ΠΑΝΕΠΙΣΤΗΜΙΟ ΚΡΗΤΗΣ - ΤΜΗΜΑ ΕΦΑΡΜΟΣΜΕΝΩΝ ΜΑΘΗΜΑΤΙΚΩΝ
Archimedes Center for Modeling, Analysis & Computation
UNIVERSITY OF CRETE - DEPARTMENT OF APPLIED MATHEMATICS
Archimedes Center for Modeling, Analysis & Computation

ACMAC's PrePrint Repository

Conditional distribution of heavy tailed random variables on large deviations of their sum

Inés Armendáriz and Michail Loulakis

Conditional Distribution of Heavy Tailed Random Variables on Large Deviations of their Sum.

Inés Armendáriz[1], Michail Loulakis[2]

ABSTRACT: It is known that large deviations of sums of subexponential random variables are most likely realised by deviations of a single random variable. In this article we give a detailed picture of how subexponential random variables are distributed when a large deviation of the sum is observed.

1 Introduction

Let $X_1, X_2, \ldots$ be a sequence of i.i.d. random variables with common distribution μ defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, and let

$$S_n = X_1 + \cdots + X_n, \qquad n \geq 1.$$

The most classical problem in large deviations is establishing asymptotic expressions for

$$\bar{F}_n(x) := \mathbb{P}\big[S_n > x\big] \tag{1.1}$$

when this quantity converges to zero. The answer depends heavily on the nature of the tails of the distribution μ. When the moment generating function is finite in a neighborhood of the origin (Cramér's condition) Cramér derived asymptotic expressions for $\bar{F}_n(x)$ valid uniformly over different ranges of x-values. These results were later refined by Petrov (cf. [15].) In this case Gibbs conditioning principle provides an answer to how a large deviation of the sum is typically realised: subject to the large deviation, the random variables $\{X_i\}$ become independent in the limit, but their marginal distribution is modified in such a way that the behavior imposed on the sum now becomes typical. In particular, no single random variable becomes excessively large compared to the others.

The situation is totally different when Cramér's condition is violated. It is known since the classical works of Heyde [10] and Nagaev [12] that large deviations of sums of independent heavy-tailed random variables are typically realised by one random variable taking a very large value. In this article we investigate the conditional distribution of the random variables $\{X_i\}_{1 \leq i \leq n}$ subject to a large deviation of their sum S_n. It turns out that as $n \to \infty$ this conditional distribution converges to a product of $n-1$ copies of μ, while the remaining variable realises the large deviation event by taking a very large value. We determine when the fluctuations around that value have a scaling limit, and we show that given the sum exceeds a large value, the maximum is asymptotically independent of the smallest variables.

---

[1]Universidad de San Andrés, Vito Dumas 284, B1644BID, Victoria, Argentina. E-mail:iarmendariz@udesa.edu.ar
[2]Department of Applied Mathematics, University of Crete, and Institute of Applied and Computational Mathematics, FORTH, Crete. Knossos Avenue, 714 09 Heraklion Crete, Greece. E-mail: loulakis@tem.uoc.gr

2 Notation and Results

Let $\{X_n\}_{n \in \mathbb{N}}$ be a sequence of i.i.d. random variables with common distribution μ defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. We denote by F their distribution function $F(x) = \mu(-\infty, x]$. We are interested in the case where F is in the class of subexponential distributions, that is

$$\lim_{x \to \infty} \frac{\bar{F}(x+y)}{\bar{F}(x)} = 1, \qquad \forall y \in \mathbb{R}, \tag{2.1}$$

with $\bar{F}(x) = \mathbb{P}[X_k > x]$, and

$$\lim_{x \to \infty} \frac{\bar{F}_n(x)}{n\bar{F}(x)} = 1, \qquad \forall n \in \mathbb{N}, \tag{2.2}$$

where $\bar{F}_n(x)$ is defined in (1.1). If μ is supported on the positive half-line then (2.1) is implied by (2.2), and in that case subexponentiality can be defined by the latter condition alone. Since it is generally true that for all $n \in \mathbb{N}$ we have

$$\lim_{x \to \infty} \frac{\mathbb{P}[\max_{1 \le k \le n} X_k > x]}{n\bar{F}(x)} = 1,$$

equation (2.2) states that the tail of the sum in a sample of independent μ-distributed random variables is determined by the tail of the largest variable. These distributions arise naturally when modelling heavy-tailed phenomena. For instance, individual claims in insurance or large interarrival times in queuing systems are usually modelled by distributions of this kind. Typical members of this class include distributions with regularly varying, lognormal-type, or Weibull-type tails. Sufficient conditions for a given distribution to be subexponential that are straightforward to check can be found in [14].

An immediate consequence of (2.2) is the existence of a sequence $d_n \to \infty$ such that

$$\lim_{n \to \infty} \sup_{x \ge d_n} \left| \frac{\bar{F}_n(x)}{n\bar{F}(x)} - 1 \right| = 0. \tag{2.3}$$

A large amount of work has been done for determining a threshold d_n, for which (2.3) holds. Interested readers can find reviews on the topic in [13, 15]. A very nice account is also provided by Mikosch and A. Nagaev in [11]. Denisov, Dieker and Schneer give an up-to-date treatment of this problem in [5].

When the distribution μ satisfies a local version of (2.2), a local version of (2.3) is valid. Let $\Delta = (0, s]$ for some $s > 0$ and denote by $x + \Delta$ the interval $(x, x+s]$. We say that μ is Δ-subexponential if $\mu[x + \Delta] > 0$ for all sufficiently large x and

$$\lim_{x \to \infty} \frac{\mu[x+y+\Delta]}{\mu[x+\Delta]} = 1, \qquad \forall y \in \mathbb{R}, \tag{2.4}$$

and

$$\lim_{x \to \infty} \frac{\mathbb{P}[S_n \in x + \Delta]}{n\mu[x+\Delta]} = 1, \qquad \forall n \in \mathbb{N}. \tag{2.5}$$

The concept of Δ-subexponentiality was introduced in [2] by Asmussen, Foss and Korshunov. A Δ-subexponential distribution is also $m\Delta$-subexponential for all $m \in \mathbb{N}$, $m\Delta = (0, ms]$, and subexponential in the sense of (2.1) and (2.2) (cf. [2].) Even though there are examples of subexponential distributions that are not Δ-subexponential for finite Δ, most distributions that are used in practice

are. This can be easily verified using the sufficient conditions for Δ-subexponentiality provided in [2]. The asymptotics for the large deviation probabilities are now given by

$$\lim_{n \to \infty} \sup_{x \geq d_n} \left| \frac{\mathbb{P}\big[S_n \in x + \Delta\big]}{n\mu\big[x + \Delta\big]} - 1 \right| = 0, \tag{2.6}$$

and sufficient conditions on d_n for (2.6) to hold can be found in [5].

