Electronic Journal of Probability

Vol. 16 (2011), Paper no. 74, pages 2039–2058.

Journal URL
http://www.math.washington.edu/~ejpecp/

Asymptotic normality of the Hill estimator for truncated data*

Arijit Chakrabarty
Statistics and Mathematics Unit
Indian Statistical Institute
7 S.J.S. Sansanwal Marg
New Delhi 110016, India
arijit@isid.ac.in

Abstract

The problem of estimating the tail index from truncated data is addressed in [2]. In that paper, a sample based (and hence random) choice of k is suggested, and it is shown that the choice leads to a consistent estimator of the inverse of the tail index. In this paper, the second order behavior of the Hill estimator with that choice of k is studied, under some additional assumptions. In the untruncated situation, asymptotic normality of the Hill estimator is well known for distributions whose tail belongs to the Hall class, see [11]. Motivated by this, we show the same in the truncated case for that class .

Key words: heavy tails, truncation, second order regular variation, Hill estimator, asymptotic normality.

AMS 2010 Subject Classification: Primary 62G32.

Submitted to EJP on February 12, 2011, final version accepted September 5, 2011.

*Research partially supported by the NSF grant "Graduate and Postdoctoral Training in Probability and its Applications" at Cornell University, the Centenary Post Doctoral Fellowship at the Indian Institute of Science and a fellowship from the National Board of Higher Mathematics.

1 Introduction

Historically, one of the most important statistical issues related to distributions with regularly varying tail is estimating the tail index. A detailed discussion on estimators of the tail index can be found in Chapter 4 of [5]. One of the most popular estimators is the Hill estimator, introduced by [12]. For a one-dimensional non-negative sample $X_1, \ldots, X_n$, the Hill statistic is defined as

$$h(k,n) := \frac{1}{k}\sum_{i=1}^{k} \log \frac{X_{(i)}}{X_{(k)}}, \tag{1.1}$$

where $X_{(1)} \geq \ldots \geq X_{(n)}$ are the order statistics of $X_1, \ldots, X_n$, and $1 \leq k \leq n$ is an user determined parameter. It is well known that if $X_1, \ldots, X_n$ are a i.i.d. sample from a distribution whose tail is regularly varying with index $-\alpha$ and (k_n) is a sequence of integers satisfying $1 \ll k_n \ll n$ (where "$a_n \ll b_n$" means "$a_n = o(b_n)$", and "$\gg$" means the obvious opposite throughout the paper), then $h(k_n, n)$ consistently estimates α^{-1}. In a sense made precise by [14], the consistency of Hill statistic is equivalent to the regular variation of the tail of the underlying distribution. Various authors have studied the second order behavior of the Hill estimator; see for example [4], [3], [10], [9], [8] and [6] among others. It is well known that if the tail of the i.i.d. random variables $X_1, \ldots, X_n$ belongs to the Hall class (defined in Assumption B below) then

$$\sqrt{k}\left(h(k,n) - \frac{1}{\alpha}\right) \Longrightarrow N\left(0, \frac{1}{\alpha^2}\right).$$

This assumption is stronger than just assuming that the tail is regularly varying with index $-\alpha$.

While there are real life phenomena that do exhibit the presence of heavy tails, in lot of the cases there is a physical upper bound on the possible values. For example most internet service providers put an upper bound on the size of a file that can be transferred using an internet connection provided by them. Clearly the natural model for such phenomena is a truncated heavy-tailed distribution, a distribution which fits a heavy-tailed distribution till a certain point and then decays significantly faster. This can be made precise in the following way. Suppose that $H, H_1, \ldots$ are i.i.d. random variables so that $P(H > \cdot)$ is regularly varying with index $-\alpha$, $\alpha > 0$ and that $L, L_1, L_2, \ldots$ are i.i.d. random variables independent of $(H, H_1, H_2, \ldots)$. All these random variables are assumed to take values in the positive half line. We observe the sample $X_1, \ldots, X_n$ given by

$$X_j := H_j \mathbf{1}(H_j \leq M_n) + (M_n + L_j)\mathbf{1}(H_j > M_n), \tag{1.2}$$

where M_n, representing the truncating threshold, is a sequence of positive numbers going to infinity. Strictly speaking, the model is actually a triangular array $\{X_{nj} : 1 \leq j \leq n\}$. However, in practice we shall observe only one row of the triangular array, and hence we denote the sample by the usual notation $X_1, \ldots, X_n$. The random variable L can be thought of to have a much lighter tail, a tail decaying exponentially fast for example. However the results of this article are true under milder assumptions. The motivation behind adding this component in the model is that there are situations where a power law fits the bulk of the data, while the tail tapers off much faster. For example, human inter contact times were studied by [13]. Their finding is that a truncated Pareto distribution, a model that "has a power law tendency at the head part and decays exponentially at the tail", is appropriate. The results of the current paper, however, are true when $L \equiv 0$, and hence the reader may choose to think of the model (1.2) without the L.

It was observed in Chakrabarty and Samorodnitsky (2009) that if the sequence M_n goes to infinity slow enough so that

$$\lim_{n\to\infty} nP(H > M_n) = \infty\,, \tag{1.3}$$

then a priori choosing a k so that the Hill estimator is consistent is a problem. The problem stems from the fact that under the assumption of hard truncation (1.3), in order to be consistent, k_n should satisfy

$$nP(H > M_n) \ll k_n \ll n\,.$$

A natural choice of k_n would be

$$k_n = nP(H > M_n)^\beta\,,$$

for some $\beta \in (0,1)$ which is meant to be specified by the user. However the problem with this choice is that the quantity $P(H > M_n)$ is unknown. In order to overcome that problem, the following sample based choice of k was suggested in that paper:

$$\hat{k}_n := \left[n \left(\frac{1}{n} \sum_{j=1}^{n} \mathbf{1}(X_j > \gamma X_{(1)}) \right)^\beta \right]\,, \tag{1.4}$$

where $\gamma \in (0,1)$ is also a user determined parameter. The intuition behind this choice is that the ration of k_n and $\hat{k}_n$ as defined above, will converge in probability, to a known constant. It has been shown in that article that this choice of $\hat{k}_n$ leads to a consistent estimator of α^{-1} when (1.3) is true, or when that limit is zero.

In this paper, we investigate the second order behavior of $h(\hat{k}_n, n)$ under the assumption (1.3) and some additional assumptions. We hope to address the case when the corresponding limit is zero in future. Knowing the second order behavior of an estimator, at least asymptotically, helps in constructing confidence intervals for the unknown parameter. While the problem is motivated by statistics, it is an interesting mathematical problem in itself. The complexity in analyzing the second order behavior of $h(\hat{k}_n, n)$ arises from the fact that now we are dealing with a random sum, and the number of summands is heavily dependent on the summands themselves. Also, a quick inspection will reveal that conditioning on the number of summands will completely destroy the i.i.d. nature of the sample, and thus make the analysis even more difficult.

In Section 2, the result that the Hill estimator with $k = \hat{k}_n$ is asymptotically normal with mean $1/\alpha$, which is the main result of the article, is stated. The result is proved in Section 3. A few examples are studied in Section 4.

