Allan Gut

Stopped Random Walks

Limit Theorems and Applications
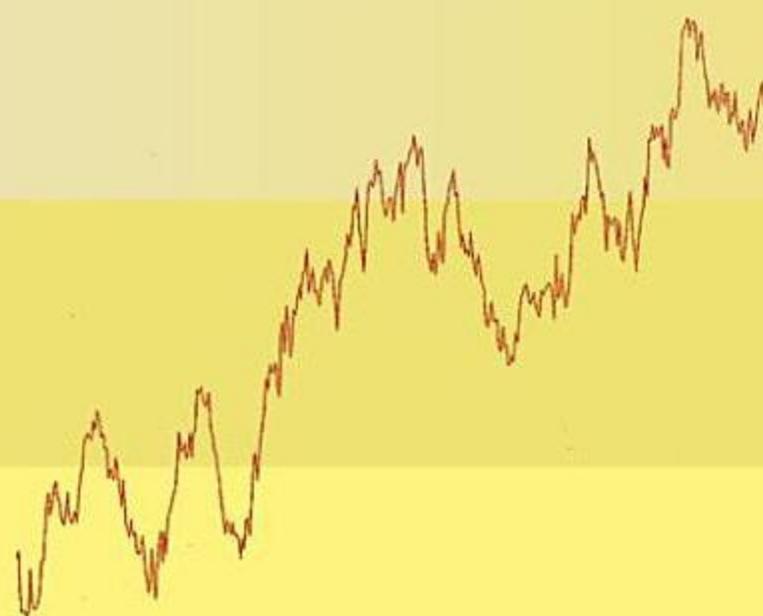
Second Edition

Springer

Springer Series in Operations Research and Financial Engineering

Allan Gut

Stopped Random Walks

Limit Theorems and Applications

Second Edition

Springer

Allan Gut
Department of Mathematics
Uppsala University
SE-751 06 Uppsala
Sweden
allan.gut@math.uu.se

Series Editors:
Thomas V. Mikosch
University of Copenhagen
Laboratory of Actuarial Mathematics
DK-1017 Copenhagen
Denmark
mikosch@act.ku.dk

Stephen M. Robinson
University of Wisconsin-Madison
Department of Industrial Engineering
Madison, WI 53706
USA
smrobins@facstaff.wise.edu

Sidney I. Resnick
Cornell University
School of Operations Research
    and Industrial Engineering
Ithaca, NY 14853
USA
sirl@cornell.edu

Preface to the 1st edition

My first encounter with renewal theory and its extensions was in 1967/68 when I took a course in probability theory and stochastic processes, where the then recent book *Stochastic Processes* by Professor N.U. Prabhu was one of the requirements. Later, my teacher, Professor Carl-Gustav Esseen, gave me some problems in this area for a possible thesis, the result of which was Gut (1974a).

Over the years I have, on and off, continued research in this field. During this time it has become clear that many limit theorems can be obtained with the aid of limit theorems for random walks indexed by families of positive, integer valued random variables, typically by families of stopping times. During the spring semester of 1984 Professor Prabhu visited Uppsala and very soon got me started on a book focusing on this aspect. I wish to thank him for getting me into this project, for his advice and suggestions, as well as his kindness and hospitality during my stay at Cornell in the spring of 1985.

Throughout the writing of this book I have had immense help and support from Svante Janson. He has not only read, but scrutinized, every word and every formula of this and earlier versions of the manuscript. My gratitude to him for all the errors he found, for his perspicacious suggestions and remarks and, above all, for what his unusual personal as well as scientific generosity has meant to me cannot be expressed in words.

It is also a pleasure to thank Ingrid Torrång for checking most of the manuscript, and for several discoveries and remarks.

Inez Hjelm has typed and retyped the manuscript. My heartfelt thanks and admiration go to her for how she has made typing into an art and for the everlasting patience and friendliness with which she has done so.

The writing of a book has its ups and downs. My final thanks are to all of you who shared the ups and endured the downs.

Uppsala   *Allan Gut*
September 1987

Preface to the 2nd edition

By now *Stopped Random Walks* has been out of print for a number of years. Although 20 years old it is still a fairly complete account of the basics in renewal theory and its ramifications, in particular first passage times of random walks. Behind all of this lies the theory of sums of a random number of (i.i.d.) random variables, that is, of *stopped random walks*.

I was therefore very happy when I received an email in which I was asked whether I would be interested in a reprint, or, rather, an updated 2nd edition of the book.

And here it is!

To the old book I have added another chapter, Chapter 6, briefly traversing nonlinear renewal processes in order to present more thoroughly the analogous theory for *perturbed random walks*, which are modeled as a random walk plus "noise", and thus behave, roughly speaking, as $\mathcal{O}(n) + o(n)$. The classical limit theorems as well as moment considerations are proved and discussed in this setting. Corresponding results are also presented for the special case when the perturbed random walk on average behaves as a continuous function of the arithmetic mean of an i.i.d. sequence of random variables, the point being that this setting is most apt for applications to exponential families, as will be demonstrated.

A short outlook on further results, extensions and generalizations is given toward the end of the chapter. A list of additional references, some of which had been overlooked in the first edition and some that appeared after the 1988 printing, is also included, whether explicitly cited in the text or not.

Finally, many thanks to Thomas Mikosch for triggering me into this and for a thorough reading of the second to last version of Chapter 6.

Uppsala   *Allan Gut*
October 2008

Contents

Notation and Symbols

$x \vee y$	$\max\{x, y\}$
$x \wedge y$	$\min\{x, y\}$
x^+	$x \vee 0$
x^-	$-(x \wedge 0)$
$[x]$	the largest integer in x, the integral part of x
$I\{A\}$	the indicator function of the set A
$\text{Card}\{A\}$	the number of elements in the set A
$X \stackrel{d}{=} Y$	X and Y are equidistributed
$X_n \xrightarrow{\text{a.s.}} X$	X_n converges almost surely to X
$X_n \xrightarrow{p} X$	X_n converges in probability to X
$X_n \xrightarrow{d} X$	X_n converges in distribution to X
$\Longrightarrow$	weak convergence
$\stackrel{J_1}{\Longrightarrow}$	weak convergence in the Skorohod J_1-topology
$\stackrel{M_1}{\Longrightarrow}$	weak convergence in the Skorohod M_1-topology
$\sigma\{X_k, 1 \le k \le n\}$	the σ-algebra generated by $X_1, X_2, \ldots, X_n$
EX exists	at least one of EX^- and EX^+ are finite
$\|X\|_r$	$(E\|X\|^r)^{1/r}$
$\Phi(x)$	$\displaystyle\int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-y^2/2} dy \quad (-\infty < x < \infty)$
$W(t)$	Brownian motion or the Wiener process
i.i.d.	independent, identically distributed
i.o.	infinitely often
iff	if and only if
$\square$	end of proof

Introduction

A *random walk* is a sequence $\{S_n, n \geq 0\}$ of random variables with independent, identically distributed (i.i.d.) increments $\{X_k, k \geq 1\}$ and $S_0 = 0$. A *Bernoulli random walk* (also called a Binomial random walk or a Binomial process) is a random walk for which the steps equal 1 or 0 with probabilities p and q, respectively, where $0 < p < 1$ and $p + q = 1$. A *simple random walk* is a random walk for which the steps equal $+1$ or -1 with probabilities p and q, respectively, where, again, $0 < p < 1$ and $p + q = 1$. The case $p = q = \frac{1}{2}$ is called the *symmetric simple random walk* (sometimes the coin-tossing random walk or the symmetric Bernoulli random walk). A *renewal process* is a random walk with nonnegative increments; the Bernoulli random walk is an example of a renewal process.

Among the oldest results for random walks are perhaps the Bernoulli law of large numbers and the De Moivre–Laplace central limit theorem for Bernoulli random walks and simple random walks, which provide information about the asymptotic behavior of such random walks. Similarly, limit theorems such as the classical law of large numbers, the central limit theorem and the Hartman–Wintner law of the iterated logarithm can be interpreted as results on the asymptotic behavior of (general) random walks.

These limit theorems all provide information about the random walks after a *fixed number of steps*. It is, however, from the point of view of applications, more natural to consider random walks evaluated at fixed or specific random *times*, and, hence, after a *random number of steps*. Namely, suppose we have some application in mind, which is modeled by a random walk; such applications are abundant. Let us just mention sequential analysis, queueing theory, insurance risk theory, reliability theory and the theory of counters. In all these cases one naturally studies the process (evolution) as *time* goes by. In particular, it is more interesting to observe the process at the time point when some "special" event occurs, such as the first time the process exceeds a given value rather than the time points when "ordinary" events occur. From this point of view it is thus more relevant to study *randomly indexed random walks*.

Let us make this statement a little more precise by briefly mentioning some examples, some of which will be discussed in Section 4.3 in greater detail. In the most classical one, sequential analysis, one studies a random walk until it leaves a finite interval and accepts or rejects the null hypothesis depending on where the random walk leaves this interval. Clearly the most interesting quantities are the *random* sample size (and, for example, the ASN, that is, the average sample number), and the random walk evaluated at the time point when the decision is made, that is, the value of the random walk when the index equals the exit time.

As for queueing theory, inventory theory, insurance risk theory or the theory of counters, the associated random walk describes the evolution of the system after a fixed number of steps, namely at the instances when the relevant objects (customers, claims, impulses, etc.) come and go. However, in real life one would rather be interested in the state of affairs at fixed or specific random *times*, that is, after a *random number of steps*. For example, it is of greater interest to know what the situation is when the queue first exceeds a given length or when the insurance company first has paid more than a certain amount of money, than to investigate the queue after 10 customers have arrived or the capital of the company after 15 claims. Some simple cases can be covered within the framework of renewal theory.

Another important application is reliability theory, where also generalizations of renewal theory come into play. In the standard example in renewal theory one considers components in a machine and assumes that they are instantly replaced upon failure. The renewal counting process then counts the number of replacements during a fixed time interval. An immediate generalization, called replacement based on age, is to replace the components at failure or at some fixed maximal age, whichever comes first. The random walk whose increments are the interreplacement times then describes the times of the first, second, etc. replacement. It is certainly more relevant to investigate, for example, the number of replacements during a fixed *time* interval or the number of replacements due to failure during a fixed *time* interval and related quantities. Further, if the replacements cause different costs depending on the reason for replacement one can study the total cost generated by these failures within this framework.

There are also applications within the theory of random walks itself. Much attention has, for example, been devoted to the theory of ladder variables, that is, the successive record times and record values of a random walk. A generalization of ladder variables and also of renewal theory and sequential analysis is the theory of first passage times across horizontal boundaries, where one considers the index of the random walk when it first reaches above a given value, t, say, that is, when it leaves the interval $(-\infty, t]$. This theory has applications in sequential analysis when the alternative is one-sided. A further generalization, which allows more general sequential test procedures, is obtained if one considers first passage times across more general (time dependent) boundaries.

These examples clearly motivate a need for a theory on the (limiting) behavior of randomly indexed random walks. Furthermore, in view of the immense interest and effort that has been spent on ordinary random walks, in particular, on the classical limit theorems mentioned earlier, it is obvious that it also is interesting from a purely theoretical point of view to establish such a theory. Let us further mention, in passing, that it has proved useful in certain cases to prove ordinary limit theorems by a detour via a limit theorem for a randomly indexed process.

We are thus led to the study of randomly indexed random walks because of the vast applicability, but also because it is a theory, which is interesting in its own right. It has, however, not yet found its way into books on probability theory.

The purpose of this book is to present the theory of limit theorems for randomly indexed random walks, to show how these results can be used to prove limit theorems for renewal counting processes, first passage time processes for random walks with positive drift and certain two-dimensional random walks and, finally, how these results, in turn, are useful in various kinds of applications.

Let us now make a brief description of the contents of the book.

Let $\{S_n, n \geq 0\}$ be a random walk and $\{N(t), t \geq 0\}$ a family of random indices. The randomly indexed random walk then is the family

$$\{S_{N(t)}, t \geq 0\}. \tag{1}$$

Furthermore, we do not make any assumption about independence between the family of indices and the random walk. In fact, in the typical case the random indices are defined in terms of the random walk; for example, as the first time some special event occurs.

An early (the first?) general limit theorem for randomly indexed families of random variables is the theorem of Anscombe (1952), where sequential estimation is considered. Later, Rényi (1957), motivated by a problem on alternating renewal processes, stated and proved a version of Anscombe's theorem for random walks, which runs as follows:

Let $\{S_n, n \geq 0\}$ be a random walk whose (i.i.d.) increments have mean 0 and positive, finite variance σ^2. Further, suppose that $\{N(t), t \geq 0\}$ is a family of positive, integer valued random variables, such that

$$\frac{N(t)}{t} \xrightarrow{p} \theta \quad (0 < \theta < \infty) \quad \text{as } t \to \infty. \tag{2}$$

Then $S_{N(t)}/\sqrt{N(t)}$ and $S_{N(t)}/\sqrt{t}$ are both asymptotically normal with mean 0 and variances σ^2 and $\sigma^2 \cdot \theta$, respectively.

There exist, of course, more general versions of this result. We are, however, not concerned with them in the present context.

A more general problem is as follows: Given a sequence of random variables $\{Y_n, n \geq 1\}$, such that $Y_n \to Y$ as $n \to \infty$ and a family of random indices

$\{N(t), t \geq 0\}$, such that $N(t) \to \infty$ as $t \to \infty$, when is it possible to conclude that

$$Y_{N(t)} \to Y \quad \text{as} \quad t \to \infty? \tag{3}$$

The convergence mode in each case may be one of the four standard ones; a.s. convergence, convergence in probability, in L^r (in r-mean) and in distribution. For example, Anscombe's theorem above is of that kind; condition (2) implies that $N(t) \xrightarrow{p} \infty$ as $t \to \infty$.

The first general investigation of this class of problems is due to Richter (1965). We begin Chapter 1 by reviewing some of his results, after which we turn our attention to randomly indexed random walks. Now, in order to prove theorems on uniform integrability and moment convergence for randomly indexed random walks under minimal conditions it turns out that it is necessary that the indices are *stopping times*, that is, that they do not depend on the future. We call a random walk thus indexed a *stopped random walk*. Since the other limit theorems hold for random walks indexed by more general families of random variables, it follows, as an unfortunate consequence, that the title of this book is a little too restrictive; on the other hand, from the point of view of applications it is natural that the stopping procedure does not depend on the future. The stopped random walk is thus what we should have to mind.

To make the treatise more selfcontained we include some general background material for renewal processes and random walks. This is done in Chapter 2. After some introductory material we give, in the first half of the chapter, a survey of the general theory for renewal processes. However, no attempt is made to give a complete exposition. Rather, we focus on the results which are relevant to the approach of this book. Proofs will, in general, only be given in those cases where our findings in Chapter 1 can be used. For more on renewal processes we refer to the books by Feller (1968, 1971), Prabhu (1965), Çinlar (1975), Jagers (1975) and Asmussen (2003). The pioneering work of Feller (1949) on recurrent events is also important in this context.

In the second half of Chapter 2 we survey some of the general theory for random walks in the same spirit as that of the first half of the chapter.

A major step forward in the theory of random walks was taken in the 1950s when classical fluctuation theory, combinatorial methods, Wiener–Hopf factorization, etc. were developed. Chung and Fuchs (1951) introduced the concepts of possible points and recurrent points and showed, for example, that either all (suitably interpreted) or none of the points are recurrent (persistent), see also Chung and Ornstein (1962). Sparre Andersen (1953a,b, 1954) and Spitzer (1956, 1960, 1976) developed fluctuation theory by combinatorial methods and Tauberian theorems. An important milestone here is Spitzer (1976), which in its first edition appeared in 1964. These parts of random walk theory are not covered in this book; we refer to the work cited above and to the books by Feller (1968, 1971), Prabhu (1965) and Chung (1974).

We begin instead by classifying random walks as transient or recurrent and then as drifting or oscillating. We introduce ladder variables, the sequences of partial maxima and partial minima and prove some general limit theorems for those sequences.

A more exhaustive attempt to show the usefulness of the results of Chapter 1 is made in Chapter 3, where we extend renewal theoretic results to random walks $\{S_n, n \geq 0\}$ on the whole real line. We assume throughout that the random walk drifts to $+\infty$. In general we assume, in addition, that the increments $\{X_k, k \geq 1\}$ have positive, finite mean (or, at least, that $E(X_1^-) < \infty$).

There are several ways of making such extensions; the most immediate one is, in our opinion, based on the family of *first passage times* $\{\nu(t), t \geq 0\}$ defined by

$$\nu(t) = \min\{n\colon S_n > t\}. \tag{4}$$

Following are some arguments supporting this point of view.

For renewal processes one usually studies the (*renewal*) *counting process* $\{N(t), t \geq 0\}$ defined by

$$N(t) = \max\{n\colon S_n \leq t\}. \tag{5}$$

Now, since renewal processes have nonnegative increments one has, in this case, $\nu(t) = N(t) + 1$ and then one may study either process and make inference about the other. However, in order to prove certain results for counting processes one uses (has to use) stopping times, in which case one introduces first passage time processes in the proofs. It is thus, mathematically, more convenient to work with first passage time processes.

Secondly, many of the problems in renewal theory are centered around the *renewal function* $U(t) = \sum_{n=1}^{\infty} P(S_n \leq t)$ $(= EN(t))$, which is finite for all t. However, for random walks it turns out that it is necessary that $E(X_1^-)^2 < \infty$ for this to be the case. An extension of the so-called elementary renewal theorem, based on $U(t)$, thus requires this additional condition and, thus, cannot hold for all random walks under consideration. A final argument is that some very important random time points considered for random walks are the ladder epochs, where, in fact, the first strong ascending ladder epoch is $\nu(0)$.

So, as mentioned above, our extension of renewal theory is based on the family of first passage times $\{\nu(t), t \geq 0\}$ defined in (4). In Chapter 3 we investigate first passage time processes and the associated families of stopped random walks $\{S_{\nu(t)}, t \geq 0\}$, thus obtaining what one might call a *renewal theory for random walks with positive drift*. Some of the results generalize the analogs for renewal processes, some of them, however, do (did) not exist earlier for renewal processes. This is due to the fact that some of the original proofs for renewal processes depended heavily on the fact that the increments were nonnegative, whereas more modern methods do not require this.

Just as for Chapter 1 we may, in fact, add that a complete presentation of this theory has not been given in books before.

Before we proceed to describe the contents of the remaining chapters, we pause a moment in order to mention something that is not contained in the book. Namely, just as we have described above that one can extend renewal theory to drifting random walks it turns out that it is also possible to do so for oscillating random walks, in particular for those whose increments have mean 0.

However, these random walks behave completely differently compared to the drifting ones. For example, when the mean equals 0 the random walk is recurrent and every (possible) finite interval is visited infinitely often almost surely, whereas drifting random walks are transient and every finite interval is only visited finitely often. Secondly, our approach, which is based on the limit theorems for stopped random walks obtained in Chapter 1, requires the relation (2). Now, for oscillating random walks with finite variance, the first passage time process belongs to the domain of attraction of a positive stable law with index $\frac{1}{2}$, that is, (2) does not hold. For drifting random walks, however, (2) holds for counting processes as well as for first passage time processes.

The oscillating random walk thus yields a completely different story. There exists, however, what might be called a renewal theory for oscillating random walks. We refer the interested reader to papers by Port and Stone (1967), Ornstein (1969a,b) and Stone (1969) and the books by Revuz (1975) and Spitzer (1976), where renewal theorems are proved for a generalized renewal function. For asymptotics of first passage time processes, see e.g. Erdős and Kac (1946), Feller (1968) and Teicher (1973).

Chapter 4 consists of four main parts, each of which corresponds to a major generalization or extension of the results derived earlier. In the first part we investigate a class of two-dimensional random walks. Specifically, we establish limit theorems of the kind discussed in Chapter 3 for the process obtained by considering the second component of a random walk evaluated at the first passage times of the first component (or vice versa). Thus, let $\{(U_n, V_n),\ n \geq 0\}$ be a two-dimensional random walk, suppose that the increments of the first component have positive mean and define

$$\tau(t) = \min\{n\colon U_n > t\} \qquad (t \geq 0). \qquad (6)$$

The process of interest then is $\{V_{\tau(t)}, t \geq 0\}$.

Furthermore, if, in particular, $\{U_n, n \geq 0\}$ is a renewal process, then it is also possible to obtain results for $\{V_{M(t)}, t \geq 0\}$, where

$$M(t) = \max\{n\colon U_n \leq t\} \qquad (t \geq 0) \qquad (7)$$

(note that $\tau(t) = M(t) + 1$ in this case).

Interestingly enough, processes of the above kind arise in a variety of contexts. In fact, the motivation for the theoretical results of the first part of the chapter (which is largely based on Gut and Janson (1983)) comes through the work on a problem in the theory of chromatography (Gut and

Ahlberg (1981)), where a special case of two-dimensional random walks was considered (the so-called alternating renewal process). Moreover, it turned out that various further applications of different kinds could be modeled in the more general framework of the first part of the chapter. In the second part of the chapter we present a number of these applications.

A special application from within probability theory itself is given by the sequence of partial maxima $\{M_n, n \geq 0\}$, defined by

$$M_n = \max\{0, S_1, S_2, \ldots, S_n\}. \tag{8}$$

Namely, a representation formula obtained in Chapter 2 allows us to treat this sequence in the setup of the first part of Chapter 4 (provided the underlying random walk drifts to $+\infty$). However, the random indices are not stopping times in this case; the framework is that of $\{V_{M(t)}, t \geq 0\}$ as defined through (7). In the third part of the chapter we apply the results from the first part, thus obtaining limit theorems for M_n as $n \to \infty$ when $\{S_n, n \geq 0\}$ is a random walk whose increments have positive mean. These results supplement those obtained earlier in Chapter 2.

In the final part of the chapter we study first passage times across time-dependent barriers, the typical case being

$$\nu(t) = \min\{n \colon S_n > tn^\beta\} \qquad (0 \leq \beta < 1, t \geq 0), \tag{9}$$

where $\{S_n, n \geq 0\}$ is a random walk whose increments have positive mean. The first more systematic investigation of such stopping times was made in Gut (1974a). Here we extend the results obtained in Chapter 3 to this case.

We also mention that first passage times of this more general kind provide a starting point for what is sometimes called *nonlinear renewal theory*, see Lai and Siegmund (1977, 1979), Woodroofe (1982) and Siegmund (1985).

Just as before the situation is radically different when $EX_1 = 0$. Some investigations concerning the first passage times defined in (9) have, however, been made in this case for the two-sided versions $\min\{n \colon |S_n| > tn^\beta\}$ ($0 < \beta \leq 1/2$). Some references are Breiman (1965), Chow and Teicher (1966), Gundy and Siegmund (1967), Lai (1977) and Brown (1969). Note also that the case $\beta = 1/2$ is of special interest here in view of the central limit theorem.

Beginning with the work of Erdős and Kac (1946) and Donsker (1951) the central limit theorem has been generalized to *functional limit theorems*, also called *weak invariance principles*. The standard reference here is Billingsley (1968, 1999). The law of the iterated logarithm has been generalized analogously into a so-called *strong invariance principle* by Strassen (1964), see also Stout (1974). In Chapter 5 we present corresponding generalizations for the processes discussed in the earlier chapters.

A final Chapter 6 is devoted to analogous results for *perturbed random walks*, which can be viewed as a random walk plus "noise", roughly speaking, as $\mathcal{O}(n) + o(n)$. The classical limit theorems as well as moment considerations

are proved and discussed in this setting. A special case is also treated and some applications to repeated significance tests are presented. The chapter closes with an outlook on further extensions and generalizations.

The book concludes with two appendices containing some prerequisites which might not be completely familiar to everyone.

1

Limit Theorems for Stopped Random Walks

1.1 Introduction

Classical limit theorems such as the law of large numbers, the central limit theorem and the law of the iterated logarithm are statements concerning sums of independent and identically distributed random variables, and thus, statements concerning *random walks*. Frequently, however, one considers random walks evaluated after a *random* number of steps. In sequential analysis, for example, one considers the time points when the random walk leaves some given finite interval. In renewal theory one considers the time points generated by the so-called renewal counting process. For random walks on the whole real line one studies first passage times across horizontal levels, where, in particular, the zero level corresponds to the first ascending ladder epoch. In reliability theory one may, for example, be interested in the total cost for the replacements made during a fixed time interval and so on.

It turns out that the limit theorems mentioned above can be extended to *random walks with random indices*. Frequently such limit theorems provide a limiting relation involving the randomly indexed sequence as well as the random index, but if it is possible to obtain a precise estimate for one of them, one can obtain a limit theorem for the other. For example, if a process is stopped when something "rather precise" occurs one would hope that it might be possible to replace the stopped process by something deterministic, thus obtaining a result for the family of random indices.

Such limit theorems seem to have been first used in the 1950s by F.J. Anscombe (see Section 1.3 below), D. Blackwell in his extension of his renewal theorem (see Theorem 3.6.6 below) and A. Rényi in his proof of a theorem of Takács (this result will be discussed in a more general setting in Chapter 4). See also Smith (1955). Since then this approach has turned out to be increasingly useful. The literature in the area is, however, widely scattered.

The aim of the first chapter of this book is twofold. Firstly it provides a unified presentation of the various limit theorems for (certain) randomly indexed random walks, which is a theory in its own right. Secondly it will serve

as a basis for the chapters to follow. Let us also, in passing, mention that it has proved useful in various contexts to prove ordinary limit theorems by first proving them for randomly indexed processes and then by some approximation procedure arrive at the desired result.

Let us now introduce the notion of a *stopped random walk*—the central object of the book. As a preliminary observation we note that the renewal counting process, mentioned above is not a family of stopping times, whereas the exit times in sequential analysis, or the first passage times for random walks are stopping times; the counting process depends on the future, whereas the other random times do not (for the definition of a stopping time we refer to Section A.2).

Now, for all limits theorems below, which do not involve convergence of moments or uniform integrability, the stopping time property is of no relevance. It is in connection with theorems on uniform integrability that the stopping time property is essential (unless one requires additional assumptions). Since our main interest is the case when the family of random indices is, indeed, a family of stopping times we call a random walk thus indexed a *stopped random walk*. We present, however, our results without the stopping time assumption whenever this is possible. As a consequence the heading of this chapter (and of the book) is a little too restrictive, but, on the other hand, it captures the heart of our material.

Before we begin our presentation of the limit theorems for stopped random walks we shall consider the following, more general problem:

Let $(\Omega, \mathscr{F}, P)$ be a probability space, let $\{Y_n, n \geq 1\}$ be a sequence of random variables and let $\{N(t), t \geq 0\}$ be a family of positive, integer valued random variables. Suppose that

$$Y_n \to Y \quad \text{in some sense as } n \to \infty \tag{1.1}$$

and that

$$N(t) \to +\infty \quad \text{in some sense at } t \to \infty. \tag{1.2}$$

When can we conclude that

$$Y_{N(t)} \to Y \quad \text{in some sense as } t \to \infty? \tag{1.3}$$

Here "in some sense" means one of the four standard convergence modes; almost surely, in probability, in distribution or in L^r.

After presenting some general answers and counterexamples when the question involves a.s. convergence and convergence in probability we turn our attention to stopped random walks. Here we shall consider all four convergence modes and also, but more briefly, the law of the iterated logarithm and complete convergence.

The first elementary result of the above kind seems to be the following:

Theorem 1.1. *Let $\{Y_n, n \geq 1\}$ be a sequence of random variables such that*

$$Y_n \xrightarrow{d} Y \quad as \quad n \to \infty. \tag{1.4}$$

Suppose further that $\{N(t), \ t \geq 0\}$ is a family of positive, integer valued random variables, independent of $\{Y_n, n \geq 1\}$ and such that

$$N(t) \xrightarrow{P} +\infty \quad as \quad t \to \infty. \tag{1.5}$$

Then

$$Y_{N(t)} \xrightarrow{d} Y \quad as \quad t \to \infty. \tag{1.6}$$

Proof. Let φ_U denote the characteristic function of the random variable U. By the independence assumption we have

$$\varphi_{Y_{N(t)}}(u) = \sum_{k=1}^{\infty} E(e^{iuY_k}|N(t) = k) \cdot P(N(t) = k)$$

$$= \sum_{k=1}^{\infty} \varphi_{Y_k}(u) \cdot P(N(t) = k).$$

Now, choose k_0 so large that $|\varphi_{Y_k}(u) - \varphi_Y(u)| \leq \varepsilon$ for $k > k_0$ and then t_0 so large that $P(N(t) \leq k_0) < \varepsilon$ for $t > t_0$. We then obtain

$$|\varphi_{Y_{N(t)}}(u) - \varphi_Y(u)| = \left| \sum_{k=1}^{\infty} (\varphi_{Y_k}(u) - \varphi_Y(u)) \cdot P(N(t) = k) \right|$$

$$\leq \sum_{k=1}^{k_0} |\varphi_{Y_k}(u) - \varphi_Y(u)| \cdot P(N(t) = k)$$

$$+ \sum_{k=k_0+1}^{\infty} |\varphi_{Y_k}(u) - \varphi_Y(u)| \cdot P(N(t) = k)$$

$$\leq 2 \cdot P(N(t) \leq k_0) + \varepsilon \cdot P(N(t) > k_0)$$

$$\leq 2 \cdot \varepsilon + \varepsilon \cdot 1 = 3\varepsilon,$$

which in view of the arbitrariness of ε proves the conclusion. $\qquad \square$

We have thus obtained a positive result under minimal assumptions provided $\{Y_n, n \geq 1\}$ and $\{N(t), t \geq 0\}$ are assumed to be independent of each other. In the remainder of this chapter we therefore make no such assumption.

1.2 a.s. Convergence and Convergence in Probability

The simplest case is when one has a.s. convergence for the sequences or families of random variables considered in (1.1) and (1.2). In the following, let $\{Y_n, n \geq 1\}$ be a sequence of random variables and $\{N(t), t \geq 0\}$ a family of positive, integer valued random variables.

Theorem 2.1. *Suppose that*

$$Y_n \xrightarrow{\text{a.s.}} Y \quad as \quad n \to \infty \quad and \quad N(t) \xrightarrow{\text{a.s.}} +\infty \quad as \quad t \to \infty. \tag{2.1}$$

Then

$$Y_{N(t)} \xrightarrow{\text{a.s.}} Y \quad as \quad t \to \infty. \tag{2.2}$$

Proof. Let $A = \{\omega\colon Y_n(\omega) \nrightarrow Y(\omega)\}$, $B = \{\omega\colon N(t,\omega) \nrightarrow +\infty\}$ and $C = \{\omega\colon Y_{N(t,\omega)}(\omega) \nrightarrow Y(\omega)\}$. Then $C \subset A \cup B$, which proves the assertion. □

The problem of what happens if one of $\{Y_n, n \geq 1\}$ and $\{N(t), t \geq 0\}$ converges in probability and the other one converges almost surely is a little more delicate. The following result is due to Richter (1965).

Theorem 2.2. *Suppose that*

$$Y_n \xrightarrow{\text{a.s.}} Y \quad as \quad n \to \infty \quad and \quad N(t) \xrightarrow{p} +\infty \quad as \quad t \to \infty. \tag{2.3}$$

Then

$$Y_{N(t)} \xrightarrow{p} Y \quad as \quad t \to \infty. \tag{2.4}$$

Proof. We shall prove that every subsequence of $Y_{N(t)}$ contains a further subsequence which converges almost surely, and hence also in probability, to Y. (This proves the theorem; see, however, the discussion following the proof.)

Since $N(t) \xrightarrow{p} \infty$ we have $N(t_k) \xrightarrow{p} \infty$ for every subsequence $\{t_k, k \geq 1\}$. Now, from this subsequence we can always select a subsequence $\{t_{k_j}, j \geq 1\}$ such that $N(t_{k_j}) \xrightarrow{\text{a.s.}} \infty$ as $j \to \infty$ (see e.g. Gut (2007), Theorem 5.3.4). Finally, since $Y_n \xrightarrow{\text{a.s.}} Y$ as $n \to \infty$ it follows by Theorem 2.1 that $Y_{N(t_{k_j})} \xrightarrow{\text{a.s.}} Y$ and, hence, that $Y_{N(t_{k_j})} \xrightarrow{p} Y$ as $j \to \infty$. □

Let $\{x_n, n \geq 1\}$ be a sequence of reals. From analysis we know that $x_n \to x$ as $n \to \infty$ if and only if each subsequence of $\{x_n\}$ contains a subsequence which converges to x. In the proof of Theorem 2.2 we used the corresponding result for convergence in probability. Actually, we did more; we showed that each subsequence of $Y_{N(t)}$ contains a subsequence which, in fact, is almost surely convergent. Yet we only concluded that $Y_{N(t)}$ converges in probability.

To clarify this further we first observe that, since $Y_n \xrightarrow{p} Y$ is equivalent to

$$E \frac{|Y_n - Y|}{1 + |Y_n - Y|} \to 0 \quad as \quad n \to \infty$$

(see e.g. Gut (2007), Section 5.7) it follows that $Y_n \xrightarrow{p} Y$ as $n \to \infty$ iff for each subsequence of $\{Y_n\}$ there exists a subsequence converging in probability to Y.

However, the corresponding result is not true for almost sure convergence as is seen by the following example, given to me by Svante Janson.

Example 2.1. Let $\{Y_n, n \geq 1\}$ be a sequence of independent random variables such that $Y_n \in Be(1/n)$, that is, $P(Y_n = 1) = 1/n$ and $P(Y_n = 0) = 1 - 1/n$. Clearly $Y_n \to 0$ in probability but not almost surely as $n \to \infty$. Nevertheless, for each subsequence we can select a subsequence which converges almost surely to 0.

This still raises the question whether the conclusion of Theorem 2.2 can be sharpened or not. The following example shows that it cannot be.

Example 2.2. Let $\Omega = [0, 1]$, $\mathscr{F}$ = the σ-algebra of measurable subsets of Ω and P the Lebesgue measure. Set

$$
Y_n(\omega) = \begin{cases} \frac{1}{m+1}, & \text{if } \frac{j}{2^m} \leq \omega < \frac{j+1}{2^m}, \\ 0, & \text{otherwise,} \end{cases} \quad \text{where } n = 2^m + j, \ 0 \leq j \leq 2^m - 1,
$$

and define

$$
N(t, \omega) = \begin{cases} 1, & \text{if } \frac{s}{2^r} \leq \omega < \frac{s+1}{2^r}, \\ \min\{k \colon k \geq 2^t \text{ and } Y_k(\omega) > 0\}, & \text{otherwise,} \end{cases}
$$

where $t = 2^r + s$, $0 \leq s \leq 2^r - 1$.

Then $Y_n \xrightarrow{\text{a.s.}} 0$ as $n \to \infty$ and $N(t) \xrightarrow{p} \infty$ as $t \to \infty$. It follows that

$$
Y_{N(t,\omega)}(\omega) = \begin{cases} 1, & \text{if } \frac{s}{2^r} \leq \omega < \frac{s+1}{2^r}, \\ \frac{1}{t+1}, & \text{otherwise,} \end{cases} \quad \text{where } t = 2^r + s, \ 0 \leq s \leq 2^r - 1,
$$

and it is now easy to see that $Y_{N(t)}$ converges to 0 in probability but, since $P(Y_{N(t)} = 1 \text{ i.o.}) = 1$, $Y_{N(t)}$ does not converge almost surely as $t \to \infty$.

The above example, due to Richter (1965), is mentioned here because of its close connection with Example 2.4 below. The following, simpler, example with the same conclusion as that of Example 2.2, is due to Svante Janson.

Example 2.3. Let $P(Y_n = 1/n) = 1$. Clearly $Y_n \xrightarrow{\text{a.s.}} 0$ as $n \to \infty$. For any family $\{N(t), t \geq 0\}$ of positive, integer valued random variables we have $Y_{N(t)} = 1/N(t)$, which converges a.s. (in probability) to 0 as $t \to \infty$ iff $N(t) \to \infty$ a.s. (in probability) as $t \to \infty$.

These examples thus demonstrate that Theorem 2.2 is sharp. In the remaining case, that is when $Y_n \xrightarrow{p} Y$ as $n \to \infty$ and $N(t) \xrightarrow{\text{a.s.}} +\infty$ as $t \to \infty$, there is no general theorem as the following example (see Richter (1965)) shows.

Example 2.4. Let the probability space be the same as that of Example 2.2, set

$$Y_n = \begin{cases} 1, & \text{if } \frac{j}{2^m} \le \omega < \frac{j+1}{2^m}, \\ 0, & \text{otherwise,} \end{cases} \quad \text{where } n = 2^m + j, \ 0 \le j \le 2^m - 1,$$

and let

$$N(t) = \min\{k\colon k \ge 2^t \text{ and } Y_k > 0\}.$$

As in Example 2.2, we find that Y_n converges to 0 in probability but not almost surely as $n \to \infty$. Also $N(t) \xrightarrow{\text{a.s.}} +\infty$ as $t \to \infty$.

As for $Y_{N(t)}$ we find that $Y_{N(t)} = 1$ a.s. for all t, that is, no limiting result like those above can be obtained.

In the following theorem we present some applications of Theorem 2.1, which will be of use in the sequel.

Theorem 2.3. *Let $\{X_k, k \ge 1\}$ be i.i.d. random variables and let $\{S_n, n \ge 1\}$ be their partial sums. Further, suppose that $N(t) \xrightarrow{\text{a.s.}} +\infty$ as $t \to \infty$.*

(i) *If $E|X_1|^r < \infty, r > 0$, then*

$$\frac{X_{N(t)}}{(N(t))^{1/r}} \xrightarrow{\text{a.s.}} 0 \quad as \quad t \to \infty. \tag{2.5}$$

If, moreover,

$$\frac{N(t)}{t} \xrightarrow{\text{a.s.}} \theta \quad (0 < \theta < \infty) \quad as \quad t \to \infty, \tag{2.6}$$

then

$$\frac{X_{N(t)}}{t^{1/r}} \xrightarrow{\text{a.s.}} 0 \quad as \quad t \to \infty. \tag{2.7}$$

(ii) *If $E|X_1|^r < \infty$ $(0 < r < 2)$ and $EX_1 = 0$ when $1 \le r < 2$, then*

$$\frac{S_{N(t)}}{(N(t))^{1/r}} \xrightarrow{\text{a.s.}} 0 \quad as \quad t \to \infty. \tag{2.8}$$

If, furthermore, (2.6) holds, then

$$\frac{S_{N(t)}}{t^{1/r}} \xrightarrow{\text{a.s.}} 0 \quad as \quad t \to \infty. \tag{2.9}$$

(iii) *If $E|X_1| < \infty$ and $EX_1 = \mu$, then*

$$\frac{S_{N(t)}}{N(t)} \xrightarrow{\text{a.s.}} \mu \quad as \quad t \to \infty. \tag{2.10}$$

If, furthermore, (2.6) holds, then

$$\frac{S_{N(t)}}{t} \xrightarrow{\text{a.s.}} \mu \cdot \theta \quad as \quad t \to \infty. \tag{2.11}$$

Proof. (i) By assumption we have

$$\sum_{n=1}^{\infty} P(|X_1| > \varepsilon n^{1/r}) < \infty \quad \text{for all } \varepsilon > 0, \tag{2.12}$$

which in view of the stationarity is equivalent to

$$\sum_{n=1}^{\infty} P(|X_n| > \varepsilon n^{1/r}) < \infty \quad \text{for all } \varepsilon > 0, \tag{2.13}$$

which in view of independence and the Borel–Cantelli lemma is equivalent to

$$\frac{X_n}{n^{1/r}} \xrightarrow{\text{a.s.}} 0 \quad as \quad n \to \infty. \tag{2.14}$$

An application of Theorem 2.1 concludes the proof of the first half and the second half is immediate.

(ii) By the Kolmogorov–Marcinkiewicz–Zygmund strong law of large numbers (see e.g. Gut (2007), Theorem 6.7.1, or Loève (1977), p. 255) we have

$$\frac{S_n}{n^{1/r}} \xrightarrow{\text{a.s.}} 0 \quad as \quad n \to \infty, \tag{2.15}$$

which together with Theorem 2.1 yields the first statement, from which the second one follows immediately.

(iii) The strong law of large numbers and Theorem 2.1 together yield (2.10). As for (2.11) we have, by (2.9) and (2.6),

$$\frac{S_{N(t)}}{t} = \frac{S_{N(t)} - \mu N(t)}{t} + \frac{\mu N(t)}{t} \xrightarrow{\text{a.s.}} 0 + \mu\theta = \mu\theta \quad as \quad t \to \infty \tag{2.16}$$

and the proof is complete. □

The final result of this section is a variation of Theorem 2.1. Here we assume that $N(t)$ converges a.s. to an a.s. finite random variable as $t \to \infty$.

Theorem 2.4. *Suppose that*

$$N(t) \xrightarrow{\text{a.s.}} N \quad as \quad t \to \infty, \tag{2.17}$$

where N is an a.s. finite random variable. Then, for any sequence $\{Y_n, n \geq 1\}$,

$$Y_{N(t)} \xrightarrow{\text{a.s.}} Y_N \quad as \quad t \to \infty. \tag{2.18}$$

Proof. Let $A = \{\omega\colon N(t,\omega) \to N(\omega)\}$ and let $\omega \in A$. Since all indices are integer valued it follows that

$$N(t,\omega) = N(\omega) \quad \text{for all } t > t_0(\omega) \tag{2.19}$$

and hence that, in fact,

$$Y_{N(t,\omega)}(\omega) = Y_{N(\omega)}(\omega) \quad \text{for } t > t_0(\omega), \tag{2.20}$$

which proves the assertion.   □

1.3 Anscombe's Theorem

In this section we shall be concerned with a sequence $\{Y_n, n \geq 1\}$ of random variables converging in distribution, which is indexed by a family $\{N(t), t \geq 0\}$ of random variables. The first result *not* assuming independence between $\{Y_n, n \geq 1\}$ and $\{N(t), t \geq 0\}$ is due to Anscombe (1952) and can be described as follows: Suppose that $\{Y_n, n \geq 1\}$ is a sequence of random variables converging in distribution to Y. Suppose further that $\{N(t), t \geq 0\}$ are positive, integer valued random variables such that $N(t)/n(t) \xrightarrow{P} 1$ as $t \to \infty$, where $\{n(t)\}$ is a family of positive numbers tending to infinity. Finally, suppose that

(A)
$$\begin{cases} \text{Given } \varepsilon > 0 \text{ and } \eta > 0 \text{ there exist } \delta > 0 \text{ and } n_0, \text{ such that} \\ P\left(\max_{\{m:\ |m-n|<n\delta\}} |Y_m - Y_n| > \varepsilon\right) < \eta, \quad \text{for all } n > n_0. \end{cases}$$

Then $Y_{N(t)}$ converges in distribution to Y as $t \to \infty$.

Anscombe calls condition (A) uniform continuity in probability of $\{Y_n\}$; it is now frequently called "the Anscombe condition."

This theorem has been generalized in various ways. Here we shall confine ourselves to stating and proving the theorem for the case which will be useful for our purposes, namely the case when

(a) Y_n equals a normalized sum of i.i.d. random variables (a normalized random walk) with finite variance,
(b) $n(t) = t$;

this yields a central limit theorem for stopped random walks.

The following version of Anscombe's theorem was given by Rényi (1957), who also presented a direct proof of the result. Note that condition (A) is not assumed in the statement of the theorem. The proof below is a slight modification of Rényi's original proof, see also Chung (1974), pp. 216–217 or Gut (2007), Theorem 7.3.2. The crucial estimate, which is an application of Kolmogorov's inequality yields essentially the estimate required to prove that condition (A) is automatically satisfied in this case.

Theorem 3.1. *Let $\{X_k, k \geq 1\}$ be a sequence of i.i.d. random variables with mean 0 and variance σ^2 $(0 < \sigma^2 < \infty)$ and let $\{S_n, n \geq 1\}$ denote their partial sums. Further, assume that*

$$\frac{N(t)}{t} \xrightarrow{p} \theta \quad (0 < \theta < \infty) \quad as \quad t \to \infty. \tag{3.1}$$

Then

(i)
$$\frac{S_{N(t)}}{\sigma\sqrt{N(t)}} \xrightarrow{d} N(0,1) \quad as \quad t \to \infty,$$

(ii)
$$\frac{S_{N(t)}}{\sigma\sqrt{t\theta}} \xrightarrow{d} N(0,1) \quad as \quad t \to \infty.$$

Proof. We assume, without restriction, that $\sigma^2 = 1$. Set $n_0 = [\theta t]$. Then

$$\frac{S_{N(t)}}{\sqrt{N(t)}} = \left(\frac{S_{n_0}}{\sqrt{n_0}} + \frac{S_{N(t)} - S_{n_0}}{\sqrt{n_0}} \right) \cdot \sqrt{\frac{n_0}{N(t)}}. \tag{3.2}$$

Since $S_{n_0}/\sqrt{n_0}$ (also) converges in distribution to the standard normal distribution and since the right-most factor converges in probability to 1 as $t \to \infty$ it remains to show that

$$\frac{S_{N(t)} - S_{n_0}}{\sqrt{n_0}} \xrightarrow{p} 0 \quad as \quad t \to \infty. \tag{3.3}$$

To this end, let $\varepsilon, 0 < \varepsilon < 1/3$, be given and set $n_1 = [n_0(1 - \varepsilon^3)] + 1$ and $n_2 = [n_0(1 + \varepsilon^3)]$. Now,

$$P(|S_{N(t)} - S_{n_0}| > \varepsilon\sqrt{n_0})$$

$$= P(|S_{N(t)} - S_{n_0}| > \varepsilon\sqrt{n_0}, \; N(t) \in [n_1, n_2])$$

$$+ P(|S_{N(t)} - S_{n_0}| > \varepsilon\sqrt{n_0}, \; N(t) \notin [n_1, n_2])$$

$$\leq P\left(\max_{n_1 \leq n \leq n_0} |S_n - S_{n_0}| > \varepsilon\sqrt{n_0} \right) + P\left(\max_{n_0 \leq n \leq n_2} |S_n - S_{n_0}| > \varepsilon\sqrt{n_0} \right)$$

$$+ P(N(t) \notin [n_1, n_2]).$$

By Kolmogorov's inequality (cf. e.g. Gut (2007), Theorem 3.1.6) the first two probabilities in the right most member are majorized by $(n_0 - n_1)/\varepsilon^2 n_0$ and $(n_2 - n_0)/\varepsilon^2 n_0$ respectively, that is, by ε, and, by assumption, the last probability is smaller than ε for t sufficiently large. Thus, if t_0 is sufficiently large we have

$$P(|S_{N(t)} - S_{n_0}| > \varepsilon\sqrt{n_0}) < 3\varepsilon \quad \text{for all } t > t_0, \tag{3.4}$$

which concludes the proof of (i). Since (ii) follows from (i) and Cramér's theorem we are done.   □

We conclude this section by stating, without proof, a version of Anscombe's theorem for the case when the distribution of the summands belongs to the domain of attraction of a stable distribution.

Theorem 3.2. *Let $\{X_k, k \geq 1\}$ be a sequence of i.i.d. random variables with mean 0 and let $S_n = \Sigma_{k=1}^n X_k$, $n \geq 1$. Suppose that $\{B_n, n \geq 1\}$ is a sequence of positive normalizing coefficients such that*

$$\frac{S_n}{B_n} \xrightarrow{d} G_\alpha \quad as \quad n \to \infty, \tag{3.5}$$

where G_α is (a random variable distributed according to) a stable law with index α, $1 < \alpha \leq 2$. Further, assume that

$$\frac{N(t)}{t} \xrightarrow{p} \theta \quad (0 < \theta < \infty) \quad as \quad t \to \infty. \tag{3.6}$$

Then

(i)
$$\frac{S_{N(t)}}{B_{N(t)}} \xrightarrow{d} G_\alpha \quad as \quad t \to \infty,$$

(ii)
$$\frac{S_{N(t)}}{B_{[\theta t]}} \xrightarrow{d} G_\alpha \quad as \quad t \to \infty.$$

1.4 Moment Convergence in the Strong Law and the Central Limit Theorem

It is well known that, for example, almost sure convergence and convergence in distribution do not imply that moments converge. An additional requirement for this to be the case is that the sequence of random variables also is uniformly integrable; recall that a sequence $\{Y_n, n \geq 1\}$ of random variables is uniformly integrable iff

$$\lim_{\alpha \to \infty} E|Y_n|I\{|Y_n| > \alpha\} = 0 \quad \text{uniformly in } n \tag{4.1}$$

(see Section A.1).

In Section 1.7 we shall prove that we have moment convergence in the strong laws of large numbers and the central limit theorem for stopped random walks. Since in our applications the random indices are stopping times, we present our results in this framework. Without the stopping time assumption one needs higher moments for the summands (see Lai (1975) and Chow, Hsiung and Lai (1979)).

Before we present those results we show in this section that we have moment convergence in the strong law of large numbers and in the central

limit theorem in the classical setting under appropriate moment assumptions. In Sections 1.5 and 1.6 we present some relations between the existence of moments and properties of uniform integrability for stopping times and stopped random walks.

Whereas the classical convergence results themselves are proved in most textbooks, results about moment convergence are not. The ideas and details of the proofs of the results in Section 1.6 and 1.7 are best understood before the complication of the index being random appears.

The Strong Law

Theorem 4.1. *Let* $\{X_k, k \geq 1\}$ *be i.i.d. random variables such that* $E|X_1|^r < \infty$ *for some* $r \geq 1$ *and set* $S_n = \Sigma_{k=1}^n X_k$, $n \geq 1$. *Then*

$$\frac{S_n}{n} \to EX_1 \quad \text{a.s. and in } L^r \quad \text{as} \quad n \to \infty. \tag{4.2}$$

Proof. The a.s. convergence is well known. To prove L^r-convergence we shall use Lemma A.1.1 to show that

$$\left\{ \left| \frac{S_n}{n} \right|^r, n \geq 1 \right\} \quad \text{is uniformly integrable,} \tag{4.3}$$

from which the conclusion follows by Theorem A.1.1.

Let $x_1, \ldots, x_n$ be positive reals. By convexity we have

$$\left(\frac{1}{n} \sum_{k=1}^n x_k \right)^r \leq \frac{1}{n} \sum_{k=1}^n x_k^r. \tag{4.4}$$

Thus,

$$E \left| \frac{S_n}{n} \right|^r \leq E \left(\frac{1}{n} \sum_{k=1}^n |X_k| \right)^r \leq E \frac{1}{n} \sum_{k=1}^n |X_k|^r = E|X_1|^r, \tag{4.5}$$

which proves uniform boundedness of $\{E|S_n/n|^r, \; n \geq 1\}$, that is condition (i) in Lemma A.1.1 is satisfied.

To prove (ii), we observe that for any $\varepsilon > 0$ there exists $\delta > 0$ such that $E|X_1|^r I\{A\} < \varepsilon$ for all A with $P(A) < \delta$. Let A be an arbitrary set of this kind. Then

$$E \left| \frac{S_n}{n} \right|^r I\{A\} \leq E \frac{1}{n} \sum_{k=1}^n |X_k|^r I\{A\} < \varepsilon \tag{4.6}$$

independently of n. Thus (ii) is also satisfied and the proof is complete. $\square$

The Central Limit Theorem

Theorem 4.2. *Let $\{X_k, k \geq 1\}$ be i.i.d. random variables and set $S_n = \Sigma_{k=1}^n X_k$, $n \geq 1$. Assume that $EX_1 = 0$, that $\mathrm{Var}\ X_1 = \sigma^2 < \infty$ and that $E|X_1|^r < \infty$ for some $r \geq 2$. Then*

$$E\left|\frac{S_n}{\sqrt{n}}\right|^p \to E|Z|^p \quad as \quad n \to \infty \quad for\ all\ p, 0 < p \leq r, \qquad (4.7)$$

where Z is a normal random variable with mean 0 and variance σ^2.

Proof. Since $S_n/\sqrt{n}$ converges to Z in distribution it remains, in view of Theorem A.1.1 and Remark A.1.1, to show that

$$\left\{\left|\frac{S_n}{\sqrt{n}}\right|^r, n \geq 1\right\} \quad \text{is uniformly integrable.} \qquad (4.8)$$

We follow Billingsley (1968), p. 176, where the case $r = 2$ is treated.

Let $\varepsilon > 0$ and choose $M > 0$ so large that $E|X_1|^r I\{|X_1| > M\} < \varepsilon$. Set, for $k \geq 1$ and $n \geq 1$, respectively,

$$X_k' = X_k I\{|X_k| \leq M\} - EX_k I\{|X_k| \leq M\}, \quad X_k'' = X_k - X_k',$$

$$S_n' = \sum_{k=1}^n X_k' \quad \text{and} \quad S_n'' = \sum_{k=1}^n X_k''. \qquad (4.9)$$

Note that $EX_k' = EX_k'' = 0$.

(i)

$$E\left|\frac{S_n'}{\sqrt{n}}\right|^r I\left\{\left|\frac{S_n'}{\sqrt{n}}\right| > \alpha\right\} \leq \frac{B_{2r}(2M)^{2r}}{\alpha^r}. \qquad (4.10)$$

To see this, let U be a positive random variable. Then

$$EU^r I\{U > \alpha\} = \int_\alpha^\infty u^r dF_U(u) \leq \alpha^{-r} \int_\alpha^\infty u^{2r} dF_U(u) \leq \alpha^{-r} EU^{2r}. \quad (4.11)$$

Thus,

$$E\left|\frac{S_n'}{\sqrt{n}}\right|^r I\left\{\left|\frac{S_n'}{\sqrt{n}}\right| > \alpha\right\} \leq \alpha^{-r} E\left|\frac{S_n'}{\sqrt{n}}\right|^{2r} = (\alpha n)^{-r} E|S_n'|^{2r}. \qquad (4.12)$$

Now, since $\{S_n', n \geq 1\}$ are sums of uniformly bounded i.i.d. random variables with mean 0 we can apply the Marcinkiewicz–Zygmund inequality as given in formula (A.2.3) of any order; we choose to apply it to the order $2r$. It follows that

$$E|S_n'|^{2r} \leq B_{2r} n^r E|X_1'|^{2r} \leq B_{2r} n^r (2M)^{2r}, \qquad (4.13)$$

which, together with (4.12), yields (4.10) and thus step (i) has been proved.

(ii)

$$E\left|\frac{S_n''}{\sqrt{n}}\right|^r I\left\{\left|\frac{S_n''}{\sqrt{n}}\right| > \alpha\right\} < B_r \cdot 2^r \varepsilon. \tag{4.14}$$

Here we apply the inequality (A.2.3) of order r (≥ 2) to obtain

$$E\left|\frac{S_n''}{\sqrt{n}}\right|^r I\left\{\left|\frac{S_n''}{\sqrt{n}}\right| > \alpha\right\} \leq E\left|\frac{S_n''}{\sqrt{n}}\right|^r$$

$$\leq n^{-r/2} \cdot B_r \cdot E|X_1''|^r \cdot n^{r/2} < B_r \cdot 2^r \varepsilon,$$

which proves (4.14) and thus step (ii).

(iii)

$$E\left|\frac{S_n}{\sqrt{n}}\right|^r I\left\{\left|\frac{S_n}{\sqrt{n}}\right| > 2\alpha\right\} < \delta \quad \text{for any } \delta > 0 \quad \text{provided } \alpha > \alpha_0(\delta). \tag{4.15}$$

To see this, let $\delta > 0$ be given. By the triangle inequality, Lemma A.1.2, (i) and (ii) we obtain

$$E\left|\frac{S_n}{\sqrt{n}}\right|^r I\left\{\left|\frac{S_n}{\sqrt{n}}\right| > 2\alpha\right\}$$

$$\leq E\left(\left|\frac{S_n'}{\sqrt{n}}\right| + \left|\frac{S_n''}{\sqrt{n}}\right|\right)^r I\left\{\left|\frac{S_n'}{\sqrt{n}}\right| + \left|\frac{S_n''}{\sqrt{n}}\right| > 2\alpha\right\}$$

$$\leq 2^r E\left|\frac{S_n'}{\sqrt{n}}\right|^r I\left\{\left|\frac{S_n'}{\sqrt{n}}\right| > \alpha\right\} + 2^r E\left|\frac{S_n''}{\sqrt{n}}\right|^r I\left\{\left|\frac{S_n''}{\sqrt{n}}\right| > \alpha\right\}$$

$$< \left(\frac{2}{\alpha}\right)^r \cdot B_{2r}(2M)^{2r} + 2^{2r}\varepsilon B_r < \delta$$

provided we first choose M so large that ε is so small that $2^{2r}\varepsilon B_r < \delta/2$ and then α_0 so large that $(2/\alpha_0)^r B_{2r}(2M)^{2r} < \delta/2$. $\qquad\square$

We remark here that alternative proofs of Theorem 4.1 can be obtained by martingale methods (see Problem 11.4) and by the method used to prove Theorem 4.2. The latter method can also be modified in such a way that a corresponding result is obtained for the Marcinkiewicz–Zygmund law, a result which was first proved by Pyke and Root (1968) by a different method.

1.5 Moment Inequalities

In the proof of Theorem 4.2 the Marcinkiewicz–Zygmund inequalities, as given in formula (A.2.3) played an important role. These inequalities provide estimates of the moments of S_n in terms of powers of n. In order to

prove "Anscombe versions" of Theorem 4.1 and 4.2 we need generalizations of the Marcinkiewicz–Zygmund inequalities, which apply to stopped sums. These inequalities will then provide estimates of the moments of $S_{N(t)}$ in terms of moments of $N(t)$ (and X_1). The proofs of them are based on inequalities by Burkholder and Davis given in Appendix A; Theorems A.2.2 and A.2.3.

In this section we consider a random walk $\{S_n, n \geq 0\}$, with $S_0 = 0$ and i.i.d. increments $\{X_k, k \geq 1\}$. We further are given a *stopping time*, N, with respect to an increasing sequence of sub-σ-algebras, $\{\mathscr{F}_n, n \geq 1\}$, such that X_n is $\mathscr{F}_n$-measurable and independent of $\mathscr{F}_{n-1}$ for all n, that is, for every $n \geq 1$, we have

$$\{N = n\} \in \mathscr{F}_n. \tag{5.1}$$

The standard case is when $\mathscr{F}_n = \sigma\{X_1, \ldots, X_n\}$ and $\mathscr{F}_0 = \{\emptyset, \Omega\}$ (cf. Appendix A).

Theorem 5.1. *Suppose that $E|X_1|^r < \infty$ for some r $(0 < r < \infty)$ and that $EX_1 = 0$ when $r \geq 1$. Then*

(i) $E|S_N|^r \leq E|X_1|^r \cdot EN$ *for* $0 < r \leq 1$;

(ii) $E|S_N|^r \leq B_r \cdot E|X_1|^r \cdot EN$ *for* $1 \leq r \leq 2$;

(iii) $E|S_N|^r \leq B_r((EX_1^2)^{r/2} \cdot EN^{r/2} + E|X_1|^r \cdot EN)$

$\leq 2B_r \cdot E|X_1|^r \cdot EN^{r/2}$ *for* $r \geq 2$,

where B_r is a numerical constant depending on r only.

Proof. (i) Although we shall not apply this case we present a proof because of its simplicity and in order to make the result complete.

Define the bounded stopping time

$$N_n = N \wedge n. \tag{5.2}$$

By the c_r-inequality (Loève (1977), p. 157 or Gut (2007), Theorem 2.2.2) we obtain

$$E|S_{N_n}|^r \leq E \sum_{k=1}^{N_n} |X_k|^r. \tag{5.3}$$

Now, $\{|X_k|^r - E|X_k|^r, k \geq 1\}$ is a sequence of i.i.d. random variables with mean 0 and, by Doob's optional sampling theorem (Theorem A.2.4, in particular formula (A.2.7)), we therefore have

$$E \sum_{k=1}^{N_n} (|X_k|^r - E|X_k|^r) = 0 \tag{5.4}$$

or, equivalently,

$$E \sum_{k=1}^{N_n} |X_k|^r = EN_n \cdot E|X_1|^r. \tag{5.5}$$

By inserting this into (5.3) we obtain

$$E|S_{N_n}|^r \leq EN_n \cdot E|X_1|^r \leq EN \cdot E|X_1|^r. \tag{5.6}$$

An application of Fatou's lemma (cf. e.g. Gut (2007), Theorem 2.5.2) now yields (i).

(ii) Define N_n as in (5.2). By Theorem A.2.2, the c_r-inequality and (5.5) (which is valid for all $r > 0$) we have

$$E|S_{N_n}|^r \leq B_r \cdot E\left(\sum_{k=1}^{N_n} X_k^2\right)^{r/2} \leq B_r \cdot E\sum_{k=1}^{N_n} |X_k|^r = B_r \cdot EN_n \cdot E|X_1|^r$$

$$\leq B_r \cdot EN \cdot E|X_1|^r$$

and, again, an application of Fatou's lemma concludes the proof.

(iii) We proceed similarly, but with Theorem A.2.3 instead of Theorem A.2.2. It follows that

$$E|S_{N_n}|^r \leq B_r \cdot E\left(\sum_{k=1}^{N_n} E(X_k^2|\mathscr{F}_{k-1})\right)^{r/2} + B_r \cdot E\left(\sup_{1\leq k\leq N_n} |X_k|^r\right). \tag{5.7}$$

By independence we have

$$E\left(\sum_{k=1}^{N_n} E(X_k^2|\mathscr{F}_{k-1})\right)^{r/2} = E\left(\sum_{k=1}^{N_n} EX_k^2\right)^{r/2} = EN_n^{r/2} \cdot (EX_1^2)^{r/2}$$

$$\leq EN^{r/2}(EX_1^2)^{r/2} \tag{5.8}$$

and, by (5.5), it follows that

$$E\left(\sup_{1\leq k\leq N_n} |X_k|^r\right) \leq E\left(\sum_{k=1}^{N_n} |X_k|^r\right) = EN_n \cdot E|X_1|^r \leq EN \cdot E|X_1|^r. \tag{5.9}$$

By combining (5.7), (5.8) and (5.9) and Fatou's lemma we obtain the first inequality in (iii). The other follows from Lyapounov's inequality (see e.g. Gut (2007), Theorem 3.2.5) and the fact that $N \geq 1$ (which yields $N \leq N^{r/2}$). $\square$

Remark 5.1. For $r \geq 1$ this result has been proved for first passage times in Gut (1974a,b). The above proof follows Gut (1974b).

Remark 5.2. For $r = 1$ and $r = 2$ there is an overlap; this makes later reference easier.

Remark 5.3. Note that inequalities (A.2.3) are a special case of the theorem.
If $r \geq 1$ and we do not assume that $EX_1 \ (= \mu) = 0$ we have

$$|S_N| \leq |S_N - N\mu| + |\mu|N \tag{5.10}$$

and we can combine Theorem 5.1(ii) and (iii) and the c_r-inequalities to obtain

$$E|S_N|^r \leq 2^{r-1}(2B_r \cdot E|X_1|^r \cdot EN^{(r/2)\vee 1} + |\mu|^r \cdot EN^r)$$

$$\leq 2^{r-1}(2B_r \cdot E|X_1|^r + |\mu|^r)EN^r.$$

Since $|\mu|^r \leq E|X_1|^r$ the following result emerges.

Theorem 5.2. *Suppose that $E|X_1|^r < \infty$ for some r ($1 \leq r < \infty$). There exists a numerical constant B'_r depending on r only such that*

$$E|S_N|^r \leq B'_r \cdot E|X_1|^r \cdot EN^r. \tag{5.11}$$

Remark 5.4. In words this result means that if X_1 and N have finite moments of order $r \geq 1$ then so has S_N. However, if $EX_1 = 0$, then Theorem 5.1 tells us that the weaker condition $EN^{(r/2)\vee 1} < \infty$ suffices.

Just as in the context of sequential analysis one can obtain equalities for the first and second moments of the stopped sums.

Theorem 5.3.

(i) *If $EX_1 = \mu$ and $EN < \infty$, then*

$$ES_N = \mu \cdot EN. \tag{5.12}$$

(ii) *If, furthermore, $\sigma^2 = Var\, X_1 < \infty$, then*

$$E(S_N - N\mu)^2 = \sigma^2 \cdot EN. \tag{5.13}$$

Proof. (i) Again, let $N_n = N \wedge n$. Since

$$\{(S_n - n\mu, \mathscr{F}_n), n \geq 1\} \quad \text{is a martingale,} \tag{5.14}$$

it follows from the optional sampling theorem (Theorem A.2.4) that

$$ES_{N_n} = \mu \cdot EN_n. \tag{5.15}$$

Now, let $n \to \infty$. Since $N_n \to N$ monotonically, it follows from the monotone convergence theorem (cf. e.g. Gut (2007), Theorem 2.5.1) that

$$EN_n \to EN \quad \text{as} \quad n \to \infty, \tag{5.16}$$

which, together with (5.15), implies that

$$ES_{N_n} \to \mu \cdot EN < \infty \quad \text{as} \quad n \to \infty. \tag{5.17}$$

Moreover, $S_{N_n} \xrightarrow{\text{a.s.}} S_N$ as $n \to \infty$ (Theorem 2.4) and

$$|S_{N_n}| \leq \sum_{k=1}^{N_n} |X_k| \xrightarrow{\text{a.s.}} \sum_{k=1}^{N} |X_k| \quad \text{monotonically as } n \to \infty \tag{5.18}$$

(Theorem 2.4 again), which, together with (5.17) (applied to $\{|X_k|, k \geq 1\}$) and the montone convergence theorem, yields

$$E \sum_{k=1}^{N} |X_k| = EN \cdot E|X_1| < \infty. \tag{5.19}$$

An application of the dominated convergence theorem (cf. e.g. Gut (2007), Theorem 2.5.3) now shows that

$$ES_{N_n} \to ES_N \quad \text{as} \quad n \to \infty, \tag{5.20}$$

which, together with (5.15) and (5.16), proves (5.12).

(ii) Since

$$\{((S_n - n\mu)^2 - n\sigma^2, \mathscr{F}_n), n \geq 1\} \quad \text{is a martingale} \tag{5.21}$$

(cf. Problem A.5.7), we can apply the optional sampling theorem to obtain

$$E(S_{N_n} - \mu N_n)^2 = \sigma^2 \cdot EN_n. \tag{5.22}$$

Now, let $m \geq n$. Since martingales have orthogonal increments we have, in view of (5.22) and (5.16),

$$E(S_{N_m} - \mu N_m - (S_{N_n} - \mu N_n))^2$$
$$= E(S_{N_m} - \mu N_m)^2 - E(S_{N_n} - \mu N_n)^2 = \sigma^2 \cdot (EN_m - EN_n) \to 0$$

as $n, m \to \infty$, that is,

$$S_{N_n} - \mu N_n \quad \text{converges in } L^2 \quad \text{as} \quad n \to \infty. \tag{5.23}$$

However, since we already know that $S_{N_n} - \mu N_n \xrightarrow{\text{a.s.}} S_N - \mu N$ as $n \to \infty$, it follows that

$$E(S_{N_n} - \mu N_n)^2 \to E(S_N - \mu N)^2 \quad \text{as} \quad n \to \infty, \tag{5.24}$$

which, together with (5.22) and (5.16), concludes the proof. $\qquad\square$

Remark 5.5. The second half of Theorem 5.3 is due to Chow, Robbins and Teicher (1965). For a beautiful proof of (i) we refer to Blackwell (1946), Theorem 1 (see also De Groot (1986), pp. 42–43) and Problem 11.7 below.

Remark 5.6. It is also possible to prove the theorem by direct computations.

Remark 5.7. Just as variance formulas extend to covariance formulas it is possible to extend (5.13) as follows: Suppose that $\{X_k, k \geq 1\}$ and $\{Y_k, k \geq 1\}$ are two sequences of i.i.d. random variables and that $\{Y_k, k \geq 1\}$ has the same measurability properties as $\{X_k, k \geq 1\}$, described at the beginning of this

section. Suppose, further, that $EN < \infty$ and that both sequences have finite variances, σ_x^2 and σ_y^2, respectively and set $\mu_x = EX_1$ and $\mu_y = EY_1$. Then

$$E\left(\sum_{k=1}^{N} X_k - N\mu_x\right)\left(\sum_{k=1}^{N} Y_k - N\mu_y\right) = \text{Cov}(X_1, Y_1) \cdot EN. \qquad (5.25)$$

To prove this one uses either the relation $4xy = (x + y)^2 - (x - y)^2$ together with (5.13) or one proceeds as in the proof of Theorem 5.3(ii) with (5.21) replaced by

$$\left\{\left(\sum_{k=1}^{n}(X_k - \mu_x) \cdot \sum_{k=1}^{n}(Y_k - \mu_y) - n\text{Cov}(X_1, Y_1)\right), n \geq 1\right\} \text{ is a martingale.}$$

Remark 5.8. Note that $E(S_N - N\mu)^2$ is not the same as $\text{Var}(S_N) = E(S_N - EN \cdot \mu)^2$ when $\mu \neq 0$.

Similar equalities can also be obtained for higher moments, see Neveu (1975), Chapter IV. We also refer to Chow, Robbins and Siegmund (1971) for further results of this kind.

In sequential analysis one considers a stopping time, N, which is a first passage time out of a finite interval. One can show that N has finite moments of all orders. In contrast to this case, the following example shows that the stopped sum may have finite expectation whereas the stopping time has not.

Example 5.1. Consider the symmetric simple random walk, that is $\{S_n, n \geq 0\}$, where $S_0 = 0$, $S_n = \sum_{k=1}^{n} X_k$, $n \geq 1$, and $P(X_k = 1) = P(X_k = -1) = \frac{1}{2}$. Define

$$N_+ = \min\{n: S_n = +1\}. \qquad (5.26)$$

Clearly,

$$P(S_{N_+} = 1) = 1, \quad \text{in particular } ES_{N_+} = 1. \qquad (5.27)$$

Now, suppose that $EN_+ < \infty$. Then (5.12) holds, but, since $\mu = EX_1 = 0$, we have $1 = 0 \cdot EN_+$, which is a contradiction. Thus $EN_+ = +\infty$.

The fact that $EN_+ = +\infty$ is proved by direct computation of the distribution of N_+ in Feller (1968), Chapter XIII. Our proof is a nice alternative. (In fact N_+ has no moment of order $\geq \frac{1}{2}$.)

Note, however, that if we set $N_n = N_+ \wedge n$ in Example 5.1, then we have, of course, that $ES_{N_n} = 0$. Note also that the same arguments apply to $N_- = \min\{n: S_n = -1\}$.

Finally, set $N = N_+ \wedge N_-$. Then

$$P(N = 1) = 1 \quad \text{and thus } EN = 1. \qquad (5.28)$$

Furthermore,

$$P(|S_N| = 1) = 1. \qquad (5.29)$$

It follows from (5.12) that

$$ES_N = 1 \cdot 0 = 0 \tag{5.30}$$

as expected.

Converses

The above results and Example 5.1 lead to the question of possible converses for Theorems 5.1 and 5.2. The former states that

$$E|X_1|^r < \infty, EX_1 = 0 \quad \text{when } r \geq 1 \text{ and } EN^{(r/2)\vee 1} < \infty$$

$$\implies \quad E|S_N|^r < \infty \quad (r > 0), \tag{5.31}$$

and the latter states that

$$E|X_1|^r < \infty \quad \text{and} \quad EN^r < \infty \quad \implies \quad E|S_N|^r < \infty \quad (r \geq 1). \tag{5.32}$$

On the other hand, Example 5.1 shows that S_N may have moments of all orders without N even having finite expectation.

Let us first consider the case of positive random variables. The following result shows that the converse of (5.32) holds in this case.

Theorem 5.4. *Let $r \geq 1$ and suppose that $P(X_1 \geq 0) = 1$ and that $P(X_1 > 0) > 0$. Then*

$$ES_N^r < \infty \tag{5.33}$$

iff

$$EX_1^r < \infty \quad \text{and} \quad EN^r < \infty. \tag{5.34}$$

Proof. In view of Theorem 5.2 we have to prove that (5.33) $\implies$ (5.34). We proceed in steps.

$$\text{(i)} \quad EX_1^r < \infty \quad \text{and} \quad \mu = EX_1 < \infty$$

Since the summands are positive we have

$$S_N \geq X_1 \tag{5.35}$$

and thus

$$\infty > ES_N^r \geq EX_1^r. \tag{5.36}$$

If $r > 1$ we further note that, by Lyapounov's inequality,

$$\mu = EX_1 \leq \|X_1\|_r < \infty, \tag{5.37}$$

which completes the proof of (i).

$$\text{(ii)} \quad EN < \infty$$

Since $\mu < \infty$ we can apply the optional sampling theorem, in particular formula (5.15), to obtain

$$ES_{N_n} = \mu EN_n. \tag{5.38}$$

Furthermore, $S_{N_n} \to S_N$ as $n \to \infty$ and it follows from monotone convergence that

$$\lim_{n \to \infty} ES_{N_n} = ES_N \le \|S_N\|_r < \infty, \tag{5.39}$$

which, by monotone convergence, proves (ii).

Next we note that (cf. (5.10))

$$\mu N \le |S_N - N\mu| + S_N. \tag{5.40}$$

$$\text{(iiia)} \quad E|S_N - N\mu|^r < \infty$$

$$\text{(iiib)} \quad EN^r < \infty$$

To prove this we proceed by induction on r through the powers of 2 (cf. Gut (1974a,b)).

Let $1 < r \le 2$. By Theorem 5.1(ii) we have

$$E|S_N - N\mu|^r \le B_r \cdot E|X_1 - \mu|^r \cdot EN, \tag{5.41}$$

which is finite by step (ii), so (iiia) holds for this case, from which (iiib) follows by (5.40) and Minkowski's inequality or the c_r-inequality.

Next, suppose that $2 < r \le 2^2$. Since $1 < r/2 \le 2$ we know from what has just been proved that (iiib) holds with r replaced by $r/2$. This, together with Theorem 5.1(iii), shows that

$$E|S_N - N\mu|^r \le 2B_r \cdot E|X_1 - \mu|^r \cdot EN^{r/2} < \infty, \tag{5.42}$$

and another application of (5.40) shows that $EN^r < \infty$. Thus (iiia) and (iiib) hold again.

In general, if $2^k < r \le 2^{k+1}$ for some $k > 2$ we repeat the same procedure from $r/2^k$ to $r/2^{k-1}$ etc. until r is reached and the conclusion follows.   □

Now consider a general random walk and suppose that $EX_1 > 0$. In Chapter 3 we shall study first passage times across horizontal barriers for such random walks and it will follow that the moments of the stopped sum are linked to the moments of the positive part of the summands (Theorem 3.3.1(ii)). Thus, a general converse to Theorem 5.2 cannot hold. The following result is, however, true.

Theorem 5.5. *Suppose that $EX_1 \ne 0$. If, for some $r \ge 1$,*

$$E|X_1|^r < \infty \quad and \quad E|S_N|^r < \infty \tag{5.43}$$

then

$$EN^r < \infty. \tag{5.44}$$

Proof. Suppose, without restriction, that $\mu = EX_1 > 0$.

$$\text{(i)} \quad EN < \infty$$

To prove this we shall use a trick due to Blackwell (1953), who used it in the context of ladder variables. (The trick is, in fact, a variation of the proof of Theorem 5.3(i) given in Blackwell (1946).) Let $\{N_k, k \geq 1\}$ be independent copies of N, constructed as follows: Let $N_1 = N$. Restart after N_1, i.e. consider the sequence $X_{N_1+1}, X_{N_1+2}, \ldots$, and let N_2 be a stopping time for this sequence. Restart after $N_1 + N_2$ to obtain N_3, and so on. Thus, $\{N_k, k \geq 1\}$ is a sequence of i.i.d. random variables distributed as N, and $\{S_{N_1+\cdots+N_k}, k \geq 1\}$ is a sequence of partial sums of i.i.d. random variables distributed as S_N and, by assumption, with finite mean, ES_N.

Now

$$\frac{N_1 + \cdots + N_k}{k} = \frac{N_1 + \cdots + N_k}{S_{N_1+\cdots+N_k}} \cdot \frac{S_{N_1+\cdots+N_k}}{k}. \tag{5.45}$$

Clearly $N_1 + \cdots + N_k \to +\infty$ as $k \to \infty$. Thus, by the strong law of large numbers and Theorem 2.3 it follows that

$$\frac{S_{N_1+\cdots+N_k}}{N_1 + \cdots + N_k} \xrightarrow{\text{a.s.}} \mu \quad \text{as} \quad k \to \infty \tag{5.46}$$

and that

$$\frac{S_{N_1+\cdots+N_k}}{k} \xrightarrow{\text{a.s.}} ES_N \quad \text{as} \quad k \to \infty. \tag{5.47}$$

Consequently,

$$\frac{N_1 + \cdots + N_k}{k} \xrightarrow{\text{a.s.}} \mu^{-1} \cdot ES_N \quad \text{as} \quad k \to \infty, \tag{5.48}$$

from which it follows that

$$EN = EN_1 < \infty \tag{5.49}$$

by the converse of the Kolmogorov strong law of large numbers (cf. e.g. Gut (2007), Theorem 6.6.1(b)) . Thus (i) holds.

$$\text{(ii)} \quad EN^r < \infty$$

Since the summands may take negative values we first have to replace (5.40) by

$$\mu N \leq |S_N - N\mu| + |S_N|. \tag{5.50}$$

An inspection of the proof of Theorem 5.4 shows that the positivity there was only used to conclude that the summands had a finite moment of order r and that the stopping time had finite expectation. Now, in the present result the first fact was assumed and the second fact has been proved in step (i). Therefore, step (iii) from the previous proof, with (5.40) replaced by (5.50) carries over verbatim to the present theorem. We can thus conclude that (5.44) holds and the proof is complete. $\qquad\square$

The case $E|S_N|^r < \infty$ and $EX_1 = 0$ has been dealt with in Example 5.1, where it was shown that no converse was possible.

Before closing we wish to point to the fact that the hardest part in the proofs of Theorems 5.4 and 5.5 in some sense was to show that $EN < \infty$, because once this was done the fact that $EN^r < \infty$ followed from Theorem 5.1, that is, once we knew that the first moment was finite, we could, more or less automatically, conclude that a higher moment was finite. For some further results of the above kind as well as for one sided versions we refer to Gut and Janson (1986).

1.6 Uniform Integrability

In Section 1.6–1.8 we consider a random walk $\{S_n, n \geq 0\}$ with i.i.d. increments $\{X_k, k \geq 1\}$ as before, but now we consider a *family of stopping times*, $\{N(t), t \geq 0\}$, with respect to $\{\mathscr{F}_n, n \geq 0\}$ as given in Section 1.5.

Before we proceed two remarks are in order. Firstly, the results below also hold for sequences of stopping times, for example, $\{N(n), n \geq 1\}$. Secondly, to avoid certain trivial technical problems we only consider families of stopping times for values of $t \geq$ some t_0, where, for convenience, we assume that $t_0 \geq 1$.

The following theorem is due to Lai (1975).

Theorem 6.1. *Let $r \geq 1$ and suppose that $E|X_1|^r < \infty$. If, for some $t_0 \geq 1$, the family*

$$\left\{ \left(\frac{N(t)}{t} \right)^r, t \geq t_0 \right\} \quad \text{is uniformly integrable,} \tag{6.1}$$

then the family

$$\left\{ \left| \frac{S_{N(t)}}{t} \right|^r, t \geq t_0 \right\} \quad \text{is uniformly integrable.} \tag{6.2}$$

Proof. As in the proof of Theorem 4.2 we let $\varepsilon > 0$ and choose M so large that $E|X_1|^r I\{|X_1| > M\} < \varepsilon$. Define the truncated variables without centering, that is, set

$$X_k' = X_k I\{|X_k| \leq M\}, \qquad X_k'' = X_k I\{|X_k| > M\} \quad (k \geq 1),$$

$$S_n' = \sum_{k=1}^{n} X_k', \qquad S_n'' = \sum_{k=1}^{n} X_k'' \quad (n \geq 1). \tag{6.3}$$

(i) $E \left| \frac{S_{N(t)}'}{t} \right|^r I\left\{ \left| \frac{S_{N(t)}'}{t} \right| > \alpha \right\} \leq M^r E \left(\frac{N(t)}{t} \right)^r I\left\{ \frac{N(t)}{t} > \alpha/M \right\}.$

Since $|X'_k| \leq M$ for all $k \geq 1$ it follows that the arithmetic means are bounded by M, in particular, we have

$$\left| \frac{S'_{N(t)}}{N(t)} \right| \leq M \tag{6.4}$$

and thus, that

$$\left| \frac{S'_{N(t)}}{t} \right| = \left| \frac{S'_{N(t)}}{N(t)} \right| \cdot \left| \frac{N(t)}{t} \right| \leq M \frac{N(t)}{t}, \tag{6.5}$$

from which (i) follows.

For the other sum we proceed as in the proof of Theorem 4.2, but with (the Burkholder inequalities given in) Theorem 5.2 replacing the inequalities from (A.2.3).

$$\text{(ii)} \qquad E \left| \frac{S''_{N(t)}}{t} \right|^r I \left\{ \left| \frac{S''_{N(t)}}{t} \right| > \alpha \right\} < B'_r \varepsilon E \left(\frac{N(t)}{t} \right)^r.$$

By Theorem 5.2 we have

$$E|S''_{N(t)}|^r \leq B'_r E|X''_1|^r E(N(t))^r < B'_r \varepsilon E(N(t))^r, \tag{6.6}$$

from which (ii) follows.

(iii) (6.2) holds.

This last step is now achieved by combining (i), (ii) and Lemma A.1.2 as in the proof of Theorem 4.2. It follows that for any given $\delta > 0$ we have

$$E \left| \frac{S_{N(t)}}{t} \right|^r I \left\{ \left| \frac{S_{N(t)}}{t} \right| > 2\alpha \right\}$$

$$< 2^r \left(M^r E \left(\frac{N(t)}{t} \right)^r I \left\{ \frac{N(t)}{t} > \alpha/M \right\} + B'_r \varepsilon E \left(\frac{N(t)}{t} \right)^r \right) < \delta, \tag{6.7}$$

provided we first choose M so large that ε is small enough to ensure that $2^r B'_r \varepsilon E(N(t)/t)^r < \delta/2$ (this is possible, because $E(N(t)/t)^r \leq$ constant, uniformly in t) and then α so large that

$$(2M)^r E \left(\frac{N(t)}{t} \right)^r I \left\{ \frac{N(t)}{t} > \alpha/M \right\} < \delta/2$$

(which is possible by (6.1)).   □

In the proof we used Theorem 5.2, that is, we estimated moments of $S_{N(t)}$ by moments of $N(t)$ of *the same order*. Since there is the sharper result, Theorem 5.1, for the case $EX_1 = 0$, there is reason to believe that we can obtain a better result for that case here too. The following theorems show that

this is indeed the case. Also, just as Theorem 6.1 is related to the strong law of large numbers without any assumption about the mean, Theorem 6.4 below corresponds to the case $EX_1 = 0$. Further, Theorem 6.2 (with $0 < r < 2$) and Theorem 6.3 are related to the Marcinkiewicz–Zygmund law and the central limit theorem, respectively.

