Oberwolfach
Preprints

OWP 2009 - 13

SERGEY FOSS, DMITRY KORSHUNOV AND STAN ZACHARY

An Introduction to Heavy-tailed and Subexponential Distributions

Mathematisches Forschungsinstitut Oberwolfach gGmbH
Oberwolfach Preprints (OWP) ISSN 1864-7596

Oberwolfach Preprints (OWP)

Starting in 2007, the MFO publishes a preprint series which mainly contains research results related to a longer stay in Oberwolfach. In particular, this concerns the Research in Pairs-Programme (RiP) and the Oberwolfach-Leibniz-Fellows (OWLF), but this can also include an Oberwolfach Lecture, for example.

A preprint can have a size from 1 - 200 pages, and the MFO will publish it on its website as well as by hard copy. Every RiP group or Oberwolfach-Leibniz-Fellow may receive on request 30 free hard copies (DIN A4, black and white copy) by surface mail.

Of course, the full copy right is left to the authors. The MFO only needs the right to publish it on its website *www.mfo.de* as a documentation of the research work done at the MFO, which you are accepting by sending us your file.

In case of interest, please send a **pdf file** of your preprint by email to *rip@mfo.de* or *owlf@mfo.de*, respectively. The file should be sent to the MFO within 12 months after your stay as RiP or OWLF at the MFO.

There are no requirements for the format of the preprint, except that the introduction should contain a short appreciation and that the paper size (respectively format) should be DIN A4, "letter" or "article".

On the front page of the hard copies, which contains the logo of the MFO, title and authors, we shall add a running number (20XX - XX).

We cordially invite the researchers within the RiP or OWLF programme to make use of this offer and would like to thank you in advance for your cooperation.

Imprint:

Mathematisches Forschungsinstitut Oberwolfach gGmbH (MFO)
Schwarzwaldstrasse 9-11
77709 Oberwolfach-Walke
Germany

Tel +49 7834 979 50
Fax +49 7834 979 55
Email admin@mfo.de
URL www.mfo.de

The Oberwolfach Preprints (OWP, ISSN 1864-7596) are published by the MFO.

An Introduction to Heavy-Tailed and Subexponential Distributions

Sergey Foss, Dmitry Korshunov and Stan Zachary

April 30, 2009

Preface

This text studies heavy-tailed distributions in probability theory, and especially convolutions of such distributions. The mail goal is to provide a complete and comprehensive introduction to the theory of long-tailed and subexponential distributions which includes many novel elements and, in particular, is based on the regular use of the principle of a single big jump. Much of the material appears for the first time in text form, including

– the establishment of new relations between known classes of subexponential distributions and the introduction of important new classes;
– the development of some important new concepts, including those of h-insensitivity and local subexponentiality;
– the presentation of new and direct probabilistic proofs of known asymptotic results.

A number of recent textbooks and monographs contain some elements of the present theory, notably those by Asmussen [1, 2], Embrechts, Kluppelberg, and Mikosch [23], Rolski, Teugels, Schmidli, and Schmidt [39], and Borovkov and Borovkov [10]. Further, the monograph by Bingham, Goldie, and Teugels [8] comprehensively develops the theory of regularly varying functions and distributions; the latter form an important subclass of the subexponential distributions.

Chapters 2 and 3 of the present monograph deal comprehensively with the properties of heavy-tailed, long-tailed and subexponential distributions, and give applications to random sums. Chapter 4 develops concepts of local subexponentiality and gives further applications. Finally, Chapter 5 studies the distribution of the maximum of a random walk with negative drift and heavy-tailed increments; notably it contains new and short probabilistic proofs for the tail asymptotics of this distribution for both finite and infinite time horizons. The study of heavy-tailed distributions in more general probability models—for example, Markov-modulated models, those with dependencies, and continuous-time models—is postponed until such future date as the authors may again find some spare time. Nevertheless, the same basic principles apply there as are developed in the present text.

We are thankful to many colleagues for helpful discussions and contributions, most notably to Søren Asmussen, Denis Denisov and Andrew Richards. All errors should be blamed on one of these individuals.

A first version of this manuscript was finished while the authors stayed at the Mathematisches Forschungsinstitut Oberwolfach, under the Research in Pairs programme from March 23 to April 5, 2008. We thank the Institute for its great hospitality and support.

A list of errata, together with complements and updates to this manuscript, will be maintained at http://www.ma.hw.ac.uk/∼stan/heavytails/.

April 2009 S.F., D.K., S.Z.

Notation and conventions

Intervals (x, y) is an open, $[x, y]$ a closed interval; half-open intervals are denoted by $(x, y]$ and $[x, y)$.

Integrals $\int_x^y$ is the integral over the interval $(x, y]$.

$\mathbb{R}$, $\mathbb{R}^+$, $\mathbb{R}^s$ stand for the real line, the positive real half-line $[0, \infty)$, and s-dimensional Cartesian space.

$\mathbb{Z}$, $\mathbb{Z}^+$ stand for the set of integers and for the set $\{0, 1, 2, \ldots\}$.

$\mathbb{I}(A)$ stands for the indicator function of A, that is $\mathbb{I}(A) = 1$ if A holds and $\mathbb{I}(A) = 0$ otherwise.

O, o, and $\sim$ Let u and v depend on a parameter x which tends, say, to infinity. Assuming that v is positive we write

$$\begin{aligned} u = O(v) &\quad \text{if} \quad \limsup_{x \to \infty} |u|/v < \infty \\ u = o(v) &\quad \text{if} \quad u/v \to 0 \text{ as } x \to \infty \\ u \sim v &\quad \text{if} \quad u/v \to 1 \text{ as } x \to \infty. \end{aligned}$$

$\mathbb{P}\{B\}$ stands for the probability (on some appropriate space) of the event B.

$\mathbb{E}\xi$ stands for the mean of the random variable ξ.

$\mathbb{E}\{\xi; B\}$ stands for the mean of ξ over the event B, that is, for $\mathbb{E}\xi\mathbb{I}(B)$.

$F * G$ stands for the convolution of the distributions F and G.

F^{*n} stands for the n-fold convolution of the distribution F with itself.

ξ^+, F^+ for any random variable ξ on $\mathbb{R}$ with distribution F, the random variable $\xi^+ = \max(\xi, 0)$ and F^+ denotes its distribution.

$:=$ The quantity on the left is defined to be equal to the quantity on the right.

$=:$ The quantity on the right is defined to be equal to the quantity on the left.

$\square$ indicates the end of a proof.

Contents

Chapter 1

Introduction

Heavy-tailed distributions (probability measures) play a major role in the analysis of many stochastic systems. For example, they are frequently necessary to accurately model inputs to computer and communications networks, they are an essential component of the description of many risk processes, and they occur naturally in models of epidemiological spread. Important examples are Pareto distributions (and other essentially power-law distributions), lognormal distributions, and Weibull distributions (with parameter less than 1). Indeed most heavy-tailed distributions used in practice belong to one of these families, which are defined, along with others, in Chapter 2. We also consider the Weibull distribution at the end of this chapter.

Since the inputs to systems such as those described above are frequently cumulative in their effects, the analysis of the corresponding models typically features convolutions of heavy-tailed distributions. The properties of such convolutions depend on their satisfying certain regularity conditions. From the point of view of applications practically all heavy-tailed distributions may be considered to be long-tailed, and indeed to possess the stronger property of subexponentiality (see below for definitions).

In this monograph we study convolutions of long-tailed and subexponential distributions on the real line. Our aim is to prove some important new results, and to do so through a simple, coherent and systematic approach. It turns out that all the standard properties of such convolutions are then obtained as easy consequences of these results. Thus we also hope to provide further insight into these properties, and to dispel some of the mystery which still seems to surround the phenomenon of subexponentiality in particular.

We define the *tail function* $\overline{F}$ of a distribution F on $\mathbb{R}$ to be given by $\overline{F}(x) = F(x, \infty)$ for all x. We describe as a *tail property* of F any property which depends only on $\{\overline{F}(x) : x \geq x_0\}$ for any (finite) x_0. We further say that F has *right-unbounded support* if $\overline{F}(x) > 0$ for all x.

Heavy-tailed distributions. A distribution F on $\mathbb{R}$ is said to be *(right-) heavy-tailed* if

$$\int_{-\infty}^{\infty} e^{\lambda x} \, F(dx) = \infty \quad \text{for all } \lambda > 0, \tag{1.1}$$

that is, if and only if F fails to possess any positive exponential moment. Otherwise F is said to be *light-tailed*. We shall show in Chapter 2 that the distribution F is heavy-tailed if and only if its tail function $\overline{F}$ fails to be bounded by any exponentially decreasing function.

It follows that for a distribution F to be heavy-tailed is a tail property of F, and of course that any heavy-tailed distribution has right-unbounded support.

We mention briefly at this point the connection with hazard rates. Let F be a distribution on $\mathbb{R}$ which is absolutely continuous with density f with respect to Lebesgue measure. Such a distribution is often characterised in terms of its *hazard rate* $r(x) = f(x)/\overline{F}(x)$, most naturally in the case where F is concentrated on the positive half-line $\mathbb{R}^+$. We then have

$$\overline{F}(x) = \exp\left(-\int_{-\infty}^{x} r(y)\,dy\right).$$

It follows easily from (1.1) that if $\lim_{x\to\infty} r(x) = 0$ then the distribution F is heavy-tailed, whereas if $\liminf_{x\to\infty} r(x) > 0$ then F fails to be heavy-tailed (indeed the integral in (1.1) is finite for any λ such that $\liminf_{x\to\infty} r(x) > \lambda$). In the final case where $\liminf_{x\to\infty} r(x) = 0$ but in which the limit itself fails to exist then both possibilities for F exist.

Long-tailed distributions. A distribution F on $\mathbb{R}$ is said to be *long-tailed* if F has right-unbounded support and, for any fixed $y > 0$,

$$\frac{\overline{F}(x+y)}{\overline{F}(x)} \to 1 \qquad \text{as } x \to \infty. \tag{1.2}$$

Clearly to be long-tailed is again a tail property of a distribution. Further, it is fairly easy to see that a long-tailed distribution is also heavy-tailed. However, the condition (1.2) implies a degree of smoothness in the tail function $\overline{F}$ which is not possessed by every heavy-tailed distribution.

Subexponential distributions. In order to make good progress with heavy-tailed distributions, we require slightly a stronger regularity condition than the requirement that such a distribution be long-tailed. This will turn out to be satisfied by all heavy-tailed distributions likely to be encountered in practice.

We consider first distributions on the positive half-line $\mathbb{R}^+$. Let F be any distribution on $\mathbb{R}^+$ and let $\xi_1, \dots, \xi_n$ be independent random variables with the common distribution F. Then

$$\begin{aligned}\mathbb{P}\{\xi_1 + \ldots + \xi_n > x\} &\geq \mathbb{P}\{\max(\xi_1, \dots, \xi_n) > x\} \\ &= 1 - F^n(x) \\ &\sim n\overline{F}(x) \quad \text{as } x \to \infty. \end{aligned} \tag{1.3}$$

(Here and throughout we use "$\sim$" to mean that the ratio of the quantities on either side of this symbol converges to one; we further frequently omit, especially in proofs, the qualifier "as $x \to \infty$", as unless otherwise indicated all our limits will be of this form.)

Taking $n = 2$ it follows in particular that

$$\liminf_{x\to\infty} \frac{\overline{F*F}(x)}{\overline{F}(x)} \geq 2, \tag{1.4}$$

where as usual, for any two distributions F and G, by $F * G$ we denote their *convolution*, i.e. the distribution of the random variable $\xi + \eta$ where the random variables ξ and η are independent with distributions F and G.

A considerably deeper result (proved in Chapter 2) than the above inequality (1.4) is that if F is *heavy-tailed* then the relation (1.4) holds with *equality*. (We remark that there are also examples of light-tailed distributions on $\mathbb{R}^+$ for which (1.4) holds with equality.) The distribution F on $\mathbb{R}^+$ is said to be *subexponential* if

$$\lim_{x\to\infty} \frac{\overline{F * F}(x)}{\overline{F}(x)} = 2. \tag{1.5}$$

It turns out that the above condition is now sufficient to ensure that F is heavy-tailed—and indeed that F is long-tailed. Thus a distribution F on $\mathbb{R}^+$ is subexponential if and only if is heavy-tailed and sufficiently regular that the limit on the left side of (1.5) exists; this limit is then equal to 2. It is therefore not surprising that the various examples of heavy-tailed distributions on $\mathbb{R}^+$ mentioned at the start of this chapter all turn out to be subexponential. Indeed those heavy-tailed distributions which do not possess this property are all distinctly pathological in character.

We shall see that subexponentiality as defined above is also a *tail property* of the distribution F. Inductive arguments (see Chapter 3) now show that if a distribution F on $\mathbb{R}^+$ is subexponential then the relation (1.5) generalises to

$$\lim_{x\to\infty} \frac{\overline{F^{*n}}(x)}{\overline{F}(x)} = n \quad \text{for all integer } n \geq 1$$

(where F^{*n} denotes the n-fold convolution of the distribution F with itself). It follows from this and from the argument leading to (1.3) that subexponentiality of F is equivalent to the requirement that

$$\mathbb{P}\{\max(\xi_1, \ldots, \xi_n) > x\} \sim \mathbb{P}\{\xi_1 + \cdots + \xi_n > x\} \quad \text{as } x \to \infty, \tag{1.6}$$

The interpretation of the condition (1.6) is that the only significant way in which the sum $\xi_1 + \cdots + \xi_n$ can exceed some large value x is that the maximum of one of the individual random variables $\xi_1, \ldots, \xi_n$ also exceeds x. This is the *principle of a single big jump* which underlies the probabilistic behaviour of sums of independent subexponential random variables.

Since subexponentiality is a tail property of a distribution, it is natural, and important for many applications, to extend the concept to a distribution F on the entire real line $\mathbb{R}$. This may be done *either* by requiring that F have the same tail as that of a subexponential distribution on $\mathbb{R}^+$ (it is natural to consider the distribution F^+ given by $F^+(x) = F(x)$ for $x \geq 0$ and $F^+(x) = 0$ for $x < 0$) *or*, equivalently as it turns out, by requiring that F is long-tailed and again satisfies the condition (1.5)—the latter condition on its own no longer being sufficient for the subexponentiality of F. We explore these matters further in Chapter 3.

We develop also similar concepts of subexponentiality for local probabilities and densities (see Chapter 4).

Further examples of heavy-tailed distributions which are of use in practical applications, e.g. the modelling of insurance claim sizes, are given by Embrechts et al. [23]. These, and the examples mentioned above, are all well-behaved in a manner we shall shortly make precise. However, mathematically there is a whole range of further possible distributions, and one of our aims is to provide a firm basis for excluding those which are in some sense pathological and to study the properties of those which remain.

Example: the Weibull distribution In order to understand better the typical behaviour of heavy-tailed distributions, that is, the single big jump phenomenon—as opposed to the behaviour of distributions which are light-tailed—we study the Weibull distribution F_α given by its tail function

$$\overline{F}_\alpha(x) = e^{-x^\alpha}, \quad x \geq 0,$$

and hence density $f_\alpha(x) = \alpha x^{\alpha-1} e^{-x^\alpha}$, $x \geq 0$, for some *shape* parameter $\alpha > 0$. This is a heavy-tailed distribution if and only if $\alpha < 1$. Note that in the case $\alpha = 1$ we have the *exponential distribution*. All moments of the Weibull distribution are finite.

Let ξ_1 and ξ_2 be independent random variables with common distribution function F_α. We consider the distribution of the random variable ξ_1/d conditional on the sum $\xi_1 + \xi_2 = d$ for varying values of d and the shape parameter α. This conditional distribution has density $g_{\alpha,d}$ where

$$g_{\alpha,d}(z) = c[z(1-z)]^{\alpha-1} e^{-d^\alpha (z^\alpha + (1-z)^\alpha)}. \tag{1.7}$$

for the appropriate normalising constant c. Clearly this conditional density is symmetric about $1/2$. The left panel of Figure 1.1 plots the density for $d = 10$ and for each of the three cases $\alpha = 0.5$, $\alpha = 1$, and $\alpha = 2$, while the right panel plots the density for $d = 25$ and for each of the same three values of α. We see that in the heavy-tailed case $\alpha = 0.5$, conditional on the fixed value d of the sum $\xi_1 + \xi_2$, the value of ξ_1/d tends to be either close to 0 or close to 1; further this effect is more pronounced for the larger value of d. For the case $\alpha = 1$ and for any value of d, the above conditional density is uniform. For the case $\alpha = 2$, we see that the conditional density of ξ_1/d is concentrated in a neighbourhood of $1/2$, and that again this concentration is more pronounced for the larger value of d.

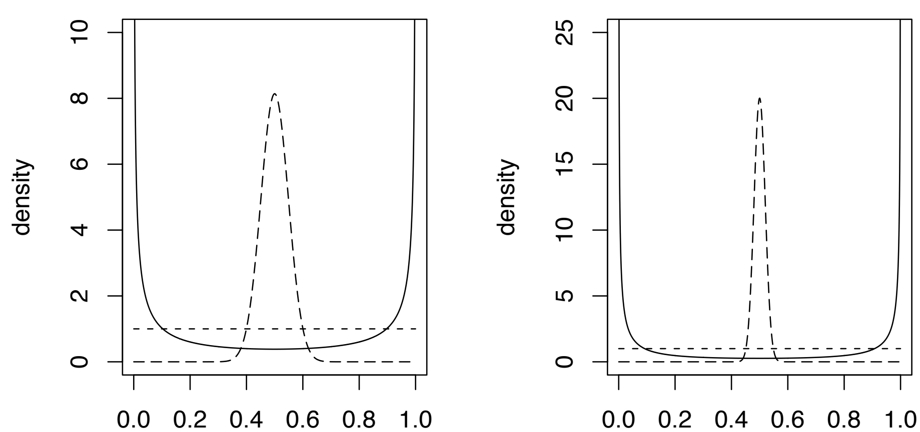

Figure 1.1: Density of ξ_1/d conditional on $\xi_1 + \xi_2 = d$, for $d = 10$ (left panel) and $d = 25$ (right panel), and for $\alpha = 0.5$ (solid line), $\alpha = 1$ (short-dashed line) and $\alpha = 2$ (long-dashed line).

These observations are readily verified from (1.7). Indeed it follows from that expression that, for $\alpha < 1$ and as $d \to \infty$, the distribution of ξ_1/d conditional on $\xi_1 + \xi_2 = d$ converges to that which assigns probability $1/2$ to each of the points 0 and 1. For $\alpha = 1$ and for all d, the distribution is uniform. Finally, for $\alpha > 1$ and as $d \to \infty$, the distribution converges to that which is concentrated on the single point $1/2$.

Chapter 2

Heavy- and long-tailed distributions

In this chapter we are interested in *(right-) tail properties* of distributions, i.e. in properties of a distribution which, for any x, depend only on the restriction of the distribution to (x, ∞). More generally it is helpful to consider tail properties of functions.

Recall that for any distribution F on $\mathbb{R}$ we define the *tail function* $\overline{F}$ by

$$\overline{F}(x) = F(x, \infty), \quad x \in \mathbb{R}.$$

We start with characteristic properties of heavy-tailed distributions, that is, of distributions all of whose positive exponential moments are infinite. The main result here concerns lower limits for convolution tails, see Section 2.3.

Then we study different properties of long-tailed distributions, that is, of distributions whose tails are asymptotically self-similar under shifting by a constant. Of particular interest are convolutions of long-tailed distributions. Our approach is based on a simple decomposition for such convolutions, and on the concept of 'h-insensitivity' for a long-tailed distribution with respect to some (slowly) increasing function h. In Section 2.8, we present useful characterisations of h-insensitive distributions.

2.1 Heavy-tailed distributions

The usage of the term "heavy-tailed distribution" varies according to the area of interest, but is frequently taken to correspond to an absence of (positive) exponential moments. In the following definitions—which, for completeness here, repeat some of those made in the Introduction—we follow this tradition.

Definition 2.1. A distribution F on $\mathbb{R}$ is said to have *right-unbounded support* if $\overline{F}(x) > 0$ for all x.

Definition 2.2. We define a *distribution* F to be (right-) *heavy-tailed* if and only if

$$\int_{\mathbb{R}} e^{\lambda x} F(dx) = \infty \quad \text{for all } \lambda > 0. \tag{2.1}$$

It will follow from Theorem 2.6 that to be heavy-tailed is indeed a tail property of a distribution. As a counterpart we give also the following definition.

Definition 2.3. A *distribution* F is called *light-tailed* if and only if

$$\int_{\mathbb{R}} e^{\lambda x} F(dx) < \infty \quad \text{for some } \lambda > 0. \tag{2.2}$$

Clearly, for any light-tailed distribution F on the positive half-line $\mathbb{R}^+ = [0, \infty)$, all moments are finite, that is, $\int_0^\infty x^k F(dx) < \infty$ for all $k > 0$.

We shall say that a nonnegative *function* (usually tending to zero) is *heavy-tailed* if it fails to be bounded by a decreasing exponential function. More precisely we make the following definition.

Definition 2.4. We define a *function* $f \geq 0$ to be *heavy-tailed* if and only if

$$\limsup_{x\to\infty} f(x) e^{\lambda x} = \infty \quad \text{for all } \lambda > 0. \tag{2.3}$$

For a function to be heavy-tailed is clearly a tail-property of that function. Theorem 2.6 shows in particular that a *distribution* is heavy-tailed if and only if its tail function is a heavy-tailed function. First we make the following definition.

Definition 2.5. For any distribution F, the function $R(x) := -\ln \overline{F}(x)$ is called the *hazard function*. If the hazard function is differentiable then its derivative $r(x) = R'(x)$ is called the *hazard rate*.

The hazard rate, when it exists, has the usual interpretation discussed in the Introduction.

Theorem 2.6. *For any distribution F the following assertions are equivalent:*

(i) *F is a heavy-tailed distribution;*

(ii) *the function $\overline{F}$ is heavy-tailed;*

(iii) *the hazard function satisfies $\liminf_{x\to\infty} R(x)/x = 0$;*

(iv) *for any fixed $T > 0$, the function $F(x, x+T]$ is heavy-tailed;*

(v) *for some fixed $T > 0$, the function $F(x, x+T]$ is heavy-tailed.*

Proof. (i)⇒(iv). Suppose that $F(x, x+T]$ is not heavy-tailed. Then

$$c := \sup_{x\in\mathbb{R}} F(x, x+T] e^{\lambda' x} < \infty \quad \text{for some } \lambda' > 0,$$

and, therefore, for all $\lambda < \lambda'$

$$\begin{aligned}\int_0^\infty e^{\lambda x} F(dx) &\leq \sum_{n=0}^\infty e^{\lambda(n+1)T} F(nT, nT+T] \\ &\leq c \sum_{n=0}^\infty e^{\lambda(n+1)T} e^{-\lambda' nT} = c e^{\lambda T} \sum_{n=0}^\infty e^{(\lambda-\lambda')nT} < \infty.\end{aligned}$$

It follows that the integral defined in (2.1) is finite for all $\lambda \in (0, \lambda')$, which contradicts heavy-tailedness of the distribution F.

(v)⇒(ii). This implication follows from the inequality $\overline{F}(x) \geq F(x, x+T]$.

(ii)⇒(iii). Suppose that, on the contrary, '$\liminf$' in (iii) is positive. Then there exist $x_0 > 0$ and $\varepsilon > 0$ such that $R(x) \geq \varepsilon x$ for all $x \geq x_0$ which implies that $\overline{F}(x) \leq e^{-\varepsilon x}$ in contradiction of (ii).

(iii)⇒(i). Suppose that, on the contrary, F is light-tailed. It then follows from (2.2) (for example by the exponential Chebyshev inequality) that, for some $\lambda > 0$ and $c > 0$, we have $\overline{F}(x) \leq ce^{-\lambda x}$ for all x. This implies that $\liminf_{x\to\infty} R(x)/x \geq \lambda$ which contradicts (iii). □

Lemma 2.7. *Let the distribution F be absolutely continuous with density function f. If F is a heavy-tailed distribution, then the function $f(x)$ is heavy-tailed too.*

Proof. If $f(x)$ is not heavy-tailed, then there exist $\lambda' > 0$ and x_0 such that

$$c := \sup_{x>x_0} f(x)e^{\lambda' x} < \infty,$$

and, therefore, for all $\lambda \in (0, \lambda')$

$$\int_{\mathbb{R}} e^{\lambda x} F(dx) \leq e^{\lambda x_0} + c \int_{x_0}^{\infty} e^{\lambda x} e^{-\lambda' x} dx < \infty.$$

It follows that the integral defined in (2.1) is finite for all λ such that $0 < \lambda < \lambda'$, which contradicts heavy-tailedness of the distribution F. □

We give an example to show that the converse assertion is not in general true. Consider the following piecewise continuous density function:

$$f(x) = \sum_{n=1}^{\infty} \mathbb{I}\{x \in [n, n + 2^{-n}]\}.$$

We have $\limsup_{x\to\infty} f(x)e^{\lambda x} = \infty$ for all $\lambda > 0$, so that f is heavy-tailed. On the other hand, for all $\lambda \in (0, \ln 2)$,

$$\int_0^{\infty} e^{\lambda x} f(x)dx < \sum_{n=1}^{\infty} e^{\lambda(n+2^{-n})} 2^{-n} = \sum_{n=1}^{\infty} e^{\lambda(n+2^{-n})-n\ln 2} < \infty,$$

so that F is light-tailed.

For lattice distributions we have the following result.

Lemma 2.8. *Let F be a distribution on some lattice $\{a + hn, n \in \mathbb{Z}\}$, $a \in \mathbb{R}$, $h > 0$, with probabilities $F\{a + hn\} = p_n$. Then F is heavy-tailed if and only if the sequence $\{p_n\}$ is heavy-tailed, that is,*

$$\limsup_{n\to\infty} p_n e^{\lambda n} = \infty \quad \textit{for all } \lambda > 0. \tag{2.4}$$

Proof. The result follows from Theorem 2.6 with $T = h$. □

Examples of heavy-tailed distributions We conclude this section with a number of examples.

- The *Pareto distribution* on $\mathbb{R}^+$. This has tail function $\overline{F}$ given by

  $$\overline{F}(x) = \left(\frac{\kappa}{x+\kappa}\right)^{\alpha}$$

  for some scale parameter $\kappa > 0$ and shape parameter $\alpha > 0$. Clearly we have $\overline{F}(x) \sim (x/\kappa)^{-\alpha}$ as $x \to \infty$, and for this reason the Pareto distributions are sometimes referred to as the *power law distributions*. The Pareto distribution has all moments of order $\gamma < \alpha$ finite, while all moments of order $\gamma \geq \alpha$ are infinite.

- The *Burr distribution* on $\mathbb{R}^+$. This has tail function $\overline{F}$ given by

  $$\overline{F}(x) = \left(\frac{\kappa}{x^{\tau}+\kappa}\right)^{\alpha}$$

  for parameters $\alpha, \kappa, \tau > 0$. The Burr distribution has all moments of order $\gamma < \alpha\tau$ finite, while all moments of order $\gamma \geq \alpha\tau$ are infinite.

- The *Cauchy distribution* on $\mathbb{R}$. This is most easily given by its density function f where

  $$f(x) = \frac{1}{\pi((x-a)^2+1)}$$

  for some location parameter a. All moments of order $\gamma < 1$ are finite. The first moment does not exist.

- The *lognormal distribution* on $\mathbb{R}^+$. This is again most easily given by its density function f where

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

  for parameters μ and $\sigma > 0$. All moments of the lognormal distribution are finite. Note that a (positive) random variable ξ has a lognormal distribution with parameters μ and σ if and only if $\log \xi$ has a *normal distribution* with mean μ and variance σ^2. For this reason the distribution is natural in many applications.

- The *Weibull distribution* on $\mathbb{R}^+$. This has tail function $\overline{F}$ given by

  $$\overline{F}(x) = e^{-(x/\lambda)^{\alpha}}.$$

  for some scale parameter $\lambda > 0$ and shape parameter $\alpha > 0$. This is a heavy-tailed distribution if and only if $\alpha < 1$. Note that in the case $\alpha = 1$ we have the *exponential distribution*. All moments of the Weibull distribution are finite.

Another useful class of heavy-tailed distributions is that of dominated-varying distributions. We say that F is a *dominated-varying distribution* (and write $F \in \mathcal{D}$) if there exists $c > 0$ such that

$$\overline{F}(2x) \geq c\overline{F}(x) \quad \text{for all } x.$$

Any intermediate regularly varying distribution (see Section 2.8) belongs to $\mathcal{D}$. Other examples may be constructed using the following scheme. Let G be a distribution with a regularly varying tail (again see Section 2.8). Then a distribution F belongs to the class $\mathcal{D}$, provided $c_1\overline{G}_1(x) \leq \overline{F}(x) \leq c_2\overline{G}(x)$ for some $0 < c_1 < c_2 < \infty$ and for all sufficiently large x.

2.2 Characterisation of heavy-tailed distributions in terms of generalised moments

In this section we consider an important characterisation of (the tails of) heavy-tailed distributions on $\mathbb{R}^+$, which is both of interest in itself and essential to the consideration of convolutions in the following section. In very approximate terms, for any such distribution we seek the existence of a monotone concave function h such that the function $e^{-h(\cdot)}$ characterises the tail of the distribution.

If a distribution F on the positive half-line $\mathbb{R}^+$ is such that not all of its moments are finite, that is, $\int_0^\infty x^k F(dx) = \infty$ for some k, then F is heavy-tailed. In this case we can find such $k \geq 1$ that the kth moment is infinite, while the $(k-1)$th moment is finite. That is

$$\int_0^\infty x e^{(k-1)\ln x} dx = \infty \quad \text{and} \quad \int_0^\infty e^{(k-1)\ln x} dx < \infty. \tag{2.5}$$

Note that here the power of the exponent is a concave function. This observation can be generalised onto the whole class of heavy-tailed distributions as follows.

Theorem 2.9. *Let $\xi \geq 0$ be a random variable with a heavy-tailed distribution. Let the function $g(x)$ be such that $g(x) \to \infty$ as $x \to \infty$. Then there exists a monotone concave function $h : \mathbb{R}^+ \to \mathbb{R}^+$ such that $h(x) = o(x)$ as $x \to \infty$, $\mathbb{E}e^{h(\xi)} < \infty$, and $\mathbb{E}e^{h(\xi)+g(\xi)} = \infty$.*

Now, (2.5) is a particular example of the latter theorem with $g(x) = \ln x$. If not all moments of ξ are finite, then the concave function $h(x)$ may be taken as $(k-1)\ln x$ for k as defined above. But, actually, Theorem 2.9 is sharper; it guarantees the existence of a concave function h for any g, which may be taken as thin as we please.

As an example of the assertion at the beginning of this section, note that if ξ has a Weibull distribution with tail function $\overline{F}(x) = e^{-x^\alpha}$, $\alpha \in (0,1)$, and if $g(x) = \ln x$, then one can choose $h(x) = (x+c)^\alpha - \ln(x+c)$, with $c > 0$ sufficiently large.

Note also that Theorem 2.9 provides a characteristic property of heavy-tailed distributions; it fails for any light-tailed distribution. Indeed, consider any non-negative random variable ξ having a light-tailed distribution, that is, $\mathbb{E}e^{\lambda\xi} < \infty$ for some $\lambda > 0$. Take $g(x) = \ln x$. If $h(x) = o(x)$ as $x \to \infty$, then $h(x) \leq c + \lambda x/2$ for some $c < \infty$ and, hence,

$$\mathbb{E}e^{h(\xi)+g(\xi)} \leq \mathbb{E}\xi e^{c+\lambda\xi/2} < \infty.$$

Proof of Theorem 2.9. We will construct a piecewise linear function $h(x)$. To do so we construct two positive sequences $x_n \uparrow \infty$ and $\varepsilon_n \downarrow 0$ as $n \to \infty$ and let

$$h(x) = h(x_{n-1}) + \varepsilon_n(x - x_{n-1}) \quad \text{if } x \in (x_{n-1}, x_n],\ n \geq 1.$$

This function is monotone, since $\varepsilon_n > 0$. Moreover, this function is concave, due to the monotonicity of ε_n.

Put $x_0 = 0$ and $h(0) = 0$. Since ξ is heavy-tailed and $g(x) \to \infty$, we can choose x_1 so large that $e^{g(x)} \geq 2^1$ for all $x > x_1$ and

$$\mathbb{E}\{e^\xi; \xi \in (x_0, x_1]\} + e^{x_1}\overline{F}(x_1) > \overline{F}(x_0) + 1.$$

Choose $\varepsilon_1 > 0$ so that

$$\mathbb{E}\{e^{\varepsilon_1 \xi}; \xi \in (x_0, x_1]\} + e^{\varepsilon_1 x_1}\overline{F}(x_1) = \overline{F}(0) + 1/2,$$

which is equivalent to

$$\mathbb{E}\{e^{h(\xi)}; \xi \in (x_0, x_1]\} + e^{h(x_1)}\overline{F}(x_1) = e^{h(x_0)}\overline{F}(0) + 1/2.$$

By induction we construct an increasing sequence x_n and a decreasing sequence $\varepsilon_n > 0$ such that $e^{g(x)} \geq 2^n$ for all $x > x_n$ and

$$\mathbb{E}\{e^{h(\xi)}; \xi \in (x_{n-1}, x_n]\} + e^{h(x_n)}\overline{F}(x_n) = e^{h(x_{n-1})}\overline{F}(x_{n-1}) + 1/2^n$$

for any $n \geq 2$. For $n = 1$ this is already done. Make the induction hypothesis for some $n \geq 2$. Due to the heavy-tailedness of ξ and to the convergence $g(x) \to \infty$, there exists x_{n+1} so large that $e^{g(x)} \geq 2^{n+1}$ for all $x > x_{n+1}$ and

$$\mathbb{E}\{e^{\varepsilon_n(\xi - x_n)}; \xi \in (x_n, x_{n+1}]\} + e^{\varepsilon_n(x_{n+1}-x_n)}\overline{F}(x_{n+1}) > 2;$$

As a function of ε_{n+1}, the sum

$$\mathbb{E}\{e^{\varepsilon_{n+1}(\xi - x_n)}; \xi \in (x_n, x_{n+1}]\} + e^{\varepsilon_{n+1}(x_{n+1}-x_n)}\overline{F}(x_{n+1})$$

is continuously decreasing to $\overline{F}(x_n)$ as $\varepsilon_{n+1} \downarrow 0$. Therefore, we can choose $\varepsilon_{n+1} \in (0, \varepsilon_n)$ so that

$$\mathbb{E}\{e^{\varepsilon_{n+1}(\xi - x_n)}; \xi \in (x_n, x_{n+1}]\} + e^{\varepsilon_{n+1}(x_{n+1}-x_n)}\overline{F}(x_{n+1}) = \overline{F}(x_n) + 1/(2^{n+1}e^{h(x_n)}).$$

By the definition of $h(x)$ this is equivalent to the following equality:

$$\mathbb{E}\{e^{h(\xi)}; \xi \in (x_n, x_{n+1}]\} + e^{h(x_{n+1})}\overline{F}(x_{n+1}) = e^{h(x_n)}\overline{F}(x_n) + 1/2^{n+1}.$$

Our induction hypothesis now holds with $n+1$ in place of n as required.

