $\int_0^\infty u(s)\,ds$ in Lemma 3.25. (*Ans.* (iv) By interchanging order of integration,

$$\begin{aligned}\int_0^\infty u(s)\,ds &= \int_0^\xi \left(\int_0^r ds\right) dF(r) + \int_\xi^\infty \left(\int_{r-\xi}^r ds\right) dF(r) \\ &= \int_0^\xi (1 - F(r))dr.\end{aligned}$$

(v) Similarly $\int_0^\infty u(s)\,ds = \int_0^\infty \left(\int_r^{r+\xi} ds\right) dF(r) = \xi$.)

**3.3.8.** Consider the delayed renewal process $N^{(D)}$ as in Definition 3.17. Let $F$ be nonlattice and have finite expectation $\lambda^{-1}$. Let $G$ be the initial distribution, and $H^{(D)}$ denote the delayed renewal function
(i) Show that

$$H^{(D)}(t+h) - H^{(D)}(t) = \int_0^{t+h} [H(t+h-s) - H(t-s)]\,dG(s).$$

(ii) Show that $\lim_{t\to\infty} H^{(D)}(t+h) - H^{(D)}(t) = \lambda\, h,\ h > 0$. Thus Blackwell's theorem holds for the delayed renewal process. This also means that the renewal rate tends to a constant, and that the constant rate is independent of the initial distribution $G$.
(iii) Using Exercise 3.2.5 (iii) and the key renewal theorem give another proof of part (ii) above.

**3.3.9.** (a) Let the notation be as in Theorem 3.28. Set $J(t) = \int_0^t [F(t-x+\xi) - F(t-x)]\,d(G*H)(x),\ t \geq 0$ where $\xi \geq 0$ is fixed. Using Exercises 3.2.2 (b), 3.3.8 (ii), 3.3.7 (b), Lemma 3.25 (iv), and the proof of (a) $\Rightarrow$ (b) in Theorem 3.26 conclude that $\lim_{t\to\infty} J(t) = F_I(\xi)$.

(b) With notation as in (a) above show that $J(\cdot)$ is the locally

bounded solution to the renewal equation (3.28) with

$$u(s) = \int_0^s [F(s+\xi-y) - F(s-y)]\, dG(y), \quad s \geq 0$$

using Exercise 3.2.1. Next using Exercise 3.3.3 (a) show that $u$ is directly Riemann integrable; (see also the proof of Lemma 3.25 (iv)). Show that

$$\begin{aligned} \int_0^\infty u(s)\, ds &= \int_0^\infty \int_y^\infty [F(s+\xi-y) - F(s-y)\, ds\, dG(y) \\ &= \int_0^\infty \int_0^\infty [F(r+\xi) - F(r)]\, dr\, dG(y). \end{aligned}$$

Applying the key renewal theorem show that $\lim_{t\to\infty} J(t) = F_I(\xi)$. Thus we have another way of proving the result in (a).

**3.3.10.** Let $M \geq 1$ be an integer. Let $\{A_i : i \geq 1\}$ be a sequence of independent nonnegative random variables such that $A_i$ has distribution function $G_i$, $i = 1, 2, \ldots, M$, and each $A_j$, $j \geq M+1$ has the distribution function $F$. Define $\tilde{T}_0 \equiv 0$, $\tilde{T}_n = A_1 + A_2 + \ldots + A_n$, $n \geq 1$. $\{\tilde{T}_n\}$ may be called the *$M$-delayed renewal sequence.* Let $\tilde{N}$ denote the associated *$M$-delayed renewal process*, $\tilde{H}$ the *$M$-delayed renewal function* associated with $\tilde{N}$.

(a) Find the renewal equation satisfied by the function $\tilde{L}(t) = \tilde{H}(t) - \sum_{\ell=1}^{M-1} (G_1 * \ldots * G_\ell)(t)$, $t \geq 0$. Show that $\tilde{L}$ is the renewal function associated with a delayed renewal process.

(b) Find the analogue of (3.53) for $\tilde{N}$.

(c) Show that analogues of Theorems 3.6, 3.21 and 3.28 hold for the process $\tilde{N}$.

## Exercises on Section 3.4

**3.4.1.** (a) In the Sparre Andersen model assume that $E(X_1) < \infty$. Show that $E(S(t))$ exists and $E(S(t)) = E(X_1)E(N(t))$.

(b) Complete the proof of Theorem 3.31.

**3.4.2.** Assume that $E(X_1^2) < \infty$ in the Sparre Andersen model. Show that

$$\mathrm{Var}(S(t)) = \mathrm{Var}(X_1)\, E(N(t)) + (E(X_1))^2 \mathrm{Var}(N(t)).$$

(Hint: Consider $E[(S(t))^2 \mid N(t) = n]$.)

**3.4.3.** Formulate a CLT for the total claim amount process in the Cramér-Lundberg model.

**3.4.4.** Using the approach in the proof of Theorem 3.33 give a proof of Theorem 3.7 (CLT for the renewal process $N(\cdot)$).

**3.4.5.** Let $\alpha > 0$ and consider the standard deviation principle for Sparre Andersen model.

a) Show that

$$\lim_{t\to\infty} P(S(t) - p_{SD}(t) > 0) = 1 - \Phi(\alpha),$$

where $\Phi$ denotes the distribution function of the standard normal distribution.

b) Indicate how (a) above gives a justification for the standard deviation principle.

c) Show that $\frac{p_{SD}(t)}{p_{\text{net}}(t)} \to 1$ as $t \to \infty$.

So one may say that in the standard deviation principle, as in the net premium principle, there is no gain in the long run for the insurer. What advantage does the standard deviation principle have over the net principle? (Note that the premium in the former is higher; so, how to justify it?)

**3.4.6.** Let $\alpha > 0$ and consider the variance principle for Sparre Andersen model.

(a) What is $\lim_{t\to\infty} P(S(t) - p_{\text{Var}}(t) > 0)$?

(b) Show that $\frac{p_{\text{Var}}(t)}{p_{EV}(t)}$ converges to a constant as $t \to \infty$.

(c) Using (a), (b) above indicate a justification for the variance principle, with $\alpha$ serving as a safety loading factor.

**3.4.7.** Consider the Cramér-Lundberg model with appropriate assumptions. Suppose $p(t) = ct$, $t \geq 0$ where $c > 0$ is a constant; $c$ is called the *premium rate*. Determine $c$ by each of the five premium principles discussed in this section.

**3.4.8.** Consider the same exponential utility function for the insurer and the insured; that is

$$v(x) = V(x) = -a\, e^{-ax}, \quad x \in \mathbb{R}$$

where $a > 0$ is the risk aversion factor. Formulate the utility equilibrium equations for the insurer and the insured respectively. Show that $p^+(t) = p^-(t) = p_{\text{Exp}}(t)$ given by (3.70). This gives a justification for the exponential premium principle.

**3.4.9.** In actuarial context, some of the desirable theoretical requirements on the premium principles are:

(a) *Nonnegative loading*: $p(t) \geq E(S(t))$.

(b) *Consistency*: If the premium for $S(t)$ is $p(t)$, then the premium for $S(t) + c$ is $p(t) + c$.

(c) *Additivity*: For independent total claim amounts $S(t)$, $S'(t)$ with respective premiums $p(t)$, $p'(t)$, the premium for $S(t) + S'(t)$ should be $p(t) + p'(t)$.

(d) *Homogeneity*: For $c > 0$, the premium for $cS(t)$ should be $cp(t)$.

Now consider the Cramér-Lundberg model with appropriate assumptions.

(i) Show that $p_{\text{Net}}(\cdot)$ satisfies all the four requirements.

(ii) Show that $p_{EV}(\cdot)$ satisfies (a), (c), (d) but not (b).

(iii) Show that $p_{\text{Var}}(\cdot)$ satisfies (a), (b), (c) but not (d).

(iv) Show that $p_{SD}(\cdot)$ satisfies (a), (b), (d) but not (c).

(v) Show that $p_{\text{Exp}}(\cdot)$ satisfies (a), (b), (c) but not (d).

**3.4.10.** For the linear utility function $v(x) = x$, find the zero utility premium.

# Chapter 4

# Claim size distributions

In principle, any distribution supported on $[0, \infty)$ can be used to model claim size. However, in actuarial practice a clear distinction is made between "well behaved" distributions and "dangerous" distributions. As we shall see later there are very good reasons for this distinction.

## 4.1 Light-tailed distributions

Let $F$ be a distribution function supported on $[0, \infty)$, with $X$ denoting a generic random variable with distribution function $F$. Let

$$\hat{m}_F(s) = E(e^{sX}) = \int_0^\infty e^{sx} dF(x), \ s \in \mathbb{R} \tag{4.1}$$

denote the *moment generating function of* $F$, whenever the integral is finite. Let

$$s_F = \sup\{s \geq 0 : \hat{m}_F(s) < \infty\}. \tag{4.2}$$

**Proposition 4.1** *If* $s_F > 0$ *then there exists* $C > 0, 0 < \lambda < s_F$ *such that*

$$1 - F(x) \leq C \ e^{-\lambda x}, \ \textit{for all } x \geq 0. \tag{4.3}$$

*Conversely, if (4.3) holds for some* $C > 0, \ \lambda > 0$ *then* $\hat{m}_F(s) < \infty$ *for all* $s < \lambda$, *and hence* $s_F > 0$.

**Proof:** Let $s_F > 0$. So there is $\lambda > 0$ such that $\hat{m}_F(\lambda) < \infty$. Now for any $x \geq 0$

$$\begin{aligned} 1 - F(x) &= \int_x^\infty dF(y) \leq \int_x^\infty e^{\lambda(y-x)} dF(y) \\ &\leq e^{-\lambda x} \int_0^\infty e^{\lambda y} dF(y) = [\hat{m}_F(\lambda)]\, e^{-\lambda x} \end{aligned}$$

establishing (4.3).

To prove the converse, note that for $s > 0$

$$\hat{m}_F(s) - 1 = \int_0^\infty (e^{sx} - 1)\, dF(x) = \int_0^\infty \int_0^x s e^{sy} dy\, dF(x)$$

$$= s \int_0^\infty \left( \int_y^\infty dF(x) \right) e^{sy} dy = s \int_0^\infty (1 - F(y))\, e^{sy} dy \qquad (4.4)$$

by Tonnelli's theorem, in the sense that if one side of (4.4) is infinite then the other is also infinite, and in case of their being finite the two sides are equal. Now suppose (4.3) holds for some $\lambda > 0$, $C > 0$. Then for $0 < s < \lambda$ we get

$$\frac{\hat{m}_F(s) - 1}{s} \leq C \int_0^\infty e^{-\lambda x} e^{sx} dx = \frac{C}{(\lambda - s)}.$$

Thus $\hat{m}_F(s) < \infty$ for all $0 \leq s < \lambda$. As $F$ is supported on $[0, \infty)$ it is clear that $\hat{m}_F(s) < \infty$ for all $s < 0$. Hence $s_F > 0$, completing the proof. □

**Definition 4.2** Let $F$ be a distribution function supported on $[0, \infty)$. $F$ is said to be *light-tailed* if there exist constants $C > 0$, $\lambda > 0$ such that (4.3) holds for all $x \geq 0$, or equivalently by Proposition 4.1 $\hat{m}_F(\cdot)$ exists in a neighbourhood of 0.

The above is generally regarded as the proper notion of a well behaved distribution.

The following are examples of light-tailed distributions:
(i) Exponential distribution, Exp($\lambda$), with parameter $\lambda > 0$.

(ii) Gamma distribution, $\Gamma(\alpha, \lambda)$, with parameters $\lambda > 0,\ \alpha > 0$.
(iii) Any distribution with compact support.
(iv) *Truncated standard normal*, that is, $F(x) = P(|N(0,1)| \leq x),\ x \in \mathbb{R}$. Note that $F(x) = 2\Phi(x) - 1,\ x \geq 0$ where $\Phi$ is the distribution function of the standard normal distribution.

To any light tailed distribution a useful class of associated distributions can be defined. Let $F$ be a light tailed distribution. Then $s_F > 0$. Let $t < s_F$. Define

$$\tilde{F}_t(x) = \begin{cases} \frac{1}{\hat{m}_F(t)} \int\limits_{[0,x]} e^{ty} dF(y), & \text{if } x \geq 0 \\ 0, & \text{if } x < 0. \end{cases} \tag{4.5}$$

The distribution function $\tilde{F}_t$ is called an *Esscher transform of* $F$ or an *exponential tilting of* $F$. (Some authors call it a *Cramér transform.* Using Esscher transform one can get some information on the tail behaviour of $F^{*(n)}$ as given below.

**Proposition 4.3** *Let $F$ be a light tailed distribution function supported on $[0, \infty)$. Then for $t < s_F$ and $n = 1, 2, \ldots$*

$$(\hat{m}_F(t))^n\ d\tilde{F}_t^{*(n)}\ (y) = e^{ty}\ dF^{*(n)}\ (y) \tag{4.6}$$

*and hence for all $x \geq 0$*

$$1 - F^{*(n)}(x) = (\hat{m}_F(t))^n \int\limits_x^\infty e^{-ty}\ d\tilde{F}_t^{*(n)}\ (y). \tag{4.7}$$

**Proof:** Note that $s_F > 0$. By Exercise 4.1.4 below we know that the moment generating function of $\tilde{F}_t$ exists in a neighbourhood of 0 and that for $0 < s < s_F - t,\ n = 1, 2, \ldots$

$$(\hat{m}_F(t))^n\ (\hat{m}_{\tilde{F}_t}(s))^n = (\hat{m}_F(t+s))^n.$$

As $(\hat{m}_{\tilde{F}_t}(\cdot))^n$, $(\hat{m}_F(\cdot))^n$ are respectively the moment generating functions of $\tilde{F}_t^{*(n)}$, $F^{*(n)}$ we now get from the above

$$(\hat{m}_F(t))^n \int\limits_{[0,\infty)} e^{sy}\ d\tilde{F}_t^{*(n)}\ (y) = \int\limits_{[0,\infty)} e^{(t+s)y}\ dF^{*(n)}\ (y).$$

The assertion (4.6) follows now by uniqueness of moment generating function. (4.7) is an easy consequence of (4.6). □

For a light tailed distribution $F$, it is clear that the classical central limit theorem is applicable to $X_1 + \ldots + X_n$, $n \geq 1$ where $\{X_i\}$ is a sequence of i.i.d.'s with distribution $F$. So asymptotic normality can sometimes be used as an approximation when time parameter is fairly large.

## 4.2 Regularly varying functions and distributions

Any distribution function $F$ supported on $[0, \infty)$ such that $\hat{m}_F(s) = \infty$ for all $s > 0$ may be considered a *heavy-tailed distribution* for actuarial purposes. However, one would like to identify smaller subclasses, parametric families if possible, which are useful to get reasonable models and are mathematically amenable.

For example, if $F$ does not have finite variance and belongs to the domain of attraction of a non-Gaussian stable law, then it is known that the tail behaviour of $F$ is related to a power law; see Appendix A.2. This suggests that one may try to define heavy tailed behaviour in terms of a power law behaviour of the right tail. In this section we try to make this heuristic workable. Most of the concepts discussed in this section had been developed by Karamata in the context of Tauberian theorems in classical analysis. Their importance to probability theory had been recognised by Feller. An authoritative account of all these and much more is given in [BGT]. We have borrowed heavily from [BGT], [Fe], [EKM]. We begin with a notion that has played a crucial role in the central limit problem

**Definition 4.4** Let $L(\cdot)$ be a positive measurable function on $(0, \infty)$. If

$$\lim_{x \to \infty} \frac{L(cx)}{L(x)} = 1, \text{ for all } c > 0 \tag{4.8}$$

then $L(\cdot)$ is said to be *slowly varying at* $\infty$. □

Obvious examples of slowly varying functions are: constants, logarithmic function, powers of logarithm $|\log x|^\alpha$, $\alpha > 0$, iterates of logarithm $\log(\log x)$, $\log(\log(\log x))$, etc.

The next result gives a useful class of slowly varying functions

**Proposition 4.5** *Let the function $L$ be of the form*

$$L(x) = c(x) \exp\left\{\int_{x_0}^{x} \frac{1}{t}\,\epsilon(t)\,dt\right\},\ x > x_0 \tag{4.9}$$

*where $x_0 > 0$, $\epsilon(t) \to 0$ as $t \to \infty$, $c(\cdot) > 0$, $c(x) \to c \in (0,\infty)$ as $x \to \infty$, and $c(\cdot), \epsilon(\cdot)$ are measurable. Then $L$ is slowly varying at $\infty$.*

**Proof:** It is enough to prove that

$$\lim_{x\to\infty} \int_{x}^{ax} \frac{1}{t}\,\epsilon(t)\,dt = 0, \text{ if } a > 1 \tag{4.10}$$

$$\lim_{x\to\infty} \int_{ax}^{x} \frac{1}{t}\,\epsilon(t)\,dt = 0, \text{ if } a < 1. \tag{4.11}$$

Let $a > 1$. For $\delta > 0$ there exists $t_0 > 0$ such that $|\epsilon(t)| < \delta$ for all $t \geq t_0$. Without loss of generality $t_0 < x < ax$. Then

$$\left|\int_{x}^{ax} \frac{1}{t}\,\epsilon(t)\,dt\right| \leq \delta \log a, \text{ for all } x > t_0$$

whence (4.10) follows. Similarly if $a < 1$, then

$$\left|\int_{ax}^{x} \frac{1}{t}\,\epsilon(t)\,dt\right| \leq \delta \log\left(\frac{1}{a}\right), \text{ for all } x > ax > t_0$$

and hence (4.11) is evident. □

**Proposition 4.6** *Let $L$ be a slowly varying function of the form (4.9). Then for any $a > 0$*

$$\lim_{x\to\infty} x^{-a}L(x) = 0, \quad \lim_{x\to\infty} x^{a}L(x) = \infty. \tag{4.12}$$

*That is, $L$ is "small" compared to any power function $x^a$.*

**Proof:** Let $a > 0$ be fixed. For any $\eta \in (0,a)$ there exists $t_0 > 0$ such that $-\eta < \epsilon(t) < \eta$ for all $t > t_0$. Then for $x > t_0$

$$\exp\left(\int_{x_0}^{x} \frac{1}{t}\,\epsilon(t)\,dt\right) \leq C_1 \exp\left(\eta \log\frac{x}{t_0}\right) = C_1\left(\frac{x}{t_0}\right)^{\eta}.$$

So, as $c(t) \to c > 0$ we now get $x^{-a}L(x) \leq C\ t_0^{-\eta}x^{-(a-\eta)}$ for $x > t_0$, where $C > 0$ is a constant. As $a > \eta$ the first assertion now follows.

In a similar fashion, for $x > t_0$ one can show that $x^a L(x) \geq C\ t_0^{\eta}x^{(a-\eta)}$; the second assertion follows now since $a > \eta$. □

Our next objective is to show that (4.9) gives all slowly varying functions. For this we need a very basic result due to Karamata.

**Theorem 4.7** *(Uniform convergence theorem for slowly varying functions): Let $L$ be a positive measurable slowly varying function on $(0, \infty)$. Then for any compact set $K \subset (0, \infty)$*

$$\lim_{x\to\infty} \frac{L(cx)}{L(x)} = 1 \text{ uniformly over } c \in K. \tag{4.13}$$

**Proof:** Write $h(x) \triangleq \log L(e^x)$, $x \in \mathbb{R}$. As $L$ is slowly varying note that

$$\lim_{x\to\infty} h(x+u) - h(x) = 0, \text{ for any } u \in \mathbb{R}. \tag{4.14}$$

We need to prove that the convergence in (4.14) is uniform over $u \in K'$, $K'$ any arbitrary compact set in $\mathbb{R}$. In this case it is clear that it is enough to consider $K' = [0, A]$, $A > 0$.

Let $\epsilon > 0$; without loss of generality $\epsilon < A$. For $x > 0$, set $I_x = [x, x + 2A]$, $E_x = \{t \in I_x : |h(t) - h(x)| \geq \frac{1}{2}\epsilon\}$, $E_x^* = \{t \in [0, 2A] : |h(x + t) - h(x)| \geq \frac{1}{2}\epsilon\}$. As $L$ and hence $h$ are measurable, it follows that $E_x$, $E_x^*$ are measurable sets. Clearly $E_x^* = E_x - x$ and hence $|E_x^*| = |E_x|$, where $|\cdot|$ denotes the Lebesgue measure. By (4.14) note that $\lim_{x\to\infty} I_{E_x^*}(t) = 0$ for each $t$. As $I_{E_x^*}(\cdot) \leq I_{[0,2A]}(\cdot)$ for all $x \geq 0$, by the dominated convergence theorem it follows that $\lim_{x\to\infty} |E_x^*| = 0$. So there is $x_0 > 0$ such that

$$|E_x^*| < \frac{1}{2}\ \epsilon \text{ for all } x \geq x_0. \tag{4.15}$$

Next let $u \in [0, A]$. Clearly $|I_{x+u} \cap I_x| = 2A - u \geq A$ for any $x$. For $x \geq x_0$, by (4.15)

$$\begin{aligned} |E_x \cup E_{x+u}| &\leq |E_x| + |E_{x+u}| = |E_x^*| + |E_{x+u}^*| \\ &< \epsilon < A. \end{aligned}$$

Hence for $u \in [0, A]$, $x \geq x_0$

$$|(I_x \cap I_{x+u}) \backslash (E_x \cup E_{x+u})| \geq 2A - u - \epsilon > 0.$$

In particular $(I_x \cap I_{x+u})\backslash(E_x \cup E_{x+u}) \neq \phi$. For any $t \in (I_x \cap I_{x+u})\backslash(E_x \cup E_{x+u})$, we see that $|h(t) - h(x)| < \frac{1}{2}\epsilon$ and $|h(t) - h(x+u)| < \frac{1}{2}\epsilon$. Therefore for all $u \in [0, A]$, $x \geq x_0$ we get $|h(x+u) - h(x)| < \epsilon$, completing the proof.

□

An immediate consequence is

**Lemma 4.8** *Let $L$ be a positive, measurable, slowly varying function on $(0, \infty)$. Let $h$ be defined by $h(x) \triangleq \log L(e^x)$, $x \in \mathbb{R}$. Then there is $A > 0$ such that $L$ is locally bounded, and hence locally integrable, in $[A, \infty)$. Also there is $\hat{A} > 0$ such that $h$ is locally bounded, and hence locally integrable in $[\hat{A}, \infty)$.*

**Proof:** Enough to prove the result for $h$. By the uniform convergence theorem, Theorem 4.7, one can find $\hat{A} > 0$ such that

$$|h(x+t) - h(x)| < 1 \text{ for all } x \geq \hat{A},\ t \in [0, 1].$$

Hence $|h(x)| \leq 1 + |h(\hat{A})|$ on $[\hat{A}, \hat{A}+1]$. By induction $|h(x)| \leq n + |h(\hat{A})|$ on $[\hat{A}, \hat{A}+n]$, for $n = 1, 2, \ldots$ The required conclusion is now immediate. □

**Theorem 4.9** *(Karamata representation theorem): Let $L$ be a positive measurable function on $(0, \infty)$. Then $L$ is slowly varying at $\infty$ if and only if it can be written in the form (4.9) as in Proposition 4.5.*

**Proof:** In view of Proposition 4.5, it is enough to prove the "only if" part. Let $h$ be as in the proof of Theorem 4.7 or Lemma 4.8. Note that getting $L$ in the form (4.9) is equivalent to writing $h$ in the form

$$h(x) = d(x) + \int_b^x e(v)\, dv,\ x > b \tag{4.16}$$

where $d$, $e$ are measurable functions on $\mathbb{R}$ such that $d(x) \to d \in \mathbb{R}$, $e(x) \to 0$ as $x \to \infty$, and $b \in \mathbb{R}$; to see this take $d(x) = \log c(e^x)$, $e(v) = \epsilon(e^v)$, $v \in \mathbb{R}$, $b = \log x_0$. Therefore to prove the result we need to show that (4.14) in the proof of Theorem 4.7 implies (4.16).

By Lemma 4.8, $h$ is locally bounded and locally integrable on $[A, \infty)$ for some $A > 0$. Choosing a fairly large $b > A > 0$ we can write

$$h(x) = \int_x^{x+1} (h(x) - h(t))\, dt + \int_b^x (h(t+1) - h(t))\, dt + \int_b^{b+1} h(t)\, dt,\ x \geq b. \tag{4.17}$$

The last term on the r.h.s. of (4.17) is a constant, $d$ say. Put $e(x) = h(x+1) - h(x)$; then $e(x) \to 0$ as $x \to \infty$ by (4.14). The first term on the r.h.s. of (4.17) $= \int_0^1 (h(x) - h(x+s))\, ds \to 0$ as $x \to \infty$ by the uniform convergence theorem (Theorem 4.7). Thus (4.16) holds with $d(x) = \int_b^{b+1} h(t)\, dt + \int_0^1 (h(x) - h(x+s))\, ds$, $e(x) = h(x+1) - h(x)$. □

Proposition 4.6 leads naturally to the following definition, again due to Karamata. This also plays an important role in the central limit problem; see Appendix A.2.

**Definition 4.10** (i) Let $f$ be a measurable function on $(0, \infty)$ such that $f(x) = x^{\delta} L(x)$, $x > 0$ where $L$ is a slowly varying function and $\delta \in \mathbb{R}$. Then $f$ is said to be *regularly varying with index* $\delta$.

(ii) A nonnegative random variable $X$, or equivalently its distribution function $F$, is said to be *regularly varying with tail index* $\alpha \geq 0$ if

$$P(X > x) = 1 - F(x) = x^{-\alpha} L(x),\ x > 0 \tag{4.18}$$

where $L$ is slowly varying. We shall denote $F \in \mathcal{R}_{-\alpha}$ or $X \in \mathcal{R}_{-\alpha}$.

Note that regular variation is one way to describe small deviations from exact power law behaviour.

Our next result is a *uniform convergence theorem for regularly varying functions with negative index.*

**Theorem 4.11** *Let $\alpha > 0$ and $f$ be a regularly varying function with index $(-\alpha)$. Then for any $A > 0$*

$$\lim_{x \to \infty} \frac{f(\lambda x)}{f(x)} = \frac{1}{\lambda^{\alpha}},\ \textit{uniformly over } \lambda \in [A, \infty). \tag{4.19}$$

**Proof:** By Karamata representation theorem (Theorem 4.9) note that

$$f(x) = x^{-\alpha}c(x)\ \exp\left(\int_a^x \frac{1}{t}\epsilon(t)\ dt\right),\ x > a \tag{4.20}$$

where $c(\cdot)$ is positive, $\lim c(x) = c \in (0,\infty)$, and $\lim \epsilon(x) = 0$; moreover as $L, c, \epsilon$ can be modified in any finite interval of the form $[0,\sigma]$, we may take $a = 1$.

Now let $\epsilon > 0$. Choose $a_1 > 1$ such that $4\lambda^{-\alpha/2} < \epsilon$ for all $\lambda \geq a_1$; (without loss of generality we may assume $\epsilon < 4$). Note that there is $x_0 > 0$ such that $\frac{1}{2} \leq \frac{c(x)}{c} \leq 2$, $\epsilon(x) < \frac{1}{2}\alpha$ for all $x \geq x_0$. For $\lambda \geq a_1 > 1$ note that $\lambda x \geq x_0$ if $x \geq x_0$. Therefore by (4.20), for $\lambda \geq a_1$, $x \geq x_0$ we get

$$\frac{f(\lambda x)}{f(x)} \leq 4\lambda^{-\alpha} \exp\left(\int_x^{\lambda x} \frac{1}{t}\ \frac{\alpha}{2}\ dt\right) = 4\lambda^{-\alpha/2}. \tag{4.21}$$

Note that $4\lambda^{-\alpha/2} > \lambda^{-\alpha}$ and $\lambda > 1$. Hence, as $f(\cdot) > 0$ and $4\lambda^{-\alpha/2} < \epsilon$, we get for $\lambda \geq a_1$, $x \geq x_0$ by (4.21)

$$\left|\frac{f(\lambda x)}{f(x)} - \frac{1}{\lambda^\alpha}\right| \leq 4\lambda^{-\alpha/2} < \epsilon,$$

proving uniform convergence in (4.19) over $[a_1,\infty)$.

Finally, by the uniform convergence theorem for slowly varying functions, uniform convergence in (4.19) over the compact set $[A, a_1] \subset (0,\infty)$ can be obtained. Combining the two we get the desired uniformity. □

**Note:** If $f$ is regularly varying with index $\delta > 0$, and $f$ is bounded on each interval of the form $(0,a]$, then it can be shown that $\frac{f(\lambda x)}{f(x)} \to \lambda^\delta$ as $x \to \infty$, uniformly over $\lambda \in (0,A]$ for any $A > 0$. See [BGT].

Our next objective is to investigate the asymptotics of truncated moments with respect to regularly varying functions. We have the following result.

**Theorem 4.12** *Let $f$ be a regularly varying function with index $\delta \in \mathbb{R}$. Let $K > 0$ be such that $f$ is locally bounded on $[K,\infty)$; by Lemma 4.8 such a $K$ exists. Denote $\overline{I}_p(x) = \int_K^x t^p f(t)\ dt$, $\underline{I}_p(x) = \int_x^\infty t^p f(t)\ dt$, $x \in (0,\infty)$, $p \in \mathbb{R}$ whenever the integrals exist. Then the following hold.*

*(i) If* $p > -(\delta+1)$ *then* $\overline{I}_p(x)$ *diverges as* $x \to \infty$, *and*

$$\lim_{x\to\infty} x^{p+1} f(x)/\overline{I}_p(x) = (p+\delta+1). \tag{4.22}$$

*In particular* $\overline{I}_p(x)$ *is regularly varying with index* $(p+\delta+1)$.

*(ii) If* $p = -(\delta+1)$ *then* $\overline{I}_p(x)$ *may converge or diverge, but*

$$\lim_{x\to\infty} \overline{I}_p(x)/(x^{p+1} f(x)) = \infty. \tag{4.23}$$

*Moreover* $\overline{I}_p(x)$ *is slowly varying.*

*(iii) If* $p < -(\delta+1)$ *then* $\overline{I}_p(x)$ *converges as* $x \to \infty$. *Also*

$$\lim_{x\to\infty} x^{p+1} f(x)/\underline{I}_p(x) = -(p+\delta+1), \tag{4.24}$$

*that is,* $\underline{I}_p(x)$ *is regularly varying with index* $(p+\delta+1)$.

*(iv) If* $p = -(\delta+1)$ *and* $\int_K^\infty t^p f(t)\, dt < \infty$, *then*

$$\lim_{x\to\infty} \underline{I}_p(x)/(x^{p+1} f(x)) = \infty \tag{4.25}$$

*and* $\underline{I}_p(x)$ *is slowly varying.* □

For proving the theorem it is enough to consider the case $\delta = 0$. In this case $f$ is slowly varying and we denote $L(x) = f(x)$, $x > 0$. The proof is in several steps. We begin with a lemma

**Lemma 4.13** *(Potter's bound): Let* $L$ *be slowly varying at* $\infty$. *Then for any* $A > 1$, $\eta > 0$ *there exists* $K = K(A, \eta)$ *such that*

$$\frac{L(y)}{L(x)} \le A \max\left\{\left(\frac{y}{x}\right)^\eta, \left(\frac{x}{y}\right)^\eta\right\}, \forall\ x,\ y \ge K. \tag{4.26}$$

**Proof:** By Karamata representation (4.9), for $y \ge x \ge x_0$

$$\frac{L(y)}{L(x)} = \frac{c(y)}{c(x)} \exp\left(\int_x^y \frac{1}{t}\epsilon(t)\, dt\right).$$

As $c(\cdot) \to c \in (0, \infty)$ and $\epsilon(\cdot) \to 0$, it is now clear that we can choose $K(A, \eta) > x_0$ such that

$$\frac{L(y)}{L(x)} \leq A \exp\left(\eta \log \frac{y}{x}\right) = A\left(\frac{y}{x}\right)^{\eta}$$

for all $y \geq x \geq K(A, \eta)$. The case $x \geq y \geq K(A, \eta)$ is similar. □

**Proof of Theorem 4.12:** We consider only the case when $L(x) = f(x)$ is slowly varying.

**Proof of (i):** Note that $p > -1$. Take $0 < \eta < p + 1$. With $K(2, \eta)$ as in Lemma 4.13 denote $K' = \max\{K, K(2, \eta)\}$. Clearly

$$\begin{aligned} \frac{1}{x^{p+1}L(x)} \int_{K'}^{x} t^p L(t)\, dt &= \frac{1}{L(x)} \int_{K'/x}^{1} s^p L(sx)\, ds \\ &= \int_0^1 s^p \frac{L(sx)}{L(x)} I_{[\frac{1}{x}K', 1]}(s)\, ds. \end{aligned}$$

By Lemma 4.13 the integrand on r.h.s. of the above is dominated by $2s^{p-\eta}$ for all $x \geq K'$. Also as $x \to \infty$ the integrand converges to $s^p$. So by the dominated convergence theorem

$$\lim_{x \to \infty} \frac{1}{x^{p+1}L(x)} \int_{K'}^{x} t^p L(t)\, dt = \frac{1}{p+1}. \tag{4.27}$$

As $L$ is bounded on $[K, K']$ it is clear that

$$\lim_{x \to \infty} \frac{1}{x^{p+1}L(x)} \int_{K}^{K'} t^p L(t)\, dt = 0. \tag{4.28}$$

Note that (4.27), (4.28) imply (4.22). Divergence of $\overline{I}_p(x)$ follows from (4.22) and Proposition 4.6.

**Proof of (ii):** Here $p = -1$; write $\tilde{L}(x) := \overline{I}_p(x) = \int_K^x \frac{1}{t} L(t)\, dt$. Let $0 < \theta < 1$ be arbitrary. Then for $x \geq K$ such that $\theta x \geq K$

$$\frac{\tilde{L}(x)}{L(x)} \geq \int_{\theta}^{1} \frac{1}{t} \frac{L(xt)}{L(x)}\, dt.$$

By the uniform convergence given in Theorem 4.7 it now follows that

$$\liminf_{x\to\infty} \frac{\tilde{L}(x)}{L(x)} \geq \log\left(\frac{1}{\theta}\right).$$

As $\theta > 0$ can be arbitrarily small it is now clear that $\tilde{L}(x)/L(x)$ goes to $\infty$ as $x \to \infty$, proving (4.23).

To complete the proof of (ii) we need to show that $\tilde{L}(\cdot)$ is slowly varying. As $L(\cdot) > 0$ note that $\tilde{L}$ is strictly increasing. So it is clear that $\tilde{L}$ is locally bounded and is bounded away from 0 on $[K_1, \infty)$ for any $K_1 \geq K$. Denote $\tilde{\epsilon}(x) = L(x)/\tilde{L}(x)$, $x \geq K_1$. Note that $\tilde{\epsilon}$ is well defined, and from the earlier part of the proof it follows that $\tilde{\epsilon}(x) \to 0$ as $x \to \infty$.

As $\tilde{L}(\cdot)$ is absolutely continuous, note that $\tilde{L}'(x) = \frac{1}{x}L(x)$ $= \frac{1}{x}\,\tilde{\epsilon}(x)\,\tilde{L}(x)$, a.e. on $[K_1, \infty)$ and hence $\tilde{L}'(x)/\tilde{L}(x) = \frac{1}{x}\,\tilde{\epsilon}(x)$, a.e. on $[K_1, \infty)$. Consequently $\log \tilde{L}$ is absolutely continuous and

$$\log \tilde{L}(x) - \log \tilde{L}(K_1) = \int_{K_1}^{x} \frac{1}{t}\,\tilde{\epsilon}(t)\,dt,\ x \geq K_1.$$

Therefore

$$\tilde{L}(x) = \tilde{L}(K_1) \exp\left(\int_{K_1}^{x} \frac{1}{t}\,\tilde{\epsilon}(t)\,dt\right),\ x \geq K_1. \tag{4.29}$$

By Karamata representation the required conclusion follows.

**Proof of (iii):** Note that $p < -1$. Convergence of $\overline{I}_p(x)$ can easily be established using Proposition 4.6.

Now set $\rho = \frac{1}{2}(p+1)$; clearly $\rho < 0$. Define $g(x) = x^{\rho}L(x)$. Then $g$ is a regularly varying function with index $\rho < 0$. Observe that

$$\begin{aligned}\frac{1}{x^{p+1}L(x)} \int_x^{\infty} t^p L(t)\,dt + \frac{1}{p+1} &= \int_1^{\infty} s^p\,\frac{L(sx)}{L(x)}\,ds - \int_1^{\infty} s^p ds \\ &= \int_1^{\infty} \left(\frac{g(sx)}{g(x)} - s^{\rho}\right) s^{\rho-1} ds.\end{aligned}$$

By Theorem 4.11 note that $\left(\frac{g(sx)}{g(x)} - s^{\rho}\right) \to 0$ as $x \to \infty$ uniformly over $s \in [1, \infty)$. And as $s^{\rho-1}$ is integrable over $[1, \infty)$, by the dominated convergence theorem (4.24) now follows.

**Proof of (iv):** Now $p = -1$, and $\int_K^\infty \frac{1}{t} L(t)\, dt < \infty$. Write $\hat{L}(x) := \underline{I}_p(x) = \int_x^\infty \frac{1}{t} L(t)\, dt$. We proceed as in the proof of (ii). Let $\theta > 1$ be arbitrary. Then for $x > K$

$$\frac{\hat{L}(x)}{L(x)} \geq \int_1^\theta \frac{1}{s} \frac{L(sx)}{L(x)}\, ds.$$

By Theorem 4.7 we now have

$$\liminf_{x\to\infty} \frac{\hat{L}(x)}{L(x)} \geq \log \theta.$$

As $\theta > 1$ can be arbitrarily large, (4.25) is now clear.

By assumption $\hat{L}(x) \in (0, \infty)$ for all $x$. As $\frac{1}{t} L(t) > 0$, it is clear that $\hat{L}$ is strictly decreasing. So on any bounded interval contained in $[K, \infty)$ it is clear that $\hat{L}$ is bounded and bounded away from 0. Set $\hat{\epsilon}(x) = -L(x)/\hat{L}(x)$, $x \geq K_1$ where $K_1 > K$. As in the proof of part (ii) we can establish that (analogous to (4.29))

$$\hat{L}(x) = \hat{L}(K_1) \exp\left( \int_{K_1}^x \frac{1}{t} \hat{\epsilon}(t)\, dt \right), \quad x \geq K_1 \tag{4.30}$$

with $\lim_{x\to\infty} \hat{\epsilon}(x) = 0$. Again by Karamata representation we conclude that $\hat{L}$ is slowly varying. This completes the proof. □

The next result interprets the above for random variables that are regularly varying with tail index $\alpha \geq 0$.

**Theorem 4.14** *Let $X$ be a nonnegative random variable with distribution function $F$. Suppose $X \in \mathcal{R}_{-\alpha}$. Then the following hold.*

*(i) If $\alpha \geq 0$ then*

$$\begin{aligned} E(X^p) &< \infty, \quad \text{for } 0 \leq p < \alpha \\ &= \infty, \quad \text{for } p > \alpha. \end{aligned} \tag{4.31}$$

*(ii) Let $\alpha > 1$, and $F_I$ denote the corresponding integrated tail distribution given by (3.34); that is,*

$$dF_I(t) = I_{(0,\infty)}(t) \frac{1}{E(X)} (1 - F(t))\, dt. \tag{4.32}$$

*Then $F_I \in \mathcal{R}_{-(\alpha-1)}$.*

**Proof:** (i) For $p > 0$ note that

$$E(X^p) = \int_0^\infty ps^{p-1}(1 - F(s))\, ds \tag{4.33}$$

in the sense that if one side is finite then so is the other and the equality of the two. The first part of the theorem is now easily obtained by (i), (iii) of Theorem 4.12.

(ii) As $\alpha > 1$ note that $E(X) < \infty$ and hence $F_I$ makes sense. Observe that

$$1 - F_I(x) = \frac{1}{E(X)} \int_x^\infty (1 - F(t))\, dt,\ x > 0$$

which is a regularly varying function with index $(0+(-\alpha)+1) = -(\alpha-1)$ by part (iii) of Theorem 4.12. This completes the proof. □

**Example 4.15** *Pareto distribution.* Let $\alpha > 0$, $\kappa > 0$. Let the distribution function $F$, supported on $[0, \infty)$, be such that its right tail is given by

$$\overline{F}(x) = 1 - F(x) = \frac{\kappa^\alpha}{(\kappa + x)^\alpha},\ x \geq 0. \tag{4.34}$$

This distribution is called the *Pareto* $(\alpha, \kappa)$ *distribution.* The parameter $\alpha > 0$ is called the *tail parameter*, and $\kappa > 0$ is called the *scale parameter.* As the tail parameter is the important one, by usual abuse of notation we will suppress the scale parameter, and may just write *Pareto* $(\alpha)$ distribution to denote (4.34). It is clear that Pareto $(\alpha) \in \mathcal{R}_{-\alpha}$, $\alpha > 0$. By Exercise 4.2.12 it follows that Pareto $(\alpha)$ is heavy-tailed, that is, its moment generating function does not exist in any neighbourhood of 0. By Exercise 4.2.13, if $\alpha > 1$, $F_I$ has Pareto $(\alpha - 1)$ distribution and $F_I \in \mathcal{R}_{-(\alpha-1)}$.

From the discussion in Appendix A.2, for $0 < \alpha < 2$ note that Pareto $(\alpha)$ belongs to the domain of normal attraction of an $\alpha$-stable distribution.

Pareto distribution is perhaps the most important heavy tailed distribution. It is one of the few distributions that has been found to give a good fit to observed data, especially in actuarial context. In economics,

it has been conventionally used to describe income distribution. Because of its very simple description (4.34) it is mathematically amenable. This, together with the major role played in the central limit problem, perhaps explain its versatility. □

The next result shows that $\mathcal{R}_{-\alpha}$ is closed under convolutions for any $\alpha \geq 0$.

