Methodology and Computing in Applied Probability, 7, 139–148, 2005

Iterative Estimation of the Extreme Value Index*

SAMUEL MÜLLER samuel.mueller@stat.unibe.ch
Department of Mathematical Statistics and Actuarial Science, University of Bern, 3012 Bern, Switzerland

JÜRG HÜSLER juerg.huesler@stat.unibe.ch
Department of Mathematical Statistics and Actuarial Science, University of Bern, 3012 Bern, Switzerland

Received May 15, 2004; Revised November 19, 2004; Accepted January 28, 2005

Abstract. Let $\{X_n,\ n \geq 1\}$ be a sequence of independent random variables with common continuous distribution function F having finite and unknown upper endpoint. A new iterative estimation procedure for the extreme value index γ is proposed and one implemented iterative estimator is investigated in detail, which is asymptotically as good as the uniform minimum varianced unbiased estimator in an ideal model. Moreover, the superiority of the iterative estimator over its non iterated counterpart in the non asymptotic case is shown in a simulation study.

Keywords: extreme value theory, tail index estimation, iterative estimator

AMS 2000 Subject Classification: 62G32

1. Introduction

Let $\{X_n,\ n \geq 1\}$ be a sequence of independent random variables with common distribution function F, such that F belongs to the max-domain of attraction of G, denoted by $F \in \mathcal{D}(G)$ i.e., there exist constants $a_n > 0$, $b_n \in \mathbb{R}$ such that for $x \in \mathbb{R}$

$$G(x) = \lim_{n\to\infty} \mathrm{P}\left(a_n^{-1} \cdot [\max(X_1, \ldots, X_n) + b_n] \leq x\right) = \lim_{n\to\infty} F^n(a_n \cdot x - b_n)$$

$$\Leftrightarrow \sup_{x\in\mathbb{R}} |F^n(a_n \cdot x - b_n) - G(x)| \to 0, \text{ as } n \to \infty.$$

From Gnedenko (1943) it is known that $F \in \mathcal{D}(G)$ if and only if $G \in \left\{G_\gamma : \gamma \in \mathbb{R}\right\}$, where

$$G_\gamma(x) = \exp\left(-(1+\gamma x)^{-1/\gamma}\right),\ 1 + \gamma x > 0,$$

and $G_\gamma(\cdot)$ is called an extreme value distribution. Since

$$(1+\gamma x)^{-1/\gamma} \to \exp(-x), \quad \text{for } \gamma \to 0,$$

interpret $G_0(x)$ as $\exp(-e^{-x})$.

*Supported by Swiss National Science foundation.

We consider the case that the distribution function F has finite upper endpoint $\omega(F) := \sup\{x: F(x) < 1\} < \infty$ and that F is continuous in the left neighborhood of ω. Hence $\gamma \le 0$.

Estimators for the extreme value tail index γ based on maximum-likelihood go back to the work of Hall (1982) and Smith (1985, 1987) and Smith and Weissman (1985) and it is well known that these estimators are not consistent for $\gamma < -1/2$. In the last three decades the estimation of γ was intensively studied and to list all relevant articles would go beyond the scope of this article. Very recent contributions among others are Ferreira et al. (2003), Müller (2003), and Paulauskas (2003).

The paper is organized as follows: In Section 2 an iterative procedure for the estimation of the extreme tail index is derived and the definition of the iterated tail index estimator is given. In Section 3 the main theorem is given, which shows that the defined estimator is asymptotically as good as the uniform minimum varianced unbiased estimator in an ideal model. Moreover, the superiority of the iterated estimator over its non iterated counterpart in the non asymptotic case is shown in a simulation study. The proof of the main theorem is given in Section 4.