Next, for any N,

$$\begin{aligned}
\mathbb{E}\{e^{h(\xi)}; \xi \leq x_{N+1}\} &= \sum_{n=0}^{N} \mathbb{E}\{e^{h(\xi)}; \xi \in (x_n, x_{n+1}]\} \\
&= \sum_{n=0}^{N} \Big(e^{h(x_n)}\overline{F}(x_n) - e^{h(x_{n+1})}\overline{F}(x_{n+1}) + 1/2^{n+1}\Big) \\
&\leq e^{h(x_0)}\overline{F}(x_0) + 1.
\end{aligned}$$

Hence, $\mathbb{E}e^{h(\xi)}$ is finite. On the other hand, since $e^{g(x)} \geq 2^n$ for all $x > x_n$,

$$\begin{aligned}
\mathbb{E}\{e^{h(\xi)+g(\xi)}; \xi > x_n\} &\geq 2^n \mathbb{E}\{e^{h(\xi)}; \xi > x_n\} \\
&\geq 2^n \Big(\mathbb{E}\{e^{h(\xi)}; \xi \in (x_n, x_{n+1}]\} + e^{h(x_{n+1})}\overline{F}(x_{n+1})\Big) \\
&= 2^n \Big(e^{h(x_n)}\overline{F}(x_n) + 1/2^{n+1}\Big).
\end{aligned}$$

Then $\mathbb{E}\{e^{h(\xi)+g(\xi)}; \xi > x_n\} \geq 1/2$ for any n, which implies $\mathbb{E}e^{h(\xi)+g(\xi)} = \infty$. Note also that necessarily $\lim_{n\to\infty} \varepsilon_n = 0$; otherwise $\liminf_{x\to\infty} h(x)/x > 0$ and ξ is light tailed. $\square$

The latter theorem can be strengthened in the following way (for the proof see [19]):

Theorem 2.10. *Let $\xi \geq 0$ be a random variable with a heavy-tailed distribution. Let $f : \mathbb{R}^+ \to \mathbb{R}$ be a concave function such that $\mathbb{E}e^{f(\xi)} = \infty$. Let the function $g : \mathbb{R}^+ \to \mathbb{R}$ be such that $g(x) \to \infty$ as $x \to \infty$. Then there exists a concave function $h : \mathbb{R}^+ \to \mathbb{R}^+$ such that $h \leq f$, $\mathbb{E}e^{h(\xi)} < \infty$, and $\mathbb{E}e^{h(\xi)+g(\xi)} = \infty$.*

2.3 Lower limit for tails of convolutions

Recall that the *convolution* $F * G$ of any two distributions F and G is given by, for any Borel set B,

$$(F * G)(B) = \int_{-\infty}^{\infty} F(B-y)G(dy) = \int_{-\infty}^{\infty} G(B-y)F(dy),$$

where $B - y = \{x - y : x \in B\}$. If, on some probability space with probability measure $\mathbb{P}$, ξ and η are independent random variables with respective distributions F and G, then $(F * G)(B) = \mathbb{P}\{\xi + \eta \in B\}$. The tail function of the convolution, the *convolution tail*, of F and G is then given by, for any $x \in \mathbb{R}$,

$$\overline{F * G}(x) = \mathbb{P}\{\xi + \eta > x\} = \int_{-\infty}^{\infty} \overline{F}(x-y)G(dy) = \int_{-\infty}^{\infty} \overline{G}(x-y)F(dy).$$

Now let F be a distribution on $\mathbb{R}^+$. In the present section we discuss the following lower limit:

$$\liminf_{x\to\infty} \frac{\overline{F * F}(x)}{\overline{F}(x)},$$

in the case where F is heavy-tailed. We start with the following result.

Theorem 2.11. *For any distributions $F_1, \ldots, F_n$ on $\mathbb{R}^+$ with unbounded supports,*

$$\liminf_{x\to\infty} \frac{\overline{F_1 * \ldots * F_n}(x)}{\overline{F}_1(x) + \ldots + \overline{F}_n(x)} \geq 1.$$

Proof. Let $\xi_1, \ldots, \xi_n$ be independent random variables with respective distributions $F_1, \ldots, F_n$. Since the events $\{\xi_k > x, \xi_j \in [0, x] \text{ for all } j \neq k\}$ are disjoint for different k, the convolution tail can be estimated from below in the following way:

$$\begin{aligned}
\overline{F_1 * \ldots * F_n}(x) &\geq \sum_{k=1}^{n} \mathbb{P}\{\xi_k > x, \xi_j \in [0, x] \text{ for all } j \neq k\} \\
&= \sum_{k=11}^{n} \overline{F}_k(x) \prod_{j\neq k} F_j(x) \\
&\sim \sum_{k=1}^{n} \overline{F}_k(x) \quad \text{as } x \to \infty,
\end{aligned}$$

which implies the desired statement. Additionally note that we have heavily used the condition $F_k(\mathbb{R}^+) = 1$; for distributions on the whole real line the conclusion in general fails. □

It follows in particular that, for any distribution F on $\mathbb{R}^+$ with unbounded support and for any $n \geq 2$,

$$\liminf_{x\to\infty} \frac{\overline{F^{*n}}(x)}{\overline{F}(x)} \geq n. \tag{2.6}$$

In particular,

$$\liminf_{x\to\infty} \frac{\overline{F*F}(x)}{\overline{F}(x)} \geq 2. \tag{2.7}$$

As already discussed in the Introduction, in the light-tailed case the limit given by the left side of (2.7) is typically greater than 2. For example, for an exponential distribution it equals infinity. Thus we may ask under what conditions do we have equality in (2.7). We show that heavy-tailedness of F is sufficient.

Theorem 2.12. *For any heavy-tailed distribution F on $\mathbb{R}^+$,*

$$\liminf_{x\to\infty} \frac{\overline{F*F}(x)}{\overline{F}(x)} = 2. \tag{2.8}$$

Proof. By the lower bound (2.7), it remains to prove the upper bound only,

$$\liminf_{x\to\infty} \frac{\overline{F*F}(x)}{\overline{F}(x)} \leq 2.$$

Assume the contrary, i.e., there exist $\delta > 0$ and x_0 such that

$$\overline{F*F}(x) \geq (2+\delta)\overline{F}(x) \quad \text{for all } x > x_0. \tag{2.9}$$

Applying Theorem 2.9 with $g(x) = \ln x$, we can choose an increasing concave function $h : \mathbb{R}^+ \to \mathbb{R}^+$ such that $\mathbb{E}e^{h(\xi)} < \infty$ and $\mathbb{E}\xi e^{h(\xi)} = \infty$. For any positive $b > 0$, consider the concave function

$$h_b(x) := \min\{h(x), bx\}.$$

Since F is heavy-tailed, $h(x) = o(x)$ as $x \to \infty$; therefore, for any fixed b there exists x_1 such that $h_b(x) = h(x)$ for all $x > x_1$. Hence, $\mathbb{E}e^{h_b(\xi)} < \infty$ and $\mathbb{E}\xi e^{h_b(\xi)} = \infty$.

For any x, we have the convergence $h_b(x) \downarrow 0$ as $b \downarrow 0$. Then $\mathbb{E}e^{h_b(\xi_1)} \downarrow 1$ as $b \downarrow 0$. Thus there exists b such that

$$\mathbb{E}e^{h_b(\xi_1)} \leq 1 + \delta/4. \tag{2.10}$$

For any real a and t, put $a^{[t]} = \min\{a, t\}$. Then

$$\mathbb{E}(\xi_1^{[t]} + \xi_2^{[t]})e^{h_b(\xi_1+\xi_2)} = 2\mathbb{E}\xi_1^{[t]}e^{h_b(\xi_1+\xi_2)} \leq 2\mathbb{E}\xi_1^{[t]}e^{h_b(\xi_1)+h_b(\xi_2)},$$

by the concavity of the function h_b. Hence,

$$\begin{aligned}\frac{\mathbb{E}(\xi_1^{[t]} + \xi_2^{[t]})e^{h_b(\xi_1+\xi_2)}}{\mathbb{E}\xi_1^{[t]}e^{h_b(\xi_1)}} &\leq 2\frac{\mathbb{E}\xi_1^{[t]}e^{h_b(\xi_1)}\mathbb{E}e^{h_b(\xi_2)}}{\mathbb{E}\xi_1^{[t]}e^{h_b(\xi_1)}} \\ &= 2\mathbb{E}e^{h_b(\xi_2)} \leq 2 + \delta/2, \end{aligned} \tag{2.11}$$

by (2.10). On the other hand, since $(\xi_1+\xi_2)^{[t]} \le \xi_1^{[t]} + \xi_2^{[t]}$,

$$\begin{aligned}\frac{\mathbb{E}(\xi_1^{[t]}+\xi_2^{[t]})e^{h_b(\xi_1+\xi_2)}}{\mathbb{E}\xi_1^{[t]}e^{h_b(\xi_1)}} &\ge \frac{\mathbb{E}(\xi_1+\xi_2)^{[t]}e^{h_b(\xi_1+\xi_2)}}{\mathbb{E}\xi_1^{[t]}e^{h_b(\xi_1)}}\\ &= \frac{\int_0^\infty x^{[t]}e^{h_b(x)}(F*F)(dx)}{\int_0^\infty x^{[t]}e^{h_b(x)}F(dx)}. \end{aligned}\tag{2.12}$$

The right side, after integration by parts, is equal to

$$\frac{\int_0^\infty \overline{F*F}(x)d(x^{[t]}e^{h_b(x)})}{\int_0^\infty \overline{F}(x)d(x^{[t]}e^{h_b(x)})}.$$

Since $\mathbb{E}\xi_1 e^{h_b(\xi_1)} = \infty$, in the latter fraction both the integrals in the numerator and the denominator tend to infinity as $t\to\infty$. For the *increasing* function $h_b(x)$, together with the assumption (2.9) this implies that

$$\liminf_{t\to\infty}\frac{\int_0^\infty \overline{F*F}(x)d(x^{[t]}e^{h_b(x)})}{\int_0^\infty \overline{F}(x)d(x^{[t]}e^{h_b(x)})} \ge 2+\delta.$$

Substituting this into (2.12) we get a contradiction to (2.11) for sufficiently large t. □

It turns out that the '$\liminf$' given by the left side of (2.7) is equal to 2 not only for heavy-tailed, but also for some light-tailed, distributions. Here is an example. Let F be an atomic distribution at the points x_n, $n=0,1,\dots$, with masses p_n, i.e., $F\{x_n\}=p_n$. Suppose that $x_0=1$ and that $x_{n+1}>2x_n$ for every n. Then the tail of the convolution $F*F$ at the point x_n-1 is equal to

$$\begin{aligned}\overline{F*F}(x_n-1) &= (F\times F)([x_n,\infty)\times\mathbb{R}^+) + (F\times F)([0,x_{n-1}]\times[x_n,\infty))\\ &\sim 2\overline{F}(x_n-1) \quad \text{as } n\to\infty.\end{aligned}$$

Hence,

$$\lim_{n\to\infty}\frac{\overline{F*F}(x_n-1)}{\overline{F}(x_n-1)} = 2.$$

From this equality and from (2.7),

$$\liminf_{x\to\infty}\frac{\overline{F*F}(x)}{\overline{F}(x)} = 2. \tag{2.13}$$

Take now $x_n=3^n$, $n=0,1,\dots$, and $p_n=ce^{-3^n}$, where c is the normalising constant. Then F is a light-tailed distribution satisfying the relation (2.13).

We conclude this section with the following result for convolutions of non-identical distributions.

Theorem 2.13. *Let F_1 and F_2 be two distributions on $\mathbb{R}^+$. If the distribution F_1 is heavy-tailed, then*

$$\liminf_{x\to\infty}\frac{\overline{F_1*F_2}(x)}{\overline{F}_1(x)+\overline{F}_2(x)} = 1. \tag{2.14}$$

Proof. By Theorem 2.11, the left side of (2.14) is at least 1. Assume now that it is strictly greater than 1. Then there exists $\varepsilon > 0$ such that, for all sufficiently large x,

$$\frac{\overline{F_1 * F_2}(x)}{\overline{F}_1(x) + \overline{F}_2(x)} \geq 1 + 2\varepsilon. \tag{2.15}$$

Consider the distribution $G = (F_1 + F_2)/2$. This distribution is heavy-tailed. By Theorem 2.12 we get

$$\liminf_{x\to\infty} \frac{\overline{G * G}(x)}{\overline{G}(x)} = 2. \tag{2.16}$$

On the other hand, (2.15) and Theorem 2.11 imply that, for all sufficiently large x,

$$\begin{aligned}
\overline{G * G}(x) &= \frac{\overline{F_1 * F_1}(x) + \overline{F_2 * F_2}(x) + 2\overline{F_1 * F_2}(x)}{4} \\
&\geq \frac{2(1-\varepsilon)\overline{F}_1(x) + 2(1-\varepsilon)\overline{F}_2(x) + 2(1+2\varepsilon)(\overline{F}_1(x) + \overline{F}_2(x))}{4} \\
&= 2(1+\varepsilon/2)\overline{G}(x),
\end{aligned}$$

which contradicts (2.16). □

2.4 Long-tailed functions and their properties

Our plan is to introduce and to study the subclass of heavy-tailed distributions which are *long-tailed*. Later on we will study also long-tailedness properties of other characteristics of distributions. Therefore, we find it reasonable to start with a discussion of some generic properties of long-tailed functions.

Definition 2.14. An ultimately positive function f is *long-tailed* if and only if

$$\lim_{x\to\infty} \frac{f(x+y)}{f(x)} = 1, \quad \text{for all } y > 0. \tag{2.17}$$

Clearly if f is long-tailed then we may also replace y by $-y$ in (2.17).

The following result makes a useful connection.

Lemma 2.15. *The function f is long-tailed if and only if $g(x) := f(\log x)$ (defined for positive x) is slowly varying at infinity, that is, for any fixed $a > 0$,*

$$\frac{g(ax)}{g(x)} \to 1 \ \ \textit{as } x \to \infty.$$

Proof. The proof is immediate from the definition of g since

$$\frac{g(ax)}{g(x)} = \frac{f(\log x + \log a)}{f(\log x)}.$$

□

If f is long-tailed then we also have uniform convergence in (2.17) over y in compact intervals. This is obvious for monotone functions, but in the general case the result follows from the Uniform Convergence Theorem for functions slowly varying at infinity, see Theorem 1.2.1 in [8]. Thus, for any $a > 0$, we have

$$\sup_{|y|\le a} |f(x) - f(x+y)| = o(f(x)) \text{ as } x \to \infty. \tag{2.18}$$

We give some quite basic closure properties for the class of long-tailed functions. We shall make frequent use of these—usually without further comment.

Lemma 2.16. *Suppose that the functions $f_1, \ldots, f_n$ are all long-tailed. Then*

(i) *the function $f_1(c_1 + c_2 x)$ is long-tailed;*
(ii) *if $f \sim \sum_{k=1}^n c_k f_k$ where $c_1, \ldots, c_n > 0$, then f is long-tailed;*
(iii) *the product function $f_1 \cdots f_n$ is long-tailed;*
(iv) *the function $\min(f_1, \ldots, f_n)$ is long-tailed;*
(v) *the function $\max(f_1, \ldots, f_n)$ is long-tailed.*

Proof. The proofs of (i)–(iii) are routine from the definition of long-tailedness.

For (iv) observe that, for any $a > 0$ and any x, we have

$$\begin{aligned}\min\left(\frac{f_1(x+a)}{f_1(x)}, \frac{f_2(x+a)}{f_2(x)}\right) &\le \frac{\min(f_1(x+a),\ f_2(x+a))}{\min(f_1(x),\ f_2(x))}\\ &\le \max\left(\frac{f_1(x+a)}{f_1(x)}, \frac{f_2(x+a)}{f_2(x)}\right),\end{aligned}$$

Since f_1, f_2 are long-tailed the result now follows for the case $n = 2$. The result for general n follows by induction.

For (v) observe that, analogously to the argument for (iv) above, for any $a > 0$ and any x, we have

$$\begin{aligned}\min\left(\frac{f_1(x+a)}{f_1(x)}, \frac{f_2(x+a)}{f_2(x)}\right) &\le \frac{\max(f_1(x+a),\ f_2(x+a))}{\max(f_1(x),\ f_2(x))}\\ &\le \max\left(\frac{f_1(x+a)}{f_1(x)}, \frac{f_2(x+a)}{f_2(x)}\right),\end{aligned}$$

and the result now follows as before. □

We now have the following result.

Lemma 2.17. *Any long-tailed function f is heavy-tailed and, moreover, satisfies the following relation: for every $\lambda > 0$,*

$$\lim_{x\to\infty} f(x)e^{\lambda x} = \infty.$$

Proof. Fix $\lambda > 0$. Since f is long-tailed, $f(x+y) \sim f(x)$ as $x \to \infty$ uniformly in $y \in [0, 1]$. Hence, there exists x_0 such that, for all $x \ge x_0$ and $y \in [0, 1]$,

$$f(x+y) \ge f(x)e^{-\lambda/2}.$$

Then $f(x_0+n+y) \ge f(x_0)e^{-\lambda(n+1)/2}$ for all $n \ge 1$ and $y \in [0,1]$, and, therefore,

$$\liminf_{x\to\infty} f(x)e^{\lambda x} \ge f(x_0) \lim_{n\to\infty} e^{-\lambda(n+1)/2}e^{\lambda n} = \infty,$$

so that the lemma now follows. □

However, it is not difficult to construct a heavy-tailed function f which is not long-tailed. Put

$$f(x) = \sum_{n=1}^{\infty} 2^{-n}\mathbb{I}\{2^{n-1} < x \le 2^n\}.$$

Then, for any $\lambda > 0$,

$$\limsup_{x\to\infty} f(x)e^{\lambda x} \ge \limsup_{n\to\infty} 2^{-n}e^{\lambda 2^n} = \infty,$$

so that f is heavy-tailed. On the other hand,

$$\liminf_{x\to\infty} \frac{f(x+1)}{f(x)} \le \liminf_{n\to\infty} \frac{f(2^n+1)}{f(2^n)} = \frac{1}{2},$$

which shows that f is not long-tailed.

h-insensitivity. We now introduce a very important concept of which we shall make frequent subsequent use.

Definition 2.18. Given a strictly positive non-decreasing function h, an ultimately positive function f is called *h-insensitive* (or *h-flat*) if

$$\sup_{|y|\le h(x)} |f(x+y)-f(x)| = o(f(x)) \quad \text{as } x\to\infty, \text{ uniformly in } |y| \le h(x). \tag{2.19}$$

It is clear that the relation (2.19) implies that the function f is long-tailed, and conversely that any long-tailed function is h-insensitive for any constant function h. The following lemma gives a strong converse result, which we shall use repeatedly in Section 2.7 and subsequently throughout the monograph.

Lemma 2.19. *Suppose that the function f is long-tailed. Then there exists a function h such that $h(x)\to\infty$ as $x\to\infty$ and f is h-insensitive.*

Proof. For any integer $n \ge 1$, by (2.18), we can choose x_n such that

$$\sup_{|y|\le n} |f(x+y)-f(x)| \le f(x)/n \quad \text{for all } x > x_n.$$

Without loss of generality we may assume that the sequence $\{x_n\}$ is increasing to infinity. Put $h(x) = n$ for $x \in [x_n, x_{n+1}]$. Since $x_n \to \infty$ as $n\to\infty$, we have $h(x)\to\infty$ as $x\to\infty$. By the construction we have

$$\sup_{|y|\le h(x)} |f(x+y)-f(x)| \le f(x)/n$$

for all $x > x_n$, which completes the proof. □

One important use of h-insensitivity is the following. The "natural" definition of long-tailedness of a function f is that of h-insensitivity with respect to any constant function $h(x) = a$ for all x and some $a > 0$. The use of this property in this form would then require that both the statements and the proofs of many results would involve a double limiting operation in which first x was allowed to tend to infinity, with the use of the relation (2.18), and following which a was allowed to tend to infinity. The replacement of the constant a by a function h itself increasing to infinity, but sufficiently slowly that the long-tailed function f is h-insensitive, not only enables two limiting operations to be replaced with a single one in proofs, but also permits simpler, cleaner and more insightful presentations of many results (a typical example is the all-important Lemma 2.33 in Section 2.7).

Now observe that if a long-tailed function f is h-insensitive for some function h and if a further positive non-decreasing function $\hat{h}$ is such that $\hat{h}(x) \le h(x)$ for all x, then (by definition) f is also $\hat{h}$-insensitive. Two trivial, but important (and frequently used), consequences of the combination of this observation with Lemma 2.19 are given by the following proposition.

Proposition 2.20.

(i) *Given a finite collection of long-tailed functions $f_1, \ldots, f_n$, we may choose a* single *function h with respect to which each of the functions f_i is h-insensitive;*

(ii) *given any long-tailed function f and any positive non-decreasing function $\hat{h}$, we may choose a function h such that $h(x) \le \hat{h}(x)$ for all x and f is h-insensitive.*

Proof. For (i), note that for each i we may choose a function h_i such that f_i is h_i-insensitive, and then define h by $h(x) = \min_i h_i(x)$.

For (ii), note that we may take $h(x) = \min(\hat{h}(x), \bar{h}(x))$ where $\bar{h}$ is such that f is $\bar{h}$-insensitive. □

Finally we note that a further important use of h-insensitivity is the following. For any given positive function h, increasing to ∞, we may consider the class of those distributions whose (necessarily long-tailed) tail functions are h-insensitive. For varying h, this gives a powerful method for the classification of such distributions, which we explore in detail in Section 2.8.

2.5 Long-tailed distributions

As we discussed in the Introduction, all heavy-tailed distributions likely to be encountered in practical applications are sufficiently regular as to be long-tailed, and it is the latter property, as applied to distributions, which we study in this section.

First, for any distribution F on $\mathbb{R}$, recall that we denote by R the hazard function $R(x) := -\ln \overline{F}(x)$. By definition, R is always a non-decreasing function and

$$R(x+1) - R(x) = -\ln \frac{\overline{F}(x+1)}{\overline{F}(x)}.$$

Definition 2.21. A distribution F on $\mathbb{R}$ is called *long-tailed* if $\overline{F}(x) > 0$ for all x and, for any fixed $y > 0$,

$$\overline{F}(x+y) \sim \overline{F}(x) \qquad \text{as } x \to \infty \tag{2.20}$$

That is, the distribution F is long-tailed if and only if its tail function $\overline{F}$ is a long-tailed function. Note that in (2.20) we may again replace y by $-y$. Further, for a distribution F to be long-tailed it is sufficient to require (2.20) to hold for any one non-zero value of y. Note also that the convergence in (2.20) is again uniform over y in compact intervals.

We shall write $\mathcal{L}$ for the class of long-tailed distributions on $\mathbb{R}$. Clearly $F \in \mathcal{L}$ is a *tail property* of the distribution F, since it depends only on $\{\overline{F}(x) : x \geq x_0\}$ for any finite x_0. Further, it follows from Lemma 2.17 that if the distribution F is long-tailed ($F \in \mathcal{L}$) then $\overline{F}$ is a heavy-tailed function, and so, by Theorem 2.6, F is also a heavy-tailed distribution. However, as the example following Lemma 2.17 shows, a heavy-tailed distribution need not be long-tailed.

The following lemma gives some readily verified equivalent characterisations of long-tailedness.

Lemma 2.22. *Let F be a distribution on $\mathbb{R}$ with right-unbounded support, and let ξ be a random variable with distribution F. Then the following are equivalent:*

(i) *the distribution F is long-tailed ($F \in \mathcal{L}$);*
(ii) *for any fixed $y > 0$, $F(x, x+y] = o(\overline{F}(x))$ as $x \to \infty$;*
(iii) *for any fixed $y > 0$, $\mathbb{P}\{\xi > x + y \,|\, \xi > x\} \to 1$ as $x \to \infty$;*
(iv) *the hazard function $R(x)$ satisfies $R(x+1) - R(x) \to 0$ as $x \to \infty$.*

Analogously to Lemma 2.16 we further have the following result.

Lemma 2.23. *Suppose that the distributions $F_1, \ldots, F_n$ are all long-tailed (i.e. belong to the class $\mathcal{L}$) and that $\xi_1, \ldots, \xi_n$ are random variables with distributions $F_1, \ldots, F_n$ respectively. Then*

(i) *for any constants c_1 and $c_2 > 0$, the distribution of $c_2\xi_1 + c_1$ is long tailed;*
(ii) *if $\overline{F}(x) \sim \sum_{k=1}^{n} c_k \overline{F}_k(x)$ where $c_1, \ldots, c_n > 0$, then F is long-tailed;*
(iii) *if $F(x) = \min(F_1(x), \ldots, F_n(x))$, then F is long-tailed;*
(iv) *if $F(x) = \max(F_1(x), \ldots, F_n(x))$, then F is long-tailed;*
(v) *the distribution of $\min(\xi_1, \ldots, \xi_n)$ is long-tailed;*
(vi) *the distribution of $\max(\xi_1, \ldots, \xi_n)$ is long-tailed.*

Proof. The proofs follow from the application of Lemma 2.16 to the corresponding tail functions. In particular (v) and (vi) follow from (i) and (iii) of Lemma 2.16. □

2.6 Long-tailed distributions and integrated tails

In the study of random walks in particular, a key role is played by the integrated tail distribution, the fundamental properties of which we introduce in this section.

Definition 2.24. For any distribution F on $\mathbb{R}$ such that

$$\int_0^\infty \overline{F}(y)\,dy < \infty, \tag{2.21}$$

(and hence $\int_x^\infty \overline{F}(y)\,dy < \infty$ for any finite x) we define the *integrated tail distribution* F_I via its tail function by

$$\overline{F}_I(x) = \min\Big(1, \int_x^\infty \overline{F}(y)dy\Big). \tag{2.22}$$

Note that if ξ is a random variable with distribution F then

$$\int_x^\infty \overline{F}(y)dy = \mathbb{E}\{\xi; \xi > x\} - x\mathbb{P}\{\xi > x\} = \mathbb{E}\{\xi - x; \xi > x\}. \tag{2.23}$$

The following characterisation will frequently be useful.

Lemma 2.25. *Suppose that the distribution F is such that (2.21) holds. Then F_I is long-tailed if and only if $\overline{F}(x) = o(\overline{F}_I(x))$ as $x \to \infty$.*

Proof. The integrated tail distribution F_I is long-tailed ($F_I \in \mathcal{L}$) if and only if $\overline{F}_I(x) - \overline{F}_I(x+1) = o(\overline{F}_I(x))$, or, equivalently, $\overline{F}_I(x) - \overline{F}_I(x+1) = o(\overline{F}_I(x+1))$. The required result now follows from the inequalities

$$\overline{F}(x+1) \le \overline{F}_I(x) - \overline{F}_I(x+1) \le \overline{F}(x),$$

valid for all sufficiently large x. □

Lemma 2.26. *Suppose that the distribution F is long-tailed ($F \in \mathcal{L}$) and such that (2.21) holds. Then F_I is long-tailed as well ($F_I \in \mathcal{L}$) and $\overline{F}(x) = o(\overline{F}_I(x))$ as $x \to \infty$.*

Proof. The long-tailedness of F_I follows from the relations, as $x \to \infty$,

$$\overline{F}_I(x+t) = \int_x^\infty \overline{F}(x+t+y)dy \sim \int_x^\infty \overline{F}(x+y)dy = \overline{F}_I(x),$$

for any fixed t. That $\overline{F}(x) = o(\overline{F}_I(x))$ as $x \to \infty$ now follows from Lemma 2.25. □

The converse assertion, that is, that long-tailedness of F_I implies long-tailedness of F, is not in general true. This is illustrated by the following example.

Example 2.27. Let the distribution F be such that $\overline{F}(x) = 2^{-2n}$ for $x \in [2^n, 2^{n+1})$. Then F is not long-tailed since $\overline{F}(2^n - 1)/\overline{F}(2^n) = 4$ for any n, so that $\overline{F}(x-1)/\overline{F}(x) \not\to 1$ as $x \to \infty$. But we have $x^{-2} \le \overline{F}(x) \le 4x^{-2}$ for any $x > 0$. In particular, $\overline{F}_I(x) \ge x^{-1}$ and thus $\overline{F}(x) = o(\overline{F}_I(x))$ as $x \to \infty$. Thus, by Lemma 2.25, F_I is long-tailed.

We now formulate a more general result which will be needed in the theory of random walks with heavy-tailed increments, and is also of some interest in its own right. Let F be a distribution on $\mathbb{R}$ and μ a non-negative measure on $\mathbb{R}^+$ such that

$$\int_0^\infty \overline{F}(t)\mu(dt) < \infty. \tag{2.24}$$

We may then define the distribution F_μ on $\mathbb{R}^+$ given by

$$\overline{F}_\mu(x) := \min\Big(1, \int_0^\infty \overline{F}(x+t)\mu(dt)\Big), \quad x \ge 0. \tag{2.25}$$

If μ is the Lebesgue measure, then F_μ is the integrated tail distribution. We can formulate the same question as for F_I: what type of conditions on F imply long-tailedness of F_μ? The answer is given by the following theorem.

Theorem 2.28. *If F is a long-tailed distribution, then F_μ is a long-tailed distribution and, for any fixed $y > 0$,*

$$\overline{F}_\mu(x+y) \sim \overline{F}_\mu(x)$$

as $x \to \infty$ uniformly in all μ satisfying (2.24), that is,

$$\inf_\mu \inf_{x>x_0} \frac{\overline{F}_\mu(x+y)}{\overline{F}_\mu(x)} \to 1 \quad \text{as } x_0 \to \infty. \tag{2.26}$$

If, in addition, $\overline{F}(x \pm h(x)) \sim \overline{F}(x)$ as $x \to \infty$, for some positive function h, then (2.26) holds with $\pm h(x)$ in place of y.

Proof. Fix $\varepsilon > 0$. Since $\overline{F}(x+y+u) \sim \overline{F}(x+u)$ as $x \to \infty$ uniformly in $u \geq 0$, there exists x_0 such that

$$\overline{F}_\mu(x+y) \geq (1-\varepsilon)\overline{F}_\mu(x) \quad \text{for all } x > x_0.$$

Then, for all $x > x_0$ and μ,

$$\overline{F_\mu}(x+y) = \int_0^\infty \overline{F}(x+y+u)\mu(dy) \geq (1-\varepsilon)\int_0^\infty \overline{F}(x+u)\mu(du) = \overline{F_\mu}(x).$$

Letting $\varepsilon \to 0$ we obtain the desired result. The same argument holds when y is replaced by $\pm h(x)$. □

2.7 Convolutions of long-tailed distributions

We know from Theorem 2.11 that for any distributions F and G on the positive half-line $\mathbb{R}^+$

$$\liminf_{x\to\infty} \frac{\overline{F*G}(x)}{\overline{F}(x)+\overline{G}(x)} \geq 1. \tag{2.27}$$

In order to get the same result for distributions on the whole real line $\mathbb{R}$, we assume some of distributions involved to be long-tailed. The assumption of the theorem below seems to be the weakest possible.

Theorem 2.29. *Let the distributions $F_1, \ldots, F_n$ on $\mathbb{R}$ be such that the function $\overline{F}_1(x) + \ldots + \overline{F}_n(x)$ is long-tailed. Then, for any distribution G,*

$$\liminf_{x\to\infty} \frac{\overline{F_1 * \ldots * F_n * G}(x)}{\overline{F}_1(x)+\ldots+\overline{F}_n(x)} \geq 1. \tag{2.28}$$

In particular (2.28) holds whenever each of the distributions F_i is long-tailed.

Proof (cf Theorem 2.11). Let $\xi_1, \ldots, \xi_n$ and η be independent random variables with respective distributions $F_1, \ldots, F_n$ and G. For any fixed $a > 0$, we have the following lower bound:

$$\begin{aligned}\overline{F_1 * \ldots * F_n * G}(x) &\geq \sum_{k=1}^n \mathbb{P}\{\xi_k > x+na, \xi_j \in (-a,x] \text{ for all } j \neq k, \eta > -a\} \\ &= \overline{G}(-a)\sum_{k=1}^n \overline{F}_k(x+na)\prod_{j\neq k} F_j(-a,x]. \end{aligned} \tag{2.29}$$

For every $\varepsilon > 0$ there exists a such that $F_j(-a, a] \geq 1 - \varepsilon$ for all j and $\overline{G}(-a) > 1 - \varepsilon$. Thus, for all $x > a$,

$$\overline{F_1 * \ldots * F_n * G}(x) \geq (1-\varepsilon)^n \sum_{k=1}^{n} \overline{F}_k(x + na).$$

Since the function $\overline{F}_1 + \ldots + \overline{F}_n$ is long-tailed,

$$\liminf_{x\to\infty} \frac{\overline{F_1 * \ldots * F_n * G}(x)}{\overline{F}_1(x) + \ldots + \overline{F}_n(x)} \geq (1-\varepsilon)^n.$$

The required result (2.28) now follows by letting $\varepsilon \to 0$. □

For identical distributions, Theorem 2.29 yields the following corollaries.

Corollary 2.30. *Let the distribution F on $\mathbb{R}$ be long-tailed ($F \in \mathcal{L}$). Then, for any $n \geq 2$,*

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

Corollary 2.31. *Let the distribution F on $\mathbb{R}$ be F is long-tailed ($F \in \mathcal{L}$) and let the distribution G be such that $\overline{G}(a) = 0$ for some a. Then $\overline{F * G}(x) \sim \overline{F}(x)$ as $x \to \infty$.*

Proof. Since $\overline{G}(a) = 0$, we have $\overline{F * G}(x) \leq \overline{F}(x - a)$. Thus since F is long-tailed we have

$$\limsup_{x\to\infty} \frac{\overline{F * G}(x)}{\overline{F}(x)} \leq 1.$$

Combining this result with the lower bound of Lemma 2.29 in the case $n = 1$, we obtain the desired equivalence. □

In order to further study convolutions of long-tailed distributions, we make repeated use of two fundamental decompositions. Let $h > 0$ and let ξ and η be independent random variables with distributions F and G respectively. Then the tail function of the convolution of F and G possesses the following decomposition: for $x > 0$,

$$\overline{F * G}(x) = \mathbb{P}\{\xi + \eta > x, \xi \leq h\} + \mathbb{P}\{\xi + \eta > x, \xi > h\}. \tag{2.30}$$

If in addition $h \leq x/2$ then

$$\begin{aligned}&\overline{F * G}(x)\\ &= \mathbb{P}\{\xi + \eta > x, \xi \leq h\} + \mathbb{P}\{\xi + \eta > x, \eta \leq h\} + \mathbb{P}\{\xi + \eta > x, \xi > h, \eta > h\},\end{aligned} \tag{2.31}$$

since if $\xi \leq h$ and $\eta \leq h$ then $\xi + \eta \leq 2h \leq x$.

Note that

$$\mathbb{P}\{\xi + \eta > x, \xi \leq h\} = \int_{-\infty}^{h} \overline{G}(x - y) F(dy), \tag{2.32}$$

while the probability of the event $\{\xi + \eta > x, \xi > h, \eta > h\}$ is symmetric in F and G and

$$\begin{aligned}\mathbb{P}\{\xi + \eta > x, \xi > h, \eta > h\} &= \int_{h}^{\infty} \overline{F}(\max(h, x - y)) G(dy)\\ &= \int_{h}^{\infty} \overline{G}(\max(h, x - y)) F(dy).\end{aligned} \tag{2.33}$$

Definition 2.32. Given a strictly positive non-decreasing function h, a distribution F on $\mathbb{R}$ is called h-*insensitive* if its tail function $\overline{F}$ is an h-insensitive function (see Definition 2.18). Since $\overline{F}$ is monotone, this reduces to the requirement that $\overline{F}(x \pm h(x)) \sim \overline{F}(x)$ as $x \to \infty$.

Recall from the results for h-insensitive functions that a distribution F is long-tailed if and only if there exists a function h as above with respect to which F is h-insensitive.

For long-tailed distributions F and G we shall now make particular use of the decomposition (2.31) in which the constant h is replaced by a function h increasing to infinity (and with $h(x) < x/2$ for all x) and such that both F and G are h-insensitive.

The following three lemmas are the keys to everything that follows later in this section.

Lemma 2.33. *Suppose that the distribution G on $\mathbb{R}$ is long-tailed ($G \in \mathcal{L}$) and that the positive function h is such that $h(x) \to \infty$ as $x \to \infty$ and G is h-insensitive. Then, for any distribution F, as $x \to \infty$,*

$$\int_{-\infty}^{h(x)} \overline{G}(x-y)F(dy) \sim \overline{G}(x),$$
$$\int_{x-h(x)}^{\infty} \overline{F}(x-y)G(dy) \sim \overline{G}(x).$$

Proof. The existence of the function h is guaranteed by Lemma 2.19. We now have

$$\int_{-\infty}^{h(x)} \overline{G}(x-y)F(dy) \le \overline{G}(x-h(x)).$$

On the other hand we also have,

$$\begin{aligned}\int_{-\infty}^{h(x)} \overline{G}(x-y)F(dy) &\ge \int_{-h(x)}^{h(x)} \overline{G}(x-y)F(dy)\\ &\ge F(-h(x),h(x)]\overline{G}(x+h(x))\\ &\sim \overline{G}(x+h(x)) \quad \text{as } x \to \infty,\end{aligned}$$

where the last equivalence follows since $h(x) \to \infty$ as $x \to \infty$. The first result now follows from the choice of the function h. The second result follows entirely similarly: the integral is again bounded from above by $\overline{G}(x-h(x))$ and from below by $F(-h(x),h(x)]\overline{G}(x+h(x))$ and the result follows as previously. □

Remark 2.34. Note the crucial role played by the monotonicity of the tail function $\overline{G}$ in the proof of Lemma 2.33—something which is not available to us in considering, for example, densities in Chapter 4.

We now prove a version of Lemma 2.33 which is symmetric in the distributions F and G, and which allows us to get many important results for convolutions—see the further discussion below.

