A Bootstrap-based Method to Achieve Optimality in Estimating the Extreme-value Index

L. de Haan, L. Peng
Erasmus University Rotterdam
T.T. Pereira
University of Lisbon

August 18, 1997

Abstract

Estimators of the extreme-value index are based on a set of upper order statistics. We present an adaptive method to choose the number of order statistics involved in an optimal way, balancing variance and bias components. Recently this has been achieved for the similar but somewhat less involved case of regularly varying tails (Drees and Kaufmann(1997); Danielsson et al.(1997)). The present paper follows the line of proof of the last paper.

Key words & phrases: Moment estimator, Pickands estimator, bootstrap, mean squared error.

1 Introduction

Suppose we have i.i.d. observations $X_1, X_2, \cdots, X_n$ whose common distribution function F is in the domain of attraction of an extreme-value distribution. The shape parameter $\gamma \in R$ of this extreme-value distribution (functional form: $\exp(-(1+\gamma x)^{-1/\gamma})$) can be estimated in various ways starting from the sample $X_1, X_2, \cdots, X_n$. Two popular estimators are Pickands' estimator (in its generalized form see e.g. Pereira (1993)):

$$ {}^{(P)}\hat{\gamma}_{n,\theta}(k) := (-\log\theta)^{-1} \log \frac{X_{n,n-[k\theta^2]} - X_{n,n-[k\theta]}}{X_{n,n-[k\theta]} - X_{n,n-k}} \tag{1.1} $$

($\theta \in (0,1)$) where $X_{n,1} \leq \cdots \leq X_{n,n}$ are the order statistics of $X_1, \cdots, X_n$ and $[z]$ denotes the largest integer which is not larger than z, and the moment estimator

$$ {}^{(M)}\hat{\gamma}_n(k) := M_n^{(1)} + 1 - \frac{1}{2}(1 - \frac{(M_n^{(1)})^2}{M_n^{(2)}})^{-1} \tag{1.2} $$

with $M_n^{(j)} := \frac{1}{k}\sum_{i=0}^{k-1}(\log X_{n,n-i} - \log X_{n,n-k})^j$. For this estimator we have to require that the right end point of the distribution is positive.

If the underlying probability distribution is known, the asymptotically optimal value of k can be determined (Dekkers and de Haan (1993)) in minimal mean squared error sense by assuming some kind of second order condition. However the asymptotically optimal value of k depends on the unknown parameter γ and on the unknown second order function. We shall develop a bootstrap procedure that gives the asymptotically optimal value of k adaptively. Results for moment estimator and for Pickands' estimator are given in section 2 and section 3 respectively. All the proofs are postponed till section 4. In appendix we shall explain why we use different second order conditions in section 2 and section 3.

2 Main results for moment estimator

Throughout this section we assume $U(\infty) > 0$ and the following second order conditions:

$$ \lim_{t\to\infty} \frac{\frac{\log U(tx) - \log U(t)}{a(t)/U(t)} - \frac{x^{\gamma\wedge 0}-1}{\gamma\wedge 0}}{A(t)} = H(x) \tag{2.1} $$

where $U(t)$ is the inverse function of the function $1/(1-F)$, $a(t)$ is positive and A not changing sign eventually. The function $H(x)$ is assumed not to be a multiple of $(x^\gamma - 1)/\gamma$ and takes the form (supposing the function a and A are chosen properly)

$$ H(x) = \frac{1}{\rho}[\frac{x^{\rho+\gamma\wedge 0}-1}{\rho+\gamma\wedge 0} - \frac{x^{\gamma\wedge 0}-1}{\gamma\wedge 0}], \tag{2.2} $$

depending on a second order parameter $\rho \leq 0$ (see de Haan and Stadtmüller, relation (2.9) page 387).

First we restate in slightly greater generality a result from Dekkers and de Haan (1993) providing the optimal number of order statistics for the moment estimator as a function of γ, ρ and the function A.

Theorem 2.1. *Suppose (2.1) and (2.2) hold for $\rho < 0$. Let*

$$k_0(n) := arg \inf_k E({}^{(M)}\hat{\gamma}_n(k) - \gamma)^2. \tag{2.3}$$

Then

$$k_0(n) \sim n(\frac{V^2(\gamma)}{b^2(\gamma,\rho)})^{\frac{1}{1-2\rho}}(s^{-}(\frac{1}{n}))^{-1} \tag{2.4}$$

where

$$V^2(\gamma) = \gamma_+^2 + (1-\gamma_-)^2(1-2\gamma_-)\{4 - 8\frac{1-2\gamma_-}{1-3\gamma_-} + \frac{(5-11\gamma_-)(1-2\gamma_-)}{(1-3\gamma_-)(1-4\gamma_-)}\} \tag{2.5}$$

with $\gamma_+ = 0 \vee \gamma$ and $\gamma_- = 0 \wedge \gamma$ (the variance component) and

$$b(\gamma,\rho) = \frac{\gamma_+}{(1-\gamma_-)(\bar{\rho}-\gamma_-)} + \frac{(1-\gamma_-)(1-2\gamma_-)}{(\bar{\rho}-\gamma_-)(\bar{\rho}-2\gamma_-)} \tag{2.6}$$

with $\bar{\rho} = 1-\rho$ (the bias component). The function s^- is the inverse function of the decreasing function s satisfying

$$A^2(t) = (1+o(1))\int_t^\infty s(u)\,du. \tag{2.7}$$

We are going to turn the formula in (2.3) into something we can handle adaptively, the first step is to replace the unknown γ in the formula by an alternative estimator for γ. The alternative estimator is

$$\hat{\gamma}_n(k) := \sqrt{M_n^{(2)}/2} + 1 - \frac{2}{3}(1 - \frac{M_n^{(1)}M_n^{(2)}}{M_n^{(3)}})^{-1}.$$

The proof of the following Theorem is very similar to that of Theorem 2.1(involving more lengthy calculations) and will be omitted.

Theorem 2.2. *Assume the conditions of Theorem 2.1. Let*

$$\bar{k}_0(n) := arg \inf_k E({}^{(M)}\hat{\gamma}_n(k) - \hat{\gamma}_n(k))^2.$$

Then

$$\bar{k}_0(n) \sim n(\frac{\bar{V}^2(\gamma)}{\bar{b}^2(\gamma,\rho)})^{\frac{1}{1-2\rho}}(s^{-}(\frac{1}{n}))^{-1}$$

where

$$\begin{aligned}
& (1-\gamma_-)^{-2}\bar{V}^2(\gamma) \\
= \; & (\gamma_+ - 2(1-2\gamma_-) + \tfrac{1}{2}(1-3\gamma_-))^2\tfrac{1}{1-2\gamma_-} \\
& +(-\tfrac{\gamma_+}{4} + \tfrac{(1-2\gamma_-)^2}{2} + \tfrac{(1-2\gamma_-)(1-3\gamma_-)}{4})^2\tfrac{4(5-11\gamma_-)}{(1-2\gamma_-)^2(1-3\gamma_-)(1-4\gamma_-)} \\
& +\tfrac{3(1-3\gamma_-)^2(146\gamma_-^2-105\gamma_-+19)}{(1-4\gamma_-)(1-5\gamma_-)(1-6\gamma_-)} \\
& +\tfrac{8(\gamma_+-2(1-2\gamma_-)+\frac{1-3\gamma_-}{2})(-\frac{\gamma_+}{4}+\frac{(1-2\gamma_-)^2}{2}+\frac{(1-2\gamma_-)(1-3\gamma_-)}{4})}{(1-2\gamma_-)(1-3\gamma_-)} \\
& +\tfrac{3(\gamma_+-2(1-2\gamma_-)+\frac{1-3\gamma_-}{2})(1-3\gamma_-)}{1-4\gamma_-} \\
& -\tfrac{6(-\frac{\gamma_+}{4}+\frac{(1-2\gamma_-)^2}{2}+\frac{(1-2\gamma_-)(1-3\gamma_-)}{4})(1-3\gamma_-)(7\gamma_--3)}{(1-2\gamma_-)(1-4\gamma_-)(1-5\gamma_-)}
\end{aligned}$$

and

$$\begin{aligned}
\bar{b}(\gamma,\rho) \; = \; & \tfrac{\gamma_+(2\bar{\rho}-1)}{4\bar{\rho}^2} + \tfrac{(1-\gamma_-)(5\gamma_--3)}{2(\bar{\rho}-\gamma_-)} \\
& +\tfrac{(1-\gamma_-)(3-7\gamma_-)(2\bar{\rho}+1-4\gamma_)}{2(\bar{\rho}-\gamma_-)(\bar{\rho}-2\gamma_-)} \\
& -\tfrac{(1-\gamma_-)(1-3\gamma_-)(1-2\bar{\rho}+3\bar{\rho}^2+15\bar{\rho}\gamma_--7\gamma_-+18\gamma_-^2)}{2(\bar{\rho}-\gamma_-)(\bar{\rho}-2\gamma_-)(\bar{\rho}-3\gamma_-)}.
\end{aligned}$$