As for the question of the necessity of our assumptions we shall discuss this in the later half of this section.

Theorem 6.2. *Let $0 < r \leq 2$. Suppose that $E|X_1|^r < \infty$ and that $EX_1 = 0$ when $r \geq 1$. If, for some $t_0 \geq 1$, the family*

$$\left\{ \left(\frac{N(t)}{t} \right), t \geq t_0 \right\} \quad \text{is uniformly integrable,} \tag{6.8}$$

then the family

$$\left\{ \left| \frac{S_{N(t)}}{t^{1/r}} \right|^r, t \geq t_0 \right\} \quad \text{is uniformly integrable.} \tag{6.9}$$

Proof.
$$\text{The Case } 0 < r < 1$$

Let ε and M be given as before. We use the (noncentered) truncation from the proof of Theorem 6.1 (cf. (6.3)).

(i)
$$E \left| \frac{S'_{N(t)}}{t^{1/r}} \right|^r I \left\{ \left| \frac{S'_{N(t)}}{t^{1/r}} \right|^r > \alpha \right\} \leq M^r E \left(\frac{N(t)}{t} \right) I \left\{ \frac{N(t)}{t} > \alpha/M^r \right\}.$$

Since $|S'_{N(t)}| \leq N(t) \cdot M$ it follows that

$$\left| \frac{S'_{N(t)}}{t^{1/r}} \right|^r \leq M^r \frac{(N(t))^r}{t} \leq M^r \frac{N(t)}{t} \tag{6.10}$$

and (i) follows.

(ii)
$$E \left| \frac{S''_{N(t)}}{t^{1/r}} \right|^r I \left\{ \left| \frac{S''_{N(t)}}{t^{1/r}} \right|^r > \alpha \right\} \leq E \left| \frac{S''_{N(t)}}{t^{1/r}} \right|^r < \varepsilon E \frac{N(t)}{t}.$$

By Theorem 5.1(i) we have

$$E|S''_{N(t)}|^r \leq EN(t)E|X''_1|^r < \varepsilon EN(t), \tag{6.11}$$

which yields (ii).

The conclusion now follows from (i), (ii) and Lemma A.1.2 as before.

The Case $1 \le r \le 2$

Here the situation is more delicate in that we also have to truncate the stopping times. We follow the truncation argument in Yu (1979).

Let ε and M be given as before. This time we use the (centered) truncation defined in (4.9). Further, set, for some $A \ge 1$,

$$N'(t) = \min\{N(t), A_t\}, \quad \text{where } A_t = [At]. \tag{6.12}$$

We first observe that $N'(t)$ is a stopping time and that

$$|S_{N(t)}| \le |S'_{N(t)}| + |S'_{N(t)} - S'_{N'(t)}| + |S''_{N(t)}|. \tag{6.13}$$

(iii)

$$E \left| \frac{S'_{N'(t)}}{t^{1/r}} \right|^r I\left\{ \left| \frac{S'_{N'(t)}}{t^{1/r}} \right| > \alpha \right\} \le \alpha^{-r} \cdot 2B_{2r}(2M)^{2r} A^r.$$

By arguing as in the proof of (4.10) we first have

$$E \left| \frac{S'_{N'(t)}}{t^{1/r}} \right|^r I\left\{ \left| \frac{S'_{N'(t)}}{t^{1/r}} \right| > \alpha \right\} \le \alpha^{-r} \cdot E \left| \frac{S'_{N'(t)}}{t^{1/r}} \right|^{2r}. \tag{6.14}$$

Furthermore, by Theorem 5.1(iii)

$$E|S'_{N'(t)}|^{2r} \le 2B_{2r} \cdot E|X'_1|^{2r} \cdot E(N'(t))^r \le 2B_{2r}(2M)^{2r}(At)^r, \tag{6.15}$$

which together with (6.14) yields (iii).

(iv)

$$E \left| \frac{S'_{N(t)} - S'_{N'(t)}}{t^{1/r}} \right|^r I\left\{ \left| \frac{S'_{N(t)} - S'_{N'(t)}}{t^{1/r}} \right| > \alpha \right\}$$

$$\le B_r(2M)^r E\left(\frac{N(t)}{t} \right) I\left\{ \frac{N(t)}{t} > A - 1 \right\}.$$

Observe that, for each fixed t, $N(t) - N'(t) = (N(t) - A_t)^+$ is a stopping time relative to $\{\mathscr{F}_{n+A_t}, n \ge 0\}$, that is,

$$\{N(t) - A_t = n\} \in \mathscr{F}_{n+A_t} \quad \text{for all } n \ge 0. \tag{6.16}$$

An application of Theorem 5.1(ii) therefore yields

$$E|S'_{N(t)} - S'_{N'(t)}|^r = E \left| \sum_{n=1}^{(N(t)-A_t)^+} X'_{n+A_t} \right|^r$$

$$\le B_r E|X'_1|^r E(N(t) - A_t)^+$$

$$\leq B_r(2M)^r E(N(t) - A_t)I\{N(t) > A_t\}$$

$$\leq B_r(2M)^r EN(t)I\{N(t) > (A - 1)t\},$$

from which (iv) follows.

(v)
$$E\left|\frac{S''_{N(t)}}{t^{1/r}}\right|^r I\left\{\left|\frac{S''_{N(t)}}{t^{1/r}}\right| > \alpha\right\} < B_r \cdot 2^r \varepsilon E\frac{N(t)}{t}.$$

By Theorem 5.1(ii) we have

$$E|S''_{N(t)}|^r \leq B_r \cdot E|X''_1|^r \cdot EN(t) < B_r \cdot 2^r \varepsilon EN(t), \qquad (6.17)$$

from which (v) follows.

(vi) (6.9) holds.

By generalizing Lemma A.1.2 to 3 random variables we note that 2^r is replaced by 3^r and $\alpha/2$ by $\alpha/3$. Thus, by combining (iii), (iv), and (v) with this extension we obtain

$$E\left|\frac{S_{N(t)}}{t^{1/r}}\right|^r I\left\{\left|\frac{S_{N(t)}}{t^{1/r}}\right| > 3\alpha\right\} < 3^r\left(\alpha^{-r}2B_{2r}(2M)^{2r}A^r\right.$$

$$+B_r(2M)^r E\left(\frac{N(t)}{t}\right)I\left\{\frac{N(t)}{t} > A - 1\right\} + 2^r B_r \varepsilon E\frac{N(t)}{t}\right). \qquad (6.18)$$

In order to make this smaller than a given small δ we first choose M so large that ε is small enough to ensure that $3^r 2^r B_r \varepsilon E(N(t)/t) < \delta/3$. We then choose A large enough to ensure that $3^r B_r(2M)^r E(N(t)/t)I \{(N(t)/t) > A - 1\} < \delta/3$ and finally α so large that $3^r \alpha^{-r} 2B_{2r}(2M)^{2r}A^r < \delta/3$. This proves the theorem. □

Theorem 6.3. *Let $r \geq 2$. Suppose that $E|X_1|^r < \infty$ and that $EX_1 = 0$. If, for some $t_0 \geq 1$, the family*

$$\left\{\left(\frac{N(t)}{t}\right)^{r/2}, t \geq t_0\right\} \quad \text{is uniformly integrable,} \qquad (6.19)$$

then the family

$$\left\{\left|\frac{S_{N(t)}}{\sqrt{t}}\right|^r, t \geq t_0\right\} \quad \text{is uniformly integrable.} \qquad (6.20)$$

Proof. We can use most of the estimates obtained in the proof of Theorem 6.2 for the case $1 \leq r \leq 2$. Let ε and M be given as before and define the (centered) truncated variables X'_n, X''_n etc. as before. Also, define $N'(t) = \min\{N(t), A_t\}$ for some $A \geq 1$, where $A_t = [At]$ (cf. (6.12)).

Now $r \geq 2$ throughout. By proceeding as in the estimates used for steps (iii) and (iv) in the proof of Theorem 6.2 we obtain

(i)

$$E\left|\frac{S'_{N'(t)}}{\sqrt{t}}\right|^r I\left\{\left|\frac{S'_{N'(t)}}{\sqrt{t}}\right| > \alpha\right\} \le \alpha^{-r} \cdot E\left|\frac{S'_{N'(t)}}{\sqrt{t}}\right|^{2r} \le \alpha^{-r} \cdot 2B_{2r}(2M)^{2r}A^r$$

and

$$E|S'_{N(t)} - S'_{N'(t)}|^r \le 2B_r(2M)^r E(N(t) - A_t)^{r/2} I\{N(t) > (A-1)t\}$$

and thus that
(ii)

$$E\left|\frac{S'_{N(t)} - S'_{N'(t)}}{\sqrt{t}}\right|^r I\left\{\left|\frac{S'_{N(t)} - S'_{N'(t)}}{\sqrt{t}}\right| > \alpha\right\}$$

$$\le 2B_r(2M)^r E\left(\frac{N(t)}{t}\right)^{r/2} I\left\{\frac{N(t)}{t} > A-1\right\}.$$

Finally,
(iii)

$$E\left|\frac{S''_{N(t)}}{\sqrt{t}}\right|^r I\left\{\left|\frac{S''_{N(t)}}{\sqrt{t}}\right| > \alpha\right\} < 2^{r+1}B_r\varepsilon E\left(\frac{N(t)}{t}\right)^{r/2},$$

because, by Theorem 5.1(iii) we have

$$E|S''_{N(t)}|^r \le 2B_r E|X''_1|^r E(N(t))^{r/2} \le 2^{r+1}B_r\varepsilon E(N(t))^{r/2}, \qquad (6.21)$$

and (iii) follows the usual way.

By combining (i), (ii), (iii), and Lemma A.1.2 as in the proof of Theorem 6.2 we obtain (6.20) and the proof is complete. □

Remark 6.1. Theorems 6.2 and 6.3 overlap at $r = 2$, but, since $r/2 = 1$ then, they coincide at that point.

Remark 6.2. Theorems 6.3 is due to Yu (1979). Earlier Chow, Hsiung and Lai (1979) proved the result under the stronger assumption that

$$\left\{\left(\frac{N(t)}{t}\right)^r, t \ge t_0\right\} \quad \text{is uniformly integrable.} \qquad (6.22)$$

They also called their result an L^r-analogue of Anscombe's theorem (recall Theorem 3.1 above).

Chang and Hsiung (1979) use the truncation of Yu (1979) to prove Theorem 6.2.

We can now easily derive the stronger result for the strong law of large numbers when $EX_1 = 0$, which was alluded to in the discussion following Theorem 6.1 (recall Remark 5.4, where the corresponding situation was discussed concerning the existence of moments for the stopped random walk).

Theorem 6.4. *Let $r \geq 1$, suppose that $E|X_1|^r < \infty$ and that $EX_1 = 0$. If, for some $t_0 \geq 1$, the family*

$$\left\{ \left(\frac{N(t)}{t} \right)^{(r/2)\vee 1}, t \geq t_0 \right\} \quad \text{is uniformly integrable,} \qquad (6.23)$$

then the family

$$\left\{ \left| \frac{S_{N(t)}}{t} \right|^r, t \geq t_0 \right\} \quad \text{is uniformly integrable.} \qquad (6.24)$$

Proof. The proof follows immediately from Theorems 6.2 and 6.3 and the observations that

$$\left| \frac{S_{N(t)}}{t} \right|^r \leq \left| \frac{S_{N(t)}}{t^{1/r}} \right|^r \quad \text{for } r \geq 1 \qquad (6.25)$$

and

$$\left| \frac{S_{N(t)}}{t} \right|^r \leq \left| \frac{S_{N(t)}}{\sqrt{t}} \right|^r, \qquad (6.26)$$

by using (6.25) and Theorem 6.2 when $1 \leq r < 2$ and (6.26) and Theorem 6.3 when $r \geq 2$. □

Remark 6.3. If $\{N(t), t \geq 0\}$ are *not* stopping times, Lai (1975) shows that Theorem 6.1 holds under the additional assumption that $E|X_1|^{r+1} < \infty$. Also, Chow, Hsiung and Lai (1979) show that Theorem 6.3 holds (but under condition (6.22) instead of (6.19)) under the additional assumption that $E|X_1|^{r+1} < \infty$.

Converses

In Theorems 6.1–6.4 we made assumptions about the uniform integrability of $\{(N(t)/t)^s, t \geq t_0\}$ with $s = r, 1, r/2$ and $(r/2) \vee 1$ in order to conclude that families of suitably normalized stopped sums had certain degrees of uniform integrability.

In the immediate context in which we shall use these assumptions such conditions will be verified whenever required. In the remainder of this section we shall be concerned with their necessity (cf. Section 1.5). To this end, assume first that $\{X_k, k \geq 1\}$ is a sequence of *positive* i.i.d. random variables. The following converse of Theorem 6.1 then holds (cf. Theorem 5.4).

Theorem 6.5. *Let $r \geq 1$, suppose that $P(X_1 \geq 0) = 1$ and that $P(X_1 > 0) > 0$. If, for some $t_0 \geq 1$, the family*

$$\left\{ \left(\frac{S_{N(t)}}{t} \right)^r, t \geq t_0 \right\} \quad \text{is uniformly integrable,} \tag{6.27}$$

then the family

$$\left\{ \left(\frac{N(t)}{t} \right)^r, t \geq t_0 \right\} \quad \text{is uniformly integrable.} \tag{6.28}$$

Proof. The case $r = 1$ is the most difficult one (cf. the remarks at the end of Section 1.5) and for its proof we refer to Gut and Janson (1986). Thus, suppose that $r > 1$. Since (6.27), in particular, implies that

$$E \left(\frac{S_{N(t)}}{t} \right)^r \leq \text{constant, uniformly in } t \geq t_0, \tag{6.29}$$

it follows that $ES_{N(t)}^r < \infty$ for all t and thus, by Theorem (5.4) that $EX_1^r < \infty$ and $E(N(t))^r < \infty$. Set $\mu = EX_1$. By Theorem 5.1 we have, for $1 \leq r \leq 2$,

$$E \left| \frac{S_{N(t)} - N(t)\mu}{t} \right|^r \leq t^{1-r} \cdot B_r E |X_1 - \mu|^r \cdot E \frac{N(t)}{t}$$

$$\leq B_r \cdot E|X_1 - \mu|^r \cdot \mu^{-1} \cdot E \frac{S_{N(t)}}{t} \leq \text{constant} \tag{6.30}$$

and, for $r \geq 2$,

$$E \left| \frac{S_{N(t)} - N(t)\mu}{t} \right|^r \leq 2 B_r E |X_1 - \mu|^r E \left(\frac{N(t)}{t} \right)^{r/2}. \tag{6.31}$$

Now, by using (6.30) and (6.31) together with (5.38) we can rework the induction procedure of the proof of Theorem 5.4 (cf. step (iii) there) and conclude that

$$E \left(\frac{N(t)}{t} \right)^r \leq \text{constant, uniformly in } t \geq t_0. \tag{6.32}$$

(So far everything holds also for $r = 1$.)

Now, since $r > 1$, it follows from (6.32) that

$$\left\{ \left(\frac{N(t)}{t} \right)^{(r/2) \vee 1}, t \geq t_0 \right\} \quad \text{is uniformly integrable.} \tag{6.33}$$

We can thus apply Theorem 6.4 to the sequence $\{X_k - \mu, k \geq 1\}$ and conclude that

$$\left\{ \left| \frac{S_{N(t)} - N(t)\mu}{t} \right|^r, t \geq t_0 \right\} \quad \text{is uniformly integrable,} \tag{6.34}$$

which together with (6.27), (5.40) and Lemma A.1.3 proves the theorem (for $r > 1$). $\qquad\square$

Remark 6.4. The reason why the case $r > 1$ is easier is that we always have $(r/2) \vee 1 < r$ in that case, which makes the transition from (6.32) to (6.33) possible then.

For the general case with $EX_1 \neq 0$ we can modify Theorem 6.5 for $r > 1$ in precisely the same way as Theorem 5.5 was derived from Theorem 5.4, and just as for Theorem 6.5 the case $r = 1$ needs special care, see Gut and Janson (1986). The following result is obtained.

Theorem 6.6. *Suppose that $EX_1 \neq 0$. If, for some $r \geq 1$ and $t_0 \geq 1$,*

$$E|X_1|^r < \infty \tag{6.35}$$

and the family

$$\left\{ \left| \frac{S_{N(t)}}{t} \right|^r, t \geq t_0 \right\} \quad \text{is uniformly integrable,} \tag{6.36}$$

then the family

$$\left\{ \left(\frac{N(t)}{t} \right)^r, t \geq t_0 \right\} \quad \text{is uniformly integrable.} \tag{6.37}$$

Finally, suppose that $EX_1 = 0$. Recall from Example 5.1 that the situation for this case is completely different. By expanding that example a little it is easily seen that (6.36) may very well hold for all $r > 0$ and yet (6.37) does not hold for any $r \geq 1$ ($r \geq \frac{1}{2}$).

Example 6.1. We thus consider the example from Section 1.5, that is, let $\{X_k, k \geq 1\}$ be a sequence of i.i.d. random variables such that $P(X_k = 1) = P(X_k = -1) = \frac{1}{2}$ and set $S_n = \sum_{k=1}^{n} X_k$, $n \geq 0$. Define

$$N(t) = \min\{n: S_n \geq [t]\} = \min\{n: S_n = [t]\} \quad (t \geq 0). \tag{6.38}$$

Clearly, $S_{N(t)} = [t]$, and so

$$(0 \leq) \ \frac{S_{N(t)}}{t} = \frac{[t]}{t} \leq 1 \quad \text{for all } t > 0, \tag{6.39}$$

that is, (6.36) holds trivially for all $r > 0$. On the other hand, we know from random walk theory (see e.g. Example 5.1) that $E(N(0))^r = +\infty$ for all $r \geq 1$ and, since $N(t) \geq N(0)$, (6.37) cannot hold for any $r \geq 1$.

Remark 6.5. In the previous example we also observe that it follows from Theorem 2.3(iii) (since $N(t) \xrightarrow{\text{a.s.}} +\infty$ as $t \to \infty$) that (2.10) holds, that is, that

$$\frac{S_{N(t)}}{N(t)} \xrightarrow{\text{a.s.}} 0 \quad \text{as} \quad t \to \infty. \tag{6.40}$$

On the other hand, (2.11) does not hold, since (6.39) implies that

$$\frac{S_{N(t)}}{t} \xrightarrow{\text{a.s.}} 1 \neq 0 \quad \text{as} \quad t \to \infty. \tag{6.41}$$

This shows the importance of condition (2.6); in the present example we have

$$\lim_{t \to \infty} P(N(t) \leq xt^2) = G_{1/2}(x) = \sqrt{\frac{2}{\pi}} \int_{1/\sqrt{x}}^{\infty} e^{-y^2/2} dy, \tag{6.42}$$

(see Feller (1968), p. 90), that is, condition (2.6) does not hold.

Concerning a converse to Theorem 6.3 we can modify Example 6.1 so that $N(t) = \min\{n \colon S_n = [\sqrt{t}]\}$, $t \geq 0$, and the same phenomenon as before will appear. Note also that, since $S_{N(t)}/\sqrt{t} \xrightarrow{\text{a.s.}} 1$ as $t \to \infty$, Anscombe's theorem (Theorem 3.1(ii)) fails. Moreover, by (6.42) we have

$$\lim_{t \to \infty} P(N(t) \leq x(\sqrt{t})^2) = G_{1/2}(x), \tag{6.43}$$

that is, $N(t)/t$ *converges in distribution* to the strictly asymmetric stable distribution with index 1/2, from which it follows that $S_{N(t)}/\sqrt{N(t)} = [\sqrt{t}]/\sqrt{N(t)}$ converges in distribution to $|Z|$, where Z is standard normal. Thus, Theorem 3.1(i) does not hold either. Now, in generalized versions of Anscombe's theorem $N(t)/t$ may converge *in probability* to a positive random variable, but in the present example $N(t)/t$ converges in distribution only.

As for positive converses when $EX_1 = 0$ the situation seems harder.

1.7 Moment Convergence

We are now ready to establish our results on moment convergence in the strong laws of large numbers and the central limit theorem. The proofs are immediate consequences of Theorems 6.1–6.3 and Theorems 2.3 and 3.1, respectively and Theorem A.1.1 (recall Remarks A.1.1 and A.1.2). We thus assume that $\{S_n, n \geq 0\}$ is a random walk with i.i.d. increments $\{X_k, k \geq 1\}$ and that $\{N(t), t \geq 0\}$ is a family of *stopping times* as described before.

Theorem 7.1. *Let $r \geq 1$ and suppose that $E|X_1|^r < \infty$. Set $\mu = EX_1$. If*

$$\frac{N(t)}{t} \xrightarrow{\text{a.s.}} \theta \quad (0 < \theta < \infty) \quad \text{as} \quad t \to \infty \tag{7.1}$$

and, for some $t_0 \geq 1$, the family

$$\left\{ \left(\frac{N(t)}{t} \right)^r, t \geq t_0 \right\} \quad \text{is uniformly integrable,} \tag{7.2}$$

then

$$\frac{S_{N(t)}}{t} \to \mu \cdot \theta \quad \text{a.s. and in } L^r \quad \text{as} \quad t \to \infty \tag{7.3}$$

and

$$E \left| \frac{S_{N(t)}}{t} \right|^p \to (|\mu|\theta)^p \quad \text{as} \quad t \to \infty \quad \text{for all } p, 0 < p \leq r. \tag{7.4}$$

Theorem 7.2. *Let $0 < r < 2$, suppose that $E|X_1|^r < \infty$ and that $EX_1 = 0$ when $1 \leq r < 2$. If*

$$\frac{N(t)}{t} \xrightarrow{\text{a.s.}} \theta \quad (0 < \theta < \infty) \quad \text{as} \quad t \to \infty \tag{7.5}$$

and, for some $t_0 \geq 1$, the family

$$\left\{ \frac{N(t)}{t}, t \geq t_0 \right\} \quad \text{is uniformly integrable,} \tag{7.6}$$

then

$$\frac{S_{N(t)}}{t^{1/r}} \to 0 \quad \text{a.s. and in } L^r \quad \text{as} \quad t \to \infty \tag{7.7}$$

and

$$E \left| \frac{S_{N(t)}}{t^{1/r}} \right|^p \to 0 \quad \text{as} \quad t \to \infty \quad \text{for all } p, 0 < p \leq r. \tag{7.8}$$

Theorem 7.3. *Let $r \geq 2$, suppose that $E|X_1|^r < \infty$ and $EX_1 = 0$ and set $\sigma^2 = \text{Var} X_1$ $(0 < \sigma^2 < \infty)$. If*

$$\frac{N(t)}{t} \xrightarrow{p} \theta \quad (0 < \theta < \infty) \quad \text{as} \quad t \to \infty \tag{7.9}$$

and, for some $t_0 \geq 1$, the family

$$\left\{ \left(\frac{N(t)}{t} \right)^{r/2}, t \geq t_0 \right\} \quad \text{is uniformly integrable,} \tag{7.10}$$

then

$$E \left| \frac{S_{N(t)}}{\sqrt{t}} \right|^p \to E|Z|^p \quad \text{as} \quad t \to \infty \quad \text{for all } p, 0 < p \leq r \tag{7.11}$$

and

$$E \left(\frac{S_{N(t)}}{\sqrt{t}} \right)^k \to 0 \quad \text{as} \quad t \to \infty \quad \text{for } k = \text{odd integer} \leq r, \tag{7.12}$$

where Z is a normal random variable with mean 0 and variance $\theta \sigma^2$.

We have preferred to prove moment convergence via uniform integrability. In some cases, however, one can prove moment convergence by studying the moments directly, in particular when they converge to 0. Let us consider Theorem 7.2 (cf. Gut (1974c)).

By collecting the estimates for the moments in the case $0 < r < 1$ (cf. (6.10) and (6.11)) we find that

$$E\left|\frac{S_{N(t)}}{t^{1/r}}\right|^r \le E\left|\frac{S'_{N(t)}}{t^{1/r}}\right|^r + E\left|\frac{S''_{N(t)}}{t^{1/r}}\right|^r < t^{r-1}M^r E\left(\frac{N(t)}{t}\right)^r + \varepsilon E\frac{N(t)}{t},$$

from which it follows that

$$\limsup_{t\to\infty} E\left|\frac{S_{N(t)}}{t^{1/r}}\right|^r < \varepsilon \cdot \sup_{t\ge t_0} E\frac{N(t)}{t}. \tag{7.13}$$

For the case $1 \le r < 2$ we set $N_n(t) = N(t) \wedge n$ and apply Theorem A.2.2 to obtain

$$E|S'_{N_n(t)}|^r \le B_r \cdot E\left|\sum_{k=1}^{N_n(t)} (X'_k)^2\right|^{r/2} \le B_r \cdot E(N_n(t)(2M)^2)^{r/2}$$

$$\le B_r \cdot (2M)^r E(N(t))^{r/2},$$

which by Fatou's lemma yields

$$E|S'_{N(t)}|^r \le B_r \cdot (2M)^r E(N(t))^{r/2}. \tag{7.14}$$

This, together with (6.17) and the c_r-inequalities, yields

$$E\left|\frac{S_{N(t)}}{t^{1/r}}\right|^r < 2^{r-1}\left(t^{-1+(r/2)}B_r \cdot (2M)^r E\left(\frac{N(t)}{t}\right)^{r/2} + 2^r B_r \cdot \varepsilon E\frac{N(t)}{t}\right)$$

and thus

$$\limsup_{t\to\infty} E\left|\frac{S_{N(t)}}{t^{1/r}}\right|^r < \text{constant} \cdot \varepsilon. \tag{7.15}$$

By the arbitrariness of ε we obtain the following result from (7.13) and (7.15). Note also that we actually only used boundedness of $\{E(N(t)/t)\}$.

Theorem 7.4. *Let $0 < r < 2$, let $\{X_k, k \ge 1\}$ be as in Theorem 7.2 and suppose that*

$$\left\{E\frac{N(t)}{t}, t \ge t_0\right\} \quad \text{is bounded.} \tag{7.16}$$

Then

$$\frac{S_{N(t)}}{t^{1/r}} \to 0 \quad \text{in } L^r \text{ and (hence) in probability} \quad \text{as } t \to \infty, \tag{7.17}$$

that is

$$E\left|\frac{S_{N(t)}}{t^{1/r}}\right|^r \to 0 \quad as \quad t \to \infty. \tag{7.18}$$

The differences between Theorems 7.2 and 7.4 are

(a) since we did not assume (7.5) in the latter we cannot expect a.s. convergence and (and this is more important)
(b) since the assumption (7.16) is weaker than (7.6) we cannot conclude that $\{|S_{N(t)}/t^{1/r}|^r\}$ is uniformly integrable without any extra assumptions, such as, for example, that $\{N(t)\}$ is strictly increasing as a function of t (for almost all ω).
(c) since we studied *moments* and not *tails* in Theorem 7.4 we did not have to introduce the truncation (6.12) of the stopping times (for $1 \le r < 2$).

Remark 7.1. Theorems 4.1 and 4.2 are special cases of Theorems 7.1 and 7.3 respectively.

Remark 7.2. In view of Remark 6.3 it is also possible to state results when $\{N(t), t \ge 0\}$ are not stopping times (by assuming the existence of one extra moment of X_1).

1.8 The Stopping Summand

Since we, in general, will be dealing with stopping times it may be of special interest to study the (asymptotic) behavior of the stopping summand, that is, the summand which causes the stopping of the process.

As an example we may look at the theory of first passage times, see Chapter 3, where one considers the stopping times

$$\nu(t) = \min\{n: S_n > t\} \quad (t \ge 0) \tag{8.1}$$

for the sequence of partial sums of the i.i.d. random variables $\{X_k, k \ge 1\}$ with positive mean. The object of interest then is $X_{\nu(t)}$. An immediate observation is that, although a general summand assumes negative values with positive probability, we must have

$$P(X_{\nu(t)} > 0) = 1. \tag{8.2}$$

In what follows, we let $\{X_k, k \ge 1\}$ and $\{N(t), t \ge 0\}$ be as before. The following lemma is Gut (1974a), Lemma 2.1. It gives a rather crude, but all the same, frequently sufficiently powerful, estimate on the moments of $X_{N(t)}$.

Lemma 8.1. *Let $r \ge 1$. If $E|X_1|^r < \infty$ and $EN(t) < \infty$, then*

$$E|X_{N(t)}| \le (E|X_{N(t)}|^r)^{1/r} \le (EN(t) \cdot E|X_1|^r)^{1/r}. \tag{8.3}$$

Proof. The first inequality is Lyapounov's inequality. As for the second one we apply the useful relation

$$|X_{N(t)}|^r \leq |X_1|^r + \cdots + |X_{N(t)}|^r \tag{8.4}$$

and Theorem 5.3 to the sequence of i.i.d. random variables $\{|X_k|^r, k \geq 1\}$ and the stopping times $\{N(t), t \geq 0\}$ to obtain

$$E|X_{N(t)}|^r \leq E \sum_{k=1}^{N(t)} |X_k|^r = EN(t) \cdot E|X_1|^r. \tag{8.5}$$

□

Remark 8.1. Note that (8.4) and (8.5) hold for all $r > 0$.

From (8.5) it follows, for example, that

$$E\frac{|X_{N(t)}|^r}{t} \leq E\left(\frac{N(t)}{t}\right) \cdot E|X_1|^r \tag{8.6}$$

and that if $\{E(N(t)/t), t \geq t_0\}$ is bounded, then so is $\{E(|X_{N(t)}|^r/t), t \geq t_0\}$. To improve this to uniform integrability we can proceed as in Section 1.6, but we can, in fact, also appeal to Theorem 6.1 for the following reason.

By viewing $\{|X_k|^r, k \geq 1\}$ as a sequence of i.i.d. random variables with finite mean we can invoke Theorem 6.1 for first moments to conclude that

$$\left\{\frac{1}{t} \sum_{k=1}^{N(t)} |X_k|^r, t \geq t_0\right\} \quad \text{is uniformly integrable} \tag{8.7}$$

provided $\{(N(t)/t), t \geq t_0\}$ is. In view of (8.4) and Theorem 2.3(i) the following result emerges.

Theorem 8.1. *Let $r > 0$ and suppose that $E|X_1|^r < \infty$. If, for some $t_0 \geq 1$, the family*

$$\left\{\frac{N(t)}{t}, t \geq t_0\right\} \quad \text{is uniformly integrable,} \tag{8.8}$$

then the family

$$\left\{\frac{|X_{N(t)}|^r}{t}, t \geq t_0\right\} \quad \text{is uniformly integrable.} \tag{8.9}$$

If, moreover,

$$\frac{N(t)}{t} \xrightarrow{\text{a.s.}} \theta \quad (0 < \theta < \infty) \quad as \quad t \to \infty, \tag{8.10}$$

Then

$$C\left(\left\{\frac{S_{N(t)}}{\sqrt{2\sigma^2 N(t)\log^+\log^+ N(t)}}, t\geq 1\right\}\right) = [-1,1]\ \text{a.s.} \qquad (9.5)$$

and

$$C\left(\left\{\frac{S_{N(t)}}{\sqrt{2\sigma^2 t\theta\log\log t}}, t\geq 3\right\}\right) = [-1,1]\ \text{a.s.} \qquad (9.6)$$

In particular,

$$\limsup_{t\to\infty}\left(\liminf_{t\to\infty}\right)\frac{S_{N(t)}}{\sqrt{2\sigma^2 N(t)\log\log N(t)}} = \overset{+}{(-)}1\ \text{a.s.} \qquad (9.7)$$

and

$$\limsup_{t\to\infty}\left(\liminf_{t\to\infty}\right)\frac{S_{N(t)}}{\sqrt{2\sigma^2 t\theta\log\log t}} = \overset{+}{(-)}1\ \text{a.s.} \qquad (9.8)$$

Such results are generally obtained as particular cases of a strong invariance principle, in which the whole summation process is considered. For an elementary direct proof of Theorem 9.1 (in discrete time) we refer to Torrång (1987).

1.10 Complete Convergence and Convergence Rates

A sequence of random variables $\{U_n, n\geq 1\}$ is said to *converge completely* to a constant c if

$$\sum_{n=1}^{\infty} P(|U_n - c| > \varepsilon) < \infty \quad \text{for all } \varepsilon > 0. \qquad (10.1)$$

This definition was introduced by Hsu and Robbins (1947), who proved that the sequence of arithmetic means of i.i.d. random variables converges completely to the expected value of the summands, provided the variance is finite. Later Erdős (1949, 1950) proved the converse (cf. also Gut (2007), Theorem 6.11.2).

Such a result can also be interpreted as a statement concerning the rate of convergence in the law of large numbers. In this context it has been extended by several authors. For the most general version we refer to Baum and Katz (1965); cf. also Gut (2007), Theorem 6.12.1.

The Hsu and Robbins result has been generalized to stopped random walks by Szynal (1972) and the corresponding generalization of Baum and Katz (1965) (see also Katz (1963)) has been established in Gut (1983b). Here we quote Theorem 2.1 and Gut (1983b) in a form relevant to our considerations in Chapter 3. Note that $\{N(t), t\geq 0\}$ need not be stopping times.

Theorem 10.1. *Let $\alpha r \geq 1$ and $\alpha > 1/2$. Suppose that $E|X_1|^r < \infty$ and that $EX_1 = 0$ when $r \geq 1$. If, for some θ $(0 < \theta < \infty)$ and some δ $(0 < \delta < \theta)$,*

$$\sum_{n=1}^{\infty} n^{\alpha r - 2} P(|N(n) - n\theta| > n\delta) < \infty, \tag{10.2}$$

then

$$\sum_{n=1}^{\infty} n^{\alpha r - 2} P(|S_{N(n)}| > (N(n))^{\alpha} \varepsilon) < \infty \quad \text{for all } \varepsilon > 0 \tag{10.3}$$

and

$$\sum_{n=1}^{\infty} n^{\alpha r - 2} P(|S_{N(n)}| > n^{\alpha} \varepsilon) \quad \text{for all } \varepsilon > 0. \tag{10.4}$$

Outline of Proof. The proof of (10.3) is based on the relation

$$P(|S_{N(n)}| > (N(n))^{\alpha} \varepsilon)$$

$$= P(\{|S_{N(n)}| > (N(n))^{\alpha} \varepsilon\} \cap \{|N(n) - n\theta| > n\delta\})$$

$$+ P(\{|S_{N(n)}| > (N(n))^{\alpha} \varepsilon\} \cap \{|N(n) - n\theta| \leq n\delta\}), \tag{10.5}$$

after which the first probability in the RHS is taken care of by (10.2) and the last term is taken care of by the results for the classical case. For details we refer to the original paper. The proof of (10.4) is similar. $\qquad\square$

Remark 10.1. For $\alpha = 1, r = 2$ (Szynal (1972)), a slightly weaker result can be stated in terms of complete convergence; namely, if $\{(N(n)/n), n \geq 1\}$ converges completely to a positive, finite constant, then so do $\{|S_{N(n)}/N(n)|, n \geq 1\}$ and $\{|S_{N(n)}/n|, n \geq 1\}$ provided the variance is finite (and the mean is 0).

For the limiting case $\alpha = 1/2$ the probabilities $P(|S_n| > \sqrt{n} \cdot \varepsilon)$ do not converge to 0 as $n \to \infty$ in view of the central limit theorem, so there cannot be a general theorem for that case. However, by modifying the normalization slightly one can still obtain positive results. The following theorem is related to the law of the iterated logarithm for stopped random walks.

Theorem 10.2. *Suppose that $EX_1 = 0$ and that $EX_1^2 = \sigma^2 < \infty$. If, for some θ $(0 < \theta < \infty)$,*

$$\sum_{n=1}^{\infty} \frac{1}{n} P(|N(n) - n\theta| > n\delta) < \infty \quad \text{for all } \delta < 0 \tag{10.6}$$

then

$$\sum_{n=1}^{\infty} \frac{1}{n} P(|S_{N(n)}| > \varepsilon \sqrt{N(n) \log^+ \log^+ N(n)}) < \infty \quad \text{for } \varepsilon > \sigma\sqrt{2} \tag{10.7}$$

and

$$\sum_{n=3}^{\infty} \frac{1}{n} P(|S_{N(n)}| > \varepsilon \sqrt{n \log \log n}) < \infty \quad \text{for } \varepsilon > \sigma\sqrt{2\theta}. \tag{10.8}$$

Remark 10.2. The results have been stated for $\{N(n), n \geq 1\}$. It is, however, also possible to state them for the whole family $\{N(t), t \geq 0\}$ with the sums replaced by integrals (see Gut (1983b), Remark 5.D).

1.11 Problems

1. Prove that Theorem 2.3(i) and (ii) remain valid if (2.6) is replaced by

$$\limsup_{t \to \infty} \frac{N(t)}{t} < \infty \text{ a.s.} \tag{11.1}$$

2. Prove Anscombe's theorem (Theorem 3.1) by showing that the Anscombe condition (A) is satisfied.
3. Prove Theorem 3.2.
4. A martingale proof of Theorem 4.1. Set $\mathscr{F}_n = \sigma\{S_k, k \geq n\}$, $n \geq 1$.
   (a) Suppose first that $E|X_1| < \infty$ (and that $EX_1 = \mu$). Show that $\{(S_n/n, \mathscr{F}_n), n \geq 1\}$ is a reversed martingale and hence, that

$$\frac{S_n}{n} \to \mu \quad \text{a.s. and in } L^1 \quad \text{as} \quad n \to \infty. \tag{11.2}$$

   (b) Suppose, in addition, that $E|X_1|^r < \infty$ $(r \geq 1)$. Extend the arguments in (a) to show that

$$\frac{S_n}{n} \to \mu \quad \text{a.s. and in } L^r \quad \text{as} \quad n \to \infty. \tag{11.3}$$

5. Prove Theorem 4.1 by using the method to prove Theorem 4.2.
6. Prove the version of Theorem 4.1 corresponding to the Marcinkiewicz–Zygmund strong law mentioned at the end of Section 1.4.
7. In Remark 5.5 we mentioned Blackwell (1946), where another proof of Theorem 5.3(i) is given. The idea there is closely related to that of step (i) in the proof of Theorem 5.5, that is, the crucial formula is

$$\frac{S_{N_1+\cdots+N_k}}{k} = \frac{S_{N_1+\cdots+N_k}}{N_1+\cdots+N_k} \cdot \frac{N_1+\cdots+N_k}{k} \tag{11.4}$$

   (cf. formula (5.45)). Complete the proof.
8. Prove Theorem 5.3 by direct computations; cf. Remark 5.6. For (ii) it simplifies to assume (without restriction) that $EX_1 = 0$.
9. Verify formula (5.25).
10. Verify Remark 8.2.
11. Prove (8.9) by proceeding as in Section 1.6.
12. Complete the proof of Theorem 10.1.

2

Renewal Processes and Random Walks

2.1 Introduction

In the first chapter we stated and proved various limit theorems for stopped random walks. These limit theorems shall, in subsequent chapters, be used in order to obtain results for random walks stopped according to specific stopping procedures as well as for the families of stopping times (random indices) themselves. However, before doing so we shall, in this chapter, survey some of the basic facts about random walks.

Our emphasis will be on that part of the theory which is most relevant for this book. Classical fluctuation theory, the combinatorial formulas, Wiener–Hopf factorization etc. are therefore excluded in our presentation; we refer the reader to the existing literature cited below. We furthermore assume that the reader already has some familiarity with much of the material, so its character will rather be a review than a through exposition. As a consequence proofs will not always be given; in general only when they are short or in the spirit of the present treatise.

We begin our survey by considering an important class of random walks which has attended special interest; the class of *renewal processes*. Their special feature is that they are concentrated on $[0, \infty)$, that is, the steps are non-negative. A similar theory exists, of course, for random walks concentrated on $(-\infty, 0]$. Sections 2.2–2.7 are devoted to the study of renewal processes.

In the remaining sections of the chapter we treat random walks on $(-\infty, \infty)$ (and such that they are not concentrated on either half axis). We give a characterization of the three possible kinds of random walks, introduce ladder variables, the partial maxima and present some general limit theorems.

We close this section by introducing some notation. Throughout $(\Omega, \mathscr{F}, P)$ is a probability space on which everything is defined, $\{S_n, n \geq 0\}$ is a random walk with i.i.d. increments $\{X_k, k \geq 1\}$ and $S_0 = 0$, or, equivalently, $\{X_k, k \geq 1\}$ is a sequence of i.i.d. random variables with partial sums $S_n = \sum_{k=1}^{n} X_k, \ n \geq 0$. We let F denote the (common) distribution of

the increments (summands). To avoid trivialities we assume throughout that $P(X_1 \neq 0) > 0$.

Depending on the support of the distribution function F we shall distinguish between two kinds of random walks (renewal processes). We say that the random walk (renewal process) is *arithmetic* if F has its support on $\{0, \pm d, \pm 2d, \dots\}$ ($\{0, d, 2d, \dots\}$) for some $d > 0$. The largest d with this property is called the *span*. A random walk or renewal process which is *arithmetic with span d* will also be called *d-arithmetic*. If F is not of this kind (for any d) we say that the random walk or renewal process is *nonarithmetic*.

2.2 Renewal Processes; Introductory Examples

Example 2.1. Consider some electronic device, in particular, one specific component. As soon as the component breaks down it is automatically and instantly replaced by a new identical one, which, when it breaks down, is replaced similarly etc. Let $\{X_k, k \geq 1\}$ denote the successive lifetimes and set $S_n = \sum_{k=1}^{n} X_k$, $n \geq 0$. With this setup S_n denotes the (random) total lifetime of the n first components. If, for example, $\{X_k, k \geq 1\}$ is a sequence of i.i.d. exponentially distributed random variables, then S_n has a gamma distribution.

In practice, however, it is more likely that one is interested in the *random number of components* required to keep the device alive during a *fixed time interval* rather than being interested in the accumulated *random lifetime* for a *fixed number* of components.

In the exponential case, the stochastic process which counts the number of components that are replaced during fixed time intervals is the well known Poisson process.

Example 2.2. In a chemical substance the movement of the molecules can be modeled in such a way that each molecule is subject to repeated (i.i.d.) random displacements at random times. In analogy with the previous example, rather than being interested in the accumulated displacement of a molecule after a fixed number of steps, it is more natural to consider the location of a molecule at a fixed time point.

Example 2.3 (The Bernoulli random walk). Here $\{X_k, k \geq 1\}$ are i.i.d. Be(p)-distributed random variables, that is, $P(X_k = 1) = 1 - P(X_k = 0) = p$, and $S_n = \sum_{k=1}^{n} X_k$, $n \geq 0$, denotes the random number of "successes" after n trials, which has a Binomial distribution. If, instead, we consider the random number of performances required to obtain a given number of successes we are lead to the Negative Binomial process.

A common feature in these examples thus is that rather than studying the random value of a sum of a fixed number of random variables one investigates the random number of terms required in order for the sum to attain a

certain (deterministic) value. For nonnegative summands we are lead to the part of probability theory called *renewal theory* and the summation process $\{S_n, n \geq 0\}$ is called a *renewal process*. Examples 2.1 and 2.3 are exactly of this kind and, under appropriate additional assumptions, also Example 2.2.

We now proceed to give stringent definitions and present a survey of the most important results for renewal processes.

Some fundamental papers on renewal theory are Feller (1949), Doob (1948) and Smith (1954, 1958). The most standard book references are Feller (1968), Chapter XIII for the arithmetic case (where the treatment is in the context of recurrent events) and Feller (1971), Chapters VI and XI for the nonarithmetic case. Some further references are Prabhu (1965), Chapter 5, Cox (1967), Çinlar (1975), Chapter 9, Jagers (1975), Chapter 5 and Asmussen (2003).

2.3 Renewal Processes; Definition and General Facts

Let $\{X_k, k \geq 1\}$ be i.i.d. nonnegative random variables and let $\{S_n, n \geq 0\}$ be the partial sums. The sequence $\{S_n, n \geq 0\}$ thus defined is a *renewal process*.

We further let F denote the common distribution function of the summands and let F_n be the distribution function of S_n, $n \geq 0$. Thus

$$F_0(x) = \begin{cases} 0, & x < 0, \\ 1, & x \geq 0, \end{cases} \tag{3.1a}$$

$$F_1(x) = F(x), \tag{3.1b}$$

$$F_n(x) = F^{n*}(x) \qquad (n \geq 1), \tag{3.1c}$$

that is, F_n equals the n-fold convolution of F with itself.

To avoid trivialities we assume throughout that $P(X_1 > 0) > 0$.

The main object of interest in the study of renewal processes (in renewal theory) is the *renewal counting process* $\{N(t), t \geq 0\}$, defined by

$$N(t) = \max\{n\colon S_n \leq t\}. \tag{3.2}$$

An alternative interpretation is $N(t) =$ the number of renewals in $(0, t] =$ Card$\{n \geq 1\colon S_n \leq t\}$.

Remark 3.1. A case of particular importance is when $F(x) = 1 - e^{-\lambda x}$, $x \geq 0$, that is, when the lifetimes are exponentially distributed. In this case $\{N(t), t \geq 0\}$ is a *Poisson process* with intensity λ (recall Example 2.1). We also observe that the renewal process defined in Example 2.1 is nonarithmetic. The renewal process defined in Example 2.3 is, however, arithmetic with span 1 (1-arithmetic).

Remark 3.2. The definition (3.2) is not the only existing definition of the renewal counting process. Some authors prefer Card$\{n \geq 0: S_n \leq t\}$, thus counting $n = 0$ as a renewal.

We are now ready for our first result.

Theorem 3.1.

(i) $P(N(t) < \infty) = 1$;
(ii) $E(N(t))^r < \infty$ *for all* $r > 0$;
(iii) *There exists* $s_0 > 0$ *such that* $Ee^{sN(t)} < \infty$ *for all* $s < s_0$.

Proof. By assumption there exists $x_0 > 0$ such that $P(X_1 \geq x_0) > 0$. Since scaling does not affect the conclusions we may, without loss of generality, assume that $x_0 = 1$. Now, define, for $k \geq 1$,

$$\bar{X}_k = \begin{cases} 0, & \text{if } X_k < 1, \\ 1, & \text{if } X_k \geq 1, \end{cases} \tag{3.3}$$

$\bar{S}_n = \sum_{k=1}^{n} \bar{X}_k$, $n \geq 0$, and $\bar{N}(t) = \max\{n: \bar{S}_n \leq t\}$. Then, clearly, $\bar{X}_k \leq X_k$, $k \geq 1$, $\bar{S}_n \leq S_n$, $n \geq 0$, and, hence, $\bar{N}(t) \geq N(t)$, $t \geq 0$.

But, $\{\bar{N}(t), t \geq 0\}$ is the Negative Binomial process from Example 2.3 (cf. also Remark 3.1), for which the theorem is well known, and the conclusions follow. $\qquad \square$

An important relation upon which several proofs are based is the inverse relationship between renewal processes and counting processes, namely

$$\{N(t) \geq n\} = \{S_n \leq t\}. \tag{3.4}$$

As an immediate example we have

$$EN(t) = \sum_{n=1}^{\infty} P(N(t) \geq n) = \sum_{n=1}^{\infty} P\{S_n \leq t\} = \sum_{n=1}^{\infty} F_n(t). \tag{3.5}$$

Next we define the *renewal function* $U(t) = \sum_{n=1}^{\infty} F_n(t)$ and conclude, in view of (3.5), that

$$U(t) = \sum_{n=1}^{\infty} F_n(t) = EN(t). \tag{3.6}$$

Remark 3.3. Just as there is no unique way of defining counting processes (see Remark 3.2) there is also some ambiguity concerning the renewal function. If one defines counting processes in such a way that $n = 0$ is also counted as a renewal, then it is more natural to define the renewal function as $\sum_{n=0}^{\infty} F_n(t)$ (for example, in order for the analog of (3.6) to remain true). In fact, some authors *begin* by defining $U(t)$ this way and *then* define $N(t)$ so that $U(t) = EN(t)$, that is, such that $N(t) = \text{Card}\{n \geq 0: S_n \leq t\}$.

By using the first equality in (3.6) twice we obtain

$$U(t) = F_1(t) + \sum_{n=1}^{\infty} F_{n+1}(t) = F(t) + \sum_{n=1}^{\infty} (F_n * F)(t)$$

$$= F(t) + (U * F)(t),$$

which is one half of the following theorem.

Theorem 3.2 (The Integral Equation for Renewal Processes). *The renewal function $U(t)$ satisfies the integral equation*

$$U(t) = F(t) + (U * F)(t) \tag{3.7a}$$

or, equivalently,

$$U(t) = F(t) + \int_0^t U(t-s)dF(s). \tag{3.7b}$$

Moreover, $U(t)$ is the unique solution of (3.7), which is bounded on finite intervals.

Remark 3.4. If the renewal process is d-arithmetic, then

$$U(nd) = u_0 + u_1 + \cdots + u_n \qquad (n \geq 0), \tag{3.8a}$$

where

$$u_k = \sum_{j=1}^{\infty} P(S_j = kd) \qquad (k \geq 0). \tag{3.8b}$$

Moreover, with $f_k = P(X_1 = kd)$, $k \geq 0$, we obtain the discrete convolution formula

$$u_n = f_n + \sum_{k=0}^{n} u_{n-k} f_k. \tag{3.9}$$

Remark 3.5. By defining indicator variables $\{I_j, j \geq 1\}$ by $I_j = I\{S_j = kd\}$ we note that $\sum_{j=1}^{\infty} I\{S_j = kd\}$ equals the *actual number* of partial sums equal to kd, $k \geq 0$, and that u_k equals the *expected number* of partial sums which are equal to kd, $k \geq 0$. (Observe that S_0 is not included in the count).

A mathematically important fact is that $N(t)$ *is not a stopping time* (with respect to the renewal process); for the definition of a stopping time we refer to the end of Section A.2. To see this intuitively, we note that we cannot determine whether or not the event $\{N(t) = n\}$ has occurred without looking into the future. It is therefore convenient to introduce the *first passage time process* $\{\nu(t), t \geq 0\}$, defined by

$$\nu(t) = \min\{n \colon S_n > t\}, \tag{3.10}$$

because $\nu(t)$ *is a stopping time* for all $t > 0$. Moreover, it turns out that first passage time processes are the natural processes to consider in the context of renewal theory for random walks (see Chapter 3 below).

We now observe that

$$\nu(t) = N(t) + 1, \tag{3.11}$$

from which it, for example, follows that the conclusions of Theorem 3.1 also hold for the first passage times.

The following formula, corresponding to (3.6) also follows.

$$E\nu(t) = 1 + EN(t) = 1 + \sum_{n=1}^{\infty} F_n(t) = \sum_{n=0}^{\infty} F_n(t). \tag{3.12}$$

However, whereas $0 \leq S_{N(t)} \leq t$, that is, $S_{N(t)}$ has moments of all orders we cannot conclude that $S_{\nu(t)}$ has any moments without further assumptions. We have, in fact,

$$E(S_{\nu(t)})^r < \infty \quad \Longleftrightarrow \quad E(X_1)^r < \infty \qquad (r > 0). \tag{3.13}$$

This follows from Theorems 3.1, 1.5.1 and 1.5.2. We give no details, since such an equivalence will be established for general random walks with $EX_1 = \mu > 0$ (and $r \geq 1$) in Chapter 3; see Theorem 3.3.1.

In the following section we present some asymptotic results for renewal counting processes, so-called renewal theorems, which, due to their nature and in view of (3.11), also hold for first passage time processes. In fact, the proofs of some of them are normally first given for the latter processes (because of the stopping time property) after which one uses (3.11).

2.4 Renewal Theorems

Much of the early work on renewal processes was devoted to the study of the renewal function $U(t)$, in particular to asymptotics. In this section we present some of these limit theorems.

Theorem 4.1 (The Elementary Renewal Theorem). *Let* $0 < \mu = EX_1 \leq \infty$. *Then*

$$\frac{U(t)}{t} \rightarrow \frac{1}{\mu} \quad as \quad t \rightarrow \infty, \tag{4.1}$$

the limit being 0 when $\mu = +\infty$.

Proof. Suppose first that $0 < \mu < \infty$. Since $N(t)$ is not a stopping time we consider $\nu(t)$ in order to apply Theorem 1.5.3. It follows that

$$U(t) = EN(t) = E\nu(t) - 1 = \frac{1}{\mu}ES_{\nu(t)} - 1$$

$$= \frac{t}{\mu} + \frac{1}{\mu}E(S_{\nu(t)} - t) - 1,$$

and, hence, that

$$\frac{U(t)}{t} = \frac{1}{\mu} + \frac{E(S_{\nu(t)} - t)}{\mu t} - \frac{1}{t}. \tag{4.2}$$

Since $S_{\nu(t)} - t \geq 0$ we obtain

$$\liminf_{t \to \infty} \frac{U(t)}{t} \geq \frac{1}{\mu}. \tag{4.3}$$

Next we note that

$$S_{\nu(t)} - t \leq S_{\nu(t)} - S_{N(t)} = X_{\nu(t)}. \tag{4.4}$$

Suppose that $P(X_k \leq M) = 1$ for some $M > 0$, $k \geq 1$. Then (4.2) and (4.4) together imply that

$$\frac{U(t)}{t} \leq \frac{1}{\mu} + \frac{M}{t} \tag{4.5}$$

and thus that

$$\limsup_{t \to \infty} \frac{U(t)}{t} \leq \frac{1}{\mu}, \tag{4.6}$$

which, together with (4.3), proves (4.1) for that case.

For arbitrary $\{X_k, k \geq 1\}$ the conclusion follows by a truncation procedure. We define a new renewal process $\{S'_n, n \geq 0\}$, by defining $X'_k = X_k I\{X_k \leq M\}$, $k \geq 1$. By arguing as in the proof of Theorem 3.1 we now obtain $U(t) \leq U'(t)$ and, hence, that

$$\limsup_{t \to \infty} \frac{U(t)}{t} \leq \frac{1}{\mu'} = \frac{1}{EX_1 I\{X_1 \leq M\}}. \tag{4.7}$$

The conclusion follows by letting $M \to \infty$.

Finally, if $\mu = +\infty$ we truncate again and let $M \to \infty$. $\qquad \square$

The following result is a refinement of Theorem 4.1. The arithmetic case is due to Kolmogorov (1936) and Erdős, Feller and Pollard (1949) and the nonarithmetic case is due to Blackwell (1948).

Theorem 4.2.

(i) *For nonarithmetic renewal processes we have*

$$U(t) - U(t - h) \to \frac{h}{\mu} \quad as \quad t \to \infty. \tag{4.8}$$

(ii) *For d-arithmetic renewal processes we have*

$$u_n = \sum_{k=1}^{\infty} P(S_k = nd) \to \frac{d}{\mu} \quad as \quad n \to \infty. \tag{4.9}$$

The limits are 0 when $\mu = +\infty$.

Theorem 4.2 is closely related to the integral equation (3.7). The following theorem gives a connection.

Theorem 4.3 (The Key Renewal Theorem).

(i) *Suppose that the renewal process is nonarithmetic. If $G(t)$, $t \geq 0$, is a bounded, nonnegative, nonincreasing function, such that $\int_0^\infty G(t)dt < \infty$, then*

$$\int_0^t G(t-s)dU(s) \rightarrow \frac{1}{\mu} \int_0^\infty G(s)ds \quad as \quad t \rightarrow \infty. \qquad (4.10)$$

(ii) *Suppose that the renewal process is d-arithmetic. If $G(t)$, $t \geq 0$, is nonnegative and $\sum_{n=0}^\infty G(nd) < \infty$, then*

$$\sum_{k=0}^n G(nd-kd)u_{kd} \rightarrow \frac{d}{\mu} \sum_{k=0}^\infty G(kd) \quad as \quad n \rightarrow \infty. \qquad (4.11)$$

If $\mu = +\infty$ the limits in (i) *and* (ii) *are 0.*

If, in particular, $\{X_k, k \geq 1\}$ are exponentially distributed with mean λ^{-1}, that is, if $\{N(t), t \geq 0\}$ is a Poisson process with intensity λ (recall Example 2.1 and Remark 3.1), then $\mu = \lambda^{-1}$ and it follows, in particular, that $U(t) = EN(t) = \lambda t = t/\mu$ and that $U(t) - U(t-h) = h/\mu$. Thus, there is *equality* in Theorems 4.1 and 4.2 for all t in this case.

The renewal process in Example 2.3 (which is arithmetic with span 1) runs as follows: First there is a geometric (with mean q/p) number of zeroes, then a one, then a geometric (with mean q/p) number of zeroes followed by a one and so on. It follows that the actual number of partial sums equal to k (recall Remark 3.5) equals $1 +$ a geometric (with mean q/p) number. To see this we observe that the first partial sum that equals n is obtained for S_k where k is such that $S_{k-1} = n - 1$ and $X_k = 1$ (and thus, $S_k = n$). After this there is a geometric number of zeroes before the next one (which brings the sum to $n + 1$) appears. We thus obtain $u_n = 1 + (q/p) = 1/p = 1/\mu$ (since $\mu = p$). Thus, formula (4.9) is exact for all n in this case.

Remark 4.1. In the classical proofs of Theorems 4.2 and 4.3 a lemma due to Choquet and Deny (1960) plays an important role. In Doob, Snell and Williamson (1960) a proof, based on martingale theory, is given; see also Meyer (1966), pp. 192–193 or Jagers (1975), p. 107.

Remark 4.2. In the 1970s an old method due to Doeblin, called coupling enjoyed a big revival. We only mention some references where proofs of renewal theorems using coupling can be found; Lindvall (1977, 1979, 1982, 1986), Arjas, Nummelin and Tweedie (1978), Athreya, McDonald and Ney (1978), Ney (1981) and Thorisson (1987).

Remark 4.3. The key renewal theorem (Theorem 4.3) has been generalized by Mohan (1976) in the nonarithmetic case and by Niculescu (1979) in the *d*-arithmetic case. They assume that G varies regularly with exponent α ($0 \leq \alpha < 1$) and conclude (essentially) that the LHS's in (4.10) and (4.11) vary regularly with exponent α. For a slightly different generalization, see Erickson (1970).

2.5 Limit Theorems

In this section we present the strong law and the central limit theorem for renewal counting processes. The strong law is due to Doob (1948) (a.s. convergence) and Feller (1941) and Doob (1948) (convergence of the mean) and Hatōri (1959) (convergence of moments of order > 1). The central limit theorem is due to Feller (1949) in the arithmetic case and to Takács (1956) in the nonarithmetic case. Here we present a proof based on Anscombe's theorem (Theorem 1.3.1). The asymptotic expansions of mean and variance are due to Feller (1949) in the arithmetic case and Smith (1954) in the nonarithmetic case.

Let us, however, first observe that, by (3.4), we have

$$N(t) \to +\infty \quad \text{as} \quad t \to \infty. \tag{5.1}$$

By (5.1) and (3.11) we also have $\nu(t) \to \infty$ as $t \to \infty$.

Theorem 5.1 (The Strong Law for Counting Processes). *Let $0 < \mu = EX_1 \leq \infty$. Then*

(i) $$\frac{N(t)}{t} \xrightarrow{\text{a.s.}} \frac{1}{\mu} \quad as \quad t \to \infty;$$

(ii) $$E \left(\frac{N(t)}{t} \right)^r \to \frac{1}{\mu^r} \quad as \quad t \to \infty \quad for \ all \ r > 0.$$

For $\mu = +\infty$ the limits are 0.

Proof. Suppose that $0 < \mu < \infty$.
(i) The relation

$$S_{N(t)} \leq t < S_{\nu(t)} \tag{5.2}$$

and (3.11) together yield

$$\frac{S_{N(t)}}{N(t)} \leq \frac{t}{N(t)} < \frac{S_{\nu(t)}}{\nu(t)} \cdot \frac{N(t) + 1}{N(t)}. \tag{5.3}$$

An application of Theorem 1.2.3 now shows that the extreme members in (5.3) both converge almost surely to μ as $t \to \infty$, which shows that

$$\frac{t}{N(t)} \xrightarrow{\text{a.s.}} \mu \quad \text{as} \quad t \to \infty \tag{5.4}$$

and (i) follows.
(ii) We wish to show that

$$\left\{ \left(\frac{N(t)}{t} \right)^r, t \geq 1 \right\} \quad \text{is uniformly integrable for all } r > 0, \tag{5.5}$$

because (ii) then follows from (i), (5.5) and Theorem A.1.1.
    We prefer, however, to prove the equivalent fact (recall (3.11)) that

$$\left\{ \left(\frac{\nu(t)}{t} \right)^r, t \geq 1 \right\} \quad \text{is uniformly integrable for all } r > 0. \tag{5.6}$$

This result is, essentially, a consequence of the subadditivity of first passage time processes. Namely, let $t, s > 0$ and consider $\nu(t + s)$. In order to reach the level $t + s$ we must first reach the level t. When this has been done the process starts afresh. Since $S_{\nu(t)} > t$ the remaining distance for the process to climb is at most equal to s, and thus, the required number of steps to achieve this is majorized by a random quantity distributed as $\nu(s)$. More formally,

$$\nu(t + s) \leq \nu(t) + \min\{k - \nu(t): S_k - S_{\nu(t)} > s\}$$

$$= \nu(t) + \nu_1(s), \tag{5.7}$$

where $\nu_1(s)$ is distributed as $\nu(s)$.
    Now, let $n \geq 1$ be an integer. By induction we have

$$\nu(n) \leq \nu_1(1) + \cdots + \nu_n(1), \tag{5.8}$$

where $\{\nu_k(1), k \geq 1\}$ are distributed as $\nu(1)$. This, together with Minkowski's inequality (see e.g. Gut (2007), Theorem 3.2.6) and Theorem 3.1(ii), shows that

$$\|\nu(n)\|_r \leq n\|\nu(1)\|_r < \infty. \tag{5.9}$$

Finally, since $\nu(t) \leq \nu([t] + 1)$, we have, for $t \geq 1$,

$$\frac{\nu(t)}{t} \leq 2\frac{\nu([t] + 1)}{[t] + 1} \tag{5.10}$$

and thus, in view of (5.9), that

$$\|\nu(t)/t\|_r \leq 2\|\nu([t] + 1)/([t] + 1)\|_r \leq 2\|\nu(1)\|_r < \infty. \tag{5.11}$$

Since the bound is uniform in t it follows that $\{(\nu(t)/t)^p, t \geq 1\}$ is uniformly integrable for all $p < r$. Since r was arbitrary, (5.6) (and hence (5.5)) follows.

This concludes the proof for the case $0 < \mu < \infty$. For $\mu = +\infty$ the conclusion follows by the truncation procedure used in the proof of Theorem 4.1. We omit the details. $\qquad\square$

Remark 5.1. Theorem 5.1 also holds for $\{\nu(t), t \geq 0\}$. This is immediate from Theorem 5.1, (3.11) and (5.6).

Theorem 5.2 (The Central Limit Theorem for Counting Processes).
Suppose that $0 < \mu = EX_1 < \infty$ and $\sigma^2 = \mathrm{Var} X_1 < \infty$. Then

(i)
$$\frac{N(t) - t/\mu}{\sqrt{\frac{\sigma^2 t}{\mu^3}}} \xrightarrow{d} N(0,1) \quad as \quad t \to \infty,$$

(ii) *If the renewal process is nonarithmetic, then*

$$EN(t) = \frac{t}{\mu} + \frac{\sigma^2 - \mu^2}{2\mu^2} + o(1) \quad as \quad t \to \infty \qquad (5.12)$$

$$\mathrm{Var}\, N(t) = \frac{\sigma^2 t}{\mu^3} + o(t) \quad as \quad t \to \infty. \qquad (5.13)$$

If the renewal process is d-arithmetic, then

$$EN(nd) = \frac{nd}{\mu} + \frac{\sigma^2 - \mu^2}{2\mu^2} + \frac{d}{2\mu} + o(1) \quad as \quad n \to \infty \qquad (5.14)$$

$$\mathrm{Var}\, N(nd) = \frac{\sigma^2 nd}{\mu^3} + o(n) \quad as \quad n \to \infty. \qquad (5.15)$$

Proof. (i) We first observe that

$$\frac{S_{N(t)} - N(t)\mu}{\sigma\sqrt{t/\mu}} \leq \frac{t - N(t)\mu}{\sigma\sqrt{t/\mu}} < \frac{S_{\nu(t)} - \nu(t)\mu}{\sigma\sqrt{t/\mu}} + \frac{\mu}{\sigma}\sqrt{\frac{\mu}{t}} \qquad (5.16)$$

by (3.11) and (5.2). In view of Theorem 5.1(i) and Remark 5.1 we now apply Anscombe's theorem (Theorem 1.3.1(ii)) with $\theta = \mu^{-1}$ to the extreme members of (5.16) and conclude that they both converge in distribution to the standard normal distribution. Thus

$$\frac{t - N(t)\mu}{\sigma\sqrt{t/\mu}} \xrightarrow{d} N(0,1) \quad as \quad t \to \infty. \qquad (5.17)$$

which, in view of the symmetry of the normal distribution, proves (i).

(ii) Formulas (5.12) and (5.13) follow, essentially, from repeated use of the key renewal theorem (Theorem 4.3(i)). Formulas (5.14) and (5.15) follow similarly. $\qquad\square$

Remark 5.2. The classical proofs of Theorem 5.2(i) are based on the ordinary central limit theorem and the inversion formula (3.4).

Remark 5.3. Theorem 5.2 (with obvious modifications) also holds for $\{\nu(t), t \geq 0\}$ in view of (3.11).

To prove moment convergence in Theorem 5.1 we proceeded via uniform integrability and Theorem A.1.1 (and the remarks there). For Theorem 5.2 we referred to proofs based on direct computations. Now, in order to conclude, from Theorem 5.2 that

$$\left\{ \left(\frac{N(t) - t/\mu}{\sqrt{t}} \right)^2, t \geq 1 \right\} \quad \text{is uniformly integrable} \qquad (5.18)$$

we observe that this does not follow immediately from Theorem A.1.1 since we are concerned with a *family* of random variables; recall Remark A.1.2. It does, however, follow that

$$\left\{ \left(\frac{N(n) - n/\mu}{\sqrt{n}} \right)^2, n \geq 1 \right\} \quad \text{is uniformly integrable,} \qquad (5.19)$$

which, together with the monotonicity of $\{N(t), t \geq 0\}$ proves (5.18).

Moreover, (5.18) and (3.11) together imply that (5.18) also holds for first passage time processes, that is, that

$$\left\{ \left(\frac{\nu(t) - t/\mu}{\sqrt{t}} \right)^2, t \geq 1 \right\} \quad \text{is uniformly integrable.} \qquad (5.20)$$

Let us now consider the Poisson process and the Negative Binomial process.

In the former case (recall the notation from above) we have $\mu = \lambda^{-1}$, $\sigma^2 = \lambda^{-2}$, $EN(t) = \lambda t$ and $\operatorname{Var} N(t) = \lambda t$. The validity of the central limit theorem is trivial. As for formulas (5.12) and (5.13) we find that

$$\frac{t}{\mu} + \frac{\sigma^2 - \mu^2}{2\mu^2} = \lambda t \quad \text{and that} \quad \frac{\sigma^2 t}{\mu^3} = \lambda t,$$

that is, we have equalities (without remainder) as expected.

In the latter case we have $\mu = p, \sigma^2 = pq$ (and $d = 1$). Here, however, it is easier to see that $E\nu(n) = (n+1) \cdot (1/p)$ and that $\operatorname{Var} \nu(n) = (n+1) \cdot (q/p^2)$, since $\nu(n)$ equals the number of performances required in order to succeed more than n times (that is, $n+1$ times). Since $N(n) = \nu(n) - 1$ it follows that $EN(n) = ((n+1)/p) - 1$ and that $\operatorname{Var} N(n) = (n+1)q/p^2$. (Alternatively $N(n)$ equals the sum of n independent geometric variables with mean $1/p$ and one geometric variable with mean q/p.) Again the central limit theorem presents no problem and it is easy to check that

$$\frac{n}{\mu} + \frac{\sigma^2 - \mu^2}{2\mu^2} + \frac{1}{2\mu} = \frac{n+1}{p} - 1.$$

Finally,

$$\frac{\sigma^2 n}{\mu^3} = \frac{pqn}{p^3} = \frac{(n+1)q}{p^2} - \frac{q}{p^2},$$

that is, here we only have asymptotic equality with the leading term (but with a remainder which is much smaller than $o(n)$).

Finally, some remarks for the case when $\text{Var } X_1 = +\infty$.

A generalization of (5.12) and (5.14) under the assumption that $E|X_1|^r < \infty$ for some r $(1 < r < 2)$ has been proved in Täcklind (1944). The remainder then is $o(t^{2-r})$ as $t \to \infty$.

For nonarithmetic renewal processes it has been shown by Mohan (1976) that $EN(t) - (t/\mu)$ $(= U(t) - (t/\mu))$ is regularly varying with exponent $2 - \alpha$ iff F belongs to the domain of attraction of a stable law with index α $(1 < \alpha \le 2)$. Moreover, a generalization of (5.13) is obtained, see also Teugels (1968).

2.6 The Residual Lifetime

In view of the fact that the sequence $\{X_k, k \ge 1\}$ frequently is interpreted as a sequence of lifetimes it is natural to consider, in particular, the object (component) that is alive at time t. Its total lifetime is, of course, $X_{N(t)+1} = X_{\nu(t)}$. Of special interest is also *the residual lifetime*

$$R(t) = S_{N(t)+1} - t = S_{\nu(t)} - t. \tag{6.1}$$

In this section we present some asymptotic results.

As a first observation we note that, if $\text{Var } X_1 = \sigma^2 < \infty$, then, by Theorem 1.5.3, we have

$$ER(t) = ES_{\nu(t)} - t = \mu \left(E\nu(t) - \frac{t}{\mu} \right). \tag{6.2}$$

By combining this with Theorem 5.2(ii) and (3.11) the following result emerges.

Theorem 6.1. *Suppose that* $\text{Var } X_1 = \sigma^2 < \infty$.

(i) *If the renewal process is nonarithmetic, then*

$$ER(t) \to \frac{\sigma^2 + \mu^2}{2\mu} \quad as \quad t \to \infty. \tag{6.3}$$

(ii) *If the renewal process is d-arithmetic, then*

$$ER(nd) \to \frac{\sigma^2 + \mu^2}{2\mu} + \frac{d}{2} \quad as \quad n \to \infty. \tag{6.4}$$

This result indicates that $R(t)$, under appropriate conditions, may converge without normalization. That this is, indeed, the case is shown next.

Theorem 6.2. *Suppose that $0 < EX_1 = \mu < \infty$.*

(i) *If the renewal process is nonarithmetic, then, for $x > 0$, we have*

$$\lim_{t \to \infty} P(R(t) \leq x) = \frac{1}{\mu} \int_0^x (1 - F(s))ds. \qquad (6.5)$$

(ii) *If the renewal process is d-arithmetic, then, for $k = 1, 2, 3, \ldots$, we have*

$$\lim_{n \to \infty} P(R(nd) \leq kd) = \frac{d}{\mu} \sum_{j=0}^{k-1} (1 - F(jd)) \qquad (6.6)$$

or, equivalently,

$$\lim_{n \to \infty} P(R(nd) = kd) = \frac{d}{\mu} P(X_1 \geq kd). \qquad (6.7)$$

Proof. We only prove (i), the proof of (ii) being similar.

It is more convenient to consider the tail of the distribution. We have

$$P(R(t) > x) = \sum_{n=1}^{\infty} P(S_{n-1} \leq t, S_n > t + x)$$

$$= P(X_1 > t + x) + \sum_{n=2}^{\infty} \int_0^t P(X_n > t + x - s)dF_{n-1}(s)$$

$$= 1 - F(t + x) + \int_0^t (1 - F(t + x - s))dU(s)$$

$$\to 0 + \frac{1}{\mu} \int_0^\infty (1 - F(x + s))ds$$

$$= \frac{1}{\mu} \int_x^\infty (1 - F(s))ds \quad \text{as} \quad t \to \infty.$$

For the convergence we use the key renewal theorem (Theorem 4.3(i)) with $G(t) = 1 - F(t + x)$, $t > 0$, and the fact that $\int_0^\infty G(t)dt = \int_x^\infty (1 - F(s))ds \leq EX_1 < \infty$. □

Before we proceed, let us, as in Sections 2.4 and 2.5, consider the Poisson process and the Negative Binomial process.

Due to the lack of memory property it is clear that $R(t)$ is exponential with mean $\mu = \lambda^{-1}$ for all t. It is now easy to check that $\mu^{-1} \int_0^x (1 - F(s))ds = 1 - e^{-\lambda x}$ for $x > 0$ and that $(\sigma^2 + \mu^2)/2\mu = \lambda^{-1}$, that is, we have equality for all t in (6.3) and (6.5) as expected.

For the Negative Binomial process $R(n) = 1$ a.s. for all n. Since $\mu = p$ and $\sigma^2 = pq$ we have $\mu^{-1}P(X_1 \geq 1) = 1$ and $(\sigma^2 + \mu^2)/(2\mu) + \frac{1}{2} = 1$, that is, there is equality in (6.7) and (6.4) for all n.

In Theorem 6.1 we have seen how the expected value of the residual lifetime behaves asymptotically. The following result extends Theorem 6.1 to arbitrary moments; however, we confine ourselves to the nonarithmetic case. The proof is, essentially, due to Lai (1976), p. 65.

Theorem 6.3. *Suppose that $EX_1^r < \infty$ for some $r > 1$. If the renewal process is nonarithmetic, then*

$$E(R(t))^{r-1} \to \frac{1}{r\mu}EX_1^r \quad as \quad t \to \infty. \tag{6.8}$$

Proof. We have

$$R(t) = \sum_{n=1}^{\infty} I\{S_{n-1} \leq t\} \cdot (S_{n-1} + X_n - t)^+ \quad (t > 0), \tag{6.9}$$

where all terms but one equal 0. We can thus raise the sum to any power termwise. An elementary computation then shows that

$$E(R(t))^{r-1} = E((X_1 - t)^+)^{r-1} + \int_0^t G(t-s)dU(s), \tag{6.10}$$

where

$$G(y) = \int_y^{\infty} (u - y)^{r-1}dF(u) \tag{6.11}$$

and U is the renewal function. Since

$$\int_0^{\infty} G(y)dy = \frac{1}{r}EX_1^r < \infty, \tag{6.12}$$

an application of the key renewal theorem (Theorem 4.3(i)) yields

$$E(R(t))^{r-1} \to \frac{1}{\mu}\int_0^{\infty} G(y)dy \quad as \quad t \to \infty, \tag{6.13}$$

which, in view of 6.12, proves the theorem.   □

Remark 6.1. Note that, for $r = 2$, we rediscover (6.3).

Another quantity of interest is *the age* of the object that is alive at time t, $A(t) = t - S_{N(t)}$. However, since this object will not be considered in the sequel (it does not carry over in a reasonable sense to random walks), we only mention here that it can be shown that, in the nonarithmetic case, the limit distribution of the age is the same as that of the residual lifetime given in Theorem 6.2 (a minor modification is necessary in the arithmetic case).

We also note that $R(t)$ is, mathematically, a more pleasant object, since it is expressed in terms of a random quantity indexed by a stopping time, whereas $A(t)$ is not.

We finally mention that it is also possible to obtain the asymptotic distribution of the lifetime, $X_{\nu(t)} = A(t) + R(t)$, itself. One can, in fact, show that

$$\lim_{t\to\infty} P(X_{\nu(t)} \le x) = \frac{1}{\mu} \int_0^x s\,dF(s). \qquad (6.14)$$

2.7 Further Results

An exhaustive exposition of renewal theory would require a separate book. As mentioned in the introduction of this chapter our aim is to present a review of the main results, with emphasis on those which are most relevant with respect to the contents of this book. In this section we shall, however, briefly mention some further results and references.

2.7.1

Just as for the classical limit theorems one can ask for limit theorems for the case when the variance is infinite or even when the mean is infinite. However, one must then make more detailed assumptions about the tail of the distribution function F.

In Feller (1971), pp. 373–74 it is shown that a limit distribution for $N(t)$ exists if and only if F belongs to some domain of attraction—the limit laws are the stable distributions.

Concerning renewal theorems and expansions of $U(t)$ $(= EN(t))$ and Var $N(t)$, suppose first that Var $X_1 = +\infty$. Then Teugels (1968) shows, essentially, that if $1 - F(x) = x^{-\alpha} L(x)$, where $1 < \alpha < 2$ and L is slowly varying, then $U(t) - t/\mu = EN(t) - t/\mu$ varies regularly with exponent $2 - \alpha$ and Var $(N(t))$ varies regularly with exponent $3 - \alpha$ (see his Section 3). Mohan (1976) improves these results. He also considers the case $\alpha = 2$.

For the case $EX_1 = +\infty$ Teugels (1968) obtains asymptotic expansions for the renewal function under the assumption that the tail $1 - F$ satisfies $1 - F(x) = x^{-\alpha} \cdot L(x)$ as $x \to \infty$, where $0 \le \alpha \le 1$ and L is slowly varying; the result, essentially, being that U varies regularly with exponent α (see his Theorem 1). For $L(x) = $ const and $0 < \alpha < 1$ this was obtained by Feller (1949) in the arithmetic case. Erickson (1970) generalizes the elementary renewal theorem, Blackwell's renewal theorem and the key renewal theorem (Theorems 4.1–4.3 above) to this situation. For a further contribution, see Anderson and Athreya (1987).

Garsia and Lamperti (1962/63) and Williamson (1968) study the arithmetic case, the latter also in higher dimensions.

Finally, if the mean is infinite and F belongs to the domain of attraction of a (positive) stable law with index α, $0 < \alpha < 1$, then the limit distribution for $R(t)/t$ and $A(t)/t$ (the normalized residual lifetime and age, respectively) is given by a so-called generalized arc sine distribution, see Feller (1971), Section XIV.3; see also Dynkin (1955) and Lamperti (1958, 1961). For the case $\alpha = 1$, see Erickson (1970).

2.7.2

There are also other kinds of limit theorems which could be of interest for renewal processes, such as the Marcinkiewicz-Zygmund strong law, the law of the iterated logarithm, convergence of higher moments in the central limit theorem (Theorem 5.2(i)), remainder term estimates in the central limit theorem etc.

It turns out that several such results have not been established separately for renewal processes; the historic development was such that renewal theory had been extended to renewal theory for random walks (on the whole real line) in such a way that proofs for the random walk case automatically also covered the corresponding theorems for renewal processes. In other words, the generalization to random walks came first.

We remark, however, that Feller (1949) presents a law of the iterated logarithm for renewal counting processes in the arithmetic case.

The following Berry–Esseen theorem for renewal counting processes, that is, a remainder term estimate in the central limit theorem, Theorem 5.2(i), is due to Englund (1980).

Theorem 7.1. *Suppose that $\mu = EX_1, \sigma^2 = \mathrm{Var}\ X_1$ and $\gamma^3 = E|X_1 - \mu|^3$ are all finite. Then*

$$\sup_n \left| P(N(t) < n) - \Phi\left(\frac{(n\mu - t)\sqrt{\mu}}{\sigma\sqrt{t}} \right) \right| \leq 4 \left(\frac{\gamma}{\sigma} \right)^3 \sqrt{\frac{\mu}{t}}. \tag{7.1}$$

This result is mentioned mainly because it has not yet been extended to the random walk case.

Some large deviation results, that is, limit theorems for the ratio of the tail of the normalized distribution function of $N(t)$ and the tail of the standard normal distribution, are obtained in Serfozo (1974).

2.7.3

In this subsection we briefly mention two generalizations.

Consider Example 2.1. It is completely reasonable to assume that the initial component is not new, but has been used before. This amounts to the assumption that X_1 has a distribution different from F. Such a process is called a *delayed renewal process*. The renewal processes discussed above are then called *pure renewal processes*.

Another generalization is to allow defective distributions, that is, distributions such that $F(\infty) < 1$. The defect, $1 - F(\infty)$, corresponds to the probability of termination. Such processes are called *terminating*, or *transient renewal processes*. There are important applications of such processes, for example in insurance risk theory, see e.g. Asmussen (2000, 2003).

2.7.4

It is possible to develop a renewal theory for random walks (on the whole real line). We shall, in fact, do so in Chapter 3 (for the case $EX_1 > 0$).

2.7.5

Some attention has also been devoted to renewal theory for Markov chains, or Markov renewal theory; see Smith (1955), Çinlar (1975), Chapter 10 and Asmussen (2003). Some further references in this connection are Kemperman (1961), Spitzer (1965), Kesten (1974), Lalley (1984b, 1986) and papers by Alsmeyer and coauthors listed in the bibliography. In some of these references more general state spaces are considered. For a few lines on this topic, see Section 6.13 below.

2.7.6

We conclude by mentioning multidimensional or multivariate renewal theory. Some references are Chung (1952), Farrell (1964, 1966), Bickel and Yahav (1965), Doney (1966), Stam (1968, 1969, 1971), Hunter (1974a,b, 1977), Nagaev (1979), Gafurov (1980), Carlsson (1982) and Carlsson and Wainger (1984).

2.8 Random Walks; Introduction and Classifications

Let $\{X_k, k \geq 1\}$ be a sequence of i.i.d. random variables and set $S_n = \sum_{k=1}^n X_k$, $n \geq 0$, (where $S_0 = 0$). The sequence $\{S_n, n \geq 0\}$ is called a *random walk*. Throughout we ignore the trivial case $X_k = 0$ a.s., $k \geq 1$.

One of the basic problems concerning random walks is to investigate the asymptotics as $n \to \infty$, that is, to investigate "where" the random walk is after a "long time." All the classical limit theorems provide some answer to this problem. Another way of studying random walks is to investigate "how often" a given point or interval is visited, in particular, if there are finitely many visits or infinitely many visits and if the answer differs for different points or intervals.

Example 8.1. The simple random walk. Here $P(X_k = +1) = p$ and $P(X_k = -1) = q = 1 - p$ $(0 \leq p \leq 1)$.

Set $u_n = P(S_n = 0)$. It is well known (and/or easy to see) that

$$u_{2n} = \binom{2n}{n} p^n q^n \quad \text{and} \quad u_{2n-1} = 0 \quad (n \geq 1). \tag{8.1}$$

Thus, if $p \neq q$, then $\sum_{n=1}^{\infty} u_n < \infty$ and it follows from the Borel–Cantelli lemma that

$$P(S_n = 0 \text{ i.o.}) = 0 \quad \text{for } p \neq q. \tag{8.2}$$

If $p = q = \frac{1}{2}$, then $\sum_{n=1}^{\infty} u_n = +\infty$ and it follows (but *not* from the Borel–Cantelli lemma) that

$$P(S_n = 0 \text{ i.o.}) = 1 \quad \text{for } p = q = \frac{1}{2}. \tag{8.3}$$

A similar computation would show that the same dichotomy holds for returns to any integer point on the real line.

