Subexponential Distributions

Charles M. Goldie and Claudia Klüppelberg

Abstract

We survey the properties and uses of the class of subexponential probability distributions, paying particular attention to their use in modelling heavy-tailed data such as occurs in insurance and queueing applications. We give a detailed summary of the core theory and discuss subexponentiality in various contexts including extremes, random walks and Lévy processes with negative drift, and sums of random variables, the latter extended to cover random sums, weighted sums and moving averages.

1. Definition and first properties

Subexponential distributions are a special class of heavy–tailed distributions. The name arises from one of their properties, that their tails decrease more slowly than any exponential tail; see (1.4). This implies that large values can occur in a sample with non–negligible probability, and makes the subexponential distributions candidates for modelling situations where some extremely large values occur in a sample compared to the mean size of the data. Such a pattern is often seen in insurance data, for instance in fire, wind–storm or flood insurance (collectively known as catastrophe insurance). Subexponential claims can account for large fluctuations in the surplus process of a company, increasing the risk involved in such portfolios. This situation is treated in Section 2.

Subexponentials play a similar role in queueing models. Situations with extreme service times, modelled by a subexponential distribution, result in huge waiting times in the system (see Example 2.7). The workload process also shows large fluctuations (see Example 6.4).

Linear models are widely used as simple models for (or first order approximations to) dependent data. Extremely large values in the innovations, modelled by subexponential distributions, have immediate consequences for the single observation. Moreover, they cause effects in larger parts of the sample, determined by the linear filter.

In all these models a few large values may determine the long–term behaviour of a system. This can be made very precise by describing the sample path behaviour of resulting stochastic processes as the surplus process in insurance or the workload process of a queue, since the latter models have been the most fully investigated. This is reviewed in Section 6.

Heavy tails are just one of the consequences of the defining property of subexponential distributions, which is designed specially to work well with the probabilistic models commonly employed in the above–mentioned areas of application. The subexponential concept has just the right level of generality to be usable in these models while including as wide a range of distributions as possible. It includes all distributions with regularly varying tails (domains of attraction of sum– or max–stable laws) but is considerably wider (see Table 3.7). Hence it encompasses many more types of behaviour in the extremes (see Section 4).

Subexponential distributions were first studied in 1964 by Chistyakov. Research during the seventies was centred around applications in insurance, queueing and branching processes, based on the Pollaczek–Khinchin formula (2.2), linking a subexponential input df and an output df of interest. In a simple insurance model this output df may be the ruin probability, while in a simple queueing model it may be the df of the stationary waiting time. Methods were rather more analytic than probabilistic at that time. Properties of subexponential moment generating functions, necessary and sufficient conditions for subexponentiality, and closure properties were investigated.

Extensions to more general models followed: renewal arrival streams replaced Poisson arrivals. Modelling in that generality required the tracing of subexponential input distributions through a Wiener–Hopf factorisation. Use of random Markov environments required tracing different input distributions (light– and heavy–tailed), by means of matrix algebra.

Recently, more probabilistic methods have entered the field. Questions like "how does ruin happen?" or "when is ruin most likely to happen?" given it happens at all, or "what does the workload process at a high level look like?" were asked and answered. They necessitated novel methods to investigate path properties using the regenerative structure of models, as well as excursion theory for Markov processes and extreme value theory.

Against this background we present two defining properties of subexponential distributions. The first, more analytic one, is motivated by the Pollaczek–Khinchin formula (2.2) below, while the second probabilistic one provides a more intuitive interpretation of subexponentiality.

Definition 1.1. (Subexponential distribution function)
Let $(X_i)_{i\in\mathbb{N}}$ be iid positive rvs with df F such that $F(x) < 1$ for all $x > 0$. Denote

$$\overline{F}(x) = 1 - F(x)\,, \quad x \geq 0\,,$$

the tail of F and

$$\overline{F}^{n*} = 1 - F^{n*}(x) = P(X_1 + \cdots + X_n > x)$$

the tail of the n–fold convolution of F. F is a subexponential df ($F \in \mathcal{S}$) *if one of the following equivalent conditions holds:*

$$(a)\ \lim_{x\to\infty} \frac{\overline{F}^{n*}(x)}{\overline{F}(x)} = n \text{ for some (all) } n \geq 2\,,$$

$$(b)\ \lim_{x\to\infty} \frac{P(X_1 + \cdots + X_n > x)}{P(\max(X_1, \ldots, X_n) > x)} = 1 \text{ for some (all) } n \geq 2. \qquad \square$$

Remarks 1) Definition (a) goes back to Chistyakov (1964). He proved that the limit (a) holds for all $n \geq 2$ if and only if it holds for $n = 2$. It was shown in Embrechts and Goldie (1982) that (a) holds for $n = 2$ if it holds for some $n \geq 2$.

2) The equivalence of (a) and (b) was shown in Embrechts and Goldie (1980). A proof goes as follows:

$$P(\max(X_1, \dots, X_n) > x) = 1 - F^n(x) = \overline{F}(x) \sum_{k=0}^{n-1} F^k(x) \sim n\overline{F}(x)\,, \quad x \to \infty\,,$$

($\sim$ means that the quotient of lhs and rhs tends to 1). Hence

$$\frac{P(X_1 + \cdots + X_n > x)}{P(\max(X_1, \dots, X_n) > x)} \sim \frac{\overline{F}^{n*}(x)}{n\overline{F}(x)} \to 1 \quad \Longleftrightarrow \quad F \in \mathcal{S}\,.$$

3) Definition (b) provides a physical interpretation of subexponentiality: the sum of n iid subexponential rvs is likely to be large if and only if their maximum is likely to be large. This accounts for extremely large values in a subexponential sample.

4) From Definition (a) and the fact that $\mathcal{S}$ is closed with respect to tail–equivalence (see Definition 3.3) we conclude that

$$F \in \mathcal{S} \quad \Longrightarrow \quad F^{n*} \in \mathcal{S}\,, \quad n \in \mathbb{N}. \tag{1.1}$$

Furthermore, from Definition (b) and the fact that F^n is the df of the maximum of n iid rvs with df F, we conclude that

$$F \in \mathcal{S} \quad \Longrightarrow \quad F^{n} \in \mathcal{S}\,, \quad n \in \mathbb{N}.$$

Hence $\mathcal{S}$ is closed with respect to taking sums and maxima of iid rvs. The relationship of subexponentials and maxima will be further investigated in Section 4. Various generalisations of (1.1) will be considered in Section 5.

5) Definition (b) demonstrates the heavy–tailedness of subexponential dfs. It is further substantiated by the implications (first proved by Chistyakov (1964))

$$F \in \mathcal{S} \quad \Longrightarrow \quad \lim_{x \to \infty} \frac{\overline{F}(x-y)}{\overline{F}(x)} = 1 \quad \forall\, y \in \mathbb{R} \tag{1.2}$$

$$\Longrightarrow \quad \int_0^\infty e^{\varepsilon x}\, dF(x) = \infty \quad \forall\, \varepsilon > 0 \tag{1.3}$$

$$\Longrightarrow \quad \overline{F}(x)/e^{-\varepsilon x} \to \infty \quad \forall\, \varepsilon > 0\,. \tag{1.4}$$

Property (1.4) accounts for the name subexponential df: the tail of F decreases more slowly than any exponential tail. Property (1.3) shows that subexponential dfs have no exponential moments. This prevents any method being applicable that requires the existence of exponential moments. □

2. The supremum of a random walk with negative drift

Subexponential dfs traditionally play an important role in continuous time models with a random walk skeleton. We choose a class of insurance risk models for demonstrating the general method.

The *classical insurance risk process* is defined as

$$R(t) = u + ct - \sum_{i=1}^{N(t)} X_i\,, \quad t \geq 0\,,$$

where $u \geq 0$ is the *initial capital* (or *risk reserve*), and $c > 0$ is the *premium rate*, i.e. premiums are linear in time. $(N(t))_{t\geq 0}$ is a *homogeneous Poisson process with intensity* $\lambda > 0$, counting the number of claims up to time t. $(X_i)_{i\in\mathbb{N}}$ are iid positive claims, independent of $(N(t))$, with df F, finite mean μ and *integrated tail df*

$$F_I(x) = \frac{1}{\mu}\int_0^x \overline{F}(t)\,dt\,, \quad x \geq 0\,. \tag{2.1}$$

Denote by $\psi(u)$ the *ruin probability*, given a risk reserve u, i.e.

$$\psi(u) = P\big(R(t) < 0 \quad \text{for some} \quad t > 0\big).$$

The risk process $(R(t))$ has two important features: the *inter–arrival times* are iid exponential rvs $(E_i)_{i\in\mathbb{N}}$ with mean $1/\lambda$, and ruin can occur only at claim times. Hence if we define

$$S_0 = 0\,, \quad S_n = \sum_{i=1}^{n}(X_i - cE_i)\,, \quad n \in \mathbb{N},$$

then

$$\begin{aligned}\psi(u) &= P\big(S(t) > u \quad \text{for some} \quad t > 0\big)\\ &= P\big(S_n > u \quad \text{for some} \quad n \in \mathbb{N}\big)\\ &= P\Big(\max_{n\geq 1} S_n > u\Big).\end{aligned}$$

