

# LIMIT THEOREMS FOR SUPER-HEAVY TAILED RANDOM VARIABLES WITH TRUNCATION: APPLICATION TO THE SUPER-PETERSBURG GAME

TOSHIO NAKATA

Department of Mathematics, University of Teacher Education Fukuoka, Akama-Bunkyo-machi, Munakata, Fukuoka, 811-4192, JAPAN.  
E-mail: nakata@fukuoka-edu.ac.jp

## Abstract

Motivated by the *super-Petersburg game*, we consider the *super-heavy tailed* independent and identically distributed (iid) random variables whose tail are characterized by slow variation. This article explores strong laws of large numbers and central limit theorems for a class of super-heavy tailed random variables with two types of truncations, respectively. We apply our results to the logPareto distributions and the super-Petersburg distributions.

## 1. Introduction

### 1.1. Notation

Functions  $\log(\cdot)$  and  $\lg(\cdot)$  denote logarithms whose bases are  $e$  and  $2$ , respectively. For positive functions  $f(x)$  and  $g(x)$  the symbol  $f(x) \asymp g(x)$  means that there exist positive constants  $C_1, C_2$  and  $a$  satisfying that

$$C_1 \leq f(x)/g(x) \leq C_2 \quad \text{for } x \geq a.$$

Similarly,  $f(x) \sim g(x)$  and  $f(x) = o(g(x))$  denote  $\lim_{x \rightarrow \infty} f(x)/g(x) = 1$  and  $\lim_{x \rightarrow \infty} f(x)/g(x) = 0$ , respectively. Moreover, for positive sequences

---

Received January 26, 2020 and in revised form March 11, 2020.

AMS Subject Classification: 60G50, 60F15.

Key words and phrases: Super-Petersburg game, super-heavy tail, truncation, strong law of large numbers, central limit theorem, logPareto distribution.

This research was supported by KAKENHI 19K03622 of Japan Society for the Promotion of Science.

$\{a_n\}$  and  $\{b_n\}$  we use the same notation  $a_n \asymp b_n$ ,  $a_n \sim b_n$  and  $a_n = o(b_n)$ , respectively.

## 1.2. Super-heavy tailed distributions

The distribution for the *St. Petersburg game* is known as one of the simplest examples with infinite expectation. The game is the following. Peter tosses a fair coin repeatedly until it falls heads. If this happens at the  $k$ th trial, he gives Paul  $2^k$  ducats for  $k \in \mathbb{N} = \{1, 2, 3, \dots\}$ . Letting  $X^{\text{St}}$  be a payoff, we have

$$\mathbb{P}(X^{\text{St}} = 2^k) = 2^{-k} \quad \text{for } k \in \mathbb{N}. \quad (1)$$

Although the fee of the game can be considered as the expectation, which is calculated as  $\mathbb{E}X^{\text{St}} = 2 \cdot \frac{1}{2} + 2^2 \cdot \frac{1}{2^2} + \dots = \infty$ , it does not fit our intuition (see [15, Section X.4]). The distribution of Eq. (1) is sometimes called the *Peter and Paul distribution* (see [7, Section 8.8.2, page 372]), but let us call it the *St. Petersburg distribution*. In 1738, D. Bernoulli [5] proposed  $\mathbb{E} \log X^{\text{St}} < \infty$  instead of  $\mathbb{E}X^{\text{St}}$  as the “moral expectation”. However, it is pointed out that considering the logarithmic utility function in this situation is *ad hoc* by giving the following *super-Petersburg game* constructed by K. Menger (see [12, page 30], [21, page 635], [28, Section II, page 32] and [31, Section 4]). If Peter tosses the coin repeatedly, and it falls heads at the  $k$ th trial, then Paul receives  $2^{2^k}$  ducats for  $k \in \mathbb{N}$ . Letting  $X^{\text{sP}}$  be a payoff, we have

$$\mathbb{P}(X^{\text{sP}} = 2^{2^k}) = 2^{-k} \quad \text{for } k \in \mathbb{N}, \quad (2)$$

and call it the *super-Petersburg distribution*, since Samuelson [28] used the word “super-Petersburg paradox”. Note that the same calculation of  $\mathbb{E}X^{\text{St}}$  yields that  $\mathbb{E} \log X^{\text{sP}}$  is also infinite.

Let us consider the distribution of  $X^{\text{sP}}$  by comparison with the distribution of  $X^{\text{St}}$ . They are both *heavy-tailed* in the sense of [16, Equation (1.1), page 2], and  $\mathbb{E}X^{\text{St}} = \mathbb{E}X^{\text{sP}} = \infty$ . However, we can see that  $\mathbb{E}X^{\text{St}}$  *barely* diverges to infinity but  $\mathbb{E}X^{\text{sP}}$  *certainly* diverges to infinity. Indeed, the decay orders for tail functions  $x \mapsto \mathbb{P}(X^{\text{St}} > x)$  and  $x \mapsto \mathbb{P}(X^{\text{sP}} > x)$

are different. Namely, the tail probabilities of  $X^{\text{St}}$  and  $X^{\text{sP}}$  are

$$\frac{1}{x} \leq \mathbb{P}(X^{\text{St}} > x) = 2^{-\lfloor \lg x \rfloor} = \frac{2^{\{\lg x\}}}{x} < \frac{2}{x} \quad \text{for } x > 2 \quad (3)$$

and

$$\frac{1}{\lg x} \leq \mathbb{P}(X^{\text{sP}} > x) = 2^{-\lfloor \lg \lg x \rfloor} = \frac{2^{\{\lg \lg x\}}}{\lg x} < \frac{2}{\lg x} \quad \text{for } x > 4, \quad (4)$$

where  $\lfloor x \rfloor$  and  $\{x\}$  are the integer and the fractional parts of  $x$  for  $x > 0$ , respectively. The distribution of Eq. (3) is in a class of *power laws*, i.e.

$$\mathbb{P}(X > x) \asymp x^{-\alpha} \quad \text{for } 0 < \alpha \leq 1. \quad (5)$$

There are many results of limit theorems for iid random variables of this type distribution, in particular it is interesting when  $\alpha = 1$  (see [1], [2], [3], [22], [25], [26] and references therein).