We are interested in the conditional distribution of the variables $\{X_n\}$ subject to a large deviation of their sum. Assuming that (2.6) holds for some interval Δ that may be finite or infinite, we would like to determine the asymptotic behaviour of

$$\mu_{n,x}^{\Delta}\big[A\big] = \mathbb{P}\big[(X_1, \ldots, X_n) \in A \mid S_n \in x + \Delta\big]$$

when $n \to \infty$ and $x \geq d_n$. Note that when $\Delta = (0, \infty)$ the definition of Δ-subexponentiality reduces to the standard definition of subexponentiality and (2.6) reduces to (2.3). This allows to treat rare events of the form $\{S_n \in (x, x+s]\}$ or $\{S_n > x\}$ simultaneously in Theorem 1 below.

A related question was raised in [9] where certain subexponential families μ of lattice type are considered under $\{S_n = x(n)\}$, and it is shown that the finite dimensional marginals of the conditional distribution converge to a product of copies of μ.

We will denote by $T : \cup_{n \in \mathbb{N}} \mathbb{R}^n \to \cup_{n \in \mathbb{N}} \mathbb{R}^n$ the operator that exchanges the last and the maximum component of a finite sequence:

$$T(x_1, \ldots, x_n)_k = \begin{cases} \max_{1 \leq i \leq n} x_i & \text{if } k = n, \\ x_n & \text{if } x_k > \max_{1 \leq i < k} x_i \text{ and } x_k = \max_{i \geq k} x_i, \\ x_k & \text{otherwise.} \end{cases}$$

Theorem 1. *Suppose μ is Δ-subexponential. There exists a sequence q_n such that*

$$\lim_{n \to \infty} \sup_{x \geq q_n} \sup_{A \in \mathcal{B}(\mathbb{R}^{n-1})} \left| \mu_{n,x}^{\Delta} \circ T^{-1}\big[A \times \mathbb{R}\big] - \mu^{n-1}\big[A\big] \right| = 0.$$

The sequence q_n in the statement can be easily computed from the sequence d_n in (2.6) and F, and in most cases turns out to be d_n itself. Note that $\mu_{n,x}^{\Delta} \circ T^{-1}\big[A \times \mathbb{R}\big]$ is the measure assigned to $A \in \mathbb{R}^{n-1}$ by the conditional distribution of the $n-1$ smallest variables. In other words, Theorem 1 states that under (2.3), conditioning on $\{S_n \in x + \Delta\}$ affects only the maximum in the limit, and the $n-1$ smallest variables become asymptotically independent. Such a result is rather uncommon, and when μ satisfies Cramér's condition an analogous statement is not true. Now, any limit theorem for i.i.d. random variables with distribution μ can be cast in this setting. For instance, we could obtain conditional limit theorems for the statistics of any order $k > 1$: the k-th order statistic of $(X_1, \ldots, X_n)$ subject to the condition $S_n \in x + \Delta$, $x \geq q_n$ asymptotically behaves like the $(k-1)$-th order statistic of an independent sample.

Unlike the asymptotic independence of the smallest variables, the fluctuations of the maximum $M_n = \max_{1 \leq i \leq n} X_i$ and its dependence on the smallest variables are influenced by the form of conditioning. When $\Delta = (0, s]$ the condition we impose on the sum is very restrictive and the fluctuations of the maximum are determined by the fluctuations of the sum of the smallest variables. This can be easily seen since $\mu_{n,x}^{\Delta}\big[M_n + \sum_{j=1}^{n-1}(TX)_i \in (x, x+s]\big] = 1$ by definition. Therefore,

3

if the (unconditioned) distribution of S_{n-1}/b_n converges to a stable law H, it follows immediately from Theorem 1 that under $\mu_{n,x}^{\Delta}$ we have

$$\frac{M_n - x}{b_n} \xrightarrow{d} -H. \tag{2.7}$$

Note that the converse is also true. In particular, the fluctuations of the conditional maximum are typically two-sided and they have a non trivial scaling limit if and only if μ is attracted to a stable distribution. In [1] Theorem 1 is proved for a particular family of lattice distributions subject to $\{S_n = x\}$ and this observation is used to obtain a limit theorem for the fluctuations of the maximum in a system of interacting particles.

On the other hand when we condition on $\{S_n > x\}$ it turns out that the maximum coordinate is asymptotically independent of the smallest variables, its fluctuations around x are one-sided, and they have a non trivial scaling limit if and only if μ is in the maximum domain of attraction of an extreme value distribution. For ease of notation, we will now drop $\Delta = (0, \infty]$ from the notation,

$$\mu_{n,x}[A] = \mathbb{P}[(X_1, \ldots, X_n) \in A \mid S_n > x].$$

Let ν_x stand for the conditional distribution of X_i subject to $X_i > x$. That is,

$$\nu_x[A] = \mathbb{P}[X_i \in A \mid X_i > x] = \frac{\mu[A \cap (x, \infty)]}{\bar{F}(x)}.$$

We will use $\|\nu\|_{t.v.}$ to denote the total variation norm of a signed Borel measure on $\mathbb{R}^n$. That is

$$\|\nu\|_{t.v.} = \sup_{A \in \mathcal{B}(\mathbb{R}^n)} |\nu(A)|.$$

Theorem 2. *Suppose μ is subexponential. Then*

$$\lim_{n \to \infty} \sup_{x \geq q_n} \left\| \mu_{n,x} \circ T^{-1} - (\mu^{n-1} \times \nu_x) \right\|_{t.v.} = 0, \tag{2.8}$$

where q_n is the sequence appearing in Theorem 1.

Since the distribution of $(X_1, \ldots, X_n)$ subject to $\{S_n > x\}$ is clearly exchangeable, the position of the maximum coordinate is uniformly distributed among $1, \ldots, n$. Theorem 2 states that the conditional distribution of the maximum coordinate becomes asymptotically a randomly located ν_x, while the law of the remaining $n-1$ variables is the product μ^{n-1} as was established in Theorem 1.

It is interesting to examine whether (2.8) entails a limit theorem for the fluctuations of the maximum around x, that is, whether there exists a scaling function $\psi(\cdot)$ such that under $\mu_{n,x}$ we have

$$\frac{M_n - x}{\psi(x)} \xrightarrow{d} \Lambda, \tag{2.9}$$

for some non trivial distribution Λ. In view of Theorem 2 this is equivalent to asking when

$$\nu_x[(x + u\psi(x), \infty)] = \frac{\bar{F}((x + u\psi(x))}{\bar{F}(x)} \tag{2.10}$$

converges as $x \to \infty$ to a nontrivial function of u. This is precisely the subject of [3], where Balkema and de Haan determine all possible scaling limits of residual life times as the survival time goes to infinity, and the corresponding domains of attraction. It follows from their results (Theorems 1, 3 and 4 there) that nontrivial limits in the right hand side of (2.9) can only be of two types.

4

1. An exponential distribution of rate 1 if and only if μ is in the maximum domain of attraction of the Gumbel distribution. In this case ψ can be determined by requiring the expression in (2.10) to converge to e^{-u}.

2. A Pareto destribution on $\mathbb{R}_+$ with $\bar{\Lambda}(u) = (1+u)^{-\alpha}$ and $\alpha > 0$, if and only if μ has regularly varying tails with index $-\alpha$, that is $\bar{F}(x) = x^{-\alpha}L(x)$, as $x \to \infty$, and L is a slowly varying function. Note that this is equivalent to μ being in the maximum domain of attraction of the Fréchet distribution with index α (cf. [4].) In this case $\psi(x) = x$.