Lemma 2.35. *Suppose that the distributions F and G on $\mathbb{R}$ are such that the sum $\overline{F}+\overline{G}$ of their tail functions is a long-tailed function (equivalently the measure $F+G$ is long-tailed in the obvious sense) and that the positive function h is such that $h(x) \to \infty$ as $x \to \infty$ and $\overline{F}+\overline{G}$ is h-insensitive. Then*

$$\int_{-\infty}^{h(x)} \overline{G}(x-y)F(dy) + \int_{-\infty}^{h(x)} \overline{F}(x-y)G(dy) \sim \overline{G}(x)+\overline{F}(x) \quad \text{as } x \to \infty.$$

Proof. The proof is simply a two-sided version of that for the first assertion of Lemma 2.33. The existence of the function h is again guaranteed by Lemma 2.19. Now note first that

$$\int_{-\infty}^{h(x)} \overline{G}(x-y)F(dy) + \int_{-\infty}^{h(x)} \overline{F}(x-y)G(dy) \le \overline{G}(x-h(x)) + \overline{F}(x-h(x)),$$

and second that

$$\begin{aligned}\int_{-\infty}^{h(x)} &\overline{G}(x-y)F(dy) + \int_{-\infty}^{h(x)} \overline{F}(x-y)G(dy) \\ &\ge \int_{-h(x)}^{h(x)} \overline{G}(x-y)F(dy) + \int_{-h(x)}^{h(x)} \overline{F}(x-y)G(dy) \\ &\ge F(-h(x),h(x)]\overline{G}(x+h(x)) + G(-h(x),h(x)]\overline{F}(x+h(x)) \\ &\sim \overline{G}(x+h(x)) + \overline{F}(x+h(x)) \quad \text{as } x\to\infty,\end{aligned}$$

where the last equivalence follows since $h(x)\to\infty$ as $x\to\infty$. The required result now follows from the choice of the function h. □

Note that special cases under which $\overline{F}+\overline{G}$ is long-tailed are (a) F and G are both long-tailed—in which case Lemma 2.35 (almost) follows from 2.33, and (b) F is long-tailed and $\overline{G}(x) = o(\overline{F}(x))$ as $x\to\infty$.

In various calculations we need to estimate the "internal" part of the convolution. The following result will be useful.

Lemma 2.36. *Let h be any increasing function on $\mathbb{R}^+$ such that $h(x)\to\infty$. Then, for any distributions F_1, F_2, G_1 and G_2 on $\mathbb{R}$,*

$$\limsup_{x\to\infty} \frac{\mathbb{P}\{\xi_1+\eta_1>x, \xi_1>h(x), \eta_1>h(x)\}}{\mathbb{P}\{\xi_2+\eta_2>x, \xi_2>h(x), \eta_2>h(x)\}} \le \limsup_{x\to\infty}\frac{\overline{F_1}(x)}{\overline{F_2}(x)}\cdot\limsup_{x\to\infty}\frac{\overline{G_1}(x)}{\overline{G_2}(x)},$$

where ξ_1, ξ_2, η_1, and η_2 are independent random variables with respective distributions F_1, F_2, G_1, and G_2.

In particular, in the case where the limits of the ratios $\overline{F_1}(x)/\overline{F_2}(x)$ and $\overline{G_1}(x)/\overline{G_2}(x)$ exist, we have

$$\lim_{x\to\infty} \frac{\mathbb{P}\{\xi_1+\eta_1>x, \xi_1>h(x), \eta_1>h(x)\}}{\mathbb{P}\{\xi_2+\eta_2>x, \xi_2>h(x), \eta_2>h(x)\}} = \lim_{x\to\infty}\frac{\overline{F_1}(x)}{\overline{F_2}(x)}\cdot\lim_{x\to\infty}\frac{\overline{G_1}(x)}{\overline{G_2}(x)}.$$

Proof. It follows from (2.33) that

$$\begin{aligned}\mathbb{P}\{\xi_1&+\eta_1>x, \xi_1>h(x), \eta_1>h(x)\} \\ &\le \sup_{z>h(x)}\frac{\overline{F_1}(z)}{\overline{F_2}(z)}\int_{h(x)}^{\infty}\overline{F}_2(\max(h(x),x-y))G_1(dy) \\ &= \sup_{z>h(x)}\frac{\overline{F_1}(z)}{\overline{F_2}(z)}\int_{h(x)}^{\infty}\overline{G}_1(\max(h(x),x-y))F_2(dy).\end{aligned}$$

Similarly,

$$\begin{aligned}
\int_{h(x)}^{\infty} & \overline{G}_1(\max(h(x), x-y))F_2(dy) \\
&\le \sup_{z>h(x)} \frac{\overline{G_1}(z)}{\overline{G_2}(z)} \int_{h(x)}^{\infty} \overline{G}_2(\max(h(x), x-y))F_2(dy) \\
&= \sup_{z>h(x)} \frac{\overline{G_1}(z)}{\overline{G_2}(z)} \mathbb{P}\{\xi_2+\eta_2 > x, \xi_2 > h(x), \eta_2 > h(x)\}.
\end{aligned}$$

Combining these results and recalling that $h(x) \to \infty$ as $x \to \infty$, we obtain the desired conclusion. □

Definition 2.37. Two distributions F and G with right-unbounded supports are said to *tail-equivalent* if $\overline{F}(x) \sim \overline{G}(x)$ as $x \to \infty$ (i.e. $\lim_{x\to\infty} \overline{F}(x)/\overline{G}(x) = 1$).

In the next two theorems we provide conditions under which a random shifting preserves tail equivalence.

Theorem 2.38. *Suppose that F_1, F_2, and G are distributions on $\mathbb{R}$ such that $\overline{F}_1(x) \sim \overline{F}_2(x)$ as $x \to \infty$. If G is long-tailed then $\overline{F_1 * G}(x) \sim \overline{F_2 * G}(x)$ as $x \to \infty$.*

Proof. By Lemma 2.19 we can find a function h such that $h(x) \to \infty$ and

$$\overline{G}(x \pm h(x)) \sim \overline{G}(x) \quad \text{as } x \to \infty,$$

i.e. G is h-insensitive. We use the following decomposition: for $k = 1, 2$,

$$\overline{F_k * G}(x) = \Big(\int_{-\infty}^{x-h(x)} + \int_{x-h(x)}^{\infty}\Big)\overline{F}_k(x-y)G(dy). \tag{2.34}$$

It follows from the tail equivalence of F_1 and F_2 that $\overline{F}_1(x-y) \sim \overline{F}_2(x-y)$ as $x \to \infty$ uniformly in $y < x - h(x)$. Thus,

$$\int_{-\infty}^{x-h(x)} \overline{F}_1(x-y)G(dy) \sim \int_{-\infty}^{x-h(x)} \overline{F}_2(x-y)G(dy) \tag{2.35}$$

as $x \to \infty$. Next, by Lemma 2.33, for $k = 1, 2$,

$$\int_{x-h(x)}^{\infty} \overline{F}_k(x-y)G(dy) \sim \overline{G}(x) \quad \text{as } x \to \infty. \tag{2.36}$$

Substituting (2.35) and (2.36) into (2.34) we obtain the required equivalence $\overline{F_1 * G}(x) \sim \overline{F_2 * G}(x)$. □

Theorem 2.39. *Suppose that F_1, F_2, G_1, and G_2 are distributions on $\mathbb{R}$ such that $\overline{F}_1(x) \sim \overline{F}_2(x)$ and $\overline{G}_1(x) \sim \overline{G}_2(x)$ as $x \to \infty$. If the function $\overline{F}_1 + \overline{G}_1$ is long-tailed then $\overline{F_1 * G_1}(x) \sim \overline{F_2 * G_2}(x)$ as $x \to \infty$.*

Proof. The conditions of the theorem imply that the function $\overline{F}_2 + \overline{G}_2$ is similarly long-tailed. By Lemma 2.19 and the following remark we can choose a function h such that $h(x) \to \infty$ as $x \to \infty$, $h(x) \le x/2$ and, for $k = 1, 2$,

$$\overline{F}_k(x \pm h(x)) + \overline{G}_k(x \pm h(x)) \sim \overline{F}_k(x) + \overline{G}_k(x) \quad \text{as } x \to \infty,$$

i.e. $\overline{F}_k + \overline{G}_k$ is h-insensitive. We use the following decomposition which follows from (2.31)–(2.33):

$$\begin{aligned}\overline{F_k * G_k}(x) &= \int_{-\infty}^{h(x)} \overline{F}_k(x-y)G_k(dy) + \int_{-\infty}^{h(x)} \overline{G}_k(x-y)F_k(dy) \\ &\quad + \int_{h(x)}^{\infty} \overline{F}_k(\max(h(x), x-y))G_k(dy). \end{aligned} \tag{2.37}$$

Since F_1 and F_2 are tail equivalent and G_1 and G_2 are tail equivalent, it follows from Lemma 2.36 that, as $x \to \infty$,

$$\int_{h(x)}^{\infty} \overline{F}_1(\max(h(x), x-y))G_1(dy) \sim \int_{h(x)}^{\infty} \overline{F}_2(\max(h(x), x-y))G_2(dy). \tag{2.38}$$

Further, by Lemma 2.35, for $k = 1, 2$ and as $x \to \infty$,

$$\int_{-\infty}^{h(x)} \overline{F}_k(x-y)G_k(dy) + \int_{-\infty}^{h(x)} \overline{G}_k(x-y)F_k(dy) \sim \overline{F}_k(x) + \overline{G}_k(x). \tag{2.39}$$

Substituting (2.38) and (2.39) into (2.37) we obtain the required equivalence $\overline{F_1 * G_1}(x) \sim \overline{F_2 * G_2}(x)$. □

We now use Theorem 2.39 to show that the class $\mathcal{L}$ is closed under convolutions. This is a corollary of the following result.

Theorem 2.40. *Suppose that the distributions F and G are such that F is long-tailed and the measure $F + G$ is also long-tailed (i.e. the sum $\overline{F} + \overline{G}$ of the tail functions of the two distributions is long-tailed). Then the convolution $F * G$ is also long-tailed.*

Proof. Fix $y > 0$. Take $F_1 = F$ and F_2 to be equal to F shifted by $-y$, that is, $\overline{F}_2(x) = \overline{F}(x+y)$. Then $F_2 * G$ is equal to $F * G$ shifted by $-y$. Since F is long-tailed, $\overline{F}_1(x) \sim \overline{F}_2(x)$. Since also $\overline{F}_1 + \overline{G}$ is long-tailed, it follows from Theorem 2.39 with $G_1 = G_2 = G$ that $\overline{F_1 * G}(x) \sim \overline{F_2 * G}(x)$. Hence $\overline{F * G}(x) \sim \overline{F * G}(x+y)$ as $x \to \infty$. □

Both the following corollaries are now immediate from Theorem 2.40 since in each case the measure $F + G$ is long-tailed.

Corollary 2.41. *Let the distributions F and G be long-tailed. Then the convolution $F * G$ is also long-tailed.*

Corollary 2.42. *Suppose that F and G are distributions and that F is long-tailed. If $\overline{G}(x) = o(\overline{F}(x))$ as $x \to \infty$, then $F * G$ is long-tailed.*

Finally in this section we have the following converse result.

Lemma 2.43. *Let F and G be two distributions on $\mathbb{R}^+$ such that F has unbounded support and G is non-degenerate at 0. If $\overline{G}(x) \leq c\overline{F}(x)$ for some $c < \infty$ and*

$$\limsup_{x\to\infty} \frac{\overline{F * G}(x)}{\overline{F}(x) + \overline{G}(x)} \leq 1, \tag{2.40}$$

then F is long-tailed.

Proof. Take any a such that $G(a, \infty) > 0$ which is possible because G is not concentrated at 0. Since for any two distributions on $\mathbb{R}^+$

$$\overline{F * G}(x) = \int_0^x \overline{F}(x-y)G(dy) + \overline{G}(x),$$

it follows from the condition (2.40) that

$$\begin{aligned}\int_0^x \overline{F}(x-y)G(dy) &\leq \overline{F}(x) + o(\overline{F}(x) + \overline{G}(x)) \\ &= \overline{F}(x) + o(\overline{F}(x)) \quad \text{as } x \to \infty,\end{aligned}$$

due to the condition $\overline{G}(x) \leq c\overline{F}(x)$. This implies that

$$\begin{aligned}\int_0^x F(x-y, x]G(dy) &= \int_0^x (\overline{F}(x-y) - \overline{F}(x))G(dy) \\ &= o(\overline{F}(x)) \quad \text{as } x \to \infty.\end{aligned}$$

The left side is not less than $F(x-a, x]G(a, x]$, hence $F(x-a, x] = o(\overline{F}(x))$ as $x \to \infty$. The latter relation is equivalent to $\overline{F}(x-a) \sim \overline{F}(x)$ which completes the proof. □

2.8 h-insensitive distributions

Let F be a long-tailed distribution ($F \in \mathcal{L}$), i.e. a distribution whose tail function $\overline{F}$ is such that for some (and hence for all) non-zero y, we have $\overline{F}(x+y) \sim \overline{F}(x)$ as $x \to \infty$. We saw in Lemma 2.19 that we can then find a non-decreasing positive function h such that $h(x) \to \infty$ as $x \to \infty$ and

$$\overline{F}(x+y) \sim \overline{F}(x) \quad \text{uniformly in } |y| \leq h(x), \tag{2.41}$$

i.e. such that the distribution F is h-insensitive (see Definition 2.32).

In this section we turn this process around: we fix a positive function h which is increasing to infinity, and seek to identify those long-tailed distributions which are h-insensitive. By varying the choice of h, we then have an important technique for classifying long-tailed distributions according to the heaviness of their tails and for establishing characteristic properties of various classes of these distributions.

As a first example, consider the function h given by $h(x) = \varepsilon x$ for some $\varepsilon > 0$; then the class of h-insensitive distributions coincides with the class of distributions whose tails are *slowly varying at infinity*, that is, for any $\varepsilon > 0$,

$$\frac{\overline{F}((1+\varepsilon)x)}{\overline{F}(x)} \to 1 \quad \text{as } x \to \infty. \tag{2.42}$$

These distributions are extremely heavy; in particular they do not possess any finite positive moments, that is, $\int x^\gamma F(dx) = \infty$ for any $\gamma > 0$. Examples are given by distributions F with the following tail functions:

$$\overline{F}(x) \sim 1/\ln^\gamma x, \quad \overline{F}(x) \sim 1/(\ln\ln x)^\gamma \text{ as } x \to 0, \quad \gamma > 0.$$

Regularly varying distributions We introduce here the well-known class of regularly varying distributions, and consider their insensitivity properties.

Definition 2.44. An ultimately positive function f is called *regularly varying at infinity with index* $\alpha \in \mathbb{R}$ if, for any fixed $c > 0$,

$$f(cx) \sim c^\alpha f(x) \quad \text{as } x \to \infty. \tag{2.43}$$

A distribution F on $\mathbb{R}$ is called *regularly varying at infinity with index* $-\alpha < 0$ if $\overline{F}(cx) \sim c^{-\alpha}\overline{F}(x)$ as $x \to \infty$, that is, $\overline{F}(x)$ is regularly varying at infinity with index $-\alpha < 0$.

Particular examples of regularly varying distributions which were introduced in Section 2.1 are the the *Pareto*, *Burr*, and *Cauchy* distributions.

If a distribution F on $\mathbb{R}^+$ is regularly varying at infinity with index $-\alpha < 0$, then all moments of order $\gamma < \alpha$ are finite, while all moments of order $\gamma > \alpha$ are infinite. The moment of order $\gamma = \alpha$ may be finite or infinite depending on the particular behaviour of the corresponding slowly varying function (see below).

If a function f is regularly varying at infinity with index α then we have $f(x) = x^\alpha l(x)$ for some slowly varying function l. Hence it follows from the discussion of Section 2.4 that, for any positive function h such that $h(x) = o(x)$ as $x \to \infty$, we have $f(x+y) \sim f(x)$ as $x \to \infty$ uniformly in $|y| \le h(x)$; we shall then say that f is $o(x)$-*insensitive*. Similarly we shall say that a distribution F is $o(x)$-*insensitive* if its tail function $\overline{F}$ is $o(x)$-insensitive. Thus distributions which are regularly varying at infinity are $o(x)$-insensitive.

Intermediate regularly varying distributions It turns out that the property of $o(x)$-insensitivity characterises a slightly wider class of distributions than that of distributions whose tails are regularly varying, and we now discuss this.

Definition 2.45. A distribution F on $\mathbb{R}$ is called *intermediate regularly varying* if

$$\lim_{\varepsilon\downarrow 0}\liminf_{x\to\infty}\frac{\overline{F}(x(1+\varepsilon))}{\overline{F}(x)} = 1. \tag{2.44}$$

Any regularly varying distribution is intermediate regularly varying. But the latter class is richer. We provide first a simple example. Take any density function g which is regularly varying at infinity with index $-\alpha < -1$. Then, by Karamata's Theorem, the corresponding distribution G will be regularly varying with index $-\alpha + 1 < 0$. Now consider any density function f such that $c_1 g(x) \le f(x) \le c_2 g(x)$, for some $0 < c_1 < c_2 < \infty$ and for all x. The corresponding distribution F is intermediate regularly varying because

$$F(x, x(1+\varepsilon)] \le c_2 G(x, x(1+\varepsilon)] \quad \text{and} \quad \overline{F}(x) \ge c_1\overline{G}(x).$$

On the other hand, F is not necessarily a regularly varying distribution. We now have the following characterisation result.

Theorem 2.46. *A distribution F on $\mathbb{R}$ is intermediate regularly varying if and only if, for any positive function h such that $h(x) = o(x)$ as $x \to \infty$,*

$$\overline{F}(x + h(x)) \sim \overline{F}(x), \tag{2.45}$$

i.e. if and only if F is $o(x)$-insensitive.

Proof. It is straightforward that if F is intermediate regularly varying then it is $o(x)$-insensitive. Hence it only remains to prove the reverse implication. Assume, on the contrary, that this implication fails. Thus let F be a distribution which is $o(x)$-insensitive but which fails to be intermediate regularly varying. The function

$$l(\varepsilon) := \liminf_{x\to\infty} \frac{\overline{F}(x(1+\varepsilon))}{\overline{F}(x)}$$

decreases in $\varepsilon > 0$, due to the monotonicity of $\overline{F}$. Therefore, the failure of (2.44) implies that there exists a positive δ such that $l(\varepsilon) \le 1 - 2\delta$ for any $\varepsilon > 0$. Hence, for any positive integer n, we can find x_n such that

$$\overline{F}(x_n(1 + 1/n)) \le (1-\delta)\overline{F}(x_n)$$

Without loss of generality we may assume the sequence $\{x_n\}$ to be increasing. Now put $h(x) = x/n$ for $x \in [x_n, x_{n+1})$. Then $h(x) = o(x)$ as $x \to \infty$. However,

$$\begin{aligned}\liminf_{x\to\infty} \frac{\overline{F}(x+h(x))}{\overline{F}(x)} &\le \liminf_{n\to\infty} \frac{\overline{F}(x_n + h(x_n))}{\overline{F}(x_n)} \\ &= \liminf_{n\to\infty} \frac{\overline{F}(x_n(1+1/n))}{\overline{F}(x_n)} \\ &\le 1-\delta,\end{aligned}$$

which contradicts the $o(x)$-insensitivity of F. □

We now give an attractive probabilistic characterisation of intermediate regularly varying distributions.

Theorem 2.47. *A distribution F on $\mathbb{R}$ is intermediate regularly varying if and only if, for any sequence of independent identically distributed random variables ξ_1, ξ_2, … with finite positive mean,*

$$\frac{\overline{F}(S_n)}{\overline{F}(n\mathbb{E}\xi_1)} \to 1 \quad \text{as } n \to \infty \tag{2.46}$$

with probability 1, where $S_n = \xi_1 + \ldots + \xi_n$.

Proof. We suppose first that F is intermediate regularly varying; let ξ_1, ξ_2, … be any sequence of independent identically distributed random variables with finite positive mean, and, for each n, let $S_n = \xi_1 + \ldots + \xi_n$; we show that then the relation (2.46) holds. Let $a = \mathbb{E}\xi_1$. Fix any $\varepsilon > 0$. It follows from the definition of intermediate regular variation that there is n_0 and a $\delta > 0$ such that

$$\sup_{n\ge n_0} \left| \frac{\overline{F}(n(a\pm\delta))}{\overline{F}(na)} - 1 \right| \le \varepsilon.$$

By the Strong Law of Large Numbers, with probability 1, there exists a random number N such that $|S_n - na| \le n\delta$ for all $n \ge N$. Then, for $n \ge \max\{N, n_0\}$,

$$\left|\frac{\overline{F}(S_n)}{\overline{F}(na)} - 1\right| \le \sup_{n \ge n_0}\left|\frac{\overline{F}(n(a \pm \delta))}{\overline{F}(na)} - 1\right| \le \varepsilon.$$

Since $\epsilon > 0$ is arbitrary, this implies the convergence (2.46).

We now prove the converse implication. Assume that the distribution F is not intermediate regularly varying. It is sufficient to construct a sequence of independent identically distributed random variables $\xi_1, \xi_2, \dots$ with mean 1, such that the relation (2.46) fails to hold (where again $S_n = \xi_1 + \dots + \xi_n$). By Theorem 2.46 F fails to be $o(x)$-insensitive, and so there exist an $\varepsilon > 0$, an increasing sequence n_k and an increasing function h with $h(x) = o(x)$ such that

$$\overline{F}(n_k + h(n_k)) \le (1 - \varepsilon)\overline{F}(n_k) \quad \text{for all } k. \tag{2.47}$$

Since $h(x)/x \to 0$, we can choose an increasing subsequence n_{k_m} such that

$$\sum_{m=1}^{\infty} \frac{h(n_{k_m})}{n_{k_m}} < \infty. \tag{2.48}$$

Since h is increasing it follows also that $\sum_{m=1}^{\infty} n_{k_m}^{-1} < \infty$, and so we can define a random variable ξ taking values on $\{1 \pm h(n_{k_m}), m = 1, 2, \dots\}$ with probabilities

$$\mathbb{P}\{\xi = 1 - h(n_{k_m})\} = \mathbb{P}\{\xi = 1 + h(n_{k_m})\} = c/n_{k_m}$$

(where c is the appropriate normalising constant). It further follows from (2.48) that the random variable ξ has a finite mean; moreover, this mean equals 1. Define the sequence of independent random variables $\xi_1, \xi_2, \dots$ to each have the same distribution as ξ. We shall show that

$$\liminf_{m \to \infty} \mathbb{P}\{S_{n_{k_m}} \ge n_{k_m} + h(n_{k_m})\} > 0. \tag{2.49}$$

From this and from (2.47), and since also $\overline{F}$ is non-increasing, it will follow that

$$\liminf_{m \to \infty} \mathbb{P}\{\overline{F}(S_{n_{k_m}}) \le (1 - \varepsilon)\overline{F}(n_k)\} > 0,$$

so that (2.46) cannot hold.

To show (2.49), fix m and consider the events

$$A_j = \bigcap_{i \le n_{k_m}, i \ne j} \{\xi_i \ne 1 \pm h(n_{k_m})\}, \qquad j = 1, \dots n_{k_m}.$$

Then the events $A_j \cap \{\xi_j = 1 + h(n_{k_m})\}$ are disjoint. Therefore,

$$\begin{aligned}
&\mathbb{P}\{S_{n_{k_m}} \ge n_{k_m} + h(n_{k_m})\} \\
&\ge \sum_{j=1}^{n_{k_m}} \mathbb{P}\{S_{n_{k_m}} \ge n_{k_m} + h(n_{k_m}) \,|\, A_j,\, \xi_j = 1 + h(n_{k_m})\}\mathbb{P}\{A_j,\, \xi_j = 1 + h(n_{k_m})\} \\
&= n_{k_m}\mathbb{P}\{S_{n_{k_m}} - n_{k_m} \ge h(n_{k_m}) \,|\, A_1,\, \xi_1 - 1 = h(n_{k_m})\}\mathbb{P}\{A_1\}\mathbb{P}\{\xi_1 = 1 + h(n_{k_m})\} \\
&= c\mathbb{P}\{S_{n_{k_m}} - n_{k_m} \ge h(n_{k_m}) \,|\, A_1,\, \xi_1 - 1 = h(n_{k_m})\}\mathbb{P}\{A_1\},
\end{aligned}$$

where the final equality follows from the definition of the distribution of ξ_1. Using again the independence of the random variables ξ_i, we have

$$\mathbb{P}\{S_{n_{k_m}} - n_{k_m} \geq h(n_{k_m}) \,|\, A_1, \xi_1 - 1 = h(n_{k_m})\} \\ = \mathbb{P}\{S_{n_{k_m}} - n_{k_m} - (\xi_1 - 1) \geq 0 \,|\, A_1\} \geq 1/2,$$

where the final inequality follows from the symmetry about 1 of the common distribution of the random variables ξ_i. In addition,

$$\begin{aligned}\mathbb{P}\{A_1\} &= \left(\mathbb{P}\{\xi_i \neq 1 \pm h(n_{k_m})\}\right)^{n_{k_m}-1} \\ &= (1 - 2c/n_{k_m})^{n_{k_m}-1} \to e^{-2c} \quad \text{as } m \to \infty.\end{aligned}$$

We thus finally obtain that

$$\liminf_{m\to\infty} \mathbb{P}\{S_{n_{k_m}} \geq n_{k_m} + h(n_{k_m})\} \geq ce^{-2c}/2,$$

so that (2.49) follows. □

Other heavy-tailed distributions We proceed now to heavy-tailed distributions with thinner tails. For the *lognormal distribution*, one can take $h(x) = o(x/\ln x)$ in order to have h-insensitivity. For the *Weibull distribution* with parameter $\alpha \in (0, 1)$, one can take $h(x) = o(x^{1-\alpha})$.

In many practical situations, the class of so-called $\sqrt{x}$-*insensitive distributions* —those which are h-insensitive for the function $h(x) = x^{1/2}$—is of special interest. Among these are intermediate regularly-varying distributions (in particular regularly-varying distributions), lognormal distributions and Weibull distributions with shape parameter $\alpha < 1/2$. The reason for interest in this quite broad class is explained by the following theorem, which should be compared with Theorem 2.47.

Theorem 2.48. *For any distribution F on $\mathbb{R}$, the following assertions are equivalent:*

(i) *F is $\sqrt{x}$-insensitive;*

(ii) *for any sequence of independent identically distributed random variables ξ_1, ξ_2, ... with positive mean and with finite positive variance,*

$$\frac{\overline{F}(S_n)}{\overline{F}(n\mathbb{E}\xi_1)} \to 1 \quad \text{as } n \to \infty \tag{2.50}$$

in probability, where $S_n = \xi_1 + \ldots + \xi_n$;

(iii) *for some sequence of independent identically distributed random variables ξ_1, ξ_2, ... with positive mean and with finite positive variance,* (2.50) *holds.*

Proof. To show (i)⇒(ii) suppose that the distribution F is $\sqrt{x}$-insensitive and that the independent identically distributed random variables ξ_1, ξ_2, ... have common mean $a > 0$ and finite variance. Fix $\varepsilon > 0$. By the Central Limit Theorem, there exist N and A such that $\mathbb{P}\{|S_n - na| \leq A\sqrt{n}\} \geq 1 - \varepsilon$ for all $n \geq N$. It follows from the definition of $\sqrt{x}$-insensitivity that there is n_0 such that

$$\left|\frac{\overline{F}(na \pm A\sqrt{n}))}{\overline{F}(na)} - 1\right| \leq \varepsilon \quad \text{for all } n \geq n_0.$$

Then, for $n \geq \max\{N, n_0\}$,

$$\mathbb{P}\left\{\left|\frac{\overline{F}(S_n)}{\overline{F}(na)} - 1\right| \leq \varepsilon\right\} \geq \mathbb{P}\{|S_n - na| \leq A\sqrt{n}\} \geq 1 - \varepsilon,$$

which completes the proof of (2.50).

It remains to prove implication (iii)⇒(i). Assume that the independent identically distributed random variables $\xi_1, \xi_2, \ldots$ have common mean $a > 0$ and finite variance $\sigma^2 > 0$, but that the distribution F fails to be $\sqrt{x}$-insensitive. Then there exists $\varepsilon > 0$ and an increasing sequence n_k such that, for all k,

$$\overline{F}(n_k a + \sqrt{n_k \sigma^2}) \leq (1 - \varepsilon)\overline{F}(n_k a).$$

Therefore,

$$\mathbb{P}\left\{\left|\frac{\overline{F}(S_{n_k})}{\overline{F}(n_k a)} - 1\right| \geq \varepsilon\right\} \geq \mathbb{P}\{S_{n_k} - n_k a \geq \sqrt{n_k \sigma^2}\} \to \int_1^\infty \frac{e^{-u^2/2}}{\sqrt{2\pi}} du > 0,$$

which contradicts (2.50). □

We finish this section by observing that the exponential distribution, while itself light-tailed, is, in an obvious sense, on the boundary of the class of such distributions. We may construct examples of long-tailed (and hence heavy-tailed) distributions on $\mathbb{R}^+$, say, whose tails are arbitrarily close to that of the exponential distribution. For example, the distribution with tail function

$$\overline{F}(x) = e^{-cx/\ln^\alpha x},\ \alpha > 0,\ c > 0,$$

is very close to the exponential distribution, but is still long-tailed; indeed one can take the function h of Lemma 2.19 to be any such that $h(x) = o(\ln^\alpha x)$ as $x \to \infty$. Further, if we replace the logarithmic function by the mth iterated logarithm, we obtain again a long-tailed distribution.

2.9 Comments

The lower bound (2.7) may be found in Chistyakov [12] and in Pakes [34].

The class of long-tailed distributions (but not the term itself) was introduced by Chistyakov in [12], in the context of branching processes.

Theorem 2.39 generalises a result of Cline [15] where the case $F_1, F_2, G_1, G_2 \in \mathcal{L}$ was considered.

Corollary 2.41 is well-known from Embrechts and Goldie [20].

Chapter 3

Subexponential distributions

As we stated in the Introduction, all those heavy-tailed distributions likely to be of use in practical applications are not only long-tailed but possess the additional regularity property of subexponentiality. Essentially this corresponds to good tail behaviour under the operation of convolution. In this chapter, following established tradition, we introduce first subexponential distributions on the positive half-line $\mathbb{R}^+$. It is not immediately obvious from the definition, but it nevertheless turns out, that subexponentiality is a tail property of a distribution. It is thus both natural, and important for many applications, to extend the concept to distributions on the entire real line $\mathbb{R}$. We also study the very useful subclass of subexponential distributions which is called $\mathcal{S}^*$, and which again contains all those heavy-tailed distributions likely to be encountered in practice.

Different sufficient and necessary conditions for subexponentiality may be found in Sections 3.5 and 3.6. We also discuss the questions of why not every long-tailed distribution is subexponential and why the subexponentiality of a distribution does not imply subexponentiality of the integrated tail distribution.

In Section 3.9 we consider closure properties for the class of subexponential distributions. We conclude with the fundamental uniform upper bound for the tail of the nth convolution of a subexponential distribution known as Kesten's estimate.

3.1 Subexponential distributions on the positive half-line

In the previous chapter we showed in (2.7) that, for any distribution F on $\mathbb{R}^+$ with unbounded support,

$$\liminf_{x\to\infty} \frac{\overline{F*F}(x)}{\overline{F}(x)} \geq 2.$$

It was then proved in Theorem 2.12 that, for any heavy-tailed distribution F on $\mathbb{R}^+$,

$$\liminf_{x\to\infty} \frac{\overline{F*F}(x)}{\overline{F}(x)} = 2.$$

In particular, if F is heavy-tailed on $\mathbb{R}^+$ and if

$$\frac{\overline{F*F}(x)}{\overline{F}(x)} \to c \quad \text{as } x\to\infty,$$

where $c \in (0, \infty]$, then necessarily $c = 2$. This observation leads naturally to the following definition.

Definition 3.1. Let F be a distribution on $\mathbb{R}^+$ with unbounded support. We say that F is *subexponential*, and write $F \in \mathcal{S}$, if

$$\overline{F * F}(x) \sim 2\overline{F}(x) \quad \text{as } x \to \infty. \tag{3.1}$$

Now let ξ_1 and ξ_2 be independent random variables on $\mathbb{R}^+$ with common distribution F. Then the above definition is equivalent to stating that F is subexponential if

$$\mathbb{P}\{\xi_1 + \xi_2 > x\} \sim 2\mathbb{P}\{\xi_1 > x\} \quad \text{as } x \to \infty.$$

This last relation may be rewritten as

$$\mathbb{P}\{\xi_1 > x | \xi_1 + \xi_2 > x\} \to 1/2 \quad \text{as } x \to \infty.$$

Further, since we always have the equivalence

$$\mathbb{P}\{\max(\xi_1, \xi_2) > x\} = 1 - (1 - \mathbb{P}\{\xi_1 > x\})^2 \sim 2\mathbb{P}\{\xi_1 > x\}$$

as $x \to \infty$, it follows that F is a subexponential distribution if and only if

$$\mathbb{P}\{\xi_1 + \xi_2 > x\} \sim \mathbb{P}\{\max(\xi_1, \xi_2) > x\} \quad \text{as } x \to \infty.$$

Finally, since ξ_1, ξ_2 are non-negative, the inequality $\max(\xi_1, \xi_2) > x$ implies also that $\xi_1 + \xi_2 > x$, and so the subexponentiality of their distribution is equivalent to the following relation:

$$\mathbb{P}\{\xi_1 + \xi_2 > x, \max(\xi_1, \xi_2) \le x\} = o(\mathbb{P}\{\xi_1 > x\}) \quad \text{as } x \to \infty. \tag{3.2}$$

That is, for large x, the only significant way in which $\xi_1 + \xi_2$ can exceed x is that either ξ_1 or ξ_2 should itself exceed x. This is the well-known "principle of a single big jump" for sums of subexponentially distributed random variables.

Lemma 2.43 with $G = F$ implies immediately the following result.

Lemma 3.2. *Any subexponential distribution on $\mathbb{R}^+$ is long-tailed. In particular, any subexponential distribution is heavy-tailed.*

The converse is not true; there exist some long-tailed distributions on $\mathbb{R}^+$ which are not subexponential; see Section 3.7 for more detail.

Since a long-tailed distribution F satisfies $\overline{F}(x)e^{\lambda x} \to \infty$ as $x \to \infty$, for all $\lambda > 0$ (see Lemma 2.17), it is this property that suggested the name subexponential. However, the name is used in the slightly more restrictive sense that we have defined.

In the class of distributions on the positive half-line, subexponentiality is a *tail property*, as are both heavy- and long-tailedness. To see this, observe that if a distribution F_1 on $\mathbb{R}^+$ is subexponential (and therefore long-tailed) and if a distribution F_2 on $\mathbb{R}^+$ is such that, for some x_0, we have $\overline{F}_1(x) = \overline{F}_2(x)$ for all $x \ge x_0$, then by Theorem 2.39 $\overline{F_1 * F_1}(x) \sim \overline{F_2 * F_2}(x)$ as $x \to \infty$, which implies the subexponentiality of F_2.

3.2 Subexponential distributions on the whole real line

In the previous section we defined subexponential distributions on the positive half-line $\mathbb{R}^+$. We showed there that subexponentiality was a tail property of a distribution. Thus, as remarked at the beginning of this chapter, it is both natural and desirable to extend the concept to distributions on the entire real line $\mathbb{R}$.

The problem is now that of extending the definition appropriately. It turns out that, for a distribution F on the entire real line $\mathbb{R}$, the condition (3.1) no longer defines a tail property of that distribution, nor even implies that the distribution is long-tailed. This is illustrated by the following example.

Example 3.3. For $A \geq 0$, consider the distribution F on the interval $[-A, \infty)$ with the tail function

$$\overline{F}(x) = (x + A + 1)^{-2} e^{-(x+A)}, \quad x \geq -A.$$

The convolution tail is given by

$$\begin{aligned}
\overline{F * F}(x) &= \int_{-\infty}^{\infty} \overline{F}(x-y) F(dy) \\
&= \int_{-\infty}^{x/2} \overline{F}(x-y) F(dy) - \int_{x/2}^{\infty} \overline{F}(x-y) d\overline{F}(y) \\
&= 2 \int_{-\infty}^{x/2} \overline{F}(x-y) F(dy) + (\overline{F}(x/2))^2,
\end{aligned}$$

after integration by parts. We thus have that, as $x \to \infty$,

$$\begin{aligned}
\overline{F * F}(x) &\sim 2e^{-x-A} \int_{-A}^{x/2} (x-y)^{-2} e^{y} F(dy) + o(\overline{F}(x)) \\
&\sim 2x^{-2} e^{-x-A} \int_{-A}^{x/2} e^{y} F(dy) + o(\overline{F}(x)) \\
&\sim 2\overline{F}(x) \int_{-A}^{\infty} e^{y} F(dy).
\end{aligned}$$

Take A such that $\int_{-A}^{\infty} e^y F(dy) = 1$. Then $\overline{F * F}(x) \sim 2\overline{F}(x)$, but F is not long-tailed and indeed F is light-tailed.