2 Asymptotic normality of the Hill estimator

Suppose that we have a one-dimensional non-negative sample $X_1, \ldots, X_n$ given by (1.2). The following assumption makes precise the idea that the random variable L has "a much lighter tail".
Assumption A: There exists a sequence (ε_n) such that

$$\begin{aligned} \lim_{n\to\infty} P(H > M_n)^{-(1-\beta)}\varepsilon_n &= 0\,, && (2.1)\\ \text{and } \lim_{n\to\infty} nP(H > M_n)P(L > \varepsilon_n M_n) &= 0\,. && (2.2) \end{aligned}$$

Our next assumption is that the tail of H is in the Hall class (see [11]).
Assumption B: As $x \to \infty$,

$$P(H > x) = Cx^{-\alpha} + O\left(x^{-\bar{\alpha}}\right),$$

for some $C > 0$ and $0 < \alpha < \bar{\alpha}$.

The main result of this paper, Theorem 2.1 below, describes the second order behavior of $h(\hat{k}_n, n)$, where $h(\cdot,\cdot)$ and $\hat{k}_n$ are as defined in (1.1) and (1.4) respectively, under suitable bounds on the growth rate of M_n. The proof is postponed until Section 3.

Theorem 2.1. *Suppose that assumptions A and B hold, β is in the range*

$$\frac{\alpha}{2\bar{\alpha} - \alpha} < \beta < 1,$$

and M_n satisfies

$$M_n^\alpha \ll n \ll \min\left\{(\log M_n)^{-2} M_n^{\alpha(2-\beta)}, M_n^{\beta(2\bar{\alpha}-\alpha)}\right\}, \tag{2.3}$$

as $n \to \infty$. Then,

$$\sqrt{\hat{k}_n}\left\{h(\hat{k}_n, n) - \frac{1}{\alpha}\right\} \Longrightarrow N\left(0, \frac{1}{\alpha^2}\right). \tag{2.4}$$

We next point out that for random variables whose tails are second order regularly varying as defined in, for example, (2.3.24) of [5], then Assumption B is automatically satisfied; see the discussion following Example 2.3.11 on page 49 of the same reference. By the tail being second order regularly varying, we mean that there is a function $A : (0, \infty) \longrightarrow \mathbb{R}$ which is regularly varying with index $\rho\alpha$ where $\rho < 0$, such that

$$\lim_{t\to\infty} \frac{\frac{P(H>tx)}{P(H>t)} - x^{-\alpha}}{A(t)} = x^{-\alpha} \frac{x^{\rho\alpha} - 1}{\rho/\alpha} \tag{2.5}$$

for all $x > 0$.

Corollary 1. *Suppose that Assumption A holds, and the tail of H satisfies* (2.5). *Let*

$$(1 - 2\rho)^{-1} < \beta < 1,$$

and M_n's satisfy

$$M_n^\alpha \ll n \ll \min\left\{(\log M_n)^{-2} M_n^{\alpha(2-\beta)}, M_n^{\alpha\beta(1-2\rho)}\right\}.$$

Then (2.4) *holds.*

We end this section with a couple of remarks.

Remark 1

Assumption B and (2.5) are not equivalent. Example 3 below satisfies the former but not the latter.

Remark 2

It would be nice if Corollary 1 could be extended to cases with $\rho = 0$. Unfortunately, the author has not been able to achieve that.

3 Proof of Theorem 2.1

We start with a brief outline of the plan of the proof. Define

$$\begin{aligned} U_n &:= \sum_{j=1}^{n} \mathbf{1}(X_j > \gamma M_n), \\ V_n &:= \sum_{j=1}^{n} \mathbf{1}(X_j > \gamma X_{(1)}), \\ \tilde{k}_n &:= \left[n^{1-\beta} U_n^{\beta} \right]. \end{aligned}$$

Note that

$$\hat{k}_n := \left[n^{1-\beta} V_n^{\beta} \right].$$

Since we are dealing with a random sum, a natural way of proceeding is conditioning on the number of summands. However, as commented earlier, conditioning on V_n or $\hat{k}_n$ destroys the i.i.d. nature of the sample. Hence, we condition on $U_n = u_n$, where (u_n) is any sequence of integers satisfying $u_n \sim nP(H > \gamma M_n)$. Lemma 3.1 is a general result, which allows us to claim weak convergence of the unconditional distribution based on that of the conditional distribution. Clearly, by conditioning on U_n, $h(\tilde{k}_n, n)$ becomes the Hill statistic with a deterministic k applied to a triangular array. The second order behavior of that is studied in Lemma 3.4. In view of Lemma 3.1, this translates to second order behavior of (the unconditional distribution of) $h(\tilde{k}_n, n)$. In order to argue the claim of Theorem 2.1, all we need is showing that $h(\tilde{k}_n, n)$ and $h(\hat{k}_n, n)$ are not very far apart, and that is done in Lemma 3.5. For Lemma 3.4 and Lemma 3.5, we need that the tail empirical process, after suitable centering and scaling, converge to a Brownian Motion. This is shown in Lemma 3.3.

We now proceed to execute the plan described above. Throughout this section, the hypotheses of Theorem 2.1 will be assumed, even it is not explicitly mentioned.

Lemma 3.1. *Suppose that $(B_n : n \geq 1)$ is a sequence of discrete random variables satisfying*

$$\frac{B_n}{b_n} \xrightarrow{P} 1,$$

for some deterministic sequence (b_n). Assume that $(A_n : n \geq 1)$ is a family of random variables such that whenever $\hat{b}_n$ is any deterministic sequence satisfying $\hat{b}_n \sim b_n$ as $n \longrightarrow \infty$ and $P(B_n = \hat{b}_n) > 0$, it follows that

$$P(A_n \leq \cdot | B_n = \hat{b}_n) \Longrightarrow F(\cdot), \tag{3.1}$$

for some c.d.f. F. Then $A_n \Longrightarrow F$.

Proof. It suffices to show that every subsequence of (A_n) has a further subsequence that converges weakly to F. Since every sequence that converges in probability has a subsequence that converges almost surely, we can assume without loss of generality that

$$\frac{B_n}{b_n} \longrightarrow 1 \text{ a.s.}. \tag{3.2}$$

Fix a continuity point x of F and define a function $f_n : \mathbb{R} \longrightarrow [0,1]$ by

$$f_n(u) = \begin{cases} \frac{P(A_n \le x, B_n = u)}{P(B_n = u)}, & \text{if } P(B_n = u) > 0 \\ 0, & \text{otherwise.} \end{cases}$$

