The Annals of Statistics
1991, Vol. 19, No. 3, 1547–1569

ON TAIL INDEX ESTIMATION USING DEPENDENT DATA[1]

By Tailen Hsing

Texas A & M University

Let $X_1, X_2, \dots$ be possibly dependent random variables having the same marginal distribution. Consider the situation where $\bar{F}(x) := P[X_1 > x]$ is regularly varying at ∞ with an unknown index $-\alpha < 0$ which is to be estimated. In the i.i.d. setting, it is well known that Hill's estimator is consistent for α^{-1}, and is asymptotically normally distributed. It is the purpose of this paper to demonstrate that such properties of Hill's estimator extend considerably beyond the independent setting. In addition to some basic results derived under very general conditions, the case where the observations are strictly stationary and satisfy a certain mixing condition is considered in detail. Also a finite moving average sequence is studied to illustrate the results.

1. Introduction. Suppose $\{X_i\}$ is a sequence of random variables having the same marginal distribution function F, where $\bar{F} := 1 - F$ is regularly varying at ∞, namely there exists an $\alpha > 0$, such that

$$\bar{F}(tx)/\bar{F}(x) \to t^{-\alpha} \quad \text{as } x \to \infty \text{ for all } t > 0, \tag{1.1a}$$

or equivalently

$$\bar{F}(x) = x^{-\alpha}L(x), \qquad x > 0 \quad \text{for some slowly varying function } L. \tag{1.1b}$$

This will be denoted by $\bar{F} \in RV_{-\alpha}$. $-\alpha$ is called the regular variation index of $\bar{F}$, and more conveniently in this context, the tail index of F. The class of distributions having the tail behavior (1.1) is infinitely large, and is known to coincide with the maximum domain of attraction of the extreme value distribution $\exp(-x^{-\alpha})$, $x > 0$. See Bingham, Goldie and Teugels (1987), de Haan (1970), Feller (1971) and Seneta (1976) for details on the notion of regular variation and its applications in statistics and probability.

We are interested in the estimation of α when observing $X_1, \dots, X_n$. It is intuitively clear that if little or no additional structural information on F is available, which we assume to be the case, any inference on α should be made with the tail portion of the empirical distribution of the sample. Thus we can assume without any loss of generality that F is supported on $(0, \infty)$. Define

$$F^{-1}(y) = \inf\{x\colon F(x) \geq y\}, \qquad 0 < y < 1.$$

Assume that $\bar{F}(x)/\bar{F}(x-) \to 1$ as $x \to \infty$ [cf. Leadbetter, Lindgren and

Received March 1989; revised June 1990.
[1]Research supported by NSF Grant DMS-88-14006.
AMS 1980 *subject classifications*. Primary 62F10; secondary 62G05.
Key words and phrases. Order statistics, regular variation, parameter estimation.

Rootzén (1983), Theorem 1.7.13]. Under this, it is easily seen that

$$\bar{F}(b(t)) \sim t^{-1} \quad \text{as } t \to \infty, \tag{1.2}$$

where

$$b(t) := \bar{F}^{-1}(1 - t^{-1}), \qquad t > 1.$$

Later in the paper, regularity conditions on the slowly varying component L in (1.1b) will be introduced to control the rate of approximation in (1.2). For $1 \le j \le n$, write $X_{(j)} = X_{(n:\, j)}$ for the jth largest value of $X_1, \dots, X_n$. For $x > 0$, x^* denotes $\log x$, and for $x \in \mathbb{R}$, x_+ denotes $\max(x, 0)$ and x_- denotes $\max(-x, 0)$.

Papers that discuss the estimation of α in the i.i.d. setting include Beirlant and Teugels (1989), Csörgő, Deheuvels and Mason (1985), Davis and Resnick (1984), Haeusler and Teugels (1985), Hall (1982), Hall and Welsh (1984), Hill (1975), Mason (1982) and Smith (1987), to mention a few. One ingredient common to these papers is the consideration of the so-called Hill's estimator

$$H_n := m^{-1} \sum_{j=1}^{m} X_{(j)}^* - X_{(m+1)}^*, \tag{1.3}$$

which was first proposed by Hill (1975) in a slightly different manner. Asymptotic properties of H_n, including consistency and asymptotic normality, were studied by letting m vary with n such that

$$m \to \infty \quad \text{and} \quad m/n \to 0 \quad \text{as } n \to \infty. \tag{1.4}$$

The idea behind Hill's estimator is easily understood. By dominated convergence and Potter's theorem [cf. Bingham, Goldie and Teugels (1987)], the kth moment of $(X_1^* - b^*(n/m))_+$, the excess of $\log X_1$ above $\log b(n/m)$, is evaluated as

$$\begin{aligned} \mathscr{E}(X_1^* - b^*(n/m))_+^k &= \int_0^\infty P\left[(X_1^* - b^*(n/m))^k > u\right] du \\ &= \int_0^\infty \bar{F}\left(e^{u^{1/k}} b(n/m)\right) du \\ &= \bar{F}(b(n/m)) \int_0^\infty \bar{F}\left(e^{u^{1/k}} b(n/m)\right) / \bar{F}(b(n/m))\, du \\ &\sim \frac{m}{n} \int_0^\infty e^{-\alpha u^{1/k}}\, du = \frac{m}{n} \frac{k!}{\alpha^k}. \end{aligned} \tag{1.5}$$

In particular, $\mathscr{E}(X_1^* - b^*(n/m))_+ \sim (m/n)\alpha^{-1}$. If $\{X_i\}$ is i.i.d., clearly $X_{(m+1)}^*$ estimates $b^*(n/m)$ and hence

$$H_n = m^{-1} \sum_{i=1}^{n} (X_i^* - X_{(m+1)}^*)_+ \approx m^{-1} \sum_{i=1}^{n} (X_i^* - b^*(n/m))_+, \tag{1.6}$$

showing that H_n is essentially a method of moments estimator of α^{-1}.

A powerful tool for studying the asymptotics of Hill's estimator H_n in the i.i.d. setting is representing the order statistics $X_{(j)}^*$ as weighted sums of i.i.d.

random variables. The technique is known as Rényi's representation [cf. Davis and Resnick (1984)]. Unfortunately, the representation does not work when dependence is present. The goal of this paper is to generalize certain results on H_n in the i.i.d. case by dropping independence. Specifically, Section 2 shows that, by formalizing the idea of approximation in (1.6), the consistency and asymptotic normality of H_n can be extended considerably beyond the independent setting. Section 3 specializes the results in Section 2 to sequences $\{X_i\}$ which are strictly stationary and which satisfy a certain mixing condition. To give a more concrete demonstration of the results in Sections 2 and 3, the finite moving average sequence is considered in Section 4.

Without repeated reference, the assumptions and notation introduced in this section, including the ones in (1.1), (1.2), (1.3), (1.4) and (1.6), are used throughout this paper. The dependence of m on n in (1.4) is suppressed except in Section 3.

2. General results.

In this section, we study some asymptotics of Hill's estimator H_n without requiring $\{X_i\}$ to be i.i.d. At the outset we make no assumptions on the X_i's other than that they have the common marginal distribution as described in Section 1.

Define two quantities $\tilde{H}_n$ and H_n^+ by

$$\tilde{H}_n = m^{-1} \sum_{j=1}^{m} \left(X_{(j)}^* - b^*(n/m) \right),$$

$$H_n^+ = m^{-1} \sum_{i=1}^{n} \left(X_i^* - b^*(n/m) \right)_+ .$$

Clearly, H_n, $\tilde{H}_n$ and H_n^+ are intimately related. The following lemma describes the relationship as $n \to \infty$.