The above example shows that the satisfaction of the condition (3.1) *is not* a tail property for the class of distributions on the whole real line $\mathbb{R}$—for otherwise the condition would be satisfied by the distribution F^+ (given, as in the Introduction, by $F^+(x) = F(x)$ for $x \geq 0$ and $F^+(x) = 0$ for $x < 0$), and Lemma 3.2 would then guarantee that F^+ was long-tailed in contradiction to the result F is not long-tailed.

Thus the most usual way to define the subexponentiality of a distribution F on the whole real line R is to require that the distribution F^+ on $\mathbb{R}^+$ be subexponential. The condition (3.1) then continues to hold—it is simply no longer sufficient for subexponentiality. This approach has the advantage of making it immediately clear that subexponentiality remains a tail property, but the disadvantage of requiring a two-stage definition. We shall see below that an equivalent definition, which we shall make formally, is to require that the distribution F, in additional to satisfying (3.1), is also long-tailed. The asserted equivalence follows from the following lemma.

Lemma 3.4. *Let F be a distribution on $\mathbb{R}$ and let ξ be a random variable with distribution F. Then the following are equivalent:*

(i) *F is long-tailed and $\overline{F*F}(x) \sim 2\overline{F}(x)$ as $x \to \infty$;*

(ii) *the distribution F^+ of ξ^+ is subexponential;*

(iii) *the conditional distribution $G(B) := \mathbb{P}\{\xi \in B \,|\, \xi \geq 0\}$ is subexponential.*

Proof. Let ξ_1 and ξ_2 be two independent copies of ξ.

(i)⇒(ii). Suppose that F is long-tailed. Fix $A > 0$. On the event $\{\xi_k > -A\}$, we have $\xi_k^+ \leq \xi_k + A$. Thus, for $x \geq 0$,

$$\begin{aligned}\mathbb{P}\{\xi_1^+ + \xi_2^+ > x\} &\leq \mathbb{P}\{\xi_1 + \xi_2 > x - 2A, \xi_1 > -A, \xi_2 > -A\} \\ &\quad + \mathbb{P}\{\xi_2 > x, \xi_1 \leq -A\} + \mathbb{P}\{\xi_1 > x, \xi_2 \leq -A\} \\ &\leq \mathbb{P}\{\xi_1 + \xi_2 > x - 2A\} + 2\overline{F}(x)F(-A).\end{aligned}$$

Hence, since F is long-tailed,

$$\begin{aligned}\limsup_{x\to\infty} \frac{\mathbb{P}\{\xi_1^+ + \xi_2^+ > x\}}{\overline{F}(x)} &\leq \lim_{x\to\infty} \frac{\mathbb{P}\{\xi_1 + \xi_2 > x - 2A\}}{\overline{F}(x - 2A)} + 2F(-A) \\ &= 2 + 2F(-A).\end{aligned}$$

Since A can be chosen as large as we please,

$$\limsup_{x\to\infty} \frac{\mathbb{P}\{\xi_1^+ + \xi_2^+ > x\}}{\overline{F}(x)} \leq 2.$$

Together with (2.7) this implies that $\overline{F^+ * F^+}(x) \sim 2\overline{F^+}(x)$ as $x \to \infty$, i.e. that the distribution F^+ of ξ^+ is subexponential.

(ii)⇒(i). Suppose now that the distribution F^+ of ξ^+ is subexponential. That F^+ and hence F is long-tailed follows from Lemma 3.2. We further have $\xi_1 + \xi_2 \leq \xi_1^+ + \xi_2^+$, so that

$$\overline{F*F}(x) \leq \overline{F^+ * F^+}(x) \sim 2\overline{F}(x)$$

as $x \to \infty$, again by the subexponentiality of F^+. Together with the lower estimate for the '$\liminf$' provided by Corollary 2.30, we get the required tail asymptotics $\overline{F*F}(x) \sim 2\overline{F}(x)$ as $x \to \infty$.

(ii)⇔(iii). We show now the equivalence of the conditions (ii) and (iii). Define first $p = \mathbb{P}\{\xi < 0\}$ and observe that, for $x \geq 0$,

$$\begin{aligned}\mathbb{P}\{\xi_1^+ + \xi_2^+ > x\} &= 2\mathbb{P}\{\xi_1 < 0, \xi_2 > x\} + \mathbb{P}\{\xi_1 + \xi_2 > x, \xi_1 \geq 0, \xi_2 \geq 0\} \\ &= 2p\overline{F}(x) + (1-p)^2\mathbb{P}\{\xi_1 + \xi_2 > x \,|\, \xi_1 \geq 0, \xi_2 \geq 0\} \\ &= 2p\overline{F}(x) + (1-p)^2\overline{G*G}(x).\end{aligned}$$

Since also, for $x \ge 0$, we have $\overline{F}(x) = (1-p)\overline{G}(x)$, the subexponentiality of G is equivalent to the condition that, as $x \to \infty$,

$$\begin{aligned}\mathbb{P}\{\xi_1^+ + \xi_2^+ > x\} &\sim 2p\overline{F}(x) + 2(1-p)\overline{F}(x)\\ &= 2\mathbb{P}\{\xi^+ > x\},\end{aligned}$$

i.e. to the subexponentiality of F^+. □

The above lemma allows us to make the following definition of subexponentiality on the whole real line $\mathbb{R}$.

Definition 3.5. Let F be a distribution on $\mathbb{R}$ with right-unbounded support. We say that F is *subexponential on the whole line* , and write $F \in \mathcal{S}_{\mathbb{R}}$, if F is long-tailed and

$$\overline{F * F}(x) \sim 2\overline{F}(x) \quad \text{as } x \to \infty.$$

Equivalently, a random variable ξ has a subexponential distribution on the whole line if ξ^+ has a subexponential distribution.

Thus subexponentiality on the whole real line $\mathbb{R}$ generalises the concept of subexponentiality on the positive half-line $\mathbb{R}^+$ and any distribution which is subexponential on $\mathbb{R}^+$ or $\mathbb{R}$ is long-tailed, i.e. $\mathcal{S} \subseteq \mathcal{S}_{\mathbb{R}} \subseteq \mathcal{L}$.

We now have the following theorem which provides the foundation for our results on convolutions of subexponential distributions.

Theorem 3.6. *Let the distribution F on $\mathbb{R}$ be long-tailed ($F \in \mathcal{L}$) and let ξ_1, ξ_2 be two independent random variables with distribution F. Let the function h be such that $h(x) \to \infty$ as $\to \infty$ and F is h-insensitive (see Definition 2.32). Then F is subexponential on the whole line ($F \in \mathcal{S}_{\mathbb{R}}$) if and only if*

$$\mathbb{P}\{\xi_1 + \xi_2 > x, \xi_1 > h(x), \xi_2 > h(x)\} = o(\overline{F}(x)) \text{ as } x \to \infty. \tag{3.3}$$

Proof. We assume first that additionally $h(x) < x/2$ for all x. Then, for any x,

$$\begin{aligned}&\mathbb{P}\{\xi_1 + \xi_2 > x\}\\ &\quad= \mathbb{P}\{\xi_1 + \xi_2 > x,\ \xi_1 \le h(x)\} + \mathbb{P}\{\xi_1 + \xi_2 > x,\ \xi_2 \le h(x)\}\\ &\qquad + \mathbb{P}\{\xi_1 + \xi_2 > x,\ \xi_1 > h(x),\ \xi_2 > h(x)\}.\end{aligned} \tag{3.4}$$

Since F is long-tailed, it follows from (2.32), the given conditions on h and Lemma 2.33 that, for $i = 1, 2$,

$$\mathbb{P}\{\xi_1 + \xi_2 > x,\ \xi_i \le h(x)\} \sim \overline{F}(x) \qquad \text{as } x \to \infty. \tag{3.5}$$

Again, since F is long-tailed, the subexponentiality of F is equivalent to the requirement that $\mathbb{P}\{\xi_1 + \xi_2 > x\} \sim 2\overline{F}(x)$ as $x \to \infty$, and the equivalence of this to the condition (3.3) now follows from (3.4) and (3.5).

In the case where we do not have $h(x) < x/2$ for all x, small variations are required to the above proof. If F is subexponential, then we may consider instead the function $\hat{h}$ given by $\hat{h}(x) = \min(h(x), x/2)$. Since F is then also $\hat{h}$-insensitive, the relation (3.3) holds with h replaced by $\hat{h}$, and so also in its original form. Conversely, if (3.3) holds, then that F is subexponential follows as before, except only that we now have "≤" instead of equality in (3.4), which does not affect the argument. □

Theorem 3.6 implies that, as in the case of non-negative subexponential summands, the most probable way for large deviations of the sum $\xi_1 + \xi_2$ to occur is that one summand is small and the other is large; for (very) large x, the main contribution to the probability $\mathbb{P}\{\xi_1 + \xi_2 > x\}$ is made by the probabilities of the events $\{\xi_1 + \xi_2 > x,\, \xi_i \le h(x)\}$ for $i = 1, 2$.

We now give what is almost a restatement of Theorem 3.6 in terms of integrals, in a form which will be of use in various of our subsequent calculations.

Theorem 3.7. *Let the distribution F on $\mathbb{R}$ be long-tailed. Then the following are equivalent:*

(i) *F is subexponential on the whole line, i.e. $F \in \mathcal{S}_{\mathbb{R}}$;*

(ii) *for every function h with $h(x) < x/2$ for all x and such that $h(x) \to \infty$ as $x \to \infty$,*

$$\int_{h(x)}^{x-h(x)} \overline{F}(x-y)F(dy) = o(\overline{F}(x)) \quad \text{as } x \to \infty; \tag{3.6}$$

(iii) *there exists a function h with $h(x) < x/2$ for all x, such that $h(x) \to \infty$ as $x \to \infty$ and F is h-insensitive, and the relation (3.6) holds.*

Proof. As remarked above, the theorem is only a slight variation on Theorem 3.6. Let ξ_1 and ξ_2 again be independent random variables with common distribution F, and let h be any function such that $h(x) < x/2$, $h(x) \to \infty$ and F is h-insensitive (note that since $F \in \mathcal{L}$ there is always at least one such function h); the difference between the left side of (3.3) and the left side of (3.6) is

$$\mathbb{P}\{\xi_1 > x - h(x),\, \xi_2 > h(x)\} = \overline{F}(x-h(x))\overline{F}(h(x)) \sim \overline{F}(x)\overline{F}(h(x)) = o(\overline{F}(x))$$

as $x \to \infty$. The theorem thus follows immediately from Theorem 3.6, except only that it is necessary to observe that the reason why, in the statement (ii), we do not require any restriction to functions h such that F is h-insensitive follows from Proposition 2.20(ii). □

In the succeeding sections, we will make use of the following result.

Lemma 3.8. *Suppose that F is subexponential on the whole line and that the function h is such that $h(x) \to \infty$ as $x \to \infty$. Let the distributions G_1, G_2 be such that, for $i = 1$, 2, we have $\overline{G_i}(x) = O(\overline{F}(x))$ as $x \to \infty$. If η_1 and η_2 are independent random variables with distributions G_1 and G_2, then*

$$\mathbb{P}\{\eta_1 + \eta_2 > x, \eta_1 > h(x), \eta_2 > h(x)\} = o(\overline{F}(x)) \quad \text{as } x \to \infty.$$

Proof. Let ξ_1 and ξ_2 be two independent random variables with distribution F. Since $\overline{G}_i(x) = O(\overline{F}(x))$, it follows from Lemma 2.36 that, for some $c < \infty$,

$$\mathbb{P}\{\eta_1 + \eta_2 > x, \eta_1 > h(x), \eta_2 > h(x)\} \le c\mathbb{P}\{\xi_1 + \xi_2 > x, \xi_1 > h(x), \xi_2 > h(x)\}.$$

The subexponentiality of F and Theorem 3.6, together with the immediately preceding remark, imply that

$$\mathbb{P}\{\xi_1 + \xi_2 > x, \xi_1 > h(x), \xi_2 > h(x)\} = o(\overline{F}(x)).$$

Hence the result follows. □

3.3 Subexponentiality and weak tail-equivalence

We start with the definition of weak tail-equivalence and then use this property to establish a number of powerful results.

Definition 3.9. Two distributions F and G with right-unbounded supports are called *weakly tail-equivalent* if there exist $c_1 > 0$ and $c_2 < \infty$ such that, for any $x > 0$,

$$c_1 \le \frac{\overline{F}(x)}{\overline{G}(x)} \le c_2.$$

This is equivalent to the condition

$$0 < \liminf_{x\to\infty} \frac{\overline{F}(x)}{\overline{G}(x)} \le \limsup_{x\to\infty} \frac{\overline{F}(x)}{\overline{G}(x)} < \infty.$$

Lemma 3.10. *Let F and G be weakly tail-equivalent distributions on $\mathbb{R}$. Suppose that either (i) both F and G are long-tailed, or (ii) both F and G are concentrated on $\mathbb{R}^+$, and suppose further that*

$$\limsup_{x\to\infty} \frac{\overline{F * G}(x)}{\overline{F}(x) + \overline{G}(x)} \le 1. \tag{3.7}$$

Then both F and G are subexponential.

Proof. It follows from Lemma 2.43 that, in both the cases considered, both F and G are long-tailed.

Now let h be any function such that $h(x) < x/2$, $h(x) \to \infty$, and both F and G are h-insensitive (recall that the existence of such a function is guaranteed by the results of Section 2.4). Let ξ and η be independent random variables with distributions F and G respectively. It follows from the decomposition (2.31) (with $h(x)$ in place of h), Lemma 2.33, and the condition (3.7) that

$$\mathbb{P}\{\xi + \eta > x, \xi > h(x), \eta > h(x)\} = o(\overline{F}(x) + \overline{G}(x)) \qquad \text{as } x \to \infty. \tag{3.8}$$

Let ξ' be an additional random variable, independent of ξ, with distribution F. Then, from (3.8), the weak tail-equivalence of F and G, and Lemma 2.36,

$$\begin{aligned}\mathbb{P}\{\xi + \xi' > x, \xi > h(x), \xi' > h(x)\} &= o(\overline{F}(x) + \overline{G}(x)) \\ &= o(\overline{F}(x)) \qquad \text{as } x \to \infty,\end{aligned}$$

where the second line in the above display again follows from the weak tail-equivalence of F and G. Hence, by Theorem 3.6, F is subexponential. $\square$

Now we prove that the class of subexponential distributions is closed under the weak tail-equivalence relation.

Theorem 3.11. *Suppose that F is subexponential on the whole line, that G is long-tailed, and that F and G are weakly tail-equivalent. Then $G \in \mathcal{S}_{\mathbb{R}}$.*

Proof. Choose a function h such that $h(x) \to \infty$ and G is h-insensitive. Let η_1 and η_2 be independent random variables with distribution G. Then, from Lemma 3.8 and the given weak tail-equivalence,

$$\mathbb{P}\{\eta_1 + \eta_2 > x, \eta_1 > h(x), \eta_2 > h(x)\} = o(\overline{F}(x)) = o(\overline{G}(x)).$$

Hence it follows from Theorem 3.6 that $G \in \mathcal{S}_{\mathbb{R}}$. □

Definition 3.12. Two distributions F and G with right-unbounded supports are said to be *proportionally tail-equivalent* if there exists a constant $c > 0$ such that $\overline{F}(x) \sim c\overline{G}(x)$ as $x \to \infty$.

Theorem 3.11 has the following corollary.

Corollary 3.13. *Let the distributions F and G be proportionally tail-equivalent. If $F \in \mathcal{S}_{\mathbb{R}}$ then $G \in \mathcal{S}_{\mathbb{R}}$.*

We now turn to convolutions of many distributions.

Theorem 3.14. *Let (a reference distribution) $F \in \mathcal{S}_{\mathbb{R}}$. Suppose that distributions $G_1, \ldots, G_n$ are such that, for each i, the function $\overline{F} + \overline{G_i}$ is long-tailed and $\overline{G_i}(x) = O(\overline{F}(x))$ as $x \to \infty$. Then*

$$\overline{G_1 * \cdots * G_n}(x) = \overline{G_1}(x) + \ldots + \overline{G_n}(x) + o(\overline{F}(x)) \quad \text{as } x \to \infty.$$

Proof. Note first that it follows from the conditions of the theorem that, for each i and for any constant a,

$$\begin{aligned}\overline{F}(x+a) + \overline{G_i}(x+a) &= \overline{F}(x) + \overline{G_i}(x) + o(\overline{F}(x) + \overline{G_i}(x)) \\ &= \overline{F}(x) + \overline{G_i}(x) + o(\overline{F}(x)).\end{aligned}$$

Hence from the representation $F + \sum_{i=1}^k G_i = \sum_{i=1}^k (F + G_i) - (k-1)F$ and since F is also long-tailed, for each k and for any constant a,

$$\overline{F}(x+a) + \sum_{i=1}^k \overline{G_i}(x+a) = \overline{F}(x) + \sum_{i=1}^k \overline{G_i}(x) + o(\overline{F}(x)),$$

and so the measure $F + \sum_{i=1}^k G_i$ (i.e. the function $\overline{F} + \sum_{i=1}^k \overline{G_i}$) is also long-tailed. Note also that for each k we have $\sum_{i=1}^k \overline{G_i}(x) = O(\overline{F}(x))$. It now follows that it is sufficient to prove the theorem for case $n = 2$, the general result then following by induction.

By Lemma 2.19 and Proposition 2.20 there exists a function h such that $h(x) \to \infty$, $h(x) \le x/2$, and F, $F + G_1$ and $F + G_2$ are all h-insensitive. It then follows from Lemma 2.33 that, as $x \to \infty$,

$$\begin{aligned}\int_{-\infty}^{h(x)} \overline{G_1}(x-y)G_2(dy) &= \int_{-\infty}^{h(x)} (\overline{G_1} + \overline{F})(x-y)G_2(dy) - \int_{-\infty}^{h(x)} \overline{F}(x-y)G_2(dy) \\ &= \overline{G_1 + F}(x) - \overline{F}(x) + o(\overline{G_1}(x) + \overline{F}(x)) \\ &= \overline{G_1}(x) + o(\overline{F}(x)),\end{aligned} \tag{3.9}$$

and similarly

$$\int_{-\infty}^{h(x)} \overline{G_2}(x-y)G_1(dy) = \overline{G_2}(x) + o(\overline{F}(x)). \tag{3.10}$$

Further, from Lemma 3.8,

$$\int_{h(x)}^{\infty} \overline{G_1}(\max(h(x), x-y)G_2(dy) = o(\overline{F}(x)). \tag{3.11}$$

The required result now follows from the decomposition (2.31) (where ξ and η are independent random variables with distributions G_1 and G_2) and from (3.9)–(3.11). ☐

Theorem 3.15. *Suppose again that the conditions of Theorem 3.14 hold, and that additionally G_1 satisfies the stronger condition (than (i) of Theorem 3.14) that $G_1 \in \mathcal{L}$ and that G_1 is weakly tail equivalent to F. Then $G_1 * \cdots * G_n \in \mathcal{S}_{\mathbb{R}}$, and additionally $G_1 * \cdots * G_n$ is weakly tail equivalent to F.*

Proof. It follows from Theorem 3.11 that $G_1 \in \mathcal{S}_{\mathbb{R}}$. Further the weak tail equivalence of F and G_1 implies that, for each k, $\overline{G_k}(x) = O(\overline{G_1}(x))$. Hence by Theorem 3.14 with $F = G_1$, the distribution $G_1 * G_2 * \cdots * G_n$ is long-tailed and weakly tail equivalent to G_1 and so also to F. In particular, again by Theorem 3.11, $G_1 * \cdots * G_n \in \mathcal{S}_{\mathbb{R}}$. ☐

We have the following corollaries of Theorems 3.14 and 3.15.

Corollary 3.16. *Suppose that distributions F and G are such that $F \in \mathcal{S}_{\mathbb{R}}$, that $\overline{F} + \overline{G}$ is long-tailed and that $\overline{G}(x) = O(\overline{F}(x))$ as $x \to \infty$. Then $F * G \in \mathcal{S}_{\mathbb{R}}$ and*

$$\overline{F * G}(x) = \overline{F}(x) + \overline{G}(x) + o(\overline{F}(x)) \quad \textit{as } x \to \infty.$$

Proof. This result follows from Theorems 3.14 and 3.15 in the case $n = 2$ with G_1 replaced by F and G_2 by G. ☐

Corollary 3.17. *Assume that F, $G \in \mathcal{S}_{\mathbb{R}}$. If F and G are weakly tail-equivalent, then $F * G \in \mathcal{S}_{\mathbb{R}}$.*

Corollary 3.18. *Assume that $F \in \mathcal{S}_{\mathbb{R}}$. If $\overline{G}(x) = o(\overline{F}(x))$ as $x \to \infty$, then $F * G \in \mathcal{S}_{\mathbb{R}}$ and $\overline{F * G}(x) \sim \overline{F}(x)$.*

Corollary 3.19. *Suppose that $F \in \mathcal{S}_{\mathbb{R}}$. Let $G_1, \ldots, G_n$ be distributions such that $\overline{G}_i(x)/\overline{F}(x) \to c_i$ as $x \to \infty$, for some constants $c_i \geq 0$, $i = 1, \ldots, n$. Then*

$$\frac{\overline{G_1 * \ldots * G_n}(x)}{\overline{F}(x)} \to c_1 + \ldots + c_n \quad \textit{as } x \to \infty.$$

If $c_1 + \ldots + c_n > 0$, then $G_1 * \ldots * G_n \in \mathcal{S}_{\mathbb{R}}$.

Proof. The first statement of the corollary is immediate from Theorem 3.14. If $c_1 + \ldots + c_n > 0$, we may assume without loss of generality that $c_1 > 0$, so that the second statement follows from Theorem 3.15. ☐

The following result is a special case of Corollary 3.19.

Corollary 3.20. *Assume that $F \in \mathcal{S}_{\mathbb{R}}$. Then for any $n \geq 2$, $\overline{F^{*n}}(x)/\overline{F}(x) \to n$ as $x \to \infty$. In particular, $F^{*n} \in \mathcal{S}_{\mathbb{R}}$.*

The following converse result follows.

Theorem 3.21. *Let a distribution F on $\mathbb{R}^+$ with unbounded support be such that $\overline{F^{*n}}(x) \sim n\overline{F}(x)$ for some $n \geq 2$. Then F is subexponential.*

Proof. Take $G := F^{*(n-1)}$. For any x we have the inequality $\overline{G}(x) \geq \overline{F}(x)$. On the other hand, $\overline{G}(x) \leq \overline{F^{*n}}(x) \sim n\overline{F}(x)$. Hence the distributions F and G are weakly tail-equivalent. Thus by Theorem 2.11, as $x \to \infty$,

$$\begin{aligned}\overline{F * G}(x) &\geq (1 + o(1))(\overline{F}(x) + \overline{G}(x)) \\ &= \overline{F}(x) + \overline{G}(x) + o(\overline{F}(x)).\end{aligned}$$

Recalling that $\overline{F * G}(x) = \overline{F^{*n}}(x) \sim n\overline{F}(x)$, we deduce the following upper estimate

$$\overline{F^{*(n-1)}}(x) = \overline{G}(x) \leq (n - 1 + o(1))\overline{F}(x).$$

Together with lower estimate (2.6) this implies that $\overline{F^{*(n-1)}}(x) \sim (n-1)\overline{F}(x)$ as $x \to \infty$. By induction we deduce then that $\overline{F^{*2}}(x) \sim 2\overline{F}(x)$, which completes the proof. □

3.4 The class $\mathcal{S}^*$

We have already observed that a heavy-tailed distribution F on $\mathbb{R}^+$ is subexponential if and only if it is long-tailed and its tail is sufficiently regular that $\lim_{x\to\infty} \overline{F * F}(x)/\overline{F}(x)$ exists (and that this limit is then equal to 2). Thus subexponentiality, with all its important properties for the tails of convolutions, is effectively guaranteed for all those heavy-tailed distributions likely to be encountered in practice.

However, some applications, for example, those concerned with the behaviour of the maxima of random walks with heavy-tailed increments, require a very slightly stronger regularity condition with respect to their tails—that of membership of the class $\mathcal{S}^*$ which we introduce below. We shall see that membership of $\mathcal{S}^*$ is again a tail property of a distribution and that $\mathcal{S}^*$ is a subclass of the class $\mathcal{S}_{\mathbb{R}}$ of distributions which are subexponential on $\mathbb{R}$.

For any distribution F on $\mathbb{R}$ with right-unbounded support, we have the inequality

$$\begin{aligned}\int_0^x \overline{F}(x-y)\overline{F}(y)dy &= 2\int_0^{x/2} \overline{F}(x-y)\overline{F}(y)dy \\ &\geq 2\overline{F}(x)\int_0^{x/2} \overline{F}(y)dy.\end{aligned}$$

Therefore, always

$$\liminf_{x\to\infty} \frac{1}{\overline{F}(x)} \int_0^x \overline{F}(x-y)\overline{F}(y)dy \geq 2m,$$

where $m = \mathbb{E}\xi^+$ and ξ has distribution F. If F is heavy-tailed, then (see [26], Lemma 4])

$$\liminf_{x\to\infty} \frac{1}{\overline{F}(x)} \int_0^x \overline{F}(x-y)\overline{F}(y)dy = 2m. \tag{3.12}$$

These observations provide a motivation for the following definition.

Definition 3.22. Let F be a distribution on $\mathbb{R}$ with right-unbounded support and finite mean on the positive half line. We say that F belongs to the class $\mathcal{S}^*$ if

$$\int_0^x \overline{F}(x-y)\overline{F}(y)dy \sim 2m\overline{F}(x) \quad \text{as } x \to \infty,$$

where $m = \mathbb{E}\xi^+$ and ξ has distribution F.

It follows from the observation (3.12) that a distribution F on $\mathbb{R}$ belongs to the class $\mathcal{S}^*$ if and only if it is heavy-tailed and sufficiently regular that $\lim_{x\to\infty} \frac{1}{\overline{F}(x)} \int_0^x \overline{F}(x-y)\overline{F}(y)dy$ exists. Thus it is again the case that most heavy-tailed distributions likely to be of use in practical applications belong to the class $\mathcal{S}^*$. This includes all those named distributions introduced in Section 2.1, i.e. the Pareto, Burr, Cauchy, lognormal, and Weibull (with shape parameter $\alpha < 1$) distributions.

We shall see in Section 4.2 that the condition $F \in \mathcal{S}^*$ is equivalent to the requirement that the density f on $\mathbb{R}^+$ given by $f(x) := \overline{F}(x)/m$ be subexponential in the sense defined there.

We show first that the class $\mathcal{S}^*$ is a subclass of the class $\mathcal{L}$ of long-tailed distributions on $\mathbb{R}$.

Theorem 3.23. *Let the distribution F on $\mathbb{R}$ belong to $\mathcal{S}^*$. Then F is long-tailed.*

Proof. Since

$$\int_0^x \overline{F}(x-y)\overline{F}(y)dy \geq 2\overline{F}(x)\int_0^1 \overline{F}(y)dy + 2\overline{F}(x-1)\int_1^{x/2} \overline{F}(y)dy,$$

the inclusion $F \in \mathcal{S}^*$ implies

$$\begin{aligned}(\overline{F}(x-1) - \overline{F}(x))\int_1^{x/2} \overline{F}(y)dy &\leq \frac{1}{2}\int_0^x \overline{F}(x-y)\overline{F}(y)dy - \overline{F}(x)\int_0^{x/2} \overline{F}(y)dy \\ &= o(\overline{F}(x)) \qquad \text{as } x \to \infty.\end{aligned}$$

It thus follows from Lemma 2.22 that F is long-tailed. □

We now have the following analogue to the conditions for subexponentiality given by Theorem 3.7.

Theorem 3.24. *Let F be a distribution on $\mathbb{R}$. Then the following are equivalent:*

(i) *$F \in \mathcal{S}^*$;*

(ii) *F is long-tailed, and for every function h with $h(x) < x/2$ for all x and such that $h(x) \to \infty$ as $x \to \infty$,*

$$\int_{h(x)}^{x-h(x)} \overline{F}(x-y)\overline{F}(y)dy = o(\overline{F}(x)) \ \text{ as } x \to \infty; \tag{3.13}$$

(iii) *there exists a function h with $h(x) < x/2$ for all x, such that $h(x) \to \infty$ as $x \to \infty$ and F is h-insensitive, and the relation (3.13) holds.*

Note that it follows in particular from Theorem 3.24 that membership of the class $\mathcal{S}^*$ is a tail property.

Proof of Theorem 3.24. Note first that each of the conditions (i)–(iii) implies that F is long-tailed. This follows in the case of (i) from Theorem 3.23, and in the case of (iii) from the existence of an increasing function with respect to which F is h-insensitive. Hence we assume without loss of generality that F is long-tailed ($F \in \mathcal{L}$).

Let h be any function with $h(x) < x/2$, such that $h(x) \to \infty$ as $x \to \infty$ and F is h-insensitive. (Note as usual that since F is assumed long-tailed there exists at least one such function h.) Then, for any $x \geq 0$,

$$\int_0^x \overline{F}(x-y)\overline{F}(y)\,dy = 2\int_0^{h(x)} \overline{F}(x-y)\overline{F}(y)\,dy + \int_{h(x)}^{x-h(x)} \overline{F}(x-y)\overline{F}(y)\,dy.$$

The h-insensitivity of F implies that

$$\int_0^{h(x)} \overline{F}(x-y)\overline{F}(y)\,dy \sim m\overline{F}(x) \qquad \text{as } x \to \infty,$$

where again $m = \mathbb{E}\xi^+$ and ξ has distribution F. It thus follows that the condition $F \in \mathcal{S}^*$ is equivalent to (3.13). The theorem now follows on noting that, as in the proof of Theorem 3.7, the reason why, in the statement (ii), we do not require any restriction to functions h such that F is h-insensitive follows from Proposition 2.20(ii). □

We now have the following theorem and its important corollary.

Theorem 3.25. *Suppose that $F \in \mathcal{S}^*$, that G is long-tailed, and that F and G are weakly tail-equivalent. Then $G \in \mathcal{S}^*$.*

Proof. Let h be a function such that $h(x) < x/2$, $h(x) \to \infty$ and G is h-insensitive. Then, from Theorem 3.24 and the given weak tail-equivalence,

$$\begin{aligned}\int_{h(x)}^{x-h(x)} \overline{G}(x-y)\overline{G}(y)\,dy &= O\left(\int_{h(x)}^{x-h(x)} \overline{F}(x-y)\overline{F}(y)\,dy\right)\\ &= o(\overline{F}(x))\\ &= o(\overline{G}(x)) \text{ as } x \to \infty.\end{aligned}$$

Again from Theorem 3.24, it now follows that $G \in \mathcal{S}^*$. □

Corollary 3.26. *Let distributions F and G be proportionally tail-equivalent. If $F \in \mathcal{S}^*$ then $G \in \mathcal{S}^*$.*

The following theorem asserts in particular that $\mathcal{S}^*$ is a subclass of $\mathcal{S}_{\mathbb{R}}$.

Theorem 3.27. *If $F \in \mathcal{S}^*$, then $F \in \mathcal{S}_{\mathbb{R}}$ and $F_I \in \mathcal{S}$.*

We do not provide a proof for this result now. Instead of that we recall the notion of an integrated weighted tail distribution and state sufficient conditions for its tail to be subexponential. Then Theorem 3.27 is a particular case of Theorem 3.28.

Let F be a distribution on $\mathbb{R}$ and let μ a non-negative measure on $\mathbb{R}^+$ such that

$$\int_0^\infty \overline{F}(t)\,\mu(dt) \quad \text{is finite.} \tag{3.14}$$

Then we can define the distribution F_μ on $\mathbb{R}^+$ by its tail:

$$\overline{F}_\mu(x) := \min\Big(1, \int_0^\infty \overline{F}(x+t)\,\mu(dt)\Big), \quad x \geq 0. \tag{3.15}$$

We may now ask the following question: what type of conditions on F imply the subexponentiality of F_μ?

For any $b > 0$, define the class $\mathcal{M}_b$ of all non-negative measures μ on $\mathbb{R}^+$ such that $\mu(x, x+1] \leq b$ for all x.

Theorem 3.28. *Let $F \in \mathcal{S}^*$ and $\mu \in \mathcal{M}_b$, $b \in (0,\infty)$. Then $F_\mu \in \mathcal{S}$. Moreover, $\overline{F_\mu * F_\mu}(x) \sim 2\overline{F}_\mu(x)$ as $x \to \infty$ uniformly in $\mu \in \mathcal{M}_b$.*

Here are two examples of such measures μ: (i) if $\mu(B) = \mathbb{I}\{0 \in B\}$, then F_μ is F restricted to $\mathbb{R}^+$; (ii) if $\mu(dt) = dt$ is Lebesgue measure on $\mathbb{R}^+$, then $F_\mu = F_I$. These examples give a proof of Theorem 3.27.

Proof of Theorem 3.28. First, recall that Theorem 2.28 states that if F is long-tailed, then F_μ is long-tailed uniformly in $\mu \in \mathcal{M}_b$. Thus, it is sufficient to show that, for any $h(x) \to \infty$,

$$\lim_{x\to\infty} \sup_{\mu\in\mathcal{M}_b} \frac{1}{\overline{F}_\mu(x)} \int_{h(x)}^{x-h(x)} \overline{F}_\mu(x-y) F_\mu(dy) = 0, \tag{3.16}$$

see Theorem 3.7. For any $\mu \in \mathcal{M}_b$,

$$F_\mu(y, y+1] \leq \int_y^{y+1} \overline{F}(t)\mu(dt) \leq \overline{F}(y)\mu(y, y+1] \leq b\overline{F}(y).$$

Therefore, (3.16) holds if and only if

$$\lim_{x\to\infty} \sup_{\mu\in\mathcal{M}_b} \frac{1}{\overline{F}_\mu(x)} \int_{h(x)}^{x-h(x)} \overline{F}_\mu(x-y)\overline{F}(y)dy = 0. \tag{3.17}$$

Since $F \in \mathcal{S}^*$, as $x \to \infty$,

$$\int_{h(x)}^{x-h(x)} \overline{F}(x-u)\overline{F}(u)du = o(\overline{F}(x)),$$

by Theorem 3.24. Then,

$$\begin{aligned}
\int_{h(x)}^{x-h(x)} \overline{F}_\mu(x-y)\overline{F}(y)dy &= \int_{h(x)}^{x-h(x)} \Big(\int_0^\infty \overline{F}(x+t-y)\,\mu(dt)\Big)\overline{F}(y)dy \\
&\leq \int_0^\infty \Big(\int_{h(x)}^{x+t-h(x)} \overline{F}(x+t-y)\overline{F}(y)dy\Big)\mu(dt) \\
&= \int_0^\infty o(\overline{F}(x+t))\mu(dt) = o(\overline{F}_\mu(x))
\end{aligned}$$

and we get (3.17). □

3.5 Sufficient conditions for subexponentiality

We formulate and prove here two results. The first may be applied to very heavy distributions such as Pareto distributions, while the second one may be applied to lighter distributions of the Weibull-type.

Theorem 3.29. *Let F be a long-tailed distribution on $\mathbb{R}$ ($F \in \mathcal{L}$) and suppose that there exists $c > 0$ such that $\overline{F}(2x) \geq c\overline{F}(x)$ for all x (that is, F belongs to the class $\mathcal{D}$ of dominated-varying distributions introduced in Section 2.1). Then*

(i) *F is subexponential on the whole line;*

(ii) *$F \in \mathcal{S}^*$, provided F has a finite mean on the positive half line;*

(iii) *$F_\mu \in \mathcal{S}$, for all μ satisfying (3.14).*

Note in particular that the statement (i) of Theorem 3.29 asserts that $\mathcal{D} \cap \mathcal{L} \subseteq \mathcal{S}_{\mathbb{R}}$.

Proof. It follows from the comment after Theorem 3.28 that (iii) implies (i) and (ii). So we prove (iii). The inequality $\overline{F}(2x) \geq c\overline{F}(x)$ yields, for those values of x such that the integrals below are less than 1,

$$\begin{aligned}\overline{F}_\mu(2x) &= \int_0^\infty \overline{F}(2x+y)\mu(dy)\\ &\geq c\int_0^\infty \overline{F}(x+y/2)\mu(dy)\\ &\geq c\int_0^\infty \overline{F}(x+y)\mu(dy) = c\overline{F}_\mu(x). \qquad (3.18)\end{aligned}$$

Now let h be any function such that $h(x) < x/2$ and $h(x) \to \infty$. We have the following estimate:

$$\begin{aligned}\int_{h(x)}^{x-h(x)} \overline{F}_\mu(x-y)F_\mu(dy) &= \int_{h(x)}^{x/2} \overline{F}_\mu(x-y)F_\mu(dy) + \int_{x/2}^{x-h(x)} \overline{F}_\mu(x-y)F_\mu(dy)\\ &\leq \overline{F}_\mu(x/2)\overline{F}_\mu(h(x)) + \overline{F}_\mu(h(x))\overline{F}_\mu(x/2).\end{aligned}$$

Therefore, by (3.18),

$$\int_{h(x)}^{x-h(x)} \overline{F}_\mu(x-y)F_\mu(dy) \leq 2\overline{F}_\mu(x)\overline{F}_\mu(h(x))/c = o(\overline{F}_\mu(x)) \text{ as } x \to \infty.$$

Applying now Theorem 3.7(ii), we conclude that $F_\mu \in \mathcal{S}$. □

The Pareto distribution, and more generally any regularly varying or indeed intermediate regularly varying distribution, satisfies conditions of Theorem 3.29 (i.e. belongs to $\mathcal{D}$) and is, therefore, subexponential. All of the above distributions whose means are finite also belong to the to class $\mathcal{S}^*$.