That regularly varying distributions satisfy our main assumption (2.3) is a long known fact (cf. [10, 12].) In particular, if $\alpha > 2$ one can choose $d_n = \sqrt{tn \log n}$ for any $t > \alpha - 2$ (cf. [13].) The articles [5, 11] are excellent references for subexponential distributions in the maximum domain of attraction of the Gumbel distribution and the corresponding sequences d_n for which (2.3) holds.

Remarks
1. Theorem 2 and the discussion following it generalise a result of Mikosch and Nagaev (Proposition 4.4 in [11]) where they prove (2.9) for some of the most commonly used subexponential distributions.

2. Theorem 2 also generalises an old result by Richard Durrett [6]. In that article it is proved that if μ is regularly varying with index $\alpha < -2$ and $\mathbb{E}[X_1] = -b < 0$, then

$$\Big(\frac{S_{[n\cdot]}}{n} \mid S_n > 0 \Big) \Rightarrow J_{\alpha,b} 1_{\{U \le \cdot\}} - b \cdot,$$

where U is uniform in [0,1] and $J_{\alpha,b}$ is independent of U with Pareto distribution. As in Corollary 3 in [1], Theorem 2 also establishes a conditional invariance principle for the sum of the random variables cut-off at a level εn.

Simple modifications of the proofs give ramifications of Theorems 1 and 2. For instance, the results remain true if n is fixed and we only let $x \to \infty$. Theorem 1 remains valid if we let the size of Δ grow to infinity with n and x, and a variant of Theorem 2 is satisfied if Δ grows fast enough. This enables to explore precisely how the fluctuations of the maximum switch from a stable to a residual life time nature. These ramifications are discussed in section 4, after the proof of the Theorems that follows.

3 Proof of the Theorems

Given a vector $\mathbf{x} = (x_1, \ldots, x_n) \in \mathbb{R}^n$ we denote by $M_{\mathbf{x}}$ the coordinate of maximum size and by $m_{\mathbf{x}}$ its position. Precisely,

$$M_{\mathbf{x}} = \max_{1 \le k \le n} x_i \qquad \text{and} \qquad m_{\mathbf{x}} = k \Leftrightarrow x_k > x_j, \ j < k, \ \text{and} \ x_k \ge x_j, \ j \ge k.$$

We will denote by σ^j the operator that exchanges the j-th and the last coordinate of $\mathbf{x}$, that is

$$\sigma^j(x_1, \ldots, x_j, \ldots, x_{n-1}, x_n) = (x_1, \ldots, x_n, \ldots, x_{n-1}, x_j).$$

With this notation we may write $T\mathbf{x} = \sigma^{m_{\mathbf{x}}}\mathbf{x}$. Let also $\mathbf{X} = (X_1, \ldots, X_n)$, $\mathbf{X}^{n-1} = (X_1, \ldots, X_{n-1})$. We begin with some elementary observations that will be useful for both proofs.

The convergence in (2.4) is in fact uniform over compact y–sets. This follows from the uniform convergence theorem for slowly varying functions (see [4], Theorem 1.2.1), as (2.4) implies that

$x \mapsto \mu\big[\log x + \Delta\big]$ is slowly varying. In particular, if b_n is any sequence growing to infinity there exists a sequence $m_n \to \infty$ such that

$$\lim_{n \to \infty} \sup_{x \geq m_n} \sup_{0 \leq y \leq b_n} \left| \frac{\mu[x - y + \Delta]}{\mu[x + \Delta]} - 1 \right| = 0. \tag{3.1}$$

This in turn implies that there exists a sequence $\ell_n \gg b_n$ such that

$$D_n(L) := \sup_{x \geq \ell_n} \sup_{|y| \leq L b_n} \left(1 - \frac{\mu[x - y + \Delta]}{\mu[x + \Delta]} \right) \longrightarrow 0, \qquad \text{as } n \to \infty, \qquad \forall L > 0. \tag{3.2}$$

To see this, iterate (3.1) using the fact the limit is uniform in $x \geq m_n$ to get

$$\lim_{n \to \infty} \sup_{x \geq m_n} \sup_{-L b_n \leq y \leq b_n} \left| \frac{\mu[x - y + \Delta]}{\mu[x + \Delta]} - 1 \right| = 0. \tag{3.3}$$

Now, if ρ_n is any sequence increasing to infinity we may choose $\ell_n = m_n + \rho_n b_n$.

The sequence q_n in the statement of the theorems can be constructed as follows. Take a sequence b_n such that S_{n-1}/b_n is tight and choose ℓ_n so that (3.2) holds. We may then choose $q_n = d_n \vee \ell_n$. Very often, in fact in all cases we are aware of where a threshold d_n in (2.3) or (2.6) is explicitly known, and certainly for the d_n constructed in [5], we can choose $\ell_n \leq d_n$ so the supremum in the theorems is taken for $x \geq d_n$. Finally, note that for any $B \in \mathcal{B}(\mathbb{R}^n)$ we have

$$\mathbb{P}\big[T\mathbf{X} \in B, S_n \in x + \Delta\big] = \sum_{j=1}^{n} \mathbb{P}\big[T\mathbf{X} \in B, S_n \in x + \Delta, \, m_{\mathbf{X}} = j\big]$$

$$\geq \sum_{j=1}^{n} \mathbb{P}\big[\sigma^j \mathbf{X} \in B, S_n \in x + \Delta, \, m_{\sigma^j \mathbf{X}} = n\big]$$

$$= n\,\mathbb{P}[\mathbf{X} \in B, S_n \in x + \Delta, \, m_{\mathbf{X}} = n]. \tag{3.4}$$

The last equality holds because $\mathbb{P}$ is invariant under σ^j. The penultimate inequality holds because $m_{\sigma^j \mathbf{X}} = n \Rightarrow m_{\mathbf{X}} = j$ (notice however that if μ is atomless this inequality and (3.5) below are in fact equalities.) In view of (3.4) we have

$$\mu_{n,x}^{\Delta} \circ T^{-1}\big[B\big] \geq \frac{n\,\mathbb{P}[\mathbf{X} \in B, \, S_n \in x + \Delta, \, m_{\mathbf{X}} = n]}{\mathbb{P}\big[S_n \in x + \Delta\big]}. \tag{3.5}$$

Proof of Theorem 1. Consider $A \in \mathcal{B}(\mathbb{R}^{n-1})$ as in the statement of the theorem and fix $L \in \mathbb{N}$. We have

$$\mathbb{P}\big[\mathbf{X} \in A \times \mathbb{R}, \, S_n \in x + \Delta, \, m_{\mathbf{X}} = n\big] \geq \mathbb{P}\big[\mathbf{X} \in A \times \mathbb{R}, \, S_n \in x + \Delta, \, |S_{n-1}| < L b_n, \, m_{\mathbf{X}} = n\big]$$

$$\geq \mathbb{P}\big[\mathbf{X} \in A \times \mathbb{R}, \, S_n \in x + \Delta, \, |S_{n-1}| < L b_n, \, M_{\mathbf{X}^{n-1}} \leq x - L b_n\big]$$

$$= \int_{\mathbf{X}^{n-1} \in A \cap G} \mu\big[x - S_{n-1} + \Delta\big]\, d\mathbb{P},$$

where

$$G = G(n, L, x) = \Big\{ \mathbf{u} \in \mathbb{R}^{n-1} : \big| \sum_{i=1}^{n-1} u_i \big| < L b_n, \, M_{\mathbf{u}} \leq x - L b_n \Big\}. \tag{3.6}$$