LEMMA 2.1. *If $X_{([\rho m])}^* - b^*(n/\rho m) \to_P 0$ for all ρ in I, some neighborhood of 1, then*

$$\left| H_n - \tilde{H}_n \right| + \left| H_n - H_n^+ \right| + \left| \tilde{H}_n - H_n^+ \right| \to_P 0.$$

If, in addition, the collection of probability distributions of $\sqrt{m}\,(X_{(m+1)}^* - b^*(n/m))$, $n \geq 1$, is tight, then

$$\sqrt{m} \left(\tilde{H}_n - H_n^+ \right) \to_P 0.$$

PROOF. We first prove the second claim which is the more difficult of the two. Let ε be a fixed small positive constant such that $(1 - \varepsilon, 1 + \varepsilon) \subset I$. We can write

$$\begin{aligned}\sqrt{m}\left(H_n^+ - \tilde{H}_n\right) &= \frac{1}{\sqrt{m}} \sum_{j=1}^{[(1-\varepsilon)m]} \left(X_{(j)}^* - b^*(n/m)\right)_- \\ &\quad + \frac{1}{\sqrt{m}} \sum_{j=[(1-\varepsilon)m]+1}^{m} \left(X_{(j)}^* - b^*(n/m)\right)_- \\ &\quad + \frac{1}{\sqrt{m}} \sum_{j=m+1}^{[(1+\varepsilon)m]} \left(X_{(j)}^* - b^*(n/m)\right)_+ \\ &\quad + \frac{1}{\sqrt{m}} \sum_{j=[(1+\varepsilon)m]+1}^{n} \left(X_{(j)}^* - b^*(n/m)\right)_+ \\ &=: A_n^{(\varepsilon)} + B_n^{(\varepsilon)} + C_n^{(\varepsilon)} + D_n^{(\varepsilon)}.\end{aligned}$$

Clearly, $P[A_n^{(\varepsilon)} > 0] \le P[X_{([(1-\varepsilon)m])}^* < b^*(n/m)]$. Since $X_{([(1-\varepsilon)m])}^* - b^*(n/(1-\varepsilon)m) \to_P 0$ and $\liminf_{n\to\infty}\{b^*(n/(1-\varepsilon)m) - b^*(n/m)\} > 0$, $P[A_n^{(\varepsilon)} > 0] \to 0$. The same argument shows that $P[D_n^{(\varepsilon)} > 0] \to 0$. Further it can be seen that $B_n^{(\varepsilon)} \le \varepsilon\sqrt{m}\,(X_{(m+1)}^* - b^*(n/m))_-$, and hence for $\eta > 0$,

$$\begin{aligned}&\lim_{\varepsilon\to 0}\limsup_{n\to\infty} P\left[B_n^{(\varepsilon)} > \eta\right] \\ &\quad \le \lim_{\varepsilon\to 0}\limsup_{n\to\infty} P\left[\sqrt{m}\left(X_{(m+1)}^* - b^*(n/m)\right)_- > \varepsilon^{-1}\eta\right] = 0\end{aligned}$$

by virtue of the assumption that $\sqrt{m}\,(X_{(m+1)}^* - b^*(n/m))$, $n \ge 1$, is tight. Essentially the same argument can be applied to $C_n^{(\varepsilon)}$ to get

$$\lim_{\varepsilon\to 0}\limsup_{n\to\infty} P\left[C_n^{(\varepsilon)} > \eta\right] = 0, \qquad \eta > 0.$$

Thus for $\eta > 0$,

$$\begin{aligned}&\lim_{n\to\infty} P\left[\sqrt{m}\left(H_n^+ - \tilde{H}_n\right) > \eta\right] \\ &\quad \le \lim_{n\to\infty}\left(P\left[A_n^{(\varepsilon)} > \eta/4\right] + P\left[D_n^{(\varepsilon)} > \eta/4\right]\right) \\ &\qquad + \lim_{\varepsilon\to 0}\limsup_{n\to 0}\left(P\left[B_n^{(\varepsilon)} > \eta/4\right] + P\left[C_n^{(\varepsilon)} > \eta/4\right]\right) = 0.\end{aligned}$$

This proves the second claim of the lemma. To prove the first claim of the lemma, it suffices to show $\tilde{H}_n - H_n^+ \to_P 0$ and $H_n - \tilde{H}_n \to_P 0$. The first convergence can be established through steps that are quite similar to the ones in the previous part and we therefore omit its proof. The second convergence follows from the identity.

$$H_n - \tilde{H}_n = -\left(X_{(m+1)}^* - b^*(n/m)\right),$$

where the right-hand side can easily be seen to converge to 0 under the conditions specified. □

Since H_n is basically a moment estimator of α^{-1}, as explained in Section 1, one should be able to show its consistency by showing certain "laws of large numbers." The next result does just that.

THEOREM 2.2. *Suppose*

$$m^{-1}\sum_{i=1}^{n}(T_{ni} - \mathscr{E}T_{ni}) \to_P 0 \tag{2.1}$$

for $T_{ni} = (X_i^* - b^*(n/m))_+$ *and* $I(X_i^* - b^*(n/\rho m) > \varepsilon)$ *for every* $\varepsilon \in \mathbb{R}$ *and* ρ *in some neighborhood* I *of* 1. *Then* H_n, $\tilde{H}_n$ *and* H_n^+ *all converge to* α^{-1} *in probability.*

PROOF. Since $(n/m)\mathscr{E}(X_i^* - b^*(n/m))_+ \to \alpha^{-1}$, that (2.1) holds for $T_{ni} = (X_i^* - b^*(n/m))_+$ implies that H_n^+ converges to α^{-1} in probability. To show that $\tilde{H}_n$ and H_n converge in probability to the same limit, it suffices to show, by Lemma 2.1, that

$$X^*_{([\rho m])} - b^*\left(\frac{n}{\rho m}\right) \to_P 0, \qquad \rho \in I.$$

In other words, it suffices to show that

$$P\left[X^*_{([\rho m])} - b^*\left(\frac{n}{\rho m}\right) > \varepsilon\right] \to 0$$

and

$$P\left[X^*_{([\rho m])} - b^*\left(\frac{n}{\rho m}\right) < -\varepsilon\right] \to 0, \qquad \varepsilon > 0, \rho \in I.$$

Since the two convergence statements have virtually the same proof, we only show the first one. Writing $I_{ni} = I(X_i^* - b^*(n/\rho m) > \varepsilon)$, we have

$$\begin{aligned} P&\left[X^*_{([\rho m])} - b^*\left(\frac{n}{\rho m}\right) > \varepsilon\right] \\ &= P\left[\sum_{i=1}^{n} I_{ni} \geq [\rho m]\right] \\ &= P\left[m^{-1}\sum_{i=1}^{n}(I_{ni} - \mathscr{E}I_{ni}) \geq m^{-1}([\rho m] - n\mathscr{E}I_{ni})\right]. \end{aligned}$$

Since $m^{-1}\sum_{i=1}^{n}(I_{ni} - \mathscr{E}I_{ni}) \to_P 0$ and $m^{-1}([\rho m] - n\mathscr{E}I_{ni}) \to \rho(1 - e^{-\alpha\varepsilon}) > 0$, the above probability tends to 0. This concludes the proof. □

To consider the asymptotic distribution of H_n, it is necessary to model the tail behavior of F more closely. For that we make use of the notion of slow variation with remainder introduced by Goldie and Smith (1987). The present treatment is also heavily influenced by Haeusler and Teugels (1985).

The condition (SR1) is said to hold for the slowly varying function L if there exists a positive measurable function g on $(0, \infty)$ such that

$$\text{(SR1)} \qquad \text{for all } \lambda > 1, \quad L(\lambda x)/L(x) - 1 = \mathscr{O}(g(x)) \quad \text{as } x \to \infty.$$

The condition (SR2) is said to hold for L if there exists a function k and a positive measurable function g on $(0, \infty)$ such that

$$\text{(SR2)} \qquad \text{for all } \lambda > 1, \quad L(\lambda x)/L(x) - 1 \sim k(\lambda) g(x) \quad \text{as } x \to \infty.$$

To make these conditions meaningful, we also assume that the function g in (SR1) and (SR2) tends to 0 as $x \to \infty$. g is said to have bounded increase if there exist $C, x_0, \tau < \infty$ such that

$$g(\lambda x)/g(x) \le C\lambda^{\tau}, \qquad \lambda \ge 1, x \ge x_0.$$

THEOREM 2.3 [Goldie and Smith (1987), Corollary 2.2.1 and Propositions 2.5.1 and 2.5.2]. *If the slowly varying function L satisfies* (SR1) *and g has bounded increase with $\tau \le 0$, then* (SR1) *holds uniformly on every compact λ-set in* $[1, \infty)$, *and*

$$\int_1^{\infty} y^{-\alpha-1} \frac{L(xy)}{L(x)}\, dy - \alpha^{-1} = \mathscr{O}(g(x)) \quad \text{as } x \to \infty.$$

If L satisfies (SR2) *and $g \in RV_{\gamma}$, $\gamma \le 0$, and $k(y) = K\int_1^y t^{\gamma-1}\, dt$, $|K| < \infty$, then* (SR2) *holds uniformly on every compact λ-set in* $[1, \infty)$, *and*

$$\int_1^{\infty} y^{-\alpha-1} \frac{L(xy)}{L(x)}\, dy - \alpha^{-1} \sim \frac{K}{\alpha(\alpha-\gamma)} g(x) \quad \text{as } x \to \infty.$$

The assumptions in Theorem 2.3 are very general and hold for a wide class of slowly varying functions. See Goldie and Smith (1987) and Haeusler and Teugels (1985) for details and examples.