However, the lognormal distribution and the Weibull distribution do not satisfy the conditions of Theorem 3.29 and we need a different technique for proving their subexponentiality.

Recall that we denote by R the hazard function given by $R(x) := -\ln \overline{F}(x)$ and by r the hazard rate function given by $r(x) = R'(x)$, provided the hazard function is differentiable.

Theorem 3.30. *Let F be a long-tailed distribution on $\mathbb{R}$ ($F \in \mathcal{L}$). Assume that there exist $\gamma < 1$ and $A < \infty$ such that the hazard function $R(x)$ satisfies the following inequality:*

$$R(x) - R(x-y) \leq \gamma R(y) + A, \tag{3.19}$$

for all $x > 0$ and $y \in [0, x/2]$. If the function $e^{-(1-\gamma)R(x)}$ is integrable over $\mathbb{R}^+$, then $F \in \mathcal{S}^*$. In particular, F is subexponential on the whole line ($F \in \mathcal{S}_{\mathbb{R}}$).

Proof. For any $h < x/2$,

$$\begin{aligned}\int_h^{x-h} \overline{F}(x-y)\overline{F}(y)dy &= 2\int_h^{x/2} \overline{F}(x-y)\overline{F}(y)dy \\ &= 2\overline{F}(x)\int_h^{x/2} e^{R(x)-R(x-y)-R(y)}dy.\end{aligned}$$

It follows from (3.19) that

$$\int_h^{x/2} e^{R(x)-R(x-y)-R(y)}dy \leq e^A \int_h^{\infty} e^{-(1-\gamma)R(y)}dy \to 0 \quad \text{as } h \to \infty,$$

since the function $e^{-(1-\gamma)R(x)}$ is integrable. Hence, if h is now any function such that $h(x) \to \infty$ then

$$\int_{h(x)}^{x-h(x)} \overline{F}(x-y)\overline{F}(y)dy = o(\overline{F}(x)).$$

Hence, by Theorem 3.24, we have $F \in \mathcal{S}^*$. □

We note briefly that, for $0 < \gamma < 1$, the function $e^{-(1-\gamma)R(x)}$ is integrable if F has a finite moment of order $\frac{1}{1-\gamma} + \varepsilon$ on the positive half line $\mathbb{R}^+$ for some $\varepsilon > 0$. To see this note that the tail of F may then be bounded from above by $cx^{-1/(1-\gamma)-\varepsilon}$ (by the Chebyshev inequality):

$$e^{-(1-\gamma)R(x)} = (\overline{F}(x))^{1-\gamma} \leq c^{1-\gamma}x^{-1-(1-\gamma)\varepsilon}.$$

The heavy-tailed Weibull distribution, with tail function $\overline{F}$ given by $\overline{F}(x) = e^{-x^\alpha}$ for some $\alpha \in (0,1)$, satisfies the conditions of Theorem 3.30. Indeed, since the function $R(x) = x^\alpha$ is concave for $\alpha \in (0,1)$, we have for $y \leq x/2$ (so that $x - y \geq y$ and $R'(x-y) \leq R'(y)$)

$$R(x) - R(x-y) \leq yR'(x-y) \leq yR'(y) = \alpha R(y).$$

Similarly, it may be checked that the lognormal distribution satisfies conditions of Theorem 3.30 and, therefore, belongs to the class $\mathcal{S}^*$.

3.6 Conditions for subexponentiality in terms of truncated exponential moments

Note that some heavy-tailed distributions, for example those with tail functions of the form $e^{-x/\log x}$ do not satisfy the conditions of Theorem 3.30 (in this case $\gamma = 1$) and we need a more advanced technique for proving the subexponentiality of such distributions. The next two theorems, due to Pitman [36], relate the classes $\mathcal{S}$ and $\mathcal{S}^*$ to the asymptotic behaviour of truncated exponential moments with special indices.

Theorem 3.31. *Let F be a distribution on $\mathbb{R}^+$. Suppose that the hazard rate function r exists, is eventually non-increasing and that $r(x) \to 0$ as $x \to \infty$. Then F is subexponential ($F \in \mathcal{S}$) if and only if*

$$\int_0^x e^{yr(x)} F(dy) \to 1 \quad \text{as } x \to \infty. \tag{3.20}$$

Further, a sufficient condition for subexponentiality is that the function of y given by $e^{yr(y)-R(y)}r(y)$ is integrable over $\mathbb{R}^+$. Here $R(x) = \int_0^x r(y)dy$ defines the corresponding hazard function.

Note that the integral in (3.20) is equal to $\mathbb{E}\{e^{\xi r(x)}; \xi \le x\}$, where ξ is a random variable with distribution F.

Proof. The proof consists of two steps. First, we show that it may be assumed without loss of generality that the function $r(x)$ is non-increasing for all x (and that the condition (3.20) is a *tail property*), and then we prove the results under this assumption.

Suppose first that $r(x)$ may increase in a neighbourhood of 0 but is non-increasing for all $x \ge x_*$. Define the non-increasing hazard rate

$$r_*(x) = \begin{cases} r(x_*) & \text{for } x \le x_*, \\ r(x) & \text{for } x > x_*, \end{cases}$$

and put $R_*(x) = \int_0^x r_*(y)dy$. Define also the distribution F_* by $\overline{F_*}(x) = e^{-R_*(x)}$. Then, for all $x \ge x_*$,

$$R_*(x) - R(x) = \int_0^{x_*} (r(x_*) - r(y))dy =: c_*,$$

and hence, by Theorem 3.11, either both F_* and F are subexponential or both are not. We now prove that the functions r and r_* either both satisfy, or else both fail to satisfy, the condition (3.20). Note first that $\overline{F}_*(x_*) = e^{-c_*}\overline{F}(x_*)$. Note that $r(x) \to 0$ as $x \to \infty$ implies that $e^{yr(x)} \to 1$ as $x \to \infty$, uniformly in $0 \le y \le x_*$. Hence

$$\begin{aligned} \int_0^x e^{yr(x)} F(dy) &= \left(\int_0^{x_*} + \int_{x_*}^x\right) e^{yr(x)} F(dy) \\ &= F[0, x_*] + o(1) + \int_{x_*}^x e^{yr(x)} F(dy). \end{aligned}$$

It now follows that F satisfies the condition (3.20) if and only if

$$\int_{x_*}^x e^{yr(x)} F(dy) = \int_{x_*}^x F(dy) + o(1) = \overline{F}(x_*) + o(1).$$

Note also that $F_*(dx) = e^{-c_*} F(dx)$ for $x > x_*$. Therefore, as $x \to \infty$,

$$\begin{aligned} \int_0^x e^{yr_*(x)} F_*(dy) &= \int_0^{x_*} e^{yr_*(x)} F_*(dy) + \int_{x_*}^x e^{yr(x)} F_*(dy) \\ &= F_*[0, x_*] + o(1) + e^{-c_*} \int_{x_*}^x e^{yr(x)} F(dy) \\ &= F_*[0, x_*] + o(1) + e^{-c_*}(\overline{F}(x_*) + o(1)) \\ &= F_*[0, x_*] + \overline{F}_*(x_*) + o(1) = 1 + o(1). \end{aligned}$$

where the equality in the second line follows since $r(x) \to 0$ and from the definitions, and equality in the third line holds if and only if and only if F satisfies (3.20). Thus, we have shown that F satisfies (3.20) if and only if F_* satisfies the analogous condition, with F_* in place of F and r_* in place of r. In other words, without loss of generality, we may assume from the very beginning that $r(x)$ is non-increasing for all $x \geq 0$.

Subexponentiality is equivalent to the convergence: as $x \to \infty$,

$$\int_0^x e^{R(x)-R(x-y)} F(dy) := \int_0^x e^{R(x)-R(x-y)-R(y)} r(y)dy \to 1. \tag{3.21}$$

Since $r(x) = R'(x)$ is non-increasing, $R(x)$ is concave and

$$R(x) - R(x-y) \geq yr(x) \quad \text{for any } y \in [0, x].$$

Hence, subexponentiality in the form (3.21) implies

$$\limsup_{x\to\infty} \int_0^x e^{yr(x)} F(dy) \leq 1.$$

Together with the fact that the integral in the above expression is at least $F[0, x]$, this implies (3.20).

Now suppose that (3.20) holds. We make use of the following representation:

$$\begin{aligned} \int_0^x e^{R(x)-R(x-y)} F(dy) &= \left(\int_0^{x/2} + \int_{x/2}^x\right) e^{R(x)-R(x-y)-R(y)} r(y)dy \\ &= \int_0^{x/2} e^{R(x)-R(x-y)-R(y)} r(y)dy \\ &\quad + \int_0^{x/2} e^{R(x)-R(x-y)-R(y)} r(x-y)dy \\ &=: I_1 + I_2. \end{aligned}$$

The first integral is not less than $F[0, x/2]$ which tends to 1 as $x \to \infty$. On the other hand, for $y \leq x/2$, and therefore $x - y \geq x/2$,

$$R(x) - R(x-y) \leq yr(x-y) \leq yr(x/2). \tag{3.22}$$

Thus,

$$I_1 \leq \int_0^{x/2} e^{yr(x/2)} F(dy),$$

which tends to 1 as $x \to \infty$ by (3.20). Thus $I_1 \to 1$.

On noting that, for any fixed y, $e^{R(x)-R(x-y)-R(y)} r(y) \to e^{-R(y)} r(y)$ as $x \to \infty$ and that $\int_0^\infty e^{-R(y)} r(y)dy = 1$, we obtain that the family (in x) of functions (in y)

$$z_x(y) = e^{R(x)-R(x-y)-R(y)} r(y) \mathbb{I}\{y \leq x/2\}$$

is uniformly integrable in the sense that

$$\sup_x \int_A^\infty z_x(y)dy \to 0 \quad \text{as } A \to \infty.$$

Since $r(x-y) \le r(y)$ for all $y \le x/2$, the integrand in I_2 is dominated by $z_x(y)$. It follows that $I_2 \to 0$ as $x \to \infty$, since also $e^{R(x)-R(x-y)-R(y)} r(x-y) \le r(x-y) \to 0$ for any fixed y. Thus (3.21) holds; that is, the condition (3.20) implies subexponentiality.

The second part of the theorem follows by dominated convergence, since, for all sufficiently large $y < x$, we have $r(x) \le r(y)$. □

As an example, consider a distribution F such that, for some $\alpha > 0$ and for all sufficiently large x,

$$\overline{F}(x) = e^{-x/\log^\alpha x} \tag{3.23}$$

Then, again for sufficiently large x, the hazard rate function r is given by $r(x) = 1/\log^\alpha x - \alpha/\log^{\alpha+1} x$ and the function

$$e^{xr(x)-R(x)} r(x) = e^{-\alpha x/\log^{\alpha+1} x} r(x)$$

(where, as usual, R is the corresponding hazard function) is integrable over $\mathbb{R}^+$. Therefore, by Theorem 3.31, F is subexponential.

In the following theorem we give an applicable necessary and sufficient condition for membership of the class $\mathcal{S}^*$.

Theorem 3.32. *Let F be a distribution on $\mathbb{R}^+$ with finite mean m. Suppose that the hazard rate function r exists, is eventually non-increasing, and that $r(x) \to 0$ as $x \to \infty$. Then $F \in \mathcal{S}^*$ if and only if*

$$\int_0^x e^{yr(x)} \overline{F}(y) dy \to m \quad \text{as } x \to \infty. \tag{3.24}$$

Further, a sufficient condition for $F \in \mathcal{S}^*$ is that (the function of y given by) $e^{yr(y)}\overline{F}(y)$ is integrable over $\mathbb{R}^+$.

Proof. Arguments similar to those used in the proof of Theorem 3.31 show that without loss of generality we may assume that the corresponding hazard function R satisfies $R(0) = 0$ and that the hazard rate function r is non-increasing over all of $\mathbb{R}^+$. The distribution F belongs to the class $\mathcal{S}^*$ if and only if, as $x \to \infty$,

$$\int_0^{x/2} e^{R(x)-R(x-y)} \overline{F}(y) dy := \int_0^{x/2} e^{R(x)-R(x-y)-R(y)} dy \to m. \tag{3.25}$$

Since $r(x) = R'(x)$ is non-increasing, $R(x)$ is concave and

$$R(x) - R(x-y) \ge yr(x) \quad \text{for any } y \in [0, x].$$

Suppose first that the condition (3.25) holds. Then

$$\limsup_{x\to\infty} \int_0^x e^{yr(x)} \overline{F}(y) dy \le m.$$

However, we also have that the latter integral is at least $\int_0^x \overline{F}(y)dy$ which tends to m as $x \to \infty$. Hence the condition (3.24) follows.

Now suppose instead that the condition (3.24) holds. It follows from (3.22) that

$$\int_0^{x/2} e^{R(x)-R(x-y)}\overline{F}(y)dy \le \int_0^{x/2} e^{yr(x/2)}\overline{F}(y)dy,$$

and from (3.24) that this latter integral tends to m as $x \to \infty$.

The second part of the theorem follows by dominated convergence, since, again for all sufficiently large $y < x$, we have $r(x) \le r(y)$. □

As an example we again consider a distribution F whose tail is such that, for some $\alpha > 0$ and for all sufficiently large x, the relation (3.23) holds. We now have that $F \in \mathcal{S}^*$, since in this case the function

$$e^{xr(x)-R(x)} = e^{-\alpha x/\log^{\alpha+1} x}$$

(where the hazard rate function r is given as previously and R is again the corresponding hazard function) is integrable over $\mathbb{R}^+$.

3.7 $\mathcal{S}$ is a proper subset of $\mathcal{L}$

In this section we use Theorem 3.31 to construct a distribution F which is long-tailed but not subexponential. Fix any decreasing sequence $\alpha_n \to 0$ as $n \to \infty$. The corresponding hazard function $R(x)$ will be defined as continuous and piecewise linear so that the hazard rate function $r(x) := R'(x) = \alpha_n$ for $x \in (x_{n-1}, x_n]$. Since on the interval $y \in (x_{n-1}, x_n]$

$$yr(x_n) - R(y) = y\alpha_n - [R(x_{n-1}) + \alpha_n(y - x_{n-1})] > -R(x_{n-1}),$$

we have the following lower bound for the integral on the left side of (3.20):

$$\begin{aligned}\int_{x_{n-1}}^{x_n} e^{yr(x_n)-R(y)}r(y)dy &\ge \int_{x_{n-1}}^{x_n} e^{-R(x_{n-1})}\alpha_n dy \\ &= \alpha_n(x_n - x_{n-1})e^{-R(x_{n-1})}.\end{aligned}$$

Now choose $x_0 = 0$, $R(x_0) = 0$, and the x_n so that

$$\alpha_n(x_n - x_{n-1})e^{-R(x_{n-1})} = 2.$$

For this we take $x_n = x_{n-1} + 2\alpha_n^{-1}e^{R(x_{n-1})}$ and then

$$\begin{aligned}R(x_n) &= R(x_{n-1}) + \alpha_n(x_n - x_{n-1}) \\ &= R(x_{n-1}) + 2e^{R(x_{n-1})}.\end{aligned}$$

Clearly $R(x) \to \infty$ as $x \to \infty$. Since $\alpha_n \to 0$ as $n \to \infty$, $r(x) \to 0$ as $x \to \infty$; thus, $\overline{F}(x) = e^{-R(x)}$ is long-tailed (see Section 2.5). On the other hand, by the above construction,

$$\int_{x_{n-1}}^{x_n} e^{yr(x_n)-R(y)}r(y)dy \ge 2 \quad \text{for all } n,$$

so that

$$\int_0^{x_n} e^{yr(x_n)-R(y)} r(y)dy$$

does not converge to 1 as $n \to \infty$. It now follows from Theorem 3.31 that F is not subexponential.

The idea in this example is that the tail $\overline{F}$ is a piecewise exponential function; the indexes of the exponents tend to zero and the lengths of the intervals of exponentiality grow very fast.

3.8 Does $F \in \mathcal{S}$ imply that $F_I \in \mathcal{S}$?

It is natural to consider the following question: May the assumption $F \in \mathcal{S}^*$ of Theorem 3.28 be weakened to $F \in \mathcal{S}$? In the case of Lebesgue measure μ, i.e. where $F_\mu = F_I$, this question is raised in [23, Section 1.4.2].

In this section, we answer the above question in the negative by giving an example of a distribution $F \in \mathcal{S}$ with finite mean such that $F_I \notin \mathcal{S}$. This example is based on the following construction.

Define $R_0 = 0$, $R_{n+1} = e^{\gamma R_n}$, where $\gamma \in (1/2, 1)$. Since $e^{\gamma x} > x$ for all $x \geq 1$, the sequence R_n is increasing and

$$R_{n+1}/R_n \to \infty \quad \text{as } n \to \infty. \tag{3.26}$$

Put $t_n = R_n^2$. Define the hazard function $R(x) := -\ln \overline{F}(x)$ as

$$R(x) = R_n + r_n(x - t_n) \quad \text{for } x \in (t_n, t_{n+1}],$$

where

$$r_n = \frac{R_{n+1} - R_n}{t_{n+1} - t_n} = \frac{1}{R_{n+1} + R_n} \tag{3.27}$$

$$\sim \frac{1}{R_{n+1}} \quad \text{as } n \to \infty \tag{3.28}$$

by (3.26). In other words, the hazard rate $r(x) = R'(x)$ is defined as $r(x) = r_n$ for $x \in (t_n, t_{n+1}]$, where r_n is given by (3.27). By construction, we have

$$\overline{F}(t_n) = e^{-\sqrt{t_n}},$$

so that at the points t_n the tail function $\overline{F}$ of the distribution F behaves like that of the Weibull distribution with parameter $1/2$.

We shall prove that $F \in \mathcal{S}$ and has finite mean, but that $F_I \notin \mathcal{S}$. Let

$$J_n := F_I(t_n, t_{n+1}] = \int_{t_n}^{t_{n+1}} \overline{F}(u)du = \int_{t_n}^{t_{n+1}} e^{-R(u)}du.$$

Since by (3.28)

$$\begin{aligned} J_n &= r_n^{-1}(e^{-R_n} - e^{-R_{n+1}}) \\ &\sim r_n^{-1} e^{-R_n} \sim R_{n+1} e^{-R_n} = e^{-(1-\gamma)R_n} \quad \text{as } n \to \infty, \end{aligned} \tag{3.29}$$

and $\gamma < 1$, we have $\sum_n J_n < \infty$; hence F has finite mean.

It follows from (3.27) that $r(x)$ is eventually decreasing and tends to 0, and we can thus apply Theorem 3.31 to show that F is subexponential. By that theorem, F is subexponential provided the function $e^{yr(y)-R(y)}r(y)$ is integrable over $\mathbb{R}^+$. We estimate the integral of this function. Put

$$I_n = \int_{t_n}^{t_{n+1}} e^{yr(y)-R(y)} r(y) dy.$$

Then

$$I_n = r_n \int_{t_n}^{t_{n+1}} e^{yr_n - R_n - r_n(y-t_n)} dy \le r_n e^{-R_n + r_n t_n} t_{n+1}.$$

Since, as $n \to \infty$,

$$r_n t_{n+1} = r_n R_{n+1}^2 \sim R_{n+1} \tag{3.30}$$

by (3.28) and

$$r_n t_n = r_n R_n^2 \sim R_n^2 / R_{n+1} = R_n^2 e^{-\gamma R_n} \to 0, \tag{3.31}$$

we get, for n sufficiently large,

$$I_n \le 2R_{n+1} e^{-R_n} \sim 2e^{-(1-\gamma)R_n}.$$

Therefore,

$$\int_0^\infty e^{yr(y)-R(y)} r(y) dy = \sum_{n=0}^\infty I_n < \infty,$$

and F is indeed subexponential.

For $x \in (t_n, t_{n+1}]$, it follows from (3.31) that the density of F_I may be estimated as follows:

$$F_I'(x) = \overline{F}(x) = e^{-R_n - r_n(x-t_n)} \sim e^{-R_n - r_n x} \qquad \text{as } n \to \infty. \tag{3.32}$$

For $x \in (t_n, t_{n+1}]$, define $J_n(x) = F_I[x, t_{n+1})$. We have

$$\begin{aligned} J_n(x) &= e^{-R_n + r_n t_n} \int_x^{t_{n+1}} e^{-r_n y} dy \\ &= r_n^{-1} e^{-R_n + r_n t_n} (e^{-r_n x} - e^{-r_n t_{n+1}}) \\ &\sim J_n e^{-r_n x} (1 - e^{-r_n (t_{n+1} - x)}) \qquad \text{as } n \to \infty, \end{aligned} \tag{3.33}$$

by (3.29) and (3.31). We also have from (3.33) and (3.30) that, for $x \in (t_n, t_{n+1}/2]$,

$$\overline{F}_I(x) \ge J_n(x) \sim J_n e^{-r_n x} \qquad \text{as } n \to \infty. \tag{3.34}$$

For any $x \in (t_n, t_{n+1}]$,

$$\overline{F}_I(x) = J_n(x) + J_{n+1} + J_{n+2} + \dots .$$

By (3.29), $J_{k+1} = o(J_k)$. Fix any $\alpha \in (0, 1-\gamma)$. Then by (3.34), (3.29), and (3.30), as $n \to \infty$,

$$\frac{J_{n+1}}{J_n(\alpha t_{n+1})} \sim \frac{J_{n+1}}{J_n} e^{\alpha r_n t_{n+1}} = e^{(\gamma-1+\alpha)R_{n+1}+o(R_{n+1})} \to 0.$$

Summing and using (3.34), we get

$$\overline{F}_I(\alpha t_{n+1}) \sim J_n(\alpha t_{n+1}) \sim J_n e^{-\alpha r_n t_{n+1}} \quad \text{as } n \to \infty.$$

On the other hand, by (3.32), for n sufficiently large,

$$\begin{aligned}\overline{F_I * F_I}(\alpha t_{n+1}) &\geq \int_{t_n}^{\alpha t_{n+1}-t_n} \overline{F}_I(\alpha t_{n+1} - y) F_I(dy) \\ &\geq (1+o(1)) \int_{t_n}^{\alpha t_{n+1}-t_n} \overline{F}_I(\alpha t_{n+1} - y) e^{-R_n - r_n y} dy.\end{aligned}$$

Applying now (3.34), we get

$$\begin{aligned}\overline{F_I * F_I}(\alpha t_{n+1}) &\geq (1+o(1)) J_n \int_{t_n}^{\alpha t_{n+1}-t_n} e^{-r_n(\alpha t_{n+1}-y)-R_n-r_n y} dy \\ &\sim J_n \alpha t_{n+1} e^{-R_n - \alpha r_n t_{n+1}}.\end{aligned}$$

Then the ratio

$$\frac{\overline{F_I * F_I}(\alpha t_{n+1})}{\overline{F}_I(\alpha t_{n+1})}$$

is asymptotically not less than $\alpha t_{n+1} e^{-R_n} = \alpha e^{R_n(2\gamma-1)}$ which tends to infinity as $n \to \infty$ since $\gamma > 1/2$. Thus, $F \in \mathcal{S}$ and has finite mean, but $F_I \notin \mathcal{S}$.

3.9 Closure properties of the class of subexponential distributions

In this section, we discuss the following question: is the class $\mathcal{S}_{\mathbb{R}}$ closed under convolution? It is well-known that the class of regularly varying distributions, which is a subclass of the class $\mathcal{S}_{\mathbb{R}}$ of subexponential distributions, is closed under convolution. Indeed if F and G are regularly varying, the result that $F * G$ is also regularly varying is straightforwardly obtained from Theorem 3.14 by taking the "reference" distribution of that theorem to be $(F+G)/2$. It is also known that the class $\mathcal{S}_{\mathbb{R}}$ does not possess this closure property. However, if distributions F, $G \in \mathcal{S}_{\mathbb{R}}$, then it follows from Corollary 3.16 that a sufficient condition for $F * G \in \mathcal{S}_{\mathbb{R}}$ is given by $\overline{G}(x) = O(\overline{F}(x))$ as $x \to \infty$. (Indeed, as the corollary shows, G may satisfy weaker conditions than that of being subexponential.) Further it follows that under this condition we have that, for any function h such that $h(x) \to \infty$ and both F and G are h-insensitive,

$$\mathbb{P}\{\xi + \eta > x, \xi > h(x), \eta > h(x)\} = o(\overline{F}(x) + \overline{G}(x)) \quad \text{as } x \to \infty, \tag{3.35}$$

where ξ and η are independent random variables with respective distributions F and G. (See, for example, the proof of Theorem 3.14 above.) The following result is therefore not surprising: if F, $G \in \mathcal{S}_{\mathbb{R}}$, the condition (3.35) is *necessary and sufficient* for $F * G \in \mathcal{S}_{\mathbb{R}}$.

Theorem 3.33. *Suppose that the distributions F and G on $\mathbb{R}$ are subexponential. Then the following conditions are equivalent:*

(i) $\overline{F * G}(x) \sim \overline{F}(x) + \overline{G}(x)$ *as* $x \to \infty$*;*

(ii) $F * G \in \mathcal{S}_{\mathbb{R}}$*;*

(iii) *the mixture $pF + (1-p)G$ belongs to $\mathcal{S}_{\mathbb{R}}$ for all p satisfying $0 < p < 1$;*

(iv) *the mixture $pF + (1-p)G$ belongs to $\mathcal{S}_{\mathbb{R}}$ for some p satisfying $0 < p < 1$;*

(v) *the relation (3.35) holds for any function h such that $h(x) \to \infty$ as $x \to \infty$ and both F and G are h-insensitive;*

(vi) *the relation (3.35) holds for some function h such that $h(x) \to \infty$ as $x \to \infty$ and both F and G are h-insensitive.*

Proof. Let h be any function such that $h(x) \to \infty$ as $x \to \infty$ and both F and G are h-insensitive. We show that each of the conditions (i)–(iv) is equivalent to (3.35). The equivalence of the conditions (i)–(vi) of the theorem is then immediate. First, since F and G are subexponential, and hence long-tailed, it follows from the decomposition (2.18) and Lemma 2.33 that

$$\begin{aligned}&\overline{F * G}(x)\\&\quad = \overline{F}(x) + o(\overline{F}(x)) + \overline{G}(x) + o(\overline{G}(x)) + \mathbb{P}\{\xi + \eta > x, \xi > h(x), \eta > h(x)\}.\end{aligned} \tag{3.36}$$

Hence the condition (i) and (3.35) are equivalent.

To show the equivalence of (ii) and (3.35) observe first that subexponentiality of F and G implies that

$$\overline{F^{*2}}(x) \sim 2\overline{F}(x), \qquad \overline{G^{*2}}(x) \sim 2\overline{G}(x), \tag{3.37}$$

and thus in particular, from Lemma 2.36, that

$$\begin{aligned}&\mathbb{P}\{\xi_1 + \xi_2 + \eta_1 + \eta_2 > x, \xi_1 + \xi_2 > h(x), \eta_1 + \eta_2 > h(x)\}\\&\qquad\qquad \sim 4\mathbb{P}\{\xi + \eta > x, \xi > h(x), \eta > h(x)\}.\end{aligned} \tag{3.38}$$

Further, since $(F * G)^{*2} = F^{*2} * G^{*2}$ and since both F^{*2} and G^{*2} are h-insensitive, $\overline{(F * G)^{*2}}(x)$ may be estimated as in (3.36) with F^{*2} and G^{*2} replacing F and G. Hence, using also (3.37) and (3.38),

$$\begin{aligned}&\overline{(F * G)^{*2}}(x)\\&\quad = (2 + o(1))(\overline{F}(x) + \overline{G}(x)) + (4 + o(1))\mathbb{P}\{\xi + \eta > x, \xi > h(x), \eta > h(x)\}.\end{aligned} \tag{3.39}$$

Now since subexponentiality of F and G also implies, by Corollary 2.41, that $F * G \in \mathcal{L}$, the condition (ii) is equivalent to the requirement that

$$\begin{aligned}\overline{(F * G)^{*2}}(x) &= (2 + o(1))\overline{F * G}(x)\\&= (2 + o(1))(\overline{F}(x) + \overline{G}(x)) + (2 + o(1))\mathbb{P}\{\xi + \eta > x, \xi > h(x), \eta > h(x)\},\end{aligned} \tag{3.40}$$

where (3.40) follows from (3.36). However, the equalities (3.39) and (3.40) hold simultaneously if and only if (3.35) holds.

Finally, to show the equivalence of (iii) (and (iv)) and (3.35), fix p such that $0 < p < 1$ and note first that $pF+(1-p)G$ is h-insensitive. Hence, by Theorem 3.6, subexponentiality of $pF+(1-p)G$ is equivalent to

$$\int_{h(x)}^{\infty} \overline{pF+(1-p)G}(\max(h(x), x-y))(pF+(1-p)G)(dy) = o(\overline{F}(x)+\overline{G}(x)).$$

The left side is equal to

$$p^2\mathbb{P}\{\xi_1+\xi_2 > x, \xi_1 > h(x), \xi_2 > h(x)\} + (1-p)^2\mathbb{P}\{\eta_1+\eta_2 > x, \eta_1 > h(x), \eta_2 > h(x)\} \\ +2p(1-p)\mathbb{P}\{\xi+\eta > x, \xi > h(x), \eta > h(x)\}.$$

By subexponentiality of F and G and again by Theorem 3.6, $\mathbb{P}\{\xi_1 + \xi_2 > x, \xi_1 > h(x), \xi_2 > h(x)\} = o(\overline{F}(x))$ and $\mathbb{P}\{\eta_1+\eta_2 > x, \eta_1 > h(x), \eta_2 > h(x)\} = o(\overline{G}(x))$. The equivalence of (iii) and (3.35) now follows. □

In general, the class $\mathcal{S}_{\mathbb{R}}$ is not closed under convolutions. An example of two subexponential distributions F_1 and F_2 such that $F_1 * F_2$ is not subexponential was constructed by Leslie in [33].

3.10 Kesten's estimate

We know that if a distribution F on $\mathbb{R}$ is subexponential ($F \in \mathcal{S}_{\mathbb{R}}$) then $\overline{F^{*n}}(x)/\overline{F}(x) \to n$ as $x \to \infty$. However, for many purposes, e.g. the application of the dominated convergence theorem, an upper bound for $\overline{F^{*n}}(x)/\overline{F}(x)$ is required. One such is given by the theorem below, known as Kesten's estimate.

Theorem 3.34. *Suppose that $F \in \mathcal{S}_{\mathbb{R}}$. Then, for any $\varepsilon > 0$, there exists $c(\varepsilon) > 0$ such that, for any $x \geq 0$ and $n \geq 1$,*

$$\overline{F^{*n}}(x) \leq c(\varepsilon)(1+\varepsilon)^n\overline{F}(x).$$

Proof. Since for any random variable ξ we have $\xi \leq \xi^+$, it is sufficient to prove the theorem for distributions on the positive half-line $\mathbb{R}^+$. Let $\xi_1, \xi_2, \ldots$ be a sequence of independent random variables with common distribution F, and, for each n, let $S_n = \sum_{i=1}^n \xi_i$. For $x_0 > 0$ and $k \geq 1$, put

$$A_k := A_k(x_0) = \sup_{x > x_0} \frac{\overline{F^{*k}}(x)}{\overline{F}(x)}.$$

Take $\varepsilon > 0$. It follows from subexponentiality that there exists x_0 such that, for any $x > x_0$,

$$\begin{aligned} \mathbb{P}\{\xi_1+\xi_2 > x,\ \xi_2 \leq x\} &= \mathbb{P}\{\xi_1+\xi_2 > x\} - \mathbb{P}\{\xi_2 > x\} \\ &\leq (1+\varepsilon/2)\overline{F}(x). \end{aligned}$$

We have the following decomposition

$$\begin{aligned} \mathbb{P}\{S_n > x\} &= \mathbb{P}\{S_n > x,\ \xi_n \leq x-x_0\} + \mathbb{P}\{S_n > x,\ \xi_n > x-x_0\} \\ &=: P_1(x) + P_2(x). \end{aligned}$$

By the definitions of A_{n-1} and x_0, for any $x > x_0$,

$$\begin{aligned} P_1(x) &= \int_0^{x-x_0} \mathbb{P}\{S_{n-1} > x-y\}\mathbb{P}\{\xi_n \in dy\} \\ &\le A_{n-1}\int_0^{x-x_0} \overline{F}(x-y)\mathbb{P}\{\xi_n \in dy\} \\ &= A_{n-1}\mathbb{P}\{\xi_1+\xi_n > x,\ \xi_n \le x-x_0\} \\ &\le A_{n-1}(1+\varepsilon/2)\overline{F}(x). \end{aligned} \tag{3.41}$$

Further, for any $x > x_0$,

$$P_2(x) \le \mathbb{P}\{\xi_n > x-x_0\} \le L\overline{F}(x), \tag{3.42}$$

where

$$L = \sup_y \frac{\overline{F}(y-x_0)}{\overline{F}(y)}.$$

Since F is long-tailed, L is finite. It follows from (3.41) and (3.42) that $A_n \le A_{n-1}(1+\varepsilon/2)+L$ for $n > 1$. Therefore, an induction argument yields:

$$A_n \le A_1(1+\varepsilon/2)^{n-1} + L\sum_{l=0}^{n-2}(1+\varepsilon/2)^l \le Ln(1+\varepsilon/2)^{n-1}.$$

This implies the conclusion of the theorem. □

It is straightforward to check that the above proof depends on F only through the quantity $|\overline{F*F}(x)/\overline{F}(x)-2|$. We hence obtain immediately the following uniform version of Kesten's estimate.

Theorem 3.35. *Suppose that the family $\mathcal{F}$ is uniformly subexponential, that is,*

$$\sup_{F\in\mathcal{F}}\Big|\frac{\overline{F*F}(x)}{\overline{F}(x)}-2\Big| \to 0 \quad \textit{as } x\to\infty, \tag{3.43}$$

and, in addition, for any $y > 0$,

$$\sup_{F\in\mathcal{F}}\sup_x \frac{\overline{F}(x-y)}{\overline{F}(x)} < \infty. \tag{3.44}$$

Then, for any $\varepsilon > 0$, there exists $c(\varepsilon) > 0$ such that, for any $F \in \mathcal{F}$, $x \ge 0$ and $n \ge 1$,

$$\overline{F^{*n}}(x) \le c(\varepsilon)(1+\varepsilon)^n\overline{F}(x).$$

Recall from Section 3.4 that, for any $b > 0$, we define the class $\mathcal{M}_b$ to consist of all non-negative measures μ on $\mathbb{R}^+$ such that $\mu(x, x+1] \le b$ for all x. As before, we define the distribution F_μ on $\mathbb{R}^+$ by its tail:

$$\overline{F}_\mu(x) := \min\Big(1, \int_0^\infty \overline{F}(x+t)\mu(dt)\Big), \quad x \ge 0.$$

We now have the following corollary to Theorem 3.35.

Corollary 3.36. *Assume that $F \in \mathcal{S}^*$ and $b > 0$. Then, for any $\varepsilon > 0$, there exists $c(\varepsilon) > 0$ such that, for all $\mu \in \mathcal{M}_b$, $x \geq 0$ and $n \geq 1$,*

$$\overline{F_\mu^{*n}}(x) \leq c(\varepsilon)(1+\varepsilon)^n \overline{F}_\mu(x).$$

Proof. We check the conditions of Theorem 3.35. The uniform subexponentiality follows from Theorem 3.28. Fix $y > 0$. Since $F \in \mathcal{S}^*$, F is long-tailed and, therefore, there exists $c < \infty$ such that $\overline{F}(x-y) \leq c\overline{F}(x)$ for all x. Then

$$\int_0^\infty \overline{F}(x-y+t)\mu(dt) \leq c\int_0^\infty \overline{F}(x+t)\mu(dt),$$

and so also the condition (3.44) holds.

□

3.11 Subexponentiality and randomly stopped sums

In this Section we study tail asymptotics for the distribution of a sum (of independent identically distributed random variables) stopped at a random time which is independent of the summands. These results may be used in a variety of areas including the theory of random walks, branching processes, infinitely divisible laws, etc.

