Limit laws for the norms of extremal samples

Péter Kevei, Lillian Oluoch, László Viharos

Bolyai Institute, University of Szeged, Aradi vértanúk tere 1, 6720, Szeged, Hungary

Abstract

Denote $S_n(p) = k_n^{-1} \sum_{i=1}^{k_n} (\log(X_{n+1-i,n}/X_{n-k_n,n}))^p$, where $p > 0$, $k_n \leq n$ is a sequence of integers such that $k_n \to \infty$ and $k_n/n \to 0$, and $X_{1,n} \leq \ldots \leq X_{n,n}$ are the order statistics of iid random variables $X_1, \ldots, X_n$ with regularly varying upper tail of index $1/\gamma$. The estimator $\widehat{\gamma}(n) = (S_n(p)/\Gamma(p+1))^{1/p}$ is an extension of the Hill estimator. We investigate the asymptotic properties of $S_n(p)$ and $\widehat{\gamma}(n)$ both for fixed $p > 0$ and for $p = p_n \to \infty$. We prove consistency for $\widehat{\gamma}(n)$ and limit theorem for $\widehat{\gamma}(n) - \gamma$ under appropriate assumptions. We obtain both Gaussian and non-Gaussian (stable) limit depending on the growth rate of the power sequence p_n. Applied to real data we find that for larger p the estimator is less sensitive to the change in k_n than the Hill estimator.

Keywords: tail index, Hill estimator, residual estimator, regular variation, power sum, stable distribution
2010 MSC: 62G32, 60F05

1. Introduction

Let $X, X_1, X_2, \ldots$ be independent identically distributed (iid) random variables with common distribution function $F(x) = \mathbb{P}(X \leq x)$, $x \in \mathbb{R}$. For each $n \geq 1$, let $X_{1,n} \leq \ldots \leq X_{n,n}$ denote the order statistics of the sample $X_1, \ldots, X_n$. Assume that

$$1 - F(x) = x^{-1/\gamma} L(x),$$

where L is a slowly varying function at infinity and $\gamma > 0$. This is equivalent to the condition

$$Q(1-s) = s^{-\gamma} \ell(s), \tag{1}$$

where $Q(s) = \inf\{x : F(x) \geq s\}$, $s \in (0,1)$, stands for the quantile function, and ℓ is a slowly varying function at 0. For $p > 0$ introduce the notation

$$S_n(p) = \frac{1}{k_n} \sum_{i=1}^{k_n} \left(\log \frac{X_{n+1-i,n}}{X_{n-k_n,n}} \right)^p . \tag{2}$$

Email addresses: kevei@math.u-szeged.hu (Péter Kevei), oluoch@math.u-szeged.hu (Lillian Oluoch), viharos@math.u-szeged.hu (László Viharos)

Preprint submitted to Elsevier *June 16, 2021*

The main object of the present paper is the estimate

$$\widehat{\gamma}(n)=\left(\frac{S_n(p)}{\Gamma(p+1)}\right)^{\frac{1}{p}} \tag{3}$$

of the tail index, where Γ is the usual gamma function. In what follows we always assume that $1\leq k_n\leq n$ is a sequence of integers such that $k_n\to\infty$ and $k_n/n\to 0$.

As a special case for $p=1$ we obtain the well-known Hill estimator of the tail index $\gamma>0$ introduced by Hill in 1975 [21]. For $p=2$ the estimator was suggested by Dekkers et al. [15], where they proved that $S_n(2)\to 2\gamma^2$ a.s. or in probability, depending on the assumptions on k_n, and they proved asymptotic normality of the estimator as well. For general $p>0$ the properties of the estimator $\widehat{\gamma}(n)$ in (3) were investigated by Gomes and Martins [19]. Under second-order regular variation assumption they proved weak consistency and asymptotic normality of the estimator $\widehat{\gamma}(n)$. Segers [28] considered more general estimators of the form

$$\frac{1}{k_n}\sum_{i=1}^{k_n} f\left(\frac{X_{n+1-i,n}}{X_{n-k_n,n}}\right), \tag{4}$$

for a nice class of functions f, called *residual estimators*. Segers proved weak consistency and asymptotic normality under general conditions. More recently, Ciuperca and Mercadier [7] investigated weighted version of (2). The residual estimator of Segers was further analyzed for special function classes. Paulauskas and Vaičiulis [23] considered estimators of the form (4) with $f(x)=x^r(\log x)^p$. The classical Hill estimator can be considered as the logarithm of the geometric mean of the variables $X_{n+1-i,n}/X_{n-k,n}$. Based on this interpretation, Brilhante et al. [5] introduced the *mean of order p tail index estimator*, Beran et al. [2] introduced the *harmonic moment tail index estimator*, while very recently Penalva et al. [24] introduced the *Lehmer mean-of-order-p extreme value index estimator*. For a general overview on the generalizations of the Hill estimator we refer to [24].

To the best of our knowledge the possibility $p=p_n\to\infty$ in (3) was not considered before, which is the main focus of our paper. The estimate $\widehat{\gamma}(n)$ can be considered as $p_n\to\infty$ as the limit law for the norm of the extremal sample. In this direction Schlather [27], Bogachev [4], and Janßen [22] proved limit theorems for norms of iid samples.

In the present paper we investigate the asymptotic properties of $S_n(p_n)$ and $\widehat{\gamma}(n)$ both for $p>0$ fixed and for $p=p_n\to\infty$. Although the focus of the paper is to obtain asymptotics for large p, in the course we obtain new results for p fixed. In Section 2 in Theorem 2.1 we prove strong consistency of the estimator for p fixed. Strong consistency was only obtained by Dekkers et al. [15] for $p=1$ and $p=2$, thus our result is new for general p. Asymptotic normality was obtained in several papers for different generalizations of the Hill estimator, see e.g. Gomes and Martins [19], Segers [28], Paulauskas and Vaičiulis [23], and

Penalva et al. [24] for more general estimators. In all these results second-order regular variation is assumed. In Theorem 2.4 our assumptions on the slowly varying function ℓ are weaker, therefore the asymptotic normality in this generality is new. Our main results are contained in Section 3, where we obtain weak consistency and asymptotic normality when $p = p_n \to \infty$. Under appropriate assumptions on the power sequence p_n we prove non-Gaussian stable limit theorems. Section 4 contains a small simulation study and data analysis. Here we show that for larger values of p the estimator is not so sensitive to the choice of k_n, which is a critical property in applications. The use of larger p values was already suggested in [19] for $p > 0$ fixed. We illustrate this property on the well-known dataset of Danish fire insurance claims, see Resnick [26] and Embrechts et al. [17, Example 6.2.9]. Some auxiliary results and the proofs for fixed p are contained in Section 5. In Section 6 we analyze the asymptotic behavior of the power sums. These results are extensions of Bogachev's results in [4], and are needed to prove the limit properties of $\widehat{\gamma}(n)$ as $p_n \to \infty$. Finally, the proofs of the large p asymptotics are in Section 7.

2. Results for fixed p

In what follows, $U, U_1, U_2, \ldots$ are iid uniform$(0,1)$ random variables, and $U_{1,n} \leq U_{2,n} \leq \ldots \leq U_{n,n}$ stand for the corresponding order statistics. To ease notation we frequently suppress the dependence on n and simply write $k = k_n$. Define $X = Q(1-U)$, $X_i = Q(1-U_i)$ for $i = 1, 2, \ldots$. According to the well-known quantile representation, $X, X_1, X_2, \ldots$ is an iid sequence with common distribution function F, which implies that S_n in (2) can be written as

$$S_n(p) = \frac{1}{k} \sum_{i=1}^{k} \left(\log \frac{Q(1-U_{i,n})}{Q(1-U_{k+1,n})} \right)^p \quad \text{for each } n \geq 1, \text{ a.s.} \tag{5}$$

First we show strong consistency for $S_n(p)$. Our assumption on the sequence k_n is the same as in Theorem 2.1 in [15]. This is not far from the optimal condition $k_n / \log\log n \to \infty$, which was obtained by Deheuvels et al. [14] for $p = 1$. In what follows any nonspecified limit is meant as $n \to \infty$.

Theorem 2.1. *Assume that* (1) *holds and* $k_n/n \to 0$, $(\log n)^\delta / k_n \to 0$ *for some* $\delta > 0$. *Then* $S_n(p) \to \gamma^p \Gamma(p+1)$ *a.s., that is for* $p > 0$ *fixed the estimator* $\widehat{\gamma}(n)$ *is strongly consistent.*

Weak consistency holds under weaker assumptions on k_n. The following result is a special case of Theorem 2.1 in [28], and it follows from the representation (5) and from the law of large numbers.

Theorem 2.2. *Assume that* (1) *holds, and the sequence* (k_n) *is such that* $k_n \to \infty$, $k_n/n \to 0$. *Then* $S_n(p) \xrightarrow{\mathbb{P}} \gamma^p \Gamma(p+1)$, *that is for* $p > 0$ *fixed the estimator* $\widehat{\gamma}(n)$ *is weakly consistent.*

To prove asymptotic normality we use representation (5) where the summands are independent and identically distributed conditioned on $U_{k+1,n}$. Indeed, conditioned on $U_{k+1,n}$

$$(U_{1,n}, \ldots, U_{k,n}) \stackrel{\mathcal{D}}{=} \left(\widetilde{U}_{1,k} U_{k+1,n}, \ldots, \widetilde{U}_{k,k} U_{k+1,n} \right), \tag{6}$$

where $\widetilde{U}_1, \widetilde{U}_2, \ldots$ are iid uniform$(0,1)$ random variables, independent of $U_{k+1,n}$, and $\widetilde{U}_{1,k} < \ldots < \widetilde{U}_{k,k}$ stands for the order statistics of $\widetilde{U}_1, \ldots, \widetilde{U}_k$.

To state the result, we need some notation. Introduce the variable for $v \in (0,1)$

$$Y(v) = \log \frac{Q(1-Uv)}{Q(1-v)}, \tag{7}$$

where U is uniform$(0,1)$, and $Y(0) = -\gamma \log U$. Note that $Y(v)$ is 'continuous' in v at 0, that is $Y(0) = \lim_{v \downarrow 0} Y(v)$, since for the slowly varying function ℓ in (1) we have $\lim_{v \downarrow 0} \ell(vU)/\ell(v) = 1$ a.s. Define

$$m_{p,\gamma}(v) = m_p(v) = \mathbb{E}\left[(Y(v))^p \right], \quad \sigma^2_{p,\gamma}(v) = \sigma^2_p(v) = \mathbb{V}\mathrm{ar}\left[(Y(v))^p \right], \tag{8}$$

and the corresponding limiting quantities

$$\begin{aligned} m_p &= m_{p,\gamma} = \mathbb{E}\left[(-\gamma \log U)^p \right] = \gamma^p \Gamma(p+1), \\ \sigma^2_p &= \sigma^2_{p,\gamma} = \mathbb{V}\mathrm{ar}[(-\gamma \log U)^p] = \gamma^{2p} \left(\Gamma(2p+1) - \Gamma^2(p+1) \right). \end{aligned} \tag{9}$$

Note that these quantities depend on the parameter γ. However, since the value $\gamma > 0$ is fixed, to ease notation we suppress γ.

A central limit theorem with random centering was obtained in Theorem 4.1 in [28]. Next, we spell out this result in our case. In the special case $p = 1$ we obtain Theorem 1.6 by Csörgő and Mason [9]. The key observation in the proof is representation (6). Recall the definition of the centering sequence from (8).

Theorem 2.3. *Assume that* (1) *holds, and* $k_n \to \infty$, $k_n/n \to 0$. *Then*

$$\sqrt{k_n} \left(S_n(p) - m_p(U_{k+1,n}) \right) \xrightarrow{\mathcal{D}} N(0, \sigma^2_p).$$

To obtain asymptotic normality for the estimator, that is, to change the random centering $m_p(U_{k+1,n})$ to m_p, we have to show that

$$\sqrt{k_n}(m_p(U_{k+1,n}) - m_p) \xrightarrow{\mathbb{P}} 0.$$

Since $U_{k+1,n} n/k \to 1$ in probability, this is the same as the deterministic convergence

$$\sqrt{k_n}(m_p(k/n) - m_p) \to 0;$$

see the proof of Theorem 2.4 for the precise version. In the case of the Hill estimator ($p = 1$) Csörgő and Viharos [10] obtained optimal conditions under which the random centering $m_p(U_{k+1,n})$ in Theorem 2.3 can be replaced by the deterministic one, $m_p(k/n)$. For the general residual estimator this was obtained in Theorem 4.2 in [28]. In Theorem 4.5 in [28] assuming that the slowly varying function ℓ belongs to the de Haan class Π, conditions were obtained which ensure that the random centering can be replaced by the limit m_p. Our assumptions are weaker, but some second-order conditions are necessary.

Assume that there exist a regularly varying function a and a Borel set $B \subset [0,1]$ of positive measure such that

$$\lim_{v \downarrow 0} \frac{a(v)}{\ell(v)} = 0, \quad \limsup_{v \downarrow 0} \frac{|\ell(uv) - \ell(v)|}{a(v)} < \infty \quad \text{for } u \in B. \tag{10}$$

By Theorem 3.1.4 in Bingham et al. [3] condition (10) implies that the limsup in (10) is finite uniformly on any compact subset of $(0,1]$. However, in general, uniformity cannot be extended to $[0,1]$.