2. Motivation and Definition

Falk (1994, 1995) considered the estimation problem in the setting of independent random variables $X_1, \ldots, X_n$ having distribution function F with a finite right endpoint and under the assumption that F possesses a density f in a left neighborhood of $\omega(F)$, which satisfies for some $\delta > 0$ the expansion

$$f(x) = \frac{\exp(-b/\gamma)}{-\gamma}(\omega - x)^{-\frac{1}{\gamma}-1}\Big(1 + O\Big((\omega - x)^{-\delta/\gamma}\Big)\Big) \tag{1}$$

as x tends to ω from below for some $\gamma < 0$, $\delta > 0$, $b \in \mathbb{R}$. Let $X_{(1,n)} \le X_{(2,n)} \le \ldots \le X_{(n,n)}$ denote the order statistics of $X_1, \ldots, X_n$ and let $(k_n, n \ge 1)$ be an intermediate sequence of integers, i.e., $k_n \to \infty$ and $k_n/n \to 0$ as $n \to \infty$. Falk (1994) showed that

$$\widehat{\gamma}_\omega := \frac{1}{k_n}\sum_{j=1}^{k_n}\log\left(\frac{\omega - X_{(n-j+1,n)}}{\omega - X_{(n-k_n,n)}}\right) \tag{2}$$

is in an ideal model an uniform minimum varianced unbiased estimator (UMVUE) if the endpoint ω is known. Falk (1995) replaced the endpoint ω of F by the sample maximum $X_{(n,n)}$ in the case of unknown ω and it turned out that in the case $\gamma < -1/2$

$$\widehat{\gamma}_{\text{Falk}} := \frac{1}{k_n - 1}\sum_{j=2}^{k_n}\log\left(\frac{X_{(n,n)} - X_{(n-j+1,n)}}{X_{(n,n)} - X_{(n-k_n,n)}}\right) \tag{3}$$

is asymptotically as good as his nonrandom counterpart with known endpoint ω (Theorem 1.2. and 1.3. in Falk (1995)).

Replacing the endpoint ω by the sample maximum $X_{(n,n)}$ in Equation (2) is not the only way to get an estimate for the tail index, actually any endpoint estimator could be used instead. From an algorithmic point of view the following iterative procedure has the potential to produce better estimates for finite sample size than the tail index estimator $\widehat{\gamma}_{\mathrm{Falk}}$ defined in Equation (3).

Step 1: Estimate γ by $\widehat{\gamma}_1$, where $\widehat{\gamma}_1$ is any tail estimator for negative γ.
Step 2: Estimate ω by an endpoint estimator $\widehat{\omega}(\widehat{\gamma}_1)$.
Step 3: Re-estimate γ replacing ω in Equation (2) by $\widehat{\omega}(\widehat{\gamma}_1)$.

There exists numerous ways of implementing this iterative procedure. In this paper we will consider only one iterated estimator in detail. For the first step, let $\widehat{\gamma}_1$ be any consistent estimator for γ. For the second step we take the endpoint estimator of Hall (1982), which is based on a linear combination of the m largest order statistics which has the property that for known γ its asymptotic distribution has mean ω and a variance which is a minimum among all such linear combinations if $\gamma < -1/2$. Thus our endpoint estimator used in the second step is

$$\widehat{\omega}_{\mathrm{Hall},\,m}(\widehat{\gamma}_1) := \sum_{j=1}^{m} a_j(\widehat{\gamma}_1) X_{(n-j+1,\,n)}, \tag{4}$$

where the weights $\boldsymbol{a} = (a_1, \ldots, a_m)^T$ are given by

$$\boldsymbol{a} = \frac{\boldsymbol{\Lambda}^{-1}\left[(\boldsymbol{v}^T\boldsymbol{\Lambda}^{-1}\boldsymbol{v})\mathbf{1}_m - \left(\mathbf{1}_m^T\boldsymbol{\Lambda}^{-1}\boldsymbol{v}\right)\boldsymbol{v}\right]}{(\boldsymbol{v}^T\boldsymbol{\Lambda}^{-1}\boldsymbol{v})\left(\mathbf{1}_m^T\boldsymbol{\Lambda}^{-1}\mathbf{1}_m\right) - \left(\mathbf{1}_m^T\boldsymbol{\Lambda}^{-1}\boldsymbol{v}\right)^2}, \tag{5}$$