Notice that when $\mathbf{u} \in G$ and $x \geq \ell_n$ we have

$$\mu\big[x - \sum_{i=1}^{n-1} u_i + \Delta\big] \geq \big(1 - D_n(L)\big)\mu\big[x + \Delta\big],$$

so that (3.5) can be reinforced to

$$\mu_{n,x}^{\Delta} \circ T^{-1}\big[A \times \mathbb{R}\big] \geq \big(1 - D_n(L)\big) \frac{n\mu\big[x + \Delta\big]}{\mathbb{P}\big[S_n \in x + \Delta\big]} \, \mathbb{P}\big[\mathbf{X}^{n-1} \in A \cap G\big]$$

giving the estimate

$$\mu_{n,x}^{\Delta} \circ T^{-1}\big[A \times \mathbb{R}\big] - \mathbb{P}\big[\mathbf{X}^{n-1} \in A\big] \;\geq\; -\left(\mathbb{P}\big[\mathbf{X}^{n-1} \notin G\big] + D_n(L) + \left|\frac{n\mu\big[x + \Delta\big]}{\mathbb{P}\big[S_n \in x + \Delta\big]} - 1\right|\right).$$

Denote the expression in the parenthesis on the right hand side above by $R(n, L, x)$. We can get an upper bound by applying the same estimate for $\mathbb{R}^{n-1} \setminus A$, the complement of A. Combining the two bounds we get

$$\big| \mu_{n,x}^{\Delta} \circ T^{-1}\big[A \times \mathbb{R}\big] - \mathbb{P}\big[\mathbf{X}^{n-1} \in A\big] \big| \leq R(n, L, x).$$

Now the sequence S_{n-1}/b_n is tight, so we have

$$\lim_{L \to \infty} \sup_n \mathbb{P}\big[|S_{n-1}| \geq Lb_n\big] = 0. \tag{3.7}$$

On the other hand, it is known (see [7], Section lX.7) that

$$\lim_{L \to \infty} \sup_n n[F(-Lb_n) + \bar{F}(Lb_n)] = 0, \tag{3.8}$$

and since $\ell_n \gg b_n$ we have

$$\sup_{x \geq \ell_n} \mathbb{P}\big[M_{\mathbf{X}^{n-1}} > x - Lb_n\big] = 1 - \inf_{x \geq \ell_n}\big(1 - \bar{F}(x - Lb_n)\big)^{n-1} \longrightarrow 0, \qquad \text{as } n \to \infty.$$

Combining this limit with (3.7) we see that $\mathbb{P}\big[\mathbf{X}^{n-1} \notin G\big]$ vanishes uniformly on $x \geq \ell_n$, as $n \to \infty$ and then $L \to \infty$. The result now follows from (2.6) and (3.2). $\square$

Proof of Theorem 2. The proof follows the general outline of Theorem 1.

It is sufficient to show that

$$\lim_{n \to \infty} \sup_{x \geq d_n \vee \ell_n} \sup_{U \in \mathcal{R}} \Big| \mu_{n,x} \circ T^{-1}\big[U\big] - \mu^{n-1} \times \nu_x\big[U\big] \Big| = 0, \tag{3.9}$$

where the supremum above is taken over the class $\mathcal{R}$ of finite disjoint unions of rectangles $A_j \times B_j$ with $A_j \in \mathcal{B}(\mathbb{R}^{n-1})$ and $B_j \in \mathcal{B}(\mathbb{R})$. Recall the definition of $G \subset \mathbb{R}^{n-1}$ in (3.6) and define $I = (x + Lb_n, \infty)$. Using (3.5) we have

$$\begin{aligned}
\mu_{n,x} \circ T^{-1}\Big[\bigcup_j A_j \times B_j\Big] &= \sum_j \mu_{n,x} \circ T^{-1}\big[A_j \times B_j\big] \\
&\geq \sum_j \frac{n\mathbb{P}\big[\mathbf{X}^{n-1} \in A_j,\, X_n \in B_j,\, S_n > x,\, m_{\mathbf{X}} = n\big]}{\bar{F}_n(x)} \\
&\geq \frac{n}{\bar{F}_n(x)} \sum_j \mathbb{P}\big[\mathbf{X}^{n-1} \in A_j \cap G,\, X_n \in B_j \cap I\big] \\
&= \frac{n\bar{F}(x)}{\bar{F}_n(x)} \mu^{n-1} \times \nu_x\Big[\big(\bigcup_j A_j \times B_j\big) \cap (G \times I)\Big].
\end{aligned}$$

Just as in the proof of Theorem 1 this gives that for all $U \in \mathcal{R}$ we have

$$
\mu_{n,x} \circ T^{-1}[U] - \mu^{n-1} \times \nu_x[U] \geq -\left(\mu^{n-1} \times \nu_x\left[(G \times I)'\right] + \left|\frac{n\bar{F}(x)}{\bar{F}_n(x)} - 1\right|\right)
$$

$$
\geq -\left(\mu^{n-1}[G'] + \nu_x[I'] + \left|\frac{n\bar{F}(x)}{\bar{F}_n(x)} - 1\right|\right)
$$

$$
\geq -R(n, L, x), \qquad \forall x \geq \ell_n.
$$

In the previous equation and in the following the prime symbol denotes the complement of a set in the appropriate space: $(G \times I)' = \mathbb{R}^n \setminus (G \times I)$, $I' = \mathbb{R} \setminus I$ and $G' = \mathbb{R}^{n-1} \setminus G$.

Since $\mathcal{R}$ is closed under complementation we can also get an upper bound by applying the previous inequality for U' to get

$$
\left|\mu_{n,x} \circ T^{-1}[U] - \mu^{n-1} \times \nu_x[U]\right| \leq R(n, L, x).
$$

The proof is now completed by letting $n \to \infty$, then $L \to \infty$ as before. □

Here's another proof of Theorem 2.