**Theorem 4.16** *Let $\alpha \geq 0$ and $X_1, X_2 \in \mathcal{R}_{-\alpha}$ be independent; suppose for $i = 1, 2$*

$$\overline{F}_i(x) \triangleq P(X_i > x) = x^{-\alpha} L_i(x), \quad x > 0 \tag{4.35}$$

*for (possibly different) slowly varying functions $L_i$. Then*

$$\begin{aligned} & P(X_1 + X_2 > x) \\ = \; & x^{-\alpha}[L_1(x) + L_2(x)](1 + o(1)), \quad \text{as } x \to \infty. \end{aligned} \tag{4.36}$$

*In particular $X_1 + X_2 \in \mathcal{R}_{-\alpha}$.*

**Proof:** Let $G$ denote the distribution function of $X_1 + X_2$. As $X_1, X_2$ are nonnegative, for any $x \geq 0$ note that

$$\{X_1 + X_2 > x\} \supseteq \{X_1 > x\} \cup \{X_2 > x\}.$$

Therefore by independence of $X_1, X_2$,

$$\begin{aligned} & \overline{G}(x) = P(X_1 + X_2 > x) \\ \geq \; & P(X_1 > x) + P(X_2 > x) - P(X_1 > x, X_2 > x) \\ = \; & \overline{F}_1(x) + \overline{F}_2(x) - \overline{F}_1(x) \cdot \overline{F}_2(x) \\ = \; & [\overline{F}_1(x) + \overline{F}_2(x)] \left(1 - \frac{\overline{F}_1(x) \cdot \overline{F}_2(x)}{\overline{F}_1(x) + \overline{F}_2(x)}\right). \end{aligned}$$

Clearly $\left|\frac{\overline{F}_1(x)\ \overline{F}_2(x)}{\overline{F}_1(x)+\overline{F}_2(x)}\right| \leq \overline{F}_1(x) = o(1)$, as $x \to \infty$ and hence

$$\overline{G}(x) \geq (\overline{F}_1(x) + \overline{F}_2(x))(1 + o(1)), \quad \text{as } x \to \infty. \tag{4.37}$$

Next let $0 < \delta < \frac{1}{2}$. It is easily verified that $\{X_1 + X_2 > x\} \subseteq \{X_1 > (1-\delta)x\} \cup \{X_2 > (1-\delta)x\} \cup \{X_1 > \delta x, X_2 > \delta x\}$, and hence by independence

$$\begin{aligned} \overline{G}(x) \; \leq \; & \overline{F}_1((1-\delta)x) + \overline{F}_2((1-\delta)x) + \overline{F}_1(\delta x) \cdot \overline{F}_2(\delta x) \\ = \; & [\overline{F}_1((1-\delta)x) + \overline{F}_2((1-\delta)x)] \\ & \times \left(1 + \frac{\overline{F}_1(\delta x)\ \overline{F}_2(\delta x)}{\overline{F}_1((1-\delta)x) + \overline{F}_2((1-\delta)x)}\right). \end{aligned}$$

Note that by (4.35)

$$\begin{aligned}
&\left|\frac{\overline{F}_1(\delta x)\ \overline{F}_2(\delta x)}{\overline{F}_1((1-\delta)x)+\overline{F}_2((1-\delta)x)}\right| \\
= \ &\overline{F}_1(\delta x)\ \frac{\overline{F}_2(\delta x)}{\overline{F}_1(\delta x)+\overline{F}_2(\delta x)} \cdot \frac{\overline{F}_1(\delta x)+\overline{F}_2(\delta x)}{\overline{F}_1((1-\delta)x)+\overline{F}_2((1-\delta)x)} \\
\leq \ &\overline{F}_1(\delta x)\left(\frac{\delta}{1-\delta}\right)^{-\alpha} \frac{(L_1+L_2)(\delta x)}{(L_1+L_2)((1-\delta)x)} = o(1), \text{ as } x \to \infty
\end{aligned}$$

because $L_1 + L_2$ is slowly varying. Therefore

$$\overline{G}(x) \leq [\overline{F}_1((1-\delta)x) + \overline{F}_2((1-\delta)x)](1+o(1)), \tag{4.38}$$

as $x \to \infty$. By (4.35), as $L_1 + L_2$ is slowly varying, it is clear that

$$\lim_{x\to\infty} \frac{\overline{F}_1((1-\delta)x)+\overline{F}_2((1-\delta)x)}{\overline{F}_1(x)+\overline{F}_2(x)} = (1-\delta)^{-\alpha}. \tag{4.39}$$

By (4.37), (4.38), (4.39) it now follows that

$$\begin{aligned}
1 \ &\leq \ \liminf_{x\to\infty} \frac{\overline{G}(x)}{\overline{F}_1(x)+\overline{F}_2(x)} \leq \limsup_{x\to\infty} \frac{\overline{G}(x)}{\overline{F}_1(x)+\overline{F}_2(x)} \\
&\leq \ (1-\delta)^{-\alpha}
\end{aligned}$$

for any $0 < \delta < \frac{1}{2}$. Letting $\delta \downarrow 0$ in the above we get (4.36) completing the proof. □

**Corollary 4.17** Let the i.i.d. random variables $X_1, X_2, \ldots, X_n \in \mathcal{R}_{-\alpha}$ where $\alpha \geq 0$; let $F$ denote the common distribution function. Let $S_n = X_1 + \ldots + X_n$. Then $S_n \in \mathcal{R}_{-\alpha}$ and

$$P(S_n > x) = n\ \overline{F}(x)(1+o(1)), \text{ as } x \to \infty. \tag{4.40}$$

## 4.3 Subexponential distributions

Let $X_1, X_2, \ldots$ be an i.i.d. sequence. Let $F$ denote the common distribution function. Denote by $S_n = X_1 + \ldots + X_n$, $M_n = \max\{X_1, \ldots, X_n\}$, $n \geq 2$, respectively the *partial sum* and the *partial maximum*. For $n \geq 2$

note that

$$\begin{aligned} P(M_n > x) &= 1 - P(X_1 \leq x, X_2 \leq x, \ldots, X_n \leq x) \\ &= 1 - (F(x))^n = 1 - (1 - \overline{F}(x))^n \\ &= 1 - \sum_{k=0}^{n} \binom{n}{k} (-1)^k (\overline{F}(x))^k \\ &= n\,\overline{F}(x)\,[1 + o(1)], \text{ as } x \to \infty. \end{aligned} \tag{4.41}$$

Comparing (4.40) and (4.41) we see that if $F \in \mathcal{R}_{-\alpha}$, $\alpha \geq 0$ then for $n = 2, 3, \ldots$

$$\lim_{t \to \infty} \frac{P(S_n > t)}{P(M_n > t)} = 1. \tag{4.42}$$

Thus, for regularly varying distributions, the tail behaviour of the partial sum $S_n$ is essentially determined by the tail behaviour of the partial maximum $M_n$.

In situations involving "dangerous risks", it has been observed by actuaries that whenever the total amount is very large, a very few of the largest claims (in some cases may be even a single largest claim) make up almost the entire bulk of the total claim. This is precisely the scenario envisaged in the above discussion. So the following definition is the next natural step.

**Definition 4.18** Let $F$ be a distribution function supported on $(0, \infty)$. Let $\{X_i\}$ be a sequence of i.i.d. random variables with distribution function $F$. For $n \geq 2$ let $S_n = X_1 + \ldots + X_n$, $M_n = \max\{X_1, \ldots, X_n\}$. If for all $n \geq 2$

$$P(S_n > t) = P(M_n > t)(1 + o(1)), \text{ as } t \to \infty \tag{4.43}$$

then $F$ is said to be *subexponential.* The class of subexponential distributions is denoted by $\mathcal{S}$.

As the above notion conforms to the intuitive idea of a dangerous risk, it is more or less synonymous with "heavy tailedness" in insurance mathematics.

From (4.40) - (4.42) it is clear that $\mathcal{R}_{-\alpha} \subset \mathcal{S}$ for any $\alpha > 0$. In particular Pareto distributions are subexponential.

Let $\{X_i\}, F, S_n, M_n$ be as in the above definition. We know that the $n$-fold convolution $F^{*(n)}$ is the distribution function of $S_n$. Therefore (4.41) and (4.43) imply that $F \in \mathcal{S} \Leftrightarrow$ for $n = 2, 3, \ldots$

$$\lim_{t\to\infty} \frac{\overline{F^{*(n)}}(t)}{\overline{F}(t)} = \lim_{t\to\infty} \frac{1 - F^{*(n)}(t)}{1 - F(t)} = n. \tag{4.44}$$

We want to establish that subexponentiality is equivalent to (4.44) holding for $n = 2$. For this we need some preliminary results, which are also interesting.

It is clear that any $F$ satisfying (4.44) for $n = 2$ cannot have compact support; see Exercise 4.3.3.

**Theorem 4.19** *(a) Let $F$ be a distribution supported on $(0, \infty)$ such that (4.44) holds for $n = 2$. Then for any $y \geq 0$*

$$\lim_{x\to\infty} \frac{\overline{F}(x - y)}{\overline{F}(x)} = 1. \tag{4.45}$$

*Moreover the convergence in (4.45) is uniform over $y \in K$, for any compact set $K$.*

*(b) Let $F$ be a distribution supported on $(0, \infty)$ such that (4.45) holds for any $y \geq 0$. Then*
*(i) $L(z) \triangleq \overline{F}(\log z)$ is slowly varying at $\infty$;*
*(ii) for any $a > 0$*

$$\lim_{x\to\infty} e^{ax}\overline{F}(x) = \infty. \tag{4.46}$$

*(iii) $\hat{m}_F(s) = \infty$ for all $s > 0$; so moment generating function does not exist in any neighbourhood of 0.*

*In particular, any subexponential distribution $F$ is heavy tailed, and $\overline{F}(\log x)$ is slowly varying.*

**Proof:** (a) For *any* distribution supported on $(0, \infty)$ note that

$$\overline{F^{*(2)}}(x) = 1 - \int_0^x F(x - t)\, dF(t) = \overline{F}(x) + \int_0^x \overline{F}(x - t)\, dF(t)$$

and hence

$$\frac{\overline{F^{*(2)}}(x)}{\overline{F}(x)} = 1 + \int_0^x \frac{\overline{F}(x - t)}{\overline{F}(x)}\, dF(t). \tag{4.47}$$

Therefore for any $0 \leq y \leq x$ we get

$$\begin{aligned}\frac{\overline{F^{*}(2)}(x)}{\overline{F}(x)} &= 1 + \int_0^y \frac{\overline{F}(x-t)}{\overline{F}(x)}\, dF(t) + \int_y^x \frac{\overline{F}(x-t)}{\overline{F}(x)}\, dF(t) \\ &\geq 1 + F(y) + \frac{\overline{F}(x-y)}{\overline{F}(x)}\,(F(x) - F(y)).\end{aligned}$$

Let $y \geq 0$ be fixed. Under our hypothesis, for all sufficiently large $x$ note that $F(x) - F(y) > 0$. See Exercise 4.3.3 below. So from the above

$$1 \leq \frac{\overline{F}(x-y)}{\overline{F}(x)} \leq \left(\frac{\overline{F^{*(2)}}(x)}{\overline{F}(x)} - 1 - F(y)\right) \frac{1}{(F(x) - F(y))}. \tag{4.48}$$

By (4.44) for $n = 2$ we see that r.h.s. of (4.48) converges to 1 as $x \to \infty$. This proves (4.45). Uniform convergence over compact sets is clear by monotonicity of $\overline{F}$.

(b) Assertion (i) follows easily by part (a). Consequently $z^a L(z) \to \infty$ by Proposition 4.6. Putting $z = e^x$ we get (4.46). Now (4.4) and (4.46) imply (iii). □

**Lemma 4.20** *Let $F$ be a distribution supported on $(0, \infty)$ such that (4.44) holds for $n = 2$. Let $G$ be a distribution supported on $(0, \infty)$ such that*

$$\lim_{x \to \infty} \frac{\overline{G}(x)}{\overline{F}(x)} = a \in (0, \infty). \tag{4.49}$$

*Then*

$$\lim_{x \to \infty} \frac{\overline{F * G}(x)}{\overline{F}(x)} = 1 + a. \tag{4.50}$$

**Proof:** Similar to the derivation of (4.47) one can get (see Exercise 4.3.4)

$$\frac{\overline{F * G}(x)}{\overline{F}(x)} = 1 + \int_0^x \frac{\overline{G}(x-t)}{\overline{F}(x)}\, dF(t). \tag{4.51}$$

So it is enough to prove

$$\lim_{x \to \infty} \frac{1}{\overline{F}(x)} \int_0^x \overline{G}(x-t)\, dF(t) = a. \tag{4.52}$$

Let $\epsilon > 0$ be arbitrary. By (4.49) there exists $x_0$ such that $\overline{G}(x) \leq (a+\epsilon)\ \overline{F}(x)$ for all $x \geq x_0$. Therefore

$$\begin{aligned}
&\frac{1}{\overline{F}(x)} \int_0^x G(x-t)\ dF(t) \\
= &\frac{1}{\overline{F}(x)} \int_0^{x-x_0} \overline{G}(x-t)\ dF(t) + \frac{1}{\overline{F}(x)} \int_{x-x_0}^x \overline{G}(x-t)\ dF(t) \\
\leq &(a+\epsilon)\frac{1}{\overline{F}(x)} \int_0^{x-x_0} \overline{F}(x-t)\ dF(t) + \frac{F(x)-F(x-x_0)}{\overline{F}(x)} \\
\leq &(a+\epsilon)\frac{1}{\overline{F}(x)} \int_0^x \overline{F}(x-t)\ dF(t) + \frac{\overline{F}(x-x_0)-\overline{F}(x)}{\overline{F}(x)}.
\end{aligned}$$

By Theorem 4.19(a) and Exercise 4.3.2 we see that the r.h.s. of the above converges to $(a+\epsilon)$ as $x \to \infty$. Hence

$$\limsup_{x\to\infty} \frac{1}{\overline{F}(x)} \int_0^x \overline{G}(x-t)\ dF(t) \leq a,$$

as $\epsilon > 0$ is arbitrary. In an entirely analogous manner it can be shown that

$$\liminf_{x\to\infty} \frac{1}{\overline{F}(x)} \int_0^x \overline{G}(x-t)\ dF(t) \geq a.$$

Thus (4.52) and hence the lemma is proved. □

**Theorem 4.21** *Let $F$ be a distribution supported on $(0,\infty)$. Then $F$ is subexponential $\Leftrightarrow$ (4.44) holds for $n = 2$.*

**Proof:** It is enough to prove that if (4.44) holds for $n = 2$ then it holds for all $n \geq 2$. The required implication is now proved by induction on $n$, using Lemma 4.20. □

We next present a very useful result attributed to Kesten.

**Lemma 4.22** *(Kesten): Let $F \in \mathcal{S}$. Then for $\epsilon > 0$ there exists a constant $K \in (0, \infty)$ such that*

$$\frac{\overline{F^{*(n)}}(t)}{\overline{F}(t)} \leq K(1+\epsilon)^n, \ t \geq 0, \ n \geq 2. \tag{4.53}$$

**Proof:** Put $\alpha_n = \sup_{t \geq 0} \overline{F^{*(n)}}(t)/\overline{F}(t)$; by subexponentiality $\alpha_n < \infty$ for all $n$. By (4.51) applied to $G = F^{*(n)}$ we see that for any $a < \infty$,

$$\begin{aligned}
\alpha_{n+1} &= 1 + \sup_{x \geq 0} \int_0^x \frac{\overline{F^{*(n)}}(x-t)}{\overline{F}(x)} \, dF(t) \\
&\leq 1 + \sup_{0 \leq x \leq a} \int_0^x \frac{\overline{F^{*(n)}}(x-t)}{\overline{F}(x)} \, dF(t) \\
&\quad + \sup_{x \geq a} \int_0^x \frac{\overline{F^{*(n)}}(x-t)}{\overline{F}(x)} \, dF(t) \\
&\leq 1 + \frac{1}{\overline{F}(a)} + \alpha_n \sup_{x \geq a} \int_0^x \frac{\overline{F}(x-t)}{\overline{F}(x)} \, dF(t) \\
&= 1 + \frac{1}{\overline{F}(a)} + \alpha_n \sup_{x \geq a} \frac{\overline{F^{*(2)}}(x) - \overline{F}(x)}{\overline{F}(x)}
\end{aligned}$$

by (4.47). As $F$ is subexponential for $\epsilon > 0$ we can choose $a > 0$ such that

$$\frac{\overline{F^{*(2)}}(x) - \overline{F}(x)}{\overline{F}(x)} \leq (1+\epsilon) \text{ for all } x \geq a.$$

So one can get $a > 0$ such that

$$\alpha_{n+1} \leq 1 + \frac{1}{\overline{F}(a)} + \alpha_n(1+\epsilon).$$

As $a > 0$ is independent of $n$, iterating the above, we get

$$\alpha_n \leq \left(1 + \frac{1}{\overline{F}(a)}\right) \frac{1}{\epsilon} \, (1+\epsilon)^n,$$

completing the proof. □

An immediate consequence of Kesten's lemma and the dominated convergence theorem is

**Proposition 4.23** *Let* $p_i \geq 0$ *for* $i = 0, 1, 2, \ldots$ *and* $\sum_{i=0}^{\infty} p_i = 1$; *so* $\{p_i\}$ *defines a probability distribution on* $\{0, 1, 2, \ldots\}$. *Let* $F \in \mathcal{S}$. *Define* $G(x) = \sum_{k=0}^{\infty} p_k F^{*(k)}(x)$. *If* $\sum_{n=0}^{\infty} p_n (1+\epsilon)^n < \infty$ *for some* $\epsilon > 0$, *then*

$$\lim_{x \to \infty} \frac{\overline{G}(x)}{\overline{F}(x)} = \sum_{k=0}^{\infty} k p_k.$$

□

We say that two distributions $F$, $G$ on $(0, \infty)$ are *tail equivalent* (denoted $F \sim^t G$) if there exists $0 < a < \infty$ such that

$$\lim_{x \to \infty} \frac{\overline{G}(x)}{\overline{F}(x)} = a. \tag{4.54}$$

This is an equivalence relation on $\mathcal{S}$ as shown below.

**Lemma 4.24** *Let* $F \sim^t G$. *Then* $F \in \mathcal{S} \Leftrightarrow G \in \mathcal{S}$.

**Proof:** Let $F \in \mathcal{S}$ and (4.54) hold. By analogue of (4.47) or (4.51), and Exercise 4.3.1 it is enough to prove that

$$\limsup_{x \to \infty} \int_0^x \frac{\overline{G}(x-t)}{\overline{G}(x)}\, dG(t) \leq 1. \tag{4.55}$$

Let $x_0 > 0$, to be chosen suitably later, be fixed. Write

$$\begin{aligned} & \int_0^x \frac{\overline{G}(x-t)}{\overline{G}(x)}\, dG(t) \\ = & \int_0^{x-x_0} \frac{\overline{G}(x-t)}{\overline{G}(x)}\, dG(t) + \int_{x-x_0}^{x} \frac{\overline{G}(x-t)}{\overline{G}(x)}\, dG(t) \\ = & \; I_1(x) + I_2(x), \text{ say, for } x > x_0. \end{aligned}$$

Clearly

$$\begin{aligned} 0 \leq & \; I_2(x) \leq \frac{G(x) - G(x-x_0)}{\overline{G}(x)} = \frac{\overline{G}(x-x_0)}{\overline{G}(x)} - 1 \\ = & \; \frac{\overline{G}(x-x_0)}{\overline{F}(x-x_0)} \frac{\overline{F}(x-x_0)}{\overline{F}(x)} \frac{\overline{F}(x)}{\overline{G}(x)} - 1. \end{aligned}$$

By Theorem 4.19(a) and (4.54) it now follows that $I_2(x) \to 0$ as $x \to \infty$.

Now let $0 < \epsilon < a$ be arbitrary, where $a$ is as in (4.54). Choose $x_0 > 0$ such that $a - \epsilon \le \overline{G}(x)/\overline{F}(x) \le a + \epsilon$ for all $x \ge x_0$. Then

$$
\begin{aligned}
I_1(x) &\le \frac{a+\epsilon}{a-\epsilon} \int_0^{x-x_0} \frac{\overline{F}(x-t)}{\overline{F}(x)}\, dG(t) \le \frac{a+\epsilon}{a-\epsilon} \int_0^{x} \frac{\overline{F}(x-t)}{\overline{F}(x)}\, dG(t) \\
&= \frac{a+\epsilon}{a-\epsilon} \frac{G(x) - \int_0^x G(x-t)\, dF(t)}{\overline{F}(x)} \\
&\quad \text{(because } F * G = G * F\text{)} \\
&= \frac{a+\epsilon}{a-\epsilon} \frac{\overline{F}(x) - \overline{G}(x) + \int_0^x \overline{G}(x-t)\, dF(t)}{\overline{F}(x)} \\
&= \frac{a+\epsilon}{a-\epsilon} \frac{1}{\overline{F}(x)} \left\{ \begin{array}{l} \overline{F}(x) - \overline{G}(x) + \int_0^{x-x_0} \overline{G}(x-t)\, dF(t) \\ + \int_{x-x_0}^{x} \overline{G}(x-t)\, dF(t) \end{array} \right\} \\
&\le \frac{a+\epsilon}{a-\epsilon} \frac{1}{\overline{F}(x)} \left\{ \begin{array}{l} \overline{F}(x) - (a-\epsilon)\overline{F}(x) \\ +(a+\epsilon) \int_0^{x-x_0} \overline{F}(x-t)\, dF(t) \\ +\overline{F}(x-x_0) - \overline{F}(x) \end{array} \right\} \\
&\le \frac{a+\epsilon}{a-\epsilon} \left\{ \begin{array}{l} -(a-\epsilon) + (a+\epsilon) \int_0^{x} \frac{\overline{F}(x-t)}{\overline{F}(x)}\, dF(t) \\ +\frac{\overline{F}(x-x_0)}{\overline{F}(x)} \end{array} \right\}.
\end{aligned}
$$

By Theorem 4.19(a) and Exercise 4.3.2 it now follows that

$$\limsup_{x\to\infty} I_1(x) \le \frac{a+\epsilon}{a-\epsilon}(1+2\epsilon).$$

As $\epsilon > 0$ can be arbitrarily small, the required conclusion (4.55) follows. □

An immediate consequence is

**Theorem 4.25** *Let the notation be as in Proposition 4.23. Assume that* $p_0 < 1$ *and* $\sum_n (1+\epsilon)^n p_n < \infty$ *for some* $\epsilon > 0$. *If* $F \in \mathcal{S}$ *then the convex combination* $G \in \mathcal{S}$.

**Proof:** By hypothesis note that $0 < \sum_{k=0}^{\infty} k\, p_k < \infty$. Apply Lemma 4.24. □

A converse of the above result is also true. See [EKM].

For a distribution $F$ supported on $(0,\infty)$ and with expectation $\mu \in (0,\infty)$ let $F_I$ denote the corresponding integrated tail distribution. An important question, that will play a major role in risk theory, is: When is $F_I$ subexponential? It is perhaps still not known if $F \in \mathcal{S}$ implies $F_I \in \mathcal{S}$. A related question is: Are there sufficient conditions that ensure that both $F$ and $F_I$ are subexponential? As an example of the latter, by Theorem 4.14 we know that if $F \in \mathcal{R}_{-\alpha}$, $\alpha > 1$ then $F_I \in \mathcal{R}_{-(\alpha-1)}$ and hence both are subexponential. In particular if $F$ is Pareto $(\alpha)$, $\alpha > 1$ then both $F$ and $F_I$ are subexponential.

Recall that we consider only distribution functions $F$ on $\mathbb{R}$ such that $F(0) = 0$, $F(x) < 1$ for all $x$. Let $F$ have finite expectation $\mu$. Rephrasing the first question in the preceding paragraph: When can we have

$$\lim_{x\to\infty} \frac{1}{\overline{F}_I(x)}\, \overline{F_I^{*(2)}}(x) = 2? \tag{4.56}$$

We know that $F_I$ has probability density function $\frac{1}{\mu}\overline{F}(x)$. Therefore $F_I^{*(2)} = F_I * F_I$ has probability density function $\left(\frac{1}{\mu}\,\overline{F}\right) * \left(\frac{1}{\mu}\,\overline{F}\right)$ and hence

$$\overline{F_I^{*(2)}}(x) = \frac{1}{\mu^2} \int_x^{\infty} \left( \int_0^t \overline{F}(t-y)\, \overline{F}(y)\, dy \right) dt.$$

Clearly (4.56) will hold if for $\epsilon > 0$ we can find $x_0 > 0$ such that

$$2(1-\epsilon)\, \overline{F}_I(x) \le F_I^{*(2)}(x) \le 2(1+\epsilon)\, \overline{F}_I(x),\ x \ge x_0;$$

that is, for $\epsilon > 0$ there is $x_0 > 0$ such that

$$\begin{aligned} 2(1-\epsilon) \int_x^{\infty} \frac{1}{\mu}\, \overline{F}(t)\, dt &\le \int_x^{\infty} \frac{1}{\mu^2} \left( \int_0^t \overline{F}(t-y)\, \overline{F}(y)\, dy \right) dt \\ &\le 2(1+\epsilon) \int_x^{\infty} \frac{1}{\mu}\, \overline{F}(t)\, dt, \text{ for all } x \ge x_0. \end{aligned}$$

The above inequalities will hold if the corresponding inequalities hold for the integrands w.r.t. the $dt$ integral. In other words, (4.56) would follow, if for $\epsilon > 0$ one can find $x_0 > 0$ such that

$$\begin{aligned} 2\mu(1-\epsilon)\,\overline{F}(t) &\leq \int_0^t \overline{F}(t-y)\,\overline{F}(y)\,dy \\ &\leq 2\mu(1+\epsilon)\,\overline{F}(t), \text{ for } t \geq x_0. \end{aligned}$$

This leads to the following definition

**Definition 4.26** $F$ is said to belong to the class $\mathcal{S}^*$ if $F$ has finite expectation $\mu$ and

$$\lim_{x\to\infty} \int_0^x \frac{\overline{F}(x-y)}{\overline{F}(x)}\,\overline{F}(y)\,dy = 2\mu. \tag{4.57}$$

The discussion preceding the definition can now be summarised as

**Proposition 4.27** *If $F \in \mathcal{S}^*$ then $F_I \in \mathcal{S}$.* □

We will prove later that $F \in \mathcal{S}^* \Rightarrow F \in \mathcal{S},\ F_I \in \mathcal{S}$; hence $\mathcal{S}^*$ is an important class of heavy tailed distribution amenable to analysis. For this we need the following auxiliary class.

**Definition 4.28** $F$ is said to belong to the class $\mathcal{L}$ if (4.45) holds for all $y \geq 0$. In such a case $F$ is said to be a *long-tailed distribution.*

By Theorem 4.19 we know that $\mathcal{S} \subset \mathcal{L}$ and that any $F \in \mathcal{L}$ is heavy tailed.

**Lemma 4.29** $\mathcal{S}^* \subset \mathcal{L}$.

**Proof:** Let $F \in \mathcal{S}^*$ with finite expectation $\mu$. Let $v > 0$ be fixed.

For $x > 2v$ note that

$$\begin{aligned}
&\int_0^{x/2} \frac{\overline{F}(x-y)}{\overline{F}(x)}\, \overline{F}(y)\, dy \\
&= \int_0^{v} \frac{\overline{F}(x-y)}{\overline{F}(x)}\, \overline{F}(y)\, dy + \int_v^{x/2} \frac{\overline{F}(x-y)}{\overline{F}(x)}\, \overline{F}(y)\, dy \\
&\geq \int_0^{v} \overline{F}(y)\, dy + \frac{\overline{F}(x-v)}{\overline{F}(x)} \int_v^{x/2} \overline{F}(y)\, dy \\
&= \mu \left\{ F_I(v) + \frac{\overline{F}(x-v)}{\overline{F}(x)} \left[ F_I\left(\frac{x}{2}\right) - F_I(v) \right] \right\}.
\end{aligned}$$

Therefore (cf. see (4.48))

$$\begin{aligned}
1 &\leq \frac{\overline{F}(x-v)}{\overline{F}(x)} \\
&\leq \frac{1}{[F_I(\frac{x}{2}) - F_I(v)]} \left\{ \frac{1}{\mu} \int_0^{x/2} \frac{\overline{F}(x-y)}{\overline{F}(x)}\, \overline{F}(y)\, dy - F_I(v) \right\}.
\end{aligned}$$

Since $F \in \mathcal{S}^*$, by Exercise 4.3.9(b), r.h.s. of the above converges to 1 as $x \to \infty$. So $F \in \mathcal{L}$. □

The next result may be compared with Lemma 4.24.

**Lemma 4.30** *Let $F, G \in \mathcal{L}$. Suppose there exist $0 < a_1,\ a_2 < \infty$ such that*

$$a_1 \leq \frac{\overline{G}(x)}{\overline{F}(x)} \leq a_2, \text{ for all } x \geq 0.$$

*Then $F \in \mathcal{S}^* \Leftrightarrow G \in \mathcal{S}^*$. In particular, if $F, G \in \mathcal{L}$ and $F \sim^t G$, then $F \in \mathcal{S}^* \Leftrightarrow G \in \mathcal{S}^*$.*

**Proof:** Suppose $F \in \mathcal{S}^*$. Then $F$ has expectation $\mu_F \in (0, \infty)$. Hence our hypothesis implies that $G$ has expectation $\mu_G \in (0, \infty)$. Next

for fixed $v > 0$, and any $x > 2v$,

$$\int_0^{x/2} \frac{\overline{G}(x-t)}{\overline{G}(x)}\,\overline{G}(t)\,dt$$

$$= \int_0^{v} \frac{\overline{G}(x-t)}{\overline{G}(x)}\,\overline{G}(t)\,dt + \int_v^{x/2} \frac{\overline{G}(x-t)}{\overline{G}(x)}\,\overline{G}(t)\,dt.$$

As $G \in \mathcal{L}$, by Exercise 4.3.13,

$$\lim_{x\to\infty} \int_0^{v} \frac{\overline{G}(x-t)}{\overline{G}(x)}\,\overline{G}(t)\,dt = \int_0^{v} \overline{G}(t)\,dt.$$

Therefore, in view of Exercise 4.3.9(b), it would follow that $G \in \mathcal{S}^*$ once we prove

$$\lim_{v\to\infty} \limsup_{x\to\infty} \int_v^{x/2} \frac{\overline{G}(x-t)}{\overline{G}(x)}\,\overline{G}(t)\,dt = 0,$$

which in turn would follow if we have

$$\lim_{v\to\infty} \limsup_{x\to\infty} \int_v^{x/2} \frac{\overline{F}(x-t)}{\overline{F}(x)}\,\overline{F}(t)\,dt = 0 \tag{4.58}$$

because of our hypothesis. Now, as $F \in \mathcal{S}^*$ $F \in \mathcal{L}$, using Exercise 4.3.9(b) and Exercise 4.3.13 we get

$$\limsup_{x\to\infty} \int_v^{x/2} \frac{\overline{F}(x-t)}{\overline{F}(x)}\,\overline{F}(t)\,dt$$

$$= \limsup_{x\to\infty} \left[ \int_0^{x/2} \frac{\overline{F}(x-t)}{\overline{F}(x)}\,\overline{F}(t)\,dt - \int_0^{v} \frac{\overline{F}(x-t)}{\overline{F}(x)}\,\overline{F}(t)\,dt \right]$$

$$= \mu_F - \int_0^{v} \overline{F}(t)\,dt.$$

By taking $v$ large enough the above can be made arbitrarily small. Hence (4.58) follows. □

To phrase the next result we need the following definition.

**Definition 4.31** Let $G$ be a distribution supported on $(0,\infty)$ with probability density function $g(\cdot)$. The function

$$q_G(x) = \frac{g(x)}{1-G(x)} = \frac{g(x)}{\overline{G}(x)},\ x > 0 \tag{4.59}$$

is called the *hazard rate* (or *failure rate*) *function.*

**Lemma 4.32** *Let $F \in \mathcal{L}$. Then there exists $G \in \mathcal{L}$ such that $F \sim^t G$, $G$ has a hazard rate function $q_G(\cdot)$ satisfying $\lim_{x\to\infty} q_G(x) = 0$.*

**Proof:** Set $M_F(x) = -\log \overline{F}(x),\ x \geq 0$. The continuous function $M_G$ is defined so that $M_G(n) = M_F(n)$, and $M_G$ is linear on $[n, n+1]$, $n = 0, 1, 2, \ldots$. Define $\overline{G}(x) = \exp(-M_G(x)),\ x \geq 0$. Clearly $M_G$ is differentiable at $x \neq n$; put $q_G(x) = M_G'(x),\ x \neq n$ and $q_G(n) = 0,\ n = 0, 1, 2, \ldots$. Note that

$$q_G(x) = M_F(n+1) - M_F(n) = \log \frac{\overline{F}(n)}{\overline{F}(n+1)},$$

$n < x < n+1,\ n = 0, 1, 2, \ldots$. Details of proof are left as exercise; see Exercise 4.3.10. □

**Theorem 4.33** *If $F \in \mathcal{S}^*$ then $F \in \mathcal{S}$ and $F_I \in \mathcal{S}$.*

**Proof:** By Lemma 4.29 we have $F \in \mathcal{L}$. So in view of Lemmas 4.32, 4.30, 4.24 and Exercise 4.3.11 we may assume without loss of generality that $F$ has a density function $f(\cdot)$ and a hazard rate function $q(\cdot)$ with

$$\lim_{x\to\infty} q(x) = \lim_{x\to\infty} \frac{f(x)}{\overline{F}(x)} = 0.$$

In view of Proposition 4.27 it just remains to show that $F$ is subexponential.

Let $v > 0$ be fixed. For $x > 2v$, using (4.47) we have

$$\begin{aligned}
\frac{\overline{F^{*(2)}}(x)}{\overline{F}(x)} &= 1 + \int_0^x \frac{\overline{F}(x-t)}{\overline{F}(x)}\, dF(t) \\
&= 1 + \int_0^v \frac{\overline{F}(x-t)}{\overline{F}(x)}\, dF(t) + \int_v^{x-v} \frac{\overline{F}(x-t)}{\overline{F}(x)}\, f(t)\, dt \\
&\quad + \int_{x-v}^x \frac{\overline{F}(x-t)}{\overline{F}(x)}\, f(t)\, dt.
\end{aligned} \tag{4.60}$$

Since $f(y) = \frac{d}{dy}\,(-\overline{F}(y))$, integration by parts gives

$$\int_{x-v}^{x} \frac{\overline{F}(x-t)}{\overline{F}(x)}\, f(t)\, dt$$

$$= \frac{\overline{F}(v)\,\overline{F}(x-v) - \overline{F}(x)}{\overline{F}(x)} + \int_0^v \frac{\overline{F}(x-t)}{\overline{F}(x)}\, f(t)\, dt.$$

So (4.60) becomes

$$\frac{\overline{F^{*(2)}}(x)}{\overline{F}(x)} = 2\int_0^v \frac{\overline{F}(x-t)}{\overline{F}(x)}\, dF(t) + \frac{\overline{F}(v)\,\overline{F}(x-v)}{\overline{F}(x)} + \int_v^{x-v} \frac{\overline{F}(x-t)}{\overline{F}(x)}\, f(t)\, dt. \tag{4.61}$$

As $F \in \mathcal{L}$, for any fixed $v > 0$, note that $\left|\frac{\overline{F}(x-t)}{\overline{F}(x)}\right| \le \frac{\overline{F}(x-v)}{\overline{F}(x)} \le 2$ for all sufficiently large $x$, uniformly over $0 \le t \le v$. So by the dominated convergence theorem

$$\limsup_{x\to\infty} \left[2\int_0^v \frac{\overline{F}(x-t)}{\overline{F}(x)}\, dF(t) + \overline{F}(v)\,\frac{\overline{F}(x-v)}{\overline{F}(x)}\right]$$

$$= 2F(v) + \overline{F}(v) \le 2. \tag{4.62}$$

In view of (4.61), (4.62) and Exercise 4.3.1, it is now enough to prove that

$$\lim_{v\to\infty} \limsup_{x\to\infty} \int_v^{x-v} \frac{\overline{F}(x-t)}{\overline{F}(x)}\, f(t)\, dt = 0. \tag{4.63}$$

Now as the hazard rate function $q(\cdot)$ converges to 0, there is $x_0 > 0$ such that $f(x) \le \overline{F}(x)$, $x \ge x_0$. So for $v \ge x_0$, and $x > 2v$ we have

$$\int_v^{x-v} \frac{\overline{F}(x-t)}{\overline{F}(x)}\, f(t)\, dt \le \int_v^{x-v} \frac{\overline{F}(x-t)}{\overline{F}(x)}\, \overline{F}(t)\, dt$$

$$= \int_0^x \frac{\overline{F}(x-t)}{\overline{F}(x)}\, \overline{F}(t)\, dt - 2\int_0^v \frac{\overline{F}(x-t)}{\overline{F}(x)}\, \overline{F}(t)\, dt.$$

Consequently as $F \in \mathcal{S}^* \subset \mathcal{L}$, by Exercise 4.3.13 we obtain

$$\limsup_{x\to\infty} \int_v^{x-v} \frac{\overline{F}(x-t)}{\overline{F}(x)} f(t)\, dt \le 2\mu - 2\int_0^v \overline{F}(t)\, dt.$$

Letting $v \to \infty$ in the above we get (4.63), completing the proof.

□

For an absolutely continuous distribution to belong to $\mathcal{S}^*$ we now give some sufficient conditions in terms of the hazard rate function.

**Theorem 4.34** *Let $F$ be an absolutely continuous distribution function on $(0,\infty)$ with finite expectation $\mu$ and hazard rate function $q(\cdot)$. If $\limsup_{x\to\infty} xq(x) < \infty$ then $F \in \mathcal{S}^*$, and hence $F$, $F_I$ are both subexponential.*

**Proof:** We shall first establish that

$$\limsup_{x\to\infty} \frac{1}{\overline{F}(x)} \overline{F}\left(\frac{x}{2}\right) < \infty. \tag{4.64}$$

By hypothesis there exist $a > 0$, $x_0 > 0$ such that $yq(y) < a$ for all $y \ge x_0$. Therefore

$$\begin{aligned} & \limsup_{x\to\infty} \left[\log \overline{F}\left(\frac{x}{2}\right) - \log \overline{F}(x)\right] \\ = & \limsup_{x\to\infty} \int_{x/2}^{x} q(y)\, dy \le a \log 2 < \infty \end{aligned}$$

whence (4.64) follows.

Note that $q$ converges to 0, and hence by Exercise 4.3.10(a) $F \in \mathcal{L}$. For $x > 2v > 0$ note that (by splitting $\int_0^x = \int_0^{x/2} + \int_{x/2}^x$, and changing the variable $t' = (x-t)$ in the second integral)

$$\begin{aligned} & \int_0^x \frac{\overline{F}(x-t)}{\overline{F}(x)} \overline{F}(t)\, dt = 2\int_0^{x/2} \frac{\overline{F}(x-t)}{\overline{F}(x)} \overline{F}(t)\, dt \\ \le & \ 2\int_0^v \frac{\overline{F}(x-t)}{\overline{F}(x)} \overline{F}(t)\, dt + 2\frac{\overline{F}(\frac{x}{2})}{\overline{F}(x)} \int_v^\infty \overline{F}(t)\, dt. \end{aligned} \tag{4.65}$$

As $F$ has finite expectation, because of (4.64), the second term on r.h.s. of (4.65) can be made arbitrarily small by choosing a large enough $v$. With such a $v > 0$ (chosen and fixed), as $F \in \mathcal{L}$, by Exercise 4.3.13, it now follows that

$$\lim_{x\to\infty} \int_0^x \frac{\overline{F}(x-t)}{\overline{F}(x)}\, \overline{F}(t)\, dt \leq 2\mu.$$

As the reverse inequality always holds (Exercise 4.3.9(a)) we now have $F \in \mathcal{S}^*$. □

The next result can be useful if $q(x) = O(\frac{1}{x})$, $x \to \infty$ does not hold.

**Theorem 4.35** *Let $F$ be an absolutely continuous distribution supported on $(0,\infty)$ with hazard rate function $q(\cdot)$. Assume that*
*(i) $q(\cdot)$ is eventually decreasing to 0; that is, there is $a \geq 0$ such that $q(x_1) \geq q(x_2)$ if $a \leq x_1 \leq x_2$ and $\lim_{x\to\infty} q(x) = 0$; and*
*(ii) $\int_0^\infty e^{xq(x)}\overline{F}(x)\, dx < \infty$.*
*Then $F \in \mathcal{S}^*$.*

**Proof:** Without loss of generality we may assume that $q(\cdot)$ is nonincreasing on $[0,\infty)$. Otherwise, one can define a hazard rate function $q_0(\cdot)$ such that $q_0(\cdot) = q(\cdot)$ on $[v_0,\infty)$, $q_0(\cdot) = q(v_0)$ on $[0,v_0)$ for a suitable $v_0 > 0$, so that $q_0(\cdot)$ is nonincreasing on $[0,\infty)$ and the corresponding distribution function $F_0$ is tail equivalent to $F$. Clearly $F$, $F_0 \in \mathcal{L}$ by Exercise 4.3.10, and hence Lemma 4.30 implies that $F \in \mathcal{S}^* \Leftrightarrow F_0 \in \mathcal{S}^*$. Also the integrability condition (ii) holds for $F$ if and only if it holds for $F_0$.

Note that (ii) implies that $F$ has finite expectation, say $\mu$. Write $Q(x) = -\log \overline{F}(x)$. So $Q'(\cdot) = q(\cdot)$. By Exercise 4.3.9(b), to prove $F \in \mathcal{S}^*$ it is enough to show that

$$\lim_{x\to\infty} \int_0^{x/2} e^{Q(x)-Q(x-y)}\, \overline{F}(y)\, dy = \mu. \tag{4.66}$$

Monotonicity of $q$ over $[0,\infty)$ gives

$$1 \leq e^{yq(x)} \leq e^{Q(x)-Q(x-y)} \leq e^{yq(x/2)}$$

for $0 \leq y \leq x/2$. This, in turn, gives

$$\begin{aligned} \int_0^{x/2} \overline{F}(y)\, dy &\leq \int_0^{x/2} e^{Q(x)-Q(x-y)} \overline{F}(y)\, dy \\ &\leq \int_0^{x/2} e^{yq(x/2)}\, \overline{F}(y)\, dy. \end{aligned} \tag{4.67}$$

Clearly l.h.s. of (4.67) tends to $\mu$ as $x \to \infty$. Note that $\lim_{x\to\infty} e^{yq(x/2)}\, \overline{F}(y) = \overline{F}(y)$ for all $y$. Also $I_{[0,x/2]}(y)\, e^{yq(x/2)}\, \overline{F}(y) \leq e^{yq(y)}\, \overline{F}(y)$ for all $x \geq 0$; by (ii), r.h.s. is integrable. So by the dominated convergence theorem

$$\lim_{x\to\infty} \int_0^{x/2} e^{yq(x/2)}\, \overline{F}(y)\, dy = \mu.$$

Hence, by (4.67), now (4.66) follows. □

The preceding theorem is apparently inspired by Pitman's criterion for subexponentiality of $F$. See [EKM] for a discussion on Pitman's result. Our treatment of $\mathcal{S}^*$ follows closely [RSST].

**Example 4.36** *Lognormal distribution.* Let $\Phi, \varphi$ denote respectively the distribution function and the density function of the $N(0,1)$ distribution. Then $X = e^Z$ has the lognormal distribution, where $Z$ has $N(0,1)$ distribution. The tail and the hazard rate function of $X$ are respectively given by $\overline{F}(x) = 1 - \Phi(\log x)$, $q(x) = \frac{\varphi(\log x)}{x[1-\Phi(\log x)]}$. Therefore

$$e^{xq(x)}\, \overline{F}(x) = [1 - \Phi(\log x)] \cdot \exp\left\{ \frac{\varphi(\log x)}{[1 - \Phi(\log x)]} \right\}.$$

By Exercise 4.1.2(a) we know that

$$\varphi(x) \sim x(1 - \Phi(x)), \text{ as } x \to \infty \tag{4.68}$$

and hence $q(x) \sim \frac{\log x}{x}$, $x \to \infty$; so $q(x) \to 0$ as $x \to \infty$. Again by (4.68)

$$\begin{aligned} e^{xq(x)}\overline{F}(x) &\sim (1 - \Phi(\log x)) \cdot \exp\left( x\, \frac{\log x}{x} \right) \\ &= x(1 - \Phi(\log x)) = \frac{x}{\log x}\{(\log x)[1 - \Phi(\log x)]\} \\ &\sim \frac{x}{\log x}\, \varphi(\log x), \text{ as } x \to \infty. \end{aligned}$$

Note that $\varphi(\log x) = \frac{1}{\sqrt{2\pi}}\, x^{-\frac{1}{2}(\log x)}$, and that

$$\int_1^\infty x\, \varphi(\log x) dx = \int_1^\infty \frac{1}{\sqrt{2\pi}}\, x^{1-\frac{1}{2}\log x}\, dx < \infty.$$

So it follows that $e^{xq(x)}\overline{F}(x)$ is integrable. Hence by Theorem 4.35 it follows that $F \in \mathcal{S}^*$. Thus $F$ and $F_I$ are both subexponential when $F$ is lognormal.

□

We conclude this section with a result concerning the distribution of total claim amount for subexponential claims. Recall that in the Sparre Andersen model the total claim is given by $S(t) = X_1 + \ldots + X_{N(t)}$, $t \geq 0$ (with, of course, $S(t) = 0$ if $N(t) = 0$). Here $\{N(t)\}$ is a renewal process independent of the i.i.d. claims $\{X_i\}$.

**Proposition 4.37** *Suppose $X_i \sim F \in \mathcal{S}$. Let $G_t(\cdot)$ denote the distribution of $S(t)$. Assume that $P(N(t) = 0) < 1$. Let $t > 0$ be fixed. Suppose $\sum_{n=0}^{\infty} (1+\epsilon)^n P(N(t) = n) < \infty$ for some $\epsilon > 0$. Then $G_t \in \mathcal{S}$.*

**Proof:** See Exercise 4.3.16. □

## 4.4 Exercises

### Exercises on Section 4.1

**4.1.1.** Let $F$ be a distribution function supported on $[0, \infty)$. Show that $F$ is light tailed if and only if

$$\limsup_{x\to\infty} \frac{1 - F(x)}{e^{-\lambda x}} < \infty \quad \text{for some } \lambda > 0.$$

**4.1.2.** (a) Let $\Phi$, $\varphi$ denote respectively the distribution function and the probability density function of the standard normal distribution. Using l'Hospital's rule show that

$$\lim_{x\to\infty} \frac{1 - \Phi(x)}{x^{-1}\varphi(x)} = 1,$$

and hence conclude that the truncated standard normal distribution is light tailed.

(b) Show that the truncated $N(\mu, \sigma^2)$ distribution is light tailed; here $F(x) = P(|N(\mu, \sigma^2)| \leq x),\ x \in \mathbb{R}$. (Of course parts (a), (b) can also be shown by proving that the moment generating function exists in a neighbourhood of 0.)

**4.1.3.** Show that the $\Gamma(\alpha, \lambda)$ distribution is light tailed for $\lambda > 0,\ \alpha > 0$.

**4.1.4.** (a) For a light tailed distribution $F$, show that the Esscher transform $\tilde{F}_t$ given by (4.5) is indeed a distribution function supported on $[0, \infty)$ for any $t < s_F$.