Under the *net–profit condition* $\rho = \lambda\mu/c < 1$, the random walk (S_n) has negative drift and $\psi(u) \to 0$ as $u \to \infty$. If we denote by

$$\tau(u) = \inf\{n \geq 0 : S_n > u\}$$

the *ruin time*, then

$$P\Big(\max_{n\geq 1} S_n > u\Big) = P\big(\tau(u) < \infty\big)\,,$$

representing the ruin problem as a problem of first hitting times. The ruin problem can be handled by an analysis of the ladder heights or by solving a renewal equation (see Asmussen(1996), Embrechts, Klüppelberg and Mikosch

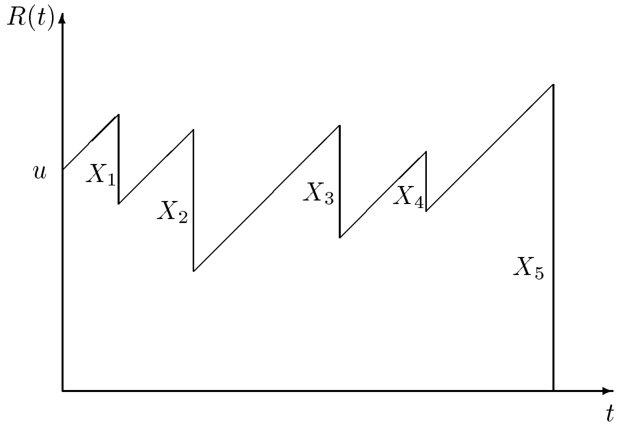

Figure 2.1. *Idealised sample path of the risk process.*

(1997), Feller (1971), Grandell (1991)), representing the non–ruin probability in terms of the *Pollaczek–Khinchin formula:*

$$1-\psi(u)=(1-\rho)\sum_{n=0}^{\infty}\rho^n F_I^{n*}(u)\,,\quad u\ge 0\,,\tag{2.2}$$

where F_I is the integrated tail df (2.1) and $F_I^{0*}=I_{[0,\infty)}$ is the df of Dirac (unit) measure at 0. In this representation ρF_I is the *ladder height df.* The infinite series on the rhs of (2.2) defines a defective renewal measure ($\rho F_I(x)\to\rho<1$ as $x\to\infty$), and the corresponding renewal process is transient: the sequence of renewals eventually stops, and at each renewal $1-\rho$ is the probability of termination then and there.

If *Cramér's condition* holds, i.e. if there exists some $\gamma>0$ such that

$$\int_0^\infty e^{\gamma x}\overline{F}(x)\,dx=\frac{c}{\lambda}\,,\tag{2.3}$$

the defect can be removed and, under the usual conditions, Smith's key renewal theorem implies that

$$\psi(u)e^{\gamma u}\to C\,,\quad u\to\infty\,,\tag{2.4}$$

where C is a non–negative constant; thus $\psi(u)$ decreases exponentially fast to 0. It is clear from (1.3) that for $F_I\in\mathcal{S}$ Cramér's condition (2.3) does not hold. But a different approach, as we now describe, shows that subexponentials form the class of *heavy–tailed* distributions that allows for ruin estimates.

We rewrite formula (2.2) in terms of the tails,

$$\psi(u)=(1-\rho)\sum_{n=1}^{\infty}\rho^n\overline{F}_I^{n*}(u)\,,\quad u\ge 0\,.$$

Dividing both sides by $\overline{F}_I(u)$, we see that Definition 1.1(a) yields an asymptotic estimate for $\psi(u)$ provided that one can safely interchange the limit and the

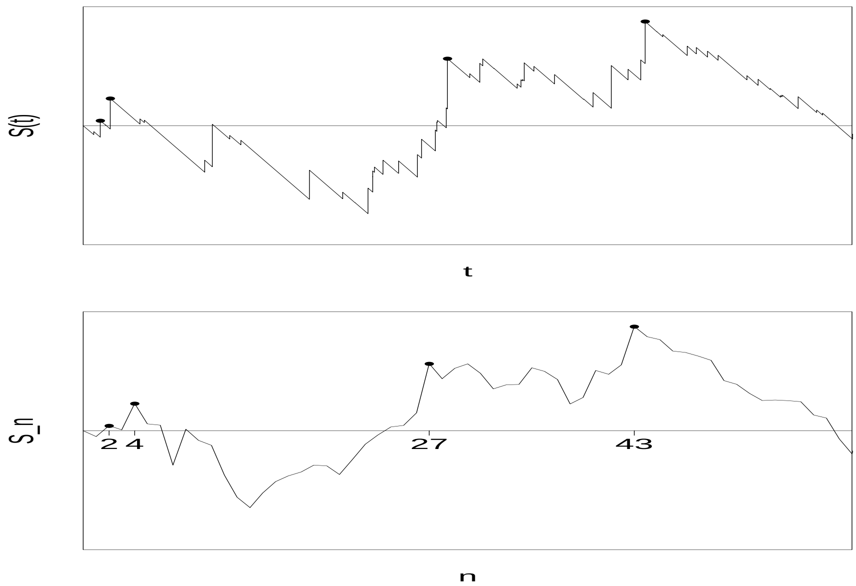

Figure 2.2. *Sample path of the process* $(S(t))$ *and its embedded random walk* (S_n)*. The ladder points are indicated by dots.*

infinite sum. This is ensured by the following lemma due to Kesten (for a proof see Athreya and Ney (1972)), and Lebesgue's dominated convergence theorem.

Lemma 2.3. *If* $F \in \mathcal{S}$*, then for every* $\varepsilon > 0$ *there exists some positive constant* $K(\varepsilon)$ *such that for all* $n \in \mathbb{N}$ *and* $x > 0$*,*

$$\frac{\overline{F}^{n*}(x)}{\overline{F}(x)} \le K(\varepsilon)(1+\varepsilon)^n . \qquad \square$$

As remarked, these considerations lead to an asymptotic evaluation of ψ. It turns out that this is not just a consequence of subexponentiality, but is characterised by it, as follows.

Theorem 2.4. (The ruin probability in the classical risk model)

$$F_I \in \mathcal{S} \iff 1-\psi \in \mathcal{S} \iff \lim_{u\to\infty} \frac{\psi(u)}{\overline{F}_I(u)} = \frac{\rho}{1-\rho} . \qquad \square$$

This theorem can be generalised by a Wiener–Hopf factorisation to the more general *Sparre Andersen model*, where the claim arrival process is an arbitrary renewal process.

Theorem 2.5. (The ruin probability in the renewal risk model)

$$1-\psi \in \mathcal{S} \iff F_I \in \mathcal{S} \implies \lim_{u\to\infty} \frac{\psi(u)}{\overline{F}_I(u)} = \frac{\rho}{1-\rho} . \qquad \square$$

The result of Theorem 2.4 has been further extended by Asmussen, Fløe Henriksen and Klüppelberg (1994) to a *Markov–modulated risk model*, where the risk process is not time–homogeneous, but evolves in an environment given by a Markov process with finite state space. A state of the Markov process defines the arrival intensity of the Poisson process and the claim–size distribution. Further results in the realm of this model have been obtained by Asmussen and Højgaard (1995) and Jelenkovič and Lazar (1996).

Asymptotic estimates for the ruin probability change when the company receives interest on its reserves. For regularly varying claim–size df F and a positive force of interest δ the corresponding ruin probability satisfies

$$\psi_\delta(u) \sim c_\delta \overline{F}(u)\,, \quad u \to \infty\,,$$

for some positive constant c_δ, i.e. it is tail–equivalent to the claim–size df itself. This has been proved in Klüppelberg and Stadtmüller (1996). The case of general subexponential claims has been treated in Asmussen (1996).

Remarks 1) The importance of subexponential dfs for insurance risk theory was recognised by Teugels (1975).

2) A textbook treatment of subexponential distributions in the context of risk theory is to be found in Embrechts, Klüppelberg and Mikosch (1997).

3) Theorem 2.4 is due to Embrechts and Veraverbeke (1982) based on work by Embrechts, Goldie and Veraverbeke (1979). Theorem 2.5 can be found in Embrechts and Veraverbeke (1982); see also Veraverbeke (1977) and Bertoin and Doney (1996). A density version of Theorem 2.4 can be found in Klüppelberg (1989a), Theorem 4.1 (since F_I has a density, so does $1-\psi$).