The different behaviors for $p \neq q$ and $p = q$ divide the simple random walks into two separate kinds. An analogous characterization can be made for all random walks. To this end we generalize the renewal function introduced in Section 2.3 as follows. We define the *renewal measure*

$$U\{I\} = \sum_{n=0}^{\infty} P(S_n \in I), \tag{8.4}$$

where $I \subset (-\infty, \infty)$ typically is an interval; here it is more convenient to let the summation start with $n = 0$ (recall Remarks 3.2 and 3.3).

For renewal processes we thus have, in particular, that $U\{[0, t]\} = 1 + U(t) < \infty$ for all t, that is $U\{I\} < \infty$ for all finite I. For random walks, which are not concentrated on one of the half axes, the series may, however, diverge even when I is a finite interval. This is far from obvious in general (see however Example 8.1 with $p = q = \frac{1}{2}$ and $I = \{0\}$) and gives, in fact, rise to the characterization to follow.

First, however, we note the following result.

Theorem 8.1.

(i) *If the random walk is nonarithmetic, then either $U\{I\} < \infty$ for every finite interval I or else $U\{I\} = +\infty$ for all intervals I.*
(ii) *If the random walk is arithmetic with span d, then either $U\{I\} < \infty$ for every finite interval I or else $U\{I\} = +\infty$ for all intervals I containing a point in the set $\{nd \colon n = 0, \pm 1, \pm 2, \dots\}$.*

This naturally leads to the following definition.

Definition 8.1. *A random walk is called transient if $U\{I\} < \infty$ for all finite intervals I and recurrent otherwise.*

Remark 8.1. Define the indicator variables $\{I_n, n \geq 0\}$ such that $I_n = 1$ on $\{S_n \in I\}$ and $I_n = 0$ otherwise and set $\Lambda\{I\} = \sum_{n=0}^{\infty} I_n$. Then $\Lambda\{I\}$ equals the *actual number of visit* to I made by the random walk (including S_0) and $U\{I\} = E\Lambda\{I\}$, that is, $U\{I\}$ equals the *expected number of visits* to I (cf. Remark 3.5).

We now return to the simple random walk from Example 8.1 and note that

$$U\{0\} = \sum_{n=0}^{\infty} u_n. \tag{8.5}$$

It thus follows from our computations there that $U\{0\} < \infty$ $(= \infty)$ when $p \neq q$ $(p = q)$. In view of Theorem 8.1 and Definition 8.1 it thus follows that the simple random walk is transient when $p \neq q$ and recurrent when $p = q = \frac{1}{2}$.

Moreover, it follows from the strong law of large numbers that $S_n \xrightarrow{\text{a.s.}} +\infty$ as $n \to \infty$ when $p > q$ and that $S_n \xrightarrow{\text{a.s.}} -\infty$ as $n \to \infty$ when $p < q$. In these cases the random walk drifts to $+\infty$ and $-\infty$, respectively. If $p = q = \frac{1}{2}$ it follows from the law of the iterated logarithm that $\limsup_{n \to \infty} S_n = +\infty$ a.s. and that $\liminf_{n \to \infty} S_n = -\infty$ a.s. In this case the random walk oscillates (between $+\infty$ and $-\infty$).

It turns out that these characterizations are not typical for simple random walks only; in fact, most of the above facts remain true for arbitrary random walks and all of them remain valid if, in addition, EX_1 is assumed to exist (that is, if $E|X_1| < \infty$ or $EX_1^- < \infty$ and $EX_1^+ = +\infty$ or $EX_1^- = \infty$ and $EX_1^+ < \infty$). We collect these facts as follows.

Theorem 8.2. *Let $\{S_n, n \geq 0\}$ be a random walk. Then exactly one of the following cases holds:*

(i) *The random walk drifts to $+\infty$; $S_n \xrightarrow{\text{a.s.}} +\infty$ as $n \to \infty$. The random walk is transient;*
(ii) *The random walk drifts to $-\infty$; $S_n \xrightarrow{\text{a.s.}} -\infty$ as $n \to \infty$. The random walk is transient;*
(iii) *The random walk oscillates between $-\infty$ and $+\infty$; $-\infty = \liminf_{n \to \infty} S_n < \limsup_{n \to \infty} S_n = +\infty$ a.s. In this case the random walk may be either transient or recurrent.*

Theorem 8.3. *Suppose, in addition, that $\mu = EX_1$ exists.*

(i) *If $0 < \mu \leq +\infty$ the random walk drifts to $+\infty$;*
(ii) *If $-\infty \leq \mu < 0$ the random walk drifts to $-\infty$;*
(iii) *If $\mu = 0$ the random walk oscillates. Moreover, in this case it is recurrent.*

Remark 8.2. Theorem 8.3 (i) and (ii) follow immediately from the strong law of large numbers. The case $EX_1 = 0$ is harder.

The development so far is based on whether or not the measure $U\{I\}$ is finite, that is whether or not the *expected number of visits* to a given interval is finite or not (recall Remark 8.1). However, it was shown in Example 8.1 that, for the simple random walk, the *actual number of visits* to 0 is a.s. finite when $p \neq q$ and a.s. infinite when $p = q$. Now, if, for a general random walk, $U\{I\} = E\Lambda\{I\} < \infty$, where $\Lambda\{I\}$ is as defined in Remark 8.1, then, obviously, we have $\Lambda\{I\} < \infty$ a.s.

However, it is, in fact, possible to show that the converse also holds. Consequently,

$$U\{I\} < \infty \quad \Longleftrightarrow \quad \Lambda\{I\} < \infty \text{ a.s. for all finite } I. \tag{8.6}$$

Theorem 8.1 thus remains true with $U\{I\}$ replaced by $\Lambda\{I\}$ and transience and recurrence can, equivalently, be defined in terms of $\Lambda\{I\}$. Moreover, since $\Lambda\{I\} < \infty$ for all finite I iff $|S_n| \to +\infty$ as $n \to \infty$ the following result emerges.

Theorem 8.4. *The random walk is transient iff* $|S_n| \to \infty$ *a.s. as* $n \to \infty$.

It is also possible to characterize random walks according to the transience or recurrence of *points* (rather than intervals). We refer to Chung and Fuchs (1951) and Chung and Ornstein (1962). Just as for irreducible Markov chains one can show that all points are of the same kind (where "all" has the obvious interpretation in the arithmetic case). Furthermore, one can show that the interval characterization and the point characterization are equivalent.

Remark 8.3. Note that we, in fact, studied the transience/recurrence of the *point* 0 for the simple random walk. However, for arithmetic random walks any finite interval consists of finitely many points of the form $\{nd; d = 0, \pm1, \dots\}$ so the distinction between points and intervals only makes sense in the nonarithmetic case.

Some book references on random walks are (in alphabetical order) Chung (1974), Feller (1968, 1971), Prabhu (1965) and Spitzer (1976).

2.9 Ladder Variables

Let $\{S_n, n \geq 0\}$ be a random walk with i.i.d. increments $\{X_k, k \geq 1\}$. Set $T_0 = 0$ and define

$$T_1 = \min\{n \colon S_n > 0\},$$

$$T_k = \min\{n > T_{k-1} \colon S_n > S_{T_{k-1}}\} \quad (k \geq 2). \tag{9.1}$$

If no such n exists we set $T_k = +\infty$. The random variables thus defined are called the (*strong*) (*ascending*) *ladder epochs*. The random variables

$Y_k = S_{T_k}$, $k \geq 0$, are the corresponding (*strong*) (*ascending*) *ladder heights* (with $Y_0 = S_{T_0} = S_0 = 0$).

Further, define, for $k \geq 1$,

$$N_k = T_k - T_{k-1} \quad \text{and} \quad Z_k = Y_k - Y_{k-1} = S_{T_k} - S_{T_{k-1}}, \qquad (9.2)$$

provided $T_k < \infty$.

It follows from the construction that $\{(N_k, Z_k), k \geq 1\}, \{N_k, k \geq 1\}$ and $\{Z_k, k \geq 1\}$ are sequences of i.i.d. random variables. Moreover, $\{T_n, n \geq 0\}$ and $\{Y_n, n \geq 0\}$ are (possibly terminating) renewal processes; the former, in fact, being arithmetic with span 1.

Similarly, set $\overline{T}_0 = 0$ and define

$$\overline{T}_1 = \min\{n : S_n \leq 0\} \quad \text{and}$$

$$\overline{T}_k = \min\{n > \overline{T}_{k-1} : S_n \leq \overline{T}_{k-1}\} \quad (k \geq 2). \qquad (9.3)$$

We call $\{\overline{T}_k, k \geq 0\}$ the (*weak*) (*descending*) *ladder epochs*. The sequences $\{\overline{Y}_k, k \geq 0\}, \{\overline{N}_k, k \geq 1\}$ and $\{\overline{Z}_k, k \geq 1\}$ are defined in the obvious manner.

Recall that a random variable is *proper* if it is finite a.s. Otherwise it is *defective*.

Theorem 9.1.

   (i) *If the random walk drifts to $+\infty$, then T_1 and Y_1 are proper and $\overline{T}_1$ and $\overline{Y}_1$ are defective. Moreover, $ET_1 < \infty$;*

  (ii) *If the random walk drifts to $-\infty$, then T_1 and Y_1 are defective and $\overline{T}_1$ and $\overline{Y}_1$ are proper. Moreover, $E\overline{T}_1 < \infty$;*

 (iii) *If the random walk oscillates, then $T_1, Y_1, \overline{T}_1$ and $\overline{Y}_1$ are proper. Moreover, $ET_1 = E\overline{T}_1 = +\infty$.*

Theorem 9.2. *If, in addition, $\mu = EX_1$ exists, then (i), (ii) and (iii) correspond to the cases $0 < \mu \leq \infty$, $-\infty \leq \mu < 0$ and $\mu = 0$, respectively. Moreover,*

$$EY_1 < \infty \quad \text{and} \quad EY_1 = \mu ET_1 \quad \text{when } 0 < \mu < \infty \qquad (9.4)$$

$$E\overline{Y}_1 > -\infty \quad \text{and} \quad E\overline{Y}_1 = \mu E\overline{T}_1 \quad \text{when } -\infty < \mu < 0. \qquad (9.5)$$

Remark 9.1. The equations in (9.4) and (9.5) are, in fact, special cases of Theorem 1.5.3.

Remark 9.2. For the case $EX_1 = 0$ we recall Example 1.5.1—the simple symmetric random walk, for which $Y_1 = 1$ a.s. and $ET_1 = +\infty$.

Remark 9.3. If $\mu = EX_1 = 0$ and, moreover, $\sigma^2 = \text{Var } X_1 < \infty$, then $EY_1 < \infty$ and $E\overline{Y}_1 < \infty$. Furthermore,

$$\sigma^2 = -2EY_1 \cdot E\overline{Y}_1 \tag{9.6}$$

and

$$EY_1 = \frac{\sigma c}{\sqrt{2}} \quad \text{and} \quad E\overline{Y}_1 = -\frac{\sigma}{c\sqrt{2}}, \tag{9.7}$$

where $0 < c = \exp\{\sum_{n=1}^{\infty} \frac{1}{n}(\frac{1}{2} - P(S_n > 0))\} < \infty$.

These results are due to Spitzer (1960, 1976). The absolute convergence of the sum was proved by Rosén (1961), see also Gut (2001), page 414.

For further results on the moments of the ladder heights we refer to Lai (1976), Chow and Lai (1979) and Doney (1980, 1982).

The case $EX_1 > 0$ will be further investigated in Chapter 3.

Remark 9.4. Weak ascending and strong descending ladder variables can be defined in the obvious manner.

2.10 The Maximum and the Minimum of a Random Walk

For a random walk $\{S_n, n \geq 0\}$ we define the partial maxima, $\{M_n, n \geq 0\}$, by

$$M_n = \max\{0, S_1, S_2, \ldots, S_n\} \tag{10.1}$$

and the partial minima, $\{m_n, n \geq 0\}$, by

$$m_n = \min\{0, S_1, S_2, \ldots, S_n\}. \tag{10.2}$$

In this section we show how these sequences can be used to characterize a random walk and in Section 2.12 we present some general limit theorems for M_n.

If the random walk drifts to $+\infty$, then, since $M_n \geq S_n$, it follows from Theorem 8.2 that $M_n \xrightarrow{\text{a.s.}} +\infty$ as $n \to \infty$. Furthermore, with probability 1, there is a last (random) epoch at which the random walk assumes a negative value. Thus

$$m = \min_{n \geq 0} S_n > -\infty \text{ a.s.} \tag{10.3}$$

Since $\{m_n, n \geq 0\}$ is nonincreasing it follows that $m_n \to m$ monotonically as $n \to \infty$. These facts, together with a similar (symmetric) argument for the case when the random walk drifts to $-\infty$, yield (i) and (ii) of the following theorem. The proof of (iii) is immediate.

Theorem 10.1.

(i) *If the random walk drifts to $+\infty$, then*

$$M_n \xrightarrow{\text{a.s.}} +\infty \quad \text{and} \quad m_n \xrightarrow{\text{a.s.}} m = \min_n S_n > -\infty \text{ a.s.} \quad as \quad n \to \infty;$$

(ii) *If the random walk drifts to $-\infty$, then*

$$M_n \xrightarrow{\text{a.s.}} M = \max_{n \geq 0} S_n < \infty \text{ a.s.} \quad \text{and} \quad m_n \xrightarrow{\text{a.s.}} -\infty \quad \text{as} \quad n \to \infty;$$

(iii) *If the random walk is oscillating, then*

$$M_n \xrightarrow{\text{a.s.}} +\infty \quad \text{and} \quad m_n \xrightarrow{\text{a.s.}} -\infty \quad \text{as} \quad n \to \infty.$$

If, in addition, $\mu = EX_1$ exists, then (i), (ii) and (iii) correspond to the cases $0 < \mu \leq +\infty$, $-\infty \leq \mu < 0$ and $\mu = 0$, respectively.

2.11 Representation Formulas for the Maximum

The sequence of partial maxima $\{M_n, n \geq 0\}$, defined in the previous section, describes the successive record values of a random walk. However, so does the sequence of strong ascending ladder heights. At every strong ascending ladder epoch there is a new record value, that is, the sequence of partial maxima and the sequence of strong ascending ladder heights both jump to a new, common, record value. Thus, each M_n equals some strong ascending ladder height.

To make this argument more stringent, let $\{N(n), n \geq 1\}$ be the counting process of the renewal process $\{T_n, n \geq 0\}$, generated by the strong ladder epochs. Thus,

$$N(n) = \max\{k : T_k \leq n\} \quad (n \geq 1), \tag{11.1}$$

or, equivalently, $N(n)$ equals the number of strong ascending ladder epochs in $[0, n]$.

The following lemma, which is due to Prabhu (1980), formalizes the relation between the sequence of partial maxima and the sequence of ladder heights.

Lemma 11.1. *We have*

$$M_n = Y_{N(n)} = \sum_{k=1}^{N(n)} Z_k = \sum_{k=1}^{N(n)} (S_{T_k} - S_{T_{k-1}}). \tag{11.2}$$

Remark 11.1. A completely analogous relation holds between the sequence of partial minima and the sequence of strong descending ladder heights.

The usefulness of this lemma lies in the fact that we have represented M_n as a sum of a random number of i.i.d. positive random variables, which permits us to apply the results obtained in Chapter 1 for such sums in order to prove results for M_n. Note, however, that $N(n)$ is not a stopping time, since $\{N(n), n \geq 1\}$ is a renewal counting process.

There is, however, another and much simpler way to derive some of these limit theorems, namely those concerning convergence in distribution or those which only involve probabilities and here the following lemma is useful.

Lemma 11.2. *We have*

$$(M_n, M_n - S_n) \overset{d}{=} (S_n - m_n, -m_n), \tag{11.3}$$

in particular,

$$M_n \overset{d}{=} S_n - m_n. \tag{11.4}$$

Remark 11.2. This is Prabhu (1980), Lemma 1.4.1. The basis for the proof is the fact that the original random walk has the same distributional properties as the so-called dual random walk. In particular,

$$(S_0, S_1, S_2, \ldots, S_n) \overset{d}{=} (S_n - S_n, S_n - S_{n-1}, \ldots, S_n - S_1, S_n - S_0). \tag{11.5}$$

Lemmas 11.1 and 11.2 are thus used in different ways. Whereas Lemma 11.1 gives an actual *representation of* M_n, Lemma 11.2 (only) defines two quantities which are *equidistributed*. Thus, if we are interested in the "sample function behavior" of $\{M_n, n \geq 0\}$ we must use Lemma 11.1, since two sequences of random variables with pairwise the same distribution need not have the same sample function behavior; for example, one may converge a.s. whereas the other does not. However, if we only need a "weak" result, then Lemma 11.2 is very convenient.

In the following section we shall use Lemmas 11.1 and 11.2 to prove (or indicate the proofs of) some general limit theorems for M_n. In Section 4.4 we shall exploit the lemmas more fully for random walks with positive drift.

We conclude this section by showing that we can obtain a representation formula for $M = \max_{n \geq 0} S_n$ for random walks drifting to $-\infty$ by letting $n \to \infty$ in (11.2).

Thus, suppose that the random walk drifts to $-\infty$; this is, for example, the case when $EX_1 < 0$. We then know from Theorem 10.1(ii) that $M_n \xrightarrow{\text{a.s.}} M = \max_{n \geq 0} S_n$ as $n \to \infty$, where M is a.s. finite. Furthermore, $N(n) \xrightarrow{\text{a.s.}} N < \infty$ a.s. as $n \to \infty$, where N equals the number of strong ascending ladder epochs (note that, in fact, $N(n) = N$ for n sufficiently large). This, together with Theorem 1.2.4, establishes the first part of the following result, see Janson (1986), Lemma 1.

Lemma 11.3. *If the random walk drifts to $-\infty$, then*

$$M = \max_{n \geq 0} S_n = Y_N = \sum_{k=1}^{N} Z_k = \sum_{k=1}^{N} (S_{T_k} - S_{T_{k-1}}), \tag{11.6}$$

where N equals the number of strong ascending ladder epochs and

$$P(N = n) = p(1 - p)^n, \ n \geq 0, \quad \text{and} \quad p = P(T_1 = +\infty) = P(M = 0) > 0.$$

Moreover, the conditional distribution of Z_k given that $N \geq k$ is independent of k and $Z_1, \ldots, Z_{k-1}$ and Z_k are independent given that $N \geq k$.

Proof. Formula (11.6) was proved above. For the remaining part of the proof we observe that $\{T_k, k \geq 1\}$ is a sequence of defective stopping times and that the random walk after T_k is independent of the random walk up to T_k on the set $\{T_k < \infty\}$. Thus

$$P(N \geq k+1 | N \geq k) = P(T_{k+1} < \infty | T_k < \infty) = P(T_1 < \infty) = 1 - p. \quad (11.7)$$

By rewriting this as

$$P(N \geq k+1) = P(T_1 < \infty)P(N \geq k), \quad (11.8)$$

it follows that N has a geometric distribution as claimed. Also, since $P(N < \infty) = 1$, we must have $p = P(T_1 = +\infty) > 0$. The final claim about independence follows as above (cf. (11.7)). $\qquad \square$

Remark 11.3. A consequence of Lemma 11.3 is that, for a random walk drifting to $-\infty$ we can represent the maximum, M, as follows:

$$M \overset{d}{=} \sum_{k=1}^{N} U_k, \quad (11.9)$$

where $\{U_k, k \geq 1\}$ is a sequence of i.i.d. random variables, which, moreover, is independent of N. Furthermore, the distribution of U_1 equals the conditional distribution of Y_1 $(= S_{T_1})$ given that $T_1 < \infty$ and the distribution of N is geometric as above.

Remark 11.4. Results corresponding to Lemmas 11.1 and 11.3 also hold for the *weak* ascending ladder epochs. Since $P(S_{T_1} > 0) \geq P(X_1 > 0) > 0$ it follows that the sums corresponding to (11.2) and (11.6) consist of a geometric number of zeroes followed by a positive term, followed by a geometric number of zeroes etc. Equivalently, the terms in (11.2) and (11.6) consist of the positive terms in the sums corresponding to the weak ascending ladder epochs.

Remark 11.5. A result analogous to Lemma 11.3 also holds for the minimum $m = \min_{n \geq 0} S_n$ if the random walk drifts to $+\infty$.

2.12 Limit Theorems for the Maximum

Having characterized random walks in various ways we shall now proceed to prove some general limit theorems for the partial maxima of a random walk. We assume throughout that the mean exists (in the sense that at least one of the tails has finite expectation). All results for partial maxima have obvious counterparts for the partial minima, which we, however, leave to the reader.

Theorem 12.1 (The Strong Law of Large Numbers). *Suppose that* $-\infty \leq \mu = EX_1 \leq \infty$. *We have*

$$\frac{M_n}{n} \xrightarrow{\text{a.s.}} \mu^+ \quad as \quad n \to \infty. \tag{12.1}$$

Proof. Suppose first that $0 \leq \mu \leq \infty$. Since $M_n \geq S_n$ it follows from the strong law of large numbers that

$$\liminf_{n \to \infty} \frac{M_n}{n} \geq \mu \text{ a.s.} \tag{12.2}$$

If $\mu = +\infty$ there is nothing more to prove, so suppose that $0 \leq \mu < \infty$. It remains to prove that

$$\limsup_{n \to \infty} \frac{M_n}{n} \leq \mu \text{ a.s.} \tag{12.3}$$

Now, let $\varepsilon > 0$ and choose $\omega \in \Omega$ such that

$$\limsup_{n \to \infty} \frac{M_n(\omega)}{n} > \mu + \varepsilon. \tag{12.4}$$

Then there exists a subsequence $\{n_k, k \geq 1\}$ tending to infinity such that

$$\frac{M_{n_k}(\omega)}{n_k} > \mu + \varepsilon \quad (k \geq 1). \tag{12.5}$$

Define $\tau_{n_k} = \min\{n \colon S_n = M_{n_k}\}$. Since $\tau_{n_k} \leq n_k$ it follows from (12.4) (suppressing ω) that

$$\frac{S_{\tau_{n_k}}}{\tau_{n_k}} = \frac{M_{n_k}}{\tau_{n_k}} \geq \frac{M_{n_k}}{n_k} > \mu + \varepsilon. \tag{12.6}$$

However, in view of the strong law of large numbers the set of ω such that this is possible must have probability 0, that is, (12.4) is a.s. impossible and (12.3) follows. This completes the proof for the case $0 \leq \mu < \infty$.

Finally, suppose that $-\infty \leq \mu < 0$. We then know from Theorem 10.1 that M_n converges a.s. to an a.s. finite random variable, M. This immediately implies that

$$\frac{M_n}{n} \xrightarrow{\text{a.s.}} 0 = \mu^+ \quad as \quad n \to \infty \tag{12.7}$$

and we are done. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

Remark 12.1. We just wish to point out that if $0 < EX_1 < \infty$ we can use Lemma 11.1 to give an alternative proof of Theorem 12.1 by arguing as follows.

By the representation (11.2) we have

$$\frac{M_n}{n} = \frac{Z_1 + \cdots + Z_{N(n)}}{n}, \tag{12.8}$$

where, again, $N(n) = \max\{k \colon T_k \leq n\}$.

Now, from Theorem 9.2 we know that $EZ_1 = EY_1 < \infty$. From renewal theory (Theorem 5.1(i)) we thus conclude that

$$\frac{N(n)}{n} \xrightarrow{\text{a.s.}} \frac{1}{ET_1} \quad \text{as} \quad n \to \infty, \tag{12.9}$$

which, together with Theorem 1.2.3(iii), yields

$$\frac{Z_1 + \cdots + Z_{N(n)}}{n} \xrightarrow{\text{a.s.}} EZ_1 \cdot \frac{1}{ET_1} \quad \text{as} \quad n \to \infty. \tag{12.10}$$

The conclusion now follows from Theorem 1.5.3 (or directly from formula (9.4)).

Remark 12.2. The same proof also works when $EX_1 = 0$ provided we also assume that Var $X_1 < \infty$. This is necessary to ensure that $EY_1 = EZ_1 < \infty$.

Next we assume that, in addition, Var $X_1 < \infty$. The resulting limit laws are different for the cases $EX_1 = 0$ and $EX_1 > 0$, respectively, and we begin with the former case.

Theorem 12.2. *Suppose that $EX_1 = 0$ and that $\sigma^2 = $ Var $X_1 < \infty$. Then*

$$\frac{M_n}{\sqrt{n}} \xrightarrow{d} |N(0,1)| \quad as \quad n \to \infty. \tag{12.11}$$

Sketch of Proof. By Lemma 11.1 we have (cf. (12.8))

$$\frac{M_n}{\sqrt{n}} = \frac{Z_1 + \cdots + Z_{N(n)}}{N(n)} \cdot \frac{N(n)}{\sqrt{n}}. \tag{12.12}$$

The idea now is that, since $EZ_1 < \infty$ (cf. Remark 9.3), it follows from Theorem 1.2.3 that the first factor in the RHS of (12.12) converges a.s. to EZ_1 as $n \to \infty$. Furthermore, one can show that

$$\frac{N(n)}{c\sqrt{2n}} \xrightarrow{d} |N(0,1)| \quad as \quad n \to \infty. \tag{12.13}$$

where c is the constant in Remark 9.3. The conclusion then follows from Cramér's theorem and the fact that $EZ_1 = EY_1 = \sigma c/\sqrt{2}$ (recall (9.7)).   □

We now turn to the case $EX_1 > 0$.

Theorem 12.3 (The Central Limit Theorem). *Suppose that $0 < \mu = EX_1 < \infty$ and that $\sigma^2 = $ Var $X_1 < \infty$. Then*

$$\frac{M_n - n\mu}{\sigma\sqrt{n}} \xrightarrow{d} N(0,1) \quad as \quad n \to \infty. \tag{12.14}$$

Proof. By Lemma 11.2 we have

$$\frac{M_n - n\mu}{\sigma\sqrt{n}} \stackrel{d}{=} \frac{S_n - n\mu}{\sigma\sqrt{n}} - \frac{m_n}{\sigma\sqrt{n}}. \tag{12.15}$$

Since $m_n \stackrel{\text{a.s.}}{\longrightarrow} m > -\infty$ a.s. as $n \to \infty$ (Theorem 10.1) it follows that

$$\frac{m_n}{\sqrt{n}} \stackrel{\text{a.s.}}{\longrightarrow} 0 \quad \text{as} \quad n \to \infty. \tag{12.16}$$

This, together with the ordinary central limit theorem and Cramér's theorem, shows that the RHS (and hence the LHS) of (12.15) converges in distribution to the standard normal distribution as $n \to \infty$, which proves the theorem. $\square$

Theorem 12.1 is part of a more general theorem (see Theorem 4.4.1(i)) due to Heyde (1966). The present proof is due to Svante Janson. Theorems 12.2 and 12.3 are due to Erdős and Kac (1946) and Wald (1947), respectively. Chung (1948) proves both results with a different method (and under additional assumptions). The proofs presented here are due to Prabhu (1980), Section 1.5.

In Section 4.4 we present further limit theorems for the maximum of random walks with positive drift.

3

Renewal Theory for Random Walks with Positive Drift

3.1 Introduction

Throughout this chapter $(\Omega, \mathscr{F}, P)$ is a probability space on which everything is defined, $\{S_n, n \geq 0\}$ is a random walk with positive drift, that is, $S_0 = 0$, $S_n = \sum_{k=1}^{n} X_k$, $n \geq 1$, where $\{X_k, k \geq 1\}$ is a sequence of i.i.d. random variables, and $S_n \xrightarrow{\text{a.s.}} +\infty$ as $n \to \infty$. We assume throughout this chapter, unless stated otherwise, that $0 < EX_1 = \mu < \infty$ (recall Theorems 2.8.2 and 2.8.3). In this section, however, no such assumption is necessary.

From Chapter 2 we recall that in renewal theory one is primarily concerned with

$$N(t) = \max\{n \colon S_n \leq t\} \quad (t \geq 0), \tag{1.1}$$

but that it sometimes is more convenient to study *the first passage time beyond the level t*, that is

$$\nu(t) = \min\{n \colon S_n > t\} \quad (t \geq 0). \tag{1.2}$$

We also found that, for renewal processes,

$$\nu(t) = N(t) + 1, \tag{1.3}$$

so the asymptotic behavior of $\nu(t)$ and $N(t)$ as $t \to \infty$ is the same. On the other hand, the first passage time is mathematically a more convenient object because it is a stopping time (relative to the sequence $\{\mathscr{F}_n, n \geq 1\}$, where $\mathscr{F}_n = \sigma\{X_1, \ldots, X_n\} = \sigma\{S_1, \ldots, S_n\}$, see Theorem 1.1(ii) below), whereas the number of renewals is not.

In the present context, when the summands also may assume negative values, relation (1.3) is no longer true, since the sequence of partial sums may very well drop below the level t after the first passage. Indeed, it turns out that in this situation $N(t)$ may be "much" larger than $\nu(t)$; a slightly more precise way to express this fact is that the requirements for existence of moments are higher for $N(t)$ than for $\nu(t)$ (see Section 3.3).

Note also that $\nu(t)$ gives information about when the random walk "reaches" t and that $N(t)$ tells us when the random walk "leaves" t on its way to $+\infty$.

Another important difference between renewal processes and random walks on the whole real line is that whereas many original proofs of limit theorems for renewal counting processes are based on inversion, that is, the relation $\{N(t) \geq n\} = \{S_n \leq t\}$ (recall formula (2.3.4) and e.g. Remark 2.5.2), this is not possible in the random walk context; just as (1.3) is no longer true. For random walks one can only conclude that

$$\{S_n \leq t\} \subset \{N(t) \geq n\}. \tag{1.4}$$

However, by considering the partial maxima introduced in Section 2.10, that is,

$$M_n = \max\{0, S_1, S_2, \ldots, S_n\} \quad (n \geq 0) \tag{1.5}$$

it is easily seen that one has an inverse relationship between first passage time processes and partial maxima, namely

$$\{\nu(t) > n\} = \{M_n \leq t\}. \tag{1.6}$$

With the aid of this observation and an appropriate limit theorem for M_n some limit theorems for $\nu(t)$ were proved, thus paralleling the approach used for renewal processes. Unfortunately, though, the limit theorems for partial maxima were not easily obtained then .... This eventually lead to the approach of this book, namely, to use results for stopped random walks. Recall that we have already used this method in Section 2.5 to prove the strong law and the central limit theorem for renewal counting processes. Since we have argued in Chapter 2 that limit theorems for partial maxima can be easily obtained we stress here, to prevent accusations for contradicting statements, the fact that the methods used there were not available (known) at that time; Lemma 2.11.2 appears for the first time in this connection in Prabhu (1980) and the approach to use Lemma 2.11.1 involves in itself results for stopped random walks.

The earliest reference for this idea seems to be Rényi (1957), who used it in a related context. The first more systematic investigation of first passage times using this method was Gut (1974a), where also more general first passage times were studied (we return to these in Section 4.5).

Note also that by using a limit theorem for first passage times together with (1.6) one obtains a(nother) proof of a corresponding limit theorem for the partial maxima; recall, however, the sometimes more powerful methods to prove results about partial maxima presented in Section 2.12 (see also Section 4.4 below).

In this chapter we shall study the first passage times defined by (1.2). Whereas $\nu(t)$ always has moments of all orders in the renewal theoretic case (Section 2.3) this is no longer true in the present context. We begin by establishing necessary and sufficient conditions for the existence of moments of $\nu(t)$

and $S_{\nu(t)}$ (Theorem 3.1). After that we derive all the classical limit theorems for $\nu(t)$, such as strong laws (Section 3.4), the central limit theorem (Theorem 5.1), convergence of moments (Section 3.8) and the law of the iterated logarithm (Theorem 11.1). We further devote Section 3.10 to the overshoot or the excess over the boundary, $S_{\nu(t)} - t$. Some of the results in this chapter thus generalize corresponding ones for renewal processes, obtained in Chapter 2 whereas others (such as the Marcinkiewicz-Zygmund law) have not been studied earlier for such processes.

In Chapter 4 we extend the results to some two-dimensional random walks and present several applications. We also study first passage times across more general boundaries. An extension to what is called perturbed random walks will be dealt with in Chapter 6.

For the proofs we shall, as mentioned above, exploit the theory of stopped random walks as developed in Chapter 1. More specifically, for several proofs concerning theorems about $\nu(t)$ we shall apply results from Chapter 1 to $S_{\nu(t)} - \nu(t)\mu = \sum_{k=1}^{\nu(t)} (X_k - \mu)$ and then prove that $S_{\nu(t)}$ is "sufficiently close" to t, from which the desired conclusion will follow.

There is also another important way to derive some of the claims, namely to combine the results from Chapter 1 with the fact that the statement in focus is already known to hold for renewal processes (see Chapter 2). The idea here is to use *ladder variables* in order to extend the validity to the general case. This method seems to have been introduced by Blackwell (1953) in order to extend his renewal theorem, Theorem 2.4.2(i), to Theorem 6.6(i) below. For a survey of how that method has been used and for some further examples we refer to Gut (1983a).

As a very first observation we have

Theorem 1.1.

(i) $P(\nu(t) < \infty) = 1$;
(ii) $\nu(t)$ *is a stopping time relative to* $\{\mathscr{F}_n, n \geq 1\}$, *where*

$$\mathscr{F}_n = \sigma\{X_k, k \leq n\} = \sigma\{S_k, k \leq n\};$$

(iii) $\nu(t) \to \infty$ *as* $t \to \infty$.

Proof. (i) follows from the fact that $S_n \xrightarrow{\text{a.s.}} \infty$ as $n \to \infty$.
(ii) holds because

$$\{\nu(t) = n\} = \{S_1 \leq t, S_2 \leq t, \ldots, S_{n-1} \leq t, S_n > t\} \in \mathscr{F}_n.$$

(iii) follows from (1.6). $\qquad\square$

Remark 1.1. Note that X_n is independent of $\mathscr{F}_{n-1}$, $n \geq 1$, where $\mathscr{F}_0 = \{\varnothing, \Omega\}$. This is essential for the application of some of the results in Chapter 1, see e.g. the introductions to Sections 1.5 and 1.6.

Before we present our results for $\nu(t)$ and $S_{\nu(t)}$ we describe the ladder variable method in the next section.

We also mention that for the proofs of the results concerning almost sure and distributional behavior of $\nu(t)$ no use is made of the fact that $\nu(t)$ is a stopping time and those statements also hold for $N(t)$ (after minor modifications of the proofs of the result for $\nu(t)$). However, to prove the results concerning uniform integrability and moment convergence it is essential that $\nu(t)$ is indeed a stopping time (recall Section 1.1).

3.2 Ladder Variables

Ladder variables of random walks were introduced in Section 2.9 and used there to obtain some classification and characterization theorems. For the sake of completeness we first recall the definitions and then proceed to make two observations which will be of great importance in the sequel.

For a random walk $\{S_n, n \geq 0\}$ we define the (*strong*) (*ascending*) *ladder epochs* $\{T_k, k \geq 0\}$ as follows: Set $T_0 = 0$ and let

$$T_k = \min\{n > T_{k-1}: S_n > S_{T_{k-1}}\} \quad (k \geq 1). \tag{2.1}$$

In particular, $T_1 = \min\{n: S_n > 0\}$. The random variables $Y_k = S_{T_k}$, $k \geq 0$, are the corresponding (*strong*) (*ascending*) *ladder heights* ($Y_0 = 0$). Note in particular that $T_1 = \nu(0)$ and that $Y_1 = S_{T_1} = S_{\nu(0)}$.

Finally, let, for $k \geq 1$,

$$N_k = T_k - T_{k-1} \quad \text{and} \quad Z_k = Y_k - Y_{k-1} = S_{T_k} - S_{T_{k-1}}. \tag{2.2}$$

It follows from the assumptions that $\{N_k, k \geq 1\}, \{Z_k, k \geq 1\}$ and $\{(N_k, Z_k), k \geq 1\}$ are sequences of i.i.d. random variables and that $\{T_n, n \geq 0\}$ and $\{Y_n, n \geq 0\}$ are *renewal processes*.

In Section 3.3 (cf. Theorem 2.9.2) we shall prove that $ET_1 < \infty$ and $EY_1 < \infty$, which, together with Theorem 1.5.3, yields

$$0 < \mu_H = \mu_T \cdot \mu < \infty, \tag{2.3}$$

where

$$\mu_H = EY_1 = EZ_1 \quad \text{and} \quad \mu_T = ET_1 = EN_1. \tag{2.4}$$

Next, we define

$$\nu_H(t) = \min\{n: Y_n > t\} \quad (t \geq 0), \tag{2.5}$$

that is, $\nu_H(t)$ is the first passage time across the level t for the ladder height process, that is, for the renewal process $\{Y_n, n \geq 0\}$.

Consequently, *all results for renewal counting processes obtained in Chapter 2 are applicable to* $\{\nu_H(t), t \geq 0\}$. Note also that $\nu_H(t) = \max\{n: Y_n \leq t\} + 1$ (cf. formula (1.3) above). This is our first important observation.

The second point is that *a first passage always must occur at a ladder epoch*.

We therefore have

$$S_{\nu(t)} = Y_{\nu_H(t)} = \sum_{k=1}^{\nu_H(t)} Z_k \quad (t \geq 0) \tag{2.6}$$

and

$$\nu(t) = T_{\nu_H(t)} = \sum_{k=1}^{\nu_H(t)} N_k \quad (t \geq 0), \tag{2.7}$$

that is, $\nu(t)$ can be expressed as a stopped random walk. Moreover, the stopping is caused by a stopping time (relative to $\{\sigma\{Z_k, k \leq n\}, n \geq 1\}$, which is such that Z_n is independent of $\sigma\{Z_k, k \leq n-1\}$, cf. Remark 1.1) and so the results from Chapter 1 apply.

Although the ladder variable method is very attractive there are limitations. In Janson (1983), Section 3, an example is given which shows that the method is not applicable in general if $\{X_k, k \geq 1\}$ is m-dependent.

3.3 Finiteness of Moments

The first result one establishes in renewal theory (and in sequential analysis) is the finiteness of all moments of the counting variable (the first passage time). In the renewal theoretic context this is achieved by truncation and comparison with a Bernoulli random walk (see Theorem 2.3.1). Here we cannot use such arguments; truncating the negative tail decreases the first passage time—the situation is more complicated. In order to prove a corresponding result we shall use results from Section 1.5.

The following theorem is due to Gut (1974a), see Theorem 2.1 there. (For integral r and $t = 0$ (i) was proved earlier in Heyde (1964).)

Recall from Theorem 1.1 that $\nu(t)$ is a proper random variable.

Theorem 3.1. *Let $r \geq 1$. We have*

(i) $$E(X_1^-)^r < \infty \quad \Longleftrightarrow \quad E(\nu(t))^r < \infty;$$

(ii) $$E(X_1^+)^r < \infty \quad \Longleftrightarrow \quad E(S_{\nu(t)})^r < \infty.$$

Proof of Sufficiencies. Let $r = 1$ and suppose that $P(X_1 \leq M) = 1$ for some $M > 0$. By definition

$$S_{\nu(t)} > t \quad \text{and} \quad S_{\nu(t)-1} \leq t \tag{3.1}$$

which, since $S_{\nu(t)} = S_{\nu(t)-1} + X_{\nu(t)}$, implies that

$$t < S_{\nu(t)} \le t + X_{\nu(t)}. \tag{3.2}$$

Next, define

$$\nu_n(t) = \nu(t) \wedge n. \tag{3.3}$$

If $\omega \in \Omega$ is such that $\nu(t) < n$ then $S_{\nu_n(t)} = S_{\nu(t)} \le t + X_{\nu(t)} \le t + M$ a.s. and if ω is such that $\nu(t) \ge n$ it follows that $S_{\nu_n(t)} \le t$. Consequently, we always have

$$S_{\nu_n(t)} \le t + M \text{ a.s.} \tag{3.4}$$

By taking expected values and applying Theorem 1.5.3 (or formula (1.5.15)), we obtain

$$\mu \cdot E\nu_n(t) = ES_{\nu_n(t)} \le t + M, \tag{3.5}$$

which, by Fatou's lemma (see e.g. Gut (2007), Theorem 2.5.2), yields

$$E\nu(t) \le \liminf_{n \to \infty} E\nu_n(t) \le \mu^{-1}(t + M) < \infty. \tag{3.6}$$

For the general case we define (cf. the proof of Theorem 2.4.1)

$$X_n' = X_n I\{X_n \le M\}, \quad S_n' = \sum_{k=1}^{n} X_k' \quad (n \ge 1) \tag{3.7}$$

and set

$$\nu'(t) = \min\{n \colon S_n' > t\} \quad (t \ge 0), \tag{3.8}$$

where $M > 0$ is so large that $\mu' = EX_1' > 0$.

Clearly $X_n' \le X_n$ and $S_n' \le S_n$ for all n. This means that the original sum always reaches the level t before the truncated sum, that is, we must have

$$\nu(t) \le \nu'(t), \tag{3.9}$$

and, in particular, that $E\nu(t) \le E\nu'(t)$, which is finite by (3.6). We have thus proved the sufficiency in (i) for $r = 1$.

The following lemma will be useful in the sequel.

Lemma 3.1. *Let $r \ge 1$. If $E(X_1^+)^r < \infty$, then*

$$EX_{\nu(t)} \le (EX_{\nu(t)}^r)^{1/r} \le (E\nu(t) \cdot E(X_1^+)^r)^{1/r} < \infty. \tag{3.10}$$

Proof. Immediate from Lemma 1.8.1 and Remark 1.8.2, since $P(X_{\nu(t)} > 0) = 1$ and $E\nu(t) < \infty$. □

The next step is to prove the sufficiency in (ii). From (3.2), Minkowski's inequality (see e.g. Gut (2007), Theorem 3.2.6) and Lemma 3.1 we have

$$\|S_{\nu(t)}\|_r \le t + \|X_{\nu(t)}\|_r \le t + (E\nu(t))^{1/r} \cdot \|X_1^+\|_r < \infty, \tag{3.11}$$

which proves the desired conclusion.

We now turn to the sufficiency in (i) for $r > 1$ and here Theorem 1.5.5 will be useful.

We first assume, as above, that $P(X_1 \le M) = 1$ for some $M > 0$. This implies, in particular, that $E|X_1|^r < \infty$ and that, by (ii), $E|S_{\nu(t)}|^r = E(S_{\nu(t)})^r < \infty$. An application of Theorem 1.5.5 now shows that $E(\nu(t))^r < \infty$.

To remove the boundedness of the upper tail we argue as in the case $r = 1$. Thus, assume that $E(X_1^-)^r < \infty$ and truncate as in (3.7). By (3.9) we then obtain

$$E(\nu(t))^r \le E(\nu'(t))^r < \infty \tag{3.12}$$

and the sufficiencies are proved.

Proof of Necessities. (i) When $r = 1$ there is nothing to prove, so suppose that $E(\nu(t))^r < \infty$ for some $r > 1$. By partial summation and the fact that

$$\{M_n \le t\} = \{\nu(t) > n\}, \tag{3.13}$$

(recall (1.6)) the assumption is the same as

$$\sum_{n=1}^{\infty} n^{r-1} P(M_n \le t) < \infty. \tag{3.14}$$

Next we observe that, for $\delta > 0$, $t > 0$, $n \ge 1$, we have

$$\{M_n \le t\} \supset \{X_1 < -n\delta, X_2 \le t + n\delta, X_2 + X_3$$

$$\le t + n\delta, \ldots, X_2 + \cdots + X_n \le t + n\delta\}, \tag{3.15}$$

that is,

$$\{M_n \le t\} \supset \left\{X_1 < -n\delta, \max_{2 \le k \le n} (S_k - X_1) \le t + n\delta\right\}. \tag{3.16}$$

By the i.i.d. assumption it therefore follows that

$$P(M_n \le t) \ge P(X_1 < -n\delta) \cdot P(M_{n-1} \le t + n\delta). \tag{3.17}$$

Since $M_n/n \xrightarrow{p} \mu$ as $n \to \infty$ (cf. Theorem 2.12.1) we can choose δ above such that

$$P(M_n \le n\delta) \ge \frac{1}{2}, \quad \text{say, for all } n \ge 1. \tag{3.18}$$

By (3.14), (3.17) and (3.18) it now follows that

$$\infty > \sum_{n=1}^{\infty} n^{r-1} P(X_1 < -n\delta) \cdot \frac{1}{2} = \frac{1}{2} \sum_{n=1}^{\infty} n^{r-1} P(-X_1 > n\delta), \qquad (3.19)$$

and since the finiteness of the last sum is equivalent to $E(X_1^-)^r < \infty$ the proof is complete.

(ii) Since

$$S_{\nu(t)} \geq X_1^+ \qquad (3.20)$$

we obtain

$$\infty > E(S_{\nu(t)})^r \geq E(X_1^+)^r \qquad (3.21)$$

and we are done. $\qquad\square$

Remark 3.1. As mentioned above, historically Theorem 3.1(i) was first proved for $t = 0$ and integral r (Heyde (1964)) by comparing with (3.14) (for $t = 0$). Later it has been shown (see the proof of Heyde (1966), Theorem 3) that if (3.14) holds for $t = 0$ and integral r, then it also holds for $t > 0$. Such conclusions are now easy, given the results from Chapter 1.5. Namely, suppose that Theorem 3.1(i) has been proved for $t = 0$ (that is, for the ladder epochs). Since $E(\nu_H(t))^r < \infty$ for all $r > 0$ (recall Theorem 2.3.1) it follows from (2.7) and Theorem 1.5.2 that, for $t > 0$, we have

$$E(\nu(t))^r = E\left(\sum_{k=1}^{\nu_H(t)} N_k\right)^r \leq B_r' E N_1^r \cdot E(\nu_H(t))^r < \infty. \qquad (3.22)$$

We have thus proved that

$$E(\nu(0))^r < \infty \quad \Longrightarrow \quad E(\nu(t))^r < \infty \quad (t > 0). \qquad (3.23)$$

In view of (3.13) this is the same as

$$\sum_{n=1}^{\infty} n^{r-1} P(M_n \leq 0) < \infty \quad \Longrightarrow \quad \sum_{n=1}^{\infty} n^{r-1} P(M_n \leq t) < \infty \quad (t > 0),$$

which establishes Theorem 3 of Heyde (1966) for all $r \geq 1$.

The converse implications are, of course, trivial.

Remark 3.2. If $E(X_1^-)^r < \infty$ and $\mu = +\infty$ the truncation procedure used in the proof of Theorem 2.4.1 (and Theorem 3.1 above) shows that $E(\nu(t))^r < \infty$ also in this case.

Remark 3.3. Concerning the necessity in (i) for $\mu = +\infty$, Smith (1967), pp. 304–305, gives an example where $E(\nu(0))^{l-1} < \infty$ for a given l and $E(X_1^-) = E(X_1^+) = +\infty$. This example can easily be modified so that, for any $p < l - 1$ we have $E(X_1^-)^r < \infty$ for $r < p$ and $E(X_1^-)^r = \infty$ for $r > p$.

In Theorem 3.1 the moments of X_1^- determined the degree of integrability for $\nu(t)$, whereas the moments of X_1^+ were the relevant object when we dealt with the integrability of $S_{\nu(t)}$. Let us just remark that this is completely reasonable. Since $S_{\nu(t)}$ lies between X_1^+ and $t + X_{\nu(t)} = t + X_{\nu(t)}^+$, it is intuitively obvious that the integrability of X_1^+, determines that of $S_{\nu(t)}$ (provided $E\nu(t) < \infty$). As for $\nu(t)$, it is clear that if X_1^+ has a low degree of integrability this does not "hurt" $\nu(t)$, since, on the contrary, very large positive X-values help to reach the level t more quickly (which is precisely the effect of the truncation procedure (3.7) (see (3.9)).

It is also possible to obtain a result like Theorem 3.1 for the moment generating function of $\nu(t)$, see Heyde (1964), Theorem 1. For renewal counting processes, recall Theorem 2.3.1(iii).

Theorem 3.2. *There exists $s_0 > 0$ such that $Ee^{s\nu(t)} < \infty$ for $|s| < s_0$ iff there exists $s_1 > 0$ such that $Ee^{sX_1^-} < \infty$ for $|s| < s_1$.*

We conclude this section with a few words about the process $\{N(t), t \geq 0\}$. Recall from above that for renewal processes $N(t) + 1 = \nu(t)$ and so knowledge about the behavior of one of the processes gives complete information about the other. Also, in the introduction of this chapter we mentioned that, in the present context, $N(t)$ may be much larger than $\nu(t)$. To make this a little more precise, suppose that $E(N(t))^r < \infty$ for some $r \geq 1$. Then, since

$$\{N(t) \geq n\} \supset \{S_n \leq t\} \tag{3.24}$$

(recall (1.4)), it follows from the assumption that

$$\infty > \sum_{n=1}^{\infty} n^{r-1} P(N(t) \geq n) \geq \sum_{n=1}^{\infty} n^{r-1} P(S_n \leq t). \tag{3.25}$$

Now, by Smith (1967), Theorem 7 (see also Janson (1986), Theorem 1), the finiteness of the last sum for $t = 0$ implies that $E(X_1^-)^{r+1} < \infty$ (provided $E|X_1| < \infty$). Thus, since $P(S_n \leq 0) \leq P(S_n \leq t)$, it follows that

$$E(N(t))^r < \infty \implies E(X_1^-)^{r+1} < \infty, \tag{3.26}$$

which implies that one extra moment of X_1^- is necessary for the integrability of $N(t)$ as compared to $\nu(t)$.

It turns out that the arrow actually goes both ways in (3.26), see e.g. Janson (1986), Theorem 1. The following result is a part of that theorem, which also treats several other related quantities.

Theorem 3.3. *Let $r > 0$. The following are equivalent:*

   (i) $E(X_1^-)^{r+1} < \infty$;
   (ii) $E(N(t))^r < \infty$;
   (iii) $E(\Lambda(t))^r < \infty$,

where

$$\Lambda(t) = \sum_{n=1}^{\infty} I\{S_n \leq t\} = \mathrm{Card}\{n \geq 1 \colon S_n \leq t\}. \tag{3.27}$$

Remark 3.4. For renewal processes, we have $N(t) = \Lambda(t)$, that is, (ii) and (iii) are, in fact, identical statements (and not only equivalent). For random walks, however, we can only conclude that $\Lambda(t) \leq N(t)$. Note that we also have $\Lambda(t) \geq \nu(t) - 1$.

Remark 3.5. After the definition in formula (2.3.2) of $\{N(t), t \geq 0\}$, the renewal counting process, we mentioned the alternative interpretation $\mathrm{Card}\{n \geq 1 \colon S_n \leq t\}$, which "justified" the word counting in the name of the process. For renewal processes the alternative interpretation thus is equivalent to the fact, mentioned in the previous remark, that $N(t) = \Lambda(t)$ in that case.

However, for random walks the counting is performed by the family $\{\Lambda(t), t \geq 0\}$. It thus seems more adequate to call this process the counting process in the present context.

Remark 3.6. A comparison with the counting measure $\Lambda\{I\}$ introduced in Remark 2.8.1 shows that $\Lambda\{(-\infty, t]\} = 1 + \Lambda(t)$ for $t \geq 0$. The extra one here is a consequence of the fact that $I\{S_0 \in (-\infty, t]\}$ is counted in the sum there but not here, that is, $\Lambda\{(-\infty, t]\} = \mathrm{Card}\{n \geq 0 \colon S_n \leq t\}$, whereas $\Lambda(t) = \mathrm{Card}\{n \geq 1 \colon S_n \leq t\}$. The extra one would not have appeared if we had defined $N(t)$ and $U(t)$ in Section 2.3 as suggested in Remarks 2.3.2 and 2.3.3. However, in that case the extra one would have popped up somewhere else.

3.4 The Strong Law of Large Numbers

Several additive processes obey a strong law of large numbers. Such laws have been extended to many subadditive processes, starting with Kingman (1968). Since $\{\nu(t), t \geq 0\}$ is subadditive (cf. formula (2.5.7)) it is reasonable to expect that a strong law holds and, indeed, this is the case. The following theorem is due to Heyde (1966).

Theorem 4.1.
$$\frac{\nu(t)}{t} \xrightarrow{\text{a.s.}} \frac{1}{\mu} \quad as \quad t \to \infty.$$

Proof. One way to prove this is to use Theorem 1.2.3 to conclude that

$$\frac{S_{\nu(t)}}{\nu(t)} \xrightarrow{\text{a.s.}} \mu \quad \text{and} \quad \frac{X_{\nu(t)}}{\nu(t)} \xrightarrow{\text{a.s.}} 0 \quad as \quad t \to \infty \tag{4.1}$$

and to combine that with (3.2). Here we shall, instead, present a proof based on ladder variables.

Consider the ladder variables as defined in Section 3.2. By Theorem 3.1 we know that $\mu_T = ET_1 < \infty$ and that $\mu_H = EY_1 = ES_{T_1} < \infty$.

Since the ladder height process, $\{Y_n, n \geq 0\}$, is a renewal process it follows from Theorem 2.5.1(i) and Remark 2.5.1 that

$$\frac{\nu_H(t)}{t} \xrightarrow{\text{a.s.}} \frac{1}{\mu_H} \quad \text{as} \quad t \to \infty, \tag{4.2}$$

which together with Theorem 1.2.3(iii) shows that

$$\frac{\nu(t)}{t} = \frac{T_{\nu_H(t)}}{t} \xrightarrow{\text{a.s.}} \mu_T \cdot \frac{1}{\mu_H} = \frac{1}{\mu} \quad \text{as} \quad t \to \infty. \qquad \square$$

Remark 4.1. With the aid of the truncation used in the proof of Theorem 2.4.1 (cf. also (3.7)) it is easy to see that the theorem remains true when $\mu = +\infty$ in the sense that the limit equals 0.

As immediate corollaries we obtain the following strong laws for the stopped sum and the stopping summand.

Theorem 4.2.
$$\frac{S_{\nu(t)}}{t} \xrightarrow{\text{a.s.}} 1 \quad as \quad t \to \infty.$$

Theorem 4.3. *If $E(X_1^+)^r < \infty$ for some $r \geq 1$, then*

$$\frac{X_{\nu(t)}}{t^{1/r}} \xrightarrow{\text{a.s.}} 0 \quad as \quad t \to \infty. \tag{4.3}$$

Proofs. Theorem 4.2 follows from Theorems 4.1 and 1.2.3(iii) and Theorem 4.3 follows from Theorems 4.1 and 1.2.3(i), since $X_{\nu(t)} = X_{\nu(t)}^+$ (recall Remark 1.8.2). $\qquad \square$

Remark 4.2. Theorem 4.3 holds, in fact, also for $0 < r < 1$ (that is, if $EX_1^- < \infty$ and $E(X_1^+)^r < \infty$ for some r, $0 < r < 1$).

Just as there is the generalization due to Marcinkiewicz and Zygmund of the Kolmogorov strong law, there is a Marcinkiewicz-Zygmund law for first passage times too, see Gut (1974a), Theorem 2.8.a (but only for $1 \leq r < 2$, since the summands are supposed to have finite mean).

Theorem 4.4. *Let $1 \leq r < 2$ and suppose that $E|X_1|^r < \infty$. Then*

$$\frac{\nu(t) - \frac{t}{\mu}}{t^{1/r}} \xrightarrow{\text{a.s.}} 0 \quad as \quad t \to \infty. \tag{4.4}$$

Proof. For $r = 1$ the theorem reduces to Theorem 4.1, so there is nothing to prove. Suppose therefore that $r > 1$.

By Theorems 4.1 and 1.2.3(ii) we have

$$\frac{S_{\nu(t)} - \nu(t)\mu}{t^{1/r}} \xrightarrow{\text{a.s.}} 0 \quad \text{as} \quad t \to \infty. \tag{4.5}$$

Furthermore, by (3.2),

$$\frac{t - \nu(t)\mu}{t^{1/r}} < \frac{S_{\nu(t)} - \nu(t)\mu}{t^{1/r}} \leq \frac{t - \nu(t)\mu}{t^{1/r}} + \frac{X_{\nu(t)}}{t^{1/r}}. \tag{4.6}$$

The conclusion now follows from (4.5), (4.6) and Theorem 4.3.   □

Remark 4.3. If we only suppose that $E|X_1|^r < \infty$ for some $r < 1$, then the Marcinkiewicz-Zygmund law tells us that $S_n/n^{1/r} \xrightarrow{\text{a.s.}} 0$ as $n \to \infty$, and the above method yields $S_{\nu(t)}/(\nu(t))^{1/r} \xrightarrow{\text{a.s.}} 0$ and $X_{\nu(t)}/(\nu(t))^{1/r} \xrightarrow{\text{a.s.}} 0$ as $t \to \infty$, which together with (3.2) proves that $t/(\nu(t))^{1/r} \xrightarrow{\text{a.s.}} 0$ as $t \to \infty$ or, equivalently, that $t^r/\nu(t) \xrightarrow{\text{a.s.}} 0$ as $t \to \infty$. We therefore only obtain an estimate for how fast $\nu(t)$ tends to infinity. To obtain something better one has to make more detailed assumptions on the tails of the distribution of the summands (cf. Section 2.7.1).

Remark 4.4. Since the random indices in Theorem 1.2.3 do not have to be stopping times it follows, by modifying the proof of Theorem 4.1, that the strong law also holds for $N(t)$. We thus have

$$\frac{N(t)}{t} \xrightarrow{\text{a.s.}} \frac{1}{\mu} \quad \text{as} \quad t \to \infty. \tag{4.7}$$

For renewal processes this is Theorem 2.5.1(i).

Remark 4.5. Suppose that $E(X_1^-)^2 = +\infty$. It follows from Theorem 3.3 that $EN(t) = +\infty$. We are thus in the position that $\nu(t)$ and $N(t)$ both are approximately equal to t/μ. Yet, $E\nu(t) < \infty$ but $EN(t) = +\infty$. This means that, although $\nu(t)$ and $N(t)$ are, *in general*, of the same order of magnitude, $N(t)$ is *occasionally* much larger than $\nu(t)$.

Remark 4.6. A strong law for $\Lambda(t)$ (defined in (3.27)) is due to Lai (1975), who shows that

$$\frac{\Lambda(t)}{t} \xrightarrow{\text{a.s.}} \frac{1}{\mu} \quad \text{as} \quad t \to \infty. \tag{4.8}$$

A remark, similar to Remark 4.5 also applies to $\Lambda(t)$.

3.5 The Central Limit Theorem

The proof of the asymptotic normality of $\nu(t)$, suitably normalized, is due to Heyde (1967a), who used the inverse relationship (1.6) between the partial maxima and the first passage times, thus generalizing the corresponding proof for renewal counting processes (cf. Remark 2.5.2). Below we present the proof of Gut (1974a) (see Theorem 2.5 there), which is based on Anscombe's theorem (Theorem 1.3.1).

Theorem 5.1. *Suppose that Var $X_1 = \sigma^2 < \infty$. Then*

$$\frac{\nu(t) - t/\mu}{\sqrt{\sigma^2 \mu^{-3} t}} \xrightarrow{d} N(0,1) \quad as \quad t \to \infty. \tag{5.1}$$

Proof. By Theorem 1.3.1(ii) applied to the sequence $\{X_k - \mu, k \geq 1\}$ and with $\theta = (1/\mu)$ (recall Theorem 4.1) we have

$$\frac{S_{\nu(t)} - \mu\nu(t)}{\sigma\sqrt{\frac{t}{\mu}}} \xrightarrow{d} |N(0,1)| \quad as \quad t \to \infty. \tag{5.2}$$

Moreover, by Theorem 4.3 we have

$$\frac{X_{\nu(t)}}{\sqrt{t}} \xrightarrow{a.s.} 0 \quad as \quad t \to \infty \tag{5.3}$$

and since, by (3.2),

$$\frac{t - \mu \cdot \nu(t)}{\sigma\sqrt{t/\mu}} < \frac{S_{\nu(t)} - \mu \cdot \nu(t)}{\sigma\sqrt{t/\mu}} \leq \frac{t - \mu \cdot \nu(t)}{\sigma\sqrt{t/\mu}} + \frac{X_{\nu(t)}}{\sigma\sqrt{t/\mu}} \tag{5.4}$$

it follows from (5.2), (5.3) and Cramér's theorem that

$$\frac{t - \mu\nu(t)}{\sigma\sqrt{\frac{t}{\mu}}} \xrightarrow{d} N(0,1) \quad as \quad t \to \infty, \tag{5.5}$$

which, by taking the symmetry of the normal distribution into account, proves the theorem. $\square$

Remark 5.1. This proof differs somewhat from the corresponding proof of the central limit theorem for renewal counting processes, Theorem 2.5.2(i). This is, essentially, due to the fact that the relation $\nu(t) = N(t) + 1$ is no longer valid now. Although we have $S_{N(t)} \leq t < S_{\nu(t)}$ now we would have to *prove* that $(\nu(t) - N(t))/\sqrt{t} \xrightarrow{p} 0$ as $t \to \infty$ (for renewal processes this is, of course, completely trivial).

Remark 5.2. Just as for the strong law (cf. Remark 4.4) it follows, after minor modifications of the proof of Theorem 5.1, that the central limit theorem also holds for $N(t)$ (provided Var $X_1 < \infty$).

Theorem 5.1 can be generalized to the case when the distribution of the summands belongs to the domain of attraction of a stable law with index α $(1 < \alpha \leq 2)$. The following result is due to Heyde (1967b). A proof using the above method, with Theorem 1.3.2 replacing Theorem 1.3.1, can be found in Gut (1974a); see Theorem 2.9 there.

Theorem 5.2. *Suppose that $\{B_n, n \geq 1\}$ are positive normalizing coefficients such that*

$$\frac{S_n}{B_n} \xrightarrow{d} G_\alpha \quad as \quad n \to \infty, \tag{5.6}$$

where G_α is (a random variable distributed according to) a stable law with index α, $1 < \alpha \leq 2$. Then

$$\frac{\nu(t) - t/\mu}{\frac{1}{\mu}B(\frac{t}{\mu})} \xrightarrow{d} -G_\alpha \quad as \quad t \to \infty. \tag{5.7}$$

Here $B(y) = B_{[y]}$ for $y > 0$.

To prove this one first uses Theorem 1.3.2 to obtain the counterpart of (5.2) and then proves (see Gut (1974a), Lemma 2.9) that

$$\frac{B_{\nu(t)}}{B\left(\frac{t}{\mu}\right)} \xrightarrow{p} 1 \quad as \quad t \to \infty \tag{5.8}$$

$$\frac{X_{\nu(t)}}{B\left(\frac{t}{\mu}\right)} \xrightarrow{p} 0 \quad as \quad t \to \infty \tag{5.9}$$

to obtain the counterpart of (5.3).

Remark 5.3. In view of the regular variation of the normalizing coefficients (see e.g. Feller (1971), Chapter IX.8 or Gut (2007), Section 9.3) the conclusion (5.7) is equivalent to

$$\frac{\nu(t) - t/\mu}{\mu^{-(1+(1/\alpha))}B(t)} \xrightarrow{d} -G_\alpha \quad as \quad t \to \infty. \tag{5.10}$$

Another class of limit theorems are the so called *local* limit theorems (in contrast to, e.g. the central limit theorem, which is a *global* limit theorem). In the present context such a theorem amounts to finding an asymptotic expression for $P(\nu(t) = n)$ as t (and n) tend to infinity. If one assumes that $\sigma^2 = \mathrm{Var} X_1 < \infty$ and lets t and n tend to infinity in a suitable manner, then one can show that $P(\nu(t) = n)$ asymptotically behaves like the density of a normal distribution with mean t/μ and variance $\sigma^2 t/\mu^3$, see e.g. Aleškevičience (1975) and Lalley (1984a).

3.6 Renewal Theorems

Among the early results in renewal theory are the elementary renewal theorem due to Doob (1948) (Theorem 2.4.1) and Blackwell's (1948) renewal theorem (Theorem 2.4.2). In this section we present random walk analogs to those results and some other extensions.

Theorem 6.1.
$$\frac{E\nu(t)}{t} \to \frac{1}{\mu} \quad as \quad t \to \infty.$$

Proof. Consider the ladder variables as defined in Section 3.2. By Theorem 3.1(i) we have

$$\mu_T = ET_1 < \infty \quad and \quad \mu_H = ES_{T_1} = EY_1 < \infty. \tag{6.1}$$

It thus follows from Theorem 1.5.3 applied to (2.7), the elementary renewal theorem (Theorem 2.4.1) and (2.3) that

$$\frac{E\nu(t)}{t} = \frac{E\nu_H(t) \cdot ET_1}{t} = \frac{E\nu_H(t)}{t} \cdot \mu_T \to \frac{1}{\mu_H} \cdot \mu_T = \frac{1}{\mu} \quad as \quad t \to \infty, \tag{6.2}$$

which proves the desired conclusion. □

Remark 6.1. For $\mu = +\infty$ the theorem remains valid in the sense that $E\nu(t)/t \to 0$ in that case (cf. Remark 4.1). A similar remark applies to all theorems in this section.

In the following, we consider

$$V(t) = \sum_{n=0}^{\infty} P(M_n \le t). \tag{6.3}$$

By (3.13) we have
$$V(t) = E\nu(t), \tag{6.4}$$

which is finite by Theorem 3.1. Theorem 6.1 is thus equivalent to the following result, which we may view as a random walk analog to the elementary renewal theorem.

Theorem 6.2.
$$\frac{V(t)}{t} \to \frac{1}{\mu} \quad as \quad t \to \infty.$$

Our next theorem is a random walk analog to Blackwell's renewal theorem (cf. Theorem 2.4.2). For the first part of the theorem we also refer to Spitzer (1960), formula (0.8).

Theorem 6.3.

(i) *For nonarithmetic random walks we have*

$$V(t) - V(t-h) = \sum_{n=0}^{\infty} P(t-h < M_n \leq t) \to \frac{h}{\mu} \quad as \quad t \to \infty. \quad (6.5)$$

(ii) *For d-arithmetic random walks we have*

$$v_n = \sum_{k=0}^{\infty} P(M_k = nd) \to \frac{d}{\mu} \quad as \quad n \to \infty. \quad (6.6)$$

Proof. We confine ourselves to proving (i), the proof of (ii) being similar, and we only consider the case $0 < \mu < \infty$. Just as in the proof of Theorem 6.1 we have, keeping (6.4) in mind, that

$$V(t) = \mu_T \cdot E\nu_H(t). \quad (6.7)$$

Moreover, Blackwell's renewal theorem applied to the renewal process $\{Y_n, n \geq 0\}$ yields

$$E\nu_H(t) - E\nu_H(t-h) \to \frac{h}{\mu_H} \quad as \quad t \to \infty. \quad (6.8)$$

Thus,

$$V(t) - V(t-h) \to \mu_T \cdot \frac{h}{\mu_H} = \frac{h}{\mu} \quad as \quad t \to \infty. \quad (6.9)$$

$\square$

In addition to these random walk analogs to Theorems 2.4.1 and 2.4.2 there exist extensions in which it is shown that these theorems, as stated for renewal processes, remain valid for random walks. These results are presented next.

In the first one we are concerned with the (extended) renewal function

$$U(t) = U\{(-\infty, t]\} = \sum_{n=0}^{\infty} P(S_n \leq t). \quad (6.10)$$

Note that summation begins with $n = 0$ here and with $n = 1$ for renewal processes.

We begin, however, by observing that we cannot apply our results from Section 2.8 to conclude that $U(t)$ is finite, since $(-\infty, t]$ is an *infinite* interval. It follows, in fact, by Theorem A of Heyde (1964) that

$$U(t) < \infty \quad \Longleftrightarrow \quad E(X_1^-)^2 < \infty. \quad (6.11)$$

The following extension of Theorem 2.4.1 follows from Heyde (1966), Theorem 1.

Theorem 6.4. *Suppose that $E(X_1^-)^2 < \infty$. Then*

$$\frac{U(t)}{t} \to \frac{1}{\mu} \quad as \quad t \to \infty. \tag{6.12}$$

Remark 6.2. The theorem thus implies that $U(t) \approx t/\mu$ for large t. Much work has been devoted to what is called remainder term estimates in the renewal theorem, that is, to estimates on the difference $U(t) - t/\mu$. Some references are Stone (1965), Stone and Wainger (1967), Essén (1973), Daley (1980) and Carlsson (1983).

In view of (6.10) (and (3.27)) we have

$$U(t) = 1 + E\Lambda(t). \tag{6.13}$$

An equivalent formulation of Theorem 6.4 thus is

Theorem 6.5. *Suppose that $E(X_1^-)^2 < \infty$. Then*

$$\frac{E\Lambda(t)}{t} \to \frac{1}{\mu} \quad as \quad t \to \infty. \tag{6.14}$$

Note that this equivalence parallels that of Theorems 6.2 and 6.1.

The analogous extension of Theorem 2.4.2 is due to Chung and Pollard (1952) and Chung and Wolfowitz (1952) in the arithmetic case and to Blackwell (1953) in the nonarithmetic case; see also Prabhu (1965), Section 6.3.

Theorem 6.6.

(i) *For nonarithmetic random walks we have*

$$U\{(t-h,t]\} = \sum_{n=0}^{\infty} P(t-h < S_n \le t) \to \begin{cases} \frac{h}{\mu}, & as \ t \to +\infty, \\ 0, & as \ t \to -\infty. \end{cases} \tag{6.15}$$

(ii) *For d-arithmetic random walks we have*

$$u_n = \sum_{k=0}^{\infty} P(S_k = nd) \to \begin{cases} \frac{d}{\mu}, & as \ n \to +\infty, \\ 0, & as \ n \to -\infty. \end{cases} \tag{6.16}$$

Remark 6.3. Observe that Theorem 6.6 is a statement concerning the *renewal measure* for finite intervals in contrast to Theorems 6.3 and 6.4 which are statements about *renewal functions*. Note also that $U\{(t-h,t]\} = E\left(\sum_{n=0}^{\infty} I\{t-h < S_n \le t\}\right)$ is the expected number of partial sums falling into the strip $(t-h,t]$, which has a *finite* width, h; recall that the additional (necessary) moment assumption in Theorem 6.4 was due to the fact that the interval $(-\infty, t]$ was infinite.

Remark 6.4. Theorem 6.6(i) has been generalized to arbitrary nonarithmetic random walks by Feller and Orey (1961), see also Spitzer (1976), pp. 288–289 and Feller (1971), Section XI.9.

We conclude this section by defining the *harmonic renewal measure*

$$U_h\{I\} = \sum_{n=1}^{\infty} \frac{1}{n} P(S_n \in I), \tag{6.17}$$

for $I \subset (-\infty, \infty)$. For some results parallel to the elementary renewal theorem and Blackwell's renewal theorem we refer to Grübel (1986), and further references given there.

3.7 Uniform Integrability

In Section 3.3 we gave necessary and sufficient conditions for the finiteness of the moments of first passage times and stopped random walks and in Sections 3.4 and 3.5 we proved convergence results. Here we shall consider uniform integrability. The proofs will be based on our findings in Chapter 1 on uniform integrability for stopped random walks (recall Remark 1.1). In some instances we shall also use the ladder variable method discussed in Section 3.2.

The following theorem and its proof are due to Lai (1975).

Theorem 7.1. *Let $r \geq 1$. If $E(X_1^-)^r < \infty$, then the family*

$$\left\{ \left(\frac{\nu(t)}{t} \right)^r, t \geq 1 \right\} \quad \text{is uniformly integrable.} \tag{7.1}$$

Proof. From formula (2.5.6) we know that the theorem holds for renewal processes. Since the ladder height process is a renewal process it thus follows that

$$\left\{ \left(\frac{\nu_H(t)}{t} \right)^r, t \geq 1 \right\} \quad \text{is uniformly integrable.} \tag{7.2}$$

Furthermore, we know from Theorem 3.1(i) that

$$ET_1^r = E(\nu(0))^r < \infty. \tag{7.3}$$

We can thus apply Theorem 1.6.1 (with $X_k \leftrightarrow N_k$ and $N(t) \leftrightarrow \nu_H(t)$) to conclude that

$$\left\{ \left(\frac{T_{\nu_H(t)}}{t} \right)^r, t \geq 1 \right\} \quad \text{is uniformly integrable.} \tag{7.4}$$

which (recall (2.7)) is the same as (7.1). □

Remark 7.1. Theorem 7.1 also holds when $\mu = +\infty$ (truncate).

Next we prove the corresponding result for the stopping summand and the stopped sum.

Theorem 7.2. *Let* $r \geq 1$. *If* $E(X_1^+)^r < \infty$, *then*

(i) $$\left\{ \left(\frac{X_{\nu(t)}}{t^{1/r}} \right)^r, t \geq 1 \right\} \quad \text{is uniformly integrable;}$$

(ii) $$\left\{ \left(\frac{S_{\nu(t)}}{t} \right)^r, t \geq 1 \right\} \quad \text{is uniformly integrable.}$$

Proof. (i) follows immediately from Theorem 1.8.1 (apply Theorem 7.1 with $r = 1$ for (1.8.8)) and (ii) follows from (i) and the fact that

$$1 < \frac{S_{\nu(t)}}{t} \leq 1 + \frac{X_{\nu(t)}}{t} \leq 1 + \frac{X_{\nu(t)}}{t^{1/r}}. \qquad \square$$

The results so far cover what we need later to prove moment convergence in the strong law. The following two theorems cover the Marcinkiewicz–Zygmund laws and the central limit theorem, respectively.

Theorem 7.3. *Suppose that* $E|X_1|^r < \infty$ *for some* r $(1 \leq r < 2)$. *Then*

(i) $$\left\{ \left| \frac{S_{\nu(t)} - \mu\nu(t)}{t^{1/r}} \right|^r, t \geq 1 \right\} \quad \text{is uniformly integrable;}$$

(ii) $$\left\{ \left| \frac{\nu(t) - t/\mu}{t^{1/r}} \right|^r, t \geq 1 \right\} \quad \text{is uniformly integrable.}$$

Theorem 7.4. *Suppose that* $E|X_1|^r < \infty$ *for some* $r \geq 2$. *Then*

(i) $$\left\{ \left| \frac{S_{\nu(t)} - \mu\nu(t)}{\sqrt{t}} \right|^r, t \geq 1 \right\} \quad \text{is uniformly integrable;}$$

(ii) $$\left\{ \left| \frac{\nu(t) - t/\mu}{\sqrt{t}} \right|^r, t \geq 1 \right\} \quad \text{is uniformly integrable.}$$

Proofs. The statements (i) in both theorems follow, in view of Theorem 7.1, from Theorems 1.6.2 and 1.6.3, respectively.

To prove the conclusions (ii) we first observe that

$$\mu|\nu(t) - t/\mu| \leq |S_{\nu(t)} - \mu\nu(t)| + |S_{\nu(t)} - t| \leq |S_{\nu(t)} - \mu\nu(t)| + X_{\nu(t)}. \quad (7.5)$$

Now,

$$\left\{ \left| \frac{S_{\nu(t)} - \mu\nu(t)}{t^{1/r}} \right|^r, t \geq 1 \right\} \quad \text{and} \quad \left\{ \left| \frac{S_{\nu(t)} - \mu\nu(t)}{\sqrt{t}} \right|^r, t \geq 1 \right\}$$

are uniformly integrable by (i) and $\{(X_{\nu(t)}/t^{1/r})^r, t \geq 1\}$ is uniformly integrable by Theorem 7.2(i). An application of Lemma A.1.3 thus terminates the proof. For Theorem 7.4(ii) we also use the fact that

$$\left(\frac{X_{\nu(t)}}{\sqrt{t}}\right)^r = t^{1-r/2}\left(\frac{X_{\nu(t)}}{t^{1/r}}\right)^r \leq \left(\frac{X_{\nu(t)}}{t^{1/r}}\right)^r \qquad \text{for } t \geq 1. \qquad \square$$

Remark 7.2. Theorem 7.4(i), which we obtained from Theorem 1.6.3, is originally due to Chow, Hsiung and Lai (1979). Note that we, in view of Theorem 7.1, were able to apply their weaker version of Theorem 1.6.3 (cf. Remark 1.6.2).

3.8 Moment Convergence

In this section we use the results from earlier sections to prove moment convergence in the strong laws and the central limit theorem. Since everything we need has been prepared, the proofs are almost immediate (cf. Section 1.7).

Theorem 8.1. *Let $r \geq 1$ and suppose that $E(X_1^-)^r < \infty$. Then*

$$E\left(\frac{\nu(t)}{t}\right)^p \to \frac{1}{\mu^p} \quad as \quad t \to \infty \qquad for\ all\ p, 0 < p \leq r. \qquad (8.1)$$

Theorem 8.2. *Let $r \geq 1$ and suppose that $E(X_1^+)^r < \infty$. Then*

(i) $$E\frac{X_{\nu(t)}^r}{t} \to 0 \quad as \quad t \to \infty;$$

(ii) $$E\left(\frac{S_{\nu(t)}}{t}\right)^p \to 1 \quad as \quad t \to \infty \quad for\ all\ p, 0 \leq p \leq r.$$

Theorem 8.3. *Let $1 \leq r < 2$ and suppose that $E|X_1|^r < \infty$. Then*

$$E\left|\frac{\nu(t) - t/\mu}{t^{1/r}}\right|^r \to 0 \quad as \quad t \to \infty. \qquad (8.2)$$

Theorem 8.4. *Let $r \geq 2$ and suppose that $E|X_1|^r < \infty$. Set $\sigma^2 = \operatorname{Var} X_1$. Then*

(i) $$E\left|\frac{\nu(t) - t/\mu}{\sqrt{t}}\right|^p \to E|Z|^p \quad as \quad t \to \infty \quad for\ all\ p, 0 < p \leq r,$$

where Z is a normal random variable with mean 0 and variance σ^2/μ^3;

(ii) $$E\left(\frac{\nu(t) - t/\mu}{\sqrt{t}}\right)^k \to 0 \quad as \quad t \to \infty \quad for\ k = odd\ integer \leq r.$$

Proofs. Theorem 8.1 follows from Theorems 4.1, 7.1 and A.1.1. Alternatively, by (1.3), (2.7), Theorem 2.5.1, Remark 2.5.1, formula (2.5.6) and Theorem 1.7.1 if we wish to use the ladder variable method.

Theorem 8.2(i) follows from Theorems 4.3, 7.2(i) and A.1.1.

Theorem 8.2(ii) follows from Theorems 4.2, 7.2(ii) and A.1.1.

Theorems 8.3 and 8.4 follow similarly.   □

Theorem 8.1 is due to Gut (1974a), see Theorem 2.3.b there. For $r = 1$ and $r = 2$, however, the result was earlier proved in Chow and Robbins (1963) and Chow (1966), respectively. The proof in Gut (1974a) is based on a direct estimation of the moments (thus without using uniform integrability). We also mention that (8.1) also holds when $\mu = +\infty$ (cf. Remark 4.1).

Theorem 8.2 is due to Gut (1974a), see Lemma 2.4 and Theorem 2.3.d there. Again, direct proofs were used. Since the direct proof of (ii) is so short we reproduce it here.

For $r = 1$ the result is a consequence of Theorems 6.1 and 1.5.3. For $r > 1$ we have, by Lemma 3.1 and Theorem 6.1,

$$1 < \left\| \frac{S_{\nu(t)}}{t} \right\|_r \leq 1 + \left\| \frac{X_{\nu(t)}}{t} \right\|_r$$

$$\leq 1 + t^{-1+(1/r)} \left(E \frac{\nu(t)}{t} \right)^{1/r} \|X_1^+\|_r \to 1 \quad \text{as} \quad t \to \infty. \qquad (8.3)$$

This is, of course, a much more elementary proof.

Theorem 8.3 is due to Gut (1974c), where direct computations of the moments were used (cf. the proof of Theorem 1.7.4).

Theorem 8.4 is due to Chow, Hsiung and Lai (1979) for $r > 2$. For $r = 2$ the statement (i) amounts to the convergence of $\text{Var}\,\nu(t)$, which was proved earlier; see Theorem 9.1 below.

Remark 8.1. We know from Remark 5.2 that $N(t)$ is asymptotically normal with mean t/μ and variance $\sigma^2 t/\mu^3$ as $t \to \infty$ provided $\sigma^2 = \text{Var}\,X_1 < \infty$. However, if $E(X_1^-)^3 = +\infty$ we know from Theorem 3.3 that $EN(t) < \infty$, but $E(N(t))^2 = +\infty$, so in this case $\text{Var}\,N(t)/t$ does not converge when $t \to \infty$ (recall also Remark 1.6.3). A similar observation was made in Remark 4.5 in connection with the strong law.

Remark 8.2. As for $\Lambda(t) = \sum_{n=1}^{\infty} I\{S_n \leq t\}$, Lai (1975) proved that

$$E \left(\frac{\Lambda(t)}{t} \right)^r \to \frac{1}{\mu^r} \quad \text{as} \quad t \to \infty \qquad (r > 0), \qquad (8.4)$$

provided $E(X_1^-)^{r+1} < \infty$. In view of Theorem 3.3 we find that the moment assumption cannot be weakened.

3.9 Further Results on $E\nu(t)$ and $\operatorname{Var}\nu(t)$

The moments most frequently dealt with are the first and the second—the mean and the variance. Because of their special interest we begin by stating the following result, which is the same as Theorem 8.4 for the case $r = 2$. We then give a different proof for the convergence of $\operatorname{Var}\nu(t)$, due to Gut (1983a), based on the ladder variable method, after which we discuss some refinements for the expected value, due to Gut (1974a), Section 2. For corresponding results for renewal processes we refer to Section 2.5.

Theorem 9.1. *If* $\operatorname{Var} X_1 = \sigma^2 < \infty$, *then*

(i)
$$E\nu(t) = \frac{t}{\mu} + o(\sqrt{t}) \quad as \quad t \to \infty;$$

(ii)
$$\operatorname{Var}\nu(t) = \frac{\sigma^2 t}{\mu^3} + o(t) \quad as \quad t \to \infty.$$

Another Proof of (ii). In view of Theorems 5.1 and A.1.1 and Remarks A.1.1 and A.1.2 it suffices to show that

$$\left\{ \left(\frac{\nu(t) - \frac{t}{\mu}}{\sqrt{t}} \right)^2, t \geq 1 \right\} \quad \text{is uniformly integrable.} \tag{9.1}$$

To this end we note that (use (2.3) and (2.7))

$$\nu(t) - \frac{t}{\mu} = \nu(t) - \nu_H(t)\mu_T + \mu_T \left(\nu_H(t) - \frac{t}{\mu_H} \right)$$

$$= \sum_{k=1}^{\nu_H(t)} (N_k - EN_k) + \mu_T \left(\nu_H(t) - \frac{t}{\mu_H} \right). \tag{9.2}$$

By Lemma A.1.3 it thus suffices to show that

$$\left\{ \left(\frac{\sum\limits_{k=1}^{\nu_H(t)} (N_k - EN_k)}{\sqrt{t}} \right)^2, t \geq 1 \right\} \quad \text{is uniformly integrable} \tag{9.3}$$

and that

$$\left\{ \left(\frac{\nu_H(t) - \frac{t}{\mu_H}}{\sqrt{t}} \right)^2, t \geq 1 \right\} \quad \text{is uniformly integrable.} \tag{9.4}$$

Now, by arguing as in the proof of (7.2) it follows that $\{(\nu_H(t)/t), t \geq 1\}$ is uniformly integrable, so (9.3) follows from Theorem 1.6.3. Finally, (9.4) follows from the fact that (9.1) holds for renewal processes (formula (2.5.20)). $\square$

Convergence of the variance was first proved in Heyde (1967a) and later by Siegmund (1969), see also Chow, Robbins and Siegmund (1971), Theorem 2.5, in all cases by direct computations.

