An adaptive optimal estimate of the tail index for MA(l) time series

J.L. Geluk *   Liang Peng †

Econometric Institute [1] Report EI-9910-A

Abstract

For samples of random variables with a regularly varying tail estimating the tail index has received much attention recently. For the proof of asymptotic normality of the tail index estimator second order regular variation is needed.

In this paper we first supplement earlier results on convolution given by Geluk et al. (1997). Secondly we propose a simple estimator of the tail index for finite moving average time series. We also give a subsampling procedure in order to estimate the optimal sample fraction in the sense of minimal mean squared error.

Key words & phrases: regular variation, tail index.

*email: jgeluk@few.eur.nl

†email: liang.peng@wintermute.anu.edu.au

[1] Erasmus University Rotterdam, P.O. Box 1738, NL-3000 DR Rotterdam, The Netherlands

1 Introduction

Assume $Z_1, Z_2, \cdots$ are independent non-negative random variables whose common distribution function F_0 has the regularly varying tail with index $\alpha > 0$, i.e.,

$$\lim_{t\to\infty} \frac{1-F_0(tx)}{1-F_0(t)} = x^{-\alpha} \text{ for all } x > 0. \tag{1}$$

For the estimation of the index α many estimators have been proposed. One of the well known estimator is Hill's estimator(Hill(1975)). It is reasonable to notice that all those estimators can only involve a part of upper order statistics. Several procedures of choosing the sample fraction can be found in Hall (1990), Danielsson et al. (1997) and Drees and Kaufmann (1997). For a general overview of regular variation theory the reader is referred to Bingham et al.(1987).

Now suppose our observations are the finite moving average time series

$$X_i = \sum_{j=1}^{l} c_j Z_{i+j-1}, \quad i = 1, 2, \cdots, \tag{2}$$

where $c_j > 0, j = 1, \cdots, l$. From Feller (1971) it follows that the $X_i's$ have an identical marginal distribution function F which satisfies

$$\lim_{t\to\infty} \frac{1-F(tx)}{1-F(t)} = x^{-\alpha} \text{ for all } x > 0. \tag{3}$$

Estimating the tail index for dependent data has received much attention recently, especially Hill's estimator has been studied (see Hsing(1991), Rootzen et al. (1997), Resnick and Starica (1995, 1996)). For the proof of asymptotic normality second order regular variation for X_i is needed. From Geluk et al. (1997) we note that the second order behaviour of X_i is different from that of Z_i. Although all the results mentioned in above references are valid for our simple model, i.e. finite moving average time series, no result concerning the choice of the sample fraction for dependent data exists in the literature up to now.

In order to obtain asymptotic normality we assume a stronger version of (**??**) namely

$$1 - F_0(x) = a_0 x^{-\alpha}[1 + b_0 x^{-\beta_0} + o(x^{-\beta})], \text{ as } x \to \infty, \tag{4}$$

where $a_0 > 0, \alpha > 0, \beta_0 > 0$ and $b_0 \neq 0$.

This is a special case of so called second order regular variation where we have

$$\lim_{t\to\infty} \frac{\frac{1-F_0(tx)}{1-F_0(t)} - x^{-\alpha}}{a(t)} = x^{-\alpha}\frac{x^{-\beta}-1}{-\beta} \tag{5}$$

for $x > 0$, where a is a function satisfying $a(t) \to 0$ $(t \to \infty)$.

In Section 2 we study the asymptotic behavior of the tail of the convolution for independent positive random variables with second order regularly varying tails in case $\alpha = 1$ and possibly different $\beta_i \geq 0$ $(i = 1, 2)$,where the mean does not exist. This result supplements earlier results on convolutions given in Geluk et al. (1997). As a consequence of the present result (Theorem 1) it follows that under condition (**??**) $X_1, \cdots, X_n$ have an identical marginal distribution function F which satisfies

$$1 - F(x) = ax^{-\alpha}[1 + bx^{-\beta} + o(x^{-\beta})] \text{ as } x \to \infty, \tag{6}$$

where $a > 0, \beta > 0$ and $b \neq 0$ are functions of $a_0, \alpha, \beta_0, b_0$ and $c_i, i = 1, \cdots, l$, unless $\alpha = 1, \beta_0 \geq 1$. See Corollary 1 below.

Let $X_{n,1} \leq \cdots \leq X_{n,n}$ be the order statistics of $X_1, \cdots, X_n$. In Section 3 we prove (for optimal k)the asymptotic normality of the simple tail index estimator

$$\hat{\alpha}_n(k) := 2(\log 2)(\log \frac{X_{n,n-2k}}{X_{n,n-4k}})^{-1} - (\log 2)(\log \frac{X_{n,n-k}}{X_{n,n-2k}})^{-1}, \tag{7}$$

and we provide a procedure to achieve the optimal choice of k, in the sense of minimal mean squared error.