4) Theorem 2.5 can be further generalised to a general discrete time or continuous time random walk or Lévy process with negative drift and increment variable S_1 with df B such that the right tail of B satisfies $\overline{B}(x) = P(S_1 > x) \sim \overline{F}(x)$ for a subexponential df F. Notice that this is in accordance with the situation for the classical risk process, where $S_1 = X_1 - cE_1$ and

$$\frac{P(X_1 - cE_1 > x)}{P(X_1 > x)} = \int_0^\infty \frac{\overline{F}(x+cy)}{\overline{F}(x)} \lambda\, e^{-\lambda y}\, dy \to 1\,, \quad x \to \infty\,.$$

($\overline{F}(x+cy) \le \overline{F}(x)$ for all $x > 0$ and the quotient tends to 1 by (1.2), hence Lebesgue dominated convergence applies.) What is needed is that the ladder height df F_I is subexponential. This also shows that in this context it is quite natural to define subexponentiality only for positive rvs. □

Similar results to those for the risk models have been derived in the context of branching processes and queueing theory.

Example 2.6. (Branching processes)
Let $(Z(t))_{t\ge 0}$ denote the population size in the Bellman–Harris model, i.e. the particles produce (independently of each other) at the end of their lifetime a random number of offspring. Let F be the lifetime df of a particle and $m < 1$ the mean number of offspring. A renewal–type argument similar to the argument

leading to equation (2.2) yields, for $\mu(t) = EZ(t)$,

$$\mu(t) = \left(\sum_{n=0}^{\infty} m^n F^{n*}\right) * (1-F)(t) = \sum_{n=0}^{\infty} m^n \left(\overline{F}^{(n+1)*}(t) - \overline{F}^{n*}(t)\right).$$

An application of Lemma 2.3 yields together with Definition 1.1 an obvious analogue of Theorem 2.4. Early references in this context are Athreya and Ney (1972), Chistyakov (1964) and Chover, Ney and Wainger (1973). □

Example 2.7. (Queueing models)
Consider a GI/G/1 queue with renewal arrival stream and general service time df F. Let F have finite mean μ and integrated tail distribution (2.1). We consider a stable queue, i.e. with traffic intensity $\rho < 1$. Then the stationary waiting time df can be represented as the df of the maximum of a random walk (Feller (1971), VI.9). Hence analogues of Theorem 2.4 (corresponding to an M/G/1 queue) and Theorem 2.5 are immediate. Early results were derived by Pakes (1975), Smith (1972) and Veraverbeke (1977). □

3. Conditions for subexponentiality

It should be clear from the definition that a characterisation of subexponential dfs or even of dfs whose integrated tail df is subexponential (as needed in the risk and queueing models) will not be possible in terms of simple expressions involving the tail.

Recall that all subexponential dfs have property (1.2), hence the class of such dfs provides potential candidates for subexponentiality. The class is named as follows.

Definition 3.1. (The class $\mathcal{L}$)
Let F be a df on $(0,\infty)$ such that $F(x) < 1$ for all $x > 0$. We say $F \in \mathcal{L}$ if

$$\lim_{x\to\infty} \frac{\overline{F}(x-y)}{\overline{F}(x)} = 1 \quad \forall\, y > 0\,. \qquad \square$$

Unfortunately, $\mathcal{S}$ is a proper subset of $\mathcal{L}$. Examples for a df in $\mathcal{L}$ but not in $\mathcal{S}$ can be found in Embrechts and Goldie (1980) and Pitman (1980).

A famous subclass of $\mathcal{S}$ is the class of dfs with regularly varying tail. For a positive measurable function f we write $f \in \mathcal{R}(\alpha)$ for $\alpha \in \mathbb{R}$ (f is *regularly varying with index* α) if

$$\lim_{x\to\infty} \frac{f(tx)}{f(x)} = t^{\alpha} \quad \forall\, t > 0\,.$$

A function $f \in \mathcal{R}(0)$ is called *slowly varying*. For further properties of regularly varying functions we refer to the monograph by Bingham, Goldie and Teugels (1989).

Example 3.2. (Distribution functions with regularly varying tails)
Let $\overline{F} \in \mathcal{R}(-\alpha)$ for $\alpha \geq 0$, then it has the representation

$$\overline{F}(x) = x^{-\alpha}\ell(x)\,, \quad x > 0\,,$$

for some $\ell \in \mathcal{R}(0)$. Notice first that $F \in \mathcal{L}$, hence it is a candidate for $\mathcal{S}$. We check Definition 1.1(a). Let X_1, X_2 be iid rvs with df F. Now use the decomposition

$$\begin{aligned} P(X_1 + X_2 > x) &= P(X_1 \leq \frac{x}{2}, X_1 + X_2 > x) + P(X_2 \leq \frac{x}{2}, X_1 + X_2 > x) \\ &\quad + P(X_1 > \frac{x}{2}, X_2 > \frac{x}{2})\,. \end{aligned}$$

Then

$$\frac{\overline{F^{2*}}(x)}{\overline{F}(x)} = 2\int_0^{x/2} \frac{\overline{F}(x-y)}{\overline{F}(x)}\,dF(y) + \frac{\overline{F}^2(x/2)}{\overline{F}(x)}\,.$$

Immediately, by the definition of $\mathcal{R}(-\alpha)$, the last term tends to 0. The integrand satisfies $\overline{F}(x-y)/\overline{F}(x) \leq \overline{F}(x/2)/\overline{F}(x)$ for $0 \leq y \leq x/2$, hence Lebesgue dominated convergence applies and, since $F \in \mathcal{L}$, the integral on the rhs tends to 1 as $x \to \infty$.

Examples of dfs with regularly varying tail are Pareto, Burr, log–gamma and stable dfs (see Table 3.7). If $\alpha > 1$ then F has finite mean and, by Karamata's theorem, $F_I \in \mathcal{R}(-(\alpha-1))$, giving $F_I \in \mathcal{S}$ as well. □

In much of the present discussion we are dealing only with the right tail of a df. This notion can be formalised, starting with the following definition.

Definition 3.3. (Tail–equivalence)
Two dfs F and G with support unbounded to the right are called tail–equivalent *if* $\lim_{x\to\infty} \overline{F}(x)/\overline{G}(x) = c \in (0,\infty)$. □

The next representation is a consequence of Theorem 1.3.1 of Bingham, Goldie and Teugels (1989) and the fact that

$$F \in \mathcal{L} \quad \Longleftrightarrow \quad \overline{F} \circ \ln \in \mathcal{R}(0)\,.$$

Lemma 3.4. (Representation of dfs in $\mathcal{L}$)
$F \in \mathcal{L}$ if and only if it has representation

$$\overline{F}(x) = c(x)\exp\left\{-\int_z^x q(t)\,dt\right\}\,, \quad x \geq z \geq 0\,,$$

where c and q are non–negative measurable functions such that $c(x) \to c \in (0,\infty)$ and $q(x) \to 0$, as $x \to \infty$, and $\int_z^\infty q(t)\,dt = \infty$. □

This implies in particular that each $F \in \mathcal{L}$ is tail–equivalent to an absolutely continuous df with hazard rate q which tends to 0 (for a definition see after

Remark 3 below). Since $\mathcal{S}$ is closed with respect to tail–equivalence (Teugels (1975)) it is of interest to find conditions on the hazard rate such that the corresponding df or/and integrated tail df is subexponential. In order to unify the problem of finding conditions for $F \in \mathcal{S}$ and $F_I \in \mathcal{S}$, the following class was introduced in Klüppelberg (1988).

Definition 3.5. (The class $\mathcal{S}^*$)
Let F be a df on $(0,\infty)$ such that $F(x) < 1$ for all $x > 0$. We say $F \in \mathcal{S}^*$ if F has finite mean μ and

$$\lim_{x\to\infty} \int_0^x \frac{\overline{F}(x-y)}{\overline{F}(x)} \overline{F}(y)\, dy = 2\mu\,. \qquad \square$$

The next result makes the class useful for applications.

Proposition 3.6. *If $F \in \mathcal{S}^*$, then $F \in \mathcal{S}$ and $F_I \in \mathcal{S}$.* $\square$

Name	*Tail $\overline{F}$ or density f*	*Parameters*
Pareto	$\overline{F}(x) = \left(\frac{\kappa}{\kappa + x}\right)^{\alpha}$	$\alpha, \kappa > 0$
Burr	$\overline{F}(x) = \left(\frac{\kappa}{\kappa + x^{\tau}}\right)^{\alpha}$	$\alpha, \kappa, \tau > 0$
Log–gamma	$f(x) = \frac{\alpha^{\beta}}{\Gamma(\beta)}(\ln x)^{\beta-1}\, x^{-\alpha-1}$	$\alpha > 1, \beta > 0$
Truncated α–stable	$\overline{F}(x) = P(\lvert X\rvert > x)$ where X is an α–stable rv	$0 < \alpha < 2$
Lognormal	$f(x) = \frac{1}{\sqrt{2\pi}\,\sigma x} e^{-(\ln x - \mu)^2/(2\sigma^2)}$	$\mu \in \mathbb{R},\ \sigma > 0$
Benktander–type–I	$\overline{F}(x) = c(\alpha + 2\beta \ln x)\, e^{-(\beta(\ln x)^2 + (\alpha+1)\ln x)}$	$c, \alpha, \beta > 0$
Benktander–type–II	$\overline{F}(x) = c\alpha x^{-(1-\beta)} e^{-\alpha\, x^{\beta}/\beta}$	$c, \alpha > 0$, $0 < \beta < 1$
Weibull	$\overline{F}(x) = e^{-x^{\tau}}$	$0 < \tau < 1$
"Almost" exponential	$\overline{F}(x) = e^{-x(\ln x)^{-\alpha}}$	$\alpha > 0$

Table 3.7. *Subexponential dfs. All of them are in $\mathcal{S}^*$ provided they have finite mean.*

Remarks 1) The class $\mathcal{S}^*$ is "almost" $\mathcal{S} \cap \{F : \mu(F) < \infty\}$, where $\mu(F)$ is the mean of F. A precise formulation can be found in Klüppelberg (1988).