Clearly, for all $n \ge 1$,

$$P(A_n \le x) = E f_n(B_n)\,.$$

By (3.1) and (3.2), it follows that

$$f_n(B_n) \longrightarrow F(x) \text{ a.s.}.$$

By the bounded convergence theorem, it follows that

$$\lim_{n\to\infty} E f_n(B_n) = F(x)\,,$$

and this completes the proof. □

Lemma 3.2. *For any sequence* (v_n) *satisfying*

$$v_n \sim nP(H > \gamma M_n)^\beta\,, \tag{3.3}$$

it holds that

$$\lim_{n\to\infty} \sqrt{v_n}\left[\frac{n}{v_n} P\left(H > b(n/v_n) y^{-1/\alpha}\right) - y\right] = 0 \tag{3.4}$$

uniformly on compact sets in $[0,\infty)$*, and*

$$\lim_{T\to\infty} \limsup_{n\to\infty} \sqrt{v_n} \int_T^\infty \left|\frac{n}{v_n} P\left(H > b(n/v_n)s\right) - s^{-\alpha}\right| \frac{ds}{s} = 0\,. \tag{3.5}$$

where

$$b(y) := \inf\left\{x : \frac{1}{P(H > x)} \ge y\right\}. \tag{3.6}$$

Proof. Assumption B means that there is $C \in (0,\infty)$ with

$$f_1(x) := x^{-\alpha} - Cx^{-\bar\alpha} \le P(H > x) \le x^{-\alpha} + Cx^{-\bar\alpha} =: f_2(x)\,,$$

for all x. The function f_2 is strictly decreasing, whereas f_1 is eventually so. Thus, for y large,

$$f_1^{-1}(1/y) \le b(y) \le f_2^{-1}(1/y)\,.$$

Clearly, for $\xi > 0$,

$$f_2^{-1}(\xi) \le \left(\xi - C\xi^{\bar\alpha/\alpha}\right)^{-1/\alpha} = \xi^{-1/\alpha} + O\left(\xi^{(\bar\alpha-\alpha-1)/\alpha}\right) \text{ as } \xi \downarrow 0\,.$$

A similar estimate holds for f_1^{-1}, and they together show that

$$b(y) = y^{1/\alpha} + O\left(y^{(1+\alpha-\bar\alpha)/\alpha}\right),$$

as $y \to \infty$. Thus, given $\delta > 0$, there is $\bar C < \infty$ such that

$$\left|P(H > x b(y)) - x^{-\alpha} y^{-1}\right| \le \bar C x^{-\alpha} y^{-\bar\alpha/\alpha} \text{ for all } x \ge \delta, y \ge \bar C\,.$$

Let (v_n) be a sequence satisfying (3.3). Thus, for n large and $x \geq \delta$,

$$\sqrt{v_n}\left|\frac{n}{v_n}P(H > b(n/v_n)x) - x^{-\alpha}\right| \leq \bar{C}x^{-\alpha}\frac{v_n^{\bar{\alpha}/\alpha-1/2}}{n^{\bar{\alpha}/\alpha-1}}\,.$$

Notice that

$$\begin{aligned}\frac{v_n^{\bar{\alpha}/\alpha-1/2}}{n^{\bar{\alpha}/\alpha-1}} &= O\left(n^{1/2}P(H > M_n)^{\beta(\bar{\alpha}/\alpha-1/2)}\right)\\ &= o(1)\,,\end{aligned}$$

the last line following from (2.3). These show (3.4) and (3.5). □

Lemma 3.3. *Suppose that* (u_n) *is a sequence of integers satisfying*

$$u_n \sim nP(H > \gamma M_n)\,, \tag{3.7}$$

and let

$$\begin{aligned}v_n &:= [n^{1-\beta}u_n^{\beta}] - u_n\,, &(3.8)\\ \tilde{M}_n &:= \gamma M_n\,. &(3.9)\end{aligned}$$

Let for $n \geq 1$, $Y_{n,1},\ldots,Y_{n,n}$ *be i.i.d. with c.d.f.* F_n*, defined as*

$$F_n(x) := P(H \leq x | H \leq \tilde{M}_n)\,.$$

Then,

$$\sqrt{v_n}\left(\frac{1}{v_n}\sum_{i=1}^{n-u_n}\delta_{Y_{n-u_n,i}/b((n-u_n)/v_n)}(y^{-1/\alpha},\infty] - y\right) \Longrightarrow W(y) \tag{3.10}$$

in $D[0,\infty)$*, where* $D[0,\infty)$ *is endowed with the topology of uniform convergence on compact sets and* W *is the standard Brownian Motion on* $[0,\infty)$*.*

Proof. For simplicity sake, denote $w_n := n - u_n$. It is easy to see by (2.3) that

$$1 \ll w_n P(H > \tilde{M}_n) \ll \sqrt{v_n} \ll \sqrt{w_n}\,. \tag{3.11}$$

Let $(\Gamma_i : i \geq 1)$ be the arrivals of a unit rate Poisson Process. Define

$$\phi_n(s) := \frac{\Gamma_{w_n+1}}{v_n}\bar{F}_n(s^{-1/\alpha}b(w_n/v_n))\,,$$

where $\bar{G} := 1 - G$ for any function G. By the discussion on page 24 in [15], it follows that

$$\lim_{n\to\infty}\frac{w_n}{v_n}P(H > b(w_n/v_n)) = 1\,. \tag{3.12}$$

It follows by (3.11) that

$$\lim_{n\to\infty}\frac{w_n}{v_n}P(H > \tilde{M}_n) = 0\,.$$

This in conjunction with (3.12) implies that

$$b(w_n/v_n) = o(\tilde{M}_n)\,.$$

It is easy to see that v_n satisfies (3.3). Hence, for n large enough,

$$\begin{aligned}
& \frac{w_n}{v_n}\bar{F}_n(s^{-1/\alpha}b(w_n/v_n)) - s \\
= \quad & \frac{1}{P(H \le \tilde{M}_n)}\left[\frac{w_n}{v_n}P\left(H > s^{-1/\alpha}b(w_n/v_n)\right) - \frac{w_n}{v_n}P(H > \tilde{M}_n)\right. \\
& \left. -s + sP(H > \tilde{M}_n)\right],
\end{aligned}$$

and hence in view of (3.4) and (3.11), it follows that for $0 < T < \infty$,

$$\lim_{n\to\infty}\sqrt{v_n}\sup_{0\le s\le T}\left|\frac{w_n}{v_n}\bar{F}_n(s^{-1/\alpha}b(w_n/v_n)) - s\right| = 0\,. \tag{3.13}$$

Also note that,

$$\begin{aligned}
& \sup_{0\le s\le T}\left|\phi_n(s) - \frac{w_n}{v_n}\bar{F}_n(s^{-1/\alpha}b(w_n/v_n))\right| \\
= \quad & \left|\frac{\Gamma_{w_n+1}}{w_n} - 1\right|\frac{w_n}{v_n}\bar{F}_n(T^{-1/\alpha}b(w_n/v_n)) \\
= \quad & O_p(w_n^{-1/2})O(1) \\
= \quad & o_p(v_n^{-1/2})\,.
\end{aligned}$$

This in conjunction with (3.13) shows that

$$\sqrt{v_n}\left(\phi_n(s) - s\right) \xrightarrow{P} 0 \tag{3.14}$$

in $D[0,\infty)$. Recall that since $1 \ll v_n \ll w_n$, in $D[0,\infty)$,

$$\sqrt{v_n}\left(\frac{1}{v_n}\sum_{i=1}^{w_n}\mathbf{1}\left(\Gamma_i \le v_n s\right) - s\right) \Longrightarrow W(s)\,;$$

see (9.7), page 294 in [15]. Hence, it follows by the continuous mapping theorem and Slutsky's theorem that

$$\sqrt{v_n}\left(\frac{1}{v_n}\sum_{i=1}^{w_n}\mathbf{1}\left(\Gamma_i \le v_n\phi_n(s)\right) - \phi_n(s)\right) \Longrightarrow W(s) \tag{3.15}$$

in $D[0,\infty)$. By similar arguments as those in the proof of Theorem 9.1 in [15], it follows that

$$\sum_{i=1}^{w_n}\delta_{Y_{w_n,i}/b(w_n/v_n)}(y^{-1/\alpha},\infty] \stackrel{d}{=} \sum_{i=1}^{w_n}\mathbf{1}\left(\Gamma_i \le v_n\phi_n(s)\right)\,.$$

