

Large deviations of means of heavy-tailed random variables with finite moments of all orders

Jaakko Lehtomaa¹

¹Department of Mathematics and Statistics, University of Helsinki, P.O. Box 68, 00014, Helsinki, Finland, email: jaakko.lehtomaa@helsinki.fi

Abstract

Logarithmic asymptotics of the mean process $\{S_n/n\}$ are investigated in the presence of heavy-tailed increments. As a consequence, full large deviations principle for means is obtained when the hazard function of an increment is regularly varying with index $\alpha \in (0, 1)$. This class includes all stretched exponential distributions. Thus the previous research of Gantert et al is extended. Furthermore, the presented proofs are more transparent than the techniques used by Nagaev et al. In addition, the novel approach is compatible with other common classes of distributions, e.g. those of lognormal-type.

Keywords: Logarithmic asymptotics; Regular variation; Large deviations; Heavy-tailed; Stretched exponential

MSC classification (2010): 60E05; 60E15; 60F10; 62G32

1 Introduction

Let X be a generic member, called *increment*, of the independent and identically distributed, abbreviated i.i.d., sequence $(X_i) := (X_i)_{i=1}^\infty$ of real valued random variables defined on (Ω, \mathcal{F}, P) . We set $S_n = X_1 + \dots + X_n$ and study the logarithmic asymptotics of $P(S_n > an)$ for $a > E(X)$. More specifically, the asymptotic behaviour of the quantity

$$\frac{-\log P(S_n > an)}{g(n)} \tag{1.1}$$

is studied in a large class of increment distributions under different choices of the *scaling function* g , as $n \rightarrow \infty$. The function g will be the *hazard function*

$$R_X(x) = R(x) = -\log P(X > x) \tag{1.2}$$

or an approximation of R .

Quantity (1.1) is extensively studied in a vast number of papers and books. However, the bulk of the research is focused into two major categories of increments: *light-tailed* and extremely *heavy-tailed*. Recall that X is light-tailed if $E(e^{sX}) < \infty$ for some $s > 0$ and heavy-tailed otherwise. Between the distributions with polynomial or heavier tails and light-tailed distributions lies a rich but relatively little understood set of distributions. They share the following two properties:

I Members are heavy-tailed

II Members have finite moments of all orders, i.e. $E((X^+)^{\alpha}) < \infty$, for all $\alpha > 0$.

This paper aims to narrow the research gap between light- and seriously heavy-tailed distributions. Specifically, the paper studies the asymptotics of the quantity (1.1) when the increment of the random walk S_n satisfies Requirements I-II.

Perhaps the best known subclass of distributions satisfying I-II is the class of *stretched exponential distributions*. A random variable X has a stretched exponential tail if

$$a_1(x)e^{-b(x)x^{\alpha}} \leq P(X > x) \leq a_2(x)e^{-b(x)x^{\alpha}} \quad (1.3)$$

holds for all large enough x , where $\alpha \in (0, 1)$ and a_1, a_2, b are slowly varying functions [12, 14]. Furthermore, Lognormal and stretched exponential distributions fulfil Requirements I-II. Random walks with increments from these and other distributions are investigated in [4, 17, 18, 19]. See also Section 6 of [3] along with its ample references.

In [12], the quantity (1.1) with stretched exponential increments is studied using the scale $g(x) = b(x)x^{\alpha}$. Under mild assumptions a limit is shown to exist in (1.1), which gives the full large deviations principle. In this paper, the previous analysis is extended by showing that the limit actually exists in much larger class of distributions. This can be done, because the method offers new insight about the interaction between the scale g and the rate function associated with the large deviations principle of the process $\{S_n/n\}$.

To begin, the non-trivial scaling functions for the upper and lower limits of (1.1) are identified. As in [9], we call the limiting behaviour non-trivial if there exist at least two separate values of a in (1.1) for which the limiting quantity is not 0 or ∞ . It is shown that the hazard function R of (1.2) is a non-trivial scale for the upper limit of (1.1). Somewhat surprisingly, the corresponding lower limit requires a different scale.

The key ingredient in the required analysis turns out to be the concept of *natural scale*. Every heavy-tailed random variable admits a concave, extended regularly varying approximation \underline{R} of R . Such a function is called a natural scale. It is important, because it produces a non-trivial lower limit in (1.1). The existence of natural scale can be used, not only to simplify existing techniques, but to produce completely new approaches under general, but previously unexplored assumptions.

Full large deviations principle for means can be obtained, if the functions R and \underline{R} are close to each other and a limit exists in (1.1). Specifically, it is the

case when R is regularly varying with index $\alpha \in (0, 1)$. This expands the known results for stretched exponential increments. Finally, a converse result is proven: under general assumptions the existence of a limit in (1.1) with scale $g = R$ implies regular variation of R with index $\alpha \in [0, 1]$.

1.1 Assumptions and Notation

A) Basic assumptions:

- i) $E((X^+)^s) < \infty$ for all $s > 0$.
- ii) $R(x) = o(x)$, as $x \rightarrow \infty$.
- iii) $E(X) = \mu = 0$.
- iv) $X^- = \max(0, -X)$ is light-tailed.

Assumptions Ai-Aii concern the growth rate of R . In fact, Ai has an alternative form $\log x = o(R(x))$, as $x \rightarrow \infty$. This is an immediate consequence of the well known relation

$$\sup \{s \geq 0 : E((X^+)^s) < \infty\} = \liminf_{x \rightarrow \infty} \frac{R(x)}{\log x}$$

found from e.g. [5]. Assumption Aiv could be relaxed in some occasions if one wanted to study the left tail of the sum S_n . However, since the results deal with the right tail of S_n , Aiv is assumed in order to make assumptions uniform.

Results of [2] are used extensively. Properties of regularly varying functions are recalled as necessary. Only the following general definitions are explicitly stated.

Definition 1.1. A positive function f is *slowly varying*, if $f(yx)/f(x) \rightarrow 1$, as $x \rightarrow \infty$ for all $y > 0$. More generally, a function f is *regularly varying* with index $\alpha \in \mathbb{R}$, denoted $f \in \mathcal{R}_\alpha$, if $f(yx)/f(x) \rightarrow y^\alpha$, as $x \rightarrow \infty$ for all $y > 0$. Following Definition (2.0.6) of [2], a function f is *extended regularly varying* if

$$y^d \leq \liminf_{x \rightarrow \infty} \frac{f(yx)}{f(x)} \leq \limsup_{x \rightarrow \infty} \frac{f(yx)}{f(x)} \leq y^c$$

for all $y \geq 1$ and some constants $c, d \in \mathbb{R}$.