(b) Show that the moment generating function $\hat{m}_{\tilde{F}_t}$ of $\tilde{F}_t$ is given by

$$\hat{m}_{\tilde{F}_t}(s) = \frac{\hat{m}_F(t+s)}{\hat{m}_F(t)}, \text{ for } s < s_F - t,$$

where $t < s_F$.

(c) Show that the expectation of $\tilde{F}_t$ is $\frac{\hat{m}_F^{(1)}(t)}{\hat{m}_F(t)}$ for $t < s_F$.

**4.1.5.** Let $F$ denote the $\Gamma(\alpha, \lambda)$ distribution, where $\alpha > 0,\ \lambda > 0$.
(i) Find $s_F$.
(ii) For $t < s_F$, find the Esscher transform of $F$.

**Exercises on Section 4.2**

**4.2.1.** Show that the following functions are slowly varying at $\infty$.
(i) $L(\cdot) > 0$, measurable with $\lim_{x \to \infty} L(x) = c > 0$.
(ii) $[\log(1+x)]^b,\ b \in \mathbb{R}$ (Hint: First consider $b =$ integer).
(iii) $\log_2 x \triangleq \log(\log x),\ \log_k x \triangleq \log(\log_{k-1} x)$.

**4.2.2.** Let $L$ be slowly varying and nondecreasing (resp. nonincreasing). Show that the function $\epsilon(\cdot)$ in Karamata representation can be taken to be nonnegative (resp. nonpositive).

**4.2.3.** If $L_1,\ L_2$ are slowly varying (at $\infty$), then so are $L_1 L_2,\ \frac{L_1}{L_2}$, $L_1 + L_2$. (Hint: Take $\max\{\epsilon_1(t), \epsilon_2(t)\}$ in Karamata representation.)

**4.2.4.** If $L$ varies slowly at $\infty$ show that

$$\lim_{x \to \infty} \frac{\log L(x)}{\log x} = 0.$$

**4.2.5.** If $L$ varies slowly at $\infty$ show that so does $(L(\cdot))^b$ for any $b \in \mathbb{R}$ (Hint: Enough to consider the case $L(x) \to \infty$ and $b =$ positive integer.)

**4.2.6.** Show that $f$ is regularly varying with index $\delta \Leftrightarrow \lim_{x\to\infty} \frac{f(cx)}{f(x)} = c^\delta$ for all $c > 0$. Here $f$ is a positive measurable function on $(0, \infty)$.

**4.2.7.** Let $f$ be a positive continuous function on $(0, \infty)$ such that $\lim_{x\to\infty} \frac{f(tx)}{f(x)} = \psi(t) < \infty$ for all $t > 0$. Assume $\psi$ is continuous.
(i) Show that $\psi(t_1 t_2) = \psi(t_1)\ \psi(t_2)$ for all $t_1, t_2 > 0$.
(ii) Show that $\psi(t) = t^\delta,\ t > 0$ for some $\delta \in \mathbb{R}$.

(Hint: Put $\xi = \log t,\ w(\xi) = \psi(e^\xi)$; then $w(\xi_1 + \xi_2) = w(\xi_1) \cdot w(\xi_2),\ \forall\ \xi_1, \xi_2 \in \mathbb{R}$ and hence $w(\xi) = (e^\delta)^\xi$ for some $\delta$ as $w$ is positive and measurable.)

**4.2.8.** Let $h(\cdot) > 0$ be a measurable function on $(0, \infty)$. Show that $h(\cdot)$ is regularly varying with index $\alpha \Leftrightarrow h(\cdot)$ can be represented as

$$h(x) = c(x) \exp\left( \int_a^x \frac{1}{t}\ \delta(t)\ dt \right), \quad x > a$$

where $c(\cdot) > 0$ is a measurable function such that $c(x) \to c_0 > 0,\ \delta(\cdot)$ is a measurable function such that $\lim_{t\to\infty} \delta(t) = \alpha$.

**4.2.9.** Let $L$ be a slowly varying function. For $\alpha \in \mathbb{R}$, let $L_\alpha(x) = \int_1^x t^\alpha L(t)\ dt,\ x > 1$. Show directly, using Proposition 4.6, that $L_\alpha(x)$ diverges as $x \to \infty$ if $\alpha > 0$, and that $L_\alpha(x)$ converges if $\alpha < -1$.

**4.2.10.** a) Let $L(t) = \frac{1}{(\log t)^2},\ t \geq e$. Show that $L$ is a slowly varying function such that $\int_K^\infty \frac{1}{t}\ L(t)\ dt < \infty$ for some $K > 0$.

b) Use (a) to get a distribution function on $(0, \infty)$ having a slowly varying tail, that is, regularly varying distribution with tail index $\alpha = 0$. Note that such a distribution does not belong to the domain of attraction of any nondegenerate distribution. (See Appendix A.2)

c) Let $X$ be a nonnegative random variable having distribution described in (b). Show directly that $E(X^r) = \infty$ for any $r > 0$.

**4.2.11.** Prove (4.33) by a Fubini argument.

**4.2.12.** Let $F \in \mathcal{R}_{-\alpha},\ \alpha \geq 0$. Show that $\hat{m}_F(s) = \infty$ for any $s > 0$. So moment generating function of $F$ does not exist in any neighbourhood of 0.

**4.2.13.** Let $F$ be supported on $[0, \infty)$ and specified by (4.34); that is, $F$ has Pareto $(\alpha)$ distribution.
(i) Show that $F$ is indeed a distribution function.
(ii) Find the expectation when $\alpha > 1$.

(iii) If $\alpha > 1$ show that $F_I \in \mathcal{R}_{-(\alpha-1)}$, where $F_I$ is the integrated tail distribution corresponding to $F$. Show also that $F_I$ has Pareto $(\alpha - 1)$ distribution.
(iv) If $X$ has Pareto $(\alpha)$ distribution where $\alpha > 0$ show that $E(X^\alpha) = \infty$.

**Exercises on Section 4.3**

**4.3.1.** (a) For any distribution $F$ supported on $(0, \infty)$ show that

$$\liminf_{x\to\infty} \frac{\overline{F^{*(2)}}(x)}{\overline{F}(x)} \geq 2.$$

(b) Let $F$ be a distribution supported on $(0, \infty)$ such that

$$\limsup_{x\to\infty} \frac{\overline{F^{*(2)}}(x)}{\overline{F}(x)} \leq 2.$$

Show that $F$ is subexponential.

**4.3.2.** If $F$ satisfies (4.44) for $n = 2$ show that

$$\lim_{x\to\infty} \int_0^x \frac{\overline{F}(x-y)}{\overline{F}(x)}\, dF(y) = 1.$$

(Hint: Use (4.47)).

**4.3.3.** Using (4.44) for $n = 2$, show that $\overline{F}(x) > 0$ for all $x \geq 0$.

**4.3.4.** Derive (4.51).

**4.3.5.** Let $p_0, p_1, \ldots, p_n \geq 0$ such that $\sum_{i=0}^{n} p_i = 1$; so $\{p_i\}$ gives a probability on $\{0, 1, \ldots, n\}$. Let $F \in \mathcal{S}$ and $G(x) = \sum_{j=0}^{n} p_j F^{*(j)}(x)$. Note that $G$ is a distribution function. Show that

$$\lim_{x\to\infty} \frac{\overline{G}(x)}{\overline{F}(x)} = \sum_{k=0}^{n} k\, p_k.$$

**4.3.6.** (a) Let $F$ be a distribution function supported on $(0, \infty)$. For any integer $m \geq 1$ show that

$$\overline{F^{*(m+1)}}(x) = \overline{F^{*(m)}}(x) + \int_0^x \overline{F}(x-t)\, dF^{*(m)}(t).$$

(b) Using a) show that

$$\limsup_{x\to\infty} \frac{\overline{F^{*(m+1)}}(x)}{\overline{F}(x)} \geq 1 + \limsup_{x\to\infty} \frac{\overline{F^{*(m)}}(x)}{\overline{F}(x)}.$$

(c) If there is an integer $n \geq 2$ such that

$$\limsup_{x\to\infty} \frac{\overline{F^{*(n)}}(x)}{\overline{F}(x)} \leq n$$

show that $F \in \mathcal{S}$. In particular if (4.44) holds for one $n \geq 2$, then it holds for all $n \geq 2$ and $F$ is subexponential.

**4.3.7.** If $F$ is subexponential, then so is $F^{*(k)}$ for any $k = 2, 3, \ldots$.

**4.3.8.** (a) Let $F, G$ be subexponential and assume that $\lim_{x\to\infty} \frac{\overline{G}(x)}{\overline{F}(x)} = c$, where $0 < c < \infty$. Let the probability distribution $Q$ be defined by $Q(x) = p\ F(x) + (1-p)\ G(x)$ where $0 < p < 1$. Show that $Q$ is subexponential.

(Hint: Note that

$$\begin{aligned} \overline{Q^{*(2)}}(x) &= p^2\ \overline{F^{*(2)}}(x) + 2p(1-p)\ \overline{F * G}(x) + (1-p)^2 \overline{G^{*(2)}}(x), \\ \overline{Q}(x) &= p\ \overline{F}(x) + (1-p)\ \overline{G}(x), \end{aligned}$$

and use Lemma 4.20).

(b) Prove that $G$ in Exercise 4.3.5 is subexponential, if $p_0 < 1$.

**4.3.9.** (a) Let $F$ have finite expectation $\mu$. Show that

$$\lim_{x\to\infty} \int_0^x \frac{\overline{F}(x-y)}{\overline{F}(x)}\ \overline{F}(y)\ dy \geq 2\mu.$$

(b) Show that $F \in \mathcal{S}^* \Leftrightarrow$

$$\lim_{x\to\infty} \int_0^{x/2} \frac{\overline{F}(x-y)}{\overline{F}(x)}\ \overline{F}(y)\ dy = \mu.$$

**4.3.10.** (a) Let $G$ be a distribution on $(0,\infty)$ with hazard rate function $q_G$. If $q_G(x) \to 0$ as $x \to \infty$ show that $G \in \mathcal{L}$.

(b) Complete the proof of Lemma 4.32.

**4.3.11.** Let $F$, $G$ be distributions supported on $(0,\infty)$. Assume that $\lim_{x\to\infty} \frac{\overline{F}(x)}{\overline{G}(x)} = c \in (0,\infty)$. Suppose $F$ has finite expectation $\mu_F$.

(i) Show that $G$ has finite expectation.
(ii) Show that

$$\lim_{x\to\infty} \frac{\overline{F}_I(x)}{\overline{G}_I(x)} = c\,\frac{\mu_G}{\mu_F}$$

where $\mu_G$ is the expectation of $G$, and $F_I$, $G_I$ respectively are the integrated tail distributions of $F$, $G$.
(iii) $F \sim^t G \Rightarrow F_I \sim^t G_I$.

**4.3.12.** A distribution $F$ on $(0,\infty)$ is said to be *dominatedly varying* if (4.64) holds; it is written $F \in \mathcal{D}$. Show that $\mathcal{R}_{-\alpha} \subset \mathcal{D}$ for $\alpha > 0$. If $F$ has finite mean and $F \in \mathcal{L} \cap \mathcal{D}$ show that $F \in \mathcal{S}^*$.

**4.3.13.** Let $F \in \mathcal{L}$. For any fixed $v > 0$ show that

$$\lim_{x\to\infty} \int_0^v \frac{\overline{F}(x-t)}{\overline{F}(x)}\,\overline{F}(t)\,dt = \int_0^v \overline{F}(t)\,dt.$$

**4.3.14.** Let $0 < \tau < 1$, $c > 0$, and $\overline{F}(x) = \exp(-cx^\tau)$, $x > 0$. This is the *heavy-tailed Weibull distribution.* Show that $F \in \mathcal{S}^*$ and hence $F$, $F_I$ are both subexponential.

**4.3.15.** Let $F$ be a distribution function supported on $(0,\infty)$ with hazard rate function $q_F(\cdot)$. Define $Q(x) = \int_0^x q_F(t)\,dt$, $x > 0$.
(a) Show that $\overline{F}(x) = \exp(-Q(x))$, $x > 0$.
(b) Using (4.47) show that

$$\frac{\overline{F^{*(2)}}(x)}{\overline{F}(x)} - 1 = \int_0^x \{\exp[Q(x) - Q(x-t) - Q(t)]\}\,q_F(t)\,dt.$$

**4.3.16.** Let the notation and hypotheses be as in Proposition 4.37.
(i) Show that $G_t(x) = \sum_{k=0}^{\infty} F^{*(k)}(x) \cdot P(N(t) = k)$.
(ii) Prove Proposition 4.37.
(iii) Show that $\overline{G}_t(x) \sim E(N(t))\,\overline{F}(x)$, $x \to \infty$.
(iv) In the case of the Cramér-Lundberg model note that the hypotheses of Proposition 4.37 are satisfied.

**4.3.17.** Give an example of a subexponential distribution having finite moments of any order.

# Chapter 5

# Ruin Problems

Ruin probability is the probability of the insurance company getting ruined in finite time. It is considered to be a reasonably objective indicator of the health of the company. So problems centering around ruin probability have a prominent place in mathematical analysis of insurance models. We shall first consider some general results in the Sparre Andersen model; later we shall specialise to the case of the Cramér-Lundberg model.

## 5.1 Risk process and an associated random walk

Recall $S(t) = \sum_{i=1}^{N(t)} X_i,\ t \geq 0$ is the total claim amount process, where the claim number process $\{N(t)\}$ is a renewal process independent of the i.i.d. claim sizes $\{X_i\}$; $X_i > 0$ a.s.; $0 = T_0 < T_1 < T_2 < \ldots < T_n < \ldots$ are claim arrival times; also $A_i = T_i - T_{i-1},\ i \geq 1$ are i.i.d. interarrival times.

We assume that the premium income $p(\cdot)$ is a deterministic linear function, that is, $p(t) = ct,\ t \geq 0$; here $c > 0$ is the *premium rate*. Define

$$R(t) = x + ct - S(t) = x + ct - \sum_{i=1}^{N(t)} X_i,\ t \geq 0. \tag{5.1}$$

$R(\cdot)$ is called the *risk* or *surplus process*; in the Sparre Andersen model it is called the *renewal risk process*; in the Cramér-Lundberg model it is called the *Lundberg risk process*. Here $x$ is the initial capital of the company. Note that $R(t)$ gives the *surplus* or *capital balance* of the company at time $t$.

The event that $R(\cdot)$ falls below zero is called *ruin*; that is, ruin is $\{R(t) < 0 \text{ for some } t \geq 0\} = \bigcup_{t \geq 0} \{R(t) < 0\}$. Define the *ruin time* $\tau_0$ by

$$\tau_0 = \inf\{t \geq 0 : R(t) < 0\} \tag{5.2}$$

and the *ruin probability* $\psi(\cdot)$ by

$$\psi(x) = P(\tau_0 < \infty | R(0) = x) \triangleq P_x(\tau_0 < \infty). \tag{5.3}$$

Finding appropriate equations, expressions, estimates for $\psi(\cdot)$ and analysing the asymptotics of $\psi(\cdot)$ form the core of "ruin problems".

By the definition of the risk process, it is clear that $R(\cdot)$ is increasing on $[T_n, T_{n+1})$, $n \geq 0$. So ruin can occur only at some $T_n$. The sequence $\{R(T_n) : n \geq 0\}$ is called the *skeleton process* of the risk process $R(\cdot)$. Now observe that, as $N(T_n) = n$,

$$\begin{aligned} \text{Ruin} &= \{R(t) < 0 \text{ for some } t \geq 0\} = \left\{ \inf_{t \geq 0} R(t) < 0 \right\} \\ &= \left\{ \inf_{n \geq 0} R(T_n) < 0 \right\} = \left\{ \inf_{n \geq 0} [x + cT_n - \sum_{i=1}^{n} X_i] < 0 \right\}. \end{aligned}$$

Write $S_0 \equiv 0$, $Z_n = X_n - cA_n$, $S_n = \sum_{i=1}^{n} Z_i$, $n \geq 1$. It is now clear that

$$\psi(x) = P\left\{ \inf_{n \geq 0} (-S_n) < -x \right\} = P\left\{ \sup_{n \geq 0} S_n > x \right\}. \tag{5.4}$$

As $\{A_i\}$, $\{X_j\}$ are i.i.d. sequences, and the sequences are mutually independent, note that $\{Z_i : i \geq 1\}$ is a sequence of real valued i.i.d. random variables. Thus $\{S_n : n \geq 0\}$ is *random walk* on $\mathbb{R}$. So the ruin probability is just the tail probability of the supremum functional of the random walk $\{S_n\}$. The main result of this section is

**Theorem 5.1** *With the notation as above, assume that* $Z_1 \not\equiv 0$, *and that* $E(Z_1)$ *exists.*
*a) If* $E(Z_1) > 0$ *then* $\lim_{n \to \infty} S_n = +\infty$, *a.s.*
*b) If* $E(Z_1) < 0$ *then* $\lim_{n \to \infty} S_n = -\infty$, *a.s.*
*c) If* $E(Z_1) = 0$ *then*

$$\limsup_{n \to \infty} S_n = +\infty, \quad \liminf_{n \to \infty} S_n = -\infty, \quad a.s.$$

□

Assertions (a) and (b) are easy consequences of the strong law of large numbers, and hence left as exercise. Though assertion (c) may not be surprising to those familiar with the symmetric simple random walk on $\mathbb{Z}$, its proof is by no means obvious. In fact we need to introduce the following before taking up the proof.

For $n = 0, 1, 2, \ldots$ let $\mathcal{G}_n \triangleq \sigma\{S_i : i \leq n\}$ denote the filtration generated by the random walk. Note that $\mathcal{G}_n = \sigma\{Z_i : i \leq n\}$. Define

$$\nu^+ = \min\{n > 0 : S_n > 0\}, \tag{5.5}$$

$$\nu^- = \min\{n > 0 : S_n \leq 0\}; \tag{5.6}$$

$\nu^+$ (resp. $\nu^-$) is called the *first entrance time of* $\{S_n\}$ *into* $(0, \infty)$ *(resp.* $(-\infty, 0]$*)*. Set $\nu_0^+ = 0$, $\nu_1^+ = \nu^+$, $\nu_0^- = 0$, $\nu_1^- = \nu^-$; for $n \geq 1$ define

$$\nu_{n+1}^+ = \min\{j > \nu_n^+ : S_j > S_{\nu_n^+}\} \tag{5.7}$$

$$\nu_{n+1}^- = \min\{j > \nu_n^- : S_j \leq S_{\nu_n^-}\}. \tag{5.8}$$

Note that one can define $S_{\nu_n^+}$ (resp. $S_{\nu_n^-}$) in an unambiguous manner whenever $\nu_n^+ < \infty$ (resp. $\nu_n^- < \infty$); see Exercise 5.1.4. The stopping time $\nu_n^+$ is called the $n^{th}$ *strong ascending ladder epoch* of $\{S_n\}$; similarly $\nu_n^-$ is called the $n^{th}$ *(weak) descending ladder epoch* of $\{S_n\}$. Note the slight asymmetry between (5.5), (5.7) on the one hand and (5.6), (5.8) on the other.

**Proof of Theorem 5.1:** In view of Exercise 5.1.3 we need to consider only case (c); so assume $E(Z_1) = 0$. We first prove

$$P(\nu^+ < \infty) = 1; \tag{5.9}$$

that is, the random walk $\{S_n\}$ enters $(0, \infty)$ in finite time with probability 1. Put $M = \min\{n : S_n = \sup_{j \geq 0} S_j\}$ = first time maximum is reached; of course, $M = \infty$ if supremum is not attained. For any $n \geq 0$, note that the events $\{S_i < S_n, i \leq n-1\}$ and $\{S_j \leq S_n, j > n\}$ are independent, because the former involves $Z_1, Z_2, \ldots, Z_n$, while the latter depends on $Z_{n+1}, Z_{n+2}, \ldots$. Clearly $(Z_1, Z_2, \ldots, Z_n) \stackrel{d}{=} (Z_n, Z_{n-1}, \ldots, Z_1)$. Therefore

$$\begin{aligned} P(S_i < S_n, i \leq n-1) &= P\left(\sum_{\ell=i+1}^{n} Z_\ell > 0, i \leq n-1\right) \\ &= P(Z_n > 0, Z_{n-1} + Z_n > 0, \ldots, Z_1 + Z_2 + \ldots + Z_n > 0) \\ &= P(Z_1 > 0, Z_1 + Z_2 > 0, \ldots, Z_1 + Z_2 + \ldots + Z_n > 0) \\ &= P(\nu^- > n) \end{aligned}$$

where the last step follows by the definition of $\nu^-$. Next note that $\sum_{k=n+1}^{j} Z_k \stackrel{d}{=} \sum_{\ell=1}^{j-n} Z_\ell$ for $j > n$. Hence using the definition of $\nu^+$ we have

$$P(S_j \leq S_n, j > n) = P\left(\sum_{k=n+1}^{j} Z_k \leq 0, j > n\right)$$

$$= P\left(\sum_{k=1}^{j} Z_k \leq 0, j > 0\right) = P(\nu^+ = \infty).$$

Consequently it now follows that

$$1 \geq P(M < \infty) = \sum_{n=0}^{\infty} P(M = n)$$

$$= \sum_{n=0}^{\infty} P(\{S_i < S_n, i \leq n-1\} \cap \{S_j \leq S_n, j > n\})$$

$$= \sum_{n=0}^{\infty} P(S_i < S_n, i \leq n-1) \cdot P(S_j \leq S_n, j > n)$$

$$= P(\nu^+ = \infty) \sum_{n=0}^{\infty} P(\nu^- > n)$$

$$= E(\nu^-) \cdot P(\nu^+ = \infty). \tag{5.10}$$

Now, if possible let $P(\nu^+ = \infty) > 0$. Then (5.10) implies $E(\nu^-) < \infty$. Consequently Wald's identity and our hypothesis give $E(S_{\nu^-}) = E(\nu^-)\, E(Z_1) = 0$. By definition of $S_{\nu^-}$ note that $S_{\nu^-} \leq 0$. Hence $S_{\nu^-} = 0$ a.s. Now $P(Z_1 = 0) < 1$, $E(Z_1) = 0$ imply $P(Z_1 < 0) > 0$. This would mean $P(S_{\nu^-} < 0) \geq P(Z_1 < 0) > 0$ which contradicts $S_{\nu^-} = 0$ a.s. Hence $P(\nu^+ = \infty) = 0$, proving (5.9).

We want to prove next that for any $m = 1, 2, \ldots$

$$P(\nu_m^+ < \infty) = 1. \tag{5.11}$$

As $\nu_1^+ = \nu^+$ we may take $m \geq 2$. Note that $\nu_m^+$ is increasing and $S_{\nu_m^+} > S_{\nu_{m-1}^+} > \ldots > S_{\nu_1^+} > 0$. Suppose (5.11) holds for $m \leq k$; we want to prove that it holds for $m = k+1$. Denote

$$\hat{M} = \min\left\{j : S_{\nu_k^+ + j} = \sup_{\ell} S_{\nu_k^+ + \ell}\right\}. \tag{5.12}$$

Then one can get the following analogue of (5.10) by proceeding similarly:

$$1 \geq P(\hat{M} < \infty) = E(\nu^-) \cdot P(\nu^+_{k+1} = \infty). \tag{5.13}$$

This in turn leads to (5.11), just like (5.9) was derived from (5.10). See Exercise 5.1.6.

By (5.11) and Exercise 5.1.5, it now follows that $\{S_{\nu^+_{n+1}} - S_{\nu^+_n} : n \geq 0\}$ is a sequence of i.i.d. random variables taking values in $(0, \infty)$. Clearly $E(S_{\nu^+_{n+1}} - S_{\nu^+_n}) = E(S_{\nu^+}) > 0$. So the strong law of large numbers implies

$$\frac{1}{n} S_{\nu^+_n} = \frac{1}{n} \sum_{i=0}^{n-1} \left( S_{\nu^+_{i+1}} - S_{\nu^+_i} \right) \to E(S_{\nu^+}) > 0 \ a.s.$$

Even if $E(S_{\nu^+}) = \infty$ the above conclusion holds since the summands are positive. Therefore $\lim_{n\to\infty} S_{\nu^+_n} = \infty$ a.s., and hence $\limsup_{n\to\infty} S_n = \infty$, a.s.

By considering $\{-S_n : n \geq 0\}$ in the place of $\{S_n\}$ with $E(-Z_1) = 0$, a similar analysis gives $\liminf_{n\to\infty} S_n = -\infty$ a.s. This completes the proof. □

By Theorem 5.1 and (5.4) we now have

**Corollary 5.2** Let $E(A_1)$, $E(X_1)$ be finite. If $E(Z_1) = E(X_1) - cE(A_1) \geq 0$ holds then $\psi(x) = 1$ for all $x > 0$; that is, if $E(Z_1) \geq 0$ then ruin occurs in finite time with probability 1 whatever be the initial capital. □

In particular, note that the net premium principle discussed in Section 3.4 leads to ruin.

**Definition 5.3** We say that the renewal risk model satisfies the *net profit condition (NPC)* if $E(X_1) - cE(A_1) < 0$.

It is clear that the premium rate has to be chosen so that NPC is satisfied, as NPC basically means that the average cashflow in the portfolio is on the positive side. This does not mean that ruin is avoided; it only means that one can hope $\psi(x) < 1$, $x > 0$.

Our proof of Theorem 5.1 is based on the approach given in [RSST].

**Note:** An interesting offshoot of ruin problem of insurance models is the branch of probability theory called *large deviations*. Let the notation

be as in Section 5.1. A question related to ruin is, what is the rate at which $P(\frac{1}{n}S_n > x)$ tends to 0 for any $x > E(X_1)$? More precisely, does there exist a nice function $I$, called the *rate function*, such that for $x > E(X_1)$

$$\frac{1}{n}\log P\left(\frac{1}{n}S_n > x\right) \to -I(x) \text{ as } n \to \infty?$$

This question was studied by Esscher and Cramér. If $\{Z_i\}$ has moment generating function, a satisfactory answer can be given, with Esscher transform (see (4.5)) playing a major role. Under suitable conditions, Esscher transform converts an event of "small" probability into one of "considerable" probability, thus facilitating analysis.

It turns out that the above question in terms of rate function is meaningful in problems arising in diverse disciplines. A far reaching abstract framework of large deviation principles has been synthesized by Indian born probabilist S.R.S. Varadhan with implications in statistics, PDE theory, physics, engineering, etc. For this outstanding work Varadhan has been awarded the prestigious Abel prize for 2007. For a quick overview, an interested reader may see M. Raussen and C. Skau: Interview with Srinivasa Varadhan, *Notices of the American Mathematical Society*, vol. 55 (2008) Feb. 2008 issue, pp. 238-246, and R. Bhatia: A Conversation with S.R.S. Varadhan, *The Mathematical Intelligencer*, vol. 30, No. 2, Spring 2008, pp. 24-42. The role of Esscher tilt is highlighted in these interviews. See also S. Ramasubramanian: Large deviations: an introduction to 2007 Abel prize. *Proceedings of the Indian Academy of Sciences (Mathematical Sciences)*, 118 (2008), 161-182.

## 5.2 Lundberg bound

We continue with the renewal risk model. We assume that the net profit condition holds; see Definition 5.3 in the preceding section. We also assume that the claim size distribution is light tailed, that is, the claim size $X_1$ has moment generating function in a neighbourhood of zero.

**Definition 5.4** Let $Z_1 = X_1 - cA_1$, and assume that the moment generating function $\hat{m}_Z(\cdot)$ of $Z_1$ exists in $(-h_0, h_0)$, where $h_0 > 0$. Suppose $h = r$ is the *unique positive* solution to the equation

$$\hat{m}_Z(h) = E[\exp(h(X_1 - cA_1))] = 1. \tag{5.14}$$

Then $r$ is called the *adjustment coefficient or Lundberg coefficient.*

**Proposition 5.5** *Assume that (i) $E(Z_1) < 0$, that is, the net profit condition holds, (ii) there exists $h_0 > 0$ such that $\hat{m}_Z(h) < \infty$ for all $|h| < h_0$, and $\lim_{h \uparrow h_0} \hat{m}_Z(h) = \infty$. Then the Lundberg coefficient is well defined.*

**Proof:** Clearly $h = 0$ is a solution to (5.14). So we want a unique solution to (5.14) in $(0, h_0)$. Put $f(h) = \hat{m}_Z(h) = E(e^{hZ_1})$, $h \in (-h_0, h_0)$. Note that $f$ has derivatives of all orders in $(-h_0, h_0)$. Also $f'(h) = E(Z_1 e^{hZ_1})$, $f''(h) = E(Z_1^2 e^{hZ_1})$, for $h \in (-h_0, h_0)$. Since (NPC) holds, $f'(0) = E(Z_1) < 0$. This together with continuity of $f$ implies that $f$ decreases in a neighbourhood of 0. But as $Z_1 \not\equiv 0$ (that is, $P(Z_1 = 0) < 1$), we see that $f''(h) = E(Z_1^2 e^{hZ_1}) > 0$; $f$ is convex. As $f(h) \to \infty$ as $h \uparrow h_0$, it now follows that there is $h_1 > 0$ such that $f'(h_1) = 0$, $f$ is increasing on $(h_1, h_0)$, and there is a unique $r \in (h_1, h_0) \subset (0, h_0)$ such that $f(r) = 1$. This completes the proof. □

The next result gives a probabilistic significance of the Lundberg coefficient.

**Proposition 5.6** *Assume that the (NPC) holds and that Lundberg coefficient $r$ exists. Let $\{S_n : n \geq 0\}$ be the random walk as in Section 5.1. Then $\{e^{rS_n} : n = 0, 1, 2, \ldots\}$ is a martingale with respect to $\{\mathcal{G}_n\}$.*

**Proof:** Integrability of $e^{rS_n}$ is clear. It is left as an exercise to the reader to show that $E[e^{rS_{n+1}} \mid \mathcal{G}_n] = e^{rS_n}$ a.s. See Exercise 5.2.2. □

**Theorem 5.7** *(Lundberg bound): Consider the renewal risk model. Assume that the (NPC) holds and that the Lundberg coefficient $r > 0$ exists. Then*

$$\psi(x) \leq \exp(-rx), \quad x > 0. \tag{5.15}$$

**Proof:** We give two proofs.

**1st Proof:** Let $\tau_0$ denote the ruin time. For $k = 1, 2, \ldots$ note that $\tau_0 \wedge k$ is a stopping time w.r.t. $\{\mathcal{G}_n\}$. So by the preceding proposition and the optional sampling theorem $E[\exp(rS_{(\tau_0 \wedge k)})] = E[e^{rS_0}] = 1$. Therefore $E[e^{rS_{\tau_0}} I_{\{\tau_0 \leq k\}} + e^{rS_k} I_{\{\tau_0 > k\}}] = 1$. Note that $S_{\tau_0} > x$, where $x$ is the initial capital. Hence

$$E[e^{rx} \cdot I_{\{\tau_0 \leq k\}}] \leq E[e^{rS_{\tau_0}} \cdot I_{\{\tau_0 \leq k\}}] \leq 1$$

whence it follows that $P_x(\tau_0 \leq k) \leq e^{-rx}$. Letting $k \uparrow \infty$ we get (5.15).

**2nd Proof:** This proof is by induction. Write $\psi_n(x) = P(\max_{1 \leq k \leq n} S_k > x)$, $x > 0$, $n = 1, 2, \ldots$. Note that $\psi_n(x) \uparrow \psi(x)$ as $n \uparrow \infty$ for every $x > 0$. So it is enough to prove that

$$\psi_n(x) \leq \exp(-rx), \ \forall \ n \geq 1, \ x > 0. \tag{5.16}$$

The case $n = 1$ is dealt with in Exercise 5.2.3. Assume that (5.16) holds for $n \leq k$. We need to show that it holds for $n = k + 1$.

Let $F_Z$ denote the distribution function of $Z_1$. As $Z_2 + \ldots + Z_i$ is independent of $Z_1$, and $Z_2 + \ldots + Z_{j+1} \stackrel{d}{=} S_j$, we obtain

$$\begin{aligned}
\psi_{k+1}(x) &= P(Z_1 > x) + P\left(Z_1 \leq x, \left(\max_{2 \leq i \leq k+1} S_i\right) > x\right) \\
&= P(Z_1 > x) \\
&\quad + \int_{(-\infty, x]} P\left(\left(\max_{2 \leq i \leq k+1} S_i\right) > x - z \mid Z_1 = z\right) \, dF_Z(z) \\
&= P(Z_1 > x) + \int_{(-\infty, x]} P\left(\max_{1 \leq j \leq k} S_j > x - z\right) \, dF_Z(z) \\
&= p_1 + p_2, \text{ say.}
\end{aligned} \tag{5.17}$$

By induction hypothesis note that

$$p_2 = \int_{(-\infty, x]} \psi_k(x - z) \, dF_Z(z) \leq \int_{(-\infty, x]} e^{-r(x-z)} \, dF_Z(z). \tag{5.18}$$

Next, by proof of Chebyshev's inequality

$$\begin{aligned}
p_1 &= P(Z_1 > x) \leq E\left[\frac{e^{rZ_1}}{e^{rx}} : \{Z_1 > x\}\right] \\
&\leq e^{-rx} \int_{(x, \infty)} e^{rz} \, dF_Z(z) = \int_{(x, \infty)} e^{-r(x-z)} \, dF_Z(z).
\end{aligned} \tag{5.19}$$

It is now clear that (5.17) - (5.19) give (5.16) for $n = k + 1$, completing the proof. □

**Example 5.8** Consider the Cramér-Lundberg model with claim sizes $\{X_i\}$ having Exp($\theta$) distribution and interarrival times $\{A_i\}$ having

$\text{Exp}(\lambda)$ distribution. It can be verified that

$$\hat{m}_Z(h) = \frac{\theta}{\theta - h} \cdot \frac{\lambda}{\lambda + ch}, \text{ if } \frac{-\lambda}{c} < h < \theta \tag{5.20}$$

and hence that $\hat{m}_Z(h) \uparrow \infty$ as $h \uparrow \theta$. Note that

$$\text{(NPC) holds } \Leftrightarrow \frac{1}{\theta} - \frac{c}{\lambda} < 0 \Leftrightarrow \theta - \frac{\lambda}{c} > 0. \tag{5.21}$$

Clearly

$$\hat{m}_Z(h) = 1,\ h > 0 \Leftrightarrow h = \theta - \frac{\lambda}{c} > 0. \tag{5.22}$$

Thus, if (NPC) holds then the Lundberg coefficient is $r = \theta - \frac{\lambda}{c}$.

Now suppose premium rate $c$ is determined using expected value principle so that

$$ct = (1 + \rho)\ E(S(t)) = (1 + \rho)\ \frac{\lambda}{\theta}\ t \tag{5.23}$$

where $\rho > 0$ is the safety loading factor. Clearly (NPC) holds as $E(Z_1) = -\frac{\rho}{\theta} < 0$. So the Lundberg coefficient is $r = \frac{\rho}{(1+\rho)}\ \theta$, and hence the Lundberg bound is

$$\psi(x) \leq \exp\left\{-\theta\ \frac{\rho}{1+\rho}\ x\right\},\ x > 0. \tag{5.24}$$

From the above we see that ruin is unlikely if the initial capital is large. Also note that $\frac{\rho}{1+\rho} \uparrow 1$ as $\rho \to \infty$; so the bound does not change significantly if $\rho$ is very large. On the other hand, for large $\rho$ the premium will be high, and hence the policy may look unattractive to a buyer.

## 5.3 An integral equation in Cramér-Lundberg model

As we know, in the Cramér-Lundberg model the claim number process $\{N(t) : t \geq 0\}$ is a homogeneous Poisson process with rate $\lambda > 0$; so the interarrival times $\{A_i\}$ is an i.i.d. sequence having $\text{Exp}(\lambda)$ distribution. We assume $X_i > 0$ a.s. and $\mu = E(X_i)$ where $\{X_i\}$ denote i.i.d. claim sizes independent of $\{N(t)\}$. Set

$$\varphi(u) = 1 - \psi(u) = P(\tau_0 = \infty \mid R(0) = u). \tag{5.25}$$

The function $\varphi$ is called the *survival probability.*

We assume the net profit condition

$$\mu - \frac{c}{\lambda} < 0 \text{ or equivalently } \frac{c}{\lambda\mu} > 1. \tag{5.26}$$

We will derive an integral equation for the survival probability $\varphi$. Let $Z_i = X_i - cA_i$, for $i \geq 1$, and $\{S_n\}$ be the random walk as in Section 5.1. Let $F_A$, $F_X$, $F_Z$ denote respectively the distribution functions of $A_i$, $X_i$, $Z_i$. As $\{Z_i\}$ is an i.i.d. sequence note that $S_n - Z_1 \stackrel{d}{=} S_{n-1}$ for $n \geq 2$. Consequently as $X_1$ and $A_1$ are independent and nonnegative, we have for $x \geq 0$

$$\begin{aligned}
& \varphi(x) = P(S_n \leq x \text{ for all } n \geq 1) \\
= \; & P(Z_1 \leq x,\ S_n - Z_1 \leq x - Z_1 \text{ for all } n \geq 2) \\
= \; & \int_{\mathbb{R}} I_{(-\infty,x]}(z) P\left(\sum_{i=2}^{n} Z_i \leq x - z,\ \forall\ n \geq 2 \mid Z_1 = z\right) dF_Z(z) \\
= \; & \int_{\mathbb{R}} I_{(-\infty,x]}(z)\ P(S_m \leq (x-z)\ \forall\ m \geq 1)\ dF_Z(z) \\
= \; & \int_0^{\infty} \int_0^{x+cw} P(S_m \leq x - (u - cw)\ \forall\ m \geq 1)\ dF_X(u)\ dF_A(w) \\
= \; & \int_0^{\infty} \int_0^{x+cw} \varphi(x - u + cw)\ dF_X(u)\ \lambda e^{-\lambda w}\ dw \\
= \; & \frac{\lambda}{c}\ e^{\frac{\lambda}{c}x} \int_x^{\infty} e^{-\frac{\lambda}{c}z} \int_0^{z} \varphi(z-u)\ dF_X(u)\ dz
\end{aligned} \tag{5.27}$$

where we have used the fact that $A_1$ has Exp($\lambda$) distribution. Hence $\varphi(\cdot)$ is absolutely continuous and

$$\varphi'(x) = \frac{\lambda}{c}\ \varphi(x) - \frac{\lambda}{c} \int_0^{x} \varphi(x-u)\ dF_X(u), \text{ for a.e. } x > 0. \tag{5.28}$$

Note that (5.28) holds a.e. with respect to the Lebesgue measure. By integration by parts

$$\int_0^{x} \varphi(x-u)\ dF_X(u) = F_X(x)\ \varphi(0) + \int_0^{x} F_X(u)\ \varphi'(x-u)\ du.$$

So integrating (5.28) over $(0,t)$ we get

$$\begin{aligned}
&\varphi(t)-\varphi(0)\\
&= \frac{\lambda}{c}\int_0^t \varphi(x)dx - \frac{\lambda}{c}\int_0^t \left[\varphi(0)F_X(x) + \int_0^x F_X(u)\varphi'(x-u)du\right] dx\\
&= \frac{\lambda}{c}\int_0^t \varphi(t-x)dx - \frac{\lambda}{c}\varphi(0)\int_0^t F_X(x)\ dx\\
&\quad -\frac{\lambda}{c}\int_0^t\int_0^x F_X(u)\varphi'(x-u)\ du\ dx\\
&= \frac{\lambda}{c}\int_0^t \varphi(t-x)dx - \frac{\lambda}{c}\varphi(0)\int_0^t F_X(x)\ dx\\
&\quad -\frac{\lambda}{c}\int_0^t [\varphi(t-u)-\varphi(0)]F_X(u)\ du\\
&= \frac{\lambda}{c}\int_0^t \varphi(t-x)(1-F_X(x))dx.
\end{aligned}$$

Consequently, as $\overline{F}_X(x) = 1 - F_X(x)$ we have

$$\varphi(t) = \varphi(0) + \frac{\lambda}{c}\int_0^t \varphi(t-x)\ \overline{F}_X(x)\ dx. \tag{5.29}$$

As the net profit condition holds, by Exercise 5.1.7 (see below), $\varphi(\infty) \triangleq \lim_{t\to\infty} \varphi(t) = 1$. So letting $t \uparrow \infty$ in (5.29) we get

$$\varphi(0) = 1 - \frac{\lambda\mu}{c}. \tag{5.30}$$

We know that the integrated tail distribution corresponding to $F_X$ is given by $dF_{X,I}(x) = \frac{1}{\mu}\ \overline{F}_X(x)\ dx$. So, using (5.30), we can now write (5.29) as

$$\varphi(t) = \left(\frac{c-\lambda\mu}{c}\right) + \frac{\lambda\mu}{c}\int_0^t \varphi(t-x)\ dF_{X,I}(x). \tag{5.31}$$

Denote $\rho = \frac{c}{\lambda\mu} - 1$; by the net profit condition (5.26) note that $\rho > 0$; this is a safety loading factor. We can summarize the above discussion in the form of the following theorem.

**Theorem 5.9** *Consider the Cramér-Lundberg model. Assume that the claim sizes $X_i > 0$ a.s., and that the net profit condition (5.26) holds. Then the survival probability $\varphi(\cdot)$ satisfies the integral equation*

$$\varphi(t) = \frac{\rho}{1+\rho} + \frac{1}{1+\rho}\int_0^t \varphi(t-x)\, dF_{X,I}(x),\ t \geq 0 \tag{5.32}$$

*where $\rho = \left(\frac{c}{\lambda\mu} - 1\right) > 0$, and $F_{X,I}$ is the integrated tail distribution of the claim size distribution $F_X$.* □

Though the integral equation (5.31), or equivalently (5.32), looks like the renewal equation (3.28) in Definition 3.15, there is a crucial difference. Note that $\frac{1}{(1+\rho)}\, dF_{X,I}(\cdot)$ is only a subprobability measure. So (5.31) or (5.32) is called a *defective renewal equation.*

Even though $\frac{1}{(1+\rho)}\, dF_{X,I}(\cdot)$ is not a probability measure, mimicking the proof of Theorem 3.16 we get the following

**Theorem 5.10** *Let the notation and hypotheses be as in Theorem 5.9. Then the unique bounded solution, supported on $[0,\infty)$, to the defective renewal equation (5.32) is given by*

$$\varphi(t) = \frac{\rho}{1+\rho}\sum_{n=0}^{\infty} \frac{1}{(1+\rho)^n}\, F_{X,I}^{*(n)}(t),\ t \geq 0 \tag{5.33}$$

*where $F_{X,I}^{*(n)}$ denotes the n-fold convolution of $F_{X,I}$. Consequently*

$$\psi(t) = 1 - \varphi(t) = \frac{\rho}{1+\rho}\sum_{n=1}^{\infty} \frac{1}{(1+\rho)^n}\, \overline{F_{X,I}^{*(n)}}(t),\ t \geq 0. \tag{5.34}$$

*(5.33) or equivalently (5.34) is called the Pollaczek-Khinchin formula.* □

The following definition gives a way of describing formulae like (5.33), (5.34).

**Definition 5.11** Let $0 < p < 1$. Let $J$ be a random variable having a geometric distribution so that $P(J = k) = (1-p)p^k$, $k = 0, 1, 2, \ldots$. Note that $p = P(J \geq 1)$. Let $\{U_i : i \geq 1\}$ be an i.i.d. sequence with common distribution function $F_U$ supported on $(0, \infty)$. Assume that $J$ and $\{U_i\}$ are independent. Define

$$V \triangleq \sum_{i=1}^{J} U_i = \begin{cases} 0 \text{ if } J = 0 \\ U_1 + \ldots + U_k, \text{ if } J = k. \end{cases} \tag{5.35}$$

Then $V$ is said to have a *compound geometric distribution* (or $V$ is *geometric compound*) with characteristics $(p, F_U)$.

In view of Exercises 5.3.4, 5.3.5, the Pollaczek-Khinchin formula can now be interpreted as follows: Survival probability $\varphi(\cdot)$ is the geometric compound with characteristics $\left(\frac{1}{1+\rho}, F_{X,I}\right)$; similarly ruin probability $\psi(\cdot)$ is the tail of the above compound geometric distribution.

## 5.4 Exact asymptotics in Cramér-Lundberg model

In this section we look at the problem of exact asymptotic behaviour of ruin probability in the Cramér-Lundberg model. First we consider light-tailed claim size distributions, and then take up the case when claim sizes are subexponential. As we will see, the qualitative behaviour in these cases are quite different.

### 5.4.1 Light-tailed claims

To obtain the Pollaczek-Khinchin formulae (5.33), (5.34) it was sufficient to follow the approach in renewal theory, as in Theorem 3.16 for solving the renewal equation. However, imitation is not good enough to get the exact asymptotics of ruin probability. The reason is the following. As $\psi(\cdot)$ satisfies a renewal type equation, one would like to use the key renewal theorem (Theorem 3.26 (b)) to get the asymptotics. We know that the key renewal theorem is equivalent to Blackwell's renewal theorem (Theorem 3.21). One ingredient in the proof of Blackwell's theorem is the Choquet-Deny theorem, viz. constants are the only solutions to the convolution equation (3.44). This conclusion, however, is not valid unless $F$ in (3.44) is a probability distribution; see Exercise 5.4.1. Our analysis below will also indicate that the key renewal theorem is not valid if $F$ is a subprobability measure, as is the case in (5.32).