Next we are going to introduce the bootstrap procedure. One takes n_1 independent drawings from the empirical distribution function of $\mathcal{X}_n := \{X_1, \cdots, X_n\}$. This results in observations $X_1^*, \cdots, X_{n_1}^*$. We form the order statistics $X_{n_1,1}^* \leq \cdots \leq X_{n_1,n_1}^*$ and define

$${}^{(j)}M_{n_1}^*(k_1) := \frac{1}{k_1}\sum_{i=1}^{k_1}(\log X_{n_1,n_1-i+1}^* - \log X_{n_1,n_1-k_1}^*)^j$$

for $k_1 < n_1$ and $j = 1, 2, 3$. Next define

$${}^{(M)}\hat{\gamma}_{n_1}^*(k_1) :={}^{(1)} M_{n_1}^* + 1 - \frac{1}{2}(1 - \frac{({}^{(1)}M_{n_1}^*)^2}{{}^{(2)}M_{n_1}^*})^{-1}$$

and

$$\hat{\gamma}^*_{n_1}(k_1) := \sqrt{{}^{(2)}M^*_{n_1}/2} + 1 - \frac{2}{3}(1 - \frac{{}^{(1)}M^*_{n_1}\,{}^{(2)}M^*_{n_1}}{{}^{(3)}M^*_{n_1}})^{-1}.$$

By bootstrapping we can now estimate

$$Q(n_1, k_1) := E(({}^{(M)}\hat{\gamma}^*_{n_1}(k_1) - \hat{\gamma}^*_{n_1}(k_1))^2|\mathcal{X}_n)$$

as well as we wish. The next Theorem connects the minimum of $Q(n_1, k_1)$ with the minimum considered in Theorem 2.2.

Theorem 2.3. *Suppose the conditions of Theorem 2.1 hold and* $n_1 = O(n^{1-\epsilon})$ *for some* $0 < \epsilon < 1$. *The random quantity* $\bar{k}^*_0(n_1)$ *is defined as follows:*

$$\bar{k}^*_0(n_1) := arg \inf_k E(({}^{(M)}\hat{\gamma}^*_{n_1}(k) - \hat{\gamma}^*_{n_1}(k))^2|\mathcal{X}_n).$$

Then

$$\bar{k}^*_0(n_1) \sim n_1(\frac{\bar{V}^2(\gamma)}{\bar{b}^2(\gamma,\rho)})^{\frac{1}{1-2\rho}}(s^-(\frac{1}{n_1}))^{-1}$$

in probability (if x_n, y_n *are two positive random sequences, we say that* $x_n \sim y_n$ *in probability if* $x_n/y_n \to 1$ *in probability).*

We now use the known quantity $\bar{k}^*_0$ to estimate $k_0(n)$ and do this via $\bar{k}_0(n)$.

Corollary 2.1. *Suppose the conditions of Theorem 2.3 hold and* $A(t) = ct^\rho$ *with* $c \neq 0$ *and* $\rho < 0$. *Then*

$$\bar{k}_0(n) \sim \bar{k}^*_0(n_1)(\frac{n}{n_1})^{\frac{-2\rho}{1-2\rho}}$$

in probability.

Next we get rid of the factor $(n/n_1)^{\frac{-2\rho}{1-2\rho}}$. We do this via a second bootstrap procedure.

Theorem 2.4. *Suppose the conditions of Corollary 2.1 hold and* $n_2 = (n_1)^2/n$. *Let*

$$\bar{k}^*_1(n_2) := arg \inf_k E(({}^{(M)}\hat{\gamma}^*_{n_2}(k) - \hat{\gamma}^*_{n_2}(k))^2|\mathcal{X}_n).$$

Then

$$\bar{k}_0(n) \sim (\bar{k}^*_0(n_1))^2/\bar{k}^*_1(n_2) \quad \textit{in probability.}$$

Corollary 2.2. *Under the conditions of Theorem 2.4,*

$$\begin{aligned} k_0(n) &\sim \bar{k}_0(n)\left(\frac{V^2(\gamma)\bar{b}^2(\gamma,\rho)}{\bar{V}^2(\gamma)b^2(\gamma,\rho)}\right)^{\frac{1}{1-2\rho}} \\ &\sim \frac{(\bar{k}_0^*(n_1))^2}{\bar{k}_1^*(n_2)}\left(\frac{V^2(\gamma)\bar{b}^2(\gamma,\rho)}{\bar{V}^2(\gamma)b^2(\gamma,\rho)}\right)^{\frac{1}{1-2\rho}} \quad (n \to \infty). \end{aligned}$$

Corollary 2.3. *Suppose the conditions of Theorem 2.4 hold. Define*

$$\hat{k}_0(n) := \frac{(\bar{k}_0^*(n_1))^2}{\bar{k}_1^*(n_2)}\left(\frac{V^2(\hat{\gamma}_n)\bar{b}^2(\hat{\gamma}_n,\hat{\rho}_n)}{\bar{V}^2(\hat{\gamma}_n)b^2(\hat{\gamma}_n,\hat{\rho}_n)}\right)^{\frac{1}{1-2\hat{\rho}_n}}$$

with $\bar{k}_0^*(n_1)$ and $\bar{k}_1^*(n_2)$ as defined in Theorem 2.3 and Theorem 2.4 respectively and with $\hat{\gamma}_n$ any consistent estimator of γ (for instance ${}^{(M)}\hat{\gamma}_n(k)$ with $k = k(n)$ any sequence with $k \to \infty, k/n \to 0$) and

$$\hat{\rho}_n := \frac{\log \bar{k}_0^*(n_1)}{-2\log n_1 + 2\log \bar{k}_0^*(n_1)}.$$

Then

$$\hat{k}_0(n) \sim k_0(n) \quad \textit{in probability,}$$

hence

$$E({}^{(M)}\hat{\gamma}_n(\hat{k}_0(n)) - \gamma)^2 \sim \inf_k E({}^{(M)}\hat{\gamma}_n(k) - \gamma)^2$$

for $n \to \infty$.

3 Main results for Pickands' estimator

Throughout this section we assume that F is in the differentiable domain of attraction of G_γ (notation: $F \in D_{\text{dif}}(G_\gamma)$), i.e., F is differentiable in a left neighborhood of $x_\infty := \sup\{x : F(x) < 1\}$ and there exist $a_n > 0$ and $b_n \in R$ such that

$$\lim_{n\to\infty} \frac{\partial}{\partial x}[F^n(a_n x + b_n)] = G'_\gamma(x) \tag{3.1}$$

locally uniformly for all $x \in R$. The differentiable domains of attraction were introduced by Pickands (1986). Clearly $F \in D_{\text{dif}}(G_\gamma)$ implies $F \in D(G_\gamma)$ for the same normalizing constants a_n and b_n. Define $U(t) := (1/(1-F))^{-}(t)$. The following proposition characterizes the differentiable domain of attraction of G_γ.

Proposition 1. *$F \in D_{dif}(G_\gamma)$ for some $\gamma \in R$ if and only if $U(t)$ is differentiable for all sufficiently large t and $U'(t) \in RV_{\gamma-1}$.*

Proof. See Pickands (1986). □

In order to get the limit distribution function of estimator ${}^{(P)}\hat{\gamma}_{n,\theta}(k)$ we have to require some kind of second order condition. Because of Proposition 1 it is quite natural to assume that there is a positive function $A^*(t)(\to 0$ as $t \to \infty)$ such that

$$\lim_{t\to\infty} \frac{\frac{U'(tx)}{U'(t)} - x^{\gamma-1}}{A^*(t)}$$

exists for every $x > 0$. In order to avoid trivialities we also assume that the limit function is not a multiple of $x^{\gamma-1}$. Then the limit function must be of the form $c'x^{\gamma-1}\frac{x^\rho-1}{\rho}$ for constants $\rho \le 0$ and $c' \ne 0$ (see Theorem 1.9 of Geluk and de Haan (1987) or Lemma 3.2.1 of Bingham et al. (1987); $(x^0-1)/0$ is defined as $\log x$). We can and will subsume the constant c' in the function A^*. So suppose there is a function A with $lim_{t\to\infty}A(t) = 0$ and not changing sigh near infinity, such that

$$\lim_{t\to\infty} \frac{\frac{U'(tx)}{U'(t)} - x^{\gamma-1}}{A(t)} = x^{\gamma-1}\frac{x^\rho - 1}{\rho} \tag{3.2}$$

for all $x > 0$. The function $|A|$ is then regularly varying with index $\rho(notation:$ $|A| \in RV_\rho)$. It can be proved (see Pereira(1993) or de Haan and Stadtmüller(1996)) that (3.2) is equivalent to

$$\lim_{t\to\infty} \frac{U(tx) - U(t) - tU'(t)\frac{x^\gamma-1}{\gamma}}{tU'(t)A(t)} = h_{\gamma,\rho}(x) := \frac{1}{\rho}[\frac{x^{\gamma+\rho}-1}{\gamma+\rho} - \frac{x^\gamma - 1}{\gamma}]. \tag{3.3}$$

First we determine the theoretically optimal value $k_0(n)$ asymptotically.