Let $\xi, \xi_1, \xi_2, \ldots$ be independent random variables with a common distribution F on $\mathbb{R}^+$. Let $S_0 = 0$ and, for $n \geq 1$, let $S_n = \xi_1 + \ldots + \xi_n$. Let the counting random variable τ be independent of the sequence $\{\xi_n\}$ and take values in $\mathbb{Z}^+$. Then the distribution of S_τ is given by

$$F^{*\tau} = \sum_{n=0}^\infty \mathbb{P}\{\tau = n\} F^{*n}. \tag{3.45}$$

The first result below says that if the random variable τ has a light-tailed distribution and the distribution F of the random variables ξ_i has a subexponential distribution, then again there holds the "principle of a single big jump" introduced in Section 3.1.

Theorem 3.37. *Suppose that $\mathbb{E}\tau < \infty$, that $F \in \mathcal{S}_\mathbb{R}$ and that $\mathbb{E}(1+\delta)^\tau < \infty$ for some $\delta > 0$. Then*

$$\frac{\mathbb{P}\{S_\tau > x\}}{\overline{F}(x)} \to \mathbb{E}\tau \quad \text{as } x \to \infty. \tag{3.46}$$

Proof. The proof is immediate from Corollary 3.20, Theorem 3.34, and the dominated convergence theorem. □

Here the result is valid for any subexponential distribution on the whole real line. For a fixed distribution F, the condition $\mathbb{E}(1+\delta)^\tau < \infty$ may be substantially weakened. We can illustrate this by the following example. Assume that there exist finite positive constants c and α such that $\overline{F}(x/n) \leq cn^\alpha \overline{F}(x)$ for all $x > 0$ and $n \geq 1$ (for instance, the Pareto distribution with parameter α satisfies this condition). Then $\mathbb{P}\{S_\tau > x\} \sim \mathbb{E}\tau \cdot \overline{F}(x)$ as

$x \to \infty$ provided $\mathbb{E}\tau^{1+\alpha}$ is finite; this follows by combining the dominated convergence with the inequalities

$$\mathbb{P}\{S_n > x\} \le \mathbb{P}\{n \cdot \max_{k\le n} \xi_k > x\} \le n\mathbb{P}\{\xi_1 > x/n\} \le n^{1+\alpha}\overline{F}(x).$$

The next result shows that subexponentiality on the positive half-line $\mathbb{R}^+$ is essentially characterised by the relation (3.46).

Theorem 3.38. *Suppose that $\mathbb{E}\tau < \infty$ and that $\mathbb{P}\{\tau > 1\} > 0$. Suppose further that the distribution F is concentrated on $\mathbb{R}^+$ and that (3.46) holds. Then $F \in \mathcal{S}$.*

Proof. For each positive integer k, let $p_k = \mathbb{P}\{\tau = k\}$; note also that, from (2.6), since F is concentrated on $\mathbb{R}^+$ and has unbounded support,

$$\liminf_{x\to\infty} \frac{\overline{F^{*k}}(x)}{\overline{F}(x)} \ge k. \tag{3.47}$$

Let $n \ge 2$ be such that $p_n > 0$. Then, from (3.45) and (3.46),

$$\begin{aligned}
\mathbb{E}\tau &= \lim_{x\to\infty} \frac{\mathbb{P}\{S_\tau > x\}}{\overline{F}(x)} \\
&\ge \liminf_{x\to\infty} \sum_{k\ne n} p_k \frac{\overline{F^{*k}}(x)}{\overline{F}(x)} + p_n \limsup_{x\to\infty} \frac{\overline{F^{*n}}(x)}{\overline{F}(x)} \\
&\ge \sum_{k\ne n} p_k \liminf_{x\to\infty} \frac{\overline{F^{*k}}(x)}{\overline{F}(x)} + p_n \limsup_{x\to\infty} \frac{\overline{F^{*n}}(x)}{\overline{F}(x)} \\
&\ge \sum_{k\ne n} p_k k + p_n \limsup_{x\to\infty} \frac{\overline{F^{*n}}(x)}{\overline{F}(x)},
\end{aligned}$$

where the third line in the above display follows from Fatou's Lemma and the last line follows from (3.47). Since also $\mathbb{E}\tau = \sum_{k\ge 0} p_k k$ and $p_n > 0$, it follows that

$$\limsup_{x\to\infty} \frac{\overline{F^{*n}}(x)}{\overline{F}(x)} \le n,$$

which, by Theorem 3.21, implies the subexponentiality of F. □

The uniform version of Kesten's estimate, see Theorem 3.35, implies the following result for families of distributions.

Theorem 3.39. *Let $\delta > 0$ and $c < \infty$. Suppose that the family of distributions $\mathcal{F}$ is uniformly subexponential, that is,*

$$\sup_{F\in\mathcal{F}} \left| \frac{\overline{F*F}(x)}{\overline{F}(x)} - 2 \right| \to 0 \quad \textit{as } x \to \infty,$$

and that, in addition, for any $y > 0$,

$$\sup_{F\in\mathcal{F}} \sup_{x} \frac{\overline{F}(x-y)}{\overline{F}(x)} < \infty.$$

Then

$$\sum_{n=0}^{\infty} \overline{F^{*n}}(x)\mathbb{P}\{\tau = n\} \sim \mathbb{E}\tau \overline{F}(x)$$

as $x \to \infty$ *uniformly in* $F \in \mathcal{F}$ *and in all* τ *such that* $\mathbb{E}(1+\delta)^{\tau} \le c$.

Together with Corollary 3.36, Theorem 3.39 implies the following uniform asymptotics.

Corollary 3.40. *Let* $b > 0$, $\delta > 0$ *and* $c < \infty$. *Suppose that* $F \in \mathcal{S}^*$. *Then*

$$\sum_{n=0}^{\infty} \overline{F_{\mu}^{*n}}(x)\mathbb{P}\{\tau = n\} \sim \mathbb{E}\tau \overline{F}_{\mu}(x)$$

as $x \to \infty$ *uniformly in* $\mu \in \mathcal{M}_b$ *and in all* τ *such that* $\mathbb{E}(1+\delta)^{\tau} \le c$.

3.12 Comments

The concept of subexponential distributions (but not the name) was introduced by Chistyakov in [12], in the context of branching processes. In the same paper, the present Lemma 3.2 was established as well as some sufficient conditions for subexponentiality. Also, Theorem 3.34 (Kesten's estimate) was proved under an additional technical assumption.

The notion of weak tail-equivalence and Theorem 3.11 go back to Klüppelberg [29].

The class $\mathcal{S}^*$ was introduced by Klüppelberg [29].

Corollary 3.18 was proved by Embrechts et al. [22].

The version of Corollary 3.16 with $G \in \mathcal{L}$ was proved by Embrechts and Goldie [20]). Corollary 3.19 is well-known (and goes back to [21] where the case $n = 2$, $G_1 = G_2$ was considered; some particular results may be found in Teugels [42] and Pakes [34], see also [4]).

Theorem 3.31 is due to Pitman [36].

Examples where F is long-tailed but not subexponential can be found in Embrechts and Goldie [20], and in Pitman [36]. Here we have followed the idea of Pitman.

The first four equivalences given by Theorem 3.33 were proved by Embrechts and Goldie in [20].

Chapter 4

Densities and local probabilities

This chapter is devoted to local long-tailedness and to local subexponentiality. First we consider densities with respect to either Lebesgue measure on $\mathbb{R}$ or counting measure on $\mathbb{Z}$. Next we study the asymptotic behaviour of the probabilities to belong to an interval of a fixed length. We give the analogues of the basic properties of the tail probabilities including two analogues of Kesten's estimate, and provide sufficient conditions for probability distributions to have these local properties..

The study of local properties of subexponentiality gives insights into the local asymptotic behaviour of sums and maxima of random variables having heavy-tailed distributions and, in particular, permits us to obtain the local asymptotics for the supremum of a random walk with negative drift, and also new results related to the renewal equation. The concept of a subexponential density on the positive line is well-known, while the broader concept of "delta"-subexponentiality has been introduced recently [4]. The theories for these two classes of distributions look similar, but there are (sometimes essential) differences in the ideas and proofs, and we therefore think that it makes sense to provide a complete treatment of both concepts.

Sections 4.1-4.3 deal with long-tailed densities, subexponential densities, and sufficient conditions for a distribution to have a subexponential density, while Sections 4.4-4.6 deal with similar topics for Δ-subexponential distributions.

4.1 Long tailed densities and their convolutions

In this section, we provide the definition and basic properties of long-tailed densities on the real line $\mathbb{R}$. Since a long-tailed density may be a non-monotone function, we cannot prove here a general result similar to Theorem 2.40 for tail distribution functions. We provide instead two separate results, Theorem 4.3 and Lemma 4.4.

Let μ be either Lebesgue measure on $\mathbb{R}$ or counting measure on $\mathbb{Z}$. We say that a distribution F on $\mathbb{R}$ is absolutely continuous with respect to μ if F has a density f with respect to μ, that is, for any Borel set $B \subseteq \mathbb{R}$,

$$F(B) = \int_B f(x)\mu(dx).$$

In what follows the argument of the density is either a real number if μ is Lebesgue measure; or an integer if μ is counting measure. If μ is Lebesgue measure, then f is a density of F

if, for any Borel set $B \subseteq \mathbb{R}$,

$$F(B) = \int_B f(x)dx.$$

If μ is counting measure, then f is a density of F if, for any $B \subseteq \mathbb{Z}$,

$$F(B) = \sum_{n \in B} f(n).$$

For two distributions F and G with densities f and g respectively, the convolution $F * G$ has density $f * g$ with respect to μ given by

$$(f * g)(x) = \int_{-\infty}^{\infty} f(x-y)G(dy) = \int_{-\infty}^{\infty} f(x-y)g(y)\mu(dy).$$

Definition 4.1. We say that a density f with respect to μ is *long-tailed* if $f(x) > 0$ for all sufficiently large x and $f(x+t) \sim f(x)$ as $x \to \infty$, for any fixed $t > 0$.

Thus a density f is long-tailed if and only if f is a long-tailed function. As pointed out in (2.18), it then follows that $f(x+t) \sim f(x)$ as $x \to \infty$ uniformly over t in compact intervals. In particular, this implies that if f is long-tailed, then $f(x) \to 0$ as $x \to \infty$. To see this, assume that, on the contrary, there exist a sequence $x_n \to \infty$ and $\varepsilon > 0$ such that $x_{n+1} > x_n + 2$ and $f(x_n) \geq 2\varepsilon$ for all n. Then, from the uniform tail-equivalence (2.18) with $a = 1$, there is N such that, for $n \geq N$ and for $x \in [x_n - 1, x_n + 1)$, $f(x) \geq \varepsilon$. Hence,

$$1 = \int_{-\infty}^{\infty} f(y)\mu(dy) \geq \sum_{n=N}^{\infty} \int_{x_n-1}^{x_n+1} f(y)\mu(dy) \geq \sum_{n=N}^{\infty} 2\varepsilon = \infty.$$

This contradiction proves that $f(x) \to 0$ as $x \to \infty$.

Every distribution F with long-tailed density f is long-tailed itself, since for any fixed y

$$\begin{aligned}\overline{F}(x+y) &= \int_0^{\infty} f(x+y+u)\mu(du) \\ &\sim \int_0^{\infty} f(x+u)\mu(du) = \overline{F}(x) \quad \text{as } x \to \infty.\end{aligned}$$

Theorem 4.2. *Let the distributions F and G on $\mathbb{R}$ have densities f and g with respect to μ. Suppose that f is long-tailed. Then the density $f * g$ satisfies*

$$\liminf_{x \to \infty} \frac{(f * g)(x)}{f(x)} \geq 1. \tag{4.1}$$

If, in addition, g is long-tailed, then

$$\liminf_{x \to \infty} \frac{(f * g)(x)}{f(x) + g(x)} = 1. \tag{4.2}$$

Proof. Fix any $a > 0$. By the uniform convergence (2.18), $f(x-y) \sim f(x)$ as $x \to \infty$ uniformly in $|y| \leq a$. Hence,

$$(f * g)(x) \geq \int_{-a}^{a} f(x-y)G(dy) \sim f(x)G[-a,a) \quad \text{as } x \to \infty.$$

Letting $a \to \infty$ we obtain (4.1).

If $g(x)$ is also long-tailed, then $g(x-y) \sim g(x)$ as $x \to \infty$ uniformly in $|y| \le a$. Thus, for all $x > 2a$,

$$\begin{aligned}(f * g)(x) &\ge \int_{-a}^{a} f(x-y)G(dy) + \int_{-a}^{a} g(x-y)F(dy) \\ &\sim f(x)G[-a,a) + g(x)F[-a,a) \qquad \text{as } x \to \infty.\end{aligned}$$

Letting $a \to \infty$ we obtain

$$\liminf_{x\to\infty} \frac{(f*g)(x)}{f(x)+g(x)} \ge 1.$$

Hence the equality (4.2) will follow if we show that

$$\liminf_{x\to\infty} \frac{(f*g)(x)}{f(x)+g(x)} \le 1.$$

To prove this, assume that, on the contrary, there exist $\varepsilon > 0$ and x_0 such that, for all $x > x_0$,

$$(f*g)(x) \ge (1+\varepsilon)(f(x)+g(x)).$$

Integrating with respect to x we obtain

$$\overline{F*G}(x) \ge (1+\varepsilon)(\overline{F}(x)+\overline{G}(x)).$$

Since the density f is long-tailed, the distribution F is also long-tailed and, therefore, heavy-tailed, and so the latter inequality contradicts Theorem 2.13. □

Theorem 4.3. *Let the distributions F and G on $\mathbb{R}$ have densities f and g with respect to μ both of which are long-tailed. Then the density $f * g$ of the convolution $F * G$ is also long-tailed.*

Proof. By Lemma 2.19 and Proposition 2.20, we can choose a function h such that $h(x) < x/2$, $h(x) \to \infty$ as $x \to \infty$ and both f and g are h-insensitive (i.e. $f(x-y) \sim f(x)$ and $g(x-y) \sim g(x)$ as $x \to \infty$ uniformly in $|y| \le h(x)$). Fix $t > 0$. Then,

$$\begin{aligned}(f*g)(x+t) &= \int_{-\infty}^{x-h(x)} f(x+t-y)G(dy) + \int_{x-h(x)}^{x+t-h(x)} f(x+t-y)G(dy) \\ &\quad + \int_{x+t-h(x)}^{\infty} f(x+t-y)g(y)\mu(dy). \qquad (4.3)\end{aligned}$$

For fixed $t > 0$, it follows from the given conditions on h that $f(x+t-y) \sim f(x-y)$ as $x \to \infty$ uniformly in $y \le x - h(x)$. Therefore, as $x \to \infty$,

$$\int_{-\infty}^{x-h(x)} f(x+t-y)G(dy) \sim \int_{-\infty}^{x-h(x)} f(x-y)G(dy). \qquad (4.4)$$

The second integral is bounded from above by

$$\begin{aligned}\sup_{y\in[h(x),h(x)+t)} f(y)G[x-h(x), x+t-h(x)) &\sim tf(h(x))g(x) \\ &= o(g(x)) = o((f*g)(x)) \qquad (4.5)\end{aligned}$$

as $x \to \infty$, by (4.2). The third integral in (4.3) is equal to

$$\int_{-\infty}^{h(x)} g(x+t-y)F(dy) \sim \int_{-\infty}^{h(x)} g(x-y)F(dy) \tag{4.6}$$

by arguments similar to that leading to (4.4). Collecting (4.4)–(4.6), we get $(f*g)(x+t) = (f*g)(x) + o((f*g)(x))$, since the sum of right sides in (4.4) and (4.6) equals $(f*g)(x)$. This completes the proof. □

Lemma 4.4. *Let the distributions F and G on $\mathbb{R}$ have densities f and g with respect to μ. Suppose that f is long-tailed and that*

$$\sup_{z \geq x} g(z) = o(f(x)) \quad \text{as } x \to \infty.$$

Then $f*g$ is also long-tailed.

Proof. Again Lemma 2.19 with Proposition 2.20 enables us to find an increasing function h such that $h(x) < x/2$, $h(x) \to \infty$ as $x \to \infty$ and f is h-insensitive. For any t, consider the following decomposition:

$$(f*g)(x+t) = \int_{-\infty}^{x-h(x)} f(x+t-y)G(dy) + \int_{x-h(x)}^{\infty} f(x+t-y)g(y)\mu(dy).$$

The first integral satisfies (4.4). The second integral is not greater than

$$\sup_{y > x-h(x)} g(y) = o(f(x-h(x))) = o(f(x)). \tag{4.7}$$

It follows from (4.4) and (4.7) that, as $x \to \infty$,

$$(f*g)(x+t) = (1+o(1))(f*g)(x) + o(f(x)).$$

Applying now the result (4.1) of Theorem 4.2, we arrive at the desired equivalence $(f*g)(x+t) \sim (f*g)(x)$ as $x \to \infty$. □

Theorems 4.2 and 4.3 imply the following corollary.

Corollary 4.5. *Suppose that f is long-tailed. Then f^{*n} is also long-tailed and*

$$\liminf_{x\to\infty} \frac{f^{*n}(x)}{f(x)} \geq n.$$

4.2 Subexponential densities

We start this Section with the introduction of the concept of a subexponential density on the whole real line $\mathbb{R}$ and with some discussion of an alternative definition, and also of the properties and relations of these two definitions (see Subsection 4.2.1). Then we move to the more classical subject of subexponential densities on the positive half-line $\mathbb{R}^+$ (see Subsection 4.2.2) where the two earlier definitions coincide.

4.2.1 Subexponential densities on the real line

In the previous chapter we showed that there are three equivalent ways to define the subexponentiality of probability distributions on the whole real line: the distribution F of a random variable ξ is subexponential if

(1) it is long-tailed and $\overline{F * F}(x) \sim 2\overline{F}(x)$ as $x \to \infty$, or, equivalently, if
(2) the conditional distribution $\mathbb{P}\{\xi \in \cdot \,\big|\, \xi \geq 0)$ on the positive half-line $\mathbb{R}^+$ is subexponential, or, equivalently, if
(3) the distribution F^+ of the random variable ξ^+ is subexponential.

We also showed that if F is a subexponential distribution on $\mathbb{R}$ and G is a distribution on $\mathbb{R}$ such that $\overline{F}(x) \sim \overline{G}(x)$ as $x \to \infty$, then G is also subexponential. Thus in particular subexponentiality on $\mathbb{R}$ is a *tail property* of distribution functions as defined in Chapter 1.

For probability densities, the analogue of the third definition above does not make much sense in general, since, in the case of Lebesgue measure, for a random variable ξ such that $\mathbb{P}\{\xi < 0\} > 0$, the random variable ξ^+ is not absolutely continuous. However, the two first definitions have natural analogues. While these two definitions agree for distributions on the positive half-line $\mathbb{R}^+$ they seem not to be generally equivalent in the case of distributions F on $\mathbb{R}$. Indeed, there is no unity among the authors about which of these definitions is better and more adequately reflects the concept of subexponentiality of densities, and about whether subexponentiality in this case should necessarily be a tail property. The first definition is formulated in terms of the distribution itself, but the left tail of the distribution may influence the right-tail asymptotics of its convolution with itself and, therefore, the tail property does not hold. The second definition (which we may call "conditional subexponentiality") depends only on the truncation of the distribution to the positive half-line and possesses then the tail property.

This section is organised as follows. First, we give the first definition and discuss its relation to subexponentiality of distributions. Then we formulate the main equivalence result for subexponential densities on the positive half-line which is needed for the two next results. Finally, we give the definition of "conditional subexponentiality" of densities and show that subexponentiality of a density implies its conditional subexponentiality, but that the converse holds only under additional assumptions.

Definition 4.6. We say that a density f on $\mathbb{R}$ with respect to μ is *subexponential* if f is long-tailed and

$$f^{*2}(x) := \int_{-\infty}^{\infty} f(x-y) f(y)\mu(dy) \sim 2f(x) \quad \text{as } x \to \infty.$$

Typical examples of subexponential densities are given by the Pareto, lognormal, and Weibull (with parameter between 0 and 1) distributions (see Section 4.3 for proofs).

Every distribution F with subexponential density f is subexponential itself, since F is then long-tailed and, as $x \to \infty$,

$$\begin{aligned}
\overline{F * F}(x) &= \int_x^{\infty} (f * f)(y)\mu(dy) \\
&\sim 2\int_x^{\infty} f(y)\mu(dy) \\
&= 2\overline{F}(x).
\end{aligned}$$

The converse result is not in general true: one can, for example, modify a density while keeping the corresponding distribution almost the same. For example, we may take any subexponential density g corresponding to a, necessarily subexponential, distribution G, and construct a new density f such that $f(x)$ is equal to $g(x)$ everywhere except the intervals $x \in [2^n, 2^n+1)$, $n \geq 1$ where we put $f(x) = 0$. To make f a probability density, we may add an appropriate mass to the interval $[0, 2]$. Then the density f is not subexponential because it is not long-tailed. On the other hand, the corresponding distribution F is subexponential, since it may easily be verified that $\overline{F}(x) \sim \overline{G}(x)$ as $x \to \infty$ and G is subexponential.

Now we formulate the basic theorem for subexponential densities on the positive half-line $\mathbb{R}^+$.

Theorem 4.7. *Suppose that the distribution F on $\mathbb{R}^+$ has a long-tailed density f with respect to μ. Then the following assertions are equivalent:*

(i) *the density f is subexponential;*

(ii) *for every function h such that $h(x) \to \infty$ as $x \to \infty$,*

$$\int_{h(x)}^{x-h(x)} f(x-y)f(y)\mu(dy) = o(f(x)) \text{ as } x \to \infty; \tag{4.8}$$

(iii) *the relation (4.8) holds for some function h such that $h(x) < x/2$, $h(x) \to \infty$ as $x \to \infty$ and f is h-insensitive.*

Proof. (i)⇒(ii). Assume that f is subexponential. Clearly it follows from the monotonicity in $h(x)$ of the integral in (4.8) that we may assume without loss of generality that $h(x) < x/2$ for all x. Then

$$f^{*2}(x) = 2\int_0^{h(x)} f(x-y)f(y)\mu(dy) + \int_{h(x)}^{x-h(x)} f(x-y)f(y)dy. \tag{4.9}$$

By Fatou's lemma,

$$\liminf_{x\to\infty} \int_0^{h(x)} \frac{f(x-y)}{f(x)} f(y)\mu(dy) \geq 1,$$

and so (4.8) follows from the subexponentiality of f.

(ii)⇒(iii). This implication is trivial on recalling that f is long-tailed.

(iii)⇒(i). Assume now that the relation (4.8) holds for some function h as given by (iii). Then (4.9) holds, and the first integral on the right of (4.9) is tail-equivalent to $f(x)$ (as $x \to \infty$) by the choice of the function h. Together with the condition (4.8) this implies the subexponentiality of f. □

Note that it follows from the above theorem that for a distribution on $\mathbb{R}^+$ subexponentiality of its density is indeed a tail property.

We now formulate the concept of conditional subexponentiality for a density on $\mathbb{R}$.

Definition 4.8. A density f on $\mathbb{R}$ is *conditionally subexponential* if the density f^+ of the corresponding conditional distribution on the positive half-line, that is

$$f^+(x) := \frac{f(x)\mathbb{I}\{x \geq 0\}}{F(\mathbb{R}^+)}, \tag{4.10}$$

is subexponential.

The next two results show the relation between subexponentiality and conditional subexponentiality of densities.

Lemma 4.9. *Suppose that the distribution F on $\mathbb{R}$ has a subexponential density f with respect to μ. Then the density f is also conditionally subexponential.*

Proof. Since the subexponentiality of f implies that it is also long-tailed, by Lemma 2.19, we can choose a function h such that $h(x) < x/2$, $h(x) \to \infty$ as $x \to \infty$ and f is h-insensitive. Then

$$\begin{aligned} f^{*2}(x) &= 2\int_{-\infty}^{-h(x)} f(x-y)f(y)\mu(dy) + 2\int_{-h(x)}^{h(x)} f(x-y)f(y)\mu(dy) \\ &\quad + \int_{h(x)}^{x-h(x)} f(x-y)f(y)\mu(dy) \qquad (4.11) \\ &\geq 2\int_{-h(x)}^{h(x)} f(x-y)f(y)\mu(dy) + \int_{h(x)}^{x-h(x)} f(x-y)f(y)\mu(dy) \\ &\sim 2f(x) + \int_{h(x)}^{x-h(x)} f(x-y)f(y)\mu(dy) \end{aligned}$$

as $x \to \infty$, by the choice of the function h. Since also f is subexponential, that is, $f^{*2}(x) \sim 2f(x)$, we obtain that

$$\int_{h(x)}^{x-h(x)} f(x-y)f(y)\mu(dy) = o(f(x)) \quad \text{as } x \to \infty.$$

Hence the density f^+ satisfies the condition (4.8) of Theorem 4.7 and so is subexponential, that is, f is conditionally subexponential. □

We now give the converse result which requires an extra condition.

Lemma 4.10. *Suppose that the distribution F on $\mathbb{R}$ has a density f with respect to μ which is conditionally subexponential. Suppose also that there exists a function h such that $h(x) < x/2$ for all x, that $h(x) \to \infty$ as $x \to \infty$, that f is h-insensitive and that*

$$\int_{-\infty}^{-h(x)} f(x-y)f(y)\mu(dy) = o(f(x)) \quad \textit{as } x \to \infty. \qquad (4.12)$$

Then the density f is also subexponential.

In particular there exists a function h satisfying the above conditions if, for some $c > 0$ and some x_0,

$$f(x+y) \leq cf(x) \quad \textit{for all } x > x_0 \textit{ and } y > 0. \qquad (4.13)$$

Proof. To prove the first part of the lemma, we let the function h be as given in its statement and make use of decomposition (4.11). From the choice of h,

$$\begin{aligned} \int_{-h(x)}^{h(x)} f(x-y)f(y)\mu(dy) &\sim f(x)\int_{-h(x)}^{h(x)} f(y)\mu(dy) \\ &\sim 2f(x) \quad \text{as } x \to \infty. \qquad (4.14) \end{aligned}$$

Since the density f is conditionally subexponential, it follows from Theorem 4.7 that

$$\int_{h(x)}^{x-h(x)} f(x-y)f(y)\mu(dy) = o(f(x)) \quad \text{as } x \to \infty. \tag{4.15}$$

The relations (4.12), (4.14), and (4.15) now imply that $f^{*2}(x) \sim 2f(x)$ as $x \to \infty$.

We now prove the second statement of the lemma. Suppose that the condition (4.13) holds. Since f is conditionally subexponential, it is long-tailed and so, by Lemma 2.19 we may choose a function h such that $h(x) < x/2$, $h(x) \to \infty$ as $x \to \infty$ and f is h-insensitive. Then

$$\begin{aligned}\int_{-\infty}^{-h(x)} f(x-y)f(y)\mu(dy) &\le cf(x+h(x))F(-h(x)) \\ &= o(f(x)) \quad \text{as } x \to \infty.\end{aligned}$$

□

We conclude this section with the following comment. In contrast to tail functions, densities are not in general decreasing functions. Moreover, a subexponential density may be not tail-equivalent to any non-decreasing function and, in particular, the condition (4.13) may fail. For this reason, there is no complete correspondence between subexponentiality and conditional subexponentiality of densities.

4.2.2 Subexponential densities on the positive half-line

In this section, we study further the properties of subexponential densities on the positive half-line $\mathbb{R}^+$, in particular giving closure properties for the class of such densities and providing also the analogue of Kesten's estimate.

Theorem 4.11. *Let f be a subexponential density on $\mathbb{R}^+$ with respect to μ. Suppose that the density g on $\mathbb{R}^+$ is long-tailed and that f and g are weakly tail-equivalent, that is,*

$$0 < \liminf_{x\to\infty} \frac{g(x)}{f(x)} \le \limsup_{x\to\infty} \frac{g(x)}{f(x)} < \infty. \tag{4.16}$$

Then g is also subexponential.

In particular, the condition (4.16) is satisfied if $g(x) \sim cf(x)$ as $x \to \infty$ for some $c \in (0, \infty)$.

Proof. The result follows from Theorem 4.7(ii) and (iii): observe that (4.16) implies that there exists $c_1 < \infty$ such that $g(x) \le c_1 f(x)$ for all sufficiently large x; hence, for any function h such that $h(x) < x/2$ for all x, $h(x) \to \infty$ as $x \to \infty$ and g is h-insensitive,

$$\begin{aligned}\int_{h(x)}^{x-h(x)} g(x-y)g(y)\mu(dy) &\le c_1^2 \int_{h(x)}^{x-h(x)} f(x-y)f(y)\mu(dy) \\ &= o(f(x)) \\ &= o(g(x)) \qquad \text{as } x \to \infty.\end{aligned}$$

□

Lemma 4.12. *Let f be a subexponential density on $\mathbb{R}^+$ with respect to μ. Let f_1, f_2 be two densities on $\mathbb{R}^+$ such that $f_1(x)/f(x) \to c_1$ and $f_2(x)/f(x) \to c_2$ as $x \to \infty$, for some constants c_1, $c_2 \geq 0$. Then*

$$\frac{(f_1 * f_2)(x)}{f(x)} \to c_1 + c_2 \quad \text{as } x \to \infty. \tag{4.17}$$

Further, if $c_1 + c_2 > 0$ then the convolution $f_1 * f_2$ is a subexponential density.

Proof. Let h be any function such that $h(x) < x/2$, $h(x) \to \infty$ as $x \to \infty$ and f is h-insensitive (note that this implies the h-insensitivity of f_1, f_2 also). Then

$$\begin{aligned} f_1 * f_2(x) &= \int_0^{h(x)} f_1(x-y) f_2(y)\mu(dy) + \int_0^{h(x)} f_2(x-y) f_1(y)\mu(dy) \\ &\quad + \int_{h(x)}^{x-h(x)} f_1(x-y) f_2(y)\mu(dy) \\ &=: I_1(x) + I_2(x) + I_3(x). \end{aligned}$$

We have $I_1(x)/f(x) \to c_1$ and $I_2(x)/f(x) \to c_2$ as $x \to \infty$. Finally,

$$I_3(x) \leq (c_1 c_2 + o(1)) \int_{h(x)}^{x-h(x)} f(x-y) f(y)\mu(dy) = o(f(x)),$$

by Theorem 4.7(ii), so that (4.17) now follows. The final assertion of the lemma follows from Theorem 4.11, □

Using induction arguments, we obtain the following corollary.

Corollary 4.13. *Assume that f is a subexponential density on $\mathbb{R}^+$ with respect to μ. Then, for any $n \geq 2$, $f^{*n}(x) \sim n f(x)$ as $x \to \infty$ and f^{*n} is a subexponential density.*

For subexponential densities we have the following analogue of Kesten's estimate.

Theorem 4.14. *Assume that f is a subexponential density on $\mathbb{R}^+$ with respect to μ. If f is bounded, then, for any $\varepsilon > 0$, there exist $x_0 = x_0(\varepsilon)$ and $c(\varepsilon) > 0$ such that, for any $x > x_0$ and for any integer $n \geq 1$,*

$$f^{*n}(x) \leq c(\varepsilon)(1+\varepsilon)^n f(x).$$

Proof. Take $c < \infty$ such that $f(x) \leq c$ for all $x \geq 0$. Then it follows from the convolution formula that

$$f^{*n}(x) \leq cF[0,x] \leq c \quad \text{for all } x \geq 0 \text{ and } n \geq 1. \tag{4.18}$$

Since f is long-tailed, there exists x_1 such that

$$\inf_{x \in [x_1, x_2]} f(x) > 0 \quad \text{for every } x_2 > x_1. \tag{4.19}$$

For $x_0 > x_1$ and $n \geq 1$, put

$$A_n(x_0) := \sup_{x > x_0} \frac{f^{*n}(x)}{f(x)}.$$

Fix any $\varepsilon > 0$. By the subexponentiality of f, there exists x_0 such that, for all $x > x_0$,

$$\int_0^{x-x_0} f(x-y)f(y)\mu(dy) \le (1+\varepsilon/2)f(x).$$

For any $n \ge 2$ and $x > 2x_0$,

$$f^{*n}(x) = \int_0^{x-x_0} f^{*(n-1)}(x-y)f(y)\mu(dy) + \int_0^{x_0} f(x-y)f^{*(n-1)}(y)\mu(dy).$$

By the definition of $A_{n-1}(x_0)$ and the choice of x_0,

$$\begin{aligned}\int_0^{x-x_0} f^{*(n-1)}(x-y)f(y)\mu(dy) &\le A_{n-1}(x_0)\int_0^{x-x_0} f(x-y)f(y)\mu(dy)\\ &\le A_{n-1}(x_0)(1+\varepsilon/2)f(x). \qquad (4.20)\end{aligned}$$

Further,

$$\int_0^{x_0} f(x-y)f^{*(n-1)}(y)\mu(dy) \le \max_{0<y\le x_0} f(x-y) \le L_1 f(x), \qquad (4.21)$$

where

$$L_1 := \sup_{0<y\le x_0, t>2x_0} \frac{f(t-y)}{f(t)}.$$

If $x_0 < x \le 2x_0$, then, by (4.18) and (4.19),

$$\frac{f^{*n}(x)}{f(x)} \le \frac{c}{\inf_{x_0<t\le 2x_0} f(t)} =: L_2 < \infty. \qquad (4.22)$$

Since f is long-tailed, we may choose x_0 so that also $L_1 < \infty$. Put $L = \max(L_1, L_2)$. It follows from (4.20)–(4.22) that, for any $x > x_0$,

$$f^{*n}(x) \le (A_{n-1}(x_0)(1+\varepsilon/2) + L)f(x).$$

Hence, $A_n(x_0) \le A_{n-1}(x_0)(1+\varepsilon/2) + L$. Therefore, an induction argument gives

$$A_n(x_0) \le A_1(x_0)(1+\varepsilon/2)^{n-1} + L\sum_{l=0}^{n-2}(1+\varepsilon/2)^l \le Ln(1+\varepsilon/2)^{n-1},$$

which implies the conclusion of the theorem. □

4.3 Sufficient conditions for subexponentiality of densities

Sufficient conditions for distributions to be subexponential were given in Section 3.5. In this section, we provide similar conditions for subexponentiality of densities.

Theorem 4.15. *Let the distribution F on $\mathbb{R}^+$ have a long-tailed density f. Suppose that there exist $c > 0$ and x_0 such that $f(y) \ge cf(x)$ for any $x > x_0$ and $y \in (x, 2x]$. Then the density f is subexponential.*

Proof. Let h be any positive function such that $h(x) \to \infty$ as $x \to \infty$ and $h(x) < x/2$ for all x. Then

$$\begin{aligned}\int_{h(x)}^{x-h(x)} f(x-y)f(y)\mu(dy) &= 2\int_{h(x)}^{x/2} f(x-y)f(y)\mu(dy)\\ &\le 2cf(x)\int_{h(x)}^{x/2} f(y)\mu(dy) = o(f(x))\end{aligned}$$

as $x \to \infty$. The subexponentiality of f now follows from Theorem 4.7(ii). □

Observe that in particular the density of the Pareto distribution satisfies the conditions of Theorem 4.15.

Theorem 4.16. *Let the distribution F on $\mathbb{R}^+$ have a long-tailed density f. Suppose that, for some x_0, the function $g(x) := -\ln f(x)$ is concave for $x \ge x_0$. Suppose further that there exists a function h such that $h(x) < x/2$ for all x, that $h(x) \to \infty$ as $x \to \infty$, that f is h-insensitive, and that $xe^{-g(h(x))} \to 0$ as $x \to \infty$. Then the density f is subexponential.*

Proof. By Theorem 4.11, without loss of generality we may assume $x_0 = 0$. Since g is concave, the minimum of the sum $g(x-y) + g(y)$ in $y \in [h(x), x-h(x)]$ is equal to $g(x-h(x)) + g(h(x))$. Therefore,

$$\int_{h(x)}^{x-h(x)} f(x-y)f(y)\mu(dy) = \int_{h(x)}^{x-h(x)} e^{-(g(x-y)+g(y))}\mu(dy) \le xe^{-(g(x-h(x))+g(h(x)))}.$$

Since $e^{-g(x-h(x))} \sim e^{-g(x)}$,

$$\int_{h(x)}^{x-h(x)} f(x-y)f(y)\mu(dy) = O(e^{-g(x)}xe^{-g(h(x))}) = o(f(x)),$$

so that the result now follows from Theorem 4.7. □

The density of the Weibull distribution with parameter $\alpha \in (0,1)$ satisfies conditions of Theorem 4.16 with $h(x) = \ln^{2/\alpha} x$. The density of the log-normal distribution satisfies the these conditions with $h(x) = \sqrt{x}$.