2) The tails of dfs in $\mathcal{S}^*$ are subexponential densities (Klüppelberg (1989a), Willekens (1986)).

3) The class $\mathcal{S}^*$ is closed with respect to tail–equivalence. □

The task of finding easily verifiable conditions for $F \in \mathcal{S}$ or/and $F_I \in \mathcal{S}$ has now been reduced to the finding of simple conditions for $F \in \mathcal{S}^*$. We formulate some of them in terms of the *hazard function* $Q = -\ln \overline{F}$ and its density q, the *hazard rate* of F. (Recall that $\mathcal{S}^* \subset \mathcal{S} \subset \mathcal{L}$, hence by Lemma 3.4 each $F \in \mathcal{S}^*$ is tail–equivalent to an absolutely continuous df whose hazard rate tends to 0.)

Proposition 3.8. (Conditions for $F \in \mathcal{S}^*$)

(*a*) *If* $\limsup_{x\to\infty} xq(x) < \infty$, *then* $F \in \mathcal{S}^*$.

(*b*) *If there exist* $\delta \in (0,1)$ *and* $v \geq 1$ *such that* $Q(xy) \leq y^\delta Q(x)$ *for all* $x \geq v\,, y \geq 1$ *and* $\liminf_{x\to\infty} xq(x) \geq (2-2^\delta)^{-1}$, *then* $F \in \mathcal{S}^*$.

(*c*) *If* q *is eventually decreasing to* 0, *then*

$$F \in \mathcal{S}^* \quad \Longleftrightarrow \quad \lim_{x\to\infty} \int_0^x e^{yq(x)}\overline{F}(y)\,dy = \mu\,. \qquad \square$$

Corollary 3.9. (More conditions for $F \in \mathcal{S}^*$)
Suppose

$$\lim_{x\to\infty} q(x) = 0 \quad \textit{and} \quad \lim_{x\to\infty} xq(x) = \infty\,.$$

If additionally one of the following conditions holds, then $F \in \mathcal{S}^*$.

(*a*) $\limsup_{x\to\infty} xq(x)/Q(x) < 1$.

(*b*) $q \in \mathcal{R}(-\delta)$ *for* $\delta \in (0,1]$.

(*c*) $Q \in \mathcal{R}(\delta)$ *for* $\delta \in (0,1)$ *and* q *is eventually decreasing.*

(*d*) $q \in \mathcal{R}(0)$, q *is eventually decreasing, and* $Q(x) - xq(x) \in \mathcal{R}(1)$. □

There are many more conditions for $F \in \mathcal{S}$ or $F_I \in \mathcal{S}$ to be found in the literature. We mention Chistyakov (1964), Cline (1986), Goldie (1978), Klüppelberg (1988), Pitman (1980), Teugels (1975); the selection above is taken from Klüppelberg (1988, 1989b).

4. Subexponentials and maxima

Definition 1.1(b) suggests subexponential dfs as appropriate models for extremal events. This immediately warrants an investigation of their relationship to classical extreme value theory. For an introduction to the latter we refer to Embrechts, Klüppelberg and Mikosch (1997), Chapter 3, or Resnick (1987).

Let $(X_n)_{n\in\mathbb{N}}$ be iid rvs with df $F \in \mathcal{S}$ and assume that there exist constants $a_n > 0$ and $b_n \in \mathbb{R}$ such that

$$a_n^{-1}\big(\max(X_1, \dots, X_n) - b_n\big) \quad \stackrel{d}{\to} \quad G\,, \quad n \to \infty\,,$$

where G is some non–degenerate df. In this case we say F is in the maximum domain of attraction of G and write $F \in \mathrm{MDA}(G)$. If $F \in \mathcal{S}$ its support is unbounded above, hence G is either the Fréchet df $\Phi_\alpha(x) = \exp\{-x^{\alpha}\}$ for $x \geq 0$, where $\alpha > 0$, or the Gumbel df $\Lambda(x) = \exp\{-e^{-x}\}$ for $x \in \mathbb{R}$. We write $F \in \mathrm{MDA}(\Phi_\alpha)$ or $F \in \mathrm{MDA}(\Lambda)$, respectively.

It is well known that $F \in \mathrm{MDA}(\Phi_\alpha)$ if and only if $\overline{F} \in \mathcal{R}(-\alpha)$. Thus it remains to investigate $\mathcal{S} \cap \mathrm{MDA}(\Lambda)$.

A good indicator for the extremal behaviour of a model is the *mean–excess function* (which exists for dfs with finite mean)

$$a(x) = E(X - x \mid X > x) = \int_x^\infty \overline{F}(y)\,dy/\overline{F}(x)\,, \quad x > 0\,.$$

From Karamata's theorem we know that F has finite mean when $\overline{F} \in \mathcal{R}(-\alpha)$ with $\alpha > 1$. Moreover, $\overline{F} \in \mathcal{R}(-(\alpha+1))$ for $\alpha > 0$ if and only if $a(x) \sim x/\alpha$. For the lognormal df we have $a(x) \sim \sigma^2 x/\ln x$, and for the Weibull df $a(x) \sim x^{1-\tau}/\tau$ in the parametrisation of Table 3.7. (Recall that $a(x)$ is constant for the exponential df and converges to 0 for the normal df.)

Necessary conditions and sufficient conditions for $F \in \mathrm{MDA}(\Lambda) \cap \mathcal{S}$ have been derived by Goldie and Resnick (1988). The following condition applies to the examples in Table 3.7.

Lemma 4.1. *Let F be a df with finite mean and assume that $a(x)$ is eventually non–decreasing and there exists some $t > 1$ such that*

$$\liminf_{x\to\infty} \frac{a(tx)}{a(x)} > 1\,.$$

Then $F \in \mathrm{MDA}(\Lambda) \cap \mathcal{S}$. □

Remarks 1) Pareto, Burr, log–gamma and stable dfs belong to $\mathrm{MDA}(\Phi_\alpha)$ for some $\alpha > 0$, while lognormal, Benktander and Weibull dfs are in $\mathrm{MDA}(\Lambda)$.

2) For $F \in \mathrm{MDA}(\Lambda)$ with infinite right endpoint we have $\lim_{x\to\infty} a(x)/x = 0$. A generalisation of Karamata's theorem ensures that $\overline{F} \in \mathcal{R}(-\infty)$, i.e.

$$\lim_{x\to\infty} \frac{\overline{F}(tx)}{\overline{F}(x)} = \begin{cases} 0 & t > 1\,, \\ \infty & t < 1\,. \end{cases}$$

3) The fact that subexponential dfs may belong to $\mathrm{MDA}(\Phi_\alpha)$ and $\mathrm{MDA}(\Lambda)$ has consequences when studying extremal events in various models with subexponential input functions; see Theorems 5.4 and 6.2 for examples. □

5. Subexponentials and sums

From (1.1) we know that $\mathcal{S}$ is closed under the operation of taking sums of iid rvs. It is also closed under convolution roots; that is, the converse to (1.1) is true (Embrechts, Goldie and Veraverbeke (1979)). In this section we investigate further closure and other properties related to sums of subexponential rvs.

Convolution closure

A question naturally emerging from (1.1) is whether $\mathcal{S}$ is in general convolution closed, i.e. if F, $G \in \mathcal{S}$, does it always follow that $F * G \in \mathcal{S}$? The (negative) answer was given by Leslie (1989), who found two subexponential dfs whose convolution is not in $\mathcal{S}$. However, this must be a rather pathological example, as the following result covers most "reasonable" cases.

Theorem 5.1. (Convolution closure properties of $\mathcal{S}$)

(*a*) *Let* $F \in \mathcal{S}$ *and* $\overline{G}_i(x) \sim c_i \overline{F}(x)$*, where* $c_i \in (0,\infty)$ *for* $i = 1, 2$*. Then* $\overline{G_1 * G_2}(x) \sim (c_1 + c_2)\overline{F}(x)$.

(*b*) *Let* $F \in \mathcal{S}$ *and* $\overline{G}(x) \sim c\overline{F}(x)$ *for* $c \in [0,\infty)$*. Then* $\overline{F * G}(x) \sim (1+c)\overline{F}(x)$.