Theorem 3.1. *Assume $F \in D_{dif}(G_\gamma)$ and (3.3) holds for $A(t) = ct^{-\rho}$ with $c \ne 0$ and $\rho < 0$. Determine $k_0(n)$ such that $E({}^{(P)}\hat{\gamma}_{n,\theta}(k) - \gamma)^2$ is minimal. Then*

$$k_0(n)/\{(\frac{(\theta^{-1}-1)(1+\theta^{-2\gamma-1})}{-2\rho c^2(\frac{1-\theta^\rho}{\rho})^2(\frac{\theta^{-\gamma-\rho}-1}{\gamma+\rho})^2\theta^{-2\rho}})^{\frac{1}{1-2\rho}} n^{\frac{-2\rho}{1-2\rho}}\} \to 1$$

as $n \to \infty$.

Next we compute the optimum with γ replaced by ${}^{(P)}\hat\gamma_{n,\theta}(k\theta^2)$.

Theorem 3.2. *Assume $F \in D_{dif}(G_\gamma)$ and (3.3) holds for $A(t) = ct^{-\rho}$ with $c \neq 0$ and $\rho < 0$. Determine $\bar k_0(n)$ such that $E({}^{(P)}\hat\gamma_{n,\theta}(k) - {}^{(P)}\hat\gamma_{n,\theta}(k\theta^2))^2$ is minimal. Then*

$$\bar k_0(n)/\{(\frac{(\theta^{-1}-1)(1+\theta^{-2\gamma-1})(1+\theta^{-2})}{-2\rho c^2(\frac{1-\theta^\rho}{\rho})^2(\frac{\theta^{-\gamma-\rho}-1}{\gamma+\rho})^2\theta^{-2\rho}(1-\theta^{-2\rho})^2})^{\frac{1}{1-2\rho}} n^{\frac{-2\rho}{1-2\rho}}\} \to 1$$

as $n \to \infty$.

Corollary 3.1. *Assume $F \in D_{dif}(G_\gamma)$ and (3.3) holds for $A(t) = ct^{-\rho}$ with $c \neq 0$ and $\rho < 0$. Determine $k_0(n)$ such that $E({}^{(P)}\hat\gamma_{n,\theta}(k) - \gamma)^2$ is minimal and $\bar k_0(n)$ such that $E({}^{(P)}\hat\gamma_{n,\theta}(k) - {}^{(P)}\hat\gamma_{n,\theta}(k\theta^2))^2$ is minimal. Then*

$$\frac{\bar k_0(n)}{k_0(n)} \to (\frac{1+\theta^{-2}}{(1-\theta^{-2\rho})^2})^{\frac{1}{1-2\rho}}$$

as $n \to \infty$.

As in Section 2, we draw resamples $\mathcal{X}^*_{n_1} = \{X^*_1, \cdots, X^*_{n_1}\}$ from $\mathcal{X}_n = \{X_1, \cdots, X_n\}$ with replacement. Let $n_1 < n$ and $X^*_{n_1,1} \le \cdots \le X^*_{n_1,n_1}$ denote the order statistics of $\mathcal{X}^*_{n_1}$ and define

$${}^{(P)}\hat\gamma^*_{n_1,\theta}(k_1) := (-\log\theta)^{-1}\log\frac{X^*_{n_1,n_1-[k_1\theta^2]} - X^*_{n_1,n_1-[k_1\theta]}}{X^*_{n_1,n_1-[k_1\theta]} - X^*_{n_1,n_1-k_1}}.$$

Then we propose to use the following bootstrap estimate of the mean square error

$$E(({}^{(P)}\hat\gamma^*_{n_1,\theta}(k_1) - {}^{(P)}\hat\gamma^*_{n_1,\theta}(k_1\theta^2))^2|\mathcal{X}_n).$$

We can prove

Theorem 3.3. *Assume $F \in D_{dif}(G_\gamma)$ and (3.3) holds for $A(t) = ct^{-\rho}$ with $c \neq 0$ and $\rho < 0$. Let $n_1 = O(n^{1-\epsilon})$ for some $\epsilon \in (0,1)$. Determine $k^*_{1,0}(n_1)$ such that $E(({}^{(P)}\hat\gamma^*_{n_1,\theta}(k_1) - {}^{(P)}\hat\gamma^*_{n_1,\theta}(k_1\theta^2))^2|\mathcal{X}_n)$ is minimal. Then*

$$k^*_{1,0}(n_1)/\{(\frac{(\theta^{-1}-1)(1+\theta^{-2\gamma-1})(1+\theta^{-2})}{-2\rho c^2(\frac{1-\theta^\rho}{\rho})^2(\frac{\theta^{-\gamma-\rho}-1}{\gamma+\rho})^2\theta^{-2\rho}(1-\theta^{-2\rho})^2})^{\frac{1}{1-2\rho}} n_1^{\frac{-2\rho}{1-2\rho}}\} \xrightarrow{p} 1$$

as $n \to \infty$.

Finally we connect $k_0(n)$ with $k_{1,0}^*$ and $k_{2,0}^*$ asymptotically.

Theorem 3.4. *Assume $F \in D_{dif}(G_\gamma)$ and (3.3) holds for $A(t)t^{-\rho} \to c$ ($\rho < 0$). Let $n_1 = O(n^{1-\epsilon})$ for some $\epsilon \in (0, 1/2)$ and $n_2 = (n_1)^2/n$. Determine $k_{i,0}^*(n_i)$ such that $E((^{(P)}\hat{\gamma}_{n_i,\theta}^*(k_i) -^{(P)} \hat{\gamma}_{n_i,\theta}^*(k_i\theta^2))^2|\mathcal{X}_n)$ is minimal ($i = 1, 2$). Define $f_\theta(\rho) = (\frac{1+\theta^{-2}}{(1-\theta^{-2\rho})^2})^{\frac{1}{1-2\rho}}$. Then*

$$\frac{(k_{1,0}^*)^2}{k_{2,0}^* f_\theta(\frac{\log k_{1,0}^*}{2(\log k_{1,0}^* - \log n_1)})} / k_0(n) \xrightarrow{p} 1$$

as $n \to \infty$.

So as before we get an estimator for $k_0(n)$ which leads to an estimator for θ which has asymptotically the lowest mean squared error.

4 Proofs

We shall give some lemmas first.

Lemma 1. *Let $Y_1, \cdots, Y_n$ be i.i.d. random variables with common distribution function $1 - x^{-1}$ ($x > 1$) and $Y_{n,1} \le \cdots \le Y_{n,n}$ be the order statistics. Assume $k \to \infty$, $k/n \to 0$. Then*

(i) $Y_{n,n-k}/\frac{n}{k} \to 1$ in probability

(ii) Define

$$\begin{cases} P_n := \frac{1}{k}\sum_{i=1}^k \frac{(Y_{n,n-i+1}/Y_{n,n-k})^{\gamma_-} - 1}{\gamma_-} - \frac{1}{1-\gamma_-} \\ Q_n := \frac{1}{k}\sum_{i=1}^k (\frac{(Y_{n,n-i+1}/Y_{n,n-k})^{\gamma_-} - 1}{\gamma_-})^2 - \frac{2}{(1-\gamma_-)(1-2\gamma_-)} \\ R_n := \frac{1}{k}\sum_{i=1}^k (\frac{(Y_{n,n-i+1}/Y_{n,n-k})^{\gamma_-} - 1}{\gamma_-})^3 - \frac{6}{(1-\gamma_-)(1-2\gamma_-)(1-3\gamma_-)}. \end{cases}$$

We have $\sqrt{k}(P_n, Q_n, R_n)$ converges in distribution to (P, Q, R), say, which is normally distributed with mean vector zero and covariance matrix

$$\begin{cases} EP^2 = \frac{1}{(1-\gamma_-)^2(1-2\gamma_-)} \\ EQ^2 = \frac{4(5-11\gamma_-)}{(1-\gamma_-)^2(1-2\gamma_-)^2(1-3\gamma_-)(1-4\gamma_-)} \\ ER^2 = \frac{36(19-105\gamma_-+146\gamma_-^2)}{(1-\gamma_-)^2(1-2\gamma_-)^2(1-3\gamma_-)^2(1-4\gamma_-)(1-5\gamma_-)(1-6\gamma_-)} \\ E(PQ) = \frac{4}{(1-\gamma_-)^2(1-2\gamma_-)(1-3\gamma_-)} \\ E(PR) = \frac{18}{(1-\gamma_-)^2(1-2\gamma_-)(1-3\gamma_-)(1-4\gamma_-)} \\ E(QR) = \frac{12(9-21\gamma_-)}{(1-\gamma_-)^2(1-2\gamma_-)^2(1-3\gamma_-)(1-4\gamma_-)(1-5\gamma_-)} \end{cases}$$

Proof. Similar to the proof of Lemma 3.1 of Dekkers et al. (1989). □

The following is an extension of a result by Drees (1995).