We emphasize that we do not need exact second-order asymptotics for ℓ, only bounds. In particular, if ℓ belongs to the de Haan class Π (defined at 0) then the condition (10) holds; see Appendix B in de Haan and Ferreira [12], or Chapter 3 in Bingham et al. [3]. Therefore, even in the special case $p = 1$, that is, for the Hill estimator, our next result is a generalization of Theorem 3.1 in [15]. The asymptotic normality of various generalizations of the Hill estimator are obtained under second-order regular variation for ℓ, see Theorem 4.5 in [28], formula (2.7) in [19], or Theorem 2 in [23]. Our conditions in the next result are weaker.

Theorem 2.4. *Assume that* (10) *holds for ℓ, and k_n is such that $k_n \to \infty$, $k_n/n \to 0$, and*

$$\sqrt{k_n} \frac{a(k_n/n)}{\ell(k_n/n)} \to 0. \tag{11}$$

Then, with $\sigma_p^2 = \gamma^{2p}(\Gamma(2p+1) - \Gamma^2(p+1))$,

$$\frac{\sqrt{k_n}}{\sigma_p} \left(S_n(p) - \gamma^p \Gamma(p+1) \right) \xrightarrow{\mathcal{D}} N(0,1),$$

and

$$\frac{p\sqrt{k_n}}{\gamma} \left(\frac{\Gamma(2p+1)}{\Gamma^2(p+1)} - 1 \right)^{-1/2} (\widehat{\gamma}(n) - \gamma) \xrightarrow{\mathcal{D}} N(0,1).$$

The asymptotic variance of $\widehat{\gamma}(n)$ is the same as obtained in (2.7) in [19].

We point out that the growth condition (11) of the subsequence is the same as in Theorem 4.5 in [29] and in the special case $p = 1$ in de Haan [11]. Assuming second-order regular variation, the asymptotic normality of the estimator was proved under the less restrictive condition

$$\sqrt{k_n} \frac{a(k_n/n)}{\ell(k_n/n)} \to \lambda, \quad \lambda \in \mathbb{R},$$

by de Haan and Peng [13, Theorem 1] (see also [17, Theorem 6.4.9]) for the Hill estimator, and in [19] for general p. For more general estimators the asymptotic normality was proved under the condition above, see Theorem 2 in [23], Theorem 2 in [2], Theorem 2 in [5].

3. Asymptotics for large p

Conditioned on $U_{k+1,n}$ the sum $k_n S_n(p_n)$ in (5) is the sum of k_n iid random variables distributed as $Y(U_{k+1,n})$. This allows us to use the appropriate uniform version of the results in [4] for power sums. These results are spelled out and proved in Section 6. As a consequence, we obtain limit theorems with *random centering and norming* for $S_n(p_n)$. In order to change to deterministic centering a precise analysis is needed.

First we need some notation. Let

$$f_v(x) = x^\gamma \ell(v/x), \quad v \in (0,1], \quad f_0(x) = x^\gamma, \quad x > 1.$$

Note that $Y(v)$ is defined for $v \in [0,1)$, while f_v is defined for $v \in [0,1]$. Then f_v is a left-continuous, nondecreasing, regularly varying function at infinity with index γ. Its inverse

$$g_v(y) = \inf\{x : f_v(x) > y\} = v g_1(y/v^\gamma), \quad v \in (0,1], \quad g_0(y) = y^{1/\gamma},$$

is regularly varying with index $1/\gamma$, see Theorem 1.5.12 in [3]. Write $f = f_1$ and $g = g_1$. Then, $g(x) = x^{1/\gamma}\widetilde{\ell}(x)$, for a slowly varying function $\widetilde{\ell}$ such that

$$\ell(1/x)^{1/\gamma}\widetilde{\ell}(x^\gamma \ell(1/x)) \sim 1 \quad \text{as } x \to \infty. \tag{12}$$

The latter follows from the fact $f(g(x)) \sim g(f(x)) \sim x$. In fact, $\widetilde{\ell}(x)^\gamma$ is the de Bruijn conjugate of $\ell(1/x^{1/\gamma})$, see [3, Section 1.7].

Using that $f_v(x) > y$ if and only if $x > g_v(y)$, for $v \in (0,1]$ fixed the tail of $Y(v)$ is

$$\begin{aligned}
\mathbb{P}(Y(v) > x) &= \mathbb{P}\left(\log U^{-\gamma}\frac{\ell(Uv)}{\ell(v)} > x\right) \\
&= \mathbb{P}(U^{-\gamma}\ell(Uv) > \ell(v)e^x) \\
&= \mathbb{P}(U^{-1} > g_v(\ell(v)e^x)) \\
&= e^{-x/\gamma}\left[\ell(v)^{1/\gamma}\widetilde{\ell}(v^{-\gamma}\ell(v)e^x)\right]^{-1},
\end{aligned}$$

and for $v = 0$ we have $\mathbb{P}(Y(0) > x) = e^{-x/\gamma}$. Thus, we obtain that the log-tail distribution function

$$h_v(x) := -\log \mathbb{P}(Y(v) > x) = \begin{cases} \frac{x}{\gamma} + \log\left(\ell(v)^{\frac{1}{\gamma}}\widetilde{\ell}(v^{-\gamma}\ell(v)e^x)\right), & v \in (0,1], \\ \frac{x}{\gamma}, & v = 0. \end{cases} \tag{13}$$

For any fixed $v \in [0,1]$ we have that $h_v(x) \sim x/\gamma$. In particular, it is regularly varying with index 1. Following [4], for $\zeta > 0$ let us define η_v as the unique solution to

$$h_v(\eta_v(x)) = \zeta x. \tag{14}$$

3.1. Weak laws and Gaussian limit

It is pointed out in [4] that the proper rate of the power sequence p_n is $\log k_n$. Let us define the parameter ζ as

$$\zeta = \liminf_{n\to\infty} \frac{\log k_n}{p_n}. \tag{15}$$

For $\zeta \le 2$ we need a precise assumption on the power sequence, and we assume that

$$k_n \sim e^{\zeta p_n}. \tag{16}$$

Note that we have different definitions for ζ depending on its range. In the results below we always state which of the two conditions we assume.

For the truncated moments for $v \in [0,1)$ put

$$\begin{aligned} m_p^1(v) &= \mathbb{E}[(Y(v))^p I(Y(v) \le \eta_v(p))] \\ \sigma_p^1(v) &= \left(\mathbb{E}\left[(Y(v))^{2p} I(Y(v) \le \eta_v(p))\right]\right)^{1/2}, \end{aligned}$$

where I stands for the indicator function. Recall (8) and define the centering and norming functions for $v \in [0,1)$

$$\widetilde{m}_p(v) = \begin{cases} 0, & \zeta \in (0,1), \\ m_p^1(v), & \zeta = 1, \\ m_p(v), & \zeta > 1, \end{cases} \qquad \widetilde{\sigma}_p(v) = \begin{cases} \sigma_p(v), & \zeta > 2, \\ \sigma_p^1(v), & \zeta = 2. \end{cases} \tag{17}$$

To ease notation put $m_p^1 = m_p^1(0)$, $\sigma_p^1 = \sigma_p^1(0)$, $\widetilde{m}_p = \widetilde{m}_p(0)$, and $\widetilde{\sigma}_p = \widetilde{\sigma}_p(0)$.

Weak consistency holds for $\zeta \ge 1$, while asymptotic normality holds for $\zeta \ge 2$. Note that in the borderline cases $\zeta = 1, 2$ the norming is different, and the condition on the subsequence p_n is stronger.

Theorem 3.1. *Assume that $k_n \to \infty$, $k_n/n \to 0$, and $p_n \to \infty$. If $\zeta > 1$ in* (15) *or $\zeta = 1$ in* (16) *then*

$$\left(\widetilde{m}_{p_n}(U_{k_n+1,n})\right)^{-1} S_n(p_n) \xrightarrow{\mathbb{P}} 1. \tag{18}$$

In both cases $\widehat{\gamma}(n)$ is weakly consistent. Furthermore, if $\zeta > 2$ in (15) *or $\zeta = 2$ in* (16) *then*

$$\frac{\sqrt{k_n}}{\widetilde{\sigma}_{p_n}(U_{k+1,n})} \left(S_n(p_n) - \widetilde{m}_{p_n}(U_{k+1,n})\right) \xrightarrow{\mathcal{D}} N(0,1), \tag{19}$$

and

$$\frac{\sqrt{k_n}\,\widetilde{m}_{p_n}(U_{k+1,n})}{\widetilde{\sigma}_{p_n}(U_{k+1,n})} p_n \left[\left(\frac{S_n(p_n)}{\widetilde{m}_{p_n}(U_{k+1,n})}\right)^{1/p_n} - 1\right] \xrightarrow{\mathcal{D}} N(0,1). \tag{20}$$

Note that both the centering and the norming are random. To change to deterministic values $\widetilde{m}_{p_n}$ and $\widetilde{\sigma}_{p_n}$ further assumptions are needed. We always assume that for the slowly varying function (10) holds. For sequences, $k_n \to \infty$, $k_n/n \to 0$, and $p_n \to \infty$ introduce the notation

$$\beta_2 = \limsup_{n\to\infty} -p_n^{-1} \log \frac{a(k_n/n)}{\ell(k_n/n)} \geq \liminf_{n\to\infty} -p_n^{-1} \log \frac{a(k_n/n)}{\ell(k_n/n)} = \beta_1, \tag{21}$$

allowing $\beta_1 = \infty$, and let

$$\beta = \begin{cases} \beta_1, & \text{if } \beta_1 \geq 1, \\ \beta_2, & \text{if } 0 < \beta_1 \leq \beta_2 \leq 1 \\ 1, & \text{otherwise.} \end{cases} \tag{22}$$

Put $a \vee b = \max\{a, b\}$, $a \wedge b = \min\{a, b\}$. Introduce the notation

$$H(u) = u - 1 - \log u, \quad u > 0, \tag{23}$$

and for $x \in (0, \infty]$

$$\nu_x = x^{-1} H(2 \vee 2x), \quad \nu_\infty = 2.$$

Then ν is decreasing on $(0, 1]$, and increasing on $[1, \infty)$.

Theorem 3.2. *Assume that for the slowly varying function ℓ (10) holds and $\beta_1 > 0$. If $\zeta > 1$ in (15) or $\zeta = 1$ in (16) then*

$$(\widetilde{m}_{p_n})^{-1} S_n(p_n) \xrightarrow{\mathbb{P}} 1. \tag{24}$$

If $\zeta > 2$ in (15) or $\zeta = 2$ in (16) then assume additionally that for some $\varepsilon > 0$

$$\limsup_{n\to\infty} p_n^{-1} \log \left(\sqrt{k_n} \left(\frac{a(k_n/n)}{\ell(k_n/n)} \right)^{(\nu_\beta - \varepsilon)\wedge 1} \right) < \log 2.$$

Then

$$\frac{\sqrt{k_n}}{\widetilde{\sigma}_{p_n}} \left(S_n(p_n) - \widetilde{m}_{p_n} \right) \xrightarrow{\mathcal{D}} N(0, 1), \tag{25}$$

and

$$\frac{\sqrt{k_n}\widetilde{m}_{p_n}}{\gamma \widetilde{\sigma}_{p_n}} p_n \left(\widehat{\gamma}(n) - \gamma \right) \xrightarrow{\mathcal{D}} N(0, 1). \tag{26}$$

If $\zeta > 2$ then $\widetilde{m}_p = m_p$ and $\widetilde{\sigma}_p = \sigma_p$, where the latter quantities are defined in (9). A simple application of the Stirling formula gives that $m_p/\sigma_p \sim 2^{-p}(p\pi)^{1/4}$ as $p \to \infty$.

Under stronger assumptions on the slowly varying function ℓ it is possible to weaken the conditions on k_n and p_n. A stronger condition on ℓ is that the limsup in (10) is finite uniformly in $u \in (0, 1]$, that is there exists a regularly varying function a such that

$$\lim_{v\downarrow 0} \frac{a(v)}{\ell(v)} = 0, \quad \limsup_{v\downarrow 0} \sup_{u\in(0,1]} \frac{|\ell(uv) - \ell(v)|}{a(v)} =: K_1 < \infty. \tag{27}$$

Theorem 3.3. *Assume that for the slowly varying function ℓ (27) hold. Furthermore, $k_n \to \infty$, $k_n/n \to 0$, and $p_n \to \infty$ such that*

$$p_n \frac{a(k_n/n)}{\ell(k_n/n)} \to 0.$$

If $\zeta > 1$ in (15) or $\zeta = 1$ in (16) then (24) holds. If $\zeta > 2$ in (15) or $\zeta = 2$ in (16), and

$$\limsup_{n\to\infty} p_n^{-1} \log \left(\sqrt{k_n} \frac{a(k_n/n)}{\ell(k_n/n)} \right) < \log 2$$

then (25) and (26) hold.

