CentER

Discussion Paper

No. 2003–08

EDGEWORTH EXPANSIONS FOR THE DISTRIBUTION FUNCTION OF THE HILL ESTIMATOR

By Anke Cuntz, Erich Haeusler, Johan Segers

January 2003

ISSN 0924-7815

TILBURG UNIVERSITY

Edgeworth expansions for the distribution function of the Hill estimator

Anke Cuntz, Erich Haeusler*

University of Giessen

Johan Segers†

Tilburg University & Catholic University Leuven

January 27, 2003

Abstract

We establish Edgeworth expansions for the distribution function of the centered and normalized Hill estimator for the positive extreme value index.

JEL: C 13; C 14

MSC: 62 G 20; 62 G 32

Key words: asymptotic normality, Edgeworth expansions, extreme value index, Hill estimator, regular variation

1 Introduction

Let $(X_i)_{i\geq 1}$ be a sequence of independent and identically distributed random variables with common distribution function F in the max-domain of attraction of a Fréchet extreme value distribution with parameter $\gamma > 0$, or equivalently

$$\lim_{t\to\infty} \frac{1-F(tx)}{1-F(t)} = x^{-1/\gamma} \quad \text{for all } x > 0, \tag{1}$$

that is, the upper tail $1-F$ is regularly varying at infinity with index $-1/\gamma$, notation $1-F \in \mathcal{R}_{-1/\gamma}$. The parameter γ determines the shape of the upper tail of F and is traditionally called the extreme value index.

*Mathematical Institute, University of Giessen, Arndtstrasse 2, D-35392 Giessen, Germany (E-mail: *erich.haeusler@math.uni-giessen.de*)

†Tilburg University, P.O. Box 90153, NL-5000 LE Tilburg, the Netherlands (E-mail: *jsegers@uvt.nl*)

Estimation of γ is often the first step in estimating tail probabilities, high quantiles and other tail quantities. One of the most popular estimators for γ is the Hill (1975) estimator

$$\hat{H}_n(k_n) = \frac{1}{k_n} \sum_{i=1}^{k_n} \log X_{n+1-i:n} - \log X_{n-k_n:n},$$

with $k_n = 1, \ldots, n-1$ and with $X_{1:n} \leq \cdots \leq X_{n:n}$ being the order statistics of $X_1, \ldots, X_n$. Without loss of generality we assume here and in the sequel that $F(0) = 0$. As customary in the literature, k_n will be an intermediate sequence of integers, that is, $k_n \to \infty$ and $k_n/n \to 0$ as $n \to \infty$. Mason (1982) showed that (1) holds if and only if $\hat{H}_n(k_n)$ is a consistent estimator for γ for all intermediate sequences k_n. Complementarily, Segers (2001) proved that, subject to an integrability condition on F ensuring existence of $E[\hat{H}_n(k_n)]$, equation (1) is equivalent to $\lim_{n\to\infty} E[\hat{H}_n(k_n)] = \gamma$ for all intermediate sequences k_n, that is, to $\hat{H}_n(k_n)$ being asymptotically unbiased.

Both characterizations are first-order results in the sense that they involve the mere convergence in (1). Asymptotic normality of the appropriately standardized Hill estimator requires information about the speed of convergence in (1), which is most conveniently discussed in terms of the tail-quantile function $V(y) = \inf\{x \in \mathbb{R} : F(x) \geq 1 - 1/y\}$ for $y > 1$. Theorem 1.5.12 in Bingham *et al.* (1987) shows that (1) is equivalent to $\lim_{t\to\infty} V(ty)/V(t) = y^\gamma$ for all $y > 0$, that is, $V \in \mathcal{R}_\gamma$. The natural second-order refinement of this condition is that there exist $C \in \mathbb{R}$, $\alpha \geq 0$, and some positive function $a \in \mathcal{R}_{-\alpha}$ with $\lim_{t\to\infty} a(t) = 0$ such that

$$\lim_{t\to\infty} \frac{1}{a(t)} \left(\frac{V(ty)}{V(t)} - y^\gamma \right) = C y^\gamma A_{-\alpha}(y) \quad \text{for all } y > 0, \tag{2}$$

where $A_{-\alpha}(y) = \int_1^y u^{-\alpha-1}\, du$. Condition (2) is called 'natural' because the mere existence, for all $y > 0$ and some positive measurable function a with $\lim_{t\to\infty} a(t) = 0$, of a limit that is not identically zero implies that $a \in \mathcal{R}_{-\alpha}$ for some $\alpha \geq 0$ as well as the analytic form of the limit function given above; see Theorem 1.9 in Geluk and de Haan (1987). Alternatively, the existence of the limit in (2) for all y in a subset of $(0, \infty)$ of positive Lebesgue measure together with the assumption that $\lim_{t\to\infty} a(t) = 0$ and $a \in \mathcal{R}_{-\alpha}$ for some $\alpha \geq 0$ also entails the given analytic form of the limit function; see Bingham *et al.* (1987), Lemma 3.2.1. For a thorough discussion on second-order conditions for V and the bias of the Hill estimator the reader is referred to Segers (2001).

Under (2) the following central limit theorem for the Hill estimator is well known; see e.g. de Haan and Peng (1998) or Haeusler and Teugels (1985): If $\lim_{n\to\infty} k_n^{1/2} a(n/k_n) = \lambda \in [0, \infty]$, then we have as $n \to \infty$

$$H_n = k_n^{1/2}\{\hat{H}_n(k_n) - \gamma\}/\gamma \begin{cases} \xrightarrow{\mathcal{D}} N(0,1) & \text{if } \lambda = 0, \\ \xrightarrow{\mathcal{D}} N\left(\{\gamma(1+\alpha)\}^{-1} C\lambda, 1\right) & \text{if } 0 < \lambda < \infty, \\ \text{has no limit law} & \text{if } \lambda = \infty, \end{cases} \tag{3}$$

where $\xrightarrow{\mathcal{D}}$ stands for convergence in distribution and $N(\mu, \sigma^2)$ denotes the normal distribution with mean μ and variance σ^2. The case $0 < \lambda < \infty$ in (3) is in some sense the more important one, because it leads to sequences k_n for which the asymptotic mean squared error of $\hat{H}_n(k_n)$ is minimal, see e.g. de Haan and Peng (1998).

The accuracy of approximations to the distribution function of H_n has been studied recently by Cheng and Pan (1998), Cheng and de Haan (2001), Guillou and Hall (2001) and Ferreira (2002). The aim of the present paper is to contribute to these studies by providing Edgeworth expansions for $P(H_n \leq x)$ in the first two cases of (3). Our two main results, stated in Section 2, correspond to these two cases and are proven in Sections 3 and 4 respectively. The Appendix contains some additional technical arguments.

2 Main results

Approximations of the distribution of H_n typically feature standardized sums of independent standard exponential random variables, and indeed our first result, for the case $\lambda = 0$ in (3), features the classical Edgeworth expansion for such sums. Formally, let $(E_i)_{i \geq 1}$ be a sequence of independent random variables, exponentially distributed with mean one. Denote the distribution function and density of the standard normal distribution by Φ and φ respectively. There exist polynomials $P_\ell(x)$, $\ell \geq 1$, such that for every integer $m \geq 1$ we have uniformly in $x \in \mathbb{R}$

$$P\left\{ n^{-1/2} \sum_{j=1}^{n} (E_j - 1) \leq x \right\} = \Phi(x) + \sum_{\ell=1}^{m} P_\ell(x) \varphi(x) n^{-\ell/2} + O(n^{-(m+1)/2}), \tag{4}$$

as $n \to \infty$, see e.g. Petrov (1975), Theorem VI.4; the polynomials P_ℓ can be computed explicitly from the moments of the exponential distribution.

Theorem 1 *Under the second-order condition (2), if* $\lim_{n\to\infty} k_n^{1/2} a(n/k_n) = 0$*, then for any integer* $m \geq 1$ *we have uniformly in* $x \in \mathbb{R}$

$$\begin{aligned} P(H_n \leq x) &= \Phi(x) + \sum_{\ell=1}^{m} P_\ell(x) \varphi(x) k_n^{-\ell/2} + O(k_n^{-(m+1)/2}) \\ &\quad - \frac{C}{\gamma(1+\alpha)} \varphi(x) k_n^{1/2} a(n/k_n) + o(k_n^{1/2} a(n/k_n)) \quad \textit{as } n \to \infty. \end{aligned}$$

The summand $\{\gamma(1+\alpha)\}^{-1} C \varphi(x) k_n^{1/2} a(n/k_n)$ in Theorem 1 reflects the influence of the bias of $\hat{H}_n(k_n)$ on the asymptotic expansion of $P(H_n \leq x)$, since by Theorem 7.1 of Segers (2001), under a suitable integrability condition on F, assumption (2) implies that $E[\hat{H}_n(k_n)] - \gamma \sim (1+\alpha)^{-1} C a(n/k_n)$ and thus $E(H_n) \sim \{\gamma(1+\alpha)\}^{-1} C k_n^{1/2} a(n/k_n)$ as

$n \to \infty$. Clearly, the order of the bias determines which m leads to a meaningful expansion; in particular, if $k_n^{-1/2} = o(k_n^{1/2} a(n/k_n))$, then there is no such m at all and the expansion reduces to $P(H_n \le x) = \Phi(x) - \{\gamma(1+\alpha)\}^{-1} C\varphi(x) k_n^{1/2} a(n/k_n) + o(k_n^{1/2} a(n/k_n))$ as $n \to \infty$.