The paper occasionally uses standard notations $f \sim g$, $f = o(g)$ and $f = O(g)$ as defined in [2]. All asymptotic relations concern the limit $x \rightarrow \infty$ or $n \rightarrow \infty$ unless otherwise stated, where $x \in \mathbb{R}$ and $n \in \mathbb{N} = \{1, 2, 3, \dots\}$. Any property that holds for *large enough* arguments means that there exist a number y_0 such that the property holds in the set $[y_0, \infty)$. Similarly, we say that a property holds *eventually* if it holds for large enough arguments. The symbol $X := Y$ means that *the definition of X is Y*. The positive part is denoted by $x^+ := \max(0, x)$. Finally, we set $\lceil x \rceil := \inf\{n \in \mathbb{Z} : x \leq n\}$ and $\lfloor x \rfloor := \sup\{n \in \mathbb{Z} : x \geq n\}$.

2 Results

A hazard function corresponding to a heavy-tailed distribution admits a concave approximation called *natural scale*. It is a smooth function that approximates R from below in an asymptotic sense. Related ideas have been previously presented in [13, 7]. However, the concavity seems to be the property that truly makes the concept of natural scale operational.

2.1 Preliminary results

A concave function \underline{R} satisfying properties 1-4 of Theorem 2.1 is called a natural scale of X . Note that the concept of natural scale does not fix a unique function. A good initial guess for finding some natural scale in practice is $R(x) = -\log P(X > x)$ itself or a dominating component of R .

The name natural scale comes from Connection (2.3). Since a random variable X is heavy-tailed if it has no finite exponential moments, it is natural to search for a renormalising function g such that exponential moments of $g(X)$ are finite again. Clearly, this is possible for any growing function g that grows slowly enough with respect to R . However, the natural scale is, in a sense, the fastest growing function that still makes the exponential moments of $\underline{R}(X)$ finite.

Theorem 2.1. *Let R be the hazard function of X . Suppose X is heavy-tailed i.e.*

$$\liminf_{x \rightarrow \infty} \frac{R(x)}{x} = 0. \quad (2.1)$$

Then there exists a strictly increasing concave function $\underline{R}: [0, \infty) \rightarrow [0, \infty)$ satisfying:

1. $\underline{R} \geq 0$
2. $\underline{R}(0) = 0$
3. $\underline{R}(x) = o(x)$, as $x \rightarrow \infty$ and
- 4.

$$\liminf_{x \rightarrow \infty} \frac{R(x)}{\underline{R}(x)} = 1. \quad (2.2)$$

Condition (2.2) is equivalent to

$$\sup \{s \geq 0 : E(e^{s\underline{R}(X)}) < \infty\} = 1. \quad (2.3)$$

The fact heavy-tailedness is equivalent with (2.1) is shown in Theorem 2.6 of [10]. Only heavy-tailed variables can have a function fulfilling Properties 3-4. Theorem 2.1 shows that it is always the case. Properties 1-2 can be obtained by a suitable linear interpolation if Properties 3-4 hold. Properties 1-2 are nevertheless demanded. They ensure, with concavity, the subadditivity of \underline{R} . The fineness of exponential moments in (2.3) is determined by the \liminf in (2.2). The corresponding \limsup does not have a similar effect to integrability.

2.2 Main Results

Theorem 2.2. *Suppose X satisfies Assumptions Ai-Aiv of Section 1.1.*

In general

$$\limsup_{n \rightarrow \infty} \frac{-\log P(S_n > n)}{R(n)} = 1. \quad (2.4)$$

If \underline{R} is a natural scale of X such that $\log x = o(\underline{R}(x))$, then

$$\liminf_{n \rightarrow \infty} \frac{-\log P(S_n > n)}{\underline{R}(n)} = 1. \quad (2.5)$$

Remark 2.3. Note that natural scales are not involved in (2.4), although its proof makes extensive use of their properties. This indicates that the concept of natural scale enables new proof techniques. The following results concern mainly regularly varying hazard functions with $\alpha \in (0, 1)$. However, Theorem 2.2 is suitable for other distribution types as well, e.g. for lognormal distributions.

Remark 2.4. If $\underline{R} \sim R$, it holds that $\log P(S_n > n) \sim \log P(X > n)$, as $n \rightarrow \infty$. At the level of logarithms one large increment dominates the tail-behaviour of S_n . This can be viewed as a manifestation of the *principle of a single big jump*, which is widely observed in heavy-tailed systems, cf. [16].

Remark 2.5. Result of Theorem 2.2 is not valid for light-tailed variables or variables that have infinite moment of some order. The light-tailed case is covered by Cramér's theorem where the scaling g always needs to be chosen as $g(n) = n$. If X does not have finite moments of all orders, but still has expectation, the natural scale can be chosen to be $\beta \log x$ for some $\beta \geq 1$ and x large enough. In this case the right hand side of (2.4) and (2.5) may not be equal to 1, but changed by a constant depending on β .

Corollary 2.6. *Let $a > 0$.*

Suppose assumptions to obtain (2.4) hold. Then

$$\limsup_{n \rightarrow \infty} \frac{-\log P(S_n > an)}{R(n)} \leq \limsup_{n \rightarrow \infty} \frac{R(an)}{R(n)}. \quad (2.6)$$

Suppose assumptions to obtain (2.5) hold. Then

$$\liminf_{n \rightarrow \infty} \frac{-\log P(S_n > an)}{\underline{R}(n)} \geq \liminf_{n \rightarrow \infty} \frac{\underline{R}(an)}{\underline{R}(n)}. \quad (2.7)$$

Lemma 2.7 shows that distributions defined by (1.3) have a regularly varying hazard function R . Since Theorem 2.8 covers regularly varying hazard functions, specifically the case of stretched exponential distributions is included.