To handle with random variables like Eq. (4), let us introduce *super-heavy tailed* distributions.

**Definition 1.1** (Super-heavy tail). A distribution of a nonnegative random variable  $X \geq 0$  is *super-heavy tailed* if there exists a slowly varying function  $L(x) > 0$  satisfying that

$$\mathbb{P}(X > x) = \frac{1}{L(x)}. \quad (6)$$

Moreover, it is *O-super-heavy tailed* if Eq. (6) is replaced by

$$\mathbb{P}(X > x) \asymp \frac{1}{L(x)}. \quad (7)$$

We usually call Eqs. (6) and (7) “*slowly varying tailed*” and “*O-slowly varying tailed*”, respectively, but we dare to use the word “*super*” to correspond to the “*super-Petersburg game*”.

For the notion of the slow or regular variation, we refer to the textbooks [7], [18, Section A.7] and [20, Appendix B]. According to Alves et al. [4, page 214], several authors gave different definitions of the super-heavy tail. The above definition of Eq. (6) is written in Falk et al. [14, Section 2.7, page 75]. Alves et al. [4] defined it only when  $L(x)$  in Eq. (6) is  $\Pi$ -*varying* (see [14, Class B, page 79], for  $\Pi$ -variation, see [7, Chapt. 3] and [20, Appendix B, page 361]). Zeevi and Glynn [32] defined it when  $X$  satisfies  $\mathbb{E} \log(1 + |X|) =$

$\infty$ , which means that the moral expectation diverges. Here, we consider the generalization from Eq. (6) to Eq. (7), since we would like to deal with the case that the tail function is not slowly varying like Eq. (4) (see Remark 4.1). Note that the class of O-super-heavy tailed distributions is a subclass of *dominatedly varying* distributions (see Embrechts et al. [13, Section 1.4.1, page 49]) and *O-subexponential* distributions (see Shimura and Watanabe [30]).

While the St. Petersburg distribution is not super-heavy tailed, the *log-Pareto distribution* (Eq. (41)) and the super-Petersburg distribution (Eq. (2)) are typical examples for super-heavy and O-super-heavy tailed distributions, respectively. It does not seem that these are well-studied. Indeed, Cormann and Reiss [10, page 94] reported that even the logPareto distribution was rarely investigated in statistical inference. In probability theory, it might be difficult to directly obtain limit theorems with respect to super-heavy tailed distributions for the following reasons.

- (i) The super-heavy tailed random variable does not have the *relatively stable* property that is also regarded as a weak law of large numbers, because it does not satisfy the condition of [7, Theorem 8.8.1, page 373].
- (ii) The super-heavy tailed random variable does not belong to the *domain of attraction* of any stable distributions, because it does not satisfy the *Doebelin-Gnedenko criterion* [18, Theorem 9.3.2, page 432].
- (iii) The super-heavy tailed random variable does not belong to the *max-domain of attraction* of any extreme value distributions. Indeed, the direct proof for the logPareto random variable (see Eq. (41)) is given in [17, Example 2.6.1, page 88]).

### 1.3. Our contributions and organization of the article

To overcome the difficulties, Darling [11] and Nagaev, Vachtel [23] studied distributional limit theorems for super-heavy tailed iid random variables using a ingenious normalization.

In this article, we try another way by considering truncations for heavy tailed random variables. We give strong laws of large numbers and central

limit theorems for sums of iid O-super-heavy random variables with truncations, respectively. Györfi and Kevei [19] studied them for the St. Petersburg distribution of Eq. (3). Their results were generalized by [24] to a class of distributions of Eq. (5). We examine them for a class of distributions which correspond to Eq. (5) by extension the case of  $\alpha = 0$ .

The organization of the article is as follows. In Section 2 after giving two types of truncations and technical assumptions, we claim main results for the strong law of large numbers (Theorem 2.1) and the central limit theorem (Theorem 2.2). Section 3 presents proofs of the theorems after two lemmas. Finally, Sections 4.1 and 4.2. illustrate examples for the logPareto distributions and the super-Petersburg distributions, respectively.

## 2. Truncations, Assumptions and Results

### 2.1. Truncations

#### 2.1.1. Definition of min-truncation and default-truncation

For a nonnegative random variable  $X \geq 0$ , we define two types of truncations with a level  $x > 0$  as follows.

$$X(x) = \min\{X, x\} = X\mathbb{I}_{\{X \leq x\}} + x\mathbb{I}_{\{X > x\}}, \quad (8)$$

$$\widehat{X}(x) = X\mathbb{I}_{\{X \leq x\}}, \quad (9)$$

where  $\mathbb{I}_A$  denotes the indicator random variable

$$\mathbb{I}_A(\omega) = \begin{cases} 1, & \text{for } \omega \in A, \\ 0, & \text{for } \omega \notin A. \end{cases}$$

For the sake of convenience, we call Eqs. (8) and (9) the *min-truncation* and the *default-truncation*, respectively.

#### 2.1.2. Interpretation of truncations

These truncations, which are given in [9, Example 6.2.2, page 182], can be interpreted as a game in the following sense. Let us suppose that Peter has  $x$  ducats. Then for the min-truncation  $X(x)$  he pays Paul all  $x$  ducats when the payoff  $X$  is greater than  $x$  ducats. By contrast, for the default-truncation  $\widehat{X}(x)$  he does not pay at all in this case.

As a result, we find that the min-truncation, which may be a *peaceful* procedure for Paul, is easier to obtain limit theorems than the default-truncation (see Remark 2.2 (i)).

## 2.2. Assumptions

If  $X$  is a O-super-heavy tailed random variable, then by Eq. (7) there exist  $C_1 > 0, C_2 > 0$  and  $x_0 > 0$  satisfying that

$$\frac{C_1}{L(x)} \leq \mathbb{P}(X > x) \leq \frac{C_2}{L(x)} \quad \text{for } x > x_0. \quad (10)$$

For the slowly varying function  $L(x)$  and  $C_1, C_2$ , we suppose the following technical assumptions.