3.2. Non-Gaussian stable limits

Next, we explore the regime $\zeta < 2$. Here we need the precise asymptotic assumption (16) on the power sequence p_n. We obtain non-Gaussian limits, where the characteristic exponent of the stable law equals ζ, coming from the growth rate of the power sequence p_n. Therefore, in what follows we use the notation $\zeta = \alpha$.

Let Z_α denote a one-sided α-stable random variable with characteristic function

$$\mathbb{E}e^{\mathrm{i}tZ_\alpha} = \begin{cases} \exp\left\{-\Gamma(1-\alpha)|t|^\alpha e^{-\mathrm{i}\frac{\pi\alpha}{2}\operatorname{sgn}(t)}\right\}, \\ \exp\left\{\mathrm{i}t(1-a) - \frac{\pi}{2}|t|\left(1+\mathrm{i}\operatorname{sgn}(t)\frac{2}{\pi}\log|t|\right)\right\}, \end{cases}$$

where $a = 0.577\ldots$ stands for the Euler–Mascheroni constant.

Theorem 3.4. *Assume that $k_n \to \infty$, $k_n/n \to 0$, and $p_n \to \infty$ such that (16) holds for some $\zeta = \alpha \in (0,2)$. Then*

$$\frac{k_n}{\eta_{U_{k_n+1,n}}(p_n)^{p_n}} \left(S_n(p_n) - \widetilde{m}_{p_n}(U_{k+1,n})\right) \xrightarrow{\mathcal{D}} Z_\alpha.$$

Moreover, for $\zeta = \alpha \in (0,1)$

$$p_n \left(\frac{[k_n S_n(p_n)]^{1/p_n}}{\eta_{U_{k_n+1,n}}(p_n)} - 1 \right) \xrightarrow{\mathcal{D}} \log Z_\alpha, \tag{28}$$

in particular,

$$\widehat{\gamma}(n) \xrightarrow{\mathbb{P}} \gamma\, \alpha e^{1-\alpha}. \tag{29}$$

While for $\alpha \in [1,2)$

$$p_n \frac{k_n \widetilde{m}_{p_n}(U_{k_n+1,n})}{\eta_{U_{k_n+1,n}}(p_n)^{p_n}} \left[\left(\frac{S_n(p_n)}{\widetilde{m}_{p_n}(U_{k_n+1,n})} \right)^{1/p_n} - 1 \right] \xrightarrow{\mathcal{D}} Z_\alpha. \tag{30}$$

In order to use deterministic norming and centering we need further assumptions on the slowly varying function. Note that by (14) with $\zeta = \alpha$ we obtain $\eta_0(x) = \alpha\gamma x$. Recall (21) and (22).

Theorem 3.5. *Assume* (16) *and that* (10) *holds. Furthermore,* $k_n \to \infty$, $k_n/n \to 0$, *and*

$$\widetilde{\ell}\,(n^\gamma \ell(k/n)) \sim \widetilde{\ell}\,((n/k)^\gamma \ell(k/n)) \tag{31}$$

and for $\alpha \in [1,2)$ *assume that*

$$\nu_\beta \beta_1 > \alpha - 1 - \log\alpha = H(\alpha). \tag{32}$$

Then for $\alpha \in (0,2)$

$$\frac{k_n}{(\alpha\gamma p_n)^{p_n}}\left(S_n(p_n) - \widetilde{m}_{p_n}\right) \xrightarrow{\mathcal{D}} Z_\alpha. \tag{33}$$

For the estimator $\widehat{\gamma}(n)$ *if* $\alpha \in (0,1)$

$$\frac{e^{\alpha-1}}{\alpha\gamma} p_n \left[\widehat{\gamma}(n)\left(1 + \frac{\log p_n}{2p_n}\right) - \gamma\alpha e^{1-\alpha}\right] \xrightarrow{\mathcal{D}} \log Z_\alpha - \frac{\log 2\pi}{2}. \tag{34}$$

while for $\alpha \in (1,2)$

$$\frac{\sqrt{2\pi}}{\gamma} e^{p_n(\alpha-1-\log\alpha)} p_n^{3/2}\left[\widehat{\gamma}(n) - \gamma\right] \xrightarrow{\mathcal{D}} Z_\alpha, \tag{35}$$

and for $\alpha = 1$

$$\frac{\sqrt{2\pi}}{2\gamma} p_n^{3/2}\left[\widehat{\gamma}(n)\left(1 + \frac{\log 2}{p_n}\right) - \gamma\right] \xrightarrow{\mathcal{D}} Z_1. \tag{36}$$

Condition (31) is rather implicit, since already $\widetilde{\ell}$ is implicit. However, from the proof it will be clear that this is exactly what is needed. In some natural special cases it can be checked. For example, (27) implies (31). Under some general growth conditions the de Bruijn conjugate, and so $\widetilde{\ell}$ can be determined explicitly, see [3, Corollary 2.3.4].

If $\beta = \beta_1$ then $\nu_\beta \beta_1 = H(2 \vee 2\beta) \geq H(2) > H(\alpha)$, that is condition (32) is automatic.

Under stronger assumptions on ℓ the result can be simplified.

Theorem 3.6. *Assume* (16) *and that* (27) *holds. Furthermore,* $k_n \to \infty$, $k_n/n \to 0$, *and for* $\alpha \in [1,2)$ *assume that* $\beta_1 > H(\alpha)$. *Then* (33), *and depending on the value* α, (34), (35), *or* (36) *hold.*

3.3. Examples

We spell out our results in three special cases. First, we consider the exact Pareto model, when $\ell \equiv 1$. Next, we consider the Hall model, when (27) holds. Finally, we consider the nonconstant slowly varying function $\ell(s) = -\log s$.

Example 1. The simplest special case is the strict Pareto model, when in (1) $\ell \equiv 1$. Then $m_p(v) \equiv m_p = \gamma^p \Gamma(p+1)$, thus the centering and norming do not depend on v. Furthermore, there is no other restriction on the sequence, only $k_n \to \infty$. In fact, $k_n = n$ is possible. Assume that $e^{\zeta p_n} \sim k_n$. Then, a direct consequence of Proposition 10.4 in [4] is that depending on the value ζ, (26), (34), (35), or (36) hold.

Example 2. Assume that the slowly varying function ℓ in (1) has the form

$$\ell(u) = c + O(u^{\delta}) \quad \text{with } \ c > 0, \delta > 0.$$

The asymptotic normality of the Hill estimator was proved for this subclass by Hall [20]. Condition (27) is satisfied with $a(u) = u^{\delta}$. By Proposition 5.5, for some $C > 0$

$$|m_{p_n}(u) - m_{p_n}| \le C\Gamma(p_n + 1)\gamma^{p_n} u^{\delta}.$$

Let $p_n = \zeta^{-1} \log k_n$. For $\zeta \ge 2$ assume

$$\limsup_{n\to\infty} \frac{1}{p_n} \log \frac{k_n^{1/2+\delta}}{n^{\delta}} < \log 2, \tag{37}$$

and for $\zeta \in [1, 2)$ assume

$$\liminf_{n\to\infty} -\frac{1}{p_n} \log \frac{k_n^{\delta}}{n^{\delta}} > H(\zeta). \tag{38}$$

Then depending on the value ζ, (26), (34), (35), or (36) hold. It is easy to see that both (37) and (38) are satisfied if $\log k_n = o(\log n)$.

Example 3. Finally, let $\ell(s) = -\log s$. Assume that $k_n = (\log n)^d$ for some $d > 0$, and $p_n = \zeta^{-1} \log k_n$. Then simple calculation shows that $\beta = \beta_1 = \beta_2 = \frac{\zeta}{d}$. Furthermore $\widetilde{\ell}(x) = (\gamma/\log x)^{1/\gamma}$, and condition (31) holds.

If $\zeta \ge 2$ assume $\zeta/2 - H(2 \vee (2\zeta/d))/\zeta < \log 2$. If $\zeta \in [1, 2)$ then condition (32) always holds. Then depending on the value ζ, (26), (34), (35), or (36) hold.

4. Simulation study

The purpose of this small simulation study is to show that understanding the behavior of $\widehat{\gamma}(n)$ for large values of p is not only a mathematical challenge. The use of larger p values is beneficial in practical situations, which was already pointed out by Gomes and Martins [19]. However, we do not intend to provide neither a theoretical nor a practical comparison of the various tail index estimators. For a comprehensive simulation study, as well as for a practical criteria for the choice of k and p, we refer to [19].

Note that for $p = 1$ we obtain the usual Hill estimator. In Theorem 5.1 Segers [28] proved the optimality of the Hill estimator among residual estimators. We also see from (26) that the asymptotic variance increases with p. However, in practical situation higher p values turns out to be useful.

In the simulations below $n = 1000$ and we repeated the simulations 5000 times. In all the figures the mean and mean squared error (MSE) are plotted as a function of k in the range $[5, 200]$. For $k \ge 200$ the estimators do not change much, and we have to estimate from a negligible portion, that is $k_n/n \to 0$. In Figures 1–3 we plotted the mean and MSE for 3 quantile functions with $\gamma = 1$ and 2 in each case, and for $p = 1$ (solid), 5 (dashed), and 10 (dotted).

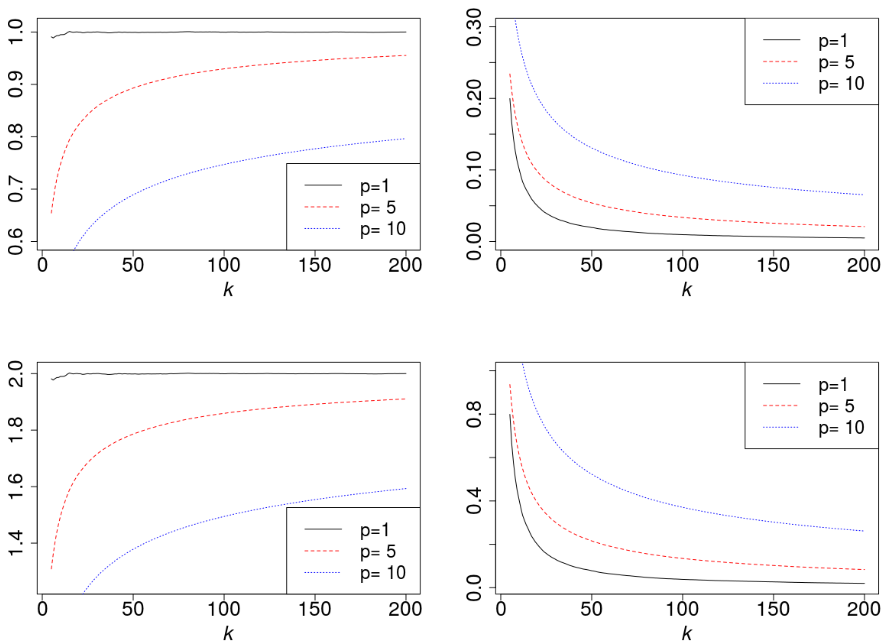

Figure 1: Mean (left) and MSE (right) for $Q(1-s) = s^{-\gamma}$ with $\gamma = 1$ (top), 2 (bottom).

In Figure 1 we see that the Hill estimator is the best in the strict Pareto model. In this case $Q(1-s) = s^{-\gamma}$. For $p = 10$ we also see that the estimator is not consistent, as $\zeta = (\log k)/10 \ll 1$. In fact we see the graph $\gamma \cdot \zeta e^{1-\zeta} = \gamma k^{-1/10} \log k\, e/10$. Note that $e^5 \approx 150$, so loosely speaking the estimator for $p = 5$ is weakly consistent only for $k \geq 150$, while $e^{10} \approx 22,000$, so asymptotic normality starts to hold for $k \geq 22,000$. Therefore, for $k \leq 200$ smaller p values should be used. We chose larger values to illustrate better the difference. We also note that for large data sets we may use larger p values.