(*c*) *Let* $F, G \in \mathcal{S}$*. Then* $F * G \in \mathcal{S}$ *if and only if* $pF + (1-p)G \in \mathcal{S}$ *for some* (*all*) $p \in (0,1)$. □

Remarks 1) It has been known for a long time (see Feller (1971)) that the subclass of dfs with regularly varying tails is convolution closed. Indeed, if $\overline{F}(x) = x^{-\alpha}\ell_1(x)$ and $\overline{G}(x) = x^{-\alpha}\ell_2(x)$, then $\overline{F * G}(x) \sim x^{-\alpha}(\ell_1(x) + \ell_2(x))$. Notice that the case of two different indices of regular variation is covered by Theorem 5.1(b) for $c = 0$.
2) For a proof of Theorem 5.1 we refer to Embrechts and Goldie (1982); see also Cline (1986). It is possible to develop a special algebra to handle convolution questions. After all everything happens in the convolution semigroup of measures on $(0,\infty)$ and subexponential dfs can be considered as idempotent elements in the factor–semigroup with respect to tail–equivalence. Cline (1987) and Klüppelberg (1990) follow such an approach. □

Random sums

Theorem 2.4, together with (2.2), can be viewed as a generalisation of (1.1) to random (geometric) sums. The following result is due to Embrechts, Goldie and Veraverbeke (1979), Embrechts and Goldie (1982), and Cline (1987).

Theorem 5.2. (Random sums of iid subexponential rvs)
Suppose (p_n) *defines a probability measure on* $\mathbb{N}_0$ *such that* $\sum_{n=0}^{\infty} p_n(1+\varepsilon)^n < \infty$ *for some* $\varepsilon > 0$ *and* $p_k > 0$ *for some* $k \geq 2$*. Let*

$$G(x) = \sum_{n=0}^{\infty} p_n F^{n*}(x)\,, \quad x > 0\,. \tag{5.1}$$

Then

$$F \in \mathcal{S} \iff \lim_{x\to\infty} \frac{\overline{G}(x)}{\overline{F}(x)} = \sum_{n=1}^{\infty} n p_n \iff G \in \mathcal{S} \text{ and } \overline{F}(x) \neq o(\overline{G}(x))\,. \quad \square$$

Remarks 3) Let $(X_i)_{i\in\mathbb{N}}$ be iid with df F and let N be a rv taking values in $\mathbb{N}_0$ with distribution (p_n). Then G is the df of the random sum $\sum_{i=1}^N X_i$ (with the convention $\sum_{i=1}^0 X_i = 0$) and the result of Theorem 5.2 translates into

$$P\Big(\sum_{i=1}^N X_i > x\Big) \sim ENP(X_1 > x)\,, \quad x \to \infty\,.$$

If (p_n) is a Poisson or geometric distribution the condition $\overline{F}(x) \neq o(\overline{G}(x))$ in (c) is unnecessary (Cline (1987)). □

A further generalisation of Theorem 5.2 is towards infinite divisibility. Let F be an infinitely divisible df on $(0,\infty)$. Then its moment generating function $\widehat{f}$ has the representation

$$\widehat{f}(s) = \exp\Big\{as - \int_0^\infty (1-e^{sx})\,d\nu(x)\Big\}\,, \quad s \geq 0\,, \tag{5.2}$$

where $a \geq 0$ is a constant and ν is the Lévy measure of F. The following result was proved by Embrechts, Goldie and Veraverbeke (1979). It is based on the representation of F as $F = F_1 * F_2$, where $F_1(x) = o(e^{-\varepsilon x})$ for all $\varepsilon > 0$ and $F_2(x)$ is compound Poisson with the normalised Lévy measure as compounding df. Then $\overline{F}(x) \sim \overline{F}_2(x)$ by Theorem 5.1(b), and the closure of $\mathcal{S}$ with respect to tail–equivalence ensures $F \in \mathcal{S} \iff F_2 \in \mathcal{S}$. From Theorem 5.2 one obtains:

Corollary 5.3. (Infinitely divisible dfs and Lévy measures)

$$F \in \mathcal{S} \iff \nu(1,x]/\nu(1,\infty) \in \mathcal{S} \iff \overline{F}(x) \sim \nu(x,\infty)\,. \qquad \square$$

Remarks 4) This result has been extended to infinitely divisible processes by Rosinski and Samorodnitsky (1993) who relate subadditive functionals of a sample path to a subexponential Lévy measure.
5) The asymptotic behaviour of high quantiles of an infinitely divisible process with regularly varying Lévy measure has been investigated by Embrechts and Samorodnitsky (1995). □

Large Deviations

A further question immediately arises from Definition 1.1, namely what happens if n varies together with x. Hence *large deviations* theory is called for. Notice that the usual "rough" large deviations machinery based on logarithms cannot be applied. Classical results for $\overline{F} \in \mathcal{R}(-\alpha)$ state that

$$P(S_n - ES_n > x) \sim P(\max(X_1,\dots,X_n) > x) \sim n\overline{F}(x)\,, \quad n \to \infty\,, \tag{5.3}$$

which relation holds uniformly for $x > \gamma n$ for every fixed $\gamma > 0$. "Uniformly" here is in a ratio sense:

$$\sup_{x\in(\gamma n,\infty)} \left| \frac{P(S_n - ES_n > n)}{n\overline{F}(x)} - 1 \right| \to 0\,, \quad n \to \infty\,;$$

see Heyde (1967a, 1967b, 1968), A.V. Nagaev (1969a, 1969b) and Vinogradov (1994). Large deviations results for so–called semi–exponential tails $\overline{F}(x) = \exp\{-x^\alpha \ell(x)\}$, for $\alpha \in (0,1)$ and $\ell \in \mathcal{R}(0)$, have been derived by S. V. Nagaev (1979); see also Rozovskii (1993). However, for such tails the x–regions, where (5.3) holds, do not in general include all the region $[\gamma n, \infty)$. A very general treatment of large deviation results for subexponentials is given in Pinelis (1985). For references and extensions of (5.3) towards random sums we refer to Klüppelberg and Mikosch (1997), where also certain applications to insurance and finance are treated. Generalisations to mixing sequences are to be found in Gantert (1996).

Weighted sums of subexponential random variables

Weighted sums are the first objects to study on the way to linear processes; they are the one–dimensional objects. The results given in Theorem 5.5 below were derived by Davis and Resnick (1985, 1988), and they have been used in combination with point–process techniques for studying the extremes of linear processes.

Assume that $(Z_j)_{j\in\mathbb{Z}}$ are iid with subexponential df F, and form the weighted sum

$$X = \sum_{j=-\infty}^{\infty} \psi_j Z_j. \tag{5.4}$$

The real sequence (ψ_j) is assumed to have properties such that X is well–defined as an almost–surely converging series. For this application it is natural to extend the notion of subexponentiality to dfs on the real line. Let F be a df on $\mathbb{R}$ and $F(x) < 1$ for all $x \in \mathbb{R}$. F is called a *subexponential df on* $\mathbb{R}$ if there exists a subexponential df G on $(0,\infty)$ such that $\overline{F}(x) \sim \overline{G}(x)$ as $x \to \infty$. In order to derive the tail behaviour of the df of the weighted sum X given in (5.4) we assume the tail balance condition

$$\overline{F}(x) \sim p\,P(|Z| > x)\,, \quad F(-x) \sim qP(|Z| > x) \tag{5.5}$$

for $p \in (0,1]$ and $q = 1-p$.

Proposition 5.4. *Assume that Z is a rv with df F, subexponential on $\mathbb{R}$.*

(*a*) *Let $\overline{F}(x) \sim px^{-\alpha}\ell(x)$ and $F(-x) \sim qx^{-\alpha}\ell(x)$. Then*

$$\begin{aligned} P(\psi_j Z > x) &= \begin{cases} P\big(Z > x/\psi_j\big) \sim \psi_j^\alpha p x^{-\alpha}\ell(x) & if \quad \psi_j > 0, \\ P\big(Z < -x/|\psi_j|\big) \sim |\psi_j|^\alpha q x^{-\alpha}\ell(x) & if \quad \psi_j < 0, \end{cases} \\ &= |\psi_j|^\alpha x^{-\alpha}\ell(x)\left(pI_{\{\psi_j>0\}} + qI_{\{\psi_j<0\}}\right). \end{aligned}$$

(*b*) *Let $\overline{F} \in \mathcal{R}(-\infty)$ and assume that (5.5) holds, then*

$$\frac{P(\psi_j Z > x)}{\overline{F}(x)} = \begin{cases} 1 & if \quad \psi_j = 1, \\ q/p & if \quad \psi_j = -1, \\ 0 & if \quad |\psi_j| < 1. \end{cases}$$

□

From Theorem 5.1(b) we derive for independent rvs $(X_i)_{|i|\le m}$ such that $P(X_i > x) \sim a_i \overline{F}(x)$, where $a_i \in [0,\infty)$, that

$$P\Big(\sum_{|i|\le m} X_i > x\Big) \sim \overline{F}(x) \sum_{|i|\le m} a_i \,.$$

This result can immediately be applied to the truncated sum

$$X^{(m)} = \sum_{|j|\le m} \psi_j Z_j \,,$$

where the Z_j are iid with subexponential df F on $\mathbb{R}$ satisfying the tail balance condition (5.5). If the sequence (ψ_j) tends sufficiently fast to 0, then the result for the truncated sum $X^{(m)}$ extends to the infinite sum (5.4).