Convergence of $E\nu(t)$ holds, of course, also when the variance does not exist (see Theorem 6.1). However, since the variance exists here, the remainder is improved to $o(\sqrt{t})$.

For renewal processes, however, Theorem 9.1(i) has been obtained with a still better remainder term, see Theorem 2.5.2(ii). Below we shall use the ladder variable method to improve the remainder term similarly (cf. Gut (1974a), Section 2). Just as in Chapter 2 we have to treat the arithmetic and the nonarithmetic cases separately.

Theorem 9.2. *Suppose that the random walk is nonarithmetic and that* $E(X_1^+)^2 < \infty$. *Then*

$$E\nu(t) = \frac{t}{\mu} + \mu_T \frac{EY_1^2}{2\mu_H^2} + o(1) \quad as \quad t \to \infty. \tag{9.5}$$

Theorem 9.3. *Suppose that the random walk is d-arithmetic and that* $E(X_1^+)^2 < \infty$. *Then*

$$E\nu(nd) = \frac{nd}{\mu} + \mu_T \frac{EY_1^2}{2\mu_H^2} + \frac{d}{2\mu} + o(1) \quad as \quad n \to \infty. \tag{9.6}$$

Proof of Theorem 9.2. From Theorem 2.5.2(ii) (and formula (2.3.11)) we have

$$E\nu_H(t) = \frac{t}{\mu_H} + \frac{EY_1^2}{2\mu_H^2} + o(1) \quad as \quad t \to \infty, \tag{9.7}$$

(since the increments of the renewal process $\{Y_n, n \geq 0\}$ have finite mean and variance (see Theorem 3.1)). Since, by (2.7) and Theorem 1.5.3, we have

$$E\nu(t) = E\nu_H(t) \cdot \mu_T \tag{9.8}$$

the conclusion follows.   □

Proof of Theorem 9.3. The proof is the same, except that (9.7) has to be replaced by

$$E\nu_H(nd) = \frac{nd}{\mu_H} + \frac{EY_1^2}{2\mu_H^2} + \frac{d}{2\mu_H} + o(1) \quad as \quad n \to \infty. \qquad □$$

Remark 9.1. Observe that the refinements are obtained under the assumption that $E(X_1^+)^2 < \infty$. It is thus possible that $\text{Var}\,X_1 = +\infty$. The assumption $E(X_1^+)^2 < \infty$ was used only to conclude that the ladder heights have a finite second moment and, recall Theorem 3.1, $E(X_1^+)^2$ is the necessary and sufficient condition for this to be the case.

Since it may be difficult to actually compute the moments of the ladder heights we can use the fact that

$$0 < Y_1 \leq X_{N_1} = X^+_{\nu(0)}, \tag{9.9}$$

and Lemma 3.1 to conclude that

$$0 \leq EY_1^2 \leq E(X_{N_1})^2 \leq E\nu(0) \cdot E(X_1^+)^2 = \mu_T E(X_1^+)^2, \tag{9.10}$$

which, together with (2.3), implies that

$$0 \leq \mu_T \frac{EY_1^2}{2\mu_H^2} \leq \frac{E(X_1^+)^2}{2\mu^2}. \tag{9.11}$$

Since, moreover,

$$\mu \cdot E\nu(t) = ES_{\nu(t)} > t \tag{9.12}$$

the following bounds on $E\nu(t)$ are obtained (Gut (1974a), p. 288).

Theorem 9.4. *Suppose that* $E(X_1^+)^2 < \infty$. *Then*

(i) $$\frac{t}{\mu} < E\nu(t) \leq \frac{t}{\mu} + \frac{E(X_1^+)^2}{2\mu^2} + o(1) \quad as \quad t \to \infty$$

  in the nonarithmetic case;

(ii) $$\frac{nd}{\mu} \leq E\nu(nd) \leq \frac{nd}{\mu} + \frac{E(X_1^+)^2}{2\mu^2} + \frac{d}{2\mu} + o(1) \quad as \quad n \to \infty$$

  in the d-arithmetic case.

It is also possible to improve the remainder $o(t^{1/r})$ (which follows from Theorem 8.3) when $E|X_1|^r < \infty$ for some r ($1 < r < 2$). For renewal processes the following result is due to Täcklind (1944). The extension using ladder variables is due to Gut (1974a), Theorem 2.8. At the end of the next section we shall present an elegant proof of this result, due to Janson (1983), who used it to obtain a bound for the moments of the overshoot, $S_{\nu(t)} - t$.

Theorem 9.5. *Let* $1 < r < 2$ *and suppose that* $E(X_1^+)^r < \infty$. *Then*

$$E\nu(t) = \frac{t}{\mu} + o(t^{2-r}) \quad as \quad t \to \infty. \tag{9.13}$$

Remark 9.2. From Theorem 8.3 and Lyapounov's inequality (cf. e.g. Gut (2007), Theorem 3.2.5) we have $E\nu(t) = t/\mu + o(t^{1/r})$ as $t \to \infty$, but, since $2 - r < 1/r$ when $1 < r < 2$, (9.13) is actually sharper in this case.

3.10 The Overshoot

For renewal processes the quantity

$$R(t) = S_{\nu(t)} - t \quad (t \geq 0) \tag{10.1}$$

is generally called *the residual lifetime* (of the individual who is alive at time t), see Section 2.6; in the first passage time context one usually calls this quantity *the overshoot*, or *the excess over the boundary*, since it defines the amount with which the random walk surpasses the level t—the level at which it was to be stopped.

Theorem 10.1. *Let $r \geq 1$. Then*

$$E(X_1^+)^r < \infty \quad \Longleftrightarrow \quad E(R(t))^r < \infty. \tag{10.2}$$

Proof. Since $R(t)$ and $S_{\nu(t)}$ differ only by a constant, t, the result follows from Theorem 3.1. □

In the remainder of this section we shall consider limit theorems and, just as in Theorems 9.2 and 9.3, the arithmetic and nonarithmetic cases have to be separated occasionally. If so, we shall sometimes confine ourselves to treating the nonarithmetic case only. Some of the results below are also discussed in Woodroofe (1982), Chapter 2.

Since

$$0 < R(t) \leq X_{\nu(t)} \tag{10.3}$$

the following result is immediate from Theorems 4.3, 7.2(i) and 8.2(i).

Theorem 10.2. *Let $r \geq 1$ and suppose that $E(X_1^+)^r < \infty$. Then,*

(i) $\dfrac{R(t)}{t^{1/r}} \xrightarrow{\text{a.s.}} 0 \quad as \quad t \to \infty$;

(ii) *The family* $\left\{ \dfrac{(R(t))^r}{t}, t \geq 1 \right\}$ *is uniformly integrable;*

(iii) $E(R(t))^r = o(t) \quad as \quad t \to \infty$;

(iv) $ER(t) = o(t^{1/r}) \quad as \quad t \to \infty$.

This is a result on how the *normalized* overshoot converges. However, for renewal processes we found that, in fact, the overshoot (that is, the residual lifetime) converges *without normalization*. We shall now see that this is true also for random walks.

Theorem 10.3.

(i) *Suppose the random walk is nonarithmetic. Then, for $x > 0$, we have*

$$\lim_{t \to \infty} P(R(t) \leq x) = \frac{1}{\mu_H} \int_0^x P(S_{T_1} > y) dy. \tag{10.4}$$

(ii) *Suppose the random walk is d-arithmetic. Then, for $k = 1, 2, 3, \ldots,$ we have*

$$\lim_{n\to\infty} P(R(nd) \le kd) = \frac{d}{\mu_H} \sum_{j=0}^{k-1} P(S_{T_1} > jd) \tag{10.5}$$

or, equivalently,

$$\lim_{n\to\infty} P(R(nd) = kd) = \frac{d}{\mu_H} P(S_{T_1} \ge kd). \tag{10.6}$$

Proof. Recall from Section 3.2 that a first passage always must occur at a ladder epoch. It follows that the overshoots of the original random walk and those of the ladder height process must be the same, that is

$$R(t) = S_{\nu(t)} - t = Y_{\nu_H(t)} - t = R_H(t) \qquad (t \ge 0). \tag{10.7}$$

Since $\mu_H < \infty$ by Theorem 3.1 and since the limiting distribution of $R_H(t)$ is given as above (cf. Theorem 2.6.2) the result follows. □

Theorem 10.3(i) with this proof is due to Gut (1983a). Woodroofe (1976) proves, under some additional assumptions, that

$$\lim_{t\to\infty} P(R(t) \le x) = \frac{1}{\mu} \int_0^x P\left\{ \inf_{n\ge 1} S_n > y \right\} dy \quad \text{as} \quad t \to \infty. \tag{10.8}$$

A look at (10.4) and (10.8) suggests that the integrands might be the same. As it stands, they are the same provided the assumptions in Woodroofe (1976) are fulfilled. However, by an approximation argument, which we omit here, it is shown in Gut (1983a) that the integrands are, in fact, equal for *all* random walks.

Theorem 10.4. *For any random walk with positive mean we have*

(i) $$\frac{1}{\mu_H} P(Y_1 > y) = \frac{1}{\mu} P\left\{ \inf_{n\ge 1} S_n > y \right\} \qquad (y \ge 0);$$

(ii) $$\mu_T = \left(P\left\{ \inf_{n\ge 1} S_n > 0 \right\} \right)^{-1} \ge (P(X_1 > 0))^{-1} = E\nu_X,$$

where $\nu_X = \inf\{n : X_n > 0\} \, (\le \nu(0) = T_1)$.

Proof. (i) is the result mentioned above and (ii) follows from (i) by choosing $y = 0$; clearly ν_X is geometric with mean $(P(X_1 > 0))^{-1}$. □

Let us now turn to the problem of convergence of the moments of the overshoot. We first note that the moment requirement of X_1^+ in Theorem 10.2(iii) is of the *same order* as that of the conclusion. In the following we shall prove moment convergence for the *unnormalized* overshoot, thus extending

the results obtained for renewal processes in Section 2.6. However, the price for not normalizing is that we now need a higher degree of integrability for X_1^+. Just as in Section 2.6 we begin by considering the expected value of the overshoot. The general result is Theorem 10.9.

Our first result follows immediately from the relation

$$ER(t) = ES_{\nu(t)} - t = \mu E(\nu(t) - t/\mu) \tag{10.9}$$

(recall formula (2.6.2)) and Theorems 9.2 and 9.3. For renewal processes, see Theorem 2.6.1.

Theorem 10.5. *Suppose that* $E(X_1^+)^2 < \infty$. *Then*

(i)
$$\lim_{t\to\infty} ER(t) = \frac{EY_1^2}{2\mu_H} \quad \text{in the nonarithmetic case;}$$

(ii)
$$\lim_{n\to\infty} ER(nd) = \frac{EY_1^2}{2\mu_H} + \frac{d}{2} \quad \text{in the d-arithmetic case.}$$

By invoking (10.9) and Theorem 9.4 we obtain the following (weaker) asymptotic inequality, which on the other hand has the advantage that it only contains moments of the original random walk.

Theorem 10.6. *Suppose that* $E(X_1^+)^2 < \infty$. *Then*

(i)
$$ER(t) \leq \frac{E(X_1^+)^2}{2\mu} + o(1) \quad \text{as} \quad t \to \infty \text{ in the nonarithmetic case;}$$

(ii)
$$ER(nd) \leq \frac{E(X_1^+)^2}{2\mu} + \frac{d}{2} + o(1) \quad \text{as} \quad n \to \infty \text{ in the d-arithmetic case.}$$

Lorden (1970), Theorem 1, obtains the (weaker) bound $E(X_1^+)^2/\mu$ in the nonarithmetic case. His bound is, on the other hand, valid for all $t > 0$. For another proof of Lorden's inequality we refer to Carlsson and Nerman (1986).

In Woodroofe (1976), Theorem 4.4, it was shown that, if $\text{Var } X_1 = \sigma^2 < \infty$ and some additional assumptions are satisfied, then

$$ER(t) \to \frac{\sigma^2 + \mu^2}{2\mu} - \sum_{k=1}^{\infty} \frac{1}{k} E(S_k^-) \quad \text{as} \quad t \to \infty. \tag{10.10}$$

Just as for Theorem 10.4(i) above one can make an approximation argument (see Gut (1983a)) to show that the limiting relations given in Theorem 10.5(i) and (10.10) coincide for all random walks with finite variance. Since, furthermore, $\inf_{n\geq1} S_n$ is a proper random variable (see Theorem 2.10.1) and

$$E\left(\inf_{n\geq1} S_n\right)^- = \sum_{k=1}^{\infty} \frac{1}{k} E(S_k^-) \tag{10.11}$$

(see e.g. Chung (1974), p. 287, with X_n replaced by $-X_n$) we obtain the following result.

Theorem 10.7. *For any random walk with positive mean and finite variance we have*

$$0 \le E\left(\inf_{n \ge 1} S_n\right)^{-} = \sum_{k=1}^{\infty} \frac{1}{k} E(S_k^-) = \frac{EX_1^2}{2EX_1} - \frac{EY_1^2}{2EY_1}. \tag{10.12}$$

Remark 10.1. Note that, if $E(X_1^+)^2 < \infty$ but $EX_1^2 = +\infty$, then Theorem 10.5(i) holds, whereas (10.10) does not.

The following result corresponds to Theorem 9.5 above.

Theorem 10.8. *Suppose that $E(X_1^+)^r < \infty$ for some $1 < r < 2$. Then*

$$ER(t) = o(t^{2-r}) \quad as \quad t \to \infty. \tag{10.13}$$

Next we extend Theorem 2.6.3 to the present context (see Gut (1983a)). The proof is inspired by Lai (1976), p. 65, where the same problem is considered for the case $EX_1 = 0$ and integral moments. For $r = 2$ we rediscover Theorem 10.5(i).

Theorem 10.9. *Suppose that the random walk is nonarithmetic and that $E(X_1^+)^r < \infty$ for some $r > 1$. Then*

$$E(R(t))^{r-1} \to \frac{1}{r\mu_H} E(Y_1)^r \quad as \quad t \to \infty. \tag{10.14}$$

Proof. We know from Theorem 3.1(ii) that the assumption $E(X_1^+)^r < \infty$ implies that $EY_1^r < \infty$. By Theorem 2.6.3 the theorem thus holds for the ladder height process, that is,

$$E(R_H(t))^{r-1} \to \frac{1}{r\mu_H} EY_1^r \quad as \quad t \to \infty, \tag{10.15}$$

from which (10.14) follows, since $R(t) = R_H(t)$ (see (10.7)). $\qquad \square$

Remark 10.2. By modifying the arguments leading from (9.9) to (9.10) and then from Theorem 10.5 to Theorem 10.6 we obtain

$$E(R(t))^{r-1} \le \frac{E(X_1^+)^r}{r\mu} + o(1) \quad as \quad t \to \infty \tag{10.16}$$

in the nonarithmetic case, provided, again, that $E(X_1^+)^r < \infty$. Lorden (1970), Theorem 3, obtains the bound $(r+1) \cdot E(X_1^+)^2/r\mu$ valid for all $t > 0$.

In Janson (1983), Theorem 3.1, it is shown that for positive, stationary, m-dependent summands, $\{X_k, k \ge 1\}$, one has $E(X_1)^r < \infty \Longrightarrow E(R(t))^{r-1} = \mathcal{O}(1)$ as $t \to \infty$ $(r > 1)$. The theorem there also covers the arithmetic case. Furthermore, a converse is given, namely it is shown that

$$E(R(t))^{r-1} = \mathcal{O}(1) \quad as \quad t \to \infty \quad \Longrightarrow \quad EX_1^r < \infty \quad (r > 1). \tag{10.17}$$

Since the ladder variable method is not applicable for m-dependent sequences in general (see Janson (1983), Example 3.1) the result there cannot, in a simple way, be extended to arbitrary summands with positive mean. In our case, however, we can use (10.17), (10.7) and Theorem 3.1(ii) to obtain the following converse to Theorem 10.9.

Theorem 10.10. *Let $r > 1$. If*

$$E(R(t))^{r-1} = \mathcal{O}(1) \quad as \quad t \to \infty \qquad (10.18)$$

then

$$E(X_1^+)^r < \infty. \qquad (10.19)$$

We conclude this section by presenting a method due to Janson (1983)—who used it to prove Theorem 10.8 for positive, stationary, m-dependent random variables—to give a different proof of Theorems 10.8 and 9.5. (Above, Theorem 10.8 was deduced from Theorem 9.5, the proof of which was not given.)

By Hölder's inequality (see e.g. Gut (2007), Theorem 3.2.4) with $p = 1/(r-1)$ and $q = 1/(2-r)$ we obtain

$$E(S_{\nu(t)} - t) = E((R(t))^{(r-1)^2} \cdot (R(t))^{r(2-r)})$$

$$\leq (E(R(t))^{r-1})^{r-1} \cdot (E(R(t))^r)^{2-r}$$

$$= (\mathcal{O}(1))^{r-1} \cdot (o(t))^{2-r} = o(t^{2-r}) \quad as \quad t \to \infty.$$

since, by Theorem 10.9 (and Problem 15.5) and Theorem 10.2(iii), we have

$$E(R(t))^{r-1} = \mathcal{O}(1) \quad and \quad E(R(t))^r = o(t) \quad as \quad t \to \infty, \qquad (10.20)$$

respectively.

This proves Theorem 10.8, and Theorem 9.5 follows immediately by (10.9).

We finally mention (without proof) the following result (due to Woodroofe (1976) under additional assumptions) see Siegmund (1985), Theorem 8.34.

Theorem 10.11. *Suppose that $\operatorname{Var} X_1 = \sigma^2 < \infty$. If the random walk is nonarithmetic, then*

$$\lim_{t \to \infty} P\left(\nu(t) - \frac{t}{\mu} \leq x\sqrt{\sigma^2\mu^{-3}t}, \ R(t) \leq y\right) = \Phi(x) \cdot \frac{1}{\mu_H} \int_0^y P(S_{T_1} > z)dz$$

for $-\infty < x < \infty$ and $y > 0$.

Thus, the first passage time (after normalization) and the overshoot are asymptotically independent. The marginal distributions are, of course, those given in Theorems 5.1 and 10.3(i).

For a local limit theorem we refer to Lalley (1984a), where the corresponding asymptotics for $P(\nu(t) = n, R(t) \leq y)$ are given.

3.11 The Law of the Iterated Logarithm

By using Theorem 1.9.1 and the method to prove the central limit theorem (Theorem 5.1) the following theorem has been obtained in Gut (1985), Section 4.

Theorem 11.1. *Suppose that* $\operatorname{Var} X_1 = \sigma^2 < \infty$. *Then*

$$C\left(\left\{\frac{\nu(t) - \frac{t}{\mu}}{\sqrt{2\frac{\sigma^2}{\mu^3} t \log \log t}}, t \geq 3\right\}\right) = [-1, 1] \text{ a.s.} \tag{11.1}$$

In particular,

$$\limsup_{t \to \infty} \left(\liminf_{t \to \infty}\right) \frac{\nu(t) - \frac{t}{\mu}}{\sqrt{2t \log \log t}} = \begin{matrix} + \\ (-) \end{matrix} \frac{\sigma}{\mu^{3/2}} \text{ a.s.} \tag{11.2}$$

Sketch of Proof. Theorems 4.1 and 1.9.1 are used to establish that

$$C\left(\left\{\frac{S_{\nu(t)} - \mu\nu(t)}{\sqrt{2\sigma^2 \frac{t}{\mu} \log \log t}}, t \geq 3\right\}\right) = [-1, 1] \text{ a.s.} \tag{11.3}$$

The finiteness of the variance is used to conclude, from Theorem 4.3, that

$$\frac{X_{\nu(t)}}{\sqrt{t \log \log t}} \xrightarrow{\text{a.s.}} 0 \quad \text{as} \quad t \to \infty. \tag{11.4}$$

By combining these facts with (3.2) and the following lemma we obtain (11.1), from which (11.2) is immediate. □

Lemma 11.1. *Let* $E \subset R$ *and suppose that* $\{Y(t), t \geq 1\}$ *is a family of random variables such that*

$$C(\{Y(t), t \geq 1\}) = E \text{ a.s.} \tag{11.5}$$

Further, let $\{\xi(t), t \geq 1\}$ *and* $\{\eta(t), t \geq 1\}$ *be families of random variables such that*

$$\xi(t) \xrightarrow{\text{a.s.}} 1 \quad \text{and} \quad \eta(t) \xrightarrow{\text{a.s.}} 0 \quad \text{as} \quad t \to \infty. \tag{11.6}$$

Then

$$C(\{\xi(t)Y(t) + \eta(t), t \geq 1\}) = E \text{ a.s.} \tag{11.7}$$

Just as in Theorem 1.9.1 the present theorem has been obtained before as special cases of functional versions. A proof using (3.13) is given in Chow and Hsiung (1976), Theorem 3.3. They first prove a law of the iterated logarithm for $\{M_n, n \geq 1\}$ (this will be done in Chapter 4 below with a different proof) and then "invert" this result.

Remark 11.1. Since we did not require the stopping to be caused by stopping times, one can show, by modifying the proof of Theorem 11.1, that the theorem also holds for $\{N(t), t \geq 0\}$.

3.12 Complete Convergence and Convergence Rates

In Section 1.10 we mentioned two results on complete convergence and convergence rates for stopped random walks, which involved a condition requiring that a similar result holds for the stopping times. The theorems below show that the first passage times statisfy this condition (cf. (1.10.2) and (1.10.6)). We refer to Gut (1983b) for proofs. For the case of positive summands we also refer to Bingham and Goldie (1982).

Theorem 12.1. *Let $r \geq 1$, $\frac{1}{2} < \alpha \leq 1$ and $\alpha r \geq 1$. If $E|X_1|^r < \infty$, then*

$$\sum_{n=1}^{\infty} n^{\alpha r - 2} \cdot P\left(\left|\nu(n) - \frac{n}{\mu}\right| > n^{\alpha}\delta\right) < \infty \quad \textit{for all } \delta > 0. \tag{12.1}$$

Remark 12.1. For $r = 2$ and $\alpha = 1$ Theorem 12.1 states that $\{(\nu(n)/n), n \geq 1\}$ converges completely to $1/\mu$.

Remark 12.2. It follows from Theorem 12.1 that Theorems 1.10.1 and 1.10.2 (with $\theta = (1/\mu)$) hold for $\{S_{\nu(n)} - \mu\nu(n), n \geq 1\}$ when $\alpha \leq 1$.

It is also possible to modify the proof of Theorem 12.1 to obtain the following result.

Theorem 12.2. *Suppose that* $\mathrm{Var}\, X_1 = \sigma^2 < \infty$. *Then*

$$\sum_{n=3}^{\infty} \frac{1}{n} P\left(\left|\nu(n) - \frac{n}{\mu}\right| > \delta\sqrt{n \log\log n}\right) < \infty \quad \textit{for all } \delta > \sqrt{2\frac{\sigma^2}{\mu^3}}. \tag{12.2}$$

Remark 12.3. The results can also be stated for the whole families of random variables considered (cf. Remark 1.10.2).

3.13 Applications to the Simple Random Walk

A very basic example in random walk theory is the simple random walk, which is characterized by

$$P(X_1 = 1) = p, \quad P(X_1 = -1) = q = 1 - p, \tag{13.1}$$

where $0 \leq p \leq 1$. Apart from being intuitively *the* random walk it can be used to give examples and counterexamples (as we have done in Sections 1.5 and 1.6) and to describe certain properties which are typical for large classes of random walks (as we have done in Section 2.8).

In this section we shall apply some of the results of this chapter to simple random walks. In some instances it is also possible to reprove the theorems by exploiting the fact that we have a specifically given random walk.

Since the random walk is supposed to drift to $+\infty$ we suppose throughout that

$$\mu = EX_1 = p - q > 0, \tag{13.2}$$

that is, that $p > q$ (cf. Section 2.8). To avoid trivialities, we also assume that $p < 1$.

The standard result in textbooks is that the generating function of the first ladder epoch is given by

$$g_{T_1}(s) = E\, s^{T_1} = \frac{1 - \sqrt{1 - 4pqs^2}}{2qs}, \tag{13.3}$$

which, by differentiation, yields

$$\mu_T = ET_1 = \frac{1}{p - q}. \tag{13.4}$$

Formula (13.3) is proved by combinatorial methods. We also note that further differentiation shows that T_1 has moments of all orders.

The crucial observation which we shall exploit in order to give different proofs of some of the results derived earlier in this chapter is that

$$\nu(0) = T_1 = \min\{n\colon S_n > 0\} = \min\{n\colon S_n = +1\} \tag{13.5}$$

and, more generally, that $\nu(n) = T_{n+1}$ and, hence, that

$$\{\nu(n), n \geq 0\} \quad \text{is a renewal process}, \tag{13.6}$$

whose i.i.d. increments are distributed like $\nu(0)$. In general we only have sub-additivity (recall formulas (2.5.7) and (2.5.8)).

Remark 13.1. The conclusions (13.5) and (13.6) depend on the fact that we are concerned with an arithmetic random walk, such that $X_1^+ = 1$ a.s. The conclusions are thus, in fact, valid for more general random walks than the simple one.

We are now in the position to give another proof of Theorem 3.1(i) by combining our knowledge about the generating function and (13.6). Namely, since we already know that $ET_1^r < \infty$ for all $r > 0$ it follows from Minkowski's inequality that

$$\|\nu(n)\|_r \leq n\|T_1\|_r < \infty. \tag{13.7}$$

Remark 13.2. For general random walks a similar conclusion was more complicated in that $\nu(t)$ then was a sum of a *random* number of ladder epoch increments; recall Remark 3.1.

Remark 13.3. Since $|X_1| = 1$ a.s. and

$$S_{T_{n+1}} = S_{\nu(n)} = n + 1 \text{ a.s.,} \tag{13.8}$$

Theorem 3.1(i) is an immediate consequence of Theorem 1.5.5, that is, we do not need to verify Theorem 3.1(ii) before proving Theorem 3.1(i) for $r > 1$. This is, on the other hand, not surprising, since Theorem 3.1(ii) is trivial in view of (13.8).

Next we give a different proof of (13.4). By Theorem 4.1(i), the strong law, we have

$$\frac{\nu(n)}{n} \xrightarrow{\text{a.s.}} \frac{1}{EX_1} = \frac{1}{p - q} \quad \text{as} \quad n \to \infty. \tag{13.9}$$

Now, by (13.6) and the ordinary strong law of large numbers we have

$$\frac{\nu(n)}{n} \xrightarrow{\text{a.s.}} E\nu(0) \quad \text{as} \quad n \to \infty, \tag{13.10}$$

which, together with (13.9) proves (13.4).

Note also that if, instead, we begin with the "known" fact that (13.4) holds and combine that with (13.6) and the ordinary strong law (13.10) we obtain an alternative proof of Theorem 4.1.

Remark 13.4. For a quick proof of (13.4) we refer to Problem 15.7.

Next we check Theorem 9.3 and compare with the bound obtained in Theorem 9.4(ii) (compare Section 2.5 for the Negative Binomial process).

Since $Y_1 = S_{T_1} = 1$ a.s. and $d = 1$, formula (9.6) becomes

$$E\nu(n) = \frac{n}{p - q} + \frac{1}{p - q} \cdot \frac{1}{2} + \frac{1}{2(p - q)} + o(1) = \frac{n + 1}{p - q} + o(1) \quad \text{as} \quad n \to \infty.$$

On the other hand a direct proof, using (13.6), gives

$$E\nu(n) = ET_{n+1} = (n + 1)ET_1 = \frac{n + 1}{p - q}, \tag{13.11}$$

that is, $o(1) \equiv 0$ in this case.

The upper bound in Theorem 9.4(i) is

$$\frac{n}{p - q} + \frac{1}{2(p - q)^2} + \frac{1}{2(p - q)} + o(1) = \frac{n + 1}{p - q} + \frac{1 - p}{(p - q)^2} + o(1), \tag{13.12}$$

as $n \to \infty$. The overestimate $(1 - p)/(p - q)^2$ increases from 0 to $+\infty$ as p decreases from 1 to $\frac{1}{2}$. However, since (13.11) is computed in terms of the original random walk there is no need for (13.12) here.

The overshoot $R(n) = S_{\nu(n)} - n = S_{T_{n+1}} - n = 1$ a.s. for all n, which proves Theorem 10.3(ii). Theorem 10.5(ii) and the bound in Theorem 10.6(ii) can easily be checked or reproved as above. We leave the details.

Finally, Theorem 10.7 yields the formula

$$E\left(\inf_{n \geq 1} S_n\right)^- = \sum_{k=1}^{\infty} \frac{1}{k} E(S_k^-) = \frac{1}{2(p - q)} - \frac{1}{2} = \frac{q}{p - q} = \frac{1 - p}{2p - 1}. \tag{13.13}$$

Observe that the function $(1-p)/(2p-1)$ ($\frac{1}{2} < p < 1$) is strictly decreasing in p and that the limits at $p = \frac{1}{2}$ and $p = 1$ are $+\infty$ and 0, respectively, in full accordance with what one might expect.

3.14 Extensions to the Non-I.I.D. Case

Some of the theorems in this chapter have been extended to sums of independent, but not necessarily identically distributed random variables and to sums of random variables fulfilling certain dependence assumptions. In this section we mention briefly some such results.

Suppose $\{S_n, n \geq 0\}$ is a sequence of sums of independent (not necessarily identically distributed) random variables. Under some (uniformity) conditions Theorem 6.1 has been extended by Chow and Robbins (1963). Also, the proof of Theorem 9.1 given in Siegmund (1969) is actually given for such more general random walks. Smith (1964) and Williamson (1965) extend Theorem 6.4 (the latter for nonnegative summands) and Cox and Smith (1953) and Maejima (1975) prove extensions of local limit theorems (recall Section 3.5).

Several theorems in this chapter are generalized to processes with independent, stationary increments (with positive mean) in Gut (1975b).

Ney and Wainger (1972) and Maejima and Mori (1984) prove renewal theorems for random walks indexed by the positive integer valued d-dimensional lattice points ($d = 2$ and $d \geq 2$ respectively).

Chow and Robbins (1963), Chow (1966) and Gut (1974b) consider moments of first passage times and the strong law for certain sums of dependent random variables. The results there extend Theorems 3.1, 4.1 and 8.1.

Berbee (1979) presents a renewal theory for random walks with stationary increments. Janson (1983) assumes that the increments of the random walk are stationary and m-dependent.

We close by mentioning "perturbed random walks", which may be interpreted as random walks + noise. Chapter 6 is devoted to renewal theory for such processes, including a final section on various extensions.

3.15 Problems

1. Prove Theorem 5.2.
2. Verify Remarks 4.4 and 5.2.
3. Prove Theorem 6.3(ii).
4. a) Let U be a positive integer valued random variable. It is well known that $EU = \sum_{k=1}^{\infty} P(U \geq k)$. Show that

$$EU(U + 1) = 2 \sum_{k=1}^{\infty} kP(U \geq k). \tag{15.1}$$

b) Use (10.5) and (15.1) to give a direct proof of Theorem 10.5(ii) for $d = 1$.

c) Generalize to the arithmetic case with span $d \geq 1$.

5. Generalize Theorem 10.9 to the arithmetic case.

6. Prove Theorem 11.1 (and Lemma 11.1) and verify Remark 11.1.

7. Consider the simple random walk of Section 3.13. Prove, by conditioning on the first step, that ET_1 satisfies the equation

$$x = 1 + 2qx, \qquad (15.2)$$

and, hence, conclude that $ET_1 = 1/(p - q)$, that is, that (13.4) holds.

4

Generalizations and Extensions

4.1 Introduction

In this chapter we present some generalizations and extensions of the results in Chapter 3. The chapter is divided into the following four parts:

• Section 4.2, where we discuss stopped two-dimensional random walks; more precisely, we consider a random walk $\{(U_n, V_n), n \geq 0\}$, which is stopped when one component reaches a level, t, say, and then evaluate the other component at that time point. Thus, assume that the increments of $\{U_n, n \geq 0\}$ have positive, finite mean and define

$$\tau(t) = \min\{n\colon U_n > t\} \qquad (t \geq 0). \tag{1.1}$$

Our interest will be focused on the process $\{V_{\tau(t)}, t \geq 0\}$. More specifically, we shall prove results, such as strong laws of large numbers, the central limit theorem, moment convergence and the law of the iterated logarithm for this process. Moreover, if the first component of the random walk is a renewal process, it turns out that the asymptotic behaviour of $V_{\tau(t)}$ is the same as that of $V_{M(t)}$, where

$$M(t) = \max\{n\colon U_n \leq t\} \qquad (t \geq 0). \tag{1.2}$$

• Section 4.3, in which we present several interesting applications, which originate from very different contexts, but, all fit into the above model.
• Section 4.4, which is concerned with the maximum of random walks with positive drift. In Chapter 2 we found that one way to describe how such random walks tend to infinity is to study the partial maxima $\{M_n, n \geq 0\}$. It was also noted that the partial maxima can be represented as sums of a random number of ladder height increments (see Lemma 2.11.1). Now, since such sums are precisely of the kind $V_{M(t)}$ as described above we can determine the asymptotic behavior of M_n as $n \to \infty$ by invoking results for $V_{M(t)}$ from Section 4.2. We also derive some further weak limit theorems with the aid of Lemma 2.11.2.

• Section 4.5, in which we study first passage times à la Chapter 3, but for more general, "time dependent," barriers. As a typical case we may consider

$$\nu(t) = \min\{n\colon S_n > t\sqrt{n}\} \qquad (t \geq 0),$$

where, as before, $\{S_n, n \geq 0\}$ is a random walk with i.i.d. increments having positive mean. Again we shall prove strong laws, asymptotic normality, etc., that is, we shall see how and to what extent the results from Chapter 3 carry over to this more general case.

As always, we are given a given probability space $(\Omega, \mathscr{F}, P)$ on which all of our random variables are defined.

4.2 A Stopped Two-Dimensional Random Walk

We consider a two-dimensional random walk $\{(U_n, V_n), n \geq 0\}$ with i.i.d. increments $\{(W_k, Z_k), k \geq 1\}$, where W_k and Z_k thus may depend on each other (and indeed they will in our applications). We assume throughout that

$$0 < EW_1 = \mu_w < \infty \quad \text{and} \quad E|Z_1| < \infty, \qquad (2.1)$$

but assume nothing about the sign of $EZ_1 = \mu_z$. Further, set, for $n \geq 1$, $\mathscr{F}_n = \sigma\{(W_k, Z_k), k \leq n\}$ and set $\mathscr{F}_0 = \{\varnothing, \Omega\}$ and $U_0 = V_0 = 0$.

Define the first passage times

$$\tau(t) = \min\{n\colon U_n > t\} \qquad (t \geq 0). \qquad (2.2)$$

Since $\mu_w > 0$ we know from Chapter 3 that $\tau(t)$ is a proper random variable for all $t \geq 0$ and that *all results derived in Chapter 3 apply to $\{\tau(t), t \geq 0\}$ and $\{U_{\tau(t)}, t \geq 0\}$*.

The object of interest in the present section is the stopped random walk

$$\{V_{\tau(t)}, t \geq 0\}. \qquad (2.3)$$

From the way we have defined our σ-algebras it follows that the process $\{V_{\tau(t)}, t \geq 0\}$ is a stopped random walk, so whenever it is appropriate we shall apply our results from Chapter 1.

Several authors have studied this process or variations of it; see Smith (1955), Cox (1967), von Bahr (1974), Siegmund (1975), Prabhu (1980), Chapter 4, Asmussen (1982) and Gut and Janson (1983).

Theorem 2.1 (The Strong Law of Large Numbers). *We have*

(i) $$\frac{V_{\tau(t)}}{t} \xrightarrow{\text{a.s.}} \frac{\mu_z}{\mu_w} \quad as \quad t \to \infty.$$

Moreover, if $E|W_1|^r < \infty$ and $E|Z_1|^r < \infty$ for some $r \geq 1$, then

(ii) $$\left\{ \left| \frac{V_{\tau(t)}}{t} \right|^r, t \geq 1 \right\} \quad \text{is uniformly integrable;}$$

(iii) $$E\left| \frac{V_{\tau(t)}}{t} \right|^p \longrightarrow \left(\frac{|\mu_z|}{\mu_w} \right)^p \quad \text{as} \quad t \to \infty \quad \text{for all } p, 0 < p \leq r.$$

Proof. (i) By Theorem 3.4.1 we have

$$\frac{\tau(t)}{t} \xrightarrow{\text{a.s.}} \frac{1}{\mu_w} \quad \text{as} \quad t \to \infty, \tag{2.4}$$

from which (i) follows by Theorem 1.2.3(iii).

(ii) Since, by Theorem 3.7.1,

$$\left\{ \left(\frac{\tau(t)}{t} \right)^r, t \geq 1 \right\} \quad \text{is uniformly integrable,} \tag{2.5}$$

the conclusion follows from Theorem 1.6.1.

(iii) Follows from (i), (ii) and Theorem A.1.1. $\qquad\square$

Just as in Chapters 1 and 3 there exists a Marcinkiewicz–Zygmund version of the strong law.

Theorem 2.2. *Suppose that $1 \leq r < 2$ and that $E|W_1|^r < \infty$ and $E|Z_1|^r < \infty$. Then*

(i) $$\frac{V_{\tau(t)} - \frac{\mu_z}{\mu_w} t}{t^{1/r}} \xrightarrow{\text{a.s.}} 0 \quad \text{as} \quad t \to \infty;$$

(ii) $$\left\{ \left| \frac{V_{\tau(t)} - \frac{\mu_z}{\mu_w} t}{t^{1/r}} \right|^r, t \geq 1 \right\} \quad \text{is uniformly integrable;}$$

(iii) $$E\left| \frac{V_{\tau(t)} - \frac{\mu_z}{\mu_w} t}{t^{1/r}} \right|^r \to 0 \quad \text{as} \quad t \to \infty.$$

Proof. (i) The proof is closely related to the proof of Theorem 3.4.4. We use a device due to Rényi (1957) in order to define the proper summands, to which we apply the results from Chapter 1. Namely, let $\{S_n, n \geq 0\}$ be a random walk (with $S_0 = 0$), whose increments $\{X_k, k \geq 1\}$ are given by

$$X_k = \mu_w Z_k - \mu_z W_k \qquad (k \geq 1). \tag{2.6}$$

The important fact is that the increments X_k, $k \geq 1$, are i.i.d. random variables with mean 0 and with the same integrability properties as the original random walk, in the sense that

$$E|W_1|^r < \infty, \qquad E|Z_1|^r < \infty \quad \Longrightarrow \quad E|X_1|^r < \infty \qquad (r > 0). \qquad (2.7)$$

With this and (2.4) in mind we can apply Theorem 1.2.3(ii) to conclude that

$$\frac{S_{\tau(t)}}{t^{1/r}} \xrightarrow{\text{a.s.}} 0 \quad \text{as} \quad t \to \infty, \qquad (2.8)$$

which is the same as

$$\frac{V_{\tau(t)} - \frac{\mu_z}{\mu_w} U_{\tau(t)}}{t^{1/r}} \xrightarrow{\text{a.s.}} 0 \quad \text{as} \quad t \to \infty. \qquad (2.9)$$

Finally, since

$$V_{\tau(t)} - \frac{\mu_z}{\mu_w} t = \left(V_{\tau(t)} - \frac{\mu_z}{\mu_w} U_{\tau(t)} \right) + \frac{\mu_z}{\mu_w} (U_{\tau(t)} - t) \qquad (2.10)$$

it only remains to prove that

$$\frac{U_{\tau(t)} - t}{t^{1/r}} \xrightarrow{\text{a.s.}} 0 \quad \text{as} \quad t \to \infty, \qquad (2.11)$$

but, since $U_{\tau(t)} - t$ is precisely an overshoot in the sense of Chapter 3, we already know from Theorem 3.10.2(i) that (2.11) holds. The proof of (i) is thus complete.

(ii) In view of (2.5) we know from Theorem 1.6.2 that, for $\{S_n, n \geq 0\}$ as above,

$$\left\{ \left| \frac{S_{\tau(t)}}{t^{1/r}} \right|^r, t \geq 1 \right\} \quad \text{is uniformly integrable,} \qquad (2.12)$$

that is we know that

$$\left\{ \left| \frac{V_{\tau(t)} - \frac{\mu_z}{\mu_w} U_{\tau(t)}}{t^{1/r}} \right|^r, t \geq 1 \right\} \quad \text{is uniformly integrable.} \qquad (2.13)$$

Furthermore,

$$\left\{ \left(\frac{U_{\tau(t)} - t}{t^{1/r}} \right)^r, t \geq 1 \right\} \quad \text{is uniformly integrable} \qquad (2.14)$$

by Theorem 3.10.2(ii), which together with (2.10) and Lemma A.1.3 proves (ii).

(iii) This is immediate from (i) and (ii). □

Remark 2.1. For the case $\mu_z = 0$ we only have to consider $V_{\tau(t)}/t^{1/r}$. The conclusions then follow directly from Theorems 1.2.3(ii) and 1.6.2.

Remark 2.2. The results above are due to Gut and Janson (1983). Theorem 2.1(i) and (iii) with $r = 1$ and $r = 2$ are very close to results in Smith (1955), Section 5.

Theorem 2.3 (The Central Limit Theorem). *Suppose that* $\operatorname{Var} W_1 = \sigma_w^2$ *and* $\operatorname{Var} Z_1 = \sigma_z^2$ *are finite, set*

$$\gamma^2 = \operatorname{Var}\left(\mu_w Z_1 - \mu_z W_1\right) = \mu_w^2 \sigma_z^2 + \mu_z^2 \sigma_w^2 - 2\mu_w \mu_z \operatorname{Cov}(W_1, Z_1) \qquad (2.15)$$

and suppose that $\gamma^2 > 0$. *Then*

(i)
$$\frac{V_{\tau(t)} - \frac{\mu_z}{\mu_w}t}{\sqrt{\gamma^2 \mu_w^{-3} t}} \xrightarrow{d} N(0,1) \quad as \quad t \to \infty.$$

If, moreover, $E|W_1|^r < \infty$ *and* $E|Z_1|^r < \infty$ *for some* $r \geq 2$, *then*

(ii)
$$\left\{ \left| \frac{V_{\tau(t)} - \frac{\mu_z}{\mu_w}t}{\sqrt{t}} \right|^r, t \geq 1 \right\} \quad \text{is uniformly integrable;}$$

(iii)
$$E \left| \frac{V_{\tau(t)} - \frac{\mu_z}{\mu_w}t}{\sqrt{t}} \right|^p \to E|Y|^p \quad as \quad t \to \infty \text{ for all } p, 0 < p \leq r,$$

where Y *is a normal random variable with mean* 0 *and variance* γ^2/μ_w^3;

(iv)
$$E \left(\frac{V_{\tau(t)} - \frac{\mu_z}{\mu_w}t}{\sqrt{t}} \right)^k \to 0 \quad as \quad t \to \infty \text{ for } k = odd \text{ integer} \leq r.$$

Proof. To prove the theorem we introduce the random walk from the proof of Theorem 2.2 and modify the proofs in Chapter 3 for $\{\nu(t), t \geq 0\}$ in the same way as those proofs were modified to obtain Theorems 2.1 and 2.2. We therefore do not give all details.

By Theorem 1.3.1 applied to $\{S_n, n \geq 0\}$ we have

$$\frac{\mu_w V_{\tau(t)} - \mu_z U_{\tau(t)}}{\sqrt{\gamma^2 \mu_w^{-1} t}} \xrightarrow{d} N(0,1) \quad as \quad t \to \infty \qquad (2.16)$$

and by Theorem 3.10.2 we have

$$\frac{U_{\tau(t)} - t}{\sqrt{t}} \xrightarrow{a.s.} 0 \quad as \quad t \to \infty. \qquad (2.17)$$

Now, (2.16), (2.17) together with (2.10) and Cramér's theorem prove (i).

As for (ii) we use Theorems 1.6.3 and 3.10.2 to conclude (cf. (2.12) and (2.14)) that

$$\left\{ \left| \frac{V_{\tau(t)} - \frac{\mu_z}{\mu_w} U_{\tau(t)}}{\sqrt{t}} \right|^r, t \geq 1 \right\} \quad \text{is uniformly integrable} \qquad (2.18)$$

and that

$$\left\{ \left| \frac{U_{\tau(t)} - t}{\sqrt{t}} \right|^r, t \geq 1 \right\} \quad \text{is uniformly integrable,} \qquad (2.19)$$

respectively. Now (ii) follows by (2.10) and Lemma A.1.3, from which (iii) and (iv) are immediate (cf. Remark A.1.1). □

Theorem 2.3 is due to Gut and Janson (1983). However, (i) has been proved for the case when $\{U_n, n \geq 0\}$ is a renewal process and $\mu_z \geq 0$ by Siegmund (1975). Smith (1955) uses characteristic function methods to prove a related result (see Section 5 there) corresponding to $\{U_n, n \geq 0\}$ being a renewal process.

Remark 2.3. As before the case $\mu_z = 0$ is easier.

Next we show that it is possible to obtain better remainders in the expressions for $EV_{\tau(t)}$ when $r = 2$ and for $\text{Var}\, V_{\tau(t)}$ when $r = 2$ and $\mu_z = 0$. It is also possible to obtain an estimate for $\text{Cov}(V_{\tau(t)}, \tau(t))$.

Theorem 2.4. *Suppose that W_1 and Z_1 have finite variance and define γ^2 as before. Then, as $t \to \infty$,*

(i)
$$EV_{\tau(t)} = \frac{\mu_z}{\mu_w} t + \mathcal{O}(1);$$

(ii)
$$\text{Var}\, V_{\tau(t)} = \frac{\gamma^2}{\mu_w^3} t + o(t);$$

(iii)
$$\text{Cov}\,(V_{\tau(t)}, \tau(t)) = \frac{\mu_z \sigma_w^2 - \mu_w \text{Cov}\,(W_1, Z_1)}{\mu_w^3} t + o(t).$$

If $\mu_z = 0$, (i) and (ii) can be improved to

(i′)
$$EV_{\tau(t)} = 0;$$

(ii′)
$$\text{Var}\, V_{\tau(t)} = \frac{\sigma_z^2}{\mu_w} t + \mathcal{O}(1) \quad as \quad t \to \infty.$$

Here $\mathcal{O}(1)$ is a constant depending on $\mu_w, \sigma_w^2, \mu_z, \sigma_z^2$ and the first two moments of the ladder heights for $\{U_n, n \geq 0\}$ and on whether the random walk is arithmetic or nonarithmetic.

Proof. The basis of the proof is Theorem 1.5.3 (and formula (1.5.25)), according to which

$$EV_{\tau(t)} = \mu_z E\tau(t) \tag{2.20}$$

$$\mathrm{Var}\,(V_{\tau(t)} - \mu_z\tau(t)) = E(V_{\tau(t)} - \mu_z\tau(t))^2 = \sigma_z^2 E\tau(t) \tag{2.21}$$

$$\mathrm{Cov}\,(U_{\tau(t)} - \mu_w\tau(t), V_{\tau(t)} - \mu_z\tau(t)) = E(U_{\tau(t)} - \mu_w\tau(t))(V_{\tau(t)} - \mu_z\tau(t))$$

$$= \mathrm{Cov}\,(W_1, Z_1)E\tau(t). \tag{2.22}$$

These results will be combined with facts from Chapter 3 concerning the asymptotic behavior of $U(\tau)$ and $\tau(t)$.

We first recall from Theorems 3.9.2 and 3.9.3 that

$$E\tau(t) = \frac{t}{\mu_w} + \mathcal{O}(1) \quad \text{as} \quad t \to \infty, \tag{2.23}$$

which in conjunction with (2.20) proves (i).

If $\mu_z = 0$, then (2.20) alone proves (i') and (2.21) together with (2.23) proves (ii').

Next we note that

$$\mathrm{Cov}\,(U_{\tau(t)} - \mu_w\tau(t), V_{\tau(t)} - \mu_z\tau(t))$$

$$= \mathrm{Cov}\,(U_{\tau(t)}, V_{\tau(t)} - \mu_z\tau(t)) - \mu_w\mathrm{Cov}(\tau(t), V_{\tau(t)} - \mu_z\tau(t)). \tag{2.24}$$

Since, by Cauchy's inequality,

$$|\mathrm{Cov}(U_{\tau(t)}, V_{\tau(t)} - \mu_z\tau(t))| \le \sqrt{\mathrm{Var}\,U_{\tau(t)} \cdot \mathrm{Var}\,(V_{\tau(t)} - \mu_z\tau(t))}, \tag{2.25}$$

it follows from (2.21), (2.23) and Theorem 3.10.2 that

$$|\mathrm{Cov}(U_{\tau(t)}, V_{\tau(t)} - \mu_z\tau(t))| \le \sqrt{\mathrm{Var}\,(U_{\tau(t)} - t) \cdot \sigma_z^2 E\tau(t)}$$

$$= \sqrt{o(t) \cdot \sigma_z^2 \cdot \left(\frac{t}{\mu_w} + \mathcal{O}(1)\right)} = o(t) \quad \text{as} \quad t \to \infty,$$

which, together with (2.24), (2.22) and (2.23) yields

$$\mathrm{Cov}\,(V_{\tau(t)} - \mu_z\tau(t), \tau(t)) = o(t) - \frac{\mathrm{Cov}\,(W_1, Z_1)E\tau(t)}{\mu_w}$$

$$= -\frac{\mathrm{Cov}(W_1, Z_1)}{\mu_w^2} \cdot t + o(t)$$

and thus (recall Theorem 3.9.1) that, as $t \to \infty$,

$$\mathrm{Cov}(V_{\tau(t)}, \tau(t)) = \mathrm{Cov}\left(V_{\tau(t)} - \mu_z \tau(t), \tau(t)\right) + \mu_z \mathrm{Var}\, \tau(t)$$

$$= -\frac{\mathrm{Cov}(W_1, Z_1)}{\mu_w^2} t + o(t) + \mu_z \cdot \left(\frac{\sigma_w^2}{\mu_w^3} t + o(t)\right)$$

$$= \frac{\mu_z \sigma_w^2 - \mu_w \mathrm{Cov}\,(W_1, Z_1)}{\mu_w^3} t + o(t),$$

which proves (iii).

As noted in connection with Theorem 1.5.3, $\mathrm{Var}\, V_{\tau(t)} = E(V_{\tau(t)} - \mu_z E\tau(t))^2$ and $E(V_{\tau(t)} - \mu_z \tau(t))^2$ are not the same if $\mu_z \neq 0$. However, it is easy to see that

$$\mathrm{Var}\, V_{\tau(t)} = \sigma_z^2 E\tau(t) - \mu_z^2 \mathrm{Var}\,\tau(t) + 2\mu_z \mathrm{Cov}\,(V_{\tau(t)}, \tau(t)). \qquad (2.26)$$

By combining this relation with (iii) and Theorem 3.9.1 we obtain

$$\mathrm{Var}\, V_{\tau(t)} = \sigma_z^2 \left(\frac{t}{\mu_w} + o(t)\right) - \mu_z^2 \left(\frac{\sigma_w^2 t}{\mu_w^3} + o(t)\right)$$

$$+ 2\mu_z \left(\frac{\mu_z \sigma_w^2 - \mu_w \mathrm{Cov}(W_1, Z_1)}{\mu_w^3} t + o(t)\right),$$

which, after routine simplifications, yields (ii).

Since (i$'$) and (ii$'$) have already been proved, the proof is complete.   □

Theorem 2.4 is due to Gut and Janson (1983) and we have followed the proof given there. In Smith (1955), Section 5, (i) and (ii) are proved with a different method in a related situation when $\{U_n, n \geq 0\}$ is a renewal process.

It is also possible to obtain a limit distribution for $V_{\tau(t)}$, when the variances are not necessarily finite. However, since two summation sequences are involved the normalization has to be chosen with care. One such result is the following.

Theorem 2.5. *Suppose that* $\{B_n, n \geq 1\}$ *are positive normalizing coefficients such that*

$$\frac{\mu_z U_n - \mu_w V_n}{B_n} \xrightarrow{d} G_\alpha \quad as \quad n \to \infty \qquad (2.27)$$

$$\frac{U_n - n\mu_w}{aB_n} \xrightarrow{d} G_\alpha \quad as \quad n \to \infty \quad for\ some \quad a > 0, \qquad (2.28)$$

where G_α *is (a random variable distributed according to) a stable law with index* α, $1 < \alpha \leq 2$. *Then*

$$\frac{V_{\tau(t)} - \frac{\mu_z}{\mu_w} t}{\frac{1}{\mu_w} B\left(\frac{t}{\mu_w}\right)} \xrightarrow{d} -G_\alpha \quad as \quad t \to \infty, \qquad (2.29)$$

where $B(y) = B_{[y]}$ *for* $y > 0$.

Remark 2.4. In particular, if $\alpha = 2$ and the variances are finite, then $B_n = \gamma\sqrt{n}$ and $a = \sigma_w/\gamma$ will do and we rediscover Theorem 2.3(i).

By inspecting the proof (of Theorem 2.3(i)) we observe that if there are different normalizing sequences in (2.27) and (2.28) such that the limits are $G_{\alpha_1}(x)$ and $G_{\alpha_2}(x)$ with $\alpha_1 < \alpha_2$ respectively, then there is still a limit theorem possible. Since these modified assumptions really mean that $\{U_n, n \geq 0\}$ and $\{V_n, n \geq 0\}$ belong to the domain of attraction of stable laws with indices α_U and α_V, where $1 < \alpha_V < \alpha_U \leq 2$ we obtain the following result.

Theorem 2.6. *Suppose that $\{B_{V,n}, n \geq 1\}$ and $\{B_{U,n}, n \geq 1\}$ are positive normalizing coefficients such that*

$$\frac{V_n - n\mu_z}{B_{V,n}} \xrightarrow{d} G_{\alpha_V} \quad as \quad n \to \infty \tag{2.30}$$

$$\frac{U_n - n\mu_w}{B_{U,n}} \xrightarrow{d} G_{\alpha_U} \quad as \quad n \to \infty, \tag{2.31}$$

where G_{α_V} and G_{α_U} are (random variables distributed according to) stable laws with indices α_V and α_U, respectively, such that $1 < \alpha_V < \alpha_U \leq 2$. Then

$$\frac{V_{\tau(t)} - \frac{\mu_z}{\mu_w}t}{B_V\left(\frac{t}{\mu_w}\right)} \xrightarrow{d} G_{\alpha_V} \quad as \quad t \to \infty, \tag{2.32}$$

where $B_V(y) = B_{V,[y]}$ for $y > 0$.

Sketch of Proof. The first step to verify (cf. (2.16)) is that

$$\frac{\mu_w V_n - \mu_z U_n}{\mu_w B_{V,n}} \xrightarrow{d} G_{\alpha_V} \quad as \quad n \to \infty. \tag{2.33}$$

To this end we note that the sequences $\{B_{V,n}, n \geq 1\}$ and $\{B_{U,n}, n \geq 1\}$ vary regularly with exponents $1/\alpha_V$ and $1/\alpha_U$, respectively, that is

$$B_{V,n} = n^{1/\alpha_V} L_V(n) \quad and \quad B_{U,n} = n^{1/\alpha_U} L_U(n) \quad as \quad n \to \infty \tag{2.34}$$

where L_U and L_V are slowly varying (see Feller (1971), Chapter IX.8 or Gut (2007), Section 9.3).

Since $\alpha_V < \alpha_U$ it follows that

$$\frac{B_{U,n}}{B_{V,n}} \to 0 \quad as \quad n \to \infty, \tag{2.35}$$

which, together with (2.31), shows that

$$\frac{U_n - n\mu_w}{B_{V,n}} = \frac{U_n - n\mu_w}{B_{U,n}} \cdot \frac{B_{U,n}}{B_{V,n}} \xrightarrow{p} 0 \quad as \quad n \to \infty. \tag{2.36}$$

By combining this with (2.30) and Cramér's theorem we obtain (2.33). The remainder follows essentially as before. $\square$

Remark 2.5. One case covered by Theorem 2.6 is when (2.30) holds and when $\operatorname{Var} W_1 = \sigma_w^2 < \infty$ and $1 < \alpha_V < 2$.

Remark 2.6. Another proof of Theorem 2.6 is obtained by using Theorem 1.3.2 to conclude that

$$\frac{V_{\tau(t)} - \tau(t) \cdot \mu_z}{B_V \frac{t}{\mu_w})} \xrightarrow{d} G_{\alpha_V} \quad \text{as} \quad t \to \infty. \tag{2.37}$$

Now,

$$V_{\tau(t)} - \frac{\mu_z}{\mu_w} t = V_{\tau(t)} - \tau(t)\mu_z + \mu_z \left(\tau(t) - \frac{t}{\mu_w} \right). \tag{2.38}$$

Since, by Theorem 3.5.2,

$$\frac{\tau(t) - t/\mu_w}{\frac{1}{\mu_w} B_U \left(\frac{t}{\mu_w} \right)} \xrightarrow{d} -G_{\alpha_U} \quad \text{as} \quad t \to \infty \tag{2.39}$$

we find (recall (2.35)) that

$$\frac{\tau(t) - t/\mu_w}{B_V \left(\frac{t}{\mu_w} \right)} \xrightarrow{p} 0 \quad \text{as} \quad t \to \infty, \tag{2.40}$$

which, together with (2.37) and (2.38), proves (2.32).

Remark 2.7. Reviewing the proofs we observe that the theorem remains true if (2.31) is replaced by the assumption that $E|W_1|^r < \infty$ for some $r > \alpha_V$.

By using Theorem 1.9.1 (and Lemma 3.11.1) with the above methods the following result, due to Gut (1985), is obtained.

Theorem 2.7 (The Law of the Iterated Logarithm). *Suppose that the variances are finite and that γ^2 as defined by (2.15) is positive. Then*

$$C \left(\left\{ \frac{V_{\tau(t)} - \frac{\mu_z}{\mu_w} t}{\sqrt{2\mu_w^{-3} \gamma^2 t \log \log t}}, t \geq 3 \right\} \right) = [-1, 1] \ a.s. \tag{2.41}$$

In particular,

$$\limsup_{t \to \infty} \left(\liminf_{t \to \infty} \right) \frac{V_{\tau(t)} - \frac{\mu_z}{\mu_w} t}{\sqrt{2t \log \log t}} = \begin{array}{c} + \\ (-) \end{array} \frac{\gamma}{\mu_w^{3/2}} \ a.s. \tag{2.42}$$

Finally, if we replace $\{Z_k, k \geq 1\}$ above by vector valued random variables

$$\{(Z_k^{(1)}, Z_k^{(2)}, \ldots, Z_k^{(d)}), k \geq 1\} \quad \text{for some } d \geq 2 \tag{2.43}$$

we can use the Cramér–Wold device (see e.g. Billingsley (1968), Theorem 7.7) and Theorem 2.3(i) applied to the i.i.d. random variables

$$\left\{ \left(W_k, \sum_{i=1}^{d} a_i Z_k^{(i)} \right), k \geq 1 \right\}, \tag{2.44}$$

where $\{a_i, 1 \leq i \leq d\}$ are arbitrary real numbers, to establish a multidimensional analogue of Theorem 2.3(i).

For the special case $d = 2$ and $(Z_k^{(1)}, Z_k^{(2)}) = (Z_k, 1)$, $k \geq 1$, we find that the limit distribution of $(V_{\tau(t)}, \tau(t))$ is jointly Gaussian as follows.

Theorem 2.8. *Under the assumptions of Theorem 2.3, $(V_{\tau(t)}, \tau(t))$ is asymptotically normal as $t \to \infty$ with asymptotic mean*

$$\left(\frac{\mu_z}{\mu_w}, \frac{1}{\mu_w} \right) t \tag{2.45}$$

and asymptotic covariance matrix given by

$$\begin{pmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{pmatrix} t, \tag{2.46}$$

where $\sigma_{11} = \gamma^2/\mu_w^3 = (\mu_z^2 \sigma_w^2 + \mu_w^2 \sigma_z^2 - 2\mu_w \mu_z \mathrm{Cov}(W_1, Z_1))/\mu_w^3$, $\sigma_{12} = \sigma_{21} = (\mu_z \sigma_w^2 - \mu_w \mathrm{Cov}(W_1, Z_1))/\mu_w^3$ and $\sigma_{22} = \sigma_w^2/\mu_w^3$.

Some Remarks

2.8. It is also possible to prove results about complete convergence and convergence rates (cf. Sections 1.10 and 3.12). We leave the formulations and the proofs to the reader.

2.9. Define the (counting) process

$$M(t) = \max\{n: U_n \leq t\} \qquad (t \geq 0). \tag{2.47}$$

In the proofs of Theorems 2.1(i), 2.2(i), 2.3(i), 2.5, 2.6, 2.7 and 2.8 we used results from Chapter 1 which did not require the random indices to be a family of stopping times. Therefore, these theorems also hold for $\{V_{M(t)}, t \geq 0\}$.

2.10. For renewal processes the counting variable and the first passage time differ by unity (formula (2.3.11)). Therefore, if we assume that $\{U_n, n \geq 0\}$ is a renewal process, then

$$M(t) + 1 = \tau(t) \tag{2.48}$$

and

$$V_{\tau(t)} = V_{M(t)} + Z_{\tau(t)}. \tag{2.49}$$

Since $\{Z_{\tau(t)}, t \geq 1\}$ is taken care of by Theorem 1.8.1 it follows that:

> If, in addition, $\{U_n, n \geq 0\}$ is a renewal process, then all results in this section, except Theorem 2.4, also hold for $\{V_{M(t)}, t \geq 0\}$. $\qquad$ (2.50)

As for Theorem 2.4, the same arguments show that parts (ii) and (iii) also hold for $\{V_{M(t)}, t \geq 0\}$. Furthermore, since $EZ_{\tau(t)} \neq 0$ in general, it follows that (i′) does not necessarily hold.

Finally, it turns out that (i) remains valid, because $EZ_{\tau(t)} = \mathcal{O}(1)$ as $t \to \infty$. In fact, by using the key renewal theorem (Theorem 2.4.3) and arguments like in the proof of Theorem 2.6.2 one can show (personal communication by Svante Janson) that

$$EZ_{\tau(t)} \to \frac{EW_1 Z_1}{EW_1} \quad \text{as} \quad t \to \infty, \qquad (2.51)$$

since, by assumption, $E|W_1 Z_1| < \infty$.

4.3 Some Applications

In spite of its simplicity the model discussed in the previous section is applicable in a variety of situations. In this section we describe some of these applications.

4.3.1 Chromatographic Methods

Chromatographic methods are of great importance in chemistry, biology, medicine, biotechnology etc., where they are employed for separation of mixtures of compounds either for analysis or for purification.

The basic for chromatographic separation is the distribution of a sample of molecules between two phases; one which is stationary and has a large surface and one which is mobile (liquid or gas) and percolates through the stationary bed. The sample is injected onto the column and is transported along the column by the moving phase. The sample spends part of the time sorbed in (attached to) the stationary phase, i.e. there is a reversible process transferring the molecules between the phases. The separation of compounds is caused by their different times spent in the phases. The molecules spread during their travel along the column and the compounds appear as zones (or bands). This zone-spreading originates from multiple paths taken by the molecules during their passage through the column packing, molecular diffusion and mass transfer (reversible transfer of molecules) between the phases.

In the following we shall only consider the zone-spreading due to mass-transfer. For further details, see Gut and Ahlberg (1981), Section 5.

The basis for the stochastic model then is that the chromatographic process is viewed as a series of sorptions and desorptions as the particle travels along the column, such that the probability of a particle changing phase during a short time interval is essentially proportional to the time interval, the proportionality constant being the transition rate constant, and that the motions of a molecule during different time intervals are independent. We are thus led to define a two-state Markov process (or a birth and death process with two states).

Let $\{X_k, k \geq 1\}$ ($\{Y_k, k \geq 1\}$) denote the successive times a molecule spends in the mobile (stationary) phase. It follows from the model assumptions that these times have exponential distributions (with mean μ^{-1} (and λ^{-1}), say). Further,

$$T_n = X_1 + Y_1 + X_2 + Y_2 + \cdots + X_n + Y_n \tag{3.1}$$

denotes the time that has elapsed when a molecule has been n times in each phase. By assuming that the longitudinal velocity in the mobile phase is constant, v, it follows that

$$v \cdot S_n = v(X_1 + \cdots + X_n) \tag{3.2}$$

denotes the longitudinal distance the molecule has travelled at time T_n.

We now define the first passage time processes

$$\tau(t) = \min\{n\colon T_n > t\} \qquad (t > 0) \tag{3.3}$$

and

$$\nu(L) = \min\{n\colon v \cdot S_n > L\} \qquad (L > 0). \tag{3.4}$$

This means that

> $\tau(t)$ becomes the number of visits into each phase made by a molecule in order to pass beyond time t,

and that

> $v \cdot S_{\tau(t)}$ equals the longitudinal distance a molecule has travelled when it has been $\tau(t)$ times in each phase.

Similarly, $\nu(L)$ equals the number of visits a molecule makes into the mobile phase in order to travel the fixed distance, L, e.g. the length of the column.

Let $\alpha(t)$ be the amount of time spent in the mobile phase during the time interval $[0, t]$. If we can show that $S_{\tau(t)}$ and $\alpha(t)$ are "close" for large values of t, then $S_{\tau(t)}$ can be viewed as the time spent in the mobile phase during the time interval $[0, t]$ and hence

> $v \cdot S_{\tau(t)}$ can be viewed as the longitudinal distance a molecule has travelled during the time interval $[0, t]$. $\tag{3.5}$

Similarly for $T_{\nu(L)}$.

Let us now interpret this problem in the context of Section 4.2. To this end we set

$$W_k = X_k + Y_k \quad \text{and} \quad Z_k = v \cdot X_k \qquad (k \geq 1). \tag{3.6}$$

Then $U_n = \sum_{k=1}^n W_k$ is the same as T_n in (3.1), $V_n = \sum_{k=1}^n Z_k$ is the same as $v \cdot S_n$ in (3.2) and $\tau(t)$ as defined in Section 4.2 corresponds to $\tau(t)$ given by (3.3).

In our particular case, when the distributions of $\{X_k, k \geq 1\}$ and $\{Y_k, k \geq 1\}$ are exponential (see above) the results from Section 4.2 apply. We obtain, for example, $(\mu_w = \mu^{-1} + \lambda^{-1}, \mu_z = v \cdot \mu^{-1}$ etc.) that

$v \cdot S_{\tau(t)}$ is asymptotically normal with mean $v\lambda t/(\lambda + \mu)$

and variance $2\lambda\mu v^2 t/(\lambda + \mu)^3$ as $t \to \infty$.
$$\tag{3.7}$$

Since

$$0 \leq S_{\tau(t)} - \alpha(t) \leq T_{\tau(t)} - t \tag{3.8}$$

if follows from Theorem 3.10.2 that

$$\frac{S_{\tau(t)} - \alpha(t)}{\sqrt{t}} \to 0 \quad \text{in probability and } L^2 \quad \text{as} \quad t \to \infty, \tag{3.9}$$

and consequently, from (3.7), (3.9) and Cramér's theorem, that $v \cdot \alpha(t)$ has the same asymptotic distribution as $v \cdot S_{\tau(t)}$, that is

$$\frac{v \cdot \alpha(t) - \frac{v\lambda t}{\lambda+\mu}}{\sqrt{\frac{2\lambda\mu v^2 t}{(\lambda+\mu)^3}}} \xrightarrow{d} N(0,1) \quad \text{as} \quad t \to \infty. \tag{3.10}$$

We have thus obtained the asymptotic distribution for the longitudinal distance travelled by a molecule at time t.

Similarly one can show that

$T_{\nu(L)}$ is asymptotically normal with mean $(\lambda + \mu)L/\lambda v$

and variance $2\mu L/\lambda^2 v$ as $L \to \infty$,
$$\tag{3.11}$$

which thus yields the asymptotic normality for the time required by a molecule to travel the longitudinal distance L, which typically may be the elution time (the time that has elapsed when the peak has travelled the length of the column).

In order to measure the efficiency of the column one is interested in obtaining the width of the concentration profile, which is defined as "4σ", where σ^2 is the variance of the (Gaussian) distribution which describes the profile. In particular, the width at elution time is of interest.

In view of the results above it follows that the width of the concentration profile when the peak is eluted is

$$4\sqrt{\text{Var}\left(v \cdot S_{\tau((\lambda v)^{-1}(\lambda+\mu)L)}\right)} = 4\sqrt{\frac{2v\mu}{(\lambda+\mu)^2}L} + o(\sqrt{L}) \quad \text{as} \quad L \to \infty. \tag{3.12}$$

Remark 3.1. Since the distributions of the times the molecules spend during their visits into each phase are exponential it is also possible to make direct computations as follows. Let 1 and 0 denote the mobile and stationary phases respectively and let $\{X(t), t \geq 0\}$ be the associated two-state birth and death process. We have $X(0) = 1$ and

$$P_1(X(t) = 1) = \frac{\lambda}{\lambda + \mu} + \frac{\mu}{\lambda + \mu}e^{-(\lambda+\mu)t} \qquad (t \geq 0). \qquad (3.13)$$

Since

$$\alpha(t) = \int_0^t I\{X(s) = 1\}ds \qquad (3.14)$$

we have, for example,

$$E\alpha(t) = E\int_0^1 I\{X(s) = 1\}ds = \int_0^t P(X(s) = 1)ds$$

$$= \frac{\lambda t}{\lambda + \mu} + \frac{\mu}{(\lambda + \mu)^2}(1 - e^{-(\lambda+\mu)t}),$$

that is,

$$E(v \cdot \alpha(t)) = \frac{\lambda v t}{\lambda + \mu} + \frac{\mu v}{(\lambda + \mu)^2}(1 - e^{-(\lambda+\mu)t}) \sim \frac{\lambda v t}{\lambda + \mu} \quad \text{as} \quad t \to \infty \quad (3.15)$$

as expected. Here, however, we have an exact expression for all $t \geq 0$.

4.3.2 Motion of Water in a River

The river is assumed to consist of two layers, such that the longitudinal velocity is constant in the upper layer and 0 in the lower layer. A particle moves back and forth between the two layers in a random way, the randomness being governed by Poisson (exponential) laws.

By interpreting the upper layer as a "mobile phase" and the lower layer as a "stationary phase" it is easy to see how quantities such as "the longitudinal distance a particle travels during the time interval $[0, t]$" and "the time it takes a particle to travel a prescribed distance" (e.g. the distance between two lakes) can be defined; see Kaijser (1971).

4.3.3 The Alternating Renewal Process

The applications above are related in that we in both cases consider a stochastic process with two states, where the transitions occur at the random time points

$$X_1, X_1 + Y_1, X_1 + Y_1 + X_2, X_1 + Y_1 + X_2 + Y_2, \ldots, \qquad (3.16)$$

where $\{X_k, k \geq 1\}$ and $\{Y_k, k \geq 1\}$ are two independent sequences of i.i.d. random variables. The sequence (3.16) is sometimes called an *alternating renewal process*. Moreover, in both cases the object of interest is $L(t)$, the time spent in one of the states during the time interval $[0, t]$.

In the setting of 4.3.1 $L(t)$ here corresponds to $\alpha(t)$ there and the statement corresponding to (3.7) and (3.10) is

$$\frac{L(t) - \frac{\mu_x}{\mu_x+\mu_y}t}{\sqrt{\frac{\mu_x^2\sigma_y^2+\mu_y^2\sigma_x^2}{(\mu_x+\mu_y)^3}t}} \xrightarrow{d} N(0,1) \quad \text{as} \quad t \to \infty, \tag{3.17}$$

where $\mu_x = EX_1, \mu_y = EY_1, \sigma_x^2 = \operatorname{Var} X_1$ and $\sigma_y^2 = \operatorname{Var} Y_1$.

This problem has earlier been studied by a number of authors, see e.g. Takács (1957), Rényi (1957), Plucińska (1962), Cox and Miller (1965), Barlow and Proschan (1965) and Gnedenko et al. (1969), who all deal with the problem directly, using the inherent structure. In Gut and Janson (1983) the problem is viewed as a special application of the results in Section 4.2 (cf. (3.6) and the following arguments).

Remark 3.2. For a corresponding generalization of (3.11) we refer to Gut and Janson (1983), Theorem 2 and Gnedenko et al. (1969).

4.3.4 Cryptomachines

A cryptomachine contains a number of (electronic) wheels. Before the coding begins one rotates the wheels into a random start position, which is a procedure one would like to automize. Thus, the first wheel is rotated with constant velocity during a random period of time (which is determined by a noise generator). When the wheel stops, the second wheel starts etc. When the last wheel stops, the first starts again etc. The procedure is terminated after a fixed time, t.

In order to determine "how randomly" the wheels have been placed one needs the expected value and the variance for the time a particular wheel has rotated during the time interval $[0, t]$.

It is by now clear how one can obtain the desired result. In fact, let $Z_k^{(i)}$ be the duration of the k-th rotation of the i-th wheel, and let $W_k = Z_k^{(1)} + \cdots + Z_k^{(d)}$, where d is the number of wheels. Theorem 2.8 then yields a jointly asymptotically normal distribution for the times the different wheels have rotated.

4.3.5 Age Replacement Policies

Consider the situation in renewal theory, in which a specific component of a machine has a given lifetime distribution (see Example 2.2.1). As soon as a component dies, it is immediately replaced by a new, identical one, the lifetime

of which is independent of the previous one and with the same distribution. An age replacement policy (see Barlow and Proschan (1965), Chapter 3; in particular Section 3) means that units are replaced at failure or after a fixed time, a, say, after its installation. The problem is to determine, for example, the number of replacements due to failure during the time period $[0, t]$. For some estimates of the *expected* number of replacements due to failure, see e.g. Barlow and Proschan (1965), Chapter 3.

By defining $\{X_k, k \geq 1\}$ to be the lifetimes of the components and by putting

$$W_k = \min\{X_k, a\} \quad \text{and} \quad Z_k = I\{X_k \leq a\}, \tag{3.18}$$

$\tau(t)$ becomes the number of replacements in $[0, t]$, $V_{\tau(t)}$ becomes the number of replacements due to failure in $[0, t]$ and the results in Section 4.2 can be used to describe the asymptotic behavior of $V_{\tau(t)}$ as $t \to \infty$.

Let F be the distribution function of the lifetimes. Then

$$\mu_w = \int_0^a x \, dF(x) + a(1 - F(a)),$$

$$\sigma_w^2 = \int_0^a x^2 \, dF(x) + a^2(1 - F(a)) - (\mu_w)^2,$$

$$\mu_z = F(a) \quad \text{and} \quad \sigma_z^2 = F(a)(1 - F(a)),$$

$$\mathrm{Cov}\,(W_1, Z_1) = (1 - F(a)) \left(\int_0^a x \, dF(x) - aF(a) \right),$$

$$\gamma^2 = (F(a))^2 \cdot \int_0^a x^2 \, dF(x) - F(a) \left(\int_0^a x \, dF(x) \right)^2$$

$$+ a^2 F(a)(1 - F(a)),$$

and it follows from Theorem 2.3(i) that $V_{\tau(t)}$ is asymptotically normal with mean $(\mu_z/\mu_w)t$ and variance $(\gamma^2/\mu_w^3)t$ as $t \to \infty$.

If, in particular, $F(x) = 1 - e^{-x/\theta}$, $x > 0$, an elementary computation shows that

$$\mu_w = \theta(1 - e^{-a/\theta}) \quad \text{and} \quad \sigma_w^2 = \theta^2(1 - e^{-2a/\theta}) - 2a\theta e^{-a/\theta},$$

$$\mu_z = 1 - e^{-a/\theta} \quad \text{and} \quad \sigma_z^2 = (1 - e^{-a/\theta})e^{-a/\theta},$$

$$\mathrm{Cov}\,(W_1, Z_1) = \theta e^{-a/\theta}(1 - e^{-a/\theta}) - ae^{-a/\theta} \quad \text{and} \quad \gamma^2 = \theta^2(1 - e^{-a/\theta})^3.$$

In particular, it follows that the parameters in the asymptotic normal distribution are both equal to θ^{-1} and thus independent of a (!) A second thought reveals the obvious fact, namely, that components with exponential lifetimes do not age and thus should only be replaced upon failure.

4.3.6 Age Replacement Policies; Cost Considerations

Barlow and Proschan (1965), Chapter 4 discuss replacement policies in terms of the *average* cost per unit time. The results in Section 4.2 can, however, be used to study asymptotics for the cost itself. Namely, suppose a cost c_1 is suffered for each replacement due to failure and a cost c_2 is suffered for each nonfailed component which is changed ($c_1 > c_2$). In the notation of the previous example we find that $W_k = X_k \wedge a$ is the time interval between the $(k-1)$-st and the k-th replacement, $Z_k = c_1 I\{X_k \leq a\} + c_2 I\{X_k > a\}$ is the cost of the k-th replacement and $V_{\tau(t)} = \sum_{k=1}^{\tau(t)} Z_k$ is the total cost for the replacements made in $[0, t]$. We leave the details to the reader.

4.3.7 Random Replacement Policies

Another replacement policy, called "random replacement," is discussed in Barlow and Proschan (1965), Chapter 3. With this policy units are replaced either upon failure or after a random time, which is generated by a renewal process. Thus, let $\{X_k, k \geq 1\}$ be the lifetimes and define $\{Y_k, k \geq 1\}$ to be a sequence of random variables corresponding to planned replacement intervals. All random variables are independent and within each sequence they have the same distribution. Actual replacement then is governed by the renewal sequence $\{W_k, k \geq 1\}$, where $W_k = \min\{X_k, Y_k\}$. By defining $Z_k = I\{X_k < Y_k\}, \tau(t)$ becomes the number of replacements in $[0, t]$ and $V_{\tau(t)}$ becomes the number of replacements in $[0, t]$ due to failure. Again, the asymptotic behavior of $V_{\tau(t)}$ as $t \to \infty$ can be determined by the results of Section 4.2.

4.3.8 Counter Models

Particles arrive to a counter. The interarrival times are assumed to be i.i.d. random variables (typically, exponential). A particle that arrives when the counter is free is registered, but the registration causes a dead time during which no arriving particles are registered. The first particle that arrives after the expiration of the dead time is registered and causes a new dead time and so on. It is also assumed that the successive dead times are i.i.d. random variables, which, furthermore, are independent of the interarrival times. For an early treatment of counter models we refer to Pyke (1958).

For *counters of type I* one assumes, in addition, that particles arriving during a dead time have no effect at all.

Suppose that, at time $t = 0$, the counter is free and that a particle arrives (and, hence, is registered). We first observe that, in this model, the interregistration times are also i.i.d. random variables. Let $\{W_k, k \geq 1\}$ denote these times and define $\{U_n, n \geq 0\}$ and $\{\tau(t), t \geq 0\}$ as in Section 4.2. Then $\tau(t)$ equals the number of registered particles in $[0, t]$.

Next, let $\{Z_k, k \geq 1\}$ denote the successive dead times and let $\{V_n, n \geq 0\}$ be defined as in Section 4.2. With this notation $V_{\tau(t)}$ becomes the total dead time during $[0, t]$, for which we can use our findings in Section 4.2.

Note also that by comparing with Application 4.3.1 we can identify the dead time state with the mobile phase and the free time state with the stationary phase; that is, we have an alternating renewal process, for which we are interested in quantities like $\alpha(t)$ defined in 4.3.1 (and $L(t)$ defined in 4.3.3).

For *counters of type II* the situation is more involved in that impulses arriving during a dead time prolong the dead time. However, suppose that the dead times are constant, a, say, and let $X_1, X_2, \ldots$ denote the i.i.d. interarrival times. After the first registration further registrations occur each time an interarrival time exceeds a. By applying the results form Section 4.2 with $W_k = X_k$ and $Z_k = I\{X_k > a\}$ asymptotics for the number of registered particles in $[0, t]$ $(= V_{\tau(t)})$ can be obtained.

4.3.9 Insurance Risk Theory

Consider an insurance company whose business runs as follows:

(i) Claims arrive at random time points in such a way that the times $\{W_k, k \geq 1\}$ between successive claims are i.i.d. positive random variables, that is, the epochs of claims form a renewal process (typically a Poisson process).

(ii) The amounts $\{Z_k, k \geq 1\}$ of the successive claims are i.i.d. random variables.

(iii) The company receives premiums from its policyholders at a constant rate β (called the gross risk premium rate).

(iv) The initial capital (risk reserve) is u.

The main objects of interest are $R(t) =$ the risk reserve at time t, and the probability of ruin (before time t).

For references, see Cramér (1955), Sparre Andersen (1957), Seal (1969), von Bahr (1974), Prabhu (1980) and Asmussen (1982, 1984). Some more recent references are Grandell (1991), Asmussen (2000, 2003), Mikosch (2006), Embrechts et al. (2008), and further references given there.

Set $U_n = \sum_{k=1}^{n} W_k$ and $V_n = \sum_{k=1}^{n}$, $n \geq 0$, and define

$$M(t) = \max\{n \colon U_n \leq t\} \qquad (t \geq 0), \tag{3.19}$$

that is,

$$M(t) = \text{the number of claims in } [0, t] \tag{3.20}$$

and

$$V_{M(t)} = \text{the total amount claimed in } [0, t]. \tag{3.21}$$

We thus observe that the amount claimed in $[0, t]$ is expressed in such a way that the results from Section 4.2 apply (see Remark 2.10).

It follows that the risk reserve at time t is

$$R(t) = u + \beta t - V_{M(t)} \tag{3.22}$$

and thus, for example, that

$$\frac{R(t) - u - (\beta - \frac{\mu_z}{\mu_w})t}{\sqrt{\gamma^2 \mu_w^{-3} t}} \xrightarrow{d} N(0,1) \quad \text{as} \quad t \to \infty, \tag{3.23}$$

where μ_z, μ_w and γ^2 are defined as in Section 4.2.

4.3.10 The Queueing System M/G/1

Consider the M/G/1 queueing system. Customers arrive according to a Poisson process with parameter λ ($0 < \lambda < \infty$). The interarrival times $\{W_k, k \geq 1\}$ are thus i.i.d. exponential with expectation λ^{-1}. Further, let $\{Z_k, k \geq 1\}$ be the i.i.d. service times. Set

$$M(t) = \max \left\{ n : \sum_{k=1}^{n} W_k \leq t \right\} \qquad (t \geq 0). \tag{3.24}$$

Then $V_{M(t)} = Z_1 + \cdots + Z_{M(t)}$ equals the amount of service time required to serve all customers arriving in $[0, t]$ (see Prabhu (1980), p. 67).

Moreover, if $\{X_k, k \geq 1\}$ are the busy periods and $\{Y_k, k \geq 0\}$ are the idle periods, where Y_k follows X_k, we can obtain results like (3.17) for the total busy (idle) time during $[0, t]$.