In practice it is very unusual to encounter data which fit to a nice distribution everywhere. It is more common that the large values fit to a Pareto-type distribution, while the smaller values behave as a light-tailed distribution, see e.g. [17, p.351] or Clauset et al. [8, p.662]. Consider the quantile function

$$Q(1-s) = \begin{cases} s^{-\gamma}, & \text{if } s \leq 0.1, \\ \frac{10^{\gamma}}{\log 10} \log s^{-1}, & \text{if } s \geq 0.1, \end{cases} \tag{39}$$

which is a mixture of an exponential and a strict Pareto quantile. The parameter of the exponential is chosen such that Q is continuous. Figure 2 contain the simulation results for $\gamma = 1$ and $\gamma = 2$. In this simple model we already see the advantage of larger p values. Note that the Hill estimator is very sensitive to the change of k_n for those values where the quantile function changes. Indeed, for $k_n \leq 100$ we basically have a sample from a strict Pareto distribution, and

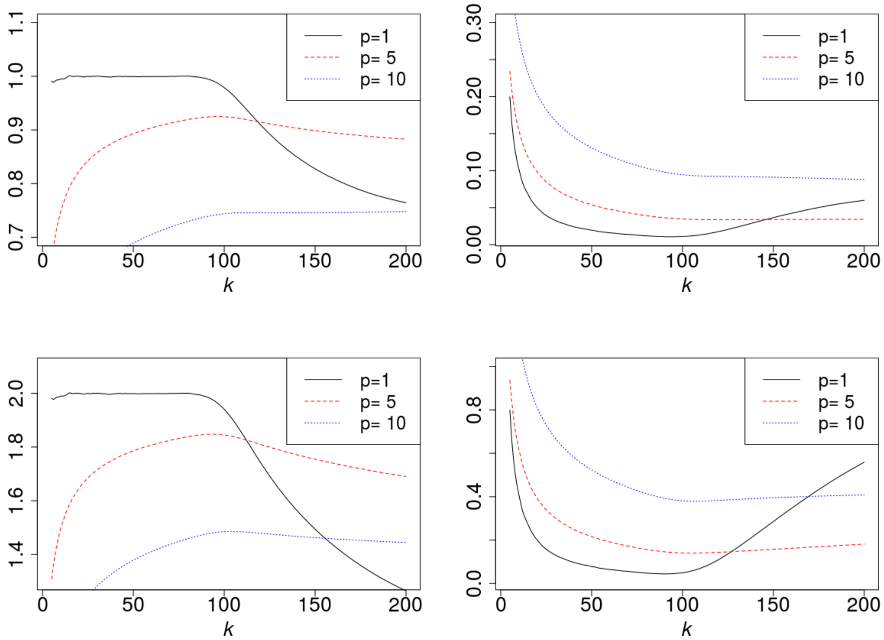

Figure 2: Mean and MSE with quantile function (39) with $\gamma = 1$ (top), 2 (bottom).

for those values the Hill estimator is the best. For $k_n = 200$ we already see the exponential part of the sample, and the Hill estimator changes drastically (for $\gamma = 1$ from 0.98 to 0.76), while for $p = 5$ the change is not as large (from 0.92 to 0.88).

Next, we further add a nonconstant slowly varying function to the quantile. A logarithmic factor in the tail of the random variable cannot be detected in practice, but it makes significantly more difficult to determine the underlying index of regular variation, see e.g. the 'Hill horror plot' on [17, p.194]. We modify the construction in (39) and consider the quantile function

$$Q(1-s) = \begin{cases} s^{-\gamma}(\log s^{-1})^3, & \text{if } s \leq 0.1, \\ 10^{\gamma}(\log 10)^2 \log s^{-1}, & \text{if } s \geq 0.1. \end{cases} \tag{40}$$

Note again that the function is continuous. We see from the simulation results in Figure 3 that in this setup the estimators with larger p values work much better than the Hill estimator. These estimators are not so sensitive for the change in the nature of the quantile function. We also see that heavier tails are in favor of larger p values.

It was pointed out in [19] that in various models under second-order regular variation for a wide range of p values (usually $p \in (1, 5]$) the estimator $\widehat{\gamma}(n)$ with p fixed is more efficient than the Hill estimator. The variance of the estimator

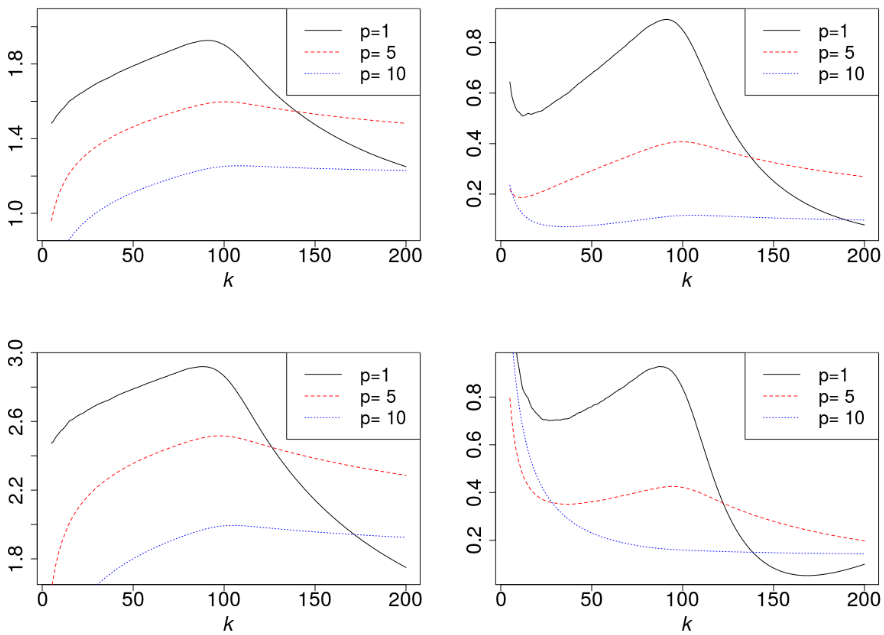

Figure 3: Mean and MSE with quantile function (40) with $\gamma = 1$ (top), 2 (bottom).

has a unique minimum at $p = 1$ (the Hill estimator), but the bias decreases in p, which is the decisive factor in some models, see Figures 3 and 4 in [19].

We also apply the estimator with different p values to real data. We chose the data set of Danish fire insurance losses, which consists of 2167 fire losses in millions of Danish Kroner. The data set is included in the R package evir, and was analyzed in [26] and in [17, Example 6.2.9]. In Figure 4 we plotted the estimate for $1/\gamma$, i.e. we plotted $1/\hat{\gamma}(n)$ against k_n, to obtain the Hill plot in [26] for $p = 1$. Resnick [26] used various techniques to obtain smoother plot. In our setting larger p values naturally produces smoother plots.

5. Proofs and auxiliary results

5.1. Strong consistency

Lemma 5.1. *Assume that* $k_n/(\log n)^\delta \to \infty$ *for some* $\delta > 0$, *and* $k_n/n \to 0$. *Then*

$$\frac{1}{k_n} \sum_{i=1}^{k_n} \left(-\log \frac{U_{i,n}}{U_{k_n+1,n}} \right)^p \longrightarrow \Gamma(p+1) \quad a.s.$$

Proof. Let F_n denote the empirical distribution function of the sample $U_1, \ldots,$

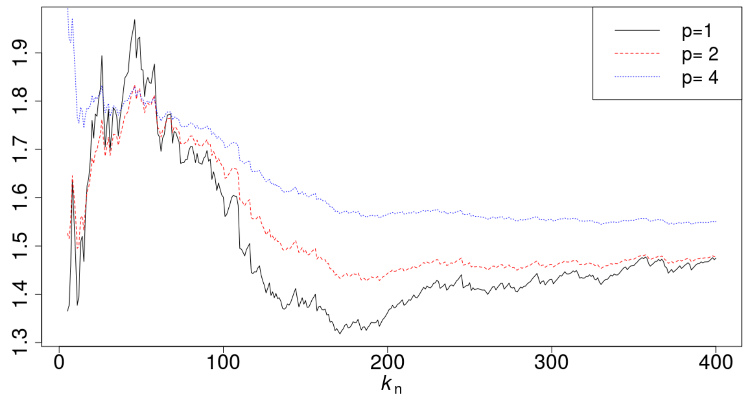

Figure 4: Hill type plots of $\hat{\gamma}(n)^{-1}$ for the Danish fire insurance claim with different p values.

U_n. Then, integrating by parts, we have

$$\frac{1}{k}\sum_{i=1}^{k}\left(-\log\frac{U_{i,n}}{U_{k+1,n}}\right)^p = \frac{n}{k}\int_{(0,U_{k,n}]}\left(-\log\frac{u}{U_{k+1,n}}\right)^p \mathrm{d}F_n(u)$$

$$= \frac{n}{k}\left[F_n(U_{k,n})\left(-\log\frac{U_{k,n}}{U_{k+1,n}}\right)^p + \int_0^{U_{k,n}} F_n(u)\frac{p}{u}\left(-\log\frac{u}{U_{k+1,n}}\right)^{p-1}\mathrm{d}u\right] \quad (41)$$

$$= \left(-\log\frac{U_{k,n}}{U_{k+1,n}}\right)^p + p\frac{n}{k}\int_0^{U_{k,n}/U_{k+1,n}} F_n(U_{k+1,n}s)(-\log s)^{p-1}\frac{1}{s}\mathrm{d}s.$$

Theorem 1 by Wellner [30] implies that

$$\frac{n}{k}U_{k,n} \to 1 \quad \text{a.s. whenever } \ k_n/\log\log n \to \infty. \quad (42)$$

Thus, the first term on the right-hand side of (41) tends to 0 a.s. For the second term

$$\frac{n}{k}\int_0^{U_{k,n}/U_{k+1,n}} F_n(U_{k+1,n}s)(-\log s)^{p-1}s^{-1}\mathrm{d}s$$

$$= \frac{n}{k}U_{k+1,n}\int_0^{U_{k,n}/U_{k+1,n}} (-\log s)^{p-1}\mathrm{d}s$$

$$+ \frac{n}{k}\int_0^{U_{k,n}/U_{k+1,n}} (F_n(U_{k+1,n}s) - U_{k+1,n}s)(-\log s)^{p-1}s^{-1}\mathrm{d}s$$

$$=: I_n + II_n.$$

Again by (42)

$$I_n \to \int_0^1 (-\log s)^{p-1}\mathrm{d}s = \Gamma(p) \quad \text{a.s.} \quad (43)$$

For the second term, choosing $\nu \in (0, 1/2)$, we have

$$
\begin{aligned}
II_n &\sim \int_0^1 \frac{F_n(U_{k+1,n}s) - U_{k+1,n}s}{U_{k+1,n}s} (-\log s)^{p-1} \mathrm{d}s \\
&= \int_0^1 \frac{F_n(U_{k+1,n}s) - U_{k+1,n}s}{(U_{k+1,n}s)^{1/2-\nu}} (-\log s)^{p-1} (U_{k+1,n}s)^{-1/2-\nu} \mathrm{d}s \\
&\leq \sup_{u \leq U_{k+1,n}} \frac{|F_n(u) - u|}{u^{1/2-\nu}} U_{k+1,n}^{-1/2-\nu} \int_0^1 (-\log s)^{p-1} s^{-1/2-\nu} \mathrm{d}s \\
&\leq C \left(\frac{\log\log n}{k} \right)^{1/2} \left[\left(\frac{n}{k} \right)^{\nu} \left(\frac{n}{\log\log n} \right)^{1/2} \sup_{u \leq 2k/n} \frac{|F_n(u) - u|}{u^{1/2-\nu}} \right],
\end{aligned}
\tag{44}
$$

where $C > 0$ is a finite constant, not depending on n, k_n. Using Theorem 1(ii) by Einmahl and Mason [16] we see that the last term in (44) is a.s. bounded, if $k_n \geq (\log n)^{(1-2\nu)/(2\nu)}$, which holds if ν is close enough to 1/2. The first term in (44) tends to 0. From (43), (44), and (41) the statement follows. □

Proof of Theorem 2.1. By the Potter bounds ([3, Theorem 1.5.6]), for any $A > 1$, $\varepsilon > 0$ there exist $x_0 = x_0(A, \varepsilon)$ such that

$$
A^{-1}(y/x)^{-\varepsilon} \leq \frac{\ell(x)}{\ell(y)} \leq A(y/x)^{\varepsilon} \quad \text{for any } 0 < x \leq y \leq x_0. \tag{45}
$$

Since $k/n \to 0$, equation (42) implies $U_{k+1,n} \to 0$ a.s. Therefore, for n large enough a.s.

$$
S_n(p) \leq \frac{1}{k} \sum_{i=1}^{k} \left(-(\gamma + \varepsilon) \log \frac{U_{i,n}}{U_{k+1,n}} + \log A \right)^p. \tag{46}
$$

First let $p \leq 1$. Using the subadditivity $(a+b)^p \leq a^p + b^p$, $a, b > 0$, by Lemma 5.1 we obtain a.s.

$$
\begin{aligned}
\limsup_{n\to\infty} S_n(p) &\leq (\gamma + \varepsilon)^p \limsup_{n\to\infty} \frac{1}{k} \sum_{i=1}^{k} \left(-\log \frac{U_{i,n}}{U_{k+1,n}} \right)^p + (\log A)^p \\
&= (\gamma + \varepsilon)^p \Gamma(p+1) + (\log A)^p.
\end{aligned}
$$

Letting $A \downarrow 1$ and $\varepsilon \downarrow 0$ we have a.s. $\limsup_{n\to\infty} S_n(p) \leq \gamma^p \Gamma(p+1)$.