The special case $m = 1$ of Theorem 1 has been proven by Cheng and Pan (1998) in their Theorem 1 under the assumption $k_n^{1/2} a(n/k_n) = \{\rho + o(1)\} k_n^{-1/2}$ as $n \to \infty$ for some $\rho \in [0, \infty)$, leading to a one-term Edgeworth expansion with a $O(k_n^{-1/2})$ correction term and a $o(k_n^{-1/2})$ remainder term. The approximations to $P(H_n \le x)$ in Theorem 1 of Cheng and de Haan (2001) and in Theorem 1 of Guillou and Hall (2001) involve versions of gamma distributions depending on k_n instead of the limiting normal distribution. These approximations are stated under extra growth conditions on k_n and in Guillou and Hall (2001) for a sub-model of (2).

In order to obtain an Edgeworth expansion in case $0 < \lambda < \infty$ in (3), we have to strengthen the second-order condition (2). Observe first that (2) is equivalent to

$$\log V(ty) - \log V(t) = \gamma \log y + CA_{-\alpha}(y) a(t) + o(a(t)) \quad \text{as } t \to \infty \quad \text{for all } y > 0. \tag{5}$$

The appropriate third-order refinement corresponding to (5) is

$$\lim_{t\to\infty} \left(\frac{\log V(ty) - \log V(t) - \gamma \log y}{a(t)} - CA_{-\alpha}(y) \right) \Big/ b(t) = B(y) \quad \text{for all } y > 0 \tag{6}$$

for some positive function $b \in \mathcal{R}_{-\beta}$ ($\beta \ge 0$) satisfying $\lim_{t\to\infty} b(t) = 0$ and for which

$$\lim_{t\to\infty} \frac{1}{b(t)} \left(\frac{a(ty)}{a(t)} - y^{-\alpha} \right) = Dy^{-\alpha} A_{-\beta}(y) \quad \text{for all } y > 0 \tag{7}$$

for some $D \in \mathbb{R}$; the limit function B is required to be of the form

$$B(y) = \begin{cases} C_1 (\log y)^2 + C_2 \log y & \text{if } \alpha = \beta = 0, \\ C_1 y^{-\alpha} \log y + C_2 A_{-\alpha}(y) & \text{if } \alpha > 0 = \beta, \\ C_1 A_{-\alpha-\beta}(y) + C_2 A_{-\alpha}(y) & \text{if } \beta > 0, \end{cases} \tag{8}$$

for some $C_1, C_2 \in \mathbb{R}$. Again, conditions (6)–(8) appear to be much more stringent than they in fact are: If the limit in (6) exists for all $y > 0$ with $C \ne 0$ and some positive measurable function b satisfying $\lim_{t\to\infty} b(t) = 0$, then (6)–(8) are implied by either of the two assumptions that (i) the limit B in (6) is not a multiple of $A_{-\alpha}$ (Theorem 1 in de Haan and Stadtmüller, 1996) or (ii) $b \in \mathcal{R}_{-\beta}$ for some $\beta \ge 0$ (personal communication by Jef Teugels, 2001). As pointed out by de Haan and Stadtmüller (1996), (6) holding with B being a multiple of $A_{-\alpha}$ is a trivial case in the sense that then (6) also holds for some function $\bar{a}(t) \sim a(t)$ as $t \to \infty$ and $\bar{B}(y) = 0$ for all $0 < y < \infty$. According to (ii) we need not exclude this case from our considerations provided that we require $b \in \mathcal{R}_{-\beta}$ for some $\beta \ge 0$ from the outset.

Denote $\lambda_n = k_n^{1/2} a(n/k_n)$, $\mu_n = \{\gamma(1+\alpha)\}^{-1} C \lambda_n$, and $B_0 = \int_0^1 B(1/u)\, du$. Further, let $V_{n,k_n} = (n/k_n) U_{k_n+1:n}$, with $U_{j:n}$ the jth ascending order statistic from n independent Uniform$(0,1)$ random variables. For integer $n \geq 1$ we have $(d^n/dx^n)\Phi(x) = \widetilde{H}_{n-1}(x)\varphi(x)$, where $(-1)^{n-1}\widetilde{H}_{n-1}(x)$ is the monic Chebyshev-Hermite polynomial of degree $n-1$.

Theorem 2 *Under the third-order conditions (6)–(8), if $\lambda_n = O(1)$, then there are polynomials $P_{\ell,j}$ and $P_{\ell,j,s}$ such that $P(H_n \leq x)$ for every integer $m \geq 1$ admits uniformly in $x \in \mathbb{R}$ the expansion*

$$\begin{aligned}
&\Phi(x-\mu_n) + \varphi(x-\mu_n) \sum_{s=1}^{2m+1} \widetilde{H}_{s-1}(x-\mu_n) \frac{\mu_n^s}{s!} E\left[\left(1 - V_{n,k_n}^{\alpha}\right)^s\right] \\
&+ \varphi(x-\mu_n) \sum_{\substack{0 \leq \ell, j \leq m \\ 1 \leq \ell + j \leq m}} P_{\ell,j}(x-\mu_n) k_n^{-(\ell+j)/2} \lambda_n^j E\left[V_{n,k_n}^{j\alpha}\right] \\
&+ \varphi(x-\mu_n) \sum_{\substack{0 \leq \ell, j \leq m \\ 1 \leq \ell + j \leq m}} k_n^{-(\ell+j)/2} \lambda_n^j \sum_{s=1}^{2m+1} \frac{\mu_n^s}{s!} P_{\ell,j,s}(x-\mu_n) E\left[\left(1 - V_{n,k_n}^{\alpha}\right)^s V_{n,k_n}^{j\alpha}\right] \\
&- \frac{B_0 \lambda_n}{\gamma} \varphi(x-\mu_n) b(n/k_n) + O\left(k_n^{-(m+1)/2}\right) + o(b(n/k_n)) \quad \textit{as } n \to \infty.
\end{aligned}$$

The expectations can be calculated using the formula

$$E[U_{k+1:n}^v] = \frac{\Gamma(n+1)\Gamma(v+k+1)}{\Gamma(k+1)\Gamma(v+n+1)} = \prod_{j=k+1}^{n} \frac{j}{v+j} \tag{9}$$

for $v \geq 0$ and integer $1 \leq k < n$, while the polynomials $P_{\ell,j}$ and $P_{\ell,j,s}$ can be reconstructed from the proof of Theorem 2 in Section 4. The expansion in Theorem 2 does no longer originate from standardized sums of standard exponential random variables but from the row sums of a triangular array of appropriately modified exponentials.

Approximations to the distribution function of H_n in the second case of (3) have been considered by Ferreira (2002) for the special case when the sequence $(k_n)_n$ is optimally chosen, that is, such that the asymptotic mean squared error of the Hill estimator is minimal. She established one-term expansions when the approximating distribution is a standardized gamma distribution (Theorem 4.B.1) or a shifted normal distribution (Corollary 4.B.1).

Example 1: Hall class. Hall (1982) studied asymptotic normality of the Hill estimator for distribution functions with upper tail

$$1 - F(x) = Ax^{-1/\gamma}\{1 + Bx^{-\tau} + o(x^{-\tau})\} \quad \text{as } x \to \infty,$$

with $\tau, A \in (0,\infty)$ and $B \in \mathbb{R}$. For these F, condition (2) is satisfied with $\alpha = \gamma\tau > 0$ and $a(t) = t^{-\gamma\tau}$. The third-order conditions (6)–(8) are satisfied under the stronger assumption

$$1 - F(x) = Ax^{-1/\gamma}\{1 + Bx^{-\tau} + Cx^{-\tau-\delta} + o(x^{-\tau-\delta})\} \quad \text{as } x \to \infty, \tag{10}$$

with $\delta \in (0,\infty)$ and $C \in \mathbb{R}$, in which case α and a are as above while $\beta = \gamma\delta > 0$ and $b(t) = t^{-\gamma\delta}$. Concrete distributions in the restricted Hall class (10) are the Fréchet, Burr, F and Student's t distributions.

Example 2: log-gamma distributions. Let X be a random variable so that $\log X$ has a $\Gamma(1/\gamma,\delta)$-distribution, i.e. with probability density $g(x) = \{\gamma^\delta \Gamma(\delta)\}^{-1} x^{\delta-1}\exp(-x/\gamma)$ for $x > 0$. The distribution function F of X satisfies

$$1 - F(x) = Ax^{-1/\gamma}(\log x)^\delta \left\{1 + B(\log x)^{-1} + C(\log x)^{-2} + O((\log x)^{-3})\right\} \quad \text{as } x \to \infty$$

with constants $A > 0$, B and C depending only on γ and δ. It can be shown that (6)–(8) are satisfied here with $\alpha = \beta = 0$, $a(t) = (\log t)^{-1}\{1 + C_1 \log\log t(\log t)^{-1} + C_2(\log t)^{-1}\}$ and $b(t) = (\log t)^{-1}$ so that Theorem 2 is again applicable.