Second Proof of Theorem 2: The measure $\mu_{n,x}$ has density with respect to μ^n given by

$$
f_1(\mathbf{x}) = \frac{\mathbb{1}\{\sum_{i=1}^n x_i > x\}}{\bar{F}_n(x)}.
$$

The measure

$$
\mu_{n,x}^* = \frac{1}{n} \sum_{j=1}^n \sigma^j\left(\mu^{n-1} \times \nu_x\right)
$$

has density with respect to μ^n given by

$$
f_2(\mathbf{x}) = \frac{1}{n} \sum_{j=1}^n \frac{\mathbb{1}\{x_j > x\}}{\bar{F}(x)} = \frac{N_n(\mathbf{x})}{n\bar{F}(x)},
$$

where $N_n(\mathbf{x})$ stands for the number of coordinates in $\mathbf{x}$ that are greater than x. Now,

$$
\|\mu_{n,x} - \mu_{n,x}^*\|_{t.v.} = \int |f_1(\mathbf{x}) - f_2(\mathbf{x})|\, d\mu(\mathbf{x}) = \int \left|\frac{\mathbb{1}\{S_n > x\}}{\bar{F}_n(x)} - \frac{N_n(\mathbf{X})}{n\bar{F}(x)}\right| d\mathbb{P}.
$$

If n is large enough so that

$$
\sup_{x \geq d_n} \left|\frac{\bar{F}_n(x)}{n\bar{F}(x)} - 1\right| < \frac{1}{2},
$$

It is a matter of a straightforward computation to see that the preceding expression becomes

$$
\|\mu_{n,x} - \mu_{n,x}^*\|_{t.v.} = 2\left(1 - \frac{\bar{F}_n(x)}{n\bar{F}(x)}\right)^+ \mu_{n,x}[N_n(\mathbf{x}) = 1] + 2\frac{\mathbb{P}[S_n > x, M_n \leq x]}{\bar{F}_n(x)}.
$$

The first term on the right hand side above clearly goes to zero, uniformly on $\{x \geq d_n\}$ as $n \to \infty$. The uniform convergence to zero of the second term can be deduced from the arguments in the proof of Theorem 4.1 in [11], but we give a proof here for the sake of completeness.

$$
\bar{F}_n(x) = \mathbb{P}[S_n > x, M_n \leq x] + \mathbb{P}[S_n > x, N_n(\mathbf{X}) \geq 1]
$$

$$
= \mathbb{P}[S_n > x, M_n \leq x] + n\bar{F}(x) \int_{\sum x_i > x} \frac{d\mu_{n,x}^*(\mathbf{x})}{N_n(\mathbf{x})}
$$

$$
= \mathbb{P}[S_n > x, M_n \leq x] + n\bar{F}(x)\left(1 - R(n, x)\right), \tag{3.10}
$$

where

$$0 \leq R(n, x) = \mu_{n,x}^* \Big[\sum x_i \leq x \Big] + \int_{\sum x_i > x} \left(1 - \frac{1}{N_n(\mathbf{x})} \right) \, d\mu_{n,x}^*(\mathbf{x})$$

$$\leq \mu_{n,x}^* \Big[\sum x_i \leq x \Big] + \mu_{n,x}^* \big[N_n(\mathbf{x}) \geq 2 \big]$$

$$= \mu_{n,x}^* \Big[\sum x_i \leq x \Big] + \mathbb{P} \big[M_{n-1} > x \big].$$

In view of (3.8), it is enough to show that $\mu_{n,x}^* \big[\sum x_i \leq x \big] \to 0$, uniformly on $x \geq d_n \vee \ell_n$. Now,

$$\mu_{n,x}^* \Big[\sum x_i \leq x \Big] = \int_{S_{n-1} < 0} \left(1 - \frac{\bar{F}(x - S_{n-1})}{\bar{F}(x)} \right) d\mathbb{P}$$

$$\leq D_n(L) + \mathbb{P} \big[S_{n-1} < -L b_n \big],$$

and this bound goes to zero if we let $n \to \infty$, then $L \to \infty$ by (3.2) and (3.7). It follows from (2.3) and (3.10) that

$$\lim_{n \to \infty} \sup_{x \geq q_n} \frac{\mathbb{P} \big[S_n > x, M_n \leq x \big]}{\bar{F}_n(x)} = 0.$$

$\square$

4 Some related results

Let us begin by observing that the assertions of Theorems 1 and 2 remain valid if we keep n fixed and let $x \to \infty$,

Proposition 1. *If μ is Δ-subexponential, then for any $n \in \mathbb{N}$*

$$\lim_{x \to \infty} \sup_{A \in \mathcal{B}(\mathbb{R}^{n-1})} \big| \mu_{n,x}^\Delta \circ T^{-1} \big[A \times \mathbb{R} \big] - \mu^{n-1} \big[A \big] \big| = 0. \tag{4.1}$$

We recall the notation $\mu_{n,x} = \mu_{n,x}^\Delta$ when $\Delta = (0, \infty)$.

Proposition 2. *If μ is subexponential and $\Delta = (0, \infty)$, then for any $n \in \mathbb{N}$*

$$\lim_{x \to \infty} \big\| \mu_{n,x} \circ T^{-1} - (\mu^{n-1} \times \nu_x) \big\|_{t.v.} = 0.$$

The proofs of Propositions 1 and 2 are essentially the same as those of Theorems 1 and 2. We can take $b_n = 1$, and instead of using (2.6) and (3.1) we may use the defining relations of (Δ)-subexponentiality. Note that (4.1) is proved in [8] for a family of discrete distributions that includes those with regularly varying tails, subject to $\{ S_n = x \}$.

So far we have assumed that the interval Δ is fixed. We would like to explore what happens when $\Delta = \big(0, s(n, x) \big]$ is a finite interval, but its size $s(n, x)$ grows to infinity with n and x. In particular, we would like to understand how the fluctuations of the maximum switch from a stable to a residual life time nature and how fast $|\Delta|$ would need to grow for the maximum to be asymptotically independent of the other variables.

Denote by $\mathcal{D}$ the family of semi–open intervals with one endpoint at 0 contained in the positive half line,

$$\mathcal{D} = \Big\{ \Delta \subseteq (0, \infty), \, \Delta = (0, z], \, z > 0 \Big\}.$$

Assume that μ satisfies (2.4) and (2.6) for some fixed interval $\Delta_0 = (0, s_0]$. Let ρ_n be any sequence increasing to infinity.

Lemma 1. *If μ satisfies (3.1) and (2.6) for some fixed Δ_0, then as $n \to \infty$ we have*

$$\sup_{\Delta \in \mathcal{D}, |\Delta| \geq \rho_n} \sup_{x \geq m_n} \sup_{0 \leq y \leq b_n} \left| \frac{\mu[x - y + \Delta]}{\mu[x + \Delta]} - 1 \right| \to 0$$

and

$$\sup_{\Delta \in \mathcal{D}, |\Delta| \geq \rho_n} \sup_{x \geq d_n \vee m_n} \left| \frac{\mathbb{P}[S_n \in x + \Delta]}{n\mu[x + \Delta]} - 1 \right| \to 0.$$

Proof: Let us first note that for any $k \in \mathbb{N} \cup \{\infty\}$

$$\sup_{x \geq m_n} \sup_{0 \leq y \leq b_n} \left| \frac{\mu[x - y + k\Delta_0]}{\mu[x + k\Delta_0]} - 1 \right| \leq \sup_{x \geq m_n} \sup_{0 \leq y \leq b_n} \left| \frac{\mu[x - y + \Delta_0]}{\mu[x + \Delta_0]} - 1 \right|. \tag{4.2}$$