with $\boldsymbol{\Lambda} = \lambda_{ij}$ the symmetric $m \times m$ matrix given by

$$\lambda_{ij} = \frac{\Gamma(-2\widehat{\gamma}_1 + i)\Gamma(-\widehat{\gamma}_1 + j)}{\Gamma(-\widehat{\gamma}_1 + i)\Gamma(j)}, \; j \leq i, \text{ and with}$$

$$\mathbf{1}_k := \underbrace{(1, \ldots, 1)^T}_{k\text{ times}},$$

$$\boldsymbol{v} := \left(\frac{\Gamma(-\widehat{\gamma}_1 + 1)}{\Gamma(1)}, \ldots, \frac{\Gamma(-\widehat{\gamma}_1 + m)}{\Gamma(m)}\right)^T,$$

where $\Gamma(\cdot)$ is the Γ-function. Then in the third step the following iterated tail estimator can be defined.

Definition 1 *The iterated tail index estimator* $\widehat{\gamma}_{m,n}$ *is defined by*

$$\widehat{\gamma}_{m,\,n} := \frac{1}{k_n - 1}\sum_{j=2}^{k_n} \log\left(\frac{\widehat{\omega}_{\mathrm{Hall},\,m} - X_{(n-j+1,\,n)}}{\widehat{\omega}_{\mathrm{Hall},\,m} - X_{(n-k_n,\,n)}}\right). \tag{6}$$

Remark 1 *From Definition 1 it follows directly that* $\widehat{\gamma}_{m,n}$ *is location and scale invariant.*

3. Results

In the following Theorem it turns out that given $\gamma < -\frac{1}{2}$ this iterated tail index estimator is asymptotically as good as the best estimator $\widehat{\gamma}_{\text{Falk}}$.

THEOREM 1 *Suppose that F satisfies* (1) *for some $\delta > 0$, $b, \omega \in \mathbb{R}$ and $\gamma < -\frac{1}{2}$. If k_n is an intermediate sequence of integers, which satisfies*

$$\frac{\log n}{\sqrt{k_n}} \to 0, \text{ as } n \to \infty,$$

then

$$\sqrt{k_n}\left|\widehat{\gamma}_{m,n} - \widehat{\gamma}_{\text{Falk}}\right| = o_P(1).$$

However, the efficiency for finite samples shows another behavior. To investigate the finite sample size performance of the iterated tail index estimator we perform simulation studies each based on 5,000 replications. We generate k upper order statistics $k \in \{1{,}000, 2{,}000, 4{,}000\}$ from a power-function distribution with tail index ranging from -1.2 to -0.4 in steps of 0.05 and $\widehat{\gamma}_{\text{Falk}}$ as well as $\widehat{\gamma}_{m,n}$ for $m = 5$ are calculated. Asymptotically the estimator is suitable for the cases $\gamma < -1/2$. But in the finite sample we can still consider this estimator for $\gamma \geq -1/2$. Figure 1 shows the simulated relative efficiencies (re), which are calculated by the ratio of the simulated mean squared errors (mse) of the iterated tail index estimator for $m = 5$ and of Falk's estimator that is

$$\text{re}\left(\widehat{\gamma}_{5,n}, \widehat{\gamma}_{\text{Falk}}\right) = \frac{\text{mse}\left(\widehat{\gamma}_{\text{Falk}}\right)}{\text{mse}\left(\widehat{\gamma}_{m,n}\right)}.$$