To deal with the situation we use an Esscher transform (or exponentially tilted transformation) of the subprobability distribution $\frac{1}{(1+\rho)}\, dF_{X,I}(\cdot)$. So we *assume* that there is $\nu > 0$ such that

$$\int_0^\infty e^{\nu x}\, \frac{1}{(1+\rho)}\, dF_{X,I}(x) = 1. \tag{5.36}$$

Then the defective renewal equation for $\psi(\cdot)$ given in Exercise 5.3.2 (see below) leads to

$$\begin{aligned} e^{\nu t}\psi(t) &= e^{\nu t}\left[\frac{1}{1+\rho}\,\overline{F}_{X,I}(t) + \frac{1}{1+\rho}\int_0^t \psi(t-x)\, dF_{X,I}(x)\right] \\ &= e^{\nu t}\frac{1}{(1+\rho)}\overline{F}_{X,I}(t) \\ &\quad + \int_0^t [e^{\nu(t-x)}\psi(t-x)]\left\{e^{\nu x}\frac{1}{(1+\rho)}dF_{X,I}(x)\right\}. \end{aligned} \tag{5.37}$$

By (5.36), $e^{\nu x}\frac{1}{(1+\rho)}\, dF_{X,I}(x)$ is a probability distribution supported on $(0,\infty)$. Therefore (5.37) is a *standard renewal equation* for the unknown function $e^{\nu t}\psi(t)$, with $e^{\nu t}\frac{1}{(1+\rho)}\,\overline{F}_{X,I}(t)$ as the given function. The assumption (5.36) is called the *Cramér-Lundberg condition*; by Exercise 5.4.2 $\nu > 0$ in (5.36) is the Lundberg coefficient as defined in Section 5.2. We now have the following result.

**Theorem 5.12** *(Cramér-Lundberg approximation) Consider the Cramér-Lundberg model. Assume the net profit condition (5.26); denote* $\rho = \frac{c}{\lambda\mu} - 1$. *Let the Cramér-Lundberg condition (5.36) hold, and*

$$\int_0^\infty x\, e^{\nu x}\overline{F}_X(x)\, dx < \infty \tag{5.38}$$

*where* $\nu > 0$ *is as in (5.36). Then*

$$\lim_{t\to\infty}\, e^{\nu t}\psi(t) = C^{-1} \tag{5.39}$$

*where*

$$C = \frac{\nu}{\rho\mu} \int_0^\infty x\, e^{\nu x} \overline{F}_X(x)\, dx. \tag{5.40}$$

*In particular the ruin probability $\psi(\cdot)$ decays exponentially.*

**Proof:** In view of the preceding discussion we would like to apply the key renewal theorem to renewal equation (5.37). For this put $u_1(t) = e^{\nu t}$, $u_2(t) = \frac{1}{(1+\rho)} \overline{F}_{X,I}(t)$, $t \geq 0$. We want to establish that $u_1(\cdot)u_2(\cdot)$ is directly Riemann integrable using Lemma 3.24; for this note that it is enough to show that $u_1(\cdot)u_2(\cdot)$ is integrable on $(0, \infty)$. Indeed

$$\begin{aligned}
\int_0^\infty u_1(t)u_2(t)\, dt &= \int_0^\infty e^{\nu t} \frac{1}{(1+\rho)} \overline{F}_{X,I}(t)\, dt \\
&= \int_0^\infty e^{\nu t} \frac{1}{(1+\rho)} \int_t^\infty \frac{1}{\mu} \overline{F}_X(x) dx\, dt \\
&= \frac{1}{\mu(1+\rho)} \int_0^\infty \overline{F}_X(x) \int_0^x e^{\nu t} dt\, dx \\
&= \frac{1}{\nu\mu(1+\rho)} \left[ \int_0^\infty e^{\nu x} \overline{F}_X(x)\, dx - \int_0^\infty \overline{F}_X(x) dx \right] \\
&= \frac{1}{\nu} \left( 1 - \frac{1}{1+\rho} \right) = \frac{\rho}{\nu(1+\rho)}
\end{aligned} \tag{5.41}$$

by (5.36). Therefore applying the key renewal theorem (Theorem 3.26 (b)) to the renewal equation (5.37) we get

$$\begin{aligned}
&\lim_{t\to\infty} e^{\nu t}\psi(t) \\
&= \left[ \int_0^\infty x\, e^{\nu x} \frac{1}{(1+\rho)\mu} \overline{F}_X(x)\, dx \right]^{-1} \left[ \int_0^\infty e^{\nu t} \frac{1}{(1+\rho)} \overline{F}_{X,I}(t)\, dt \right] \\
&= \frac{\rho}{\nu(1+\rho)} \mu(1+\rho) \left( \int_0^\infty x\, e^{\nu x} \overline{F}_X(x)\, dx \right)^{-1}
\end{aligned}$$

whence the result follows; in the above we have used (5.38), (5.41).

□

**Example 5.13** This concerns Cramér-Lundberg approximation when the claim size has an exponential distribution. In addition to the hypotheses of Theorem 5.12 assume that $X_i$ has Exp($\theta$) distribution. It is easily proved that the integrated tail distribution is

$$F_{X,I}(x) = F_X(x) = (1 - e^{-\theta x}),\ x \geq 0. \tag{5.42}$$

From Example 5.8 we know that the Lundberg coefficient is $\nu = r = \frac{\rho}{(1+\rho)}\theta$. Therefore it can easily be verified that the probability distribution in the standard renewal equation (5.37) is given by

$$dG(x) \triangleq \frac{\theta}{(1+\rho)} \exp\left\{-\frac{\theta}{(1+\rho)}x\right\} \cdot I_{(0,\infty)}(x)\ dx. \tag{5.43}$$

So, by Theorem 3.16, the solution to (5.37) can be given in terms of an integral with respect to the renewal measure $\sum_{n=0}^{\infty} G^{*(n)}(\cdot)$ corresponding to $G(\cdot)$. From (5.43) it is clear that $G(\cdot)$ is the Exp($\frac{\theta}{1+\rho}$) distribution. Hence by Theorem 2.5, the corresponding renewal process is the Poisson process with rate ($\frac{\theta}{1+\rho}$). Consequently the renewal function is $m(t) = 1 + \frac{\theta}{1+\rho}t,\ t \geq 0$. Thus the renewal measure has a jump at 0, and hence

$$\sum_{n=0}^{\infty} d\ G^{*(n)}(s) = \delta_0(ds) + \frac{\theta}{(1+\rho)}ds. \tag{5.44}$$

Therefore, by Theorem 3.16

$$\begin{aligned}
e^{\nu t}\psi(t) &= \int_0^t e^{\nu(t-x)} \frac{1}{(1+\rho)} e^{-\theta(t-x)} \left(d\sum_{n=0}^{\infty} G^{*(n)}(x)\right) \\
&= \frac{1}{1+\rho} e^{-(\theta-\nu)t} + \int_0^t \frac{1}{(1+\rho)} e^{-(\theta-\nu)(t-x)} \cdot \frac{\theta}{(1+\rho)} dx \\
&= \frac{1}{(1+\rho)} e^{-\frac{\theta}{(1+\rho)}t} + \int_0^t \frac{\theta}{(1+\rho)^2} \exp\left\{-\frac{\theta}{1+\rho}(t-x)\right\} dx \\
&= \frac{1}{1+\rho}.
\end{aligned}$$

Hence we have

$$\begin{aligned}\psi(t) &= \frac{1}{(1+\rho)}e^{-\nu t} = \frac{1}{(1+\rho)}\exp\left\{-\frac{\rho\theta}{(1+\rho)}t\right\} \\ &= \frac{\lambda}{c\theta}\exp\left\{-\left(\frac{c\theta-\lambda}{c}\right)t\right\},\ t \geq 0. \end{aligned} \tag{5.45}$$

Thus we have an explicit expression for the ruin probability in this case. □

### 5.4.2 Subexponential claims

We had already mentioned in an earlier section that the distributions that actually fit claim size data are heavy tailed distributions like the Pareto, lognormal or heavy tailed Weibull distributions. Clearly the discussion in subsection 5.4.1 does not apply to such distributions.

To get a handle on the situation we return to the Pollaczek-Khinchin formula (5.34) for the ruin probability $\psi(\cdot)$. Recall that essentially only the net profit condition is assumed here in the Cramér-Lundberg model. Dividing (5.34) by $\overline{F}_{X,I}(t)$ we get

$$\frac{\psi(t)}{\overline{F}_{X,I}(t)} = \frac{\rho}{(1+\rho)}\sum_{n=1}^{\infty}(1+\rho)^{-n}\frac{\overline{F^{*(n)}}(t)}{\overline{F}_{X,I}(t)}. \tag{5.46}$$

Now, if $F_{X,I} \in \mathcal{S}$, (that is, if $F_{X,I}$ is subexponential), then

$$\overline{F_{X,I}^{*(n)}}(t)/\overline{F}_{X,I}(t) \to n \text{ as } t \to \infty, \text{ for each } n.$$

By net profit condition, $\rho = \frac{c}{\lambda\mu} - 1 > 0$. Choose $0 < \epsilon < \rho$. Then by Kesten's lemma (Lemma 4.22) there is constant $0 < K < \infty$ such that

$$(1+\rho)^{-n}\overline{F_{X,I}^{*(n)}}(t)/\overline{F}_{X,I}(t) \leq K\left(\frac{1+\epsilon}{1+\rho}\right)^n,\ \forall\, t \geq 0.$$

As $\sum\left(\frac{1+\epsilon}{1+\rho}\right)^n < \infty$, by (5.46) and the dominated convergence theorem (applied to the counting measure) we get

$$\lim_{t\to\infty}\frac{\psi(t)}{\overline{F}_{X,I}(t)} = \frac{\rho}{1+\rho}\sum_{n=1}^{\infty}n(1+\rho)^{-n} = \frac{1}{\rho}.$$

Thus we have proved the following important result

**Theorem 5.14** *(Embrechts-Veraverbeke approximation): Consider the Cramér-Lundberg model with the net profit condition (5.26); so safety loading factor* $\rho = \frac{c}{\lambda\mu} - 1 > 0$. *Let* $F$ *denote the claim size distribution, and* $F_I$ *the corresponding integrated tail distribution. Assume that* $F_I \in \mathcal{S}$. *Then*

$$\lim_{t\to\infty} \frac{\psi(t)}{\overline{F}_I(t)} = \frac{1}{\rho}. \tag{5.47}$$

□

**Example 5.15** Consider the Cramér-Lundberg model with the net profit condition. Suppose the claim size distribution $F$ is Pareto $(\alpha)$, $\alpha > 1$. As $\alpha > 1$ we know that the claim size distribution has finite expectation. Also the integrated tail distribution $F_I$ is Pareto $(\alpha - 1)$ and hence is subexponential. So by Theorem 5.14

$$\psi(t) = \frac{1}{\rho} \frac{\kappa^{\alpha-1}}{(\kappa + t)^{\alpha-1}} (1 + o(1)), \text{ as } t \to \infty. \tag{5.48}$$

Here $\kappa > 0$ is a constant.

When claim size is Pareto $(\alpha)$, $\alpha > 1$, (5.48) indicates that the ruin probability decays only at a power rate. Exercise 5.4.5 shows that a similar behaviour is exhibited when claim size distribution is $\mathcal{R}_{-\alpha}$, with $\alpha > 1$; (in fact, Example 5.15 is a special case of Exercise 5.4.5). This is in sharp contrast to the situation in Theorem 5.12 where the ruin probability decays exponentially. So Pareto distributed risks, and risks with regularly varying tails are more dangerous than light-tailed risks.

For Cramér-Lundberg model with net profit condition $\rho > 0$, the following are, in fact, equivalent:
(a) $F_I \in \mathcal{S}$; (b) $\varphi(\cdot) = 1 - \psi(\cdot) \in \mathcal{S}$; (c) $\lim_{t\to\infty} \frac{\psi(t)}{\overline{F}_I(t)} = \rho^{-1}$.
While (a) $\Rightarrow$ (c) is proved in Theorem 5.14, (a) $\Rightarrow$ (b) is given in Exercise 5.4.4. For a proof the reader is referred to an appendix in [EKM].

Our presentation of this section is heavily influenced by [EKM].

**Remark 5.16** For the Sparre Andersen model also similar analysis is possible. For analogues of Theorems 5.10, 5.12, 5.14 see, respectively, Theorems 6.5.1, 6.5.7, 6.5.11 in Section 6.5 of [RSST]. These results make use of the so called ladder height distributions.

## 5.5 Exercises

### Exercises on Section 5.1

**5.1.1.** (a) Show that $\tau_0$ is a stopping time with respect to the filtration $\{\mathcal{F}_t : t \geq 0\}$, where $\mathcal{F}_t$ is the smallest $\sigma$-algebra making the family $\{S(r) : 0 \leq r \leq t\}$ measurable; that is, show that $\{\tau_0 \leq t\} \in \mathcal{F}_t$ for all $t$.

(Hint: $\{\tau_0 > t\} = \{R(s) \geq 0 \text{ for all } s \leq t\} = \cap\{R(s) \geq 0 : s \leq t,\ s \text{ rational}\} \cap \{R(t) \geq 0\}$. Here we are using the special structure of $R(\cdot)$, viz. it is increasing except at jumps which are negative.)

(b) Show that "Ruin" is an event.

**5.1.2.** Show that $\inf_{t \geq 0} R(t)$ is a (possibly extended real valued) random variable.

**5.1.3.** As $E(Z_1)$ exists, prove assertions (a), (b) of Theorem 5.1 using the strong law of large numbers.

**5.1.4.** (a) Let $j = 0, 1, 2, \ldots$. Show that $\{\nu_j^+ = k\},\ \{\nu_j^- = k\} \in \mathcal{G}_k,\ k = 0, 1, 2, \ldots$; that is, $\nu_j^+,\ \nu_j^-$ are stopping times w.r.t. $\{\mathcal{G}_n\}$.

(b) Show that $S_{\nu_n^+}$ (resp. $S_{\nu_n^-}$) is a random variable on $\{\nu_n^+ < \infty\}$ (resp. $\{\nu_n^- < \infty\}$).

(c) Show that $S_{\nu_{k+1}^+} - S_{\nu_k^+}$ takes values in $(0, \infty)$.

**5.1.5.** (a) For any interval $(a, b) \subset (0, \infty)$ show that

$$\begin{aligned} & P(S_{\nu_1^+} - S_{\nu_0^+} \in (a, b)) \\ = & \sum_{n=1}^{\infty} P(S_j \in (-\infty, 0],\ j \leq n-1,\ S_n \in (a, b)). \end{aligned}$$

(b) For $i = 0, 1, 2, \ldots,\ k = 0, 1, \ldots$ show that $\{\nu_k^+ = i\}$ is independent of $\{Z_{i+\ell} : \ell \geq 1\}$; in fact $\mathcal{G}_i$ and $\{Z_{i+\ell} : \ell \geq 1\}$ are independent.

(c) Let $k \geq 1$. For any interval $(a, b) \subset (0, \infty)$ show that

$$\begin{aligned} & P(S_{\nu_{k+1}^+} - S_{\nu_k^+} \in (a, b)) \\ = & \sum_{i=1}^{\infty} \left( \sum_{n=1}^{\infty} P \left\{ \begin{array}{l} \left( \sum_{\ell=1}^{j} Z_{i+\ell} \right) \in (-\infty, 0],\ j \leq n-1, \\ \left( \sum_{\ell=1}^{n} Z_{i+\ell} \right) \in (a, b) \end{array} \right\} \right) \\ & \cdot P(\nu_k^+ = i) \\ = & \sum_{n=1}^{\infty} P(S_j \in (-\infty, 0],\ j \leq n-1,\ S_n \in (a, b)). \end{aligned}$$

Thus $(S_{\nu_{k+1}^+} - S_{\nu_k^+})$ has the same distribution as $(S_{\nu_1^+} - S_{\nu_0^+})$ if $\nu_{k+1}^+ < \infty$ a.s. for $k = 0, 1, 2, \ldots$.

(d) Assume $\nu_2^+ < \infty$ a.s. Let $B_1,\ B_2 \subset (0, \infty)$ be Borel sets. Using the fact that $\{Z_q : q \geq 1\}$ is an i.i.d. family show that

$$\begin{aligned}
& P(S_{\nu_1^+} - S_{\nu_0^+} \in B_1,\ S_{\nu_2^+} - S_{\nu_1^+} \in B_2) \\
= & \sum_{k_1,k_2=1}^{\infty} P(S_j \in (-\infty, 0],\ j \leq k_1 - 1,\ S_{k_1} \in B_1) \\
& \cdot P\left(\left(\sum_{\ell=1}^{m} Z_{k_1+\ell}\right) \in (-\infty, 0], m \leq k_2 - 1, \left(\sum_{\ell=1}^{k_2} Z_{k_1+\ell}\right) \in B_2\right) \\
= & \left[\sum_{k_1=1}^{\infty} P(S_j \in (-\infty, 0],\ j \leq k_1 - 1,\ S_{k_1} \in B_1)\right] \cdot \\
& \left[\sum_{k_2=1}^{\infty} P(S_m \in (-\infty, 0],\ m \leq k_2 - 1,\ S_{k_2} \in B_2)\right] \\
= & P(S_{\nu_1^+} - S_{\nu_0^+} \in B_1) \cdot P(S_{\nu_1^+} - S_{\nu_0^+} \in B_2).
\end{aligned}$$

Hence conclude that $(S_{\nu_1^+} - S_{\nu_0^+})$ and $(S_{\nu_2^+} - S_{\nu_1^+})$ are i.i.d.'s taking values in $(0, \infty)$, provided $\nu_2^+ < \infty$ a.s.

(e) By an analogous argument, assuming $\nu_j^+ < \infty$ a.s., $j = 0, 1, 2, \ldots$ show that $\{S_{\nu_{j+1}^+} - S_{\nu_j^+} : j = 0, 1, 2, \ldots\}$ is an i.i.d. sequence taking values in $(0, \infty)$.

**5.1.6.** (a) Let $\hat{M}$ be given by (5.12) with $k \geq 1$ fixed. For $n = 0, 1, 2, \ldots$ show that $P(\hat{M} = n)$

$$\begin{aligned}
= & \sum_{N=1}^{\infty} P(\nu_k^+ = N) \cdot P\left(\sum_{\ell=N+j+1}^{N+n} Z_\ell > 0,\ \forall\ j \leq n - 1\right) \cdot \\
& P\left(\sum_{q=N+n+1}^{N+j} Z_q \leq 0,\ \forall\ j > n\right).
\end{aligned}$$

(b) Using (a), derive (5.13) and hence (5.11).

**5.1.7.** Consider the renewal risk model with the net profit condition (NPC) as given in Definition 5.3.

(a) Show that $\{\sup_{n \geq 0} S_n > x\}$ decreases to a set of measure zero as $x \uparrow \infty$.

(b) Show that $\lim_{x\to\infty} \psi(x) = 0$.

(c) Show that $R(t) \to +\infty$ a.s. as $t \uparrow \infty$.

**Exercises on Section 5.2**

**5.2.1. a)** Show that the moment generating function of $Z_1$ exists in a neighbourhood of 0 if the moment generating functions of $X_1$ and $cA_1$ do.

b) In the Cramér-Lundberg model, $\hat{m}_Z$ exists in a neighbourhood of 0 if $\hat{m}_X$ does.

**5.2.2.** Complete the details of the proof of Proposition 5.6.

**5.2.3.** Applying Chebyshev's inequality to the random variable $e^{rZ_1}$ appropriately, where $r$ is the Lundberg coefficient, establish (5.16) for $n = 1$.

**5.2.4.** Verify (5.20) - (5.22) in Example 5.8.

**Exercises on Section 5.3**

**5.3.1.** Let the claim size distribution function $F_X$ have a probability density function. Show that the survival probability function $\varphi(\cdot)$ is differentiable and that $\varphi'(\cdot)$ is given by (5.28) for all $x > 0$.

**5.3.2.** Under the hypotheses of Theorem 5.9 show that the ruin probability $\psi(\cdot)$ satisfies the integral equation

$$\psi(t) = \frac{1}{(1+\rho)} \overline{F}_{X,I}(t) + \frac{1}{(1+\rho)} \int_0^t \psi(t-x)\, dF_{X,I}(x),\ t \geq 0$$

where $\overline{F}_{X,I}$ denotes the tail of the integrated tail distribution $F_{X,I}$.

**5.3.3.** Give a proof of Theorem 5.10.

**5.3.4.** Let $V$ have a compound geometric distribution with characteristics $(p, F_U)$ where $0 < p < 1$ and $F_U$ is supported on $(0, \infty)$. Show that the distribution function of $V$ is given by

$$F_V(x) = (1-p) \sum_{k=0}^{\infty} p^k F_U^{*(k)}(x),\ x \geq 0.$$

**5.3.5. a)** Let $V$ be as in Exercise 5.3.4. Show that $F_V$ satisfies the defective renewal equation

$$F_V(t) = (1-p) + (pF_U * F_V)(t),\ t \geq 0.$$

b) In fact, $F_V$ given in Exercise 5.3.4 is the unique bounded solution supported on $(0, \infty)$ to the defective renewal equation in (a) above.

**Exercises on Section 5.4**

**5.4.1.** This exercise concerns the convolution equation (3.44) when $F$ is not a probability distribution. Let $\alpha > 0$ be a constant, and $G(\cdot)$ be the distribution function of a probability measure. Consider the convolution equation

$$\zeta(x) = \int_{\mathbb{R}} \zeta(x-y)\ \alpha\ dG(y),\ x \in \mathbb{R}.$$

(i) If $0 < \alpha < 1$ show that $\zeta(\cdot) \equiv 0$ is the only bounded solution to the convolution equation. (Hint: Iterate).

(ii) Let $0 < \alpha < 1$. Let $dG(y) = I_{(-\infty,0)}(y)\ \beta\ e^{\beta y}\ dy$; so $G$ is supported on $(-\infty, 0)$. Show that we can have solution of the form $\zeta(x) = e^{\lambda x},\ x \in \mathbb{R}$.

(iii) Let $\alpha > 1$. Let $dG(y) = I_{(0,\infty)}(y)\ \beta\ e^{-\beta y} dy$. Show that we can have solution of the form $\zeta(x) = e^{\lambda x},\ x \in \mathbb{R}$. Thus it is possible to have unbounded solutions to the convolution equation (3.44) if $F$ is not a probability distribution.

**5.4.2.** Consider the Cramér-Lundberg model with the net profit condition (5.26); denote $\rho = \frac{c}{\lambda\mu} - 1$.

(i) Show that (5.36) is equivalent to

$$\int_0^\infty e^{\nu x} \overline{F}_X(x)\ dx = \frac{c}{\lambda};$$

hence conclude that if a $\nu > 0$ satisfying (5.36) exists then it is unique.

(ii) If (5.36) holds, show that the claim size $X$ has moment generating function in a neighbourhood of 0. So $X$ is light-tailed. (Hint: Use (4.4).)

(iii) Show that (5.36) holds $\Leftrightarrow$ $\nu$ is the Lundberg coefficient as defined in Section 5.2.

**5.4.3.** Assume that the claim size $X$ has moment generating function $\hat{m}_X(\cdot)$ in $(-h_0, h_0)$, $h_0 > 0$ and that there is $\nu \in (0, h_0)$ satisfying (5.36). Show that (5.38) holds, and in fact

$$\int_0^\infty x\ e^{\nu x}\ \overline{F}_X(x)\ dx = \frac{1}{\lambda\nu}\ [\lambda \hat{m}_X^{(1)}(\nu) - c]$$

where $\hat{m}_X^{(1)}(\cdot)$ denotes the derivative of $\hat{m}_X(\cdot)$.

**5.4.4.** Under the hypotheses of Theorem 5.14 show that the survival probability $\varphi \in \mathcal{S}$.

**5.4.5.** Consider the Cramér-Lundberg model with the net profit condition. Suppose that the claim size distribution $F \in \mathcal{R}_{-\alpha}$, $\alpha > 1$, that is, regularly varying with tail index $\alpha > 1$. Show that the ruin probability satisfies

$$\psi(t) \sim t^{-(\alpha-1)} \, L(t), \text{ as } t \to \infty,$$

where $L$ is slowly varying.

# Chapter 6

# Lundberg risk process with investment

Around the same time Lundberg formulated his risk model, French mathematician Louis Bachelier worked out a quantitative theory of Brownian motion from a study of stock price fluctuations. (It is interesting to note that Bachelier's pioneering work predated the famous work of Einstein on Brownian motion by five years!) Though risk models perturbed by Brownian motion/Levy process and diffusion approximation of risk processes have been studied since 1970's, (see Chapter 13 of [RSST]), interaction between these two important stochastic models has been sporadic till recently.

Hipp and Plum [HP1, HP2] formulated a few years back a model in which part of the surplus can be invested in a risky asset, and obtained some interesting results. Such a model requires degeneracy of the diffusion at the origin. A systematic study of the set up will be very useful. This chapter is just an attempt at that. Needless to emphasize, the results are far from complete.

In this chapter we consider the Cramér-Lundberg model with the possibility of investments in a riskless bond and in a risky asset. The investment strategy may help in reducing the ruin probability.

The prerequisites needed for this chapter are substantially more than the earlier chapters. Reasonable familiarity with the Brownian motion process, stochastic calculus, stochastic differential equations, and rudiments of continuous time Markov processes will be helpful. Some of the excellent books are [O], [Du], [KS], [IW], [BW]. We shall freely use results from these sources, of course, with appropriate reference. A concise account of Browian motion and Ito integrals is given in Appendix A.4.

## 6.1 Lundberg risk process and Markov property

Recall from Chapter 5 that the Lundberg risk process is given by (5.1) where $\{N(t)\}$ is a homogeneous Poisson process with rate $\lambda > 0$. In the Cramér-Lundberg model, by Theorem 2.12 we know that the total claim amount process $\{S(t)\}$ has independent increments. Hence the Lundberg risk process $\{R(t)\}$ also has the independent increment property. That is, $R(\cdot)$ is a *Levy process.* Though the Markov property is intuitively obvious for a Levy process, we put it in a convenient form.

For $t \geq 0$ define $\mathcal{F}_t = \sigma\{R(s) : 0 \leq s \leq t\}$ = the smallest $\sigma$-algebra making the family $\{R(s) : 0 \leq s \leq t\}$ measurable; here $R(t) = x + ct - S(t) = x + ct - \sum_{i=1}^{N(t)} X_i$, $t \geq 0$ as given by (5.1). $\{\mathcal{F}_t : t \geq 0\}$ is the *filtration* generated by $R(\cdot)$.

If $\mathcal{A}, \mathcal{C}$ are sub $\sigma$-algebras in $(\Omega, \mathcal{F}, P)$ such that $P(A \cap C) = P(A) \cdot P(C)$ for any $A \in \mathcal{A}, C \in \mathcal{C}$ we say that the *sub $\sigma$-algebras $\mathcal{A}$ and $\mathcal{C}$ are independent.* In addition, if $\mathcal{C} = \sigma(Y)$ where $Y$ is a random variable, we say that the *sub $\sigma$-algebra $\mathcal{A}$ and the random variable $Y$ are independent.* So the independent increment property of $R(\cdot)$ can also be stated as: for any $0 \leq s \leq t$, the sub $\sigma$-algebra $\mathcal{F}_s$ and the random variable $R(t) - R(s)$ are independent.

**Lemma 6.1** *Let $\mathcal{A}, \mathcal{B}, \mathcal{C}$ be sub $\sigma$-algebras of $\mathcal{F}$ such that $\mathcal{B} \subseteq \mathcal{A}$, and $\mathcal{A}$ and $\mathcal{C}$ are independent. Let $\mathcal{B} \vee \mathcal{C} = \sigma\{\mathcal{B} \cup \mathcal{C}\}$ be the smallest $\sigma$-algebra containing $\mathcal{B} \cup \mathcal{C}$. If $Y$ is a $(\mathcal{B} \vee \mathcal{C})$-measurable integrable random variable then $E(Y \mid \mathcal{A}) = E(Y \mid \mathcal{B})$ a.s.*

**Proof:** Let

$$\mathcal{J} = \left\{ \bigcup_{i=1}^{n} (B_i \cap C_i) : B_i \in \mathcal{B},\ C_i \in \mathcal{C},\ 1 \leq i \leq n, n = 1, 2, \ldots \right\}.$$

It is easily seen that $\mathcal{J}$ is an algebra, and hence $\mathcal{B} \vee \mathcal{C} = \sigma(\mathcal{J})$. So by standard measure theoretic arguments it is enough to consider $Y = Y_1 Y_2$ where $Y_1$ is bounded $\mathcal{B}$-measurable function and $Y_2$ is bounded $\mathcal{C}$-measurable function. As $\mathcal{A}$ and $\mathcal{C}$ are independent note that $E(Y_2 \mid \mathcal{A}) = E(Y_2)$ a.s. Note that $\mathcal{B}$ and $\mathcal{C}$ are independent and hence $E(Y_2 \mid \mathcal{B}) = E(Y_2)$ a.s. It is now easily proved that $E(Y_1 Y_2 \mid \mathcal{A}) = Y_1 E(Y_2 \mid \mathcal{A}) = Y_1 E(Y_2) = Y_1 E(Y_2 \mid \mathcal{B}) = E(Y_1 Y_2 \mid \mathcal{B})$ a.s. □

For any random variables $Y, X$ we shall denote $E(Y \mid X) \triangleq E(Y \mid \sigma(X))$.

**Proposition 6.2** *(*Markov property*): Let $t \geq 0$. For $k = 1, 2, \ldots$, $t < t_1 < t_2 < \ldots < t_k$, $B \in \mathcal{B}(\mathbb{R}^k)$ we have*

$$\begin{aligned} & P((R(t_1), \ldots, R(t_k)) \in B \mid \mathcal{F}_t) \\ = \; & P((R(t_1), \ldots, R(t_k)) \in B \mid R(t)) \text{ a.s.} \end{aligned} \tag{6.1}$$

**Proof:** Put $\mathcal{E}^t = \sigma\{R(u) : u \geq t\}$. In Lemma 6.1 take $\mathcal{A} = \mathcal{F}_t$, $\mathcal{B} = \sigma(R(t))$, $\mathcal{C} = \{R(u) - R(t) : u \geq t\}$. Note that $\mathcal{E}^t = \mathcal{B} \vee \mathcal{C}$, and that $\mathcal{A}$ and $\mathcal{C}$ are independent. □

Let $k \geq 1$, $0 \leq s_1 < s_2 < \ldots < s_k$, $B \in \mathcal{B}(\mathbb{R}^k)$; take $t_i = t + s_i$, $1 \leq i \leq k$ in (6.1). As the r.h.s. of (6.1) is $\sigma(R(t))$-measurable, there is a measurable function $g$ on $\mathbb{R}$ such that

$$P((R(t + s_1), \ldots, R(t + s_k)) \in B \mid R(t)) = g(R(t)).$$

It is clear that

$$\begin{aligned} g(y) &= P((R(t + s_1), \ldots, R(t + s_k)) \in B \mid R(t) = y) \\ &= P((R(s_1), \ldots, R(s_k)) \in B \mid R(0) = y) \\ &= E[I_B(R(s_1), \ldots, R(s_k)) \mid R(0) = y] \\ &=: E_y[I_B(R(s_1), \ldots, R(s_k))] \end{aligned} \tag{6.2}$$

where we have used time homogeneity of the Levy process $R(\cdot)$.

Using (6.1), (6.2) and standard arguments one can prove

**Proposition 6.3** *(*Markov property*): Let $t \geq 0$. For $k \geq 1$, $0 \leq s_1 < s_2 < \ldots < s_k$ and any bounded measurable function $f$ on $\mathbb{R}^k$*

$$\begin{aligned} & E[f(R(t + s_1), \ldots, R(t + s_k)) \mid \mathcal{F}_t] \\ = \; & E[f(R(t + s_1), \ldots, R(t + s_k)) \mid R(t)] \\ = \; & E_y[f(R(s_1), \ldots, R(s_k))]\big|_{y=R(t)}, \end{aligned} \tag{6.3}$$

*where r.h.s. of (6.3) stands for $h(R(t))$ with $h$ denoting the measurable function*

$$\begin{aligned} h(y) &= E_y[f(R(s_1), \ldots, R(s_k))] \\ &:= E[f(R(s_1), \ldots, R(s_k)) \mid R(0) = y]. \end{aligned} \tag{6.4}$$

**Note:** The right hand side of (6.3) is often written as $E_{R(t)}[f(R(s_1), \ldots, R(s_k))]$. This may be confusing to a beginner. A

better way of writing is to use the "inhomogeneous" notation. With $s_i$, $f$ as above define

$$\begin{aligned} h(t,y) &= E_{t,y}[f(R(t+s_1),\ldots,R(t+s_k))] \\ &= E[f(R(t+s_1),\ldots,R(t+s_k)) \mid R(t)=y]. \end{aligned} \tag{6.5}$$

In this case, by time homogeneity of $R(\cdot)$ it follows that $h(t,y) = h(0,y)$ for all $t \geq 0$. Thus there is no inconsistency. □

Using the inhomogeneous notation, define the *transition probability function* by

$$\begin{aligned} P(s,x;t,A) &= P(R(t) \in A \mid R(s)=x) \\ &=: P_{s,x}(R(t) \in A) = E_{s,x}[I_A(R(t))] \end{aligned} \tag{6.6}$$

for $0 \leq s \leq t$, $x \in \mathbb{R}$, $A \in \mathcal{B}(\mathbb{R})$. Note the connection with (6.2), (6.3), (6.5). As the notation indicates $E_{s,x}$ denotes expectation w.r.t. $P_{s,x}$. It is easily seen that $(s,x,t) \mapsto P(s,x;t,A)$ is a bounded measurable function for fixed $A \in \mathcal{B}(\mathbb{R})$, and $A \mapsto P(s,x;t,A)$ is a probability measure on $\mathcal{B}(\mathbb{R})$ for fixed $s,x,t$.

**Proposition 6.4** *The transition probability function satisfies the* Chapman-Kolmogorov equation, *viz. for* $s \leq t \leq r$, $x \in \mathbb{R}$, $A \in \mathcal{B}(\mathbb{R})$

$$P(s,x;r,A) = \int_{\mathbb{R}} P(t,y;r,A)\ P(s,x;t,dy). \tag{6.7}$$

**Proof:**

$$\begin{aligned} P(s,x;r,A) &= E_{s,x}[I_A(R(r))] \\ &= E_{s,x}[E_{s,x}[I_A(R(r)) \mid \sigma(R(t))]] \\ &= E_{s,x}[E_{t,R(t)}(I_A(R(r))] \text{ (by (6.3), (6.5))} \\ &= E_{s,x}[P(t,R(t);r;A)] \text{ (by (6.6))} \\ &= \int_{\Omega} P(t,R(t,\omega);r,A)\ dP_{s,x}(\omega) \\ &= \int_{\mathbb{R}} P(t,y;r,A)\ P(s,x;t,dy) \text{ (by (6.6))}. \end{aligned}$$

□

Time homogeneity of $R(\cdot)$ implies that for $r, r' \geq 0,\ x \in \mathbb{R},\ A \in \mathcal{B}(\mathbb{R})$

$$P(r', x; r + r', A) = P(0, x; r, A). \tag{6.8}$$

Also the Chapman-Kolmogorov equation can be written as

$$P(0, x; t + s, A) = \int_{\mathbb{R}} P(0, y; t, A)\ P(0, x; s, dy) \tag{6.9}$$

for $s, t \geq 0,\ x \in \mathbb{R},\ A \in \mathcal{B}(\mathbb{R})$.

Our objective now is to identify the infinitesimal generator of the Markov process $R(\cdot)$. For this we define the following semigroup of operators. For any bounded measurable function $f$ on $\mathbb{R}$, $t \geq 0$ define

$$T_t f(x) \triangleq E_x[f(R(t))] = E[f(R(t)) \mid R(0) = x],\ x \in \mathbb{R}. \tag{6.10}$$

Note that $T_0 f \equiv f$. See Exercise 6.1.3 for elementary properties of $\{T_t : t \geq 0\}$.

For any function $f$ on $\mathbb{R}$, we shall denote by $\frac{d}{dx}f(x)$, $\frac{d^+}{dx}f(x)$ respectively the derivative, right derivative of $f$ at $x$ whenever they exist.

**Definition 6.5** We shall denote by $\mathcal{D}$ the collection of all functions $f : \mathbb{R} \to \mathbb{R}$ such that the following hold:
(i) $f$ is a bounded right continuous function with 0 as the only possible discontinuity point;
(ii) $f(y) =$ constant, $y \in (-\infty, 0)$;
(iii) the function $f'$ is well defined, bounded and right continuous with 0 as the only possible discontinuity point, where

$$f'(x) = \begin{cases} \frac{d}{dx}f(x), & x > 0 \\ \frac{d^+}{dx}f(x), & x = 0 \\ 0, & x < 0. \end{cases} \tag{6.11}$$

□

If $f \in \mathcal{D}$, then $f$ restricted to $[0, \infty)$ is an element of $C_b^1[0, \infty)$, that is, $f, f'$ restricted to $[0, \infty)$ are bounded continuous functions.

As indicated in Chapter 5, ruin probability plays an important role. However, ruin probability as a function of the initial capital, is not left continuous at 0; see (6.21), (6.22) below and (5.26), (5.30). This is the reason for accommodating certain discontinuous functions in the above definition.

**Theorem 6.6** *For any $f \in \mathcal{D}$ we have*

$$\begin{aligned}\lim_{t\downarrow 0} \frac{1}{t}\,[T_t f(x) - f(x)] &= cf'(x) + \lambda E[f(x-X) - f(x)] \\ &= cf'(x) + \lambda \int_0^\infty (f(x-z) - f(x))\,dF_X(z) \\ &\triangleq Lf(x), \quad x \in \mathbb{R}. \end{aligned} \tag{6.12}$$

*Here $F_X$ denotes the common distribution function of claim size $X_i$; and $X$ denotes a generic random variable with distribution function $F_X$.*

**Proof:** Note that the r.h.s. of (6.12) is well defined for any $f \in \mathcal{D}$. Now conditioning with respect to $N(t)$, we see that for $t \geq 0$, $x \in \mathbb{R}$

$$\begin{aligned} & T_t f(x) - f(x) = E[f(x+ct-S(t))] - f(x) \\ = & \sum_{n=0}^{\infty} E[\{f(x+ct-S(t)) - f(x)\} \mid N(t) = n] P(N(t) = n) \\ = & [f(x+ct) - f(x)] P(N(t) = 0) \\ & + E[\{f(x+ct-X_1) - f(x)\} \mid N(t) = 1] P(N(t) = 1) \\ & + \sum_{n\geq 2} E[\{f(x+ct-S(t)) - f(x)\} \mid N(t) = n] P(N(t) = n) \\ = & e^{-\lambda t}[f(x+ct) - f(x)] \\ & + \lambda t\, e^{-\lambda t} E[f(x+ct-X_1) - f(x)] + o(t) \\ = & J_1(t) + J_2(t) + o(t), \text{ say,} \end{aligned} \tag{6.13}$$

as $t \downarrow 0$. Here we have used the independence of $\{X_i\}$ and $\{N(t)\}$, and the fact that $P(N(t) \geq 2) = o(t)$ for a Poisson process.

Let $x \geq 0$. Then by mean value theorem

$$\frac{1}{t}\, J_1(t) = e^{-\lambda t}\, \frac{ct}{t}\, f'(\xi) \to cf'(x), \text{ as } t \downarrow 0 \tag{6.14}$$

where $0 \leq x < \xi < x + ct$, for $t > 0$ with $\xi$ possibly depending on $t$. If $x < 0$ then $x + ct < 0$ for all sufficiently small $t$. As $f$ is constant on $(-\infty, 0)$, we now have $J_1(t) = 0$ for all small $t$. So (6.14) again holds, with $f'(x) = 0$.

Next, by right continuity of $f$, and as $c > 0$, note that $f(x + ct - z) \to f(x - z)$ as $t \downarrow 0$, for all $z$, for any $x \in \mathbb{R}$. So by

the dominated convergence theorem we get

$$\begin{aligned} \frac{1}{t}\, J_2(t) &= \lambda\, e^{-\lambda t} \int [f(x+ct-z) - f(x)]\, dF_X(z) \\ &\to \lambda \int [f(x-z) - f(x)]\, dF_X(z) \end{aligned} \tag{6.15}$$

as $t \downarrow 0$. The required conclusion now follows from (6.13) - (6.15).

□

The integro-differential operator $L$ defined by (6.12) is called the *infinitesimal generator* of the Markov process $R(\cdot)$. The preceding result says that $Lf(x) = \frac{\partial^+}{\partial t}\, T_t f(x)\,|_{t=0}$ for $f \in \mathcal{D}$, $x \in \mathbb{R}$. The next result is an extension of this idea.

**Theorem 6.7** *For any $f \in \mathcal{D}$, $x \in \mathbb{R}$ the function $t \mapsto T_t f(x)$ has right derivative at any $t \geq 0$, and $\frac{\partial^+}{\partial t}\, T_t f(x) = T_t\, Lf(x)$. Moreover for any $0 \leq s \leq t$,*

$$T_t f(x) - T_s f(x) = \int_s^t T_r\, Lf(x)\, dr. \tag{6.16}$$

**Proof:** By Exercise 6.1.8 note that $Lf$ is a bounded right continuous function on $\mathbb{R}$, and $r \mapsto T_r\, Lf(x)$ is a bounded continuous function for $f \in \mathcal{D}$, $x \in \mathbb{R}$. So r.h.s. of (6.16) is differentiable in $t$. Also by Exercise 6.1.7, l.h.s. of (6.16) is continuous in $t$.

Let $f \in \mathcal{D}$, $x \in \mathbb{R}$. For any $t \geq 0$ note that

$$\frac{T_{t+h} f(x) - T_f f(x)}{h} = T_t \left( \frac{T_h f(x) - f(x)}{h} \right) \to T_t\, Lf(x),$$

as $h \downarrow 0$, by the semigroup property and Theorem 6.6. This proves the first assertion of the theorem.

Next let $s \geq 0$ be fixed. For $t \geq s$ put

$$H(t) = T_t f(x) - T_s f(x) - \int_s^t T_r f(x)\, dr.$$

By the preceding part of the proof, $t \mapsto H(t)$ is continuous, has right derivative, and $\frac{d^+ H(t)}{dt} = 0$, $t \geq s$. Also note that $H(s) = 0$. We now claim that $H(t) = 0$, $t \geq s$. To prove the claim we proceed as

in the proof of Lemma 0.8, pp.235-236 of [Dy, Vol.2]. If possible let $H(t_1) \neq 0$ for some $t_1 > 0$; without loss of generality we may assume $H(t_1) > 0$. So $H(t_1) - H(s) > 0$; hence there is $\epsilon > 0$ such that $H(t_1) - H(s) > \epsilon(t_1 - s)$. Set $\xi(r) = H(r) - H(s) - \epsilon(r - s)$, $r \geq s$, $\Gamma = \{r \geq s : \xi(r) \leq 0\}$, and $r_0 = \sup \Gamma$. Clearly $s \in \Gamma$, $t_1 \notin \Gamma$. So continuity of $H$ implies that $r_0 \neq t_1$. Hence $r_0 < t_1$. For $r \in (r_0, t_1]$ note that $0 < \xi(r) \leq \xi(r) - \xi(r_0) = H(r) - H(r_0) - \epsilon(r - r_0)$. Consequently $\frac{H(r)-H(r_0)}{r-r_0} > \epsilon$ for all $r \in (r_0, t_1]$. This contradicts $\frac{d^+H}{dt}(t) = 0$, $t \geq s$. Therefore the claim and hence (6.16) is proved.