4.3.11 The Waiting Time in a Roulette Game

The point of departure for Gut and Holst (1984) was the following roulette problem presented in Statistica Neerlandica (1982), Problem 127 (see also Statistica Neerlandica (1983)):

On a symmetric roulette with the numbers $0, 1, 2, \ldots, m - 1$ one plays until all numbers $1, 2, \ldots, k$ have appeared once in such a way that no 0 has appeared. If 0 appears before $1, 2, \ldots, k$ all have appeared the procedure starts again from the beginning. Show that the expected duration of the game is $k \cdot m$ and find its probability generating function.

Following is a mathematically more convenient problem for the case $k = m - 1$:

What is the total number of performances required to obtain each of the numbers $0, 1, 2, \ldots, m - 1$ at least once in such a way that the last "new" number to appear is 0? Again, if 0 appears "too early" the procedure starts from the beginning.

Let us now see that the total number of performances in the second version of the problem can be described as a stopped sum of the kind discussed in Section 4.2.

After each performance of the experiment we can divide the sequence of outcomes into a (random) number of cycles, where the first cycle contains the outcomes until the first 0 appears, the second cycle contains the outcomes between the first 0 and up to and including the second 0 etc., i.e. the n-th cycle contains the outcomes between the $(n-1)$-st and n-th occurrence of 0 (in the (...] sense).

Let us define $\{(W_n, Z_n), n \geq 1\}$ so that Z_n equals the length of the n-th cycle and W_n equals 1 if the n-th cycle (which ends with a 0) contains each of the numbers $0, 1, 2, \ldots, m - 1$ at least once (and 0 exactly once) and W_n equals 0 otherwise.

Since the occurrence of a 0 is a recurrent phenomenon (cf. Feller (1968), Chatper XIII) it follows that $\{(W_n, Z_n), n \geq 1\}$ are i.i.d. two-dimensional random variables. Furthermore, it is easy to see that

$$P(Z_1 = i) = m^{-1}(1 - m^{-1})^{i-1} \quad (i \geq 1) \tag{3.25}$$

and that

$$P(W_1 = 1) = 1 - P(W_1 = 0) = m^{-1}. \tag{3.26}$$

Set $V_n = \sum_{j=1}^{n} Z_j$, $n \geq 0$, and define the stopping time

$$\tau = \min\{n \colon W_n = 1\}. \tag{3.27}$$

Then V_τ is the quantity we look for.

To compare with Section 4.2 we set $U_n = \sum_{j=1}^{n} W_j$, $n \geq 0$, and observe that

$$\tau = \min\{n \colon U_n = 1\} = \min\{n \colon U_n > 0\}, \tag{3.28}$$

since $W_\tau = U_\tau = 1$ a.s. We also have

$$P(\tau = i) = m^{-1}(1 - m^{-1})^{i-1} \quad (i \geq 1). \tag{3.29}$$

It follows, for example, by using Theorem 1.5.3 that

$$EV_\tau = E\tau \cdot EZ_1 = m \cdot m. \tag{3.30}$$

Note, however, that there is a difference between the two formulations of the problem in that the final geometric number of performances required to obtain the very last 0 is included in the second formulation but not in the first one, that is, the expected value of the sought quantity is

$$m^2 - m = m(m - 1). \tag{3.31}$$

For further results and generalizations, see Gut and Holst (1984), where also nonsymmetric roulettes are considered.

4.3.12 A Curious (?) Problem

Let $\{X_k, k \geq 1\}$ be independent random variables, uniformly distributed on [0,1]. How many X values are *greater* than 1/2 when the sum of all of them reaches t? How many of them are *smaller* than 1/2?

By symmetry the obvious guess is "half of them" in both cases and this is true in the sense that the expected number of each kind is approximately t as $t \to \infty$. However, the asymptotic variances are not the same.

Set, for $k \geq 1$, $W_k = X_k$, $Z_k = I\{X_k > 1/2\}$ and let U_n and V_n be as always. If $\tau(t) = \min\{n\colon U_n > t\}$, $t \geq 0$, Theorem 2.3(i) shows that

> The number of X values greater than 1/2 is asymptotically
> normal with mean t and variance $t/6$ as $t \to \infty$.   (3.32)

If, on the other hand, $Z_k = I\{X_k < 1/2\}$ it follows that

> The number of X values smaller than 1/2 is asymptotically
> normal with mean t and variance $7t/6$ as $t \to \infty$.   (3.33)

An intuitive explanation of the fact that the asymptotic variances are not the same is the following. If there happen to be "too many" or "too few" small values, this does not influence the total sum so much, whereas too many or too few large values do.

The following extreme variation of this problem, suggested by David Heath, explains this phenomenon more stringently.

Suppose that, instead of being uniformly distributed, X_k, $k \geq 1$, are i.i.d. random variables such that $P(X_1 = 1) = p$ and $P(X_1 = 0) = q = 1 - p$; we thus only have "large" and "small" values. In this case the sum is determined by the number of ones; the zeroes do not contribute.

Set, for $k \geq 1$, $W_k = X_k$ and $Z_k = I\{X_k > 1/2\} = I\{X_k = 1\}$. Then $\tau(t)$ has a Negative Binomial distribution with mean $([t] + 1)/p$ and variance $([t] + 1)q/p^2$ and $V_{\tau(t)} = [t] + 1$ a.s., in particular, $\mathrm{Var}(V_{\tau(t)}) = 0$.

If, on the other hand, we set $W_k = X_k$ and $Z_k = I\{X_k < 1/2\} = I\{X_k = 0\}$, $k \geq 1$, then $V_{\tau(t)} = \tau(t) - ([t] + 1)$ (with $\tau(t)$ as above); in particular, $V_{\tau(t)}$ is asymptotically normal with mean $([t] + 1)q/p$ and variance $([t] + 1)q/p^2$ as $t \to \infty$.

4.4 The Maximum of a Random Walk with Positive Drift

In Chapter 3 we studied the behavior of random walks with positive drift in terms of first passage times across horizontal barriers; we studied the random walk when it reaches the level t, how close to t it is at that moment and, to some extent, when it leaves t on its way to infinity. Here we shall investigate

another quantity which describes how random walks drift towards infinity, namely, the sequence of partial maxima.

For convenience we recall the notations and definitions from before.

Thus, $\{S_n, n \geq 0\}$ is a random walk with positive drift, that is, $S_0 = 0$, $S_n = \sum_{k=1}^{n} X_k$, $n \geq 1$, and $S_n \xrightarrow{\text{a.s.}} +\infty$ as $n \to \infty$. We assume, in addition, unless otherwise stated, that $\mu = EX_1 > 0$ (see Section 2.8). The sequence of partial maxima $\{M_n, n \geq 0\}$ is defined by

$$M_n = \max\{0, S_1, \ldots, S_n\} \tag{4.1}$$

(Section 2.10). Some general results about $\{M_n, n \geq 0\}$ have already been obtained in Section 2.12 (also for the case $EX_1 \leq 0$).

Similarly, the sequence of partial minima $\{m_n, n \geq 0\}$, is defined by

$$m_n = \min\{0, S_1, \ldots, S_n\}. \tag{4.2}$$

The reason for putting this more thorough investigation of the partial maxima here is that several of the theorems to follow are immediate consequences of the results obtained in Section 4.2. In order to see this we shall show that the representation obtained in Lemma 2.11.1 is of the appropriate kind. To make the section more self contained we begin by recalling this lemma.

Lemma 4.1. *Let* $N(n)$, $n \geq 1$, *denote the number of ladder epochs in* $[0, n]$, *that is,*

$$N(n) = \max\{k \colon T_k \leq n\}. \tag{4.3}$$

Then

$$M_n = Y_{N(n)} = Z_1 + \cdots + Z_{N(n)}. \tag{4.4}$$

Recall from Section 2.11 that the sequence $\{N(n), n \geq 1\}$ is the counting process of the renewal process $\{T_n, n \geq 0\}$ (the ladder epoch process) and that therefore the results obtained for counting processes in Chapter 2 apply to $\{N(n), n \geq 1\}$. Moreover, (recall Section 3.2) the ladder height increments, $\{Z_k, k \geq 1\}$ are i.i.d. positive random variables, which, in view of Theorem 3.3.1, have positive, finite mean. The representation (4.4) of M_n is therefore of precisely the kind discussed in Chapter 1. We note, however, that $N(n)$ is *not* a stopping time.

Now, define the sequence $\{\mathscr{F}_n, n \geq 1\}$ of increasing sub-σ-algebras as follows. Let

$$\mathscr{F}_n = \sigma\{(N_k, Z_k), 1 \leq k \leq n\} = \sigma\{(T_k, Y_k), 1 \leq k \leq n\} \tag{4.5}$$

and set $\mathscr{F}_0 = \{\emptyset, \Omega\}$ and $T_0 = Y_0 = 0$.

A further consequence of the fact that the ladder epochs $\{T_n, n \geq 0\}$ form a renewal process is that

$$\nu_T(n) = \min\{k \colon T_k > n\} = N(n) + 1 \quad (n \geq 1). \tag{4.6}$$

Since $\{(N_k, Z_k), k \geq 1\}$ are i.i.d. random variables (recall Section 2.9), a comparison with Section 4.2 shows that, if we set

$$Y_{\nu_T(n)} = Z_1 + \cdots + Z_{\nu_T(n)} \quad (n \geq 1), \tag{4.7}$$

then $\{Y_{\nu_T(n)}, n \geq 1\}$ is mathematically precisely an example (in discrete time) of the stopped sums $\{V_{\tau(t)}, t \geq 0\}$ discussed there. Moreover, since the sequence $\{T_n, n \geq 0\}$ is a renewal process, it follows, in view of Remark 2.10, that all results from Section 4.2 are applicable to $\{Y_{N(n)}, n \geq 1\}$.

We also recall from Section 2.11 that Lemma 4.1 is useful for studying the "pathwise" behavior of $\{M_n, n \geq 0\}$ and that Lemma 2.11.2, which we quote next, frequently is more convenient for proving "weak" (distributional) results, since it was a statement about *equality in distribution*.

Lemma 4.2. *We have*

$$(M_n, M_n - S_n) \overset{d}{=} (S_n - m_n, -m_n), \tag{4.8}$$

in particular,

$$M_n \overset{d}{=} S_n - m_n. \tag{4.9}$$

The idea now is that, since $EX_1 > 0$, M_n and S_n should be close in some sense, so that M_n should obey the same limit law as S_n. Now, since $m_n \overset{a.s.}{\longrightarrow} m$ as $n \to \infty$, where m is finite a.s. (Theorem 2.10.1(i)) it follows, from Lemma 4.2, that, asymptotically, M_n and S_n differ, *in distribution*, by an a.s. finite random variable and that therefore any normalization tending to infinity which yields a limit law for S_n as $n \to \infty$ also must provide a limit law for M_n as $n \to \infty$. (This was, actually, how the proof of Theorem 2.12.3 worked.)

We are now ready to state our results.

Theorem 4.1 (The Strong Law of Large Numbers).
We have

(i)
$$\frac{M_n}{n} \overset{a.s.}{\longrightarrow} \mu \quad as \quad n \to \infty.$$

Moreover, if $E|X_1|^r < \infty$ for some $r \geq 1$, then

(ii)
$$\left\{ \left(\frac{M_n}{n} \right)^r, n \geq 1 \right\} \quad is\ uniformly\ integrable;$$

(iii)
$$E \left(\frac{M_n}{n} \right)^p \to \mu^p \quad as \quad n \to \infty \quad for\ all\ p, 0 < p \leq r.$$

Theorem 4.2. *Let $1 \le r < 2$ and suppose that $E|X_1|^r < \infty$. Then*

(i)
$$\frac{M_n - n\mu}{n^{1/r}} \xrightarrow{\text{a.s.}} 0 \quad \text{as} \quad n \to \infty;$$

(ii)
$$\left\{ \left| \frac{M_n - n\mu}{n^{1/r}} \right|^r, n \ge 1 \right\} \quad \text{is uniformly integrable;}$$

(iii)
$$E \left| \frac{M_n - n\mu}{n^{1/r}} \right|^r \to 0 \quad \text{as} \quad n \to \infty.$$

Theorem 4.3 (The Central Limit Theorem). *Suppose that $\operatorname{Var} X_1 = \sigma^2 < \infty$. Then*

(i)
$$\frac{M_n - n\mu}{\sigma \sqrt{n}} \xrightarrow{d} N(0, 1) \quad \text{as} \quad n \to \infty.$$

Moreover, if $E|X_1|^r < \infty$ for some $r \ge 2$, then

(ii)
$$\left\{ \left| \frac{M_n - n\mu}{\sqrt{n}} \right|^r, n \ge 1 \right\} \quad \text{is uniformly integrable;}$$

(iii)
$$E \left| \frac{M_n - n\mu}{\sqrt{n}} \right|^p \to E|Z|^p \quad \text{as} \quad n \to \infty \quad \text{for all } p, 0 < p \le r,$$

where Z is a normal random variable with mean 0 and variance σ^2;

(iv)
$$E \left(\frac{M_n - n\mu}{\sqrt{n}} \right)^k \to 0 \quad \text{as} \quad n \to \infty \quad \text{for } k = \text{odd integer} \le r.$$

In particular, if $r = 2$, then

(v)
$$EM_n = n\mu + o(\sqrt{n}) \quad \text{and} \quad \operatorname{Var} M_n = n\sigma^2 + o(n) \quad \text{as} \quad n \to \infty.$$

Theorem 4.4. *Suppose that $\{B_n, n \ge 1\}$ are positive normalizing coefficients such that*
$$\frac{S_n - n\mu_\alpha}{B_n} \xrightarrow{d} G_\alpha \quad \text{as} \quad n \to \infty, \tag{4.10}$$
where $\mu_\alpha = 0$ when $0 < \alpha < 1$ and $\mu_\alpha = \mu$ when $1 < \alpha \le 2$ and where G_α is (a random variable distributed according to) a stable law with index α, $0 < \alpha \le 2$, $\alpha \ne 1$, such that, in addition, $G_\alpha(0) = 0$ when $0 < \alpha < 1$. Then
$$\frac{M_n - n\mu_\alpha}{B_n} \xrightarrow{d} G_\alpha \quad \text{as} \quad n \to \infty. \tag{4.11}$$

Theorem 4.5 (The Law of the Iterated Logarithm). *Suppose that* $\mathrm{Var}\, X_1 = \sigma^2 < \infty$. *Then*

$$C\left(\left\{\frac{M_n - n\mu}{\sqrt{2\sigma^2 n \log\log n}}, n \geq 3\right\}\right) = [-1, 1] \text{ a.s.} \tag{4.12}$$

In particular,

$$\limsup_{n\to\infty}\left(\liminf_{n\to\infty}\right)\frac{M_n - n\mu}{\sqrt{2\sigma^2 n \log\log n}} = \begin{matrix} + \\ (-) \end{matrix} 1 \text{ a.s.} \tag{4.13}$$

Theorem 4.6. *Let* $r \geq 1$, $\alpha r \geq 1$, $\alpha > 1/2$ *and suppose that* $E|X_1|^r < \infty$. *Then*

$$\sum_{n=1}^{\infty} n^{\alpha r - 2} P(|M_n - n\mu| > n^\alpha \varepsilon) < \infty \quad \text{for all } \varepsilon > 0. \tag{4.14}$$

Theorem 4.7. *Suppose that* $\mathrm{Var}\, X_1 = \sigma^2 < \infty$. *Then*

$$\sum_{n=3}^{\infty} \frac{1}{n} P(|M_n - n\mu| > \varepsilon\sqrt{n\log\log n}) < \infty \quad \text{for all } \varepsilon > \sigma\sqrt{2}. \tag{4.15}$$

Remark 4.1. In most cases it is not too difficult to see how the various results follow. As for Theorem 4.6 and 4.7 it is not as obvious. However, the former follows, for $1/2 < \alpha \leq 1$, by applying Lemma 4.1 and, for $\alpha > 1$, by Lemma 4.2 and some additional work (which involves showing a similar result for $\{m_n, n \geq 0\}$). Theorem 4.7 follows by using Lemma 4.2.

The easiest way to prove these theorems is, however, to use the corresponding results for $\{S_n, n \geq 0\}$, together with the fact that

$$S_n - n\mu \leq M_n - n\mu \leq \max_{0 \leq k \leq n} (S_k - k\mu) = \max_{1 \leq k \leq n} (S_n - k\mu)^+.$$

Remark 4.2. Theorems 4.1(i) and 4.2(i) are due to Heyde (1966) who treats the case $0 < r < 2$ with completely different methods. For Theorem 4.1(i) we also refer to Theorem 2.12.1 (which also covers the case $EX_1 \leq 0$). As for Theorem 4.3(i) we refer to Theorem 2.12.3 and the references given in that section. The results on moment convergence in Theorems 4.1–4.3 seem to be new. Theorem 4.4 is due to Heyde (1967b), who also used the result, together with (3.3.13), to obtain limit theorems for the first passage times studied in Chapter 3; for $1 < \alpha \leq 2$ see Theorem 3.5.2. Concerning Theorem 4.5, (4.12) has been obtained before as a special case of functional versions of the law of the iterated logarithm. Formula (4.13) has been proved directly by Chow and Hsiung (1976) (see their Corollary 2.5(ii)), who used the result, together with (3.3.13), to prove the second part of Theorem 3.11.1. Theorems 4.6 and 4.7 have been obtained earlier for the case $EX_1 = 0$, in which case there exist several results of that kind (see Section 1.10 and references mentioned there).

In view of Remark 2.10 it follows from Theorem 2.4(i) that the remainder for $E(M_n - n\mu)$ given in Theorem 4.3(v) can be improved to $\mathcal{O}(1)$ (when $\text{Var} X_1 = \sigma^2 < \infty$). We shall use Lemma 4.2 to prove the following result.

Theorem 4.8. *We have*

(i)
$$EM_n = n\mu + \sum_{k=1}^n \frac{1}{k} E(S_k^-).$$

Moreover, if $\text{Var} X_1 = \sigma^2 < \infty$, *then*

(ii) $EM_n = n\mu + \sum_{k=1}^\infty \frac{1}{k} E(S_k^-) + o(1) = n\mu + \dfrac{EX_1^2}{2\mu} - \dfrac{EY_1^2}{2\mu_H} + o(1) \quad as \quad n \to \infty.$

Proof. Since $M_n - S_n$ and $-m_n$ have the same distribution they have, in particular, the same expectation. Thus

$$EM_n - n\mu = E(S_n - n\mu) - Em_n = -Em_n = \sum_{k=1}^n \frac{1}{k} E(S_k^-) \qquad (4.16)$$

(cf. formula (3.10.11)). This proves (i), from which (ii) follows from Theorem 3.10.7. $\qquad \square$

We conclude this section with some comments on $m = \min_{n \geq 0} S_n$. Since $EX_1 > 0$ we know from Theorem 2.10.1 that m is a.s. finite. We also recall from Section 2.11 that m has the same distribution as a certain sum of a geometric number of i.i.d. random variables. Moreover, it is possible to show that, under certain additional assumptions, m possesses higher moments. The following result is due to Kiefer and Wolfowitz (1956), see also Janson (1986), Theorem 1.

Theorem 4.9. *Let* $r > 0$. *Then*

$$E(X_1^-)^{r+1} < \infty \qquad \Longleftrightarrow \qquad Em^r < \infty. \qquad (4.17)$$

4.5 First Passage Times Across General Boundaries

In this section we generalize the results obtained in Chapter 3 in that we consider more general boundary crossings.

Let, as before, $\{S_n, n \geq 0\}$ be a random walk with i.i.d. increments $\{X_k, k \geq 1\}$ and suppose that $0 < \mu = EX_1 < \infty$.

The object of interest in this section is the family of stopping times $\{\nu(t), t \geq 0\}$, defined by

$$\nu(t) = \min\{n: S_n > t \cdot a(n)\}, \qquad (5.1)$$

where $a(y)$, $y \geq 0$, is a positive, continuous function and $a(y) = o(y)$ as $y \to \infty$. The typical case is $a(y) = y^\beta$ for some $\beta \in (0,1)$ and the case $a(y) \equiv 1$ corresponds to the case treated in Chapter 3.

The family $\{\nu(t), t \geq 0\}$ thus defined is still a family of first passage times, but more general than in Chapter 3, in that the levels to be crossed may change with time here. These, more general, first passage times are of importance in certain statistical applications. They also provide a starting point for what is sometimes called *nonlinear renewal theory*, a topic we shall return to in Chapter 6.

We shall now derive results for $\nu(t)$ and $S_{\nu(t)}$ corresponding to those in Chapter 3. The basis for the material is taken from Gut (1974a), Section 3.

Finiteness of Moments

Theorem 5.1. *Let $r \geq 1$.*

(i) $$E(X_1^-)^r < \infty \quad \Longleftrightarrow \quad E(\nu(t))^r < \infty;$$

(ii) $$E|X_1|^r < \infty \quad \Longrightarrow \quad E(S_{\nu(t)})^r < \infty.$$

Proof. (i) If $t = 0$ there is nothing to prove (see Theorem 3.3.1), so suppose that $t > 0$.

First, suppose that $E(X_1^-)^r < \infty$. By assumption there exist δ $(0 < \delta < \mu)$ and a_0, such that

$$t \cdot a(y) \leq a_0 + \delta y \quad (y \geq 0). \tag{5.2}$$

Define

$$\nu_*(t) = \min\{n: S_n > a_0 + \delta n\} \quad (t \geq 0). \tag{5.3}$$

Then

$$\nu(t) \leq \nu_*(t). \tag{5.4}$$

But, since $\nu_*(t) = \min\{n: \sum_{k=1}^n (X_k - \delta) > a_0\}$, where $E(X_1 - \delta) = \mu - \delta > 0$, it follows from Theorem 3.3.1 that $E(\nu_*(t))^r < \infty$ and hence, by (5.4), that $E(\nu(t))^r < \infty$ (and, in particular, that $\nu(t)$ is a.s. finite).

Now, suppose that $E(\nu(t))^r < \infty$ and let $T_1 = \min\{n: S_n > 0\}$ (recall Section 3.2). Since $T_1 \leq \nu(t)$ we have

$$\infty > E(\nu(t))^r \geq ET_1^r \tag{5.5}$$

and it follows from Theorem 3.3.1 that $E(X_1^-)^r < \infty$.

(ii) For $r = 1$ we obtain from (i) and Theorem 1.5.3 that $ES_{\nu(t)} = \mu E\nu(t) < \infty$, so assume that $r > 1$.

Corresponding to (3.3.2) we now have

$$t \cdot a(\nu(t)) < S_{\nu(t)} \leq t \cdot a(\nu(t) - 1) + X_{\nu(t)}^+ \tag{5.6}$$

(note that $X_{\nu(t)}$ is not necessarily positive here). By (5.2) we now obtain

$$S_{\nu(t)} \leq a_0 + \delta(\nu(t) - 1) + X_{\nu(t)} \leq a_0 + \delta\nu(t) + X_{\nu(t)}^+. \tag{5.7}$$

Finally, (5.7), together with Minkowski's inequality (cf. e.g. Gut (2007), Theorem 3.2.6), Lemma 1.8.1 and (i), show that

$$\|S_{\nu(t)}\|_r \leq a_0 + \delta\|\nu(t)\|_r + (E\nu(t))^{1/r}\|X_1^+\|_r < \infty. \qquad \square$$

Remark 5.1. Since the class of boundaries considered is rather large, there is no general converse in (ii). If $a(y) \equiv 1$, see, however, Theorem 3.3.1. Also, if $a(y) = y^\beta$ $(0 \leq \beta < 1)$, then (5.6) states that

$$t \cdot (\nu(t))^\beta < S_{\nu(t)} \leq t \cdot (\nu(t))^\beta + X_{\nu(t)} \tag{5.8}$$

and it follows that

$$E(S_{\nu(t)})^r < \infty \iff E(X_1^+)^r < \infty \quad \text{and} \quad E(X_1^-)^{\beta r} < \infty. \tag{5.9}$$

Here we have also used the facts that $X_{\nu(t)}$ is positive and that $S_{\nu(t)} \geq X_1^+$, which follow from the fact that $a(y)$ is nondecreasing.

It is also possible to prove a theorem for the moment generating function with the aid of (5.2), see Gut (1974a), Theorem 3.2.

Asymptotics

In the following we shall study asymptotics as $t \to \infty$. For this we need some further assumptions on $a(y)$, $y \geq 0$. Thus, in addition to the above, assume that

> for some $A \geq 0, a(y)$, $y \geq A$, is nondecreasing, concave
> and (for simplicity) differentiable and regularly varying   (RV)
> with exponent β $(0 \leq \beta < 1)$.

For some general facts about regularly varying functions we refer to Feller (1971), de Haan (1970), Seneta (1976), Bingham et al. (1987), Gut (2007), Section A.7 and to Appendix B.

One consequence of the assumption (RV) is that

$$a(y) = y^\beta \cdot L(y) \quad (0 \leq \beta < 1), \tag{5.10}$$

where $L(y)$ is slowly varying. Typical examples are

$$a(y) = y^\beta, \ y^\beta \cdot \log y, \ \log y, \ \arctan y. \tag{5.11}$$

Following Siegmund (1967) we define $\lambda(t) > A$ as a solution of the equation $t \cdot a(y) = \mu y$, that is,

$$t \cdot a(\lambda(t)) = \mu\lambda(t). \tag{5.12}$$

If t is large then $\lambda(t)$ exists and is unique and $\lambda(t) \to \infty$ when $t \to \infty$.

We also observe that

$$\text{if } a(y) = y^\beta, \quad \text{then } \lambda(t) = (t/\mu)^{1/(1-\beta)}, \tag{5.13}$$

in particular, if $a(y) \equiv 1$, then $\lambda(t) = t/\mu$. In fact, $\lambda(t)$ here will play the role of t/μ in Chapter 3.

To avoid some uninteresting technical trivialities, we suppose in the following that A, introduced in (RV), is equal to 0. We can then replace (5.6) by

$$t \cdot a(\nu(t)) < S_{\nu(t)} \le t \cdot a(\nu(t)) + X_{\nu(t)}. \tag{5.14}$$

The Strong Law

Theorem 5.2. $\dfrac{\nu(t)}{\lambda(t)} \xrightarrow{\text{a.s.}} 1 \quad as \ t \to \infty.$

Proof. By Theorem 1.2.3 (cf. also (3.4.1)) and (5.14) it follows that

$$\frac{t \cdot a(\nu(t))}{\nu(t)} \xrightarrow{\text{a.s.}} \mu \quad \text{as} \quad t \to \infty, \tag{5.15}$$

which, in view of (5.12), is the same as

$$\frac{a(\nu(t))}{a(\lambda(t))} \cdot \frac{\lambda(t)}{\nu(t)} \xrightarrow{\text{a.s.}} 1 \quad \text{as} \quad t \to \infty, \tag{5.16}$$

and hence, by (5.10), the same as

$$\left(\frac{\nu(t)}{\lambda(t)}\right)^{\beta-1} \cdot \frac{L(\nu(t))}{L(\lambda(t))} \xrightarrow{\text{a.s.}} 1 \quad \text{as} \quad t \to \infty. \tag{5.17}$$

For $a(y) \equiv 1$ (5.16) reduces to Theorem 3.4.1. If $a(y) \not\equiv 1$ the conclusion follows from (5.17) and Lemma B.2.2 (since $\nu(t) \xrightarrow{\text{a.s.}} +\infty$ as $t \to \infty$). $\qquad \square$

Remark 5.2. For $0 < \beta < 1$ the result is due to Siegmund (1967), Lemma 4.

Following are a.s. convergence results for the stopped sum and the stopping summand.

Theorem 5.3.

$$\frac{S_{\nu(t)}}{ta(\lambda(t))} = \frac{S_{\nu(t)}}{\mu\lambda(t)} \xrightarrow{\text{a.s.}} 1 \quad as \quad t \to \infty.$$

Theorem 5.4. *If $E(X_1^+)^r < \infty$ for some $r \ge 1$, then*

$$\frac{X_{\nu(t)}}{(\lambda(t))^{1/r}} \xrightarrow{\text{a.s.}} 0 \quad as \quad t \to \infty. \tag{5.18}$$

The proofs are the same as those of Theorems 3.4.2 and 3.4.3 with Theorem 3.4.1 replaced by Theorem 5.2. We omit the details.

We now turn to the Marcinkiewicz–Zygmund law for $\nu(t)$.

Theorem 5.5. *Let* $1 \leq r < 2$ *and suppose that* $E|X_1|^r < \infty$. *Then*

$$\frac{\nu(t) - \lambda(t)}{(\lambda(t))^{1/r}} \xrightarrow{\text{a.s.}} 0 \quad as \quad t \to \infty. \tag{5.19}$$

Proof. By proceeding exactly as in the proof of Theorem 3.4.4 for $r > 1$ (for $r = 1$ nothing remains to prove) we obtain, omitting all details,

$$\frac{ta(\nu(t)) - \nu(t)\mu}{(\lambda(t))^{1/r}} \xrightarrow{\text{a.s.}} 0 \quad as \quad t \to \infty. \tag{5.20}$$

For the case $a(y) \equiv 1$ this reduces to Theorem 3.4.4. To finish off the proof for the general case we use Taylor expansion of $a(y)$ at the point $a(\lambda(t))$ (see Gut (1974a), page 298). It follows (recall (5.12)) that

$$\nu(t)\mu - ta(\nu(t))$$

$$= \nu(t)\mu - t[a(\lambda(t)) + (\nu(t) - \lambda(t)) \cdot a'(\lambda(t) + \rho(\nu(t) - \lambda(t)))]$$

$$= \mu(\nu(t) - \lambda(t)) \left[1 - \frac{\lambda(t) \cdot a'(\lambda(t) + \rho(\nu(t) - \lambda(t)))}{a(\lambda(t))} \right],$$

where $0 \leq \rho = \rho(t, \omega) \leq 1$. Set

$$Y_{\nu(t)} = (1 - \beta)^{-1} \left[1 - \frac{\lambda(t) \cdot a'(\lambda(t) + \rho(\nu(t) - \lambda(t)))}{a(\lambda(t))} \right]. \tag{5.21}$$

In view of the above we thus have

$$\nu(t)\mu - ta(\nu(t)) = (\nu(t) - \lambda(t)) \cdot \mu(1 - \beta)Y_{\nu(t)} \tag{5.22}$$

and, by (5.20), the conclusion follows from the following result, which is Lemma 3.3 of Gut (1974a). □

Lemma 5.1. *Let* $0 \leq \beta < 1$ *and let* $\{Y_{\nu(t)}, t \geq 0\}$ *be defined by (5.21). Then*

$$Y_{\nu(t)} \xrightarrow{\text{a.s.}} 1 \quad as \quad t \to \infty. \tag{5.23}$$

Proof. It suffices to show that

$$\beta(t) = \frac{\lambda(t) \cdot a'(\lambda(t) + \rho(\nu(t) - \lambda(t)))}{a(\lambda(t))} \xrightarrow{\text{a.s.}} \beta \quad as \quad t \to \infty. \tag{5.24}$$

To this end, set

$$B = \{\omega: (\lambda(t))^{-1} \cdot \nu(t) \to 1 \quad as \quad t \to \infty\}. \tag{5.25}$$

By Theorem 5.2 we have $P(B) = 1$. Let $\omega \in B$ and $\varepsilon > 0$. There exists t_0, such that $|\nu(t) - \lambda(t)| < \lambda(t)\varepsilon$ and, hence, such that

$$\lambda(t)(1 - \varepsilon) < \lambda(t) + \rho(\nu(t) - \lambda(t)) < \lambda(t)(1 + \varepsilon) \quad \text{for } t \geq t_0. \tag{5.26}$$

Since $a(y)$ is concave and nondecreasing it follows that $a'(y)$ is nonincreasing, which, together with (5.26), yields

$$\frac{\lambda(t) \cdot a'(\lambda(t)(1 + \varepsilon))}{a(\lambda(t))} < \beta(t) < \frac{\lambda(t)a'(\lambda(t)(1 - \varepsilon))}{a(\lambda(t))}. \tag{5.27}$$

By invoking Lemma B.2.3 it follows that, for any $\gamma > 0$, we have

$$\lim_{t \to \infty} \frac{\lambda(t)a'(\lambda(t)\gamma)}{a(\lambda(t))} = \lim_{s \to \infty} \frac{s \cdot a'(s\gamma)}{a(s)} = \gamma^{-1} \lim_{s \to \infty} \frac{s\gamma a'(s\gamma)}{a(s\gamma)} \cdot \frac{a(s\gamma)}{a(s)}$$

$$= \gamma^{-1} \cdot \beta \cdot \gamma^{\beta} = \beta\gamma^{\beta-1}, \tag{5.28}$$

and hence, in view of (5.27), that

$$\beta(1 + \varepsilon)^{\beta-1} \leq \liminf_{t \to \infty} \beta(t) \leq \limsup_{t \to \infty} \beta(t) \leq \beta(1 - \varepsilon)^{\beta-1}. \tag{5.29}$$

Since ε was arbitrarily chosen and $P(B) = 1$, (5.24) follows. $\square$

The Central Limit Theorem

The following result is Gut (1974a), Theorem 3.5. The case $a(y) = y^{\beta}$ $(0 \leq \beta < 1)$ was earlier proved by Siegmund (1968) with a different method.

Theorem 5.6. *Suppose that* $\operatorname{Var} X_1 = \sigma^2 < \infty$. *Then*

$$\frac{\nu(t) - \lambda(t)}{\sigma(\mu(1 - \beta))^{-1}\sqrt{\lambda(t)}} \xrightarrow{d} N(0,1) \quad as \quad t \to \infty. \tag{5.30}$$

Proof. By Anscombe's theorem (Theorem 1.3.1), Theorem 5.4 and (5.14) (cf. the proof of Theorem 3.5.1) we obtain

$$\frac{t \cdot a(\nu(t)) - \nu(t)\mu}{\sigma\sqrt{\lambda(t)}} \xrightarrow{d} N(0,1) \quad as \quad t \to \infty, \tag{5.31}$$

which, in view of (5.22), is the same as

$$\frac{\big(\nu(t) - \lambda(t)\big)\mu(1 - \beta)}{\sigma\sqrt{\lambda(t)}} Y_{\nu(t)} \xrightarrow{d} -N(0,1) \quad as \quad t \to \infty. \tag{5.32}$$

By invoking the symmetry of the normal distribution, Lemma 5.1 and Cramér's theorem we finally obtain the desired conclusion. $\square$

Remark 5.3. A different approach has been used in Chow and Hsiung (1976) for the case $a(y) = y^\beta$, $y \geq 0$, $(0 \leq \beta < 1)$. They use the fact that

$$\nu(t) = \min\left\{n\colon \frac{S_n}{n^\beta} > t\right\} \quad (t \geq 0), \tag{5.33}$$

that is, $\{\nu(t), t \geq 0\}$ is the first passage time process for $\{(S_n/n^\beta), n \geq 1\}$. Moreover,

$$\{\nu(t) > n\} = \{S_k \leq tk^\beta, 1 \leq k \leq n\} = \left\{\max_{1 \leq k \leq n} \frac{S_k}{k^\beta} \leq t\right\}, \tag{5.34}$$

that is $\{\nu(t), t \geq 0\}$ and $\{\max_{1 \leq k \leq n}(S_k/k^\beta), n \geq 1\}$ are inverse processes (cf. (3.3.13)) for the case $\beta = 0$). A central limit theorem for the latter process follows from their Corollary 2.5(i). This result is then inverted to yield (5.30) (their Theorem 3.1). As a corollary of our Theorem 5.6 we can thus (via (5.34)) reobtain their result by inversion (cf. Teicher (1973)).

We also mention, without proof, that it is possible to obtain convergence to stable distributions under suitable conditions as follows.

Theorem 5.7. *Let $0 \leq \beta < 1$. If the assumptions of Theorem 3.5.2 are satisfied, then*

(i)
$$\frac{B_{\nu(t)}}{B(\lambda(t))} \xrightarrow{P} 1 \quad \text{and} \quad \frac{X_{\nu(t)}}{B(\lambda(t))} \xrightarrow{P} 0 \quad \text{as} \quad t \to \infty,$$

(ii)
$$\frac{\nu(t) - \lambda(t)}{(\mu(1-\beta))^{-1}B(\lambda(t))} \xrightarrow{d} -G_\alpha \quad \text{as} \quad t \to \infty,$$

where G_α is (a random variable distributed according to) a stable law with index α, $1 < \alpha \leq 2$.

For proofs and details, see Gut (1974a), Lemma 3.5 and Theorem 3.8.

Uniform Integrability

Theorem 5.8. *Let $r \geq 1$. There exists $t_0 > 0$ such that*

(i) $E(X_1^-)^r < \infty$ $\implies$ $\left\{\left(\frac{\nu(t)}{\lambda(t)}\right)^r, t \geq t_0\right\}$ *is uniformly integrable;*

(ii) $E(X_1^+)^r < \infty$ $\implies$ $\left\{\left(\frac{X_{\nu(t)}}{(\lambda(t))^{1/r}}\right)^r, t \geq t_0\right\}$ *is uniformly integrable;*

(iii) $E|X_1|^r < \infty$ $\implies$ $\left\{\left(\frac{S_{\nu(t)}}{\lambda(t)}\right)^r, t \geq t_0\right\}$ *is uniformly integrable.*

Proof. (i) Following Siegmund (1967) (see also Gut (1974a), p. 295) we introduce the auxiliary first passage times

$$\nu^*(t) = \min\{n: S_n > \mu\beta^*(\lambda(t))n + \mu\lambda(t)(1 - \beta^*(\lambda(t)))\} \quad (t \geq 0), \quad (5.35)$$

where

$$\beta^*(y) = \frac{y \cdot a'(y)}{a(y)} \quad (y \geq 0). \quad (5.36)$$

By Lemma B.2.3 we have

$$\lim_{t\to\infty} \beta^*(t) = \beta \quad (5.37)$$

(cf. also (5.28) with $\gamma = 1$). We can therefore choose t_1 such that $0 \leq \beta^*(t) < 1$ for $t > t_1$.

Now, $\mu\beta^*(\lambda(t))y + \lambda(t)\mu(1 - \beta^*(\lambda(t)))$ is a line support to $ta(y)$ ($= \mu\lambda(t)a(y)/a(\lambda(t))$, cf. (5.12)) at the point $(\lambda(t), \mu\lambda(t))$ and because of the concavity of $a(y)$ it follows that

$$\nu(t) \leq \nu^*(t). \quad (5.38)$$

We further note that

$$\nu^*(t) = \min\left\{n: \sum_{k=1}^n \frac{X_k - \mu\beta^*(\lambda(t))}{\mu(1 - \beta^*(\lambda(t)))} > \lambda(t)\right\}, \quad (5.39)$$

that is, $\nu^*(t)$ is the first passage time of a random walk whose increments have mean 1. However, the random walk itself depends on t and we shall therefore introduce a second family of auxiliary first passage times.

In view of (5.37) we choose ε ($0 < \varepsilon < 1 - \beta$) and $t_2 \geq t_1$ such that $t > t_2$ implies that $|\beta^*(t) - \beta| < \varepsilon$. Following Gut (1974a), p. 296 we now define

$$\nu^{**}(t) = \min\{n: S_n > \mu(\beta + \varepsilon)n + \lambda(t)\mu(1 - \beta + \varepsilon)\}. \quad (5.40)$$

By construction we have $\nu^*(t) \leq \nu^{**}(t)$ for $t \geq t_2$, which together with (5.38) shows that

$$\nu(t) \leq \nu^*(t) \leq \nu^{**}(t) \quad (\text{for } t \geq t_2). \quad (5.41)$$

Now we observe that

$$\nu^{**}(t) = \min\left\{n: \sum_{k=1}^n \frac{X_k - \mu(\beta + \varepsilon)}{\mu(1 - \beta + \varepsilon)} > \lambda(t)\right\}, \quad (5.42)$$

that is, $\nu^{**}(t)$ is the first passage time across the level $\lambda(t)$ of a random walk whose increments are i.i.d. with mean $(1 - \beta - \varepsilon)/(1 - \beta + \varepsilon) > 0$. It follows from Theorem 3.7.1 (see also Gut and Janson (1986), Section 9 and Remark 5.5 below) that

$$\left\{ \left(\frac{\nu^{**}(t)}{\lambda(t)} \right)^r, t \geq t_2 \right\} \quad \text{is uniformly integrable} \tag{5.43}$$

and thus, by (5.41), that

$$\left\{ \left(\frac{\nu(t)}{\lambda(t)} \right)^r, t \geq t_2 \right\} \quad \text{is uniformly integrable,} \tag{5.44}$$

which proves (i) (with $t_0 = t_2$).

(ii) Same as the proof of Theorem 3.7.2(i), except for obvious modifications.

(iii) By (5.7) we have, for large t,

$$0 \leq \frac{S_{\nu(t)}}{\lambda(t)} \leq \frac{a_0}{\lambda(t)} + \delta \cdot \frac{\nu(t)}{\lambda(t)} + \frac{X_{\nu(t)}}{\lambda(t)} \leq \frac{a_0}{\lambda(t)} + \delta \cdot \frac{\nu(t)}{\lambda(t)} + \frac{X_{\nu(t)}}{(\lambda(t))^{1/r}} \tag{5.45}$$

(since $P(X_{\nu(t)} > 0) = 1$ now).

By Lemma A.1.3 the conclusion thus follows if $\{(\nu(t)/\lambda(t))^r, t \geq t_0\}$ and $\{(X_{\nu(t)}/(\lambda(t))^{1/r})^r, t \geq t_0\}$ are uniformly integrable, something that, however, is ensured by (i) and (ii) since $E|X_1|^r < \infty$. The proof of the theorem is thus complete. □

Remark 5.4. Theorem 5.8 in this form seems to be new. In Gut (1974a) the convergence of $E(\nu(t)/\lambda(t))^r$ and $E((X_{\nu(t)})^r/\lambda(t))$ as $t \to \infty$ was studied by direct computation. Siegmund (1967) proved that $E(\nu(t)/\lambda(t)) \to 1$ as $t \to \infty$.

Remark 5.5. For (5.43) we used Theorem 3.7.1 which, as stated, presupposes a normalization by t. Behind that was Theorem 1.6.1, in which this normalization was used in (1.6.1) to conclude that (1.6.2) holds. Now, if we had used any other increasing normalization in (1.6.1) for the stopping time we would have obtained a result for the stopped sum with that very normalization. Another way to view it is to say that, as long as the normalization is a continuous nondecreasing function of t the problem is, in essence, a matter of how the normalizing function is parameterized (see also Gut and Janson (1986), Section 9.

A similar remark applies to (ii) and (iii) and the results below.

Remark 5.6. In view of Theorem 5.1(i) the assumption in (i) is the weakest possible. For (ii) this is obvious. For (iii) we used the cruder (5.7) instead of (5.14) and the result is therefore not necessarily the strongest one. For example, if $a(y) = y^\beta$, (5.12) and (5.14) yield (recall Remark 5.1), for large t,

$$\mu \left(\frac{\nu(t)}{\lambda(t)} \right)^\beta = \frac{ta(\nu(t))}{\lambda(t)} < \frac{S_{\nu(t)}}{\lambda(t)} \leq \frac{ta(\nu(t))}{\lambda(t)} + \frac{X_{\nu(t)}}{\lambda(t)} \leq \mu \left(\frac{\nu(t)}{\lambda(t)} \right)^\beta + \frac{X_{\nu(t)}}{(\lambda(t))^{1/r}}$$

and it follows that $E(X_1^+)^r < \infty$ and $E(X_1^-)^{\beta r} < \infty$ together imply that $\{(S_{\nu(t)}/\lambda(t))^r, t \geq t_0\}$ is uniformly integrable. Moreover, it follows from (5.9) that the assumptions are also necessary.

It is also possible to obtain results like Theorem 5.8 corresponding to the Marcinkiewicz–Zygmund law or the central limit theorem, but everything becomes technically much more complicated.

The first step is to proceed as in the proof of Theorems 3.7.3 and 3.7.4. Thus, suppose first that $E|X_1|^r < \infty$ for some r $(1 < r < 2)$. Then Theorem 5.8(i), formula (3.7.6), with t replaced by $t \cdot a(\nu(t))$, Theorem 1.6.2, Theorem 5.8(ii) and Lemma A.1.3 (recall Remark 5.5) together show that

$$\left\{ \left| \frac{\nu(t)\mu - ta(\nu(t))}{(\lambda(t))^{1/r}} \right|^r, t \geq t_0 \right\} \quad \text{is uniformly integrable.} \tag{5.46}$$

To obtain the desired conclusion we must do some additional work, sketched below, corresponding to Taylor expansion and Lemma 5.1 above. This, finally, leads to the following result for the case $a(y) = y^\beta$, $y \geq 0$, $(0 \leq \beta < 1)$ (and $\lambda(t) = (t/\mu)^{1/(1-\beta)}$, $t \geq 0$).

Theorem 5.9. *Suppose that $a(y) = y^\beta$, $y \geq 0$, for some β $(0 \leq \beta < 1)$. If $E|X_1|^r < \infty$ for some r $(1 < r < 2)$, then*

$$\left\{ \left| \frac{\nu(t) - \frac{t}{\mu}(\nu(t))^\beta}{(\lambda(t))^{1/r}} \right|^r, t \geq t_0 \right\} \quad \text{is uniformly integrable.} \tag{5.47}$$

Moreover, when $0 \leq \beta \leq 1/r$,

$$\left\{ \left| \frac{\nu(t) - \lambda(t)}{(\lambda(t))^{1/r}} \right|^r, t \geq t_0 \right\} \quad \text{is uniformly integrable.} \tag{5.48}$$

If, in addition,

$$(\lambda(t))^{r-1} \cdot P(X_1 \geq t) \to 0 \quad as \quad t \to \infty, \tag{5.49}$$

then (5.48) also holds when $1/r < \beta < 1$.

As for the central limit theorem one obtains similarly, by using Theorems 5.8(i), 1.6.3 and 5.8(ii) and Lemma A.1.3 that if $E|X_1|^r < \infty$ for some $r \geq 2$, then

$$\left\{ \left| \frac{\nu(t)\mu - ta(\nu(t))}{\sqrt{\lambda(t)}} \right|^r, t \geq t_0 \right\} \quad \text{is uniformly integrable,} \tag{5.50}$$

which, again with $a(y) = y^\beta$, $y \geq 0$, $(0 \leq \beta < 1)$, leads to the following.

Theorem 5.10. *Suppose that $a(y) = y^\beta$, $y \geq 0$, for some β $(0 \leq \beta < 1)$. If $E|X_1|^r < \infty$ for some $r \geq 2$, then*

$$\left\{ \left| \frac{\nu(t) - \frac{t}{\mu}(\nu(t))^\beta}{\sqrt{\lambda(t)}} \right|^r, t \geq t_0 \right\} \quad \text{is uniformly integrable.} \tag{5.51}$$

Moreover, when $0 \leq \beta \leq 1/2$,

$$\left\{ \left| \frac{\nu(t) - \lambda(t)}{\sqrt{\lambda(t)}} \right|^r, t \geq t_0 \right\} \quad \text{is uniformly integrable.} \qquad (5.52)$$

If, in addition,

$$(\lambda(t))^{r/2} \cdot P(X_1 \geq t) \to 0 \quad as \quad t \to \infty, \qquad (5.53)$$

then (5.52) also holds when $1/2 < \beta < 1$.

The proofs of Theorems 5.9 and 5.10 are based on (5.46) and (5.50) respectively and the following lemmas. Recall that $a(y) = y^\beta$, $y \geq 0$, for some β $(0 \leq \beta < 1)$ and thus that $\lambda(t) = (t/\mu)^{1/(1-\beta)}$, $t \geq 0$.

Lemma 5.2. *For all $t > 0, x \geq 0$, we have*

$$y \geq \lambda(t) + (\lambda(t))^{1/r} x \quad \implies \quad y - (t/\mu)y^\beta \geq (1 - \beta)(\lambda(t))^{1/r} x. \qquad (5.54)$$

Lemma 5.3. *Given γ $(0 < \gamma < 1 - \beta)$ there exists θ $(0 < \theta < 1)$ such that for all $t > 0$ and $0 \leq x \leq \theta(\lambda(t))^{1-1/r}$ we have*

$$(1 - \theta)\lambda(t) \leq y \leq \lambda(t) - (\lambda(t))^{1/r} x \quad \implies \quad y - (t/\mu)y^\beta \leq -(\lambda(t))^{1/r}\gamma x. \qquad (5.55)$$

Lemma 5.4. *Let $q > 1$. If*

$$P(X_1 \geq x) = o(x^{-q}) \quad as \quad x \to \infty, \qquad (5.56)$$

then, for every θ $(0 < \theta < 1)$ we have, as $t \to \infty$,

$$P(\nu(t) \leq (1 - \theta)\lambda(t)) = \begin{cases} o((\lambda(t))^{-q(1-\beta)}), & \text{if } \beta q > 1, \\ o((\lambda(t))^{-q(1-\beta)} \log t), & \text{if } \beta q = 1, \\ o((\lambda(t))^{-(q-1)}), & \text{if } \beta q < 1. \end{cases} \qquad (5.57)$$

Moreover, if $\beta q = 1$ and (5.56) is sharpened to

$$\lim_{x \to \infty} \int_x^{x^\rho} u^{q-1} P(X_1 \geq u)du = 0 \quad \text{for some } \rho > 1, \qquad (5.58)$$

then, as $t \to \infty$,

$$P(\nu(t) \leq (1 - \theta)\lambda(t)) = o((\lambda(t))^{-q(1-\beta)}). \qquad (5.59)$$

Lemmas 5.2 and 5.3 with $r = 2$ and Lemma 5.4 are Lemmas 2, 3 and 4 of Chow, Hsiung and Lai (1979) respectively. We refer to their paper for the proofs. The proofs of Lemmas 5.2 and 5.3 for $1 < r < 2$ are essentially the same as the proofs for the case $r = 2$.

Proof of Theorem 5.9. We use Lemmas 5.2 and 5.3 to conclude that, for $M > 0$,

$$\int_M^\infty x^{r-1} \cdot P(|\nu(t) - \lambda(t)| > x(\lambda(t))^{1/r})dx$$

$$\leq \int_M^{M \vee (\lambda(t))^{1-1/r}} x^{r-1} \cdot P(\nu(t) \leq (1 - \theta)\lambda(t))dx$$

$$+ \int_M^{M \vee \theta(\lambda(t))^{1-1/r}} x^{r-1} \cdot P\left(\nu(t) - \frac{t}{\mu}(\nu(t))^\beta \leq -\gamma x(\lambda(t))^{1/r}\right) dx$$

$$+ \int_M^\infty x^{r-1} \cdot P\left(\nu(t) - \frac{t}{\mu}(\nu(t))^\beta \geq (1 - \beta)x(\lambda(t))^{1/r}\right) dx$$

$$\leq \frac{((\lambda(t))^{r-1} - M^r)^+}{r} P(\nu(t) \leq (1 - \theta)\lambda(t))$$

$$+ \int_M^\infty x^{r-1} \cdot P\left(\left|\nu(t) - \frac{t}{\mu}(\nu(t))^\beta\right| \geq \gamma x(\lambda(t))^{1/r}\right) dx. \tag{5.60}$$

From Lemma 5.4 with $q = r$ when $0 \leq \beta \leq 1/r$ and $q = (r-1)/(1-\beta) \,(> r)$ when $1/r < \beta < 1$ we conclude that

$$(\lambda(t))^{r-1} \cdot P(\nu(t) \leq (1 - \theta)\lambda(t)) \to 0 \quad \text{as} \quad t \to \infty, \tag{5.61}$$

from which it follows that the first term in (5.60) tends to 0 uniformly in t as $M \to \infty$. The second term in (5.60) tends to 0 uniformly in t as $M \to \infty$ by (5.47). We thus conclude that (5.48) holds. $\qquad\square$

Proof of Theorem 5.10. By the same reasoning as above, together with (5.51), it follows that the conclusion holds provided

$$(\lambda(t))^{r/2} \cdot P(\nu(t) \leq (1 - \theta)\lambda(t)) \to 0 \quad \text{as} \quad t \to \infty, \tag{5.62}$$

and, again, this follows from Lemma 5.4. $\qquad\square$

Remark 5.7. Theorem 5.10 is due to Chow, Hsiung and Lai (1979) and we refer to their paper for the remaining details. Theorem 5.9 is new.

Remark 5.8. The following computations show that (5.48), in fact, implies (5.49). Namely, suppose that (5.48) holds. It follows, in particular, that given $\varepsilon > 0, \gamma \,(0 < \gamma < 1)$ and t_0 sufficiently large we have, for $t > t_0$,

$$\varepsilon > (\lambda(t))^{-1} \cdot E(\lambda(t) - \nu(t))^r I\{\nu(t) \leq \gamma\lambda(t)\}$$
$$\geq (1 - \gamma)^r (\lambda(t))^{-1+r} \cdot P(\nu(t) \leq \gamma\lambda(t)) \geq (1 - \gamma)^r (\lambda(t))^{r-1} \cdot P(\nu(t) = 1)$$
$$= (1 - \gamma)^r (\lambda(t))^{r-1} \cdot P(X_1 > t),$$

that is, it follows that

$$(\lambda(t))^{r-1} \cdot P(X_1 > t) \to 0 \quad \text{as} \quad t \to \infty, \tag{5.63}$$

which proves that (5.49) is necessary as claimed. Now, if $\beta \leq 1/r$ then $1/(1-\beta) \leq r/(r-1)$ and $(\lambda(t))^{r-1} \leq (t/\mu)^r$, that is, (5.49) is automatically satisfied. If $1/r < \beta < 1$, then (5.49) is a necessary extra assumption.

Remark 5.9. The same kind of argument shows that (5.52) $\Longrightarrow$ (5.53) and that (5.53) is automatically satisfied when $0 \leq \beta \leq \frac{1}{2}$. When $1/2 < \beta < 1$ (5.53) thus is a necessary additional assumption.

Remark 5.10. Note that the final part of the proof amounted to showing that (5.49) and (5.53) were satisfied and that these relations are precisely those mentioned in Remarks 5.8 and 5.9.

Moment Convergence

By using the above results in the same way as the corresponding results in Chapter 3 were used in Section 3.8 we obtain the following theorems for the convergence of moments in the strong laws and the central limit theorem.

Theorem 5.11. *Let $r \geq 1$. Then*

(i) $E(X_1^-)^r < \infty \quad \Longrightarrow \quad E\left(\dfrac{\nu(t)}{\lambda(t)}\right)^p \to 1 \quad \text{as} \quad t \to \infty \;$ *for all* $p, 0 < p \leq r$;

(ii) $E(X_1^+)^r < \infty \quad \Longrightarrow \quad E\dfrac{X_{\nu(t)}^r}{\lambda(t)} \to 0 \quad \text{as} \quad t \to \infty$;

(iii) $E|X_1|^r < \infty \quad \Longrightarrow \quad E\left(\dfrac{S_{\nu(t)}}{\lambda(t)}\right)^p \to \mu^p \quad \text{as} \quad t \to \infty \;$ *for all* $p, 0 < p \leq r$.

Theorem 5.12. *Let $1 < r < 2$ and suppose that $a(y) = y^\beta$, $y \geq 0$, for some β $(0 \leq \beta \leq 1/r)$. If $E|X_1|^r < \infty$, then*

$$E\left|\frac{\nu(t) - \lambda(t)}{(\lambda(t))^{1/r}}\right|^r \to 0 \quad \text{as} \quad t \to \infty. \tag{5.64}$$

If $1/r < \beta < 1$ the conclusion remains true provided, in addition,

$$(\lambda(t))^{r-1} \cdot P(X_1 > t) \to 0 \quad \text{as} \quad t \to \infty. \tag{5.65}$$

Theorem 5.13. *Suppose that $a(y) = y^\beta$, $y \geq 0$, for some β $(0 \leq \beta \leq 1/2)$. Set $\sigma^2 = \mathrm{Var}\, X_1$. If $E|X_1|^r < \infty$ for some $r \geq 2$, then*

(i) $\quad E\left|\dfrac{\nu(t) - \lambda(t)}{\sqrt{\lambda(t)}}\right|^p \to E|Z|^p \quad \text{as} \quad t \to \infty \quad \text{for all } p, 0 < p \leq r,$

where Z *is normal with mean 0 and variance* $\sigma^2/\mu^2(1-\beta)^2$;

(ii) $E\left(\dfrac{\nu(t) - \lambda(t)}{\sqrt{\lambda(t)}}\right)^{k} \to 0$ as $t \to \infty$ for $k = $ odd integer $\leq r$.

If $1/2 < \beta < 1$ the conclusions remain true provided, in addition,

$$(\lambda(t))^{r/2} \cdot P(X_1 > t) \to 0 \quad as \quad t \to \infty. \tag{5.66}$$

Remark 5.11. Theorem 5.11(i) and (ii) were proved by direct estimates in Gut (1974a), Section 3; it was shown that

$$E\left(\frac{\nu^{**}(t)}{\lambda(t)}\right)^{r} \to \left(\frac{1 - \beta + \varepsilon}{1 - \beta - \varepsilon}\right)^{r} \quad as \quad t \to \infty$$

with $\nu^{**}(t)$ defined by (5.40) and then that a corresponding result holds for a related family of first passage times $\{\nu_{**}(t)\}$, from which it was concluded that

$$E\left(\frac{\nu^{*}(t)}{\lambda(t)}\right)^{r} \to 1 \quad as \quad t \to \infty$$

with $\nu^{*}(t)$ defined by (5.35). By (5.38) it followed that

$$\limsup_{t\to\infty} E\left(\frac{\nu(t)}{\lambda(t)}\right)^{r} \leq 1.$$

The conclusion then followed from Theorem 5.2 and Fatou's lemma.

Remark 5.12. Theorem 5.13 is due to Chow, Hsiung and Lai (1979). In particular, if $\operatorname{Var} X_1 = \sigma^2 < \infty$, the theorem shows that the relation

$$\operatorname{Var}\nu(t) = \frac{\sigma^2}{\mu^2(1-\beta)^2}\left(\frac{t}{\mu}\right)^{1/(1-\beta)} + o(t^{1/(1-\beta)}) \quad as \quad t \to \infty \tag{5.67}$$

always holds when $\beta \leq \frac{1}{2}$, but only under the additional assumption

$$t^{1/(1-\beta)} \cdot P(X_1 > t) \to 0 \quad as \quad t \to \infty \tag{5.68}$$

when $1/2 < \beta < 1$.
    Theorem 5.12 is new.

The Overshoot

The quantity of interest here is

$$R(t) = S_{\nu(t)} - t a(\nu(t)) \quad (t \geq 0). \tag{5.69}$$

Since $0 \leq R(t) \leq X_{\nu(t)}$ (cf. (3.10.3)) the following result is easily established.

Theorem 5.14. *Let $r \geq 1$ and suppose that $E(X_1^+)^r < \infty$. Then*

(i) $E(R(t))^r < \infty$;

(ii) $\frac{R(t)}{(\lambda(t))^{1/r}} \xrightarrow{\text{a.s.}} 0 \quad as \quad t \to \infty$;

(iii) $\left\{ \frac{(R(t))^r}{\lambda(t)}, t \geq t_0 \right\}$ *is uniformly integrable for some $t_0 \geq 1$*;

(iv) $E(R(t))^r = o(\lambda(t))$ *as $t \to \infty$.*

Proof. By Lemma 1.8.1 we have $E(R(t))^r \leq E(X_{\nu(t)})^r \leq E\nu(t) \cdot E(X_1^+)^r < \infty$, which proves (i), (ii) follows from Theorem 5.4, (iii) follows from Theorem 5.8(ii) and (iv) follows from (ii) and (iii). $\qquad \square$

The Law of the Iterated Logarithm

By using the ideas of the present section the following law of the iterated logarithm was proved in Gut (1985), Theorem 5.

Theorem 5.15. *Suppose that $Var X_1 = \sigma^2 < \infty$. Then*

$$C\left(\left\{ \frac{\nu(t) - \lambda(t)}{\sqrt{2 \frac{\sigma^2}{\mu^2 (1-\beta)^2} \lambda(t) \log \log t}}, t \geq 3 \right\} \right) = [-1, 1] \text{ a.s.} \qquad (5.70)$$

In particular,

$$\limsup_{t \to \infty} \left(\liminf_{t \to \infty} \right) \frac{\nu(t) - \lambda(t)}{\sqrt{2\lambda(t) \log \log t}} = \frac{+}{(-)} \frac{\sigma}{\mu(1 - \beta)} \text{ a.s.} \qquad (5.71)$$

Outline of Proof. To prove (5.70) one first establishes that

$$C\left(\left\{ \frac{ta(\nu(t)) - \nu(t)\mu}{\sqrt{2\sigma^2 \lambda(t) \log \log t}}, t \geq 3 \right\} \right) = [-1, 1] \text{ a.s.} \qquad (5.72)$$

(cf. (5.20) and (5.31)). One then uses Taylor expansion, introduces $Y_{\nu(t)}$ defined in (5.21) and applies Lemmas 5.1 and 3.11.1. $\qquad \square$

Remark 5.13. For the case $a(y) = y^\beta$, $y \geq 0$, $(0 \leq \beta < 1)$ Chow and Hsiung (1976) have used the approach described in Remark 5.3 above to prove a law of the iterated logarithm for $\max_{1 \leq k \leq n} S_k / k^\beta$ (see their Corollary 2.5(ii)) and inverted this result to obtain (5.71). Similarly, we can invert (5.71) to obtain their result.

Extensions to the Non-I.I.D. Case

Just as for Chapter 3 (see Section 3.14) several results in this section have been extended to random walks whose increments do not have a common distribution or are subject to certain dependence assumptions.

The results in Siegmund (1967, 1968), which correspond to Theorems 5.2, 5.11(i) with $r = 1$ and 5.6, respectively, are proved without assuming that the increments have a common distribution. Theorems 5.1, 5.2 and 5.11 are proved under certain dependence assumptions in Gut (1974b). Several results in this section are extended to processes with independent, stationary increments (with positive mean) in Gut (1975b). Finally, we recall nonlinear renewal theory mentioned in the beginning of this section, and the more general perturbed random walks which are the topic of Chapter 6.

5

Functional Limit Theorems

5.1 Introduction

The classical central limit theorem was generalized to a functional central limit theorem by Donsker (1951) (see Theorem A.3.2). In words the result means that one considers the partial sums $\{S_0, S_1, \ldots, S_n\}$ of i.i.d. variables jointly for each n and shows that if the mean and variance are finite then the (polygonal) process obtained by normalization (and linear interpolation), behaves, asymptotically, like Brownian motion.

In this chapter we shall begin by presenting corresponding generalizations of Anscombe's theorem (Theorem 1.3.1) and of the results about asymptotic normality of the first passage times discussed in Chapter 3 (see Theorem 3.5.1). We shall also deal with the stable case, that is when the variance is not necessarily finite, and in this case our point of departure is Theorem A.3.3. The corresponding one-dimensional results are Theorems 1.3.2 and 3.5.2. We then present functional (central) limit theorems for the various processes treated in Chapter 4. In the final section we discuss functional versions of the law of the iterated logarithm for some of the processes. The starting point here is Strassen (1964), see Theorem A.4.1.

We also mention (cf. Appendix A) that results of this kind sometimes are called *invariance principles*; results like Donsker's theorem are called weak invariance principles and results like Strassen's are called strong invariance principles. Some background material and notation for this chapter can be found in Sections A.3 and A.4.

5.2 An Anscombe–Donsker Invariance Principle

Let $\{\xi_k, k \geq 1\}$ be i.i.d. random variables, suppose that $E\xi_1 = 0, \operatorname{Var} \xi_1 = \sigma^2 < \infty$, set $S_n = \sum_{k=1}^{n} \xi_k$, $n \geq 0$, $(S_0 = 0)$ and define

$$X_n(t, \omega) = \frac{1}{\sigma\sqrt{n}} S_{[nt]}(\omega) \qquad (t \geq 0). \tag{2.1}$$

Donsker's theorem states that $X_n \stackrel{J_1}{\Longrightarrow} W$ as $n \to \infty$, where W is the Wiener measure on $(C, \mathscr{C})$—the space of continuous functions on $[0, \infty)$ (cf. Theorem A.3.2 and Remark A.3.8).

Now, suppose that $\{N(t), t \geq 0\}$ is a nondecreasing, right-continuous family of positive, integer valued random variables and define

$$Y_n(t, \omega) = \frac{1}{\sigma\sqrt{n}} S_{N(nt,\omega)}(\omega) \qquad (t \geq 0). \tag{2.2}$$

Theorem 2.1. *Suppose that*

$$\frac{N(t)}{t} \xrightarrow{\text{a.s.}} \theta \quad (0 < \theta < \infty) \quad as \quad t \to \infty. \tag{2.3}$$

Then

$$\theta^{-1/2} \cdot Y_n \stackrel{J_1}{\Longrightarrow} W \quad as \quad n \to \infty. \tag{2.4}$$

Proof. Define the random change of time $\Phi_n \in D_0$ by

$$\Phi_n(t, \omega) = \frac{N(nt, \omega)}{n} \qquad (t \geq 0), \tag{2.5}$$

let $\varphi(t) = t\theta, \ t \geq 0$, and note that $\varphi \in D_0 \cap C$.

Lemma 2.1. $\Phi_n \xrightarrow{P} \varphi(J_1)$ *as* $n \to \infty$.

Proof. We wish to show that $\sup_{0 \leq t \leq b} |\Phi_n(t, \omega) - \varphi(t)|$ tends to 0 in probability as $n \to \infty$ for every $b > 0$.

Now, by (2.3),

$$\frac{N(nt)}{n} - t\theta \xrightarrow{\text{a.s.}} 0 \quad as \quad n \to \infty \text{ for each } \textit{fixed } t. \tag{2.6}$$

Let Q be the set of rationals in $[0, \infty)$. It follows that

$$\frac{N(nt)}{n} \to t\theta \quad as \quad n \to \infty \text{ for all } t \in Q \quad \text{a.s.} \tag{2.7}$$

The conclusion of Lemma 2.1 now follows from the following result, the proof of which we leave to the reader.   □

Lemma 2.2. *Suppose that* $\{f_n(t), n \geq 1\}$ *is a sequence of real valued functions, such that* $f_n(t)$, *for all* $n \geq 1$, *is nondecreasing as a function of* $t \ (0 \leq t \leq 1)$ *and suppose that* $f_n(t) \to t$ *as* $n \to \infty$ *for all* $t \in Q$. *Then the convergence is uniform, that is,*

$$\sup_{0 \leq t \leq 1} |f_n(t) - t| \to 0 \quad as \quad n \to \infty. \tag{2.8}$$

The proof of Lemma 2.1 thus is complete and we return to the proof of Theorem 2.1.

By combining Donsker's theorem and Lemma 2.1 with Theorem A.3.9 we now otain

$$(X_n, \Phi_n) \overset{J_1}{\Longrightarrow} (W, \varphi) \quad \text{as} \quad n \to \infty, \tag{2.9}$$

to which we can apply the continuous mapping theorem (Theorem A.3.5) and conclude that

$$Y_n = X_n \circ \Phi_n \overset{J_1}{\Longrightarrow} W \circ \varphi \quad \text{as} \quad n \to \infty \tag{2.10}$$

in view of the following lemma.

Lemma 2.3. *Let* $\psi \colon D \times D_0 \to D$ *be defined by*

$$\psi(x, \varphi) = x \circ \varphi. \tag{2.11}$$

Then ψ *is continuous at* (x, φ) *for* $x \in C$ *and* $\varphi \in C \cap D_0$.

Proof. Let $b > 0$ and let $\{x_n, n \geq 1\}$ be a sequence of elements of D, such that $x_n \to x(J_1)$, where $x \in C$. Further, let $\{\varphi_n, n \geq 1\}$ belong to D_0 and suppose that $\varphi_n \to \varphi(J_1)$, where $\varphi \in C \cap D_0$. Then there exists $B < \infty$, such that $\varphi_n(b) \leq B$ for all $n \geq 1$. It follows that

$$\sup_{0 \leq t \leq b} |x_n(\varphi_n(t)) - x(\varphi(t))|$$

$$\leq \sup_{0 \leq t \leq b} |x_n(\varphi_n(t)) - x(\varphi_n(t))| + \sup_{0 \leq t \leq b} |x(\varphi_n(t)) - x(\varphi(t))|$$

$$\leq \sup_{0 \leq t \leq B} |x_n(t) - x(t)| + \sup_{0 \leq t \leq b} |x(\varphi_n(t)) - x(\varphi(t))| \to 0 \quad \text{as} \quad n \to \infty;$$

the first term tends to 0 as $n \to \infty$ since $x_n \to x$ uniformly on $[0, B]$ as $n \to \infty$ (recall that $x \in C$) and the second term tends to 0 as $n \to \infty$ since x is uniformly continuous on compact intervals and $\varphi_n \to \varphi$ uniformly on $[0, b]$ as $n \to \infty$. ☐

Since (2.10) is the same as (2.4), the proof of the theorem is complete. ☐

Billingsley (1968), Section 17 (Billingsley (1999), Section 14) studies the process

$$Y_n'(t, \omega) = \frac{1}{\sigma \sqrt{N(n)}} S_{[t \cdot N(n, \omega)]}(\omega) \qquad (0 \leq t \leq 1), \tag{2.12}$$

where now $\{N(n), n \geq 1\}$ is a sequence of positive, integer valued random variables and $\{S_n, n \geq 0\}$ is a sequence of partial sums of arbitrary random variables. It is shown there that if $\{a_n, n \geq 1\}$ are constants tending to infinity and

$$\frac{N(n)}{a_n} \overset{p}{\to} \theta \quad (0 < \theta < \infty) \quad \text{as} \quad n \to \infty, \tag{2.13}$$

then

$$Y_n' \overset{J_1}{\Longrightarrow} W \circ \varphi \quad \text{as} \quad n \to \infty, \tag{2.14}$$

provided $X_n \overset{J_1}{\Longrightarrow} W$ as $n \to \infty$ (where now W is the Wiener measure on $C[0,1]$).

To prove this result it is first shown that

$$Y_n'' \overset{J_1}{\Longrightarrow} W \circ \varphi \quad \text{as} \quad n \to \infty, \tag{2.15}$$

where

$$Y_n''(t,\omega) = \frac{1}{\sigma\sqrt{[a_n]}} S_{[t \cdot N(n,\omega)]}(\omega) \qquad (0 \le t \le 1), \tag{2.16}$$

and φ is as before.

The proof of (2.15) follows the same basic pattern as the proof of Theorem 2.1 (or rather vice versa), but the part corresponding to Lemma 2.1 is easier, since

$$\sup_{0 \le t \le 1} \left| t \cdot \frac{N(n,\omega)}{a_n} - t\theta \right| = \left| \frac{N(n,\omega)}{a_n} - \theta \right| \overset{p}{\to} 0 \quad \text{as} \quad n \to \infty \tag{2.17}$$

in view of (2.13).

The proof of Lemma 2.1 is adapted from Gut (1973), where the random indices are the first passage times studied in Chapter 3 and Section 4.5 (with $L(y) \equiv 1$). For first passage times of renewal processes, see Billingsley (1968), Theorem 17.3 (Billingsley (1999), Theorem 14.6). A further reference is Serfozo (1975).

Remark 2.1. It is also possible to use more general normalizations than linear ones in Theorem 2.1 (cf. Remark 4.5.5).

Remark 2.2. The assumption (2.3) was used in the proof of Lemma 2.1. We observe that what we really need is that

$$\sup_{0 \le t \le b} \left| \frac{N(nt)}{n} - t\theta \right| \overset{p}{\to} 0 \quad \text{as} \quad n \to \infty \quad (0 < \theta < \infty) \quad \text{for every } b > 0. \tag{2.18}$$

But (2.3) is (as we have seen) sufficient for this to hold and will, in fact, be satisfied in the applications to follow.

Remark 2.3. Since the projections from C to R^k are continuous it follows from the continuous mapping theorem that the finite-dimensional distributions of Y_n converge to multidimensional normal distributions, which yields multi-dimensional versions of Anscombe's theorem. In particular, the case $k = 1$ and $t = 1$ yields Theorem 1.3.1(ii) (under slightly stronger assumptions).

The following is a stable version of the previous result.

Let $\{\xi_k, k \geq 1\}$ be i.i.d. random variables with mean 0 and set $S_n = \sum_{k=1}^{n} \xi_k,\ n \geq 0$. Suppose that $\{B_n, n \geq 1\}$ are positive normalizing coefficients such that

$$\frac{S_n}{B_n} \xrightarrow{d} G_\alpha \quad \text{as} \quad n \to \infty, \tag{2.19}$$

where G_α is (a random variable distributed according to) a stable law with index α $(1 < \alpha \leq 2)$ and let X be the stable process whose one-dimensional marginal distribution at $t = 1$ is G_α. Finally, let $\{N(t), t \geq 0\}$ be a non-decreasing, right-continuous family of positive, integer valued random variables as above and define

$$X_n(t, \omega) = \frac{S_{[nt]}(\omega)}{B_n} \quad \text{and} \quad Y_n(t, \omega) = \frac{S_{N(nt,\omega)}(\omega)}{B_n} \qquad (t \geq 0). \tag{2.20}$$

Theorem 2.2. *If*

$$\frac{N(t)}{t} \xrightarrow{\text{a.s.}} \theta \quad (0 < \theta < \infty) \quad as \quad t \to \infty \tag{2.21}$$

then

$$\theta^{-1/\alpha} \cdot Y_n \overset{J_1}{\Longrightarrow} X \quad as \quad n \to \infty. \tag{2.22}$$

Proof. Consider, again, the random time change (2.5). Since (2.19) is equivalent to $X_n \overset{J_1}{\Longrightarrow} X$ as $n \to \infty$ (see Gikhman and Skorohod (1969), Chapter IX.6) we conclude, in view of Lemma 2.1, (cf. (2.9)) that

$$(X_n, \Phi_n) \overset{J_1}{\Longrightarrow} (X, \varphi) \quad \text{as} \quad n \to \infty. \tag{2.23}$$

Thus, so far the proof is essentially the same as the proof of Theorem 2.1. For the next step we need a counterpart to Lemma 2.3, that is, we need to know that composition is continuous at points (x, φ) with $x \in D$ and strictly increasing $\varphi \in C \cap D_0$. That this is, indeed, the case is proved in Whitt (1980), Theorem 3.1. It therefore follows from Theorem A.3.5 that

$$Y_n = X_n \circ \Phi_n \overset{J_1}{\Longrightarrow} X \circ \varphi \quad \text{as} \quad n \to \infty, \tag{2.24}$$

and by checking e.g. the characteristic function of a stable process it is easily seen that (2.24) is the same as (2.22), which completes the proof. $\qquad\square$

Remark 2.4. Remarks 2.1–2.3 all apply here too. The corresponding one-dimensional result is Theorem 1.3.2.

Remark 2.5. Observe that convergence of the finite-dimensional distributions was not used in the proofs of Theorems 2.1 and 2.2, since we used continuous mapping results. Had we proved the result directly we would have had to prove convergence of the finite-dimensional distributions plus tightness (cf. Theorem A.3.1).

5.3 First Passage Times for Random Walks with Positive Drift

We shall now use the results from Section 5.2 to derive functional limit theorems for the first passage times discussed in Chapter 3. Theorem 2.1 will here play the role of Anscombe's theorem (cf. the proof of Theorem 3.5.1).

For renewal processes the following result is due to Billingsley (1968), see Theorem 17.3 there (or Billingsley (1999), Theorem 14.6), and for random walks the result has been proved in Vervaat (1972a,b), Basu (1972) and Gut (1973); the two latter proofs are closely related to Billingsley's.

Theorem 3.1. Let $\{\xi_k, k \geq 1\}$ be i.i.d. *random variables with mean* μ $(0 < \mu < \infty)$ *and variance* σ^2 $(0 < \sigma^2 < \infty)$. *Set* $S_n = \sum_{k=1}^n \xi_k$, $n \geq 0$, *and define*

$$\nu(t) = \min\{n\colon S_n > t\} \qquad (t \geq 0) \tag{3.1}$$

and

$$Z_n(t, \omega) = \frac{\nu(nt, \omega) - nt/\mu}{\sigma\mu^{-3/2}\sqrt{n}} \qquad (t \geq 0). \tag{3.2}$$

Then

$$Z_n \overset{J_1}{\Longrightarrow} W \quad as \quad n \to \infty. \tag{3.3}$$

Proof. Set

$$Y_n(t, \omega) = \frac{1}{\sigma\sqrt{n}} \sum_{k=1}^{\nu(nt,\omega)} (\xi_k(\omega) - \mu) \qquad (t \geq 0). \tag{3.4}$$

Since, $\nu(t)/t \overset{\text{a.s.}}{\longrightarrow} 1/\mu$ as $t \to \infty$ (Theorem 3.4.1) we can apply Theorem 2.1 (with $\theta = \mu^{-1}$) to assert that

$$\mu^{1/2} \cdot Y_n \overset{J_1}{\Longrightarrow} W \quad as \quad n \to \infty. \tag{3.5}$$

Next we observe (cf. (3.3.2)) that

$$\frac{nt - \mu\nu(nt, \omega)}{\sigma\sqrt{n}} < Y_n(t, \omega) \leq \frac{nt - \mu\nu(nt, \omega)}{\sigma\sqrt{n}} + \frac{\xi_{\nu(nt,\omega)}(\omega)}{\sigma\sqrt{n}}. \tag{3.6}$$

Define

$$Z_n'(t, \omega) = \mu^{1/2} \cdot \frac{nt - \mu\nu(nt, \omega)}{\sigma\sqrt{n}} \qquad (t \geq 0). \tag{3.7}$$

Since $\xi_n/\sqrt{n} \overset{\text{a.s.}}{\longrightarrow} 0$ as $n \to \infty$ it follows that $\max_{1 \leq k \leq n} \xi_k/\sqrt{n} \overset{\text{a.s.}}{\longrightarrow} 0$ as $n \to \infty$ and, hence, in particular that

$$\sup_{0 \leq t \leq b} \xi_{\nu(nt)}/\sqrt{n} \overset{P}{\to} 0 \quad as \quad n \to \infty \quad for\ every\ b > 0, \tag{3.8}$$

which, together with (3.6) and Theorem A.3.8, implies that

$$Z_n' \overset{J_1}{\Longrightarrow} W \quad as \quad n \to \infty. \tag{3.9}$$

The conclusion now follows from the symmetry of W. $\qquad\qquad\square$

Before proceeding to the case where the variance is not necessarily finite we mention that the ladder variable method can be used to extend Theorem 3.1 for positive summands (Billingsley (1968, 1999)) to the case of general summands with positive mean, see Gut (1973), Section 3.

Next, we consider the case when the variance need not be finite, that is, when the limits are stable processes—see Theorem 3.5.2 for the one-dimensional result.

We thus assume that $\{\xi_k, k \geq 1\}$ are i.i.d. random variables with mean μ ($0 < \mu < \infty$). Set $S_n = \sum_{k=1}^{n} \xi_k$, $n \geq 0$, and suppose that $\{B_n, n \geq 1\}$ are positive normalizing coefficients such that

$$\frac{S_n - n\mu}{B_n} \xrightarrow{d} G_\alpha \quad \text{as} \quad n \to \infty, \tag{3.10}$$

where G_α is (a random variable distributed according to) a stable law with index α ($1 < \alpha \leq 2$) and let X be the stable process whose one-dimensional marginal distribution at $t = 1$ is G_α. Finally, set, for $t \geq 0$,

$$X_n(t, \omega) = \frac{S_{[nt]}(\omega) - [nt]\mu}{B_n} \quad \text{and} \quad Z_n(t, \omega) = \frac{\nu(nt, \omega) - nt/\mu}{B_n} \mu^{1+(1/\alpha)}.$$

Theorem 3.2.

(i) *Under the above assumptions we have*

$$Z_n \overset{M_1}{\Longrightarrow} -X \quad as \quad n \to \infty. \tag{3.11}$$

(ii) *If $1 < \alpha < 2$ and, moreover, X is a strictly asymmetric stable process without positive jumps, then*

$$Z_n \overset{J_1}{\Longrightarrow} -X \quad as \quad n \to \infty. \tag{3.12}$$

Proof. (ii) The proof of the second half of the theorem is very much like the proof of Theorem 3.1, so we begin with that part. We follow Gut (1975a). Set

$$Y_n(t, \omega) = \frac{1}{B_n} \sum_{k=1}^{\nu(nt,\omega)} (\xi_k(\omega) - \mu) \quad (t \geq 0). \tag{3.13}$$

It follows from Theorem 2.2 (cf. (3.5)) that

$$\mu^{1/\alpha} \cdot Y_n \overset{J_1}{\Longrightarrow} X \quad as \quad n \to \infty. \tag{3.14}$$

Set

$$Z'_n(t, \omega) = \mu^{1/\alpha} \cdot \frac{nt - \mu\nu(nt, \omega)}{B_n} \quad (t \geq 0). \tag{3.15}$$

By (3.14) and (3.3.2) we then have

$$Z_n' \overset{J_1}{\Longrightarrow} X \quad \text{as} \quad n \to \infty, \tag{3.16}$$

provided we can show (cf. (3.8)) that

$$\frac{1}{B_n} \cdot \sup_{0 \le t \le b} \xi_{\nu(nt)} \overset{p}{\to} 0 \quad \text{as} \quad n \to \infty \quad \text{for every } b > 0. \tag{3.17}$$

It suffices to prove (3.17) for $b = 1$. To this end we define the largest positive jump function, $J^+ \colon D \to R$, as follows;

$$J^+(x) = \sup_{0 \le t \le 1} (x(t) - x(t-)). \tag{3.18}$$

Lemma 3.1. J^+ is J_1-continuous on $D[0, 1]$.

Proof. Let $\{x_n, n \ge 1\}$ and x be elements of D, such that

$$x_n \to x(J_1) \quad \text{as} \quad n \to \infty. \tag{3.19}$$

This means (cf. Definition A.3.4) that there exists a sequence $\{\lambda_n, n \ge 1\}$ in $\Lambda = \{\lambda \colon \lambda$ is a strictly increasing, continuous mapping of $[0, 1]$ onto itself, $\lambda(0) = 0, \lambda(1) = 1\}$, such that

$$\sup_{0 \le t \le 1} |(x_n \circ \lambda_n)(t) - x(t)| \to 0 \quad \text{and} \quad \sup_{0 \le t \le 1} |\lambda_n(t) - t| \to 0 \quad \text{as} \quad n \to \infty. \tag{3.20}$$

Now, because of the uniform convergence of $x_n \circ \lambda_n$ to x it follows that

$$J^+(x_n \circ \lambda_n) \to J^+(x)(J_1) \quad \text{as} \quad n \to \infty. \tag{3.21}$$

However, since $J^+(x_n \circ \lambda_n) = J^+(x_n)$, we conclude that

$$J^+(x_n) \to J^+(x)(J_1) \quad \text{as} \quad n \to \infty, \tag{3.22}$$

which proves the lemma.   □

We now apply Lemma 3.1 and the continuous mapping theorem (Theorem A.3.5) to (3.14) restricted to $[0, 1]$. Since, by assumption,

$$J^+(X) = 0, \tag{3.23}$$

we have

$$J^+(Y_n) \overset{p}{\to} 0(J_1) \quad \text{as} \quad n \to \infty, \tag{3.24}$$

and (3.17) follows.