This along with (3.14) and (3.15) shows (3.10). □

Lemma 3.4. *Let (u_n) be a sequence of integers satisfying* (3.7) *and let (v_n) and $(\tilde{M}_n)$ be as defined in* (3.8) *and* (3.9) *respectively. Then,*

$$\sqrt{v_n}\left(\frac{1}{v_n}\sum_{i=1}^{v_n}\log\frac{Y_{(n-u_n,i)}}{Y_{(n-u_n,v_n)}}-\frac{1}{\alpha}\right)\Longrightarrow N\left(0,\frac{1}{\alpha^2}\right),$$

where $Y_{(n,1)}\geq\ldots\geq Y_{(n,n)}$ are the order statistics of $Y_{n,1},\ldots,Y_{n,n}$, and the latter is as defined in Lemma 3.3.

Proof. Once again, let us denote $w_n:=n-u_n$. An application of Vervaat's lemma (Proposition 3.3 in [15]) to (3.10) shows that

$$\sqrt{v_n}\left[\left\{\frac{Y_{(w_n,v_n)}}{b(w_n/v_n)}\right\}^{-\alpha}-1\right]\Longrightarrow -W(1) \tag{3.16}$$

jointly with (3.10). This in particular, shows that

$$\left(\sqrt{v_n}\left\{\frac{1}{v_n}\sum_{i=1}^{w_n}\delta_{Y_{w_n,i}/b(w_n/v_n)}(x,\infty]-x^{-\alpha}\right\},\frac{Y_{(w_n,v_n)}}{b(w_n/v_n)}\right)$$
$$\Longrightarrow (W(x^{-\alpha}),1),$$

in $D(0,\infty]\times\mathbb{R}$, jointly with (3.16), where $D(0,\infty]$ is also endowed with the topology of uniform convergence on compact sets. Using the continuous mapping theorem, it follows that

$$\sqrt{v_n}\left\{\frac{1}{v_n}\sum_{i=1}^{w_n}\delta_{Y_{w_n,i}/Y_{(w_n,v_n)}}(x,\infty]-x^{-\alpha}\frac{Y^{-\alpha}_{(w_n,v_n)}}{b(w_n/v_n)^{-\alpha}}\right\}$$
$$\Longrightarrow W(x^{-\alpha}), \tag{3.17}$$

in $D(0,\infty]$, jointly with (3.16). As in the proof of Proposition 9.1 in [15], we shall apply the map ψ from $D(0,\infty]$ to $\mathbb{R}$, defined by

$$\psi(f):=\int_1^\infty f(s)\frac{ds}{s},$$

to conclude that

$$\sqrt{v_n}\left\{\frac{1}{v_n}\sum_{i=1}^{v_n}\log\frac{Y_{(w_n,i)}}{Y_{(w_n,v_n)}}-\frac{1}{\alpha}\frac{Y^{-\alpha}_{(w_n,v_n)}}{b(w_n/v_n)^{-\alpha}}\right\}\Longrightarrow\int_1^\infty W(x^{-\alpha})\frac{dx}{x}, \tag{3.18}$$

jointly with (3.16). This implies that

$$\sqrt{v_n}\left\{\frac{1}{v_n}\sum_{i=1}^{v_n}\log\frac{Y_{(n,i)}}{Y_{(n,v_n)}}-\frac{1}{\alpha}\right\}\Longrightarrow\int_1^\infty W(x^{-\alpha})\frac{dx}{x}-\frac{1}{\alpha}W(1)$$

as desired. Thus, it suffices to show (3.18).

To that end, note that for $1<T<\infty$, the map ψ_T, defined by

$$\psi_T(f):=\int_1^T f(s)\frac{ds}{s}$$

is continuous and has compact support. Also, as $T \longrightarrow \infty$,

$$\psi_T(W(s^{-\alpha})) \Longrightarrow \psi(W(s^{-\alpha}))\,.$$

Some calculations will show that ψ applied to the left hand side of (3.17) gives the left hand side of (3.18). Thus, all that needs to be done is justifying the application of ψ to (3.17), and for that, it suffices to check that for all $\epsilon > 0$,

$$\lim_{T\to\infty}\limsup_{n\to\infty} P\Bigg[\sqrt{v_n}\int_T^\infty \Bigg|\frac{1}{v_n}\sum_{i=1}^{w_n}\delta_{Y_{w_n,i}/Y_{(w_n,v_n)}}(x,\infty] \\ -x^{-\alpha}\frac{Y_{(w_n,v_n)}^{-\alpha}}{b(w_n/v_n)^{-\alpha}}\Bigg|\frac{dx}{x} > \epsilon\Bigg] = 0\,.$$

Note that on the set $\{Y_{(w_n,v_n)}/b(w_n/v_n) > 1/2\}$,

$$\begin{aligned} &\int_T^\infty \Bigg|\frac{1}{v_n}\sum_{i=1}^{w_n}\delta_{Y_{w_n,i}/Y_{(w_n,v_n)}}(x,\infty] - x^{-\alpha}\frac{Y_{(w_n,v_n)}^{-\alpha}}{b(w_n/v_n)^{-\alpha}}\Bigg|\frac{dx}{x} \\ = &\int_{TY_{(w_n,v_n)}/b(w_n/v_n)}^\infty \Bigg|\frac{1}{v_n}\sum_{i=1}^{w_n}\delta_{Y_{w_n,i}/b(w_n/v_n)}(u,\infty] - u^{-\alpha}\Bigg|\frac{du}{u} \\ \le &\int_{T/2}^\infty \Bigg|\frac{1}{v_n}\sum_{i=1}^{w_n}\delta_{Y_{w_n,i}/b(w_n/v_n)}(u,\infty] - u^{-\alpha}\Bigg|\frac{du}{u}\,. \end{aligned}$$

Since $P[Y_{(w_n,v_n)}/b(w_n/v_n) \le 1/2]$ goes to zero, it suffices to show that

$$\lim_{T\to\infty}\limsup_{n\to\infty} P\Bigg[\sqrt{v_n}\int_{T/2}^\infty \Bigg|\frac{1}{v_n}\sum_{i=1}^{w_n}\delta_{Y_{w_n,i}/b(w_n/v_n)}(u,\infty] \\ -u^{-\alpha}\Bigg|\frac{du}{u} > \epsilon\Bigg] = 0\,. \tag{3.19}$$