□

**Theorem 6.8** *For $f \in \mathcal{D}$ we have*

$$M_f(t) \triangleq f(R(t)) - f(R(0)) - \int_0^t Lf(R(r))\, dr, \quad t \geq 0 \tag{6.17}$$

*is a martingale w.r.t. $\{\mathcal{F}_t\}$.*

**Proof:** It is easily verified that $M_f(t)$ is $\mathcal{F}_t$-measurable and integrable for any $t \geq 0$. It remains to prove

$$E[M_f(t) - M_f(s) \mid \mathcal{F}_s] = 0 \ a.s. \tag{6.18}$$

By Markov property (Proposition 6.3) note that

$$\begin{aligned} E(f(R(t)) \mid \mathcal{F}_s) &= E_y(f(R(t-s)))\,|_{y=R(s)} \\ &= T_{t-s}f(y)\,|_{y=R(s)} \ \ a.s. \end{aligned} \tag{6.19}$$

Again by the Markov property we have

$$\begin{aligned} E\left[\int_s^t Lf(R(u))du \mid \mathcal{F}_s\right] &= \int_s^t E(Lf(R(u)) \mid \mathcal{F}_s)du \\ = \left(\int_s^t E_y(Lf(R(u-s)))\, du\right)\Big|_{y=R(s)} \\ = \left(\int_s^t T_{u-s}Lf(y)\, du\right)\Big|_{y=R(s)} \quad a.s. \end{aligned} \tag{6.20}$$

By (6.19), (6.20) we now obtain

$$\begin{aligned} & E[M_f(t) - M_f(s) \mid \mathcal{F}_s] \\ = & \left( T_{t-s}f(y) - f(y) - \int_s^t T_{u-s}Lf(y)\, du \right)\Big|_{y=R(s)} \\ = & \left( T_{t-s}f(y) - f(y) - \int_0^{t-s} T_rLf(y)\, dr \right)\Big|_{y=R(s)} \\ = & \ 0 \ \ a.s. \end{aligned}$$

by Theorem 6.7, proving (6.18). □

Our next objective is to connect the ruin/survival probability with the infinitesimal generator. Building on the definitions (5.3), (5.25) given in Chapter 5, set

$$\psi(x) = \begin{cases} P(\tau_0 < \infty \mid R(0) = x) = P_x(\tau_0 < \infty), & \text{if } x \geq 0 \\ 1, \ \text{if } x < 0 \end{cases} \tag{6.21}$$

$$\varphi(x) = 1 - \psi(x), \ x \in \mathbb{R}. \tag{6.22}$$

**Theorem 6.9** *Assume that the claim size distribution $F_X$ has a probability density function. Assume also the net profit condition (NPC) (5.26). Then $\psi, \varphi \in \mathcal{D}$. Moreover*
*(i) $\psi$ solves the problem:*

$$\left. \begin{array}{c} L\psi(x) = 0, \ x \geq 0, \ \psi(x) = 1, \ x < 0, \\ \lim\limits_{x\to\infty} \psi(x) = 0 \end{array} \right\}; \tag{6.23}$$

*(ii) $\varphi$ solves the problem*

$$\left. \begin{array}{c} L\varphi(x) = 0, \ x \geq 0, \ \varphi(x) = 0, \ x < 0 \\ \lim\limits_{x\to\infty} \varphi(x) = 1 \end{array} \right\}. \tag{6.24}$$

**Proof:** It is enough to prove the assertions concerning $\varphi$. As $F_X$ has a probability density function, by Exercise 5.3.1, (5.27), (5.28) it follows that $\varphi$ is $C_b^1$ on $(0, \infty)$. The same arguments also imply that $\varphi$ has right derivative at $x = 0$, and that $\frac{d^+\varphi}{dx}$ is right continuous at $x = 0$. It is now clear that $\varphi \in \mathcal{D}$.

By continuity of $\varphi'$ on $[0,\infty)$, the expression (5.28) holds for all $x \geq 0$. As $\varphi(\cdot) = 0$ on $(-\infty, 0)$, and $F_X$ is supported on $(0,\infty)$ it is now easily seen from (5.28) that $L\varphi(x) = 0$, $x \geq 0$. Finally, as the (NPC) holds, by Exercise 5.1.7 (b) it follows that $\lim_{x\to\infty} \varphi(x) = 1$. □

**Theorem 6.10** *Assume the net profit condition in the Cramér-Lundberg model. Let $g \in \mathcal{D}$ be such that*

$$Lg(x) = 0 \text{ for } x \geq 0, \; g(x) = 1 \text{ for } x < 0, \; \lim_{x\to\infty} g(x) = 0. \tag{6.25}$$

*Then $g(x) = \psi(x) = P_x(\tau_0 < \infty)$.*

**Proof:** As $g$ is constant on $(-\infty, 0)$ and $F_X$ is supported on $(0,\infty)$, note that $Lg(y) = 0$, $y < 0$. So $Lg(\cdot) = 0$ on $\mathbb{R}$. Let $x \geq 0$. If $R(0) = x$, by Theorem 6.8 it follows that $g(R(t)) - g(x)$ is a martingale. So by the optional sampling theorem $g(x) = E_x[g(R(\tau_0 \wedge t))]$, for all $t \geq 0$.

By Exercise 6.1.10, $R(\tau_0) < 0$ and hence $g(R(\tau_0)) = 1$, whenever $\tau_0 < \infty$. Therefore it follows that

$$\begin{aligned} g(x) &= \lim_{t\uparrow\infty} E_x[g(R(\tau_0 \wedge t))] \\ &= \lim_{t\uparrow\infty} \{P_x(\tau_0 < t) + E_x[I_{\{\tau_0 \geq t\}} g(R(t))]\}. \end{aligned} \tag{6.26}$$

As the net profit condition holds, by Exercise 5.1.7 (c), $R(t) \to +\infty$ a.s. as $t \to \infty$. So on the set $\{\tau_0 = \infty\}$, $R(t) \to +\infty$ a.s. Since $\lim_{y\uparrow\infty} g(y) = 0$, we now get $\lim_{t\uparrow\infty} E_x[I_{\{\tau_0 \geq t\}} g(R(t))] = 0$. Consequently from (6.26) we now get $g(x) = \lim_{t\uparrow\infty} P_x(\tau_0 \leq t) = P_x(\tau_0 < \infty) = \psi(x)$ as required. □

Thus, if $F_X$ has a density function, and if (NPC) holds, by Theorems 6.9, 6.10, it follows that the ruin probability $\psi(\cdot)$ is the unique solution in $\mathcal{D}$ to the problem (6.25). So (6.25) can be used to find the ruin probability. We illustrate this in the case of exponentially distributed claim sizes. We follow [H].

**Example 6.11** Let $F_X = \text{Exp}\,(\theta)$, $\theta > 0$; so $\mu = E(x) = \frac{1}{\theta}$. The net profit condition here means $c > \lambda\mu = \frac{\lambda}{\theta}$. Put $h(x) = E[\psi(x - X)]$, $x \in \mathbb{R}$. Since $\psi(\cdot) = 1$ on $(-\infty, 0)$ note that $h(x) = 1$, $x < 0$. For $x \geq 0$ it

is easily seen that

$$\begin{aligned} h(x) &= \int_0^\infty \psi(x-z)\ \theta\ e^{-\theta z} dz \\ &= e^{-\theta x} + \int_0^x \psi(z)\ \ \theta\ e^{-\theta(x-z)} dz. \end{aligned} \tag{6.27}$$

Clearly $h(\cdot)$ is bounded continuous on $\mathbb{R}$. As $\psi$ is continuous on $[0,\infty)$, from (6.27) it follows that $h(\cdot)$ is differentiable on $(0,\infty)$ and that $\frac{d^+h}{dx}(\cdot)$ exists at $x=0$; moreover

$$\begin{aligned} h'(x) &= -\theta e^{-\theta x} - \theta \int_0^x \psi(z)\ \theta e^{-\theta(x-z)} dz + \theta\psi(x) \\ &= \theta[\psi(x) - h(x)],\ x \geq 0 \end{aligned} \tag{6.28}$$

and hence $h'$ is bounded continuous on $[0,\infty)$; (that is $h$ restricted to $[0,\infty)$ is a $C_b^1$-function). From $L\psi(x)=0,\ x\geq 0$ we get

$$-c\ \psi'(x) = \lambda\ h(x) - \lambda\psi(x),\ x \geq 0. \tag{6.29}$$

As the r.h.s. of (6.29) is $C^1$ on $[0,\infty)$, it now follows that $\psi$ is $C^2$ on $[0,\infty)$. So differentiating (6.29), and using (6.28), (6.29) we obtain $\psi''(x) = \frac{\lambda - c\theta}{c}\ \psi'(x),\ x \geq 0$. That is $\psi(\cdot)$ satisfies the ODE:

$$\psi''(x) + \left(\frac{c\theta - \lambda}{c}\right)\psi'(x) = 0,\ x \geq 0. \tag{6.30}$$

Here note that $\psi''(0) = \left[\frac{d^+}{dx}\left(\frac{d^+}{dx}\psi(x)\right)\right]\Big|_{x=0}$. The general solution to (6.30) is of the form

$$\psi(x) = K_1 + K_2 \exp\left[-\left(\frac{c\theta-\lambda}{c}\right)x\right],\ x \geq 0$$

where $K_1,\ K_2$ are constants. Since $c\ \theta > \lambda$, and $\lim_{x\to\infty} \psi(x) = 0$ is a requirement, it follows that $K_1 = 0$. So $\psi(0) = K_2$. As $\psi$ has right derivative at 0 we get $\psi'(0) = -K_2\left(\frac{c\theta-\lambda}{c}\right)$. Plugging the above values in (6.29) with $x=0$, and as $h(0)=1$ we obtain $K_2 = \frac{\lambda}{c\theta}$. Therefore

$$\psi(x) = \frac{\lambda}{c\theta} \exp\left\{-\left(\frac{c\theta-\lambda}{c}\right)x\right\},\ x \geq 0$$

which agrees with (5.45) in Example 5.13. □

## 6.2 Risk process with investment

We continue with the Cramér-Lundberg model. Suppose the insurance company invests part of its surplus in a riskless bond, say a bank deposit, and the remaining part in a risky asset, say a stock. Let $\zeta \geq 0$ denote the force of interest (i.e. interest rate operating continuously). We assume that the price $Z(\cdot)$ of one unit of the risky asset is given by the classical Samuelson model in terms of a geometric Brownian motion, viz.

$$dZ(t) = bZ(t)\ dt + \sigma\ Z(t)\ dB(t),\ t > 0 \tag{6.31}$$

with initial value $Z(0) = z_0 > 0$, where $\sigma, b > 0$ are constants, and $\{B(t) : t \geq 0\}$ is a standard one dimensional Brownian motion independent of the total claim amount process $S(\cdot)$. Let $A(x)$ denote the amount invested in the risky asset when the surplus is $x$; so $A(R(t))$ would represent the amount invested in the risky asset at time $t$. The remaining amount $R(t) - A(R(t))$ is invested in the bond. (If we assume $A(x) \leq x$ that would be a budget constraint.) The function $A(\cdot)$ will be treated as a control; it will be assumed to be a reasonably well behaved function. We will write $R^{(A)}(\cdot)$ to denote the risk/surplus process when investment is made according to $A(\cdot)$ as above.

The two (random) "observables" of the system are the total claim amount $S(\cdot)$, and the unit price $Z(\cdot)$ of the risky asset. We *assume* that the processes $\{S(s) : s \geq 0\}$ and $\{Z(t) : t \geq 0\}$ are independent; this means that the sub $\sigma$-algebras $\sigma\{S(s) : s \geq 0\}$ and $\sigma\{Z(t) : t \geq 0\}$ are independent.

Note that (6.31) stands for the stochastic integral equation

$$Z(t) = z_0 + b \int_0^t Z(s)\ ds + \sigma \int_0^t Z(s)\ dB(s),\ t \geq 0$$

where the last term on the r.h.s. is an Ito integral. The "dispersion" or "volatility" factor $\sigma > 0$ is indicative of fluctuations in the price of the risky asset; the "drift" factor $b > 0$ indicates the rate at which the risky asset would appreciate in the absence of any volatility.

It is intuitively clear that the SDE for the current surplus $R^{(A)}(\cdot)$ corresponding to the control $A(\cdot)$ can be written as (see Exercise 6.2.2)]

$$\begin{aligned} dR^{(A)}(t) &= c\ dt - dS(t) + A(R^{(A)}(t-))\ [b\ dt + \sigma\ dB(t)] \\ &\quad + \zeta \cdot [R^{(A)}(t) - A(R^{(A)}(t-))]\ dt. \end{aligned} \tag{6.32}$$

Note that $dS(t)$ is meaningful as sample paths of $S(\cdot)$ are piecewise constant. Since $\sigma\{N(s) : 0 \leq s < t\} \neq \sigma\{N(s) : 0 \leq s \leq t\}$, observe that knowledge of $\{S(s) : 0 \leq s < t\}$ does not help predict if there would be a claim at time $t$. So $A(R^{(A)}(t-))$ indicates that investment decision at time $t$ can depend only on information available before $t$ (see [H], [HP1, HP2]). However we know that the process $N(\cdot)$, and hence the processes $S(\cdot)$ and $R^{(A)}(\cdot)$, can have only finitely many jumps in any finite time interval. So by continuity of Brownian paths and non-atomicity of $dt$ measure, it follows that the SDE (6.32) is the same (with probability one) as the SDE

$$\begin{aligned} dR^{(A)}(t) &= c\,dt - dS(t) + A(R^{(A)}(t))\,[b\,dt + \sigma\,dB(t)] \\ &\quad + \zeta \cdot [R^{(A)}(t) - A(R^{(A)}(t))]\,dt. \end{aligned} \tag{6.33}$$

In (6.33) it is apparent how "drift" and "dispersion" are defined only for $x \geq 0$, that is, when surplus is nonnegative. We need, however, to specify these coefficients for all $x$ to have a meaningful framework. A natural way of doing this in our context is given below.

Let $A : [0, \infty) \to \mathbb{R}$ be a nice function. Define

$$\sigma^{(A)}(x) = \begin{cases} \sigma A(x), & x \geq 0 \\ \sigma^{(A)}(0), & x < 0 \end{cases} \tag{6.34}$$

$$b^{(A)}(x) = \begin{cases} c + bA(x) + \zeta x - \zeta A(x), & x \geq 0 \\ b^{(A)}(0), & x < 0 \end{cases} \tag{6.35}$$

where $c,\ b,\ \sigma > 0,\ \zeta \geq 0$. The functions $\sigma^{(A)},\ b^{(A)}$ are respectively the dispersion and drift coefficients corresponding to the control $A(\cdot)$. The stochastic process $R^{(A)}(\cdot)$ is required to satisfy the SDE

$$\begin{aligned} R^{(A)}(t) &= R^{(A)}(0) + \int_0^t \sigma^{(A)}\,(R^{(A)}(s))\,dB(s) \\ &\quad + \int_0^t b^{(A)}\,(R^{(A)}(s))\,ds - S(t),\ t \geq 0. \end{aligned} \tag{6.36}$$

Let $\{B(r) : r \geq 0\}$, $\{S(r) : r \geq 0\}$, $R^{(A)}(0)$ be independent denoting respectively a standard one dimensional Brownian motion, a compound

Poisson process (total claim amount process), a real valued random variable. We assume that $\{\mathcal{G}_t : t \geq 0\}$ is a right continuous filtration such that $\sigma\{R^{(A)}(0),\ Z(s),\ S(r) : 0 \leq s,\ r \leq t\} \subseteq \mathcal{G}_t,\ t \geq 0$. We also assume that for each $t \geq 0$, $\{B(t+r)-B(t), S(t+\alpha)-S(t) : r, \alpha \geq 0\}$ is independent of $\mathcal{G}_t$. Here 'right continuous filtration' means that $\mathcal{G}_t$ is a nondecreasing family of sub $\sigma$-algebras with the property $\mathcal{G}_t = \mathcal{G}_{t+h} \triangleq \bigcap_{h>0} \mathcal{G}_{t+h}$ for each $t \geq 0$. This is an important technical assumption making arguments involving stopping times simpler; see [KS] and references therein for details and nuances. By a *strong solution to the SDE* (6.36) we mean a $\{\mathcal{G}_t\}$-adapted process $\{R^{(A)}(t) : t \geq 0\}$ with r.c.l.l. sample paths satisfying the equation (6.36) with probability one.

Note that the corresponding continuous diffusion with drift $b^{(A)}$ and dispersion $\sigma^{(A)}$ is the solution to the Brownian driven SDE

$$\begin{aligned} Y^{(A)}(t) &= Y^{(A)}(0) + \int_0^t \sigma^{(A)}(R^{(A)}(s))\, dB(s) \\ &\quad + \int_0^t b^{(A)}(R^{(A)}(s))\, ds,\ t \geq 0. \end{aligned} \tag{6.37}$$

It is well known that a unique strong solution of (6.37) exists when $\sigma^{(A)}$, $b^{(A)}$ are Lipschitz continuous. Moreover the process $Y^{(A)}(\cdot)$ is a strong Markov process with continuous sample paths. See [KS], [IW] for more details and information; see also Appendix A.4.

**Theorem 6.12** *Let $A : [0,\infty) \to \mathbb{R}$ be a Lipschitz continuous function; that is, there is a constant $K > 0$ such that $|A(x) - A(y)| \leq K|x - y|$ for all $x, y \geq 0$. Define $\sigma^{(A)}$, $b^{(A)}$ respectively by (6.34), (6.35). Let $R^{(A)}(0)$ be a random variable independent of the processes $Z(\cdot)$, $S(\cdot)$. Then there exists a unique strong solution to the SDE (6.36).*

**Proof:** We essentially follow the approach in [IW], which can be consulted for more general results. (Thanks to Mrinal Ghosh for pointing this out, replacing an earlier long winded argument.)

Clearly $\sigma^{(A)}$, $b^{(A)}$ are also Lipschitz continuous and satisfy a linear growth condition. Note that discontinuities in $R^{(A)}(\cdot)$ occur only from $S(\cdot)$. Let $\{\tau_i\}$ denote the stopping times (w.r.t. the filtration generated by $S(\cdot)$, and hence w.r.t. $\{\mathcal{G}_t\}$) at which jumps of $S(\cdot)$ occur; these are

the claim arrival times. We know that $0 < \tau_1 < \tau_2 < \ldots < \tau_n < \ldots$ and $\tau_n \uparrow \infty$ a.s.

On the random interval $[0, \tau_1)$ our (6.36) is just the Brownian driven SDE (6.37) with $Y^{(A)}(0) = R^{(A)}(0)$; so $Y^{(A)}(\cdot)$ is well defined. Now put

$$R^{(A)}(t) = \begin{cases} Y^{(A)}(t), \ 0 \le t < \tau_1 \\ Y^{(A)}(\tau_1 -) - X_1, \ \text{if } t = \tau_1 \end{cases}$$

where $X_1$ is the jump at $\tau_1$; note that $X_1$ is the claim size for the first claim. Clearly the above is the unique solution on $[0, \tau_1]$. By the strong Markov property of $B(\cdot)$, $S(\cdot)$ note that $R^{(A)}(\tau_1)$ is independent of $\{B(\tau_1 + s) - B(\tau_1), \ S(\tau_1 + r) - S(\tau_1) : s, r \ge 0\}$; see [IW], [KS] for detailed discussion on strong Markov property; a brief account is given in Appendix A.4. So, with $R^{(A)}(\tau_1)$ playing the role of $R^{(A)}(0)$, repeat the above procedure on $[\tau_1, \tau_2)$, using Brownian driven SDE on $[\tau_1, \tau_2)$ to get $R^{(A)}(t)$ for $\tau_1 \le t < \tau_2$ (and hence for $t \in [0, \tau_2)$). Then put $R^{(A)}(\tau_2) = R^{(A)}(\tau_2 -) - X_2$. This determines the solution uniquely on $[0, \tau_2]$.

Proceed inductively to get the unique strong solution on $[0, \tau_n]$ for each $n$. As $\tau_n \uparrow \infty$ we are done. □

A useful comparison result, which may be of independent interest, is given next. Note that no moment conditions on claim sizes are assumed

**Theorem 6.13** *(Comparison result): Let the notation and hypotheses be as in the preceding theorem.*
*(i) If* $R^{(A)}(0) \le Y^{(A)}(0)$ *a.s. then*

$$P(R^{(A)}(t) \le Y^{(A)}(t) \ \textit{for all } 0 \le t < \infty) = 1.$$

*(ii) Let* $\xi_1, \xi_2$ *be random variables independent of* $B(\cdot)$, $S(\cdot)$ *such that* $\xi_1 \le \xi_2$ *a.s. Let* $R^{(i)}$ *be the solution to the SDE (6.36) with initial value* $\xi_i$, $i = 1, 2$. *Then* $P(R^{(1)}(t) \le R^{(2)}(t)$ *for all* $0 \le t < \infty) = 1$.

*(iii) Let* $R^{(A)}(0) \le Y^{(A)}(0)$. *Then for any* $T > 0$ *there exist constants* $C_1, C_2 > 0$ *such that*

$$E\{[I_{[0,\infty)}(R^{(A)}(t)) \ R^{(A)}(t)]^2\} \le C_1[1 + E(Y^{(A)}(0))^2] \ e^{C_2 t}$$

*for all* $0 \le t \le T$.

**Proof:** (i) On $[0, \tau_1)$ both $R^{(A)}$, $Y^{(A)}$ are one dimensional Brownian driven SDE's. So by the comparison result on pp. 293-294 of

[KS], $R^{(A)}(t) \leq Y^{(A)}(t)$ for $0 \leq t < \tau_1$ a.s. Therefore, as $X_1 \geq 0$ and $Y^{(A)}$ is continuous, $R^{(A)}(\tau_1) = R^{(A)}(\tau_1-) - X_1 \leq Y^{(A)}(\tau_1)$. Next, on $(\tau_1, \tau_2)$ again $R^{(A)}(\cdot)$, $Y^{(A)}(\cdot)$ are one dimensional Brownian driven SDE's. So invoking the comparison result from [KS] as before we get $R^{(A)}(\cdot) \leq Y^{(A)}(\cdot)$ on $[0, \tau_2)$. Proceed thus, and as $\tau_n \uparrow \infty$ a.s., the required conclusion follows.
(ii) Similar to (i).
(iii) By part (i), note that $I_{[0,\infty)}(R^{(A)}(t)) \cdot R^{(A)}(t) \leq I_{[0,\infty)}(Y^{(A)}(t)) \cdot Y^{(A)}(t) \leq |Y^{(A)}(t)|$. The assertion now follows by standard results on Brownian driven SDE's; see p.289 of [KS]. □

One of our objectives is to establish the Markov property and the strong Markov property of the stochastic process $R^{(A)}(\cdot)$, as well as to identify the associated infinitesimal generator.

We shall write $E_x[f(R^{(A)}(t))] \triangleq E[f(R^{(A)}(t))|R^{(A)}(0) = x]$ to denote expectation of $f(R^{(A)}(t))$ when the initial value $R^{(A)}(0) = x$.

The following result giving the Feller property of the process $R^{(A)}(\cdot)$ is an important continuity property.

**Theorem 6.14** *(Feller property): Let $A$, $\sigma^{(A)}$, $b^{(A)}$ be as in Theorem 6.12.*
*(i) Let $x_n \to x_0$. For $n = 0, 1, 2, \ldots$ let $R^{(n)}(\cdot)$ denote the solution to the SDE (6.36) with initial value $x_n$. Then for any $T > 0$*

$$\lim_{n\to\infty} E\left[\sup_{0\leq t\leq T} |R^{(n)}(t) - R^{(0)}(t)|^2\right] = 0. \tag{6.38}$$

*(ii) Let $R^{(A)}(\cdot)$ be as in Theorem 6.12. For any bounded continuous function $f$ on $\mathbb{R}$, $t > 0$ the function $x \mapsto E_x[f(R^{(A)}(t))]$ is bounded continuous.*

**Proof:** (i) We first claim that for $T > 0$

$$E\int_0^T |\sigma^{(A)}(R^{(n)}(s))|^2\, ds < \infty,\ n = 0, 1, 2, \ldots$$

Indeed, denoting by $Y^{(n)}$ the solution to the Brownian driven SDE (6.37) with initial value $x_n$, and using Theorem 6.13 (iii), we have

$$
\begin{aligned}
& E \int_0^T |\sigma^{(A)}(R^{(n)}(s))|^2 ds \\
= \; & |\sigma^{(A)}(0)|^2 \; E \int_0^T I_{(-\infty,0)}(R^{(n)}(s)) \; ds \\
& + E \int_0^T I_{[0,\infty)}(R^{(n)}(s)) \cdot |\sigma^{(A)}(R^{(n)}(s))|^2 ds \\
\leq \; & |\sigma^{(A)}(0)|^2 T + 2K_1^2 E \int_0^T I_{[0,\infty)}(R^{(n)}(s)) \; [1 + |R^{(n)}(s)|^2] \; ds \\
\leq \; & (|\sigma^{(A)}(0)|^2 + 2K_1^2) T + 2K_1^2 C_1 (1 + E|Y^{(n)}(0)|^2) \int_0^T e^{C_2 t} \; dt \\
< \; & \infty
\end{aligned}
$$

where $K_1 > 0$ is such that $|\sigma^{(A)}(y)| \leq K_1[1 + |y|]$, $y \geq 0$.

By the claim $\int_0^t \sigma^{(A)}(R^{(n)}(s)) \; dB(s)$, $t \geq 0$ is a continuous square integrable martingale for $n = 0, 1, 2, \ldots$. It also follows that $R^{(n)}(t) - R^{(0)}(t)$, $t \geq 0$ is a continuous square integrable process, for $n = 1, 2, \ldots$. (Note that the above hold, though the claim sizes $X_i$ may not even be integrable.)

Now fix $T > 0$, and write

$$
\varphi_n(t) = E \left[ \sup_{0 \leq s \leq t} |R^{(n)}(s) - R^{(0)}(s)|^2 \right], \; t \in [0, T], \; n = 1, 2, \ldots
$$

By Doob's maximal inequality, Schwartz inequality and the Lipschitz

condition we get

$$\begin{aligned}\varphi_n(t) &\leq 3|x_n - x_0|^2 + 12E \int_0^t |\sigma^{(A)}(R^{(n)}(r)) - \sigma^{(A)}(R^{(0)}(r))|^2 dr \\ &\quad +3t\ E \int_0^t |b^{(A)}(R^{(n)}(r)) - b^{(A)}(R^{(0)}(r))|^2 dr \\ &\leq 3|x_n - x_0|^2 + (12 + 3T)K^2 \int_0^t \varphi_n(s)\ ds.\end{aligned}$$

So by Gronwall's inequality (see [O], [KS]) we get

$$\varphi_n(t) \leq 3|x_n - x_0|^2\ \exp[(12 + 3T)K^2 t],\ 0 \leq t \leq T$$

whence (6.38) follows.

(ii) Let $x_n \to x_0$. We need to prove that

$$E[f(R^{(n)}(t))] \to E[f(R^{(0)}(t))]$$

where $t > 0$, $f$ a bounded continuous function. By (i), $R^{(n)}(t) \to R^{(0)}(t)$ in probability. So there is a subsequence $\{n_k\}$ such that $R^{(n_k)}(t) \to R^{(0)}(t)$ a.s. and hence $f(R^{(n_k)}(t)) \to f(R^{(0)}(t))$ a.s. Apply the dominated convergence theorem to conclude that $E[f(R^{(n_k)}(t))] \to E[f(R^{(0)}(t))]$. This argument, in fact, shows that for any subsequence $\{x_{n_k}\}$ of $\{x_n\}$ there is a further sequence $\{y_p\}$ such that $E[f(R^{(p)}(t))] \to E[f(R^{(0)}(t))]$. So the required conclusion follows. □

Next we look at the Markov property of the process $R^{(A)}(\cdot)$.

Fix $s \geq 0$; let $\xi$ be a random variable independent of $\{B(t) - B(s) : t \geq s\}$ and $\{S(t) - S(s) : t \geq s\}$. Denote $\mathcal{G}_t^s = \sigma\{\xi, B(r) - B(s),\ S(r) - S(s) : s \leq r \leq t\}$, $t \geq s$. Consider the SDE

$$d\bar{R}(t) = \sigma^{(A)}(\bar{R}(t))\ dB(t) + b^{(A)}(\bar{R}(t))\ dt - dS(t),\ t \geq s \qquad (6.39)$$

with initial value $\bar{R}(s) = \xi$. By an easy adaptation of Theorem 6.12, existence of a unique $\{\mathcal{G}_t^s\}$-adapted strong solution to (6.39) can be established; see Exercise 6.2.3. Consequently $\{R^{(A)}(t) : t \geq s\}$ is the unique solution to (6.39) with $\bar{R}(s) = R^{(A)}(s)$. Hence for $t \geq s$, $R^{(A)}(t)$ is measurable w.r.t. $\sigma\{R^{(A)}(s),\ B(r) - B(s),\ S(r) - S(s) : s \leq r \leq t\}$.

Let $t \geq 0$ be fixed. Take $\mathcal{A} = \mathcal{G}_t$, $\mathcal{B} = \sigma\{R^{(A)}(t)\}$, $\mathcal{C} = \sigma\{B(r) - B(t),\ S(r) - S(t) : r \geq t\}$. Then $\mathcal{B} \subseteq \mathcal{A}$, and $\mathcal{A}$ and $\mathcal{C}$ are independent. Put $\mathcal{E}^t = \sigma\{R^{(A)}(r) : r \geq t\}$. By the preceding paragraph $\mathcal{E}^t \subseteq (\mathcal{B} \vee \mathcal{C})$. Using Lemma 6.1, and proceeding as in Propositions 6.2, 6.3 the following result can be proved.

**Theorem 6.15** *(Markov property): Let $A$, $\{R^{(A)}(t)\}$, $\{\mathcal{G}_t\}$ be as in Theorem 6.12. Fix $t \geq 0$ and let $\mathcal{A}$, $\mathcal{B}$, $\mathcal{C}$ be as above. Then for any $(\mathcal{B} \vee \mathcal{C})$-measurable, integrable random variable $V$*

$$E(V \mid \mathcal{G}_t) = E(V \mid R^{(A)}(t)) \ a.s. \tag{6.40}$$

*In particular, for $k \geq 1$, $0 \leq s_1 < s_2 < \ldots < s_k$ and any bounded measurable function $f$ on $\mathbb{R}^k$*

$$\begin{aligned} & E[f(R^{(A)}(t+s_1), \ldots, R^{(A)}(t+s_k)) \mid \mathcal{G}_t] \\ = \ & E[f(R^{(A)}(t+s_1), \ldots, R^{(A)}(t+s_k)) \mid R^{(A)}(t)] \\ = \ & h(R^{(A)}(t)) \ a.s. \end{aligned} \tag{6.41}$$

*where $h$ is the measurable function on $\mathbb{R}$ defined by*

$$\begin{aligned} h(y) \ &\triangleq \ E_y[f(R^{(A)}(s_1), \ldots, R^{(A)}(s_k))] \\ &:= \ E[f(R^{(A)}(s_1), \ldots, R^{(A)}(s_k)) \mid R^{(A)}(0) = y]. \end{aligned} \tag{6.42}$$

□

In fact, a stronger property holds, which is also very useful. Recall that a random variable $\tau$ taking values in $[0, \infty] \triangleq [0, \infty) \cup \{+\infty\}$ is called a stopping time w.r.t. $\{\mathcal{G}_t\}$ if $\{\tau \leq t\} \in \mathcal{G}_t$ for all $t \geq 0$. Define the pre-$\tau$ $\sigma$-algebra

$$\mathcal{G}_\tau = \{E \in \mathcal{F} : E \cap \{\tau \leq t\} \in \mathcal{G}_t \ \text{ for all } t \geq 0\}.$$

Then $\mathcal{G}_\tau$ is a $\sigma$-algebra that can be considered as the repository of all information about the process $R^{(A)}$ upto time $\tau$. (See [KS], [SV2].) It can be proved that $\tau$ and $R^{(A)}(\tau)$ are $\mathcal{G}_\tau$-measurable; here the random variable $R^{(A)}(\tau)$ is defined only on $\{\tau < \infty\}$ by $R^{(A)}(\tau)(\omega) = R^{(A)}(\tau(\omega), \omega)$. (See Exercise 6.2.5.)

Now let $\tau$ be a finite stopping time w.r.t. $\{\mathcal{G}_t\}$. Using conditional characteristic functions, it can be shown that $\{\tilde{B}(s) \triangleq B(\tau + s) - B(\tau) : s \geq 0\}$ is a standard Brownian motion independent of $\mathcal{G}_\tau$. Similarly it

can be shown that $\{\tilde{S}(s) \triangleq S(\tau+s) - S(\tau) : s \geq 0\}$ is a compound Poisson process independent of $\mathcal{G}_\tau$, and that the processes $\tilde{S}(\cdot)$ and $S(\cdot)$ are identical in law. See [KS]. Of course $\tilde{B}(\cdot)$, $\tilde{S}(\cdot)$ are independent processes. We have already used these facts in the proof of Theorem 6.12. Observe that

$$\begin{aligned} R^{(A)}(\tau+t) &= R^{(A)}(\tau) + \int_\tau^{\tau+t} \sigma^{(A)}(R^{(A)}(r))\, dB(r) \\ &\quad + \int_\tau^{\tau+t} b^{(A)}(R^{(A)}(r))\, dr - (S(\tau+t) - S(\tau)) \\ &= R^{(A)}(\tau) + \int_0^t \sigma^{(A)}(R^{(A)}(\tau+r))\, d\tilde{B}(r) \\ &\quad + \int_0^t b^{(A)}(R^{(A)}(\tau+r))\, dr - \tilde{S}(t). \end{aligned}$$

So by uniqueness of solutions to SDE's of the form (6.36) w.r.t. $\tilde{B}(\cdot)$, $\tilde{S}(\cdot)$, it follows that for any $k \geq 1$, $0 \leq s_1 < s_2 < \ldots < s_k < \infty$ the $k$-dimensional random variable $(R^{(A)}(\tau+s_1), \ldots, R^{(A)}(\tau+s_k))$ is measurable w.r.t. $\sigma\{\tau, R^{(A)}(\tau), \tilde{B}(r), \tilde{S}(r) : r \geq 0\}$.

**Theorem 6.16** *(Strong Markov property): Let $A, R^{(A)}, \{\mathcal{G}_t\}$ be as in Theorem 6.12. Let $\tau$ be a finite stopping time w.r.t. $\{\mathcal{G}_t\}$. Let $\tilde{\mathcal{B}} = \sigma\{\tau, R^{(A)}(\tau)\}$, $\tilde{\mathcal{C}} = \sigma\{B(\tau+r) - B(\tau), S(\tau+r) - S(\tau) : r \geq 0\}$. Then for any $(\tilde{\mathcal{B}} \vee \tilde{\mathcal{C}})$-measurable integrable random variable $V$*

$$E[V \mid \mathcal{G}_\tau] = E[V \mid \sigma\{\tau, R^{(A)}(\tau)\}] \ a.s. \tag{6.43}$$

*In particular, for any $k \geq 1$, $0 \leq s_1 < \ldots < s_k < \infty$ and any bounded measurable function $f$ on $\mathbb{R}^k$*

$$\begin{aligned} & E[f(R^{(A)}(\tau+s_1), \ldots, R^{(A)}(\tau+s_k)) \mid \mathcal{G}_\tau] \\ = \ & h(R^{(A)}(\tau)) \ \ a.s. \end{aligned} \tag{6.44}$$

*where $h$ is given by (6.42).*

**Proof:** Clearly $\tilde{\mathcal{B}} \subseteq \mathcal{G}_\tau$, and $\mathcal{G}_\tau$ and $\tilde{\mathcal{C}}$ are independent. So, in view of the discussion preceding the theorem, (6.43) follows by Lemma

6.1. As $\sigma\{R^{(A)}(\tau+r) : r \geq 0\} \subseteq (\tilde{\mathcal{B}} \vee \tilde{\mathcal{C}})$, from (6.43) it is clear that l.h.s. of (6.44) is $\sigma\{\tau, R^{(A)}(\tau)\}$-measurable. Hence there is a measurable function $g : [0,\infty) \times \mathbb{R} \to \mathbb{R}$ such that l.h.s. of (6.44) $= g(\tau, R^{(A)}(\tau))$ a.s. Since $\sigma^{(A)}$, $b^{(A)}$ do not depend on the time variable, and since $B(\cdot)$, $S(\cdot)$ have stationary increments, time homogeneity now implies

$$\begin{aligned} & g(t,y) = E[f(R^{(A)}(t+s_1),\ldots,R^{(A)}(t+s_k)) \mid R^{(A)}(t) = y] \\ = \; & E[f(R^{(A)}(s_1),\ldots,R^{(A)}(s_k)) \mid R^{(A)}(0) = y] = h(y) \end{aligned} \tag{6.45}$$

completing the proof. □

**Remark:** Actually, it can be proved that any Markov process with r.c.l.l. sample paths having the Feller property is strong Markov.

As in Section 6.1 we want to investigate the connection between the ruin probability of the Markov process $R^{(A)}$ and the generator. Fix $A$ satisfying the hypothesis of Theorem 6.12. It will transpire that we need to impose additional conditions on the way. For $t \geq 0$ and measurable function $f$ on $\mathbb{R}$ define

$$T_t^{(A)} f(x) \triangleq E_x[f(R^{(A)}(t))] := E[f(R^{(A)}(t)) \mid R^{(A)}(0) = x] \tag{6.46}$$

whenever r.h.s. makes sense. If $f$ is bounded measurable, clearly $T_t^{(A)} f$ is well defined, $T_{t+s}^{(A)} f = T_t^{(A)}(T_s^{(A)} f) = T_s^{(A)}(T_t^{(A)} f)$, and $\sup_x |T_t^{(A)} f(x)| \leq \sup_x |f(x)|$. Thus $\{T_t^{(A)} : t \geq 0\}$ is a contraction semigroup of linear operators on the space of bounded measurable functions (under the "sup" norm). A function space suitable for our purposes is given below.

**Definition 6.17** We shall denote by $\mathcal{D}_1$ the collection of all functions $f : \mathbb{R} \to \mathbb{R}$ such that the following hold:
(i) $f \in \mathcal{D}$; (that is, (i) - (iii) of Definition 6.5 hold);
(ii) the function $f''$ is well defined, bounded and right continuous, with 0 as the only possible discontinuity point, where

$$f''(x) = \begin{cases} \frac{d^2 f}{dx^2}(x), & x > 0 \\ \frac{d^+}{dx}\left(\frac{d^+}{dx} f(x)\right), & x = 0 \\ 0, & x < 0. \end{cases} \tag{6.47}$$

Note that $f''(0)$ is the right derivative of the right derivative of $f$ at $x = 0$.

□

We know that the continuous process $Y^{(A)}$ defined by (6.37) is the one dimensional diffusion with drift $b^{(A)}(\cdot)$ and dispersion $\sigma^{(A)}(\cdot)$. By Ito's formula, for $g \in C_b^2(\mathbb{R})$

$$\begin{aligned} & g(Y^{(A)}(t)) - g(Y^{(A)}(0)) - \int_0^t L_{cts}^{(A)} g(Y^{(A)}(s))\, ds \\ = \quad & \text{a continuous martingale,} \end{aligned} \tag{6.48}$$

where

$$L_{cts}^{(A)} g(x) = \frac{1}{2}(\sigma^{(A)}(x))^2 \frac{d^2 g}{dx^2}(x) + b^{(A)}(x) \frac{dg}{dx}(x), \; x \in \mathbb{R}. \tag{6.49}$$

The martingale on r.h.s. of (6.48) is an Ito integral w.r.t. Brownian motion. Consequently for $g \in C_b^2(\mathbb{R})$, $x \in \mathbb{R}$

$$\begin{aligned} & \lim_{t \downarrow 0} \frac{1}{t} E[g(Y^{(A)}(t)) - g(Y^{(A)}(0)) \mid Y^{(A)}(0) = x] \\ = \quad & L_{cts}^{(A)}\, g(x). \end{aligned} \tag{6.50}$$

In particular, the second order differential operator $L_{cts}^{(A)}$ is the infinitesimal generator of the diffusion $Y^{(A)}$. For proofs of these assertions and detailed discussions about the nuances see [KS], [BW], [Du] and references therein.

Define the *ruin time* of the controlled risk process $R^{(A)}(\cdot)$ by

$$\tau_0^{(A)} = \inf\{t > 0 : R^{(A)}(t) < 0\}. \tag{6.51}$$

**Remark 6.18** Suppose $\sigma^{(A)}(\cdot)$ is nondegenerate at 0; that is, $(\sigma^{(A)}(0))^2 > 0$. Assume that $R^{(A)}(0) = 0$, which means that the initial capital is 0. By continuity $\inf\{(\sigma^{(A)}(y))^2 : y \in K\} > 0$ for some compact neighbourhood of 0. Since the first jump time (= arrival time of first claim) $\tau_1 > 0$ a.s., the behaviour of $R^{(A)}(t)$ for small $t$ is like $Y^{(A)}(t)$, which in turn is like the Brownian motion due to nondegeneracy at 0. Then by the "smudging effect" of Brownian motion it follows that $\tau_0^{(A)} = 0$ a.s.; see Exercise 6.2.6. This is not a desirable situation as we will explain now. If $A \equiv 0$, that is, there is no investment in risky asset, our process is basically like the Lundberg risk process $R(\cdot)$. For such a process starting at 0, we know that the ruin time is strictly positive, and that the ruin probability is strictly less than one at least

when the net profit condition hold; see (5.30). See Exercise 6.2.7. So, having a nondegnerate investment plan at 0 increases the ruin probability. This is contrary to our objective, and hence nondegeneracy at 0 is not desirable.

**Lemma 6.19** *Assume* $A(x) \geq 0$ *for* $x \geq 0$, $A(0) = 0$, *and* $b \geq \zeta$. *Let* $\sigma^{(A)}(\cdot)$, $b^{(A)}(\cdot)$, $Y^{(A)}(\cdot)$ *be as in (6.37). Let* $Y^{(A)}(0) \geq 0$, *and* $\tilde{\tau}_0 = \inf\{t \geq 0 : Y^{(A)}(t) < 0\}$ = *first hitting time of* $(-\infty, 0)$ *by the diffusion* $Y^{(A)}$. *Then* $P(\tilde{\tau}_0 = \infty) = 1$ *or equivalently*

$$P(Y^{(A)}(t) \geq 0 \text{ for all } t \geq 0 \mid Y^{(A)}(0) \geq 0) = 1.$$

**Proof:** By continuity of sample paths of $Y^{(A)}$, it is enough to consider the case $Y^{(A)}(0) \equiv 0$. Let $P(Y_0^{(A)})^{-1}$ denote the distribution of the process $Y^{(A)}$ starting at 0; this is a probability measure on the function space $C_0[0, \infty)$. Denote

$$\begin{aligned} H : \; = \; & \{h : [0, \infty) \to \mathbb{R} : h \text{ solves the ODE} \\ & h(t) = 0 + \int_0^t b^{(A)}(h(s))\, ds + \int_0^t \sigma^{(A)}(h(s))\dot{\phi}(s)\, ds \\ & \text{for some } \phi \in C^\infty(\mathbb{R})\}. \end{aligned}$$

By the support theorem of Stroock and Varadhan (see [SV1]), support of $P(Y_0^{(A)})^{-1}$ = closure of $H$ in $C_0[0, \infty)$. As $\sigma^{(A)}(0) = 0$, $b^{(A)}(\cdot) \geq c > 0$ and $Y^{(A)}(0) = 0$, any $h \in H$ will satisfy $h(t) > 0$ for all sufficiently small $t > 0$. So for any $w$ in the support of $P(Y_0^{(A)})^{-1}$ we have $w(t) \geq 0$ for all sufficiently small $t \geq 0$. Hence $P(\tilde{\tau}_0 > 0) = 1$.

Next suppose $P(\tilde{\tau}_0 < \infty) > 0$. By continuity of sample paths $Y^{(A)}(\tilde{\tau}_0(w)) = 0$ for any $w$ with $\tilde{\tau}_0(w) < \infty$. Then by the strong Markov property of $Y^{(A)}$ and the preceding paragraph, it follows that $Y^{(A)}(\tilde{\tau}_0(w) + t) \geq 0$ for all sufficiently small $t \geq 0$ for a.e.$\omega$. This would contradict the definition of $\tilde{\tau}_0(w)$. Hence $P(\tilde{\tau}_0 = \infty) = 1$. □

**Proposition 6.20** *Let* $A$, $\sigma^{(A)}$, $b^{(A)}$ *satisfy the hypotheses of Theorem 6.12 and Lemma 6.19. Let the ruin time* $\tau_0^{(A)}$ *of* $R^{(A)}$ *be defined by (6.51). If* $R^{(A)}(0) \geq 0$ *then ruin can occur only at a jump time. Moreover* $R^{(A)}(\tau_0^{(A)}) < 0$ *a.s. on* $\{\tau_0^{(A)} < \infty\}$ *for any starting point.*

**Proof:** We need to consider only $R^{(A)}(0) \geq 0$. Suppose $\tau_0^{(A)}(\omega) < \infty$. The three possibilities are:

(i) $\tau_0^{(A)}(\omega)$ is a continuity point of $R^{(A)}(\cdot,\omega)$;
(ii) $\tau_0^{(A)}(\omega)$ is a point of discontinuity and $R^{(A)}(\tau_0^{(A)}(\omega)) = 0$;
(iii) $\tau_0^{(A)}(\omega)$ is a discontinuity point and $R^{(A)}(\tau_0^{(A)}(\omega)) < 0$.