We have thus verified that (3.16) holds, from which (3.12) is immediate.

(i) Reviewing the proof of (ii) we find that the assumption that there were no positive jumps was crucial for (3.17). In the setting of (i) we would,

by the same procedure, arrive at $J^+(Y_n) \xrightarrow{P} J^+(X)$ as $n \to \infty$, where the convergence is in the J_1-topology (and, hence, also in the M_1-topology), but, since $J^+(X)$ is nonzero with positive probability we cannot proceed by that method (see also Remark 3.5 below).

Instead of studying the first passage times directly we proceed via the supremum. Define, for $n \geq 0$,

$$S_n^* = \max_{0 \leq k \leq n} S_k \quad \text{and} \quad X_n^*(t, \omega) = \frac{S_{[nt]}^*(\omega) - nt\mu}{B_n} \qquad (t \geq 0). \qquad (3.25)$$

Lemma 3.2. *Let $\{x_n, n \geq 1\}$ and x be elements of D and set $x_n^*(t) = \sup_{0 \leq s \leq t} x_n(s)$ $(t \geq 0)$. If $(x_n - c_n I) \to x(M_1)$ as $n \to \infty$ then*

$$(x_n^* - c_n I) \to x(M_1) \quad as \quad n \to \infty, \qquad (3.26)$$

where I is the identity $(I(t) = t)$ and $\{c_n, n \geq 1\}$ are constants tending to $+\infty$ as $n \to \infty$.

This is Theorem 6.3(ii) of Whitt (1980). In Whitt (1980), Theorem 6.2(ii) it is shown that the lemma can be strengthened to hold in the J_1-topology provided x has no negative jumps.

Let us now use Lemma 3.2 to conclude that

$$X_n^* \xRightarrow{M_1} X \quad as \quad n \to \infty. \qquad (3.27)$$

This is seen as follows. Since $X_n \xRightarrow{J_1} X$ as $n \to \infty$ and since the M_1-topology is weaker than the J_1-topology it follows that

$$X_n \xRightarrow{M_1} X \quad as \quad n \to \infty. \qquad (3.28)$$

Now, since B_n varies regularly with exponent α $(1 < \alpha \leq 2)$ it follows that $n\mu/B_n \to +\infty$ as $n \to \infty$. Thus, by applying Theorem A.3.7, Lemma 3.2 and arguments like those of Example A.3.1 to (3.28) we obtain (3.27).

Next we observe that $\{\nu(t), t \geq 0\}$ is also the first passage time process for the sequence of suprema, $\{S_n^*, n \geq 0\}$, but, since the latter is nondecreasing we can interpret the first passage time process and the supremum process as (generalized) inverses of each other. The following lemma states that inversion together with translation is an M_1-continuous functional.

Lemma 3.3. *Let $\{x_n, n \geq 1\}$ be random elements of D_0, which are unbounded above and such that $x_n(0) \geq 0$ and set $z_n(t) = \inf\{s: x_n > t\}$ $(t \geq 0)$. If $c_n(x_n - I) \to x(M_1)$ as $n \to \infty$ with $x(0) = 0$, then*

$$c_n(z_n - I) \to -x(M_1) \quad as \quad n \to \infty, \qquad (3.29)$$

where I and $\{c_n, n \geq 1\}$ are as in Lemma 3.2.

This is (essentially) Theorem 7.5 of Whitt (1980). In the case when x has no positive jumps the statement can be strengthened to a result relative to the J_1-topology (see Whitt (1980), Theorem 7.3). In that case we do, however, not assume that the elements are nondecreasing; as above, the proof uses composition, not inversion. (Theorem 7.3 of Whitt (1980) actually combines Lemmas 3.2 and 3.3.)

In order to apply Lemma 3.3 we first have to identify our process with the notation of the lemma. We first observe that

$$X_n^*(t, \omega) = \frac{n\mu}{B_n} \left(\frac{S_{[nt]}^*(\omega)}{n\mu} - t \right) \quad (t \geq 0). \tag{3.30}$$

Furthermore,

$$\frac{\nu(nt\mu)}{n} = \frac{1}{n} \cdot \inf\{u : S_{[u]} > nt\mu\} = \frac{1}{n} \cdot \inf\{u : S_{[u]}^* > nt\mu\}$$

$$= \inf\{u : S_{[nu]}^* > nt\mu\} = \inf \left\{ u : \frac{S_{[nu]}^*}{n\mu} > t \right\}.$$

Next, set

$$\tilde{Z}_n(t, \omega) = \frac{\nu(nt\mu, \omega) - nt}{B_n} \mu = \frac{n\mu}{B_n} \left(\frac{\nu(nt\mu, \omega)}{n} - t \right) \quad (t \geq 0), \tag{3.31}$$

and recall from above that $n\mu/B_n \to +\infty$ as $n \to \infty$.

By (3.27), Theorem A.3.7, Lemma 3.3 and the arguments of Example A.3.1 we now obtain

$$\tilde{Z}_n \overset{M_1}{\Longrightarrow} -X \quad \text{as} \quad n \to \infty. \tag{3.32}$$

We finally make a change of variable, $t \to t/\mu$. The conclusion then follows from the fact that $b \cdot X(t)$ and $X(b^\alpha t)$ have the same distribution (cf. the end of the proof of Theorem 2.2). $\qquad\square$

Theorem 3.2(i) is due to Bingham (1973), Theorem 1b, who also treats the case $0 < \alpha < 1$. Theorem 3.2(ii) is due to Gut (1975a). We also mention Iglehart and Whitt (1971), Vervaat (1972a,b) and Whitt (1980). If $\alpha = 2$ and $\text{Var}\xi_1 = \sigma^2 < \infty$ we rediscover Theorem 3.1.

Remark 3.1. Since the projections $\pi_{t_1,\ldots,t_k}$ from D to R^k are continuous at every x such that x is continuous at $t_1, \ldots, t_k$, the finite-dimensional distributions converge. In particular, Theorem 3.5.2 follows.

Remark 3.2. Note that (3.27) is a functional version of Theorem 4.4.4. We shall return to this result below, see Theorem 5.2.

Remark 3.3. The first passage time function is not J_1-continuous in general. It is, however, J_1-continuous at strictly increasing elements and M_1-continuous at elements, x, such that $x \not\equiv 0$ on $[0,\delta]$ for any $\delta > 0$. For increasing elements

the M_1-continuity follows from the symmetric role played by the coordinates in the parametric representation; for the general case we also use the continuity of the supremum function (cf. Whitt (1980), Section 7).

Remark 3.4. It follows immediately from Lemma 3.1 that if $\{x_n, n \geq 1\}$ are elements of D without positive jumps and such that $x_n \to x(J_1)$ as $n \to \infty$, where $x \in D$, then x has no positive jumps. Similarly for negative jumps.

Remark 3.5. A first consequence of the previous remark is that, since Z_n as defined in Theorem 3.2 has no negative jumps (i.e., $-Z_n$ has no positive jumps), we can obtain J_1-convergence in Theorem 3.2(ii) only when the limiting stable process has no positive jumps, that is, only for the case considered there.

Remark 3.6. Secondly, if $\{x_n, n \geq 1\}$ all belong to C and if $x_n \to x(J_1)$ as $n \to \infty$, then $x \in C$ and, hence, $x_n \to x(U)$ as $n \to \infty$.

Remark 3.7. A further consequence of Remark 3.4 is that if we assume that the conclusion of Lemma 3.2 is true relative to the J_1-topology, then x has no negative jumps; namely, since $x_n^* - c_n I$ has no negative jumps, neither has x. Similarly for Lemma 3.3. Also, if both lemmas hold relative to the J_1-topology, then $x \in C$ (see also Whitt (1980), Theorem 7.4).

5.4 A Stopped Two-Dimensional Random Walk

Let $\{(\xi_k, \eta_k), k \geq 1\}$ be i.i.d. two-dimensional random variables such that $\mu_\eta = E\eta_1$ exists and $0 < \mu_\xi = E\xi_1 < \infty$. Moreover, set $U_n = \sum_{k=1}^{n} \xi_k$ and $V_n = \sum_{k=1}^{n} \eta_k$, $n \geq 1$, and let all random variables with index 0 be equal to 0. Finally, $\tau(t) = \min\{n : U_n > t\}$ $(t \geq 0)$.

We first suppose that the variances σ_ξ^2 and σ_η^2 are finite. Define

$$Z_n(t, \omega) = \frac{V_{\tau(nt, \omega)}(\omega) - \frac{\mu_\eta}{\mu_\xi} nt}{\gamma \mu_\xi^{-3/2} \sqrt{n}} \quad (t \geq 0), \tag{4.1}$$

where $\gamma^2 = \mathrm{Var}(\mu_\xi \eta_1 - \mu_\eta \xi_1)$ is assumed to be positive.

Theorem 4.1. $Z_n \overset{J_1}{\Longrightarrow} W$ *as* $n \to \infty$.

The proof consists of modifying the proof of Theorem 3.1 in the same way as the proof of Theorem 4.2.3(i) was obtained by modifying the proof of Theorem 3.5.1. □

By making a corresponding modification of the proof of Theorem 3.2 we obtain the following result.

Theorem 4.2. *Suppose that* $\{B_n, n \geq 1\}$ *are positive normalizing coefficients such that*

$$\frac{\mu_\eta U_n - \mu_\xi V_n}{B_n} \xrightarrow{d} G_\alpha \quad as \quad n \to \infty \tag{4.2}$$

$$\frac{U_n - n\mu_\xi}{aB_n} \xrightarrow{d} G_\alpha \quad as \quad n \to \infty \quad for\ some \quad a > 0, \tag{4.3}$$

where G_α is (a random variable distributed according to) a stable law with index α, $1 < \alpha \leq 2$, and let X be a strictly asymmetric stable process without positive jumps, whose one-dimensional distribution at $t = 1$ is G_α. Set

$$Z_n(t, \omega) = \frac{V_{\tau(nt,\omega)}(\omega) - \frac{\mu_\eta}{\mu_\xi}nt}{B_n} \cdot \mu_\xi^{1+1/\alpha} \qquad (t \geq 0). \tag{4.4}$$

Then

$$Z_n \xRightarrow{J_1} -X \quad as \quad n \to \infty. \tag{4.5}$$

Theorem 4.1 is due to Gut and Janson (1983). We also mention Högfeldt (1977) for related results.

Theorems 4.2.3(i) and 4.2.5 are special cases of the above results and, as before, multidimensional versions follows. However, a result like Theorem 4.2.8 is, of course, not included in these multidimensional results. A weak convergence result including Theorem 4.2.8 as a special case is the following (see Gut and Janson (1983), for further details).

For $t \geq 0$ we define

$$X_n^U(t, \omega) = \frac{U_{[nt]}(\omega) - [nt]\mu_\xi}{\sqrt{n}}, \qquad X_n^V(t, \omega) = \frac{V_{[nt]}(\omega) - [nt]\mu_\eta}{\sqrt{n}},$$

$$Z_n^\tau(t, \omega) = \frac{\tau(nt, \omega) - nt/\mu_\xi}{\sqrt{n}}, \qquad Z_n^{V(t)}(t, \omega) = \frac{V_{\tau(nt,\omega)}(\omega) - nt\mu_\eta/\mu_\xi}{\sqrt{n}}.$$

Furthermore, let $\Lambda_{\xi,\eta}$ denote the covariance matrix of (ξ, η).

By our previous results each of the four processes converges to a Wiener process in $D[0, \infty)$. Also, by Donsker's theorem (X_n^U, X_n^V) converges to a two-dimensional Wiener process, whose covariance matrix at $t = 1$ equals $\Lambda_{\xi,\eta}$. However, the following is true (see Gut and Janson (1983), Theorem 5).

Theorem 4.3.

$$(X_n^U(\cdot/\mu_\xi), X_n^V(\cdot/\mu_\eta), Z_n^\tau(\cdot), Z_n^{V(\tau)}(\cdot))$$

$$\xRightarrow{J_1} \left(W_1, W_2, -\frac{1}{\mu_\xi}W_1, W_2 - \frac{\mu_\eta}{\mu_\xi}W_1\right) \quad as \quad n \to \infty,$$

where the limit process is a two-dimensional Wiener process in R^4. The covariance matrix of (W_1, W_2) at $t = 1$ equals $\mu_\xi^{-1}\Lambda_{\xi,\eta}$.

By defining

$$\nu(t) = \min\{n : V_n > t\} \quad \text{and} \quad Z_n^\nu(t,\omega) = \frac{\nu(nt,\omega) - nt/\mu_\eta}{\sqrt{n}} \quad (t \geq 0), \quad (4.6)$$

the following result (see Gut and Janson (1983), Theorem 6) can be obtained.

Theorem 4.4.

$$(\mu_\xi Z_n^\tau(\mu_\xi\cdot), \mu_\eta Z_n^\nu(\mu_\eta\cdot)) \overset{J_1}{\Longrightarrow} (W_1, W_2) \quad as \quad n \to \infty,$$

where (W_1, W_2) is a two-dimensional Wiener process, whose covariance matrix at $t = 1$ equals $\Lambda_{\xi,\eta}$. In particular, the stopping times $\tau(t)$ and $\nu(t)$ are jointly asymptotically normally distributed as $t \to \infty$.

5.5 The Maximum of a Random Walk with Positive Drift

This quantity has already been considered in the proof of Theorem 3.2(i). By using Lemma 4.4.1 and Theorem 4.1 (recall Remark 4.2.10) the following functional version of Theorem 4.4.3(i) is immediate.

Let $\{\xi_k, k \geq 1\}$ be i.i.d. random variables with mean μ $(0 < \mu < \infty)$ and variance σ^2 $(0 < \sigma^2 < \infty)$ and set

$$X_n(t,\omega) = \frac{S_{[nt]}(\omega) - [nt]\mu}{\sigma\sqrt{n}} \quad \text{and} \quad X_n^*(t,\omega) = \frac{S_{[nt]}^*(\omega) - nt\mu}{\sigma\sqrt{n}} \quad (t \geq 0),$$

$$(5.1)$$

where $S_n = \sum_{k=1}^n \xi_k$, $S_n^* = \max_{0\leq k\leq n} S_k$ and $S_0 = 0$.

Theorem 5.1. $X_n^* \overset{J_1}{\Longrightarrow} W$ as $n \to \infty$.

Remark 5.1. Note that we can subtract $[nt]\mu$ or $nt\mu$ in (5.1) without affecting the conclusion.

The functional version of Theorem 4.4.4 follows similarly.

Theorem 5.2.

(i) *Under the assumptions of Theorem 3.2(i) we have*

$$X_n^* \overset{M_1}{\Longrightarrow} X \quad as \quad n \to \infty. \qquad (5.2)$$

(ii) *If $1 < \alpha < 2$ and, moreover, X is a strictly asymmetric stable process without negative jumps, then*

$$X_n^* \overset{J_1}{\Longrightarrow} X \quad as \quad n \to \infty. \qquad (5.3)$$

Remark 5.2. Note that (5.2) is the same as (3.27), but that the proof is different.

Remark 5.3. By recalling Remark 3.4 (and switching X to $-X$) it follows that J_1-convergence can only hold in Theorem 5.2 for the case considered, in view of the fact that X_n^* has no negative jumps (cf. Remark 3.5).

5.6 First Passage Times Across General Boundaries

Next we consider extensions of Theorems 4.5.6 and 4.5.7 when $a(y) = y^\beta$ $(0 \leq \beta < 1)$. Let $\{\xi_k, k \geq 1\}$ be i.i.d. random variables with mean μ $(0 < \mu < \infty)$, set $S_n = \sum_{k=1}^{n} \xi_k$, $n \geq 0$, and define, for some β $(0 \leq \beta < 1)$

$$\nu(t) = \min\{n : S_n > t \cdot n^\beta\} \quad (t \geq 0). \tag{6.1}$$

The case $\beta = 0$ corresponds to the situation discussed in Section 5.3. We define $\lambda(t)$ as in Section 4.5, that is, we have $\lambda(t) = (t/\mu)^{1/(1-\beta)}$, $t \geq 0$, (since the slowly varying part is identically equal to one).

Suppose that $\mathrm{Var}\xi_1 = \sigma^2 < \infty$ and set

$$Z_n(t, \omega) = \frac{\nu(nt, \omega) - \lambda(nt)}{\sigma(\mu(1 - \beta))^{-1}\sqrt{\lambda(n)}} \quad (t \geq 0). \tag{6.2}$$

Theorem 6.1. $Z_n \overset{J_1}{\Longrightarrow} W'$ as $n \to \infty$, where $W'(t) = W(t^{1/(1-\beta)})$ $(t \geq 0)$.

Outline of Proof. We follow the pattern of the proof of Theorem 3.1. For simplicity we treat n as a continuous parameter.

Define the random time change

$$\Phi_n(t, \omega) = \frac{\nu((nt)^{1-\beta}, \omega)}{n} \quad (t \geq 0) \tag{6.3}$$

and set

$$Y_n(t, \omega) = \frac{1}{\sigma\sqrt{n}} \sum_{k=1}^{\nu((nt)^{1-\beta}, \omega)} (\xi_k(\omega) - \mu) \quad (t \geq 0). \tag{6.4}$$

By modifying the proof of Theorem 2.1 (recall Theorem 4.5.2) we obtain (cf. (3.5))

$$\mu^{1/2(1-\beta)} \cdot Y_n \overset{J_1}{\Longrightarrow} W \quad \text{as} \quad n \to \infty. \tag{6.5}$$

Next we define

$$Z_n^*(t, \omega) = \frac{(nt)^{1-\beta} \cdot \nu^\beta((nt)^{1-\beta}, \omega) - \mu \cdot \nu((nt)^{1-\beta}, \omega)}{\sigma\sqrt{n}} \quad (t \geq 0). \tag{6.6}$$

By continuing as in the last part of the proof of Theorem 3.1 it follows that

$$-\mu^{1/2(1-\beta)} Z_n^* \overset{J_1}{\Longrightarrow} W \quad \text{as} \quad n \to \infty. \tag{6.7}$$

For the case $\beta = 0$ this reduces to Theorem 3.1. For the case $0 < \beta < 1$ we first change the time scale, which yields

$$Z_n^{**} \overset{J_1}{\Longrightarrow} W' \quad \text{as} \quad n \to \infty, \tag{6.8}$$

where

$$Z_n^{**}(t,\omega) = \mu^{1/2(1-\beta)} \cdot \frac{\mu\nu(nt,\omega) - nt\nu^\beta(nt,\omega)}{\sigma n^{1/2(1-\beta)}}$$

$$= \frac{\nu(nt,\omega) - \lambda(nt)^{1-\beta}\nu^\beta(nt,\omega)}{\sigma\mu^{-1}\sqrt{\lambda(n)}} \quad (t \geq 0). \tag{6.9}$$

Finally, by using Taylor expansion (cf. Section 4.5) it is possible to show that

$$\sup_{0 \leq t \leq b} |Z_n^{**}(t) - Z_n(t)| \xrightarrow{P} 0 \quad \text{as} \quad n \to \infty \text{ for every } b > 0, \tag{6.10}$$

which, together with (6.8) and Theorem A.3.8, proves the conclusion. □

Remark 6.1. Theorem 6.1 is (for $0 < \beta < 1$) due to Gut (1973). Since the general idea of the proof is as before, but the technicalities are much more complicated we have preferred to give only an outline of the proof and we refer to the original paper for the details.

The following theorem, in which the variance need not be finite is due to Gut (1975a). Since the details get still more involved we confine ourselves to stating the theorem. Very briefly speaking, the proof consists of generalizing the proof of Theorem 3.2 along the same lines as in the proof of the previous result.

We thus assume that (3.10) holds, that $\{X_n, n \geq 1\}$ is given as in Theorem 3.2 and define

$$Z_n(t,\omega) = \frac{\nu(nt,\omega) - \lambda(nt)}{B(n^{1/(1-\beta)})} \cdot (1-\beta)\mu^{1+(1/\alpha(1-\beta))} \quad (t \geq 0), \tag{6.11}$$

where $B(y) = B_{[y]}$ for $y > 0$.

Theorem 6.2.

(i) *For $1 < \alpha \leq 2$ we have*

$$Z_n \xrightarrow{M_1} -X' \quad \text{as} \quad n \to \infty, \tag{6.12}$$

 where $X'(t) = X(t^{1/(1-\beta)})$, $t \geq 0$.

(ii) *If $1 < \alpha < 2$ and, moreover, X is a strictly asymmetric stable process without positive jumps, then*

$$Z_n \xrightarrow{J_1} -X' \quad \text{as} \quad n \to \infty. \tag{6.13}$$

For a generalization of Theorem 6.2 to more general processes we refer to Lindberger (1978), Theorem 2.

5.7 The Law of the Iterated Logarithm

We close this chapter by presenting some results related to Strassen's (1964) strong invariance principle. For preliminary material we refer to Section A.4 and the references given there.

Let $\{\xi_k, k \geq 1\}$ be i.i.d. random variables with mean 0 and variance 1 (for simplicity) and define, for $n \geq 3$,

$$X_n(t, \omega) = \frac{1}{\sqrt{2n \log \log n}} S_{[nt]}(\omega) \quad (t \geq 0), \tag{7.1}$$

where $S_n = \sum_{k=1}^{n} \xi_k, \ n \geq 0$.

Strassen's famous result asserts that, with probability 1, the sequence $\{X_n, n \geq 3\}$ is relatively compact in the uniform topology and that the set of its limit points coincides with K, where

$$K = \left\{ x \in AC[0, \infty) : x(0) = 0 \text{ and } \int_0^\infty (x'(t))^2 dt \leq 1 \right\} \tag{7.2}$$

(cf. Theorem A.4.1 and Remark A.4.2).

Stopped Random Walks

The first result below corresponds to Theorem 2.1, which was an Anscombe-version of Donsker's theorem. Thus, suppose, as in Section 5.2, that $\{N(t), t \geq 0\}$ is a nondecreasing family of positive, integer valued random variables such that

$$\frac{N(t)}{t} \xrightarrow{\text{a.s.}} \theta \ (0 < \theta < \infty) \quad \text{as} \quad t \to \infty. \tag{7.3}$$

Further, define, for $n \geq 3$,

$$Y_n(t, \omega) = \frac{1}{\sqrt{2n \log \log n}} S_{N(nt,\omega)}(\omega) \quad (t \geq 0). \tag{7.4}$$

Theorem 7.1. *The sequence $\{\theta^{-1/2} \cdot Y_n, n \geq 3\}$ is, with probability 1, relatively compact (in the uniform topology) and the set of its limit points coincides with K.*

This is a special case of Serfozo (1975), Corollary 2.3.

Just as in Section 5.2 we can also stop the random walk at $t \cdot N(n)$ ($0 \leq t \leq 1$) instead of at $N(nt)$ ($0 \leq t \leq 1$). We then suppose that (7.3) holds and define, for $n \geq 3$,

$$Y_n'(t, \omega) = \frac{1}{\sqrt{2N(n, \omega) \log^+ \log^+ N(n, \omega)}} S_{[t \cdot N(n,\omega)]}(\omega) \quad (0 \leq t \leq 1) \tag{7.5}$$

and

$$Y_n''(t, \omega) = \frac{1}{\sqrt{2\theta n \log \log n}} S_{[t \cdot N(n,\omega)]}(\omega) \quad (0 \leq t \leq 1). \tag{7.6}$$

Theorem 7.2. *The sequences $\{Y_n', n \geq 3\}$ and $\{Y_n'', n \geq 3\}$ are, with probability 1, relatively compact (in the uniform topology) and the set of their limit points coincides with K as defined in (A.4.4).*

For this result we refer to Huggins (1985) and Chang and Hsiung (1983), who treat (7.5) and (7.6) (with $\theta = 1$), respectively—see also Hall and Heyde (1980), Chapter 4. Note that it is easy to see that the set of limit points is *contained in K*. The difficulty is to prove *equality*.

First Passage Times for Random Walks with Positive Drift

Consider the setup of Chapter 3 and Section 5.3 of the present chapter, that is, let $\{\xi_k, k \geq 1\}$ be i.i.d. random variables with positive, finite mean μ and variance σ^2. Set, for $n \geq 3$,

$$X_n(t, \omega) = \frac{1}{\sqrt{2\sigma^2 n \log \log n}} (S_{[nt]}(\omega) - [nt]\mu) \quad (t \geq 0), \qquad (7.7)$$

let $\nu(t) = \min\{n : S_n > t\}$, $t \geq 0$, and define

$$Z_n(t, \omega) = \frac{\nu(nt, \omega) - nt/\mu}{\sqrt{2n\sigma^2 \mu^{-3} \log \log n}} \quad (t \geq 0). \qquad (7.8)$$

Theorem 7.3. *The sequence $\{Z_n, n \geq 3\}$ is, with probability 1, relatively compact and the set of its limit points coincides with K.*

Theorem 7.3 is due to Vervaat (1972a,b), who uses continuous mapping results. We also mention Horváth (1984a), where strong approximation results are used (under higher moment assumptions). This means that the theorem is first proved for Brownian motion, after which it is shown that the uniform distance between Z_n and the corresponding object for Brownian motion tends to 0 sufficiently rapidly.

For the case of first passage times across more general boundaries (cf. Sections 4.5 and 5.6) we refer to Horváth (1984b), where also more general stituations are investigated.

A Stopped Two-Dimensional Random Walk

We use the notations and assumptions of Section 5.4 above. Further, define, for $n \geq 3$,

$$Z_n(t, \omega) = \frac{V_{\tau(nt, \omega)}(\omega) - \frac{\mu_\eta}{\mu_\xi} nt}{\sqrt{2\gamma^2 \mu_\xi^{-3} n \log \log n}} \quad (t \geq 0). \qquad (7.9)$$

Theorem 7.4. *The sequence $\{Z_n, n \geq 3\}$ is, with probability 1, relatively compact and the set of its limit points coincides with K.*

For strong approximation theorems in the vector valued case (cf. (4.2.44)) we refer to Horváth (1984c, 1986).

The Maximum of a Random Walk with Positive Drift

Consider the random walk of Section 5.5, but set, for $n \geq 3$,

$$X_n^*(t, \omega) = \frac{S_{[nt]}^*(\omega) - nt\mu}{\sqrt{2\sigma^2 n \log \log n}} \quad (t \geq 0), \tag{7.10}$$

where $S_n^* = \max_{0 \leq k \leq n} S_k$.

Theorem 7.5. *The sequence* $\{X_n^*, n \geq 3\}$ *is, with probability 1, relatively compact and the set of its limit points coincides with* K.

This result follows immediately from the previous one (recall Remark 4.2.10), but it can also be deduced from Strassen's result, see Vervaat (1972a,b) (recall Remark 5.2).

5.8 Further Results

There exist more general functional limit theorems for the supremum functional and the inverse functional. Note, for example, that Lemmas 3.2 and 3.3 are statements about elements in $D[0, \infty)$ and $D_0[0, \infty)$ and that no i.i.d. assumptions are required in order to apply them. As a consequence one can prove that whenever a random process in $D[0, \infty)$, suitably normalized, converges weakly, say, then so does the supremum functional, properly normalized. Similarly, whenever a random process in $D_0[0, \infty)$ converges so does its inverse (again, with proper normalizations). This also implies that there is equivalence between convergence for a (normalized) increasing process and convergence for the (normalized) inverse process (or, equivalently, the first passage time process). By employing such arguments it follows that, for renewal processes, there is *equivalence* between Donsker's theorem and the conclusion of Theorem 3.1 (although, in this case, the converse might not seem very useful). For more on this we refer to Iglehart and Whitt (1971), Vervaat (1972a,b), Serfozo (1975), Lindberger (1978), Whitt (1980) and further references given there.

6

Perturbed Random Walks

Renewal theory can be generalized in various ways. The first generalization is to leave the assumption of nonnegativity of the summands. This topic was covered in Chapter 3 and onwards. The next one was considered in Section 4.5, namely "time dependent" boundaries. Since the appearance of the first edition of this book in 1988 research in the area has moved on. In this chapter we present some of the post-1988 development.

Recalling the model in Section 4.5, we note that the processes, that is, the random walks, are linear whereas the boundaries to be crossed are nonlinear. A generalization of this setting was initiated by Lai and Siegmund (1977, 1979), see also Woodroofe (1982) and Siegmund (1985), who introduced the concept of *nonlinear renewal theory*, where instead the processes are nonlinear and the boundaries are linear. This model will be briefly touched upon in Section 6.1. Since, as it turns out, the conditions in nonlinear renewal theory involve finite variance (as well as a number of technical assumptions), the remainder of the chapter will be devoted to a further extension—*perturbed random walks*—for which we present results corresponding to those presented earlier in this book. Basic sources for these are Gut (1992, 1997). The chapter closes with a brief outlook on further extensions and generalizations and some problems.

6.1 Introduction

We thus begin by introducing nonlinear renewal theory, after which we turn to perturbed random walks which are the main topic of the chapter.

The setup as described in the seminal papers Lai and Siegmund (1977, 1979) is the family of first passage times

$$\nu(t) = \min\{n : Z_n > t\}, \quad t \geq 0,$$

where

$$Z_n = S_n + \xi_n, \quad n \geq 1, \tag{1.1}$$

and where, in turn, $\{S_n,\ n \geq 1\}$ is a random walk and $\{\xi_n,\ n \geq 1\}$ a sequence of random variables that is

1. *slowly changing*, i.e.,

$$n^{-1} \max_{1 \leq k \leq n} |\xi_k| \xrightarrow{P} 0 \quad \text{as} \quad n \to \infty; \tag{1.2}$$

2. uniformly continuous in probability, viz., for every $\varepsilon > 0$ there exists $\delta > 0$, such that

$$P(\max_{1 \leq k \leq n\delta} |\xi_{n+k} - \xi_n| > \varepsilon) < \varepsilon \quad \text{for all} \quad n \geq n_0 \quad \text{some} \quad n_0; \tag{1.3}$$

3. component-wise independent of the future of the random walk, viz.,

$$\xi_n \quad \text{is independent of} \quad \{X_k,\ k > n\} \quad \text{for all } n. \tag{1.4}$$

Remark 1.1. Condition (1.3) is equivalent to the Anscombe condition (A) in Section 1.3.

The additional assumptions and conditions in these general results are rather technical and we refer to the original sources for details.

As for an important special case, suppose that $Y_1, Y_2, \ldots$ are i.i.d. random variables with positive finite mean, θ, and finite variance, τ^2, and that g is a positive function, that is twice continuously differentiable in some neighborhood of θ. Finally, set

$$Z_n = n \cdot g(\bar{Y}_n), \quad n \geq 1, \tag{1.5}$$

where $\bar{Y}_n = \frac{1}{n} \sum_{k=1}^{n} Y_k$, $n \geq 1$, and consider the family of first passage times

$$\nu(t) = \min\{n : n \cdot g(\bar{Y}_n) > t\}, \quad t \geq 0.$$

Although this case is less general it covers many important applications—some examples will be given in Section 6.9. On the other hand, the conditions involved in these examples are fewer and simpler.

To see that $\{Z_n,\ n \geq 1\}$ as just defined fulfills the requirements of a nonlinear renewal process one makes a Taylor expansion of g at θ. This yields

$$Z_n = n \cdot g(\theta) + n \cdot g'(\theta)(\bar{Y}_n - \theta) + n \cdot \frac{g''(\rho_n)}{2}(\bar{Y}_n - \theta)^2, \tag{1.6}$$

where $\rho_n = \rho_n(\omega)$ lies between θ and $\bar{Y}_n$.

Next, set $X_k = g(\theta) + g'(\theta)(Y_k - \theta)$, $k \geq 1$. Then $X_k,\ k \geq 1$, are i.i.d. random variables with mean $\mu = g(\theta) + g'(\theta) \cdot 0 = g(\theta) > 0$ and variance $\sigma^2 = \tau^2(g'(\theta))^2$. The random walk component, $\{S_n,\ n \geq 1\}$, is given by

$$S_n = \sum_{k=1}^{n} X_k = \sum_{k=1}^{n} \big(g(\theta) + g'(\theta)(Y_k - \theta)\big), \quad n \geq 1.$$

Furthermore, in view of the continuity of g'' and the strong law of large numbers,

$$\frac{1}{n} \cdot \frac{n \cdot g''(\rho_n)}{2}(\bar{Y}_n - \theta)^2 \xrightarrow{\text{a.s.}} 0 \quad \text{as} \quad n \to \infty, \tag{1.7}$$

that is, with

$$\xi_n = \frac{ng''(\rho_n)}{2}(\bar{Y}_n - \theta)^2,$$

the sequence $\{\xi_n, n \geq 1\}$ satisfies (1.2). The verification of (1.3) amounts to verifying Anscombe's condition (recall Remark 1.1). For details, see Woodroofe (1982), page 42, Siegmund (1985), page 190. Since, in the following, we shall focus on *perturbed random walks* we refer to the bibliography for additional references related to nonlinear renewal theory.

In the general case (1.2) implies, roughly speaking, that $\xi_n = o(n)$ as $n \to \infty$. Now, the law of large numbers implies that $S_n = \mathcal{O}(n)$ as $n \to \infty$, which means that Z_n as defined in (1.1) is of the form

$$Z_n = \mathcal{O}(n) + o(n) \quad \text{as} \quad n \to \infty, \tag{1.8}$$

which, in turn, suggests the following natural definition.

Definition 1.1. *A process $\{Z_n, n \geq 1\}$, such that*

$$Z_n = S_n + \xi_n, \quad n \geq 1,$$

is called a perturbed random walk *if $\{S_n, n \geq 1\}$ is a random walk, whose increments have positive finite mean, and $\{\xi_n, n \geq 1\}$ is a sequence of random variables, such that,*

$$\frac{\xi_n}{n} \xrightarrow{\text{a.s.}} 0 \quad \text{as} \quad n \to \infty. \tag{1.9}$$

Remark 1.2. Note that we neither assume that the elements of the perturbing process are independent of the future of the random walk (1.4) nor that the Anscombe condition (1.3) is satisfied in the definition.

Remark 1.3. For the special case (1.5) it is necessary that $\operatorname{Var} Y_1 < \infty$ for $\{\xi_n, n \geq 1\}$ to be slowly changing (in order to satisfy the Anscombe condition). For, say, the strong law of large numbers to hold this is, obviously, a superfluous assumption and, as we shall see below, it will not be needed in order to prove strong laws for the first passage times.

Remark 1.4. Introducing the increments of the perturbing process

$$\eta_1 = \xi_1, \quad \eta_n = \xi_n - \xi_{n-1}, \quad n \geq 2, \tag{1.10}$$

we mention for later use that, since

$$|\eta_1| = |\xi_1| \quad \text{and} \quad |\eta_n| \leq |\xi_n| + |\xi_{n-1}|, \quad n \geq 2, \tag{1.11}$$

it follows from (1.9), in particular, that

$$\frac{\eta_n}{n} \xrightarrow{\text{a.s.}} 0 \quad \text{as} \quad n \to \infty.$$

Our aim is to study first passage times of perturbed random walks across "time dependent" boundaries, that is the family $\{\nu(t),\, t \geq 0\}$ defined by

$$\nu(t) = \min\{n : Z_n > t \cdot a(n)\},$$

where $\{Z_n,\, n \geq 1\}$ is a perturbed random walk as described in Definition 1.1 and $a(y)$, for $y \geq A$ for some $A > 0$, is nondecreasing, concave and regularly varying at infinity with exponent β, $0 \leq \beta < 1$; cf. Section 4.5.

The plan of this chapter is as follows. In Section 6.2 we present Kolmogorov and Marcinkiewicz–Zygmund strong laws, a central limit theorem, a result on convergence to stable laws and a law of the iterated logarithm for the first passage times. Results for stopped processes and the overshoot are also given. The theorems are of the kind that if the perturbing process is sufficiently nice, the corresponding results are the same as those for random walks. The proofs follow closely the method of stopped random walks described earlier in this book; cf. also Gut (1974a).

Earlier papers in the area are Chow, Hsiung and Yu (1980, 1983), who consider a special form of perturbation and Horváth (1985), who proves strong approximation results.

In order to make the results more concrete we study, in Section 6.3, the special case when $\{Z_n,\, n \geq 1\}$ is of the form (1.5). However, since we only require that the perturbing process satisfies (1.9), we may use less restrictive conditions than in earlier work in the area.

In Section 6.4 we present an example of a convergence rate result. Sections 6.5 and 6.6 are devoted to results on finiteness of moments in the general case and the special case, respectively, and in Sections 6.7 and 6.8 we consider uniform integrability and moment convergence. Here, however, we must, in addition, assume that, for each n, ξ_n is independent of $\{X_k,\, k > n\}$, since we need to invoke results on moments and uniform integrability of stopped random walks, for which it is essential that $\{\nu(t),\, t \geq 0\}$, are *stopping times*.

In Section 6.9 we provide some examples related to sequential analysis and repeated significance tests, after which, in Sections 6.10 and 6.11, we extend the results from Section 4.2 (Gut and Janson (1983)) to perturbed random walks, which, in turn, are followed, in Section 6.12, by an application to repeated significance tests in two-parameter exponential families. In particular, we mention an example involving the normal distribution where some interesting relations between marginal one-parameter tests and joint tests are briefly hinted at. The basic sources here are Gut and Schwabe (1996, 1999).

The chapter closes with an outlook on further extensions and results and some problems.

6.2 Limit Theorems; the General Case

Let $X_1, X_2, \ldots$ be i.i.d. random variables with positive, finite mean μ and partial sums $\{S_n,\, n \geq 1\}$, let $\{\xi_n,\, n \geq 1\}$ be a sequence of random variables,

such that

$$\frac{\xi_n}{n} \xrightarrow{\text{a.s.}} 0 \quad \text{as} \quad n \to \infty$$

and let $\{\eta_k,\ k \geq 1\}$ denote the increments of $\{\xi_n,\ n \geq 1\}$. The present section is devoted to the study of perturbed random walks $\{Z_n,\ n \geq 1\}$, defined by

$$Z_n = S_n + \xi_n, \quad n \geq 1,$$

and the associated first passage time processes $\{\nu(t),\ t \geq 0\}$, defined by

$$\nu(t) = \min\{n : Z_n > t \cdot a(n)\},$$

where $a(y)$, for $y \geq A$ for some $A > 0$, is nondecreasing, concave and regularly varying at infinity with exponent β, $0 \leq \beta < 1$. The typical cases to have in mind are $a(y) = y^\beta$ and $a(y) = y^\beta \cdot \log y$. The general case is $a(y) = y^\beta \cdot L(y)$, where $L(y)$ is slowly varying at infinity. We recall that necessary prerequisites and further references to the literature on regular variation can be found in Appendix B.

One final piece of notation is needed. Following Siegmund (1967), cf. also Section 4.5, we define $\lambda(t)$ to be the solution of the equation $t \cdot a(y) = \mu y$, that is,

$$t \cdot a(\lambda(t)) = \mu \cdot \lambda(t).$$

If t is large the solution is unique. Moreover, $\lambda(t) \to \infty$ as $t \to \infty$. If $a(y) = y^\beta$ for some $\beta \in [0, 1)$, then $\lambda(t) = (t/\mu)^{1/(1-\beta)}$. In particular, if $a(y) \equiv 1$, then $\lambda(t) = t/\mu$.

Theorem 2.1.

$$\frac{\nu(t)}{\lambda(t)} \xrightarrow{\text{a.s.}} 1 \quad as \quad t \to \infty.$$

Theorem 2.2. *Suppose that* $E|X_1|^r < \infty$ *for some* r, $1 \leq r < 2$. *If*

$$\frac{\xi_n}{n^{1/r}} \xrightarrow{\text{a.s.}} 0 \quad as \quad n \to \infty,$$

then

$$\frac{\nu(t) - \lambda(t)}{(\lambda(t))^{1/r}} \xrightarrow{\text{a.s.}} 0 \quad as \quad t \to \infty.$$

Theorem 2.3. *Suppose that* $\sigma^2 = \mathrm{Var}\, X_1 < \infty$. *If*

$$\frac{\xi_n}{\sqrt{n}} \xrightarrow{\text{a.s.}} 0 \quad as \quad n \to \infty, \tag{2.1}$$

or if

$$\frac{\xi_n}{\sqrt{n}} \xrightarrow{P} 0 \quad as \quad n \to \infty \quad and \tag{2.2}$$

$$\left\{ \frac{\xi_n}{\sqrt{n}},\ n \geq 1 \right\} \quad satisfies\ Anscombe's\ condition, \tag{2.3}$$

then

$$\frac{\nu(t) - \lambda(t)}{\frac{\sigma}{(1-\beta)\mu}\sqrt{\lambda(t)}} \xrightarrow{d} N(0,1) \quad as \quad t \to \infty.$$

Theorem 2.4. *Suppose that* $\{B_n, \ n \geq 1\}$ *are normalizing constants, such that*

$$\frac{S_n - n\mu}{B_n} \xrightarrow{d} G_\alpha \quad as \quad n \to \infty,$$

where G_α *is (a random variable distributed according to) a stable law with index* α $(1 < \alpha \leq 2)$, *and let* $B(y) = B_{[y]}$ *for* $y > 0$. *If*

$$\frac{\xi_n}{B_n} \xrightarrow{p} 0 \quad as \quad n \to \infty$$

and $\{B(n)^{-1} \cdot \xi_n, \ n \geq 1\}$ *satisfies Anscombe's condition, then*

$$\frac{\nu(t) - \lambda(t)}{\frac{1}{(1-\beta)\mu}B(\lambda(t))} \xrightarrow{d} -G_\alpha \quad as \quad t \to \infty,$$

or, equivalently,

$$\frac{\nu(t) - \lambda(t)}{(1-\beta)^{-1}\mu^{-(1+1/(\alpha(1-\beta)))}B(t^{1/(1-\beta)})} \xrightarrow{d} -G_\alpha \quad as \quad t \to \infty.$$

Recall that $C(\{x_n\})$ denotes the cluster set of the sequence $\{x_n\}$.

Theorem 2.5. *Suppose that* $\sigma^2 = \operatorname{Var} X_1 < \infty$. *If*

$$\frac{\xi_n}{\sqrt{n \log \log n}} \xrightarrow{a.s.} 0 \quad as \quad n \to \infty,$$

then

$$C\left(\left\{\frac{\nu(t) - \lambda(t)}{\frac{\sigma}{(1-\beta)\mu}\sqrt{2\lambda(t)\log\log t}}, \ t \geq 3\right\}\right) = [-1, 1] \quad a.s.$$

In particular,

$$\limsup_{t\to\infty}(\liminf_{t\to\infty})\frac{\nu(t) - \lambda(t)}{\sqrt{2\lambda(t)\log\log t}} = \frac{+}{(-)}\frac{\sigma}{(1-\beta)\mu} \quad a.s.$$

Proof. The proofs proceed very much like those for random walks; cf. Section 4.5 (or Gut (1974a), Section 3, Gut (1985)). We therefore confine ourselves to giving an outline of the proof of Theorem 2.3.

By Anscombe's theorem, (Theorem 1.3.1), we have

$$\frac{S_{\nu(t)} - \nu(t)\mu}{\sigma\sqrt{\nu(t)}} \xrightarrow{d} N(0,1) \quad as \quad t \to \infty.$$

Now, if (2.1) holds, then this and the fact that $\nu(t) \xrightarrow{\text{a.s.}} \infty$ as $t \to \infty$, together imply (Theorem 1.2.1) that

$$\frac{\xi_{\nu(t)}}{\sqrt{\nu(t)}} \xrightarrow{\text{a.s.}} 0 \quad \text{as} \quad t \to \infty,$$

and if (2.2) and (2.3) hold, an application of Anscombe's theorem yields

$$\frac{\xi_{\nu(t)}}{\sqrt{\nu(t)}} \xrightarrow{p} 0 \quad \text{as} \quad t \to \infty.$$

In either case it thus follows that

$$\frac{Z_{\nu(t)} - \nu(t)\mu}{\sigma\sqrt{\nu(t)}} \xrightarrow{d} N(0,1) \quad \text{as} \quad t \to \infty.$$

Next we observe that

$$ta(\nu(t)) < Z_{\nu(t)} \leq ta(\nu(t)) + X_{\nu(t)} + \eta_{\nu(t)} \leq ta(\nu(t)) + X_{\nu(t)}^+ + \eta_{\nu(t)}^+. \quad (2.4)$$

Now, by Theorem 1.2.3, we have

$$\frac{X_{\nu(t)}^+}{\sqrt{\nu(t)}} \xrightarrow{\text{a.s.}} 0 \quad \text{as} \quad t \to \infty.$$

Furthermore, in view of (1.11) it follows that assumptions (2.1) and (2.2), respectively, also hold for $\{\eta_k, \ k \geq 1\}$ and, hence, that

$$\frac{\eta_{\nu(t)}^+}{\sqrt{\nu(t)}} \xrightarrow{\text{a.s.}} 0 \quad \text{as} \quad t \to \infty$$

or

$$\frac{\eta_{\nu(t)}^+}{\sqrt{\nu(t)}} \xrightarrow{p} 0 \quad \text{as} \quad t \to \infty,$$

respectively. The above facts, together with an application of Theorem 2.1, now yield

$$\frac{t \cdot a(\nu(t)) - \nu(t)\mu}{\sigma\sqrt{\lambda(t)}} \xrightarrow{d} N(0,1) \quad \text{as} \quad t \to \infty.$$

To finish off we perform a Taylor expansion of $a(y)$ at the point $\lambda(t)$, which yields

$$\nu(t)\mu - t \cdot a(\nu(t)) = (\nu(t) - \lambda(t))\mu(1 - \beta) \cdot \gamma_{\nu(t)},$$

where $\gamma_{\nu(t)}$ is some random variable which tends to 1 a.s. as $t \to \infty$; for details, see Lemma 4.5.1. This part of the proof is, in fact, identical to the corresponding one for random walks, since it is purely a matter of the asymptotic behavior of the function $a(y)$ and the fact that Theorem 2.1 holds. The desired conclusion follows. $\qquad\square$

In the same sense as Chapter 5 contains functional central limit theorems related to the results in earlier chapters, one can prove such results for perturbed random walks. The following one, which extends Theorem 5.3.1, is due to Larsson-Cohn (2000a). We refer to the original paper for details.

Theorem 2.6. *Let $a(y) \equiv 1$. Suppose that $\sigma^2 = \operatorname{Var} X_1 < \infty$, and set*

$$Z_n(t, \omega) = \frac{\nu(nt, \omega) - nt/\mu}{\sigma \mu^{-3/2}\sqrt{n}} \quad (t \geq 0).$$

If

$$\frac{\max_{1 \leq k \leq n} \xi_k}{\sqrt{n}} \xrightarrow{P} 0 \quad as \quad n \to \infty,$$

then

$$Z_n \overset{J_1}{\Longrightarrow} W \quad as \quad n \to \infty.$$

It is now easy to obtain the following results for the stopped perturbed random walk, $\{Z_{\nu(t)}, t \geq 0\}$, and the family of overshoots, $\{R(t), t \geq 0\}$, defined by

$$R(t) = Z_{\nu(t)} - a(\nu(t)).$$

Note (recall (2.4)) that

$$0 \leq R(t) \leq X_{\nu(t)} + \eta_{\nu(t)} \leq X_{\nu(t)}^+ + \eta_{\nu(t)}^+. \tag{2.5}$$

Theorem 2.7.
$$\frac{Z_{\nu(t)}}{\lambda(t)} \xrightarrow{\text{a.s.}} \mu \quad as \quad t \to \infty.$$

Theorem 2.8. *Let $r \geq 1$. If $E(X_1^+)^r < \infty$ and if*

$$\frac{\eta_n^+}{n^{1/r}} \xrightarrow{\text{a.s.}} 0 \quad as \quad n \to \infty,$$

then

$$\frac{R(t)}{(\lambda(t))^{1/r}} \xrightarrow{\text{a.s.}} 0 \quad as \quad t \to \infty.$$

Remark 2.1. For nonlinear renewal processes (recall (1.1)–(1.2)), the limiting distribution of $R(t)$ as $t \to \infty$ (without normalization) was obtained by Lai and Siegmund (1977), Theorem 1, see also Woodroofe (1982), Theorem 4.1.

In the following section we apply our results to the case $Z_n = n \cdot g(\bar{Y}_n)$, $n \geq 1$, mentioned in the introduction, under suitable conditions on the function g.

6.3 Limit Theorems; the Case $Z_n = n \cdot g(\bar{Y}_n)$

Let $Y_1, Y_2, \ldots$ be i.i.d. random variables with positive, finite mean θ and set $\bar{Y}_n = \frac{1}{n} \sum_{k=1}^{n} Y_k$, $n \geq 1$. In this section we consider the particular case

$$Z_n = n \cdot g(\bar{Y}_n), \quad n \geq 1,$$

of a perturbed random walk, where g belongs to the following class $\mathcal{G}$ of functions.

Definition 3.1. *The function* $g \in \mathcal{G}$ *if* g *is positive and continuous in some neighborhood of* θ *and nonnegative elsewhere on* $\mathbb{R}$.

Our concern is the family of first passage times, $\{\nu(t), \ t \geq 0\}$, defined by

$$\nu(t) = \min\{n : n \cdot g(\bar{Y}_n) > t \cdot a(n)\},$$

where $a(y)$, $y \geq 0$, is as before. In the definition of $\lambda(t)$, however, μ is replaced by $g(\theta)$, that is, $\lambda(t)$ is the solution of the equation $t \cdot a(y) = g(\theta)y$, viz.

$$t \cdot a(\lambda(t)) = g(\theta) \cdot \lambda(t).$$

In particular, if $a(y) = y^\beta$ for some β $(0 \leq \beta < 1)$, then $\lambda(t) = (t/g(\theta))^{1/(1-\beta)}$. For $\beta = 0$, i.e., $a(y) \equiv 1$, and $g(y) = y^+$ the relation reduces to the familiar $\lambda(t) = t/\theta$.

We remind the reader that, although this case is less general it covers many important applications, some of which will be given in Section 6.9.

The computations leading to (1.6) and (1.7) (which proved that the special class of processes is a subclass of the more general class) amounted to Taylor expansion, which produced a random walk component and a perturbing component. The proofs of the theorems below amount to doing the same and to checking that the perturbing process satisfies the conditions of the theorems. Recall, however, that our assumptions are weaker than those of the traditional setup.

As a final preparation we need the following

Definition 3.2. *The function* g *is Lipschitz continuous with exponent* α *at the point* c *if, for some constant* $C < \infty$,

$$|g(x) - g(c)| \leq C \cdot |x - c|^\alpha \quad \text{in some neighborhood of} \quad c.$$

Notation: $g \in \text{Lip}\,(\alpha, c)$.

In the theorems to follow we first state the general result, after which the corresponding result is given for the case $a(y) = y^\beta$ for some β $(0 \leq \beta < 1)$.

Theorem 3.1.

$$\frac{\nu(t)}{\lambda(t)} \xrightarrow{\text{a.s.}} 1 \quad \text{as} \quad t \to \infty.$$

In particular, if $a(y) = y^\beta$ for some β $(0 \le \beta < 1)$, then

$$\frac{\nu(t)}{t^{1/(1-\beta)}} \xrightarrow{a.s.} \frac{1}{(g(\theta))^{1/(1-\beta)}} \quad as \quad t \to \infty.$$

Theorem 3.2. *Suppose that $E|Y_1|^r < \infty$ for some r, $1 \le r < 2$. If $g \in$ Lip $(1, \theta)$, then*

$$\frac{\nu(t) - \lambda(t)}{(\lambda(t))^{1/r}} \xrightarrow{a.s.} 0 \quad as \quad t \to \infty.$$

In particular, if $a(y) = y^\beta$ for some β $(0 \le \beta < 1)$, then

$$\frac{\nu(t) - (t/g(\theta))^{1/(1-\beta)}}{t^{1/r(1-\beta)}} \xrightarrow{a.s.} 0 \quad as \quad t \to \infty.$$

Theorem 3.3. *Suppose that $\tau^2 = \operatorname{Var} Y_1 < \infty$ and that $g'(\theta) \ne 0$. If g' is continuous at θ, then*

$$\frac{\nu(t) - \lambda(t)}{\frac{\tau g'(\theta)}{(1-\beta)g(\theta)}\sqrt{\lambda(t)}} \xrightarrow{d} N(0, 1) \quad as \quad t \to \infty.$$

In particular, if $a(y) = y^\beta$ for some β $(0 \le \beta < 1)$, then

$$\frac{\nu(t) - (t/g(\theta))^{1/(1-\beta)}}{t^{1/2(1-\beta)}} \xrightarrow{d} N(0, (1-\beta)^{-2}(g(\theta))^{-\frac{3-2\beta}{1-\beta}}\tau^2(g'(\theta))^2) \quad as \quad t \to \infty.$$

Theorem 3.4. *Suppose that $\{B_n, n \ge 1\}$ are normalizing constants, such that*

$$\frac{n(\bar{Y}_n - \theta)}{B_n} \xrightarrow{d} G_\alpha \quad as \quad n \to \infty,$$

where G_α is (a random variable distributed according to) a stable law with index α $(1 < \alpha \le 2)$, and let $B(y) = B_{[y]}$ for $y > 0$. Suppose, further, that $g'(\theta) \ne 0$. If g' is continuous at θ, then

$$\frac{\nu(t) - \lambda(t)}{\frac{1}{(1-\beta)g(\theta)}B(\lambda(t))} \xrightarrow{d} -G_\alpha \quad as \quad t \to \infty.$$

In particular, if $a(y) = y^\beta$ for some β $(0 \le \beta < 1)$, then

$$\frac{\nu(t) - (t/g(\theta))^{1/(1-\beta)}}{(1-\beta)^{-1}(g(\theta))^{-(1+1/(\alpha(1-\beta)))}B(t^{1/(1-\beta)})} \xrightarrow{d} -G_\alpha \quad as \quad t \to \infty.$$

Theorem 3.5. *Suppose that $\tau^2 = \operatorname{Var} Y_1 < \infty$ and that $g'(\theta) \ne 0$. If g' is continuous at θ, then*

$$C\left(\left\{\frac{\nu(t) - \lambda(t)}{\frac{\tau g'(\theta)}{(1-\beta)g(\theta)}\sqrt{2\lambda(t)\log\log t}}, t \ge 3\right\}\right) = [-1, 1] \quad a.s.$$

and

$$\limsup_{t\to\infty}(\liminf_{t\to\infty})\frac{\nu(t) - \lambda(t)}{\sqrt{2\lambda(t)\log\log t}} = \begin{matrix} + \\ (-) \end{matrix} \frac{\tau g'(\theta)}{(1-\beta)g(\theta)} \quad a.s.$$

In particular, if $a(y) = y^\beta$ for some β $(0 \le \beta < 1)$, then

$$C\left(\left\{\frac{\nu(t) - (t/g(\theta))^{1/(1-\beta)}}{\sqrt{2t^{1/(1-\beta)}\log\log t}}, t \ge 3\right\}\right)$$

$$= \left[-\frac{\tau g'(\theta)}{(1-\beta)(g(\theta))^{\frac{3-2\beta}{2(1-\beta)}}}, \frac{\tau g'(\theta)}{(1-\beta)(g(\theta))^{\frac{3-2\beta}{2(1-\beta)}}}\right] \quad a.s.$$

and

$$\limsup_{t\to\infty}(\liminf_{t\to\infty})\frac{\nu(t) - (t/g(\theta))^{1/(1-\beta)}}{\sqrt{2t^{1/(1-\beta)}\log\log t}} = \begin{matrix} + \\ (-) \end{matrix} \frac{\tau g'(\theta)}{(1-\beta)(g(\theta))^{\frac{3-2\beta}{2(1-\beta)}}} \quad a.s.$$

Theorem 3.6.

$$\frac{Z_{\nu(t)}}{\lambda(t)} \xrightarrow{a.s.} g(\theta) \quad as \quad t \to \infty.$$

In particular, if $a(y) = y^\beta$ for some β $(0 \le \beta < 1)$, then

$$\frac{Z_{\nu(t)}}{t^{1/(1-\beta)}} \xrightarrow{a.s.} \frac{1}{(g(\theta))^{\beta/(1-\beta)}} \quad as \quad t \to \infty.$$

Theorem 3.7. *Suppose that $E|Y_1|^r < \infty$ for some $r \ge 1$. Suppose, in addition, that $g \in \text{Lip}(1, \theta)$ when $1 < r < 2$ and that $g' \in \text{Lip}(\alpha, \theta)$ for some $\alpha > 1 - \frac{2}{r}$ when $r \ge 2$. Then*

$$\frac{R(t)}{(\lambda(t))^{1/r}} \xrightarrow{a.s.} 0 \quad as \quad t \to \infty.$$

In particular, if $a(y) = y^\beta$ for some β $(0 \le \beta < 1)$, then

$$\frac{R(t)}{t^{1/r(1-\beta)}} \xrightarrow{a.s.} 0 \quad as \quad t \to \infty.$$

Remark 3.1. Theorems 3.1 and 3.3 have been proved in the sources cited above (see, in particular, Woodroofe (1982) and Siegmund (1985)) under the additional assumptions that $a(y) \equiv 1$ and that g'' is continuous (at θ).

Remark 3.2. For $g(y) = y^+$, we rediscover results from Section 4.5 (where further references can be found) on first passage times for random walks across curved boundaries. If, in addition, $a(y) \equiv 1$ we are back to renewal theory for random walks, i.e., Chapter 3.

As mentioned above, the proofs amount to defining the relevant perturbed random walk, that is, to show that the process under consideration actually is a perturbed random walk, and to checking the relevant conditions in the corresponding theorems of Section 6.2.

The following lemmas provide the necessary tools.

Lemma 3.1. *Let $\{U_n, n \geq 1\}$ be a sequence of random variables, such that $U_n \xrightarrow{\text{a.s.}} c$ as $n \to \infty$ for some constant, c, and suppose that the function $h \in$ Lip (α, c), $\alpha > 0$. Then*

$$h(U_n) \xrightarrow{\text{a.s.}} h(c) \quad as \quad n \to \infty.$$

Proof. The assumptions imply that, for almost all ω we have, for n sufficiently large, (suppressing ω) that

$$|h(U_n) - h(c)| \leq C \cdot |U_n - c|^\alpha \to 0 \quad as \quad n \to \infty. \qquad \square$$

Lemma 3.2. *Suppose that g is continuous at θ. Set, for $n \geq 1$,*

$$S_n = ng(\theta) \quad and \quad \xi_n = n \cdot (g(\bar{Y}_n) - g(\theta)),$$

and $Z_n = S_n + \xi_n$. Then

$$\frac{\xi_n}{n} \xrightarrow{\text{a.s.}} 0 \quad as \quad n \to \infty,$$

that is, $\{Z_n, n \geq 1\}$ is a perturbed random walk.

Proof. Immediate. $\qquad \square$

Lemma 3.3. *Suppose, in addition to the above, that $g \in$ Lip $(1, \theta)$ and that $E|Y_1|^r < \infty$ for some $r \in [1, 2)$. Then*

$$\frac{\xi_n}{n^{1/r}} \xrightarrow{\text{a.s.}} 0 \quad as \quad n \to \infty.$$

Proof. Apply Lemma 3.1 and the Marcinkiewicz–Zygmund strong law of large numbers (Gut (2007), Theorem 6.7.1). $\qquad \square$

Lemma 3.4. *Suppose that g' is continuous at θ and that $\tau^2 = \text{Var } Y_1 < \infty$. Set, for $n \geq 1$,*

$$S_n = ng(\theta) + ng'(\theta)(\bar{Y}_n - \theta) \quad and \quad \xi_n = n(g'(\rho_n) - g'(\theta))(\bar{Y}_n - \theta),$$

where $|\rho_n - \theta| \leq |\bar{Y}_n - \theta|$, and $Z_n = S_n + \xi_n$. Then $\{Z_n, n \geq 1\}$ is a perturbed random walk. The increments of the random walk component $\{S_n, n \geq 1\}$ are the i.i.d. random variables $\{g(\theta) + g'(\theta)(Y_k - \theta), k \geq 1\}$, which have positive, finite mean $\mu = g(\theta)$ and positive, finite variance $\sigma^2 = \tau^2 g'(\theta)^2$. Furthermore,

$$\frac{\xi_n}{\sqrt{n}} \xrightarrow{P} 0 \quad as \quad n \to \infty, \tag{3.1}$$

$\{n^{-1/2}\xi_n, \, n \geq 1\}$ satisfies Anscombe's condition, and

$$\frac{\xi_n}{\sqrt{n \log \log n}} \xrightarrow{a.s.} 0 \quad as \quad n \to \infty. \tag{3.2}$$

Proof. Since the properties of $\{S_n, \, n \geq 1\}$ are immediate we only have to show that $n^{-1}\xi_n \xrightarrow{a.s.} 0$ as $n \to \infty$ in order for $\{Z_n, \, n \geq 1\}$ to be a perturbed random walk. This, however, follows from the strong law of large numbers and the fact that $g \in \mathcal{G}$.

As for (3.1),

$$\frac{\xi_n}{\sqrt{n}} = (g'(\rho_n) - g'(\theta)) \cdot \sqrt{n}(\bar{Y}_n - \theta) \xrightarrow{P} 0 \quad as \quad n \to \infty, \tag{3.3}$$

since the first factor in the RHS tends to 0 almost surely as $n \to \infty$ (the strong law of large numbers and the continuity of g') and the second factor converges in distribution to the normal distribution with mean 0 and variance τ^2.

Next we note that, since the second factor in the RHS satisfies Anscombe's condition and the first factor tends to 0 a.s. as $n \to \infty$, it follows that $\{n^{-1/2}\xi_n, \, n \geq 1\}$ satisfies Anscombe's condition (cf. Woodroofe (1982), Lemma 1.4).

Finally, to prove (3.2) we note that

$$\frac{\xi_n}{\sqrt{n \log \log n}} = (g'(\rho_n) - g'(\theta)) \cdot \sqrt{\frac{n}{\log \log n}}(\bar{Y}_n - \theta),$$

where the first factor in the RHS tends to 0 a.s. as $n \to \infty$ as before and the limit superior of the absolute value of the second factor equals $\tau\sqrt{2}$ a.s. in view of the law of the iterated logarithm. The conclusion follows. $\qquad\square$

Lemma 3.5. *Suppose that g' is continuous at θ and that $\{B_n, \, n \geq 1\}$ are normalizing constants, such that the conditions of Theorem 3.4 are satisfied. Let, for $n \geq 1$, $Z_n = S_n + \xi_n$, where S_n and ξ_n are as in Lemma 3.4. Then*

$$\frac{\xi_n}{B_n} \xrightarrow{P} 0 \quad as \quad n \to \infty, \tag{3.4}$$

and $\{B_n^{-1}\xi_n, \, n \geq 1\}$ satisfies Anscombe's condition.

Proof. The proof is similar to that of Lemma 3.4. For example,

$$\frac{\xi_n}{B_n} = (g'(\rho_n) - g'(\theta)) \cdot \frac{n(\bar{Y}_n - \theta)}{B_n} \xrightarrow{P} 0 \quad as \quad n \to \infty,$$

since the first factor in the RHS tends to 0 almost surely as $n \to \infty$ and the second factor converges in distribution to a stable distribution. $\qquad\square$

Lemma 3.6. *Let $r \geq 2$. Suppose that $g' \in \mathrm{Lip}\,(\alpha, \theta)$ for some $\alpha > 1 - \frac{2}{r}$ and that $E|Y_1|^r < \infty$ and let, for $n \geq 1$, $Z_n = S_n + \xi_n$, where S_n and ξ_n are defined as in Lemma 3.4. Then*

$$\frac{\xi_n}{n^{1/r}} \xrightarrow{\mathrm{a.s.}} 0 \quad as \quad n \to \infty. \tag{3.5}$$

Proof. Use Lemma 3.1 and the Marcinkiewicz–Zygmund strong law (Gut (2007), Theorem 6.7.1)—note that $\frac{r(1+\alpha)}{r\alpha+1} < 2$. □

Remark 3.3. Note that, in view of (1.11), the conclusions of Lemmas 3.2–3.6 also hold for the sequence $\{\eta_k,\ k \geq 1\}$.

Proof of Theorems 3.1–3.6. It follows from the lemmas that $\{Z_n,\ n \geq 1\}$ is a perturbed walk and that the perturbing process in each case satisfies the smallness conditions required in the corresponding theorem of Section 6.2. □

Among the standard assumptions in nonlinear renewal theory are that g is twice continuously differentiable and that $\mathrm{Var}\,Y_1 < \infty$. To complete our treatment we add some results, which provide asymptotics for the perturbing process under these, traditional, assumptions and may be used to prove theorems of the above kind in that case.

Lemma 3.7. *Suppose that g'' is continuous at θ and that $\tau^2 = \mathrm{Var}\,Y_1 < \infty$. Set, for $n \geq 1$,*

$$S_n = ng(\theta) + ng'(\theta)(\bar{Y}_n - \theta) \quad and \quad \xi_n = n\frac{g''(\rho_n)}{2}(\bar{Y}_n - \theta)^2,$$

where $|\rho_n - \theta| \leq |\bar{Y}_n - \theta|$, and $Z_n = S_n + \xi_n$. Then $\{Z_n,\ n \geq 1\}$ is a perturbed random walk. The increments of the random walk component $\{S_n,\ n \geq 1\}$ are the i.i.d. random variables $\{g(\theta) + g'(\theta)(Y_k - \theta),\ k \geq 1\}$, which have positive, finite mean $\mu = g(\theta)$ and finite variance $\sigma^2 = \tau^2(g'(\theta))^2$. Furthermore,

$$\xi_n \xrightarrow{d} \frac{\tau^2 g''(\theta)}{2}\chi_1^2 \quad as \quad n \to \infty, \tag{3.6}$$

where χ_1^2 is a χ^2-distributed random variable with one degree of freedom.
If, in addition, $E|Y_1|^r < \infty$ for some $r \geq 2$, then

$$\frac{\xi_n}{n^{1/r}} \xrightarrow{\mathrm{a.s.}} 0 \quad as \quad n \to \infty. \tag{3.7}$$

Proof. The properties of $\{S_n,\ n \geq 1\}$ are the same as those of Lemma 3.4. The fact that $n^{-1}\xi_n \xrightarrow{\mathrm{a.s.}} 0$ as $n \to \infty$ follows from the strong law of large numbers and the continuity of g'' at θ. Thus, the process $\{Z_n,\ n \geq 1\}$ is a perturbed random walk.

The conclusion (3.6), follows from the central limit theorem and the continuity of g'' at θ.

The proof of (3.7) is similar to that of (3.5); the important ingredients being the Marcinkiewicz–Zygmund strong law, according to which (note that $\frac{2r}{r+1} < 2$)

$$n^{-1/r}\xi_n = \frac{1}{2}g''(\rho_n)\left(\frac{n(\bar{Y}_n - \theta)}{n^{(r+1)/2r}}\right)^2 \xrightarrow{a.s.} 0 \quad \text{as} \quad n \to \infty,$$

and the continuity of g'' at θ.   □

If, in Lemma 3.5, we only assume that the variance of the summands is finite, then (3.6) and (1.11) alone do not suffice in order to obtain any reasonable information about the asymptotic behavior of η_n, *unnormalized*, as $n \to \infty$. However, a direct investigation shows that, in fact, the following lemma holds.

Lemma 3.8. *Under the assumptions of Lemma 3.7 we have*

$$\eta_n \xrightarrow{P} 0 \quad \text{as} \quad n \to \infty.$$

Proof. We have

$$2\eta_n = ng''(\rho_n)(\bar{Y}_n - \theta)^2 - (n-1)g''(\rho_{n-1})(\bar{Y}_{n-1} - \theta)^2$$

$$= \big(g''(\rho_n) - g''(\theta)\big)n(\bar{Y}_n - \theta)^2 - \big(g''(\rho_{n-1}) - g''(\theta)\big)(n-1)(\bar{Y}_{n-1} - \theta)^2$$

$$+ g''(\theta) \cdot n\big((\bar{Y}_n - \theta)^2 - (\bar{Y}_{n-1} - \theta)^2\big) + g''(\theta)(\bar{Y}_{n-1} - \theta)^2$$

$$= \big(g''(\rho_n) - g''(\theta)\big)n(\bar{Y}_n - \theta)^2 - \big(g''(\rho_{n-1}) - g''(\theta)\big)(n-1)(\bar{Y}_{n-1} - \theta)^2$$

$$- \frac{n-1}{n}g''(\theta)(\bar{Y}_{n-1} - \theta)^2 + g''(\theta)\frac{(Y_n - \theta)^2}{n}$$

$$+ 2\frac{n-1}{n}g''(\theta)(Y_n - \theta)(\bar{Y}_{n-1} - \theta) \xrightarrow{P} 0 \quad \text{as} \quad n \to \infty.$$

The convergence to 0 is a consequence of the strong law of large numbers, the central limit theorem and the continuity of g'' at θ.   □

Remark 3.4. Since, under the assumptions of Lemma 3.7, $\{\xi_n, n \geq 1\}$, and, hence, $\{\eta_n, n \geq 1\}$, satisfy the traditional assumptions, it follows, in view of (3.6), Lemma 3.8, Theorem 3.1, and Anscombe's theorem, that, under those assumptions, we also have

$$\xi_{\nu(t)} \xrightarrow{d} \frac{\tau^2 g''(\theta)}{2}\chi_1^2 \quad \text{as} \quad t \to \infty,$$

and

$$\eta_{\nu(t)} \xrightarrow{P} 0 \quad \text{as} \quad t \to \infty,$$

where, as before, χ_1^2 is a χ^2-distributed random variable with one degree of freedom.

6.4 Convergence Rates

As a complement to a limit theorem one frequently tries to find results on the *rate* of convergence. For such theorems connected with the ordinary law of large numbers, see Katz (1963) and Baum and Katz (1965). For results on the rate of convergence in the strong law of large numbers and the law of the iterated logarithm for first passage times in the context of renewal theory on the real line, see Section 3.12; for details, see Gut (1983b). Here we confine our attention to one such theorem, related to Theorem 3.1, under the additional assumption that $a(y) \equiv 1$. Note that we have no assumption on g beyond the standard one.

Theorem 4.1. *Let $r \geq 1$. If $E|Y_1|^r < \infty$, then*

$$\sum_{n=1}^{\infty} n^{r-2} P(|\nu(n) - n/g(\theta)| > n\varepsilon) < \infty \quad \text{for all} \quad \varepsilon > 0.$$

Proof. The continuity of g implies that for every $\varepsilon > 0$ there exists $\delta > 0$ such that

$$\{|g(\bar{Y}_n) - g(\theta)| > \varepsilon\} \subset \{|\bar{Y}_n - \theta| > \delta\}. \tag{4.1}$$

It follows that

$$\sum_{n=1}^{\infty} n^{r-2} P(|Z_n - ng(\theta)| > n\varepsilon) \leq \sum_{n=1}^{\infty} n^{r-2} P(|\bar{Y}_n - \theta| > \delta) < \infty,$$

since the last sum, in fact, converges for every $\delta > 0$; see Baum and Katz (1965). Consequently the first sum also converges for all $\varepsilon > 0$. Similarly one may obtain

$$\sum_{n=1}^{\infty} n^{r-2} P(\max_{k \leq n} |Z_k - kg(\theta)| > n\varepsilon) < \infty \quad \text{for all} \quad \varepsilon > 0 \tag{4.2}$$

and

$$\sum_{n=1}^{\infty} n^{r-2} P\left(\sup_{k \geq n} \left|\frac{Z_k}{k} - g(\theta)\right| > \varepsilon\right) < \infty \quad \text{for all} \quad \varepsilon > 0 \tag{4.3}$$

where, however, (4.3) only holds for $r > 1$.

The conclusion of the theorem now follows by "inverting" (4.2) in the same fashion as in the proof of Gut (1983b), Theorem 3.1. We omit the details. $\square$

6.5 Finiteness of Moments; the General Case

We now return to the setting of Section 6.2 in order to establish results on the existence of moments of $\nu(t)$ and $Z_{\nu(t)}$.

Our first result deals with the moments of $\nu(t)$.

Theorem 5.1. *Let $r \geq 1$. Suppose that $E|X|^{r+1} < \infty$ and suppose that, for some ε, $0 < \varepsilon < \mu$,*

$$\sum_{n=1}^{\infty} n^{r-1} P(\xi_n \leq -n\varepsilon) < \infty. \qquad (5.1)$$

Then

$$E(\nu(t))^r < \infty.$$

Proof. We follow the ideas of Woodroofe (1982), page 46, formula (4.8). However, we first note that for $\delta > 0$, arbitrary small, $t \cdot a(n) \leq \delta n$ for n sufficiently large.

By modifying the just cited formula, it follows that, for $n \geq n_0$, large, that

$$P(\nu(t) > n) \leq P(S_n + \xi_n \leq n\delta) \leq P(S_n \leq n\delta + n\varepsilon) + P(\xi_n \leq -n\varepsilon)$$

$$\leq P(|S_n - n\mu| \geq n\gamma) + P(\xi_n \leq -n\varepsilon),$$

where $0 < \gamma = \mu - \delta - \varepsilon < \mu$. Consequently,

$$\sum_{n=n_0}^{\infty} n^{r-1} P(\nu(t) > n) \leq \sum_{n=n_0}^{\infty} n^{r-1} P(|S_n - n\mu| \geq n\gamma)$$

$$+ \sum_{n=1}^{\infty} n^{r-1} P(\xi_n \leq -n\varepsilon) < \infty,$$

since the first sum in the RHS is finite by Theorem 1 of Katz (1963) and the second sum is finite by assumption. □

Remark 5.1. The only property of $a(y)$, $y > 0$, that we actually used in the proof was $a(y) = o(y)$ as $y \to \infty$.

Remark 5.2. Assumption (5.1) is connected with the moments of the random variables

$$N(\varepsilon) = \text{Card} \left\{ n : \frac{\xi_n}{n} \leq -\varepsilon \right\} \quad \text{and} \quad L(\varepsilon) = \sup \left\{ n : \frac{\xi_n}{n} \leq -\varepsilon \right\}.$$

We know from (1.9) that both random variables are a.s. finite. For $r = 1$ (5.1) is exactly the same as the assumption that $E\, N(\varepsilon) < \infty$.

For the corresponding result for pure random walks we refer to Theorem 4.5.1, see also Gut (1974a), Theorem 3.1. In this case we also note that the moment assumption is $E|X|^r < \infty$, i.e., weaker than here. However, in that case the first passage times are stopping times (relative to the random walk), whereas here we have not imposed any such conditions (recall Remark 1.2). In the remainder of the section we do impose such conditions. More precisely, let, for each n, ξ_n be independent of $\{X_k, \, k > n\}$, and let $\mathcal{F}_n$ denote the

σ-algebra generated by $\{X_k, 1 \leq k \leq n\}$ and ξ_n (and let $\mathcal{F}_0$ be the trivial σ-algebra). Then $\{\nu(t), t \geq 0\}$ is a family of stopping times with respect to $\{\mathcal{F}_n, n \geq 1\}$ and it follows that any limit theorem for stopped random walks is (potentially) applicable to the family of stopped random walks

$$\{S_{\nu(t)}, t \geq 0\},$$

for which we can apply our results from Chapter 1.

Since the conditions involved in the most general results are fairly unpleasant we confine ourselves, at times, to considering the case $a(y) \equiv 1$.

The following result is inspired by Woodroofe (1982), Corollary 4.1.

Theorem 5.2. *Assume that $a(y) \equiv 1$. Let $r \geq 1$ and suppose that $E(X^-)^r < \infty$. If*

$$E\left((R(t) - \xi_{\nu(t)})^+\right)^r < \infty, \tag{5.2}$$

then

$$E(\nu(t))^r < \infty.$$

Proof. Since

$$Z_{\nu(t)} = S_{\nu(t)} + \xi_{\nu(t)} = t + R(t),$$

it follows that (5.2) is equivalent to the assumption that

$$E(S_{\nu(t)}^+)^r < \infty.$$

The conclusion now follows from Gut and Janson (1986), Theorem 3.3.   □

Remark 5.3. A second moment assumption for $R(t) - \xi_{\nu(t)}$ has been used in Woodroofe (1982), Chapter 4 in oder to obtain asymptotics for $E(\nu(t) - t/\mu)^2$ as $t \to \infty$ (i.e. for $a(y) \equiv 1$).

Condition (5.2) is not a very pleasant one. It is, however, possible to simplify the assumption in special cases (which, of course, in turn, involve other special assumptions). For example, if the perturbing process is nonnegative, it follows immediately that

$$\nu(t) \leq \nu^*(t), \tag{5.3}$$

where $\nu^*(t) = \min\{n : S_n > t\}$, $t \geq 0$, since, for all n, $S_n \leq Z_n$. In view of Theorem 3.3.1 we obtain the following result.

Theorem 5.3. *Let $r \geq 1$. Suppose that $P(\xi_n \geq 0) = 1$ for all n. If $E(X^-)^r < \infty$, then*

$$E(\nu(t))^r < \infty.$$

Remark 5.4. Under the assumptions of Theorem 5.3 it also follows that if X^- possesses a moment generating function then so does $\nu(t)$ (cf. Theorem 3.3.2 and Gut (1974a), Theorem 3.2).

Note also that here we actually do not need the stopping time assumption (since $\nu^*(t)$ is a stopping time with respect to the random walk).

Remark 5.5. For perturbed *renewal processes* (i.e., the random walk compo-
nents live on $[0, \infty)$), we only need to assume the validity of (5.1) for the
conclusion to hold.

Next we present a result for stopped processes. For details, see also
Larsson-Cohn (2001a).

Theorem 5.4. *Assume that $a(y) \equiv 1$. Let $r \geq 1$ and suppose that $E\,\nu(t) < \infty$.
Then*

(a) $E(X^+)^r < \infty$ and $E(\eta^+_{\nu(t)})^r < \infty$ $\implies$ $E(Z_{\nu(t)})^r < \infty$;

(b) $E(\eta_1^-)^r < \infty$ and $E(Z_{\nu(t)})^r < \infty$ $\implies$ $E(X^+)^r < \infty$;

(c) $E(X^-)^r < \infty$ and $E(Z_{\nu(t)})^r < \infty$ $\implies$ $E(\eta_1^+)^r < \infty$.

Proof. (a) The proof is similar to the proof of Theorem 3.3.1. In view of
Lemma 1.8.1 we have

$$E(X^+_{\nu(t)})^r \leq E\nu(t) \cdot E(X^+)^r < \infty. \tag{5.4}$$

Next, let $\| \cdot \|_r$ denote $(E| \cdot |^r)^{1/r}$. Then, by (5.4), (2.4) and Minkowski's
inequality (see e.g. Gut (2007), Theorem 3.2.6) we obtain (cf. formula (3.3.11)),

$$\|Z_{\nu(t)}\|_r \leq t + \|X^+_{\nu(t)}\|_r + \|\eta^+_{\nu(t)}\|_r \leq t + (E\nu(t))^{1/r}\|X^+\|_r + \|\eta^+_{\nu(t)}\|_r, \tag{5.5}$$

which proves (a).

Conclusion (b) follows from the inequalities

$$Z_{\nu(t)} \geq X_1 + \eta_1 \geq X_1 - \eta_1^-,$$

so that

$$X_1 \leq Z_{\nu(t)} + \eta_1^- \quad \implies \quad X_1^+ \leq Z_{\nu(t)} + \eta_1^-,$$

and (c) follows via

$$\eta_1^+ \leq (X_1 + \eta_1 - X_1)^+ \leq (X_1 + \eta_1)^+ + (-X_1)^+$$

$$= (X_1 + \eta_1)^+ + X_1^- \leq Z_{\nu(t)} + X_1^-. \qquad \square$$

Remark 5.6. As a corollary we note that if $E\,\nu(t) < \infty$ and

$$E \sup_k |\eta_k|^r < \infty, \tag{5.6}$$

then

$$E(X^+)^r < \infty \quad \Longleftrightarrow \quad E(Z_{\nu(t)})^r < \infty.$$

In particular, in the pure random walk case the conclusion reduces to
Theorem 3.3.1(ii). Moreover, in a specific application it is likely that (5.6)
is more easily checked than condition $E(\eta^+_{\nu(t)})^r < \infty$.

Remark 5.7. For general boundaries, such that $a(y) = o(y)$ as $y \to \infty$, we have, for δ arbitrary small, that $ta(n) \le a + \delta n$ for all n, which implies that (5.5) may be replaced by

$$\|Z_{\nu(t)}\|_r \le a + \delta\|\nu(t)\|_r + \|X^+_{\nu(t)}\|_r + \|\eta^+_{\nu(t)}\|_r$$

$$\le a + \delta\|\nu(t)\|_r + (E\nu(t))^{1/r}\|X^+\|_r + \|\eta^+_{\nu(t)}\|_r.$$

This shows that $E(Z_{\nu(t)})^r < \infty$ if the assumptions of Theorem 5.4 are fulfilled and if $E(\nu(t))^r < \infty$, for which conditions are presented above. For example, $E|X|^{r+1} < \infty$, (5.1) and (5.6) together imply the desired conclusion.

There is also the (trivial) conclusion that $E(S_{\nu(t)})^r < \infty$ and $E(\xi_{\nu(t)})^r < \infty$ together imply that $E(Z_{\nu(t)})^r < \infty$. Now, if the former assumption can be verified via the theory of stopped random walks and the latter can be verified by direct computation, then the implication can be used. For example, in view of Remark 5.4, it follows that

$$E(X^+)^r < \infty, \ E(\nu(t))^r < \infty, \ E(\xi_{\nu(t)})^r < \infty \quad \Longrightarrow \quad E(Z_{\nu(t)})^r < \infty.$$

This will be exploited in an example in the next section.

In view of (5.6), the previous lines and Remark 1.4 one might be tempted to use the, possibly simpler, condition

$$E \sup_n |\xi_n|^r < \infty. \tag{5.7}$$

However, in the same example below we shall see that (5.6) may be verified, whereas (5.7) does not hold (in spite of the fact that $E(\xi_{\nu(t)})^r < \infty$).

We close this section with a result for the overshoot. The proof consists of combining (2.5) and (5.4).

Theorem 5.5. *Let $r \ge 1$ and suppose that $E\nu(t) < \infty$ (thus, for example, that $\operatorname{Var} X < \infty$ and that (5.1) holds with $r=1$). If $E(X^+)^r < \infty$ and $E(\eta^+_{\nu(t)})^r < \infty$, then*

$$E(R(t))^r < \infty.$$

6.6 Finiteness of Moments; the Case $Z_n = n \cdot g(\bar{Y}_n)$

In this section we return to the model of Section 6.3, that is, $Y_1, Y_2, \ldots$ are i.i.d. random variables with positive, finite mean θ and $g \in \mathcal{G}$; recall Definition 3.1. Set $Z_n = n \cdot g(\bar{Y}_n)$, $n \ge 1$, where, again $\bar{Y}_n = \sum_{k=1}^n Y_k$, $n \ge 1$, and let

$$\nu(t) = \min\{n : n \cdot g(\bar{Y}_n) > t \cdot a(n)\} \quad (t \ge 0),$$

where $a(y)$, $y \ge 0$, is ultimately nondecreasing, concave and regularly varying at infinity with exponent β $(0 \le \beta < 1)$.