Lemma 2.7. *Suppose X satisfies (1.3), i.e. X is a random variable with a stretched exponential distribution.*

Then $R \in \mathcal{R}_\alpha$ with $\alpha \in (0, 1)$, where \mathcal{R}_α is defined as in Definition 1.1

Theorem 2.8. *Suppose Assumptions Aiii-Aiv hold and $R \in \mathcal{R}_\alpha$ with $\alpha \in (0, 1)$ so that also Ai-Aii hold. Then*

$$\lim_{n \rightarrow \infty} \frac{-\log P(S_n > an)}{R(n)} = \begin{cases} 0 & : a < 0 \\ a^\alpha & : a \geq 0. \end{cases} \quad (2.8)$$

Furthermore, the process $\{S_n/n\}$ satisfies full large deviations principle with good rate function

$$I(x) = \begin{cases} \infty & : x < 0 \\ x^\alpha & : x \geq 0 \end{cases} \quad (2.9)$$

and scale R , i.e.

$$-\inf_{y \in A^c} I(y) \leq \liminf_{n \rightarrow \infty} \frac{\log P(S_n/n \in A)}{R(n)} \leq \limsup_{n \rightarrow \infty} \frac{\log P(S_n/n \in A)}{R(n)} \leq -\inf_{y \in \text{cl}(A)} I(y) \quad (2.10)$$

for all Borel sets $A \subset \mathbb{R}$.

In Theorem 2.8 the rate function I is concave instead of convex on the positive half line. This is fundamentally different from what is observed in light-tailed theory. See e.g. [1] for other non-standard rates and [20, 6] for classical results. Theorem 2.9 gives a partial converse of Theorem 2.8. Under general assumptions the necessary limit can not exist unless the hazard function is regularly varying. If more smoothness is assumed from R , e.g. if it is almost Weibullian, then the result can be proved in a more straightforward way via estimation arguments or application of the results of [14].

Theorem 2.9. *Suppose R satisfies Assumptions Ai-Aiv. Assume further that $\underline{R} \sim R$ for some natural scale \underline{R} of R . If*

$$\liminf_{n \rightarrow \infty} \frac{-\log P(S_n > an)}{R(n)} = \limsup_{n \rightarrow \infty} \frac{-\log P(S_n > an)}{R(n)}$$

for all a in some set of positive Lebesgue measure $\mathcal{A} \subset (0, \infty)$, then $R \in \mathcal{R}_\alpha$ for some $\alpha \in [0, 1]$.

3 Comments and Counterexamples

Example 3.1 demonstrates why the scaling $g = \underline{R}$ must be used instead of $g = R$ in (2.5). In addition, some regularity properties of the central functions are provided.

There exist increasing concave functions which are not regularly varying. See [15], page 512 for explicit examples. Since such a function could be a hazard function or a natural scale of a random variable, the function \underline{R} may have oscillations that ruin regular variation. Nevertheless, some degree of regularity is available directly from the definition of the natural scale, namely the extended regular variation property and subadditivity. By Lemma 4.4 it is immediate that for $\lambda \geq 1$

$$1 \leq \liminf_{x \rightarrow \infty} \frac{R(\lambda x)}{\underline{R}(x)} \leq \limsup_{x \rightarrow \infty} \frac{R(\lambda x)}{\underline{R}(x)} \leq \lambda,$$

so that $\underline{R} \in ERV$ with $d = 0$ and $c = 1$. Subadditivity property $\underline{R}(a+b) \leq \underline{R}(a) + \underline{R}(b)$ for $a, b \geq 0$ is clear from the definition of concavity and Property 2 of Theorem 2.1.

The function

$$g(a) := \liminf_{n \rightarrow \infty} \frac{\underline{R}(an)}{\underline{R}(n)},$$

where $a \geq 0$, is clearly a concave and increasing function. This shows that the concavity of \underline{R} induces some regularity to the lower limit of (2.7).

Example 3.1. The function R cannot generally replace the function \underline{R} in (2.5). To see this, let h_1 and h_2 be concave, strictly increasing functions such that $h_2(x) = o(x)$ and $\log x = o(h_1(x))$, as $x \rightarrow \infty$. Assume further that $h_i(0) = 0$, $i = 1, 2$ and

$$h_2(x) > (1 + \eta)h_1(x)$$

for some $\eta > 0$ and all $x > 0$. Now, one can construct the hazard function R of a random variable X using the functions h_1 and h_2 so that

$$\liminf_{n \rightarrow \infty} \frac{-\log P(S_n > n)}{R(n)} < 1. \quad (3.1)$$

We define the right tail of X by defining its hazard function R .

First, define the members of the sequence $(y_n) \subset \mathbb{R}$ by setting $y_1 = 1$ and demanding that $h_1(y_{n+1}) = h_2(y_n)$ for all $n \in \mathbb{N}$. The function R is constructed recursively from the sequence (y_n) by setting $R(x) = h_2(y_n)$, $x \in [y_n, y_{n+1})$, $n \in \mathbb{N}$. The function R is a hazard function and fixes the right tail of the random variable X in set $[1, \infty)$. To fix the full distribution of X , set a single point mass on the negative half axis so that $E(X) = 0$. The function R has a jump at every point of the sequence (y_n) , which is real-valued. Set $z_n := \lceil y_n \rceil$ to obtain an integer-valued sequence (z_n) . Now, for a fixed n , it holds that $P(S_{z_n} > z_n) \geq P(X > z_n - 2)P(X > 2)P(S_{z_n-2} \geq 0)$.

Application of the central limit theorem shows that the last factor tends to $1/2$, as $n \rightarrow \infty$. Therefore

$$\begin{aligned} \liminf_{n \rightarrow \infty} \frac{-\log P(S_n > n)}{R(n)} &\leq \liminf_{z_n \rightarrow \infty} \frac{-\log P(S_{z_n} > z_n)}{R(z_n)} \\ &\leq \liminf_{z_n \rightarrow \infty} \frac{-\log P(X > z_n - 2)}{R(z_n)} \\ &= \liminf_{z_n \rightarrow \infty} \frac{R(z_n - 2)}{R(z_n)} \end{aligned}$$

The growth rate of h_2 on intervals of constant size must diminish to zero by assumption $h_2(x) = o(x)$. For large enough n inequality

$$R(z_n - 2)/R(z_n) \leq h_1(z_n)/(h_2(z_n) - 1)$$

holds and thus

$$\liminf_{n \rightarrow \infty} \frac{-\log P(S_n > n)}{R(n)} \leq \frac{1}{1 + \eta} < 1,$$

which is the claim of (3.1). This shows that Formula (2.5) is not generally valid if \underline{R} is replaced by R .