4.4 Δ-Long-tailed distributions and their convolutions

This section and the next deal with local properties of long-tailedness and subexponentiality which may be considered as intermediate properties of a distribution between that of being long-tailed/subexponential and that of having a long-tailed/subexponential density, and are formulated in terms of the probability for a random variable to belong to an interval of a fixed length when the location of the interval is tending to infinity.

Define $\Delta = (0, T]$ for some finite $T > 0$. For any x and for any nonnegative integer n, define also $x + \Delta := (x, x+T]$ and $n\Delta := (0, nT]$.

We now introduce the following definition.

Definition 4.17. A distribution F on $\mathbb{R}$ is called Δ-*long-tailed* if $F(x+\Delta)$ is a long-tailed function, that is, for any fixed $y>0$,

$$\frac{F(x+y+\Delta)}{F(x+\Delta)} \to 1 \quad \text{as } x\to\infty.$$

By the property (2.18) of long-tailed functions, the latter convergence holds uniformly over all y in any compact set. We write also $\mathcal{L}_\Delta$ for the class of Δ-long-tailed distributions. We consider here only finite intervals Δ, but if we allowed the interval to be infinite, $\Delta = (0,\infty)$, we would have $\mathcal{L}_\Delta = \mathcal{L}$, the class of long-tailed distributions.

It follows from the definition that, if $F\in\mathcal{L}_\Delta$ for some interval $\Delta=(0,T]$, then $F\in\mathcal{L}_{n\Delta}$ for any $n=2,3,\dots$ and also $F\in\mathcal{L}$. To see this observe that, for any fixed $y>0$ and any $n\in\{2,3,\dots,\infty\}$,

$$\begin{aligned} F(x+y+n\Delta) &= \sum_{k=0}^{n-1} F(x+kT+y+\Delta) \\ &\sim \sum_{k=0}^{n-1} F(x+kT+\Delta) = F(x+n\Delta). \end{aligned}$$

Note that any distribution F on the integer lattice with $F\{n+1\}\sim F\{n\}$ as $n\to\infty$ (i.e. with a long-tailed density with respect to counting measure) may be also viewed as a member of $\mathcal{L}_\Delta$ with $\Delta=(0,1]$.

In earlier chapters we dealt with tail functions of distributions (for any distribution F and for any x, the tail $\overline{F}(x) = F(x+\Delta)$ with $\Delta=(0,\infty)$). Tail functions are monotone non-increasing, and this allowed us to prove Theorem 2.11. For finite intervals Δ, there is in general no such monotonicity, and we need further restrictions, given by Theorem 4.18, to obtain the inequality

$$\liminf_{x\to\infty} \frac{(F*G)(x+\Delta)}{F(x+\Delta)+G(x+\Delta)} \ge 1. \tag{4.23}$$

Theorem 4.18. *Let the distributions F and G belong to the class $\mathcal{L}_\Delta$, where $\Delta=(0,T]$ for some finite T. Then*

$$\liminf_{x\to\infty} \frac{(F*G)(x+\Delta)}{F(x+\Delta)+G(x+\Delta)} = 1. \tag{4.24}$$

Proof. Let ξ and η be two independent random variables with respective distributions F and G. Fix any $a>0$. For $x>2a$, we have the following lower bound:

$$(F*G)(x+\Delta) \ge \mathbb{P}\{\xi+\eta\in x+\Delta,\ |\xi|\le a\} + \mathbb{P}\{\xi+\eta\in x+\Delta,\ |\eta|\le a\}.$$

We also have the tail equivalences $F(x+y+\Delta)\sim F(x+\Delta)$ and $G(x+y+\Delta)\sim G(x+\Delta)$ as $x\to\infty$ uniformly in $|y|\le a$. Therefore, as $x\to\infty$,

$$\begin{aligned} \mathbb{P}\{\xi+\eta\in x+\Delta,\ |\xi|\le a\} &= \int_{[-a,a]} G(x-y+\Delta)F(dy) \\ &\sim G(x+\Delta)\int_{[-a,a]} F(dy) \\ &\sim G(x+\Delta)F[-a,a], \end{aligned}$$

and similarly

$$\mathbb{P}\{\xi+\eta \in x+\Delta,\ |\eta|\le a\} \sim F(x+\Delta)G[-a,a].$$

Letting $a \to \infty$ implies the lower bound (4.23). Now assume that, on the contrary, the equality (4.24) does not hold, that is,

$$\liminf_{x\to\infty} \frac{(F*G)(x+\Delta)}{F(x+\Delta)+G(x+\Delta)} > 1.$$

Then there exist $\varepsilon > 0$ and x_0 such that, for all $x > x_0$ and $n \ge 0$,

$$(F*G)(x+nT+\Delta) \ge (1+\varepsilon)(F(x+nT+\Delta)+G(x+nT+\Delta)).$$

Summing over $n \ge 0$, we obtain

$$\overline{F*G}(x) \ge (1+\varepsilon)(\overline{F}(x)+\overline{G}(x)).$$

However, since $F \in \mathcal{L}_\Delta \subseteq \mathcal{L}$, it follows that the distribution F is heavy-tailed, and therefore the latter inequality contradicts Theorem 2.13. □

In the next theorem we prove that, for any Δ, the class $\mathcal{L}_\Delta$ is closed under convolutions.

Theorem 4.19. *Let the distributions F and G belong to the class $\mathcal{L}_\Delta$ for some finite interval $\Delta = (0,T]$. Then $F*G \in \mathcal{L}_\Delta$.*

Proof. Let ξ and η be two independent random variables with respective distributions F and G. By Lemma 2.19 and Proposition 2.20 there exists a function h such that $h(x) < x/2$, $h(x) \to \infty$ and both $F(x+\Delta)$ and $G(x+\Delta)$ are h-insensitive.

Consider the event $B(x,t) = \{\xi+\eta \in x+t+\Delta\}$. In order to prove that $F*G \in \mathcal{L}_\Delta$, we need to check that, for any $t > 0$, $\mathbb{P}\{B(x,t)\} \sim \mathbb{P}\{B(x,0)\}$ as $x \to \infty$. Since the events $\{\xi \le x-h(x)\}$ and $\{\eta \le h(x)\}$ together imply $\{\xi+\eta \le x\}$, we have the following decomposition:

$$\begin{aligned}\mathbb{P}\{B(x,t)\} &= \mathbb{P}\{B(x,t),\ \xi \le x-h(x)\}\\ &\quad + \mathbb{P}\{B(x,t),\ \eta \le h(x)\} + \mathbb{P}\{B(x,t),\ \xi > x-h(x),\ \eta > h(x)\}.\end{aligned} \tag{4.25}$$

For fixed $t > 0$, $G(x+t-y+\Delta) \sim G(x-y+\Delta)$ as $x \to \infty$ uniformly in $y \le x-h(x)$, since $h(x) \to \infty$. Therefore, as $x \to \infty$,

$$\begin{aligned}\mathbb{P}\{B(x,t),\ \xi \le x-h(x)\} &= \int_{-\infty}^{x-h(x)} G(x+t-y+\Delta)F(dy)\\ &\sim \int_{-\infty}^{x-h(x)} G(x-y+\Delta)F(dy)\\ &= \mathbb{P}\{B(x,0),\ \xi \le x-h(x)\}.\end{aligned} \tag{4.26}$$

A similar argument shows that

$$\begin{aligned}\mathbb{P}\{B(x,t),\ \eta \le h(x)\} &= \int_{-\infty}^{h(x)} F(x+t-y+\Delta)G(dy)\\ &\sim \int_{-\infty}^{h(x)} F(x-y+\Delta)G(dy)\\ &= \mathbb{P}\{B(x,0),\ \eta \le h(x)\}.\end{aligned} \tag{4.27}$$

Finally,

$$\begin{aligned}&\mathbb{P}\{B(x,t),\ \xi > x-h(x),\ \eta > h(x)\}\\&\quad= \mathbb{P}\{B(x,t),\ \xi \in (x-h(x), x-h(x)+t+T],\ \eta > h(x)\}.\end{aligned}$$

The value of the latter probability is at most

$$\overline{G}(h(x))F(x-h(x)+(0,t+T]) = o(F(x-h(x)+(0,t+T])) \quad \text{as } x\to\infty.$$

Without loss of generality, we can assume that $t < T$. Then,

$$\begin{aligned}F(x-h(x)+(0,t+T]) &\le F(x-h(x)+(0,T]) + F(x-h(x)+T+(0,T])\\&= o(F(x-h(x)+\Delta)) + o(F(x-h(x)+T+\Delta))\end{aligned}$$

as $x\to\infty$. Both terms on the right side of the above expression are of the order $o(F(x+\Delta))$, by the choice of the function h. Thus, as $x\to\infty$,

$$\mathbb{P}\{B(x,t),\ \xi > x-h(x),\ \eta > h(x)\} = o(F(x+\Delta)). \tag{4.28}$$

Combining (4.25)–(4.28) we conclude that

$$\mathbb{P}\{B(x,t)\} = (1+o(1))\mathbb{P}\{B(x,0)\} + o(F(x+\Delta))$$

as $x\to\infty$. The conclusion of the theorem now follows on applying Theorem 4.18. □

By induction arguments we obtain the following corollary to Theorems 4.18 and 4.19.

Corollary 4.20. *If $F \in \mathcal{L}_\Delta$, then, for all $n \ge 2$, $F^{*n} \in \mathcal{L}_\Delta$ and*

$$\liminf_{x\to\infty} \frac{F^{*n}(x+\Delta)}{F(x+\Delta)} \ge n.$$

4.5 Δ-Subexponential distributions

We continue our study of local properties by introducing the concept of Δ-subexponentiality.

As in the case of subexponential densities, there are two slightly different approaches to the definition of Δ-subexponentiality of a distribution on the whole real line $\mathbb{R}$, and these approaches coincide if the support of the distribution is restricted to the positive half-line $\mathbb{R}^+$. As earlier, the first definition—which is that we make below—seems in some respects more natural, but does not possess the tail property. However, for distributions on $\mathbb{R}^+$, Δ-subexponentiality *is* a tail property. The alternative definition is that a distribution on R should be Δ-subexponential (to avoid confusion we might say *conditionally* Δ-subexponential) if and only if the distribution conditioned on $\mathbb{R}^+$ is Δ-subexponential, and here of course the tail property is preserved. Here we only touch the distributions on the whole line R and provide only the first definition of Δ-subexponentiality; we then concentrate on distributions on $\mathbb{R}^+$, where the dilemma indicated above does not arise.

Definition 4.21. Let F be a distribution on $\mathbb{R}$ with right-unbounded support. For any fixed $\Delta = (0,T]$ for some finite $T>0$ we say that F is Δ-*subexponential* if $F\in\mathcal{L}_\Delta$ and

$$(F*F)(x+\Delta) \sim 2F(x+\Delta) \quad \text{as } x\to\infty.$$

Equivalently, a random variable ξ has a Δ-subexponential distribution if the function $\mathbb{P}\{\xi \in x+\Delta\}$ is long-tailed and, for two independent copies ξ_1 and ξ_2 of ξ,

$$\mathbb{P}\{\xi_1+\xi_2 \in x+\Delta\} \sim 2\mathbb{P}\{\xi \in x+\Delta\} \quad \text{as } x \to \infty.$$

In this and the following sections, we always consider finite intervals Δ. But if we allowed the interval to be infinite, $\Delta = (0,\infty)$, then the class of $(0,\infty)$-subexponential distributions would be none other than the standard class $\mathcal{S}_{\mathbb{R}}$ of subexponential distributions on the whole real line $\mathbb{R}$. For all finite Δ, the typical examples of Δ-subexponential distributions are the same—in particular the Pareto, lognormal, and Weibull (with parameter between 0 and 1) distributions, as we shall show in Section 4.6. Also, many properties of Δ-subexponential distributions with finite Δ are very close to those of subexponential distributions, as we shall show below. However, we have to repeat (see the previous section) that, for any distribution F, in contrast to the tail function $\overline{F}$, the function $F(x+\Delta)$ may be non-monotone. This leads to extra challenges in the study of Δ-subexponentiality (see the example on non-monotonicity at the end of Section 4.6).

Note that, for any $\Delta = (0,T)$, any distribution F with subexponential density f is Δ-subexponential since

$$\begin{aligned}(F*F)(x+\Delta) &= \int_x^{x+T} (f*f)(y)\mu(dy) \\ &\sim 2\int_x^{x+T} f(y)\mu(dy) = 2F(x+\Delta) \quad \text{as } x\to\infty.\end{aligned}$$

Further, it follows from the definition that, if F is Δ-subexponential, then F is $n\Delta$-subexponential for any $n = 2,3,\ldots$ and $F \in \mathcal{S}_{\mathbb{R}}$. To see this observe that, for any $n \in \{2,3,\ldots,\infty\}$ and as $x \to \infty$,

$$\begin{aligned}\mathbb{P}\{\xi_1+\xi_2 \in x+n\Delta\} &= \sum_{k=0}^{n-1} \mathbb{P}\{\xi_1+\xi_2 \in x+kT+\Delta\} \\ &\sim 2\sum_{k=0}^{n-1} \mathbb{P}\{\xi \in x+kT+\Delta\} \\ &= 2\mathbb{P}\{\xi \in x+n\Delta\}.\end{aligned}$$

Thus we have in particular that, for any Δ, the class of Δ-subexponential distributions is a subclass of $\mathcal{S}_{\mathbb{R}}$.

Note also that if we consider the distributions concentrated on the integers, then the class of $(0,1]$-subexponential distributions consists of all distributions F such that $F\{n+1\} \sim F\{n\}$ and $F^{*2}\{n\} \sim 2F\{n\}$ as $n \to \infty$, and so coincides with the class of distributions with subexponential densities.

We now have the following theorem which characterises Δ-subexponential distributions on the positive half-line $\mathbb{R}^+$ and which is analogous to Theorem 3.6 or Theorem 3.7 for subexponential distributions on $\mathbb{R}$ and to Theorem 4.7 for subexponential densities on $\mathbb{R}^+$. Then Theorem 4.24 shows that, as for subexponentiality of densities, Δ-subexponentiality is a tail property for distributions on $\mathbb{R}^+$. This latter result can also be deduced from Theorem 4.22.

Theorem 4.22. *Suppose that the distribution F on R_+ is such that $F \in \mathcal{L}_\Delta$ for some Δ. Let ξ_1 and ξ_2 be two independent random variables with common distribution F. Then the following assertions are equivalent:*

(i) *F is Δ-subexponential;*

(ii) *for every function h such that $h(x) \to \infty$ as $x \to \infty$,*

$$\mathbb{P}\{\xi_1 + \xi_2 \in x + \Delta, \xi_1 > h(x), \xi_2 > h(x)\} = o(F(x + \Delta)); \tag{4.29}$$

(iii) *there exists a function h such that $h(x) < x/2$, $h(x) \to \infty$ as $x \to \infty$, the function $F(x + \Delta)$ is h-insensitive and the relation (4.29) holds.*

Proof. (i)⇒(ii). Suppose first that F is Δ-subexponential. Define the event $B = \{\xi_1 + \xi_2 \in x + \Delta\}$. Note that if (4.29) is valid for some function h, then it is valid for any function h_1 such that $h_1 \geq h$. Hence we may assume without loss of generality that $h(x) < x/2$ for all x. Then

$$\begin{aligned}\mathbb{P}(B) &= \mathbb{P}\{B, \xi_1 \leq h(x)\} + \mathbb{P}\{B, \xi_2 \leq h(x)\} + \mathbb{P}\{B, \xi_1 > h(x),\ \xi_2 > h(x)\} \\ &= 2\mathbb{P}\{B, \xi_1 \leq h(x)\} + \mathbb{P}\{B, \xi_1 > h(x),\ \xi_2 > h(x)\}.\end{aligned} \tag{4.30}$$

By Fatou's lemma,

$$\liminf_{x\to\infty} \frac{\mathbb{P}\{B, \xi_1 \leq h(x)\}}{F(x+\Delta)} = \liminf_{x\to\infty} \int_0^{h(x)} \frac{F(x-y+\Delta)}{F(x+\Delta)} F(dy) \geq 1. \tag{4.31}$$

Hence from (4.30), (4.31) and the Definition 4.21 of Δ-subexponentiality, we obtain (4.29).

That (ii) implies (iii) is trivial since the condition $F \in \mathcal{L}_\Delta$ implies the existence of a function with respect to which $F(x + \Delta)$ is h-insensitive.

(iii)⇒(i). Now suppose that the condition (iii) holds for some function h. We again use the decomposition (4.30) for B as defined above. Then

$$\begin{aligned}\mathbb{P}\{B, \xi_1 \leq h(x)\} &= \int_0^{h(x)} F(x-y+\Delta) F(dy) \\ &\sim F(x+\Delta) \int_0^{h(x)} F(dy) \\ &\sim F(x+\Delta) \qquad \text{as } x \to \infty,\end{aligned}$$

and so (4.29) together with (4.30) implies the Δ-subexponentiality of F. □

Next we prove the result which shows in particular that the subclass of Δ-subexponential distributions on the positive half-line $\mathbb{R}^+$ is closed under the natural local tail-equivalence relation.

Theorem 4.23. *Let F be a Δ-subexponential distribution for some Δ. Suppose that the distribution G on $\mathbb{R}^+$ belongs to $\mathcal{L}_\Delta$ and the functions $F(x+\Delta)$ and $G(x+\Delta)$ are weakly tail-equivalent, that is,*

$$0 < \liminf_{x\to\infty} \frac{G(x+\Delta)}{F(x+\Delta)} \leq \limsup_{x\to\infty} \frac{G(x+\Delta)}{F(x+\Delta)} < \infty. \tag{4.32}$$

Then G is also Δ-subexponential. In particular, G is Δ-subexponential provided $G(x + \Delta) \sim cF(x + \Delta)$ as $x \to \infty$ for some $c > 0$.

Proof. Choose a function h such that $h(x) < x/2$ for all x, $h(x) \to \infty$ as $x \to \infty$ and the function $g(x + \Delta)$ is h-insensitive. Let ξ_1, ξ_2, ζ_1, ζ_2 be independent random variables such that ξ_1 and ξ_2 have common distribution F, and ζ_1 and ζ_2 have common distribution G. By Theorem 4.22, it is sufficient to prove that

$$\mathbb{P}\{\zeta_1 + \zeta_2 \in x + \Delta, \zeta_1 > h(x), \zeta_2 > h(x)\} = o(G(x + \Delta)).$$

The probability on the left side of the above expression is not greater than

$$\int_{h(x)}^{x-h(x)+T} G(x - y + \Delta)G(dy) =: I.$$

By the condition (4.32), for some $c_1 < \infty$ and for all sufficiently large x,

$$\begin{aligned} I &\le c_1 \int_{h(x)}^{x-h(x)+T} F(x - y + \Delta)G(dy) \\ &\le c_1 \mathbb{P}\{\zeta_1 + \xi_2 \in x + \Delta, \zeta_1 > h(x), \xi_2 > h(x) - T\} \\ &\le c_1 \int_{h(x)-T}^{x-h(x)+2T} G(x - y + \Delta)F(dy). \end{aligned}$$

A repetition of the above argument now gives that

$$\begin{aligned} I &\le c_1^2 \int_{h(x)-T}^{x-h(x)+2T} F(x - y + \Delta)F(dy) \\ &\le c_1^2 \mathbb{P}\{\xi_1 + \xi_2 \in x + \Delta, \xi_1 \ge h(x) - 2T, \xi_2 \ge h(x) - T\} \\ &= o(F(x + \Delta)) \\ &= o(G(x + \Delta)). \end{aligned}$$

as required, where the third line in the above display again follows from Theorem 4.22. □

Theorem 4.24 shows in particular that the convolution of tail-equivalent Δ-subexponential distributions on the positive half-line is also Δ-subexponential.

Theorem 4.24. *Suppose that the distribution F is Δ-subexponential, for some Δ. Let G_1, G_2 be two distributions on $\mathbb{R}^+$ such that $G_1(x+\Delta)/F(x+\Delta) \to c_1$ and $G_2(x+\Delta)/F(x+\Delta) \to c_2$ as $x \to \infty$, for some constants c_1, $c_2 \ge 0$. Then*

$$\frac{(G_1 * G_2)(x + \Delta)}{F(x + \Delta)} \to c_1 + c_2 \quad \text{as } x \to \infty. \tag{4.33}$$

Further, if $c_1 + c_2 > 0$ then the convolution $G_1 * G_2$ is Δ-subexponential.

Proof. Let ζ_1 and ζ_2 be independent random variables with distributions G_1 and G_2 respectively. Let h be a function such that $h(x) < x/2$ for all x, $h(x) \to \infty$ as $x \to \infty$, and the function $F(x + \Delta)$ is h-insensitive. Define also the event $B = \{\zeta_1 + \zeta_2 \in x + \Delta\}$. Then

$$\mathbb{P}\{B\} = \mathbb{P}\{B, \zeta_1 \le h(x)\} + \mathbb{P}\{B, \zeta_2 \le h(x)\} + \mathbb{P}\{B, \zeta_1 > h(x), \zeta_2 > h(x)\}. \tag{4.34}$$

As in the last lines of the proof of Theorem 4.22, one can show that

$$\mathbb{P}\{B, \zeta_1 \le h(x)) \sim G_2(x + \Delta)$$

as $x \to \infty$, and then

$$\frac{\mathbb{P}\{B, \zeta_1 \le h(x)\}}{F(x+\Delta)} \to c_2, \qquad \frac{\mathbb{P}\{B, \zeta_2 \le h(x)\}}{F(x+\Delta)} \to c_1. \tag{4.35}$$

Following the same argument as that in the proof of Theorem 4.23, we obtain also that

$$\mathbb{P}\{B, \zeta_1 > h(x), \zeta_2 > h(x)\} = o(F(x+\Delta)). \tag{4.36}$$

The result (4.33) now follows from (4.34)– (4.36).

The final assertion of the theorem follows from Theorem 4.23. □

By induction, Theorem 4.24 implies the following corollary.

Corollary 4.25. *Suppose that the distribution F is Δ-subexponential, for some Δ. Let G be a distribution on $\mathbb{R}^+$ such that $G(x+\Delta)/F(x+\Delta) \to c \ge 0$ as $x \to \infty$. Then, for any $n \ge 2$, $G^{*n}(x+\Delta)/F(x+\Delta) \to nc$ as $x \to \infty$. If $c > 0$, then G^{*n} is Δ-subexponential.*

We conclude this Section with the result which provides of Kesten's upper bound for the class of Δ-subexponential distributions.

Theorem 4.26. *Suppose that the distribution F is Δ-subexponential, for some $\Delta = (0, T]$. Then, for any $\varepsilon > 0$, there exist $x_0 = x_0(\varepsilon) > 0$ and $c(\varepsilon) > 0$ such that, for any $x > x_0$ and for any $n \ge 1$,*

$$F^{*n}(x+\Delta) \le c(\varepsilon)(1+\varepsilon)^n F(x+\Delta).$$

Proof. Let $\{\xi_n\}$ be a sequence of independent non-negative random variables with common distribution F. Put $S_n = \xi_1 + \ldots + \xi_n$. For $x_0 \ge 0$ and $k \ge 1$, put

$$A_n := A_n(x_0) = \sup_{x > x_0} \frac{F^{*n}(x+\Delta)}{F(x+\Delta)}.$$

Take any $\varepsilon > 0$. Appealing to Theorem 4.22, we conclude that x_0 may be chosen such that, for any $x > x_0$,

$$\mathbb{P}\{\xi_1 + \xi_2 \in x + \Delta, \xi_2 \le x - x_0\} \le (1 + \varepsilon/2) F(x+\Delta).$$

For any $n > 1$ and $x > x_0$,

$$\begin{aligned} \mathbb{P}\{S_n \in x+\Delta\} &= \mathbb{P}\{S_n \in x+\Delta, \xi_n \le x - x_0\} + \mathbb{P}\{S_n \in x+\Delta, \xi_n > x - x_0\} \\ &=: P_1(x) + P_2(x), \end{aligned}$$

where, by the definition of A_{n-1} and x_0,

$$\begin{aligned} P_1(x) &= \int_0^{x-x_0} \mathbb{P}\{S_{n-1} \in x - y + \Delta\} F(dy) \\ &\le A_{n-1} \int_0^{x-x_0} F(x-y+\Delta) F(dy) \\ &= A_{n-1} \mathbb{P}\{\xi_1 + \xi_n \in x+\Delta, \xi_n \le x - x_0\} \\ &\le A_{n-1}(1+\varepsilon/2) F(x+\Delta). \end{aligned} \tag{4.37}$$

Further,

$$
\begin{aligned}
P_2(x) &= \int_0^{x_0+T} \mathbb{P}\{\xi_n \in x - y + \Delta, \xi_n > x - x_0\}\mathbb{P}\{S_{n-1} \in dy\} \\
&\le \sup_{0<t\le x_0} F(x - t + \Delta).
\end{aligned}
$$

Thus, if $x > 2x_0$, then

$$P_2(x) \le L_1 F(x + \Delta),$$

where

$$L_1 = \sup_{0<t\le x_0,\ y>2x_0} \frac{F(y - t + \Delta)}{F(y + \Delta)}.$$

If $x_0 < x \le 2x_0$, then $P_2(x) \le 1$ implies

$$\frac{P_2(x)}{F(x + \Delta)} \le \frac{1}{\inf_{x_0<x\le 2x_0} F(x + \Delta)} =: L_2.$$

Since $F \in \mathcal{L}_\Delta$, both L_1 and L_2 are finite for x_0 sufficiently large. Put $L = \max(L_1, L_2)$. Then, for any $x > x_0$,

$$P_2(x) \le LF(x + \Delta). \tag{4.38}$$

It follows from (4.37) and (4.38) that $A_n \le A_{n-1}(1 + \varepsilon/2) + L$ for $n > 1$. Therefore, the induction argument yields

$$A_n \le A_1(1 + \varepsilon/2)^{n-1} + L\sum_{l=0}^{n-2}(1 + \varepsilon/2)^l \le Ln(1 + \varepsilon/2)^{n-1}.$$

This implies the conclusion of the theorem. □

4.6 Sufficient conditions for Δ-subexponentiality

In this Section, we give sufficient conditions for a distribution to be Δ-subexponential. There is much similarity between these conditions and the conditions given earlier for subexponentiality.

Theorem 4.27. *Let the distribution F on $\mathbb{R}^+$ belong to the class $\mathcal{L}_\Delta$ where $\Delta = (0, T]$ for some finite $T > 0$. Suppose that there exist $c > 0$ and $x_0 < \infty$ such that $F(x + t + \Delta) \ge cF(x + \Delta)$ for any $t \in (0, x]$ and $x > x_0$. Then F is Δ-subexponential.*

Proof. Let the function h be such that $h(x) < x/2$ for all x and $h(x) \to \infty$ as $x \to \infty$. Then

$$
\begin{aligned}
\mathbb{P}\{\xi_1 + \xi_2 \in x+\Delta, \xi_1 > h(x), \xi_2 > h(x)\} &\le 2\int_{h(x)}^{x/2+T} F(x - y + \Delta)F(dy) \\
&\le 2(c + o(1))F(x + \Delta)\int_{h(x)}^{x/2+T} F(dy) \\
&= o(F(x + \Delta))
\end{aligned}
$$

as $x \to \infty$. Applying now Theorem 4.22(ii) we conclude that F is Δ-subexponential. □

The Pareto distribution (with the tail $\overline{F}(x) = x^{-\alpha}$, $\alpha > 0$, $x \ge 1$) satisfies the conditions of Theorem 4.27. The same is true for any distribution F such that the function $F(x + \Delta)$ is regularly varying at infinity.

Theorem 4.28. *Suppose that the distribution F on $\mathbb{R}^+$ belongs to the class $\mathcal{L}_\Delta$ for some finite $\Delta = (0, T]$. Suppose also that for some x_0 the function $g(x) := -\ln F(x + \Delta)$ is concave for $x \ge x_0$. Suppose finally that there exists a function h such that $h(x) \to \infty$ as $x \to \infty$, $F(x + \Delta)$ is h-insensitive, and $xF(h(x) + \Delta) \to 0$ as $x \to \infty$. Then F is Δ-subexponential.*

Proof. By Theorem 4.23, we may assume without loss of generality that $x_0 = 0$. Since $g(x)$ is concave, the minimum of the sum $g(x - y) + g(y)$ on the interval $y \in [h(x), x - h(x)]$ is equal to $g(x - h(x)) + g(h(x))$.

Choose any $c > 1$ and x so large than

$$\sup_{t \ge h(x) - T} \sup_{u,v \in \Delta} F(t + v + \Delta)/F(t + u + \Delta) \le c.$$

Then, for any x sufficiently large and for any $x - h(x) > t > h(x)$,

$$\int_{t-T}^{t} F(x-y+\Delta)F(dy) \le c^2 F(x-t+\Delta)F(t+\Delta) \le c^4/T \int_{t-T}^{t} F(x-y+\Delta)F(y+\Delta)dy.$$

Let $c_1 = c^4/T$. We may assume for simplicity that $x - 2h(x)$ is a multiple of T and let $k = (x - 2h(x))/T$. Then

$$\begin{aligned}
\int_{h(x)}^{x-h(x)} F(x - y + \Delta)F(dy) &= \sum_{i=0}^{k-1} \int_{h(x)+iT}^{h(x)+(i+1)T} F(x - y + \Delta)F(dy) \\
&\le c_1 \int_{h(x)}^{x-h(x)} F(x - y + \Delta)F(y + \Delta)dy \\
&= c_1 \int_{h(x)}^{x-h(x)} e^{-(g(x-y)+g(y))}dy \\
&\le c_1 x e^{-(g(x-h(x))+g(h(x)))}.
\end{aligned}$$

Since $e^{-g(x-h(x))} \sim e^{-g(x)}$, it follows that

$$\int_{h(x)}^{x-h(x)} F(x - y + \Delta)F(dy) = O(e^{-g(x)} x e^{-g(h(x))}) = o(F(x + \Delta)). \tag{4.39}$$

We now have

$$\begin{aligned}
&\mathbb{P}\{\xi_1 + \xi_2 \in x + \Delta,\, \xi_1 > h(x),\, \xi_2 > h(x)\} \\
&= \int_{h(x)}^{x-h(x)} F(x - y + \Delta)F(dy) + \mathbb{P}\{\xi_1 > x - h(x),\, \xi_1 + \xi_2 \in x + \Delta,\, \xi_2 > h(x)\}.
\end{aligned}$$

The second term on the right side of the above expression is not bigger than

$$\mathbb{P}\{\xi_1 \in x - h(x) + \Delta\} \cdot \mathbb{P}\{\xi_2 > h(x)\} \sim F(x+\Delta)\overline{F}(h(x)) = o(F(x+\Delta)) \qquad \text{as } x \to \infty.$$

Hence, using also (4.39), we have

$$\mathbb{P}\{\xi_1 + \xi_2 \in x + \Delta,\, \xi_1 > h(x),\, \xi_2 > h(x)\} = o(F(x+\Delta)) \qquad \text{as } x \to \infty,$$

and so the Δ-subexponentiality of F follows from Theorem 4.22. □

To show the applicability of the latter theorem, we consider two examples. First, we consider the Weibull distribution F on the positive half-line $\mathbb{R}^+$ with tail function given by $\overline{F}(x) = e^{-x^\beta}$, $x \geq 0$, $\beta \in (0,1)$, and let $\Delta = (0, T]$, for some finite T. Then it can be deduced that

$$F(x+\Delta) \sim \beta T x^{\beta-1} \exp(-x^\beta) \quad \text{as } x \to \infty.$$

From this, we want to show that $F(x+\Delta)$ is asymptotically equivalent to a function which satisfied the condition of Theorem 4.28. Indeed, consider the distribution $\widehat{F}$ with the tail function given by $\overline{\widehat{F}}(x) = \min(1, x^{\beta-1}e^{-x^\beta})$. Let x_0 be the unique positive solution to the equation $x^{1-\beta} = e^{-x^\beta}$. Then the function $\widehat{g}(x) = -\ln \widehat{F}(x+\Delta)$ is concave for $x \geq x_0$, and the conditions of Theorem 4.28 are satisfied with $h(x) = x^\gamma$, $\gamma \in (0, 1-\beta)$. Therefore, $\widehat{F}$ is Δ-subexponential and, by Theorem 4.23, F is also Δ-subexponential.

Second, we consider the lognormal distribution F with the density f given by $f(x) = e^{-(\ln x - \ln a)^2/2\sigma^2}/x\sqrt{2\pi\sigma^2}$ and note that the function $g(x) = -\ln(x^{-1}e^{-(\ln x - \ln a)^2/2\sigma^2}) = \ln x + (\ln x - \ln a)^2/2\sigma^2$ is eventually concave. Since, for any fixed $\Delta = (0, T]$,

$$F(x+\Delta) \sim Tf(x)$$

as $x \to \infty$, the conditions of Theorem 4.28 are satisfied for any function h such that $h(x) = o(x)$. Thus, F is Δ-subexponential.

We now show by two examples that the classes of Δ-subexponential distributions differ for different Δ and also the complexity of the relations between these classes. The first example deals with lattice distributions. Let the random variable ξ be positive and integer-valued, with $\mathbb{P}\{\xi = 2k\} = \gamma/k^2$ and $\mathbb{P}\{\xi = 2k+1\} = \gamma/2^k$, where γ is the appropriate normalizing constant. Then ξ has a lattice distribution F with span 1. By Theorem 4.27, F is $(0,2]$-subexponential. But it cannot be $(0,a]$-subexponential if a is not an even integer or infinity.

In the second example, we consider absolutely continuous distributions. Assume that ξ is the sum of two independent random variables: $\xi = \eta + \zeta$ where η is distributed uniformly on $(-1/8, 1/8)$ and $\mathbb{P}\{\zeta = k\} = \gamma/k^2$ for $k = 1, 2, \ldots$ where γ is the appropriate normalising constant. Then the distribution F of ξ is absolutely continuous. It may be verified that F is $(0,1]$-subexponential, but cannot be $(0,a]$-subexponential if a is not an integer or infinity.

Finally, recall that in the previous section we undertook to provide an example of a Δ-subexponential distribution F where the function $F(x+\Delta)$ is not asymptotically equivalent to any non-decreasing function. Consider first a long-tailed function f such that $f(x) \in [1/x^2, 2/x^2]$ for all $x > 0$. Choose the function f in such a way that f is not asymptotically equivalent to a non-increasing function. For instance, one can define f as follows. Consider the increasing sequence $x_n = 2^{n/4}$. Put $f(x_{2n}) = 1/x_{2n}^2$ and $f(x_{2n+1}) = 2/x_{2n+1}^2$. Then assume that f is linear between any two consecutive members of the above sequence. Consider now the lattice distribution F on the set of natural numbers with $F\{n\} = f(n)$ for all sufficiently large integers n. Then by Theorem 4.23, F is Δ-subexponential, but $f(n) = F(n-1, n]$ is not asymptotically equivalent to a non-increasing function.

4.7 Local asymptotics for a randomly stopped sum

In this section, we give local analogues, both for subexponential densities and for Δ-subexponential distributions, of results which were given in Section 3.11 for subexponential distributions. We show that a random sum preserves a local subexponential property of independent identically distributed summands provided that the counting variable has a light-tailed distribution. We also establish the corresponding characteristic properties.

As for the results obtained in Section 3.11, the results from this section are needed in a variety of models in which random sums may appear, including random walks, branching processes, and infinitely divisible laws.

We again consider a sequence $\xi, \xi_1, \xi_2, \dots$ of independent random variables with a common distribution F on $\mathbb{R}^+$ and their partial sums $S_0 = 0$, $S_n = \xi_1 + \dots + \xi_n$ for each $n \geq 1$, together a counting random variable τ which is independent of the sequence $\{\xi_n\}$ and takes values in $\mathbb{Z}^+$.

4.7.1 Density of a randomly stopped sum

Let μ be either Lebesgue measure on $\mathbb{R}$ or counting measure on $\mathbb{Z}$. Throughout this section, the argument x of the density function f is either a real number if μ is Lebesgue measure; or an integer if μ is counting measure.