Clearly,

$$\begin{aligned} &\int_{T/2}^\infty \Bigg|\frac{1}{v_n}\sum_{i=1}^{w_n}\delta_{Y_{w_n,i}/b(w_n/v_n)}(u,\infty] - u^{-\alpha}\Bigg|\frac{du}{u} \\ \le &\int_{T/2}^\infty \Bigg|\frac{1}{v_n}\sum_{i=1}^{w_n}\delta_{Y_{w_n,i}/b(w_n/v_n)}(u,\infty] - \frac{w_n}{v_n}\bar{F}_n\left(ub(w_n/v_n)\right)\Bigg|\frac{du}{u} \\ &+\frac{w_n}{v_n}\int_{T/2}^\infty \left|\bar{F}_n\left(ub(w_n/v_n)\right) - P\left(H > ub(w_n/v_n)\right)\right|\frac{du}{u} \\ &+\int_{T/2}^\infty \left|\frac{w_n}{v_n}P\left(H > ub(w_n/v_n)\right) - u^{-\alpha}\right|\frac{du}{u} \end{aligned}$$

$$
\begin{aligned}
= \quad & \int_{T/2}^{\infty} \left| \frac{1}{v_n} \sum_{i=1}^{w_n} \delta_{Y_{w_n,i}/b(w_n/v_n)}(u,\infty] - \frac{w_n}{v_n} \bar{F}_n \left(ub(w_n/v_n) \right) \right| \frac{du}{u} \\
& + \frac{w_n}{v_n} \int_{T/2}^{\tilde{M}_n/b(w_n/v_n)} \left| \bar{F}_n \left(ub(w_n/v_n) \right) - P\left(H > ub(w_n/v_n) \right) \right| \frac{du}{u} \\
& + \frac{w_n}{v_n} \int_{\tilde{M}_n}^{\infty} P(H > u) \frac{du}{u} \\
& + \int_{T/2}^{\infty} \left| \frac{w_n}{v_n} P\left(H > ub(w_n/v_n) \right) - u^{-\alpha} \right| \frac{du}{u} \\
=: \quad & I_1 + I_2 + I_3 + I_4 \,.
\end{aligned}
$$

Since v_n is defined by (3.8), (3.3) holds. By (3.5), it follows that

$$
\lim_{T\to\infty} \limsup_{n\to\infty} \sqrt{v_n} I_4 = 0 \,.
$$

Karamata's theorem (Theorem VIII.9.1, page 281 in [7]) implies that

$$
I_3 = O\left(\frac{w_n}{v_n} P(H > \tilde{M}_n) \right) = o\left(v_n^{-1/2} \right) ,
$$

the second equality following from (3.11). For I_2, note that

$$
\begin{aligned}
& \bar{F}_n \left(ub(w_n/v_n) \right) - P\left(H > ub(w_n/v_n) \right) \\
= \quad & -\frac{P(H > \tilde{M}_n) P\left(H \le ub(w_n/v_n) \right)}{P(H \le \tilde{M}_n)} \,.
\end{aligned}
$$

Also, it is easy to see from (2.3) that

$$
\lim_{n\to\infty} \frac{w_n P(H > \tilde{M}_n)}{\sqrt{v_n}} \log \left\{ \frac{\tilde{M}_n}{b(w_n/v_n)} \right\} = 0 \,. \tag{3.20}
$$

Thus,

$$
I_2 = O\left(\frac{w_n}{v_n} P(H > \tilde{M}_n) \log \frac{\tilde{M}_n}{b(w_n/v_n)} \right) = o\left(v_n^{-1/2} \right) ,
$$

the second equality following from (3.20).

Thus, all that remains is showing

$$
\lim_{T\to\infty} \limsup_{n\to\infty} P[\sqrt{v_n} I_1 > \epsilon] = 0 \,. \tag{3.21}
$$

Notice that

$$
E\left[\frac{1}{v_n} \sum_{i=1}^{w_n} \delta_{Y_{w_n,i}/b(w_n/v_n)}(u,\infty] \right] = \frac{w_n}{v_n} \bar{F}_n \left(ub(w_n/v_n) \right) \,.
$$

Letting C to be a finite positive constant independent of n, whose value may change from line to line, observe that

$$
\begin{aligned}
& P[\sqrt{v_n} I_1 > \epsilon] \\
\leq\ & \frac{\sqrt{v_n}}{\epsilon} E(I_1) \\
=\ & C\sqrt{v_n}\int_{T/2}^{\infty} E\left|\frac{1}{v_n}\sum_{i=1}^{w_n}\delta_{Y_{w_n,i}/b(w_n/v_n)}(u,\infty] - \frac{w_n}{v_n}\bar{F}_n\left(ub(w_n/v_n)\right)\right|\frac{du}{u} \\
\leq\ & C\sqrt{v_n}\int_{T/2}^{\infty}\mathrm{Var}\left[\frac{1}{v_n}\sum_{i=1}^{w_n}\delta_{Y_{w_n,i}/b(w_n/v_n)}(u,\infty]\right]^{1/2}\frac{du}{u} \\
\leq\ & C\frac{\sqrt{w_n}}{\sqrt{v_n}}\int_{T/2}^{\infty}\bar{F}_n\left(ub(w_n/v_n)\right)^{1/2}\frac{du}{u} \\
\leq\ & C\int_{T/2}^{\infty}\frac{\sqrt{w_n}}{\sqrt{v_n}}P\left(H > ub(w_n/v_n)\right)^{1/2}\frac{du}{u}.
\end{aligned}
$$

By (3.12), the integrand clearly converges to $u^{-\alpha/2}$ as $n \longrightarrow \infty$. By (3.6), the integrand is bounded above by

$$
\left[\frac{P(H > ub(w_n/v_n))}{P(H > b(w_n/v_n))}\right]^{1/2},
$$

which by the Potter bounds (Proposition 2.6 in [15]) is bounded above by $2u^{-\alpha/3}$ for n large enough. An appeal to the dominated convergence theorem shows (3.21) and thus completes the proof. □

Lemma 3.5. *As $n \longrightarrow \infty$,*

$$
\sqrt{\tilde{k}_n}\left\{h(\tilde{k}_n, n) - h(\hat{k}_n, n)\right\} \xrightarrow{P} 0. \tag{3.22}
$$

Proof. We start with showing that

$$
\sqrt{\hat{k}_n}\left[\frac{\hat{k}_n}{\tilde{k}_n} - 1\right] \xrightarrow{P} 0. \tag{3.23}
$$

In the proof of Theorem 3.2 in Chakrabarty and Samorodnitsky (2009), it has been shown that under the assumption of hard truncation (left inequality in (2.3)),

$$
\frac{U_n}{nP(H > \gamma M_n)} \xrightarrow{P} 1, \tag{3.24}
$$

$$
\frac{V_n}{nP(H > \gamma M_n)} \xrightarrow{P} 1, \tag{3.25}
$$

$$
\text{and } \frac{\hat{k}_n}{nP(H > \gamma M_n)^{\beta}} \xrightarrow{P} 1. \tag{3.26}
$$

In view of (3.26), it suffices to show that

$$
n^{1/2}P(H > M_n)^{\beta/2}\left[\frac{\hat{k}_n}{\tilde{k}_n} - 1\right] \xrightarrow{P} 0.
$$