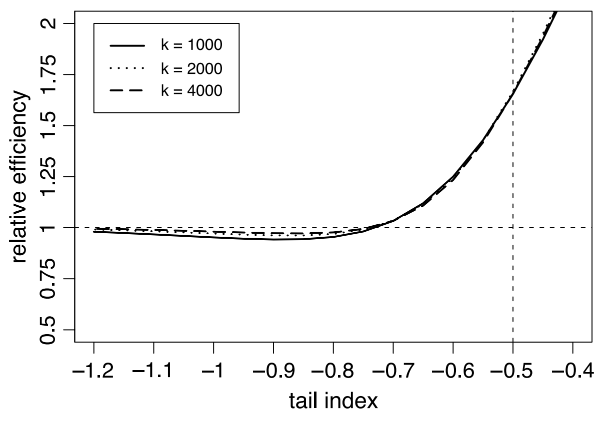

Figure 1. Relative efficiencies of $\widehat{\gamma}_{m,n} = \widehat{\gamma}\left(\widehat{\omega}_{\text{Hall},5}\left(\widehat{\gamma}_{\text{Falk}}\right)\right)$ and Falk's estimator.

Table 1. Relative efficiencies for $\gamma = -0.6$ in dependence of k.

k	$\mathrm{mse}(\widehat{\gamma}_{\mathrm{Falk}})$	$\mathrm{mse}(\widehat{\gamma}_{5,n})$	$\mathrm{re}(\widehat{\gamma}_{5,n}, \widehat{\gamma}_{\mathrm{Falk}})$
1,000	$8.1271 \cdot 10^{-4}$	$6.5031 \cdot 10^{-4}$	1.2497
4,000	$1.8412 \cdot 10^{-4}$	$1.4924 \cdot 10^{-4}$	1.2337
16,000	$0.3968 \cdot 10^{-4}$	$0.3236 \cdot 10^{-4}$	1.2262
64,000	$0.0912 \cdot 10^{-4}$	$0.0754 \cdot 10^{-4}$	1.2095

The efficiency curves suggest, that the iterated tail index is superior in the left neighborhood of $\gamma = -0.5$ and slightly less efficient for γ in a neighborhood of -0.9. Moreover, a closer look reveals that for increasing n the efficiency curve is approximating the horizontal line of equal efficiency. From Tables 1 and 2 we conclude that this approximation process is very slow. The simulated values for $k = 16{,}000$ and $k = 64{,}000$ are both based on 5,000 replications.

The influence of the number of upper order statistics used for estimating ω is rather important as can be seen in Figure 2, which shows the simulated efficiencies for $k = 2{,}000$ in dependence of $m \in \{2, 3, 4, 5\}$. It indicates that the efficiency is increasing with m.

Overall the simulation results suggest that the proposed iterative procedure is superior than its non iterated counterpart.

REMARK 2 *Theorem 1 is still holding, if $\widehat{\gamma}_1$ is replaced by any other consistent estimator for γ.*

Moreover, it is interesting to note that even if a rather inefficient initial tail index estimator is used the superiority of the iterated tail index estimator over Falk's estimator is still holding. For example let $\widehat{\gamma}_1$ be Pickands estimator (Pickands, 1975) defined by

$$\widehat{\gamma}_P(n,r) := \frac{1}{\log 2}\log\left(\frac{X_{(n-r,n)} - X_{(n-2r,n)}}{X_{(n-2r,n)} - X_{(n-4r,n)}}\right),$$

which is well known to be less efficient than Falk's estimator if $\gamma \leq -1/2$. For sample size $k = 1{,}000$ we continue the simulation study from above. Based on 5,000 simulation runs we calculate Falk's estimator, Pickands estimator for $r = \lfloor k/4 \rfloor$, the iterated tail index estimator using Falk's estimator (Falk–Falk), as well as Pickands estimator

Table 2. Relative efficiencies for $\gamma = -0.9$ in dependence of k.

k	$\mathrm{mse}(\widehat{\gamma}_{\mathrm{Falk}})$	$\mathrm{mse}(\widehat{\gamma}_{5,n})$	$\mathrm{re}(\widehat{\gamma}_{5,n}, \widehat{\gamma}_{\mathrm{Falk}})$
1,000	$8.0910 \cdot 10^{-4}$	$8.5870 \cdot 10^{-4}$	0.9422
4,000	$2.0248 \cdot 10^{-4}$	$2.0812 \cdot 10^{-4}$	0.9729
16,000	$0.4887 \cdot 10^{-4}$	$0.4952 \cdot 10^{-4}$	0.9869
64,000	$0.1247 \cdot 10^{-4}$	$0.1251 \cdot 10^{-4}$	0.9968