3 Proof of Theorem 1

The proof is a refinement of the proof of Theorem 1 in Cheng and Pan (1998), so we borrow most of our notation from them. Let $(U_i)_{i\geq 1}$ be a sequence of independent Uniform$(0,1)$ random variables, and for every positive integer n, let $U_{1:n} \leq \cdots \leq U_{n:n}$ denote the order statistics of $U_1,\ldots,U_n$. Set $a_n = 1 - k_n/(n+1)$ and $b_n = \{a_n(1-a_n)/(n+1)\}^{1/2} = (k_n a_n)^{1/2}/(n+1)$ for every $n \geq 2$, and in addition denote for all $-a_n/b_n < u < (1-a_n)/b_n$

$$\begin{aligned}
(u)_n &= a_n + b_n u, \quad (\overline{u})_n = \{1 - (u)_n\}^{-1} \quad \text{and} \quad W_i = (1-U_i)^{-1} \\
Y_{n,i}(u) &= \log V((\overline{u})_n W_i) - \log V((\overline{u})_n) \quad \text{for all } i = 1,\ldots,k_n, \\
H_n(u) &= k_n^{-1/2}\gamma^{-1}\sum_{i=1}^{k_n}\{Y_{n,i}(u) - \gamma\}.
\end{aligned}$$

Finally, let ϕ_n denote the density of $(U_{n-k_n:n} - a_n)/b_n$.

The following representation of the distribution function of H_n is crucial; see (2.8) in Cheng and Pan (1998). For all $n \geq 2$ and $x \in \mathbb{R}$ we have

$$P(H_n \leq x) = \int_{-a_n/b_n}^{(1-a_n)/b_n} P\{H_n(u) \leq x\}\phi_n(u)\,du.$$

Set $T_n = k_n^{1/5}$, which is $o(k_n^{1/2})$, as required in the proof in Cheng and Pan (1998). Then for all $x \in \mathbb{R}$ we have

$$P(H_n \leq x) = \left(\int_{-T_n}^{T_n} + \int_{|u|>T_n}\right) P\{H_n(u) \leq x\}\phi_n(u)\,du = I_n + II_n,$$

say. The contribution from II_n is negligible, since for all $\tau > 0$ we have

$$\begin{aligned}
II_n &\leq \int_{|u|>T_n} \phi_n(u)\,du \\
&= P(U_{n-k_n:n} > a_n + b_n T_n) + P(U_{n-k_n:n} < a_n - b_n T_n) = O(k_n^{-\tau}) \qquad (11)
\end{aligned}$$

by a straightforward application of standard bounds on binomial tail probabilities; see e.g. Shorack and Wellner (1986), p. 440.

For the analysis of the main term I_n we rewrite (5) as $\log V(ty) - \log V(t) = \gamma \log y + R(y,t)$ so that $Y_{n,i}(u) = \gamma \log W_i + R_{n,i}(u)$ with $R_{n,i}(u) = R(W_i, (\overline{u})_n)$. Now we can write

$$H_n(u) = S_n + R_n(u) + r_n(u) \tag{12}$$

with

$$\begin{aligned} S_n &= k_n^{-1/2} \sum_{i=1}^{k_n} (\log W_i - 1), \\ R_n(u) &= \gamma^{-1} k_n^{-1/2} \sum_{i=1}^{k_n} \{R_{n,i}(u) - E[R_{n,i}(u)]\}, \\ r_n(u) &= \gamma^{-1} k_n^{-1/2} \sum_{i=1}^{k_n} E[R_{n,i}(u)] = \gamma^{-1} k_n^{1/2} E[R_{n,1}(u)]. \end{aligned}$$

Set $c_n = k_n^{3/8} a(n/k_n)$, which is $o(k_n^{1/2} a(n/k_n))$ since $k_n \to \infty$. For $x \in \mathbb{R}$ and $u \in [-T_n, T_n]$ set $x_{n,\pm}(u) = x - r_n(u) \pm c_n$. From (12) we get

$$\begin{aligned} P\{S_n \le x_{n,-}(u)\} - P\{|R_n(u)| \ge c_n\} &\le P\{H_n(u) \le x\} \\ &\le P\{S_n \le x_{n,+}(u)\} + P\{|R_n(u)| \ge c_n\}. \end{aligned} \tag{13}$$

Our first aim is to show that $P\{|R_n(u)| \ge c_n\}$ is negligible. Observe that Lemma 2.3 in Cheng and Pan (1998) is formulated and proven for any $T_n = o(k_n^{1/2})$. Consequently, all parts of their proof which do not involve their condition (1.4) on the k_n-sequence apply to our $T_n = k_n^{1/5}$. Thus, by (2.3) in Cheng and Pan (1998), for all $\varepsilon > 0$ there exists $K_\varepsilon \in (0, \infty)$ such that for all large n we have

$$\left| \frac{R_{n,1}(u)}{a((\overline{u})_n)} \right| \le K_\varepsilon (1 - U_1)^{-\varepsilon} = K_\varepsilon W_1^\varepsilon \quad \text{for all } u \in [-T_n, T_n].$$

Moreover, the uniform convergence theorem for the regularly varying function a together with $(\overline{u})_n \sim n/k_n$ uniformly in $u \in [-T_n, T_n]$ gives

$$a((\overline{u})_n) \sim a(n/k_n) \quad \text{uniformly in } u \in [-T_n, T_n] \text{ as } n \to \infty. \tag{14}$$

Therefore, using the classical moment bound $E[|\sum_{i=1}^n Z_i|^p] \le M_p n^{p/2} E[|Z_1|^p]$ for independent, identically distributed and mean-zero random variables $Z_1, \dots, Z_n$, which is valid for all $p \ge 2$ with $M_p \in (0, \infty)$ depending only on p and which follows from the Marzinkiewicz-Zygmund inequality and Hölder's inequality, we obtain uniformly in $u \in [-T_n, T_n]$

$$\begin{aligned} P\{|R_n(u)| \ge c_n\} &\le c_n^{-p} E[|R_n(u)|^p] \le 2^p c_n^{-p} M_p \gamma^{-p} E[|R_{n,1}(u)|^p] \\ &\le 2^p M_p \gamma^{-p} K_\varepsilon^p E(W_1^{\varepsilon p}) k_n^{-3p/8} \{1 + o(1)\}. \end{aligned}$$

Now $E(W_1^{\varepsilon p}) < \infty$ as long as $\varepsilon > 0$ is chosen so small that $\varepsilon p < 1$. We obtain that

$$P\{|R_n(u)| \geq c_n\} = O(k_n^{-\tau}) \quad \text{uniformly in } u \in [-T_n, T_n], \text{ for all } \tau > 0. \tag{15}$$

Our next aim is to obtain the precise order of magnitude of $r_n(u)$. From the proof of Lemma 2.3 in Cheng and Pan (1998) we infer that, uniformly in $u \in [-T_n, T_n]$,

$$\lim_{n\to\infty} E\left(\frac{R_{n,1}(u)}{a((\overline{u})_n)}\right) = \frac{C}{1+\alpha};$$

see lines 4–6 on page 722. Hence, by (14) we have, uniformly in $u \in [-T_n, T_n]$,

$$r_n(u) = \gamma^{-1} k_n^{1/2} a((\overline{u})_n) E\left(\frac{R_{n,1}(u)}{a((\overline{u})_n)}\right) = \frac{C}{\gamma(1+\alpha)} k_n^{1/2} a(n/k_n)\{1+o(1)\}. \tag{16}$$

Now we can complete the proof along the lines of the proof in Cheng and Pan (1998). Observe that S_n is a standardized sum of k_n independent exponential random variables with mean one. Hence by (4) we have for each positive integer m, uniformly in $x \in \mathbb{R}$,

$$P(S_n \leq x) = \Phi(x) + \sum_{\ell=1}^{m} P_\ell(x)\varphi(x) k_n^{-\ell/2} + O(k_n^{-(m+1)/2}) \quad \text{as } n \to \infty.$$

Consequently, from the right-hand side inequality of (13) and the bound in (15) we obtain, uniformly in $x \in \mathbb{R}$ and $u \in [-T_n, T_n]$,

$$P\{H_n(u) \leq x\} \leq \Phi(x_{n,+}(u)) + \sum_{\ell=1}^{m} P_\ell(x_{n,+}(u))\varphi(x_{n,+}(u)) k_n^{-\ell/2} + O(k_n^{-(m+1)/2}).$$