Note that,

$$\frac{n^{1-\beta}V_n^\beta}{n^{1-\beta}U_n^\beta+1} \le \frac{\hat{k}_n}{\tilde{k}_n} \le \frac{n^{1-\beta}V_n^\beta+1}{n^{1-\beta}U_n^\beta},$$

$$\frac{n^{1-\beta}V_n^\beta}{n^{1-\beta}U_n^\beta+1} \le \left(\frac{V_n}{U_n}\right)^\beta \le \frac{n^{1-\beta}V_n^\beta+1}{n^{1-\beta}U_n^\beta},$$

and

$$\begin{aligned}\frac{n^{1-\beta}V_n^\beta+1}{n^{1-\beta}U_n^\beta} - \frac{n^{1-\beta}V_n^\beta}{n^{1-\beta}U_n^\beta+1} &= \frac{n^{1-\beta}V_n^\beta+n^{1-\beta}U_n^\beta+1}{n^{1-\beta}U_n^\beta(n^{1-\beta}U_n^\beta+1)} \\ &= O_p\left(n^{-1}P(H>M_n)^{-\beta}\right) \\ &= o_p\left(n^{-1/2}P(H>M_n)^{-\beta/2}\right),\end{aligned}$$

the equality in the second line following from (3.24) and (3.25), and that in the third line following from (2.3). Thus, it suffices to show that

$$n^{1/2}P(H>M_n)^{\beta/2}\left[\left(\frac{V_n}{U_n}\right)^\beta - 1\right] \xrightarrow{P} 0.$$

By the mean value theorem, it follows that as $x \longrightarrow 1$,

$$x^\beta - 1 = O(|x-1|).$$

Hence, in view of the fact that V_n/U_n converges to 1 in probability, it suffices to show that

$$n^{1/2}P(H>M_n)^{\beta/2}\left(\frac{V_n}{U_n}-1\right) \xrightarrow{P} 0.$$

Using (3.24) once again, all that needs to be shown is

$$V_n - U_n = o_p\left(n^{-1/2}P(H>M_n)^{-(1-\beta/2)}\right).$$

Note that on the set $\{M_n \le X_{(1)} \le M_n(1+\varepsilon_n)\}$, where ε_n is chosen to satisfy Assumption A,

$$0 \le U_n - V_n \le \sum_{j=1}^n \mathbf{1}\left(\gamma M_n < X_j \le \gamma M_n(1+\varepsilon_n)\right) =: T_n.$$

Thus, it suffices to show that

$$\lim_{n\to\infty} P(X_{(1)} \le M_n(1+\varepsilon_n)) = 1, \tag{3.27}$$

$$\lim_{n\to\infty} P(X_{(1)} \ge M_n) = 1, \tag{3.28}$$

$$\text{and } T_n = o_p\left(n^{-1/2}P(H>M_n)^{-(1-\beta/2)}\right). \tag{3.29}$$

For (3.27), note that as $n \longrightarrow \infty$,

$$P(X_{(1)} \le M_n(1+\varepsilon_n)) = \left(1 - P(H>M_n)P(L>\varepsilon_n M_n)\right)^n \longrightarrow 1,$$

the convergence following from (2.2) in Assumption A. This shows (3.27). For (3.28), observe that

$$P(X_{(1)} < M_n) \le \left(1 - P(H > M_n)\right)^n .$$

By (2.3), the right hand side converges to zero, and hence (3.28) holds. Clearly, (3.29) will follow if it can be shown that

$$E(T_n) = o\left(n^{-1/2} M_n^{\alpha(1-\beta/2)}\right) ,$$

and

$$\mathrm{Var}(T_n) = o\left(n^{-1} M_n^{\alpha(2-\beta)}\right) .$$

Set

$$p_n := P(\gamma M_n < X_1 \le \gamma(1+\varepsilon_n) M_n) ,$$

and note that,

$$\begin{aligned} n M_n^{-\alpha(2-\beta)} \mathrm{Var}(T_n) &\le n^2 M_n^{-\alpha(2-\beta)} p_n \\ &\ll n^{3/2} M_n^{-\alpha(1-\beta/2)} p_n = n^{1/2} M_n^{-\alpha(1-\beta/2)} E(T_n) \\ &\ll p_n M_n^{\alpha(2-\beta)} , \end{aligned}$$

the inequalities in the last two lines following by (2.3). Thus, for (3.29), it suffices to show that

$$p_n = o(P(H > M_n)^{2-\beta}) . \tag{3.30}$$

For n large enough so that $\gamma(1+\varepsilon_n) < 1$, notice that

$$\begin{aligned} p_n &= P(H > \gamma M_n) - \gamma^{-\alpha} M_n^{-\alpha} (1+\varepsilon_n)^{-\alpha} l\left(\gamma M_n (1+\varepsilon_n)\right) \\ &= \gamma^{-\alpha} M_n^{-\alpha} l\left(\gamma M_n (1+\varepsilon_n)\right) \left\{1 - (1+\varepsilon_n)^{-\alpha}\right\} \\ &\quad + P(H > \gamma M_n) \left\{1 - \frac{l\left(\gamma M_n (1+\varepsilon_n)\right)}{l(\gamma M_n)}\right\} , \end{aligned}$$

where

$$l(x) := x^\alpha P(H > x) .$$

The first term on the right hand side is clearly $O(\varepsilon_n P(H > M_n))$, which by (2.1), is $o\left(P(H > M_n)^{2-\beta}\right)$. For the second term, notice that by Assumption B,

$$\begin{aligned} \frac{l\left(\gamma M_n (1+\varepsilon_n)\right)}{l(\gamma M_n)} - 1 &= O\left(M_n^{\alpha-\tilde{\alpha}} \left\{(1+\varepsilon_n)^{\alpha-\tilde{\alpha}} - 1\right\}\right) \\ &= O(\varepsilon_n) \\ &= o(P(H > M_n)^{1-\beta}) , \end{aligned}$$

the last step following from (2.1). This shows (3.30), and thus completes the proof of (3.23).

Next, we show that for all $\eta \in \mathbb{R}$, as $n \longrightarrow \infty$,

$$\sqrt{\tilde{k}_n} \log \frac{X_{(n,[\tilde{k}_n + \eta \tilde{k}_n^{1/2}])}}{X_{(n,\tilde{k}_n)}} \xrightarrow{P} -\frac{\eta}{\alpha} . \tag{3.31}$$

Let (u_n) be a sequence of positive integers satisfying (3.7) For n large enough so that $1 \le u_n < [n^{1-\beta}u_n^\beta] \le n$ and $1 \le u_n < [n^{1-\beta}u_n^\beta] + \eta[n^{1-\beta}u_n^\beta]^{1/2} \le n$, the conditional distribution of $\left(X_{(\tilde{k}_n)}, X_{([\tilde{k}_n+\eta\tilde{k}_n^{1/2}])}\right)$ given that $U_n = u_n$ is same as the (unconditional) distribution of

$$\left(Y_{(n-u_n,[n^{1-\beta}u_n^\beta]-u_n)}, Y_{(n-u_n,[n^{1-\beta}u_n^\beta]+\eta[n^{1-\beta}u_n^\beta]^{1/2}-u_n)}\right),$$

where $\{Y_{(n,j)} : 1 \le j \le n\}$ is as defined in Lemma 3.3, with $\tilde{M}_n$ as in (3.9). Define v_n as in (3.8) By Lemma 3.3, it follows that

$$\sqrt{v_n}\left(\frac{1}{v_n}\sum_{i=1}^n \delta_{Y_{n-u_n,i}/b((n-u_n)/v_n)}(y^{-1/\alpha},\infty] - y\right) \Longrightarrow W(y)$$

in $D[0,\infty)$. Using Vervaat's lemma, it follows that

$$\sqrt{v_n}\left[\left(\frac{Y_{(n-u_n,[v_n x])}}{b((n-u_n)/v_n)}\right)^{-\alpha} - x\right] \Longrightarrow -W(x) \tag{3.32}$$