Define

$$\text{(2.2)} \qquad S_n^{(\zeta)} = m^{-1} \sum_{i=1}^{n} I\big(X_i^* > b^*(n/m) + \zeta/\sqrt{m}\big), \qquad \zeta \in \mathbb{R}.$$

THEOREM 2.4. *Assume that the slowly varying component L in* (1.1b) *satisfies either*

(2.3a) (SR1) *with g having bounded increase with $\tau \le 0$ and* $\sqrt{m}\, g(b(n/m)) \to 0$

or

(2.3b) (SR2) *with $g \in RV_{\gamma}$, $\gamma \le 0$, $k(y) = K\int_1^y t^{\gamma-1}\, dt$, $|K| < \infty$ and* $\sqrt{m}\, g(b(n/m)) \to A \in \mathbb{R}$.

Suppose further that there exists a random vector (Y, Z) such that for all $\zeta \in \mathbb{R}$,

$$\text{(2.4)} \qquad \sqrt{m}\left(H_n^+ - \mathscr{E}H_n^+, \alpha^{-1}\big(S_n^{(\zeta)} - \mathscr{E}S_n^{(\zeta)}\big)\right) \to_d (Y, Z).$$

Then

$$\sqrt{m}\left(H_n^+ - \mathscr{E}H_n^+, X_{(m+1)}^* - b^*(n/m)\right) \to_d (Y, Z). \tag{2.5}$$

If, in addition, for all $\varepsilon \in \mathbb{R}$ *and* p *in some neighborhood* I *of* 1,

$$m^{-1}\sum_{i=1}^{n}\left(I\left(X_i^* > b^*\left(\frac{n}{\rho m}\right) + \varepsilon\right) - P\left[X_i^* > b^*\left(\frac{n}{\rho m}\right) + \varepsilon\right]\right) \to_P 0, \tag{2.6}$$

then we have

$$\sqrt{m}\left(\tilde{H}_n - \mathscr{E}H_n^+, X_{(m+1)}^* - b^*(n/m)\right) \to_d (Y, Z), \tag{2.7}$$

$$\sqrt{m}\left(H_n - \alpha^{-1}\right) \to_d \begin{cases} Y - Z, & \text{under (2.3a)}, \\ Y - Z + \dfrac{KA}{\alpha(\alpha-\gamma)}, & \text{under (2.3b)}. \end{cases} \tag{2.8}$$

PROOF. First assume that (2.3a) holds. First fix $\zeta \in \mathbb{R}$ and observe that $\sqrt{m}\,(X_{(m+1)}^* - b^*(n/m)) \le \zeta$ iff $S_n^{(\zeta)} \le 1$, and hence iff

$$\begin{aligned} &\alpha^{-1}\sqrt{m}\left(S_n^{(\zeta)} - \mathscr{E}S_n^{(\zeta)}\right) \\ &\le \alpha^{-1}\sqrt{m}\left(1 - \frac{n}{m}\bar{F}\left(e^{\zeta/\sqrt{m}}b(n/m)\right)\right) \\ &= \alpha^{-1}\sqrt{m}\left(1 - \frac{n}{m}\bar{F}(b(n/m))\bar{F}\left(e^{\zeta/\sqrt{m}}b(n/m)\right)/\bar{F}(b(n/m))\right). \end{aligned} \tag{2.9}$$

By arguments leading up to (2.2) in Smith (1982),

$$\bar{F}(b(n/m)) = \frac{m}{n}\left[1 + \mathscr{O}(g(b(n/m)))\right]. \tag{2.10}$$

Also by Theorem 2.3, uniformly in n,

$$\begin{aligned} &\bar{F}\left(e^{\zeta/\sqrt{m}}b(n/m)\right)/\bar{F}(b(n/m)) \\ &\quad = e^{-\alpha\zeta/\sqrt{m}}L\left(e^{\zeta/\sqrt{m}}b(n/m)\right)/L(b(n/m)) \\ &\quad = e^{-\alpha\zeta/\sqrt{m}}\left(1 + \mathscr{O}(g(b(n/m)))\right), \end{aligned}$$

since $\{e^{\zeta/\sqrt{m}}: m \ge 1\} \subset [1, e^{\zeta}]$. Thus (2.9) is equivalent to

$$\begin{aligned} \alpha^{-1}\sqrt{m}\left(S_n^{(\zeta)} - \mathscr{E}S_n^{(\zeta)}\right) &\le \alpha^{-1}\sqrt{m}\left[1 - e^{-\alpha\zeta/\sqrt{m}}\left(1 + \mathscr{O}(g(b(n/m)))\right)\right] \\ &= \alpha^{-1}\sqrt{m}\left[\alpha\zeta/\sqrt{m} + \mathscr{O}(1/m) + \mathscr{O}(g(b(n/m)))\right] \\ &= \zeta + o(1), \end{aligned}$$

since $\sqrt{m}\,g(b(n/m)) \to 0$. Consequently, (2.4) implies that for each $(y, \zeta) \in \mathbb{R}^2$

at which $P[Y \le y, Z \le \zeta]$ is continuous,

$$\lim_{n\to\infty} P\left[\sqrt{m}\,(H_n^+ - \mathscr{E}H_n^+) \le y, \sqrt{m}\left(X^*_{(m+1)} - b^*(n/m)\right) \le \zeta\right]$$
$$= \lim_{n\to\infty} P\left[\sqrt{m}\,(H_n^+ - \mathscr{E}H_n^+) \le y, \alpha^{-1}\sqrt{m}\left(S_n^{(\zeta)} - \mathscr{E}S_n^{(\zeta)}\right) \le \zeta + o(1)\right]$$
$$= P[Y \le y, Z \le \zeta].$$

This shows (2.5). It follows from (2.6) and the proof of Theorem 2.2 that $X^*_{([\rho m])} - b^*(n/\rho m) \to_P 0$, $\rho \in I$. Thus (2.5) and Lemma 2.1 imply (2.7). Moreover since $H_n = \tilde{H}_n - (X^*_{(m+1)} - b^*(n/m))$, (2.7) and the continuous mapping theorem implies $\sqrt{m}\,(H_n - \mathscr{E}H_n^+) \to_d Y - Z$. By (2.10) and Theorem 2.3 we have

$$\sqrt{m}\left(\mathscr{E}H_n^+ - \alpha^{-1}\right) = \sqrt{m}\left(\frac{n}{m}\bar{F}(b(n/m))\int_1^\infty y^{-1}\frac{\bar{F}(yb(n/m))}{\bar{F}(b(n/m))}\,dy - \alpha^{-1}\right)$$
$$= \sqrt{m}\left((1 + \mathscr{O}(g(b(n/m))))\right.$$
$$\left.\times\int_1^\infty y^{-\alpha-1}\frac{L(yb(n/m))}{L(b(n/m))}\,dy - \alpha^{-1}\right)$$
$$= \sqrt{m}\,\mathscr{O}(g(b(n/m))) \to 0,$$

showing (2.8) under (2.3a). Next suppose (2.3b) holds. Then the above proof still applies with the following modification. By arguments leading up to (3.1) of Smith (1982),

$$\bar{F}(b(n/m)) = \frac{m}{n}\left[1 + o(g(b(n/m)))\right].$$

Applying this and Theorem 2.3, the conclusions of the theorem can be established in a similar manner as in the proof under (2.3a). □

3. Some limit theorems for stationary sequences. In this section, we assume that $\{X_j\}$ is strictly stationary with "short range" dependence; that is to say that the finite-dimensional distributions of $\{X_j\}$ are invariant under shifts, and the dependence between observations from $\{X_j\}$ becomes weaker as time separation becomes larger. The dependence structure will be described in more detail shortly. The goal of this section is to show, in this setting, how the sufficient conditions in Theorems 2.2 and 2.4 can be verified. In dealing with dependence, we follow the style of Ibragimov and Linnik (1969), Section 18.4. The approach that we take is by no means optimal for every specific situation. In practice, any information on $\{X_j\}$ should be taken into account when studying problems of this nature. Further it is useful to note that the setting being investigated in this section is not the only dependent setting where the results in Section 2 are applicable; indeed, even the stationarity assumption can be relaxed. The content of this section should therefore be seen as a broad description of what can be expected and a general guideline of how to proceed in the problem of tail index estimation when dependence is present.