(A1) For the function  $L(x)$  in Eq. (10) there exist  $c > 0$  and  $a_0 \geq x_0$  such that

$$L(x) = c \exp \left\{ \int_{a_0}^x \frac{\varepsilon(u)}{u} du \right\} \quad \text{for } x \geq a_0, \quad (11)$$

where  $\varepsilon(x)$  is positive, continuous, slowly varying and  $o(1)$ .

(A2) For  $i = 1$  and  $2$ , there exists a random variable  $\widetilde{X}_i$  whose distribution function  $F_i(x)$  with a constant  $a_i > 0$  is

$$F_i(x) = \mathbb{P}(\widetilde{X}_i \leq x) = \begin{cases} 1 - \frac{C_i}{L(x)}, & \text{for } x \geq a_i, \\ 0, & \text{for } x < a_i. \end{cases} \quad (12)$$

### 2.2.1. Remarks on Assumptions

- (i) Eq.(11) is a *normalized slowly varying* function (see [7, Equation (1.3.4), page 15]), which is from Representation Theorem (see [7, Theorem 1.3.1, page 12] and [29, Theorem 1.2, page 2]).
- (ii) Assumption (A1), which is required by some technical reasons, does not affect the tail behavior of  $X$ . Although it seems to be strong (see Remark 3.2), we suppose it for clarity and simplicity. Since Assumption (A1) implies the differentiability of  $L(x)$ , the function  $\varepsilon(x)$  is calculated

by

$$\varepsilon(x) = \frac{xL'(x)}{L(x)}, \quad (13)$$

and we have

$$f_i(x) = \begin{cases} \frac{C_i \varepsilon(x)}{xL(x)}, & \text{for } x \geq a_i, \\ 0, & \text{for } x < a_i, \end{cases} \quad (14)$$

where  $f_i(x)$  is the probability density function of  $\widetilde{X}_i$  for  $i = 1$  and  $2$ . Note that  $x \mapsto \varepsilon(x)/L(x)$  in Eq. (14) is a slowly varying function by the assumption of slowly variation for  $\varepsilon(x)$ . Therefore, [7, Sec. 3.0, page 127] or [20, Theorem B.2.12, page 377] yields  $F_i(x)$  in (A2) is  $\Pi$ -varying. Hence so is  $L(x)$ .

- (iii) It is true that Assumption (A2) yields that  $\varepsilon(x) \geq 0$  by [29, Lemma 1.7, page 15], but we suppose  $\varepsilon(x) > 0$  which means  $L'(x) > 0$ .

### 2.3. Results

Let  $X_1, X_2, \dots$  be iid random variables whose distribution is given by Eq. (10). We define

$$S_n(x) = \sum_{i=1}^n X_i(x) \quad \text{and} \quad \widehat{S}_n(x) = \sum_{i=1}^n \widehat{X}_i(x),$$

and call them the *min-truncated sum* and the *default-truncated sum*, respectively. Letting  $l_n$  be the number of trials of the game, we suppose

$$l_n \in \mathbb{N} \quad \text{and} \quad \lim_{n \rightarrow \infty} l_n = \infty. \quad (15)$$

Moreover, letting  $c_n$  be the level of the truncation, for nontriviality we also suppose

$$c_n > 0 \quad \text{and} \quad \lim_{n \rightarrow \infty} c_n = \infty. \quad (16)$$

Then we have the next statements.

**Theorem 2.1** (Strong law of large numbers). *Under Assumptions (A1) and (A2) and Eqs. (15) and (16), the following claims hold.*

(i) For the min-truncated sum, it follows that

$$\lim_{n \rightarrow \infty} \frac{S_{l_n}(c_n)}{\mathbb{E}S_{l_n}(c_n)} = 1 \quad a.s. \quad (17)$$

if

$$L(c_n) = o\left(\frac{l_n}{\log n}\right). \quad (18)$$

(ii) For the default-truncated sum, it follows that

$$\lim_{n \rightarrow \infty} \frac{\widehat{S}_{l_n}(c_n)}{\mathbb{E}\widehat{S}_{l_n}(c_n)} = 1 \quad a.s. \quad (19)$$

if

$$\frac{L(c_n)}{\varepsilon(c_n)} = o\left(\frac{l_n}{\log n}\right). \quad (20)$$

**Remark 2.1.** It is true that Conditions (18) and (20) may be strong, but it does not seem to be so easy to find some kind of converse to these results.

**Theorem 2.2** (Central limit theorem). *Under Assumptions (A1) and (A2) and Eqs. (15) and (16), the following claims hold.*

(i) For the min-truncated sum it follows that

$$\lim_{n \rightarrow \infty} \frac{S_{l_n}(c_n) - \mathbb{E}S_{l_n}(c_n)}{\sqrt{\text{Var}S_{l_n}(c_n)}} = N(0, 1) \quad \text{in distribution}, \quad (21)$$

if and only if

$$L(c_n) = o(l_n), \quad (22)$$

where  $N(0, 1)$  is the standard normal distribution.

(ii) For the default-truncated sum it follows that

$$\lim_{n \rightarrow \infty} \frac{\widehat{S}_{l_n}(c_n) - \mathbb{E}\widehat{S}_{l_n}(c_n)}{\sqrt{\text{Var}\widehat{S}_{l_n}(c_n)}} = N(0, 1) \quad \text{in distribution}, \quad (23)$$

if and only if

$$\frac{L(c_n)}{\varepsilon(c_n)} = o(l_n). \quad (24)$$



**Remark 2.2.**

- (i) It is easy to see that Eqs. (18) and (20) yield Eqs. (22) and (24), respectively. Since  $\varepsilon(x) = o(1)$ , Eqs. (20) and (24) also yield Eqs. (18) and (22), respectively.
- (ii) Conditions of Eqs. (18) and (22) for O-super-heavy tailed distributions correspond to [24, Eqs. (13) and (23)] for power laws with index  $\alpha$ , respectively.

**3. Proofs****3.1. Lemmas**

We use two lemmas to prove Theorems 2.1 and 2.2.