In order to see this we split the interval $k\Delta_0 = (0, ks_0]$ into k disjoint intervals of length $|\Delta_0|$, $\Delta_0^i = ((i-1)s_0, is_0] = (i-1)s_0 + \Delta_0$, $1 \leq i \leq k$. We get

$$\left| \frac{\mu[x - y + k\Delta_0]}{\mu[x + k\Delta_0]} - 1 \right| = \frac{1}{\mu[x + k\Delta_0]} \left| \sum_{i=1}^{k} \mu[x - y + \Delta_0^i] - \mu[x + \Delta_0^i] \right|$$

$$= \frac{1}{\mu[x + k\Delta_0]} \left| \sum_{i=1}^{k} \mu[x + \Delta_0^i] \left(\frac{\mu[x + (i-1)s_0 - y + \Delta_0]}{\mu[x + (i-1)s_0 + \Delta_0]} - 1 \right) \right|$$

$$\leq \frac{1}{\mu[x + k\Delta_0]} \left(\sum_{i=1}^{k} \mu[x + \Delta_0^i] \sup_{x \geq m_n} \sup_{0 \leq y \leq b_n} \left| \frac{\mu[x - y + \Delta_0]}{\mu[x + \Delta_0]} - 1 \right| \right)$$

$$= \sup_{x \geq m_n} \sup_{0 \leq y \leq b_n} \left| \frac{\mu[x - y + \Delta_0]}{\mu[x + \Delta_0]} - 1 \right|.$$

Now, for an arbitrary interval $\Delta \in \mathcal{D}$ with $|\Delta| \geq \rho_n$, let $k = [|\Delta|/|\Delta_0|]$ and $\bar{s} = |\Delta| - k|\Delta_0| \leq s_0$. If $\bar{\Delta} = (0, \bar{s}]$ we have

$$\frac{\mu[x - y + \Delta]}{\mu[x + \Delta]} = \frac{\mu[x - y + \bar{\Delta}] + \mu[x + \bar{s} - y + k\Delta_0]}{\mu[x + \bar{\Delta}] + \mu[x + \bar{s} + k\Delta_0]}.$$

By dividing both terms of the fraction in the right hand side by $\mu[x + \bar{s} + k\Delta_0]$, and using (3.1) and (4.2) we see that in order to prove the first assertion of the lemma it suffices to show that

$$\frac{\mu[x + \Delta_0]}{\mu[x + k\Delta_0]} \to 0 \tag{4.3}$$

uniformly on $x \geq m_n$, as $k \to \infty$. For this, we may assume without loss of generality that $k|\Delta_0| \leq b_n$ (as otherwise the denominator would be even larger), in which case a computation similar to the previous one yields

$$\mu[x + k\Delta_0] = \sum_{i=1}^{k} \mu[x + (i-1)s_0 + \Delta_0] \geq k\mu[x + \Delta_0](1 - \delta_n),$$

10

where δ_n is the supremum in the right hand side of (4.2), and (4.3) follows.

For the second assertion we may check again that for any $k \in \mathbb{N} \cup \{\infty\}$

$$\sup_{x \geq d_n} \left| \frac{\mathbb{P}\big[S_n \in x + k\Delta_0\big]}{n\mu\big[x + k\Delta_0\big]} - 1 \right| \leq \sup_{x \geq d_n} \left| \frac{\mathbb{P}\big[S_n \in x + \Delta_0\big]}{n\mu\big[x + \Delta_0\big]} - 1 \right|$$

and use (2.6) and (4.3) to conclude the proof. $\square$

In view of Lemma 1 we can repeat the argument in the proof of Theorem 1 to establish the following.

Proposition 3. *Suppose μ is Δ_0-subexponential for some finite interval Δ_0. Then,*

$$\lim_{n \to \infty} \sup_{\Delta \in \mathcal{D}, |\Delta| \geq \rho_n} \sup_{x \geq q_n} \sup_{A \in \mathcal{B}(\mathbb{R}^{n-1})} \left| \mu_{n,x}^{\Delta} \circ T^{-1}\big[A \times \mathbb{R}\big] - \mu^{n-1}\big[A\big] \right| = 0,$$

where the sequence q_n is the same appearing in Theorem 1.

Theorem 2 also admits a generalisation in this case. Denote by ν_x^{Δ} the conditional distribution of X_i subject to $X_i \in x + \Delta$, that is

$$\nu_x^{\Delta}\big[A\big] = \frac{\mu\big[A \cap (x + \Delta)\big]}{\mu\big[x + \Delta\big]}.$$

Proposition 4. *Suppose μ is Δ_0-subexponential for some finite interval Δ_0, and let b_n be a sequence such that S_{n-1}/b_n is tight. Then*

$$\lim_{n \to \infty} \sup_{\Delta \in \mathcal{D}, |\Delta| \geq \rho_n b_n} \sup_{x \geq q_n} \left\| \mu_{n,x}^{\Delta} \circ T^{-1} - (\mu^{n-1} \times \nu_x^{\Delta}) \right\|_{t.v.} = 0. \tag{4.4}$$

The proof is the same as that of Theorem 2, if we substitute $\bar{F}(x)$, $\bar{F}_n(x)$ and I by $\mu\big[x + \Delta\big]$, $\mathbb{P}\big[S_n \in x + \Delta\big]$ and $\tilde{I} = (x + Lb_n, x + |\Delta| - Lb_n]$ respectively. To show that $R(n, L, x) \to 0$ we need to use Lemma 1 and the fact that $\nu_x^{\Delta}\big[\tilde{I}'\big] \to 0$, which can be proved just as (4.3). A direct consequence is that the maximum becomes asymptotically independent from the rest of the variables as long as $|\Delta|$ grows faster than b_n.

We now discuss the fluctuations of the maximum $M_n = \max_{1 \leq i \leq n} X_i$ subject to $S_n \in x + \Delta$, when the distribution μ is both in the sum-domain of attraction of a stable law H, that is, there exists a sequence $b_n \uparrow \infty$ such that $S_{n-1}/b_n \to H$, and the max-domain of attraction of an extreme value distribution (necessarily Gumbel or Fréchet). In most cases these can be easily derived from the preceding results. The following diagram illustrates the typical behaviour for locally subexponential distributions arising in applications. We assume throughout that $x \geq q_n$, and ρ_n is a sequence growing to infinity.
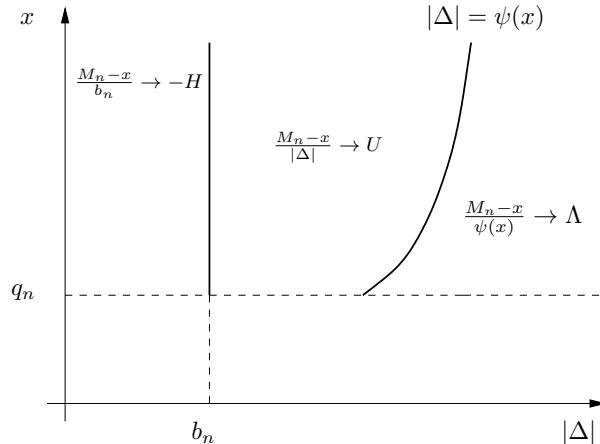
When $|\Delta| \leq b_n/\rho_n$ the maximum fluctuations are the same as the ones that arise when conditioning on a finite interval. This holds for all Δ_0-subexponential μ in the domain of attraction of a stable law. It follows from Proposition 3 and the fact that $\mu_{n,x}^{\Delta}\big[M_n + \sum_{j=1}^{n-1}(TX)_j \in x + \Delta\big] = 1$.