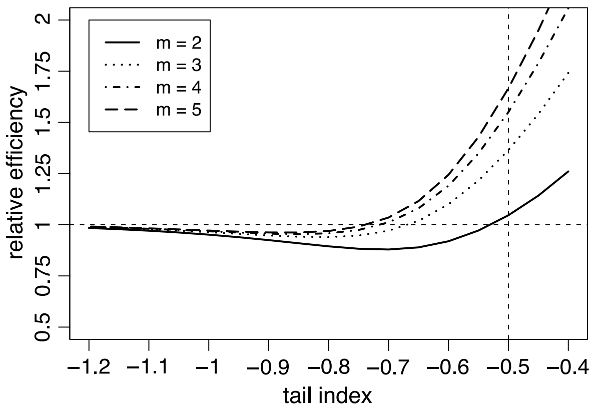

Figure 2. Relative efficiencies of $\widehat{\gamma}_{m,n} = \widehat{\gamma}\big(\widehat{\omega}_{\mathrm{Hall},m}(\widehat{\gamma}_{\mathrm{Falk}})\big)$ and $\widehat{\gamma}_{\mathrm{Falk}}$ for k = 2,000.

(Falk–Pickands) as initial tail index estimator. For comparison purpose we also calculate the shift and scale invariant tail-index moment estimator from Ferreira et al. (2003, Section 2.3) defined by

$$\widehat{\gamma}_{\mathrm{FdHP}}(k) = \frac{N_n^{(2)} - 2\big(N_n^{(1)}\big)^2}{2\Big(N_n^{(1)}\Big)^2 - 2N_n^{(2)}}, \tag{7}$$

with $N_n^{(j)} = \frac{1}{k}\sum_{i=0}^{k-1}\big(X_{(n-i,n)} - X_{(n-k,n)}\big)^j$, j = 1, 2. The relative efficiencies are shown in Figure 3 which underscores the superiority of the iterated tail index estimator.

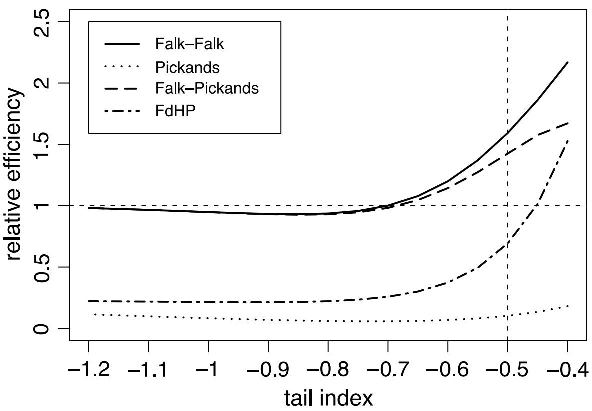

Figure 3. Relative efficiencies of Falk's estimator compared to Pickands, the shift and scale invariant moment estimator, and two versions of the iterated tail index estimator for m = 5 and k = 1,000.

REMARK 3 *In the simulation study we generated the k upper order statistics directly from the power-function distribution and therefore, we do not have to specify n. However, for real data the choice of k has to be addressed. One possibility is to apply the automatic choice procedure of Reiss and Thomas (2001, p. 149) that is the optimal k* is given by*

$$k^* = \underset{k}{\arg\min} \frac{1}{k} \sum_{i \le k} i^{\beta} \left| \widehat{\gamma}_{i,n} - \text{med}\left\{ \widehat{\gamma}_{i,n}, \ldots, \widehat{\gamma}_{k,n} \right\} \right|$$

with $0 \le \beta < 1/2$, *where* $\widehat{\gamma}_{i,n}$ *is some tail-index estimator based on the upper i order statistics.*