Similarly, an arbitrary choice of \underline{R} cannot be used in the upper limit (2.4). Different choices of \underline{R} may have regions of slow growth ruining the desired asymptotic relation. In conclusion, generally different functions are needed for the upper and lower limits of (2.4) and (2.5), respectively.

4 Proofs

4.1 Proofs of Section 2.1

The following lemma is needed for the proof of Theorem 2.1.

Lemma 4.1. *Suppose X is a random variable. Assume further that function h is strictly increasing, continuous and $h(x) \rightarrow \infty$, as $x \rightarrow \infty$.*

Then $\liminf_{x \rightarrow \infty} R(x)/h(x) = \sup\{s \geq 0 : E(e^{sh(X)}) < \infty\}$.

Proof. It is known from [5] that $\liminf_{x \rightarrow \infty} R_X(x)/x = \sup\{s \geq 0 : E(e^{sX}) < \infty\}$. The claim follows from the previous connection when it is applied to the random variable $h(X)$ while recalling that h is invertible and continuous. \square

Proof of Theorem 2.1. The existence of a natural scale is proved by explicit construction. Suppose (2.1) holds. We will construct a function h that is piecewise linear and whose slope is decreasing. The idea is to define h first for $x_n \in a\mathbb{Z}^+$ and then extend the function to the continuum using linear interpolation.

Let $a, r > 0$ be fixed numbers such that $r > a$. Define the sequence $(x_n)_{n=0}^\infty$ by formula $x_n := an$. Next, to give a precise description of the interpolation procedure, we will find a sequence of slopes $(\epsilon_n)_{n=0}^\infty$ such that $\epsilon_n \downarrow 0$, as $n \rightarrow \infty$. This allows us to define the function h by conditions $h(0) = 0$ and

$$h(x) = h(x_n) + \epsilon_n(x - x_n),$$

for $x \in (x_n, x_{n+1}]$ and $n = 0, 1, 2, \dots$, where the sequence of slopes $(\epsilon_n)_{n=0}^\infty$ is not yet fixed. In order to fix the sequence, set $\epsilon_n = (h(x_{n+1}) - h(x_n))/a$.

Now, it remains to define the value of h at points x_n . In order to account for the case where there is no probability mass near the origin, set $n_0 := \inf\{n \in \mathbb{N} : n > 0, R(x_n) > 2\}$ and define h by $h(x) = x/x_{n_0}$ when $x \in [0, x_{n_0}]$ so that $h(x_{n_0}) = 1$. The rest of the function h is defined inductively.

Suppose the function h is defined in the interval $[0, x_n]$ for some $n \geq n_0$. Set $\epsilon_{n_0-1} := 1/x_{n_0}$ and define the function g_1 by

$$g_1(x) := h(x_n) + \epsilon_{n-1}(x - x_n), \quad x \in [x_n, x_n + r].$$

The function g_1 copies the slope of h from the previous step.

The following rule consisting of two parts is then used to obtain $h(x_{n+1})$:

1° If $g_1(x) \leq R(x)$ for all $x \in [x_n, x_n + r)$, set $h(x_{n+1}) := g_1(x_{n+1})$.

2° If $g_1(x) > R(x)$ for some $x \in [x_n, x_n + r)$, define

$$\hat{\epsilon} := \sup\{\epsilon > 0 : h(x_n) + \epsilon(x - x_n) \leq R(x) \text{ for all } x \in [x_n, x_n + r]\}$$

and

$$g_2(x) := h(x_n) + \hat{\epsilon}(x - x_n), \quad x \in [x_n, x_n + r].$$

The value of function h is then defined as $h(x_{n+1}) = g_2(x_{n+1})$.

The function h defined as above can not be bounded from above by any constant. More specifically, even in the worst case scenario where R is constant in an interval of arbitrary length, the function h keeps a strictly positive, albeit possibly very small, slope. Since $\liminf_{x \rightarrow \infty} R(x)/x = 0$, a linear function of any positive slope below R must eventually cross with R . This being the case, there always exists a subsequence (x_{n_k}) so that x_{n_k} is close to R . Since R grows to infinity, the linear interpolation function h must also do the same.

So, it is possible to choose $\underline{R} = h$. The function h is concave, because $\epsilon_n \downarrow 0$, as $n \rightarrow \infty$. The part $\epsilon_n \downarrow 0$ can be seen from the fact that the function R can not remain indefinitely above any linear function of positive slope. Properties 1-2 also hold by construction. Since

$$h(x) \leq R(x) \tag{4.1}$$

for large enough x and $\epsilon_n \downarrow 0$, Property 3 holds. Since $\epsilon_n \downarrow 0$, it follows that corresponding to any fixed $C > 0$ there exists a subsequence $(x_{n_k})_{k=1}^{\infty}$ such that

$$|R(x_{n_k}) - h(x_{n_k})| < C. \tag{4.2}$$

Inequality (4.2) implies Property 4.

Equivalence of (2.2) and (2.3) follows directly from Lemma 4.1. \square

Remark 4.2. The proof of Theorem 2.1 shows that some additional properties can be obtained for the concave approximation h . Namely, Properties (4.1) and (4.2). In addition, assuming $\log(x) = o(\underline{R}(x))$ in Theorem 2.2 is not a limiting requirement, because such scales can always be chosen under the made assumptions. For example, one can add any non-negative concave function r so that eventually $\log x \leq r(x) \leq R(x)$, $\log x = o(r(x))$ and $r(x) = o(R(x))$ to any natural scale h to obtain a suitable function which is a natural scale after scaling and translation.

4.1.1 Basic Properties of Natural Scales

The following lemma concerns concave functions in general, but will be used mainly with natural scales.

Lemma 4.3. Let $h: [0, \infty) \rightarrow [0, \infty)$ be an increasing concave function with $h(0) \geq 0$. Then for all $x \geq 0$:

$$\begin{cases} h(cx) \geq ch(x), & \text{if } 0 < c < 1 \text{ and} \\ h(cx) \leq ch(x), & \text{if } c > 1. \end{cases}$$

Proof. Assume $x > 0$. Suppose first that $0 < c < 1$. Using the definition of concavity we have $h(cx) = h(cx + (1-c)0) \geq ch(x) + (1-c)h(0) \geq ch(x)$, which proves the first part.