One of the reasons why we choose the estimator $\hat{\alpha}_n(k)$ instead of other estimators, for example, the simple estimator

$$\bar{\alpha}_n(k) := (\log 2)(\log \frac{X_{n,n-k}}{X_{n,n-2k}})^{-1},$$

or Hill's estimator (Hill (1975)) , etc., is due to the fact that we don't have to estimate the c_i in (**??**) in our bootstrap procedure. The other reason is that among the commonly used Hill's estimator, Pickands' estimator (Pickands (1975)) and the moment estimator (Dekkers et al. (1989)), with optimal sample fraction (in the sense of minimal mean squared error) none of the estimators is better than the others for all values of the parameters. See also de Haan and Peng (1998).

2 Second order behaviour for convolution

Theorem 1 *Let Y_1 and Y_2 be independent non-negative random variables with distribution functions G_1 and G_2, respectively. Assume there exist functions $a_i(t) \to 0$ as $t \to \infty$ such that*

$$\lim_{t\to\infty} \frac{\frac{1-G_i(tx)}{1-G_i(t)} - x^{-1}}{a_i(t)} = x^{-1}\frac{x^{-\beta_i} - 1}{-\beta_i} \tag{8}$$

for $x > 0$, where $\beta_i \geq 0,\ i = 1, 2$.

In case $\beta_i = 0$ read $\frac{x^{-\beta_i}-1}{-\beta_i} = \log x$.

Then as $x \to \infty$

$$\begin{aligned} P(Y_1+Y_2>x) &= P(Y_1>x)+P(Y_2>x)+ \\ &+ (1+o(1))\frac{P(Y_1>x)}{x}\int_0^x P(Y_2>y)dy+ \\ &+ (1+o(1))\frac{P(Y_2>x)}{x}\int_0^x P(Y_1>y)dy+ \\ &+ o((a_1(x)P(Y_1>x))\wedge(a_2(x)P(Y_2>x))). \end{aligned} \tag{9}$$

The following result (see de Haan, Pereira (1998)) is needed for the proof of the theorem.

Lemma 1 *Let f be a measurable function and for some function $a_1(t)>0$ we have*

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

for all $x>0$ where γ is a real parameter. Then there exists a positive function a with $a(t)\sim a_1(t)$ $(t\to\infty)$ with the property that for every $\varepsilon,\varepsilon'>0$ there exists a t_0 such that for $t\geq t_0, tx\geq t_0$

$$x^{-\gamma}e^{-\varepsilon'|\log x|}\left|\frac{f(tx)-f(t)}{a(t)}-\frac{x^\gamma-1}{\gamma}\right|<\varepsilon.$$

Proof of Theorem 1 Define $\overline{G_1}(x):=P(Y_1>x)$ and $\overline{G_2}(x):=P(Y_2>x)$.
Note that

$$\begin{aligned} P(Y_1+Y_2 &> x)-P(Y_1>x)-P(Y_2>x)+P(Y_1>x)P(Y_2>x)= \\ \textstyle\int_0^{x/2} &(\overline{G_1}(x-y)-\overline{G_1}(x))dG_2(y)+\int_0^{x/2}(\overline{G_2}(x-y)-\overline{G_2}(x))dG_1(y)+ \\ &+(\overline{G_1}(x/2)-\overline{G_1}(x))(\overline{G_2}(x/2)-\overline{G_2}(x)). \end{aligned} \tag{10}$$

It follows that

$$\begin{aligned} P(Y_1 &+ Y_2>x)-P(Y_1>x)-P(Y_2>x)+P(Y_1>x)P(Y_2>x)= \\ &\overline{G_1}(x)\int_0^{x/2}((1-\frac{y}{x})^{-1}-1)dG_2(y)+\overline{G_2}(x)\int_0^{x/2}((1-\frac{y}{x})^{-1}-1)dG_1(y)+ \\ &+ \overline{G_1}(x)\int_0^{x/2}\left(\frac{\overline{G_1}(x-y)}{\overline{G_1}(x)}-(1-\frac{y}{x})^{-1}\right)dG_2(y)+ \\ &+ \overline{G_2}(x)\int_0^{x/2}\left(\frac{\overline{G_2}(x-y)}{\overline{G_2}(x)}-(1-\frac{y}{x})^{-1}\right)dG_1(y)+ \\ &+ (\overline{G_1}(x/2)-\overline{G_1}(x))(\overline{G_2}(x/2)-\overline{G_2}(x))=: \\ &I_1+I_2+I_3+I_4+I_5. \end{aligned} \tag{11}$$

Note that

$$\begin{aligned}\frac{I_1}{\overline{G_1}(x)} &= \int_0^{x/2}((1-\frac{y}{x})^{-1}-1)dG_2(y) = \\ &= \frac{1}{x}\int_0^{x/2} y dG_2(y) + \frac{1}{x}\int_0^{x/2}\frac{y^2}{x-y}dG_2(y) = J_1 + J_2.\end{aligned}$$