in $D[0,\infty)$. From here, we conclude that

$$\left(\sqrt{v_n}\left[\left(\frac{Y_{(n-u_n,[v_n s_n])}}{b((n-u_n)/v_n)}\right)^{-\alpha} - s_n\right], \sqrt{v_n}\left[\left(\frac{Y_{(n-u_n,v_n)}}{b((n-u_n)/v_n)}\right)^{-\alpha} - 1\right]\right)$$
$$\Longrightarrow (-W(1), -W(1)),$$

where $s_n := 1 + \eta v_n^{-1}[n^{1-\beta}u_n^\beta]^{1/2}$. Since the limit process is $C[0,\infty) \times C[0,\infty)$ valued, this can be done using Skorohod's Theorem (Theorem 2.2.2 in [1]). Using the Delta method with $x \mapsto -\frac{1}{\alpha}\log x$, it follows that

$$\left(\sqrt{v_n}\left\{\log\frac{Y_{(n-u_n,[v_n s_n])}}{b((n-u_n)/v_n)} + \frac{1}{\alpha}\log s_n\right\}, \sqrt{v_n}\log\frac{Y_{(n-u_n,v_n)}}{b((n-u_n)/v_n)}\right)$$
$$\Longrightarrow \left(\frac{1}{\alpha}W(1), \frac{1}{\alpha}W(1)\right).$$

Since,

$$\lim_{n\to\infty}\sqrt{v_n}\log s_n = \eta,$$

it follows that

$$\sqrt{v_n}\log\frac{Y_{(n-u_n,[n^{1-\beta}u_n^\beta]-u_n)}}{Y_{(n-u_n,[n^{1-\beta}u_n^\beta]+\eta[n^{1-\beta}u_n^\beta]^{1/2}-u_n)}} \xrightarrow{P} -\frac{\eta}{\alpha}.$$

What we have shown is that whenever (u_n) is a sequence satisfying (3.7), the conditional distribution of the left hand side of (3.31) given $U_n = u_n$ converges weakly to $-\eta/\alpha$. By an appeal to Lemma 3.1, this shows (3.31).

Coming to the proof of (3.22), note that

$$\begin{aligned}
&\sqrt{\tilde{k}_n}\left[h(\hat{k}_n, n) - h(\tilde{k}_n, n)\right] \\
= &\frac{1}{\sqrt{\tilde{k}_n}}\left[\sum_{i=1}^{\hat{k}_n}\log\frac{X_{(i)}}{X_{(\hat{k}_n)}} - \sum_{i=1}^{\tilde{k}_n}\log\frac{X_{(i)}}{X_{(\tilde{k}_n)}}\right] + \frac{\hat{k}_n}{\sqrt{\tilde{k}_n}}\log\frac{X_{(\tilde{k}_n)}}{X_{(\hat{k}_n)}} \\
&+\sqrt{\tilde{k}_n}\left(\frac{1}{\hat{k}_n} - \frac{1}{\tilde{k}_n}\right)\sum_{i=1}^{\hat{k}_n}\log\frac{X_{(i)}}{X_{(\hat{k}_n)}} \\
=: \ &A + B + C\,.
\end{aligned}$$

Clearly,

$$C = \sqrt{\tilde{k}_n}\left(1 - \frac{\hat{k}_n}{\tilde{k}_n}\right)h(\hat{k}_n, n) \xrightarrow{P} 0\,,$$

the convergence in probability following from (3.23) and the fact that

$$h(\hat{k}_n, n) \xrightarrow{P} 1/\alpha\,,$$

which has been shown in [2]. For showing that $B \xrightarrow{P} 0$, fix $\epsilon > 0$ and let $\eta := \epsilon\alpha/6$. Note that

$$\begin{aligned}
&P(|B| > \epsilon) \\
\le P\left[\frac{\hat{k}_n}{\tilde{k}_n} > 2\right] + P\left[\sqrt{\tilde{k}_n}\left|\frac{\hat{k}_n}{\tilde{k}_n} - 1\right| > \eta\right] + P\left[\sqrt{\tilde{k}_n}\log\frac{X_{(\tilde{k}_n - \eta\tilde{k}_n^{1/2})}}{X_{(\tilde{k}_n + \eta\tilde{k}_n^{1/2})}} > 3\frac{\eta}{\alpha}\right]\,.
\end{aligned}$$

By (3.23) and (3.31), it follows that $B \xrightarrow{P} 0$. Since for $0 < \epsilon < 1$,

$$P(|A| > \epsilon) \le P\left[\sqrt{\tilde{k}_n}\left|\frac{\hat{k}_n}{\tilde{k}_n} - 1\right| > \epsilon\right] + P\left[\log\frac{X_{(\tilde{k}_n - \tilde{k}_n^{1/2})}}{X_{(\tilde{k}_n + \tilde{k}_n^{1/2})}} > 1\right]\,,$$

it is immediate that $A \xrightarrow{P} 0$. This completes the proof. □

Proof of Theorem 2.1. In view of Lemma 3.5, it suffices to show that

$$\sqrt{\tilde{k}_n}\left(h(\tilde{k}_n, n) - \frac{1}{\alpha}\right) \Longrightarrow N\left(0, \frac{1}{\alpha^2}\right)\,. \tag{3.33}$$

Define

$$\begin{aligned}
S_1 &:= \sum_{i=1}^{U_n}\log\frac{X_{(i)}}{X_{(\tilde{k}_n)}} \\
S_2 &:= \sum_{i=U_n+1}^{\tilde{k}_n}\log\frac{X_{(i)}}{X_{(\tilde{k}_n)}}
\end{aligned}$$

and note that on the set $\{U_n \le \tilde{k}_n\}$,

$$h(\tilde{k}_n, n) = \frac{1}{\tilde{k}_n}(S_1 + S_2).$$

Let u_n be a sequence of integers satisfying (3.7) and define v_n and $\tilde{M}_n$ as in (3.8) and (3.9). For n large enough, note that

$$[S_2 | U_n = u_n] \stackrel{d}{=} \sum_{i=1}^{v_n} \log \frac{Y_{(n-u_n,i)}}{Y_{(n-u_n,v_n)}} =: \tilde{S}_2,$$

where $\{Y_{(n,j)} : 1 \le j \le n\}$ is as defined in the statement of Lemma 3.4. By Lemma 3.4, it follows that

$$\sqrt{v_n}\left(\frac{1}{v_n}\tilde{S}_2 - \frac{1}{\alpha}\right) \Longrightarrow N\left(0, \frac{1}{\alpha^2}\right).$$

This along with the fact that

$$\sqrt{v_n}\tilde{S}_2\left(\frac{1}{[n^{1-\beta}u_n^\beta]} - \frac{1}{v_n}\right) = -\frac{\tilde{S}_2}{[n^{1-\beta}u_n^\beta]}\frac{u_n}{\sqrt{v_n}} = O_p(1)o(1),$$

shows that

$$\left[\sqrt{\tilde{k}_n}\left(\frac{1}{\tilde{k}_n}S_2 - \frac{1}{\alpha}\right)\middle| U_n = u_n\right] \Longrightarrow N\left(0, \frac{1}{\alpha^2}\right).$$

Since this is true for all sequence of integers (u_n) satisfying (3.7), by Lemma 3.1 it follows that

$$\sqrt{\tilde{k}_n}\left(\frac{1}{\tilde{k}_n}S_2 - \frac{1}{\alpha}\right) \Longrightarrow N\left(0, \frac{1}{\alpha^2}\right).$$

On the set $\{1 \le X_{(1)} \le 2M_n\}$,

$$\begin{aligned} \frac{S_1}{\sqrt{\tilde{k}_n}} &\le \frac{U_n \log(2M_n)}{\sqrt{\tilde{k}_n}} \\ &= O_p\left(n^{1/2}P(H > M_n)^{1-\beta/2}\log M_n\right) \\ &= o_p(1). \end{aligned}$$

Since the probability of that set converges to one, it follows that

$$\frac{S_1}{\sqrt{\tilde{k}_n}} \xrightarrow{P} 0.$$

This completes the proof. □

4 Examples

In this section, we study some specific examples of c.d.f.s for which Theorem 2.1 hold.