Lemma 2. *Let f be a measurable function. Suppose there exist a real parameter α and functions $a_1(t) > 0$ and $A_1(t) \to 0$ such that for all $x > 0$*

$$\lim_{t\to\infty} \frac{\frac{f(tx)-f(t)}{a_1(t)} - \frac{x^\alpha - 1}{\alpha}}{A_1(t)} = H_1(x)$$

where

$$H_1(x) = \frac{1}{\beta}[\frac{x^{\alpha+\beta}-1}{\alpha+\beta} - \frac{x^\alpha - 1}{\alpha}] \quad (\beta \le 0).$$

Then for any $\epsilon > 0$ there exists $t_0 > 0$ such that for all $t \ge t_0$, $tx \ge t_0$,

$$|\frac{\frac{f(tx)-f(t)}{a_1(t)} - \frac{x^\alpha - 1}{\alpha}}{A_1(t)} - H_1(x)| \le \epsilon[1 + x^\alpha + 2x^{\alpha+\beta} e^{\epsilon|\log x|}].$$

Proof. Suppose $\alpha \neq 0$. Then from relation (2.2) of Theorem 1 of de Haan and Stadmüller (1996), we have

$$\frac{(tx)^{-\alpha} a_1(tx) - t^{-\alpha} a_1(t)}{t^{-\alpha} a_1(t) A_1(t)} \to \frac{x^\beta - 1}{\beta}.$$

Hence

$$\begin{aligned} & \frac{f(tx)-a_1(tx)/\alpha - (f(t)-a_1(t)/\alpha)}{a_1(t)A_1(t)/\alpha} \\ = & \frac{f(tx)-f(t)-a_1(t)\frac{x^\alpha-1}{\alpha}}{a_1(t)A_1(t)/\alpha} \\ & -x^\alpha \frac{(tx)^{-\alpha}a_1(tx)-t^{-\alpha}a_1(t)}{t^{-\alpha}a_1(t)A_1(t)} \\ \to & \alpha H_1(x) - x^\alpha \frac{x^\beta-1}{\beta} = -\frac{x^{\alpha+\beta}-1}{\alpha+\beta}. \end{aligned}$$

Similar to the proof of Lemma 2.2 of de Haan and Peng (1996), we get

$$\begin{aligned} & |x^\alpha \frac{(tx)^{-\alpha}a_1(tx)-t^{-\alpha}a_1(t)}{t^{-\alpha}a_1(t)A_1(t)/\alpha} - x^\alpha \frac{x^\beta-1}{\beta}| \\ \le & x^\alpha \epsilon[1 + x^\beta e^{\epsilon|\log x|}] \end{aligned}$$

and

$$\begin{aligned} & |\frac{f(tx)-a_1(tx)/\alpha-(f(t)-a_1(t)/\alpha)}{a_1(t)A_1(t)/\alpha} + \frac{x^{\alpha+\beta}-1}{\alpha+\beta}| \\ \le & \epsilon[1 + x^{\alpha+\rho} e^{\epsilon|\log x|}]. \end{aligned}$$

Hence

$$\begin{aligned}
&|\frac{\frac{f(tx)-f(t)}{a_1(t)}-\frac{x^\alpha-1}{\alpha}}{A_1(t)}-H_1(x)|\\
= & |\frac{f(tx)-a_1(tx)/\alpha-(f(t)-a_1(t)/\alpha)}{a_1(t)A_1(t)}+\frac{x^{\alpha+\beta}-1}{\alpha(\alpha+\beta)}\\
&+x^\alpha\frac{(tx)^{-\alpha}a_1(tx)/\alpha-t^{-\alpha}a_1(t)/\alpha}{t^{-\alpha}a_1(t)A_1(t)}-x^\alpha\frac{x^\beta-1}{\alpha\beta}|\\
\leq & \frac{\epsilon}{|\alpha|}[1+x^\alpha+2x^{\alpha+\beta}e^{\epsilon|\log x|}].
\end{aligned}$$

Suppose $\alpha=0$ and $\beta<0$. Then from the proof of Theorem 2 (iii) of de Haan and Stadtmüller (1996) we have $a_1(t)\to c_0\in(-\infty,0)\cup(0,\infty)$ and $\frac{c_0-a_1(t)}{a_1(t)A_1(t)}\to -1/\beta$. Hence

$$\frac{f(tx)-c_0\log(tx)-(f(t)-c_0\log t)}{a_1(t)A(t)}\to\frac{1}{\beta}\frac{x^\beta-1}{\beta}.$$

Similarly as above we know the lemma holds for $\alpha=0$ and $\beta\neq 0$.

Suppose $\alpha=\beta=0$. Write

$$g(t):=f(t)-\frac{1}{t}\int_0^t f(s)\,ds$$

which implies

$$f(t)=g(t)+\int_0^t\frac{g(s)}{s}\,ds$$

(see Corollary 1.2.1 of de Haan (1970)). From Omey and Willekens (1988) we have

$$\frac{g(tx)-g(t)}{a_1(t)A_1(t)}\to\log x.$$

Note that

$$\begin{aligned}
&\frac{f(tx)-f(t)-a_1(t)\log x}{a_1(t)A_1(t)}\\
= & \frac{g(tx)-g(t)}{a_1(t)A_1(t)}+\int_1^x\frac{g(ts)-a_1(t)}{sa_1(t)A_1(t)}\,ds.
\end{aligned}$$

Hence

$$\frac{g(tx)-a_1(t)}{a_1(t)A_1(t)}\to\log x-1.$$

Furthermore

$$\frac{g(t)-a_1(t)}{a_1(t)A_1(t)} \to -1.$$

Using Proposition 1.19.4 of Geluk and de Haan (1987), we can easily see the lemma holds. Thus we complete the proof. $\square$

Let F_n denote the empirical distribution function of $\mathcal{X}_n$ and $U_n = (\frac{1}{1-F_n})^{-}$.

Lemma 3. *If (2.1) and (2.2) hold and $n_1 = O(n^{1-\epsilon_0})$ for some $\epsilon_0 \in (0,1)$. Then for any $0 < \epsilon < 1$ there exists $t_0 > 0$ such that for all $t_0 \le t \le n_1(\log n_1)^2$ and $t_0 \le tx \le n_1(\log n_1)^2$*

$$\begin{aligned} & \left|\frac{\frac{\log U_n(tx)-\log U_n(t)}{a(t)/U(t)} - \frac{x^{\gamma_-}-1}{\gamma_-}}{A(t)} - H(x)\right| \\ \le\ & [\tfrac{\sqrt{tx}\log n}{n} + \epsilon] d(\gamma_-,\rho) x^\rho e^{\epsilon|\log x|} \\ & +[\tfrac{\sqrt{t}\log n}{n} + \epsilon] d(\gamma_-,\rho) \\ & +\epsilon[1 + x^{\gamma_-} + 2x^{\gamma_-+\rho} e^{\epsilon|\log x|}] \\ & +\tfrac{d(\gamma_-,\rho)}{|A(t)|}\tfrac{\log n}{n}[\sqrt{tx}+\sqrt{t}] \end{aligned} \qquad (4.1)$$

where $d(\gamma_-,\rho) > 0$ is a constant which only depends on γ_- and ρ.

Proof. Let G_n denote the empirical distribution function of n independent, uniformly distributed random variables. As n is large enough and $n_1 = O(n^{1-\epsilon_0})$, we have

$$1/2 \le \sup_{t \le n_1(\log n_1)^2} |tG_n^{-}(\frac{1}{t})| \le 2 \quad \text{a.s.} \qquad (4.2)$$

and

$$\sup_{t\ge 2} |\sqrt{t}(G_n(\frac{1}{t}) - \frac{1}{t})| \le \frac{\log n}{\sqrt{n}} \quad \text{a.s.}$$

(see equations (10) and (17) of Chapter 10.5 of Shorack and Wellner (1986)). Hence

$$\sup_{4 \le t \le n_1(\log n_1)^2} \sqrt{\frac{1}{G_n^{-}(\frac{1}{t})}} |G_n(G_n^{-}(\frac{1}{t})) - G_n^{-}(\frac{1}{t})| \le \frac{\log n}{\sqrt{n}} \quad \text{a.s.}$$

Therefore for all $4 \le t \le n_1(\log n_1)^2$

$$|tG_n(\frac{1}{t}) - 1| \le \frac{2\sqrt{t}\log n}{\sqrt{n}} \quad \text{a.s.} \tag{4.3}$$