Since $\overline{G_2}$ is regularly varying with exponent -1 it follows that $J_1 \sim \frac{1}{x}\int_0^{x/2}\overline{G_2}(y)dy$. Moreover $\frac{1}{x^2}\int_0^{x/2} y^2 dG_2(y) \leq J_2 \leq 2\frac{1}{x^2}\int_0^{x/2} y^2 dG_2(y)$ and $\int_0^{x/2} y^2 dG_2(y) \sim \frac{1}{2}x^2\overline{G_2}(x)$ as $x \to \infty$. Regular variation of $\overline{G_2}$ now gives $J_2 = o(J_1)$, hence $I_1 \sim \frac{\overline{G_1}(x)}{x}\int_0^{x/2}\overline{G_2}(y)dy$. I_2 can be treated similarly.

Next we estimate I_3. Note that

$$\begin{aligned}&| \int_0^{x/2}\frac{\frac{\overline{G_1}(x-y)}{\overline{G_1}(x)} - (1-\frac{y}{x})^{-1}}{a_1(x)}dG_2(y)| \leq \\ \leq\ & |\int_0^{x/2}\left(\frac{x(1-\frac{y}{x})\overline{G_1}(x(1-\frac{y}{x})) - x\overline{G_1}(x)}{x a_1(x)\overline{G_1}(x)} - \frac{(1-\frac{y}{x})^{-\beta_1}-1}{-\beta_1}\right)(1-\frac{y}{x})^{-1}dG_2(y)| + \\ &+ |\int_0^{x/2}\frac{(1-\frac{y}{x})^{-\beta_1}-1}{-\beta_1}(1-\frac{y}{x})^{-1}dG_2(y)|\end{aligned}$$

can be dominated with (in view of Lemma 1)

$$\int_0^{x/2}\left(\varepsilon(1-\frac{y}{x})^{-\beta_1-\varepsilon-1} + \frac{(1-\frac{y}{x})^{-\beta_1}-1}{-\beta_1(1-\frac{y}{x})}\right)dG_2(y).$$

Using dominated convergence it follows that the last integral has limit ε as $x \to \infty$, hence $I_3 = o(a_1(x)\overline{G_1}(x))$. Similarly we find $I_4 = o(a_2(x)\overline{G_2}(x))$.

Note that since $1 - G_i$ $(i = 1, 2)$ is regularly varying with exponent -1, hence $I_5 \sim \overline{G_1}(x)\overline{G_2}(x)$. It follows that $I_5 = o(I_1 + I_2)$. The statement follows by combination of the above estimates.

A special case is the situation where the tail function behaves like a power and a remainder term. The case $\alpha = 1$ below is a consequence of the above result. The other cases can be proved using similar arguments as in [8]. We omit the details.

Corollary 1 *Let Y_1 and Y_2 be independent non-negative random variables with distribution functions G_1 and G_2, respectively. Assume that*

$$1 - G_i(x) = a_i x^{-\alpha}[1 + b_i x^{-\beta_i} + o(x^{-\beta_i})] \text{ as } x \to \infty,$$

where $a_i > 0, \alpha > 0, \beta_i > 0,\ b_i \neq 0\ (i = 1, 2)$.

i) If $\alpha < 1$, then as $x \to \infty$

$$P(Y_1 + Y_2 > x) = (a_1 + a_2)x^{-\alpha}[1 + b_0(x^{-(\alpha\wedge\beta_1\wedge\beta_2)}) + o(x^{-(\alpha\wedge\beta_1\wedge\beta_2)})],$$

where

$$b_0 = \begin{cases} \frac{a_1 a_2}{a_1+a_2}(2\alpha \int_0^{1/2} y^{-\alpha}(1-y)^{-\alpha-1}\,dy - 2^{2\alpha}) & \text{if } \alpha \le (\beta_1 \wedge \beta_2) \\ \frac{a_2 b_2}{a_1+a_2} & \text{if } \alpha > \beta_1 \ge \beta_2 \\ \frac{a_1 b_1}{a_1+a_2} & \text{if } \alpha > \beta_2 \ge \beta_1. \end{cases}$$

ii) If $\alpha = 1$, then as $x \to \infty$

$$P(Y_1+Y_2>x) = (a_1+a_2)x^{-1}[1 + b_0(\frac{\log x}{x} \vee x^{-(\beta_1\wedge\beta_2)}) + o((\frac{\log x}{x} \vee x^{-(\beta_1\wedge\beta_2)}))],$$