Suppose Y_{ni} is a functional of X_i; for example, Y_{ni} may be $(X_i^* - b^*(n/m_n))_+$ or $I(X_i^* > b^*(n/m_n) + \varepsilon)$, etc. Let $\mathscr{F}_a^b\{Y_{ni}\}$ be the σ-field $\sigma\{Y_{ni}: a \le i \le b\}$, and for $1 \le l \le n-1$, let

$$\beta(l; \{Y_{ni}\}) := \sup\big\{|P(A \cap B) - P(A)P(B)|: \\ A \in \mathscr{F}_1^k\{Y_{ni}\},\, B \in \mathscr{F}_{k+l}^n\{Y_{ni}\},\, 1 \le k \le n-l\big\}.$$

We first prove a weak law of large numbers under some quite general conditions.

THEOREM 3.1. *Suppose $\{r_n = o(n)\}$ is a sequence of positive integers and S_{nk} is a random variable measurable with respect to $\mathscr{F}_{(k-1)r_n+1}^{kr_n}\{Y_{ni}\}$, $1 \le k \le k_n$, where $k_n = [n/r_n]$. Assume that*

(a) $k_n\beta(r_n; \{Y_{ni}\}) \to 0$,
(b) $m_n^{-1}\sum_{k=1}^{k_n}\mathscr{E}|S_{nk}|I(|S_{nk}| > m_n) \to 0$,
(c) $m_n^{-2}\sum_{k=1}^{k_n}\mathscr{E}S_{nk}^2 I(|S_{nk}| \le m_n) \to 0$.

Then

$$m_n^{-1}\sum_{k=1}^{k_n}(S_{nk} - \mathscr{E}S_{nk}) \to_P 0.$$

PROOF. Write

$$m_n^{-1}\sum_{k=1}^{k_n}(S_{nk} - \mathscr{E}S_{nk}) = m_n^{-1}\sum_{k=1,\, k \text{ odd}}^{k_n}(S_{nk} - \mathscr{E}S_{nk}) \\ + m_n^{-1}\sum_{k=1,\, k \text{ even}}^{k_n}(S_{nk} - \mathscr{E}S_{nk}).$$

We shall only prove that the first piece tends to 0, since the proof for the second piece tending to 0 is the same. Denote by $\mathscr{O}_n$ the set of odd numbers in $\{1, 2, \dots, k_n\}$. It follows from Theorem 17.2.1 of Ibragimov and Linnik (1969) and the triangle inequality that

$$\left|\mathscr{E}\exp\left(\frac{\mathbf{i}t}{m_n}\sum_{k\in\mathscr{O}_n}S_{nk}\right) - \prod_{k\in\mathscr{O}_n}\mathscr{E}\exp\left(\frac{\mathbf{i}t}{m_n}S_{nk}\right)\right| \\ \le \frac{k_n+1}{2}16\beta(r_n; \{Y_{ni}\}),$$

which tends to 0 by (a), where $\mathbf{i}$ denotes the imaginary unit. This implies that one can proceed by assuming that S_{nk}, $k \in \mathscr{O}_n$, are independent random variables, which we do from now on. Define

$$\tilde{S}_{nk} := S_{nk}I(|S_{nk}| \le m_n), \qquad 1 \le k \le k_n.$$

For each $\varepsilon > 0$,

$$P\left[m_n^{-1}\left|\sum_{k\in\mathcal{O}_n}\left(S_{nk}-\mathscr{E}\tilde{S}_{nk}\right)\right|>\varepsilon\right]$$

$$\leq P\left[S_{nk}\neq\tilde{S}_{nk}\text{ for some }k\in\mathcal{O}_n\right]+P\left[m_n^{-1}\left|\sum_{k\in\mathcal{O}_n}\left(\tilde{S}_{nk}-\mathscr{E}\tilde{S}_{nk}\right)\right|>\varepsilon\right]$$

$$\leq\sum_{k\in\mathcal{O}_n}P[|S_{nk}|>m_n]+m_n^{-2}\varepsilon^{-2}\sum_{k\in\mathcal{O}_n}\operatorname{Var}\left(\tilde{S}_{nk}\right).$$

By (b) and (c) the right-hand side tends to 0, and hence the result follows from (b) which implies

$$m_n^{-1}\sum_{k=1}^{k_n}\left(\mathscr{E}S_n-\mathscr{E}\tilde{S}_n\right)\to 0. \qquad \square$$

THEOREM 3.2. *All three quantities H_n, $\tilde{H}_n$ and H_n^+ converge to α^{-1} in probability under the following conditions.*

(i) *There exists a sequence $\{r_n = o(n)\}$ of positive integers such that $k_n\beta(r_n;\{Y_{ni}\})\to 0$, where $k_n=[n/r_n]$ and $Y_{ni}:=(X_i^*-b^*(n/m_n))_+$, and that* (b) *and* (c) *of Theorem* 3.1 *hold for $S_{nk}:=\sum_{i=(k-1)r_n+1}^{kr_n}Y_{ni}$.*

(ii) *For each $\varepsilon\in\mathbb{R}$ and ρ in some interval containing* 1, *there exists a sequence $\{r_n=o(n)\}$ of positive constants such that $k_n\beta(r_n;\{I_{ni}\})\to 0$, where $k_n=[n/r_n]$ and $I_{ni}:=I(X_i^*>b^*(n/\rho m_n)+\varepsilon)$, and that* (b) *and* (c) *of Theorem* 3.1 *hold for $S_{nk}:=\sum_{i=(k-1)r_n+1}^{kr_n}I_{ni}$.*

PROOF. It follows from Theorem 3.1 that the condition (i) implies that

$$m_n^{-1}\sum_{k=1}^{k_n}\sum_{i=(k-1)r_n+1}^{kr_n}(Y_{ni}-\mathscr{E}Y_{ni})=m_n^{-1}\sum_{i=1}^{k_nr_n}(Y_{ni}-\mathscr{E}Y_{ni})\to_P 0.$$

Since $m_n^{-1}\sum_{i=k_nr_n+1}^{n}Y_{ni}$ is a positive quantity having an expectation tending to 0, we conclude that $m_n^{-1}\sum_{i=1}^{n}(Y_{ni}-\mathscr{E}Y_{ni})\to_P 0$. In precisely the same manner, (ii) implies that $m_n^{-1}\sum_{i=1}^{n}(I_{ni}-\mathscr{E}I_{ni})\to_P 0$ for all $\varepsilon\in\mathbb{R}$ and ρ in some neighborhood containing 1. The conclusion of the Theorem thus follows readily from Theorem 2.2. $\square$

Let

$$Y_{ni}=\left(X_i^*-b^*(n/m_n)\right)_+\quad\text{and}\quad Y_{ni}^{(\zeta)}=I\left(X_i^*-b^*(n/m_n)>\zeta/\sqrt{m_n}\right).$$

THEOREM 3.3. *Assume that the slowly varying component L in* (1.1b) *satisfies either* (2.3a) *or* (2.3b) *in Theorem* 2.4. *Suppose there exist positive integers r_n with $r_n\to\infty$ and $n/r_n^2\to\infty$, and constants χ, ψ and ω such that for* *all* $\zeta \in \mathbb{R}$ *and* $0 < \varepsilon < 1$,

(a) $$k_n \beta([\varepsilon r_n]; \{Y_{ni}, Y_{ni}^{(\zeta)}\}) \to 0,$$

(b) $$\frac{2\alpha^2 k_n}{\varepsilon m_n} \sum\sum_{1 \le i < j \le [\varepsilon r_n]}^{*} \operatorname{Cov}(Y_{ni}, Y_{nj}) \to \chi,$$