4. Proofs

Let $X_i = F^{-1}(1 - U_i)$, $i = 1, 2, \ldots,$ where $U_1, U_2, \ldots$ are iid $\mathcal{U}(0,1)$ random variables. From Proposition 1.1 (ii) in Falk (1994) the expansion

$$F^{-1}(1-q) = const \cdot q^{-\gamma}\left(1 + O\left(q^{\delta}\right)\right)$$

is obtained as $q \to 0$. With the Renyi representation

$$\left(U_{(i,n)}\right)_{i=1}^{n} \overset{\mathcal{L}}{=} \left(\frac{S_i}{S_{n+1}}\right)_{i=1}^{n},$$

where $S_j = \xi_1 + \ldots + \xi_j$ is the sum of j independent and standard exponential random variables (see Corollary 1.6.9 in Reiss (1989)) we get the following results.

PROPOSITION 1 *For* $1 \le j \le k_n$ *the following relation holds uniformly for every* $1 \le j \le k_n$

$$\frac{F^{-1}\left(1 - U_{(1,n)}\right)}{F^{-1}\left(1 - U_{(j+1,n)}\right)} = \frac{U_{(1,n)}^{-\gamma}\left(1 + O\left(U_{(1,n)}^{\delta}\right)\right)}{U_{(j+1,n)}^{-\gamma}\left(1 + O\left(U_{(j+1,n)}^{\delta}\right)\right)}$$

$$\overset{\mathcal{L}}{=} \left(\frac{S_1}{S_{j+1}}\right)^{-\gamma} \frac{1 + O\left((S_1/S_{n+1})^{\delta}\right)}{1 + O\left((S_{j+1}/S_{n+1})^{\delta}\right)} = O_P((1/j)^{-\gamma}).$$

Proof of Proposition 1: The $O_P((1/j)^{-\gamma})$ part is taken from page 119 in Falk (1995). The relation is holding uniformly since S_j/j is obviously more and more concentrating around 1 as j is increasing. More precisely, we can simply apply Chebychev inequality to give a uniform bound for the probability $P(S_j/j \le w)$ for $j \ge j_0$ and small w. ■

Proof of Theorem 1: From Theorem 1.3 in Falk (1995) it is known under the same conditions as in our theorem $\sqrt{k_n}|\widehat{\gamma}_{\text{Falk}} - \widehat{\gamma}_{\omega}| = o_P(1)$. Hence $\sqrt{k_n}\left|\widehat{\gamma}_{m,n} - \widehat{\gamma}_{\text{Falk}}\right| = o_P(1)$ holds if and only if

$$\sqrt{k_n}\left|\widehat{\gamma}_{m,n} - \widehat{\gamma}_{\omega}\right| = o_P(1). \tag{8}$$

Moreover, by the definition of Hall's endpoint estimator it follows that $\widehat{\omega}_{\text{Hall},m}$ is greater than $X_{(n,n)}$. By Definition 1 the iterated tail index estimator $\widehat{\gamma}_{m,n}$ is bounded from below by $\widehat{\gamma}_{\text{Falk}}$

$$\widehat{\gamma}_{m,n} \geq \frac{1}{k-1}\sum_{j=2}^{k}\log\left(\frac{X_{(n,n)} - X_{(n-j+1,n)}}{X_{(n,n)} - X_{(n-k,n)}}\right) = \widehat{\gamma}_{\text{Falk}}, \tag{9}$$

which follows from the inequality

$$\frac{z_1}{z_2} \leq \frac{z_1+\varepsilon}{z_2+\varepsilon}, \quad \text{if } \varepsilon > 0 \text{ and } 0 < z_1 \leq z_2. \tag{10}$$

Hence, for $\widehat{\omega}_{\text{Hall},m} \in [X_{(n,n)}, \omega]$ the results follows immediately from Theorem 1.3 in Falk (1995) and from the inequalities (9) and (10). Thus it suffices to prove the Theorem for the case that

$$\widehat{\omega}_{\text{Hall},m} > \omega. \tag{11}$$

Because of the location invariance of the iterated tail index estimator we assume without loss of generality that

$$\omega = 0. \tag{12}$$

Then we consider

$$\widehat{\gamma}_{m,n} - \widehat{\gamma}_{\omega} = \widehat{\gamma}_{m,n} - \frac{1}{k_n}\sum_{j=2}^{k_n}\log\left(\frac{X_{(n-j+1,n)}}{X_{(n-k_n,n)}}\right) - \underbrace{\frac{1}{k_n}\log\left(\frac{X_{(n,n)}}{X_{(n-k_n,n)}}\right)}_{=:A_n}. \tag{13}$$