Suppose then that $c > 1$. Set $y := cx$. Applying the first part at point y we get $h((1/c)y) \geq (1/c)h(y)$ so that $h(cx) \leq ch(x)$ holds and we are done. \square

Lemma 4.4. Let \underline{R} be a natural scale. Then

1. $\lambda \underline{R}(x) \leq \underline{R}(\lambda x)$ for $0 \leq \lambda \leq 1$.
2. $\underline{R}(\lambda x) \leq \lambda \underline{R}(x)$ for $\lambda \geq 1$.
3. $\underline{R}(x + \lambda) \sim \underline{R}(x)$ as $x \rightarrow \infty$ for $\lambda \in \mathbb{R}$.

Proof. For 1-2, apply Lemma 4.3 to function \underline{R} . Part 3 is standard since \underline{R} is a subadditive function due to its concavity for $\lambda \geq 0$. The remaining case where $\lambda < 0$ can be reduced to this by setting $y = x + \lambda$ and considering the limit $y \rightarrow \infty$. \square

4.2 General Lemmas for the Proofs of Section 2.2

The proofs of the upper and lower limits require different techniques. For the convenience of the reader these estimates are discussed in different sections. The most technical parts are presented in a series of separate lemmas.

The following lemma proves to be useful for lower estimates concerning probabilities of the form $P(S_n > an)$.

Lemma 4.5. Let $\xi, \xi_1, \xi_2 \dots$ be i.i.d. variables with $E(|\xi|^2) < \infty$ and $E(\xi) = 0$.

Then there exists a constant $c > 0$, depending on the fixed number $a > 0$, such that for all $n = 2, 3, \dots$

$$P(S_n > na) \geq cP(\xi > na)$$

holds.

Proof. Independence implies $P(S_n > na) \geq P(\xi > na)P(S_{n-1} \geq 0)$. Now, application of the central limit theorem yields $P(S_{n-1} \geq 0) \rightarrow 1/2$ in the limit $n \rightarrow \infty$. Choosing $c := \inf_{n \in \mathbb{N}} P(S_n \geq 0)$ proves the claim. \square

The following lemma turns out to be of central importance.

Lemma 4.6. Assume Ai-Aiv of Section 1.1. Suppose \underline{R} is a natural scale of X such that $\log x = o(\underline{R}(x))$ and $0 < \eta < 1$ is fixed.

Then

$$E(e^{(1-\eta)\frac{\underline{R}(n)}{n}X} \mathbf{1}(X \leq n)) \leq 1 + o\left(\frac{\underline{R}(n)}{n}\right), \quad n \rightarrow \infty. \quad (4.3)$$

Proof. Let $A := A(n)$ be a sequence of positive numbers such that $A(n) \rightarrow \infty$ and $(\underline{R}(n)/n)A(n) \rightarrow 0$, as $n \rightarrow \infty$. For any $A > 0$, we write

$$E(e^{(1-\eta)\frac{R(n)}{n}X}\mathbf{1}(X \leq n)) = E(e^{(1-\eta)\frac{R(n)}{n}X}\mathbf{1}(X \leq A)) \quad (4.4)$$

$$+ E(e^{(1-\eta)\frac{R(n)}{n}X}\mathbf{1}(A < X \leq n)) \quad (4.5)$$

and estimate the two terms (4.4) and (4.5) separately.

For (4.4), one can use Taylor expansion and the fact that $\mu = 0$ to obtain

$$E(e^{(1-\eta)\frac{R(n)}{n}X}\mathbf{1}(X \leq A)) \leq P(X \leq A) + o\left(\frac{R(n)}{n}\right). \quad (4.6)$$

For (4.5), integration by parts gives

$$\begin{aligned} & E(e^{(1-\eta)\frac{R(n)}{n}X}\mathbf{1}(A < X \leq n)) \\ &= e^{(1-\eta)\frac{R(n)}{n}A}P(X > A) - e^{(1-\eta)\frac{R(n)}{n}n}P(X > n) \end{aligned} \quad (4.7)$$

$$+ (1-\eta)\frac{R(n)}{n} \int_A^n e^{(1-\eta)\frac{R(n)}{n}y}P(X > y) dy. \quad (4.8)$$

The second term of (4.7) can be omitted and the first term estimated by inequality

$$e^{(1-\eta)\frac{R(n)}{n}A}P(X > A) \leq P(X > A) + O\left(\frac{R(n)}{n}\right)AP(X > A). \quad (4.9)$$

Note that $xP(X > x) \rightarrow 0$, as $x \rightarrow \infty$, because X has finite moments of all orders.

Term (4.8) can be bounded in the following way. We may assume A to be so large that $P(X > y) \leq e^{-(1-\eta/2)\underline{R}(y)}$ holds for all $y \geq A$. Now,

$$(1-\eta)\frac{R(n)}{n} \int_A^n e^{(1-\eta)\frac{R(n)}{n}y}P(X > y) dy \leq \frac{R(n)}{n} \int_A^n e^{(1-\eta)\left(-\underline{R}(y)+\frac{R(n)}{n}y\right)-\frac{\eta}{2}\underline{R}(y)} dy.$$

Concavity of \underline{R} ensures $\underline{R}(y) \geq \frac{R(n)}{n}y$ so that

$$\frac{R(n)}{n} \int_A^n e^{(1-\eta)\left(-\underline{R}(y)+\frac{R(n)}{n}y\right)-\frac{\eta}{2}\underline{R}(y)} dy \leq \frac{R(n)}{n} \int_A^n e^{-\frac{\eta}{2}\underline{R}(y)} dy, \quad (4.10)$$

where the last integral is convergent due to assumption $\log x = o(\underline{R}(x))$.

Combining estimates (4.6), (4.9) and (4.10) ends the proof. \square

The following estimation method is widely known. It is recalled here as a lemma for the reader's convenience.