(c) $$\frac{\alpha k_n}{\varepsilon m_n} \sum\sum_{1 \le i < j \le [\varepsilon r_n]} \left(\operatorname{Cov}(Y_{ni}, Y_{nj}^{(\zeta)}) + \operatorname{Cov}(Y_{ni}^{(\zeta)}, Y_{nj})\right) \to \psi,$$

(d) $$\frac{2k_n}{\varepsilon m_n} \sum\sum_{1 \le i < j \le [\varepsilon r_n]} \operatorname{Cov}(Y_{ni}^{(\zeta)}, Y_{nj}^{(\zeta)}) \to \omega,$$

(e) $$\frac{k_n}{m_n} \mathscr{E}(W_n(\delta))^2 I\left(|W_n(\delta)| > \tau\sqrt{m_n}\right) \to 0 \quad \text{for all } \delta, \tau > 0,$$

where $k_n = [n/r_n]$ *and* $W_n(\delta) = \sum_{i=1}^{[\varepsilon r_n]}[(Y_{ni} - \mathscr{E}Y_{ni}) + (\delta/\alpha)(Y_{ni}^{(\zeta)} - \mathscr{E}Y_{ni}^{(\zeta)})]$. *Then for all* $\zeta \in \mathbb{R}$,

$$\begin{aligned} &\sqrt{m_n}\left(H_n^+ - \mathscr{E}H_n^+, \alpha^{-1}(S_n^{(\zeta)} - \mathscr{E}S_n^{(\zeta)})\right) \\ &\qquad \to_d N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \alpha^{-2}\begin{pmatrix} 2+\chi & 1+\psi \\ 1+\psi & 1+\omega \end{pmatrix}\right), \end{aligned} \tag{3.1}$$

where $S_n^{(\zeta)} := m_n^{-1}\sum_{i=1}^{n} Y_{ni}^{(\zeta)}$.

Proof. It clearly suffices to show

$$\begin{aligned} &\sqrt{m_n}\left((H_n^+ - \mathscr{E}H_n^+) + \frac{\delta}{\alpha}(S_n^{(\zeta)} - \mathscr{E}S_n^{(\zeta)})\right) \\ &\qquad \to_d N\left(0, \alpha^{-2}((2+\chi) + 2\delta(1+\psi) + \delta^2(1+\omega))\right) \end{aligned}$$

for each δ. Let $0 < \varepsilon < 1$ be a fixed constant. Define

$$\begin{aligned} \mathscr{I}_{ni} &= \{(i-1)r_n + 1, \dots, (i-1)r_n + [(1-\varepsilon)r_n]\}, && 1 \le i \le k_n, \\ \mathscr{I}'_{ni} &= \{(i-1)r_n + [(1-\varepsilon)r_n] + 1, \dots, ir_n\}, && 1 \le i \le k_n, \\ \mathscr{I}_n &= \{k_n r_n + 1, \dots, n\}. \end{aligned}$$

Define also

$$\begin{aligned} U_{ni} &= \frac{1}{\sqrt{m_n}} \sum_{j \in \mathscr{I}_{ni}} \left((Y_{nj} - \mathscr{E}Y_{nj}) + \frac{\delta}{\alpha}(Y_{nj}^{(\zeta)} - \mathscr{E}Y_{nj}^{(\zeta)})\right), && 1 \le i \le k_n, \\ U'_{ni} &= \frac{1}{\sqrt{m_n}} \sum_{j \in \mathscr{I}'_{ni}} \left((Y_{nj} - \mathscr{E}Y_{nj}) + \frac{\delta}{\alpha}(Y_{nj}^{(\zeta)} - \mathscr{E}Y_{nj}^{(\zeta)})\right), && 1 \le i \le k_n, \\ R_n &= \frac{1}{\sqrt{m_n}} \sum_{j \in \mathscr{I}_n} \left((Y_{nj} - \mathscr{E}Y_{nj}) + \frac{\delta}{\alpha}(Y_{nj}^{(\zeta)} - \mathscr{E}Y_{nj}^{(\zeta)})\right). \end{aligned}$$

Thus,

$$(3.2)\quad \sqrt{m_n}\left((H_n^+ - \mathscr{E}H_n^+) + \frac{\delta}{\alpha}\left(S_n^{(\zeta)} - \mathscr{E}S_n^{(\zeta)}\right)\right) = \sum_{i=1}^{k_n} (U_{ni} + U'_{ni}) + R_n.$$

We first consider the convergence of $\sum_{j=1}^{k_n} U_{ni}$. By a simple argument in conjunction with condition (a) (cf. the proof of Theorem 3.1), we can proceed as if the U_{ni} are i.i.d. r.v.'s. The sum of the variances of the U_{ni} is derived as follows.

$$k_n \operatorname{Var}(U_{ni}) = \frac{k_n}{m_n}\left(\operatorname{Var}\sum_{j=1}^{[(1-\varepsilon)r_n]} Y_{nj} + \frac{\delta^2}{\alpha^2}\operatorname{Var}\sum_{j=1}^{[(1-\varepsilon)r_n]} Y_{nj}^{(\zeta)} + 2\frac{\delta}{\alpha}\operatorname{Cov}\left(\sum_{j=1}^{[(1-\varepsilon)r_n]} Y_{nj}, \sum_{j=1}^{[(1-\varepsilon)r_n]} Y_{nj}^{(\zeta)}\right)\right),$$

where, by (1.5), (b), (c) and (d),

$$\frac{k_n}{m_n}\operatorname{Var}\sum_{j=1}^{[(1-\varepsilon)r_n]} Y_{nj} = \frac{k_n}{m_n}\left([(1-\varepsilon)r_n]\operatorname{Var}(Y_{n1}) + 2\sum\sum_{1\le i<j\le[(1-\varepsilon)r_n]} \operatorname{Cov}(Y_{ni}, Y_{nj})\right)$$
$$\to \alpha^{-2}(1-\varepsilon)(2+\chi),$$

$$\frac{\delta^2 k_n}{\alpha^2 m_n}\operatorname{Var}\sum_{j=1}^{[(1-\varepsilon)r_n]} Y_{nj}^{(\zeta)} = \frac{\delta^2 k_n}{\alpha^2 m_n}\left([(1-\varepsilon)r_n]\operatorname{Var}(Y_{n1}^{(\zeta)}) + 2\sum\sum_{1\le i<j\le[(1-\varepsilon)r_n]} \operatorname{Cov}\left(Y_{ni}^{(\zeta)}, Y_{nj}^{(\zeta)}\right)\right)$$
$$\to \delta^2\alpha^{-2}(1-\varepsilon)(1+\omega),$$

$$2\frac{\delta k_n}{\alpha m_n}\operatorname{Cov}\left(\sum_{j=1}^{[(1-\varepsilon)r_n]} Y_{nj}, \sum_{j=1}^{[(1-\varepsilon)r_n]} Y_{nj}^{(\zeta)}\right) = 2\frac{\delta k_n}{\alpha m_n}\left([(1-\varepsilon)r_n]\operatorname{Cov}\left(Y_{n1}, Y_{n1}^{(\zeta)}\right) + \sum\sum_{1\le i<j\le[(1-\varepsilon)r_n]}\left(\operatorname{Cov}\left(Y_{ni}, Y_{nj}^{(\zeta)}\right) + \operatorname{Cov}\left(Y_{ni}^{(\zeta)}, Y_{nj}\right)\right)\right)$$
$$\to 2\delta\alpha^{-2}(1-\varepsilon)(1+\psi).$$

Thus $k_n \operatorname{Var}(U_{ni}) \to \alpha^{-2}(1-\varepsilon)((2+\chi) + 2\delta(1+\psi) + \delta^2(1+\omega))$. Using this and (e), we conclude readily from the Lindeberg central limit theorem that

$$(1-\varepsilon)^{-1/2}\sum_{i=1}^{k_n} U_{ni} \to_d N\big(0, \alpha^{-2}((2+\chi)+2\delta(1+\psi)+\delta^2(1+\omega))\big).$$