Now we use Lemma 2, (4.2), (4.3),

$$|y^\gamma - 1| \le |\gamma|(2^{\gamma-1} \vee 2^{\gamma+1})|y-1| \quad \text{for} \quad 1/2 \le y \le 2$$

and $U_n \stackrel{d}{=} U(\frac{t}{tG_n^-(\frac{1}{t})})$. It follows that for any $\epsilon \in (0,1)$ there exists $t_0 > 4$ such that for all $t_0 \le t \le n_1(\log n_1)^2$ and $t_0 \le t \le n_1(\log n_1)^2$

$$\begin{aligned}
& |\frac{\frac{\log U_n(tx) - \log U_n(t)}{a(t)/U(t)} - \frac{x^{\gamma_-}-1}{\gamma_-}}{A(t)} - H(x)| \\
\stackrel{d}{=} \; & |\frac{\log U(\frac{tx}{txG_n^-(\frac{1}{tx})}) - \log U(tx) - \frac{a(tx)}{U(tx)}\frac{(txG_n^-(\frac{1}{tx}))^{-\gamma_-}-1}{\gamma_-}}{A(tx)a(tx)/U(tx)} \frac{A(tx)a(tx)/U(tx)}{A(t)a(t)/U(t)} \\
& - \frac{\log U(\frac{t}{tG_n^-(\frac{1}{t})}) - \log U(t) - \frac{a(t)}{U(t)}\frac{(tG_n^-(\frac{1}{t}))^{-\gamma_-}-1}{\gamma_-}}{A(t)a(t)/U(t)} \\
& + \frac{\log U(tx) - \log U(t) - \frac{a(t)}{U(t)}\frac{x^{\gamma_-}-1}{\gamma_-}}{A(t)a(t)/U(t)} - H(x) \\
& + \frac{\frac{a(tx)}{U(tx)}\frac{(txG_n^-(\frac{1}{tx}))^{-\gamma_-}-1}{\gamma_-}}{A(t)a(t)/U(t)} - \frac{(tG_n^-(\frac{1}{t}))^{-\gamma_-}-1}{\gamma_- A(t)}| \\
\le \; & \{|H(\frac{1}{txG_n^-(\frac{1}{tx})})| + \epsilon[1 + (txG_n^-(\frac{1}{tx}))^{-\gamma_-} \\
& + 2(txG_n^-(\frac{1}{tx}))^{-\gamma_- + \rho} e^{\epsilon|\log(txG_n^-(\frac{1}{tx}))|}\}(1+\epsilon)x^\rho e^{\epsilon|\log x|} \\
& + |H(\frac{1}{tG_n^-(\frac{1}{t})})| + \epsilon[1 + (tG_n^-(\frac{1}{t}))^{-\gamma_-} \\
& + 2(tG_n^-(\frac{1}{t}))^{-\gamma_- + \rho} e^{\epsilon|\log(tG_n^-(\frac{1}{t}))|}] \\
& + \epsilon[1 + x^{\gamma_-} + 2x^{\gamma_- + \rho} e^{\epsilon|\log x|}] \\
& + (1+\epsilon)|\frac{(txG_n^-(\frac{1}{tx}))^{-\gamma_-}-1}{\gamma_- A(t)}| + |\frac{(tG_n^-(\frac{1}{t}))^{-\gamma_-}-1}{\gamma_- A(t)}| \\
\le \; & [d(\gamma_-,\rho)\frac{\sqrt{tx}\log n}{\sqrt{n}} + \epsilon d(\gamma_-,\rho)]x^\rho e^{\epsilon \log x|} \\
& + d(\gamma_-,\rho)\frac{\sqrt{t}\log n}{\sqrt{n}} + \epsilon d(\gamma_-,\rho) \\
& + \epsilon[1 + x^{\gamma_-} + 2x^{\gamma_- + \rho} e^{\epsilon|\log x|}] \\
& + \frac{d(\gamma_-,\rho)}{|A(t)|}\frac{\sqrt{tx}\log n}{\sqrt{n}} + \frac{d(\gamma_-,\rho)}{|A(t)|}\frac{\sqrt{t}\log n}{\sqrt{n}}
\end{aligned}$$

where $d(\gamma_-, \rho) > 0$ is a constant only depending on γ_- and ρ. The lemma follows. □

Proof of Theorem 2.1. A full proof of a somewhat restricted case has been given in Dekkers and de Haan (1993). We shall give a sketch of the proof.

By Lemma 2, for any $\epsilon > 0$ there exists $t_0 > 0$ such that for all $t \geq t_0$, $tx \geq t_0$

$$\left|\frac{\frac{\log U(tx)-\log U(t)}{a(t)/U(t)} - \frac{x^{\gamma_-}-1}{\gamma_-}}{A(t)} - H(x)\right| \leq \epsilon[1 + x^{\gamma_-} + 2x^{\gamma_-+\rho+\rho}e^{\epsilon|\log x|}].$$

Applying this relation with t replaced by $Y_{n,n-k}$ and x by $Y_{n,n-i}/Y_{n,n-k}$, adding the inequalities for $i = 0, 1, \cdots, k-1$ and dividing by k we get

$$\begin{array}{rl} & \frac{M_n^{(1)}}{a(Y_{n,n-k})/U(Y_{n,n-k})} \\ \leq & \frac{1}{1-\gamma_-} + P_n + A(Y_{n,n-k})\frac{1}{k}\sum_{i=1}^k H(Y_{n,n-i+1}/Y_{n,n-k}) \\ & +\epsilon A(Y_{n,n-k})\frac{1}{k}\sum_{i=1}^k\{1 + (Y_{n,n-i+1}/Y_{n,n-k})^{\gamma_-} \\ & +2(Y_{n,n-i+1}/Y_{n,n-k})^{\gamma_-+\rho}e^{\epsilon|\log(Y_{n,n-i+1}/Y_{n,n-k})|}\} \end{array}$$

Note that $\{Y_{n,n-i+1}/Y_{n,n-k}\}_{i=1}^k \stackrel{d}{=} \{Y_i'\}_{i=1}^k$ with $Y_1', \cdots, Y_k'$ i.i.d. with common distribution function $1 - 1/x \quad (x > 1)$. We apply the law of large numbers to the third term. Also note that $\frac{k}{n}Y_{n,n-k} \to 1$ in probability, so that since $|A|$ is regularly varying, we have $(A(n/k))^{-1}A(Y_{n,n-k}) \to 1$ in probability. As a result

$$\frac{M_n^{(1)}}{a(Y_{n,n-k})/U(Y_{n,n-k})} = \frac{1}{1-\gamma_-} + P_n + \frac{A(n/k)}{(1-\gamma_-)(\bar{\rho}-\gamma_-)} + o(A(n/k)).$$

Hence

$$\begin{array}{rl} & \frac{(M_n^{(1)})^2}{a^2(Y_{n,n-k})/U^2(Y_{n,n-k})} \\ = & \frac{1}{(1-\gamma_-)^2} + \frac{2P_n}{1-\gamma_-} + \frac{2A(n/k)}{(1-\gamma_-)^2(\bar{\rho}-\gamma_-)} + o(A(n/k)). \end{array}$$

Similarly

$$\begin{array}{rl} & \frac{M_n^{(2)}}{a^2(Y_{n,n-k})/U^2(Y_{n,n-k})} \\ = & \frac{2}{(1-\gamma_-)(1-2\gamma_-)} + Q_n + \frac{2A(n/k)(2\bar{\rho}+1-4\gamma_-)}{(1-\gamma_-)(1-2\gamma_-)(\bar{\rho}-\gamma_-)(\bar{\rho}-2\gamma_-)} + o(A(n/k)). \end{array}$$

Combining these expansions we get (since $a(Y_{n,n-k})/U(Y_{n,n-k}) \to \gamma_+$ in probability),

$$\begin{array}{rl} & {}^{(M)}\hat{\gamma}_n(k) \\ = & M_n^{(1)} + \frac{M_n^{(2)}-2(M_n^{(1)})^2}{2M_n^{(2)}-2(M_n^{(1)})^2} \\ = & (\gamma_+ + o_p(1))[\frac{1}{1-\gamma_-} + P_n + \frac{A(n/k)}{(1-\gamma_-)(\bar{\rho}-\gamma_-)}] \\ & +\gamma_- + (1-\gamma_-)^2(1-2\gamma_-)[(1/2-\gamma_-)Q_n - 2P_n] \\ & +A(n/k)\frac{(1-\gamma_-)(1-2\gamma_-)}{(\bar{\rho}-\gamma_-)(\bar{\rho}-2\gamma_-)} + o(A(n/k)). \end{array}$$

Consequently, by Lemma 1, we have

$$\begin{aligned} & E({}^{(M)}\hat{\gamma}_n(k)-\gamma)^2 \\ = \; & (V^2(\gamma)/k + b^2(\gamma,\rho)A^2(n/k))(1+o(1)) \\ = \; & (V^2(\gamma)r/n + b^2(\gamma,\rho)A^2(r))(1+o(1)) \end{aligned}$$

with $r := k/n$. One obtains the minimum with respect to r by using (2.5) and equating the derivative to zero (for details see Dekkers and de Haan (1993)). The theorem follows. □