where

$$b_0 = \begin{cases} \frac{2a_1 a_2}{a_1+a_2} & \text{if } (\beta_1 \wedge \beta_2) \ge 1 \\ \frac{a_2 b_2}{a_1+a_2} & \text{if } 1 > \beta_1 \ge \beta_2 \\ \frac{a_1 b_1}{a_1+a_2} & \text{if } 1 > \beta_2 \ge \beta_1. \end{cases}$$

iii) If $\alpha > 1$, then as $x \to \infty$

$$P(Y_1+Y_2>x) = (a_1+a_2)x^{-\alpha}[1 + b_0 x^{-(1\wedge\beta_1\wedge\beta_2)} + o(x^{-(1\wedge\beta_1\wedge\beta_2)})],$$

where

$$b_0 = \begin{cases} \frac{\alpha a_2 EY_1 + \alpha a_1 EY_2}{a_1+a_2} & \text{if } (\beta_1 \wedge \beta_2) \ge 1 \\ \frac{a_2 b_2}{a_1+a_2} & \text{if } 1 > \beta_1 \ge \beta_2 \\ \frac{a_1 b_1}{a_1+a_2} & \text{if } 1 > \beta_2 \ge \beta_1. \end{cases}$$

3 Main results

Before we state the theorem we prove two Lemmas.

Lemma 2 *Assume (**??**) and (**??**) hold. Then for any intermediate integer sequence k=k(n), i.e., $k \to \infty, k/n \to 0$, we have*

$$\sqrt{k}(\frac{1}{k}\sum_{i=1}^{n} \mathbf{1}(X_i > (1-F)^{-}(\frac{k}{n}x)) - x) \xrightarrow{d} W(x), \tag{12}$$

in the sense of convergence of all finite dimensional distributions, where $(1-F)^{-}$ denotes the inverse function of $1-F$ and $W(x)$ is a Gaussian process with mean zero and covariance structure

$$EW(x)W(y) = x \wedge y + \frac{2\sum_{i=1}^{l-1}\sum_{j=1}^{l-i}(c_i^\alpha x) \wedge (c_{j+i}^\alpha y)}{\sum_{i=1}^{l} c_i^\alpha} := \gamma(x,y).$$

Proof Note that

$$E(\frac{1}{k}\sum_{i=1}^{n}\mathbf{1}(X_i > (1-F)^{-}(\frac{k}{n}x)) - x)(\frac{1}{k}\sum_{j=1}^{n}\mathbf{1}(X_j > (1-F)^{-}(\frac{k}{n}y)) - y)$$

$$=: S_1 + S_2 \ ,$$

where

$$S_1 = \frac{1}{k^2}E[\sum_{i=1}^{n}(\mathbf{1}(X_i > (1-F)^{-}(\frac{k}{n}x)) - \frac{k}{n}x)][\mathbf{1}(X_i > (1-F)^{-}(\frac{k}{n}y)) - \frac{k}{n}y)] =$$

$$= \frac{1}{k^2}\sum_{i=1}^{n}[\frac{k}{n}(x\wedge y) - \frac{k^2}{n^2}xy] = \frac{1}{k}(x\wedge y)(1-\frac{k}{n}(x\vee y))$$

and

$$S_2 = \frac{2}{k^2}\sum_{i=1}^{n-1}\sum_{j=i+1}^{n}E[\mathbf{1}(X_i > (1-F)^{-}(\frac{k}{n}x)) - \frac{k}{n}x][\mathbf{1}(X_j > (1-F)^{-}(\frac{k}{n}y)) - \frac{k}{n}y] =$$

$$\frac{2}{k^2}\sum_{i=1}^{n-1}\sum_{j=i+1}^{n}E[\mathbf{1}(X_1 > (1-F)^{-}(\frac{k}{n}x)) - \frac{k}{n}x][\mathbf{1}(X_{j-i+1} > (1-F)^{-}(\frac{k}{n}y)) - \frac{k}{n}y].$$

The last expression equals

$$\frac{2}{k^2}\sum_{i=1}^{n-1}\sum_{j=1}^{n-i}E[\mathbf{1}(X_1 > (1-F)^{-}(\frac{k}{n}x)) - \frac{k}{n}x][\mathbf{1}(X_{j+1} > (1-F)^{-}(\frac{k}{n}y)) - \frac{k}{n}y].$$

Changing the order of summation we have

$$S_2 = \frac{2}{k^2}\sum_{j=1}^{l-1}(n-j)E\mathbf{1}(X_1 > (1-F)^{-}(\frac{k}{n}x))\mathbf{1}(X_{j+1} > (1-F)^{-}(\frac{k}{n}y)) - \frac{2}{k^2}\sum_{j=1}^{l-1}(n-j)\frac{k^2}{n^2}xy.$$

Note that the terms with $j > l-1$ are zero since the X_i satisfy the equation (**??**).