In the first two cases $R^{(A)}(\tau_0^{(A)}(\omega)) = 0$. As the claims arrive according to a Poisson process, note that $\eta^{(A)} > \tau_0^{(A)}$ where $\eta^{(A)}$ denotes the next jump time after $\tau_0^{(A)}$ of $R^{(A)}$. Also, as $\tau_0^{(A)}$ is a stopping time, by the strong Markov property, on $[\tau_0^{(A)}, \eta^{(A)})$ the process $R^{(A)}$ behaves like $Y^{(A)}(\cdot)$ starting at $R^{(A)}(\tau_0^{(A)})$. Therefore by strong Markov property and Lemma 6.19,

$$\begin{aligned} & P\left(R^{(A)}(t) < 0 \text{ for some } t \in \left(\tau_0^{(A)}, \eta^{(A)}\right) \middle| R^{(A)}\left(\tau_0^{(A)}\right) = 0\right) \\ = \; & P\left(R^{(A)}(t) < 0 \text{ for some } 0 < t < \tau_1 \middle| R^{(A)}(0) = 0\right) \\ \leq \; & P\left(Y^{(A)}(t) < 0 \text{ for some } t > 0 \middle| Y^{(A)}(0) = 0\right) = 0. \end{aligned}$$

Thus the first two cases can occur only with probability zero. □

The next result identifies the infinitesimal generator of the Markov process $R^{(A)}(\cdot)$.

**Theorem 6.21** *Let $A$, $\sigma^{(A)}$, $b^{(A)}$ satisfy the hypotheses of Theorem 6.12 and Lemma 6.19. Assume that i.i.d. claim sizes $X_i$ are continuous random variables. Let the semigroup $\{T_t^{(A)}\}$ be given by (6.46). Then for $f \in \mathcal{D}_1$*

$$\begin{aligned} & \lim_{t \downarrow 0} \frac{1}{t} \left[T_t^{(A)} f(x) - f(x)\right] := L^{(A)} f(x) \\ \triangleq \; & \frac{1}{2} \left(\sigma^{(A)}(x)\right)^2 f''(x) + b^{(A)}(x) f'(x) \\ & + \lambda \int_{[0,\infty)} [f(x-z) - f(x)] \, dF_X(z), \; x \in \mathbb{R} \end{aligned} \tag{6.52}$$

*where $F_X$ is the distribution function of claim size, and $f'$, $f''$ are as in Definitions 6.5, 6.17.*

We need two lemmas.

**Lemma 6.22** *To prove Theorem 6.21 it is enough to consider $f \in \mathcal{D}_1$ such that $\exists\ K > 0$ with $f(y) = 0$ for all $y \geq K$,*

**Proof:** See Exercise 6.2.8. □

**Lemma 6.23** *Let the hypotheses of Theorem 6.21 hold. Let $f \in \mathcal{D}_1$ be as in Lemma 6.22. Let $\tau_1$ denote the arrival time of the first claim. Then $\exists$ a constant $C$ independent of $y$, $t$ such that*

$$\left| E_y \left[ \left\{ f\left(R^{(A)}(t)\right) - f(y) \right\} \cdot I_{\{\tau_1 > t\}} \right] \right| \leq Ct\ e^{-\lambda t}. \tag{6.53}$$

**Proof:** By independence of $Y^{(A)}(\cdot)$ and $N(\cdot)$ note that for any $y \in \mathbb{R}$, $t \geq 0$

$$\begin{aligned} & E_y \left[ \left\{ f\left(R^{(A)}(t)\right) - f(y) \right\} \cdot I_{\{\tau_1 > t\}} \right] \\ = \ & E_y \left[ \left\{ f\left(Y^{(A)}(t)\right) - f(y) \right\} \cdot I_{\{\tau_1 > t\}} \right] \\ = \ & e^{-\lambda t}\ E_y \left[ f\left(Y^{(A)}(t)\right) - f(y) \right]. \end{aligned}$$

So it is enough to prove that

$$\left| E_y \left[ f\left(Y^{(A)}(t)\right) - f(y) \right] \right| \leq Ct \tag{6.54}$$

where $C$ is independent of $y$, $t$.

Let $y \geq 0$. Then by Lemma 6.19, $Y^{(A)}(t) \geq 0$ for all $t$. By our assumption on $f$, note that $L^{(A)}_{cts}\ f$ is bounded. Hence, as $f$ is $C^2$ on $[0, \infty)$, by (6.48) we get (extending $f$ to a $C^2_b$-function on $\mathbb{R}$ if necessary),

$$E_y \left[ f\left(Y^{(A)}(t)\right) - f(y) \right] = E_y \left[ \int_0^t L^{(A)}_{cts} f\left(Y^{(A)}(r)\right) dr \right] \tag{6.55}$$

whence (6.54) follows for $y \geq 0$.

Next let $y < 0$, and $\eta = \inf\{s \geq 0 : Y^{(A)}(s) \geq 0\}$. Since $f$ is constant on $(-\infty, 0)$, $Y^{(A)}(\eta) = 0$, by strong Markov property of $Y^{(A)}$ we have, using (6.55)

$$\begin{aligned} & E_y \left[ f\left(Y^{(A)}(t)\right) - f(y) \right] \\ = \ & E_y \left[ I_{\{\eta \leq t\}} \cdot E_0 \left[ f\left(Y^{(A)}(t-\eta)\right) - f(0) \right] \right] \\ = \ & E_y \left[ I_{\{\eta \leq t\}}\ E_0 \left[ \int_0^{t-\eta} L^{(A)}_{cts}\ f\left(Y^{(A)}(r)\right) dr \right] \right]. \end{aligned} \tag{6.56}$$

Boundedness of $L^{(A)}_{cts}\ f$, and $t - \eta \leq t$ now imply that (6.54) holds for $y < 0$.

□

**Proof of Theorem 6.21:** We may assume that $f$ is as in Lemma 6.22; and by Exercise 6.2.9 it is enough to consider $x \geq 0$. Fix $x \geq 0$.

Note that for $t \geq 0$,

$$\begin{aligned} T^{(A)}_t\ f(x) - f(x) &= E_x\left[\left\{f\left(R^{(A)}(t)\right) - f(x)\right\} \cdot I_{\{N(t)=0\}}\right] \\ &\quad + E_x\left[\left\{f\left(R^{(A)}(t)\right) - f(x)\right\} \cdot I_{\{N(t)=1\}}\right] \\ &\quad + E_x\left[\left\{f\left(R^{(A)}(t)\right) - f(x)\right\} \cdot I_{\{N(t)\geq 2\}}\right] \\ &= J_0(t) + J_1(t) + J_2(t), \quad \text{say.} \end{aligned} \tag{6.57}$$

Clearly $J_2(t) = o(t)$, $t \downarrow 0$ as $f$ is bounded and $P(N(t) \geq 2) = o(t)$, $t \downarrow 0$. On $\{N(t) = 0\}$, we know that $R^{(A)}(s) = Y^{(A)}(s) \geq 0$, $S(s) = 0$, $0 \leq s \leq t$. So as (in the derivation of (6.55)) in the proof of Lemma 6.23, using (6.48) and independence of $Y^{(A)}$ and $N$ we get

$$\begin{aligned} &\lim_{t\downarrow 0} \frac{1}{t}\ J_0(t) = \lim_{t\downarrow 0} \frac{1}{t}\ E_x\left[f\left(Y^{(A)}(t)\right) - f(x)\right] \cdot e^{-\lambda t} \\ = &\lim_{t\downarrow 0} \frac{1}{t}\ e^{-\lambda t}\ E_x\left[\int_0^t L^{(A)}_{cts}\ f\left(Y^{(A)}(r)\right) dr\right] \\ = &\ L^{(A)}_{cts}\ f(x) = \frac{1}{2}\ \left(\sigma^{(A)}(x)\right)^2 f''(x) + b^{(A)}(x)\ f'(x). \end{aligned} \tag{6.58}$$

Here we have used the fact that $L^{(A)}_{cts}\ f$ is a bounded continuous function when restricted to $[0, \infty)$, because $f$ is as in Lemma 6.22.

Next, on the set $\{N(t) = 1\}$ clearly $S(t) = X_1$, the first claim size. Therefore

$$\begin{aligned} J_1(t) &= E_x\left[\left\{f\left(R^{(A)}(t)\right) - f(x)\right\} \cdot I_{\{\tau_1\leq t\}} \cdot I_{\{\tau_2>t\}}\right] \\ &= E_x\left[\left\{f\left(R^{(A)}(t)\right) - f\left(R^{(A)}(\tau_1)\right)\right\} \cdot I_{\{\tau_1\leq t\}} \cdot I_{\{\tau_2>t\}}\right] \\ &\quad + E_x\left[\left\{f\left(R^{(A)}(\tau_1)\right) - f(x - X_1)\right\} \cdot I_{\{\tau_1\leq t\}} \cdot I_{\{\tau_2>t\}}\right] \\ &\quad + E_x\left[\{f(x - X_1) - f(x)\} \cdot I_{\{\tau_1\leq t\}} \cdot I_{\{\tau_2>t\}}\right] \\ &= J_{11}(t) + J_{12}(t) + J_{13}(t), \quad \text{say.} \end{aligned} \tag{6.59}$$

As $X_1$ and $N(\cdot)$ are independent we now obtain

$$\begin{aligned} & \frac{1}{t}\, J_{13}(t) = \frac{1}{t}\, E_x\left[\{f(x-X_1) - f(x)\}\,|N(t)=1\right]\ \lambda t\ e^{-\lambda t} \\ = \ & \lambda e^{-\lambda t}\, E_x[f(x-X_1) - f(x)] \\ \to \ & \lambda \int\limits_{[0,\infty)} [f(x-z) - f(x)]\, dF_X(z), \quad \text{as } t \downarrow 0. \end{aligned} \tag{6.60}$$

In view of (6.57) - (6.60), to complete the proof of the theorem we just need to establish

$$\lim_{t\downarrow 0}\ \frac{1}{t}\ J_{11}(t) = 0, \quad \text{and} \quad \lim_{t\downarrow 0}\ \frac{1}{t}\ J_{12}(t) = 0. \tag{6.61}$$

With $t \geq 0$ fixed, for $0 \leq s \leq t,\ y \in \mathbb{R}$ denote

$$\varphi(s,y) = E_y\left[\left\{f\left(R^{(A)}(t-s)\right) - f(y)\right\} \cdot I_{\{\tau_1 > t-s\}}\right].$$

Then by Lemma 6.23 there is a constant $C$ independent of $y,\ s,\ t$ such that $|\varphi(s,y)| \leq C(t-s)$. Therefore using the strong Markov property of $R^{(A)}$ we get

$$\begin{aligned} & |J_{11}(t)| \\ = \ & \left|E_x\left[E\left[\left\{f\left(R^{(A)}(t)\right) - f\left(R^{(A)}(\tau_1)\right)\right\} I_{\{\tau_1 \leq t\}} I_{\{\tau_2 > t\}} \middle| \mathcal{G}_{\tau_1}\right]\right]\right| \\ = \ & \left|E_x\left[I_{\{\tau_1 \leq t\}}\ \varphi\left(\tau_1, R^{(A)}(\tau_1)\right)\right]\right| \\ \leq \ & C\ E_x\left[(t-\tau_1)\ I_{\{\tau_1 \leq t\}}\right] \leq Ct\left(1 - e^{-\lambda t}\right) \end{aligned}$$

and hence $\frac{1}{t}\ J_{11}(t) \to 0$ as $t \downarrow 0$.

We know, by Proposition 2.7, that the conditional distribution of $\tau_1$ given $N(t) = 1$ is the uniform distribution over $(0,t)$. Consequently

$$\begin{aligned} & \frac{1}{t}\ J_{12}(t) \\ = \ & \frac{1}{t}\ E_x\left[\left\{f\left(R^{(A)}(\tau_1)\right) - f(x-X_1)\right\}\middle| N(t)=1\right] \lambda t\ e^{-\lambda t} \\ = \ & \lambda e^{-\lambda t} E_x\left[\left\{f\left(Y^{(A)}(\tau_1) - X_1\right) - f(x-X_1)\right\}\middle| N(t)=1\right] \\ = \ & \lambda e^{-\lambda t} \int\limits_0^t E_x\left[f\left(Y^{(A)}(s) - X_1\right) - f(x-X_1)\right]\ \frac{1}{t}\, ds. \end{aligned}$$

Note that $P(X_1 = x) = 0$. Since $f$ is continuous except possibly at 0 it follows that $f(Y^{(A)}(s) - X_1) \to f(x - X_1)$ a.e $P_x$ as $s \downarrow 0$. As $f$ is bounded we now have $E_x[f(Y^{(A)}(s) - X_1) - f(x - X_1)] \to 0$ as $s \downarrow 0$. Hence from the above it follows that $\frac{1}{t}\, J_{12}(t) \to 0$ as $t \downarrow 0$. Thus both the assertions in (6.61) are now established, completing the proof. □

**Corollary 6.24** *Hypotheses as in Theorem 6.21. Let $f \in \mathcal{D}_1$. Then (i) $L^{(A)}f$ is a right continuous function with 0 as the only possible discontinuity point, (ii) $L^{(A)}f(x) = 0$, $x < 0$, and (iii) $|L^{(A)}f(x)| \leq C(1 + x^2)$, $x \geq 0$ where $C$ is a constant. In addition, if $f'(0) = 0$ then $L^{(A)}f$ is continuous.* □

**Proposition 6.25** *Let hypotheses of Theorem 6.21 hold. Let $g$ be a right continuous function with 0 as the only possible discontinuity point, $g(x) = g_0$, a constant, for $x < 0$, and $|g(x)| \leq K[1 + x^2]$, $x \geq 0$. Then (i) $T_t^{(A)}g(x)$ is well defined for $t \geq 0$, $x \in \mathbb{R}$; also for $t_0 > 0$, there exist constants $C_1$, $C_2$ such that $|T_t^{(A)}g(x)| \leq C_1 + C_2(x \vee 0)^2$, for all $x \in \mathbb{R}$, $0 \leq t \leq t_0$; (ii) $\lim_{t \downarrow 0} T_t^{(A)}g(x) = g(x)$, $x \in \mathbb{R}$; (iii) for fixed $x \in \mathbb{R}$, $t \mapsto T_t^{(A)}g(x)$ is right continuous on $[0, \infty)$; (iv) for fixed $t \geq 0$, $x \mapsto T_t^{(A)}g(x)$ is a right continuous function. In particular, putting $g = L^{(A)}f$, with $f \in \mathcal{D}_1$, we see that $T_t^{(A)}L^{(A)}f(x)$ has the above properties.*

**Proof:** (i) We will prove the estimate in (i); well definedness will be evident on the way. Let $t_0 > 0$ be fixed. For any event $U$ we denote $P_x(U) \triangleq P(U \mid R^{(A)}(0) = x)$. Consider first $x \geq 0$. By our hypotheses on $g$, Theorem 6.13 and known estimates concerning Brownian driven SDE's we get for $0 \leq t \leq t_0$

$$\begin{aligned}
\left|T_t^{(A)}g(x)\right| &\leq \left|g_0 \cdot P_x\left(R^{(A)}(t) < 0\right)\right| \\
&\quad + \left|E_x\left[I_{[0,\infty)}\left(R^{(A)}(t)\right) \cdot g\left(R^{(A)}(t)\right)\right]\right| \\
&\leq |g_0| + KE_x\left[I_{[0,\infty)}\left(R^{(A)}(t)\right)\left\{1 + \left(R^{(A)}(t)\right)^2\right\}\right] \\
&\leq |g_0| + K + K\, E_x\left[\sup_{0 \leq t \leq t_0}\left|Y^{(A)}(t)\right|^2\right] \\
&\leq C_1 + C_2\, x^2 \qquad (6.62)
\end{aligned}$$

where $C_1$, $C_2$ are constants depending only on $t_0$, $K$, the Lipschitz and growth constants of $\sigma$, $b$.

Next let $x < 0$, and $\eta = \inf\{t \geq 0 : R^{(A)}(t) \geq 0\}$. Since jumps are negative it is clear that $R^{(A)}(\eta) = 0$ on $\{\eta < \infty\}$. Therefore by the strong Markov property (Theorem 6.16) we obtain for $0 \leq t \leq t_0$

$$\begin{aligned} \left|T_t^{(A)} g(x)\right| &= \left|g_0 P_x(\eta > t) + E_x\left[I_{\{\eta \leq t\}} g\left(R^{(A)}(t)\right)\right]\right| \\ &= \left|g_0 P_x(\eta > t) + E_x\left[I_{\{\eta \leq t\}} E_{R^{(A)}(\eta)}\left(g\left(R^{(A)}(t-\eta)\right)\right)\right]\right| \\ &\leq |g_0| + C_1. \end{aligned} \tag{6.63}$$

In the last step we have used (6.62) with 0 as the starting point. Thus part (i) follows from (6.62) and (6.63).

(ii) Let $x < 0$. As $\sigma^{(A)}(y) = 0$, $b^{(A)}(y) = c > 0$ for $y < 0$, it is clear that $R^{(A)}(t) < 0$ and hence $g(R^{(A)}(t)) = g(x)$ if $t < \frac{1}{c}|x|$. So the required conclusion is evident.

Next let $x \geq 0$. By right continuity of $S(\cdot)$ and continuity of Ito integrals in $t$, we see that $R^{(A)}(t) \to x$ a.s. as $t \downarrow 0$. At any $x > 0$, $g$ is continuous and hence $g(R^{(A)}(t)) \to g(x)$ as $t \downarrow 0$ for $x > 0$. Suppose $x = 0$. Since $\tau_1 > 0$ a.s., note that $R^{(A)}(t,\omega) = Y^{(A)}(t,\omega)$ for small $t$, a.e.$\omega$. So by Lemma 6.19, $R^{(A)}(t,\omega) \geq 0$ for small $t$, for a.e.$\omega$. Hence right continuity of $g$ at $x = 0$ now implies $g(R^{(A)}(t)) \to g(0)$ a.e. Thus $g(R^{(A)}(t)) \to g(x)$ a.s. as $t \downarrow 0$, for any $x \geq 0$. Now the quadratic growth condition on $g$ and the comparison result (Theorem 6.13) imply

$$\left|I_{[0,\infty)}\left(R^{(A)}(t)\right)\ g\left(R^{(A)}(t)\right)\right| \leq K\left[1 + \sup_{0 \leq s \leq 1}\ \left|Y^{(A)}(s)\right|^2\right]$$

for all $0 \leq t \leq 1$; cf. (6.62). As the r.h.s. of the above is known to be integrable, an application of the dominated convergence theorem completes the proof of (ii).

(iii) See Exercise 6.2.11.

(iv) Let $t > 0$ be fixed; (for $t = 0$ there is nothing to prove). Let $x_n \downarrow x_0$. For $n = 0, 1, 2, \ldots$ let $R^{(n)}(\cdot)$ denote the solution to the SDE (6.36) with initial value $x_n$. Let $y > 0$ be such that $x_0 \leq x_n < y$ for all $n$. Let $Y(\cdot)$ be the solution to the Brownian driven SDE (6.37) with initial value $y$. By the comparison result (Theorem 6.13) we have

$$R^{(0)}(t) \leq R^{(n)}(t) \leq Y(t) \text{ a.s. for all } n. \tag{6.64}$$

By Theorem 6.14 (i) we have $R^{(n)}(t) \to R^{(0)}(t)$ in probability. So, for any subsequence $\{x_{n'}\}$ of $\{x_n\}$ there is a further subsequence $\{x_{n''}\}$ such

that $R^{(n'')} \to R^{(0)}(t)$ a.s. As $g$ is right continuous, (6.64) now implies that $g(R^{(n'')}(t)) \to g(R^{(0)}(t))$ a.s. By our hypothesis on $g$ and (6.64) we see that

$$\left|g\left(R^{(n'')}(t)\right)\right| \leq C_1 + C_2\ |Y(t)|^2, \text{ for all } n''.$$

Since $Y(t)$ is square integrable, by dominated convergence theorem $E[g(R^{(n'')}(t))] \to E[g(R^{(0)}(t))]$ as $n'' \to \infty$. So any subsequence has a further subsequence converging to $E[g(R^{(0)}(t))]$. Therefore $E[g(R^{(n)}(t))] \to E[g(R^{(0)}(t))]$ as $n \to \infty$; that is $T_t^{(A)}\ g(x_n) \to T_t^{(A)}\ g(x)$ whenever $x_n \downarrow x_0$. Hence $T_t^{(A)}\ g(\cdot)$ is right continuous, completing the proof. □

**Lemma 6.26** *Let the hypotheses of Theorem 6.21 hold. For $f \in \mathcal{D}_1$, $x \in \mathbb{R}$, $t \mapsto T_t^{(A)}\ f(x)$ is a continuous function on $[0, \infty)$.*

**Proof:** Though we need to prove only left continuity at $t > 0$, our proof will also give continuity at $t > 0$. We follow the approach in [Dy], Vol.I, p.35; we just sketch the proof.

Let $0 < a < b < t,\ 0 < s < t - h,\ s < t,\ h$ small. As $\{T_r^{(A)} : r \geq 0\}$ is a contraction semigroup,

$$\begin{aligned}
\left\|T_{t-h}^{(A)} f - T_t^{(A)} f\right\| &:= \sup_{x \in \mathbb{R}} \left|T_{t-h}^{(A)} f(x) - T_t^{(A)} f(x)\right| \\
&= \sup_{x \in \mathbb{R}} \left|T_s^{(A)} \left[T_{t-h-s}^{(A)} f(x) - T_{t-s}^{(A)} f(x)\right]\right| \\
&\leq \left\|T_{t-h-s}^{(A)} f - T_{t-s}^{(A)} f\right\|. \qquad (6.65)
\end{aligned}$$

Integrating the above w.r.t. $s$ from $(t-b)$ to $(t-a)$,

$$\begin{aligned}
(b-a)\left\|T_{t-h}^{(A)} f - T_t^{(A)} f\right\| &\leq \int_{t-b}^{t-a} \left\|T_{t-h-s}^{(A)} f - T_{t-s}^{(A)} f\right\|\ ds \\
&= \int_a^b \left\|\left(T_{r-h}^{(A)} - T_r^{(A)}\right) f\right\| dr. \qquad (6.66)
\end{aligned}$$

Since $r \mapsto \|T_r^{(A)} f\|$ is integrable over $[0, t+1]$, by Exercise 6.1.6, r.h.s. of (6.66) tends to 0 as $h \to 0$. In fact we have proved $t \mapsto T_t^{(A)}\ f$ is continuous in the 'sup' norm. □

In the proof above, note that $f'$, $f''$ do not enter the picture; only boundedness of $f$ is used. This leads to the next lemma

**Lemma 6.27** *In addition to the hypotheses of Theorem 6.21, let $A(\cdot)$ be bounded. Then $t \mapsto T_t^{(A)}\ L^{(A)}\ f$ is continuous on $[0,\infty)$ in the sup norm, for any $f \in \mathcal{D}_1$.*

**Proof:** Note that $L^{(A)}f$ is bounded in this case. So the proof of preceding lemma applies to $L^{(A)}f$. □

**Proposition 6.28** *Assume the hypotheses of Theorem 6.21. For any $f \in \mathcal{D}_1$, $x \in \mathbb{R}$ the function $t \mapsto T_t^{(A)}f(x)$ has right derivative at any $t \geq 0$, and $\frac{\partial^+}{\partial t}\ T_t^{(A)}f(x) = T_t^{(A)}L^{(A)}f(x)$; also the right derivative is right continuous on $[0,\infty)$. Moreover for $0 \leq s \leq t$*

$$T_t^{(A)}f(x) - T_s^{(A)}f(x) = \int_s^t T_r^{(A)}\ L^{(A)}\ f(x)\ dr. \tag{6.67}$$

**Proof:** By the semigroup property and Theorem 6.21

$$\begin{aligned} \frac{T_{t+h}^{(A)}f(x) - T_t^{(A)}f(x)}{h} &= T_t^{(A)}\left\{\frac{1}{h}\left[T_h^{(A)}f(x) - f(x)\right]\right\} \\ &\to T_t^{(A)}L^{(A)}f(x), \text{ as } h \downarrow 0 \end{aligned} \tag{6.68}$$

proving the first assertion. As $T_t^{(A)}f(x)$ and $T_t^{(A)}L^{(A)}f(x)$ are respectively continuous and right continuous functions of $t$ (by Lemma 6.26 and Proposition 6.25), by Lemma 0.10 on p.237 of [Dy], vol.2, (6.67) and (6.68) are equivalent. □

The next result is the analogue of Theorem 6.8.

**Theorem 6.29** *Under the hypotheses of Theorem 6.21, for any $f \in \mathcal{D}_1$, any nonnegative square integrable $R^{(A)}(0)$,*

$$\begin{aligned} M_f^{(A)}(t) &= f\left(R^{(A)}(t)\right) - f\left(R^{(A)}(0)\right) \\ &\quad - \int_0^t L^{(A)}f\left(R^{(A)}(s)\right)\ ds,\ t \geq 0 \end{aligned} \tag{6.69}$$

*is a martingale w.r.t. $\{\mathcal{G}_t\}$.*

**Proof:** By Corollary 6.24 and Theorem 6.13, for $0 \leq s \leq t$ we have

$$\begin{aligned}\left|L^{(A)}f\left(R^{(A)}(s)\right)\right| &\leq I_{[0,\infty)}\left(R^{(A)}(s)\right)\left|L^{(A)}f\left(R^{(A)}(s)\right)\right| \\ &\leq K\, I_{[0,\infty)}\left(R^{(A)}(s)\right)\left[1+\left|R^{(A)}(s)\right|^2\right] \\ &\leq K\left[1+\left|Y^{(A)}(s)\right|^2\right] \\ &\leq K\left[1+\sup_{0\leq r\leq t}\left|Y^{(A)}(r)\right|^2\right]\end{aligned}$$

where $Y^{(A)}(0) = R^{(A)}(0)$. It follows that $L^{(A)}f(R^{(A)}(s))$, $0 \leq s \leq t$ and $M_f^{(A)}(t)$ are integrable. In view of the preceding proposition, imitating the proof of Theorem 6.8 the required result can now be established. □

Next, define the *ruin probability* of the process $R^{(A)}(\cdot)$ by

$$\begin{aligned}\psi^{(A)}(x) &\triangleq P_x\left(\tau_0^{(A)} < \infty\right) = P\left(\tau_0^{(A)} < \infty \,\middle|\, R^{(A)}(0) = x\right), x \geq 0 \\ &= 1, x < 0 \end{aligned} \tag{6.70}$$

and the *survival probability* by

$$\varphi^{(A)}(x) \triangleq 1 - \psi^{(A)}(x),\ x \in \mathbb{R}. \tag{6.71}$$

A connection between survival probability and the infinitesimal generator is indicated below

**Theorem 6.30** *In addition to the hypotheses of Theorem 6.21, assume that for any* $x \in \mathbb{R}$

$$\begin{aligned}&P_x\left(R^{(A)}(t) \to +\infty \text{ as } t \uparrow \infty\right) \\ =\ &P\left(\lim_{t\to\infty} R^{(A)}(t) = +\infty \,\middle|\, R^{(A)}(0) = x\right) = 1.\end{aligned} \tag{6.72}$$

*Let* $u \in \mathcal{D}_1$ *be such that*

$$\left.\begin{array}{ll}(i) & L^{(A)}u(x) = 0, \text{ if } x \geq 0 \\ (ii) & u(x) = 0, \text{ if } x < 0 \text{ and} \\ (iii) & \lim\limits_{x\to\infty} u(x) = 1\end{array}\right\}. \tag{6.73}$$

*Then* $u(x) = \varphi^{(A)}(x)$, $x \in \mathbb{R}$.

**Proof:** It is easily checked that $L^{(A)}u(x) = 0$ for $x < 0$; thus $L^{(A)}u(\cdot) \equiv 0$ on $\mathbb{R}$. Now fix $x \geq 0$. By Theorem 6.29 it follows that

$$u\left(R^{(A)}(t)\right) - u(x) = P_x - \text{ martingale w.r.t. } \{\mathcal{G}_t\}.$$

So by the optional sampling theorem

$$u(x) = E_x\left[u\left(R^{(A)}\left(\tau_0^{(A)} \wedge t\right)\right)\right] \text{ for all } t \geq 0.$$

It may be noted that $\tau_0^{(A)}$ is a stopping time w.r.t. $\{\mathcal{G}_t\}$ as the filtration is assumed to be right continuous. As $R^{(A)}(\tau_0^{(A)}) < 0$ on $\{\tau_0^{(A)} < \infty\}$ by Proposition 6.20, and $u(\cdot) = 0$ on $(-\infty, 0)$ we now get

$$\begin{aligned} u(x) &= \lim_{t\to\infty} E_x\left[u\left(R^{(A)}\left(\tau_0^{(A)} \wedge t\right)\right)\right] \\ &= \lim_{t\to\infty} E_x\left\{\left[I_{(0,t)}\left(\tau_0^{(A)}\right) \cdot u\left(R^{(A)}\left(\tau_0^{(A)}\right)\right)\right]\right. \\ &\quad \left. + \left[I_{[t,\infty)}\left(\tau_0^{(A)}\right) \cdot u\left(R^{(A)}(t)\right)\right]\right\} \\ &= \lim_{t\to\infty} E_x\left[I_{[t,\infty)}\left(\tau_0^{(A)}\right) \cdot u\left(R^{(A)}(t)\right)\right]. \end{aligned} \tag{6.74}$$

Now by (6.72), (iii) of (6.73) and (6.74) it follows that $u(x) = P_x(\tau_0^{(A)} = \infty) = \varphi^{(A)}(x)$. □

The requirement (6.72) means that the process $R^{(A)}(\cdot)$ wanders off to $+\infty$ with probability one. Recall that the analogous property of the Lundberg risk process $R(\cdot)$ played a key role earlier; the net profit condition ensured this property for $R(\cdot)$ by Theorem 5.1 and Exercise 5.1.7. We want to give a sufficient condition for (6.72) to hold. For this we need a lemma

**Lemma 6.31** *Assume the hypotheses of Theorem 6.21. In addition assume that the net profit condition holds, that is,* $(c - \lambda\ E(X_1)) > 0$*; (in particular, it is assumed that the claim size distribution has finite expectation* $E(X_1)$*). Then* $E_x(\eta_r^{(A)}) < \infty$ *for any* $r$, $x \in \mathbb{R}$, *where* $\eta_r^{(A)} = \inf\{t \geq 0 : R^{(A)}(t) \geq r\}$. *Consequently the process* $R^{(A)}(\cdot)$ *hits* $[r, \infty)$ *in finite time with probability one.*

**Proof:** The result holds trivially if $x \geq r$; so it is enough to consider $x < r$. Now take constant $h_1 > 0$ so that $h_1(c - \lambda E(X_1)) \geq 1$. Put $f(x) = h_1 x$, $x \in \mathbb{R}$. By Exercise 6.2.13,

$f(R^{(A)}(t)) - f(x) - \int_0^t L^{(A)} f(R^{(A)}(s))\, ds$, $t \geq 0$ is a $P_x$-martingale. Fix $r > 0$. Then for $x < r$, $t > 0$ by optional sampling theorem

$$\begin{aligned} E_x\left[\eta_r^{(A)} \wedge t\right] &\leq E_x\left[\int_0^{\eta_r^{(A)} \wedge t} L^{(A)} f\left(R^{(A)}(s)\right)\, ds\right] \\ &= E_x\left[f\left(R^{(A)}\left(\eta_r^{(A)} \wedge t\right)\right)\right] - f(x) \leq h_1(r-x). \end{aligned}$$

In the above we have used $R^{(A)}(\eta_r^{(A)}) = r$ since the jumps are negative. Letting $t \uparrow \infty$ we have the desired result. □

The next result is motivated by results concerning transience of continuous diffusions; proof follows that of Proposition 4.1 of [BR].

**Theorem 6.32** *Let the hypotheses of Theorem 6.21 and the net profit condition (5.26) hold. Suppose there exist $r_0 > 0$, $u \in \mathcal{D}_1$ such that (i) $u \geq 0$ and non-increasing, (ii) $u$ is strictly decreasing on $[r_0, \infty)$, and (iii) $L^{(A)}u(\cdot) \leq 0$ on $[r_0, \infty)$. Then (6.72) holds.*

**Proof:** Let $r_1 > r_0$ be arbitrary; let $\tau_1^{(A)} = \inf\{t \geq 0 : R^{(A)}(t) \leq r_1\}$. We claim that $P_x(\tau_1^{(A)} < \infty) < 1$ for all $x > r_1$. Suppose not; then there is $x > r_1$ with $P_x(\tau_1^{(A)} < \infty) = 1$. By Theorem 6.29, optional sampling theorem and (iii) above, for any $t > 0$

$$\begin{aligned} u(x) &\geq E_x\left[u\left(R^{(A)}\left(\tau_1^{(A)} \wedge t\right)\right)\right] \\ &\geq E_x\left[I_{[0,t)}\left(\tau_1^{(A)}\right) \cdot u\left(R^{(A)}\left(\tau_1^{(A)}\right)\right)\right]. \end{aligned}$$

Consequently

$$u(x) \geq \lim_{t\to\infty} E_x\left[I_{[0,t)}\left(\tau_1^{(A)}\right) \cdot u\left(R^{(A)}\left(\tau_1^{(A)}\right)\right)\right] \geq u(r_1)$$

as $R^{(A)}(\tau_1^{(A)}) \leq r_1$ and $u$ nonincreasing. This contradicts assumption (ii) of the theorem. Hence the claim holds.

Now let $r_1$, $r_2$ be arbitrary such that $r_0 < r_1 < r_2$. By the above paragraph and part (ii) of Theorem 6.13 it follows that

$$\begin{aligned} \sup\{P_x\left(\tau_1^{(A)} < \infty\right) : x \geq r_2\} &= P_{r_2}\left(\tau_1^{(A)} < \infty\right) \\ &\triangleq \delta < 1. \end{aligned} \tag{6.75}$$

Let $x \in \mathbb{R}$. Define $\zeta_1 = \inf\{t \geq 0 : R^{(A)}(t) \geq r_2\}$, and inductively for $i = 1, 2, \ldots$.

$$\begin{aligned} \zeta_{2i} &= \inf\left\{t > \zeta_{2i-1} : R^{(A)}(t) \leq r_1\right\}, \\ \zeta_{2i+1} &= \inf\left\{t > \zeta_{2i} : R^{(A)}(t) \geq r_2\right\}. \end{aligned}$$

By the strong Markov property, (6.75) and Lemma 6.31 we have for $i \geq 1$

$$\begin{aligned} & P_x\left(R^{(A)}(t) \leq r_1 \text{ for a sequence of } t's \uparrow \infty\right) \\ \leq\ & P_x\left(\zeta_{2i} < \infty\right) = E_x\left[I_{\{\zeta_{2i-1}<\infty\}} \cdot P_{R^{(A)}(\zeta_{2i-1})}\left(\tau_1^{(A)} < \infty\right)\right] \\ \leq\ & \delta\, P_x\left(\zeta_{2i-1} < \infty\right) = \delta\, P_x\left(\zeta_{2i-2} < \infty\right) \leq \delta^i. \end{aligned}$$

As $\delta < 1$ it now follows that

$$P_x\left(\liminf_{t\to\infty}\ R^{(A)}(t) > r_1\right) = 1.$$

Since $r_1 > r_0$ is arbitrary, the required result now follows. □

**Example 6.33** In addition to the hypotheses of Theorem 6.21, assume the following

(i) $A(0) = 0,\ A(x) > 0,\ x > 0,\ A(\cdot)$ nondecreasing;

(ii) $\zeta \geq 0,\ b > \zeta$, and there exist constants $\sigma_1 > 0,\ b_1 > 0$ such that for all $x \geq 1,\ \sigma^2(A(x))^2 \leq \sigma_1^2$ and $[\zeta x + (b - \zeta)A(x))] \geq b_1$;

(iii) Claim sizes $X_i,\ i \geq 1$ are i.i.d. random variables having exponential distribution with parameter $\theta > 0$;

(iv) net profit condition (5.26) holds; so $(c - \frac{\lambda}{\theta}) > 0$.

Note that $\sigma^{(A)}(\cdot)$ is bounded, while $b^{(A)}(\cdot)$ will be bounded only if $\zeta = 0$.

For suitable function $f$ write for $x > 0$

$$\begin{aligned} L^{(A)}_{\text{diff}} f(x) &= \frac{1}{2}\,\sigma^2(A(x))^2 f''(x) + [\zeta x + (b - \zeta)A(x)]f'(x), \\ L_{\text{jump}} f(x) &= cf'(x) + \lambda \int_0^\infty (f(x - z) - f(x))\, dF_X(z), \end{aligned}$$

so that

$$L^{(A)} f(x) = L^{(A)}_{\text{diff}} f(x) + L_{\text{jump}} f(x),\ x > 0.$$

Now choose a constant $K_0$ such that

$$0 < K_0 < \min\left\{\frac{2b_1}{\sigma_1^2}, \theta, \frac{\theta}{c}\left(c - \frac{\lambda}{\theta}\right)\right\}. \tag{6.76}$$

Define

$$u(x) = \begin{cases} \exp(-K_0 x), & x \geq 0 \\ 1, & x \leq 0 \end{cases}$$

Then it can be shown that $L^{(A)}_{\text{diff}} u(x) \leq 0$, $L_{\text{jump}} u(x) \leq 0$ for $x \geq 1$; hence $L^{(A)} u(x) \leq 0$ for $x \geq 1$. Consequently by Theorem 6.32 it follows that $R^{(A)}(t) \to +\infty$ with probability one. See Exercise 6.2.14. □

**Remark 6.34** We shall continue with the set up as in Example 6.33. Suppose $u \in \mathcal{D}_1$ solves the problem (6.73); in such a case, by Theorem 6.30, $u(x) = \varphi^{(A)}(x)$, $x \in \mathbb{R}$, the survival probability. As the claim size has an exponential distribution, we shall try to mimic the approach in Example 6.11. Put

$$\begin{aligned} h^{(A)}(x) &= E[u(x - X)] = E[\varphi^{(A)}(x - X)] \\ &= \int_0^\infty \varphi^{(A)}(x - z)\theta\, e^{-\theta z}\, dz, \ x \in \mathbb{R}. \end{aligned}$$

Since $\varphi^{(A)}(y) = 0$ for $y < 0$, it is easily seen that

$$h^{(A)}(x) = e^{-\theta x} \int_0^x \varphi^{(A)}(y)\ \theta\ e^{\theta y} dy, \quad x \geq 0. \tag{6.77}$$

As $u$ is continuous on $[0, \infty)$, by (6.77), $h^{(A)}$ is continuously differentiable on $[0, \infty)$; here we mean right differentiability at $x = 0$. Hence

$$\left(h^{(A)}\right)'(x) = \theta\left[\varphi^{(A)}(x) - h^{(A)}(x)\right], \quad x > 0. \tag{6.78}$$

Since $L^{(A)} u(x) = L^{(A)} \varphi^{(A)}(x) = 0$, $x > 0$ we get

$$\begin{aligned} &\left\{\frac{1}{2}\sigma^2 (A(x))^2 (\varphi^{(A)})''(x) + [\zeta x + (b - \zeta) A(x)](\varphi^{(A)})'(x)\right. \\ &\left. + c(\varphi^{(A)})'(x) + \lambda\, h^{(A)}(x) - \lambda\, \varphi^{(A)}(x)\right\} \\ = \ &0, \ x > 0. \end{aligned} \tag{6.79}$$

Now assume in addition that $A(\cdot)$ *is continuously differentiable on* $[0,\infty)$. Then (6.78), (6.79) imply that $\varphi^{(A)}$ is thrice continuously differentiable on $(0,\infty)$. Differentiate (6.79), and use (6.78), (6.79) again to obtain

$$\begin{aligned} & \frac{1}{2}\sigma^2(A(x))^2\ (\varphi^{(A)})'''(x) \\ & + \left\{ \begin{array}{l} \frac{\theta}{2}\sigma^2(A(x))^2 + \sigma^2 A(x)\ A'(x) \\ +\zeta x + (b-\zeta)A(x) + c \end{array} \right\} (\varphi^{(A)})''(x) \\ & + \left\{ \begin{array}{l} [\zeta + (b-\zeta)A'(x) - \lambda] \\ +\theta[c + \zeta x + (b-\zeta)A(x)] \end{array} \right\} (\varphi^{(A)})'(x) \\ = \ & 0,\ x > 0. \end{aligned} \tag{6.80}$$

Note that (6.80) is a third order ODE. However, if we put $w(x) \equiv w^{(A)}(x) = (\varphi^{(A)})'(x)$, $x > 0$, clearly (6.80) is a second order ODE for $w(\cdot)$. We want a nonconstant solution to (6.80). In fact $w(\cdot)$ should satisfy, as it is a density function on $(0,\infty)$,

$$w(x) > 0,\ x > 0 \ \text{ and } \ \int_0^\infty w(x)\ dx = 1 - \varphi^{(A)}(0). \tag{6.81}$$

Since $A(0) = 0$ note that $x = 0$ is a singular point for the ODE (6.80). In general the singularity can be irregular; (see [CL] for a discussion on singular points of ordinary differential equations). Even if $A(x) \sim \alpha x$, $x \downarrow 0$, the point $x = 0$ is not a regular singular point. It is not clear how to handle the ODE (6.80) so that (6.81) is satisfied. □

Further investigations concerning the process $R^{(A)}(\cdot)$ and the operator $L^{(A)}(\cdot)$ are needed to get sharper results. Even good examples will be helpful.

Hipp and Plum have considered the problem of optimal investment policy that minimizes the ruin probability; see [HP1], [HP2]. This is, however, beyond the scope of this introductory text.

## 6.3 Assorted comments

In this section we briefly comment on some topics not (adequately) discussed in the text.

Primary among them are insurance models based on finite state-space Markov chains. Life insurance models, pension funds, bonus-malus

models in motor insurance fall under this category. Time horizon in such models may be much longer compared to collective risk theory; moreover, there may be absorbing states. See [RSST], [KGDD] and appropriate references therein.

Reinsurance, which concerns insuring insurance companies, has not been touched upon. There are various types of reinsurance, like proportional reinsurance, stop-loss reinsurance, excess-loss reinsurance, etc. Usually reinsurance is very expensive so that the primary insurer is discouraged from shifting the entire liability to the reinsurer. [RSST], [EKM], [M] are good starting points.

Ruin problems for renewal risk models continue to be investigated, especially with interarrival times being Erlang distributions or phase-type distributions, and claim sizes being heavy tailed. Finite time ruin probability, deficit at ruin, surplus just before ruin are useful quantities. Semi-Markov models are also studied. For example see [As], [AB].

Heavy tailed distributions, especially subexponential distributions, attract attention for their own sake. [MN], [DDS] are recent examples.

Methods from optimization theory play significant role in insurance modelling. Strong suggestions to such an interplay were given forty years back by Borch in his seminal paper [Bc]; however, only recently some of these are being realized. As discussed in Section 6.2, insurance models with possibility of investment are being considered only now even for the Cramér-Lundberg framework. Amount invested in risky asset can be considered as a control, and one may seek to minimize ruin probability or maximize dividend payment. This naturally brings in ideas from control theory. See [T], [H], [HP1], [HP2]. Another instance of such an interaction concerns reinsurance involving two or more interacting companies with each trying to get an optimal deal. Ideas from game theory can not be far away; see [Aa], [Ra].

For multidimensional risk models see [BG], [Ra] and references therein.

There is more to insurance besides mathematical modelling. There are statistical questions like estimation of various parameters in the models, fitting appropriate claim size (or loss) distributions, etc. Some of these are addressed in [Bd]. Certain underpinning ideas from economics and finance have been indicated respectively in Sections 3.4.1 and 6.2; see [Bc], [Aa] for more of the former, and [H], [T] and references therein for more of the latter. Many practical issues are discussed in [BCHJKPR]. Of late, many methods from insurance are being applied to advantage in risk management involving banking and other financial

institutions; an authoritative account is given in [MFE].

## 6.4 Exercises

### Exercises on Section 6.1

**6.1.1.** For $t \geq 0$ let $\mathcal{E}^t$ be as in the proof of Proposition 6.2. For any $\mathcal{E}^t$-measurable integrable random variable $Y$ show that

$$E(Y \mid \mathcal{F}_t) = E(Y \mid R(t)) \ a.s.$$

This is, in fact, equivalent to (6.1).

**6.1.2.** Let $A \in \mathcal{F}_t$, $C \in \mathcal{E}^t$. Show that

$$P(A \cap C \mid R(t)) = P(A \mid R(t)) \cdot P(C \mid R(t)) \ \text{ a.s.}$$

(Hint: $E[I_A I_C \mid R(t)] = E[E[I_A I_C \mid \mathcal{F}_t] \mid R(t)]$ a.s.). We know that $\mathcal{F}_t$ can be regarded as the repository of all information concerning the "past", $\mathcal{E}^t$ as that of the "future", and $\sigma(R(t))$ as the "present". So the exercise means that conditional on the "present" (that is, given the "present"), the "past" and the "future" are independent. In fact this notion can be shown to be equivalent to the Markov property.