Next, let $p > 1$. The convexity of the function x^p implies that for any $\varepsilon' > 0$, for $a, b > 0$

$$
\begin{aligned}
(a+b)^p &\leq (1+\varepsilon')a^p + \left(1 - (1+\varepsilon')^{-1/(p-1)} \right)^{-(p-1)} b^p \\
&=: (1+\varepsilon')a^p + C_{\varepsilon'} b^p.
\end{aligned}
$$

Therefore, using Lemma 5.1 and (46), we obtain a.s.

$$
\begin{aligned}
\limsup_{n\to\infty} S_n(p) &\leq (\gamma + \varepsilon)^p (1+\varepsilon') \limsup_{n\to\infty} \frac{1}{k} \sum_{i=1}^{k} \left(-\log \frac{U_{i,n}}{U_{k+1,n}} \right)^p + C_{\varepsilon'} (\log A)^p \\
&= (\gamma + \varepsilon)^p (1+\varepsilon') \Gamma(p+1) + C_{\varepsilon'} (\log A)^p.
\end{aligned}
$$

As $A \downarrow 1$, $\varepsilon \downarrow 0$, $\varepsilon' \downarrow 0$, we have a.s. $\limsup_{n\to\infty} S_n(p) \le \gamma^p \Gamma(p+1)$.

With the analogous lower bound, the proof is complete. □

5.2. Moment bounds

In order to replace the random centering sequence $m_p(U_{k+1,n})$ in Theorem 2.3, and the random centering $\widetilde{m}_{p_n}(U_{k+1,n})$ and norming $\widetilde{\sigma}_{p_n}(U_{k+1,n})$ and $\eta_{U_{k+1,n}}(p_n)$ in Theorems 3.1 and 3.4 with the corresponding deterministic counterparts, we need a bound for the difference $|m_p(v) - m_p|$ as $v \downarrow 0$, both for p fixed and for $p = p_v \to \infty$. These highly technical bounds are given in Propositions 5.5 and 5.6. First we need three simple auxiliary lemmas.

Lemma 5.2. *For $a \in (0, 1/2)$, $b \in (-1/2, 1/2)$, and $a + b > 0$ we have*

$$|(a+b)^p - a^p| \le \begin{cases} p|b|, & p \ge 1, \\ 2|b|a^{p-1}, & p \le 1. \end{cases}$$

Proof. Simply $(a+b)^p - a^p = bp\xi^{p-1}$, with ξ being between a and $a+b$. If $b > -a/2$ then $\xi \in [a/2, 1]$, thus

$$|(a+b)^p - a^p| \le |b|p\left((a/2)^{p-1} \vee 1\right).$$

If $b < -a/2$ then $\xi \le a$, thus $\xi^{p-1} \le a^{p-1}$ for $p \ge 1$, and

$$|(a+b)^p - a^p| \le |b|pa^{p-1}.$$

While if $b < -a/2$ and $p < 1$

$$\begin{aligned} |(a+b)^p - a^p| &= (a - |b| + |b|)^p - (a - |b|)^p \le |b|^p \\ &= |b||b|^{p-1} \le |b|(a/2)^{p-1}. \end{aligned}$$

□

Lemma 5.3. *For $x \ge p > 0$ we have*

$$\int_x^\infty e^{-y} y^p \mathrm{d}y \le x^{p+1} e^{-x} (x-p)^{-1}.$$

Proof. Simple calculation gives that

$$\begin{aligned} \int_x^\infty e^{-y} y^p \mathrm{d}y &= x^{p+1} e^{-x} \int_1^\infty e^{-x(u-1) + p \log u} \mathrm{d}u \\ &= x^{p+1} e^{-x} \int_1^\infty e^{-(x-p)(u-1) - p(u-1-\log u)} \mathrm{d}u \\ &\le x^{p+1} e^{-x} \int_1^\infty e^{-(x-p)(u-1)} \mathrm{d}u \\ &= x^{p+1} e^{-x} (x-p)^{-1}. \end{aligned}$$

□

In the borderline cases $\alpha = 1$ and $\alpha = 2$ the centering and norming sequences are different.

Lemma 5.4. *For the truncated moments as $p \to \infty$ we have for $\zeta = 1$ that*

$$m_p^1 \sim \left(\frac{\gamma p}{e}\right)^p \frac{\sqrt{p\pi}}{\sqrt{2}},$$

and for $\zeta = 2$ that

$$\sigma_p^1 \sim \left(\frac{2\gamma p}{e}\right)^p (p\pi)^{1/4}.$$

Proof. Since $\eta_0(p) = \zeta\gamma p$, for $\zeta = 1$

$$m_p^1 = \gamma^p \int_0^p y^p e^{-y} \mathrm{d}y.$$

We have

$$\int_0^p y^p e^{-y} \mathrm{d}y = p^{p+1} e^{-p} \int_0^1 e^{-p(x-1-\log x)} \mathrm{d}x.$$

The exponent is negative and $x - 1 - \log x \sim (x-1)^2/2$ as $x \uparrow 1$. Thus

$$\int_0^1 e^{-p(x-1-\log x)} \mathrm{d}x \sim \sqrt{\pi/(2p)},$$

and the result follows.

For $\zeta = 2$ using the definition and the previous result

$$(\sigma_p^1)^2 = \gamma^{2p} \int_0^{2p} y^{2p} e^{-y} \mathrm{d}y \sim \left(\frac{2\gamma p}{e}\right)^{2p} \sqrt{p\pi},$$

as claimed. □

In what follows p_v is a positive function of v.

Proposition 5.5. *Assume* (27) *and that*

$$\lim_{v \downarrow 0} p_v \frac{a(v)}{\ell(v)} = 0. \tag{47}$$

Then there exists $v_0 > 0$ such that for all $v \in (0, v_0)$

$$|m_{p_v}(v) - m_{p_v}| \leq 2K_1 \frac{a(v)}{\ell(v)} \gamma^{p_v - 1} \Gamma(p_v + 1).$$

Proof. To ease notation put

$$\eta(u, v) = \left(-\gamma \log u + \log \frac{\ell(uv)}{\ell(v)}\right)^p - (-\gamma \log u)^p. \tag{48}$$

We have by (1)

$$
\begin{aligned}
m_p(v) - m_p &= \mathbb{E}\left[\left(\log \frac{Q(1-Uv)}{Q(1-v)}\right)^p - (-\gamma \log U)^p\right] \\
&= \mathbb{E}\left[\left(-\gamma \log U + \log \frac{\ell(Uv)}{\ell(v)}\right)^p - (-\gamma \log U)^p\right] \\
&= \int_0^1 \eta(u,v) \mathrm{d}u =: I_1(\delta) + I_2(\delta),
\end{aligned}
$$

where I_1, I_2 are the integrals on $(0, 1-\delta)$, $(1-\delta, 1)$, with $\delta \in (0, 1/2)$.

First we deal with the integral on $(0, 1-\delta)$. By (45), for any $\varepsilon > 0$, $A > 1$, there is $v_0 > 0$ such that for $v \le v_0$, $u \in (0,1)$

$$
A^{-1} u^{\varepsilon} \le \frac{\ell(uv)}{\ell(v)} \le A u^{-\varepsilon}, \tag{49}
$$

implying that uniformly on $u \in (0, 1-\delta]$

$$
\frac{\log \frac{\ell(uv)}{\ell(v)}}{-\log u} \to 0 \quad \text{as } v \downarrow 0. \tag{50}
$$

Writing

$$
\frac{\ell(uv) - \ell(v)}{\ell(v)} = \frac{a(v)}{\ell(v)} \frac{\ell(uv) - \ell(v)}{a(v)},
$$

by (27) we see that the first factor tends to 0 and the second factor is bounded. Therefore, uniformly in $u \in [0,1]$

$$
\log \frac{\ell(uv)}{\ell(v)} \sim \frac{a(v)}{\ell(v)} \frac{\ell(uv) - \ell(v)}{a(v)} \quad \text{as } v \downarrow 0. \tag{51}
$$

By (50) and (51), if (47) holds then, uniformly on $u \in [0, 1-\delta]$,

$$
\left(1 + \frac{\log \frac{\ell(uv)}{\ell(v)}}{-\gamma \log u}\right)^p - 1 \sim p\,(-\gamma \log u)^{-1} \, \frac{a(v)}{\ell(v)} \frac{\ell(uv) - \ell(v)}{a(v)}. \tag{52}
$$

Thus,

$$
I_1(\delta) \le p \frac{a(v)}{\ell(v)} \, \frac{3}{2} K_1 \gamma^{p-1} \int_0^{1-\delta} (-\log u)^{p-1} \, \mathrm{d}u. \tag{53}
$$

Next, we turn to I_2. Note that (51) holds, but (50) does not, because $\log u$ can be small. Choosing $\delta > 0$ small enough we can achieve that $-\gamma \log(1-\delta) \in (0, 1/2)$ and by (51) also that $\log \ell(uv)/\ell(v) \in (-1/2, 1/2)$ for v small and

$u \in [1-\delta, 1]$. Therefore, we can apply Lemma 5.2 with $a = -\gamma \log u$ and $b = \log(\ell(uv)/\ell(v))$ together with (51) and (27), and we obtain for $p \leq 1$

$$\begin{aligned} |\eta(u,v)| &\leq 2 \left| \log \frac{\ell(uv)}{\ell(v)} \right| (-\gamma \log u)^{p-1} \\ &\leq \frac{a(v)}{\ell(v)} 2K_1 (-\gamma \log u)^{p-1}. \end{aligned}$$

While, for $p \geq 1$

$$|\eta(u,v)| \leq p \left| \log \frac{\ell(uv)}{\ell(v)} \right| \leq p \frac{a(v)}{\ell(v)} K_1.$$

Summarizing,

$$I_2(\delta) \leq \begin{cases} \frac{a(v)}{\ell(v)} 2K_1 \gamma^{p-1} \int_{1-\delta}^1 (-\log u)^{p-1} \mathrm{d}u, & p \leq 1, \\ p \frac{a(v)}{\ell(v)} K_1 \delta, & p \geq 1. \end{cases} \tag{54}$$

The bounds (53) and (54) imply the statement. □

Proposition 5.6. *Assume* (10) *and let*

$$\beta_2 := \limsup_{v \downarrow 0} \frac{-\log \frac{a(v)}{\ell(v)}}{p_v} \geq \liminf_{v \downarrow 0} \frac{-\log \frac{a(v)}{\ell(v)}}{p_v} := \beta_1, \tag{55}$$

allowing $\beta_1 = \infty$. Assume either $\beta_1 \geq 1$ or $\beta_2 \leq 1$, and define β as in (22). *Then for any $\varepsilon > 0$ there exists a $K > 0$ such that for v small enough*

$$|m_{p_v}(v) - m_{p_v}| \leq K \left(\frac{a(v)}{\ell(v)} \right)^{(\nu_\beta - \varepsilon) \wedge 1} (\gamma + \varepsilon)^{p_v} \Gamma(p_v + 1).$$

Note that if $p > 0$ is fixed then $\beta = \infty$ and we obtain the same bound as in Proposition 5.5.

Proof. The difference compared to the previous proof is that (27) does not hold uniformly in $[0,1]$, which implies that the integral of $\eta(u,v)$ in (48) on the interval $[0,\delta]$ has to be treated differently.

By Theorem 3.1.4 in [3] (translating the results from infinity to zero, by defining $\overline{\ell}(x) = \ell(x^{-1})$, $\overline{a}(x) = a(x^{-1})$)

$$\limsup_{v \downarrow 0} \sup_{u \in [\delta, 1]} \frac{|\ell(uv) - \ell(v)|}{a(v)} =: K_1(\delta) < \infty.$$

This implies that the bound (54) on $[1-\delta, 1]$ remains true and on $[\delta, 1-\delta]$ as in (53) we have

$$\int_\delta^{1-\delta} \eta(u,v) \mathrm{d}u \leq p \frac{a(v)}{\ell(v)} \frac{3}{2} K_1 \gamma^{p-1} \int_\delta^{1-\delta} (-\log u)^{p-1} \mathrm{d}u. \tag{56}$$

Recall (48) and let

$$J_1 = \int_0^{b(v)} \eta(u,v)\mathrm{d}u, \quad J_2 = \int_{b(v)}^{\delta} \eta(u,v)\mathrm{d}u, \tag{57}$$

where

$$b(v) = \left(\frac{a(v)}{\ell(v)}\right)^2 \wedge e^{-2p}. \tag{58}$$

By Theorem 3.1.4 in [3] for any $\varepsilon > 0$ there is $v_0(\varepsilon) > 0$ and $K_2(\varepsilon) > 0$ such that

$$\frac{|\ell(uv) - \ell(v)|}{a(v)} \le K_2(\varepsilon) u^{-\varepsilon} \quad \text{for all } u \le 1, v \le v_0(\varepsilon). \tag{59}$$

By (58) and (55) for $\varepsilon_1 > 0$ small enough

$$p\frac{a(v)}{\ell(v)} b(v)^{-\varepsilon_1} \to 0. \tag{60}$$

Using (59), for $u \ge b(v)$

$$\frac{|\ell(uv) - \ell(v)|}{\ell(v)} \le K_2(\varepsilon_1)\frac{a(v)}{\ell(v)} u^{-\varepsilon_1} \le K_2(\varepsilon_1)\frac{a(v)}{\ell(v)} b(v)^{-\varepsilon_1} \to 0,$$

therefore

$$\left|\log\frac{\ell(uv)}{\ell(v)}\right| \sim \frac{|\ell(uv) - \ell(v)|}{\ell(v)} \le K_2(\varepsilon_1)\frac{a(v)}{\ell(v)} u^{-\varepsilon_1}.$$