The first term can be evaluated using Lemma 5.1 and Lemma 5.2 of Datta and McCormick (1995). Combining the above results it follows that

$$kE[\frac{1}{k}\sum_{i=1}^{n}\mathbf{1}(X_i > (1-F)^{-}(\frac{k}{n}x)) - x][\frac{1}{k}\sum_{i=1}^{n}\mathbf{1}(X_i > (1-F)^{-}(\frac{k}{n}y)) - y] \to \gamma(x,y),$$

hence

$$\begin{aligned}
&\operatorname{var}\textstyle\sum_{i=1}^{n}(\frac{n}{k}\mathbf{1}(X_i > (1-F)^{-}(\frac{k}{n}x)) - x) = E[\sum_{i=1}^{n}\left(\frac{n}{k}\mathbf{1}(X_i > (1-F)^{-}(\frac{k}{n}x)) - x\right)]^2 \\
&= \tfrac{n^2}{k}kE[\tfrac{1}{k}\textstyle\sum_{i=1}^{n}\mathbf{1}(X_i > (1-F)^{-}(\frac{k}{n}x)) - x]^2 \\
&\sim \tfrac{n^2}{k}\gamma(x,x) \to \infty \ (n\to\infty).
\end{aligned}$$

Application of theorem 18.5.1 in Ibragimov and Linnik (1971) completes the proof.

Lemma 3 *Assume that (??) holds. Then for any intermediate integer sequence $k = k(n)$ satisfying*

$$\lim_{n\to\infty} \sqrt{k}(n/k)^{-\beta/\alpha} = \lambda \in [0,\infty), \tag{13}$$

we have

$$\sqrt{k}\{\frac{X_{n,n-[kx]}}{(k/(na))^{-1/\alpha}} - x^{-1/\alpha}\} \xrightarrow{d} \frac{1}{\alpha}x^{-1-1/\alpha}W(x) + \frac{\lambda a^{-\beta/\alpha}b}{\alpha}x^{-1/\alpha+\beta/\alpha},$$

where $W(x)$ is defined in Lemma 2.

Proof By Vervaat's lemma (Vervaat(1979)) and Lemma 2, we have

$$\sqrt{k}\{\frac{n}{k}(1-F(X_{n,n-[kx]})) - x\} \xrightarrow{d} -W(x).$$

Using (??) we have

$$\sqrt{k}\{\frac{n}{k}aX_{n,n-[kx]}^{-\alpha} - x\} \xrightarrow{d} -W(x) - \lambda a^{-\beta/\alpha}bx^{1+\beta/\alpha}.$$

Hence the lemma follows.

Next we give the optimal sample fraction k_0 for the estimator $\hat{\alpha}_n(k)$ defined in (??) and establish the asymptotic normality of the optimal estimator $\hat{\alpha}_n(k_0)$.

Theorem 2 *Assume that (??) holds. Let*

$$k_0 := argmin_k E(\hat{\alpha}_n(k) - \alpha)^2.$$

Then

$$k_0/\{(\frac{\alpha(13\gamma(1,1) - 12\gamma(1,2) + 2\gamma(1,4))}{4\beta a^{-2\beta/\alpha}b^2(2^{1+\beta/\alpha}-1)^2(2^{\beta/\alpha}-1)^2})^{\frac{\alpha}{\alpha+2\beta}} n^{\frac{2\beta}{\alpha+2\beta}}\} \to 1. \tag{14}$$

Furthermore

$$\begin{aligned} &\sqrt{k_0}(\hat{\alpha}_n(k_0) - \alpha) \\ \xrightarrow{d}\quad & N(\tfrac{\alpha\sqrt{\alpha(13\gamma(1,1)-12\gamma(1,2)+2\gamma(1,4))}}{2\sqrt{\beta}\log 2}, \tfrac{\alpha^2(13\gamma(1,1)-12\gamma(1,2)+2\gamma(1,4))}{2(\log 2)^2}). \end{aligned}$$

Proof In view of Lemma 3 we have $\frac{X_{n,n-2k}}{X_{n,n-k}} \xrightarrow{p} 2^{-1/\alpha}$, hence

$$\begin{aligned} & (\log 2)(\log \tfrac{X_{n,n-k}}{X_{n,n-2k}})^{-1} - \alpha \\ = \quad & \alpha(\log \tfrac{X_{n,n-k}}{X_{n,n-2k}})^{-1} \log \tfrac{2^{1/\alpha}X_{n,n-2k}}{X_{n,n-k}} \\ = \quad & \alpha(\log \tfrac{X_{n,n-k}}{X_{n,n-2k}})^{-1} \tfrac{2^{1/\alpha}X_{n,n-2k}-X_{n,n-k}}{X_{n,n-k}}(1+o(1)) \\ = \quad & \alpha(\log \tfrac{X_{n,n-k}}{X_{n,n-2k}})^{-1} \tfrac{(k/(na))^{-1/\alpha}}{X_{n,n-k}}\{2^{1/\alpha}(\tfrac{X_{n,n-2k}}{(k/(na))^{-1/\alpha}} - 2^{-1/\alpha}) \\ & -(\tfrac{X_{n,n-k}}{(k/(na))^{-1/\alpha}} - 1)\}(1+o(1)). \end{aligned} \tag{15}$$