Proof of Theorem 2.3. Given $\mathcal{X}_n := \{X_1, \cdots, X_n\}$, we have

$$^{(1)}M^*_{n_1}(k_1) \stackrel{d}{=} \frac{1}{k_1}\sum_{i=1}^{k_1} \log U_n(Y_{n_1,n_1-i+1}) - \log U_n(Y_{n_1,n_1-k_1})$$

with $\{Y_{n_1,i}\}_{i=1}^{n_1}$ the order statistics from a distribution function $1-1/x \quad (x > 1)$ and independent of $\mathcal{X}_n$. By the same arguments as in the proof of Theorem 2.1 using Lemma 3 instead of Lemma 2 we get

$$\begin{aligned} & \frac{{}^{(1)}M^*_{n_1}(k_1)}{a(Y_{n_1,n_1-k_1})/U(Y_{n_1,n_1-k_1})} \\ = \; & \frac{1}{1-\gamma_-} + P_{n_1} + \frac{A(n_1/k_1)}{(1-\gamma_-)(\bar{\rho}-\gamma_-)} + o(A(n_1/k_1)) + O(\frac{\sqrt{n_1/k_1}\log n}{\sqrt{n}}). \end{aligned}$$

Note that $\frac{\sqrt{n_1/k_1}\log n}{\sqrt{n}} = o(1/\sqrt{k_1})$, so that the last term can be absorbed into the second one. The expansion for ${}^{(1)}M^*_{n_1}(k_1)$ is the same as for $M^{(1)}_{n_1}(k_1)$ given $\mathcal{X}_n$. Similarly for ${}^{(2)}M^*_{n_1}(k_1)$ and ${}^{(3)}M^*_{n_1}(k_1)$. The Theorem follows. □

Proof of Corollary 2.1. Note that $A(t) = ct^\rho$ implies

$$s^-(1/t) = (-2c^2\rho)^{\frac{1}{1-2\rho}} t^{\frac{1}{1-2\rho}}.$$

The Corollary easily follows. □

Proof of Theorem 2.4. This follows by combining the results of Corollary 2.1 for $\bar{k}^*_0(n_1)$ and $\bar{k}^*_1(n_2)$. □

Proof of Corollary 2.3. We only have to prove that $\hat{\rho}_n$ is a consistent estimator of ρ. By Theorem 2.3 the sequence $\bar{k}^*_0(n_1)$ is asymptotic to $c_1 n_1^{\frac{-2\rho}{1-2\rho}}$. Hence

$$\log \bar{k}^*_0(n_1)/\log n_1 \to \frac{-2\rho}{1-2\rho}$$

in probability. This gives the consistency. □

Lemma 4. *If $F \in D_{dif}(G_\gamma)$, then (3.1) holds for $a_n = nU'(n)$ and $b_n = U(n)$ and for any $k \to \infty$, $k/n \to 0$ and $\theta \in (0, 1]$, the stochastic process*

$$W_{n,k}(\theta) := \sqrt{k}\frac{X_{n,n-[k\theta]} - U(\frac{n}{k\theta})}{\frac{n}{k}U'(\frac{n}{k})}$$

converges (in the sense of convergence of all finite marginal distributions) to a gaussian process $w(\theta)$ which has mean zero and covariance structure

$$Cov(w(\theta_1), w(\theta_2)) = \theta_1^{-\gamma}\theta_2^{-\gamma-1}, \quad 0 < \theta_1 \le \theta_2 \le 1.$$

Proof. See Theorem 2.3 of Cooil (1985). □

Lemma 5. *If (3.3) holds and $n_1 = O(n^{1-\epsilon_0})$ for some $\epsilon_0 \in (0, 1)$. Then for any $0 < \epsilon < 1$ there exists $t_0 > 0$ such that for all $t_0 \le t \le n_1(\log n_1)^2$ and $t_0 \le tx \le n_1(\log n_1)^2$*

$$\begin{aligned} & |\frac{\frac{U_n(tx)-U_n(t)}{a(t)} - \frac{x^\gamma - 1}{\gamma}}{A(t)} - h_{\gamma,\rho}(x)| \\ \le \ & [\frac{\sqrt{tx}\log n}{n} + \epsilon]D(\gamma,\rho)x^{\gamma+\rho}e^{\epsilon|\log x|} \\ & +[\frac{\sqrt{t}\log n}{n} + \epsilon]D(\gamma,\rho) \\ & +\epsilon[1 + x^\gamma + 2x^{\gamma+\rho}e^{\epsilon|\log x|}] \\ & +\frac{D(\gamma,\rho)}{|A(t)|}\frac{\sqrt{t}\log n}{n}[\sqrt{x} + 1] \end{aligned} \tag{4.4}$$

where $D(\gamma, \rho) > 0$ is a constant which only depends on γ and ρ.

Proof. Similar to the proof of Lemma 3. □

Proof of Theorem 3.1. By Lemma 4 we have

$$
\begin{aligned}
&\sqrt{k}({}^{(P)}\hat{\gamma}_{n,\theta}(k)-\gamma)\\
= \;& \sqrt{k}(\tfrac{1}{-\log\theta}\log\tfrac{X_{n,n-[k\theta^2]}-X_{n,n-[k\theta]}}{X_{n,n-[k\theta]}-X_{n,n-k}}-\gamma)\\
= \;& \tfrac{\sqrt{k}}{-\log\theta}\log(1+\theta^{\gamma}\tfrac{X_{n,n-[k\theta^2]}-X_{n,n-[k\theta]}}{X_{n,n-[k\theta]}-X_{n,n-k}}-1)\\
\stackrel{d}{=}\;& \tfrac{\sqrt{k}}{-\log\theta}\tfrac{X_{n,n-[k\theta^2]}-X_{n,n-[k\theta]}-\theta^{-\gamma}(X_{n,n-[k\theta]}-X_{n,n-k})}{\theta^{-\gamma}(X_{n,n-[k\theta]}-X_{n,n-k})}(1+o(1))\\
= \;& [\tfrac{\sqrt{k}}{-\log\theta}\tfrac{X_{n,n-[k\theta]}-U(\frac{n}{k\theta^2})-(1+\theta^{-\gamma})(X_{n,n-[k\theta]}-U(\frac{n}{k\theta}))+\theta^{-\gamma}(X_{n,n-k}-U(\frac{n}{k}))}{\theta^{-\gamma}(X_{n,n-[k\theta]}-X_{n,n-k})}\\
&+\tfrac{\sqrt{k}}{-\log\theta}\tfrac{U(\frac{n}{k\theta^2})-(1+\theta^{-\gamma})U(\frac{n}{k\theta})+\theta^{-\gamma}U(\frac{n}{k})}{\theta^{-\gamma}(X_{n,n-[k\theta]}-X_{n,n-k})}](1+o(1))\\
&(\text{ note }\tfrac{X_{n,n-[k\theta]}-X_{n,n-k}}{\frac{n}{k}U'(\frac{n}{k})}\stackrel{p}{\to}\tfrac{\theta^{-\gamma}-1}{\gamma})\\
\stackrel{d}{=}\;& [\tfrac{\sqrt{k}}{-\log\theta}\tfrac{X_{n,n-[k\theta^2]}-U(\frac{n}{k\theta^2})-(1+\theta^{-\gamma})(X_{n,n-[k\theta]}-U(\frac{n}{k\theta}))+\theta^{-\gamma}(X_{n,n-k}-U(\frac{n}{k}))}{\frac{n}{k}U'(\frac{n}{k})\theta^{-\gamma}\frac{\theta^{-\gamma}-1}{\gamma}}\\
&+\tfrac{\sqrt{k}}{-\log\theta}\tfrac{U(\frac{n}{k\theta^2})-(1+\theta^{-\gamma})U(\frac{n}{k\theta})+\theta^{-\gamma}U(\frac{n}{k})}{\frac{n}{k}U'(\frac{n}{k})\theta^{-\gamma}\frac{\theta^{-\gamma}-1}{\gamma}}](1+o(1))\\
\stackrel{d}{=}\;& \tfrac{1}{-\log\theta}\tfrac{1}{\theta^{-\gamma}\frac{\theta^{-\gamma}-1}{\gamma}}(w(\theta^2)-(1+\theta^{-\gamma})w(\theta)+\theta^{-\gamma}w(1))+o(1)\\
&+\tfrac{\sqrt{k}}{-\log\theta}\tfrac{U(\frac{n}{k\theta^2})-(1+\theta^{-\gamma})U(\frac{n}{k\theta})+\theta^{-\gamma}U(\frac{n}{k})}{\frac{n}{k}U'(\frac{n}{k})\theta^{-\gamma}\frac{\theta^{-\gamma}-1}{\gamma}}(1+o(1)),
\end{aligned}
$$

thus the variance of $\sqrt{k}({}^{(P)}\hat{\gamma}_{n,\theta}(k)-\gamma)$ asymptotically equals to

$$\frac{\gamma^2(\theta^{-1}-1)(1+\theta^{-2\gamma-1})}{(\log\theta)^2(\theta^{-\gamma}-1)^2}$$

and the bias of $\sqrt{k}({}^{(P)}\hat{\gamma}_{n,\theta}(k)-\gamma)$ asymptotically equals to

$$\sqrt{k}A(\frac{n}{k})\frac{\theta^{-\rho}}{-\log\theta}\frac{\gamma}{\theta^{-\gamma}-1}\frac{1-\theta^{\rho}}{\rho}\frac{\theta^{-\gamma-\rho}-1}{\gamma+\rho}.$$