Theorem 6.1. *Let* $r \geq 1$. *Suppose that* $E|Y_1|^{r+1} < \infty$. *Then*

$$E(\nu(t))^r < \infty.$$

Proof. Let $\{Z_n, n \geq 1\}$ be given as in Lemma 3.2, that is, set, for $n \geq 1$, $S_n = ng(\theta)$ and $\xi_n = n \cdot (g(\bar{Y}_n) - g(\theta))$ and $Z_n = S_n + \xi_n$. Further, let $0 < \varepsilon < \mu$. The continuity of g implies (recall (4.1)) that, for an appropriate $\delta > 0$,

$$\sum_{n=1}^{\infty} n^{r-1} P(\xi_n \leq -n\varepsilon) \leq \sum_{n=1}^{\infty} n^{r-1} P(|g(\bar{Y}_n) - g(\theta)| \geq \varepsilon)$$

$$\leq \sum_{n=1}^{\infty} n^{r-1} P(|\bar{Y}_n - \theta| \geq \delta) < \infty,$$

by Katz (1963), Theorem 1, which verifies (5.1). ☐

Remark 6.1. In this case it follows from the construction that ξ_n is independent of $\{Y_k, k > n\}$, for each n, and, hence, that $\{\nu(t), t \geq 0\}$ are stopping times as desired.

The remark makes it tempting to guess that Theorem 6.1 might hold under weaker moment assumptions. This is, at least, the case in the next result, which relates to Theorem 5.3.

Theorem 6.2. *Let* $r \geq 1$ *and suppose that* g *is convex and twice continuously differentiable. If* $E((\text{sign } g'(\theta) \cdot Y_1)^{-})^r < \infty$, *then* $E(\nu(t))^r < \infty$.

Proof. We use the decomposition (1.6); cf. also Lemma 3.7. Then $g'' \geq 0$, i.e., the perturbing process is nonnegative. The conclusion follows. ☐

Remark 6.2. Another tractable case is $g(x) \geq x^+$, because existence of moments for the first passage times is easily established by domination (à la (5.3)).

Analogs to Theorem 5.4 are less pleasant to state. We leave these to the reader. Let us, however, consider the case $g(x) = x^2$, which is of special importance, as will be further elaborated in Section 6.9.

We first observe (cf. also Theorem 5.4.(b),(c)) that if $E(Z_{\nu(t)})^r < \infty$, then, since $Z_{\nu(t)} \geq Z_1^+ = (1 \cdot g(Y_1))^+ = Y_1^2$, it is necessary that

$$E|Y_1|^{2r} < \infty. \tag{6.1}$$

Thus, assume that this condition is met. Now, $g''(x) \equiv 2$ and $\xi_n = n \cdot (\bar{Y}_n - \theta)^2$ and, hence,

$$E(\xi_{\nu(t)})^r = E\big(\nu(t)(\bar{Y}_{\nu(t)} - \theta)^2\big)^r \leq E\left(\sum_{k=1}^{\nu(t)} (Y_k - \theta)\right)^{2r}$$

$$\leq B_{2r}\, E(\nu(t))^r\, E|Y_1 - \theta|^{2r} < \infty$$

by Theorems 1.5.1 and 6.2. Here B_{2r} is a numerical constant depending on r only.

Similarly,

$$E(S_{\nu(t)})^r = E\big(\nu(t)g(\theta) + \nu(t)g'(\theta)(\bar{Y}_{\nu(t)} - \theta)\big)^r$$

$$= E\big(\nu(t)\theta^2 + 2\nu(t)\theta(\bar{Y}_{\nu(t)} - \theta)\big)^r$$

$$\leq 2^{r-1}\theta^{2r} E(\nu(t))^r + 2^{2r-1}\theta^r E\left(\sum_{k=1}^{\nu(t)} (Y_k - \theta)\right)^r$$

$$\leq 2^{r-1}\theta^{2r} E(\nu(t))^r + 2^{2r-1}\theta^r B_r E(\nu(t))^{\frac{r}{2}\vee 1} E|Y_1 - \theta|^r < \infty.$$

We have thus shown that (6.1) in this case is necessary as well as sufficient for $E(Z_{\nu(t)})^r < \infty$ to hold.

Remark 6.3. Since we require a moment of order $2r$ in this example it might at first sight appear to be a weaker conclusion than the corresponding results for (pure) random walks. To see that this is not the case we observe that

$$\nu(t) = \min\{n : ng(\bar{Y}_n) > t\} = \min\{n : n(\bar{Y}_n)^2 > t\}$$

$$= \min\left\{n : \left|\sum_{k=1}^{n} Y_k\right| > \sqrt{t}\sqrt{n}\right\}.$$

Thus, in the present context we stop the *square* of a random walk, i.e., moments of order r of the stopped process involve moments of the summands of order $2r$. Now, as we have just seen, this stopping problem is equivalent to stopping the random walk itself (at another level), in which case moments of the process of order r involve moments of the summands of order r. So, although it is the same stopping problem, we consider different processes, which is what causes the, seemingly, different results.

Let us also check whether or not (5.7), i.e., $E\sup_n(\xi_n)^r < \infty$, holds. This amounts to studying the quantity $E\sup_n n^r(\bar{Y}_n - \theta)^{2r} = E\sup_n(\sqrt{n}(\bar{Y}_n - \theta))^{2r}$. However, in view of the law of the iterated logarithm this expectation is always infinite.

Next we turn our attention to (5.6). We have (cf. the proof of Lemma 3.8)

$$\eta_n = n(\bar{Y}_n - \theta)^2 - (n-1)(\bar{Y}_{n-1} - \theta)^2$$

$$= -\frac{n-1}{n}(\bar{Y}_{n-1} - \theta)^2 + \frac{(Y_n - \theta)^2}{n} + 2\frac{n-1}{n}(Y_n - \theta)(\bar{Y}_{n-1} - \theta). \quad (6.2)$$

First, let $r > 1$ and assume that the (trivially) necessary assumption $E|Y_1|^{2r} < \infty$ is satisfied. Since $\{\bar{Y}_n - \theta, n \geq 1\}$ is a reversed martingale (see e.g. Gut (2007), p. 545), it follows from Doob's maximal inequality (see e.g. Gut (2007), Theorem 10.9.4) that

$$E \sup_n |\bar{Y}_n - \theta|^{2r} \leq \left(\frac{2r}{2r-1}\right)^{2r} E|Y_1 - \theta|^{2r} < \infty.$$

Further,

$$E \sup_n \left(\frac{(Y_n - \theta)^2}{n}\right)^r \leq \sum_{n=1}^{\infty} \frac{E|Y_1 - \theta|^{2r}}{n^r} < \infty. \quad (6.3)$$

Next, by independence and the Marcinkiewicz–Zygmund (1937) inequalities (see e.g. Gut (2007), Theorem 3.8.1 or formula (A.2.2)), we have

$$E \sup_n \left|\frac{n-1}{n}(Y_n - \theta)(\bar{Y}_{n-1} - \theta)\right|^{2r}$$

$$\leq \sum_{n=2}^{\infty} \left(\frac{n-1}{n}\right)^{2r} E|Y_n - \theta|^{2r} \cdot |\bar{Y}_{n-1} - \theta|^{2r}$$

$$= E|Y_1 - \theta|^{2r} \cdot \sum_{n=1}^{\infty} n^{-2r} E \left|\sum_{k=1}^{n}(Y_k - \theta)\right|^{2r}$$

$$\leq E|Y_1 - \theta|^{2r} \cdot \sum_{n=1}^{\infty} n^{-2r} B_{2r} n^r E|Y_1 - \theta|^{2r}$$

$$= B_{2r} \left(E|Y_1 - \theta|^{2r}\right)^2 \sum_{n=1}^{\infty} n^{-r} < \infty.$$

Again, B_{2r} is a numerical constant depending on r only. By Lyapounov's inequality it finally follows that

$$\left\|\sup_n \frac{n-1}{n}(Y_n - \theta)(\bar{Y}_{n-1} - \theta)\right\|_r \leq \left\|\sup_n \frac{n-1}{n}(Y_n - \theta)(\bar{Y}_{n-1} - \theta)\right\|_{2r} < \infty.$$

Thus, if $E|Y_1|^{2r} < \infty$, then (5.6) is satisfied. Note that we, in fact, just have shown that

$$E \sup_k |\eta_k|^r < \infty \quad \Longrightarrow \quad E|Y_1|^{2r} < \infty$$

in this case.

For $r = 1$, $E \sup_n n^{-r}(Y_n - \theta)^{2r} < \infty$ if and only if $E Y_1^2 \log^+ |Y_1| < \infty$ (Burkholder (1962)), that is, we need a little more than finite variance in order to ensure the finiteness of the LHS of (6.3).

6.7 Moment Convergence; the General Case

We now return to the general model, defined in Sections 6.2 and 6.5 under the additional assumption that ξ_n is independent of $\{X_k, \ k > n\}$ for each n. Moreover, we restrict ourselves to considering the case $a(y) \equiv 1$. As always, moment convergence follows from convergence in probability or in distribution, together with uniform integrability. Since we have already established the former, the essential task is to prove uniform integrability, the proof of which is based on an induction procedure, which has been used before in the random walk context; see Sections 1.5 and 1.6 (cf. also Gut (1974a)).

Theorem 7.1. *Let $r \geq 1$. Suppose that $E|X|^{r \vee 2} < \infty$ and that*

$$\sum_{n=1}^{\infty} P(\xi_n \leq -n\varepsilon) < \infty \tag{7.1}$$

If

$$\left\{ \left| \frac{R(t) - \xi_{\nu(t)}}{t} \right|^r, \ t \geq 1 \right\} \quad \text{is uniformly integrable,} \tag{7.2}$$

then

(a) $\left\{ \left(\frac{\nu(t)}{t} \right)^r, \ t \geq 1 \right\}$ *is uniformly integrable;*

(b) $E \left(\frac{\nu(t)}{t} \right)^p \to \frac{1}{\mu^p} \quad \text{as} \quad t \to \infty \quad \text{for all} \quad p, \ 0 < p \leq r.$

Proof. For the case $r = 1$, see Woodroofe (1982), Theorem 4.4.

Since (b) follows from (a) and Theorem 2.1, we only have to prove the first part. Toward this end,

$$\mu \cdot \frac{\nu(t)}{t} \leq \frac{|S_{\nu(t)} - \mu\nu(t)|}{t} + 1 + \frac{|R(t) - \xi_{\nu(t)}|}{t}. \tag{7.3}$$

First, let $1 < r \leq 2$. It follows from Theorem 1.6.4 that

$$\left\{ \left| \frac{S_{\nu(t)} - \mu\nu(t)}{t} \right|^r, \ t \geq 1 \right\} \quad \text{is uniformly integrable.}$$

This, together with (7.2) and (7.3), proves the conclusion for this r-interval.

Next, if $2 < r \leq 2^2$ the conclusion follows in the same way, via what we just proved, etc. $\qquad\square$

Remark 7.1. The moment assumption about finite variance is needed (in the proof) for the interval $1 < r \leq 2$, since we rely on Woodroofe (1982).

Next we present the analog of Theorem 5.3.

Theorem 7.2. *Let $r \geq 1$. Suppose that $P(\xi_n \geq 0) = 1$ for all n. If $E(X^-)^r < \infty$, then the conclusions of Theorem 7.1 hold.*

Proof. The conclusions follow from (5.3) and the corresponding results in the pure random walk case; Theorem 3.7.1. $\qquad\square$

The corresponding result related to the central limit theorem runs as follows.

Theorem 7.3. *Suppose that $E|X|^r < \infty$ for some $r \geq 2$ and set $\sigma^2 = \operatorname{Var} X$. Further, suppose that (7.1) is satisfied. If*

$$\left\{ \left| \frac{R(t) - \xi_{\nu(t)}}{\sqrt{t}} \right|^r , t \geq 1 \right\} \quad \text{is uniformly integrable,} \qquad (7.4)$$

then

(a) $\left\{ \left| \dfrac{\nu(t) - t/\mu}{\sqrt{t}} \right|^r , t \geq 1 \right\}$ *is uniformly integrable;*

(b) $E \left| \dfrac{\nu(t) - t/\mu}{\sqrt{t}} \right|^r \to E|N|^r$ *as $t \to \infty$ for all p, $0 < p \leq r$. Here N is a normal random variable with mean 0 and variance $\mu^{-3}\sigma^2$;*

(c) $E \left(\dfrac{\nu(t) - t/\mu}{\sqrt{t}} \right)^k \to 0$ *as $t \to \infty$ for $k =$ odd integer $\leq r$.*

Proof. The proof follows the same pattern as that of Theorem 7.1 with (7.3) replaced by

$$\mu \cdot \frac{|\nu(t) - t/\mu|}{\sqrt{t}} \leq \frac{|S_{\nu(t)} - \mu\nu(t)|}{\sqrt{t}} + \frac{|R(t) - \xi_{\nu(t)}|}{\sqrt{t}},$$

and the appeal to Theorem 1.6.4, replaced by an appeal to Theorem 1.6.3 there. We omit the details. $\qquad\square$

Remark 7.2. The corresponding result for the Marcinkiewicz–Zygmund law (Theorem 2.2) is not relevant here, since we would need finite variance in the proof.

For the stopped process and the overshoot (recall Theorems 5.4 and 5.5, respectively) we have the following result.

Theorem 7.4. *If, for some $r \geq 1$, $E(X^+)^{r \vee 2} < \infty$,*

$$\left\{ \left(\frac{\nu(t)}{t} \right)^r , t \geq 1 \right\} \quad and \quad \left\{ \left(\frac{\eta^+_{\nu(t)}}{t} \right)^r , t \geq 1 \right\} \quad are\ uniformly\ integrable,$$

(7.5)

then

(a) $\left\{ \left(\frac{Z_{\nu(t)}}{t} \right)^r , t \geq 1 \right\}$ *is uniformly integrable;*

(b) $E \left(\frac{Z_{\nu(t)}}{t} \right)^p \to 1 \quad as \quad t \to \infty \quad for\ all\ p,\ 0 < p \leq r.$

Proof. We have (cf. (5.5)), for $t \geq 1$,

$$1 < \frac{Z_{\nu(t)}}{t} \leq 1 + \frac{X^+_{\nu(t)}}{t} + \frac{\eta^+_{\nu(t)}}{t} \leq 1 + \frac{X^+_{\nu(t)}}{t^{1/r}} + \frac{\eta^+_{\nu(t)}}{t},$$

(7.6)

which, together with Theorem 1.8.1 and (7.5), proves (a), from which (b) follows in view of Theorem 2.6 (cf. also Larsson-Cohn (2001a)).   □

Theorem 7.5. *If, for some $r \geq 1$, $E(X^+)^{r \vee 2} < \infty$ and*

$$\left\{ \left(\frac{\eta^+_{\nu(t)}}{t^{1/r}} \right)^r , t \geq 1 \right\} \quad is\ uniformly\ integrable,$$

then

(a) $\left\{ \left(\frac{R(t)}{t^{1/r}} \right)^r , t \geq 1 \right\}$ *is uniformly integrable;*

(b) $E \left(\frac{R(t)}{t^{1/r}} \right)^r \to 0 \quad as \quad t \to \infty.$

Proof. The proof follows the same pattern as the previous one. We leave the obvious modifications to the reader.   □

6.8 Moment Convergence; the Case $Z_n = n \cdot g(\bar{Y}_n)$

Consider, again, the model described in Sections 6.3 and 6.6. Here we present some results corresponding to those of Section 6.7; recall that $a(y) \equiv 1$. We further restrict our attention to the case when g is convex and twice continuously differentiable at θ.

The first result is an immediate consequence of Theorem 7.2 (and Theorem 3.1).

Theorem 8.1. *If* $E((\operatorname{sign} g'(\theta) \cdot Y)^-)^r < \infty$, *then*

(a) $\left\{ \left(\dfrac{\nu(t)}{t} \right)^r, t \geq 1 \right\}$ *is uniformly integrable;*

(b) $E \left(\dfrac{\nu(t)}{t} \right)^p \rightarrow \dfrac{1}{g(\theta)^p}$ *as* $t \to \infty$ *for all* $p,\ 0 < p \leq r$.

Since no extra moment assumptions are needed in this case, we also provide a result related to the Marcinkiewicz–Zygmund strong law (Theorem 3.2). The proof follows the lines of that of Theorem 7.3. Theorem 8.3 below is immediate.

Theorem 8.2. *Suppose that* $E|Y_1|^r < \infty$ *for some* $r,\ 1 < r < 2$. *If*

$$\left\{ \left| \frac{R(t) - \xi_{\nu(t)}}{t^{1/r}} \right|^r, t \geq 1 \right\} \quad \text{is uniformly integrable,}$$

then

(a) $\left\{ \left| \dfrac{\nu(t) - t/g(\theta)}{t^{1/r}} \right|^r, t \geq 1 \right\}$ *is uniformly integrable;*

(b) $E \left| \dfrac{\nu(t) - t/g(\theta)}{t^{1/r}} \right|^r \rightarrow 0$ *as* $t \to \infty$.

Theorem 8.3. *Suppose that* $E|Y_1|^r < \infty$ *for some* $r \geq 2$ *and set* $\tau^2 = \operatorname{Var} Y_1$. *If (7.4) is satisfied, then*

(a) $\left\{ \left| \dfrac{\nu(t) - t/g(\theta)}{\sqrt{t}} \right|^r, t \geq 1 \right\}$ *is uniformly integrable;*

(b) $E \left| \dfrac{\nu(t) - t/g(\theta)}{\sqrt{t}} \right|^r \rightarrow E|N|^r$ *as* $t \to \infty$ *for all* $p,\ 0 < p \leq r$. *Here*
 N *is a normal random variable with mean 0 and variance* $g(\theta)^{-3}(\tau g'(\theta))^2$;

(c) $E \left(\dfrac{\nu(t) - t/g(\theta)}{\sqrt{t}} \right)^k \rightarrow 0$ *as* $t \to \infty$ *for* $k = $ *odd integer* $\leq r$.

The analogs for the stopped process and the overshoot are identical to Theorems 7.4 and 7.5, respectively, except that the moment assumption becomes $E|Y_1|^r < \infty$ for some $r \geq 1$.

We conclude this section by checking condition (7.4) for the case $g(x) = x^2$. By proceeding as in Section 6.6, starting with (6.2), we obtain

$$|\eta_{\nu(t)}| \leq \sup_n (\bar{Y}_n - \theta)^2 + \sup_n \frac{(Y_n - \theta)^2}{n} + 2|Y_{\nu(t)} - \mu| \sup_n |\bar{Y}_n - \theta|. \qquad (8.1)$$

Now, suppose that $E|Y_1|^{2r} < \infty$ for some $r \geq 2$ (since we wish to check the validity of Theorem 8.3). We know from Section 6.6 that the first two terms of the RHS of (8.1) are in L^r, so to verify (7.4) it suffices to check the last term, properly normalized.

Via the inequality $2|xy| \leq x^2 + y^2$ we have, for $t \geq 1$,

$$2\left|\frac{Y_{\nu(t)} - \theta}{\sqrt{t}}\right|^r \sup_n |\bar{Y}_n - \theta|^r \leq \left|\frac{Y_{\nu(t)} - \theta}{\sqrt{t}}\right|^{2r} + \sup_n |\bar{Y}_n - \theta|^{2r}$$

$$\leq \frac{|Y_{\nu(t)} - \theta|^{2r}}{t} + \sup_n |\bar{Y}_n - \theta|^{2r}. \qquad (8.2)$$

Now, since $\{\frac{|Y_{\nu(t)} - \theta|^{2r}}{t}, \ t \geq 1\}$ is uniformly integrable (Theorem 1.8.1) and the second term in the RHS of (8.2) is integrable, it follows that (7.4) holds and we are done. Theorem 8.3 thus applies under the (sole) assumption that $E|Y_1|^{2r} < \infty$.

6.9 Examples

In this section we present some examples of the kind discussed in Sections 6.3, 6.6 and 6.8.

Example 9.1. The obvious first example is $g(x) = x^+$ and $a(y) \equiv 1$, which reduces the model to renewal theory for random walks, i.e., Chapter 3.

Example 9.2. The case $g(x) = x^+$ yields $\nu(t) = \min\{n : \sum_{k=1}^n Y_k > t \cdot a(n)\}$, $t \geq 0$, that is, first passage times across general boundaries, i.e., Section 4.5.

Example 9.3. Next, let $g(x) = (x^+)^{\frac{1}{1-\beta}}$ $(0 < \beta < 1)$ and $a(y) \equiv 1$. Then

$$\nu(t) = \min\{n : n \cdot (\bar{Y}_n)^{\frac{1}{1-\beta}} > t\} = \min\left\{n : \sum_{k=1}^n Y_k > t^{1-\beta} \cdot n^\beta\right\},$$

and a change of boundary parameter shows that the family $\{\nu(t), \ t \geq 0\}$ corresponds to the family

$$\min\left\{n : \sum_{k=1}^n Y_k > tn^\beta\right\} \quad (t \geq 0).$$

More general examples are obtained with more general boundaries.

Example 9.4. Let $g(x)$ be as in Example 9.3, but suppose that $\beta > 1$, and let, again, $a(y) \equiv 1$. Computations like those above show that $\{\nu(t), \ t \geq 0\}$ corresponds to

$$\min\left\{n : \sum_{k=1}^n Y_k < tn^\beta\right\} = \min\left\{n : \frac{\sum_{k=1}^n Y_k}{n^\beta} < t\right\} \quad (t \geq 0).$$

This model is called *inverse renewal theory*; see Goldmann (1984). More general boundaries yield extensions of his results.

Let, for example, $\beta = 2$ (that is, $g(x) = 1/x^+$) and suppose that $E|Y_1|^r < \infty$, for some r, $1 \le r < 2$. Then $\lambda(t) = 1/(\theta t)$, so that Theorem 3.2 provides the Marcinkiewicz–Zygmund law

$$t^{1/r}\big(\nu(t) - 1/(\theta t)\big) \xrightarrow{\text{a.s.}} 0 \quad \text{as} \quad t \to 0,$$

which is not contained in Goldmann (1984).

Example 9.5. The choices $g(x) = |x|$ and $g(x) = |x|^{\frac{1}{1-\beta}}$ with $0 < \beta < 1$ or $\beta > 1$ yield two-sided versions of the above. The example $g(x) = x^2$ and $a(y) \equiv 1$ corresponds to $\min\{n : |\sum_{k=1}^n Y_k| > t\sqrt{n}\}$, $t > 0$; "the square root boundary".

Example 9.6. In connection with a sequential procedure, Chernoff (1972), page 80, considers the family

$$\min\left\{n : \left|\sum_{k=1}^n Y_k\right| + n/2 > t\sqrt{n}\right\} \quad (t \ge 0),$$

which corresponds to $g(x) = (|x| + \frac{1}{2})^2$ and $a(y) \equiv 1$; see also Lai and Siegmund (1977), page 946.

The examples so far were obtained by relatively simple choices of the function g. Our final example is more sophisticated. For details beyond those given here and for further examples we refer to Woodroofe (1982), Chapters 6 and 7, Siegmund (1985), Wijsman (1991), Section 3 and references given therein. The example appears in the context of repeated significance tests in one-parameter exponential families.

Example 9.7. Consider the family of distributions

$$G_\theta(dx) = \exp\{\theta x - \psi(\theta)\}\lambda(dx), \quad \theta \in \Theta,$$

where λ is a nondegenerate, σ-finite measure on $(-\infty, \infty)$ and Θ a nondegenerate real interval. Furthermore, let $Y, Y_1, Y_2, \ldots$ be i.i.d. random variables with distribution function G_θ for some $\theta \in \Theta$, where θ is unknown, and suppose that we wish to test the hypothesis

$$H_0 : \theta = \theta_0 \quad \text{vs} \quad H_1 : \theta \ne \theta_0.$$

For simplicity we assume that $\theta_0 = 0$ and that $\psi(0) = \psi'(0) = 0$ (this can otherwise be achieved by a renormalization).

As for some facts we mention that the moment generating function $E_\theta \exp\{tY\} = \exp\{\psi(\theta + t) - \psi(\theta)\}$ exists for t-values such that $t + \theta \in \Theta$, that ψ is convex, and, assuming that ψ is twice differentiable in a neighborhood of $\theta_0 = 0$, it follows that

$$E_\theta Y = \psi'(\theta) \quad \text{and} \quad \text{Var}_\theta Y = \psi''(\theta) > 0,$$

(in particular, G_θ is not degenerate).

The log-likelihood ratio is

$$T_n = \sup_{\theta \in \Theta} \log \prod_{k=1}^{n} \exp\{\theta Y_k - \psi(\theta)\} = \sup_{\theta \in \Theta} \left\{ \theta \sum_{k=1}^{n} Y_k - n\psi(\theta) \right\}$$

$$= n \cdot \sup_{\theta \in \Theta} \{\theta \bar{Y}_n - \psi(\theta)\} = n \cdot g(\bar{Y}_n),$$

where

$$g(x) = \sup_{\theta} \left(\theta x - \psi(\theta) \right), \quad -\infty < x < \infty,$$

is the convex (Fenchel) conjugate of ψ. Moreover, g is convex and twice continuously differentiable. This shows that T_n can be rewritten in the desired form and that, in fact, all traditional assumptions are satisfied.

The sequential test procedure is to reject H_0 as soon as $T_n > c$, that is, the object of interest is

$$N(t) = \min\{n : T_n > t\} \quad (t \geq 0). \tag{9.1}$$

This yields a so-called *open-ended test*. A *repeated significance test* amounts to rejecting H_0 as soon as $T_{N \wedge m} > t$, where m is some given number or, more generally, to define $N^*(t) = \min\{n \geq m_0 : T_n > t\}$ with the same rejection criterion. Here m_0 represents an initial sample size and m the maximal sample size.

As a complement to the previous example we close with some applications.

• The typical example is $\psi(\theta) = \frac{1}{2}\theta^2$, for which

$$E_\theta \exp\{tY\} = \exp\left\{ t\theta + \frac{1}{2}\theta^2 \right\},$$

which means that $Y \in N(\theta, 1)$. Then $g(x) = \frac{1}{2}x^2$ and

$$T_n = \frac{1}{2}n \cdot (\bar{Y}_n)^2 = \frac{1}{2n} \left(\sum_{k=1}^{n} Y_k \right)^2 \tag{9.2}$$

(recall the example in Sections 6.6 and 6.8). The stopping procedure (9.1) corresponds to the stopping times

$$N(t) = \min \left\{ n : \left| \sum_{k=1}^{n} Y_k \right| > t\sqrt{n} \right\} \quad (t \geq 0),$$

(with another t), that is, we rediscover the square root boundary problem. As an example, an application of Theorems 3.3 and 8.3 shows that

$$\frac{\nu(t) - 2t/\theta^2}{\frac{2}{\theta^2}\sqrt{2t}} \to N(0, 1) \quad \text{in distribution and in } L^r \text{ for all } r \quad \text{as} \quad t \to \infty.$$

- For the case $P(Y_1 = 1) = 1 - P(Y_1 = 0) = p$, $0 < p < 1$ (cf. Woodroofe (1982), Siegmund (1985)), a direct computation (without normalization) shows that $g(x) = x \log x + (1 - x) \log(1 - x) + \log 2$, which calls for the stopping times

$$N(t) = \min\{n : n(\bar{Y}_n \log \bar{Y}_n + (1 - \bar{Y}_n) \log(1 - \bar{Y}_n) + \log 2) > t\} \quad (t \geq 0).$$

  An application of the same theorems yields

$$\frac{\nu(t) - (p \log p + (1 - p) \log(1 - p) + \log 2)^{-1} \cdot t}{\sqrt{t}} \to N \quad \text{as} \quad t \to \infty$$

  in distribution and in L^r for all r. Here N is a normal random variable with mean 0 and variance

$$\frac{p(1 - p)\big(\log(p/1 - p)\big)^2}{\big(p \log p + (1 - p) \log(1 - p) + \log 2\big)^3}.$$

- For the exponential distribution with mean $1/(1 - \theta)$ and $H_0 : \theta = 0$ we have $g(x) = x - 1 - \log x$ and

$$N(t) = \min\{n : n(\bar{Y}_n - 1 - \log \bar{Y}_n) > t\} \quad (t \geq 0).$$

  We leave it to the reader to formulate relevant results for this case.

6.10 Stopped Two-Dimensional Perturbed Random Walks

In the same vein as the results in Chapter 3 have been extended to the case of perturbed random walks earlier in this chapter, we shall, in this section, extend the results from Section 4.2 to the case of *stopped two-dimensional perturbed random walks*. In the following section we consider the case $Z_n = n \cdot g(\bar{Y}_n)$, after which, in Section 6.12, we present an application to repeated significance tests in two-parameter exponential families and conclude with an example for normally distributed random variables. For more about the latter, see Gut and Schwabe (1996, 1999).

Thus, let $\{(U_n^{(x)}, U_n^{(y)}), n \geq 1\}$, with i.i.d. increments $\{(X_k, Y_k), k \geq 1\}$, be a two-dimensional random walk and suppose that $\mu_y = E Y_1 > 0$ and that $\mu_x = E X_1$ exists, finite. Further, let $\{\xi_n^{(x)}, n \geq 1\}$ and $\{\xi_n^{(y)}, n \geq 1\}$ be sequences of random variables, such that (cf. (1.9) in the one-dimensional case)

$$\frac{\xi_n^{(x)}}{n} \xrightarrow{\text{a.s.}} 0 \quad \text{and} \quad \frac{\xi_n^{(y)}}{n} \xrightarrow{\text{a.s.}} 0 \quad \text{as} \quad n \to \infty.$$

As in Section 4.2, we do not assume any independence between the components X_k and Y_k for any k, and, as earlier in this chapter, no independence between the perturbations and the future of the random walk.

Next we define the two-dimensional perturbed random walk

$$(Z_n^{(x)}, Z_n^{(y)}) = (U_n^{(x)} + \xi_n^{(x)}, U_n^{(y)} + \xi_n^{(y)}), \quad n \geq 1,$$

and the first passage time process

$$\tau(t) = \min\{n : Z_n^{(y)} > t\} \quad (t \geq 0).$$

Clearly, the first passage times are stopping times (relative to the sequence of σ-algebras generated by the perturbed random walk). Moreover, since $\mu_y > 0$ the results from the earlier sections of this chapter apply to $\{\tau(t), t \geq 0\}$ and $\{Z_{\tau(t)}^{(y)}, t \geq 0\}$ (note also that $a(y) \equiv 1$ throughout in this section).

The object of interest is the family of stopped perturbed random walks

$$\{Z_{\tau(t)}^{(x)}, \quad t \geq 0\}.$$

As a sample we state the following extensions of the results from Section 4.2. The proofs follow rather much the, by now, usual pattern. We therefore confine ourselves to indicating proofs and pointing to differences and novelties.

Theorem 10.1. *We have*

$$\frac{Z_{\tau(t)}^{(x)}}{t} \xrightarrow{\text{a.s.}} \frac{\mu_x}{\mu_y} \quad as \quad t \to \infty.$$

Theorem 10.2. *Let $1 \leq r < 2$ and suppose that $E|X|^r < \infty$ and $E|Y|^r < \infty$. If*

$$\frac{\xi_n^{(x)}}{n^{1/r}} \xrightarrow{\text{a.s.}} 0 \quad and \quad \frac{\xi_n^{(y)}}{n^{1/r}} \xrightarrow{\text{a.s.}} 0 \quad as \quad n \to \infty,$$

then

$$\frac{Z_{\tau(t)}^{(x)} - \frac{\mu_x}{\mu_y} t}{n^{1/r}} \xrightarrow{\text{a.s.}} 0 \quad as \quad t \to \infty.$$

Theorem 10.3. *Suppose that $\sigma_x^2 = \text{Var}\, X_1 < \infty$, $\sigma_y^2 = \text{Var}\, Y_1 < \infty$ and that $\gamma^2 = \text{Var}\,(\mu_y X_1 - \mu_x Y_1) > 0$. If*

$$\frac{\xi_n^{(x)}}{\sqrt{n}} \xrightarrow{\text{a.s.}} 0 \quad and \quad \frac{\xi_n^{(y)}}{\sqrt{n}} \xrightarrow{\text{a.s.}} 0 \quad as \quad n \to \infty,$$

or if

$$\frac{\xi_n^{(x)}}{\sqrt{n}} \xrightarrow{p} 0 \quad and \quad \frac{\xi_n^{(y)}}{\sqrt{n}} \xrightarrow{p} 0 \quad as \quad n \to \infty$$

and

$$\left\{\frac{\xi_n^{(x)}}{\sqrt{n}},\ n\geq 1\right\}\quad\text{and}\quad\left\{\frac{\xi_n^{(y)}}{\sqrt{n}},\ n\geq 1\right\}\quad\text{satisfy Anscombe's condition,}$$

then

$$\frac{Z_{\tau(t)}^{(x)}-\frac{\mu_x}{\mu_y}t}{\mu_y^{-3/2}\gamma\sqrt{t}}\xrightarrow{d}N(0,1)\quad\text{as}\quad t\to\infty.$$

Theorem 10.4. *Suppose that* $\sigma_x^2=\operatorname{Var}X_1<\infty$, $\sigma_y^2=\operatorname{Var}Y_1<\infty$ *and that* $\gamma^2=\operatorname{Var}(\mu_yX_1-\mu_xY_1)>0$. *If*

$$\frac{\xi_n^{(x)}}{\sqrt{n\log\log n}}\xrightarrow{\text{a.s.}}0\quad\text{and}\quad\frac{\xi_n^{(y)}}{\sqrt{n\log\log n}}\xrightarrow{\text{a.s.}}0\quad\text{as}\quad n\to\infty,$$

then

$$C\left(\left\{\frac{Z_{\tau(t)}^{(x)}-\frac{\mu_x}{\mu_y}t}{\mu_y^{-3/2}\gamma\sqrt{2t\log\log t}},\ t\geq 3\right\}\right)=[-1,1]\quad\text{a.s.}$$

In particular,

$$\limsup_{t\to\infty}(\liminf_{t\to\infty})\frac{Z_{\tau(t)}^{(x)}-\frac{\mu_x}{\mu_y}t}{\sqrt{2t\log\log t}}=\frac{+}{(-)}\frac{\gamma}{\mu_y^{3/2}}\quad\text{a.s.}$$

We confine ourselves to proving Theorem 10.3 leaving the others to the reader; it is just a matter of modifying the proofs in Sections 4.2 and those given earlier in this chapter in the natural manner.

Proof of Theorem 10.3. Following the proof of Theorem 4.2.3 we set

$$S_n=\mu_yU_n^{(x)}-\mu_xU_n^{(y)},\quad n\geq 1.$$

The random walk $\{S_n,\ n\geq 1\}$ thus defined has increments with mean 0 and finite variance γ^2. It follows from the central limit theorem, Theorem 2.1 and Anscombe's theorem (Theorem 1.3.1) that

$$\frac{S_{\tau(t)}}{\gamma\sqrt{\tau(t)}}\xrightarrow{d}N(0,1)\quad\text{as}\quad t\to\infty.$$

Next we note that

$$\frac{\xi_{\tau(t)}^{(x)}}{\sqrt{\tau(t)}}\xrightarrow{p}0\quad\text{and}\quad\frac{\xi_{\tau(t)}^{(y)}}{\sqrt{\tau(t)}}\xrightarrow{p}0\quad\text{as}\quad t\to\infty$$

by the assumptions on the perturbing processes and Theorem 1.2.1 or Anscombe's theorem, respectively. In either case it follows that

$$\frac{S_{\tau(t)} + \mu_y \xi_{\tau(t)}^{(x)} - \mu_x \xi_{\tau(t)}^{(y)}}{\gamma \sqrt{\tau(t)}} \xrightarrow{d} N(0,1) \quad \text{as} \quad t \to \infty,$$

which is the same as

$$\frac{\mu_y Z_{\tau(t)}^{(x)} - \mu_x Z_{\tau(t)}^{(y)}}{\gamma \sqrt{\tau(t)}} \xrightarrow{d} N(0,1) \quad \text{as} \quad t \to \infty.$$

Combining this with Theorem 2.8 and Remark 1.4, which take care of the overshoots $R(t) = Z_{\tau(t)}^{(y)} - t$, $t \geq 0$, we arrive at

$$\frac{\mu_y Z_{\tau(t)}^{(x)} - \mu_x t}{\gamma \sqrt{\tau(t)}} \xrightarrow{d} N(0,1) \quad \text{as} \quad t \to \infty.$$

A final application of Theorem 10.1 and some trivial rearrangements finish the proof.   □

Remark 10.1. As in Section 4.2 it is possible to allow for $\{Z_n^{(x)}, n \geq 1\}$ to be multidimensional and, for example, obtain a multidimensional version of Theorem 10.3. In particular, one can show that $(Z_{\tau(t)}^{(x)}, \tau(t))$ is asymptotically normal with the same (asymptotic) parameters as in the pure random walk case.

We close this section by stating Theorem 2 of Larsson-Cohn (2000a). The result is the two-dimensional analog of Theorem 2.6 and, thus, provides the extension of Theorem 10.3 to a functional central limit theorem.

Theorem 10.5. *Let $a(y) \equiv 1$. Suppose that $\sigma_x^2 = \operatorname{Var} X_1 < \infty$, $\sigma_y^2 = \operatorname{Var} Y_1 < \infty$, and set*

$$Z_n^{(x)}(t, \omega) = \frac{Z_{\tau(nt, \omega)}^{(x)}(\omega) - \frac{nt\mu_x}{\mu_y}}{\gamma \mu_y^{-3/2} \sqrt{n}} \quad (t \geq 0),$$

where $\gamma^2 = \operatorname{Var}(\mu_y X_1 - \mu_x Y_1) > 0$. If

$$\frac{\max_{1 \leq k \leq n} \xi_k}{\sqrt{n}} \xrightarrow{P} 0 \quad \text{and} \quad \frac{\max_{1 \leq k \leq n} \xi_k}{\sqrt{n}} \xrightarrow{P} 0 \quad \text{as} \quad n \to \infty,$$

then

$$Z_n^{(x)} \xrightarrow{J_1} W \quad \text{as} \quad n \to \infty.$$

6.11 The Case $Z_n = n \cdot g(\bar{Y}_n)$

The present section is devoted to a brief discussion of two-dimensional analogs of the special case $Z_n = n \cdot g(\bar{Y}_n)$. In the following section we mention an application to repeated significance tests in exponential families, together with an example involving the normal distribution.

Thus, suppose that $X_1, X_2, \ldots$ and $Y_1, Y_2, \ldots$ are sequences of i.i.d. random variables with finite means θ_x and θ_y, respectively—no independence is assumed *between* the sequences. Suppose, further, that $g_x \in \mathcal{G}$ and $g_y \in \mathcal{G}$, and consider

$$(Z_n^{(x)},\, Z_n^{(y)}) = (n \cdot g_x(\bar{X}_n),\, n \cdot g_y(\bar{Y}_n)), \quad n \geq 1,$$

the first passage time process

$$\tau(t) = \min\{n : Z_n^{(y)} > t\} \quad (t \geq 0),$$

and the stopped family

$$\{Z_{\tau(t)}^{(x)}, \quad t \geq 0\}.$$

The strong law, the Marcinkiewicz–Zygmund strong law, and the central limit theorem are as follows.

Theorem 11.1. *We have*

$$\frac{Z_{\tau(t)}^{(x)}}{t} \xrightarrow{\text{a.s.}} \frac{g_x(\theta_x)}{g_y(\theta_y)} \quad as \quad t \to \infty.$$

Theorem 11.2. *Let $1 \leq r < 2$ and suppose that $E|X|^r < \infty$ and $E|Y|^r < \infty$. If $g_x' \in \mathrm{Lip}\,(1, \theta_x)$ and $g_y' \in \mathrm{Lip}\,(1, \theta_y)$, then*

$$\frac{Z_{\tau(t)}^{(x)} - \frac{g_x(\theta_x)}{g_y(\theta_y)}t}{t^{1/r}} \xrightarrow{\text{a.s.}} 0 \quad as \quad t \to \infty.$$

Theorem 11.3. *Suppose that $\mathrm{Var}\, X_1 < \infty$, $\mathrm{Var}\, Y_1 < \infty$, and that g_x' and g_y' are continuous at θ_x and θ_y, respectively. Then*

$$\frac{Z_{\tau(t)}^{(x)} - \frac{g_x(\theta_x)}{g_y(\theta_y)}t}{(g_y(\theta_y))^{-3/2}\gamma\sqrt{t}} \xrightarrow{d} N(0,1) \quad as \quad t \to \infty,$$

where

$$\gamma^2 = \mathrm{Var}\left(g_y(\theta_y)g_x'(\theta_x)X_1 - g_x(\theta_x)g_y'(\theta_y)Y_1\right)$$

is assumed to be positive.

We only prove Theorem 11.3.

Proof of Theorem 11.3. Taylor expansion of the components shows that $\{(Z_n^{(x)}, Z_n^{(y)}), \; n \geq 1\}$ as defined is, indeed, a two-dimensional perturbed random walk satisfying the conditions of Theorem 10.3, where X_k there corresponds to $g_x(\theta_x) + g'_x(\theta_x)(X_k - \theta_x)$ here, μ_x there to $g_x(\theta_x)$ here, and σ_x^2 there to $(g'_x(\theta_x))^2 \mathrm{Var}\, X_1$ here. Similarly for the other random walk component. The remainders in the expansion correspond to the perturbing parts. Finally, γ^2 there corresponds to the expression in the statement above. These facts all follow from what we already have proved in the one-dimensional case earlier in this chapter, see in particular Lemma 3.4. We omit the details.

The conclusion now follows from Theorem 10.3. □

An interesting variant is when the second component is two-dimensional; this is the case in our application below.

Suppose that $\{X_k, \; k \geq 1\}$ and g_x are as before and that $(Y_k^{(1)}, Y_k^{(2)}), \; k \geq 1$, are i.i.d. two-dimensional random variables with mean vector $(\theta_y^{(1)}, \theta_y^{(2)})$, that $g_y(y_1, y_2)$ is positive and continuous in a neighborhood of $(\theta_y^{(1)}, \theta_y^{(2)})$ and non-negative elsewhere, and consider

$$(Z_n^{(x)}, Z_n^{(y)}) = (n \cdot g_x(\bar{X}_n), \; n \cdot g_y(\bar{Y}_n^{(1)}, \bar{Y}_n^{(2)})), \quad n \geq 1,$$

the first passage time process

$$\tau(t) = \min\{n : Z_n^{(y)} > t\} \quad (t \geq 0),$$

and the stopped family

$$\{Z_{\tau(t)}^{(x)}, \quad t \geq 0\}.$$

The generalizations of Theorems 11.1 and 11.3 run as follows.

Theorem 11.4. *We have*

$$\frac{Z_{\tau(t)}^{(x)}}{t} \xrightarrow{\text{a.s.}} \frac{g_x(\theta_x)}{g_y(\theta_y^{(1)}, \theta_y^{(2)})} \quad as \quad t \to \infty.$$

Theorem 11.5. *Suppose that* $\mathrm{Var}\, X_1 < \infty$, *Cov* $\mathbf{Y}$ *is positive definite and that* g'_x, $\frac{\partial g_y}{\partial y_1}$ *and* $\frac{\partial g_y}{\partial y_2}$ *are continuous at* θ_x *and* $(\theta_y^{(1)}, \theta_y^{(2)})$, *respectively. Then*

$$\frac{Z_{\tau(t)}^{(x)} - \frac{g_x(\theta_x)}{g_y(\theta_y^{(1)}, \theta_y^{(2)})} t}{\gamma \big(g_y(\theta_y^{(1)}, \theta_y^{(2)})\big)^{-3/2} \sqrt{t}} \xrightarrow{d} N(0, 1) \quad as \quad t \to \infty,$$

where

$$\gamma^2 = \mathrm{Var} \bigg(g_y(\theta_y^{(1)}, \theta_y^{(2)}) g'_x(\theta_x) X_1$$

$$- g_x(\theta_x) \bigg\{ \frac{\partial g_y}{\partial y_1}(\theta_y^{(1)}, \theta_y^{(2)}) Y_1^{(1)} + \frac{\partial g_y}{\partial y_2}(\theta_y^{(1)}, \theta_y^{(2)}) Y_1^{(2)} \bigg\} \bigg)$$

is assumed to be positive.

The results are proved along the lines of the previous ones, the essential difference being that Taylor expansion in the second component introduces partial derivatives.

Remark 11.1. If, in particular, $X_k = 1$ a.s. for all k the results yield asymptotics for $\tau(t)$ as $t \to \infty$.

Remark 11.2. One can also prove multidimensional versions analogous to those cited in Remark 10.1.

6.12 An Application

The application below is related to repeated significance tests in two-parameter exponential families. Interestingly enough it is not merely an extension from the one-parameter family to the two-parameter family case; the results can be used to provide additional insights, in particular relations between marginal one-parameter tests and joint tests. For the one-dimensional case we refer back to Section 6.9 and references given there, and for the multidimensional case to Woodroofe (1982), Chapter 8, and Siegmund (1985). Woodroofe (1978), Lalley (1983) and Hu (1988) treat related aspects of the model. Gut and Schwabe (1996, 1999) consider some statistical aspects that are consequences of the results below. Namely, it may for example happen that, whereas the two-dimensional test statistic falls into the (two-dimensional) critical region, none of the (one-dimensional) marginal test statistics fall into theirs, which means that one may conclude that "something is wrong somewhere" but not where or what.

In order to put this into mathematics, consider the family of distributions

$$G_{\theta_1,\theta_2}(dy_1, dy_2) = \exp\{\theta_1 y_1 + \theta_2 y_2 - \psi(\theta_1, \theta_2)\}\lambda(dy_1, dy_2), \quad (\theta_1, \theta_2) \in \Theta,$$

where λ is a nondegenerate, σ-finite measure on $\mathbb{R}^2$, Θ a convex subset of $\mathbb{R}^2$ and ψ is, for simplicity, strictly convex and twice differentiable. The corresponding moment generating function equals

$$E_\theta \exp\{t_1 Y_1 + t_2 Y_2\} = \exp\{\psi(\theta_1 + t_1, \theta_2 + t_2) - \psi(\theta_1, \theta_2)\},$$

and the corresponding mean vector equals

$$E\begin{pmatrix} Y_1 \\ Y_2 \end{pmatrix} = \begin{pmatrix} \frac{\partial \psi}{\partial \theta_1}(\theta_1, \theta_2) \\ \frac{\partial \psi}{\partial \theta_2}(\theta_1, \theta_2) \end{pmatrix}.$$

Let $(Y_k^{(1)}, Y_k^{(2)})$, $k \geq 1$, be i.i.d. two-dimensional random variables with distribution function G_{θ_1,θ_2}, where the parameters are unknown. Suppose we wish to test the hypothesis

$$H_0 : \theta_1 = \theta_{01}, \, \theta_2 = \theta_{02} \quad \text{vs.} \quad H_1 : \theta_1 \neq \theta_{01} \text{ or } \theta_2 \neq \theta_{02},$$

where w.l.o.g. we assume that $(\theta_{01}, \theta_{02}) = (0, 0) \in \Theta$ and that $\psi(0, 0) = \frac{\partial \psi}{\partial \theta_1}(0, 0) = \frac{\partial \psi}{\partial \theta_2}(0, 0) = 0$ (cf. Example 9.7 for the univariate case).

The log-likelihood ratio is

$$
\begin{aligned}
T_n &= \sup_{(\theta_1,\theta_2)\in\Theta} \log \prod_{k=1}^{n} \exp\{\theta_1 Y_k^{(1)} + \theta_2 Y_k^{(2)} - \psi(\theta_1,\theta_2)\} \\
&= \sup_{(\theta_1,\theta_2)\in\Theta} \left\{ \theta_1 \sum_{k=1}^{n} Y_k^{(1)} + \theta_2 \sum_{k=1}^{n} Y_k^{(2)} - n\psi(\theta_1,\theta_2) \right\} \\
&= n \cdot \sup_{(\theta_1,\theta_2)\in\Theta} \{\theta_1 \bar{Y}_n^{(1)} + \theta_2 \bar{Y}_n^{(2)} - \psi(\theta_1,\theta_2)\} \\
&= n \cdot g\big(\bar{Y}_n^{(1)}, \bar{Y}_n^{(2)}\big),
\end{aligned}
$$

where

$$
g(y_1,y_2) = \sup_{\theta_1,\theta_2} \big(\theta_1 y_1 + \theta_2 y_2 - \psi(\theta_1,\theta_2)\big), \quad -\infty < y_1, y_2 < \infty,
$$

is the convex conjugate of ψ. Moreover, g is strictly convex and twice continuously differentiable. It follows that $\{T_n, n \geq 1\}$ is a perturbed random walk of the special form considered above. Note also that all traditional assumptions are satisfied.

For sequential test procedures a typical object of interest is

$$
\tau = \tau(t) = \min\{n : T_n > t\} \quad (t > 0),
$$

for which asymptotics can be obtained from above.

However, now we may, in addition, consider T_n, $n \geq 1$, as a second component of a two-dimensional perturbed random walk as treated in Section 6.11 and apply our results from there. As a first component we might, for example, consider $\sum_{k=1}^{n} Y_k^{(1)}$ and $\sum_{k=1}^{n} Y_k^{(2)}$, $n \geq 1$, or, via the multidimensional versions of the results above, both of them jointly, in order to obtain asymptotics about the size of the marginal sums at the time of rejection of the two-dimensional null hypothesis. Another possible limit theorem concerns the joint asymptotics of the stopping time and one (or both) of $\sum_{k=1}^{n} Y_k^{(1)}$ and $\sum_{k=1}^{n} Y_k^{(2)}$, $n \geq 1$.

Example 12.1. The natural and simplest first example normally is the normal distribution. Suppose that $(Y_k^{(1)}, Y_k^{(2)})$, $k \geq 1$, are i.i.d. with independent normal components with means θ_1 and θ_2, respectively, and common variance 1. It follows that $\psi(\theta_1,\theta_2) = \frac{1}{2}(\theta_1^2 + \theta_2^2)$, that $g(y_1,y_2) = \frac{1}{2}(y_1^2 + y_2^2)$ and, hence, that

$$
T_n = \frac{n}{2}\left(\big(\bar{Y}_n^{(1)}\big)^2 + \big(\bar{Y}_n^{(2)}\big)^2 \right) = \frac{1}{2n}\left(\left(\sum_{k=1}^{n} Y_k^{(1)}\right)^2 + \left(\sum_{k=1}^{n} Y_k^{(2)}\right)^2 \right)
$$

$$
= \frac{1}{2n}\big((\Sigma_n^{(1)})^2 + (\Sigma_n^{(2)})^2\big),
$$

which is a natural analog of (9.2) in the one-dimensional case.

With $\mathbf{\Sigma}_n = (\Sigma_n^{(1)}, \Sigma_n^{(2)})'$ and $\| \cdot \|$ denoting Euclidean distance in $\mathbb{R}^2$, the relevant stopping time becomes

$$\tau(t) = \min\{n : \|\mathbf{\Sigma}_n\| > \sqrt{2tn}\} \quad (t \geq 0),$$

which might be interpreted as a generalization of the square root boundary problem.

Finally, here are some conclusions under relevant alternatives concerning this setup, which might be used for comparisons between marginal and two-dimensional tests.

1. By Theorem 11.4 with $X_k = 1$ a.s. for all k we have $g(x) \equiv 1$ and

$$\frac{\tau(t)}{t} \xrightarrow{\text{a.s.}} \frac{2}{\theta_1^2 + \theta_2^2} \quad \text{as} \quad t \to \infty.$$

2. The corresponding strong laws for the marginal tests are

$$\frac{\tau_i(t)}{t} \xrightarrow{\text{a.s.}} \frac{2}{\theta_i^2} \quad \text{as} \quad t \to \infty, \quad i = 1, 2.$$

3. An application of Theorem 11.5 with $X_k = 1$ a.s. for all k tells us that

$$\frac{\tau(t) - \frac{2}{\theta_1^2 + \theta_2^2} t}{\frac{1}{\theta_1^2 + \theta_2^2} \sqrt{8t}} \xrightarrow{d} N(0,1) \quad \text{as} \quad t \to \infty.$$

4. For the marginal tests we quote from the end of Section 6.9,

$$\frac{\tau_i(t) - \frac{2}{\theta_i^2} t}{\frac{1}{\theta_i^2} \sqrt{8t}} \xrightarrow{d} N(0,1) \quad \text{as} \quad t \to \infty, \quad i = 1, 2.$$

5. With $X_k = Y_k^{(1)}$ for all k we have $g_x(x) = x$, which, together with Theorem 11.4, yields

$$\frac{\sum_{k=1}^{\tau(t)} Y_k^{(1)}}{t} \xrightarrow{\text{a.s.}} \frac{2\theta_1}{\theta_1^2 + \theta_2^2} \quad \text{as} \quad t \to \infty,$$

and, together with Theorem 11.5, yields

$$\frac{\sum_{k=1}^{\tau(t)} Y_k^{(1)} - \frac{2\theta_1}{\theta_1^2 + \theta_2^2} t}{\sqrt{\frac{2t}{\theta_1^2 + \theta_2^2}}} \xrightarrow{d} N(0,1) \quad \text{as} \quad t \to \infty.$$

6. Alternatively, if we let $T_{i,n}$, $i = 1, 2$, denote the marginal log-likelihood ratios, and set $g_x(x) = \frac{1}{2}x^2$, we obtain

$$\frac{T_{i,\tau(t)}}{t} \xrightarrow{\text{a.s.}} \frac{\theta_i^2}{\theta_1^2 + \theta_2^2} \quad \text{as} \quad t \to \infty, \quad i = 1, 2,$$

and, for $\theta_1\theta_2 \neq 0$,

$$\frac{T_{i,\tau(t)} - \frac{\theta_i^2}{\theta_1^2+\theta_2^2}t}{\frac{\theta_1\theta_2}{\theta_1^2+\theta_2^2}\sqrt{2t}} \xrightarrow{d} N(0,1) \quad \text{as} \quad t \to \infty, \quad i = 1,2.$$

We close by reminding the reader that statistical conclusions of the above results may be found in Gut and Schwabe (1996, 1999).

6.13 Remarks on Further Results and Extensions

The results in this book all concern stopped (perturbed) random walks and first passage times for (perturbed) random walks. Clearly,

- some additional problems have been considered in the literature;
- most results can be extended to more general models.

In this section we provide some glimpses into some of these matters.

The difference between two stopping times

The perturbed random walk $\{Z_n, n \geq 1\}$ contains two parts; the random walk itself $\{S_n, n \geq 0\}$ and the perturbation $\{\xi_n, n \geq 1\}$. We have seen that the results for the random walk case in earlier chapters remain true for perturbed random walks under suitable smallness conditions on the perturbation. The following natural question therefore arises: How large is the difference between the first passage times of a random walk and those of its perturbed version?

Technically, in addition to the first passage time process $\{\nu(t), t \geq 0\}$ for the perturbed random walk, where in the following $a(y) \equiv 1$, we introduce the first passage times

$$\tau(t) = \min\{n : S_n > t\} \quad (t \geq 0),$$

where, as before, the increments $X_1, X_2, \ldots$ of the random walk have positive, finite mean μ. The question we thus address here is to find estimates or limit theorems for $\nu(t) - \tau(t)$. This problem was investigated in Larsson-Cohn (2001b) and in Alsmeyer (2001b).

In addition to the usual setup it is also assumed that ξ_n is independent of $\{X_k, k > n\}$ for all n—recall (1.4)—and note that this is automatically satisfied in the case $Z_n = n \cdot g(\bar{Y}_n)$ in view of (1.6)–(1.7).

The following weak law is due to Larsson-Cohn (2001b), Theorem 3.1.

Theorem 13.1. *Let ψ be regularly varying with exponent ρ $(0 \leq \rho < 1)$, and such that $\psi(x) \to \infty$ as $x \to \infty$. Furthermore, suppose that*

$$\frac{\xi_n}{\psi(n)} \xrightarrow{\text{a.s.}} 0 \quad \text{as} \quad n \to \infty,$$

or that

$$\frac{\xi_n}{\psi(n)} \xrightarrow{p} 0 \quad as \quad n \to \infty \quad and$$

$$\left\{ \frac{\xi_n}{\psi(n)}, \, n \geq 1 \right\} \quad satisfies \; Anscombe's \; condition.$$

Then

$$\frac{\nu(t) - \tau(t)}{\psi(t)} \xrightarrow{p} 0 \quad as \quad t \to \infty.$$

The corresponding results for the case $Z_n = n \cdot g(\bar{Y}_n)$ is Larsson-Cohn (2001b), Theorem 5.1.

Theorem 13.2. (a) *Let* $0 < \rho < 1$ *and suppose that* $E|Y|^{\frac{2}{1+\rho}} < \infty$. *Then*

$$\frac{\nu(t) - \tau(t)}{t^\rho} \xrightarrow{p} 0 \quad as \quad t \to \infty. \tag{13.1}$$

In particular, if $E|Y|^r < \infty$ *for every* $r < 2$, *then (13.1) holds for all* $\rho > 0$.

(b) *If* $EY^2 < \infty$ *and* ψ *is slowly varying and such that*

$$\frac{\psi(n)}{\log\log n} \to \infty \quad as \quad n \to \infty,$$

then

$$\frac{\nu(t) - \tau(t)}{\psi(t)} \xrightarrow{p} 0 \quad as \quad t \to \infty.$$

As for strong laws it follows immediately from Theorems 3.4.4 and 2.2 that, if $E|X|^r < \infty$ for some $r \in [1, 2)$, and $\xi_n/n^{1/r} \xrightarrow{\text{a.s.}} 0$ as $n \to \infty$, then

$$\frac{\nu(t) - \tau(t)}{t^{1/r}} \xrightarrow{\text{a.s.}} 0 \quad as \quad t \to \infty.$$

For further results in the general case we refer to Larsson-Cohn (2001b). The following strong law for the case $Z_n = n \cdot g(\bar{Y}_n)$ is Larsson-Cohn (2001b), Theorem 5.4.

Theorem 13.3. *Let* $0 < \rho < 1$ *and suppose that* $E|Y|^r < \infty$ *for some* r *such that*

$$\begin{cases} r > \frac{2(1-\rho)}{\rho}, & when \quad 0 < \rho \leq \frac{1}{2}, \\ r = \frac{1}{\rho}, & when \quad \rho = \frac{1}{2}. \end{cases}$$

Then

$$\frac{\nu(t) - \tau(t)}{t^\rho} \xrightarrow{\text{a.s.}} 0 \quad as \quad t \to \infty.$$

Remark 13.1. Note, however, that $r \geq 2$, that is, finite variance is assumed, in all cases.

In addition to these results, Alsmeyer (2001b) proved the following results concerning distributional convergence in the nonarithmetic case under the usual conditions in nonlinear renewal theory mentioned in Section 6.1. However, before stating them we need the following piece of notation.

Suppose that the random walk is nonarithmetic. Let T_1 denote the first strong ascending ladder epoch and, thus, S_{T_1} the first strong ascending ladder height (Section 2.9). Since $0 < \mu = E X < \infty$ we know from Chapter 3 that $E S_{T_1} < \infty$, and that

$$R(t) = S_{\tau(t)} - t \xrightarrow{d} Z \quad \text{as} \quad t \to \infty,$$

where

$$F_Z(x) = \frac{1}{E S_{T_1}} \int_{-\infty}^{x} (1 - F_{S_{T_1}}(y)) \, dy.$$

Finally, set

$$G_a(x) = 1 - F_Z(a) + \int_{0}^{\infty} P(\tau(a - y) \leq x) \, dF_Z(y).$$

Theorem 13.4. *Suppose that the random walk is nonarithmetic, that conditions (1.2)–(1.4) are satisfied, and, furthermore, that*

$$\xi_n \xrightarrow{d} \xi \quad as \quad n \to \infty.$$

Then

$$\tau(t) - \nu(t) \xrightarrow{d} V \text{ as } t \to \infty, \quad where \quad F_V(x) = \int_{-\infty}^{\infty} G_a(x) \, dF_\xi(a).$$

Theorem 13.5. *Suppose that the random walk is nonarithmetic, that conditions (1.2)–(1.4) are satisfied, and, let ψ be regularly varying with exponent ρ $(0 \leq \rho < 1)$, and such that $\psi(x) \to \infty$ as $x \to \infty$. If*

$$\frac{\xi_n}{\psi(n)} \xrightarrow{d} \xi \quad as \quad n \to \infty,$$

where F_ξ is continuous at 0, then

$$\frac{\mu}{\psi(t/\mu)} \big(\tau(t) - \nu(t)\big) \xrightarrow{d} \xi \quad as \quad t \to \infty.$$

Erdős-Rényi laws

The classical limit laws concern *partial sums*, typically of i.i.d. random variables. The so-called Erdős-Rényi laws focus on the increments or windows of such processes. The name stems from the seminal paper Erdős and Rényi

(1970). In the setting above this means that the interest centers around quantities such as

$$\max_{1\leq j\leq n-k}(S_{j+k} - S_j) \quad \text{and} \quad \max_{1\leq j\leq n-k}\max_{1\leq i\leq k}(S_{j+i} - S_i),$$

where, for example, in the first expression k may or may not depend on n.

As an example we quote Theorem 1 from Steinebach (1986) where an analog is proved for first passage time processes.

Theorem 13.6. *Let $X_1, X_2, \ldots$ be i.i.d. random variables with positive mean μ and positive, finite variance σ^2, and suppose that, for some $s_1 \in [-\infty, 0)$,*

$$\psi(t) = E\exp\{sX_1\} < \infty \quad \text{for} \quad s_1 < s \leq 0.$$

Set

$$I = \left\{\frac{\psi'(s)}{\psi(s)} : s_1 < s < 0\right\} \quad \text{and} \quad \rho(1/a) = \inf_s(\psi(s)\exp\{-s/a\}).$$

Finally, if $0 < 1/a < \mu$ with $1/a \in I$, and $C(a)$ is such that $\exp\{-1/C(a)\} = \rho^a(1/a)$, then, for the first passage time process $\{\nu(t),\, t \geq 0\}$, we have

$$\left|\frac{\sup_{0\leq t\leq T-K}(\nu(t+K) - \nu(t))}{C(a)\log T} - a\right| = \mathcal{O}\left(\frac{\log\log T}{\log T}\right) \quad \text{as} \quad T \to \infty,$$

i.e.,

$$\sup_{0\leq t\leq T-K}(\nu(t+K) - \nu(t)) - aC(a)\log T = \mathcal{O}(\log\log T) \quad \text{as} \quad T \to \infty.$$

Strong invariance principles

Donsker's theorem and its extensions, so-called *weak invariance principles* that were discussed in Chapter 5, tell us that summation processes and first passage time processes can be approximated by Brownian motion (Wiener processes) or stable processes in a distributional sense; the results deal with convergence in distribution, in this context called *weak* convergence. Another means of approximation is pathwise approximation, which is called *strong* approximation. Such results are therefore called *strong invariance principles*.

Two path breaking papers in this context are Komlós, Major and Tusnády (1975, 1976) (see also e.g. Csörgő and Révész (1981)) who provided strong approximation results for sums of i.i.d. random variables with mean zero and a finite moment generating function—later under certain moment conditions. They proved that

$$\max_{0\leq k\leq n}|S_n - W(n)| \overset{\text{a.s.}}{=} o(g(n)) \quad \text{as} \quad n \to \infty,$$

where S_n, $n \geq 1$, are partial sums, $W(t)$, $t \geq 0$, is a suitably rescaled Wiener process, and g some function depending on the integrability of the original summation process.

The analog for counting processes (first passage times) was obtained by Csörgő, Horváth and Steinebach (1987), who proved e.g. the following result.

Theorem 13.7. *Suppose that $X_1, X_2, \ldots$ be i.i.d. random variables with positive mean μ, positive, finite variance σ^2, and $E|X_1|^r < \infty$ for some $r > 2$. Let, as before, $\{\nu(t),\, t \geq 0\}$ be the associated first passage time process. Then one can define a standard Wiener process $\{W(t),\, t \geq 0\}$, such that*

$$\sup_{0 \leq t \leq n} |\nu(t) - W(t)| \overset{\text{a.s.}}{=} o(n^{1/r}) \quad as \quad n \to \infty.$$

We also remark that the authors obtain results under assumptions weaker than i.i.d. summands.

A strong approximation theorem in nonlinear renewal theory is given in Horváth (1985).

Higher order expansions

After strong laws one has the central limit theorem and Berry-Esseen type results (cf. e.g. Gut (2007), Section 7.6). Higher order expansions for sums are often treated under the heading Edgeworth expansions. Analogous result also exist in the renewal theoretic context. We confine ourselves at this point to referring to Siegmund (1985).

Local limit theorems

Local limit theorems, i.e., results concerning the approximation of $P(\nu(t) = n)$ with a normal density do not seem to arouse a great amount of interest. In addition to the references in Section 3.5 we mention Wang (1992).

Processes with independent stationary increments

There exist natural extensions to first passage times of stochastic processes with independent stationary increments with positive drift; see e.g. Gut (1974a, 1975b, 1996, 1997).

Markov renewal theory

An abbreviated description of the setup in Markov renewal theory—for further details, see papers by Alsmeyer and coauthors in the list of references—consists of measurable spaces $(S, \mathcal{S})$ and $(Y, \mathcal{Y})$ with countably generated σ-fields and a transition kernel $\mathbb{P} : S \times (\mathcal{S} \otimes \mathcal{B} \otimes \mathcal{Y}) \to [0, 1]$, where $\mathcal{B}$ is the Borel-σ-algebra on $\mathbb{R}$. Further, $\{(M_n, X_n, Y_n),\, n \geq 0\}$ is an associated Markov chain, defined on a probability space $(\Omega, \mathcal{A}, P)$, with state space $S \times \mathbb{R} \times \mathcal{Y}$, i.e.,

$$P(M_{n+1} \in A, X_{n+1} \in B, Y_{n+1} \in C \mid M_n, X_n, Y_n)$$

$$= \mathbb{P}(M_n,\, A \times B \times C) \quad \text{a.s.}$$

for all $n \geq 0$ and $A \in \mathcal{S}, B \in \mathcal{B}, C \in \mathcal{Y}$. Thus $(M_{n+1}, X_{n+1}, Y_{n+1})$ depends on the past only through M_n. This produces a Markov chain.

Markov renewal theory deals with certain asymptotic properties of the *Markov random walk* $\{(M_n, S_n),\, n \geq 0\}$, where $S_n = X_0 + X_1 + \cdots + X_n$ for $n \geq 0$, and related processes. An important assumption is that of, what is called *Harris recurrence*, which induces a regenerative structure on the full sequence $\{(M_n, X_n, Y_n),\, n \geq 0\}$, which, in turn, conveniently generates a strictly increasing sequence of stopping times (with respect to an appropriate σ-algebra), called *regeneration times*. The essential point is that the regeneration times divide the process under consideration into one-dependent cycles (blocks), which are stationary (except for the first two).

Some of the results in Section 4.2 (Gut and Janson (1983)) have been extended to this setting in Alsmeyer and Gut (1999). An important point there was that the application to chromatographic processes (Subsection 4.3.1) was generalized to include more than two phases, which amounts to allowing for more than two layers in the river (Subsection 4.3.2).

In addition to the papers by Alsmeyer and coauthors in the list of references we mention Melfi (1992, 1994) and to Su (1993) for more on this topic.

Records and record times

In many applications, such as the strength of materials, flooding or "shocks" of various kinds, *extremes* rather than sums are of importance. A special case concerns the theory of *records*, that is, extremes at their first appearance. Records are by no means connected to renewal theory as we have encountered it here, but the technique of using stopped sums in order to prove laws of large numbers, central limit theorems, and so on has proved useful (see Gut (1990a, 2007)), the point being that "the number of records so far" and "the record times" are inverses of each other in the same way as a renewal process and its counting process. In this subsection we provide a brief hint on the procedure. For details we refer to the cited sources and further references given there.

Thus, let $X, X_1, X_2, \ldots$ be independent, identically distributed, and in order to avoid ties, continuous random variables. The *record times* are $L(1) = 1$ and, recursively,

$$L(n) = \min\{k : X_k > X_{L(n-1)}\}, \quad n \geq 2,$$

the *record values* are $X_{L(n)}$, $n \geq 1$, and the associated *counting process* $\{\mu(n),\, n \geq 1\}$ is

$$\mu(n) = \#\text{records among } X_1, X_2, \ldots, X_n = \max\{k : L(k) \leq n\}.$$

If we let

$$I_k = \begin{cases} 1, & \text{if } X_k \text{ is a record,} \\ 0, & \text{otherwise,} \end{cases}$$

it follows that $\mu(n) = \sum_{k=1}^n I_k$, $n \geq 1$, where, as one can show, the indicators are independent random variables, so that

$$m_n = E\,\mu(n) = \sum_{k=1}^n \frac{1}{k} = \log n + \gamma + o(1) \quad \text{as} \quad n \to \infty,$$

$$\text{Var}\,\mu(n) = \sum_{k=1}^n \frac{1}{k}\left(1 - \frac{1}{k}\right) = \log n + \gamma - \frac{\pi^2}{6} + o(1) \quad \text{as} \quad n \to \infty,$$

where $\gamma = 0.5772156649\ldots$ is Euler's constant. The original paper in this context is Rényi (1962), where the connection between the counting process and the record times was based on the inversion

$$\{L(n) \geq k\} = \{\mu(k) \leq n\}.$$

With the renewal context in mind it seems tempting to try our findings in Chapter 1; recall, however, that the indicators are independent but *not* identically distributed.

Here we illustrate the method via the law of large numbers.

Theorem 13.8. *We have*

$$\frac{\mu(n)}{\log n} \xrightarrow{\text{a.s.}} 1 \quad as \quad n \to \infty;$$

$$\frac{\log L(n)}{n} \xrightarrow{\text{a.s.}} 1 \quad as \quad n \to \infty.$$

Proof. Since (centered) indicators are uniformly bounded the conclusion for the counting process follows from the Kolmogorov convergence criterion and Kronecker's lemma (for details check Rényi (1962) or e.g. Gut (2007), page 308).

In order to prove the second half of the theorem we note that $L(n) \xrightarrow{\text{a.s.}} \infty$ as $n \to \infty$ (there is always room for another record), so that Theorem 1.2.1 tells us that

$$\frac{\mu(L(n))}{\log L(n)} \xrightarrow{\text{a.s.}} 1 \quad \text{as} \quad n \to \infty.$$

The nominator obviously corresponds to the usual stopped sum which means that the next step would be a kind of sandwich inequality (such as (3.3.2) or (2.4)). However, in this case,

$$\mu(L(n)) = n,$$

in other words, there is no overshoot or excess over the boundary. Inserting this into our previous relation yields

$$\frac{n}{\log L(n)} \xrightarrow{\text{a.s.}} 1 \quad \text{as} \quad n \to \infty,$$

and the conclusion follows.   □

6.14 Problems

1. Prove Theorem 7.5.
2. State (and prove) the theorems that emerge from the applications at the end of Section 6.9.
3. Prove Theorems 10.1 and 10.2.
4. Determine the limit distribution of $(Z^{(x)}_{\tau(t)}, \tau(t))$ mentioned in Remark 10.1.
5. Formulate (and prove) the missing Marcinkiewicz–Zygmund analog to Theorem 10.2 in Section 6.11.
6. Formulate (and prove) the analogs of Theorem 3.4 for Sections 6.10 and 6.11.
7. Formulate (and prove) the missing laws of the iterated logarithm in Section 6.11.
8. Formulate the analogs of Theorems 2.6 and 10.5 for the case $Z_n = n \cdot g(\bar{Y}_n)$.
9. Combine Theorem 13.4 and (3.6) in order to formulate the corresponding result for the case $Z_n = n \cdot g(\bar{Y}_n)$.

A

Some Facts from Probability Theory

A.1 Convergence of Moments. Uniform Integrability

In all probability courses we are told that convergence in probability and convergence in distribution do not imply that moments converge (even if they exist).

Example 1.1. The standard example is $\{X_n, n \geq 1\}$ defined by

$$P(X_n = 0) = 1 - \frac{1}{n} \quad \text{and} \quad P(X_n = n) = \frac{1}{n}. \tag{1.1}$$

Then $X_n \xrightarrow{p} 0$ as $n \to \infty$, but, for example,

$$EX_n \to 1 \quad \text{and} \quad \text{Var } X_n \to \infty \quad \text{as} \quad n \to \infty,$$

that is, the expected value converges, but not to the expected value of the limiting random variable, and the variance diverges.

The reason for this behavior is that the mass escapes to infinity in a forbidden way.

The adequate concept in this context is the notion of uniform integrability. A sequence of random variables, $\{X_n, n \geq 1\}$, is said to be *uniformly integrable* if

$$\lim_{\alpha \to \infty} E|X_n| I\{|X_n| > \alpha\} = 0 \quad \text{uniformly in } n. \tag{1.2}$$

It is now easy to see that the sequence defined by (1.1) is not uniformly integrable.

Another way to check uniform integrability is given by the following criterion (see e.g. Gut (2007), Theorem 5.4.1).

Lemma 1.1. *A sequence $\{Y_n, n \geq 1\}$ is uniformly integrable iff*

(i) $$\sup_n E|Y_n| < \infty.$$

(ii) *For every $\varepsilon > 0$ there exists $\delta > 0$, such that for all events A with $P(A) < \delta$ we have*

$$E|Y_n|I\{A\} < \varepsilon \quad \text{for all } n. \tag{1.3}$$

The following is an important result connecting uniform integrability and moment convergence.

Theorem 1.1. *Let $0 < r < \infty$, suppose that $E|X_n|^r < \infty$ for all n and that $X_n \xrightarrow{P} X$ as $n \to \infty$. The following are equivalent:*

(i) $X_n \to X$ *in L^r as $n \to \infty$;*

(ii) $E|X_n|^r \to E|X|^r < \infty$ *as $n \to \infty$;*

(iii) $\{|X_n|^r, n \geq 1\}$ *is uniformly integrable.*
    Furthermore, if $X_n \xrightarrow{P} X$ and one of (i)–(iii) hold, then

(iv) $E|X_n|^p \to E|X|^p$ *as $n \to \infty$ for all $p, 0 < p \leq r$.*

For proofs, see e.g. Gut (2007), Section 5.5, Loève (1977), pp. 165–166, where this result (apart from (iv)) is called an L^r-convergence theorem, or Chung (1974), Theorem 4.5.4.

Remark 1.1. The theorem remains obviously true if one assumes convergence almost surely instead of convergence in probability, but also (except for (i)) if one assumes convergence in distribution. For the latter, see Billingsley (1968), Theorem 5.4, cf. also Billingsley (1999), Theorems 3.5 and 3.6 and Gut (2007), Theorem 5.5.9.

Remark 1.2. The implications (iii) $\Longrightarrow$ (i), (ii), (iv) remain true for families of random variables.

Lemma 1.2. *Let U and V be positive random variables, such that $EU^r < \infty$ and $EV^r < \infty$ for some $r > 0$. Then*

$$E(U+V)^r I\{U+V > \alpha\} \leq 2^r EU^r I\left\{U > \frac{\alpha}{2}\right\} + 2^r EV^r I\left\{V > \frac{\alpha}{2}\right\} (\alpha > 0).$$

Proof. $E(U+V)^r I\{U+V > \alpha\} \leq E(\max\{2U, 2V\})^r I\{\max\{2U, 2V\} > \alpha\} \leq 2^r EU^r I\{U > \alpha/2\} + 2^r EV^r I\{V > \alpha/2\}.$ $\qquad\square$

Lemma 1.3. *Let $\{U_n, n \geq 1\}$ and $\{V_n, n \geq 1\}$ be sequences of positive random variables such that, for some $p > 0$,*

$$\{U_n^p, n \geq 1\} \quad and \quad \{V_n^p, n \geq 1\} \quad \text{are uniformly integrable.} \tag{1.4}$$

Then

$$\{(U_n + V_n)^p, n \geq 1\} \quad \text{is uniformly integrable.} \tag{1.5}$$

Proof. We have to show that

$$\lim_{\alpha \to \infty} E(U_n + V_n)^p I\{U_n + V_n > \alpha\} = 0 \quad \text{uniformly in } n. \qquad (1.6)$$

But this follows from Lemma 1.2 and (1.4).   □

Remark 1.3. Lemmas 1.2 and 1.3 can be extended to more than two random variables (sequences) in an obvious way.

A.2 Moment Inequalities for Martingales

The most technical convergence results are those where moments of stopped sums are involved, in particular, convergence of such moments. For those results we need some moment inequalities for martingales. The study of such inequalities was initiated in Burkholder (1966). Two further important references are Burkholder (1973) and Garsia (1973).

Just as martingales in some sense are generalizations of sequences of partial sums of independent random variables with mean zero, the martingale inequalities we consider below are generalizations of the corresponding inequalities for such sequences of partial sums of independent random variables obtained by Marcinkiewicz and Zygmund (1937, 1938).

As a preparation for the proofs of the moment convergence results for stopped random walks we proved moment convergence in the ordinary central limit theorem (see Theorem 1.4.2). Since the martingale inequalities below, in the proofs of moment convergence for stopped random walks, will play the role played by the Marcinkiewicz–Zygmund inequalities in the proof of moment convergence in the ordinary central limit theorem, we present the latter inequalities before we derive the former.

Finally, since there is equality for the first and second moments of stopped sums just as in the classical case (Theorem 1.5.3), we close by quoting an application of Doob's optional sampling theorem to stopped sums, which will be useful.

However, we begin, for completeness, with the definition of a martingale and the most basic convergence result.

Thus, let $(\Omega, \mathscr{F}, P)$ be a probability space, let $\{\mathscr{F}_n, n \geq 1\}$ be an increasing sequence of sub-σ-algebras of $\mathscr{F}$ and set $\mathscr{F}_\infty = \sigma\{\cup_{n=1}^\infty \mathscr{F}_n\}$ and $\mathscr{F}_0 = \{\emptyset, \Omega\}$. Further, let $\{X_n, n \geq 1\}$ be a sequence of integrable random variables.

We say that $\{(X_n, \mathscr{F}_n), n \geq 1\}$ is a *martingale* if

(i) $\{X_n, n \geq 1\}$ is adapted, that is, if X_n is $\mathscr{F}_n$-measurable for all n,
(ii) $E(X_n | \mathscr{F}_m) = X_m$ for $m \leq n$.

Recall that a simple example of a martingale is obtained by $X_n = \sum_{k=1}^n Y_k$, where $\{Y_k, k \geq 1\}$ are independent random variables with mean 0 and $\mathscr{F}_n = \sigma\{Y_k, k \leq n\}$.

The classical reference for martingale theory is Doob (1953). For more, see e.g. Chung (1974), Neveu (1975) and Gut (2007), Chapter 10.

We say that $\{(X_n, \mathscr{G}_n), n \geq 1\}$ is a *reversed martingale* if $\{\mathscr{G}_n, n \geq 1\}$ is a *decreasing* sequence of sub-σ-algebras of $\mathscr{F}$ and

(i) X_n is $\mathscr{G}_n$-measurable for all n,
(ii) $E(X_n | \mathscr{G}_m) = X_m$ for $n \leq m$.

Alternatively one can define a reversed martingale exactly like a martingale, but with the index set being the negative integers.

A common example of a reversed martingale is $X_n = (1/n) \sum_{k=1}^{n} Y_k$, where $\{Y_k, k \geq 1\}$ are i.i.d. random variables with finite mean and $\mathscr{G}_n = \sigma\{X_k, k \geq n\}$.

If the equality sign in (ii) of the definitions is replaced by $\geq$ ($\leq$) we have sub(super)martingales.

The classical martingale convergence theorems state that an L^1-bounded martingale is a.s. convergent and that all reversed martingales converge a.s. and in L^1.

The following is an important result on moment convergence for martingales (see e.g. Gut (2007), Theorem 10.12.1).

Theorem 2.1. *Let $\{(X_n, \mathscr{F}_n), n \geq 1\}$ be a martingale. The following are equivalent:*

(i) *$\{X_n, n \geq 1\}$ is uniformly integrable;*
(ii) *X_n converges in L^1 as $n \to \infty$;*
(iii) *$X_n \xrightarrow{\text{a.s.}} X_\infty$ as $n \to \infty$, where X_∞ is integrable and such that $\{(X_n, \mathscr{F}_n), 1 \leq n \leq \infty\}$ is a martingale;*
(iv) *there exists a random variable Y such that $X_n = E(Y|\mathscr{F}_n)$ for all n.*

A similar result is true for reversed martingales, but, since ever reversed martingale is, in fact, uniformly integrable, all corresponding conditions are automatically satisfied (check e.g. Gut (2007), Section 16).

Moment Inequalities for Sums of Independent Random Variables

Let $\{X_k, k \geq 1\}$ be a sequence of random variables with the same distribution, such that $E|X_1|^r < \infty$ for some $r \geq 1$, and let $\{S_n, n \geq 1\}$ be their partial sums. It is easy to see, by applying Minkowski's inequality, that

$$\|S_n\|_r \leq n\|X_1\|_r, \tag{2.1}$$

which gives an upper bound for $E|S_n|^r$ of the order of magnitude of n^r. Now, if, for example, the summands are i.i.d. with mean 0 and finite variance, σ^2, the central limit theorem states that

$$\frac{S_n}{\sqrt{n}} \xrightarrow{d} N(0,1) \quad \text{as} \quad n \to \infty.$$

In view of the previous section it would therefore follow that $E|S_n/\sqrt{n}|^r$ converges as $n \to \infty$ provided $\{S_n/\sqrt{n}|^r, n \geq 1\}$ is uniformly integrable, in which case $E|S_n|^r$ would be $\mathcal{O}(n^{r/2})$ instead of $\mathcal{O}(n^r)$. Below we shall see that this is, indeed, the correct order of magnitude when $r \geq 2$.

By starting with Khintchine's inequality and Rademacher functions Marcinkiewicz and Zygmund (1937, 1938) proved (without using the central limit theorem) the following celebrated inequalities (see also Gut (2007), Theorem 3.8.1):

Let $\{X_k, k \geq 1\}$ be independent (not necessarily identically distributed) random variables with mean 0 and a finite moment of some order $r \geq 1$. Let $\{S_n, n \geq 1\}$ denote the partial sums. Then there exist numerical constants A_r and B_r, depending on r only such that

$$A_r E \left(\sum_{k=1}^n X_k^2 \right)^{r/2} \leq E|S_n|^r \leq B_r E \left(\sum_{k=1}^n X_k^2 \right)^{r/2}. \qquad (2.2)$$

If, in particular, the summands are i.i.d. the right most inequality can be estimated by Minkowski's inequality if $r/2 \geq 1$ and by the c_r-inequality if $r/2 \leq 1$ and one obtains

$$E|S_n|^r \leq \begin{cases} B_r n E|X_1|^r, & \text{for } 1 \leq r \leq 2, \\ B_r n^{r/2} E|X_1|^r, & \text{for } r \geq 2. \end{cases} \qquad (2.3)$$

For $r = 2$ we have, of course, $ES_n^2 = nEX_1^2$ and, for $r = 1, E|S_n| \leq nE|X_1|$.