and

$$\begin{array}{rl} & 2(\log 2)(\log \frac{X_{n,n-2k}}{X_{n,n-4k}})^{-1} - 2\alpha \\ = & 2\alpha(\log \frac{X_{n,n-2k}}{X_{n,n-4k}})^{-1} \frac{(k/(na))^{-1/\alpha}}{X_{n,n-2k}} \{2^{1/\alpha}(\frac{X_{n,n-4k}}{(k/(na))^{-1/\alpha}} - 4^{-1/\alpha}) \\ & -(\frac{X_{n,n-2k}}{(k/(na))^{-1/\alpha}} - 2^{-1/\alpha})\}(1+o(1)) \end{array} \tag{16}$$

Thus by Lemma 3 the variance of $\sqrt{k}(\hat{\alpha}_n(k) - \alpha)$ asymptotically equals

$$\begin{array}{rl} & E\{2\alpha\frac{\alpha}{\log 2}\frac{1}{2^{-1/\alpha}}(2^{1/\alpha}\frac{1}{\alpha}4^{-1-1/\alpha}W(4) - \frac{1}{\alpha}2^{-1-1/\alpha}W(2)) \\ & -\frac{\alpha^2}{\log 2}(2^{1/\alpha}\frac{1}{\alpha}2^{-1-1/\alpha}W(2) - \frac{1}{\alpha}W(1))\}^2 \\ = & \frac{\alpha^2}{(log 2)^2}E\{2^{-1}W(4) - W(2) - 2^{-1}W(2) + W(1)\}^2 \\ = & \frac{\alpha^2[13\gamma(1,1) - 12\gamma(1,2) + 2\gamma(1,4)]}{2(\log 2)^2}, \end{array} \tag{17}$$

where γ is as defined in Lemma 2.

Using Lemma 3 the bias of $\sqrt{k}(\hat{\alpha}_n(k) - \alpha)$ asymptotically equals

$$\begin{array}{rl} & 2\frac{\alpha}{\log 2}\alpha\frac{1}{2^{-1/\alpha}}(2^{1/\alpha}\frac{a^{-\beta/\alpha}b}{\alpha}4^{-1/\alpha+\beta/\alpha}\sqrt{k}(n/k)^{-\beta/\alpha} - \frac{a^{-\beta/\alpha}b}{a}2^{-1/\alpha+\beta/\alpha}\sqrt{k}(n/k)^{-\beta/\alpha}) \\ & -\alpha\frac{\alpha}{\log 2}[2^{1/\alpha}\frac{a^{-\beta/\alpha}b}{\alpha}2^{-1/\alpha+\beta/\alpha}\sqrt{k}(n/k)^{-\beta/\alpha} - \frac{a^{-\beta/\alpha}b}{\alpha}\sqrt{k}(n/k)^{-\beta/\alpha}] \\ = & \frac{\alpha a^{-\beta/\alpha}b}{\log 2}\sqrt{k}(n/k)^{-\beta/\alpha}(2^{1+\beta/\alpha} - 1)(2^{\beta/\alpha} - 1). \end{array} \tag{18}$$

By the same minimization argument as in Dekkers and de Haan (1993), (**??**) follows easily. Moreover minimizing the sum of the variance and the bias squared gives

$$\lim_{n\to\infty}\sqrt{k_0}(n/k_0)^{-\beta/\alpha} = \sqrt{\frac{\alpha(13\gamma(1,1) - 12\gamma(1,2) + 2\gamma(1,4))}{4\beta a^{-2\beta/\alpha}b^2(2^{1+\beta/\alpha} - 1)^2(2^{\beta/\alpha} - 1)^2}}.$$

Thus

$$\begin{array}{rl} & \sqrt{k_0}(\hat{\alpha}_n(k_0) - \alpha) \\ \xrightarrow{d} & N(\frac{\alpha\sqrt{\alpha(13\gamma(1,1) - 12\gamma(1,2) + 2\gamma(1,4))}}{2\sqrt{\beta}\log 2}, \frac{\alpha^2(13\gamma(1,1) - 12\gamma(1,2) + 2\gamma(1,4))}{2(\log 2)^2}). \end{array}$$

In order to implement our bootstrap procedure and avoid choosing the initial estimator which involves the optimal choice of sample fraction, we are going to bootstrap $E(\hat{\alpha}_n(k) - \bar{\alpha}_n(k))^2$ instead of $E(\hat{\alpha}_n(k) - \alpha)^2$. The following result specifies the optimal k_0 for the minimization of $E(\hat{\alpha}_n(k) - \bar{\alpha}_n(k))^2$.