**Lemma 3.1.** *Under Assumptions (A1) and (A2), the random variable  $X$  with Eq. (10) satisfies*

$$\mathbb{P}(\widetilde{X}_2 > x) \geq \mathbb{P}(X > x) \geq \mathbb{P}(\widetilde{X}_1 > x) \quad \text{for } x \geq a_*, \quad (25)$$

where

$$a_* = \max\{a_0, a_1, a_2\}.$$

Moreover, it follows that

$$\mathbb{E}X = \infty. \quad (26)$$

**Proof.** Eq. (25) holds by combining Eqs. (10) and (12). To obtain Eq. (26), we check  $\mathbb{E}\widetilde{X}_1 = \infty$ . This is given by  $\mathbb{E}\widetilde{X}_1 = \lim_{x \rightarrow \infty} \int_0^x \mathbb{P}(\widetilde{X}_1 > t) dt$  and  $\int_0^x \mathbb{P}(\widetilde{X}_1 > t) dt \sim \frac{x}{L(x)} \rightarrow \infty$  as  $x \rightarrow \infty$  because the function  $x \mapsto 1/L(x)$  is slowly varying.  $\square$

Using Lemma 3.1 and Eq. (13) we have the following statement.

**Lemma 3.2** (Estimations of truncated moments). *Under Assumptions (A1) and (A2), the random variable  $X$  with Eq. (10) satisfies*

$$\mathbb{E}(X(x))^k \asymp \frac{1}{L(x)} x^k \quad \text{for } k \in \mathbb{N}, \quad (27)$$

$$\mathbb{E}(\widehat{X}(x))^k \asymp \frac{\varepsilon(x)}{kL(x)} x^k \quad \text{for } k \in \mathbb{N}, \quad (28)$$

where “ $\asymp$ ” in Eqs. (27) and (28) holds uniformly  $k \in \mathbb{N}$ , respectively, i.e., for Eq. (27) there exist  $C'_1 > 0, C'_2 > 0$  and  $x_0 > 0$  satisfying that

$$C'_1 \frac{x^k}{L(x)} \leq \mathbb{E}(X(x))^k \leq C'_2 \frac{x^k}{L(x)} \quad \text{for } x > x_0 \text{ and } k \in \mathbb{N}.$$

**Remark 3.1.** Since  $\varepsilon(x) = o(1)$ , it follows that

$$\mathbb{E}(\widehat{X}(x))^k = o\left(\mathbb{E}(X(x))^k\right) \quad \text{for each } k \in \mathbb{N}. \quad (29)$$

**Proof of Lemma 3.2.** Fix  $i \in \{1, 2\}$ . By Karamata’s theorem ([7, Proposition 1.5.8, 26 page]) we have

$$\mathbb{E}(\widetilde{X}_i(x))^k = k \int_0^x y^{k-1} \mathbb{P}(\widetilde{X}_i > y) dy = \int_0^x (y^k)' \frac{C_i}{L(y)} dy \sim \frac{C_i}{L(x)} x^k \quad (30)$$

for sufficient large  $x \geq a_i$ . Since Eqs. (25) and (26) tell us

$$\mathbb{E}(\widetilde{X}_1(x))^k \leq \mathbb{E}(X(x))^k \leq \mathbb{E}(\widetilde{X}_2(x))^k \quad (31)$$

for sufficient large  $x > a_*$ , we have Eq. (27) by inserting Eq. (30) to Eq. (31).

Next, since  $\varepsilon(x)/L(x)$  is slowly varying from Sec. 2.2.1 (ii), it follows that

$$\begin{aligned} \mathbb{E}(\widehat{X}_i(x))^k &= \mathbb{E} \widetilde{X}_i^k \mathbb{I}_{\{\widetilde{X}_i \leq x\}} = \int_{a_i}^x t^k f_i(t) dt = C_i \int_{a_i}^x \frac{\varepsilon(t)}{L(t)} t^{k-1} dt \\ &\sim \frac{C_i \varepsilon(x)}{kL(x)} x^k \end{aligned} \quad (32)$$

for sufficiently large  $x > a_*$ , which implies Eq. (28).  $\square$

**Remark 3.2.** To obtain (27), we use only properties that  $L(x)$  is slowly varying, and  $\lim_{x \rightarrow \infty} L(x) = \infty$ . Note that  $L(x)$  does not have to be normalized.

### 3.2. Proof of Theorem 2.1

- (i) Although we can prove by a similar method of the proof of [24, Theorem 1.2 (ii)], here we use a *bypass* method. Let us fix  $\delta > 0$ . By Bennett’s

inequality (see [8, Theorem 2.9, page 35] and [27, Problem 2.6.3, page 78]), we have

$$\begin{aligned} & \mathbb{P} \left( \frac{\sum_{i=1}^{l_n} \{X_i(c_n) - \mathbb{E}X_i(c_n)\}}{\mathbb{E}S_{l_n}(c_n)} > \delta \right) \\ & \leq \exp \left( -\frac{\text{Var}S_{l_n}(c_n)}{c_n^2} h \left( \frac{c_n \mathbb{E}S_{l_n}(c_n) \delta}{\text{Var}S_{l_n}(c_n)} \right) \right) \leq \exp \left( -\frac{K_1 l_n}{L(c_n)} \right) \\ & = n^{-\frac{K_1 l_n}{(\log n)L(c_n)}} \leq n^{-2}, \end{aligned}$$

where  $h(u) = (1+u)\log(1+u) - u > 0$  for  $u > 0$  and  $K_1 > 0$ . The second inequality follows from  $\mathbb{E}S_{l_n}(c_n) \asymp l_n c_n / L(c_n)$  and  $\text{Var}S_{l_n}(c_n) \asymp l_n c_n^2 / L(c_n)$ . The last inequality follows from Eq. (18). The same procedure implies

$$\mathbb{P} \left( \left| \frac{\sum_{i=1}^{l_n} \{(X_i(c_n)) - \mathbb{E}(X_i(c_n))\}}{\mathbb{E}S_{l_n}(c_n)} \right| > \delta \right) \leq 2n^{-2}.$$

Hence the first Borel-Cantelli lemma yields Eq. (17).