In exactly the same way, $\varepsilon^{-1/2}\sum_{i=1}^{k_n} U'_{ni}$ can be shown to converge to the same distribution. By the Cauchy–Schwarz inequality and the assumption that $r_n^2/n \to 0$, $\mathrm{Var}(R_n)$ is seen to tend to 0 and hence $R_n \to_P 0$. Since $0 < \varepsilon < 1$ is arbitrary and in view of (3.2), (3.1) follows from letting $\varepsilon \to 0$. This concludes the proof. □

COROLLARY 3.4. *Under the assumptions stated in Theorem* 3.3,

$$\sqrt{m_n}\big(H_n^+ - \mathscr{E}H_n^+, X^*_{(m_n+1)} - b^*(n/m_n)\big) \to_d N\left(\begin{pmatrix}0\\0\end{pmatrix}, \alpha^{-2}\begin{pmatrix}2+\chi & 1+\psi\\ 1+\psi & 1+\omega\end{pmatrix}\right). \tag{3.3}$$

Suppose further that condition (ii) *of Theorem* 3.2 *holds. Then*

$$\sqrt{m_n}\big(\tilde{H}_n - \mathscr{E}H_n^+, X^*_{(m_n+1)} - b^*(n/m_n)\big) \to_d N\left(\begin{pmatrix}0\\0\end{pmatrix}, \alpha^{-2}\begin{pmatrix}2+\chi & 1+\psi\\ 1+\psi & 1+\omega\end{pmatrix}\right) \tag{3.4}$$

and

$$\sqrt{m_n}\big(H_n - \alpha^{-1}\big) \to_d \begin{cases} N\big(0, \alpha^{-2}(1+\chi+\omega-2\psi)\big), & \textit{under } (2.3\text{a}),\\ N\big(KA/\alpha(\alpha-\gamma), \alpha^{-2}(1+\chi+\omega-2\psi)\big), & \textit{under } (2.3\text{b}). \end{cases} \tag{3.5}$$

PROOF. (3.3) follows readily from Theorems 2.4 and 3.3. Further, condition (ii) of Theorem 3.2 implies that (2.6) holds for all $\varepsilon \in \mathbb{R}$ and ρ in some interval containing 1. Hence (3.4) and (3.5) are also simple consequences of Theorems 2.4 and 3.3. □

If additional information on $\{X_j\}$ is available, the conditions in Theorem 3.3 may be considerably simplified. The following is an example.

THEOREM 3.5. *Let $\{X_j\}$ be a strictly stationary l-dependent sequence, where l is a positive integer, and the slowly varying component L in* (1.1b) *satisfies either* (2.3a) *or* (2.3b). *Suppose there exist constants χ, ψ and ω such that for* *all* $\zeta \in \mathbb{R}$,

$$\text{(a)} \qquad 2\alpha^2 \frac{n}{m_n} \sum_{j=2}^{l} \mathscr{E} Y_{n1} Y_{nj} \to \chi,$$

$$\text{(b)} \qquad \alpha \frac{n}{m_n} \sum_{j=2}^{l} \left(\mathscr{E} Y_{n1} Y_{nj}^{(\zeta)} + \mathscr{E} Y_{n1}^{(\zeta)} Y_{nj} \right) \to \psi,$$

$$\text{(c)} \qquad 2 \frac{n}{m_n} \sum_{j=2}^{l} \mathscr{E} Y_{n1}^{(\zeta)} Y_{nj}^{(\zeta)} \to \omega.$$

Then (3.3), (3.4) *and* (3.5) *hold.*

PROOF. By Corollary 3.4 it suffices to show that the conditions (a)–(e) of Theorem 3.3 hold for any $\{r_n\}$ such that $r_n = o(m_n)$ and $r_n \to \infty$, and that condition (ii) of Theorem 3.2 holds for $r_n = l$.

First let $\{r_n\}$ be such that $r_n = o(m_n)$ and $r_n \to \infty$. Condition (a) of Theorem 3.3 holds trivially by l-dependence. To show condition (b) there, note that

$$2\alpha^2 \frac{k_n}{\varepsilon m_n} \sum_{1 \le i < j \le [\varepsilon r_n]} \sum \operatorname{Cov}(Y_{ni}, Y_{nj})$$

$$= 2\alpha^2 \frac{k_n}{\varepsilon m_n} \sum_{i=1}^{[\varepsilon r_n]} \sum_{j=i+1}^{i+l-1} \operatorname{Cov}(Y_{ni}, Y_{nj}) - 2\alpha^2 \frac{k_n}{\varepsilon m_n} \sum_{j=2}^{l} (j-1) \operatorname{Cov}(Y_{n1}, Y_{nj}),$$

where the first term is equal to

$$2\alpha^2 k_n \frac{[\varepsilon r_n]}{\varepsilon m_n} \sum_{j=2}^{l} \operatorname{Cov}(Y_{n1}, Y_{nj}) \to \chi,$$

and the second term is bounded in absolute value by [cf. (1.5)]

$$2\alpha^2 \frac{k_n}{\varepsilon m_n} (l-1)^2 \operatorname{Var}(Y_{n1}) \to 0.$$

This proves condition (b) of Theorem 3.3, and (c) and (d) there are proved in a similar manner. Next we show condition (e) of Theorem 3.3. Let $W_n(\delta)$ be as defined there. Clearly,

$$\frac{k_n}{m_n} \mathscr{E} W_n^2(\delta) I\left(|W_n(\delta)| > \tau \sqrt{m_n} \right) \le \frac{k_n}{\tau^2 m_n^2} \mathscr{E} W_n^4(\delta).$$

By l-dependence, the right-hand side is bounded by

$$\frac{k_n}{\tau^2 m_n^2} \mathscr{O}(r_n^2) \mathscr{E} \left[(Y_{n1} - \mathscr{E} Y_{n1}) + \frac{\delta}{\alpha} \left(Y_{n1}^{(\zeta)} - \mathscr{E} Y_{n1}^{(\zeta)} \right) \right]^4,$$

which tends to 0 by the choice of r_n. Next we let $r_n = l$ and observe that condition (ii) of Theorem 3.2 holds for this choice—it suffices to show that

$$\frac{k_n}{m_n^2}\mathscr{E}\left(\sum_{i=1}^{l} I\left(X_i^* > b^*\left(\frac{n}{\rho m_n}\right) + \varepsilon\right)\right)^2 \to 0$$

which is most straightforward. □

If $\{X_i\}$ is i.i.d., then the parameters χ, ψ and ω in Theorem 3.5 are equal to 0, and (3.5) coincides with a familiar result in the literature [cf. Davis and Resnick (1984)]. In view of Theorem 3.1, it is seen that χ, ψ and ω can be consistently estimated once an appropriate restriction on the dependence of $\{X_i\}$ is satisfied, making it possible to construct confidence intervals of α. To see how well the asymptotic results work, we simulated actual coverage probabilities of confidence intervals of α^{-1} for the moving average sequence

$$X_i = Z_i + 2Z_{i+1}, \qquad i = 1, 2, \dots,$$

where the Z_i are i.i.d. Cauchy (and hence $\alpha = 1$). For a large number n and some properly chosen number m, an approximate $(1-q)100\%$ confidence interval of α^{-1} suggested by Theorem 3.5 is

$$\hat{\alpha}^{-1} \pm z_{q/2}\left(\frac{1+\hat{\chi}+\hat{\omega}-2\hat{\psi}}{m\hat{\alpha}^2}\right)^{1/2}, \tag{3.6}$$

where

$$\hat{\alpha}^{-1} = H_n = m^{-1}\sum_{j=1}^{m} X_{(j)}^* - X_{(m+1)}^*,$$

$$\hat{\chi} = 2\hat{\alpha}^2 m^{-1}\sum_{i=1}^{n-1}\left(X_i^* - X_{(m+1)}^*\right)_+\left(X_{i+1}^* - X_{(m+1)}^*\right)_+,$$

$$\hat{\psi} = \hat{\alpha} m^{-1}\sum_{i=1}^{n-1}\Big\{\left(X_i^* - X_{(m+1)}^*\right)_+ I\left(X_{i+1} > X_{(m+1)}\right) + \left(X_{i+1}^* - X_{(m+1)}^*\right)_+ I\left(X_i > X_{(m+1)}\right)\Big\},$$

$$\hat{\omega} = 2m^{-1}\sum_{i=1}^{n-1} I\left(X_i > X_{(m+1)}, X_{i+1} > X_{(m+1)}\right).$$