By (60) for $u \in [b(v), \delta]$ the asymptotic equality in (52) holds, thus for J_2 in (57)

$$\begin{aligned} J_2 &\sim \int_{b(v)}^{\delta} (-\gamma \log u)^p p (-\gamma \log u)^{-1} \frac{a(v)}{\ell(v)} \frac{\ell(uv) - \ell(v)}{a(v)} \mathrm{d}u \\ &\le p\frac{a(v)}{\ell(v)} K_2(\varepsilon_1) \int_{b(v)}^{\delta} (-\gamma \log u)^{p-1} u^{-\varepsilon_1} \, \mathrm{d}u \\ &\le p\frac{a(v)}{\ell(v)} K_2(\varepsilon_1)(1-\varepsilon_1)^{-p} \gamma^{p-1} \Gamma(p), \end{aligned} \tag{61}$$

where at the last inequality we used that

$$\begin{aligned} \int_0^1 (-\log u)^{p-1} u^{-\varepsilon_1} \, \mathrm{d}u &= \int_0^\infty y^{p-1} e^{-(1-\varepsilon_1)y} \mathrm{d}y \\ &= (1-\varepsilon_1)^{-p} \, \Gamma(p). \end{aligned}$$

On $(0, b(v))$ using (49), $b(v) \to 0$, Lemma 5.3, and that $-\log b(v) - p \ge$

$(-\log b(v))/2$ we obtain for v small enough

$$\begin{aligned}
J_1 &\leq 2 \int_0^{b(v)} (-(\gamma+\varepsilon)\log u + \log A)^p \,\mathrm{d}u \\
&\leq 2(\gamma+2\varepsilon)^p \int_0^{b(v)} (-\log u)^p \,\mathrm{d}u \\
&= 2(\gamma+2\varepsilon)^p \int_{-\log b(v)}^{\infty} y^p e^{-y} \mathrm{d}y \\
&\leq 2(\gamma+2\varepsilon)^p (-\log b(v))^{p+1} e^{\log b(v)} (-\log b(v) - p)^{-1} \\
&\leq 4(\gamma+2\varepsilon)^p \,(-\log b(v))^p \, b(v).
\end{aligned} \tag{62}$$

Note that for $\log x > p$

$$\begin{aligned}
\frac{(\log x)^p}{x} \frac{e^p}{p^p} &= \exp\left\{ -p \left(\frac{\log x}{p} - 1 - \log \frac{\log x}{p} \right) \right\} \\
&= \exp\left\{ -pH\left(\frac{\log x}{p} \right) \right\},
\end{aligned}$$

with H defined in (23). Thus with $x = b(v)^{-1}$

$$\begin{aligned}
\left(\frac{e}{p} \right)^p (-\log b(v))^p \, b(v) &= \exp\left\{ -pH \left(2 \vee \frac{-2\log(a(v)/\ell(v))}{p} \right) \right\} \\
&= \left(\frac{a(v)}{\ell(v)} \right)^{\frac{p}{-\log(a(v)/\ell(v))} H\left(2 \vee \frac{-2\log(a(v)/\ell(v))}{p}\right)} .
\end{aligned}$$

The function $\nu_x = x^{-1} \, H(2 \vee 2x)$ is strictly decreasing on $(0, 1]$, and strictly increasing on $[1, \infty)$ attaining its unique minimum at 1. Continuing (62) for any $\varepsilon_2 > 0$ for v small enough

$$J_1 \leq \frac{4}{\sqrt{p\pi}} (\gamma + 2\varepsilon)^p \, \Gamma(p+1) \left(\frac{a(v)}{\ell(v)} \right)^{\nu_\beta - \varepsilon_2}.$$

Combining with (61), (56), and (54) the result follows. □

As an easy consequence of the moment bounds we show that the random centering and norming can be substituted with the deterministic one.

Proof of Theorem 2.4. The theorem is an immediate consequence of Theorem 2.3 and Proposition 5.5. Indeed, by Proposition 5.5 for some $c > 0$

$$\sqrt{k}\,|m_p(U_{k+1,n}) - m_p| \leq c\sqrt{k} \frac{a(U_{k+1,n})}{\ell(U_{k+1,n})} = c\sqrt{k} \frac{a(k/n)}{\ell(k/n)} \frac{a(U_{k+1,n})}{a(k/n)} \frac{\ell(k/n)}{\ell(U_{k+1,n})}.$$

By the assumption $\sqrt{k}a(k/n)/\ell(k/n) \to 0$, while the last two factors tends to 1, since a and ℓ are regularly varying and $U_{k+1,n} \sim k/n$.

The central limit theorem for $\widehat{\gamma}(n)$ follows from the previous result using the delta method, see Agresti [1, Section 14.1]. □

6. Limit results for power sums

In this section we assume that $p = p_n$ tends to infinity at a certain rate. We prove the analogues of Bogachev's result [4, Section 2] for the random variables $Y(v)$ uniformly in v. As the log-tail distribution function h_v in (13) is regularly varying, for each $v \in [0, 1)$ fixed all the following results are consequences of Bogachev's results. However, the main difficulty in our setup is the additional parameter v, in which we need some kind of uniformity. We apply these results to prove limit theorems for the $S_n(p_n)$ and $\widehat{\gamma}(n)$.

Recall (7). Let $Y(v), Y_1(v), Y_2(v), \dots$ be iid random variables, and put

$$Z_n(p, v) = \sum_{i=1}^{n} Y_i(v)^p.$$

First we determine the asymptotic behavior of the moments as $p \to \infty$.

Lemma 6.1. *For any $\varepsilon > 0$ there is a $p_0 > 0$ such that for $v \in [0, 1)$, $p > p_0$*

$$(\gamma - \varepsilon)^p \, \Gamma(p+1) \le m_p(v) \le (\gamma + \varepsilon)^p \, \Gamma(p+1). \tag{63}$$

In particular, as $p \to \infty$ uniformly in v

$$\frac{\log m_p(v)}{p} - \log p \to \log \gamma \, - 1.$$

Proof. First note that if X is a nonnegative random variable for which $\mathbb{P}(X > x) > 0$ for any x then for any $K > 0$

$$\mathbb{E}X^p \sim \mathbb{E}X^p I(X > K) \quad \text{as } p \to \infty.$$

This implies that for any $\varepsilon > 0$ and $a > 0$ there exist $p_0 = p_0(\varepsilon, a)$ such that for $p > p_0$

$$(1 - \varepsilon)^p \, \mathbb{E}(X + a)^p \le \mathbb{E}X^p \le (1 + \varepsilon)^p \, \mathbb{E}|X - a|^p. \tag{64}$$

Using the Potter bounds (see (49)) and (64), for any $\varepsilon > 0$ there exists $A > 1$ and $p_0 > 0$ such that for $v \in [0, 1]$, $p > p_0$

$$\begin{aligned} m_p(v) &= \mathbb{E}\left(\log\left(U^{-\gamma}\frac{\ell(Uv)}{\ell(v)}\right)\right)^p \\ &\le \mathbb{E}\left(\log\left(U^{-(\gamma+\varepsilon)}A\right)\right)^p \\ &\le (\gamma + \varepsilon)^p \mathbb{E}\left(\log U^{-1} + \frac{\log A}{\gamma + \varepsilon}\right)^p \\ &\le ((1 + \varepsilon)(\gamma + \varepsilon))^p \Gamma(p + 1). \end{aligned}$$

Together with an analogous lower bound, (63) follows. The second part simply follows from Stirling's formula. □

If $V_n(v)$ is a sequence of random variables indexed by $v \in [0,1)$, then $V_n(v)$ *converges in distribution uniformly in* v *to a random variable* W, if for each continuity point x of the distribution function of W

$$\lim_{n\to\infty} \sup_{v\in[0,1)} |\mathbb{P}(V_n(v) \le x) - \mathbb{P}(W \le x)| = 0.$$

Similarly, $V_n(v)$ *converges in probability uniformly in* v *to a random variable* W, if for each $\varepsilon > 0$

$$\lim_{n\to\infty} \sup_{v\in[0,1)} \mathbb{P}(|V_n(v) - W| > \varepsilon) = 0.$$

Similarly as in (15) and (16), we introduce the growth parameter ζ of the sequence p_n as follows. For the sequence $p = p_n$ let

$$\liminf_{n\to\infty} \frac{\log n}{p_n} = \zeta \ge 0; \tag{65}$$

for $\zeta \le 2$ we need the stronger assumption

$$n \sim e^{\zeta p_n}. \tag{66}$$

To obtain a weak law of large numbers we need that $\zeta > 1$.

Proposition 6.2. *If* $\zeta > 1$ *in* (65) *or* $\zeta = 1$ *in* (66) *then uniformly for* $v \in [0,1)$ *as* $p_n \to \infty$

$$\frac{Z_n(p_n, v) - n\widetilde{m}_{p_n}(v)}{n\widetilde{m}_{p_n}(v)} \xrightarrow{\mathbb{P}} 0.$$

Proof. Let $\zeta > 1$. We follow the proof of Theorem 2.1 in [4]. Fix $\varepsilon > 0$, and let $r \in [1,2]$. Using the Markov inequality, the Marcinkiewicz–Zygmund inequality (see e.g. [25, 2.6.18]), and the subadditivity we have with some $c_r > 0$

$$\begin{aligned}
&\mathbb{P}\left(\frac{|Z_n(p,v) - nm_p(v)|}{nm_p(v)} > \varepsilon\right) \\
&\le (\varepsilon n m_p(v))^{-r} \mathbb{E}|Z_n(p,v) - nm_p(v)|^r \\
&\le c_r (\varepsilon n m_p(v))^{-r} \mathbb{E}\left(\sum_{i=1}^n (Y_i(v)^p - m_p(v))^2\right)^{r/2} \\
&\le c_r (\varepsilon n m_p(v))^{-r} n \mathbb{E}|Y(v)^p - m_p(v)|^r \\
&\le c_r \varepsilon^{-r} n^{1-r} \frac{m_{rp}(v)}{m_p(v)^r}.
\end{aligned} \tag{67}$$

By Lemma 6.1 for any $\varepsilon_1 > 0$ we can choose $p_0 > 0$ such that for $v \in [0,1)$ and $p > p_0$

$$\frac{m_{rp}(v)}{m_p(v)^r} \le \frac{(\gamma + \varepsilon_1)^{rp} \Gamma(rp+1)}{(\gamma - \varepsilon_1)^{rp} \Gamma(p+1)^r} \le (1+\varepsilon_2)^{rp} \frac{\Gamma(rp+1)}{\Gamma(p+1)^r},$$

with $\varepsilon_2 = 2\varepsilon_2/(\gamma - \varepsilon_2)$. Thus, by the Stirling formula

$$\begin{aligned}\limsup_{p\to\infty} \frac{1}{p} \log \frac{m_{rp}(v)}{n^{r-1} m_p(v)^r} &\le r\log(1+\varepsilon_2) + r\log r - (r-1)\liminf_{p\to\infty} \frac{\log n}{p} \\ &\le r\log(1+\varepsilon_2) + r\log r - (r-1)\zeta.\end{aligned} \tag{68}$$

As $\zeta > 1$ we can choose $r \in [1, 2]$ such that $r \log r - (r-1)\zeta < 0$. Then choosing ε_1 small enough we see that the right-hand side in (68) is negative, implying that the right-hand side in (67) tends to 0.

For $\zeta = 1$ the result is a consequence of Proposition 6.6. We only need that $n\widetilde{m}_p(v)/\eta_v(p)^p \to \infty$, which follows from (74) in Lemma 6.5 with $r = \zeta = 1$. □

For the central limit theorem we need further restriction on p_n.

Proposition 6.3. *If $\zeta > 2$ in (65) or $\zeta = 2$ in (66) then uniformly on $[0, 1)$*

$$\frac{Z_n(p_n, v) - n\widetilde{m}_{p_n}(v)}{\sqrt{n}\widetilde{\sigma}_{p_n}(v)} \xrightarrow{\mathcal{D}} N(0, 1).$$

Proof. Let $\zeta > 2$. Applying the Berry–Esseen bound (see [6, Corollary 4, p.300]), there exists a universal constant C_δ such that

$$\begin{aligned}&\sup_{x\in\mathbb{R}} \left| \mathbb{P}\left(\frac{Z_n(p_n, v) - n m_{p_n}(v)}{\sqrt{n}\sigma_{p_n}(v)} \le x \right) - \Phi(x) \right| \\ &\le C_\delta \frac{n}{(\sqrt{n}\sigma_p(v))^{2+\delta}} \mathbb{E}|Y(v)^p - m_p(v)|^{2+\delta},\end{aligned}$$

where Φ is the standard normal distribution function. Therefore, it is enough to show that for some $\delta > 0$ uniformly in v

$$\frac{n}{(\sqrt{n}\sigma_p(v))^{2+\delta}} \mathbb{E}|Y(v)^p - m_p(v)|^{2+\delta} \to 0$$

as $n \to \infty$. By Lemma 6.1 $\sigma_p(v) \sim \sqrt{m_{2p}(v)}$ as $p \to \infty$. Thus we have to show that

$$\frac{m_{p(2+\delta)}(v)}{n^{\delta/2} m_{2p}(v)^{1+\delta/2}} \to 0.$$

As in the proof of Proposition 6.2

$$\limsup_{p\to\infty} \frac{1}{p} \log \frac{m_{p(2+\delta)}(v)}{n^{\delta/2} m_{2p}(v)^{1+\delta/2}} \le -\frac{\delta}{2}\zeta + \log(1+\varepsilon) + (2+\delta)\log(1+\delta/2).$$

We have to choose $\delta > 0$ such that

$$\frac{2}{\delta}(2+\delta)\log\left(1 + \frac{\delta}{2}\right) < \zeta.$$

This is possible for $\zeta > 2$.