Theorem 5.5. *Let $(Z_j)_{j\in\mathbb{Z}}$ be iid rvs with df F, subexponential on $\mathbb{R}$, and let X be the random sum given by (5.4).*

(*a*) *If $\overline{F} \in \mathcal{R}(-\alpha)$ for $\alpha \in (0,\infty)$, i.e. $P(|Z_1| > x) = x^{-\alpha}\ell(x)$, and $\sum_{j=-\infty}^{\infty} |\psi_j|^\delta < \infty$ for some $\delta \in (0,\min(\alpha,1))$, then*

$$P(X > x) \sim x^{-\alpha}\ell(x) \sum_{j=-\infty}^{\infty} |\psi_j|^\alpha \left(pI_{\{\psi_j>0\}} + qI_{\{\psi_j<0\}}\right) \,.$$

(*b*) *If $F \in \mathrm{MDA}(\Lambda)\cap\mathcal{S}$ and $\sum_{j=-\infty}^{\infty} |\psi_j|^\delta < \infty$ for some $\delta \in (0,1)$ and without loss of generality $\max_j |\psi_j| = 1$ (or else we normalise X), then*

$$P(X > x) \sim (pk^+ + qk^-)P(|Z_1| > x) \,,$$

where k^+ is the total number of times ψ_j takes the value 1 (there can only be finitely many), and k^- is the total number of times ψ_j takes the value -1. □

6. Rare events of a Lévy process with subexponential increments

Let (S_t) be a Lévy process in continuous time or a random walk in discrete time with increment S_1 having df B. Assume furthermore that (S_t) has negative drift, i.e. $-\beta = ES_1 < 0$. Then $M = \max_{t\ge 0} S_t < \infty$ a.s. and, if we define $\tau(u) = \inf\{t > 0 : S_t > u\}$, then $\{M > u\} = \{\tau(u) < \infty\}$. Furthermore, for u large, this event is rare, i.e.

$$\psi(u) = P(\tau(u) < \infty) = P(M > u)$$

is small. Typical large deviations problems are the asymptotic form of $\psi(u)$ as $u \to \infty$ (derived in Section 2), and properties of a sample path leading to an upcrossing of a high level u. Let

$$P^{(u)} = P(\cdot \mid \tau(u) < \infty) \,,$$

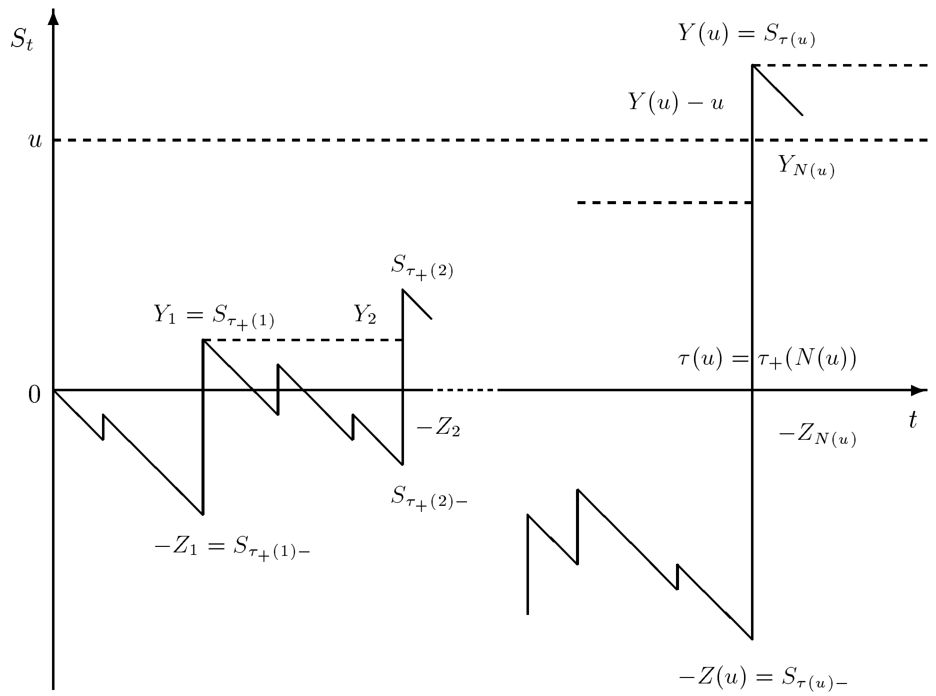

Figure 6.1. *Sample path of a classical ruin process leading to ruin.* $(Y_i)_{i\in\mathbb{N}}$ *are the ladder heights,* $(S_{\tau+(k)})_{k\in\mathbb{N}}$ *are the ladder epochs of* $(S_t)_{t\geq 0}$*. The rv* $N(u)$*, giving the number of ladder segments until ruin, is geometric.*

then we are interested in the $P^{(u)}$–distribution of the path

$$S_{[0,\tau(u))} = (S_t)_{0\leq t\leq \tau(u)}$$

leading to the occurrence of a rare event.

We are in particular interested in the following quantities:

$Y(u) = S_{\tau(u)}$	the level of the process after the upcrossing,
$Z(u) = S_{\tau(u)-}$	the level of the process just before the upcrossing,
$Y(u) - u$	the size of the overshoot,
$W(u) = Y(u) + Z(u)$	the size of the increment leading to the upcrossing.

For the sake of contrast we briefly report on the results under a Cramér condition (see Asmussen (1982) and references therein). Suppose there exists some $\gamma > 0$ such that the moment generating function $\widehat{b}(s) = Ee^{sS_1}$ has $\widehat{b}(\gamma) = 1$. Then the $P^{(u)}$–distribution of $S_{[0,\tau(u))}$ is in an appropriate sense the same as the unconditional distribution with respect to the Lévy process obtained by the exponential change of measure defined by

$$dB_\gamma(x) = e^{\gamma x} dB(x)\,.$$

The Lévy process so defined (with increment df B_γ) has positive drift. In particular, if the cumulant generating function $\kappa(s) = \ln \widehat{b}(s)$ has $\kappa'(\gamma) < \infty$

then, as $u \to \infty$,

$$\frac{\tau(u)}{u} \xrightarrow{P^{(u)}} (\kappa'(\gamma))^{-1} \quad \text{and} \quad \frac{S_{t\tau(u)}}{t\tau(u)} \xrightarrow{P^{(u)}} \kappa'(\gamma)\,.$$

We can summarise the situation under a Cramér condition as follows: the behaviour of the sample path of the Lévy process (or random walk) leading to an upcrossing is as if the increment distribution changed from B to B_γ. The main dramatic feature we see in the sample path is a change of drift causing the upcrossing. The intuitive picture is that rare events occur as a consequence of a build–up of claims over a period where the underlying parameters change by exponential change of measure. The sample path of the risk process leading to ruin exhibits a change of drift.

This picture changes radically for subexponential increment distributions (Asmussen and Klüppelberg (1996); see also Asmussen (1996)). Here an upcrossing happens as a result of one large increment whereas the process behaves in a typical way until the rare event happens. The following result describes the behaviour of the process before an upcrossing, and the upcrossing event itself.

Theorem 6.2. (Sample path leading to ruin)
Assume that the increment S_1 has df B with finite mean $-\mu < 0$. Assume furthermore that $\overline{B}(x) \sim \overline{F}(x)$ as $x \to \infty$ for some $F \in \mathcal{S}^* \cap \mathrm{MDA}(G)$ for some extreme value distribution G. Let $a(u) = \int_u^\infty \overline{F}(y)\,dy/\overline{F}(u)$. Then, as $u \to \infty$,

$$\left(\frac{Z(u)}{a(u)}, \frac{\tau(u)}{a(u)}, \frac{Y(u)-u}{a(u)}, \left(\frac{S_{t\tau(u)}}{\tau(u)}\right)_{0\le t<1}\right) \longrightarrow \left(V_\alpha, \frac{V_\alpha}{\mu}, T_\alpha, (-\mu t)_{0\le t<1}\right)$$

in $P^{(u)}$–distribution in $\mathbb{R}\times\mathbb{R}_+\times\mathbb{R}_+\times\mathbb{D}\,[0,1)$, where $\mathbb{D}\,[0,1)$ denotes the space of cadlag functions on $[0,1)$, and V_α and T_α are positive rvs with df satisfying

$$P(V_\alpha > x, T_\alpha > y) = \overline{G}_\alpha(x+y) = \begin{cases} \left(1+\dfrac{x+y}{\alpha}\right)^{-\alpha} & if \quad \overline{F} \in \mathcal{R}(-\alpha-1)\,, \\ e^{-(x+y)} & if \quad F \in \mathrm{MDA}(\Lambda)\,. \end{cases}$$