- (ii) Since  $\mathbb{E}\widehat{S}_{l_n}(c_n) \asymp l_n \varepsilon(c_n) c_n / L(c_n)$  and  $\text{Var}\widehat{S}_{l_n}(c_n) \asymp l_n \varepsilon(c_n) c_n^2 / L(c_n)$ , there exists  $K_2 > 0$  such that

$$\mathbb{P} \left( \frac{\sum_{i=1}^{l_n} \{\widehat{X}_i(c_n) - \mathbb{E}\widehat{X}_i(c_n)\}}{\mathbb{E}\widehat{S}_{l_n}(c_n)} > \delta \right) \leq n^{-\frac{K_2 l_n \varepsilon(c_n)}{(\log n)L(c_n)}} \leq n^{-2}.$$

The last inequality follows from Eq. (20). The rest of the proof is the same of (i).

### 3.3. Proof of Theorem 2.2

This is similar to the proof of [24, Theorem 1.3].

- (i) Eq. (22)  $\Rightarrow$  Eq. (21): For the sake of convenience, we put

$$s_{l_n}(c_n) = \sqrt{\text{Var}S_{l_n}(c_n)} \asymp \sqrt{\frac{l_n c_n^2}{L(c_n)}} = \sqrt{\frac{l_n}{L(c_n)}} c_n. \quad (33)$$

It is known that the Lyapunov condition ([6, Equation (27.16)], page

385)]

$$\frac{l_n}{(s_{l_n}(c_n))^3} \mathbb{E}|X_1(c_n) - \mathbb{E}X_1(c_n)|^3 = o(1) \quad (34)$$

implies the central limit theorem i.e. Eq. (21). From Eqs. (27) and (22) we have

$$\frac{l_n}{(s_{l_n}(c_n))^3} \mathbb{E}|X_1(c_n) - \mathbb{E}X_1(c_n)|^3 \asymp \sqrt{\frac{L(c_n)}{l_n}} = o(1),$$

which means Eq. (34).

Eq. (21)  $\Rightarrow$  Eq. (22): Because of the Lindeberg theorem [27, Theorem 4.1, page 114]), Eq. (21) implies that for  $\delta > 0$

$$\mathbb{E} \left( \frac{(X_1(c_n) - \mathbb{E}X_1(c_n))^2}{\text{Var}X_1(c_n)} \mathbb{I}_{\{|X_1(c_n) - \mathbb{E}X_1(c_n)| \geq \delta s_{l_n}(c_n)\}} \right) = o(1). \quad (35)$$

The expectation is bounded by

$$\begin{aligned} & \mathbb{E} \left( \frac{(X_1(c_n) - \mathbb{E}X_1(c_n))^2}{\text{Var}X_1(c_n)} \mathbb{I}_{\{|X_1(c_n) - \mathbb{E}X_1(c_n)| \geq \delta s_{l_n}(c_n)\}} \right) \\ & \geq \frac{1}{\text{Var}X_1(c_n)} \mathbb{E} \left( (X_1(c_n) - \mathbb{E}X_1(c_n))^2 \mathbb{I}_{\{|X_1(c_n) - \mathbb{E}X_1(c_n)| \geq \delta s_{l_n}(c_n), X_1 > c_n\}} \right) \\ & \geq \frac{(c_n - \mathbb{E}X_1(c_n))^2}{\text{Var}X_1(c_n)} \mathbb{P}(X_1(c_n) \geq \delta s_{l_n}(c_n) + \mathbb{E}X_1(c_n), X_1 > c_n). \end{aligned} \quad (36)$$

From Eq. (27) we see

$$(c_n - \mathbb{E}X_1(c_n))^2 \sim c_n^2 \quad \text{and} \quad \text{Var}X_1(c_n) \asymp \frac{c_n^2}{L(c_n)}.$$

This equation and Eqs. (35) and (36) assure us that

$$\mathbb{P}(c_n \geq \delta s_{l_n}(c_n) + \mathbb{E}X_1(c_n), X_1 > c_n) = o\left(\frac{1}{L(c_n)}\right).$$

Therefore, Eq. (10) implies that for sufficient large  $n$

$$c_n < \delta s_{l_n}(c_n) + \mathbb{E}X_1(c_n). \quad (37)$$

Now,  $\mathbb{E}X_1(c_n) = o(s_{l_n}(c_n))$  because of Eqs. (27) and (33). Hence

Eq. (37) implies

$$\lim_{n \rightarrow \infty} c_n / s_{l_n}(c_n) = 0. \quad (38)$$

Thus we have Eq. (22).

(ii) We use the same manner of the proof of (i). Assume Eq. (24). Then putting

$$\widehat{s}_{l_n}(c_n) = \sqrt{\text{Var} \widehat{S}_{l_n}(c_n)} \asymp c_n \sqrt{\frac{l_n \varepsilon(c_n)}{L(c_n)}}, \quad (39)$$

we can check the Lyapunov condition

$$\frac{l_n}{(\widehat{s}_{l_n}(c_n))^3} \mathbb{E} |\widehat{X}_1(c_n) - \mathbb{E} \widehat{X}_1(c_n)|^3 \asymp \sqrt{\frac{L(c_n)}{l_n \varepsilon(c_n)}} = o(1).$$

Thus we have Eq. (23). Next, assuming Eq.(23), we have  $c_n / \widehat{s}_{l_n}(c_n) = o(1)$  by the same argument. This condition means Eq. (24).

**Remark 3.3.** The calculation of [24, Proof of Theorem 1.3, page 14, line 12] corresponds to Eq. (36). Although the same result for Eq. (38) was obtained, it contains a wrong calculation. Hence it should be fixed to the argument which corresponds to Eq. (36).