Lemma 4.7. *Suppose ξ, ξ_1, ξ_2, \dots is an i.i.d. sequence of random variables. Denote $M_n := \max(\xi_1, \xi_2, \dots, \xi_n)$ and $S_n := \xi_1 + \xi_2 + \dots + \xi_n$ as usual. Let $(a_n)_{n=1}^\infty$, $(b_n)_{n=1}^\infty$ and $(c_n)_{n=1}^\infty$ be sequences of positive numbers.*

Then, for any $n \in \mathbb{N}$:

$$P(S_n > b_n, M_n \leq c_n) \leq e^{-a_n b_n} \left(E(e^{a_n \xi} \mathbf{1}(\xi \leq c_n)) \right)^n. \quad (4.11)$$

Proof. Fix n . Since $E(e^{a_n S_n} \mathbf{1}(S_n > b_n, M_n \leq c_n)) \geq e^{a_n b_n} P(S_n > b_n, M_n \leq c_n)$, independence implies the result. \square

4.3 Proof of Theorem 2.2

4.3.1 Lower bound for (2.5)

Somewhat similar approaches of estimation have been used in the proofs of [14] and [19]. However, the main arguments of the present proof are completely new and, in particular, more general.

Writing

$$P(S_n > n) = P(S_n > n, M_n \leq n) + P(S_n > n, M_n > n)$$

and applying Lemma 1.2.15 of [6] gives

$$\begin{aligned} & \limsup_{n \rightarrow \infty} \frac{\log P(S_n > n)}{\underline{R}(n)} \\ = & \max \left(\underbrace{\limsup_{n \rightarrow \infty} \frac{\log P(S_n > n, M_n \leq n)}{\underline{R}(n)}}_{=: I_1}, \underbrace{\limsup_{n \rightarrow \infty} \frac{\log P(S_n > n, M_n > n)}{\underline{R}(n)}}_{=: I_2} \right). \end{aligned}$$

For the quantity I_2 it suffices to make the standard estimate

$$P(S_n > n, M_n > n) \leq P(M_n > n) \leq nP(X_1 > n)$$

so that

$$I_2 \leq \limsup_{n \rightarrow \infty} \frac{\log P(X > n)}{\underline{R}(n)} = -1.$$

To estimate I_1 , fix a small $\eta > 0$. Application of Lemma 4.7 with $a_n = (1 - \eta)\underline{R}(n)/n$, $b_n = n$ and $c_n = n$ gives

$$P(S_n > n, M_n \leq n) \leq e^{-(1-\eta)\underline{R}(n)} \left(E(e^{(1-\eta)\frac{\underline{R}(n)}{n}X} \mathbf{1}(X \leq n)) \right)^n.$$

So,

$$I_1 \leq \underbrace{-(1-\eta) + \limsup_{n \rightarrow \infty} \frac{n \log \left(E(e^{(1-\eta)\frac{\underline{R}(n)}{n}X} \mathbf{1}(X \leq n)) \right)}{\underline{R}(n)}}_{=: I_3}. \quad (4.12)$$

It follows from Lemma 4.6 and inequality $\log x \leq x - 1$ for $x > 0$ that $I_3 \leq 0$.

Combining the previous results concerning Quantities I_1 and I_2 we obtain

$$\limsup_{n \rightarrow \infty} \frac{\log P(S_n > n)}{\underline{R}(n)} \leq -(1-\eta).$$

this proves the claim, because the argument can be repeated with an arbitrarily small $\eta > 0$.

4.3.2 Lower bound for (2.4)

For the lower bound we only need to find suitable subsequences of integers, since we are concerned with a lower bound of a limit superior.

Let $\epsilon \in (0, 1)$ be so small that $(1 + \epsilon)(1 - \eta) = (1 - \eta') < 1$ for some $\eta' \in (0, 1)$. Suppose \underline{R} is a natural scale of X such that $\log x = o(\underline{R}(x))$. Such a scale can be found by Remark 4.2. Equation (2.2) implies the existence of a sequence $(z_n) \subset \mathbb{N}$ such that $z_n \uparrow \infty$,

1° $R(z_n) \leq (1 + \epsilon)\underline{R}(z_n)$ for all $n \in \mathbb{N}$ and

2° $R(x) \geq (1 - \epsilon)\underline{R}(x)$ for all real x such that $x \geq z_1$.

Now

$$\limsup_{n \rightarrow \infty} \frac{-\log P(S_n > n)}{R(n)} \geq \limsup_{n \rightarrow \infty} \frac{-\log P(S_{z_n} > z_n)}{R(z_n)} \geq \liminf_{n \rightarrow \infty} \frac{-\log P(S_{z_n} > z_n)}{R(z_n)}.$$

To prove the claim it then suffices to show that there are subsequences that allow

$$\limsup_{n \rightarrow \infty} \frac{\log P(S_{z_n} > z_n)}{R(z_n)} \tag{4.13}$$

to be bounded from above by a number arbitrarily close to -1 .

We may rewrite calculations of Section 4.3.1 until Equation (4.12) by replacing \underline{R} with R everywhere. So, to prove the claim it is enough to show that

$$\limsup_{n \rightarrow \infty} \frac{z_n \log \left(E \left(e^{(1-\eta)\frac{R(z_n)}{z_n} X} \mathbf{1}(X \leq z_n) \right) \right)}{R(z_n)}$$

can be estimated from above by an arbitrarily small positive number.

This follows from the choice of Sequence (z_n) via Properties 1-2 and Lemma 4.6, since the effect of the light-tailed error from the left tail vanishes if we repeat the argument with arbitrarily small $\epsilon > 0$. The deduction shows (4.13) can be bounded as required thus completing the proof.

4.3.3 Upper bound for (2.4) and (2.5)

Proof of Theorem 2.2. We prove the upper limits for (2.4) and (2.5). The proof is based on application of Lemma 4.5 with $a = 1$. We see that

$$\liminf_{n \rightarrow \infty} \frac{\log P(S_n > n)}{R(n)} \geq \liminf_{n \rightarrow \infty} \frac{\log P(X > n)}{R(n)} = -1,$$

which proves the upper bound of (2.4).