Table 1 consists of coverage probabilities of (3.6) for $n = 250$, $m =$

TABLE 1

	m										
q	**50**	**60**	**70**	**80**	**90**	**100**	**110**	**120**	**130**	**140**	**150**
0.1	0.8395	0.8490	0.8490	0.8595	0.8805	0.8945	0.9080	0.9055	0.8990	0.8630	0.8005
0.05	0.8915	0.8975	0.9085	0.9245	0.9330	0.9455	0.9555	0.9625	0.9515	0.9385	0.9030
0.01	0.9435	0.9520	0.9605	0.9690	0.9755	0.9820	0.9870	0.9905	0.9950	0.9925	0.9865

TABLE 2

	m									
q	50	100	150	200	250	300	350	400	450	500
0.1	0.8441	0.8403	0.8377	0.8377	0.8499	0.8706	0.8929	0.9093	0.8968	0.8182
0.05	0.8930	0.8957	0.8954	0.8986	0.9057	0.9212	0.9432	0.9543	0.9503	0.9055
0.01	0.9478	0.9538	0.9573	0.9597	0.9652	0.9738	0.9833	0.9893	0.9915	0.9833

$50, 60, \dots, 150$ and $q = 0.1, 0.05, 0.01$, and Table 2 consists of coverage probabilities of (3.6) for $n = 1000$, $m = 50, 100, \dots, 500$ and $q = 0.1, 0.05, 0.01$. The coverage probabilities are obtained with 10,000 simulation runs.

As can be seen, the approximations are quite acceptable for a large range of values of m for both $n = 250$ and 1000. This brings up the important question of how to choose the optimal m for a given data set of size n. To deal with this, one possibility is to again resort to (3.5) which gives the asymptotic mean squared error of Hill's estimator. This is best illustrated by the works of Hall (1982) and Hall and Welsh (1985) [cf. Haeusler and Teugels (1985)] which considered related issues for the case where observations are i.i.d. with tail distribution

$$\bar{F}(x) = Cx^{-\alpha}\left[1 + Dx^{-\beta} + o(x^{-\beta})\right] \quad \text{as } x \to \infty,$$

where $C, \alpha, \beta > 0$ and $D \in \mathbb{R}$. The presence of dependence will undoubtedly make matters more complicated. A study on this topic is under way.

4. Tail index estimation for finite moving averages—an example. In this section we illustrate the central limit theorem, Theorem 3.5, by considering a finite moving average sequence. Despite the simple dependence structure of this sequence, the asymptotic variance of Hill's estimator is a highly nonlinear function of the moving average coefficients and the derivation is less straightforward than one might expect.

Let $Z, Z_1, Z_2, \dots$ be i.i.d. positive r.v.'s with $P[Z > z] = z^{-\alpha}L(z)$, where $\alpha > 0$ and L is slowly varying at ∞ and satisfies either (2.3a) or (2.3b). Define

$$X_j = \sum_{i=1}^{l} a_i Z_{i+j-1}, \qquad j = 1, 2, \dots,$$

where a_i, $1 \le i \le l$, are positive constants. $\{X_i\}$ is a strictly stationary l-dependent sequence. By a result in Section 8.8 of Feller (1971),

$$\frac{P[X_1 > t]}{P[Z > t]} \to \sum_{i=1}^{l} a_i^{\alpha} \quad \text{as } t \to \infty. \tag{4.1}$$

As in the previous sections, let $b(t)$ be the quantile function of X_1 defined by (1.2). Throughout this section, let $\{m = m_n\}$ be a fixed sequence of positive integers satisfying (1.4), and write $b_n = b(n/m)$ for convenience. Also let $\{\varepsilon_n\}$

be a collection of positive numbers such that

$$\varepsilon_n \to 0 \quad \text{and} \quad \sqrt{\frac{n}{m}}\, P[Z > \varepsilon_n b_n] \to 0. \tag{4.2}$$

Such an $\{\varepsilon_n\}$-sequence exists since, by (4.1),

$$\lim_{n\to\infty} \sqrt{\frac{n}{m}}\, P[Z > \varepsilon b_n] = \lim_{n\to\infty} \sqrt{\frac{n}{m}}\, P[X_1 > \varepsilon b_n] \Big/ \sum_{i=1}^{l} a_i^\alpha = 0$$

for any fixed $\varepsilon > 0$.

We first state a basic lemma which is crucial for deriving the covariance of $(X_1^* - b_n^*)_+$ and $(X_j^* - b_n^*)_+$, $2 \le j \le l$. The proof is tedious and is given in the appendix.

LEMMA 4.1. *For positive constant κ and λ, the following can be proved*:

(a) $$\lim_{n\to\infty} \frac{n}{m} \mathscr{E}((\kappa Z)^* - b_n^*)_+ ((\lambda Z)^* - b_n^*)_+ = \alpha^{-2}\rho_1(\kappa, \lambda),$$

(b) $$\lim_{n\to\infty} \frac{n}{m} \mathscr{E}((\kappa Z)^* - b_n^*)_+ I((\lambda Z)^* - b_n^* > \zeta/\sqrt{m}\,) = \alpha^{-1}\rho_2(\kappa, \lambda), \qquad \zeta > 0,$$

(c) $$\lim_{n\to\infty} \frac{n}{m} P\big[(\kappa Z)^* - b_n^* > \zeta/\sqrt{m}\,, (\lambda Z)^* - b_n^* > \zeta/\sqrt{m}\,\big] = \rho_3(\kappa, \lambda), \qquad \zeta > 0,$$

where ρ_1, ρ_2, ρ_3 are bivariate functions on $(0,\infty) \times (0,\infty)$ *defined by*

$$\rho_1(\kappa, \lambda) = \frac{(\kappa \wedge \lambda)^\alpha}{\sum_{i=1}^{l} a_i^\alpha} \left(2 + \alpha \log\left(\frac{\kappa \vee \lambda}{\kappa \wedge \lambda}\right)\right), \tag{4.3}$$

$$\rho_2(\kappa, \lambda) = \frac{(\kappa \wedge \lambda)^\alpha}{\sum_{i=1}^{l} a_i^\alpha} \left(1 + \alpha \left(\log \frac{\kappa}{\lambda}\right)_+\right), \tag{4.4}$$

$$\rho_3(\kappa, \lambda) = \frac{(\kappa \wedge \lambda)^\alpha}{\sum_{i=1}^{l} a_i^\alpha}. \tag{4.5}$$

By the definition of $\{X_j\}$, if $2 \le j \le l$ then the Z_i's which contribute to X_1 or X_j are Z_r, $1 \le r \le l + j - 1$, and the Z_i's that contribute to both X_1 and X_j are Z_r, $j \le r \le l$. We can write

$$\mathscr{E}(X_1^* - b_n^*)_+ (X_j^* - b_n^*)_+ = \mathscr{E}\big((X_1^* - b_n^*)_+ (X_j^* - b_n^*)_+ S_n\big), \tag{4.6}$$

where

$$S_n = I(Z_r \le \varepsilon_n b_n, 1 \le r \le l + j - 1) \tag{4.7}$$
$$+ I(Z_r > \varepsilon_n b_n, Z_s > \varepsilon_n b_n \text{ for some } r \ne s \text{ in } \{1, 2, \dots, l + j - 1\})$$
$$+ \sum_{r=1}^{l+j-1} I_{nr},$$

with

$$I_{nr} = I(Z_r > \varepsilon_n b_n, Z_s \le \varepsilon_n b_n \text{ for all } s \in \{1, 2, \dots, l + j - 1\} \setminus \{r\}).$$

LEMMA 4.2. *For* $2 \le j \le l$,

(a) $$\lim_{n \to \infty} \frac{n}{m} \mathscr{E}(X_1^* - b_n^*)_+ (X_j^* - b_n^*)_+ I_{nr} = 0,$$
$$r \in \{1, 2, \dots, l + j - 1\} \setminus \{j, \dots, l\},$$

(b) $$\lim_{n \to \infty} \frac{n}{m} \mathscr{E}(X_1^* - b_n^*)_+ (X_j^* - b_n^*)_+ I(Z_r \le \varepsilon_n b_n, 1 \le r \le l + j - 1) = 0,$$