## 4. Examples

### 4.1. logPareto distributions

In this section, we suppose  $\nu \in \mathbb{N}$ . Let  $X^{\nu\text{-IP}}$  be a random variable with

$$\mathbb{P}(X^{\nu\text{-IP}} > x) = 1 / \log_{\nu} x \quad \text{for } x \geq e_{\nu}(1), \quad (40)$$

where

$$e_{\nu}(x) = \begin{cases} x, & \text{for } \nu = 0, \\ \exp(e_{\nu-1}(x)), & \text{for } \nu \geq 1, \end{cases}$$

$$\log_{\nu} x = \begin{cases} x, & \text{for } \nu = 0, \\ \log(\log_{\nu-1} x), & \text{for } \nu \geq 1. \end{cases}$$

We call the distribution defined by Eq. (40) the  $\nu$ -logPareto distribution. In particular, the 1-logPareto distribution

$$\mathbb{P}(X^{1\text{-IP}} > x) = 1 / \log x \quad \text{for } x \geq e \quad (41)$$

is the standard *logPareto distribution* (see Galambos [17, Example 1.3.3, page 13]). Since the function  $x \mapsto 1/\log_\nu x$  is slowly varying, the function  $\nu$ -logPareto distribution is super-heavy tailed because of Eq. (6).

Let us consider the  $\nu$ -logPareto distribution. It turns out that it satisfies Assumptions (A1) and (A2) with respect to

$$L(x) = \log_\nu x \quad \text{and} \quad \varepsilon(x) = \left( \prod_{j=1}^{\nu} \log_j x \right)^{-1} \quad \text{for } x \geq e_\nu(1)$$

(c.f. [3, Lemma 5.1 (i)]),  $\widetilde{X}_1^{\nu\text{-IP}} = \widetilde{X}_2^{\nu\text{-IP}} = X^{\nu\text{-IP}}$ ,  $c = C_1 = C_2 = 1$ ,  $x_0 = a_0 = a_1 = a_2 = e_\nu(1)$  and the density function is

$$f_1(x) = f_2(x) = \left( (\log_\nu x) \prod_{j=0}^{\nu} \log_j x \right)^{-1} \quad \text{for } x \geq e_\nu(1).$$

Therefore, Theorems 2.1 and 2.2 are applicable to this setting. Here we illustrate concrete examples for each truncation as follows. For the sake of simplicity and comparison, we assume  $l_n = n$  for all the examples below. Moreover, we put  $S_n = \sum_{i=1}^n X_i^{\nu\text{-IP}}$ , where  $X_1^{\nu\text{-IP}}, \dots, X_n^{\nu\text{-IP}}$  are iid random variables with (40).

(i) min-truncation: We suppose

$$c_n = e_\nu(n^\beta) \quad \text{for } 0 < \beta < 1. \quad (42)$$

Then Eq. (18) holds because  $L(c_n) = n^\beta = o(n/\log n)$ . Since  $ES_n(e_\nu(n^\beta)) \sim n^{1-\beta} e_\nu(n^\beta)$ , it follows that

$$\lim_{n \rightarrow \infty} \frac{S_n(e_\nu(n^\beta))}{n^{1-\beta} e_\nu(n^\beta)} = 1 \quad \text{a.s.} \quad (43)$$

Eq. (42) automatically implies Eq. (22) by Remark 2.2 (i). Hence  $\sqrt{\text{Var}S_n(e_\nu(n^\beta))} \sim n^{\frac{1-\beta}{2}} e_\nu(n^\beta)$  yields

$$\lim_{n \rightarrow \infty} \frac{S_n(e_\nu(n^\beta))e^{-n^\beta} - n^{1-\beta}}{\sqrt{n^{1-\beta}}} = N(0, 1) \quad \text{in distribution.} \quad (44)$$

(ii) default-truncation: We suppose

$$c_n = \begin{cases} e^{n^\beta}, & \text{for } 0 < \beta < \frac{1}{2}, \text{ if } \nu = 1, \\ e^{n^\gamma}, & \text{for } 0 < \gamma < 1, \text{ if } \nu \geq 2. \end{cases} \quad (45)$$

(a) Consider the case of  $\nu = 1$ . Since  $L(e^{n^\beta})/\varepsilon(e^{n^\beta}) = n^{2\beta} = o(n/\log n)$  for  $0 < \beta < \frac{1}{2}$ , Eq. (20) holds. Hence we have two limit theorems which correspond to Eqs. (43) and (44), respectively, where

$$\mathbb{E}\widehat{S}_n(e^{n^\beta}) \sim n^{1-2\beta}e^{n^\beta} \quad \text{and} \quad \sqrt{\text{Var}\widehat{S}_n(e^{n^\beta})} \sim \frac{n^{\frac{1}{2}-\beta}}{\sqrt{2}}e^{n^\beta}.$$

(b) Consider the case of  $\nu \geq 2$ . Since

$$\frac{L(e^{n^\gamma})}{\varepsilon(e^{n^\gamma})} = n^\gamma(\log_{\nu-1} n^\gamma) \prod_{j=1}^{\nu-1} \log_j n^\gamma = o\left(\frac{n}{\log n}\right) \text{ for } 0 < \gamma < 1,$$

Eq. (20) holds. Hence we have two limit theorems which correspond to Eqs. (43) and (44), respectively, where

$$\mathbb{E}\widehat{S}_n(e^{n^\gamma}) \sim \frac{n^{1-\gamma}e^{n^\gamma}}{(\log_{\nu-1} n^\gamma) \prod_{j=1}^{\nu-1} \log_j n^\gamma}$$

and

$$\sqrt{\text{Var}\widehat{S}_n(e^{n^\gamma})} \sim \frac{n^{\frac{1-\gamma}{2}}}{\sqrt{2(\log_{\nu-1} n^\gamma) \prod_{j=1}^{\nu-1} \log_j n^\gamma}} e^{n^\gamma}.$$

In addition to Eq. (45), it turns out that Eq. (20) also holds for

$$c_n = \begin{cases} e^{\sqrt{n}(\log n)^{-1}}, & \text{if } \nu = 1, \\ e^{n(\log n)^{-3}}, & \text{if } \nu \geq 2. \end{cases}$$

## 4.2. Super-Petersburg distributions

In this section, we suppose  $\nu \in \mathbb{N} \cup \{0\}$ . Let  $X^{\nu\text{-sP}}$  be a random variable with

$$\mathbb{P}(X^{\nu\text{-sP}} = \varrho_{\nu+1}(k)) = 2^{-k} \quad \text{for } k \in \mathbb{N}, \quad (46)$$

where

$$\mathcal{L}_\nu(x) = \begin{cases} x, & \text{for } \nu = 0, \\ 2^{\mathcal{L}_{\nu-1}(x)}, & \text{for } \nu \geq 1. \end{cases}$$

Moreover, we also define the  $\nu$  times iterated logarithm of base 2 as follows.

$$\lg_\nu x = \begin{cases} x, & \text{for } \nu = 0, \\ \lg(\lg_{\nu-1} x), & \text{for } \nu \geq 1. \end{cases}$$

We call the distribution defined by Eq. (46) the  $\nu$ -super-Petersburg distribution. In particular,  $X^{0\text{-sP}}$  and  $X^{1\text{-sP}}$  are equivalent to the St. Petersburg random variable  $X^{\text{St}}$  and the super-Petersburg random variable  $X^{\text{sP}}$ , respectively.