Now by (16) we have, for all $x \in \mathbb{R}$ and uniformly in $u \in [-T_n, T_n]$,

$$x_{n,+}(u) = x - r_n(u) + c_n = x - \frac{C}{\gamma(1+\alpha)} k_n^{1/2} a(n/k_n) + o(k_n^{1/2} a(n/k_n))$$

so that a Taylor expansion yields, uniformly in $x \in \mathbb{R}$ and $u \in [-T_n, T_n]$,

$$\begin{aligned} &\Phi(x_{n,+}(u)) + \sum_{\ell=1}^{m} P_\ell(x_{n,+}(u))\varphi(x_{n,+}(u)) k_n^{-\ell/2} \\ &\quad = \Phi(x) - \frac{C}{\gamma(1+\alpha)} \varphi(x) k_n^{1/2} a(n/k_n) + \sum_{\ell=1}^{m} P_\ell(x)\varphi(x) k_n^{-\ell/2} + o(k_n^{1/2} a(n/k_n)), \end{aligned}$$

since the derivatives of $P_\ell(x)\varphi(x)$ for $1 \leq \ell \leq m$ are all bounded in $x \in \mathbb{R}$. This gives, uniformly in $x \in \mathbb{R}$ and $u \in [-T_n, T_n]$,

$$\begin{aligned} P\{H_n(u) \leq x\} \;\leq\; & \Phi(x) - \frac{C}{\gamma(1+\alpha)} \varphi(x) k_n^{1/2} a(n/k_n) + \sum_{\ell=1}^{m} P_\ell(x)\varphi(x) k_n^{-\ell/2} \\ & + o(k_n^{1/2} a(n/k_n)) + O(k_n^{-(m+1)/2}). \end{aligned} \tag{17}$$

Exploiting the left-hand inequality in (13) in a similar way, we see that (17) also holds with '$\leq$' replaced by '$\geq$' and hence with '$=$'. Together we have, uniformly in $x \in \mathbb{R}$,

$$\begin{aligned} I_n &= \int_{-T_n}^{T_n} \Bigg[\Phi(x) - \frac{C}{\gamma(1+\alpha)}\varphi(x)k_n^{1/2}a(n/k_n) + \sum_{\ell=1}^{m} P_\ell(x)\varphi(x)k_n^{-\ell/2} \\ &\qquad + o(k_n^{1/2}a(n/k_n)) + O(k_n^{-(m+1)/2})\Bigg]\phi_n(u)\,du \\ &= \left[\Phi(x) - \frac{C}{\gamma(1+\alpha)}\varphi(x)k_n^{1/2}a(n/k_n) + \sum_{\ell=1}^{m} P_\ell(x)\varphi(x)k_n^{-\ell/2}\right]\int_{-T_n}^{T_n}\phi_n(u)\,du \\ &\quad + o(k_n^{1/2}a(n/k_n)) + O(k_n^{-(m+1)/2}) \\ &= \Phi(x) - \frac{C}{\gamma(1+\alpha)}\varphi(x)k_n^{1/2}a(n/k_n) + \sum_{\ell=1}^{m} P_\ell(x)\varphi(x)k_n^{-\ell/2} \\ &\quad + o(k_n^{1/2}a(n/k_n)) + O(k_n^{-(m+1)/2}), \end{aligned}$$

where the last equality follows from (11) and the boundedness of the expansion. This concludes the proof of Theorem 1.

4 Proof of Theorem 2

The expansion of Theorem 2 will be derived from an Edgeworth expansion for sums of independent and identically distributed random variables which depend on an additional parameter as in the following Lemma.

Lemma 1 *Let $(\xi_i, \eta_i)_{i\geq 1}$ be a sequence of independent and identically distributed bivariate random vectors. Assume that for some integer $m \geq 1$ the $(m+3)$th moments of ξ_1 and η_1 are finite, as well as $E(\xi_1) = E(\eta_1) = 0$ and $E(\xi_1^2) = 1$. If there exists $\varepsilon_0 > 0$ such that*

$$\limsup_{|t|\to\infty} \sup_{0<\varepsilon\leq\varepsilon_0} |E[\exp\{it(\xi_1 + \varepsilon\eta_1)\}]| < 1 \tag{18}$$

then for all integer $p \geq 1$ there exist $0 < \varepsilon^* \leq \varepsilon_0$, integer $k^* \geq 1$ and a constant $C^* > 0$ such that

$$\begin{aligned} &\left| P\left\{k^{-1/2}\sum_{i=1}^{k}(\xi_i + \varepsilon\eta_i) \leq x\right\} - \Phi(x) - \varphi(x) \sum_{\substack{0\leq\ell\leq m,\, 0\leq j\leq p \\ (\ell,j)\neq(0,0)}} P_{\ell,j}(x)k^{-\ell/2}\varepsilon^j \right| \\ &\leq C^*\left(k^{-(m+1)/2} + \varepsilon^{p+1}\right) \quad \text{for all } x \in \mathbb{R},\ k \geq k^* \text{ and } 0 < \varepsilon \leq \varepsilon^*, \end{aligned}$$

where the $P_{\ell,j}(x)$ are polynomials in x depending only on the (mixed) moments of (ξ_1, η_1).

Lemma 1 can be derived from classical Edgeworth expansions for sums of independent and identically distributed random variables and is proven in Appendix A.1.

Moreover, a Potter bound pertaining to the third-order condition (5) will be crucial.

Lemma 2 *Under (6)–(8), for every $\varepsilon > 0$ there exist $K_\varepsilon > 0$ and $t_\varepsilon > 1$ such that*

$$\left| \frac{1}{b(t)} \left(\frac{\log V(ty) - \log V(t) - \gamma \log y}{a(t)} - CA_{-\alpha}(y) \right) \right| \leq K_\varepsilon y^\varepsilon \tag{19}$$

for all $1 \leq y < \infty$ and $t \geq t_\varepsilon$.

Although this bound may be derived from Lemma 2.1 in Drees (1998), we give in Appendix A.2 a short, direct proof which, unlike Drees' Lemma, treats the cases $\beta = 0$ and $\beta > 0$ simultaneously.

Throughout the proof of Theorem 2 we use the same notation as in the proof of Theorem 1, and we additionally set $\lambda_n = k_n^{1/2} a(n/k_n)$, which is $O(1)$ by assumption. The proof of Theorem 1 remains unchanged up to and including (11). For the analysis of I_n we now write $\log V(ty) - \log V(t) = \gamma \log y + CA_{-\alpha}(y)a(t) + \overline{R}(y,t)$ with

$$\overline{R}(y,t) = B(y)a(t)b(t) + o(a(t)b(t)) \quad \text{for } 0 < y < \infty \text{ as } t \to \infty \tag{20}$$

from the third-order condition (6). Since for all $u \in [-T_n, T_n]$ we have

$$Y_{n,i}(u) = \log V((\overline{u})_n W_i) - \log V((\overline{u})_n) = \gamma \log W_i + CA_{-\alpha}(W_i)a((\overline{u})_n) + \overline{R}(W_i, (\overline{u})_n),$$

we obtain, denoting $y_n(u) = \{(\overline{u})_n/(n/k_n)\}^{-\alpha}$,

$$\begin{aligned} H_n(u) &= \gamma^{-1} k_n^{-1/2} \sum_{i=1}^{k_n} \left\{ \gamma \log W_i - \gamma + CA_{-\alpha}(W_i) a((\overline{u})_n) + \overline{R}(W_i, (\overline{u})_n) \right\} \\ &= k_n^{-1/2} \sum_{i=1}^{k_n} \left\{ \log W_i - 1 + \frac{C}{\gamma} \left[A_{-\alpha}(W_i) - \frac{1}{1+\alpha} \right] y_n(u) a(n/k_n) \right\} \\ &\quad + \frac{C}{\gamma} \{ a((\overline{u})_n) - y_n(u) a(n/k_n) \} k_n^{-1/2} \sum_{i=1}^{k_n} A_{-\alpha}(W_i) + \frac{C}{\gamma(1+\alpha)} y_n(u) \lambda_n \\ &\quad + \gamma^{-1} k_n^{-1/2} \sum_{i=1}^{k_n} \left\{ \overline{R}(W_i, (\overline{u})_n) - E[\overline{R}(W_i, (\overline{u})_n)] \right\} + \gamma^{-1} k_n^{1/2} E[\overline{R}(W_1, (\overline{u})_n)]. \end{aligned} \tag{21}$$

Claim 1. *There exist $\delta_n = o(b(n/k_n))$ such that for all $\tau > 0$ we have as $n \to \infty$ and uniformly in $u \in [-T_n, T_n]$*

$$P \left\{ \left| a((\overline{u})_n) - y_n(u) a(n/k_n) \right| k_n^{-1/2} \sum_{i=1}^{k_n} A_{-\alpha}(W_i) \geq \delta_n \right\} = O(k_n^{-\tau}). \tag{22}$$

Proof of Claim 1. We have

$$|a((\overline{u})_n) - y_n(u)a(n/k_n)|\, k_n^{-1/2} \sum_{i=1}^{k_n} A_{-\alpha}(W_i) = \lambda_n \left| \frac{a((\overline{u})_n)}{a(n/k_n)} - y_n(u) \right| k_n^{-1} \sum_{i=1}^{k_n} A_{-\alpha}(W_i).$$

Since $E[A_{-\alpha}(W_1)] = (1+\alpha)^{-1}$ and since the moment-generating function of $A_{-\alpha}(W_1)$ is defined in a neighborhood of 0, there exist positive constants c_1 and c_2 such that

$$P\left\{ k_n^{-1} \sum_{i=1}^{k_n} A_{-\alpha}(W_i) \geq \frac{2}{1+\alpha} \right\} \leq c_1 \exp(-c_2 k_n).$$

Moreover, since $(\overline{u})_n \sim n/k_n$ uniformly in $u \in [-T_n, T_n]$ and since (7) holds locally uniformly in $y \in (0, \infty)$, we have as $n \to \infty$

$$\frac{1}{b(n/k_n)} \left| \frac{a((\overline{u})_n)}{a(n/k_n)} - y_n(u) \right| \to |DA_{-\beta}(1)| = 0$$

uniformly in $u \in [-T_n, T_n]$. As $\lambda_n = O(1)$, Claim 1 is proven.