When $|\Delta| \geq \rho_n\psi(x)$ the maximum fluctuations are the same as those obtained by conditioning on $\{S_n > x\}$. This holds for all Δ_0-subexponential μ in the max-domain of attraction of the Gumbel or the Fréchet distribution. Note that in view of Lemma 1, (2.4) implies that

$$\sup_{x \geq m_n} \sup_{|y| \leq b_n} \left| \frac{\bar{F}(x-y)}{\bar{F}(x)} - 1 \right| \longrightarrow 0.$$

Recall that if the function ψ and the distribution Λ are those appearing in (2.9), for all t for which $\bar{\Lambda}(t)$ is defined we have

$$\frac{\bar{F}\big(x + t\psi(x)\big)}{\bar{F}(x)} \longrightarrow \bar{\Lambda}(t), \tag{4.5}$$

so we must have $\inf_{x \geq m_n} \psi(x)/b_n \to \infty$. Since by Proposition 4

$$\mu_{n,x}^{\Delta}\left[\frac{M_n - x}{\psi(x)} \leq t\right] - \frac{\bar{F}(x) - \bar{F}\big(x + t\psi(x)\big)}{\bar{F}(x) - \bar{F}(x + |\Delta|)} \longrightarrow 0 \text{ as } n \to \infty,$$

uniformly on $x \geq q_n$, $|\Delta| \gg b_n$, it follows from (4.5) that $(M_n - x)/\psi(x)$ converges to Λ.

In the intermediate region $b_n\rho_n \leq |\Delta| \leq \psi(x)/\rho_n$ we prove below that $(M_n - x)/|\Delta|$ converges to a uniform random variable U in $[0,1]$ under some monotonicity condition on μ. Precisely, this holds if μ is Δ_0-subexponential and in the max-domain of attraction of the Gumbel or Fréchet distribution, S_{n-1}/b_n is tight, and furthermore there exists an interval $\Delta_1 = [0, s_1)$ such that $\mu[x + \Delta_1]$ is eventually decreasing in x. We then provide an example to show that the result does not hold in general when the last condition is violated. With this caveat, note that the subexponential distributions typically used in modelling do satisfy all the above conditions.

The argument is an adaptation of the proof of Theorem 3.10.11 in [4]. For any $0 < \varepsilon < \frac{1}{2}$, if we set $k = \left[\frac{\varepsilon\psi(x)}{|\Delta_1|}\right]$ we have for all sufficiently large x

$$\bar{F}(x) - \bar{F}\big(x + \varepsilon\psi(x)\big) \leq \sum_{i=0}^{k} \mu\big[x + is_1 + \Delta_1\big] \leq (k+1)\mu\big[x + \Delta_1\big],$$

by monotonicity. Likewise,

$$\bar{F}\big(x - \varepsilon\psi(x)\big) - \bar{F}\big(x + |\Delta_1|\big) \geq (k+1)\mu\big[x + \Delta_1\big].$$

Applying these estimates for $\mu\big[x + \Delta_1\big]$ and $\mu\big[x - y + \Delta_1\big]$ for any $y \in \mathbb{R}$ we have

$$\frac{\bar{F}(x-y) - \bar{F}\big(x - y + \varepsilon\psi(x)\big)}{\bar{F}\big(x - \varepsilon\psi(x)\big) - \bar{F}\big(x + |\Delta_1|\big)} \leq \frac{\mu\big[x - y + \Delta_1\big]}{\mu\big[x + \Delta_1\big]} \leq \frac{\bar{F}\big(x - y - \varepsilon\psi(x)\big) - \bar{F}\big(x - y + |\Delta_1|\big)}{\bar{F}(x) - \bar{F}\big(x + \varepsilon\psi(x)\big)}. \quad (4.6)$$

Since the left hand side in (4.5) is decreasing in t, and $\bar{\Lambda}(\cdot)$ is continuous ($\bar{\Lambda}(t) = e^{-t}$ if μ is attracted to the Gumbel distribution, and $\bar{\Lambda}(t) = (1+t)^{-\alpha}$ if μ is attracted to the Fréchet distribution), the convergence in (4.5) is uniform in $t \in [-2\varepsilon, 2\varepsilon]$. Hence,

$$\lim_{n\to\infty} \sup_{x \geq m_n} \sup_{|y| \leq \frac{\psi(x)}{\rho_n}} \left| \frac{\mu\big[x - y + \Delta_1\big]}{\mu\big[x + \Delta_1\big]} - 1 \right| \leq \frac{\bar{\Lambda}(\varepsilon) + \bar{\Lambda}(-\varepsilon) - 2\bar{\Lambda}(0)}{\bar{\Lambda}(0) - \bar{\Lambda}(\varepsilon)}.$$

We can let $\varepsilon \to 0$ now to get

$$\sup_{x \geq m_n} \sup_{|y| \leq \frac{\psi(x)}{\rho_n}} \left| \frac{\mu\big[x - y + \Delta_1\big]}{\mu\big[x + \Delta_1\big]} - 1 \right| = 0.$$

Just as in Lemma 1 this in turn implies that

$$\sup_{|y| \leq |\Delta|} \left| \frac{\mu\big[x - y + u\Delta\big]}{\mu\big[x + u\Delta\big]} - 1 \right| = 0, \quad \text{as } n \to \infty,$$

for any $u \in [0,1]$, uniformly of $x \geq m_n$, $\rho_n \leq |\Delta| \leq \psi(x)/\rho_n$, and therefore

$$\nu_x^\Delta\big[x + u\Delta\big] = \frac{\mu\big[x + u\Delta\big]}{\mu\big[x + \Delta\big]} \longrightarrow u.$$

Now by Proposition 4 the distribution of the maximum M_n is asymptotically equal to ν_x^Δ when $|\Delta| \geq \rho_n b_n$, hence $(M_n - x)/|\Delta| \xrightarrow{d} U$. $\square$

This result is not generally valid without some regularity assumption on μ. In order to see this, let ν be the Pareto distribution with density function

$$\phi(x) = \frac{\alpha}{x^{\alpha+1}} \qquad x \geq 1,$$

where $\alpha > 0$. Consider increasing sequences $c_k \to \infty$ with $c_{k+1}/c_k \to 1$, $d_k = \sum_{j=1}^k c_{j-1}$, $k \geq 1$, some fixed $0 < \epsilon < 1$, and define a distribution μ with density function

$$\bar{\phi}(x) = \sum_{k \geq 0} \frac{1}{\epsilon c_k} \nu\big[d_k, d_{k+1}\big] \, \mathbb{1}_{[d_k, d_k + \epsilon c_k)}(x).$$

In other words, μ redistributes the mass assigned by ν to the interval $[d_k, d_{k+1})$ uniformly over the sub–interval $[d_k, d_k + \epsilon c_k)$. It is easy to see that μ satisfies the conditions of Theorem 4 with H an α–stable law, and that it belongs to the max–domain of attraction of the Fréchet distribution.