(Here α is a positive parameter, the latter case when $F \in \mathrm{MDA}(\Lambda)$ being considered as the case $\alpha = \infty$.) □

Remarks 1) The normalising function $a(\cdot)$ is unique only up to asymptotic equivalence. Since

$$a(u) \sim \int_u^\infty \overline{B}(y)\,dy/\overline{B}(u)\,, \quad u \to \infty\,,$$

the rhs here is also a possible normalising function.
2) Extreme value theory is the foundation of this result: recall first that $\overline{F} \in \mathcal{R}(-(\alpha+1))$ is equivalent to $F \in \mathrm{MDA}(\Phi_{\alpha+1})$, and hence to $F_I \in \mathrm{MDA}(\Phi_\alpha)$ by Karamata's theorem. Furthermore, $F \in \mathrm{MDA}(\Lambda) \cap \mathcal{S}^*$ implies $F_I \in \mathrm{MDA}(\Lambda) \cap \mathcal{S}$. Extreme value theory then provides the form of G_α as the only possible

limit df for the excess distribution (Balkema and de Haan (1974)). G_α is called a *generalised Pareto distribution*. The normalising function $a(u)$ tends to infinity as $u \to \infty$. For $\overline{F} \in \mathcal{R}(-(\alpha+1))$ Karamata's theorem gives $a(u) \sim u/\alpha$. For $F \in \mathrm{MDA}(\Lambda)$ this is Lemma 2.1 in Goldie and Resnick (1988).
3) The limit result for $(S_{t\tau(u)})$ given in Theorem 6.2 substantiates the assertion that the process (S_t) evolves typically up to time $\tau(u)$. □

Theorem 6.2 applies in particular to the models in Section 2. Indeed, stronger results (employing total variation distance) can be obtained for these examples since (S_t) is a downwards skip–free Markov process (with paths having only upwards jumps and deterministic downwards movements). We conclude with two special results which answered questions that were open for some time, but refer to Asmussen and Klüppelberg (1996, 1997) for details.

Example 6.3. (Finite time ruin probability)
Define the ruin probability before time T by

$$\Psi(u,T) = P(\tau(u) \le T)\,.$$

From the limit result on $\tau(u)$ given in Theorem 6.2 one finds the following: if $\overline{F} \in \mathcal{R}(-(\alpha+1))$ for some $\alpha > 0$ then

$$\lim_{u\to\infty} \frac{\psi(u,uT)}{\psi(u)} = 1 - \big(1 + (1-\rho)T\big)^{-\alpha}\,,$$

and if $F \in \mathrm{MDA}(\Lambda) \cap \mathcal{S}^*$ then

$$\lim_{u\to\infty} \frac{\psi(u,a(u)T)}{\psi(u)} = 1 - e^{-(1-\rho)T}\,.$$

□

Example 6.4. (Excursions of the workload process of an M/G/1–queue)
Let $P^{(u)}$ denote the distribution of the doubly infinite version $(V_t)_{t\in\mathbb{R}}$ of the workload process, for which a stationary excursion above level u starts at time 0. Assume that $\rho = \lambda\mu < 1$, λ being the arrival rate and μ the mean service time, and let π denote the stationary distribution of (V_t) and F_I the stationary excess distribution. By the Markov property, the existence of a limit law for an excursion is equivalent to $P^{(u)}$–convergence of $V_0 - u$.

In the light–tailed case the excess $V_0 - u$ and hence the whole excursion has a limit as $u \to \infty$. However, if $F_I \in \mathcal{S}$, this limit is defective; more precisely,

$$\lim_{u\to\infty} P^{(u)}(V_0 - u \le y) = \rho F_I(y)\,, \quad y > 0\,.$$

Furthermore, if $F \in \mathcal{L}$,

$$P^{(u)}(V_0 > u + y \mid V_{0-} = z) = \frac{\overline{F}(u+y-z)}{\overline{F}(u-z)} \to 1\,, \quad u \to \infty\,,$$

for all $y,\, z > 0$. So for $F^* \in \mathcal{S}$ (then $F \in \mathcal{L}$ and $F_I \in \mathcal{S}$) there are two types of excursions.

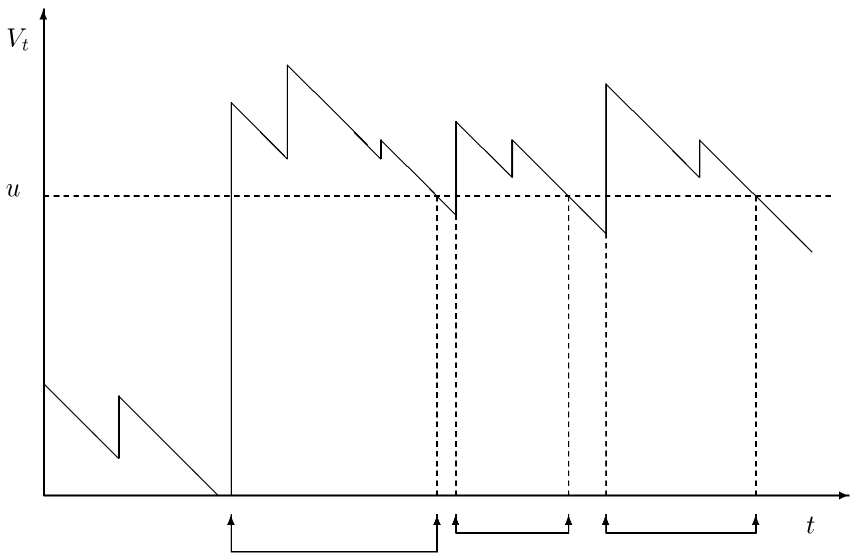

Figure 6.5. *Sample path of the workload process showing three high–level excursions.*

(1) With probability $1-\rho$ the excursion starts from $V_{0-} = O(1)$ and the excess is huge. There is one indicated in Figure 6.5, the first one.
(2) With probability ρ the excursion starts from pre–level $u - V_{0-} = O(1)$ and the excess $V_0 - u$ has df F_I. There are two indicated in Figure 6.5, namely the last two.
This can be interpreted as that the process evolves in a typical way, with negative drift, until a very large service time causes an excursion. After the overshoot the drift takes over again, but there may be some smaller excursions on the way down which can be considered as aftershocks caused mainly by the preceding large service time. □

7. Concluding remarks

Recent interest in subexponential distributions concentrates mainly on relations between heavy tails, long range dependence and self–similarity.

If for instance the input stream of a GI/M/1 queue exhibits long range dependence, then the stationary queue size and the stationary waiting time distributions are each heavy–tailed; see Resnick and Samorodnitsky (1996). They describe a special model for a long range dependent arrival stream (the inter–arrival times are stationary with a special long range dependence structure) and derive bounds for the tails of the stationary queue size and the stationary waiting time distributions. Their results are by no means as explicit as the results presented in this paper, but they derive bounds for distribution tails. As stated by the authors, “one simply needs to better understand the behaviour of queues with long range dependent input”.

An on/off model for packet transmission has been described in Willinger, Taqqu, Sherman and Wilson (1995). This model explains the slow rate of decay of the covariance function of the data, which is an indicator for long range dependence. An interesting review paper with updated references is Resnick (1996).

Vesilo and Daley (1996) consider long range dependence of point processes with queueing examples. They show for instance that certain regularly varying inter–arrival times or service times lead to a long–range dependent departure process in a queueing model.

On/off models with subexponential on–periods have been considered by Jelenkovič and Lazar (1996), using mainly regular variation arguments. There is also a paper by Heath, Resnick and Samorodnitsky (1996) on this topic.

Acknowledgments

The authors are most grateful to Thomas Mikosch for a careful reading and for suggesting many improvements and clarifications. Any remaining obscurities and infelicities are the responsibility of the authors.

CMG thanks the Department of Mathematics of the Johannes Gutenberg University, Mainz, for hospitality and support.

CK thanks the Center for Applied Probability at Columbia University for the invitation to speak at the Applied Probability Day in April 1996. This review paper is partly based on her seminar at this occasion.

References

[1] Asmussen, S. (1982) Conditioned limit theorems relating a random walk to its associate, with applications to risk reserve processes and the GI/G/1 queue. *Adv. Appl. Prob.* **14**, 143–170.

[2] Asmussen, S. (1996) Rare events in the presence of heavy tails. Preprint. Lund University.

[3] Asmussen, S. (1996) Subexponential asymptotics for stochastic processes: extremal behaviour, stationary distributions and first passage probabilities. Preprint, University of Lund.

[4] Asmussen, S. (1997) *Ruin Probabilities.* World Scientific, Singapore. To appear.

[5] Asmussen, S., Fløe Henriksen, L. and Klüppelberg, C. (1994) Large claims approximations for risk processes in a Markovian environment. *Stoch. Proc. Appl.* **54**, 29–43.

[6] Asmussen, S. and Højgaard, B. (1995) Ruin probability approximations for Markov–modulated risk processes with heavy tails. Preprint, University of Aalborg.