Claim 2. *For every $\tau > 0$ we have, as $n \to \infty$ and uniformly in $u \in [-T_n, T_n]$,*

$$P\left\{ \left| k_n^{-1/2} \sum_{i=1}^{k_n} \left\{ \overline{R}(W_i, (\overline{u})_n) - E[\overline{R}(W_i, (\overline{u})_n)] \right\} \right| \geq k_n^{-1/4} b(n/k_n) \right\} = O(k_n^{-\tau}). \tag{23}$$

Proof of Claim 2. By (19) there exists for every $\varepsilon > 0$ a constant $K_\varepsilon > 0$ such that

$$\frac{|\overline{R}(y,t)|}{a(t)b(t)} \leq K_\varepsilon y^\varepsilon \quad \text{for all } 1 \leq y < \infty \text{ and all large } t. \tag{24}$$

Setting $\overline{c}_n = k_n^{-1/4} b(n/k_n)$ and applying first the inequality $E[|\sum_{i=1}^n Z_i|^p] \leq M_p n^{p/2} E[|Z_1|^p]$ as in the proof of Theorem 1 and second the bound (24), we obtain for all $p \geq 2$, uniformly in $u \in [-T_n, T_n]$ and for all large n

$$\begin{aligned} P\{|\overline{R}_n(u)| \geq \overline{c}_n\} &\leq 2^p M_p \gamma^{-p} \overline{c}_n^{-p} [a((\overline{u})_n) b((\overline{u})_n)]^p E\left[\left| \frac{\overline{R}(W_1, (\overline{u})_n)}{a((\overline{u})_n) b((\overline{u})_n)} \right|^p \right] \\ &\leq 2^p M_p K_\varepsilon^p \gamma^{-p} \overline{c}_n^{-p} E(W_1^{\varepsilon p}) [a(n/k_n) b(n/k_n)]^p \{1 + o(1)\}, \end{aligned}$$

where, to obtain the last inequality, we also used (14) and the corresponding result

$$b((\overline{u})_n) \sim b(n/k_n) \quad \text{uniformly in } u \in [-T_n, T_n] \text{ as } n \to \infty \tag{25}$$

for the regularly varying function b. Now $E(W_1^{\varepsilon p}) < \infty$ as long as $\varepsilon p < 1$, and

$$\overline{c}_n^{-p} [a(n/k_n) b(n/k_n)]^p = k_n^{p/4} a(n/k_n)^p = k_n^{-p/4} \lambda_n^p.$$

Since $\lambda_n = O(1)$ and since p can be chosen arbitrarily large, we obtain the assertion in Claim 2.

Claim 3. *We have, as $n \to \infty$ and uniformly in $u \in [-T_n, T_n]$*

$$k_n^{1/2} E[\overline{R}(W_1, (\overline{u})_n)] \sim B_0 \lambda_n b(n/k_n). \tag{26}$$

Proof of Claim 3. By (20) we have almost surely as $n \to \infty$

$$\sup_{u \in [-T_n, T_n]} \left| \frac{\overline{R}(W_1, (\overline{u})_n)}{a((\overline{u})_n) b((\overline{u})_n)} - B(W_1) \right| \to 0.$$

Apply (24) for some $0 < \varepsilon < 1$ and use the fact that $E[B(W_1)] = B_0$ to see that by dominated convergence

$$\sup_{u \in [-T_n, T_n]} \left| \frac{E[\overline{R}(W_1, (\overline{u})_n)]}{a((\overline{u})_n) b((\overline{u})_n)} - B_0 \right| \le E \left[\sup_{u \in [-T_n, T_n]} \left| \frac{\overline{R}(W_1, (\overline{u})_n)}{a((\overline{u})_n) b((\overline{u})_n)} - B(W_1) \right| \right] \to 0$$

as $n \to \infty$. Thus, using again (14) and (25) we obtain $E[\overline{R}(W_1, (\overline{u})_n)] \sim B_0 a(n/k_n) b(n/k_n)$ uniformly in $u \in [-T_n, T_n]$ as $n \to \infty$, whence the assertion in Claim 3.

Now for $\varepsilon > 0$ and $u \in [-T_n, T_n]$ denote

$$\begin{aligned} S_n(\varepsilon) &= k_n^{-1/2} \sum_{i=1}^{k_n} \left\{ \log W_i - 1 + \varepsilon \frac{C}{\gamma} \left[A_{-\alpha}(W_i) - \frac{1}{1+\alpha} \right] \right\}, \\ \varepsilon_n(u) &= y_n(u) a(n/k_n). \end{aligned}$$

Representation (21) for $H_n(u)$ combined with (22), (23) and (26) leads to

$$H_n(u) = S_n(\varepsilon_n(u)) + \frac{C}{\gamma(1+\alpha)} y_n(u) \lambda_n + \frac{B_0}{\gamma} \lambda_n b(n/k_n) + \rho_n(u),$$

where $\rho_n(u)$ is such that there exist $\delta_n = o(b(n/k_n))$ such that for all $\tau > 0$ we have

$$P\{|\rho_n(u)| \ge \delta_n\} = O(k_n^{-\tau}) \quad \text{uniformly in } u \in [-T_n, T_n] \text{ as } n \to \infty.$$

Denote $c_n(u) = \{\gamma(1+\alpha)\}^{-1} C y_n(u) \lambda_n + \gamma^{-1} B_0 \lambda_n b(n/k_n)$ for $u \in [-T_n, T_n]$, so that $H_n(u) = S_n(\varepsilon_n(u)) + c_n(u) + \rho_n(u)$. Since for all $\tau > 0$ we have

$$P(H_n \le x) = \int_{-T_n}^{T_n} P\{H_n(u) \le x\} \phi_n(u)\, du + O(k_n^{-\tau}) \quad \text{uniformly in } x \in \mathbb{R} \text{ as } n \to \infty,$$

we obtain for all $\tau > 0$

$$\begin{aligned} &\int_{-T_n}^{T_n} P\left\{ S_n(\varepsilon_n(u)) \le x - c_n(u) - \delta_n \right\} \phi_n(u)\, du + O(k_n^{-\tau}) \\ &\le \; P(H_n \le x) \le \int_{-T_n}^{T_n} P\left\{ S_n(\varepsilon_n(u)) \le x - c_n(u) + \delta_n \right\} \phi_n(u)\, du + O(k_n^{-\tau}), \end{aligned} \tag{27}$$

uniformly in $x \in \mathbb{R}$ as $n \to \infty$.

Next, to apply Lemma 1 to $S_n(\varepsilon)$, we need to verify condition (18). Let $\varepsilon_0 > 0$ be such that $\varepsilon_0 |C|/\gamma < 1$ and write $T_\varepsilon(w) = \log(w) + \varepsilon C \gamma^{-1} A_{-\alpha}(w)$ for $\varepsilon \in [-\varepsilon_0, \varepsilon_0]$ and $w \in [1, \infty)$. The derivative $T'_\varepsilon(w) = w^{-1} + \varepsilon C \gamma^{-1} w^{-1-\alpha}$ is positive for all $w \geq 1$ and $\varepsilon \in [-\varepsilon_0, \varepsilon_0]$. Hence by an integration by parts, we have

$$\begin{aligned} & E\left[\exp\left\{it\left(\log W_1 + \varepsilon C\gamma^{-1} A_{-\alpha}(W_1)\right)\right\}\right] \\ & = \int_1^\infty \exp\{itT_\varepsilon(w)\} \frac{dw}{w^2} \\ & = \left. \frac{\exp\{itT_\varepsilon(w)\}}{itw^2 T'_\varepsilon(w)} \right|_1^\infty - \frac{1}{it} \int_1^\infty \exp\{itT_\varepsilon(w)\} \frac{d}{dw}\left(\frac{1}{w^2 T'_\varepsilon(w)}\right) dw \\ & = -\frac{1}{it(1+\varepsilon C/\gamma)} + \frac{1}{it}\int_1^\infty \exp\{itT_\varepsilon(w)\} \frac{1 + \varepsilon C\gamma^{-1}(1-\alpha) w^{-\alpha}}{(w + \varepsilon C \gamma^{-1} w^{1-\alpha})^2} dw, \end{aligned}$$

yielding

$$\sup_{|\varepsilon| \leq \varepsilon_0} \left| E\left[\exp\left\{it\left(\log W_1 + \varepsilon C\gamma^{-1} A_{-\alpha}(W_1)\right)\right\}\right]\right| = O(|t|^{-1}) \quad \text{as } |t| \to \infty. \tag{28}$$