Note that (2.3), for $r \geq 2$, can be rewritten as $E|S_n/\sqrt{n}|^r \leq B_r \cdot E|X_1|^r$. Since, by Fatou's lemma,

$$\liminf_{n \to \infty} E|S_n/\sqrt{n}|^r \geq E|Z|^r,$$

where Z is normal with mean 0 and variance σ^2, we have established the correct *order of magnitude* for these moments. That they actually converge is proved in Theorem 1.4.2, by showing that the sequence $\{|S_n/\sqrt{n}|^r, n \geq 1\}$ is uniformly integrable.

Moment Inequalities for Martingales

We now turn our attention to martingale extensions of the Marcinkiewicz–Zygmund inequalities (2.2). The first extension is due to Burkholder (1966), Theorem 9, where it is shown that (2.2) remains true for $r > 1$, if $\{X_k, k \geq 1\}$ is a sequence of martingale differences and $\{S_n, n \geq 1\}$ is a martingale. Later Davis (1970) proved that the right hand inequality also holds for $r = 1$. More precisely, the following result is true.

Theorem 2.2. *Let $\{(Z_n, \mathscr{F}_n), n \geq 1\}$ be a martingale and set $Y_1 = Z_1$ and $Y_k = Z_k - Z_{k-1}$ for $k \geq 2$. There exist numerical constants A_r and B_r $(r \geq 1)$, depending on r only, such that*

(i) $A_r E \left(\sum_{k=1}^{n} Y_k^2 \right)^{r/2} \leq E|Z_n|^r \leq B_r E \left(\sum_{k=1}^{n} Y_k^2 \right)^{r/2}$ $(r > 1)$,

(ii) $E|Z_n| \leq B_1 E \left(\sum_{k=1}^{n} Y_k^2 \right)^{1/2}$ $(r = 1)$.

The application of Theorem 2.2 which will be of interest to us is when the martingale is a stopped sum of i.i.d. random variables with mean 0 and finite moment of order r $(1 < r \leq 2)$. For $r \geq 2$ we shall use the following inequality, which is a special case of Burkholder (1973), Theorem 21.1.

Theorem 2.3. *Let* $\{(Z_n, \mathscr{F}_n), n \geq 1\}$ *be a martingale with increments* $\{Y_k, k \geq 1\}$. *There exists a numerical constant* B_r $(r \geq 2)$, *depending on* r *only, such that*

$$E|Z_n|^r \leq B_r E \left(\sum_{k=1}^{n} E(Y_k^2|\mathscr{F}_{k-1}) \right)^{r/2} + B_r E \left(\sup_{1 \leq k \leq n} |Y_k|^r \right). \qquad (2.4)$$

In Burkholder (1973) this inequality is proved for more general convex functions than $|x|^r$.

Remark 2.1. The quantities $(\sum_{k=1}^{n} Y_k^2)^{1/2}$ and $(\sum_{k=1}^{n} E(Y_k^2|\mathscr{F}_{k-1}))^{1/2}$ are called the square function and the conditional square function, respectively, and they play an important role in martingale theory as well as in other branches of mathematics where martingale theory is used. We observe that Theorem 2.2 provides us with a relation between the moments or order r of the martingale and the square function and Theorem 2.3 gives an upper bound for the moments of order r of the martingale which involves the moment of order r of the conditional square function.

An Application of Doob's Optional Sampling Theorem

Finally, suppose that $\{X_k, k \geq 1\}$ are i.i.d. random variables with mean 0, which are adapted to a sequence of increasing sub-σ-algebras, $\{\mathscr{F}_k, k \geq 1\}$, of $\mathscr{F}$ and furthermore such that X_n is independent of $\mathscr{F}_{n-1}$ for all $n \geq 1$ and set $\mathscr{F}_0 = \{\emptyset, \Omega\}$—the typical case being $\mathscr{F}_n = \sigma\{X_1, \ldots, X_n\}$. Further, set $S_n = \sum_{k=1}^{n} X_k$, $n \geq 1$, and let τ be a *stopping time* relative to $\{\mathscr{F}_n, n \geq 1\}$, that is, let τ be a positive integer valued random variable, such that

$$\{\tau = n\} \in \mathscr{F}_n \quad (n \geq 1). \qquad (2.5)$$

Since $\{(S_n, \mathscr{F}_n), n \geq 1\}$ is a martingale we can apply Doob's optional sampling theorem (see e.g. Doob (1953), Section 7.2, Chung (1974), Section 9.3 or Gut (2007), Section 10.7), which yields the following result.

Theorem 2.4. *Under the above assumptions*

$$\{(S_{\tau \wedge n}, \mathscr{F}_n), n \geq 1\} \quad \text{is a martingale.} \qquad (2.6)$$

In particular,

$$ES_{\tau \wedge n} = 0 \quad \text{for all } n. \qquad (2.7)$$

A.3 Convergence of Probability Measures

In this section we shall review some results, which are used in Chapter 5. A basic reference is Billingsley (1968, 1999).

Definition 3.1. *Let S be a metric space, let $\mathscr{S}$ be the σ-algebra generated by the open sets and let $\{P_n, n \geq 1\}$ and P be probability measures on $\mathscr{S}$. We say that P_n converges weakly to P, denoted $P_n \Longrightarrow P$, if*

$$\int_S f \, dP_n \to \int_S f \, dP \tag{3.1}$$

for all $f \in C(S)$, where $C(S)$ is the set of bounded, continuous, real valued functions on S.

The first important case is when $(S, \mathscr{S}) = (C, \mathscr{C})$, where $C = C[0, 1]$ is the space of continuous functions on $[0, 1]$ with the uniform metric. In this metric we say that, if $\{x_n, n \geq 1\}$ and x are elements of C, then x_n is *U-convergent* to x,

$$x_n \to x(U) \quad \text{as} \quad n \to \infty, \tag{3.2}$$

if

$$\sup_{0 \leq t \leq 1} |x_n(t) - x(t)| \to 0 \quad \text{as} \quad n \to \infty. \tag{3.3}$$

When working with random elements, that is, with mappings from our given probability space $(\Omega, \mathscr{F}, P)$ into S, we say that a sequence $\{X_n, n \geq 1\}$ of random elements *converges in distribution* to the random element X,

$$X_n \Longrightarrow X \quad \text{as} \quad n \to \infty, \tag{3.4}$$

if the corresponding measures converge weakly in the sense of Definition 3.1.

Remark 3.1. Sometimes it will be convenient to add a letter above the arrow, which denotes the topology considered. (For the space C this is, in general, not necessary; see further below.) We shall also use the mixture $X_n \Longrightarrow P$ as $n \to \infty$.

Now, let $x \in C$ and define $\pi_{t_1,\ldots,t_k}$ as the mapping that carries x to $(x(t_1), \ldots, x(t_k)) \in R^k$ $(k \geq 1)$.

The sets of the form $\pi_{t_1,\ldots,t_k}^{-1} H$ with H a Borel set in R^k $(k \geq 1)$ constitute the *finite-dimensional* sets. Now, the convergence of all finite-dimensional distributions does not in general imply weak convergence. In order to find additional criteria for this to be the case we introduce the concept of tightness.

Definition 3.2. *Let S be a metric space. A family π of probability measures is tight if for every $\varepsilon > 0$ there exists a compact set K such that*

$$P\{K\} > 1 - \varepsilon \quad \text{for all } P \in \pi. \tag{3.5}$$

The following theorem then holds.

Theorem 3.1. *Let* $\{P_n, n \geq 1\}$ *and* P *be probability measures on* $(C, \mathscr{C})$. *If the finite-dimensional distributions of* P_n *converge to those of* P *and if* $\{P_n, n \geq 1\}$ *is tight, then* $P_n \Longrightarrow P$ *as* $n \to \infty$.

It still remains to find criteria which allow us to apply this result. One important case of interest in the context of the present book is the *functional central limit theorem* called Donsker's theorem, which we now present.

Let $\{\xi_k, k \geq 1\}$ be i.i.d. random variables defined on a common probability space $(\Omega, \mathscr{F}, P)$ and suppose that $E\xi_1 = 0$ and $\mathrm{Var}\xi_1 = \sigma^2 < \infty$. Further, set $S_n = \sum_{k=1}^{n} \xi_k$, $n \geq 0$, $(S_0 = 0)$ and define

$$X_n(t, \omega) = \frac{1}{\sigma\sqrt{n}} S_{[nt]}(\omega) + \frac{nt - [nt]}{\sigma\sqrt{n}} \xi_{[nt]+1}(\omega) \quad (0 \leq t \leq 1). \tag{3.6}$$

Theorem 3.2. $X_n \Longrightarrow W$ *as* $n \to \infty$, *where* $W = \{W(t), 0 \leq t \leq 1\}$ *is the Wiener measure on* $(C, \mathscr{C})$.

For the case when the variance is not necessarily finite a similar result will be stated below. In this situation, however, the limiting processes, the stable processes, no longer have continuous paths.

Definition 3.3. $D = D[0, 1]$ *is the space of functions on* $[0, 1]$ *that are right-continuous and have left-hand limits. Also,* $\mathscr{D}$ *denotes the* σ-*algebra generated by the open sets in the* J_1-*topology defined below. (The open sets in the* M_1-*topology defined below actually generate the same* σ-*algebra.)*

Remark 3.2. All discontinuities of elements of D are of the first kind.

Remark 3.3. C is a subset of D.

For the space D we shall define two different topologies; the J_1-topology and the M_1-topology, which were introduced in Skorohod (1956).

Definition 3.4. *Let* Λ *denote the class of strictly increasing, continuous mappings of* $[0, 1]$ *onto itself, such that* $\lambda(0) = 0$ *and* $\lambda(1) = 1$ *for all* $\lambda \in \Lambda$. *Suppose that* $\{x_n, n \geq 1\}$ *and* x *are elements of* D. *We say that* x_n *is* J_1-*convergent to* x,

$$x_n \to x(J_1) \quad as \quad n \to \infty, \tag{3.7}$$

if there exist $\{\lambda_n, n \geq 1\}$ *in* Λ, *such that*

$$\sup_{0 \leq t \leq 1} |\lambda_n(t) - t| \to 0 \quad and \quad \sup_{0 \leq t \leq 1} |x_n(\lambda_n(t)) - x(t)| \to 0 \quad as \ n \to \infty. \tag{3.8}$$

Definition 3.5. *Define the graph* $G(x)$ *as the subset of* $R \times [0, 1]$, *which contains all pairs* (x, t), *such that, for all* $t \in [0, 1]$, *the point* x *belongs to the segment joining* $x(t-)$ *and* $x(t)$, *that is*

$$G(x) = \{(x, t) : 0 \leq t \leq 1\} \quad and \quad x(t-) \leq x \leq x(t)\}. \tag{3.9}$$

$G(x)$ *is a continuous curve in* $R \times [0, 1]$. *The pair of functions* $(y(s), t(s))$ *gives a parametric representation of the graph if* $(y(s), t(s))$ *is a continuous* $1 - 1$ *mapping of* $[0, 1]$ *onto* $G(x)$, *such that* $t(s)$ *is nondecreasing.*

Definition 3.6. *Suppose that $\{x_n, n \geq 1\}$ and x are elements of D. We say that x_n is M_1-convergent to x,*

$$x_n \to x(M_1) \quad as \quad n \to \infty, \tag{3.10}$$

if there exist parametric representations $(y_n(s), t_n(s))$ of $G(x_n)$ and $(y(s), t(s))$ of $G(x)$, such that

$$\sup_{\{s:0 \leq s \leq 1\}} (|y_n(s) - y(s)| + |t_n(s) - t(s)|) \to 0 \quad as \quad n \to \infty. \tag{3.11}$$

Remark 3.4. The M_1-topology is weaker than the J_1-topology, which, in turn, is weaker than the U-topology. If the limit $x \in C$, the J_1-topology and the M_1-topology both reduce to the U-topology.

We are now ready to state the stable counterpart of Theorem 3.2. Let $\{\xi_k, k \geq 1\}$ be i.i.d. random variables with mean 0 and set $S_n = \sum_{k=1}^{n} \xi_k$, $n \geq 0$. Suppose that $\{B_n, n \geq 1\}$ are positive, normalizing coefficients such that

$$\frac{S_n}{B_n} \xrightarrow{d} G_\alpha \quad as \quad n \to \infty, \tag{3.12}$$

where G_α is (a random variable distributed according to) a stable law with index α $(1 < \alpha \leq 2)$ and define

$$X_n(t, \omega) = \frac{S_{[nt]}(\omega)}{B_n} \quad (0 \leq t \leq 1). \tag{3.13}$$

Theorem 3.3. $X_n \overset{J_1}{\Longrightarrow} X$ *as $n \to \infty$, where $X \in D[0, 1]$ is the stable process whose one-dimensional marginal distribution at $t = 1$ is G_α.*

In connection with this result we also refer to Gikhman and Skorohod (1969), Chapter IX.6.

Remark 3.5. When $\sigma^2 = \text{Var} \, \xi_1 < \infty$ we can use $B_n = \sigma \sqrt{n}$ in (3.13) and it follows that $X_n \overset{J_1}{\Longrightarrow} W$ as $n \to \infty$, but we can actually conclude that $X_n \overset{U}{\Longrightarrow} W$ as $n \to \infty$. In this case we may define X_n by (3.6) or by (3.13), see Remark 3.4.

The processes studied in Chapter 5 are all suitable functions of $\{X_n, n \geq 1\}$ as defined by (3.6) or (3.13). Instead of verifying that Theorem 3.1 is applicable we shall use so-called *continuous mapping theorems*.

Theorem 3.4. *Let S and S' be metric spaces, let $h : S \to S'$ be measurable and define*

$$D_h = the \; set \; of \; discontinuities \; of \; h. \tag{3.14}$$

If

$$P_n \Longrightarrow P \quad as \quad n \to \infty \quad and \; P(D_h) = 0, \tag{3.15}$$

then

$$P_n h^{-1} \Longrightarrow P h^{-1} \quad as \quad n \to \infty. \tag{3.16}$$

The corresponding counterpart for random elements is the following.

Theorem 3.5. *Let S, S', h and D_h be as before and suppose that $\{X_n, n \geq 1\}$ and X are random elements of S. If*

$$X_n \Longrightarrow X \quad as \quad n \to \infty \quad and \ P(X \in D_h) = 0, \tag{3.17}$$

then

$$h(X_n) \Longrightarrow h(X) \quad as \quad n \to \infty. \tag{3.18}$$

It is also possible to obtain such a result for a sequence of functions, $\{h_n, n \geq 1\}$, which are not necessarily the same. Following is the result corresponding to Theorem 3.5.

Theorem 3.6. *Let S and S' be separable metric spaces and suppose that $\{h_n, n \geq 1\}$ and h are measurable mappings from S to S'. If*

$$X_n \Longrightarrow X \quad as \quad n \to \infty \quad and \ P(X \in E) = 0, \tag{3.19}$$

where

$$E = \{x : h_n x_n \nrightarrow hx \ \textit{for some} \ \{x_n, n \geq 1\}, \ \textit{such that} \ x_n \to x\}, \tag{3.20}$$

then

$$h_n(X_n) \Longrightarrow h(X) \quad as \quad n \to \infty. \tag{3.21}$$

Remark 3.6. If $h_n = h$ $(n \geq 1)$ then $E = D_h$.

Remark 3.7. The spaces $(C, U), (D, J_1)$ and (D, M_1) are separable (whereas (D, U) is not).

If, for example, we wish to apply Theorem 3.5 to Theorem 3.2 it follows that, if h is continuous on C, then $h(X_n) \Longrightarrow h(W)$ as $n \to \infty$. If we cannot compute the limiting distribution of $h(W)$ directly, it may be possible to compute the limit of the distributions of $h(X_n)$ as $n \to \infty$ for some *special choice* of $\{\xi_k, k \geq 1\}$ and thus, since the theorem tells us that the limit of $h(X_n)$ is $h(W)$ *irrespectively* of the sequence $\{\xi_k, k \geq 1\}$, this provides us with a method to compute the distribution of $h(W)$. Theorems 3.2 and 3.3 are therefore sometimes called (*weak*) *invariance principles*.

The continuous mapping theorems just mentioned are related to convergence in distribution. Similarly one can state such theorems for other convergence modes. However, the following result, due to Skorohod (1956, §3) (see also Vervaat (1972a,b) and Whitt (1980), Section 1 and references given there) shows that it actually suffices to consider *deterministic* continuous mapping results.

Theorem 3.7. *Let $\{X_n, n \geq 1\}$ and X belong to a separable metric space and suppose that $X_n \Longrightarrow X$ as $n \to \infty$. Then there exists a probability space supporting the random elements $\{Y_n, n \geq 1\}$ and Y such that*

$$Y_n \overset{d}{=} X_n \quad (n \geq 1) \quad and \quad Y \overset{d}{=} X \tag{3.22}$$

and such that

$$Y_n \xrightarrow{\text{a.s.}} Y \quad as \quad n \to \infty. \tag{3.23}$$

Let us, as an example, show how this result can be used to prove Theorem 3.5.

Example 3.1. Suppose that h is a continuous functional and that $X_n \Longrightarrow X$ as $n \to \infty$. Then, since $Y_n \xrightarrow{\text{a.s.}} Y$ we have $h(Y_n) \xrightarrow{\text{a.s.}} h(Y)$ as $n \to \infty$ and hence $h(Y_n) \Longrightarrow h(Y)$ as $n \to \infty$. Finally, since $h(Y_n) \overset{d}{=} h(X_n)$ $(n \geq 1)$ and $h(Y) \overset{d}{=} h(X)$, we conclude that $h(X_n) \Longrightarrow h(X)$ as $n \to \infty$.

The continuity of a number of functions is investigated in Whitt (1980).

Remark 3.8. The above theory has been generalized to the space $D[0, \infty)$, see Lindvall (1973), as follows.

Let $\{X_n, n \geq 1\}$ and X be random elements in $(D[0, \infty), \mathscr{D}[0, \infty))$ and define, for $b > 0$, $r_b : D[0, \infty) \to D[0, b]$ by

$$(r_b x)(t) = x(t) \quad (t \leq b). \tag{3.24}$$

Then

$$X_n \overset{J_1}{\Longrightarrow} X \quad as \quad n \to \infty \quad \text{iff} \quad r_b X_n \overset{J_1}{\Longrightarrow} r_b X \quad as \quad n \to \infty, \tag{3.25}$$

for every b, such that $P(X(b) \neq X(b-)) = 0$. The same is true for the other two topologies discussed above. Thus, convergence in $C[0, \infty)$ is uniform convergence on compact sets. It is easily seen that the proofs of Theorems 3.2 and 3.3 carry over to $C[0, b]$ and $D[0, b]$ for any $b > 0$, that is, the results also hold true in $C[0, \infty)$ and $D[0, \infty)$, respectively. In our problems in Chapter 5 it is more natural and also easier to work in the latter spaces.

When considering Anscombe type results we shall use random time changes (cf. Billingsley (1968), Section 17, Billingsley (1999), Section 14). To do so we also need the space $D_0 = D_0[0, \infty)$, which consists of all elements φ of $D[0, \infty)$ that are nondecreasing and satisfy $\varphi(t) \geq 0$ $(t \geq 0)$. This space is topologized by the relativized J_1-topology.

For $x \in D$ and $\varphi \in D_0$ we let $x \circ \varphi$ denote the *composition* of x and φ, which is defined through the relation

$$(x \circ \varphi)(t) = x(\varphi(t)). \tag{3.26}$$

This is a measurable mapping $D \times D_0 \to D$.

We shall also use the following results, which involve the notion of convergence in probability, defined next.

Definition 3.7. *Let (S, ρ) be a metric space. If for some $a \in S$ we have*

$$P(\rho(X_n, a) \geq \varepsilon) \to 0 \quad as \quad n \to \infty \quad for \ all \ \varepsilon > 0, \tag{3.27}$$

we say that X_n converges in probability to a as $n \to \infty$ and write $X_n \xrightarrow{p} a$ as $n \to \infty$.

Theorem 3.8. *Let (S, ρ) be a separable, metric space and suppose that $\{X_n, n \geq 1\}, \{Y_n, n \geq 1\}$ and X are random elements of S, such that X_n and Y_n have the same domain for each n. If*

$$X_n \Longrightarrow X \quad and \quad \rho(X_n, Y_n) \xrightarrow{p} 0 \quad as \quad n \to \infty, \tag{3.28}$$

then

$$Y_n \Longrightarrow X \quad as \quad n \to \infty. \tag{3.29}$$

Theorem 3.9. *Let $\{X_n, n \geq 1\}$ and X be random elements of S' and let $\{Y_n, n \geq 1\}$ be random elements of S'', where S' and S'' are separable, metric spaces. Further, assume that X_n and Y_n have the same domain for each n. If*

$$X_n \Longrightarrow X \quad and \quad Y_n \xrightarrow{p} a \quad as \quad n \to \infty, \tag{3.30}$$

then

$$(X_n, Y_n) \Longrightarrow (X, a) \quad as \quad n \to \infty. \tag{3.31}$$

A.4 Strong Invariance Principles

Just as Donsker's theorem above extends the central limit theorem, the classical Hartman–Wintner law of the iterated logarithm has been extended by Strassen (1964), see also e.g. Stout (1974), Chapter 5.

Let $\{\xi_k, k \geq 1\}$ be i.i.d. random variables with mean 0 and (for convenience) variance 1 and define, for $n \geq 3$,

$$X_n(t, \omega) = \frac{1}{\sqrt{2n \log \log n}} S_{[nt]}(\omega) + \frac{nt - [nt]}{\sqrt{2n \log \log n}} \xi_{[nt]+1}(\omega) \quad (0 \leq t \leq 1), \tag{4.1}$$

where, as always, $S_n = \sum_{k=1}^{n} \xi_k, \ n \geq 0$.

The Hartman–Wintner (1941) law (cf. also Gut (2007), Theorem 8.1.2) states that

$$\limsup_{n \to \infty} \left(\liminf_{n \to \infty} \right) X_n(1) = \overset{+}{(-)} 1 \text{ a.s.,} \tag{4.2}$$

that is, the *extreme limit points* of $\{X_n(1), n \geq 3\}$ are $+1$ and -1.

A strengthening of this result is that

$$C(\{X_n(1), n \geq 3\}) = [-1, 1] \text{ a.s.,} \tag{4.3}$$

that is, the *set of limit points* of $\{X_n(1), n \geq 3\}$ are, in fact, *all points between the extreme limit points*, see e.g. Stout (1974), Section 5.3, De Acosta (1983), Theorem 2.5 or Gut (2007), Theorem 8.6.1. ($C\{x_n\}$ denotes the cluster set of $\{x_n\}$.)

The result below describes the set of limit points of $\{X_n(t)\ (0 \leq t \leq 1)$, $n \geq 3\}$. The method in Strassen (1964) is to prove such a result for Brownian motion and then to use a strong approximation result, which states that the *pathwise* behaviors of $X_n(t)\ (0 \leq t \leq 1)$ and of Brownian motion are "close." (Note the difference with Donsker's theorem which is a distributional result.) Theorem 4.1 below is therefore sometimes called a *strong invariance principle* or an *almost sure invariance principle*.

To describe the result we let

$$ K = \left\{ x \in AC[0,1] : x(0) = 0 \text{ and } \int_0^1 (x'(t))^2 dt \leq 1 \right\}, \qquad (4.4) $$

where AC denotes the set of absolutely continuous functions. If x is considered as being the motion of a particle of mass 2 from 0 to 1, then K is the set of motions starting at 0 and with kinetic energy at most equal to one.

Lemma 4.1. K *is compact* (*in the uniform topology*).

Proof. See Stout (1974), pp. 282–284. □

We are now ready to state Strassen's result.

Theorem 4.1. *With probability one the sequence* $\{X_n, n \geq 3\}$, *defined in* (4.1), *is relatively compact* (*in the uniform topology*) *and the set of its limit points coincides with* K.

Remark 4.1. Since the projections are continuous mappings (cf. above) we obtain (4.3) and the Hartman–Wintner law (4.2) as trivial corollaries. The reason we mention De Acosta (1983) above is that (to the best of our knowledge) this is the first place where a direct proof of (4.3) based on elementary methods only is given (that is, neither Strassen's result nor Skorohod embedding is used); see also Gut (2007), Theorem 8.6.1.

For more on the law of the iterated logarithm we refer to the survey paper by Bingham (1986).

Remark 4.2. For the extension to $D[0, \infty)$ we refer to Vervaat (1972a), (see also Vervaat (1972b)).

A.5 Problems

1. Show that the sequence given in (1.1) is not uniformly integrable.
2. Let $\{Y_n, n \geq 1\}$ be such that $\sup_n E|Y_n|^r < \infty$. Show that $\{|Y_n|^p, n \geq 1\}$ is uniformly integrable for all $p, 0 < p < r$.

3. Observe the strict inequality, $p < r$, in Problem 2. Compare with the sequence (1.1) and $r = 1$. Show that $\sup_n E|X_n| < \infty$ but that $\{|X_n|, n \geq 1\}$ is not uniformly integrable.

4. Prove Lemma 1.1.

5. Let $\{U_n, n \geq 1\}$ and $\{V_n, n \geq 1\}$ be nonnegative sequences of random variables, such that $U_n \leq V_n$ a.s. Show that if $\{V_n, n \geq 1\}$ is uniformly integrable, then so is $\{U_n, n \geq 1\}$.

6. Deduce (2.3) from (2.2).

7. Let $\{X_k, k \geq 1\}$ be independent random variables, such that $EX_k = \mu_k$ and $\operatorname{Var} X_k = \sigma_k^2$ exist, finite. Set $S_n = \sum_{k=1}^n X_k, m_n = \sum_{k=1}^n \mu_k$ and $s_n^2 = \sum_{k=1}^n \sigma_k^2$ and define $\mathscr{F}_n = \sigma\{X_1, \ldots, X_n\}$, $n \geq 0$. Prove that

$$\{((S_n - m_n)^2 - s_n^2, \mathscr{F}_n), n \geq 1\} \quad \text{is a martingale.} \tag{5.1}$$

8. Let $\{X_k, k \geq 1\}$ be a sequence of i.i.d. random variables with finite mean. Set $S_n = \sum_{k=1}^n X_k$, $n \geq 1$, and define, for $n \geq 1$, $\mathscr{G}_n = \sigma\{S_k, k \geq n\}$. Show that

$$\left\{\left(\frac{S_n}{n}, \mathscr{G}_n\right), n \geq 1\right\} \quad \text{is a reversed martingale.} \tag{5.2}$$

9. Let τ be a stopping time relative to an increasing sequence of σ-algebras and set $\tau_n = \tau \wedge n$. Show that τ_n is a stopping time.

10. Let τ_1 and τ_2 be stopping times relative to an increasing sequence of σ-algebras. Show that $\tau_1 \wedge \tau_2$ and $\tau_1 \vee \tau_2$ are stopping times.

11. What about Problem 10 when τ_1 and τ_2 are stopping times relative to different increasing sequences of σ-algebras?

12. Let $\{X_n, n \geq 1\}$ and X be real valued random variables.

   a) What does it mean to say that $\{X_n, n \geq 1\}$ is tight?
   b) Show that, if X_n converges in distribution to X as $n \to \infty$, then $\{X_n, n \geq 1\}$ is tight. (Compare with Theorem 3.1; in the real valued case tightness is automatic.)

13. Show that the J_1-topology is weaker than the U-topology, that is, let $\{x_n, n \geq 1\}$ and x be elements of D and show that $x_n \to x(U) \Longrightarrow x_n \to x(J_1)$.

14. Let $\{x_n, n \geq 1\}$ be elements of D and suppose that $x \in C$. Show that $x_n \to x(J_1) \Longleftrightarrow x_n \to x(U)$.

B

Some Facts about Regularly Varying Functions

B.1 Introduction and Definitions

Regular variation was introduced by Karamata (1930). The theory of regularly varying functions has proved to be important in many branches of probability theory. In this appendix we present definitions and a few facts that are needed in this book. For further reading we suggest Feller (1971), de Haan (1970), Seneta (1976) and Bingham et al. (1987). For a short introduction to regular variation one may also consult Gut (2007), Section A.7.

Definition 1.1. *A positive, measurable function U on $(0, \infty)$ is regularly varying (at infinity) with exponent ρ $(-\infty < \rho < \infty)$ if*

$$\frac{U(tx)}{U(t)} \to x^\rho \quad \text{as } t \to \infty \tag{1.1}$$

for all $x > 0$.

A function that is regularly varying with exponent 0 is called slowly varying.

Definition 1.2. *A positive, measurable function L on $(0, \infty)$ is slowly varying (at infinity) if*

$$\frac{L(tx)}{L(t)} \to 1 \quad \text{as } t \to \infty \tag{1.2}$$

for all $x > 0$.

Suppose U is regularly varying with exponent ρ. It is then easy to see that we can write

$$U(x) = x^\rho L(x), \tag{1.3}$$

where L is slowly varying.

Remark 1.1. One can also define regular and slow variation at some finite point a. We shall, however, only deal with regular and slow variation at infinity.

Typical examples of regularly varying functions are

$$x^\rho, \ x^\rho \log x, \ x^\rho \frac{\log x}{\log \log x}, \ \text{etc.}$$

Typical examples of slowly varying functions are

$$\log x, \ \log \log x, \ \text{etc.,}$$

but also functions having a finite, positive limit as $x \to \infty$, such as arc tan x.

Remark 1.2. We tacitly assume that the examples above have been modified so that they are positive for all $x > 0$.

B.2 Some Results

We now present some results for regularly and slowly varying functions. We give no proofs (apart from one exception); they can all be found in the main references above. Also, for the "typical examples" mentioned above the results are fairly trivial.

Lemma 2.1 (A Representation Formula). *A function L varies slowly iff it is of the form*

$$L(x) = c(x) \cdot \exp \left\{ \int_1^x \frac{\varepsilon(y)}{y} dy \right\}, \tag{2.1}$$

where $\varepsilon(x) \to 0$ and $c(x) \to c$ $(0 < c < \infty)$ as $x \to \infty$.

Remark 2.1. It follows from this result that if L varies slowly then

$$x^{-\varepsilon} < L(x) < x^\varepsilon \tag{2.2}$$

for any fixed $\varepsilon > 0$ and all sufficiently large x.

Next we note that the ratio between the arguments in the definitions is constant, x. It is an easy exercise to show that if L is slowly varying and $\{a_n, n \geq 1\}$ and $\{b_n, n \geq 1\}$ are sequences of positive reals, such that

$$\frac{a_n}{b_n} \to c \quad (0 < c < \infty) \quad \text{as } n \to \infty, \tag{2.3}$$

then

$$\frac{L(a_n x)}{L(b_n)} \to 1 \quad \text{as } n \to \infty. \tag{2.4}$$

A similar result holds for regularly varying functions.

The following result, due to Sreehari (1970), is a kind of converse.

Lemma 2.2. *Let $f(t)$ and $g(t)$ be real valued functions tending to infinity as $t \to \infty$ and let $L(t)$ be a slowly varying function. If*

$$\left(\frac{f(t)}{g(t)}\right)^p \cdot \frac{L(f(t))}{L(g(t))} \to 1 \quad as\ t \to \infty \quad for\ some\ p \neq 0 \tag{2.5}$$

then

$$\frac{f(t)}{g(t)} \to 1 \quad as\ t \to \infty. \tag{2.6}$$

Proof. Set $h_* = f \wedge g$ and $h^* = f \vee g$. We use the representation formula for slowly varying functions given in Lemma 2.1, which yields

$$\left(\frac{f(t)}{g(t)}\right)^p \cdot \frac{L(f(t))}{L(g(t))} = \left(\frac{f(t)}{g(t)}\right)^p \cdot \frac{c(f(t))}{c(g(t))} \cdot \exp\left\{\operatorname{sign}(f(t) - g(t)) \cdot \int_{h_*(t)}^{h^*(t)} \frac{\varepsilon(y)}{y} dy\right\}. \tag{2.7}$$

By the mean value theorem we have

$$\int_{h_*(t)}^{h^*(t)} \frac{\varepsilon(y)}{y} dy = \tilde{\varepsilon}(t) \log \frac{h^*(t)}{h_*(t)}, \tag{2.8}$$

where

$$\inf_{h_*(t) \leq s \leq h^*(t)} \varepsilon(s) \leq \tilde{\varepsilon}(t) \leq \sup_{h_*(t) \leq s \leq h^*(t)} \varepsilon(s).$$

Since the LHS in (2.7) tends to 1 and $c(f(t))/c(g(t)) \to 1$ as $t \to \infty$ it follows that

$$\left(\frac{f(t)}{g(t)}\right)^p \cdot \left(\frac{h^*(t)}{h_*(t)}\right)^{\tilde{\varepsilon}(t) \cdot \operatorname{sign}(f(t) - g(t))} \to 1 \quad as\ t \to \infty, \tag{2.9}$$

which is the same as

$$\left(\frac{f(t)}{g(t)}\right)^{p + \tilde{\varepsilon}(t)} \to 1 \quad as\ t \to \infty. \tag{2.10}$$

The conclusion now follows from the fact that $\tilde{\varepsilon}(t) \to 0$ as $t \to \infty$. $\qquad\square$

We finally need a limit theorem for the case when U has a density, u.

Lemma 2.3. *Suppose that $\rho \geq 0$ and that U has an ultimately monotone density u. Then*

$$\frac{tu(t)}{U(t)} \to \rho \quad as\ t \to \infty. \tag{2.11}$$

For the case $\rho = 0$ we also refer to Gut (1974a), Lemma 3.1(a).

References

1. Aleškevičiene, A. (1975): On the local limit theorem for the first passage time across a barrier (in Russian). *Litovsk. Mat. Sb.* **XV**, 23–66.
2. Alsmeyer, G. (1987): On the moments of certain first passage times for linear growth processes. *Stoch. Process. Appl.* **25**, 109–136.
3. Alsmeyer, G. (1988a): Second-order approximations for certain stopped sums in extended renewal theory. *Adv. Appl. Probab.* **20**, 391–410.
4. Alsmeyer, G. (1988b): On first and last exit times for curved differentiable boundaries. *Sequential Anal.* **7**, 345–362.
5. Alsmeyer, G. (1991): Some relations between harmonic renewal measures and certain first passage times. *Statist. Probab. Lett.* **12**, 19–27.
6. Alsmeyer, G. (1992): On generalized renewal measures and certain first passage times. *Ann. Probab.* **20**, 1229–1247.
7. Alsmeyer, G. (1994a): On the Markov renewal theorem. *Stoch. Process. Appl.* **50**, 37–56.
8. Alsmeyer, G. (1994b): Blackwell's renewal theorem for certain linear submartingales. *Acta Appl. Math.* **34**, 135–150.
9. Alsmeyer, G. (1994c): Random walks with stochastically bounded increments: Renewal theory via Fourier analysis. *Yokohama Math. J.* **42**, 1–21.
10. Alsmeyer, G. (1995): Random walks with stochastically bounded increments: Renewal theory. *Math. Nachr.* **175**, 13–31.
11. Alsmeyer, G. (1997): The Markov renewal theorem and related results. *Markov Proc. Rel. Fields* **3**, 103–127.
12. Alsmeyer, G. (2000): The ladder variables of a Markov random walk. *Prob. Math. Statist.* **20**, 151–168.
13. Alsmeyer, G. (2001a): Recurrence theorems for Markov random walks. *Prob. Math. Statist.* **21**, 123–134.
14. Alsmeyer, G. (2001b): Two comparison theorems for nonlinear first passage times and their linear counterparts. *Statist. Probab. Lett.* **55**, 163–171.
15. Alsmeyer, G. (2003): On the existence of moments of stopped sums in Markov renewal theory. *Prob. Math. Statist.* **23**, 389–411.
16. Alsmeyer, G. and Fuh, C.D. (2001): Limit theorems for iterated random functions by regenerative methods. *Stoch. Process. Appl.* **96**, 123–142. Corrigenda: *ibid.* **97**, 341–345 (2002).

17. Alsmeyer, G. and Gut, A. (1999): Limit theorems for stopped functionals of Markov renewal processes. *Ann. Inst. Statist. Math.* **51**, 369–382.
18. Alsmeyer, G. and Hoefs, V. (2001): Markov renewal theory for stationary *m*-block factors. *Markov Proc. Rel. Fields* **7**, 325–348.
19. Alsmeyer, G. and Hoefs, V. (2002): Markov renewal theory for stationary $(m + 1)$-block factors: Convergence rate results. *Stoch. Process. Appl.* **98**, 77–112.
20. Alsmeyer, G. and Hoefs, V. (2004): Markov renewal theory for stationary $(m + 1)$-block factors: first passage times and overshoot. *Semi-Markov Processes: Theory and Applications. Comm. Statist. — Theory Methods* **33**, 545–568.
21. Alsmeyer, G. and Irle, A. (1986): Asymptotic expansions for the variance of stopping times in nonlinear renewal theory. *Stoch. Process. Appl.* **23**, 235–258.
22. Anderson, K.K. and Athreya, K.B. (1987): A renewal theorem in the infinite mean case. *Ann. Probab.* **15**, 388–393.
23. Anscombe, F.J. (1952): Large-sample theory of sequential estimation. *Proc. Cambridge Philos. Soc.* **48**, 600–607.
24. Arjas, E., Nummelin, E. and Tweedie, R.L. (1978): Uniform limit theorems for non-singular renewal and Markov renewal processes. *J. Appl. Probab.* **15**, 112–125.
25. Aras, G. and Woodroofe, M. (1993): Asymptotic expansions for the moments of a randomly stopped average. *Ann. Statist.* **21**, 503–519.
26. Asmussen, S. (1982): Conditioned limit theorems relating a random walk to its associate, with applications to risk reserve processes and the GI/G/1 queue. *Adv. in Appl. Probab.* **14**, 143–170.
27. Asmussen, S. (1984): Approximations for the probability of ruin within finite time. *Scand. Actuar. J.*, 31–57; Correction, *ibid.* (1985), 64.
28. Asmussen, S. (2000): *Ruin Probabilities.* World Scientific Publishing, Singapore.
29. Asmussen, S. (2003): *Applied Probabilty and Queues*, 2nd ed. Springer-Verlag, New York.
30. Athreya, K., Mc Donald, D. and Ney, P. (1978): Coupling and the renewal theorem. *Amer. Math. Monthly* **85**, 809–814.
31. Athreya, K.B., Mc Donald, D. and Ney, P. (1978): Limit theorems for semi-Markov processes and renewal theory for Markov chains. *Ann. Probab.* **6**, 788–797.
32. Bacro, J.N., Deheuvels, P. and Steinebach, J. (1987): Exact convergence rates in Erdős-Rényi type theorems for renewal processes. *Annal. Institut Henri Poincaré (B) Probabilités et Statistiques* **23**, 195–207.
33. Barlow, R.E. and Proschan, F. (1965): *Mathematical Theory of Reliability.* John Wiley, New York.
34. Basu, A.K. (1972): Invariance theorems for first passage time random variables. *Canad. Math. Bull.* **15**, 171–176.
35. Baum, L.E. and Katz, M. (1965): Convergence rates in the law of large numbers. *Trans. Amer. Math. Soc.* **120**, 108–123.
36. Berbee, H.C.P. (1979): *Random Walks with Stationary Increments and Renewal Theory.* Mathematical Centre Tracts 112, Amsterdam.
37. Bickel, P.J. and Yahav, J.A. (1965): Renewal theory in the plane. *Ann. Math. Statist.* **36**, 946–955.

38. Billingsley, P. (1968): *Convergence of Probability Measures*. John Wiley, New York.

39. Billingsley, P. (1999): *Convergence of Probability Measures*, 2nd ed. John Wiley, New York.

40. Bingham, N.H. (1973): Maxima of sums of random variables and suprema of stable processes. *Z. Wahrsch. verw. Gebiete* **26**, 273–296.

41. Bingham, N.H. (1986): Variants of the law of the iterated logarithm. *Bull. London Math. Soc.* **18**, 433–467.

42. Bingham, N.H. and Goldie, C.M. (1982): Probabilistic and deterministic averaging. *Trans. Amer. Math. Soc.* **269**, 453–480.

43. Bingham, N.H., Goldie, C.M. and Teugels, J.L. (1987): *Regular Variation.* Cambridge University Press, Cambridge.

44. Blackwell, D. (1946): On an equation of Wald. *Ann. Math. Statist.* **17**, 84–87.

45. Blackwell, D. (1948): A renewal theorem. *Duke Math. J.* **15**, 145–150.

46. Blackwell, D. (1953): Extension of a renewal theorem. *Pacific J. Math.* **3**, 315–320.

47. Breen, S., Waterman, M.S. and Zhang, N. (1985): Renewal theory for several patterns. *J. Appl. Probab.* **22**, 228–234.

48. Breiman, L. (1965): First exit times from a square root boundary. In: *Proc. Fifth Berkeley Symp. Math. Statist. and Probability*, Vol. II/II, 9–16. University of California Press, Berkeley, CA.

49. Brown, B.M. (1969): Moments of a stopping rule related to the central limit theorem. *Ann. Math. Statist.* **40**, 1236–1249.

50. Burkholder, D.L. (1962): Successive conditional expectations of an integrable function. *Ann. Math. Statist.* **33**, 887–893.

51. Burkholder, D.L. (1966): Martingale transforms. *Ann. Math. Statist.* **37**, 1494–1504.

52. Burkholder, D.L. (1973): Distribution function inequalities for martingales. *Ann. Probab.* **1**, 19–42.

53. Carlsson, H. (1982): Error estimates in d-dimensional renewal theory. *Compositio Math.* **46**, 227–253.

54. Carlsson, H. (1983): Remainder term estimates of the renewal function. *Ann. Probab.* **11**, 143–157.

55. Carlsson, H. and Nerman, O. (1986): An alternative proof of Lorden's renewal inequality. *Adv. in Appl. Probab.* **18**, 1015–1016.

56. Carlsson, H. and Wainger, S. (1984): On the multi-dimensional renewal theorem. *J. Math. Anal. Appl.* **100**, 316–322.

57. Chang, I. and Hsiung, C. (1979): On the uniform integrability of $|b^{-1/p}W_{M_b}|^p, 0 < p < 2$. Preprint, NCU Chung-Li, Taiwan.

58. Chang, I.S. and Hsiung, C.A. (1983): Strassen's invariance principle for random subsequences. *Z. Wahrsch. verw. Gebiete* **64**, 401–409.

59. Chernoff, H. (1972): *Sequential Analysis and Optimal Design. CBMS-NSF Regional Conf. Ser. in Appl. Math.* **8**. SIAM, Philadelphia, PA.

60. Choquet, G. and Deny, J. (1960): Sur l'équation de convolution $\mu = \mu * \sigma$. *C.R. Acad. Sci. Paris* **250**, 799–801.

61. Chow, Y.S. (1966): On the moments of some one-sided stopping rules. *Ann. Math. Statist.* **37**, 382–387.

62. Chow, Y.S. and Hsiung, A. (1976): Limiting behaviour of $\max_{j \le n} S_j j^{-\alpha}$ and the first passage times in a random walk with positive drift. *Bull. Inst. Math. Acad. Sinica* **4**, 35–44.

63. Chow, Y.S., Hsiung, C.A. and Lai, T.L. (1979): Extended renewal theory and moment convergence in Anscombe's theorem. *Ann. Probab.* **7**, 304–318.

64. Chow, Y.S., Hsiung, C.A. and Yu, K.F. (1980): Limit theorems for a positively drifting process and its related first passage times. *Bull. Inst. Math. Acad. Sinica* **8**, 141–172.

65. Chow, Y.S., Hsiung, C.A. and Yu, K.F. (1983): A renewal theorem and its applications to some sequential procedures. *Z. Wahrsch. verw. Gebiete* **64**, 241–250.

66. Chow, Y.S. and Lai, T.L. (1979): Moments of ladder variables for driftless random walks. *Z. Wahrsch. verw. Gebiete* **48**, 253–257.

67. Chow, Y.S. and Robbins, H. (1963): A renewal theorem for random variables which are dependent or nonidentically distributed. *Ann. Math. Statist.* **34**, 390–395.

68. Chow, Y.S., Robbins, H. and Siegmund, D. (1971): *Great Expectations: The Theory of Optimal Stopping.* Houghton-Mifflin, Boston, MA.

69. Chow, Y.S., Robbins, H. and Teicher, H. (1965): Moments of randomly stopped sums. *Ann. Math. Statist.* **36**, 789–799.

70. Chow, Y.S. and Teicher, H. (1966): On second moments of stopping rules. *Ann. Math. Statist.* **37**, 388–392.

71. Chung, K.L. (1948): Asymptotic distribution of the maximum cumulative sum of independent random variables. *Bull. Amer. Math. Soc.* **54**, 1162–1170.

72. Chung, K.L. (1952): On the renewal theorem in higher dimensions. *Skand. Aktuarietidskr.* **35**, 188–194.

73. Chung, K.L. (1974): *A Course in Probability Theory*, 2nd ed. Academic Press, New York.

74. Chung, K.L. and Fuchs, W.H.J. (1951): On the distribution of values of sums of random variables. *Mem. Amer. Math. Soc.* **6**.

75. Chung, K.L. and Ornstein, D. (1962): On the recurrence of sums of random variables. *Bull. Amer. Math. Soc.* **68**, 30–32.

76. Chung, K.L. and Pollard, H. (1952): An extension of renewal theory. *Proc. Amer. Math. Soc.* **3**, 303–309.

77. Chung, K.L. and Wolfowitz, J. (1952): On a limit theorem in renewal theory. *Ann. of Math.* (2) **55**, 1–6.

78. Çinlar, E. (1975): *Introduction to Stochastic Processes.* Prentice Hall, Englewood Cliffs, N.J.

79. Cox, D.R. (1967): *Renewal Theory.* Methuen, London.

80. Cox, D.R. and Miller, H.D. (1965): *The Theory of Stochastic Processes.* Methuen, London.

81. Cox, D.R. and Smith, W.L. (1953): A direct proof of a fundamental theorem of renewal theory. *Skand. Aktuarietidskr.* **36**, 139–150.

82. Cramér, H. (1955): Collective risk theory: A survey of the theory from the point of view of the theory of stochastic processes. In: *The 7th Jubilee Volume of Skandia Insurance Company,* Stockholm, 51–92. Also in: *Harald Cramér: Collected Works, Vol. II*, (Ed. A. Martin-Löf), 1028–1116. Springer-Verlag, Berlin (1994).

83. Csörgő, M., Deheuvels, P. and Horváth, L. (1987): An approximation of stopped sums with applications in queueing theory. *Adv. Appl. Probab.* **19**, 674–690.

84. Csörgő, M., Horváth, L. and Steinebach, J. (1987): Invariance principles for renewal processes. *Ann. Probab.* **15**, 1441–1460.

85. Csörgő, M. and Révész, P. (1981): *Strong Approximations in Probability and Statistics.* Académiai Kiadó, Budapest.

86. Daley, D.J. (1980): Tight bounds for the renewal function of a random walk. *Ann. Probab.* **8**, 615–621.

87. Davis, B. (1970): On the integrability of the martingale square function. *Israel J. Math.* **8**, 187–190.

88. De Acosta, A. (1983): A new proof of the Hartman–Wintner law of the iterated logarithm. *Ann. Probab.* **11**, 270–276.

89. De Groot, M.H. (1986): A conversation with David Blackwell. *Statistical Science* **1**, 40–53.

90. de Haan, L. (1970): *On Regular Variation and Its Application to the Weak Convergence of Sample Extremes.* Mathematical Centre Tracts 32, Amsterdam.

91. Deheuvels, P. and Steinebach, J. (1989): Sharp rates for increments of renewal processes. *Ann. Probab.* **17**, 700–722.

92. Dippon, J. and Walk, H. (2005): Simplified analytical proof of Blackwell's renewal theorem. *Statist. Probab. Lett.* **74**, 15–20.

93. Doney, R.A. (1966): An analogue of the renewal theorem in higher dimensions. *Proc. London Math. Soc.* (3) **16**, 669–684.

94. Doney, R.A. (1980): Moments of ladder heights in random walks. *J. Appl. Probab.* **17**, 248–252.

95. Doney, R.A. (1982): On the existence of the mean ladder height for random walk. *Z. Wahrsch. verw. Gebiete* **59**, 373–382.

96. Donsker, M. (1951): An invariance principle for certain probability limit theorems. *Mem. Amer. Math. Soc.* **6**.

97. Doob, J.L. (1948): Renewal theory from the point of view of the theory of probability. *Trans. Amer. Math. Soc.* **63**, 422–438.

98. Doob, J.L. (1953): *Stochastic Processes.* John Wiley, New York.

99. Doob, J.L., Snell, J.L. and Williamson, R.E. (1960): Applications of boundary theory to sums of independent random variables. In: *Contributions to Probability and Statistics.* Essays in honor of Harold Hotelling (Eds. Ingram Olkin et al.), 182–197. Stanford University Press, Stanford, CA.

100. Dynkin, E.B. (1955): Limit theorems for sums of independent random quantities. *Izv. Akad. Nauk SSSR Ser. Mat.* **19**, 247–266. *Selected Translations in Math. Stat. and Prob.* **1** (1961), 171–189.

101. Embrechts, P., Klüppelberg, C. and Mikosch, T. (2008): *Modelling Extremal Events: for Insurance and Finance*, Corr. 4th printing. Springer-Verlag, Berlin.

102. Englund, G. (1980): Remainder term estimate for the asymptotic normality of the number of renewals. *J. Appl. Probab.* **17**, 1108–1113.

103. Erdős, P. (1949): On a theorem of Hsu and Robbins. *Ann. Math. Statist.* **20**, 286–291.

104. Erdős, P. (1950): Remark on my paper "On a theorem of Hsu and Robbins." *Ann. Math. Statist.* **21**, 138.

105. Erdős, P., Feller, W. and Pollard, H. (1949): A theorem on power series. *Bull. Amer. Math. Soc.* **55**, 201–204.

106. Erdős, P. and Kac, M. (1946): On certain limit theorems of the theory of probability. *Bull. Amer. Math. Soc.* **52**, 292–302.

107. Erdős, P. and Rényi, A. (1970): On a new law of large numbers. *J. Analyse Math.* **23**, 103–111.

108. Erickson, K.B. (1970): Strong renewal theorems with infinite mean. *Trans. Amer. Math. Soc.* **151**, 263–291.

109. Essén, M. (1973): Banach algebra methods in renewal theory. *J. Analyse Math.* **26**, 303–336.
110. Farrell, R.H. (1964): Limit theorems for stopped random walks. *Ann. Math. Statist.* **35**, 1332–1343.
111. Farrell, R.H. (1966): Limit theorems for stopped random walks III. *Ann. Math. Statist.* **37**, 1510–1527.
112. Feller, W. (1941): On the integral equation of renewal theory. *Ann. Math. Statist.* **12**, 243–267.
113. Feller, W. (1949): Fluctuation theory and recurrent events. *Trans. Amer. Math. Soc.* **67**, 98–119.
114. Feller, W. (1968): *An Introduction to Probability Theory and Its Applications*, Vol. 1, 3rd ed. John Wiley, New York.
115. Feller, W. (1971): *An Introduction of Probability Theory and Its Applications*, Vol. 2, 2nd ed. John Wiley, New York.
116. Feller, W. and Orey, S. (1961): A renewal theorem. *J. Math. Mech.* **10**, 619–624.
117. Frolov, A., Martikainen, A. and Steinebach, J. (2000): On Chung's LIL and Csáki's law for renewal processes. *Period. Math. Hung.* **41**, 149–155.
118. Frolov, A., Martikainen, A. and Steinebach, J. (2001): Limit theorems for maxima of sums and renewal processes. *Zap. Nauchn. Sem. POMI* **278**, 261–274 (2001); translation in *J. Math. Sci. (N.Y.)* **118**, 5658–5666 (2003).
119. Fuh, C.D. and Zhang, C.H. (1998): Wald's equations, first passage times and moments of ladder variables in Markov random walks. *J. Appl. Probab.* **35**, 566–580.
120. Gafurov, M.U. (1980): On the first exit time of a multidimensional random walk from expanding sets. *Soviet Math. Dokl.* **22**, 492–493.
121. Garsia, A.M. (1973): *Martingale Inequalities: Seminar Notes on Recent Progress.* Benjamin, Reading. MA.
122. Garsia, A. and Lamperti, J. (1962/63): A discrete renewal theorem with infinite mean. *Comment. Math. Helv.* **37**, 221–234.
123. Gikhman, I.I. and Skorohod, A.V. (1969): *Introduction to the Theory of Random Processes.* Saunders, Philadelphia, PA.
124. Gnedenko, B.V., Belyayev, Yu. K. and Solovyev, A.D. (1969): *Mathematical Methods of Reliability Theory.* Academic Press, New York.
125. Goldmann, W. (1983): *Sequentielle Schätzer und Erneuerungstheorie. [Sequential Estimators and Renewal Theory].* Dissertation, University of Bonn.
126. Goldmann, W. (1984): Stopping times of inverse renewal theory. Preprint, University of Bonn.
127. Grandell, J. (1991): *Aspects of Risk Theory.* Springer-Verlag, New York.
128. Grübel, R. (1986): On harmonic renewal measures. *Probab. Th. Rel. Fields* **71**, 393–404.
129. Gundy, R.F. and Siegmund, D. (1967): On a stopping rule and the central limit theorem. *Ann. Math. Statist.* **38**, 1915–1917.
130. Gut, A. (1973): A functional central limit theorem connected with extended renewal theory. *Z. Wahrsch. verw. Gebiete* **27**, 123–129.
131. Gut, A. (1974a): On the moments and limit distributions of some first passage times. *Ann. Probab.* **2**, 277–308.
132. Gut, A. (1974b): On the moments of some first passage times for sums of dependent random variables. *Stoch. Process. Appl.* **2**, 115–126.
133. Gut, A. (1974c): On convergence in r-mean of some first passage times and randomly indexed partial sums. *Ann. Probab.* **2**, 321–323.

134. Gut, A. (1975a): Weak convergence and first passage times. *J. Appl. Probab.* **12**, 324–332.

135. Gut, A. (1975b): On a.s. and r-mean convergence of random processes with an application to first passage times. *Z. Wahrsch. verw. Gebiete* **31**, 333–341.

136. Gut, A. (1983a): Renewal theory and ladder variables. In: *Probability and Mathematical Statistics.* Essays in honour of Carl–Gustav Esseen (Eds. A. Gut and L. Holst), 25–39. Uppsala.

137. Gut, A. (1983b): Complete convergence and convergence rates for randomly indexed partial sums with an application to some first passage times. *Acta Math. Acad. Sci. Hungar.* **42**, 225–232; Correction, *ibid.* **45** (1985), 235–236.

138. Gut, A. (1985): On the law of the iterated logarithm for randomly indexed partial sums with two applications. *Studia Sci. Math. Hungar.* **20**, 63–69.

139. Gut, A. (1990a): Limit theorems for record times. In: *Proc. 5th Vilnius Conf. on Prob. Theory and Math. Stat., Vol. 1*, 490–503. Utrecht, The Netherlands.

140. Gut, A. (1990b): Convergence rates for record times and the associated counting process. *Stoch. Process. Appl.* **36**, 135–151.

141. Gut, A. (1992): First passage times of perturbed random walks. *Sequential Anal.* **11**, 149–179.

142. Gut, A. (1996): Stopped Lévy processes with applications to first passage times. *Statist. Probab. Lett.* **28**, 345–352.

143. Gut, A. (1997): Stopped two-dimensional perturbed random walks and Lévy processes. *Statist. Probab. Lett.* **35**, 317–325.

144. Gut, A. (2001): Convergence rates in the central limit theorem for multidimensionally indexed random variables. *Studia Sci. Math. Hungar.* **37**, 401–418.

145. Gut, A. (2007): *Probability: A Graduate Course*, Corr. 2nd printing. Springer-Verlag, New York.

146. Gut, A. and Ahlberg, P. (1981): On the theory of chromatography based upon renewal theory and a central limit theorem for randomly indexed partial sums of random variables. *Chemica Scripta* **18**, 248–255.

147. Gut, A. and Holst, L. (1984): On the waiting time in a generalized roulette game. *Statist. Probab. Lett.* **2**, 229–239.

148. Gut, A. and Janson, S. (1983): The limiting behaviour of certain stopped sums and some applications. *Scand. J. Statist.* **10**, 281–292.

149. Gut, A. and Janson, S. (1986): Converse results for existence of moments and uniform integrability for stopped random walks. *Ann. Probab.* **14**, 1296–1317.

150. Gut, A., Klesov, O. and Steinebach, J. (1997): Equivalences in strong limit theorems for renewal counting processes. *Statist. Probab. Lett.* **35**, 381–394.

151. Gut, A. and Schwabe, R. (1996): Some properties of multiple decisions following a sequential test. *Sequential Anal.* **15**, 299–310.

152. Gut, A. and Schwabe, R. (1999): Some remarks on repeated significance tests for linear contrasts. *J. Statist. Plann. Inference* **82**, 25–34.

153. Gut, A. and Steinebach, J. (2004): Convergence rates and precise asymptotics for renewal counting processes and some first passage times. In: *Proceedings of the International Conference on Asymptotic Methods in Stochastics, May 23–25, 2002, Carleton University, Ottawa. Fields Institute Communications* **44**, 205–227.

154. Hagwood, C. and Woodroofe, M. (1982): On the expansion for expected sample size in nonlinear renewal theory. *Ann. Probab.* **10**, 844–848.

155. Hall, P. and Heyde, C.C. (1980): *Martingale Limit Theory and Its Application.* Academic Press, New York.

156. Hartman, P. and Wintner, A. (1941): On the law of the iterated logarithm. *Amer. J. Math.* **63**, 169–176.

157. Hatōri, H. (1959): Some theorems in an extended renewal theory I. *Kōdai Math. Sem. Rep.* **11**, 139–146.

158. Heyde, C.C. (1964): Two probability theorems and their applications to some first passage problems. *J. Austral. Math. Soc.* **4**, 214–222.

159. Heyde, C.C. (1966): Some renewal theorems with applications to a first passage problem. *Ann. Math. Statist.* **37**, 699–710.

160. Heyde, C.C. (1967a): Asymptotic renewal results for a natural generalization of classical renewal theory. *J. Roy. Statist. Soc. Ser. B* **29**, 141–150.

161. Heyde, C.C. (1967b): A limit theorem for random walks with drift. *J. Appl. Probab.* **4**, 144–150.

162. Hinderer, K. and Walk, H. (1974): A generalization of renewal processes. In: *Progress in Statistics (European Meeting of Statisticians, Budapest, 1972), Vol. I*, 315–318. *Colloq. Math. Soc. Janos Bolyai, Vol. 9*, North-Holland, Amsterdam.

163. Högfeldt, P. (1977): On the asymptotic behaviour of first passage time processes and certain stopped stochastic processes. Abstract. In: *Proc. 2nd Vilnius Conf. on Prob. and Math. Stat.* Vol. 3, 86–87.

164. Horváth, L. (1984a): Strong approximation of renewal processes. *Stoch. Process. Appl.* **18**, 127–138.

165. Horváth, L. (1984b): Strong approximation of extended renewal processes. *Ann. Probab.* **12**, 1149–1166.

166. Horváth, L. (1984c): Strong approximation of certain stopped sums. *Statist. Probab. Lett.* **2**, 181–185.

167. Horváth, L. (1985): A strong nonlinear renewal theorem with applications to sequential analysis. *Scand. J. Statist.* **12**, 271–280.

168. Horváth, L.(1986): Strong approximations of renewal processes and their applications. *Acta Math. Acad. Sci. Hungar.* **47**, 13–28.

169. Hsu, P.L. and Robbins, H. (1947): Complete convergence and the law of large numbers. *Proc. Nat. Acad. Sci. U.S.A.* **33**, 25–31.

170. Hu, I. (1988): Repeated significance tests for exponential families. *Ann. Statist.* **16**, 1643–1666.

171. Hu, I. (1991): Nonlinear renewal theory for conditional random walks. *Ann. Probab.* **19**, 401–422.

172. Huggins, R.M. (1985): Laws of the iterated logarithm for time changed Brownian motion with an application to branching processes. *Ann. Probab.* **13**, 1148–1156.

173. Hunter, J.J. (1974a): Renewal theory in two dimensions: Basic results. *Adv. in Appl. Probab.* **6**, 376–391.

174. Hunter, J.J. (1974b): Renewal theory in two dimensions: Asymptotic results. *Adv. in Appl. Probab.* **6**, 546–562.

175. Hunter, J.J. (1977): Renewal theory in two dimensions: Bounds on the renewal function. *Adv. in Appl. Probab.* **9**, 527–541.

176. Iglehart, D.L. and Whitt, W. (1971): The equivalence of functional central limit theorems for counting processes and associated partial sums. *Ann. Math. Statist.* **42**, 1372–1378.

177. Jagers, P. (1975): *Branching Processes with Biological Applications.* John Wiley, New York.

178. Janson, S. (1983): Renewal theory for m-dependent variables. *Ann. Probab.* **11**, 558–568.

179. Janson, S. (1986): Moments for first passage and last exit times, the minimum and related quantities for random walks with positive drift. *Adv. in Appl. Probab.* **18**, 865–879

180. Kaijser, T. (1971): A stochastic model describing the water motion in a river. *Nordic Hydrology* **II**, 243–265.

181. Karamata, J. (1930): Sur une mode de croissance régulière des fonctions. *Mathematica* (Cluj) **4**, 38–53.

182. Katz, M.L. (1963): The probability in the tail of a distribution. *Ann. Math. Statist.* **34**, 312–318.

183. Keener, R. (1987): Asymptotic expansions in nonlinear renewal theory. In: *New Perspectives in Theoretical and Applied Statistics (Bilbao, 1986)*, 479–502. Wiley Ser. Probab. Math. Statist. Probab. Math. Statist., Wiley, New York.

184. Kemperman, J.H.B. (1961): *The Passage Problem for a Stationary Markov Chain.* University of Chicago Press, Chicago, IL.

185. Kesten, H. (1974): Renewal theory for functionals of a Markov chain with general state space. *Ann. Probab.* **2**, 355–386.

186. Kiefer, J. and Wolfowitz, J. (1956): On the characteristics of the general queueing process, with applications to random walk. *Ann. Math. Statist.* **27**, 147–161.

187. Kingman, J.F.C. (1968): The ergodic theory of subadditive stochastic processes. *J. Roy. Statist. Soc. Ser. B* **30**, 499–510.

188. Klesov, O., Rychlik, Z. and Steinebach, J. (2001): Strong limit theorems for general renewal processes. *Probab. Math. Statist.* **21**, 329–349.

189. Kolmogorov, A.N. (1936): Anfangsgründe der Theorie der Markoffschen Ketten mit unendlich vielen möglichen Zuständen. *Mat. Sb. (N.S.)* **1**, 607–610.

190. Komlós, J., Major, P. and Tusnády, G. (1975): An approximation of partial sums of independent r.v.'s and the sample d.f., I. *Z. Wahrsch. Verw. Gebiete* **32**, 111–131.

191. Komlós, J., Major, P. and Tusnády, G. (1976): An approximation of partial sums of independent r.v.'s and the sample d.f., II. *Z. Wahrsch. Verw. Gebiete* **34**, 33–58.

192. Lai, T.L. (1975): On uniform integrability in renewal theory. *Bull Inst. Math. Acad. Sinica* **3**, 99–105.

193. Lai, T.L. (1976): Asymptotic moments of random walks with applications to ladder variables and renewal theory. *Ann. Probab.* **4**, 51–66.

194. Lai, T.L. (1977): First exit times from moving boundaries for sums of independent random variables. *Ann. Probab.* **5**, 210–221.

195. Lai, T.L. (1988): Boundary crossing problems for sample means. *Ann. Probab.* **16**, 375–396.

196. Lai, T.L. and Siegmund, D. (1977): A nonlinear renewal theory with applications to sequential analysis. *Ann. Statist.* **5**, 946–954.

197. Lai, T.L. and Siegmund, D. (1979): A nonlinear renewal theory with applications to sequential analysis II. *Ann. Statist.* **7**, 60–76.

198. Lalley, S.P. (1982): Nonlinear renewal theory for lattice random walks. *Comm. Statist. — Sequential Analysis* **1**, 193–205.

199. Lalley, S.P. (1983): Repeated likelihood ratio tests for curved exponential families. *Z. Wahrsch. verw. Gebiete* **62**, 293–321.

200. Lalley, S.P. (1984a): Limit theorems for first-passage times in linear and nonlinear renewal theory. *Adv. in Appl. Probab.* **16**, 766–803.

201. Lalley, S.P. (1984b): Conditional Markov renewal theory I. Finite and denumerable state space. *Ann. Probab.* **12**, 1113–1148.
202. Lalley, S.P. (1986): Renewal theorem for a class of stationary sequences. *Probab. Th. Rel. Fields* **72**, 195–213.
203. Lamperti, J. (1958): Some limit theorems for stochastic processes. *J. Math. Mech.* **7**, 433–448.
204. Lamperti, J. (1961): A contribution to renewal theory. *Proc. Amer. Math. Soc.* **12**, 724–731.
205. Larsson-Cohn, L. (1999): *Some Limit and Continuity Theorems for Perturbed Random Walks.* Licentiat thesis. *Report U.U.D.M.* **1999:2**, Uppsala University, Sweden.
206. Larsson-Cohn, L. (2000a): Invariance priniples for the first passage times of perturbed random walks. *Statist. Probab. Lett.* **48**, 347–351.
207. Larsson-Cohn, L. (2000b): A global approach to first passage times. *Probab. Math. Statist.* **20**, 337–341.
208. Larsson-Cohn, L. (2001a): Some remarks on "First passage times of perturbed random walks". *Sequential Anal.* **20**, 87–90.
209. Larsson-Cohn, L. (2001b): On the difference between two first passage times. *Studia Sci. Math. Hung.* **37**, 53–67.
210. Lindberger, K. (1978): Functional limit theorems for cumulative processes and stopping times. *Z. Wahrsch. verw. Gebiete* **44**, 47–56.
211. Lindvall, T. (1973): Weak convergence of probability measures and random functions in the function space $D[0, \infty)$. *J. Appl. Probab.* **10**, 109–121.
212. Lindvall, T. (1977): A probabilistic proof of Blackwell's renewal theorem. *Ann. Probab.* **5**, 482–485.
213. Lindvall, T. (1979): On coupling of discrete renewal processes. *Z. Wahrsch. verw. Gebiete* **48**, 57–70.
214. Lindvall, T. (1982): On coupling of continuous-time renewal processes. *J. Appl. Probab.* **19**, 82–89.
215. Lindvall, T. (1986): On coupling of renewal processes with use of failure rates. *Stoch. Process. Appl.* **22**, 1–15.
216. Loève, M. (1977): *Probability Theory*, 4th ed. Springer-Verlag, New York.
217. Lorden, G. (1970): On excess over the boundary. *Ann. Math. Statist.* **41**, 520–527.
218. Maejima, M. (1975): On local limit theorems and Blackwell's renewal theorem for independent random variables. *Ann. Inst. Statist. Math.* **27**, 507–520.
219. Maejima, M. and Mori, T. (1984): Some renewal theorems for random walks in multidimensional time. *Math. Proc. Cambridge Philos. Soc.* **95**, 149–154.
220. Marcinkiewicz, J. and Zygmund, A. (1937): Sur les fonctions indépendantes. *Fund. Math.* **29**, 60–90.
221. Marcinkiewicz, J. and Zygmund, A. (1938): Quelques théorèmes sur les fonctions indépendantes. *Studia Math.* **VII**, 104–120.
222. Melfi, V.F. (1992): Nonlinear Markov renewal theory with statistical applications. *Ann. Probab.* **20**, 753–771.
223. Melfi, V.F. (1994): Nonlinear renewal theory for Markov random walks. *Stoch. Process. Appl.* **54**, 71–93.
224. Meyer, P.A. (1966): *Probability and Potential.* Waltham, Blaisdell.
225. Mikosch, T. (2006): *Non-Life Insurance Mathematics. An Introduction with Stochastic Processes*, Corr. 2nd printing. Springer-Verlag, Berlin.

226. Mohan, N.R. (1976): Teugels' renewal theorem and stable laws. *Ann. Probab.* **4**, 863–868.

227. Nagaev, A.V. (1979): Renewal theorems in R^d. *Theory Probab. Appl.* **XXIV**, 572–581.

228. Neveu, J. (1975): *Discrete-Parameter Martingales.* North-Holland, Amsterdam.

229. Ney, P. (1981): A refinement of the coupling method in renewal theory. *Stoch. Process. Appl.* **11**, 11–26.

230. Ney, P. and Wainger, S. (1972): The renewal theorem for a random walk in two-dimensional time. *Studia Math.* **XLIV**, 71–85.

231. Niculescu, Ş.P. (1979): Extension of a renewal theorem. *Bull. Math. Soc. Sci. Math. R.S. Roumanie* **23**, 289–292.

232. Ornstein, D.S. (1969a): Random walks I. *Trans. Amer. Math. Soc.* **138**, 1–43.

233. Ornstein, D.S. (1969b): Random walks II. *Trans. Amer. Math. Soc.* **138**, 45–60.

234. Peligrad, M. (1999): Convergence of stopped sums of weakly dependent random variables. *Electronic. J. Probab.* **4**, 1–13.

235. Plucińska, A. (1962): On the joint limiting distribution of times spent in particular states by a Markov process. *Colloq. Math.* **IX**, 347–360.

236. Port, S.C. and Stone, C.J. (1967): Hitting time and hitting places for nonlattice recurrent random walks. *J. Math. Mech.* **17**, 35–57.

237. Prabhu, N.U. (1965): *Stochastic Processes.* Macmillan, New York.

238. Prabhu, N.U. (1980): *Stochastic Storage Processes. Queues, Insurance Risk, and Dams.* Springer-Verlag, New York.

239. Pyke, R. (1958): On renewal processes related to type I and type II counter models. *Ann. Math. Statist.* **29**, 737–754.

240. Pyke, R. and Root, D. (1968): On convergence in r-mean for normalized partial sums. *Ann. Math. Statist.* **39**, 379–381.

241. Rényi, A. (1957): On the asymptotic distribution of the sum of a random number of independent random variables. *Acta Math. Acad. Sci. Hungar.* **8**, 193–199.

242. Rényi, A. (1962): On the extreme elements of observations. *MTA III, Oszt. Közl.* **12**, 105–121. Also in: *Collected Works* **III**, 50–66. Akadémiai Kiadó, Budapest (1976).

243. Revuz, D. (1975): *Markov Chains.* North-Holland, Amsterdam.

244. Richter, W. (1965): Limit theorems for sequences of random variables with sequences of random indices. *Theory Probab. Appl.* **X**, 74–84.

245. Robert, C.Y. (2005): Asymptotic probabilities of exceedance over renewal thresholds with an application to risk theory. *J. Appl. Probab.* **42**, 153–162.

246. Rosén, B. (1961): On the asymptotic distribution of sums of independent identically distributed random variables. *Ark. Mat.* **4**, 323–332.

247. Seal, H.L. (1969): *Stochastic Theory of a Risk Business.* John Wiley, New York.

248. Seneta, E. (1976): *Regularly Varying Functions.* Lecture Notes in Mathematics 508. Springer-Verlag, Berlin.

249. Serfozo, R.F. (1974): Large deviations of renewal process. *Stoch. Process. Appl.* **2**, 295–301.

250. Serfozo, R.F. (1975): Functional limit theorems for stochastic processes based on embedded processes. *Adv. in Appl. Probab.* **7**, 123–139.

251. Siegmund, D.O. (1967): Some one-sided stopping rules. *Ann. Math. Statist.* **38**, 1641–1646.

252. Siegmund, D.O. (1968): On the asymptotic normality of one-sided stopping rules. *Ann. Math. Statist.* **39**, 1493–1497.

253. Siegmund, D.O. (1969): The variance of one-sided stopping rules. *Ann. Math. Statist.* **40**, 1074–1077.

254. Siegmund, D. (1975): The time until ruin in collective risk theory. *Mitt. Verein Schweiz. Versicherungsmath.* **75**, 157–165.

255. Siegmund, D. (1985): *Sequential Analysis. Tests and Confidence Intervals.* Springer-Verlag, New York.

256. Skorohod, A.V. (1956): Limit theorems for stochastic processes. *Theory Probab. Appl.* **I**, 261–290.

257. Smith, W.L. (1954): Asymptotic renewal theorems. *Proc. Roy. Soc. Edinburgh Sect. A* **64**, 9–48.

258. Smith, W.L. (1955): Regenerative stochastic processes. *Proc. Roy. Soc. London Ser. A* **232**, 6–31.

259. Smith, W.L. (1958): Renewal theory and its ramifications. *J. Roy. Statist. Soc. Ser. B* **20**, 243–302.

260. Smith, W.L. (1964): On the elementary renewal theorem for nonidentically distributed variables. *Pacific J. Math.* **14**, 673–699.

261. Smith, W.L. (1967): A theorem on functions of characteristic functions and its applications to some renewal theoretic random walk problems. In: *Proc. Fifth Berkeley Symp. Math. Statist. and Probability*, Vol. II/II, 265–309. University of California Press, Berkeley, CA.

262. Sparre Andersen, E. (1953a): On sums of symmetrically dependent random variables. *Skand. Aktuarietidskr.* **36**, 123–138.

263. Sparre Andersen, E. (1953b): On the fluctuations of sums of random variables I. *Math. Scand.* **1**, 263–285.

264. Sparre Andersen, E. (1954): On the fluctuations of sums of random variables II. *Math. Scand.* **2**, 195–223.

265. Sparre Andersen, E. (1957): On the collective theory of risk in the case of contagion between claims. In: *Trans. XV Internat. Congr. Actuaries, New York* **2**, 219–227.

266. Spitzer, F. (1956): A combinatorial lemma and its application to probability theory. *Trans. Amer. Math. Soc.* **82**, 323–339.

267. Spitzer, F. (1960): A Tauberian theorem and its probability interpretation. *Trans. Amer. Math. Soc.* **94**, 150–169.

268. Spitzer, F. (1965): Renewal theory for Markov chains. In: *Proc. Fifth Berkeley Symp. Math. Statist. and Probability*, Vol. II/II, 311–320. University of California Press, Berkeley, CA.

269. Spitzer, F. (1976): *Principles of Random Walk*, 2nd ed. Springer-Verlag, New York.

270. Sreehari, M. (1970): On a class of limit distributions for normalized sums of independent random variables. *Theory Probab. Appl.* **XV**, 258–281.

271. Stam, A.J. (1968): Two theorems in r-dimensional renewal theory. *Z. Wahrsch. verw. Gebiete* **10**, 81–86.

272. Stam, A.J. (1969): Renewal theory in r dimensions I. *Compositio Math.* **21**, 383–399.

273. Stam, A.J. (1971): Renewal theory in r dimensions II. *Compositio Math.* **23**, 1–13.

274. *Statist. Neerlandica* **36** (1982), 100.

275. *Statist. Neerlandica* **37** (1983), 46–48.

276. Stefanov, V. (1991): A note on the moments of first-crossing times related to nonlinear renewal theory. *Stochastics Stochastic Rep.* **35**, 25–30.

277. Stefanov, V. (1999): Markov renewal processes and exponential families. In: *Semi-Markov Models and Applications (Compiègne 1998)*, 167–186, Kluwer Acad. Publ., Dordrecht.

278. Steinebach, J. (1978): A strong law of Erdős-Rényi type for cumulative processes in renewal theory. *J. Appl. Probab.* **15**, 96–111.

279. Steinebach, J. (1979): *Erdős-Rényi-Zuwächse bei Erneuerungsprozessen und Partialsummen auf Gittern.* Habilitationsschrift, University of Düsseldorf.

280. Steinebach, J. (1986): Improved Erdős–Rényi and strong approximation laws for increments of renewal processes. *Ann. Probab.* **14**, 547–559.

281. Steinebach, J. (1987): Almost sure convergence of delayed renewal processes. *J. London Math. Soc. (2)* **36**, 569–576.

282. Steinebach, J. (1988a): Invariance principles for renewal processes when only moments of low order exist. *J. Multivariate Anal.* **26**, 169–183.

283. Steinebach, J. (1988b): On the optimality of strong approximation rates for compound renewal processes. *Statist. Probab. Lett.* **6**, 263–267.

284. Steinebach, J. (1991): Strong laws for small increments of renewal processes. *Ann. Probab.* **19**, 1768–1776.

285. Steinebach, J. (1998): On a conjecture of Révész and its analogue for renewal processes. In: *Asymptotic Methods in Probability and Statistics (Ottawa, ON, 1997)*, 311–322, North-Holland, Amsterdam.

286. Steinebach, J. and Eastwood, V. (1995): On extreme value asymptotics for increments of renewal processes. In: *Extreme Value Theory and Applications (Villeneuve d'Ascq, 1992). J. Statist. Plann. Inference* **45**, 301–312.

287. Stone, C. (1965): On characteristic functions and renewal theory. *Trans. Amer. Math. Soc.* **120**, 327–342.

288. Stone, C. (1969): On the potential theory operator for one-dimensional recurrent random walks. *Trans. Amer. Math. Soc.* **136**, 413–426.

289. Stone, C. and Wainger, S. (1967): One-sided error estimates in renewal theory. *J. Analyse Math.* **20**, 325–352.

290. Stout, W.F. (1974): *Almost Sure Convergence.* Academic Press, New York.

291. Strassen, V. (1964): An invariance principle for the law of the iterated logarithm. *Z. Wahrsch. verw. Gebiete* **3**, 211–226.

292. Strassen, V. (1966): A converse to the law of the iterated logarithm. *Z. Wahrsch. verw. Gebiete* **4**, 265–268.

293. Strauch, J. (1986): *Taubersätze in der Nichtlinearen Erneuerungstheorie und Anwendungen in der Sequentialanalysis.* Dissertation, University of Stuttgart. *Mitt. Math. Sem. Giessen* **172**, iv + 98 pp.

294. Su, Y.T. (1993): Limit theorems for first-passage times in nonlinear Markov renewal theory. *Sequential Anal.* **12**, 235–245.

295. Su, Y.T. (1998): A renewal theory for perturbed Markov random walks with applications to sequential analysis. *Statistica Sinica* **8**, 429–443.

296. Su, Y.T. and Wang, C.T. (1998): Nonlinear renewal theory for lattice Markov random walk. *Sequential Anal.* **17**, 267–277.

297. Szynal, D. (1972): On almost complete convergence for the sum of a random number of independent random variables. *Bull. Acad. Polon. Sci. Sér. Sci. Math. Astronom. Phys.* **20**, 571–574.

298. Täcklind, S. (1944): Elementare Behandlung vom Erneuerungsproblem für den stationären Fall. *Skand. Aktuarietidskr.* **27**, 1–15.

299. Takács, L. (1956): On a probability problem arising in the theory of counters. *Proc. Cambridge Philos. Soc.* **52**, 488–498.

300. Takács, L. (1957): On certain sojourn time problems in the theory of stochastic processes. *Acta Math. Acad. Sci Hungar.* **8**, 169–191.

301. Takahashi, H. and Woodroofe, M. (1981): Asymptotic expansions in nonlinear renewal theory. *Comm. Statist. A—Theory Methods* **10**, 2113–2135.

302. Teicher, H. (1973): A classical limit theorem without invariance or reflection. *Ann. Probab.* **1**, 702–704.

303. Teugels, J.L. (1968): Renewal theorems when the first or the second moment is infinite. *Ann. Math. Statist.* **39**, 1210–1219.

304. Thorisson, H. (1987): A complete coupling proof of Blackwell's renewal theorem. *Stoch. Process. Appl.* **26**, 87–97.

305. Torrång, I. (1987): The law of the iterated logarithm—cluster points of deterministic and random subsequences. *Probab. Math. Statist.* **VIII**, 133–141.

306. Vervaat, W. (1972a): *Success Epochs in Bernoulli Trials (with Applications in Number Theory).* Mathematical Centre Tracts 42, Amsterdam.

307. Vervaat, W. (1972b): Functional central limit theorems for processes with positive drift and their inverses. *Z. Wahrsch. verw. Gebiete* **23**, 245–253.

308. von Bahr, B. (1974): Ruin probabilities expressed in terms of ladder height distributions. *Scand. Actuar. J.*, 190–204.

309. Wald, A. (1947): Limit distribution of the maximum and minimum of successive cumulative sums of random variables. *Bull. Amer. Math. Soc.* **53**, 142–153.

310. Walk, H. (1989): Nonlinear renewal theory under growth conditions. *Stoch. Process. Appl.* **32**, 289–303.

311. Wang, M. (1992): A local limit theorem for perturbed random walks. *Statist. Probab. Lett.* **15**, 109–116.

312. Wang, M. (1996): Limit theorems on occupation times for perturbed random walks. *Sequential Anal.* **15**, 103–122.

313. Whitt, W. (1980): Some useful functions for functional limit theorems. *Math. Oper. Res.* **5**, 67–81.

314. Wijsman, R.A. (1991): Stopping times: Termination, moments, distribution. In: *Handbook of Sequential Analysis* (Eds. B.K. Ghosh and P.K. Sen), 67–119. *Statist. Textbooks Monogr.*, **118**, Marcel Dekker, New York.

315. Williamson, J.A. (1965): Some renewal theorems for non-negative independent random variables. *Trans. Amer. Math. Soc.* **114**, 417–445.

316. Williamson, J.A. (1968): Random walks and Riesz kernels. *Pacific J. Math.* **25**, 393–415.

317. Woodroofe, M. (1976): A renewal theorem for curved boundaries and moments of first passage times. *Ann. Probab.* **4**, 67–80.

318. Woodroofe, M. (1978): Large deviations of likelihood ratio statistics with applications to sequential testing. *Ann. Statist.* **6**, 72–84.

319. Woodroofe, M. (1982): *Nonlinear Renewal Theory in Sequential Analysis.* CBMS–NSF Regional Conf. Ser. in Appl. Math. **39**. SIAM, Philadelphia, PA.

320. Woodroofe, M. and Keener, R. (1987): Asymptotic expansions in boundary crossing problems. *Ann. Probab.* **15**, 102–114.

321. Woodroofe, M. (1988): Asymptotic expansions for first passage times. *Stoch. Process. Appl.* **28**, 301–315.

322. Yu, K.F. (1979): On the uniform integrability of the normalized randomly stopped sums of independent random variables. Preprint, Yale University.
323. Zhang, C.-H. (1988): A nonlinear renewal theory. *Ann. Probab.* **16**, 793–824.
324. Zhang, C.-H. (1989): A renewal theory with varying drift. *Ann. Probab.* **17**, 723–736.

Index