[7] Asmussen, S. and Klüppelberg, C. (1996) Large deviation results for subexponential tails, with applications to insurance risk. *Stoch. Proc. Appl.* To appear.

[8] Asmussen, S. and Klüppelberg, C. (1997) Stationary M/G/1 excursions in the presence of heavy tails. *J. Appl. Probab.* **34**. To appear.

[9] Athreya, K.B. and Ney, P.E. (1972) *Branching Processes.* Springer, Berlin.

[10] Balkema, A.A. and de Haan, L. (1972) Residual life–time at great age. *Ann. Probab.* **96**, 792–804.

[11] Bertoin, J. and Doney, R.A. (1996) Some asymptotic results for transient random walks. *Adv. Appl. Probab.* **28**, 207–226.

[12] Bingham, N.H., Goldie, C.M. and Teugels, J.L. (1989) *Regular Variation*, revised paperback ed. Cambridge University Press, Cambridge.

[13] Chistyakov, V.P (1964) A theorem on sums of independent positive random variables and its applications to branching random processes. *Theory Probab. Appl.* **9**, 640–648.

[14] Chover, J., Ney, P. and Wainger, S. (1973) Functions of probability measures. *J. Analyse Math.* **26**, 255–302.

[15] Cline, D.B.H. (1986) Convolution tails, product tails and domains of attraction. *Probab. Theory Related Fields* **72**, 529–557.

[16] Cline, D.B.H. (1987) Convolutions of distributions with exponential and subexponential tails. *J. Austral. Math. Soc. Ser. A* **43**, 347–365.

[17] Davis, R.A. and Resnick, S.I. (1985) Limit theory for moving averages of random variables with regularly varying tail probabilities. *Ann. Probab.* **13**, 179–195.

[18] Davis, R.A. and Resnick, S.I. (1988) Extremes of moving averages of random variables from the domain of attraction of the double exponential distribution. *Stoch. Proc. Appl.* **30**, 41–68.

[19] 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.

[20] Embrechts, P. and Goldie, C.M. (1982) On convolution tails. *Stoch. Proc. Appl.* **13**, 263–278.

[21] Embrechts, P., Goldie, C.M. and Veraverbeke, N. (1979) Subexponentiality and infinite divisibility. *Z. Wahrscheinlichkeitstheorie verw. Geb.* **49**, 335–347.

[22] Embrechts, P., Klüppelberg, C. and Mikosch, T. (1997) *Modelling Extremal Events for Insurance and Finance.* Springer, Heidelberg. To appear.

[23] Embrechts, P. and Samorodnitsky, G. (1995) Sample quantiles of heavy tailed stochastic processes. *Stoch. Proc. Appl.* **59**, 217–233.

[24] Embrechts, P. and Veraverbeke, N. (1982) Estimates for the probability of ruin with special emphasis on the possibility of large claims. *Insurance: Math. Econom.* **1**, 55–72.

[25] Feller, W. (1971) *An Introduction to Probability Theory and Its Applications II.* Wiley, New York.

[26] Gantert, N. (1996) Large deviations for a heavy–tailed mixing sequence. Preprint, Technical University Berlin.

[27] Goldie, C.M. (1978) Subexponential distributions and dominated–variation tails. *J. Appl. Probab.* **15**, 440–442.

[28] Goldie, C.M. and Resnick, S.I. (1988) Distributions that are both subexponential and in the domain of attraction of an extreme–value distribution. *Adv. Appl. Probab.* **20**, 706–718.

[29] Grandell, J. (1991) *Aspects of Risk Theory.* Springer, Berlin.

[30] Heath, D., Resnick, S., and Samorodnitsky, G. (1996) Heavy tails and long range dependence in on/off processes and associated fluid models. Preprint, Cornell University.

[31] Heyde, C.C. (1967a) A contribution to the theory of large deviations for sums of independent random variables. *Z. Wahrscheinlichkeitstheorie verw. Geb.* **7**, 303–308.

[32] Heyde, C.C. (1967b) On large deviation problems for sums of random variables not attracted to the normal law. *Ann. Math. Statist.* **38**, 1575–1578.

[33] Heyde, C.C. (1968) On large deviation probabilities in the case of attraction to a nonnormal stable law *Sankhya Ser. A* **30**, 253–258.

[34] Jelenkovič, P.R. and Lazar, A.A. (1996) Subexponential asymptotics of a Markov–modulated G/G/1 queue. Preprint, Columbia University, New York.

[35] Jelenkovič, P.R. and Lazar, A.A. (1996) Multiplexing on–off sources with subexponential on periods. CTR Technical Report 457-96-23, Columbia University, New York.

[36] Klüppelberg, C. (1988) Subexponential distributions and integrated tails. *J. Appl. Probab.* **25**, 132–141.

[37] Klüppelberg, C. (1989a) Subexponential distributions and characterisations of related classes. *Probab. Theory Related Fields* **82**, 259–269.

[38] Klüppelberg, C. (1989b) Estimation of ruin probabilities by means of hazard rates. *Insurance: Math. Econom.* **8**, 279–285.

[39] Klüppelberg, C. (1990) Asymptotic ordering of distribution functions on convolution semigroup. *Semigroup Forum* **40**, 77–92.

[40] Klüppelberg, C. and Mikosch, T. (1997) Large deviations of heavy–tailed random sums with applications to insurance and finance. *J. Appl. Probab.* **34**. To appear.

[41] Klüppelberg, C. and Stadtmüller, U. (1996) Ruin probabilities in the presence of heavy–tails and interest rates. *Scand. Actuar. J.* To appear.

[42] Leslie, J.R. (1989) On the non–closure under convolution of the subexponential family. *J. Appl. Probab.* **26**, 58–66.

[43] Mikosch, T. and Nagaev, A.V. (1996) Large deviations for heavy–tailed distributions with applications to insurance. Preprint, University of Groningen.

[44] Nagaev, A.V. (1969a) Limit theorems for large deviations where Cramér's conditions are violated (in Russian). *Izv. Akad. Nauk UzSSR Ser. Fiz.–Mat. Nauk* **6**, 17–22.

[45] Nagaev, A.V. (1969b) Integral limit theorems for large deviations when Cramér's condition is not fulfilled I,II. *Theory Probab. Appl.* **14**, 51–64 and 193–208.

[46] Nagaev, S.V. (1979) Large deviations of sums of independent random variables. *Ann. Probab.* **7**, 745–789.

[47] Pakes, A.G. (1975) On the tails of waiting time distributions. *J. Appl. Probab.* **12**, 555–564.

[48] Pinelis, I.F. (1985) On the asymptotic equivalence of probabilities of large deviations for sums and maxima of independent random variables (in Russian). In: *Limit Theorems in Probability Theory*, pp. 144–173. Trudy Inst. Math. **5**. Nauka, Novosibirsk.

[49] Pitman, E.J.G. (1980) Subexponential distribution functions. *J. Austral. Math. Soc. Ser. A* **29**, 337–347.

[50] Resnick, S.I. (1987) *Extreme Values, Regular Variation, and Point Processes.* Springer, New York.

[51] Resnick, S.I. (1996) Heavy tail modelling and teletraffic data. Preprint. Cornell University.

[52] Resnick, S.I. and Samorodnitsky, G. (1996) Performance decay in a single server exponential queueing model with long range dependence. Operations Research, to appear.

[53] Rosinski, J. and Samorodnitsky, G. (1993) Distributions of subadditive functionals of sample paths of infinitely divisible processes. *Ann. Probab.* **21**, 996–1014.

[54] Rozovskii, L.V. (1993) Probabilities of large deviations on the whole axis. *Theory Probab. Appl.* **38**, 53–79.

[55] Smith, W.L. (1972) On the tails of queueing time distributions. Mimeo Series No. 830, Dept. of Statistics, University of North Carolina, Chapel Hill.

[56] Teugels, J.L. (1975) The class of subexponential distributions. *Ann. Probab.* **3**, 1001–1011.

[57] Vesilo, R. and Daley, D.J. (1996) Long range dependence of point processes, with queueing examples. Preprint.

[58] Veraverbeke, N. (1977) Asymptotic behaviour of Wiener–Hopf factors of a random walk. *Stoch. Proc. Appl.* **5**, 27–37.

[59] Vinogradov, V. (1994) *Refined Large Deviation Limit Theorems.* Longman, Harlow (England).

[60] Willekens, E. (1986) *Hogere Orde Theorie voor Subexponentiële Verdelingen.* Ph.D. Thesis, Katholieke Universiteit Leuven.

[61] Willinger, W., Taqqu, M., Sherman, R. and Wilson, D. (1995) Self–similarity in high–speed packet traffic: analysis and modelling of ethernet traffic measurements. *Statistical Science* **10**, 67–85.

Charles M. Goldie
School of Mathematical Sciences
University of Sussex
Brighton BN1 9QH, England
email: C.M.Goldie@sussex.ac.uk

Claudia Klüppelberg
Department of Mathematics
Johannes Gutenberg–Universität
D–55099 Mainz, Germany
email: cklu@mathematik.uni-mainz.de