Theorem 3 *Assume (**??**) holds. Let*

$$\bar{k}_0 := argmin_k E(\hat{\alpha}_n(k) - \bar{\alpha}_n(k))^2.$$

Then

$$\bar{k}_0/\{(\frac{\alpha(13\gamma(1,1) - 12\gamma(1,2) + 2\gamma(1,4))}{8\beta a^{-2\beta/\alpha}b^2(2^{\beta/\alpha} - 1)^4})^{\frac{\alpha}{\alpha+2\beta}}n^{\frac{2\beta}{\alpha+2\beta}}\} \to 1.$$

Proof From (??) and (??) we know that the variance of $\sqrt{k}(\hat{\alpha}_n(k) - \bar{\alpha}_n(k))$ asymptotically equals

$$\begin{aligned} & E\{2\tfrac{\alpha}{\log 2}\alpha\tfrac{1}{2^{-1/\alpha}}(2^{1/\alpha}\tfrac{1}{\alpha}4^{-1-1/\alpha}W(4) - \tfrac{1}{\alpha}2^{-1-1/\alpha}W(2)) \\ & -2\alpha\tfrac{\alpha}{\log 2}(2^{1/\alpha}\tfrac{1}{\alpha}2^{-1-1/\alpha}W(2) - \tfrac{1}{\alpha}W(1))\}^2 \\ = & \tfrac{4\alpha^2}{(\log 2)^2}E\{4^{-1}W(4) - W(2) + W(1)\}^2 \\ = & \tfrac{\alpha^2}{(\log 2)^2}\{13\gamma(1,1) - 12\gamma(1,2) + 2\gamma(1,4)\} \end{aligned}$$

and the bias of $\sqrt{k}(\hat{\alpha}_n(k) - \bar{\alpha}_n(k))$ asymptotically equals

$$\begin{aligned} & 2\tfrac{\alpha}{\log 2}\alpha\tfrac{1}{2^{-1/\alpha}}(2^{1/\alpha}\tfrac{a^{-\beta/\alpha}b}{\alpha}4^{-1/\alpha+\beta/\alpha}\sqrt{k}(n/k)^{-\beta/\alpha} - \tfrac{a^{-\beta/\alpha}b}{\alpha}2^{-1/\alpha+\beta/\alpha}\sqrt{k}(n/k)^{-\beta/\alpha}) \\ & -2\alpha\tfrac{\alpha}{\log 2}(2^{1/\alpha}\tfrac{a^{-\beta/\alpha}b}{\alpha}2^{-1/\alpha+\beta/\alpha}\sqrt{k}(n/k)^{-\beta/\alpha} - \tfrac{a^{-\beta/\alpha}b}{\alpha}\sqrt{k}(n/k)^{-\beta/\alpha}) \\ = & \tfrac{2\alpha a^{-\beta/\alpha}b}{\log 2}\sqrt{k}(n/k)^{-\beta/\alpha}(2^{\beta/\alpha} - 1)^2. \end{aligned}$$

By the same minimization argument as in Dekkers and de Haan (1993) we have

$$\bar{k}_0/\{(\frac{\alpha(13\gamma(1,1) - 12\gamma(1,2) + 2\gamma(1,4))}{8\beta a^{-2\beta/\alpha}b^2(2^{\beta/\alpha} - 1)^4})^{\frac{\alpha}{\alpha+2\beta}}n^{\frac{2\beta}{\alpha+2\beta}}\} \to 1.$$

Combining Theorem 2 and Theorem 3, the following corollary follows easily.

Corollary 2 *Assume (??) holds. Then*

$$k_0/\bar{k}_0 \to (\frac{2(2^{\beta/\alpha} - 1)^2}{(2^{1+\beta/\alpha} - 1)^2})^{\frac{1}{1+2\beta/\alpha}}.$$

Note that the above Corollary shows that k_0 can be estimated through $\bar{k_0}$ and β/α.

Next we are going to apply the non-overlapping subsampling method (see Carlstein (1986)) in order to estimate $\bar{k}_0$ and β/α.

Let $\mathcal{X}_j^i := \{X_{i+1}, \cdots, X_{i+j}\}$ be a sequence of $(X_t)_{t\geq 1}$ starting at X_{i+1} and let $\hat{\alpha}_{i,j}(k)$ and $\bar{\alpha}_{i,j}(k)$ denote the corresponding $\hat{\alpha}_n(k)$ and $\bar{\alpha}_n(k)$ which are based on the samples $\mathcal{X}_j^i$.