Let now $k(n)$ be an increasing sequence such that $c_{k(n)} \geq q_n \gg b_n$, and take $\Delta_n = (0, c_{k(n)}]$, $x_n = d_{k(n)} \gg q_n$. By Theorem 4

$$\lim_{n\to\infty} \mu_{n,x_n}^{\Delta_n} \left[\frac{M_n - x_n}{|\Delta_n|} \in [\epsilon, 1) \right] = \lim_{n\to\infty} \frac{\mu[x_n + |\Delta_n| [\epsilon, 1)]}{\mu[x_n + \Delta_n]} = 0 \neq 1 - \epsilon,$$

where ϵ is the positive value in the definition of μ, and in particular $M_n - x_n/|\Delta_n|$ does not converge to a uniform distribution. In fact, with little extra effort one can change the measure ν and choose sequences x_n and Δ_n so that $(M_n - x_n)/|\Delta_n|$ converges to any given distribution.

One may even determine the behaviour of the fluctuations on the critical scales of $|\Delta|$.

When $|\Delta| = a\psi(x)$, $(M_n - x)/\psi(x)$ *converges to* Λ *conditioned to being less than* a. This holds under the conditions required for the case $|\Delta| \gg \psi(x)$ and the proof is also the same.

Finally, if $|\Delta| = ab_n$, *then* $(M_n - x)/b_n \xrightarrow{d} aU - H$, *where the uniform random variable* U *is independent of* H. This holds whenever μ satisfies (2.4) and (2.6), and S_{n-1}/b_n converges to a stable disribution H.

To see this suppose that $|\Delta| = ab_n$ and $t \in [0, 1]$. Let us begin by observing that

$$\lim_{n\to\infty} \sup_{x \geq m_n} \left| \frac{\mu[x + t\Delta]}{\mu[x + \Delta]} - t \right| = 0. \tag{4.7}$$

This follows from (3.3). Using the fact that $\mu[x+t\Delta]$ is increasing in t we can show that convergence is uniform on $t \in [0, 1]$. The rest of the proof follows that of Theorem 1.

$$\mathbb{P}[X_n > x + ub_n, \ S_n \in x + \Delta, \ m_{\mathbf{X}} = n] \geq \mathbb{P}[\mathbf{X}^{n-1} \in G, \ X_n > x + ub_n, \ S_n \in x + \Delta]$$

$$= \int_{\mathbf{X}^{n-1} \in G} \mathbb{1}\{X_n > x + ub_n, \ x - S_{n-1} < X_n \leq x - S_{n-1} + ab_n\} \, d\mathbb{P}$$

$$= \int_{\substack{\mathbf{X}^{n-1} \in G \\ S_{n-1} < (a-u)b_n}} \left(\bar{F}(x + (ub_n \vee -S_{n-1})) - \bar{F}(x + ab_n - S_{n-1}) \right) d\mathbb{P}.$$

Now, write the integral above as the sum of the integrals I_1 and I_2 over the sets $\{S_{n-1} \leq -ub_n\}$ and $\{-ub_n < S_{n-1} < (a - u)b_n\}$, respectively. For $x \geq \ell_n$, we can estimate the first of the integrals by

$$I_1 \geq (1 - D_n(L))\mu[x + \Delta]\mathbb{P}[\mathbf{X}^{n-1} \in G, \ S_{n-1} \leq -ub_n],$$

while the second one can be estimated by

$$I_2 \geq (1 - D_n(L))\mu[x + \Delta] \int_{\substack{\mathbf{X}^{n-1} \in G \\ -ub_n < S_{n-1} < (a-u)b_n}} \frac{\bar{F}(x) - \bar{F}(x + (a - u)b_n - S_{n-1})}{\bar{F}(x) - \bar{F}(x + ab_n)} \, d\mathbb{P}.$$

Using the uniform convergence in (4.7) we may now pass to the limit to get

$$\liminf_{n\to\infty} \inf_{x \geq q_n} \mu_{n,x}^{\Delta} \left[\frac{M_n - x}{b_n} > u \right] \geq F_H(-u) + \int_{0 < \xi + u < a} \left(1 - \frac{\xi + u}{a} \right) dF_H(\xi)$$

$$= \mathbb{P}[aU - H > u],$$

where F_H above is the distribution function of H, $F_H(x) = \mathbb{P}[H \leq x]$, and U is distributed uniformly on [0,1] and is independent of H. Similarly, we can prove that

$$\liminf_{n\to\infty} \inf_{x \geq q_n} \mu_{n,x}^{\Delta} \left[\frac{M_n - x}{b_n} \leq u \right] \geq \mathbb{P}[aU - H \leq u],$$

so $(M_n - x)/b_n$ converges in distribution to $aU - H$. $\square$

References

[1] Inés Armendáriz and Michail Loulakis. Thermodynamic limit for the invariant measures in supercritical zero range processes. *Probab. Theory Related Fields*, 145(1-2):175–188, 2009.

[2] Søren Asmussen, Serguei Foss, and Dmitry Korshunov. Asymptotics for sums of random variables with local subexponential behaviour. *J. Theoret. Probab.*, 16(2):489–518, 2003.

[3] A. A. Balkema and L. de Haan. Residual life time at great age. *Ann. Probab.*, 2:792–804, 1974.

[4] N. H. Bingham, C. M. Goldie, and J. L. Teugels. *Regular variation*, volume 27 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, Cambridge, 1989.

[5] D. Denisov, A. B. Dieker, and V. Shneer. Large deviations for random walks under subexponentiality: the big-jump domain. *Ann. Probab.*, 36(5):1946–1991, 2008.

[6] Richard Durrett. Conditioned limit theorems for random walks with negative drift. *Z. Wahrsch. Verw. Gebiete*, 52(3):277–287, 1980.

[7] William Feller. *An introduction to probability theory and its applications. Vol. II.* Second edition. John Wiley & Sons Inc., New York, 1971.

[8] Pablo A. Ferrari, Claudio Landim, and Valentin V. Sisko. Condensation for a fixed number of independent random variables. *J. Stat. Phys.*, 128(5):1153–1158, 2007.

[9] Stefan Großkinsky, Gunter M. Schütz, and Herbert Spohn. Condensation in the zero range process: stationary and dynamical properties. *J. Stat. Phys.*, 113(3-4):389–410, 2003.

[10] C. C. Heyde. On large deviation probabilities in the case of attraction to a non-normal stable law. *Sankhyā Ser. A*, 30:253–258, 1968.

[11] T. Mikosch and A. V. Nagaev. Large deviations of heavy-tailed sums with applications in insurance. *Extremes*, 1(1):81–110, 1998.

[12] A. V. Nagaev. Limit theorems that take into account large deviations when Cramér's condition is violated. *Izv. Akad. Nauk UzSSR Ser. Fiz.-Mat. Nauk*, 13(6):17–22, 1969.

[13] S. V. Nagaev. Large deviations of sums of independent random variables. *Ann. Probab.*, 7(5):745–789, 1979.

[14] Jozef L. Teugels. The class of subexponential distributions. *Ann. Probab.*, 3(6):1000–1011, 1975.

[15] Vladimir Vinogradov. *Refined large deviation limit theorems*, volume 315 of *Pitman Research Notes in Mathematics Series*. Longman Scientific & Technical, Harlow, 1994.