For $\zeta = 2$ we defer the proof after Proposition 6.6. □

In the range $\zeta \in (0,2)$ we need (66), the finer assumption on the sequence p_n. For the error term in h_v in (13)

$$\begin{aligned} h_v(x) - h_0(x) &= \log\left(\ell(v)^{1/\gamma}\widetilde{\ell}(v^{-\gamma}\ell(v)e^x)\right) \\ &= \log \ell(v)^{1/\gamma}\widetilde{\ell}(v^{-\gamma}\ell(v)) + \log\frac{\widetilde{\ell}(v^{-\gamma}\ell(v)e^x)}{\widetilde{\ell}(v^{-\gamma}\ell(v))}. \end{aligned} \tag{69}$$

By the inverse relation (12) the first term is small for v small, while the second term can be bounded using the Potter bounds, thus for any $\varepsilon > 0$ there exist $x_0 > 0$ such that for $x > x_0$

$$\left|\log\left(\ell(v)^{1/\gamma}\widetilde{\ell}(v^{-\gamma}\ell(v)e^x)\right)\right| \le \varepsilon x,$$

implying that for $x > x_0$

$$|h_v(x) - x/\gamma| \le \varepsilon x.$$

Also, for η_v in (14) there exist $x_0 > 0$ such that for $x > x_0$

$$|\eta_v(x) - \gamma\zeta x| \le \varepsilon x. \tag{70}$$

Using these bounds, we can prove the uniform version of Lemma 5.4 in [4].

Lemma 6.4. *For any $K > 0$*

$$\lim_{p\to\infty}\sup_{v\in[0,1],x\in[K^{-1},K]}\left|h_v(\eta_v(p)) - h_v(\eta_v(p)x^{1/p}) + \zeta\log x\right| = 0.$$

Proof. We have by (13)

$$h_v(\eta_v(p)) - h_v(\eta_v(p)x^{1/p}) = \frac{\eta_v(p)}{\gamma}(1 - x^{1/p}) + \log\frac{\widetilde{\ell}(v^{-\gamma}\ell(v)e^{\eta_v(p)})}{\widetilde{\ell}(v^{-\gamma}\ell(v)e^{\eta_v(p)x^{1/p}})}.$$

Using (70) and that $1 - x^{1/p} \sim -p^{-1}\log x + O(p^{-2})$, we see that the first term tends to $-\zeta\log x$. This further implies, using also the uniform convergence theorem that the second term above tends to 0, proving the statement.

We also note that the argument above shows that the uniform convergence theorem for the regularly varying h_v holds uniformly in $v \in [0,1]$. □

Once we have the uniform convergence in v, and the uniform moment bound (63), the proofs of Lemma 6.1, 6.2, and 6.3 in [4] go through. We omit the proof.

Lemma 6.5. *For any $r > 0$ and $\tau > 0$ uniformly in $v \in [0,1)$*

$$\begin{aligned} &\lim_{p\to\infty}\frac{e^{\zeta p}}{\eta_v(p)^{pr}}\mathbb{E}\left[Y(v)^{rp}\left(I(Y(v) \le \eta_v(p)\tau^{1/p}) - I(Y(v) \le \eta_v(p))\right)\right] \\ &= \begin{cases} \frac{\zeta}{r-\zeta}(\tau^{r-\zeta} - 1), & r \neq \zeta, \\ \zeta\log\tau, & r = \zeta. \end{cases} \end{aligned} \tag{71}$$

For any $\tau > 0$ *and* $r > \zeta$

$$\lim_{p\to\infty} \frac{e^{\zeta p}}{\eta_v(p)^{pr}} \mathbb{E}\left[Y(v)^{rp} I(Y(v) \le \eta_v(p)\tau^{1/p}) \right] = \frac{\zeta}{r-\zeta}\tau^{r-\zeta}, \tag{72}$$

while for $\tau > 0$ *and* $r < \zeta$

$$\lim_{p\to\infty} \frac{e^{\zeta p}}{\eta_v(p)^{pr}} \mathbb{E}\left[Y(v)^{rp} I(Y(v) > \eta_v(p)\tau^{1/p}) \right] = \frac{\zeta}{\zeta-r}\tau^{r-\zeta}. \tag{73}$$

For $r = \zeta$

$$\lim_{p\to\infty} \frac{e^{\zeta p}}{\eta_v(p)^{\zeta p}} \mathbb{E}\left[Y(v)^{\zeta p} I(Y(v) \le \eta_v(p)) \right] = \infty. \tag{74}$$

Recall the notation (17). Again, if $\zeta \le 2$ then ζ equals the characteristic exponent of the limiting stable law. Therefore, we use the notation $\zeta = \alpha$.

Proposition 6.6. *Assume that* (66) *holds with* $\zeta = \alpha \in (0,2)$*. Then as* $n \to \infty$*, uniformly in* $v \in [0,1)$

$$\frac{1}{\eta_v(p_n)^{p_n}} \left[Z_n(p_n, v) - n\widetilde{m}_{p_n}(v) \right] \xrightarrow{\mathcal{D}} Z_\alpha.$$

Proof. We use the classical criteria for convergence of sums of independent random variables, see Theorem 25.1 in Gnedenko and Kolmogorov [18].

First, by (13), (14), and Lemma 6.4, uniformly in $v \in [0,1)$

$$n\mathbb{P}(Y(v)^p > \eta_v(p)^p x) = n e^{-h_v(\eta_v(p))} e^{h_v(\eta_v(p)) - h_v(\eta_v(p)x^{1/p})} \to x^{-\alpha}. \tag{75}$$

Next, applying Lemma 6.5 with $r = 2$, uniformly in $v \in [0,1)$

$$\lim_{\tau\downarrow 0} \limsup_{n\to\infty} \frac{n}{\eta_v(p)^{2p}} \mathbb{E}\left[Y(v)^{2p} I(Y(v) \le \eta_v(p)\tau^{1/p}) \right] = 0.$$

Therefore, we already have that the normed sum converges with an appropriate centering, and the limit is a one-sided α-stable law. To see that the centering is correct note that

$$\begin{aligned} &\lim_{p\to\infty} \left(\frac{n}{\eta_v(p)^p} \mathbb{E}\left[Y(v)^p I\left(Y(v) \le \eta_v(p)\tau^{1/p} \right) \right] - \frac{n\widetilde{m}_p(v)}{\eta_v(p)^p} \right) \\ &= \begin{cases} \frac{\alpha}{1-\alpha}\tau^{1-\alpha}, & \alpha \ne 1, \\ \log \tau, & \alpha = 1. \end{cases} \end{aligned}$$

Indeed, this follows from (72) for $\alpha < 1$, from (73) for $\alpha > 1$, and from (71) for $\alpha = 1$. □

We end this section with the proof of the central limit theorem in the borderline case $\alpha = 2$.

Proof of Proposition 6.3 for $\alpha = 2$. Here we use again the classical criteria [18, Theorem 25.1], but specified to the Gaussian law.

Using (74) with $\alpha = r = 2$ we obtain that uniformly in $v \in [0, 1)$

$$\sigma_p^1(v) e^p / \eta_v(p)^p \to \infty. \tag{76}$$

Thus, for any $x > 0$ fixed and $\tau > 0$ large, for n large enough

$$n\mathbb{P}(Y(v)^p > \sqrt{n}\sigma_p^1(v)x) \le n\mathbb{P}(Y(v)^p > \eta_v(p)^p \tau),$$

which by (75) converges to τ^{-2}. Thus, for any $x > 0$

$$n\mathbb{P}(Y(v)^p > \sqrt{n}\sigma_p^1(v)x) \to 0. \tag{77}$$

For the truncated variance

$$\begin{aligned} &\frac{n}{(\sigma_p^1(v)\sqrt{n})^2}\mathbb{E}\left[Y(v)^{2p} I\left(Y(v) \le (\sigma_p^1(v)\sqrt{n}\tau)^{1/p}\right)\right] \\ &= 1 + \frac{\mathbb{E}\left[Y(v)^{2p} I\left(\eta_v(p) \le Y(v) \le (\sigma_p^1(v)\sqrt{n}\tau)^{1/p}\right)\right]}{(\sigma_p^1(v))^2}. \end{aligned} \tag{78}$$

For the second term for $\delta \in (0, 1)$ by (73) with $\alpha = 2$, $r = 2 - \delta$

$$\begin{aligned} &\frac{\mathbb{E}\left[Y(v)^{2p} I\left(\eta_v(p) \le Y(v) \le (\sigma_p^1(v)\sqrt{n}\tau)^{1/p}\right)\right]}{(\sigma_p^1(v))^2} \\ &\le \frac{(\sigma_p^1(v)\sqrt{n}\tau)^\delta}{(\sigma_p^1(v))^2}\mathbb{E}\left[Y(v)^{(2-\delta)p} I(\eta_v(p) \le Y(v))\right] \\ &\sim \frac{2}{\delta}\left(\frac{\sigma_p^1(v)e^p}{\eta_v(p)^p}\right)^{-(2-\delta)} \frac{n^{\delta/2}}{e^{p\delta}}, \end{aligned}$$

which tends to 0 by (76). Furthermore, by (76)

$$\frac{\left[\mathbb{E}\left(Y(v)^p I(Y(v) \le (\sigma_p^1(v)\sqrt{n}\tau)^{1/p})\right)\right]^2}{\mathbb{E}\left(Y(v)^{2p} I(Y(v) \le (\sigma_p^1(v)\sqrt{n}\tau)^{1/p})\right)} \le \frac{(m_p(v))^2}{(\sigma_p^1(v))^2}.$$

Using Lemma 6.1 and (70) it is simple to show that the latter quantity tends to 0 uniformly in $v \in [0, 1)$. Thus, from (78) we obtain for any $\tau > 0$

$$\begin{aligned} \lim_{n\to\infty} \frac{n}{\left(\sqrt{n}\sigma_p^1(v)\right)^2} &\left\{\mathbb{E}\left[Y(v)^{2p} I\left(Y(v) \le (\sigma_p^1(v)\sqrt{n}\tau)^{1/p}\right)\right]\right. \\ &\left. - \left(\mathbb{E}\left[Y(v)^p I\left(Y(v) \le (\sigma_p^1(v)\sqrt{n}\tau)^{1/p}\right)\right]\right)^2\right\} = 1. \end{aligned} \tag{79}$$

Finally, by (73) with $\alpha = 2$, $r = 1$

$$\begin{aligned} &\frac{n}{\sigma_p^1(v)\sqrt{n}}\mathbb{E}\left(Y(v)^p I(Y(v) > (\sigma_p^1(v)\sqrt{n}\tau)^{1/p})\right) \\ &\le \frac{\sqrt{n}}{\sigma_p^1(v)}\mathbb{E}\left(Y(v)^p I(Y(v) > \eta_p(v)\tau^{1/p})\right) \sim \frac{\sqrt{n}}{\sigma_p^1(v)}\frac{\eta_v(p)^p}{e^{2p}}\tau^{-1} \end{aligned}$$

which tends to 0 by (76). Together with (77) and (79) this implies the statement. □

7. Proofs for large p

Proof of Theorem 3.1. The limiting relations (18) and (19) are immediate consequences of Propositions 6.2 and 6.3. Indeed, for the law of large numbers by representation (5)

$$\begin{aligned}&\mathbb{P}\left(|S_n(p_n)/m_{p_n}(U_{k+1,n}) - 1| > \varepsilon\right)\\&= \int_0^1 \mathbb{P}\left(|Z_n(p_n, v)/(n m_{p_n}(v)) - 1| > \varepsilon\right) \mathrm{d}\mathbb{P}(U_{k+1,n} \le v),\end{aligned}$$

which tends to 0, since the integrand tends to 0 uniformly. The proof of (19) is similar.

The weak consistency follows from Lemma 6.1 and (18).

The CLT (20) follows from Bogachev's transfer lemma (Lemma 9.1 in [4]) and Theorem 3.1. To apply Lemma 9.1 in [4] we only need to show that

$$\frac{\sqrt{k_n} m_{p_n}(U_{k+1,n})}{\sigma_{p_n}(U_{k+1,n})} \to \infty.$$

This follows easily from Lemma 6.1 as for $\varepsilon > 0$ small enough

$$\liminf_{n\to\infty} p_n^{-1} \log \frac{\sqrt{k_n} m_{p_n}(U_{k+1,n})}{\sigma_{p_n}(U_{k+1,n})} \ge \frac{\zeta}{2} - \log 2 - \log(1+\varepsilon) > 0.$$

□

Proof of Theorem 3.2. First note that $U_{k+1,n} n/k \to 1$ in probability, and since a and ℓ are regularly varying functions $U_{k+1,n}$ can be changed to k/n.