Next, we show the remaining direction of Equation (2.5). Let $\epsilon > 0$. By definition there is a sequence $(x_n) \subset \mathbb{N}$ such that $R(x_n) \leq (1 + \epsilon)\underline{R}(x_n)$ for all n . So,

$$\begin{aligned} \liminf_{n \rightarrow \infty} \frac{-\log P(S_n > n)}{\underline{R}(n)} &\leq \liminf_{x_n \rightarrow \infty} \frac{-\log P(S_{x_n} > x_n)}{\underline{R}(x_n)} \\ &\leq (1 + \epsilon) \liminf_{x_n \rightarrow \infty} \frac{-\log P(S_{x_n} > x_n)}{R(x_n)} \\ &\leq (1 + \epsilon) \limsup_{x_n \rightarrow \infty} \frac{-\log P(S_{x_n} > x_n)}{R(x_n)} \\ &\leq (1 + \epsilon) \limsup_{n \rightarrow \infty} \frac{-\log P(S_n > n)}{R(n)} \leq (1 + \epsilon). \end{aligned}$$

□

4.4 Proofs of the Remaining Results

Proof of Lemma 2.7. Equation (1.3) implies

$$-\log a_1(x) + b(x)x^\alpha \geq R(x) \geq -\log a_2(x) + b(x)x^\alpha.$$

Denote $\hat{R}(x) = b(x)x^\alpha$. Using the basic properties of slowly varying functions, see e.g. Proposition 1.3.6 of [2], we obtain $R \sim \hat{R}$. Thus we get for all $y > 0$, as required,

$$\lim_{x \rightarrow \infty} \frac{R(yx)}{R(x)} = \lim_{x \rightarrow \infty} \frac{\hat{R}(yx)}{\hat{R}(x)} = y^\alpha.$$

□

Proof of Corollary 2.6. To show (2.6), set $\hat{X} := X/a$ and $\hat{S}_n = \hat{X}_1 + \dots + \hat{X}_n$. Clearly, \hat{X} satisfies assumptions Ai-Aiv. In addition,

$$R_{\hat{X}}(n) = -\log P(X/a > n) = -\log P(X > an) = R(an).$$

Therefore

$$\begin{aligned} \limsup_{n \rightarrow \infty} \frac{-\log P(S_n > na)}{R(n)} &= \limsup_{n \rightarrow \infty} \frac{-\log P(\hat{S}_n > n)}{R(n)} \\ &= \limsup_{n \rightarrow \infty} \frac{R_{\hat{X}}(n) - \log P(\hat{S}_n > n)}{R_{\hat{X}}(n)} \\ &\leq \left(\limsup_{n \rightarrow \infty} \frac{R_{\hat{X}}(n)}{R(n)} \right) \underbrace{\left(\limsup_{n \rightarrow \infty} \frac{-\log P(\hat{S}_n > n)}{R_{\hat{X}}(n)} \right)}_{=1} = \limsup_{n \rightarrow \infty} \frac{R(an)}{R(n)}. \end{aligned}$$

Equation (2.7) can be proven by analogous arguments. □

The following two observations become useful in the proof of Proposition 2.8.

Lemma 4.8. *Suppose Aiii-Aiv hold and let R be regularly varying with index $\alpha \in (0, 1)$. Then*

$$\limsup_{n \rightarrow \infty} \frac{-\log P(S_n > na)}{R(n)} = \limsup_{n \rightarrow \infty} \frac{R(an)}{R(n)} = a^\alpha. \quad (4.14)$$

Proof. Define \hat{X} as in the proof of Corollary 2.6. Now,

$$\begin{aligned} & \limsup_{n \rightarrow \infty} \frac{-\log P(S_n > na)}{R(n)} = \limsup_{n \rightarrow \infty} \frac{-\log P(\hat{S}_n > n)}{R(n)} \\ &= \limsup_{n \rightarrow \infty} \frac{R_{\hat{X}}(n) - \log P(\hat{S}_n > n)}{R_{\hat{X}}(n)} \\ &= \left(\lim_{n \rightarrow \infty} \frac{R_{\hat{X}}(n)}{R(n)} \right) \left(\limsup_{n \rightarrow \infty} \frac{-\log P(\hat{S}_n > n)}{R_{\hat{X}}(n)} \right) = a^\alpha. \end{aligned}$$

□

Lemma 4.9. *Suppose $R \in \mathcal{R}_\alpha$, where R is a hazard function of a heavy-tailed random variable X and $\alpha \in (0, 1)$.*

Then X has a natural scale \underline{R} such that $\underline{R} \sim R$.

Proof. The proof is based on a classical result concerning regularly varying functions. Theorem 1.8.3. of [2] states that for $f \in \mathcal{R}_\alpha$, where $\alpha \neq \{0, 1, 2, \dots\}$, there exists a C^∞ function g whose derivatives of all orders are monotone and $g \sim f$.

Applying this theorem with the choice $f = R$ we see that there exists a smooth function $g \sim R$, whose second derivative g'' is a monotone function. To prove the claim it suffices to notice that g must be eventually concave. This can be shown by inspecting the possible cases in a straightforward manner, which proves the claim □

Proof of Theorem 2.8. In the light of Lemma 4.8, it suffices to consider the corresponding lower limit. Let \underline{R} be a natural scale obtained from Lemma 4.9 satisfying $\underline{R} \sim R$. Now, using the fact that $R \sim \underline{R}$ and the lower limit (2.7) we obtain

$$\begin{aligned} \liminf_{n \rightarrow \infty} \frac{-\log P(S_n > an)}{R(n)} &= \liminf_{n \rightarrow \infty} \frac{-\log P(S_n > an)}{\underline{R}(n)} \\ &\geq \liminf_{n \rightarrow \infty} \frac{R(an)}{\underline{R}(n)} \\ &= \liminf_{n \rightarrow \infty} \frac{R(an)}{R(n)} = a^\alpha, \end{aligned}$$

which implies (2.8).

The proof of the large deviation principle is standard from here. It is based on showing the following inequalities:

$$\limsup_{n \rightarrow \infty} \frac{\log P(S_n/n \in F)}{R(n)} \leq - \inf_{y \in F} I(y) \quad (4.15)$$

for all closed sets $F \subset \mathbb{R}$ and

$$\liminf_{n \rightarrow \infty} \frac{\log P(S_n/n \in G)}{R(n)} \geq - \inf_{y \in G} I(y) \quad (4.16)$$

for all open sets $G \subset \mathbb{R}$, from which the claim easily follows. \square

Proof of Theorem 2.9. Let $a \in \mathcal{A}$ and denote $\hat{X} := X/a$. Now functions $R_{\hat{X}}(x) = R(ax)$ and \underline{R} satisfy Assumptions Ai-Aiv. By Theorem 2.2 and the assumption $\underline{R} \sim R$ we deduce $\log P(\hat{S}_n > n) \sim \log P(\hat{X} > n)$, as $n \rightarrow \infty$. Now, since

$$\begin{aligned} \lim_{n \rightarrow \infty} \frac{-\log P(S_n > an)}{R(n)} &= \lim_{n \rightarrow \infty} \frac{R(an) - \log P(\hat{S}_n > n)}{R(an)} \\ &= \lim_{n \rightarrow \infty} \frac{R(an) - \log P(\hat{X} > n)}{R_{\hat{X}}(n)} \\ &= \lim_{n \rightarrow \infty} \frac{R(an)}{R(n)}, \end{aligned}$$

it is clear that the quantity $R(an)/R(n)$ has a limit, as $n \rightarrow \infty$.