Relation (28) is more than we need to apply Lemma 1 with $\xi_i = \log(W_i) - 1$ and $\eta_i = C\gamma^{-1}\left\{A_{-\alpha}(W_i) - (1+\alpha)^{-1}\right\}$. Since $\varepsilon_n(u) = y_n(u)\lambda_n k_n^{-1/2} = O(k_n^{-1/2})$ uniformly in $u \in [-T_n, T_n]$, we obtain for all integer $m \geq 1$ as $n \to \infty$

$$P\left\{S_n(\varepsilon_n(u)) \leq x\right\} = \Phi(x) + \varphi(x) \sum_{\substack{0 \leq \ell, j \leq m \\ 1 \leq \ell + j \leq m}} P_{\ell,j}(x) k_n^{-(\ell+j)/2} y_n(u)^j \lambda_n^j + O\left(k_n^{-(m+1)/2}\right), \tag{29}$$

uniformly in $x \in \mathbb{R}$ and $u \in [-T_n, T_n]$, for the polynomials $P_{\ell,j}$ of Lemma 1. Since $x^r \varphi(x)$ is bounded in $x \in \mathbb{R}$ for all non-negative integer r, we obtain from (27) and (29) that $P(H_n \leq x)$ is equal to

$$\begin{aligned} & \int_{-T_n}^{T_n} \left\{ \Phi(x - c_n(u)) + \varphi(x - c_n(u)) \sum_{\substack{0 \leq \ell, j \leq m \\ 1 \leq \ell + j \leq m}} P_{\ell,j}(x - c_n(u)) k_n^{-(\ell+j)/2} y_n(u)^j \lambda_n^j \right\} \phi_n(u)\, du \\ & + O\left(k_n^{-(m+1)/2}\right) + o(b(n/k_n)) \quad \text{uniformly in } x \in \mathbb{R} \text{ as } n \to \infty. \end{aligned}$$

Recall that $c_n(u) = d_n(u) + \gamma^{-1} B_0 \lambda_n b(n/k_n)$ where $d_n(u) = \{\gamma(1+\alpha)\}^{-1} C y_n(u) \lambda_n$ for $u \in [-T_n, T_n]$. Hence $P(H_n \leq x)$ admits the expansion

$$\begin{aligned} & \int_{-T_n}^{T_n} \left\{ \Phi(x - d_n(u)) - \frac{B_0 \lambda_n}{\gamma} b(n/k_n) \varphi(x - d_n(u)) \right. \\ & \qquad \left. + \varphi(x - d_n(u)) \sum_{\substack{0 \leq \ell, j \leq m \\ 1 \leq \ell + j \leq m}} P_{\ell,j}(x - d_n(u)) k_n^{-(\ell+j)/2} y_n(u)^j \lambda_n^j \right\} \phi_n(u)\, du \\ & + O\left(k_n^{-(m+1)/2}\right) + o(b(n/k_n)) \quad \text{uniformly in } x \in \mathbb{R} \text{ as } n \to \infty. \end{aligned}$$

Write $d_n(u) = \mu_n - h_n(u)$ where $\mu_n = \{\gamma(1+\alpha)\}^{-1} C\lambda_n$ and $h_n(u) = \{1 - y_n(u)\}\mu_n$ for $u \in [-T_n, T_n]$. Since $y_n(u) = \{(n/k_n)(1 - a_n - b_n u)\}^\alpha \to 1$ uniformly in $u \in [-T_n, T_n]$ as $n \to \infty$, we obtain for $P(H_n \le x)$ the representation

$$\int_{-T_n}^{T_n} \left\{ \Phi(x - d_n(u)) + \varphi(x - d_n(u)) \sum_{\substack{0 \le \ell, j \le m \\ 1 \le \ell + j \le m}} P_{\ell,j}(x - d_n(u)) k_n^{-(\ell+j)/2} y_n(u)^j \lambda_n^j \right\} \phi_n(u)\, du$$
$$- \frac{B_0 \lambda_n}{\gamma} \varphi(x - \mu_n) b(n/k_n) + O\left(k_n^{-(m+1)/2}\right) + o(b(n/k_n)) \tag{30}$$

uniformly in $x \in \mathbb{R}$ as $n \to \infty$.

Since $T_n = k_n^{1/5}$, a closer look at $y_n(u)$ reveals that

$$y_n(u) = (1 - 1/(n+1))^\alpha \left\{1 - (a_n/k_n)^{1/2} u\right\}^\alpha = 1 + O\left(k_n^{-3/10}\right)$$

uniformly in $u \in [-T_n, T_n]$ as $n \to \infty$. Moreover, for f any of the functions Φ or φP with P some polynomial, both f itself and any of its derivatives $f^{(j)}$ are bounded on $\mathbb{R}$, so that for any integer $\nu \ge 1$ we have

$$f(x+h) = f(x) + \sum_{s=1}^{\nu} \frac{h^s}{s!} f^{(s)}(x) + O(h^{\nu+1}) \quad \text{uniformly in } x \in \mathbb{R} \text{ as } h \to 0, \tag{31}$$

and thus

$$f(x - d_n(u)) = f(x - \mu_n) + \sum_{s=1}^{2m+1} \frac{h_n(u)^s}{s!} f^{(s)}(x - \mu_n) + O\left(k_n^{-3(m+1)/5}\right)$$

uniformly in $x \in \mathbb{R}$ and $u \in [-T_n, T_n]$ as $n \to \infty$. For integer $s \ge 1$ we can write $(d^s/dx^s)\varphi(x)P_{\ell,j}(x) = \varphi(x)P_{\ell,j,s}(x)$, with $P_{\ell,j}$ the polynomials of (30) and $P_{\ell,j,s}$ some other polynomials. Combining (30) and (31) and using the definition of the Chebyshev-Hermite polynomials, we obtain for $P(H_n \le x)$ the expansion

$$\Phi(x - \mu_n) + \varphi(x - \mu_n) \sum_{s=1}^{2m+1} \widetilde{H}_{s-1}(x - \mu_n) \frac{\mu_n^s}{s!} \int_{-T_n}^{T_n} \{1 - y_n(u)\}^s \phi_n(u)\, du$$
$$+ \varphi(x - \mu_n) \sum_{\substack{0 \le \ell, j \le m \\ 1 \le \ell + j \le m}} P_{\ell,j}(x - \mu_n) k_n^{-(\ell+j)/2} \lambda_n^j \int_{-T_n}^{T_n} y_n(u)^j \phi_n(u)\, du$$
$$+ \varphi(x - \mu_n) \sum_{\substack{0 \le \ell, j \le m \\ 1 \le \ell + j \le m}} k_n^{-(\ell+j)/2} \lambda_n^j \sum_{s=1}^{2m+1} \frac{\mu_n^s}{s!} P_{\ell,j,s}(x - \mu_n) \int_{-T_n}^{T_n} \{1 - y_n(u)\}^s y_n(u)^j \phi_n(u)\, du$$
$$- \frac{B_0 \lambda_n}{\gamma} \varphi(x - \mu_n) b(n/k_n) + O\left(k_n^{-(m+1)/2}\right) + o(b(n/k_n)) \tag{32}$$

uniformly in $x \in \mathbb{R}$ as $n \to \infty$. For integer $1 \le k < n$, denote $V_{n,k} = (n/k)(1 - U_{n-k,n}) \overset{d}{=} (n/k)U_{k+1:n}$. For any $v \ge 0$, we have

$$\begin{aligned}\int_{-T_n}^{T_n} y_n(u)^v \phi_n(u)\, du &= \int_{a_n - b_n T_n}^{a_n + b_n T_n} \{(n/k_n)(1-u)\}^{v\alpha}\, dP\left(U_{n-k_n:n} \le u\right) \\ &= E[\{(n/k_n)(1 - U_{n-k_n:n})\}^{v\alpha} \mathbf{1}(a_n - b_n T_n \le U_{n-k_n:n} \le a_n + b_n T_n)] \\ &= E[V_{n,k_n}^{v\alpha}] - E[\{(n/k_n)(1 - U_{n-k_n:n})\}^{v\alpha} \mathbf{1}(U_{n-k_n:n} \notin [a_n - b_n T_n, a_n + b_n T_n])].\end{aligned}$$

Clearly, by (11),

$$P(U_{n-k_n:n} \notin [a_n - b_n T_n, a_n + b_n T_n]) = O(k_n^{-\tau}) \quad \text{for all } \tau > 0,$$

and by (9) and Stirling's formula

$$E(V_{n,k_n}^{\beta}) = \left(\frac{n}{k_n}\right)^{\beta} \frac{\Gamma(n+1)\Gamma(\beta + k_n + 1)}{\Gamma(k_n + 1)\Gamma(\beta + n + 1)} \to 1 \quad \text{as } n \to \infty \text{ for all } \beta > 0.$$

Hence the Cauchy-Schwarz inequality yields for all $\tau > 0$

$$\int_{-T_n}^{T_n} y_n(u)^v \phi_n(u)\, du = E\left[V_{n,k_n}^{v\alpha}\right] + O(k_n^{-\tau}) \quad \text{as } n \to \infty. \tag{33}$$

A combination of (32) and (33) completes the proof of Theorem 2.