For the first result we have to show that $m_p(k/n)/m_p \to 1$. This follows from Proposition 5.6. Indeed, for any $\varepsilon > 0$

$$\left| \frac{m_p(k/n)}{m_p} - 1 \right| \le K(1+\varepsilon)^p \left(\frac{a(k/n)}{\ell(k/n)} \right)^{\nu_\beta - \varepsilon}.$$

Taking logarithm and dividing by p we see that the right-hand side above is negative for $\varepsilon > 0$ small enough.

For the central limit theorem, $\sigma_p(k/n)/\sigma_p \to 1$ follows again from Proposition 5.6, thus $\sigma_p(U_{k_n+1,n})/\sigma_p \to 1$ also follows as above. To change the centering, using again Proposition 5.6

$$\begin{aligned}\frac{\sqrt{k}}{\sigma_{p_n}} |m_p(k/n) - m_p| &= \frac{m_p \sqrt{k}}{\sigma_p} \frac{|m_p(k/n) - m_p|}{m_p}\\&\le c\sqrt{k}(1+\varepsilon)^p \frac{\Gamma(p+1)}{\sqrt{\Gamma(2p+1)}} \left(\frac{a(k/n)}{\ell(k/n)} \right)^{\widetilde{\nu}},\end{aligned} \tag{80}$$

with $\widetilde{\nu} = 1 \wedge (\nu_\beta - \varepsilon)$. Taking logarithm, dividing by p, and using the Stirling formula

$$\limsup_{p\to\infty} p^{-1} \log \left[\sqrt{k}(1+\varepsilon)^p \frac{\Gamma(p+1)}{\sqrt{\Gamma(2p+1)}} \left(\frac{a(k/n)}{\ell(k/n)} \right)^{\widetilde{\nu}} \right]$$
$$\leq \log(1+\varepsilon) - \log 2 + \limsup_{p\to\infty} p^{-1} \log \left[\sqrt{k} \left(\frac{a(k/n)}{\ell(k/n)} \right)^{\widetilde{\nu}} \right].$$

Since $\varepsilon > 0$ in (80) is as small as we wish, the result follows.

Now (26) follows from (20) using Bogachev's transfer lemma, as above. □

Proof of Theorem 3.3. The proof goes as the previous proof, but we use Proposition 5.5. □

Proof of Theorem 3.4. The first result follows from Proposition 6.6. Combining with Bogachev's transfer lemma we obtain (28) and (30). To use the transfer lemma for $\alpha \in [1, 2)$ we have to check that

$$\frac{k_n \widetilde{m}_{p_n}(U_{k+1,n})}{\eta_{U_{k+1,n}}(p_n)^{p_n}} \to \infty.$$

For $\alpha > 1$ by Lemma 6.1 and (70) the left-hand side above is at least

$$\frac{e^{\alpha p_n}(\gamma - \varepsilon)^{p_n} \Gamma(p_n + 1)}{((\alpha\gamma - \varepsilon)p_n)^{p_n}} \geq \left(\frac{e^\alpha}{\alpha} \right)^{p_n} (1-\varepsilon)^{p_n},$$

which tends to ∞ for $\varepsilon > 0$ small enough. For $\alpha = 1$ the result follows from (74) in Proposition 6.5 with $r = \alpha = 1$.

To see (29), note that as $p_n \to \infty$,

$$\left(\frac{S_n(p_n)}{\Gamma(p_n+1)} \right)^{1/p_n} (k_n \Gamma(p_n+1))^{1/p_n} \frac{1}{\eta_{U_{k+1,n}}(p_n)} = \frac{(k_n S_n(p_n))^{1/p_n}}{\eta_{U_{k+1,n}}(p_n)} \to 1.$$

Thus (29) follows from the asymptotics

$$\frac{(k_n \Gamma(p_n+1))^{1/p_n}}{\eta_{U_{k+1,n}}(p_n)} \to \frac{e^{\alpha-1}}{\alpha\gamma}.$$

□

Proof of Theorem 3.5. First we show that we can change to deterministic normalization, i.e.

$$\lim_{n\to\infty} \left(\frac{\eta_{U_{k+1,n}}(p_n)}{\alpha\gamma p_n} \right)^{p_n} = 1. \tag{81}$$

We have $|\eta_v(x) - \alpha\gamma x| = \gamma |h_v(\eta_v(x)) - h_0(\eta_v(x))|$ by the definition of η_v in (14), recalling that $\zeta = \alpha$. Therefore,

$$\left| \frac{\eta_v(p_n)}{\alpha\gamma p_n} - 1 \right| \leq \frac{1}{\alpha p_n} \left| h_v(\eta_v(p_n)) - h_0(\eta_v(p_n)) \right|,$$

from which we see that (81) follows if we show the convergence

$$\lim_{n\to\infty} \left[h_{U_{k+1,n}}(\eta_{U_{k+1,n}}(p_n)) - h_0(\eta_{U_{k+1,n}}(p_n)) \right] = 0.$$

This holds, since the first term on the right-hand side of (69) tends to 0 as $v = U_{k+1,n} \to 0$ by (12). Changing $U_{k+1,n}$ to k/n, with $v = k/n$ and $x = \eta_v(p_n) \sim \gamma \log k$, we see that the second term tends to 0 by assumption (31). Thus (33) holds for $\alpha < 1$.

For $\alpha \in [1, 2)$ we need to handle the centering as well. For $\alpha > 1$ by Proposition 5.6

$$\frac{k_n |m_{p_n} - m_{p_n}(v)|}{(\alpha\gamma p_n)^{p_n}} \le \frac{2K_1}{\gamma} (1+\varepsilon)^p \frac{e^{\alpha p_n}}{(\alpha p_n)^{p_n}} \Gamma(p_n + 1) \left(\frac{a(v)}{\ell(v)} \right)^{\widetilde{\nu}},$$

with $\widetilde{\nu} = 1 \wedge (\nu_\beta - \varepsilon)$. As before we can substitute $U_{k+1,n}$ to k/n. Thus

$$\begin{aligned} &\limsup_{n\to\infty} p_n^{-1} \log \frac{k_n |m_{p_n} - m_{p_n}(k/n)|}{(\alpha\gamma p_n)^{p_n}} \\ &\le \log(1+\varepsilon) + \alpha - 1 - \log\alpha + \widetilde{\nu} \limsup_{n\to\infty} p_n^{-1} \log \left(\frac{a(k/n)}{\ell(k/n)} \right) \\ &= \log(1+\varepsilon) + H(\alpha) - \widetilde{\nu}\beta_1, \end{aligned}$$

which is negative for $\varepsilon > 0$ small, under our assumptions. Thus (33) follows.

To prove (34), write

$$p_n \left(\frac{[k_n S_n(p_n)]^{1/p_n}}{\alpha\gamma p_n} - 1 \right) = p_n \left(\widehat{\gamma}(n) \frac{(k_n \Gamma(p_n+1))^{1/p_n}}{\alpha\gamma p_n} - 1 \right).$$

Simple calculation shows that

$$\frac{k_n^{1/p_n}}{\alpha\gamma p_n} \Gamma(p_n+1)^{1/p_n} = \frac{e^{\alpha-1}}{\alpha\gamma} \left(1 + \frac{\log(2\pi p_n)}{2p_n} + o(1/p_n) \right).$$

Thus (34) and (35) follows from Bogachev's transfer lemma. For (36) we use the asymptotics of the truncated moments in Lemma 5.4. □

Proof of Theorem 3.6. The proof goes as the previous proof, but we use Proposition 5.5. □

Acknowledgements. We are grateful to the associate editor, to the anonymous referees, and to David Mason for the helpful comments and suggestions which greatly improved the paper. PK's research was partially supported by the János Bolyai Research Scholarship of the Hungarian Academy of Sciences, by the NKFIH grant FK124141, and by the EU-funded Hungarian grant EFOP-3.6.1-16-2016-00008. PK's and LV's research was partially supported by the grant NKFIH-1279-2/2020 of the Ministry for Innovation and Technology, Hungary.

[1] A. Agresti. *Categorical data analysis*. Wiley Series in Probability and Statistics. Wiley-Interscience [John Wiley & Sons], New York, second edition, 2002.

[2] J. Beran, D. Schell, and M. Stehlík. The harmonic moment tail index estimator: asymptotic distribution and robustness. *Ann. Inst. Statist. Math.*, 66(1):193–220, 2014.

[3] 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.

[4] L. Bogachev. Limit laws for norms of IID samples with Weibull tails. *J. Theoret. Probab.*, 19(4):849–873, 2006.

[5] M. F. Brilhante, M. I. Gomes, and D. Pestana. A simple generalisation of the Hill estimator. *Comput. Statist. Data Anal.*, 57:518–535, 2013.

[6] Y. S. Chow and H. Teicher. *Probability theory*. Springer-Verlag, New York-Heidelberg, 1978. Independence, interchangeability, martingales.

[7] G. Ciuperca and C. Mercadier. Semi-parametric estimation for heavy tailed distributions. *Extremes*, 13(1):55–87, 2010.

[8] A. Clauset, C. R. Shalizi, and M. E. J. Newman. Power-law distributions in empirical data. *SIAM Rev.*, 51(4):661–703, 2009.

[9] S. Csörgő and D. M. Mason. Central limit theorems for sums of extreme values. *Math. Proc. Cambridge Philos. Soc.*, 98(3):547–558, 1985.

[10] S. Csörgő and L. Viharos. On the asymptotic normality of Hill's estimator. *Math. Proc. Cambridge Philos. Soc.*, 118(2):375–382, 1995.

[11] L. de Haan. Extreme value statistics. In *Extreme value theory and applications. Vol. 1*, pages 93–122. Kluwer Acad. Publ., Dordrecht, 1994.

[12] L. de Haan and A. Ferreira. *Extreme value theory*. Springer Series in Operations Research and Financial Engineering. Springer, New York, 2006.

[13] L. de Haan and L. Peng. Comparison of tail index estimators. *Statist. Neerlandica*, 52(1):60–70, 1998.

[14] P. Deheuvels, E. Haeusler, and D. M. Mason. Almost sure convergence of the Hill estimator. *Math. Proc. Cambridge Philos. Soc.*, 104(2):371–381, 1988.

[15] A. L. M. Dekkers, J. H. J. Einmahl, and L. de Haan. A moment estimator for the index of an extreme-value distribution. *Ann. Statist.*, 17(4):1833–1855, 1989.

[16] J. H. J. Einmahl and D. M. Mason. Laws of the iterated logarithm in the tails for weighted uniform empirical processes. *Ann. Probab.*, 16(1):126–141, 1988.

[17] P. Embrechts, C. Klüppelberg, and T. Mikosch. *Modelling extremal events*, volume 33 of *Applications of Mathematics (New York)*. Springer-Verlag, Berlin, 1997.

[18] B. V. Gnedenko and A. N. Kolmogorov. *Limit distributions for sums of independent random variables*. Addison-Wesley Publishing Company, Inc., Cambridge, Mass., 1954.

[19] M. I. Gomes and M. J. Martins. Generalizations of the Hill estimator-asymptotic versus finite sample behaviour. *J. Statist. Plann. Inference*, 93(1-2):161–180, 2001.

[20] P. Hall. On some simple estimates of an exponent of regular variation. *J. Roy. Statist. Soc. Ser. B*, 44(1):37–42, 1982.

[21] B. M. Hill. A simple general approach to inference about the tail of a distribution. *Ann. Statist.*, 3(5):1163–1174, 1975.

[22] A. Janßen. Limit laws for power sums and norms of i.i.d. samples. *Probab. Theory Related Fields*, 146(3-4):515–533, 2010.

[23] V. Paulauskas and M. Vaičiulis. A class of new tail index estimators. *Ann. Inst. Statist. Math.*, 69(2):461–487, 2017.

[24] H. Penalva, M. I. Gomes, F. Caeiro, and M. M. Neves. A couple of non reduced bias generalized means in extreme value theory: an asymptotic comparison. *REVSTAT*, 18(3):281–298, 2020.

[25] V. V. Petrov. *Limit theorems of probability theory*, volume 4 of *Oxford Studies in Probability*. The Clarendon Press, Oxford University Press, New York, 1995.

[26] S. Resnick. Discussion of the Danish data on large fire insurance losses. 27(1):139–151, 1997.

[27] M. Schlather. Limit distributions of norms of vectors of positive i.i.d. random variables. *Ann. Probab.*, 29(2):862–881, 2001.

[28] J. Segers. Residual estimators. *J. Statist. Plann. Inference*, 98(1-2):15–27, 2001.

[29] J. Segers. Abelian and Tauberian theorems on the bias of the Hill estimator. *Scand. J. Statist.*, 29(3):461–483, 2002.

[30] J. A. Wellner. Limit theorems for the ratio of the empirical distribution function to the true distribution function. *Z. Wahrsch. Verw. Gebiete*, 45(1):73–88, 1978.