Example 1

Suppose that

$$P(H > x) = Cx^{-\alpha},$$

for x large, where $C > 0$. Let at least one of the following hold:

	Condition on β	Condition on L
1.	$0 < \beta < 1$	$L \equiv 0$
2.	$\max(0, 1 - 1/\alpha) < \beta < 1$	$P(L > x) = O\left(x^{-\theta}\right)$ where $\theta := \frac{\alpha(1-\beta)}{1-\alpha(1-\beta)}$

Assume that M_n satisfies

$$M_n^{\alpha} \ll n \ll (\log M_n)^{-2} M_n^{\alpha(2-\beta)}. \tag{4.1}$$

Then (2.4) holds.

To see this, suppose that 1. holds, *i.e.*,

$$0 < \beta < 1$$

and

$$L \equiv 0.$$

Set

$$\bar{\alpha} := \frac{\alpha}{\beta}(1 + \beta/2). \tag{4.2}$$

Thus,

$$\frac{\alpha}{2\bar{\alpha} - \alpha} = \frac{\beta}{2} < \beta < 1,$$

and

$$\beta(2\bar{\alpha} - \alpha) = 2\alpha > \alpha(2 - \beta).$$

Therefore the assumptions of Theorem 2.1 hold.

Now suppose that 2. holds. Notice that by (4.1)

$$n^{1/\theta} M_n^{-(1+\alpha/\theta)} \ll M_n^{\frac{\alpha}{\theta}(2-\beta)-(1+\frac{\alpha}{\theta})} = M_n^{-\alpha(1-\beta)}.$$

Thus, if ε_n is chosen to satisfy

$$n^{1/\theta} M_n^{-(1+\alpha/\theta)} \ll \varepsilon_n \ll M_n^{-\alpha(1-\beta)},$$

then (2.1) and (2.2) hold. If $\bar{\alpha}$ is defined by (4.2), then the premise of Theorem 2.1 is satisfied.

Example 2

Suppose that

$$P(H > x) = Cx^{-\alpha} + x^{-\bar{\alpha}}(1 + o(1) \text{ as } x \to \infty,$$

for some $C > 0$ and $0 < \alpha < \bar{\alpha}$. Let at least one of the following hold:

	Condition on β	Condition on L
1.	$\frac{\alpha}{2\bar{\alpha}-\alpha} < \beta < 1$	$L \equiv 0$
2.	$\max(\frac{\alpha}{2\bar{\alpha}-\alpha}, 1-1/\alpha) < \beta < 1$	$P(L > x) = O\left(x^{-\theta}\right)$ where $\theta := \frac{\min\{\alpha(2-\beta), \beta(2\bar{\alpha}-\alpha)\}-\alpha}{1-\alpha(1-\beta)}$

Assume that M_n satisfies (3.33). Then, by similar calculations as in Example 1, (2.4) holds.

Example 3

Let

$$P(H > x) = \frac{1}{2}x^{-\alpha}\left(1 + x^{-1}\exp\sin\log x\right), \, x \geq 1 \, .$$

As mentioned in Remark 1, this example is significant in that it satisfies Assumption B but not (2.5); see Exercise 2.7, page 61 in [5]. Setting

$$\bar{\alpha} := \alpha + 1 \, ,$$

the conditions are exactly same as those of Example 2. That is, if either 1. or 2. of Example 2 holds and M_n satisfies (3.33), then (2.4) holds.

5 Acknowledgement

The author is immensely grateful to his adviser Gennady Samorodnitsky for many helpful discussions. He also acknowledges the comments of an anonymous referee, that helped improve the presentation significantly.

References

[1] V. S. Borkar. *Probability theory: an advanced course*. Springer-Verlag, New York, 1995. MR1367959

[2] A. Chakrabarty and G. Samorodnitsky. Understanding heavy tails in a bounded world or, is a truncated heavy tail heavy or not? To appear in *Stochastic Models*, preprint available at `http://arxiv.org/pdf/1001.3218`, 2009.

[3] S. Csörgő and D. M. Mason. Central limit theorems for sums of extreme values. *Mathematical Proceedings of the Cambridge Philosophical Society*, 98:547–558, 1985. MR0803614

[4] R. A. Davis and S. I. Resnick. Tail estimates motivated by extreme value theory. *Ann. Statist.*, 12:1467–1487, 1984. MR0760700

[5] L. de Haan and A. Ferreira. *Extreme Value Theory: An Introduction*. Springer, New York, 2006. MR2234156

[6] L. de Haan and S. Resnick. On asymptotic normality of the hill estimator. *Communications in Statistics. Stochastic Models*, 14:849–866, 1998. MR1631534

[7] W. Feller. *An Introduction to Probability Theory and its Applications*, volume 2. Wiley, New York, 2nd edition, 1971. MR0270403

[8] J. L. Geluk, L. de Haan, S. Resnick, and C. Stărică. Second-order regular variation, convolution and the central limit theorem. *Stochastic Processes and their Applications*, 69:139–159, 1997. MR1472948

[9] C. M. Goldie and R. M. Smith. Slow variation with remainder: Theory and applications. *The Quarterly Journal of Mathematics*, 38:45–71, 1987. MR0876263

[10] E. Haeusler and J. L. Teugels. On asymptotic normality of hill's estimator for the exponent of regular variation. *Annals of Statistics*, 13:743–756, 1985. MR0790569

[11] P. Hall. On some simple estimates of an exponent of regular variation. *J. Roy. Stat. Soc.*, 44:37–42, 1982. Series B. MR0655370

[12] B. M. Hill. A simple general approach to inference about the tail of a distribution. *Ann. Statist.*, 3:1163–1174, 1975. MR0378204

[13] S. Hong, I. Rhee, S. Kim, K. Lee, and S. Chong. Routing performance analysis of human-driven delay tolerant networks using the truncated levy walk model. In *Mobility Models '08: Proceedings of the 1st ACM SIGMOBILE workshop on Mobility models*, pages 25–32, New York, 2008. ACM.

[14] D. M. Mason. Laws of large numbers for sums of extreme values. *Annals of Probability*, 10:754–764, 1982. MR0659544

[15] S. Resnick. *Heavy-Tail Phenomena : Probabilistic and Statistical Modeling*. Springer, New York, 2007. MR2271424