Theorem 4 *Assume (??) holds. Let $n_1 = O(n^\epsilon)(0 < \epsilon < 1)$ and*

$$\hat{k}_{0,1} := argmin_k([n/n_1])^{-1}\sum_{i=0}^{[n/n_1]-1}(\hat{\alpha}_{in_1,n_1}(k) - \bar{\alpha}_{in_1,n_1}(k))^2.$$

Then

$$\frac{\hat{k}_{0,1}n_1^{-\frac{2\beta}{\alpha+2\beta}}}{\bar{k}_0 n^{-\frac{2\beta}{\alpha+2\beta}}} \xrightarrow{p} 1.$$

Proof It follows from Theorem 3 and the fact that $\{\hat{\alpha}_{in_1,n_1}(k)-\bar{\alpha}_{in_1,n_1}(k)\}_{i=0}^{[n/n_1]-1}$ are i.i.d. random variables as $n_1 > l$.

Finally we estimate k_0 by using two independent subsampling procedures providing estimators for $\bar{k}_0$ and β/α.

Theorem 5 *Assume (??) holds. Let $n_1 = O(n^\epsilon)$ for some $1/2 < \epsilon < 1$, $n_2 = [n_1^2/n]$ and*

$$\hat{k}_{0,i} := argmin_k([n/n_i])^{-1} \sum_{j=0}^{[n/n_i]-1} (\hat{\alpha}_{jn_i,n_i}(k) - \bar{\alpha}_{jn_i,n_i}(k))^2$$

for $i = 1, 2$. Then

$$\frac{\hat{k}_{0,1}^2}{\hat{k}_{0,2}\bar{k}_0} \xrightarrow{p} 1. \tag{19}$$

Furthermore

$$\frac{\hat{k}_{0,1}^2}{\hat{k}_{0,2}}\left(\frac{2(2^{\hat{\rho}}-1)^2}{(2^{1+\hat{\rho}}-1)^2}\right)^{1/(1+2\hat{\rho})}/k_0 \xrightarrow{p} 1 \tag{20}$$

where $\hat{\rho} = \frac{\log \hat{k}_{0,1}}{2(\log n_1 - \log \hat{k}_{0,1})}$.

Moreover

$$\frac{\log \hat{k}_{0,1}}{2(\log n_1 - \log \hat{k}_{0,1})} \xrightarrow{p} \beta/\alpha. \tag{21}$$

Proof (**??**) follows easily from Theorem 4 and $n_2 = [n_1^2/n]$.

Since $\log \hat{k}_{0,1} - \frac{2\beta}{\alpha+2\beta}\log n_1 \xrightarrow{p} 0$, we have (**??**).

Note that (**??**) follows from (**??**) and Corollary 2.

References

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

[2] E. Carlstein (1986). The use of subseries values for estimating the variance of a general statistic from a stationary sequence. *Ann. Statist. 14, 1171-1179.*

[3] J. Danielsson, L. de Haan, L. Peng and C.G. de Vries (1997). Using a bootstrap method to choose the sample fraction in the tail index estimation. *Technical Report. Erasmus University Rotterdam.*

[4] S. Datta and W.P. McCormick (1998). Inference for the tail parameters of a linear process with heavy tailed innovations. *Ann. Inst. Stat. Math. 50, 337-359.*

[5] 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. Statist. 17, 1833-1855.*

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

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

[8] W. Feller (1971). An Introduction to Probability Theory and Its Applications. *Volime II, John Wiley & Sons, Inc.*

[9] J. Geluk, L. de Haan, S. Resnick and C. Starica (1997). Second-order regular variation, convolution and the central limit theorem. *stochastic Processes and their Applications, 69, 139-159.*

[10] L. de Haan, L. Peng (1998). Comparison of tail index estimators, *Statistica Neerlandica, 52, 60-70.*

[11] L. de Haan and T.T. Pereira (1998). Estimating the index of a stable distribution. *Statistics and Probability Letters, to appear*

[12] P. Hall (1990). Using the bootstrap to estimate mean squared error and selecting parameter in nonparametric problems. *J. Multivariate Anal. 32, 177-203.*

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

[14] T. Hsing (1991). On tail estimation using dependent data. *Ann. Statist. 19, 1547-1569.*

[15] I.A. Ibragimov and Y.V. Linnik (1971). Independent and Stationary Sequences of Random Variables. *Wolters-Noordhoff, Groningen.*

[16] J. Pickands III(1975). Statistical inference using extreme order statistics. *Ann. Statist. 3(1), 119-131.*

[17] S. Resnick and C. Starica (1995). Consistency of Hill's estimator for dependent data. *J. Applied Probability, 32, 139-167.*

[18] S. Resnick and C. Starica (1996). Tail index estimation for dependent data. *Technical Report. Cornell University.*

[19] H. Rootzen, M. Leadbetter and L. de Haan (1998). On the distribution of tail array sums for strongly mixing stationary sequences. *Annals of Applied Prob. 8, 868-885.*

[20] W. Vervaat (1979). On a stochastic difference equation and a representation of non-negative infinitely divisible random variables. *Adv. Appl. Prob., 11, 750-783.*