By $A(t)=ct^{-\rho}$ we get in a way similar to the proof of Theorem 2.1

$$k_0(n)/\{(\frac{(\theta^{-1}-1)(1+\theta^{-2\gamma-1})}{-2\rho c^2(\frac{1-\theta^{\rho}}{\rho})^2(\frac{\theta^{-\gamma-\rho}-1}{\gamma+\rho})^2\theta^{-2\rho}})^{\frac{1}{1-2\rho}}n^{\frac{-2\rho}{1-2\rho}}\}\to 1.$$

□

Proof of Theorem 3.2. By Lemma 4 we have

$$
\begin{aligned}
& \sqrt{k}(^{(P)}\hat{\gamma}_{n,\theta}(k) -^{(P)} \hat{\gamma}_{n,\theta}(k\theta^2)) \\
= & \sqrt{k}(^{(P)}\hat{\gamma}_{n,\theta}(k) - \gamma) - \sqrt{k}(^{(P)}\hat{\gamma}_{n,\theta}(k\theta^2) - \gamma) \\
\stackrel{d}{=} & [\frac{\sqrt{k}}{-\log\theta}\frac{X_{n,n-[m\theta^2]}-U(\frac{n}{k\theta^2})-(1+\theta^{-\gamma})(X_{n,n-[k\theta]}-U(\frac{n}{k\theta}))+\theta^{-\gamma}(X_{n,n-k}-U(\frac{n}{k}))}{\theta^{-\gamma}(X_{n,-[k\theta]}-X_{n,n-k})} \\
& +\frac{\sqrt{k}}{\log\theta}\frac{U(\frac{n}{k\theta^2})-(1+\theta^{-\gamma})U(\frac{n}{k\theta})+\theta^{-\gamma}U(\frac{n}{k})}{\theta^{-\gamma}(X_{n,n-[k\theta]}-X_{n,n-k})} \\
& -\frac{\sqrt{k}}{-\log\theta}\frac{X_{n,n-[k\theta^4]}-U(\frac{n}{k\theta^4})-(1+\theta^{-\gamma})(X_{n,n-[k\theta^3]}-U(\frac{n}{k\theta^3}))+\theta^{-\gamma}(X_{n,n-[k\theta^2]}-U(\frac{n}{k\theta^2}))}{\theta^{-\gamma}(X_{n,n-[k\theta^3]}-X_{n,n-[k\theta^2]})} \\
& -\frac{\sqrt{k}}{-\log\theta}\frac{U(\frac{n}{k\theta^3})-(1+\theta^{-\gamma})U(\frac{n}{k\theta^2})+\theta^{-\gamma}U(\frac{n}{k\theta^2})}{\theta^{-\gamma}(X_{n,n-[k\theta^3]}-X_{n,n-[k\theta^2]})}](1+o(1)) \\
& (\text{ note } \frac{X_{n,n-[k\theta^3]}-X_{n,n-[k\theta^2]}}{\frac{n}{k}U'(\frac{n}{k})} \stackrel{p}{\to} \theta^{-2\gamma}\frac{\theta^{-\gamma}-1}{\gamma}) \\
\stackrel{d}{=} & \frac{1}{-\log\theta}\frac{1}{\theta^{-\gamma}\frac{\theta^{-\gamma}-1}{\gamma}}(w(\theta^2)-(1+\theta^{-\gamma})w(\theta)+\theta^{-\gamma}w(1))+o(1) \\
& +\frac{\sqrt{k}}{-\log\theta}\frac{U(\frac{n}{k\theta^2})-(1+\theta^{-\gamma})U(\frac{n}{k\theta})+\theta^{-\gamma}U(\frac{n}{k})}{\frac{n}{k}U'(\frac{n}{k})\theta^{-\gamma}\frac{\theta^{-\gamma}-1}{\gamma}}(1+o(1)) \\
& -\frac{1}{-\log\theta}\frac{1}{\theta^{-3\gamma}\frac{\theta^{-\gamma}-1}{\gamma}}(w(\theta^4)-(1+\theta^{-\gamma})w(\theta^3)+\theta^{-\gamma}w(\theta^2))+o(1) \\
& -\frac{\sqrt{k}}{-\log\theta}\frac{U(\frac{n}{k\theta^4})-(1+\theta^{-\gamma})U(\frac{n}{k\theta^3})+\theta^{-\gamma}U(\frac{n}{k\theta^2})}{\frac{n}{k}U'(\frac{n}{k})\theta^{-3\gamma}\frac{\theta^{-\gamma}-1}{\gamma}}(1+o(1)),
\end{aligned}
$$

thus the variance of $\sqrt{k}(^{(P)}\hat{\gamma}_{n,\theta}(k) -^{(P)} \hat{\gamma}_{n,\theta}(k\theta^2))$ asymptotically equals to

$$\frac{\gamma^2(1+\theta^{-2\gamma-1})(\theta^{-1}-1)(1+\theta^{-2})}{(\log\theta)^2(\theta^{-\gamma}-1)^2}$$

and the bias of $\sqrt{k}(^{(P)}\hat{\gamma}_{n,\theta}(k) -^{(P)} \hat{\gamma}_{n,\theta}(k\theta^2))$ asymptotically equals to

$$\sqrt{k}A(\frac{n}{k})\frac{\theta^{-\rho}}{-\log\theta}\frac{1-\theta^{\rho}}{\rho}\frac{\theta^{-\gamma-\rho}-1}{\gamma+\rho}\frac{\gamma}{\theta^{-\gamma}-1}(1-\theta^{-2\rho}).$$

By $A(t) = ct^{-\rho}$ we get in a way similar to the proof of Theorem 2.1

$$\bar{k}_0(n)/\{(\frac{(\theta^{-1}-1)(1+\theta^{-2\gamma-1})(1+\theta^{-2})}{-2\rho c^2(\frac{1-\theta^{\rho}}{\rho})^2(\frac{\theta^{-\gamma-\rho}-1}{\gamma+\rho})^2\theta^{-2\rho}(1-\theta^{-2\rho})^2})^{\frac{1}{1-2\rho}}n^{\frac{-2\rho}{1-2\rho}}\} \to 1.$$

□

Proof of Theorem 3.3. Similar to the proof of Theorem 2.3 by using Lemma 5 instead of Lemma 3. □

Proof of Theorem 3.4. Similar to the proof of Corollary 2.3. □

Appendix

The following theorem and remark explain why we use different second order conditions in section 2 and section 3.

Theorem A. *Assume* $U(\infty) > 0$ *and there exist functions* $a(t) > 0$ *and* $A(t) \to 0$ *such that*

$$\frac{\frac{U(tx)-U(t)}{a(t)} - \frac{x^\gamma - 1}{\gamma}}{A(t)} \to H(x)$$

where

$$H(x) = \frac{1}{\rho}\Big[\frac{x^{\gamma+\rho}-1}{\gamma+\rho} - \frac{x^\gamma - 1}{\gamma}\Big] \quad (\rho \le 0).$$

Suppose that $\gamma \neq \rho$*. Then*

$$\lim_{t\to\infty} \frac{\frac{a(t)}{U(t)} - \gamma_+}{A(t)} = c \in [-\infty, \infty]$$

where

$$c = \begin{cases} \pm\infty & \text{if } 0 < \gamma \le -\rho \quad \text{or } \rho < \gamma \le 0 \\ \text{finite} & \text{otherwise .} \end{cases}$$

Furthermore

$$\frac{\frac{\log U(tx) - \log U(t)}{a(t)/U(t)} - \frac{x^{\gamma_-} - 1}{\gamma_-}}{\tilde{A}(t)} \to \tilde{H}(x)$$

where

$$\tilde{A}(t) = \begin{cases} A(t) & \text{if } c \in (-\infty, \infty) \\ \frac{a(t)}{U(t)} - \gamma_+ & \text{otherwise ,} \end{cases}$$