**Remark 4.1.** The tail function of  $X^{\text{St}} = X^{0\text{-sP}}$  (see Eq. (3)) is *not* regularly varying with index  $-1$  (see [7, Section 8.8.2, page 373]). Similarly, putting

$$g_{\nu\text{-sP}}(x) = \mathbb{P}(X^{\nu\text{-sP}} > x) = 2^{-\lfloor \lg_{\nu+1} x \rfloor} \quad \text{for } x > 2^{\nu+1}(1), \quad (47)$$

we have  $g_{\nu\text{-sP}}(x)$  is *not* slowly varying for  $\nu \geq 1$ . In fact, if  $t = 2_{\nu+1}(n)$  for  $n \in \mathbb{N}$  and  $x = 1/2$  then

$$\frac{g_{\nu\text{-sP}}(tx)}{g_{\nu\text{-sP}}(t)} = 2^{-\lfloor \lg_{\nu+1}(tx) \rfloor + \lfloor \lg_{\nu+1}(t) \rfloor} = 2^{-\lfloor \lg_{\nu+1}(2_{\nu+1}(n)/2) \rfloor + n} \geq 2,$$

because  $x \mapsto \lg_{\nu+1}(x)$  is strictly increasing. Compare  $g_{\nu\text{-sP}}(x)$  to functions  $x \mapsto 2 + \sin(\log \log x)$  and  $x \mapsto \lfloor \log x \rfloor$  which are both slowly varying.

We consider  $\widetilde{X}_1^{\nu\text{-sP}}$  and  $\widetilde{X}_2^{\nu\text{-sP}}$  in Assumption (A2) with tractable distribution functions instead of  $X^{\nu\text{-sP}}$ . It turns out that it satisfies Assumptions (A1) and (A2) with respect to

$$L(x) = \lg_\nu x \quad \text{and} \quad \varepsilon(x) = \left( (\log 2)^\nu \prod_{j=1}^{\nu} \lg_j x \right)^{-1} \quad \text{for } x \geq e_{\nu+1}(1),$$

$c = \lg_\nu(e_{\nu+1}(1))$ ,  $C_1 = 1$ ,  $C_2 = 2$ ,  $x_0 = 2_{\nu+1}(1)$ ,  $a_0 = e_{\nu+1}(1)$ ,  $a_1 = 2_\nu(1)$ ,  $a_2 = 2_{\nu+1}(1)$  and the probability density functions for  $\widetilde{X}_1^{\nu\text{-sP}}$  and  $\widetilde{X}_2^{\nu\text{-sP}}$  are



respectively

$$f_1(x) = \begin{cases} \frac{1}{(\log 2)^\nu (\lg_\nu x) \prod_{j=0}^{\nu-1} \lg_j x}, & \text{for } x \geq \mathfrak{L}_\nu(1), \\ 0, & \text{for } x < \mathfrak{L}_\nu(1), \end{cases}$$

$$f_2(x) = \begin{cases} \frac{2}{(\log 2)^\nu (\lg_\nu x) \prod_{j=0}^{\nu-1} \lg_j x}, & \text{for } x \geq \mathfrak{L}_{\nu+1}(1), \\ 0, & \text{for } x < \mathfrak{L}_{\nu+1}(1). \end{cases}$$

We assume  $l_n = n$ , and put  $S_n = \sum_{i=1}^n X_i^{\nu\text{-sP}}$ , where  $X_1^{\nu\text{-sP}}, \dots, X_n^{\nu\text{-sP}}$  are iid random variables with (46).

(i) min-truncation: We suppose

$$c_n = \mathfrak{L}_\nu(n^\beta) \quad \text{for } 0 < \beta < 1.$$

Then Eq. (18) holds because  $L(c_n) = n^\beta = o(n/\log n)$ . Hence we have two limit theorems which correspond to Eqs. (43) and (44), respectively, where

$$\begin{aligned} \mathbb{E}S_n(\mathfrak{L}_\nu(n^\beta)) &\sim n2^{-\lfloor \beta \lg n \rfloor} \mathfrak{L}_\nu(n^\beta) \quad \text{and} \\ \sqrt{\text{Var}S_n(\mathfrak{L}_\nu(n^\beta))} &\sim \sqrt{n2^{-\lfloor \beta \lg n \rfloor}} \mathfrak{L}_\nu(n^\beta). \end{aligned} \tag{48}$$

In fact, since  $\mathbb{E}(\widehat{X^{\nu\text{-sP}}}(x))^k = o(\mathbb{E}(X^{\nu\text{-sP}}(x))^k)$  for  $k \in \mathbb{N}$  by Eq. (29), it follows from Eq. (47) that

$$\begin{aligned} \mathbb{E}(X^{\nu\text{-sP}}(x))^k &= \mathbb{E}(\widehat{X^{\nu\text{-sP}}}(x))^k + x^k \mathbb{P}(X^{\nu\text{-sP}} > x) \\ &\sim x^k \mathbb{P}(X^{\nu\text{-sP}} > x) = x^k 2^{-\lfloor \lg_{\nu+1} x \rfloor}, \end{aligned}$$

which yields Eq. (48).