Hence using Equation (8) the statement of the theorem holds if $\sqrt{k_n} \cdot A_n = o_P(1)$ and

$$\sqrt{k_n}\left|\frac{k_n-1}{k_n}\widehat{\gamma}_{m,n} - \frac{1}{k_n}\sum_{j=2}^{k_n}\log\left(\frac{X_{(n-j+1,n)}}{X_{(n-k_n,n)}}\right)\right| = o_p(1). \tag{14}$$

For (14) we have

$$\begin{aligned}
0 &\leq \frac{1}{k_n}\sum_{j=2}^{k_n}\log\left(\frac{\widehat{\omega}_{\text{Hall},m} - X_{(n-j+1,n)}}{\widehat{\omega}_{\text{Hall},m} - X_{(n-k_n,n)}}\right) - \frac{1}{k_n}\sum_{j=2}^{k_n}\log\left(\frac{X_{(n-j+1,n)}}{X_{(n-k_n,n)}}\right) \\
&= \frac{1}{k_n}\sum_{j=2}^{k_n}\log\left(\frac{\left(\widehat{\omega}_{\text{Hall},m} - X_{(n-j+1,n)}\right) \cdot X_{(n-k_n,n)}}{\left(\widehat{\omega}_{\text{Hall},m} - X_{(n-k_n,n)}\right) \cdot X_{(n-j+1,n)}}\right) \\
&\leq \frac{1}{k_n}\sum_{j=2}^{k_n}\log\left(\frac{X_{(n-j+1,n)} - \widehat{\omega}_{\text{Hall},m}}{X_{(n-j+1,n)}}\right) = \frac{1}{k_n}\sum_{j=2}^{k_n}\log\left(1 - \frac{\widehat{\omega}_{\text{Hall},m}}{X_{(n-j+1,n)}}\right).
\end{aligned}$$

Moreover from the assumption of $\widehat{\gamma}_1$ we have that $a_1, \ldots, a_m$ are $O_P(1)$. Regarding (11) and (12) Hall's endpoint estimator is bounded from above by

$$\widehat{\omega}_{\mathrm{Hall},m} \leq -dX_{(n-m+1,n)}, \tag{15}$$

where $d = O_P(1)$ is chosen properly, for example any $d \geq 1 + (-X_{(n-m+1,n)})^2$ is suitable. For any $d' \geq 1$ and $x > 0$ we have

$$d' \log(1+x) = \log\Big((1+x)^{d'}\Big) \geq \log(1+d'x) \geq \log(1+x),$$

thus for any fixed $d' > 0$ we have

$$\frac{1}{\sqrt{k_n}} \sum_{j=2}^{k_n} \log\left(1 + \frac{X_{(n-m+1,n)}}{X_{(n-j+1,n)}}\right) = o_P(1)$$

$$\Leftrightarrow \frac{1}{\sqrt{k_n}} \sum_{j=2}^{k_n} \log\left(1 + \frac{d'_{(n-m+1,n)}}{X_{(n-j+1,n)}}\right) = o_P(1).$$

Hence with $d' = d$ and inequality (15) we get

$$\frac{1}{\sqrt{k_n}} \sum_{j=2}^{k_n} \log\left(1 + \frac{dX_{(n-m+1,n)}}{X_{(n-j+1,n)}}\right) = o_P(1)$$

$$\Rightarrow \frac{1}{\sqrt{k_n}} \sum_{j=2}^{k_n} \log\left(1 + \frac{\widehat{\omega}_{\mathrm{Hall},m}}{X_{(n-j+1,n)}}\right) = o_P(1).$$