The rest of the argument follows from a classical result, Theorem 1.4.3. of [2] in the following way. First, using the fact that $R \sim \underline{R}$,

$$\frac{R(\lambda n)}{R(n)} \sim \frac{\underline{R}(\lambda n)}{\underline{R}(n)}, \quad n \rightarrow \infty$$

for any $\lambda \geq 1$. Furthermore, utilising Part 2 of Lemma 4.4 we get

$$\lim_{\lambda \downarrow 1} \left(\limsup_{x \rightarrow \infty} \frac{R(\lambda x)}{R(x)} \right) \leq 1,$$

which is the only assumption of Theorem 1.4.3. It remains to show equality

$$\lim_{n \rightarrow \infty} \frac{R(an)}{R(n)} = \lim_{x \rightarrow \infty} \frac{R(ax)}{R(x)}, \quad a \in \mathcal{A}. \quad (4.17)$$

The validity of Equation (4.17) follows directly by combining the Part 3 of Lemma 4.4 and the fact that \underline{R} is an increasing function for which $\underline{R} \sim R$. Namely, using inequalities

$$\frac{\underline{R}(a(\lceil x \rceil - 1))}{\underline{R}(\lceil x \rceil)} \leq \frac{\underline{R}(ax)}{\underline{R}(x)} \leq \frac{\underline{R}(a\lceil x \rceil)}{\underline{R}(\lceil x \rceil - 1)},$$

we see that

$$\lim_{n \rightarrow \infty} \frac{\underline{R}(an)}{\underline{R}(n)} = \lim_{x \rightarrow \infty} \frac{\underline{R}(ax)}{\underline{R}(x)}, \quad a \in \mathcal{A},$$

which implies (4.17) via relation $\underline{R} \sim R$. \square

Remark 4.10. The connection (4.17) has been thoroughly studied in [11]. However, in our case Property 3 of Lemma 4.4 is enough for the proof of Theorem 2.9.

Acknowledgements

Comments and suggestions of the referee improved and simplified the text significantly. Most notably, his/her comments made it possible to sharpen the result of Lemma 4.6. This, in turn, enabled simplification of proofs and assumptions needed for Theorem 2.2. Financial support from FDPSS and COIN is acknowledged (Academy of Finland grant number 251170). Special thanks are due to Harri Nyrhinen for his diligent guidance throughout the writing of the paper.

References

- [1] ARCONES, M. (2002). Large and moderate deviations of empirical processes with nonstandard rates *Statist. Probab. Lett.* **57**, 315–326.
- [2] BINGHAM, N., GOLDIE, C. AND TEUGELS, J. (1989). *Regular variation*. Cambridge University Press, New York.
- [3] BOROVKOV, A. AND BOROVKOV, K. (2008). *Asymptotic analysis of random walks: heavy-tailed distributions*. Cambridge University Press, New York.
- [4] BOROVKOV, A. AND MOGUL'SKII, A. (2006). Integro-local and integral theorems for sums of random variables with semiexponential distributions. *Sib. Math. J.* **47**, 990–1026.
- [5] DALEY, D. AND GOLDIE, C. (2006). The moment index of minima (ii). *Statist. Probab. Lett.* **76**, 831–837.
- [6] DEMBO, A. AND ZEITOUNI, O. (2009). *Large deviations techniques and applications*. Springer Verlag, Berlin.
- [7] DENISOV, D. (2006). On the existence of a regularly varying majorant of an integrable monotone function. *Math. Notes* **79**, 129–133.
- [8] DENISOV, D., KORSHUNOV, D. AND FOSS, S. (2008). Lower limits for distribution tails of randomly stopped sums. *Theory Probab. Appl.* **52**, 690–699.
- [9] DUFFY, K. R. AND SAPOZHNIKOV, A. (2008). The large deviation principle for the on-off Weibull sojourn process. *J. Appl. Probab.* **45**, 107–117.
- [10] FOSS, S., KORSHUNOV, D. AND ZACHARY, S. (2013). *An introduction to heavy-tailed and subexponential distributions*. Springer Verlag, New York.
- [11] GALAMBOS, J. AND SENETA, E. (1973). Regularly varying sequences. *Proc. Amer. Math. Soc.* **41**, 110–116.
- [12] GANTERT, N. (1998). Functional Erdos-Renyi laws for semiexponential random variables. *Ann. Probab.* **26**, 1356–1369.
- [13] GANTERT, N. (2000). A note on logarithmic tail asymptotics and mixing. *Statist. Probab. Lett.* **49**, 113–118.
- [14] GANTERT, N., KAVITA, R. AND FRANZ, R. (2014). Large deviations for weighted sums of stretched exponential random variables. *Electron. Commun. Probab.* **19**, 1–14.

- [15] IKSANOV, A. M. AND RÖSLER, U. (2006). Some moment results about the limit of a martingale related to the supercritical branching random walk and perpetuities. *Ukrainian Math. J.* **58**, 505–528.
- [16] LEHTOMAA, J. (2015). Limiting behaviour of constrained sums of two variables and the principle of a single big jump. *Statist. Probab. Lett.* **107**, 157–163.
- [17] MIKOSCH, T. AND NAGAEV, A. (1998). Large deviations of heavy-tailed sums with applications in insurance. *Extremes* **1**, 81–110.
- [18] NAGAEV, S. V. (1979). Large deviations of sums of independent random variables. *Ann. Probab.* **7**, 745–789.
- [19] STABILE, G. AND TORRISI, G. L. (2010). Large deviations of Poisson shot noise processes under heavy tail semi-exponential conditions. *Statist. Probab. Lett.* **80**, 1200–1209.
- [20] VARADHAN, S. R. S. (2008). Large deviations. *Ann. Probab.* **36**, 397–419.