A Proofs of the Lemmas

A.1 Lemma 1 (Edgeworth expansion)

For $\varepsilon > 0$ and $t \in \mathbb{R}$, denote

$$\sigma(\varepsilon)^2 = E\left[(\xi_1 + \varepsilon\eta_1)^2\right], \qquad \delta(\varepsilon) = \frac{\sigma(\varepsilon)^2}{12E\left[|\xi_1 + \varepsilon\eta_1|^3\right]}, \qquad f_\varepsilon(t) = E[\exp\{it(\xi_1 + \varepsilon\eta_1)\}].$$

Theorem VI.1 on page 159 of Petrov (1975) implies for all integer $k \ge 1$ and all $\varepsilon > 0$

$$\begin{aligned}&\sup_{x \in \mathbb{R}} \left| P\left\{ k^{-1/2} \sum_{i=1}^{k} \frac{\xi_i + \varepsilon\eta_i}{\sigma(\varepsilon)} \le x \right\} - \Phi(x) - \sum_{\ell=1}^{m} k^{-\ell/2} Q_\ell(\varepsilon, x) \right| \\ &\le \; C(m) \Bigg\{ \sigma(\varepsilon)^{-(m+2)} k^{-m/2} E\left[|\xi_1 + \varepsilon\eta_1|^{m+2} \mathbf{1}\{|\xi_1 + \varepsilon\eta_1| \ge \sigma(\varepsilon) k^{1/2}\}\right] \\ &\qquad + \sigma(\varepsilon)^{-(m+3)} k^{-(m+1)/2} E\left[|\xi_1 + \varepsilon\eta_1|^{m+3}\right] \\ &\qquad + \Big(\sup_{|t| \ge \delta(\varepsilon)} |f_\varepsilon(t)| + (2k)^{-1} \Big)^k k^{(m+2)(m+3)/2} \Bigg\}\end{aligned} \tag{34}$$

with $C(m)$ a finite constant depending only on m and $Q_\ell(\varepsilon, x)$ defined in terms of the first $m+2$ moments of $\xi_1 + \varepsilon\eta_1$ as in (VI.1.13) on page 139 in Petrov (1975).

Apply the elementary inequality, valid for any random variable X,

$$E[|X|^a \mathbf{1}(|X| \geq c)] \leq c^{-1} E[|X|^{a+1}] \quad \text{for all } a > 0 \text{ and } c > 0.$$

to $X = \xi_1 + \varepsilon\eta_1$, $a = m+2$ and $c = \sigma(\varepsilon)k^{1/2}$. Since, by dominated convergence, $E[|\xi_1 + \varepsilon\eta_1|^r] \to E[|\xi_1|^r] < \infty$ for all $0 < r \leq m+3$ and $\delta(\varepsilon) \to 1/(12E[|\xi_1|^3])$ as $\varepsilon \downarrow 0$, we obtain from (34) that there exist $0 < \varepsilon_1 \leq \varepsilon_0$, $\delta_1 > 0$ and $C_1 > 0$ such that

$$\begin{aligned} &\sup_{x\in\mathbb{R}} \left| P\left\{ k^{-1/2} \sum_{i=1}^k \frac{\xi_i + \varepsilon\eta_i}{\sigma(\varepsilon)} \leq x \right\} - \Phi(x) - \sum_{\ell=1}^m Q_\ell(\varepsilon, x) k^{-\ell/2} \right| \\ &\leq \ C_1 \left\{ k^{-(m+1)/2} + \left(\sup_{|t|\geq\delta_1} |f_\varepsilon(t)| + (2k)^{-1} \right)^k k^{(m+2)(m+3)/2} \right\}, \end{aligned} \tag{35}$$

for all integers $k \geq 1$ and all $0 < \varepsilon \leq \varepsilon_1$.

By assumption (18) there exist $0 < c < 1$ and $0 < t_0 < \infty$ such that

$$\sup_{0<\varepsilon\leq\varepsilon_0} \sup_{|t|\geq t_0} |f_\varepsilon(t)| \leq c.$$

Moreover, in case $0 < \delta_1 < t_0$, Theorem I.1 on page 10 of Petrov (1975) applied to $f = f_\varepsilon$ and $b = t_0$ gives

$$|f_\varepsilon(t)| \leq 1 - \frac{1-c^2}{8t_0^2}\delta_1^2 \quad \text{for } \delta_1 \leq |t| \leq t_0 \text{ and } 0 < \varepsilon \leq \varepsilon_0.$$

Together, we see that there exists $h > 0$ such that

$$\sup_{|t|\geq\delta_1} |f_\varepsilon(t)| \leq 1 - h \quad \text{for all } 0 < \varepsilon \leq \varepsilon_0.$$

In particular, we have for all $\tau > 0$

$$\sup_{0<\varepsilon\leq\varepsilon_0} \left(\sup_{|t|\geq\delta_1} |f_\varepsilon(t)| + (2k)^{-1} \right)^k = O(k^{-\tau}) \quad \text{as } k \to \infty.$$

Therefore, the bound in (35) is of the order $O(k^{-(m+1)/2})$, uniformly in $0 < \varepsilon \leq \varepsilon_1$. Hence by substituting $x/\sigma(\varepsilon)$ for x we arrive at

$$\begin{aligned} &\sup_{0<\varepsilon\leq\varepsilon_1} \sup_{x\in\mathbb{R}} \left| P\left\{ k^{-1/2} \sum_{i=1}^k (\xi_i + \varepsilon\eta_i) \leq x \right\} - \Phi(x/\sigma(\varepsilon)) - \sum_{\ell=1}^m k^{-\ell/2} Q_\ell(\varepsilon, x/\sigma(\varepsilon)) \right| \\ &= \ O\left(k^{-(m+1)/2}\right) \quad \text{as } k \to \infty. \end{aligned} \tag{36}$$

Put $\Phi^{(0)} = \Phi$ and for integer $n \geq 1$ define $\Phi^{(n)}(x) = (d^n/dx^n)\Phi(x)$ for $x \in \mathbb{R}$. Also, set $\mu_n(\varepsilon) = E\left[(\xi_1 + \varepsilon\eta_1)^n\right]$ for $n = 1, \ldots, m+2$. From equation (VI.1.13) on page 139 of Petrov (1975), we obtain for $0 < \varepsilon \leq \varepsilon_1$, $x \in \mathbb{R}$, and $\ell = 1, \ldots, m$

$$Q_\ell(\varepsilon, x) = \sum_{s=1}^{\ell} \sigma(\varepsilon)^{-\ell-2s} K_{\ell,s}(\mu_2(\varepsilon), \ldots, \mu_{\ell+2}(\varepsilon))\Phi^{(\ell+2s)}(x);$$

the $K_{\ell,s}$ are known polynomials depending only on ℓ and s [observe that (VI.1.13) of Petrov (1975) expresses Q_ℓ in terms of the cumulants of $\xi_1 + \varepsilon\eta_1$, which in turn can be expressed as polynomials in the moments $\mu_n(\varepsilon)$, see equation (I.2.3) on page 9 of Petrov (1975)]. Now, the moments $\mu_n(\varepsilon)$ are themselves polynomials in ε, while

$$\sigma(\varepsilon) = \left\{1 + 2\varepsilon E(\xi_1\eta_1) + \varepsilon^2 E(\eta_1^2)\right\}^{1/2}$$

satisfies $\sigma(0) = 1$ and is analytic with respect to ε in a neighborhood of zero. Hence

$$Q_\ell(\varepsilon, x) = \sum_{s=1}^{\ell} g_{\ell,s}(\varepsilon)\Phi^{(\ell+2s)}(x) \quad \text{for } 0 < \varepsilon \leq \varepsilon_1,\ x \in \mathbb{R},\ \ell = 1, \ldots, m, \tag{37}$$

with functions $g_{\ell,s}$ that are analytic in a neighborhood of zero. Combining (36) and (37) we see that, as $k \to \infty$,

$$\begin{aligned} \sup_{0<\varepsilon\leq\varepsilon_1} \sup_{x\in\mathbb{R}} \Bigg| P\left\{k^{-1/2}\sum_{i=1}^{k}(\xi_i + \varepsilon\eta_i) \leq x\right\} - \Phi(x/\sigma(\varepsilon)) \\ - \sum_{\ell=1}^{m} k^{-\ell/2} \sum_{s=1}^{\ell} g_{\ell,s}(\varepsilon)\Phi^{(\ell+2s)}(x/\sigma(\varepsilon))\Bigg| &= O\left(k^{-(m+1)/2}\right). \end{aligned} \tag{38}$$