Hence we have to show, that

$$\frac{1}{\sqrt{k_n}} \sum_{j=2}^{k_n} \log\left(1 + \frac{X_{(n-m+1,n)}}{X_{(n-j+1,n)}}\right) = \frac{1}{\sqrt{k_n}} \sum_{j=2}^{k_n} \log\left(1 + \frac{F^{-1}\big(1 - U_{(m,n)}\big)}{F^{-1}\big(1 - U_{(j,n)}\big)}\right)$$

$$=: B_n = o_P(1).$$

Proposition 1 implies that

$$\frac{F^{-1}\big(1 - U_{(m,n)}\big)}{F^{-1}\big(1 - U_{(j,n)}\big)} = \frac{U_{(m,n)}^{-\gamma}\Big(1 + O\Big(U_{(m,n)}^{\delta}\Big)\Big)}{U_{(j,n)}^{-\gamma}\Big(1 + O\Big(U_{(j,n)}^{\delta}\Big)\Big)}$$

$$\stackrel{\mathcal{L}}{=} \left(\frac{S_m}{S_j}\right)^{-\gamma} \frac{1 + O\Big((S_m/S_{n+1})^{\delta}\Big)}{1 + O\Big((S_j/S_{n+1})^{\delta}\Big)} = O_P((m/j)^{-\gamma})$$

holds uniformly. Using the approximation $\log(1 + z) \sim z$ as $z \to 0$

$$B_n = O_P\left(k_n^{-1/2}\sum_{j=2}^{k_n} j^{\gamma}\right) = O_P\left(k_n^{1/2+\gamma}\right) = o_P(1),$$

and also

$$\sqrt{k_n} \cdot A_n = O_P\left(\frac{1}{\sqrt{k_n}}\log k_n\right) = o_P(1).$$

■

Acknowledgments

The authors thank a referee for his constructive suggestions which lead to improvement in the readability and clarity of the paper. This research was supported by the Schweizer Nationalfonds.

References

M. Falk, "Extreme quantile estimation in δ-neighborhoods of generalized Pareto distributions," *Statistics & Probability Letters* vol. 20 pp. 9–21, 1994.

M. Falk, "Some best parameter estimates for distributions with finite endpoint," *Statistics* vol. 27 pp. 115–125, 1995.

A. Ferreira, L. de Haan, and L. Peng, "On optimizing the estimation of high quantiles of a probability distribution," *Statistics* vol. 37 pp. 401–434, 2003.

B. V. Gnedenko, "Sur la distribution limite du terme maximum d'une série aléatoire," *Annals of Mathematics* vol. 44 pp. 423–453, 1943.

P. Hall, "On estimating the endpoint of a distribution," *Annals of Statistics* vol. 10 pp. 556–568, 1982.

S. Müller, "Tail estimation based on numbers of near *m*-extremes," *Methodology and Computing in Applied Probability* vol. 5 pp. 197–210, 2003.

V. Paulauskas, "A new estimator for a tail index," *Acta Applicandae Mathematicae* vol. 79 pp. 55–67, 2003.

J. Pickands, "Statistical inference using extreme order statistics," *Annals of Statistics* vol. 3 pp. 119–131, 1975.

R.-D. Reiss, *Approximate Distributions of Order Statistics (With Applications to Nonparametric Statistics)*, Springer Series in Statistics: New York, 1989.

R.-D. Reiss and M. Thomas, *Statistical Analysis of Extreme Values*, Birkhäuser: Basel, 2001.

R. L. Smith, "Maximum likelihood estimation in a class of nonregular cases," *Biometrika* vol. 72 pp. 67–90, 1985.

R. L. Smith, "Estimating tails of probability distributions," *Annals of Statistics* vol. 15 pp. 1174–1207, 1987.

R. L. Smith and I. Weissman, "Maximum likelihood estimation of the lower tail of a probability distribution," *Journal of the Royal Statistical Society Series B* vol. 47 pp. 285–298, 1985.