**6.1.3.** Let $T_t$, $t \geq 0$ be defined by (6.10); $f, f_1, f_2$ below denote bounded measurable functions.
(i) Show that $T_t f$ is well defined, bounded and

$$T_t f(x) = \int_{\mathbb{R}} f(z) \ P(0, x; t, dz), \ x \in \mathbb{R}.$$

(ii) Show that $T_t(f_1 + f_2) = T_t f_1 + T_t f_2$.
(iii) (*Semigroup property:*) For $t, s \geq 0$ show that

$$T_{t+s} f = T_t(T_s f) = T_s(T_t f).$$

(iv) Show that $\sup_{x \in \mathbb{R}} |T_t f(x)| \leq \sup_{x \in \mathbb{R}} |f(x)|$.

**6.1.4.** For $f \in \mathcal{D}$ show that $Lf(x) = 0$, $x < 0$ and $Lf(0) = cf'(0)$. (Hint: Recall that $X_i > 0$ a.s.)

**6.1.5.** Show that (6.12) holds also for any bounded continuously differentiable function $f$ on $\mathbb{R}$.

**6.1.6.** Let $\alpha : [t_0, t_1] \to \mathbb{R}$ be an integrable function. For any $[s_0, s_1] \subset (t_0, t_1)$ show that

$$\lim_{h \to 0} \int_{s_0}^{s_1} |\alpha(s+h) - \alpha(s)| ds = 0.$$

In particular the assertion holds for any bounded measurable function. (Hint: Result holds for continuous functions. For any closed set $A \subset (t_0, t_1)$, approximate $I_A$ pointwise by bounded continuous functions.)

**6.1.7.** (a) Let $g$ be a bounded right continuous function, and $t \geq 0$. Show that $x \mapsto T_t g(x)$ is bounded and right continuous.

(b) Let $g$ be as in (a). For any $x \in \mathbb{R}$, show that $t \mapsto T_t g(x)$ is continuous on $[0, \infty)$. (Hint: By semigroup property it is enough to show that $T_t g(x) \to g(x)$ as $t \downarrow 0$. Condition w.r.t. $N(t)$, and use (a).)

**6.1.8.** (a) For $f \in \mathcal{D}$ show that $Lf$ is a bounded right continuous function on $\mathbb{R}$.

(b) Let $f \in \mathcal{D}$, $x \in \mathbb{R}$. Show that $t \mapsto T_t Lf(x)$ is continuous on $[0, \infty)$.

**6.1.9.** If, for some $f$, the convergence in (6.12) is uniform over $x$, show that $t \mapsto T_t f(x)$ is differentiable for each $x$.

**6.1.10.** Show that $R(\tau_0) < 0$ whenever $\tau_0 < \infty$. (Hint: Suppose that $R(\tau_0) = 0$. Then using right continuity of sample paths, and by taking suitable subsequences, if necessary, there exists $\{t_k\}$ such that $t_k \downarrow \tau_0$, $R(t_k) \uparrow 0$. As $R(\cdot)$ is strictly increasing in between jumps, this means there is a jump in $[t_{k+1}, t_k]$. Arrive at a contradiction).

## Exercises on Section 6.2

**6.2.1.** (i) Using stochastic calculus show that (see [KS]) $Z(\cdot)$ satisfying (6.31) is given by

$$Z(t) = z_0 \exp\left\{\left(b - \frac{1}{2}\sigma^2\right)t + \sigma B(t)\right\}, \quad t \geq 0.$$

(ii) Show that the total claim amount $S(\cdot)$ and the unit price $Z(\cdot)$ of the risky asset are independent as processes $\Leftrightarrow$ $\{X_j, N(s) : j = 1, 2, \ldots, s \geq 0\}$ and $\{B(t) : t \geq 0\}$ are independent.

(iii) For any $t \geq 0$ show that $\sigma\{Z(u), S(r) : 0 \leq r, u \leq t\} = \sigma\{B(r), S(u) : 0 \leq r, u \leq t\}$.

(iv) It will generally be assumed that $b > \zeta$. Why is this a reasonable assumption?

**6.2.2.** (i) Let $\theta(t)$ = number of units of the risky asset bought at time $t$. So $A(R^{(A)}(t-)) = \theta(t)\ Z(t)$ = amount invested in risky asset at time. (Why ?) Argue that $R^{(A)}$ is governed by the SDE

$$\begin{aligned} dR^{(A)}(t) &= c\ dt - dS(t) + \theta(t)\ dZ(t) \\ &\quad +\zeta \cdot [R^{(A)}(t) - \theta(t)\ Z(t)]\ dt, \quad t > 0. \end{aligned}$$

Using Exercise 6.2.1 (i), now obtain (6.32).

(ii) Show that the SDE's (6.32) and (6.33) are the same with probability one.

**6.2.3.** Formulate a suitable notion of a strong solution to (6.39), and establish the analogues of Theorems 6.12 - 6.14 for the SDE (6.39).

**6.2.4.** Work out the details of the proof of Theorem 6.15.

**6.2.5.** (i) Let $\tau$ be a stopping time w.r.t $\{\mathcal{G}_t\}$ and $\mathcal{G}_\tau$ the associated pre-$\tau$ $\sigma$-algebra. Show that $\mathcal{G}_\tau$ is indeed a $\sigma$-algebra, and that $\tau$ is $\mathcal{G}_\tau$-measurable. (It can be shown that $R^{(A)}(\tau)$, defined on $\{\tau < \infty\}$, is also $\mathcal{G}_\tau$-measurable; see p.9 of [KS].)

(ii) Show that the jump times $\tau_n$, $n \geq 1$ and the ruin time $\tau_0^{(A)}$ are stopping times w.r.t. $\{\mathcal{G}_t\}$; (for the second assertion you may need the assumption that the filtration $\{\mathcal{G}_t\}$ is right continuous, that is, $\mathcal{G}_t = \mathcal{G}_{t+} \triangleq \bigcap_{h>0} \mathcal{G}_{t+h}$ for all $t \geq 0$).

**6.2.6.** The purpose of this exercise is to show that nondegeneracy of the dispersion coefficient at $x = 0$ leads to immediate ruin of $R^{(A)}$. Let $A$, $\sigma^{(A)}$, $b^{(A)}$, $R^{(A)}$, $Y^{(A)}$ be as in (6.36), (6.37). Let the ruin time $\tau_0^{(A)}$ of $R^{(A)}$ be defined by (6.51) and put $\tilde{\tau}_0 = \inf\{t \geq 0 : Y^{(A)}(t) < 0\}$. Assume that $(\sigma^{(A)}(0))^2 > 0$; so by continuity, $\inf\{(\sigma^{(A)}(y))^2 : y \in [-x_0, x_0]\} > 0$ for some $x_0 > 0$.

(i) To prove $P(\tau_0^{(A)} = 0 \mid R^{(A)}(0) = 0) = 1$ it is enough to prove that $P(\tilde{\tau}_0 = 0 \mid Y^{(A)}(0) = 0) = 1$.

(ii) Show that one may assume without loss of generality that $b^{(A)} \equiv 0$. (Hint: Enough to prove that $P(\tilde{\tau}_0 \wedge T = 0 \mid Y^{(A)}(0) = 0) = 0\ \forall\ T > 0$. Then use Girsanov's theorem for continuous diffusions; see p.202 of [Du] or p.302 of [KS].)

(iii) Now we can take $Y^{(A)}(t) = \int_0^t \sigma^{(A)}(Y^{(A)}(s))\ dB(s)$, $t \geq 0$. Assume that $\inf_{y \in \mathbb{R}} (\sigma^{(A)}(y))^2 > 0$. Using Theorem 3.4.6, p.174 of [KS] (or Theorem 4.4, Chap. 3, p.113 of [Du]) conclude that $Y^{(A)}$ is a time-changed Brownian motion. Then use Blumenthal's 0-1 law (p.14 of [Du] or p.94 of [KS]) to conlcude that $P(\tilde{\tau}_0 = 0 \mid Y^{(A)}(0) = 0) = 1$.

(iv) Remove the ad hoc assumption $\inf_{y\in\mathbb{R}} (\sigma^{A)}(y))^2 > 0$.

**6.2.7.** Let $A \equiv 0$, $c > 0$, $\zeta \geq 0$. Let $R_\zeta$ denote the corresponding risk process; (this means that investment is made only in the riskless bond). If $R_\zeta(0) \geq 0$, show that the ruin time is strictly positive, and that the ruin probability is strictly less than 1.

**6.2.8.** Give a proof of Lemma 6.22.
(Hint: Fix $x \in \mathbb{R}$, $f \in \mathcal{D}_1$; let $K > 0$ which will be chosen suitably. Let $f_K \in \mathcal{D}_1$ be such that $x + K > 0$, $f_K(y) = f(y)$, $y \leq x + K$, $f_K(y) = 0$, $y \geq x + 2K$, and $\|f_K\|_\infty := \sup |f_K(y)| \leq \sup |f(y)| =: \|f\|_\infty$. By comparison result Theorem 6.13 and Chebyshev's inequality, for all $0 < t \leq 1$,

$$\begin{aligned} & \frac{1}{t} E_x \left[ f\left(R^{(A)}(t)\right) - f_K\left(R^{(A)}(t)\right)\right] \\ \leq & \frac{1}{t} 2\|f\|_\infty P_x\left(Y^{(A)}(t) - x \geq K\right) \\ \leq & \frac{1}{K^2 t} 2\|f\|_\infty E_x\left(\left|Y^{(A)}(t) - x\right|^2\right). \end{aligned}$$

Using standard arguments for SDE's driven by Brownian motion, show that for $0 \leq t \leq 1$, $E_x(|Y^{(A)}(t) - x|^2) \leq ct$ where $c$ is a constant independent of $t, x, K$.)

**6.2.9.** For $f \in \mathcal{D}_1$, $x < 0$ show that both sides of (6.52) are zero. (Hint: As the jumps are negative, note that $\eta_0 \triangleq \inf\{t \geq 0 : R^{(A)}(t) \geq 0\} = \inf\{t \geq 0 : R^{(A)}(t) = 0\}$. Since $\sigma^{(A)}(\cdot) = 0$, $b^{(A)}(\cdot) = c > 0$ on $(-\infty, 0)$, clearly $\eta_0 \geq \frac{1}{c}|x|$ a.s. $P_x$, for $x < 0$.)

**6.2.10.** Take $A(x) = ax$, $x \geq 0$ in Lemma 6.19, where $a > 0$ is a constant. Show directly (without invoking the support theorem) that the conclusion of the lemma holds.

**6.2.11.** (a) Let $g$ be as in Proposition 6.25; (assume also the hypotheses and notation of Proposition 6.25). For $h \geq 0$ set $g_h(y) = T_h^{(A)} g(y) - g(y)$, $y \in \mathbb{R}$. By (ii) of Proposition 6.25, $\lim_{h\downarrow 0} g_h(y) = 0$, $y \in \mathbb{R}$. Using (i) of Proposition 6.25 for $t \geq 0$, $0 \leq h \leq 1$, show that

$$\left|g_h\left(R^{(A)}(t)\right)\right| \leq C_1 + C_2\left|Y^{(A)}(t)\right|^2.$$

(b) Show that $T_t^{(A)} g_h(x) \to 0$ as $h \downarrow 0$, for fixed $t \geq 0$, $x \in \mathbb{R}$, to prove part (iii) of Proposition 6.25.

**6.2.12.** Supply the details of the proof of Lemma 6.26. (Justify the measurability and integrability assumptions.)

**6.2.13.** In addition to the hypotheses of Theorem 6.21, assume that the i.i.d. claim sizes $X_i$ have finite expectation. Let $h$ be a constant and $f(x) = hx,\ x \in \mathbb{R}$. Show that for any $x \in \mathbb{R}$,

$$f\left(R^{(A)}(t)\right) - f(x) - \int_0^t L^{(A)} f\left(R^{(A)}(s)\right)\, ds, t \geq 0$$

is a $P_x$-martingale w.r.t. $\{\mathcal{G}_t\}$. (Hint: Note that the given expression is the same as

$$h\left[\int_0^t \sigma^{(A)}\left(R^{(A)}(s)\right)\, dB(s) - \{S(t) - \lambda E(X_1)t\}\right].$$

Now use the proof of Theorem 6.14 (i).)

**6.2.14.** Let the notation and hypotheses be as in Example 6.33.

(a) As $K_0 < 2b_1/\sigma_1^2$, using condition (ii) in Example 6.33 show that

$$L_{\text{diff}}^{(A)}\, u(x) \leq 0,\ x \geq 1.$$

(b) Verify that

$$\int_0^\infty [u(x-z) - u(x)]\, dF_X(z) = \frac{1}{(K_0 - \theta)}\, \lambda K_0[e^{-\theta x} - e^{-K_0 x}],\quad x > 0.$$

Now using (6.76) and (NPC) show that $L_{\text{jump}}\, u(x) \leq 0$ for $x \geq 1$.

# Appendix

## A.1 Basic notions

We review here some basic notions of probability theory. Detailed accounts can be found in any text book on probability theory, for example, [Bl2], [Bo], [CT], [P2], [V1].

Let $(\Omega, \mathcal{F}, P)$ be a *probability space.* Here $\Omega$ is a nonempty set, and $\mathcal{F}$ is a nonempty collection of subsets of $\Omega$ satisfying the following:
(i) $\phi \in \mathcal{F},\ \Omega \in \mathcal{F}$;
(ii) $B \in \mathcal{F} \Rightarrow$ the complement $B^c \in \mathcal{F}$;
(iii) $B_i \in \mathcal{F};\ i = 1, 2, \ldots \Rightarrow \bigcup_{i=1}^{\infty} B_i \in \mathcal{F}$. The collection $\mathcal{F}$ is called a *$\sigma$-algebra* (or a *$\sigma$-field*) of events. $P$ is a *probability measure* on $(\Omega, \mathcal{F})$; that is, $P : \mathcal{F} \to [0, 1]$ such that
(a) $P(\phi) = 0,\ P(\Omega) = 1$, and
(b) if $B_i \in \mathcal{F},\ i = 1, 2, \ldots$ with $B_i \cap B_j = \phi$ whenever $i \neq j$, then $P\left(\bigcup_{i=1}^{\infty} B_i\right) = \sum_{i=1}^{\infty} P(B_i)$.

A function $X : \Omega \to \mathbb{R}$ is called a (*real valued*) *random variable* (abbreviated r.v.) if $X^{-1}(E) \in \mathcal{F}$ for any Borel set $E$ in $\mathbb{R}$; so a random variable is a Borel measurable function. This is equivalent to saying $X^{-1}(G) \in \mathcal{F}$ for any open set $G \subseteq \mathbb{R}$. For a real valued random variable $X$, define

$$F_X(x) \triangleq P(X^{-1}((-\infty, x])) \triangleq P(X \leq x),\ x \in \mathbb{R}. \tag{A.1.1}$$

The function $F_X$ is called the *distribution function* of the random variable $X$. It can be easily proved that $F_X$ is a nondecreasing right continuous function on $\mathbb{R}$ with $F_X(-\infty) \triangleq \lim_{x \downarrow -\infty} F_X(x) = 0$ and $F_X(+\infty) \triangleq \lim_{x \uparrow +\infty} F_X(x) = 1$. Also the set of discontinuity points of $F_X$ can be at

most countably infinite.

If there is a countable subset $C = \{x_1, x_2, \ldots\}$ of $\mathbb{R}$ such that $P(X \in C) = 1$ then $X$ is called a *discrete random variable*; in this case $C$ is the set of discontinuity points of $F_X$. For such a random variable put

$$p_X(x) \triangleq \begin{cases} P(X = x_k) \triangleq p_k, \text{ if } x = x_k \text{ for some } k, \\ 0, \text{ otherwise} \end{cases} \tag{A.1.2}$$

The function $p_X$ on $\mathbb{R}$ is called the *probability mass function* of the discrete random variable $X$. Note that $p_k \geq 0 \ \forall \ k$ and $\sum\limits_k p_k = 1$. If $C = \{0, 1, 2, \ldots\}$ then $X$ is a nonnegative integer valued r.v.

If $F_X$ is a continuous function then $X$ is said to be a *continuous random variable*; in this case, for any fixed $x \in \mathbb{R}$ note that $P(X = x) = 0$.

If there is a nonnegative integrable function $f_X$ on $\mathbb{R}$ such that

$$F_X(x) = \int_{-\infty}^{x} f_X(y)\, dy, \ x \in \mathbb{R} \tag{A.1.3}$$

then $X$ is said to be an *absolutely continuous random variable*; the function $f_X$ in (A.1.3) is called the *probability density function* (or *p.d.f.* for short) of $X$. Note that $\int\limits_{\mathbb{R}} f_X(x)\, dx = 1$.

For a real valued random variable $X$ one can define a probability measure $\nu_X$ on the real line by $\nu_X\ (B) = P(X^{-1}(B))$, for any Borel set $B \subseteq \mathbb{R}$; this probability measure $\nu_X$ may be called the *distribution of* $X$. Clearly $F_X(x) = \nu_X((-\infty, x]),\ x \in \mathbb{R}$. If $X$ represents the random characteristic under study, all the relevant information is captured by $F_X$ or equivalently by $\nu_X$. Consequently the focus of attention can be $F_X$ or $\nu_X$.

For a Borel measurable function $g$ on $\mathbb{R}$, if

$$\int\limits_{\mathbb{R}} |g(x)|\, dF_X(x) \equiv \int\limits_{\mathbb{R}} |g(x)|\, d\, \nu_X(x) < \infty$$

then we say that $E(g(X))$ exists; in such a case, we define

$$E(g(X)) \triangleq \int\limits_{\mathbb{R}} g(x)\, dF_X(x) \equiv \int\limits_{\mathbb{R}} g(x)\, d\nu_X(x). \tag{A.1.4}$$

As $F_X$ is a monotone function, Stieltjes integral w.r.t. $dF_X(\cdot)$ makes sense and is the same as integral w.r.t. the measure $d\nu_X(\cdot)$. If $X$ is a discrete r.v. then $E(g(X)) = \sum_k g(x_k)\, p_k$, provided the sum converges absolutely. If $X$ is an absolutely continuous random variable with p.d.f. $f_X$ then $E(g(X)) = \int_{-\infty}^{\infty} g(x)\, f_X(x)\, dx$, provided the integral exists. $E(X^k)$ is called the *k-th moment of* $X$; $E(X)$ is known as *expectation* or *mean* of $X$. *Variance* of $X$ is given by $\mathrm{Var}(X) \triangleq E[(X - E(X))^2]$.

In a similar fashion, an $\mathbb{R}^n$-valued measurable function on a probability space is called an $\mathbb{R}^n$-*valued random variable.* Such a random variable can be represented as $(X_1, X_2, \ldots, X_n)$ where each $X_i$ is a real valued random variable. Also if $X_i$ is a real valued random variable for $i = 1, 2, \ldots, n$ then $(X_1, X_2, \ldots, X_n)$ is an $\mathbb{R}^n$-valued random variable. An $\mathbb{R}^n$-valued r.v. $(X_1, \ldots, X_n)$ is said to be *absolutely continuous* if there is a nonnegative integrable function $f_{(X_1,\ldots,X_n)}$ on $\mathbb{R}^n$ such that for any $a_i < b_i$, $i = 1, \ldots, n$ we have

$$\begin{aligned} & P(a_1 < X_1 \le b_1,\ a_2 < X_2 \le b_2, \ldots,\ a_n < X_n \le b_n) \\ = & \int_{a_n}^{b_n} \cdots \int_{a_2}^{b_2} \int_{a_1}^{b_1} f_{(X_1,\ldots,X_n)}(x_1, x_2, \ldots, x_n)\, dx_1 \ldots dx_n. \end{aligned}$$

Here the function $f_{(X_1,\ldots,X_n)}$ is also referred to as the *joint probability density function of* $X_1, X_2, \ldots, X_n$. We may also say that $X_1, X_2, \ldots, X_n$ are jointly distributed according to $f_{(X_1,\ldots,X_n)}$.

Real valued random variables $X_1, X_2, \ldots, X_n$ are said to be *independent* if for any Borel sets $B_1, B_2, \ldots, B_n$ in $\mathbb{R}$

$$\begin{aligned} & P(X_1 \in B_1,\ X_2 \in B_2, \ldots, X_n \in B_n) \\ = & P(X_1 \in B_1) \cdot P(X_2 \in B_2) \cdots P(X_n \in B_n). \end{aligned}$$

If real valued absolutely continuous random variables $X_1, X_2, \ldots, X_n$ are independent note that the $\mathbb{R}^n$-valued random variable $(X_1, X_2, \ldots, X_n)$ is also absolutely continuous and

$$f_{(X_1,\ldots,X_n)}\,(x_1, x_2, \ldots, x_n) = f_{X_1}(x_1) \cdot f_{X_2}(x_2) \ldots f_{X_n}(x_n) \qquad \text{(A.1.5)}$$

for all $(x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$.

Let $X_1, X_2, \ldots, X_n$ have a joint probability density function $f_{(X_1,\ldots,X_n)}$.

For a Borel measurable function $g$ on $\mathbb{R}^n$ we define

$$E(g(X_1,\ldots,X_n)) = \int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} g(x_1,\ldots,x_n)\, f_{(X_1,\ldots,X_n)}(x_1,\ldots,x_n)\, dx_1\ldots dx_n$$

provided the integral exists. In particular let $X_1, X_2$ have joint p.d.f. $f_{(X_1,X_2)}$; let $X_1, X_2$ have finite expectations $m_1, m_2$ respectively. One can then define the *covariance* between $X_1, X_2$ by

$$\begin{aligned}&\text{Cov }(X_1,X_2) \triangleq E[(X_1-m_1)(X_2-m_2)]\\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(x_1-m_1)(x_2-m_2) f_{(X_1,X_2)}(x_1,x_2)\, dx_1\, dx_2 \quad \text{(A.1.6)}\end{aligned}$$

whenever the integral on the r.h.s. exists. If $X_1, X_2$ are independent r.v.'s having finite means, then $E(X_1X_2) = E(X_1)\cdot E(X_2)$ and $\text{Cov}(X_1,X_2) = 0$; the converse is not true. In general if $E(X_1^2) < \infty$, $E(X_2^2) < \infty$ then $E(X_1), E(X_2)$, Cov $(X_1,X_2)$ exist.

An important aspect of probability theory is the study of sums of independent identically distributed random variables; these arise at least as a first step of stochastic modelling in various situations.

Let $X, Y$ be independent random variables with respective distribution functions $F_X, F_Y$. Then the distribution function of $X+Y$ is given by

$$(F_X * F_Y)(z) \equiv (F_Y * F_X)(z) \triangleq \int_{\mathbb{R}} F_X(z-y)\, dF_Y(y),\ z\in\mathbb{R}. \quad \text{(A.1.7)}$$

$F_X * F_Y$ is called the *convolution* of $F_X$ and $F_Y$. Thus if $X, Y$ are independent then $F_{X+Y} = F_X * F_Y$. Let $X, Y$ be independent r.v.'s having a common distribution function $F$; in such a case the distribution of $X+Y$ is governed by $F^{*(2)} \triangleq F*F$. In the same spirit if $X_1, X_2, \ldots, X_n$ are independent identically distributed random variables (*i.i.d.r.v.'s*, for short) then the distribution function of $X_1+X_2+\ldots+X_n$ is the *n-fold convolution* $F*\ldots*F = F^{*(n)}$. In terms of distributions, (A.1.7) can be written as

$$(\nu_X * \nu_Y)(G) \equiv (\nu_Y * \nu_X)(G) \triangleq \int_{\mathbb{R}} \nu_X(G-y)\, d\nu_Y(y)$$

for any Borel set $G \subseteq \mathbb{R}$. So if $X_1, X_2, \ldots, X_n$ are i.i.d.r.v.'s with common distribution $\nu$, then the distribution of $X_1 + \ldots + X_n$ is the $n$-fold convolution $\nu^{*(n)} = \nu * \ldots * \nu$.

Various notions of convergence are possible in the context of random variables. Let $X_n$, $n \geq 1$ and $X$ be random variables. We say that $\{X_n\}$ *converges to* $X$ *almost surely* (or *with probability one*) if $\exists\, M \in \mathcal{F}$ such that $P(M) = 0$ and $X_n(\omega) \to X(\omega)$ for $\omega \notin M$; in other words $P(\lim X_n = X) = 1$. This is often denoted $X_n \to X$ a.s. or $X_n \to X$ w.p.1. We say that $\{X_n\}$ *converges to* $X$ *in probability* if for any $\epsilon > 0$

$$\lim_{n\to\infty} P(\{\omega : |X_n(\omega) - X(\omega)| \geq \epsilon\}) = 0.$$

Almost sure convergence is stronger than convergence in probability. Next, suppose $F_n$, $n \geq 1$ and $F$ denote the corresponding distribution functions. We say that $\{X_n\}$ *converges to* $X$ *in distribution* (or *in law*, or *weakly*) if $\lim_{n\to\infty} F_n(x) = F(x)$ whenever $x$ is a continuity point of $F$; this is equivalent to

$$\lim_{n\to\infty} \int_{\mathbb{R}} g(z)\, dF_n(z) = \int_{\mathbb{R}} g(z)\, dF(z)$$

for any bounded continuous function $g$, or in terms of the corresponding distributions

$$\lim_{n\to\infty} \int g(z)\, d\nu_n(z) = \int g(z)\, d\nu(z)$$

for any bounded continuous function $g$. This is denoted $X_n \xrightarrow{d} X$, $X_n \xrightarrow{w} X$, $X_n \Rightarrow X$, $F_n \Rightarrow F$ or $\nu_n \Rightarrow \nu$. (Since this notion depends only on the distribution of the random variables, $X_n$'s and $X$ need not even be defined on the same probability space.)

We now state two basic limit theorems of probability theory.

**Theorem A.1.1** *(strong law of large numbers): Let $X_1, X_2, \ldots$ be a sequence of independent identically distributed random variables with finite expectation $m$. Then*

$$\frac{1}{n}\,(X_1 + X_2 + \ldots + X_n) \to m \text{ a.s. as } n \to \infty. \tag{A.1.8}$$

*If $\{X_i\}$ are i.i.d nonnegative random variables then (A.1.8) holds even if $m = +\infty$.* □

The preceding theorem forms the philosophical basis for taking averages in many situations.

**Theorem A.1.2** *(Lindeberg-Levy form of central limit theorem): Let* $\{X_n\}$ *be a sequence of i.i.d. random variables with finite nonzero variance. Put* $S_n = X_1 + \ldots + X_n$, $Z_n = \frac{S_n - nE(X_1)}{\sqrt{n\,Var(X_1)}}$, $n \geq 1$. *Then for any* $a \leq b$

$$\lim_{n\to\infty} P(a < Z_n \leq b) = \int_a^b \frac{1}{\sqrt{2\pi}}\, e^{-\frac{1}{2}x^2} dx. \tag{A.1.9}$$

□

The above classical central limit theorem provides the mathematical basis for bulk of statistics, and explains the prevalence of the Gaussian (or normal) distribution in diverse disciplines.

Some issues related to the central limit problem are briefly outlined in Appendix A.2.

Some important one dimensional distributions are given below.

1. **Discrete Uniform distribution:** Let $x_1, \ldots, x_K$ be distinct points in $\mathbb{R}$, where $K$ is a fixed integer. The probability mass function in this case is $p_X(x_i) = \frac{1}{K}$, $1 \leq i \leq K$ and $p_X(x) = 0$ if $x \neq x_i$ for any $i$. In this case $E(X)$ is just the average of the numbers $x_1, \ldots, x_K$.

2. **Binomial distribution:** Let $n \geq 1$ be an integer and $0 < p < 1$. Let $p_X(\cdot)$ be given by

$$p_X(k) = \binom{n}{k} p^k (1-p)^{n-k},\ k = 0, 1, \ldots, n$$

and $p_X(x) = 0$ if $x \notin \{0, 1, \ldots, n\}$. This is the *binomial distribution with parameters* $n$ *and* $p$. In this case $E(X) = np$ and $\mathrm{Var}(X) = np(1-p)$.

3. **Poisson distribution:** Let $\lambda > 0$. Here the discrete random variable $X$ takes value in $\{0, 1, 2, \ldots\}$ with probabilities

$$p_X(n) = P(X = n) = e^{-\lambda}\, \frac{\lambda^n}{n!},\quad n = 0, 1, 2, \ldots.$$

This is the *Poisson distribution with parameter* $\lambda$; and $E(X) = \lambda$, $\mathrm{Var}\,(X) = \lambda$.

4. **Geometric distribution:** Let $0 < p < 1$, $X$ take value in $\{0, 1, 2, \ldots\}$ and

$$p_X(j) = P(X = j) = p(1-p)^j, \ j = 0, 1, \ldots .$$

The discrete distributions (1) - (4) above have finite moments of any order. Let $0 < p < 1$ be fixed, and $X_1, \ldots, X_k$ be independent random variables; if $X_i$ has a binomial distribution with parameters $n_i$ and $p$ for $1 \leq i \leq k$ then $X_1 + X_2 + \ldots X_k$ has a binomial distribution with parameters $n_1 + \ldots + n_k$ and $p$. Similarly, if $X_1, X_2, \ldots, X_n$ are independent random variables with $X_i$ having a Poisson distribution with parameter $\lambda_i$, $1 \leq i \leq n$, then $X_1 + X_2 + \ldots + X_n$ has a Poisson distribution with parameter $\lambda_1 + \ldots + \lambda_n$.

Next we list some absolutely continuous distributions.

5. **Uniform distribution** $U(a, b)$**:** Here $a < b$ and the probability density function is

$$f_X(x) = \frac{1}{(b-a)} \ I_{(a,b)}(x).$$

6. **Exponential distribution** Exp($\lambda$): Here $\lambda > 0$ and the p.d.f. of the *exponential distribution with parameter* $\lambda$ is given by

$$f_X(x) = \begin{cases} \lambda e^{-\lambda x}, & x > 0 \\ 0, & \text{otherwise.} \end{cases}$$

It is easily seen that $E(X) = \frac{1}{\lambda}$, $\mathrm{Var}(X) = \frac{1}{\lambda^2}$.

7. **Normal distribution** $N(\mu, \sigma^2)$: Let $\mu \in \mathbb{R}$, $\sigma^2 > 0$ be constants and

$$f_X(x) = \frac{1}{\sqrt{2\pi}} \ \frac{1}{\sigma} \ \exp\left\{-\frac{1}{2\sigma^2}(x-\mu)^2\right\}, \ x \in \mathbb{R}.$$

The above $f_X$ is called the p.d.f. of *the normal distribution with mean* $\mu$ *and variance* $\sigma^2$. When $\mu = 0$, $\sigma^2 = 1$ the distribution is called the *standard normal distribution.* If $X_i$ has $N(\mu_i, \sigma_i^2)$ distribution for $1 \leq i \leq k$, and $X_1, \ldots, X_k$ are independent then $X_1 + X_2 + \ldots + X_k$ has $N(\mu_1 + \ldots + \mu_k, \ \sigma_1^2 + \sigma_2^2 + \ldots + \sigma_k^2)$ distribution. Observe that Theorem A.1.2 basically says that $Z_n \xrightarrow{d} N(0, 1)$.

8. **Gamma distribution** $\Gamma(\alpha, \lambda)$: Let $\alpha > 0$, $\lambda > 0$ be constants, and let the p.d.f. be

$$f_X(x) = \begin{cases} \frac{\lambda^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\lambda x}, & x > 0 \\ 0, & \text{otherwise.} \end{cases}$$

If $X_1, X_2$ are independent random variables having respectively $\Gamma(\alpha_1, \lambda)$, $\Gamma(\alpha_2, \lambda)$ distributions then $X_1 + X_2$ has $\Gamma(\alpha_1 + \alpha_2, \lambda)$ distribution. Note that $\Gamma(1, \lambda) = \text{Exp}\ (\lambda)$.

All the distributions (1) - (8) are light-tailed.

## A.2 On the central limit problem

The classical central limit theorem says that if $\{X_i\}$ is an i.i.d. sequence with finite nonzero variance, then $\frac{S_n - nE(X_1)}{\sqrt{n}\sqrt{\text{Var}(X_1)}}$ is asymptotically standard normal (in the sense of convergence in distribution), where $S_n = (X_1 + X_2 + \ldots + X_n)$, $n \geq 1$ denotes the sequence of partial sums. So, if variance exists then the partial sum $S_n$ is approximately normal (or Gaussian) for large $n$; this approach has proved to be extremely useful. There are also very interesting/significant extensions giving asymptotic normality.

However, when second moments do not exist the above result is of limited use, as it does happen in actuarial context involving heavy tailed distributions. Still, a discussion of the central limit problem gives a good perspective regarding motivation for the various definitions encountered in the sequel. It may be mentioned that the central limit problem was investigated purely as a mathematical curiosity, and many outstanding results were obtained by Levy, Doeblin, Kolmogorov, Khinchin, Feller and others.

Motivated by the CLT, the following question is meaningful.
**Question:** What are the possible nondegenerate limit distributions for partial sums $S_n$ (of i.i.d.'s) when suitably centered and normalized?

This is related to the following definition.

**Definition A.2.1** Let $G$ be a distribution function on $\mathbb{R}$ such that $\mu_G \neq \delta_x$ for any $x \in \mathbb{R}$. A distribution function $F$ is said to belong to the *domain of attraction of* $G$ if there exist constants $a_n > 0$, $b_n \in \mathbb{R}$, $n = 1, 2, \ldots$ such that $\frac{S_n - b_n}{a_n} \xrightarrow{d} G$, where $S_n = X_1 + \ldots + X_n$, $n \geq 1$ with $\{X_i\}$ being a sequence of i.i.d.'s with distribution function $F$.

In view of the above definition, if $G$ has a domain of attraction the following property is not surprising.

**Definition A.2.2** A distribution $R$ is *stable* if for each $n \geq 1$ there exist constants $c_n > 0$, $\gamma_n \in \mathbb{R}$ such that

$$Y_1 + Y_2 + \ldots + Y_n \stackrel{d}{=} c_n Y + \gamma_n$$

where $Y_1, Y_2, \ldots, Y_n$ are i.i.d.'s with $R$ as distribution function, and $Y$ also has $R$ as its distribution function.

It is clear that the normal distribution is stable.

This definition immediately implies that a stable distribution is a limit law.

**Proposition A.2.3** *$R$ is stable $\Leftrightarrow$ for any real numbers $c_1, c_2$ and $Y$, $Y_1, Y_2$ independent random variables with common distribution function $R$, there exist constants $c \neq 0, \gamma$ such that $c_1 Y_1 + c_2 Y_2 \stackrel{d}{=} cY + \gamma$. Thus the class of stable distributions coincides with the class of distributions closed under convolutions and multiplication with real numbers, upto changes of location and scale.* □

The connection between our question and the definitions above is the following

**Theorem A.2.4** *A nondegenerate distribution $G$ is a limit law of partial sums of i.i.ds.'s*
*$\Leftrightarrow$ $G$ is a nondegenerate stable distribution*
*$\Leftrightarrow$ $G$ is nondegenerate and has a domain of attraction.* □

*Note:* Degenerate distributions are excluded because any sequence of partial sums $S_n$ of i.i.d.'s can be made to converge to zero after suitable scaling. Indeed let $\{X_i\}$ be i.i.d.'s. Put $\hat{S}_n = |X_1| + \ldots + |X_n|$, $n \geq 1$. Let $\{a_n\}$ be a sequence of positive numbers such that $P(|\hat{S}_n| > \sqrt{a_n}) = P(\hat{S}_n > \sqrt{a_n}) < \frac{1}{2^n}$, $n \geq 1$, and $a_n \to \infty$. Then it can be shown that $\frac{\hat{S}_n}{a_n} \to 0$ in probability. Therefore $\frac{S_n}{a_n} \to 0$ in probability, and hence in distribution. This is not an interesting situation because the scaling factor $a_n$ is too large to detect any pattern in the bahaviour of $S_n$.

In fact one can say more than the preceding theorem. A major role is played in the analysis by the function

$$V(x) = \int_{[-x,x]} y^2 \, dF(y), \quad x > 0, \tag{A.2.1}$$

where $F$ is the common distribution function of the i.i.d. summands. Also the following property proves critical.

**Definition A.2.5** A positive measurable function $L$ on $(0, \infty)$ is said to be *slowly varying at* $\infty$ if

$$\lim_{x \to \infty} \frac{L(tx)}{L(x)} = 1, \text{ for any } t > 0. \tag{A.2.2}$$

This is a concept introduced in classical analysis by Karamata in the context of Tauberian theorems. See [BGT].

**Theorem A.2.6** *(a) $F$ belongs to the domain of attraction of the normal distribution if and only if $V$ is slowly varying at $\infty$. (This is the case whenever second moment exists.)*

*(b) $F$ belongs to the domain of attraction of a nonnormal stable distribution if and only if*

$$V(x) \sim x^{2-\alpha} L(x), \text{ as } x \to \infty, \tag{A.2.3}$$

*and*

$$\frac{1 - F(x)}{1 - F(x) + F(-x)} \to p, \quad \frac{F(-x)}{1 - F(x) + F(-x)} \to (1 - p), \text{ as } x \to \infty, \tag{A.2.4}$$

*where $\alpha \in (0, 2)$, $L$ is slowly varying, $0 \leq p \leq 1$.*

*(b)′ (equivalent form of (b)): $F$ belongs to the domain of attraction of a nonnormal stable distribution if and only if*

$$1 - F(x) = \frac{c_1 + o(1)}{x^\alpha} \, L(x), \quad F(-x) = \frac{c_2 + o(1)}{x^\alpha} \, L(x), \text{ as } x \to \infty \tag{A.2.5}$$

*where $0 < \alpha < 2$, $L(\cdot)$ is slowly varying at $\infty$, $c_1 \geq 0$, $c_2 \geq 0$ are constants such that $c_1 + c_2 > 0$.*

*Note:* $(1 - F(x))$ is the right tail of the distribution while $F(-x)$ is the left tail.

**Remark A.2.7** As can be discerned from the theorem above, the parameter $\alpha \in (0, 2]$ plays a dominant role in the description of a stable distribution. To give an idea of this, let $X$ be a random variable with distribution function $F$. Suppose the corresponding Fourier transform (or characteristic function) is given by

$$\varphi_X(t) = E[e^{itX}] = \exp(-c|t|^\alpha), \ t \in \mathbb{R} \tag{A.2.6}$$

where $c > 0$, $0 < \alpha \leq 2$ are constants, and $i = \sqrt{-1}$. Then $F$ is a stable distribution. Note that $\alpha = 2$ corresponds to the Gaussian case. In this case the random variable $X$ is symmetric about the origin, and hence is also called the *symmetric $\alpha$-stable distribution.* For $0 < \alpha < 2$ it can be shown that $X$ has finite moments only of order $< \alpha$. A representation for the characteristic function/Fourier transform of a nonnormal stable distribution can be given, with $\alpha \in (0, 2)$ once again determining tail behaviour and existence of moments. See [Fe], [EKM], [L], ... In fact, for studying stable distributions, Fourier transforms seem to be most amenable to analysis. □

Suppose $F$ belongs to the domain of attraction of $G$, where $\mu_G$ is nondegenerate. Then there exist constants $a_n > 0$, $b_n$ such that $\frac{S_n - b_n}{a_n} \xrightarrow{d} G$. How can $a_n > 0$, $b_n$ be chosen?

**Remark A.2.8** (i) If $F$ has finite nonzero variance $\sigma^2$, then we know from classical CLT that $b_n = n\ E(X_1)$, $a_n = \sqrt{n}\ \sigma$.

(ii) Let $G$ denote symmetric $\alpha$-stable distribution. Let $\{X_i\}$ be an i.i.d. sequence distributed according to $G$. Let $S_n = X_1 + \ldots + X_n$, $n \geq 1$. As $X_i$'s are symmetric about 0, it is natural to take $b_n = 0$ as the centering factor. Now $S_n \stackrel{d}{=} a_n X_1$ by definition of stable distribution for some $a_n > 0$. So by (A.2.6), for $t \in \mathbb{R}$

$$\begin{aligned} & \varphi_{S_n}(t) = \varphi_{a_n X_1}(t) = E[e^{ita_n X_1}] \\ = \ & \varphi_{X_1}(a_n t) = \exp(-c \mid a_n t \mid^\alpha) = \exp(-c\ a_n^\alpha \mid t \mid^\alpha). \end{aligned} \tag{A.2.7}$$

As $S_n = X_1 + \ldots + X_n$, again by (A.2.6), for $t \in \mathbb{R}$

$$\varphi_{S_n}(t) = (\varphi_{X_1}(t))^n = \exp(-c\ n \mid t \mid^\alpha). \tag{A.2.8}$$

From (A.2.7), (A.2.8) it is clear that $a_n = n^{1/\alpha}$ is the natural choice.

In general the following choices can be made. If $E\ X_i^2 = \infty$ and $\alpha = 2$, or $\alpha \in (0,2)$ then $a_n = n^{1/\alpha}L(n)$ where $L$ is an appropriate slowly varying function. If $\alpha \in (1,2]$ then $b_n = n\ E(X_1)$. If $\alpha \in (0,1)$ then $b_n = 0$.

The above considerations lead to the following definition and the subsequent theorem.

**Definition A.2.9** We say that $F$ belongs to the *domain of normal attraction* of an $\alpha$-stable distribution $G$, if $F$ belongs to the domain of attraction of $G$ and if the normalization $a_n = c\ n^{1/\alpha}$, where $c > 0$ is a constant, can be chosen.

**Theorem A.2.10** *(a) $F$ belongs to the domain of normal attraction of Gaussian distribution $\Leftrightarrow$ $F$ has finite variance.*

*(b) Let $\alpha \in (0,2)$. Then $F$ belongs to the domain of normal attraction of an $\alpha$-stable distribution $\Leftrightarrow$*

$$F(-x) \sim c_1\ x^{-\alpha},\ (1 - F(x)) \sim c_2\ x^{-\alpha},\ \textit{as}\ x \to \infty \tag{A.2.9}$$

*for constants $c_1 \geq 0,\ c_2 \geq 0$ with $c_1 + c_2 > 0$.*

**Remark A.2.11** Even though it can be proved that stable distributions have probability density functions, closed form expressions are known only when $\alpha = 2$ (Gaussian), $\alpha = 1$ and symmetric (Cauchy), and $\alpha = \frac{1}{2}$ and one sided (inverse Gaussian). This could be one reason why stable laws might not have received the attention they deserve. On the other hand for $F$ to belong to the domain of attraction of an $\alpha$-stable distribution, as seen from the above discussion, the tail (especially the right tail in the actuarial context) should be like a power law modified by a slowly varying function. In fact for $F$ to belong to the domain of normal attraction of a stable distribution, by (A.2.9), tail behaviour should be a Pareto-like power law.

Thus there are very good probabilistic reasons, not just aesthetic ones, why power laws (especially Pareto distributions) form an important class among the heavy tailed distributions.

For proofs of the various results discussed above, extensions/ generalizations, and for more insight and applications, one may see [Fe], [EKM], [L], [ST], etc.

## A.3 Martingales

Martingale theory has become a very powerful tool in recent decades. We shall, however, recall only those aspects having a direct bearing for our purposes. For more information one may look up Chapter 1 of [KS], [V1], [V2], [CT], [BW] and the references therein.

We first need the notion of conditional expectation. Let $(\Omega, \mathcal{F}, P)$ be a probability space and $\mathcal{A} \subseteq \mathcal{F}$ a sub $\sigma$-algebra. Let $X$ be an integrable random variable. The *conditional expectation of $X$ given $\mathcal{A}$* is defined to be the $\mathcal{A}$-measurable random variable $Y$ such that

$$\int_A X(\omega)\, dP(\omega) = \int_A Y(\omega)\, dP(\omega), \text{ for all } A \in \mathcal{A}. \tag{A.3.1}$$

It is uniquely determined up to sets of measure 0 and is denoted $E(X \mid \mathcal{A})$. Existence of conditional expectation is an important consequence of the Radon-Nikodym theorem. Important properties of conditional expectation are summarised in the following proposition; all the random variables are assumed to be integrable; $\mathcal{A}, \mathcal{A}_1, \mathcal{A}_2$ are sub $\sigma$-algebras of $\mathcal{F}$.