Theorem 4.29. *Let $\{p_n\}_{n\geq 1}$ be a non-negative sequence such that $\sum_{n\geq 1} p_n = 1$ and $m_p := \sum_{n\geq 1} np_n$ is finite. Let the distribution F on $\mathbb{R}^+$ have a long-tailed density f with respect to μ. Define the density g on $\mathbb{R}^+$ by*

$$g(x) = \sum_{n\geq 1} p_n f^{*n}(x).$$

(i) *If the density f is subexponential and bounded, and if*

$$\sum_{n\geq 1} (1+\delta)^n p_n < \infty$$

for some $\delta > 0$, then

$$g(x) \sim m_p f(x) \quad \text{as } x \to \infty. \tag{4.40}$$

(ii) *If the relation (4.40) holds and $p_1 < 1$, then the density f is subexponential.*

Proof. The result (i) is immediate from Corollary 4.13, Theorem 4.14, and the dominated convergence theorem. We prove the second result. By Corollary 4.5, for any $k \geq 2$,

$$\liminf_{x\to\infty} f^{*k}(x)/f(x) \geq k.$$

If $p_1 < 1$ then $p_n > 0$ for some $n \geq 2$, and so, arguing as in the proof of Theorem 3.38, it follows from the above bound and from (4.40) that

$$\limsup_{x\to\infty} \frac{f^{*n}(x)}{f(x)} \leq n. \tag{4.41}$$

By Corollary 4.5, $f^{*(n-1)}$ is long-tailed and so, from (4.41) and Theorem 4.2,

$$n \geq \limsup_{x\to\infty} \frac{f^{*n}(x)}{f(x)} = \limsup_{x\to\infty} \frac{(f * f^{*(n-1)})(x)}{f(x)} \geq 1 + \limsup_{x\to\infty} \frac{f^{*(n-1)}(x)}{f(x)}.$$

It follows by induction from the above bound that

$$\limsup_{x\to\infty} \frac{f^{*2}(x)}{f(x)} \leq 2.$$

Again by Theorem 4.2, this implies that $\lim_{x\to\infty} f^{*2}(x)/f(x) = 2$, which implies the subexponentiality of the density f. □

4.7.2 Δ-Subexponential distributions and random sums

Analogously to Theorem 4.29, we have the following result.

Theorem 4.30. *Let $\Delta = (0, T]$ for some finite $T > 0$. Suppose that the distribution F on $\mathbb{R}^+$ is Δ-long-tailed ($F \in \mathcal{L}_\Delta$), and that the random variable τ (introduced at the start of Section 4.7) is such that $\mathbb{E}\tau < \infty$.*

(i) *If F is a Δ-subexponential distribution and if $\mathbb{E}(1+\delta)^\tau < \infty$ for some $\delta > 0$, then*

$$\frac{\mathbb{P}\{S_\tau \in x + \Delta\}}{F(x+\Delta)} \to \mathbb{E}\tau \quad \text{as } x \to \infty. \tag{4.42}$$

(ii) *If $\mathbb{P}\{\tau > 1\} > 0$ and further the relation (4.42) holds, then the distribution F is Δ-subexponential.*

Proof. The proof of (i) follows from Corollary 4.25, Theorem 4.26, and the dominated convergence theorem. We prove (ii). Since $F \in \mathcal{L}_\Delta$, it follows from Corollary 4.20 that, for any $k \geq 2$,

$$\liminf_{x\to\infty} \frac{F^{*k}(x+\Delta)}{F(x+\Delta)} \geq k. \tag{4.43}$$

If $\mathbb{P}\{\tau = n\} > 0$ for some $n \geq 2$, then, again arguing as in the proof of Theorem 3.38, it follows from the above bound and from (4.42) that

$$\limsup_{x\to\infty} \frac{F^{*n}(x+\Delta)}{F(x+\Delta)} \leq n. \tag{4.44}$$

Since $F \in \mathcal{L}_\Delta$, by Corollary 4.20 the convolution $F^{*(n-1)}$ also belongs to the class $\mathcal{L}_\Delta$. Hence, by (4.44) and Theorem 4.18,

$$\begin{aligned} n &\geq \limsup_{x\to\infty} \frac{F^{*n}(x+\Delta)}{F(x+\Delta)} \\ &= \limsup_{x\to\infty} \frac{(F * F^{*(n-1)})(x+\Delta)}{G(x+\Delta)} \\ &\geq 1 + \limsup_{x\to\infty} \frac{F^{*(n-1)}(x+\Delta)}{F(x+\Delta)}. \end{aligned}$$

It follows by induction from the above bound that

$$\limsup_{x\to\infty} \frac{F^{*2}(x+\Delta)}{F(x+\Delta)} \le 2.$$

Again by Theorem 4.18, this implies that $\lim_{x\to\infty} F^{*2}(x+\Delta)/F(x+\Delta) = 2$, which implies the Δ-subexponentiality of the distribution F. □

4.8 Comments

Local theorems for some classes of lattice distributions are given by Chover, Ney and Wainger in [13, Section 2]. Densities are considered in [13, Section 2] (requiring continuity) and by Klüppelberg in [30] who considered asymptotics of densities for a special case (see also Sgibnev [41] for some results on densities on $\mathbb{R}$).

Much of the material of this chapter is adapted from the paper by Asmussen, Foss and Korshunov [4].

Chapter 5

Maxima of random walks

In this chapter, we study a random walk whose increments have a heavy-tailed distribution with a negative mean. The maximum of such a random walk is almost surely finite and our interest is in the tail asymptotics of its distribution. We use direct probabilistic techniques and show that again, under the appropriate subexponentiality condition, the main cause for the maximum to be very large is that a single one of the increments is similarly large. We start with Section 5.1 where a number of basic auxiliary results are collected.

5.1 Approximations of sums by integrals for long-tailed functions and distributions

For any function f which is integrable at infinity, we define the function f_I by

$$f_I(x) := \int_x^\infty f(y)dy.$$

Lemma 5.1. *Suppose that the nonnegative function f is integrable at infinity. Suppose further that either (a) f is non-increasing and f_I is long-tailed, or (b) f is long-tailed. Then, for any $a > 0$,*

$$\sum_{n=0}^{\infty} f(x+na) \sim \frac{1}{a} f_I(x) \quad \text{as } x \to \infty.$$

Proof. Suppose first that f is non-increasing and f_I is long-tailed. Then, for all $n \geq 0$ and for all x,

$$\frac{1}{a}\int_{x+na}^{x+(n+1)a} f(x)dx \leq f(x+na) \leq \frac{1}{a}\int_{x+(n-1)a}^{x+na} f(x)dx. \tag{5.1}$$

Summing over n, we obtain

$$\frac{1}{a} f_I(x) \leq \sum_{n=0}^{\infty} f(x+na) \leq \frac{1}{a} f_I(x-a), \tag{5.2}$$

which implies the result since f_I is long-tailed.

Now suppose that f is long-tailed. Then, given $\varepsilon > 0$ we have, for all $n \geq 0$ and for all sufficiently large x,

$$\frac{1-\varepsilon}{a}\int_{x+na}^{x+(n+1)a} f(x)dx \leq f(x+na) \leq \frac{1+\varepsilon}{a}\int_{x+na}^{x+(n+1)a} f(x)dx.$$

Hence, for all sufficiently large x,

$$\frac{1-\varepsilon}{a} f_I(x) \leq \sum_{n=0}^{\infty} f(x+na) \leq \frac{1+\varepsilon}{a} f_I(x).$$

and the result now follows on letting $\varepsilon \to 0$. □

The following result is of general use and, in particular, will be needed in the next section.

Lemma 5.2. *Suppose that the random variable ξ has a distribution F such that the integrated tail distribution F_I (see Section 2.6) is long-tailed. Let $\{\tau_n\}_{n\geq 1}$ be a sequence of independent identically distributed nonnegative random variables with finite mean $a > 0$ and suppose that this sequence is independent of ξ. Define also $T_0 = 0$, $T_n = \sum_{i=1}^n \tau_i$ for $n \geq 1$. Then*

$$\sum_{n=0}^{\infty} \mathbb{P}\{\xi > x + T_n\} \sim \frac{1}{a}\overline{F}_I(x) \quad \text{as } x \to \infty. \tag{5.3}$$

Proof. We prove first the lower bound. By the Law of Large Numbers, given $\varepsilon > 0$, we can choose L sufficiently large that

$$\mathbb{P}\{T_n \leq L + n(a+\varepsilon)\} > 1-\varepsilon, \quad n = 0, 1, 2, \ldots.$$

Then by Lemma 5.1 applied to the tail function $\overline{F}$, and the assumed independence of ξ and $\{\tau_n\}_{n\geq 1}$, for any x,

$$\begin{aligned}\sum_{n=0}^{\infty} \mathbb{P}\{\xi > x + T_n\} &> (1-\varepsilon)\sum_{n=0}^{\infty} \mathbb{P}\{\xi > x + L + n(a+\varepsilon)\} \\ &\sim \frac{1-\varepsilon}{a+\varepsilon}\overline{F}_I(x) \quad \text{as } x \to \infty.\end{aligned}$$

Now let $\varepsilon \to 0$ to obtain that

$$\liminf_{x\to\infty} \frac{\sum_{n=0}^{\infty} \mathbb{P}\{\xi > x + T_n\}}{\overline{F}_I(x)} \geq \frac{1}{a}.$$

For the upper bound, observe that, again by Lemma 5.1, the assumed independence of ξ and $\{\tau_n\}_{n\geq 1}$, and the positivity of the random variables T_n, given $\varepsilon \in (0, a)$, for any x,

$$\begin{aligned}\sum_{n=0}^{\infty} \mathbb{P}\{\xi > x + T_n\} &\leq \sum_{n=0}^{\infty} \mathbb{P}\{\xi > x + n(a-\varepsilon)\} + \mathbb{P}\{\xi > x\}\sum_{n=0}^{\infty} \mathbb{P}\{T_n < n(a-\varepsilon)\} \\ &\sim \frac{1}{a-\varepsilon}\overline{F}_I(x) + \overline{F}(x)\sum_{n=0}^{\infty} \mathbb{P}\{T_n < n(a-\varepsilon)\} \quad \text{as } x \to \infty. \end{aligned} \tag{5.4}$$

Since $a - \tau_1 - \varepsilon$ is bounded from above by $a - \varepsilon$ and has mean $-\varepsilon$, there exists some $\lambda > 0$ such that

$$\alpha := \mathbb{E}e^{\lambda(a-\tau_1-\varepsilon)} < 1,$$

and so

$$\begin{aligned}\mathbb{P}\{T_n < n(a-\varepsilon)\} &= \mathbb{P}\{e^{-\lambda(T_n - na)} > e^{\lambda n\varepsilon}\} \\ &\leq \mathbb{E}e^{-\lambda(T_n-na)}e^{-\lambda n\varepsilon} \\ &= \alpha^n.\end{aligned}$$

It follows that

$$\sum_{n=0}^{\infty} \mathbb{P}\{T_n < n(a-\varepsilon)\} < \infty$$

and so it follows from (5.4) and Lemma 2.25 that

$$\limsup_{x\to\infty} \frac{\sum_{n=0}^{\infty} \mathbb{P}\{\xi > x + T_n\}}{\overline{F}_I(x)} \leq \frac{1}{a-\varepsilon}.$$

Now let $\varepsilon \to 0$ to obtain the upper bound. □

5.2 Asymptotics for the supremum of a random walk with a negative drift

We give an elementary probabilistic proof of the asymptotic behaviour of the distribution of the maximum of a random walk with negative drift and heavy-tailed increments (see Theorem 5.4 below). The underlying intuition of the result is that the only significant way in which a high value of the partial maximum can be attained is through "one big jump" by the random walk away from its mean path. We give here a relatively short proof from first principles which captures this intuition. It is similar in spirit to the probabilistic proof related to the ladder heights (which may also be of use for deriving local asymptotics), but by considering instead a first renewal time at which the random walk exceeds a "tilted" level, the argument becomes more elementary. In particular subsequent renewals have an asymptotically negligible probability under appropriate limits, and results from renewal theory—notably the derivation and use of the Pollaczeck-Khinchine formula—are not required.

We proceed with the proof by deriving separately the lower and the upper bounds, since no restrictions (apart of the negativeness of the mean!) are required for the former to hold while subexponentiality is needed for the latter.

Let $\xi_1, \xi_2, \ldots$ be independent identically distributed random variables with distribution function F such that $\mathbb{E}\xi_1 = -a < 0$. Let $S_0 = 0$, $S_n = \xi_1 + \ldots + \xi_n$ for $n \geq 1$. Let $M_n = \max(S_i, 0 \leq i \leq n)$ for $n \geq 0$ and let $M = \sup(S_n, n \geq 0)$. By the Strong Law of Large Numbers, $\mathbb{P}\{M < \infty\} = 1$.

We start with the lower bound which is proved by a quite elementary equilibrium identity.

Theorem 5.3. *Suppose that $\mathbb{E}\xi_1 = -a < 0$. Then, for any $x \geq 0$,*

$$\mathbb{P}\{M > x\} \geq \frac{\int_x^\infty \overline{F}(y)dy}{a + \int_x^\infty \overline{F}(y)dy},$$

and, in particular,

$$\liminf_{x\to\infty} \frac{\mathbb{P}\{M > x\}}{\overline{F}_I(x)} \geq \frac{1}{a}.$$

Proof. Let ξ be a random variable with distribution F which is independent of M. Then M has the same distribution as $(M+\xi)^+ := \max(0, M+\xi))$. Now fix $x \geq 0$. For $z > 0$ consider the function

$$L_z(y) = \begin{cases} x & \text{if } y \leq x, \\ y & \text{if } y \in (x, x+z], \\ x+z & \text{if } y > x+z. \end{cases}$$

Since this function is bounded, $\mathbb{E}L_z(M)$ is finite and $\mathbb{E}L_z(M) = \mathbb{E}L_z(M+\xi)$. Therefore,

$$\mathbb{E}(L_z(M+\xi) - L_z(M)) = 0.$$

We have $|L_z(M+\xi) - L_z(M)| \leq |\xi|$ for all z and $L_z(M+\xi) - L_z(M) \to L(M+\xi) - L(M)$ as $z \to \infty$ where

$$L(y) = \begin{cases} x & \text{if } y \leq x, \\ y & \text{if } y > x. \end{cases}$$

Hence, by the dominated convergence we obtain the equality

$$\mathbb{E}(L(M+\xi) - L(M)) = 0. \tag{5.5}$$

We make use of the following bounds. For $y \in [0, x]$,

$$L(y+\xi) - L(y) = (y+\xi-x)\mathbb{I}\{y+\xi > x\} \geq (\xi - x)\mathbb{I}\{\xi > x\},$$

and so

$$\mathbb{E}\{L(M+\xi) - L(M); M \leq x\} \geq \mathbb{E}\{\xi - x; \xi > x\}\mathbb{P}\{M \leq x\}. \tag{5.6}$$

For $y > x$,

$$L(y+\xi) - L(y) \geq \xi,$$

and so

$$\mathbb{E}\{L(M+\xi) - L(M); M > x\} \geq \mathbb{E}\xi\mathbb{P}\{M > x\}. \tag{5.7}$$

Substituting (5.6) and (5.7) into (5.5) we get the inequality

$$\mathbb{E}\{\xi - x; \xi > x\}\mathbb{P}\{M \leq x\} \leq -\mathbb{E}\xi\mathbb{P}\{M > x\}.$$

Therefore,

$$\mathbb{P}\{M > x\} \geq \frac{\mathbb{E}\{\xi - x; \xi > x\}}{a + \mathbb{E}\{\xi - x; \xi > x\}} = \frac{\int_x^\infty \overline{F}(y)dy}{a + \int_x^\infty \overline{F}(y)dy}.$$

where the final equality follows from (2.23). □

We now give our main result of this section, for the asymptotic behaviour of the tail of M.

Theorem 5.4. *Suppose that, in addition to the condition* $\mathbb{E}\xi_1 = -a < 0$, *the integrated tail distribution* F_I *is subexponential. Then*

$$\mathbb{P}\{M > x\} \sim a^{-1}\overline{F}_I(x) \quad \text{as } x \to \infty.$$

Proof. By Theorem 5.3, it is sufficient to establish the upper bound associated with the required asymptotics. Given $\varepsilon > 0$ and some (eventually large) $A > a$, define renewal times $0 =: \tau_0 < \tau_1 \le \tau_2 \le \dots$ for the process $\{S_n\}$ by

$$\tau_1 = \min\{j \ge 1 : S_j > A - j(a - \varepsilon)\} \le \infty$$

(here we make the standard convention $\min \varnothing = \infty$), and, for $k \ge 2$,

$$\begin{aligned} \tau_k &= \infty, \quad \text{if } \tau_{k-1} = \infty, \\ \tau_k &= \tau_{k-1} + \min\{j \ge 1 : S_{\tau_{k-1}+j} - S_{\tau_{k-1}} > A - j(a-\varepsilon)\}, \quad \text{if } \tau_{k-1} < \infty. \end{aligned}$$

Observe that, for any k, the joint distribution of the vectors

$$(\tau_1,\, S_{\tau_1}), (\tau_2 - \tau_1,\, S_{\tau_2} - S_{\tau_1}) \dots, (\tau_k - \tau_{k-1},\, S_{\tau_k} - S_{\tau_{k-1}}), \tag{5.8}$$

given $\tau_k < \infty$, is that of independent identically distributed vectors. Since $\mathbb{E}\xi_1 < 0$, by the Strong Law of Large Numbers,

$$\gamma := \mathbb{P}\{\tau_1 < \infty\} \to 0 \quad \text{as } A \to \infty. \tag{5.9}$$

Define also $S_\infty = -\infty$. Since $\tau_1 = n$ implies $S_{n-1} \le A - (n-1)(a - \varepsilon)$, we now have that, for all sufficiently large x,

$$\begin{aligned} \mathbb{P}\{S_{\tau_1} > x\} &= \sum_{n=1}^{\infty} \mathbb{P}\{\tau_1 = n, S_n > x\} \\ &\le \sum_{n=1}^{\infty} \mathbb{P}\{S_{n-1} \le A - (n-1)(a-\varepsilon), S_n > x\} \\ &\le \sum_{n=1}^{\infty} \mathbb{P}\{\xi_n > x - A + (n-1)(a-\varepsilon)\}. \end{aligned}$$

Therefore, again for all sufficiently large x,

$$\mathbb{P}\{S_{\tau_1} > x\} \le \sum_{n=0}^{\infty} \overline{F}(x - A + n(a-\varepsilon)) \le \frac{1}{a-\varepsilon}\overline{F}_I(x - A - a + \varepsilon), \tag{5.10}$$

where the second inequality above follows by the same argument as that leading to the right hand side of (5.2).

Let $\varphi_1, \varphi_2, \dots$ be independent identically distributed random variables such that

$$\mathbb{P}\{\varphi_1 > x\} = \mathbb{P}\{S_{\tau_1} > x \,|\, \tau_1 < \infty\}, \quad x \in \mathbb{R}.$$

Then, from (5.10) and since F_I is long-tailed,

$$\mathbb{P}\{\varphi_1 > x\} \le \overline{G}(x), \quad x \in \mathbb{R}, \tag{5.11}$$

for some distribution function G on $\mathbb{R}$ satisfying

$$\lim_{x\to\infty} \frac{\overline{G}(x)}{\overline{F}_I(x)} = \frac{1}{\gamma(a-\varepsilon)}. \tag{5.12}$$

It follows from the subexponentiality of F_I and Corollary 3.13 that the distribution G is subexponential. Thus, by applying Theorem 3.37 with a geometrically distributed independent stopping time, we have

$$(1-\gamma)\sum_{k=0}^{\infty} \gamma^k \overline{G^{*k}}(x) \sim \frac{\gamma}{1-\gamma}\overline{G}(x) \quad \text{as } x \to \infty.$$

From the stochastic majorisation (5.11) and the relation (5.12), we now get the following asymptotic upper bound:

$$\begin{aligned}\sum_{k=1}^{\infty} \gamma^k \mathbb{P}\{\varphi_1 + \ldots + \varphi_k > x\} &\le \frac{\gamma + o(1)}{(1-\gamma)^2}\overline{G}(x) \\ &\le \frac{1+o(1)}{(1-\gamma)^2(a-\varepsilon)}\overline{F}_I(x) \quad \text{as } x \to \infty. \end{aligned} \tag{5.13}$$

If $M > x$ then there exist τ_k and $j \in [\tau_k, \tau_{k+1})$ such that $S_j > x$. Then necessarily $S_{\tau_k} > x - A + a - \varepsilon$. (To see this assume that, on the contrary, $S_{\tau_k} \le x - A + a - \varepsilon < x$. In this case $\tau_k < j < \tau_{k+1}$ and $S_j - S_{\tau_k} > x - (x - A + a - \varepsilon) = A - a + \varepsilon$. Hence we have the contradiction that $\tau_{k+1} \le j$.) It follows that

$$\{M > x\} \subseteq \bigcup_{k=1}^{\infty} \{S_{\tau_k} > x - A + a - \varepsilon\}.$$

We now have (again for sufficiently large x) that

$$\begin{aligned}\mathbb{P}\{M > x\} &\le \sum_{k=1}^{\infty} \mathbb{P}\{S_{\tau_k} > x - A + a - \varepsilon\} \\ &\le \sum_{k=1}^{\infty} \gamma^k \mathbb{P}\{\varphi_1 + \ldots + \varphi_k > x - A + a - \varepsilon\},\end{aligned}$$

by (5.8) and by the construction of the random variables φ_i. Using also (5.13), we now have

$$\limsup_{x\to\infty} \frac{\mathbb{P}\{M > x\}}{\overline{F}_I(x)} \le \frac{1}{(a-\varepsilon)(1-\gamma)^2}.$$

Now let $A \to \infty$, so that $\gamma \to 0$ by (5.9), and then let $\varepsilon \to 0$ to obtain the required upper bound

$$\limsup_{x\to\infty} \frac{\mathbb{P}\{M > x\}}{\overline{F}_I(x)} \le \frac{1}{a}.$$

□

5.3 Finite horizon asymptotics

We continue to study the random walk with negative drift introduced in the previous section. Recall that $M_n = \max(S_i, 0 \le i \le n)$ is defined to be the maximum of the random walk to time n. In this section we derive asymptotics for the probability $\mathbb{P}\{M_n > x\}$ as $x \to \infty$ under heavy-tailedness assumptions. These asymptotics are uniform in n. The underlying intuition of the result is again that the only significant way in which a high value of the partial maximum can be attained is via a "big jump" of one of its increments. The proof of the lower bound is based on direct computations and requires the extra assumption of long-tailedness of the distribution F of the increments ξ_i. The proof of the upper bound is similar to that of Theorem 5.4, although the condition of that theorem that F_I be subexponential requires to be strengthened slightly to $F \in \mathcal{S}^*$ (see Section 3.4 and in particular Theorem 3.27).

Theorem 5.5.

(i) *Suppose that, in addition to the condition* $\mathbb{E}\xi = -a < 0$*, the distribution* F *is long-tailed (*$F \in \mathcal{L}$*). Then*

$$\mathbb{P}\{M_n > x\} \ge \frac{1+o(1)}{a} \int_x^{x+na} \overline{F}(y)dy \quad \text{as } x \to \infty, \text{ uniformly in } n \ge 1. \tag{5.14}$$

(ii) *Suppose that, in addition to the condition* $\mathbb{E}\xi_1 = -a < 0$*, the distribution* $F \in \mathcal{S}^*$*. Then*

$$\mathbb{P}\{M_n > x\} \sim \frac{1}{a} \int_x^{x+na} \overline{F}(y)dy \quad \text{as } x \to \infty, \text{ uniformly in } n \ge 1. \tag{5.15}$$

Proof. We prove first the lower bound given in (5.14). Since $\mathbb{E}\xi_1 < 0$, it follows from the Weak Law of Large Numbers that, given $\varepsilon > 0$ and $\delta > 0$, we can choose A sufficiently large that

$$\mathbb{P}\{S_n > -A - n(a+\varepsilon)\} \ge 1 - \delta \quad \text{for all } n \ge 0. \tag{5.16}$$

Then the following lower bound is immediate:

$$\begin{aligned}
\mathbb{P}\{M_n > x\} &= \sum_{k=0}^{n-1} \mathbb{P}\{M_k \le x, S_{k+1} > x\} \\
&\ge \sum_{k=0}^{n-1} \mathbb{P}\{M_k \le x, S_k > -A - k(a+\varepsilon), \xi_{k+1} > x + A + k(a+\varepsilon)\}.
\end{aligned}$$

By the independence of random variables ξ_i and by (5.16), we have

$$\begin{aligned}
\mathbb{P}\{M_n > x\} &\ge \sum_{k=0}^{n-1} \mathbb{P}\{M_k \le x, S_k > -A - k(a+\varepsilon)\}\mathbb{P}\{\xi_{k+1} > x + L + k(a+\varepsilon)\} \\
&\ge \sum_{k=0}^{n-1} (1-2\delta)\overline{F}(x + A + k(a+\varepsilon)),
\end{aligned}$$

where the last inequality holds for all x sufficiently large that

$$\mathbb{P}\{M > x\} \le \delta \tag{5.17}$$

(which implies that $\mathbb{P}\{M_k > x\} \le \delta$ for all k). By applying the first inequality in (5.1) to each of the terms $\overline{F}(x + a + k(a + \varepsilon))$ above and then taking their sum, we get

$$\mathbb{P}\{M_n > x\} \ge \frac{1 - 2\delta}{a + \varepsilon} \int_x^{x+na} \overline{F}(y + A)dy.$$

Since F is assumed to be long-tailed, it now follows that

$$\mathbb{P}\{M_n > x\} \ge \frac{1 - 3\delta}{a + \varepsilon} \int_x^{x+na} \overline{F}(y)dy$$

for all x sufficiently large that (5.17) holds. That the inequality (5.14) holds with the required uniformity in n now follows by letting $\delta, \varepsilon \to 0$.

We now prove (5.15). Here F is assumed to belong to the class $\mathcal{S}^*$, so it is in particular long-tailed. Hence, it is sufficient to establish the upper bound in (5.15). Given $\varepsilon > 0$ and $A > a$, define renewal times $0 =: \tau_0 < \tau_1 \le \tau_2 \le \ldots$ for the process $\{S_k\}$ as in the proof of Theorem 5.4.

Analogously to (5.10), we obtain that

$$\mathbb{P}\{S_{\tau_1 \wedge n} > x\} \le \sum_{k=0}^{n-1} \overline{F}(x - A + k(a - \varepsilon)) \le \frac{1}{a - \varepsilon} \int_x^{x+na} \overline{F}(y - A - a + \varepsilon)dy.$$

Since F is long-tailed,

$$\mathbb{P}\{S_{\tau_1 \wedge n} > x\} \le \frac{1 + \varepsilon}{a - \varepsilon} \int_x^{x+na} \overline{F}(y)dy \tag{5.18}$$

for all sufficiently large x uniformly in $n \ge 1$. This means that we can choose x_0 such that (5.18) holds for all $x \ge x_0$ and for all $n = 1, 2, \ldots$.

Let $\varphi_{n,1}, \varphi_{n,2}, \ldots$ be independent identically distributed random variables such that

$$\mathbb{P}\{\varphi_{n,1} > x\} = \mathbb{P}\{S_{\tau_1 \wedge n} > x \,|\, \tau_1 < \infty\}, \quad x \in \mathbb{R}.$$

Then, from (5.18), for $x \ge x_0$,

$$\mathbb{P}\{\varphi_{n,1} > x\} \le \int_x^{x+na} \overline{G}_n(y)dy, \quad x \in \mathbb{R}, n \ge 1, \tag{5.19}$$

for some distribution function G_n on $\mathbb{R}$ satisfying

$$\lim_{x \to \infty} \frac{\overline{G}_n(x)}{\overline{F}(x)} = \frac{1 + \varepsilon}{\gamma(a - \varepsilon)}. \tag{5.20}$$

From the condition $F \in \mathcal{S}^*$ and Corollary 3.26, we have $G_n \in \mathcal{S}^*$. We may now apply Corollary 3.40 with a geometrically distributed stopping time (which is independent of the sequence of random variables ξ_i) to obtain that

$$(1 - \gamma) \sum_{k=0}^{\infty} \gamma^k \overline{G_n^{*k}}(x) \sim \frac{\gamma}{1 - \gamma} \overline{G}_n(x)$$

as $x \to \infty$ uniformly in $n \geq 1$. Using also the conditions (5.19) and (5.20), we get the following asymptotic upper bound:

$$\sum_{k=1}^{\infty} \gamma^k \mathbb{P}\{\varphi_{n,1} + \ldots + \varphi_{n,k} > x\} \leq \frac{\gamma + o(1)}{(1-\gamma)^2} \overline{G}_n(x)$$

$$\leq \frac{1 + \varepsilon + o(1)}{(1-\gamma)^2 (a - \varepsilon)} \int_x^{x+na} \overline{F}(y) dy \qquad (5.21)$$

as $x \to \infty$ uniformly in $n \geq 1$.

If $M_n > x$ then there exist $\tau_k \leq n$ and $j \in [\tau_k, \tau_{k+1})$ such that $S_j > x$. Then, exactly as in the proof of Theorem 5.4, we have that necessarily $S_{\tau_k} > x - A + a - \varepsilon$. It follows that

$$\{M_n > x\} \subseteq \bigcup_{k=1}^{\infty} \{S_{\tau_k \wedge n} > x - A + a - \varepsilon\}.$$

Therefore,

$$\mathbb{P}\{M_n > x\} \leq \sum_{k=1}^{\infty} \mathbb{P}\{S_{\tau_k \wedge n} > x - A + a - \varepsilon\}$$

$$\leq \sum_{k=1}^{\infty} \gamma^k \mathbb{P}\{\varphi_{n,1} + \ldots + \varphi_{n,k} > x - A + a - \varepsilon\},$$

by the construction of the random variables φ_i. Using (5.21) we obtain

$$\limsup_{x\to\infty} \sup_{n\geq 1} \frac{\mathbb{P}\{M_n > x\}}{\int_x^{x+na} \overline{F}(y) dy} \leq \frac{1 + \varepsilon}{(a - \varepsilon)(1-\gamma)^2}.$$

No let first $A \to \infty$, so that $\gamma \to 0$ by (5.9). Then let $\varepsilon \to 0$ to obtain the required upper bound

$$\limsup_{x\to\infty} \sup_{n\geq 1} \frac{\mathbb{P}\{M_n > x\}}{\int_x^{x+na} \overline{F}(y) dy} \leq \frac{1}{a},$$

which, together with the lower bound (5.14), implies the required uniform asymptotics (5.15). □

5.4 Comments

Theorem 5.4 was proved for regularly varying distributions by Callaert and Cohen in [11] and by Cohen in [16]. For dominated-varying distributions, it was proved by Borovkov in [9, Sec. 22]. In its present form, it was proved by Veraverbeke in [43] and by Embrechts, Goldie and Veraverbeke in [22]. The proof given here follows an idea of Zachary [44].

Bibliography

[1] Asmussen, S. (1987). *Applied Probability and Queues*, Wiley, Chichester (2nd ed. (2001) to be published by Springer, New York).

[2] Asmussen, S. (2000). *Ruin Probabilities*, World Scientific, Singapore.

[3] Asmussen, S. (1998). A probabilistic look at the Wiener–Hopf equation. *SIAM Review* **40**, 189–201.

[4] Asmussen S., Foss S., Korshunov D. (2003). Asymptotics for sums of random variables with local subexponential behaviour. *J. Theoret. Probab.* **16**, 489–518.

[5] Asmussen, S., Kalashnikov, V., Konstantinides, D., Klüppelberg, C., and Tsitsiashvili, G. (2002). A local limit theorem for random walk maxima with heavy tails. *Statist. Probab. Letters* **56**, 399–404.

[6] Athreya, K., and Ney, P. (1972). *Branching Processes*, Springer, Berlin.

[7] Bertoin, J., Doney, R. A. (1994). On the local behaviour of ladder height distributions. *J. Appl. Prob.* **31**, 816–821.

[8] Bingham, N. H., Goldie, C. M., and Teugels J. L. (1987). *Regular variation*, Cambridge University Press, Cambridge.

[9] Borovkov, A. A. (1976). *Stochastic Processes in Queueing Theory*, Springer, Berlin.

[10] Borovkov, A. A., Borovkov, K. A. (2008). Asymptotic Analysis of Random Walks. Heavy-Tailed Distributions. Cambridge University Press, Cambridge.

[11] Callaert, H., and Cohen, J. W. (1972). A lemma on regular variation of a transient renewal function. *Z. Wahrscheinlichkeitstheorie verw. Gebiete* **24**, 275–278.

[12] Chistyakov, V. P. (1964). A theorem on sums of independent positive random variables and its application to branching random processes, *Theor. Probability Appl.* **9**, 640–648.

[13] Chover, J., Ney, P., and Wainger, S. (1973). Functions of probability measures. *J. d'Analyse Mathématique* **26**, 255–302.

[14] Chover J., Ney P., and Wainger S. (1973). Degeneracy properties of subcritical branching processes. *Ann. Probab.* **1**, 663–673.

[15] Cline, D. (1987). Convolutions of distributions with exponential and subexponential tails. *J. Aust. Math. Soc.* **43**, 347–365.

[16] Cohen, J. W. (1973). Some results on regular variation for distributions in queueing and fluctuation theory. *J. Appl. Probab.* **10**, 343–353.

[17] Denisov, D., Foss, S. and Korshunov, D. (2004). Tail asymptotics for the supremum of a random walk when the mean is not finite. *Queueing Syst.* **46**, 15–33.

[18] Denisov, D., Foss, S., and Korshunov, D. (2008). On lower limits and equivalences for distribution tails of randomly stopped sums. *Bernoulli* **14**, 391–404.

[19] Denisov, D., Foss, S., and Korshunov, D. (2008). Lower limits for distribution tails of randomly stopped sums. *Theory Probab. Appl.* **52**, 690–699.

[20] Embrechts, P. and Goldie, C. M. (1980). On closure and factorization theorems for subexponential and related distributions. *J. Austral. Math. Soc. Ser. A* **29**, 243–256.

[21] Embrechts, P. and Goldie, C. M. (1982). On convolution tails. *Stoch. Processes Appl.* **13**, 263–278.

[22] Embrechts, P., Goldie, C. M., and Veraverbeke, N. (1979). Subexponentiality and infinite divisibility. *Z. Wahrscheinlichkeitstheorie verw. Gebiete* **49**, 335–347.

[23] Embrechts, P., Klüppelberg, C., and Mikosch, T. (1997). *Modelling Extremal Events for Insurance and Finance*, Springer, Berlin.

[24] Embrechts, P., and Veraverbeke, N. (1982). Estimates for the probability of ruin with special emphasis on the possibility of large claims. *Insurance: Math. and Economics* **1**, 55–72.

[25] Feller, W. (1971). *An Introduction to Probability Theory and Its Applications* Vol. 2, Wiley, New York.

[26] Foss, S. and Korshunov, D. (2007). Lower limits and equivalences for convolution tails. *Ann. Probab.* **35**, 366–383.

[27] Foss, S., and Zachary, S. (2003). The maximum on a random time interval of a random walk with long-tailed increments and negative drift. *Ann. Appl. Prob.* **13**, 37–53.

[28] Harris, T. (1963). *The Theory of Branching Processes*, Springer, Berlin.

[29] Klüppelberg, C. (1988). Subexponential distributions and integrated tails. *J. Appl. Prob.* **25**, 132–141.

[30] Klüppelberg, C. (1989). Subexponential distributions and characterization of related classes. *Probab. Th. Rel. Fields* **82**, 259–269.

[31] Klüppelberg, C. (1991). The full solution of the convolution closure problem for convolution-equivalent distributions. *J. Math. Anal. Appl.* **160**, 79–92.

[32] Korshunov, D. (1997). On distribution tail of the maximum of a random walk. *Stoch. Proc. Appl.* **72**, 97–103.

[33] Leslie, J. R. (1989). On the non-closure under convolution of the subexponential family. *J. Appl. Probab.* **26**, 58–66.

[34] Pakes, A. G. (1975). On the tails of waiting-time distributions. *J. Appl. Probab.* **12**, 555–564.

[35] Pakes, A. G. (2004). Convolution equivalence and infinite divisibility. *J. Appl. Probab.* **41**, 407–424.

[36] Pitman, E. J. G. (1980). Subexponential distribution functions. *J. Austral. Math. Soc. Ser. A* **29**, 337–347.

[37] Reiss, R.-D., Thomas, M. (2001). *Statistical analysis of extreme values*, Birkhäuser, Basel.

[38] Rogozin, B. A. (2000). On the constant in the definition of subexponential distributions. *Theory Probab. Appl.* **44**, 409–412.

[39] Rolski, T., Schmidli, H., Schmidt, V, and Teugels, J. (1998). *Stochastic Processes for Insurance and Finance*, Wiley, Chichester.

[40] Rudin, W. (1973). Limits of ratios of tails of measures. *Ann. Probab.* **1**, 982–994.

[41] Sgibnev, M. S. Banach algebras of functions that have identical asymptotic behaviour at infinity. *Siberian Math. J.* **22**, 179–187.

[42] Teugels, J. L. (1975). The class of subexponential distributions. *Ann. Probab.* **3**, 1000–1011.

[43] Veraverbeke, N. (1977). Asymptotic behaviour of Wiener-Hopf factors of a random walk. *Stoch. Proc. Appl.* **5**, 27–37.

[44] Zachary, S. (2004). A note on Veraverbeke's theorem. *Queueing Syst.* **46**, 9–14.

Index