(ii) default-truncation: We suppose

$$c_n = \begin{cases} 2^{n^\beta}, & \text{for } 0 < \beta < \frac{1}{2}, \text{ if } \nu = 1, \\ 2^{n^\gamma}, & \text{for } 0 < \gamma < 1, \text{ if } \nu \geq 2. \end{cases} \tag{49}$$

(a) Consider the case of  $\nu = 1$ . Since  $L(2^{n^\beta})/\varepsilon(2^{n^\beta}) = (\log 2)n^{2\beta} = o(n/\log n)$  for  $0 < \beta < \frac{1}{2}$ , Eq. (20) holds. Hence we have two limit theorems which correspond to Eqs. (43) and (44), respectively. However, it is not very easy to give simple forms both  $\mathbb{E}\widehat{S}_n(2^{n^\beta})$  and  $\sqrt{\text{Var}\widehat{S}_n(2^{n^\beta})}$ .

(b) Consider the case of  $\nu \geq 2$ . Since

$$\frac{L(2^{n^\gamma})}{\varepsilon(2^{n^\gamma})} = n^\gamma (\log 2)^\nu (\lg_{\nu-1} n^\gamma) \prod_{j=1}^{\nu-1} \lg_j n^\gamma = o\left(\frac{n}{\log n}\right) \text{ for } 0 < \gamma < 1,$$

Eq. (20) holds. Hence we have two limit theorems which correspond to Eqs. (43) and (44), respectively.

In addition to Eq. (49), it turns out that Eq. (20) also holds for

$$c_n = \begin{cases} 2^{\sqrt{n}(\log n)^{-1}}, & \text{if } \nu = 1, \\ 2^{n(\log n)^{-3}}, & \text{if } \nu \geq 2. \end{cases}$$

### Acknowledgments

The author would like to thank Professor A. G. Pakes for reading an earlier version carefully and making many comments and suggestions, which are valuable and very helpful for revising and improving it.

### References

1. A. Adler, Exact Strong laws, *Bull. Inst. Math. Acad. Sinica*, **28** (2000), 141-166.
2. A. Adler and A. Pakes, On relative stability and weighted laws of large numbers, *Extremes*, **20** (2017), 1-31.
3. A. Adler and A. Pakes, Weak and one-sided strong laws for random variables with infinite mean, *Stat. Prob. Lett.*, **142** (2018), 8-16.
4. I. Alves, L. de Haan and C. Nevesc, A test procedure for detecting super-heavy tails, *J. Stat. Plan. Infer.*, **139** (2009), 213-227.
5. D. Bernoulli, Exposition of a new theory on the measurement of risk, *Econometrica*, **22**, (1954), 23-36.
6. P. Billingsley, *Probability and Measure*, Anniv. ed., Wiley, 2012.
7. N. Bingham, C. Goldie and J. Teugels, *Regular Variation*, Cambridge UP, 1989.
8. S. Boucheron, G. Lugosi and P. Massart, *Concentration inequalities, a nonasymptotic theory of independence*, Oxford University Press, 2013.
9. Y. Chow and H. Teicher, *Probability Theory*, 3rd ed., Springer, 1997.
10. U. Cormann and R. Reiss, Generalizing the Pareto to the log-Pareto model and statistical inference, *Extremes*, **12** (2009), 93-105.

11. D. A. Darling, The influence of the maximum term in the addition of independent random variables, *Trans. Amer. Math. Soc.*, **73** (1952), 95-107.
12. J. Dutka, On the St. Petersburg paradox, *Arch. Hist. Exact Sci.*, **39** (1988), 13-39.
13. P. Embrechts, C. Klüppelberg and T. Mikosch, *Modelling Extremal Events: for Insurance and Finance*, Springer, 1997.
14. M. Falk, J. Hüsler and R. Reiss, *Laws of Small Numbers: extremes and rare events*, 3rd ed., Birkhäuser/Springer, 2011.
15. W. Feller, *An Introduction to Probability Theory and Its Applications*, Vol. 1, 3rd ed., Wiley, 1968.
16. S. Foss, D. Korshunov and S. Zachary, *An Introduction to Heavy-Tailed and Subexponential Distributions*, 2nd ed., Springer, 2013.
17. J. Galambos, *The Asymptotic Theory of Extreme Order Statistics*, Wiley, 1978.
18. A. Gut, *Probability: a graduate course*, 2nd ed., Springer, 2013.
19. L. Györfi and P. Kevei, On the rate of convergence of the St. Petersburg game, *Period. Math. Hung.*, **62** (2011), 13-37.
20. L. de Haan and A. Ferreira, *Extreme Value Theory*, Springer, 2006.
21. A. Martin-Löf, A limit theorem which clarifies the "Petersburg paradox", *J. Appl. Probab.*, **22** (1985), 634-643.
22. K. Matsumoto and T. Nakata, Limit theorems for a generalized Feller game, *J. Appl. Probab.*, **50** (2013), 54-63.
23. S. Nagaev and V. Vachtel, On sums of independent random variables without power moments, *Sib. Math. J.* **49** (2008), 1091-1100.
24. T. Nakata, Limit theorems for nonnegative independent random variables with truncation, *Acta Math. Hung.*, **145** (2015), 1-16.
25. T. Nakata, Weak laws of large numbers for weighted independent random variables with infinite mean, *Stat. Prob. Lett.*, **109** (2016), 124-129.
26. T. Nakata, Exact laws of large numbers for independent Pareto random variables, *Bull. Inst. Math. Acad. Sinica (N.S.)*, **12** (2017), 325-338.
27. V. Petrov, *Limit theorems of probability theory: sequences of independent random variables*, Oxford Univ. Press, 1995.
28. P. Samuelson, St. Petersburg paradoxes: defanged, dissected, and historically described, *J. Econ. Liter.*, **15** (1977), 24-55.
29. E. Seneta, *Regularly varying functions*, Lect. Notes in Math., **508**, (1976).
30. T. Shimura and T. Watanabe, Infinite divisibility and generalized subexponentiality, *Bernoulli*, **11** (2005), 445-469.
31. G. Székely and D. Richards, The St. Petersburg paradox and the crash of high-tech stocks in 2000, *Amer. Stat.*, **58** (2004), 225-231.
32. A. Zeevi and P. Glynn, Recurrence properties of autoregressive processes with super-heavy-tailed innovations, *J. Appl. Probab.*, **41** (2004), 639-653.