**Proposition A.3.1** *(i) $E(X \mid \mathcal{A})$ is integrable and $E(E(X \mid \mathcal{A})) = E(X)$.*
*(ii) $X \geq 0$ a.s. $\Rightarrow E(X \mid \mathcal{A}) \geq 0$ a.s.. Hence $X_1 \leq X_2$ a.s. $\Rightarrow E(X_1 \mid \mathcal{A}) \leq E(X_2 \mid \mathcal{A})$ a.s.*
*(iii) $E(a_1X_1 + a_2X_2 \mid \mathcal{A}) = a_1E(X_1 \mid \mathcal{A}) + a_2E(X_2 \mid \mathcal{A})$ a.s. where $a_1, a_2 \in \mathbb{R}$; so conditional expectation is a linear operation.*
*(iv) If $X$ is $\mathcal{A}$-measurable then $E(X \mid \mathcal{A}) = X$ a.s.*
*(v) If $Y$ is a bounded $\mathcal{A}$-measurable random variable then $E(YX \mid \mathcal{A}) = YE(X \mid \mathcal{A})$ a.s.*
*(vi) If $X_n \to X$ in probability and $|X_n| \leq U$ a.s. for all $n$, where $U$ is integrable then*

$$E(X \mid \mathcal{A}) = E(\lim_n X_n \mid \mathcal{A}) = \lim_{n\to\infty} E(X_n \mid \mathcal{A}) \text{ a.s.}$$

*Thus dominated convergence theorem holds for conditional expectation.*
*(vii) If $\mathcal{A}_1 \subset \mathcal{A}_2$ then $E(E(X \mid \mathcal{A}_1) \mid \mathcal{A}_2) = E(X \mid \mathcal{A}_1) = E(E(X \mid \mathcal{A}_2) \mid \mathcal{A}_1)$ a.s.*
*(viii) (Jensen's inequality): If $\varphi$ is a continuous convex function on $\mathbb{R}$ then $E(\varphi(X) \mid \mathcal{A}) \geq \varphi(E(X \mid \mathcal{A}))$ a.s.* □

**Note:** If $X \in L^2(\Omega, \mathcal{F}, P)$, that is, if $X$ is square integrable then $E(X \mid \mathcal{A})$ turns out to be the projection of $X$ onto the closed subspace $L^2(\Omega, \mathcal{A}, P)$.

**Note:** Let $G \in \mathcal{F}$, $\mathcal{A} \subseteq \mathcal{F}$. The *conditional probability of $G$ given $\mathcal{A}$* is defined by $P(G \mid \mathcal{A}) \triangleq E[I_G \mid \mathcal{A}]$. (Note that $P(G \mid \mathcal{A})$ may not be defined on a set of measure zero, which may depend on $G$. In general these null sets can accumulate and blow up as $G$ varies. However, in our contexts such a problem will not arise.) In particular, if $X$ is a random variable and $B \subseteq \mathbb{R}$ is a Borel set then $P(X \in B \mid \mathcal{A}) \triangleq E[I_B(X) \mid \mathcal{A}] = E[I_{X^{-1}(B)} \mid \mathcal{A}]$.

Discrete parameter martingales will be taken up first. Let $\{\mathcal{F}_n : n = 1, 2, \ldots\}$ be a sequence of sub $\sigma$-algebras such that $\mathcal{F}_n \subseteq \mathcal{F}_{n+1} \subseteq \mathcal{F}$ for all $n \geq 1$. Let $\{X_n : n \geq 1\}$ be a sequence of random variables satisfying the following:

(i) $X_n$ is integrable, and measurable with respect to $\mathcal{F}_n$ for $n = 1, 2, \ldots$ (i.e., $\{X_n\}$ is $\{\mathcal{F}_n\}$-adapted and integrable);

(ii) for each $n \geq 1$, $E(X_{n+1} \mid \mathcal{F}_n) = X_n$ a.s. Then we say $\{X_n\}$ is a *discrete parameter martingale* w.r.t. $\{\mathcal{F}_n\}$, or simply $\{(X_n, \mathcal{F}_n) : n \geq 1\}$ is a martingale.

If the equality in (ii) above is replaced by $E(X_{n+1} \mid \mathcal{F}_n) \leq X_n$ a.s. for $n = 1, 2, \ldots$ then $\{(X_n, \mathcal{F}_n) : n \geq 1\}$ is called a *supermartingale.* Likewise, if in (ii) we have $E(X_{n+1} \mid \mathcal{F}_n) \geq X_n$ a.s. then $\{(X_n, \mathcal{F}_n) : n \geq 1)$ is a *submartingale.*

**Examples:** 1) Let $\{Y_i\}$ be a sequence of independent random variables with mean 0. Set $S_n = Y_1 + \ldots + Y_n$, $\mathcal{F}_n = \sigma\{Y_1, \ldots, Y_n\}$, $n \geq 1$. Then $\{(S_n, \mathcal{F}_n) : n \geq 1\}$ is a martingale.

2) Let $X$ be an integrable random variable. Let $\mathcal{F}_n \subseteq \mathcal{F}_{n+1} \subseteq \mathcal{F}$, $n \geq 1$ be an increasing sequence of sub $\sigma$-algebras. Define $X_n = E(X \mid \mathcal{F}_n)$, $n \geq 1$. Then $\{(X_n, \mathcal{F}_n) : n \geq 1\}$ is a martingale.

An immediate consequence of Jensen's inequality for conditional expectations is

**Lemma A.3.2** *If $\{(X_n, \mathcal{F}_n) : n \geq 1\}$ is a martingale and $\varphi$ is a convex function such that $\varphi(X_n)$ is integrable for each $n$, then $\{\varphi(X_n) : n \geq 1\}$ is a submartingale w.r.t. $\{\mathcal{F}_n\}$; in particular, for $p \geq 1$, $\{|X_n|^p : n \geq 1\}$ is a submartingale provided $E(|X_n|^p) < \infty$.*

□

Most of the results we present here are due to Doob. We write $E[Y : C] \triangleq \int_C Y \, dP$.

**Theorem A.3.3** *(Doob's inequalities) Let* $\{X_n : n \geq 1\}$ *be a martingale w.r.t.* $\{\mathcal{F}_n\}$. *Let* $M_n(\omega) = \sup_{1 \leq i \leq n} |X_i(\omega)|$. *Then*

*(i) for any* $\lambda > 0$

$$\begin{aligned} P(M_n \geq \lambda) &\leq \frac{1}{\lambda} E[|X_n| : \{|X_n| \geq \lambda\}] \\ &\leq \frac{1}{\lambda} E[|X_n|]; \qquad \textit{(A.3.2)} \end{aligned}$$

*(ii) for any* $p > 1$

$$E(M_n^p) \leq \left(\frac{p}{p-1}\right)^p E(|X_n|^p). \qquad \text{(A.3.3)}$$

□

If $\{X_n\}$ is a martingale, we know that $\{|X_n|\}$ is a submartingale. So the above theorem is basically about nonnegative submartingales.

In the context of martingales, an important role is played by stopping times. Let $\{\mathcal{F}_n : n \geq 0\}$ be an increasing sequence of sub $\sigma$-algebras. A measurable function $\tau : \Omega \to \{0, 1, 2, \ldots\} \cup \{\infty\}$ is said to be a *stopping time* w.r.t. $\{\mathcal{F}_n\}$ if $\{\omega : \tau(\omega) \leq n\} \in \mathcal{F}_n$ for each $n \geq 0$. For a stopping time $\tau$ w.r.t. $\{\mathcal{F}_n\}$ define

$$\mathcal{F}_\tau = \{A : A \in \mathcal{F} \text{ and } A \cap \{\omega : \tau(\omega) \leq n\} \in \mathcal{F}_n \text{ for all } n\}.$$

It can be shown that $\mathcal{F}_\tau$ is a sub $\sigma$-algebra, and that $\tau$ is $\mathcal{F}_\tau$-measurable. $\mathcal{F}_\tau$ can be thought of as the repository of all information up to time $\tau$. If $\tau_1, \tau_2$ are two stopping times w.r.t. $\{\mathcal{F}_n\}$ such that $0 \leq \tau_1 \leq \tau_2$ then $\mathcal{F}_{\tau_1} \subseteq \mathcal{F}_{\tau_2}$. Suppose $\{Y_n\}$ is a sequence of random variables such that $Y_n$ is $\mathcal{F}_n$-measurable for each $n$. If $\tau$ is a stopping time, on the set $\{\omega : \tau(\omega) < \infty\}$, define $Y_\tau(\omega) \triangleq Y_{\tau(\omega)}(\omega)$. It can be shown that $Y_\tau$ is $\mathcal{F}_\tau$-measurable.

A very useful theorem is

**Theorem A.3.4** *(Optional Stopping theorem or Optional Sampling theorem): Let* $\{X_n\}$ *be a martingale w.r.t.* $\{\mathcal{F}_n\}$; *let* $\tau_1, \tau_2$ *be two stopping times w.r.t.* $\{\mathcal{F}_n\}$ *such that* $0 \leq \tau_1 \leq \tau_2 \leq C$ *where* $C$ *is a constant. Then*

$$E[X_{\tau_2} \mid \mathcal{F}_{\tau_1}] = X_{\tau_1} \ a.s. \qquad \text{(A.3.4)}$$

□

**Remark A.3.5** Let $\{X_n\}$ be a martingale and $\tau$ a stopping time, both w.r.t. $\{\mathcal{F}_n\}$; $\tau$ need not be bounded. For any $k$, $\tau_k \triangleq \min\{\tau, k\}$ is a bounded stopping time w.r.t. $\{\mathcal{F}_n\}$. So by the optional stopping theorem $E(X_{\tau_k}) = E(X_0)$ for any fixed $k$. If $\tau < \infty$ a.s., then $\tau_k \uparrow \tau$ and $X_{\tau_k} \to X_\tau$ as $k \uparrow \infty$ (because $\tau$ takes value in the discrete set $\{0, 1, 2, \ldots\}$). In addition if the dominated or monotone convergence theorem are applicable to $\{X_{\tau_k}\}$ then we get $E(X_\tau) = E(X_0)$. This is one way the above theorem is used.

**Remark A.3.6** Suppose $\{X_n\}$ is a submartingale in Theorem A.3.4. With $\tau_1, \tau_2$ as in the theorem, the analogue of (A.3.4) is then obtained as $E[X_{\tau_2} \mid \mathcal{F}_{\tau_1}] \geq X_{\tau_1}$ a.s. □

Now we take a brief look at continuous parameter martingales and associated stopping times.

Let $(\Omega, \mathcal{F}, P)$ be a probability space. By a *filtration* we mean a family $\{\mathcal{F}_t : t \geq 0\}$ of sub $\sigma$-algebras such that $\mathcal{F}_s \subseteq \mathcal{F}_t \subseteq \mathcal{F}$ whenever $0 \leq s \leq t$. To avoid certain technical problems we shall make the following assumptions:
(i) the filtration $\{\mathcal{F}_t\}$ is right continuous, that is, $\mathcal{F}_t = \mathcal{F}_{t+} \triangleq \bigcap_{\epsilon>0} \mathcal{F}_{t+\epsilon}$ for any $t \geq 0$;
(ii) $\mathcal{F}_0$ (and hence any $\mathcal{F}_t$) contains all $P$-null sets.

Let $\{\mathcal{F}_t : t \geq 0\}$ be a filtration satisfying the above assumptions. A process $\{X(t) : t \geq 0\}$ is called a *martingale w.r.t* $\{\mathcal{F}_t\}$ if the following holds:
a) for almost all $\omega$, $t \mapsto X(t, \omega)$ is right continuous and has left limits; that is, the process $X$ has r.c.l.l. sample paths;
b) for each $t \geq 0$, $X(t)$ is $\mathcal{F}_t$-measurable and integrable; (i.e. $\{X(t)\}$ is $\{\mathcal{F}_t\}$-adapted and integrable);
c) for $0 \leq s \leq t$, $E(X(t) \mid \mathcal{F}_s) = X(s)$ a.s.

If the above hold, we also say that $\{(X(t), \mathcal{F}_t) : t \geq 0\}$ is a *continuous parameter martingale*. As in the discrete parameter case, by replacing the equality in (c) by appropriate inequality one can define continuous parameter submartingale/supermartingale.

Analogues of Lemma A.3.2 and Theorem A.3.3 hold for continuous parameter martingales.

Let $\{\mathcal{F}_t : t \geq 0\}$ be a filtration as above. A random variable $\tau : \Omega \to [0, \infty) \cup \{+\infty\}$ is called a *stopping time w.r.t.* $\{\mathcal{F}_t\}$ if $\{\omega : \tau(\omega) \leq t\} \in \mathcal{F}_t$ for all $t \geq 0$. As in the discrete parameter case, for a stopping time $\tau$

set

$$\mathcal{F}_\tau = \{A : A \in \mathcal{F} \text{ and } A \cap \{\tau \leq t\} \in \mathcal{F}_t \text{ for all } t\}.$$

As before the $\sigma$-algebra $\mathcal{F}_\tau$ is the repository of all information up to time $\tau$.

In the continuous parameter case, a stopping time can take a continuum of values, and hence greater care is needed. Useful examples of stopping times are the first time a process enters / hits an open / a closed set. Technical assumptions (i), (ii) above are made to ensure that these are stopping times w.r.t. $\{\mathcal{F}_t\}$; see [KS].

For an $\{\mathcal{F}_t\}$-adapted process $\{X(t)\}$ and a stopping time $\tau$, define the random variable $X_\tau$ on the set $\{\omega : \tau(\omega) < \infty\}$ by $X_\tau(\omega) = X(\tau(\omega), \omega)$. Then $X_\tau$ is $\mathcal{F}_\tau$-measurable.

**Theorem A.3.7** *Let $\{X(t) : t \geq 0\}$ be a submartingale w.r.t. $\{\mathcal{F}_t\}$. Let $\tau_1, \tau_2$ be two stopping times w.r.t. $\{\mathcal{F}_t\}$ such that $0 \leq \tau_1 \leq \tau_2 \leq C$, where $C$ is a constant. Then*

$$E(X_{\tau_2} \mid \mathcal{F}_{\tau_1}) \geq X_{\tau_1} \text{ a.s.} \tag{A.3.5}$$

*If $\{X(t)\}$ is a martingale, then equality holds in (A.3.5).*

**Corollary A.3.8** *Let $\{X(t)\}$ be a martingale w.r.t. $\{\mathcal{F}_t\}$ and $\tau$ a stopping time w.r.t. $\{\mathcal{F}_t\}$. Then $\{X_{\tau \wedge t} : t \geq 0\}$ is also a martingale w.r.t. $\{\mathcal{F}_t\}$.* □

## A.4 Brownian motion and Ito integrals

Brownian motion is perhaps the most important stochastic process; and Ito integral is a useful tool in stochastic analysis. We shall outline only the one-dimensional case, which would suffice for our purposes. For details, extensions and more information and insight see [BW], [IW], [KS], [D], [V2].

The following definition has been abstracted from the study of the physical phenomenon of Brownian motion. Let $(\Omega, \mathcal{F}, P)$ be a probability space with a filtration $\{\mathcal{F}_t : t \geq 0\}$. Let $\{B(t) : t \geq 0\}$ be a one-dimensional stochastic process satisfying the following:

1. For each fixed $t \geq 0$, $B(t)$ is a real valued random variable which is $\mathcal{F}_t$-measurable.

2. For a.e. $\omega \in \Omega$, $t \mapsto B(t,\omega)$ is continuous on $[0,\infty)$.

3. For $0 \leq s \leq t$, the random variable $B(t) - B(s)$ and the sub $\sigma$-algebra $\mathcal{F}_s$ are independent; that is, $B(t) - B(s)$ and any $\mathcal{F}_s$-measurable random variable are independent. In particular $\{B(t) : t \geq 0\}$ is a process with independent increments.

4. For $0 \leq s \leq t$ the random variable $B(t) - B(s)$ has Gaussian (normal) distribution with mean 0 and variance $(t-s)$; i.e., $B(t) - B(s) \sim N(0, t-s)$. Then $\{B(t) : t \geq 0\}$ is called an $\{\mathcal{F}_t\}$-*adapted one-dimensional Brownian motion.* If a process $\{B_0(t)\}$ satisfies the above and $B_0(0) \equiv 0$ a.s. then $B_0$ is called *standard Brownian motion.* With $B_0$ as above, $x \in \mathbb{R}$, the process $B(t) = x + B_0(t)$, $t \geq 0$ is the one-dimensional Brownian motion starting at $x$.

Define for $0 \leq s < t$, $x, z \in \mathbb{R}$

$$p(s,x;t,z) = \frac{1}{\sqrt{2\pi(t-s)}} \exp\left\{-\frac{1}{2(t-s)}(z-x)^2\right\}. \tag{A.4.1}$$

Note that $p(s,x;t,z) = p(s+h,x;t+h,z) = p(0,x;t-s,z)$ indicating time homogeneity. An important fact is that the Brownian motion is a Markov process so that for $0 \leq s < t$, Borel set $A \subseteq \mathbb{R}$,

$$P(B(t) \in A \mid \mathcal{F}_s) = P(B(t) \in A \mid \sigma\{B(s)\}) = \int_A p(s, B(s); t, z)dz, \tag{A.4.2}$$

This is an easy consequence of the independent increments property. Clearly with $0 \leq s < t$, $x \in \mathbb{R}$ fixed, $z \mapsto p(0,x;t-s,z)$ is the probability density function of $B(t-s)$ given $B(0) = x$; this is also the p.d.f. of $B(t)$ conditional on $B(s) = x$. Note that (A.4.2) holds even if $B(0)$ is random variable, of course, measurable w.r.t. $\mathcal{F}_0$. It is not difficult to prove that the distribution of $B(0)$, called the initial distribution, and the function $p$ given by (A.4.1) determine the distribution of the process $\{B(t) : t \geq 0\}$. The distribution of the Brownian motion process is a probability measure on the space $C[0,\infty)$ of continuous functions on $[0,\infty)$; this probability measure is referred to as *Wiener measure.*

In fact a stronger version of (A.4.2), called the *strong Markov property* holds. Let $\tau$ be a finite stopping time w.r.t. $\{\mathcal{F}_t\}$. Then the stochastic process $\{\tilde{B}(t) = B(\tau+t) - B(\tau) : t \geq 0\}$ is again distributed

like a Brownian motion independent of $\mathcal{F}_\tau$. In particular for any finite stopping time $\tau$ w.r.t. $\{\mathcal{F}_t\}$, any $s \geq 0$ and Borel set $A \subset \mathbb{R}$,

$$\begin{aligned} P(B(\tau+s) \in A \mid \mathcal{F}_\tau) &= P(B(\tau+s) \in A \mid \sigma\{\tau, B(\tau)\}) \\ &= \int_A p(0, B(\tau); s, z)dz. \end{aligned} \tag{A.4.3}$$

Another basic property of Brownian motion is that $\{B(t) : t \geq 0\}$ and $\{B^2(t) - t : t \geq 0\}$ are both martingales w.r.t. $\{\mathcal{F}_t\}$; this is easy to prove.

Since $p$ given by (A.4.1) is smooth in its arguments, simple differentiation leads to

$$\frac{\partial p}{\partial s} + \frac{1}{2}\frac{\partial^2 p}{\partial x^2} = 0, \ 0 < s < t, \ x \in \mathbb{R} \tag{A.4.4}$$

for fixed $t > 0$, $z \in \mathbb{R}$, and

$$-\frac{\partial p}{\partial t} + \frac{1}{2}\frac{\partial^2 p}{\partial z^2} = 0, \ s < t, \ z \in \mathbb{R} \tag{A.4.5}$$

for fixed $s \geq 0$, $x \in \mathbb{R}$. These are respectively the *backward* and the *forward heat equation* or *diffusion equation.* In probability literature the function $p$ is called the *transition probability density function* of the Brownian motion, because distributional properties of the process can be got from $p$ thanks to (A.4.2), (A.4.3). Because of (A.4.4), (A.4.5) and since solutions to initial value problems of heat equation can be expressed in terms of the kernel $p$, in PDE literature the function $p$ is referred to as the *fundamental solution to the heat equation.* This forms a basis for the fruitful interaction between probability theory and partial differential equations. See [KS], [D], [V2].

Brownian motion is the prototype of a diffusion process. This led Paul Levy to suggest that a diffusion process could perhaps be expressed in terms of an integral w.r.t. Brownian motion together with a time integral. However, sample paths of Brownian motion are of infinite variation over any bounded interval; so the technology of Stieltjes integration is ruled out. This problem was elegantly solved by Kiyosi Ito. We now briefly describe Ito integrals and some of its ramifications.

Fix $T > 0$. Let $f : [0, T] \times \Omega \to \mathbb{R}$ be a *simple* function in the sense that there exist $0 = t_0 < t_1 < \ldots < t_{n-1} < t_n = T$, and bounded random variables $\alpha_i$, $i = 0, 1, \ldots, n-1$ such that $\alpha_i$ is $\mathcal{F}_{t_i}$-measurable

and $f(t,\omega) = \alpha_i(\omega)$ if $t_i \le t < t_{i+1}$, $i = 0, 1, \ldots, n-1$. In this case the natural definition of integral w.r.t. Brownian motion is

$$I_t(f)(\omega) \triangleq \int_0^t f(s,\omega)\, dB(s,\omega)$$

$$\triangleq \sum_{i=0}^{n-1} \alpha_i(\omega)[B(t_{i+1} \wedge t, \omega) - B(t_i \wedge t, \omega)],\ 0 \le t \le T. \quad \text{(A.4.6)}$$

Note that for each $i$, $\alpha_i$ and $[B(t_{i+1} \wedge t) - B(t_i \wedge t)]$ are independent. Consequently it is easily proved that $\{I_t(f) : t \in [0,T]\}$ is a martingale w.r.t. $\{\mathcal{F}_t : t \in [0,T]\}$. With a little effort one can prove that

$$E\left[\left(\int_s^t f(r)\, dB(r)\right)^2 \middle| \mathcal{F}_s\right] = E\left[\int_s^t f(r)^2 dr \middle| \mathcal{F}_s\right] \quad \text{(A.4.7)}$$

for $0 \le s \le t \le T$.

Next let $g : [0,T] \times \Omega \to \mathbb{R}$ be a measurable function satisfying
(i) for each $t \in [0,T]$, the map $(s,\omega) \mapsto g(s,\omega)$, $0 \le s \le t$, $\omega \in \Omega$ is jointly measurable as a map from $([0,t] \times \Omega, \mathcal{B}[0,t] \times \mathcal{F}_t)$ to the real line; (that is, $g$ is *progressively measurable*);

(ii) $E\left(\int_0^T |f(s)|^2 ds\right) < \infty$.

(Note that any simple function has the above properties). One can find simple functions $f_n$, $n = 1, 2, \ldots$ such that $E\int_0^T |f_n(s) - g(s)|^2 ds \to 0$. Using Doob's inequality for the martingales $I_t(f_n)$ and (A.4.7) one can now show that

$$\lim_{m,n\to\infty} E\left[\sup_{0\le t\le T} |I_t(f_m) - I_t(f_n)|^2\right] = 0.$$

Now define $I_t(g) \equiv \int_0^t g(s)\, dB(s) = \lim_{n\to\infty} I_t(f_n)$ where limit is taken in $L^2(P)$. It can be shown from the above that

$$\lim_{n\to\infty} E\left[\sup_{0\le t\le T} |I_t(f_n) - I_t(g)|^2\right] = 0.$$

The process $I_t(g) = \int_0^t g(s)\, dB(s),\ t \geq 0$ is called the *Ito integral* of $g$; it is clearly linear in $g$. The following properties of the process $\{I_t(g) : t \geq 0\}$ are not difficult to establish

i) progressive measurability w.r.t. $\{\mathcal{F}_t\}$, and continuity of sample paths with probability one;

ii) $\{I_t(g) : t \geq 0\}$ is a martingale with respect to $\{\mathcal{F}_t : t \geq 0\}$

iii) $\left\{(I_t(g))^2 - \int_0^t |g(s)|^2\, ds : t \geq 0\right\}$ is also a martingale w.r.t. $\{\mathcal{F}_t\}$.

With more effort one can define Ito integral $\int_0^t g(s)\, dB(s)$ for progressively measurable functions $g$ satisfying

$$P\left\{\omega : \int_0^t |g(s,\omega)|^2\, ds < \infty\ \forall t\right\} = 1.$$

Importance of Ito integrals is mainly due to the following chain rule, known as *Ito's formula*. Let $u(\cdot)$, $v(\cdot)$ be progressively measurable processes such that

$$P\left\{\omega : \int_0^t |u(s,\omega)|ds < \infty, \int_0^t |v(s,\omega)|^2 ds < \infty,\ \forall\ 0 \leq t \leq T\right\} = 1.$$

Let

$$Y(t) = Y(0) + \int_0^t u(s)\, ds + \int_0^t v(s)\, dB(s),\ 0 \leq t \leq T \qquad \text{(A.4.8)}$$

where $Y(0)$ is $\mathcal{F}_0$-measurable. Let $f : [0,T] \times \mathbb{R} \to \mathbb{R}$ be a continuous function such that $f_0 \triangleq \frac{\partial f}{\partial t}$, $f_1 \triangleq \frac{\partial f}{\partial x}$, $f_{11} \triangleq \frac{\partial^2 f}{\partial x^2}$ are also continuous on $[0,T] \times \mathbb{R}$. Then with probability one, for $0 \leq t \leq T$

$$\begin{aligned} & f(t,Y(t)) - f(0,Y(0)) \\ = & \int_0^t f_0(s,Y(s))\, ds + \int_0^t f_1(s,Y(s))\, u(s)\, ds \\ & + \int_0^t f_1(s,Y(s))v(s)dB(s) + \frac{1}{2}\int_0^t f_{11}(s,Y(s))v^2(s)ds. \end{aligned} \qquad \text{(A.4.9)}$$

Presence of the "second order term" (the last term) on the r.h.s. of (A.4.9) is the special feature of Ito's calculus. For a proof of Ito's formula (A.4.9) and extensions, see [KS], [V2], [O], etc.

Now that Ito integrals are at our disposal we can talk about stochastic integral equations. Let $\sigma : [0,T] \times \mathbb{R} \to \mathbb{R}$, $b : [0,T] \times \mathbb{R} \to \mathbb{R}$ be functions, and $X(0)$ an $\mathcal{F}_0$-measurable random variable. Consider the *stochastic integral equation*

$$X(t) = X(0) + \int_0^t \sigma(s, X(s))\, dB(s) + \int_0^t b(s, X(s))\, ds, 0 \le t \le T. \tag{A.4.10}$$

By a *strong solution* to (A.4.10) we mean a continuous process $\{X(t) : t \in [0,T]\}$ progressively measurable w.r.t. $\{\mathcal{F}_t\}$ such that the integrals in (A.4.10) make sense and the equation (A.4.10) holds with probability one. The above equation is formally written in differential notation as

$$dX(t) = \sigma(t, X(t))\, dB(t) + (b(t, X(t))\, dt \tag{A.4.11}$$

and is referred to as a *stochastic differential equation.* As in the case of ordinary differential equations, Picard's iteration can be applied. To be more precise, assume that $\sigma$, $b$ satisfy a uniform Lipschitz condition and a linear growth condition in $x$; that is, there is a constant $K > 0$ such that for all $0 \le t \le T$, $x, y \in \mathbb{R}$

$$\begin{aligned} |\sigma(t,x) - \sigma(t,y)| + |b(t,x) - b(t,y)| &\le K|x-y|, \\ |\sigma(t,0)| + |b(t,0)| &\le K. \end{aligned}$$

Assume also that $X(0)$ is square integrable. Then it can be shown that there is a unique strong solution $\{X(t)\}$ to (A.4.10); moreover $E\int_0^T |X(s)|^2 ds < \infty$. Furthermore it can be shown that $\{X(t) : t \ge 0\}$ is a strong Markov process. In particular, for any finite stopping time $\tau$ w.r.t. $\{\mathcal{F}_t\}$, $s \ge 0$ and $A \subseteq \mathbb{R}$

$$P(X(\tau + s) \in A \mid \mathcal{F}_\tau) = P(X(\tau + s) \in A \mid \sigma\{\tau, X(\tau)\}).$$

Therefore, the process $\{X(t) : t \ge 0\}$ is also called the *one-dimensional (continuous) diffusion process with drift $b$ and dispersion $\sigma$.*

Finally, note that Ito's formula is applicable to the diffusion $X(\cdot)$ as it is in the form (A.4.8). So for any $f \in C^{1,2}([0,\infty) \times \mathbb{R})$ we have

$$f(t, X(t)) - f(0, X(0)) = \int_0^t f_1(s, X(s))\ \sigma(s, X(s))\ dB(s)$$

$$+ \int_0^t \{f_0(s, X(s)) + Lf(s, X(s))\}\ ds \qquad \text{(A.4.12)}$$

where

$$Lf(s,x) = \frac{1}{2}\ \sigma^2(s,x)\ \frac{\partial^2 f}{\partial x^2}(s,x) + b(s,x)\ \frac{\partial f}{\partial x}(s,x). \qquad \text{(A.4.13)}$$

If $\sigma \equiv 1$, $b \equiv 0$, observe that $f_0 + Lf = \left(\frac{\partial}{\partial s} + L\right) f = \left(\frac{\partial}{\partial s} + \frac{1}{2}\frac{\partial^2}{\partial x^2}\right) f$ is just the backward heat operator encountered in (A.4.4). If $f \in C_b^{1,2}([0,\infty) \times \mathbb{R})$ and $\sigma$, $b$ are also bounded, then from (A.4.12) we obtain

$$f(t, X(t)) - f(0, X(0)) - \int_0^t \{f_0(s, X(s)) + Lf(s, X(s))\} ds$$

$$= \text{ a continuous martingale w.r.t. } \{\mathcal{F}_t\}.$$

# Bibliography


[Aa] K.K. Aase: Perspectives of risk sharing. *Scandinavian Actuarial Journal* (2002) 73-128.

[AB] H. Albrecher and O.J. Boxma: On the discounted penalty function in a Markov-dependent risk model. *Insurance: Mathematics and Economics* 37 (2005) 650-672.

[As] S. Asmussen: *Ruin Probabilties.* World Scientific, Singapore, 2000.

[BR] S. Balaji and S. Ramasubramanian: Recurrence and transience of diffusions in a half space. *Bernoulli* 3 (1997) 97-119.

[BG] N. Bauerle and R. Grubel: Multivariate risk processes with interacting intensities. *Advances in Applied Probability* 40 (2008) 578-601.

[BW] R.N. Bhattacharya and E.C. Waymire: *Stochastic Processes with Applications.* Wiley, New York, 1990.

[Bl1] P. Billingsley: *Convergence of Probability Measures.* Wiley, New York, 1968.

[Bl2] P. Billingsley: *Probability and Measure.* 3rd edition. Wiley, New York, 1995.

[BGT] N.H. Bingham, C.M. Goldie and J.L. Teugels: *Regular Variation.* Cambridge University Press, Cambridge, 1987.

[Bd] P.J. Boland: *Statistical and Probabilistic Methods in Actuarial Science.* Chapman & Hall/CRC, Boca Raton, 2007.

[BCHJKPR] P. Booth, R. Chadburn, S. Haberman, D. James, Z. Khorasanee, R.H. Plumb, and B. Rickayzen: *Modern Actuarial Theory and Practice.* Chapman & Hall/CRC, Boca Raton, 2005.

[Bc] K. Borch: The Theory of Risk. *Journal of Royal Statistical Society, Series B* 29 (1967) 432-467.

[Bo] V.S. Borkar: *Probability Theory: an advanced course* Springer-Verlag, New York, 1995.

[CT] Y.S. Chow and H. Teicher: *Probability Theory: Independence, Interchangeability, Martingales.* 3rd edition. Springer (India), New Delhi, 2004.

[CL] E.A. Coddington and N. Levinson: *Theory of Ordinary Differential Equations.* Tata McGraw-Hill, New Delhi, 1985.

[DDS] D. Denisov, A.B. Dieker and V. Shneer: Large deviations for random walks under subexponentiality: the big jump domain. *Annals of Probability* 36 (2008) 1946-1991.

[Du] R. Durrett: *Stochastic Calculus: a practical introduction.* CRC Press, Boca Raton, 1996.

[Dy] E.B. Dynkin: *Markov Processes*, Vols. I & II. Springer-Verlag, New York, 1965.

[EKM] P. Embrechts, C. Kluppelberg and T. Mikosch: *Modelling Extremal Events for Insurance and Finance.* Springer, Heildelberg, 1997.

[EK] S. Ethier and T. Kurtz: *Markov Processes: Characterization and Convergence.* Wiley, New York, 1986.

[Fe] W. Feller: *An Introduction to Probability Theory and its applications*, Vol.II. Wiley-Eastern, New Delhi, 1969.

[H] C. Hipp: Stochastic Control with Application in Insurance. In *Stochastic Methods in Finance*, pp. 127-164. Lecture Notes in Mathematics, No.1856, Springer, Berlin, 2004.

[HP1] C. Hipp and M. Plum: Optional investment for insurers. *Insurance: Mathematics and Economics* 27 (2000) 215-228.

[HP2] C. Hipp and M. Plum: Optimal investment for investors with state dependent income, and for insurers. *Finance and Stochastics* 7 (2003) 299-321.

[IW] N. Ikeda and S. Watanabe: *Stochastic Differential Equations and Diffusion Processes.* North Holland, Amsterdam, 1981.

[KGDD] R. Kaas, M. Goovaerts, J. Dhaene and M. Denuit: *Modern Actuarial Risk Theory.* Kluwer, Boston, 2001.

[KS] I. Karatzas and S. Shreve: *Brownian Motion and Stochastic Calculus.* Springer, New York, 1991.

[KT] S. Karlin and H. Taylor: *A First Course in Stochastic Processes.* Academic Press, New York, 1975.

[L] E. Lukacs: *Characteristic Functions.* Hafner, New York, 1970.

[MFE] A.J. McNeil, R. Frey and P. Embrechts: *Quantitative Risk Management. Concepts, Techniques and Tools.* Princeton University Press, Princeton, 2005.

[M] T. Mikosch: *Non-Life Insurance Mathematics: an introduction with Stochastic Processes.* Springer (India), New Delhi, 2004.

[MN] T. Mikosch and A.V. Nagaev: Large deviations of heavy tailed sums with applications in insurance. *Extremes* 1 (1998) 81-110.

[O] B. Oksendal: *Stochastic Differential Equations: an introduction with applications.* Springer (India), New Delhi, 2004.

[P1] K.R. Parthasarathy: *Probability Measures on Metric Spaces.* Academic Press, New York, 1967.

[P2] K.R. Parthasarathy: *Introduction to Probability and Measure.* Hindustan Book Agency, New Delhi, 2005.

[Ra] S. Ramasubramanian: An insurance network: Nash equilibrium. *Insurance: Mathematics and Economics* 38 (2006) 374-390.

[Re] S.I. Resnick: *Adventures in Stochastic Processes*, Birkhauser, Boston, 1992.

[RSST] T. Rolski, H. Schmidli, V. Schmidt and J.L. Teugels: *Stochastic Processes for Insurance and Finance.* Wiley, Chichester, 1999.

[Ro1] S.M. Ross: *Introduction to Probability Models.* 8th edition. Elsevier India, New Delhi, 2005.

[Ro2] S.M. Ross: *Stochastic Processes.* 2nd edition. Wiley (Asia), Singapore, 2004.

[Ru] W. Rudin: *Functional Analysis.* Tata McGraw-Hill, New Delhi, 1974.

[ST] G. Samorodnitsky and M.S. Taqqu: *Stable Non-Gaussian Random Processes. Stochastic Models with Infinite Varience.* Chapman and Hall, London, 1994.

[SV1] D.W. Stroock and S.R.S. Varadhan: On the support of diffusion processes with applications to the strong maximum principle. *Proc. Sixth Berkeley Symp. Math. Statist. Prob. III*; pp. 333-359. Univ. California Press, Berkeley, 1972.

[SV2] D.W. Stroock and S.R.S. Varadhan: *Multidimensional Diffusion Processes.* Springer, New York, 1979.

[T] M. Taksar: Optimal risk and dividend distribution control models for an insurance company. *Mathematical Methods of Operations Research* 51 (2000) 1-42.

[TS] J.L. Teugels and B. Sundt (eds.): *Encyclopedia of Actuarial Science.* 3 vols. Wiley, Chichester, 2004.

[V1] S.R.S. Varadhan: *Probability Theory.* Amer. Math. Soc., Providence, 2001.

[V2] S.R.S. Varadhan: *Stochastic Processes.* Amer. Math. Soc., Providence, 2007.

# Index


Additivity, 68
adjustment coefficient, 113
Almost sure convergence, 177
arrival times, 29

Binomial distribution, 178
Blackwell's renewal theorem, 44
Brownian motion, 130, 141, 189

central limit theorem, 34, 178
Chapman-Kolmogorov equation, 133
Choquet-Deny theorem, 50
claim arrival rate, 7
claim arrival times, 2, 10
claim number process, 3
claim sizes, 2
compound geometric distribution, 119
compound Poisson process, 21
conditional expectation, 185
conditional probability, 185
Consistency, 68
convergence in probability, 177
converges weakly, 44
convolution, 28, 176
coupling technique, 53
Cramér-Lundberg approximation, 120
Cramér-Lundberg condition, 120
Cramér-Lundberg model, 21

defective renewal equation, 118
delayed renewal function, 42
delayed renewal process, 42
delayed renewal sequence, 42
diffusion equation, 191
diffusion process, 194
directly Riemann integrable, 45
discounted sum, 21
distribution, 174
distribution function, 173
domain of attraction, 180
domain of normal attraction, 184
Doob's inequalities, 186

Elementary renewal theorem, 32
Embrechts-Veraverbeke approximation, 124
equilibrium distribution, 43
Erlang distribution, 26
Esscher transform, 71
expectation, 175
Expected value principle, 58
Exponential distribution, 179
exponential distribution, 10
Exponential principle, 59
exponential tilting, 71

failure rate, 96
Feller property, 145
filtration, 131, 188
filtration generated by the process, 24

first entrance time, 109
fundamental solution to the heat equation, 191

Gamma distribution, 180
geometric Brownian motion, 141
Geometric distribution, 179

hazard rate, 96
heat equation, 191
heavy-tailed distribution, 72
heavy-tailed Weibull distribution, 106
Helly selection principle, 64
Homogeneity, 68

independent, 175
independent increments, 6
infinitesimal generator, 136, 153
integrated tail distribution, 43
integro-differential operator, 136
interarrival times, 10, 29
Ito integral, 141, 151, 189
Ito's formula, 151, 193

Jensen's inequality, 185
joint probability density function, 175

Karamata representation theorem, 75
Key renewal theorem, 46

ladder epoch, 109
Laplace-Stieltjes transform, 36
large deviations, 111
lattice distribution, 44
Levy process, 21, 131
Light-tailed distributions, 69
locally bounded, 40
Lognormal distribution, 100
long-tailed distribution, 93
Lundberg bound, 112
Lundberg coefficient, 113
Lundberg risk process, 107

Markov property, 24, 131
martingale, 33, 186
moment generating function, 69

Net or equivalence principle, 58
net profit condition (NPC), 111
nonlattice distribution, 44
Nonnegative loading, 68
Normal distribution, 179

Optional Sampling theorem, 187
optional sampling theorem, 33
Optional Stopping theorem, 187
Order statistics, 17

Pareto distribution, 82
partial maximum, 84
partial sum, 84
Poisson distribution, 178
Poisson process, 3
Poisson process with intensity (or rate) function, 8
Pollaczek-Khinchin formula, 118
premium rate, 67
probability density function, 174
probability mass function, 174
probability measure, 173
probability space, 173

random variable, 173
random walk, 29
Regularly varying functions and distributions, 72
regularly varying with index, 76
regularly varying with tail index, 76

renewal epoch, 29
renewal equation, 40
renewal function, 36
renewal measure, 39
renewal model, 4
renewal process, 29
renewal risk process, 107
renewal sequence, 29
residual life, 16, 51
right continuous filtration, 143
risk aversion factor, 59
Risk process with investment, 141
risk theory, 2
riskless bond, 141
risky asset, 141
ruin, 108
ruin probability, 108, 161
Ruin Problems, 107
ruin time, 108, 151

safety loading factor, 58
Samuelson model, 141
semigroup of operators, 134
Semigroup property, 168
skeleton process, 108
SLLN for renewal process, 31
slowly varying, 182
slowly varying functions, 72
Sparre Andersen model, 54
stable distribution, 181
Standard deviation principle, 59
stationary excess distribution, 43
stationary increments, 6
stochastic differential equation, 194
stochastic differential equations, 130
stochastic integral equation, 141, 194
stopping time, 25, 187
strong law of large numbers, 177
Strong Markov property, 149
strong Markov property, 190
strong solution, 143, 194
Subexponential distributions, 84
submartingale, 186
supermartingale, 186
support theorem, 152
surplus process, 107
survival probability, 116, 161

tail equivalent, 90
total claim amount process, 3
transition probability density function, 191
transition probability function, 133
Truncated standard normal, 71

Uniform distribution, 179
utility function, 60
utility theory, 59

Variance, 175
Variance principle, 58

waiting times, 10
Wald's identity, 33
Wiener measure, 190

zero utility premium, 61

## Texts and Readings in Mathematics

1. R. B. Bapat: Linear Algebra and Linear Models (Second Edition)
2. Rajendra Bhatia: Fourier Series (Second Edition)
3. C. Musili: Representations of Finite Groups
4. H. Helson: Linear Algebra (Second Edition)
5. D. Sarason: Complex Function Theory (Second Edition)
6. M. G. Nadkarni: Basic Ergodic Theory (Second Edition)
7. H. Helson: Harmonic Analysis (Second Edition)
8. K. Chandrasekharan: A Course on Integration Theory
9. K. Chandrasekharan: A Course on Topological Groups
10. R. Bhatia (ed.): Analysis, Geometry and Probability
11. K. R. Davidson: C* – Algebras by Example
12. M. Bhattacharjee et al.: Notes on Infinite Permutation Groups
13. V. S. Sunder: Functional Analysis — Spectral Theory
14. V. S. Varadarajan: Algebra in Ancient and Modern Times
15. M. G. Nadkarni: Spectral Theory of Dynamical Systems
16. A. Borel: Semisimple Groups and Riemannian Symmetric Spaces
17. M. Marcolli: Seiberg – Witten Gauge Theory
18. A. Bottcher and S. M. Grudsky: Toeplitz Matrices, Asymptotic Linear Algebra and Functional Analysis
19. A. R. Rao and P. Bhimasankaram: Linear Algebra (Second Edition)
20. C. Musili: Algebraic Geometry for Beginners
21. A. R. Rajwade: Convex Polyhedra with Regularity Conditions and Hilbert's Third Problem
22. S. Kumaresan: A Course in Differential Geometry and Lie Groups
23. Stef Tijs: Introduction to Game Theory
24. B. Sury: The Congruence Subgroup Problem
25. R. Bhatia (ed.): Connected at Infinity
26. K. Mukherjea: Differential Calculus in Normed Linear Spaces (Second Edition)
27. Satya Deo: Algebraic Topology: A Primer (Corrected Reprint)
28. S. Kesavan: Nonlinear Functional Analysis: A First Course
29. S. Szabó: Topics in Factorization of Abelian Groups
30. S. Kumaresan and G. Santhanam: An Expedition to Geometry
31. D. Mumford: Lectures on Curves on an Algebraic Surface (Reprint)
32. J. W. Milnor and J. D. Stasheff: Characteristic Classes (Reprint)
33. K. R. Parthasarathy: Introduction to Probability and Measure (Corrected Reprint)
34. A. Mukherjee: Topics in Differential Topology
35. K. R. Parthasarathy: Mathematical Foundations of Quantum Mechanics
36. K. B. Athreya and S. N. Lahiri: Measure Theory
37. Terence Tao: Analysis I (Second Edition)
38. Terence Tao: Analysis II (Second Edition)

39. W. Decker and C. Lossen: Computing in Algebraic Geometry
40. A. Goswami and B. V. Rao: A Course in Applied Stochastic Processes
41. K. B. Athreya and S. N. Lahiri: Probability Theory
42. A. R. Rajwade and A. K. Bhandari: Surprises and Counterexamples in Real Function Theory
43. G. H. Golub and C. F. Van Loan: Matrix Computations (Reprint of the Third Edition)
44. Rajendra Bhatia: Positive Definite Matrices
45. K. R. Parthasarathy: Coding Theorems of Classical and Quantum Information Theory
46. C. S. Seshadri: Introduction to the Theory of Standard Monomials
47. Alain Connes and Matilde Marcolli: Noncommutative Geometry, Quantum Fields and Motives
48. Vivek S. Borkar: Stochastic Approximation: A Dynamical Systems Viewpoint
49. B. J. Venkatachala: Inequalities: An Approach Through Problems
50. Rajendra Bhatia: Notes on Functional Analysis
51. A. Clebsch (ed.) Jacobi's Lectures on Dynamics (Second Revised Edition)
52. S. Kesavan: Functional Analysis
53. V. Lakshmibai and Justin Brown: Flag Varieties: An Interplay of Geometry, Combinatorics, and Representation Theory