By Taylor's Theorem and the definition of the Chebyshev-Hermite polynomials we have

$$\Phi^{(n)}(x+h) = \Phi^{(n)}(x) + \sum_{j=1}^{p} \frac{h^j}{j!}\widetilde{H}_{n+j-1}(x)\varphi(x) + \frac{h^{p+1}}{(p+1)!}\widetilde{H}_{n+p}(x+\theta h)\varphi(x+\theta h)$$

for all non-negative integer n and p and all real x and h, where $\theta = \theta(x,h) \in [0,1]$. Putting $h(\varepsilon) = \sigma(\varepsilon)^{-1} - 1$ so that $x/\sigma(\varepsilon) = x + h(\varepsilon)x$, we obtain for all non-negative integer n and p and uniformly in $x \in \mathbb{R}$,

$$\Phi^{(n)}(x/\sigma(\varepsilon)) = \Phi^{(n)}(x) + \sum_{j=1}^{p} \frac{h(\varepsilon)^j}{j!} x^j \widetilde{H}_{n+j-1}(x)\varphi(x) + O(\varepsilon^{p+1}) \quad \text{as } \varepsilon \to 0;$$

to bound the remainder term, we used that $x^n\varphi(x)$ is bounded in $x \in \mathbb{R}$ for all integer $n \geq 1$ and that $h(\varepsilon)$ is analytic in a neighborhood of zero and satisfies $h(0) = 0$. Hence for all integer $p \geq 0$ and uniformly in $x \in \mathbb{R}$ and integer $k \geq 1$,

$$\begin{aligned}
&\Phi(x/\sigma(\varepsilon)) + \sum_{\ell=1}^{m} k^{-\ell/2} \sum_{s=1}^{\ell} g_{\ell,s}(\varepsilon)\Phi^{(\ell+2s)}(x/\sigma(\varepsilon)) \\
&= \Phi(x) + \sum_{j=1}^{p} \frac{h(\varepsilon)^j}{j!} x^j \widetilde{H}_{j-1}(x)\varphi(x) \\
&\quad + \sum_{\ell=1}^{m} k^{-\ell/2} \sum_{s=1}^{\ell} g_{\ell,s}(\varepsilon)\varphi(x)\left(\widetilde{H}_{\ell+2s-1}(x) + \sum_{j=1}^{p} \frac{h(\varepsilon)^j}{j!} x^j \widetilde{H}_{\ell+2s+j-1}(x)\right) + O(\varepsilon^{p+1})
\end{aligned}$$

as $\varepsilon \to 0$. Since $g_{\ell,s}(\varepsilon)$ and $h(\varepsilon)$ are analytic in ε in a neighborhood of zero, we can write them as a polynomial of degree p in ε with error $O(\varepsilon^{p+1})$. Using again that $x^n\varphi(x)$ is bounded in $x \in \mathbb{R}$ for all integer $n \geq 1$, we conclude that for all integer $p \geq 0$ and uniformly in $x \in \mathbb{R}$ and integer $k \geq 1$,

$$\begin{aligned}
&\Phi(x/\sigma(\varepsilon)) + \sum_{\ell=1}^{m} k^{-\ell/2} \sum_{s=1}^{\ell} g_{\ell,s}(\varepsilon)\Phi^{(\ell+2s)}(x/\sigma(\varepsilon)) \\
&= \Phi(x) + \varphi(x) \sum_{\substack{0\leq \ell \leq m,\, 0 \leq j \leq p \\ (\ell,j)\neq(0,0)}} P_{\ell,j}(x) k^{-\ell/2}\varepsilon^j + O(\varepsilon^{p+1}) \quad \text{as } \varepsilon \to 0,
\end{aligned} \tag{39}$$

for some polynomials $P_{\ell,j}$. Combining (39) with (38) finishes the proof of Lemma 1.

A.2 Lemma 2 (Potter bound)

Setting $f(t) = \log V(t) - \gamma \log t$ for all $t > 1$ we have by (6)

$$\lim_{t\to\infty} \frac{1}{b(t)} \left(\frac{f(ty) - f(t)}{a(t)} - CA_{-\alpha}(y) \right) = B(y) \quad \text{for all } y > 0. \tag{40}$$

For $t_0 > 1$ large enough such that a is defined and locally integrable on $[t_0, \infty)$, define

$$\widetilde{f}(t) = C \int_{t_0}^{t} a(u) \frac{du}{u}, \quad \text{for all } t \geq t_0.$$

Since the convergence in (7) takes place locally uniformly in $y \in (0, \infty)$, we have

$$\begin{aligned}
\frac{1}{b(t)} \left(\frac{\widetilde{f}(ty) - \widetilde{f}(t)}{a(t)} - CA_{-\alpha}(y) \right) &= C \int_{1}^{y} \frac{1}{b(t)} \left(\frac{a(ut)}{a(t)} - u^{-\alpha} \right) \frac{du}{u} \\
&\to CD \int_{1}^{y} u^{-\alpha} A_{-\beta}(u) \frac{du}{u} =: \widetilde{B}(y) \quad \text{as } t \to \infty
\end{aligned}$$

for all $y > 0$. Defining $g(t) = f(t) - \widetilde{f}(t)$, we obtain that

$$\lim_{t\to\infty} \frac{g(yt) - g(t)}{a(t)b(t)} = B(y) - \widetilde{B}(y), \quad \text{for all } y > 0.$$

Since the function ab is regularly varying with index $-\alpha-\beta \le 0$, Theorem 3.1.4 of Bingham *et al.* (1987) implies the existence of $t_1 \ge t_0$ and $K_1 > 0$ such that

$$\left| \frac{g(yt) - g(t)}{a(t)b(t)} \right| \le K_1 y^\varepsilon, \quad \text{for all } y \ge 1,\ t \ge t_1.$$

Now since $f = \widetilde{f} + g$, we have for all $y \ge 1$ and $t \ge t_1$

$$\begin{aligned} &\left| \frac{1}{b(t)} \left(\frac{f(ty) - f(t)}{a(t)} - CA_{-\alpha}(y) \right) \right| \\ &\quad \le \left| \frac{1}{b(t)} \left(\frac{\widetilde{f}(ty) - \widetilde{f}(t)}{a(t)} - CA_{-\alpha}(y) \right) \right| + \left| \frac{g(yt) - g(t)}{a(t)b(t)} \right| \\ &\quad \le |C| \int_1^y u^{-\alpha} \left| \frac{(ut)^\alpha a(ut) - t^\alpha a(t)}{t^\alpha a(t)b(t)} \right| \frac{du}{u} + K_1 y^\varepsilon. \end{aligned}$$

Equation (7) implies that

$$\lim_{t\to\infty} \frac{(ut)^\alpha a(ut) - t^\alpha a(t)}{t^\alpha a(t)b(t)} = DA_{-\beta}(u) \quad \text{for all } u > 0.$$

Since the function $t \mapsto t^\alpha a(t)b(t)$ is regularly varying with index $-\beta \le 0$, a second application of Theorem 3.1.4 of Bingham *et al.* (1987) yields $t_2 \ge t_0$ and $K_2 > 0$ such that

$$\left| \frac{(ut)^\alpha a(ut) - t^\alpha a(t)}{t^\alpha a(t)b(t)} \right| \le K_2 u^\varepsilon, \quad \text{for all } u \ge 1,\ t \ge t_2.$$

Together, we obtain for $y \ge 1$ and $t \ge t_\varepsilon = \max(t_1, t_2)$

$$\left| \frac{1}{b(t)} \left(\frac{f(ty) - f(t)}{a(t)} - CA_{-\alpha}(y) \right) \right| \le |C| K_2 \int_1^y u^{-\alpha+\varepsilon-1} du + K_1 y^\varepsilon \le (\varepsilon^{-1} |C| K_2 + K_1) y^\varepsilon,$$

as desired.

Acknowledgment

This paper was partially written while Erich Haeusler was visiting the Catholic University Leuven. He is very grateful to Jef Teugels for his hospitality. He thanks Walter van Assche, Jan Beirlant and Jef Teugels for stimulating discussions. He also thanks Laurens de Haan for pointing out the paper by Drees (1998) to him. Johan Segers is a post-doctoral research fellow of the Fund for Scientific Research, Flanders.

References

Bingham, N.H., Goldie, C.M. and Teugels, J.L. (1987). *Regular Variation.* Cambridge University Press, Cambridge.

Cheng, S. and de Haan, L. (2001). Penultimate Approximation for Hill's estimator. *Scand. J. Statist.* **28**, 569–575.

Cheng, S. and Pan, J. (1998). Asymptotic expansions of estimators for the tail index with applications. *Scand. J. Statist.* **25**, 717–728.

Drees, H. (1998). On smooth statistical tail functionals. *Scand. J. Statist.* **25**, 187–210.

Ferreira, A. (2002). *Statistics of Extremes: Estimation and Optimality.* Ph.D. dissertation, Tilburg University.

Geluk, J.L. and de Haan, L. (1987). *Regular variation, extensions and Tauberian theorems.* CWI Tract 40, Amsterdam.

Guillou, A. and Hall, P. (2001). A diagnostic for selecting the threshold in extreme value analysis. *J. R. Statist. Soc., Ser. B* **63**, 293–305.

de Haan, L. and Peng, L. (1998). Comparison of tail index estimators. *Statist. Neerl.* **52**, 60–70.

de Haan, L. and Stadtmüller, U. (1996). Generalized regular variation of second order. *J. Austral. Math. Soc., Ser. A* **61**, 381–395.

Haeusler, E. and Teugels, J.L. (1985). On asymptotic normality of Hill's estimator for the exponent of regular variation. *Ann. Statist.* **13**, 743–756.

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

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

Mason, D.M. (1982). Laws of large numbers for sums of extreme values. *Ann. Probab.* **10**, 754–764.

Petrov, V.V. (1975). *Sums of Independent Random Variables.* Springer-Verlag, Berlin.

Segers, J. (2001). Abelian and Tauberian theorems for the bias of the Hill estimator. *Scand. J. Statist.* **29**, 461–483.

Shorack, G.R. and Wellner, J.A. (1986). *Empirical Processes.* Wiley, New York.