$\tilde{A}(t) \in RV_{\rho'}$,

$$\rho' = \begin{cases} -\gamma & \text{if } 0 < \gamma \le -\rho \\ \gamma & \text{if } \rho < \gamma \le 0 \\ \rho & \text{otherwise} \end{cases}$$

and

$$\tilde{H}(x) = \begin{cases} H(x) - \frac{c}{2}(\frac{x^{\gamma}-1}{\gamma})^2 & \text{if } \gamma \leq 0, c \in (-\infty, \infty) \\ -\frac{1}{2}(\frac{x^{\gamma}-1}{\gamma})^2 & \text{if } \gamma \leq 0, c = \pm\infty \\ x^{-\gamma}H(x) - \frac{c}{\gamma}(\log x + \frac{x^{-\gamma}-1}{\gamma}) & \text{if } \gamma > 0, c \in (-\infty, \infty) \\ -\frac{1}{\gamma}(\log x + \frac{x^{-\gamma}-1}{\gamma}) & \text{if } \gamma > 0, c = \pm\infty. \end{cases}$$

Proof. Suppose that $\gamma \neq 0$. Then from the proof of Lemma 2 we have

$$\frac{U(tx) - a(tx)/\gamma - (U(t) - a(t)/\gamma)}{a(t)A(t)} \to \frac{1}{\gamma}\frac{x^{\gamma+\rho} - 1}{\gamma + \rho}.$$

If $\gamma + \rho > 0$, then

$$\frac{U(t) - a(t)/\gamma}{a(t)A(t)} \to \frac{1}{\gamma(\gamma + \rho)}.$$

Hence

$$\frac{a(t)/U(t) - \gamma}{A(t)} = \frac{a(t)\gamma}{U(t)}\frac{a(t)/\gamma - U(t)}{a(t)A(t)} \to \gamma/(\gamma + \rho).$$

If $\gamma + \rho = 0$, i.e., $\gamma = -\rho > 0$, then

$$\frac{U(t) - a(t)/\gamma}{a(t)A(t)} \to \pm\infty.$$

Hence

$$\frac{a(t)/U(t) - \gamma}{A(t)} = \frac{a(t)\gamma}{U(t)}\frac{a(t)/\gamma - U(t)}{a(t)A(t)} \to \pm\infty.$$

If $\gamma + \rho < 0$, then

$$\begin{cases} U(t) - a(t)/\gamma \to c_0 \in (-\infty, 0) \cup (0, \infty) \\ \frac{U(t) - a(t)/\gamma - c_0}{a(t)A(t)} \to \frac{1}{\gamma(\gamma+\rho)}. \end{cases}$$

For $\gamma > 0$, we have

$$\frac{a(t)/U(t) - \gamma}{A(t)} = \frac{a(t)\gamma}{U(t)}(\frac{a(t)/\gamma - U(t) + c_0}{a(t)A(t)} - \frac{c_0}{a(t)A(t)}) \to \pm\infty.$$

For $\gamma < 0$, we $|a(t)/(U(t)A(t))| \in RV_{\gamma-\rho}$. Hence

$$a(t)/(U(t)A(t)) \to \begin{cases} \pm\infty & \text{if } \gamma-\rho > 0 \\ 0 & \text{if } \gamma-\rho < 0. \end{cases}$$

Suppose that $\gamma = 0$ and $\rho < 0$. Then from the proof of Lemma 2 $a(t) \to c_1 \in (-\infty, 0) \cup (0, \infty)$. Hence

$$a(t)/(U(t)A(t)) \sim c_1/(U(\infty)A(t)) \to \pm\infty.$$

We have now proved that the first part of the theorem.

Note that $a(t)/U(t) \to \gamma_+$. For $\gamma \leq 0$, we have

$$\begin{aligned} & \log \tfrac{U(tx)}{U(t)} \\ = \; & \tfrac{a(t)}{U(t)}[\tfrac{x^{\gamma_-}-1}{\gamma_-} + A(t)H(x) + o(A(t))] \\ & +(\tfrac{a(t)}{U(t)})^2[\tfrac{x^{\gamma_-}-1}{\gamma_-} + A(t)H(x) + o(A(t))]^2 + o((\tfrac{a(t)}{U(t)})^2), \end{aligned}$$

i.e.,

$$\begin{aligned} & \tfrac{\log U(tx) - \log U(t)}{a(t)/U(t)} - \tfrac{x^{\gamma_-}-1}{\gamma_-} \\ = \; & A(t)H(x) + o(A(t)) \\ & -\tfrac{a(t)}{2U(t)}[(\tfrac{x^{\gamma_-}-1}{\gamma_-})^2 + 2\tfrac{x^{\gamma_-}-1}{\gamma_-}A(t)H(x) + o(A(t))] + o(\tfrac{a(t)}{U(t)}). \end{aligned}$$

For $\gamma > 0$, we have

$$\begin{aligned} & x^{-\gamma}\tfrac{U(tx)}{U(t)} \\ = \; & x^{-\gamma} + \tfrac{a(t)}{U(t)}\tfrac{1-x^{-\gamma}}{\gamma} + x^{-\gamma}\tfrac{a(t)}{U(t)}[A(t)H(x) + o(A(t))] \\ = \; & 1 + (x^{-\gamma}-1)(1 - \tfrac{a(t)}{\gamma U(t)}) + x^{-\gamma}\tfrac{a(t)}{U(t)}[A(t)H(x) + o(A(t))], \end{aligned}$$

i.e.,

$$\begin{aligned} & \tfrac{\log U(tx) - \log U(t)}{a(t)/U(t)} - \log x \\ = \; & -(\log x + \tfrac{x^{-\gamma}-1}{\gamma})\tfrac{a(t)}{U(t)}(a(t)/U(t) - \gamma) \\ & +x^{-\gamma}A(t)H(x) + o(A(t)) + o(a(t)/U(t) - 1). \end{aligned}$$

So the second part of the theorem follows easily. □

Remark A. *It is not true that a second order condition for U always implies a second order condition for $\log U$: Let $\gamma = \rho$ and define*

$$U'(t) = t^{\gamma-1} \exp\{\int_1^t s^{\gamma-1}(2 + \sin(\log\log s))\, ds\}.$$

From the representation (2.5) of de Haan and Resnick (1996) we find

$$\frac{\frac{U(tx)-U(t)}{tU'(t)} - \frac{x^\gamma - 1}{\gamma}}{t^\gamma[2 + \sin(\log\log t)]} \to \int_1^x u^{\gamma-1} \frac{u^{-\gamma} - 1}{-\gamma}\, du.$$

Hence

$$\begin{aligned} & a(t)/(U(t)A(t)) \\ = \; & \frac{tU'(t)}{U(t)t^\gamma[2+\sin(\log\log t)]} \\ \sim \; & \frac{\exp\{\int_1^\infty s^{\gamma-1}[2+\sin(\log\log s)]\, ds}{U(\infty)[2+\sin(\log\log t)]} \end{aligned}$$

which does not have a limit.

References

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

[2] B. Cooil (1985). Limiting multivariate distributions of intermediate order statistics. *Ann. Probab. 13, 469-477.*

[3] J. Danielsson, L. de Haan, L. Peng and C.G. de Vries (1996). Using a bootstrap method to choose the sample fraction in the tail index estimation. *Submitted.*

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

[5] A.L.M. Dekkers and L. de Haan (1993). Optimal choice of sample fraction in extreme value estimation. *Journal of Multivariate Analysis 47(2), 173-195.*

[6] J. Geluk and L. de Haan (1987). Regular Variation, Extensions and Tauberian Theorems. *CWI Tract 40, Amsterdam.*

[7] L. de Haan (1970). On Regular Variation and its Application to the Weak Convergence of Sample Extremes. *CWI Tract 32, Amsterdam.*

[8] L. de Haan and L. Peng (1997). Rates of convergence for bivariate extremes. *Journal of Multivariate Analysis. 61(2), 195-230.*

[9] L. de Haan and S.I. Resnick (1996). Second order regular variation and rates of convergence in extreme value theory. *Annals of Probability 24, 97-124.*

[10] L. de Haan and U. Stadtmüller (1996). Generalized regular variation of second order. *J. Austral. Math. Soc. (Series A) 61, 381-395.*

[11] H. Drees (1995). On smooth statistical tail functionals. *Technical Report, University of Cologne.*

[12] H. Drees and E. Kaufmann (1997). Selecting the optimal sample fraction in univariate extreme value estimation. *Technical Report, University of Cologne.*

[13] E. Omey and E. Willekens (1988). π-variation with remainder. *J. London Math. Soc. 37, 105-118.*

[14] T.T. Pereira (1993). Second order behaviour of domains of attraction and the bias of generalized pickands' estimator.

[15] J. Pickands III(1975). Statistical inference using extreme order statistics. *Ann. Statis. 3, 119-131.*

[16] J. Pickands III (1986). The continuous and differentiable domains of attraction of the extreme value distributions. *Ann. Probab. 14, 996-1004.*

[17] G. Shorack and J. Wellner (1986). Empirical Processes with Applications to Statistics. *John Wiley & Sons.*