(c) $$\lim_{n \to \infty} \frac{n}{m} \mathscr{E}(X_1^* - b_n^*)_+ (X_j^* - b_n^*)_+ I(Z_r > \varepsilon_n b_n, Z_s > \varepsilon_n b_n \textit{ for some } r \ne s$$
$$\textit{in } \{1, 2, \dots, l + j - 1\}) = 0.$$

PROOF. The proof of (a) goes as follows. If $r \in \{1, 2, \dots, l + j - 1\} \setminus \{j, \dots, l\}$, I_{nr} is the event that all of the l components Z_i's contributing to either X_1 or X_j are no greater than $\varepsilon_n b_n$. Let us say that it is X_1 that has the property. Then

$$\begin{aligned}(X_1^* - b_n^*)_+ &= (\log X_1 - \log b_n)_+ \\ &\le \left(\log\left(\sum_{i=1}^{l} a_i \varepsilon_n b_n\right) - \log b_n\right)_+ \\ &= \left(\log\left(\sum_{i=1}^{l} a_i\right) + \log \varepsilon_n\right)_+,\end{aligned}$$

which is equal to 0 from some n on since $\varepsilon_n \to 0$. Thus (a) is proved. The proof of (b) uses exactly the same idea and is omitted. We now show (c) by showing that for $r \ne s$,

$$\frac{n}{m} \mathscr{E}(X_1^* - b_n^*)_+ (X_j^* - b_n^*)_+ I(Z_r > \varepsilon_n b_n, Z_s > \varepsilon_n b_n) \to 0.$$

Applying the Cauchy–Schwarz inequality twice, independence, (1.5) and (4.2), the above left-hand side is bounded by

$$\frac{n}{m}\Big(\mathscr{E}(X_1^* - b_n^*)_+^2(X_j^* - b_n^*)_+^2\Big)^{1/2}\big(\mathscr{E}I^2(Z_r > \varepsilon_n b_n, Z_s > \varepsilon_n b_n)\big)^{1/2}$$

$$= \frac{n}{m}\Big(\mathscr{E}(X_1^* - b_n^*)_+^2(X_j^* - b_n^*)_+^2\Big)^{1/2}P[Z_1 > \varepsilon_n b_n]$$

$$\le \frac{n}{m}\Big(\mathscr{E}(X_1^* - b_n^*)_+^4\mathscr{E}(X_j^* - b_n^*)_+^4\Big)^{1/4}P[Z_1 > \varepsilon_n b_n]$$

$$= \frac{n}{m}\Big(\mathscr{E}(X_1^* - b_n^*)_+^4\Big)^{1/2}P[Z_1 > \varepsilon_n b_n]$$

$$= \mathscr{O}\left(\sqrt{\frac{n}{m}}\right)P[Z_1 > \varepsilon_n b_n] \to 0.$$

This ends the proof. □

LEMMA 4.3. *For* $2 \le j \le r \le l$,

$$\lim_{n\to\infty} \alpha^2 \frac{n}{m}\mathscr{E}(X_1^* - b_n^*)_+ (X_j^* - b_n^*)_+ I_{nr} = \rho_1(a_r, a_{r-j+1}),$$

where ρ_1 *is defined by* (4.3).

PROOF. Since the a_i's and Z_i's are positive, $\liminf_{n\to\infty}(n/m)\mathscr{E}(X_1^* - b_n^*)_+(X_j^* - b_n^*)_+I_{nr}$ is bounded below by

$$\lim_{n\to\infty}\frac{n}{m}\mathscr{E}((a_rZ_r)^* - b_n^*)_+\big((a_{r-j+1}Z_r)^* - b_n^*\big)_+I_{nr}$$

$$= \lim_{n\to\infty}\frac{n}{m}\mathscr{E}((a_rZ_r)^* - b_n^*)_+\big((a_{r-j+1}Z_r)^* - b_n^*\big)_+$$

$$\times I(Z_i \le \varepsilon_n b_n \text{for all } i \in \{1,2,\dots,l+j-1\} \setminus \{r\}),$$

where the equality follows from the arguments in the proof of Lemma 4.2 (a). By independence, (4.1) and Lemma 4.1, the above quantity is equal to

$$\lim_{n\to\infty}\frac{n}{m}\mathscr{E}((a_rZ_r)^* - b_n^*)_+\big((a_{r-j+1}Z_r)^* - b_n^*\big)_+P^{l+j-2}[Z_1 \le \varepsilon_n b_n]$$

$$(4.8)\qquad = \lim_{n\to\infty}\frac{n}{m}\mathscr{E}((a_rZ_r)^* - b_n^*)_+\big((a_{r-j+1}Z_r)^* - b_n^*\big)_+$$

$$= \alpha^{-2}\rho_1(a_r, a_{r-j+1}).$$

Next $\limsup_{n\to\infty}(n/m)\mathscr{E}(X_1^* - b_n^*)_+(X_j^* - b_n^*)_+I_{nr}$ is bounded above by

$$\limsup_{n\to\infty}\frac{n}{m}\mathscr{E}((a_rZ_r + \varepsilon b_n)^* - b_n^*)_+\big((a_{r-j+1}Z_r + \varepsilon b_n)^* - b_n^*\big)_+$$

for any fixed $\varepsilon > 0$, since $\varepsilon_n \to 0$. This quantity can be evaluated as [cf. (A.1)]

$$\limsup_{n\to\infty}\frac{n}{m}\int_{x=0}^{\infty}\int_{y=0}^{\infty}P\big[a_rZ_r + \varepsilon b_n > e^x b_n, a_{r-j+1}Z_r + \varepsilon b_n > e^y b_n\big]\,dx\,dy,$$

which is bounded by

$$\limsup_{n\to\infty} \frac{n}{m} \int_{x=0}^{\infty}\int_{y=0}^{\infty} P\big[a_r Z_r > e^x(1-\varepsilon)b_n,\, a_{r-j+1}Z_r > e^y(1-\varepsilon)b_n\big]\,dx\,dy$$

$$= \limsup_{n\to\infty} \frac{n}{m}\mathscr{E}\bigg(\bigg(\frac{a_r}{1-\varepsilon}Z_r\bigg)^* - b_n^*\bigg)_+\bigg(\bigg(\frac{a_{r-j+1}}{1-\varepsilon}Z_r\bigg)^* - b_n^*\bigg)_+$$

$$= \alpha^{-2}\rho_1\bigg(\frac{a_r}{1-\varepsilon}, \frac{a_{r-j+1}}{1-\varepsilon}\bigg).$$

Since ρ_1 is continuous, as $\varepsilon \to 0$ this upper bound tends to the lower bound given in (4.8) and the result is proved. □

THEOREM 4.4. *For* $2 \le j \le l$,

(a)
$$\lim_{n\to\infty}\frac{n}{m}\mathscr{E}(X_1^* - b_n^*)_+(X_j^* - b_n^*)_+ = \alpha^{-2}\sum_{r=j}^{l}\rho_1(a_r, a_{r-j+1}),$$

$$\lim_{n\to\infty}\frac{n}{m}\mathscr{E}(X_1^* - b_n^*)_+ I\big(X_j^* - b_n^* > \zeta/\sqrt{m}\big) = \alpha^{-1}\sum_{r=j}^{l}\rho_2(a_r, a_{r-j+1}),$$

(b)
$$\lim_{n\to\infty}\frac{n}{m}\mathscr{E}(X_j^* - b_n^*)_+ I\big(X_1^* - b_n^* > \zeta/\sqrt{m}\big) = \alpha^{-1}\sum_{r=j}^{l}\rho_2(a_{r-j+1}, a_r),$$

$\zeta \in \mathbb{R}$,

(c)
$$\lim_{n\to\infty}\frac{n}{m}P\big[X_1^* - b_n^* > \zeta/\sqrt{m},\, X_j^* - b_n^* > \zeta/\sqrt{m}\big] = \sum_{r=j}^{l}\rho_3(a_r, a_{r-j+1}),$$

$\zeta \in \mathbb{R}$.

PROOF. (a) follows readily from (4.6) and Lemmas 4.2 and 4.3. To prove (b) and (c), one can write [cf. (4.6)], for $2 \le j \le l$,

$$\mathscr{E}(X_1^* - b_n^*)_+ I\big(X_j^* - b_n^* > \zeta/\sqrt{m}\big)$$
$$= \mathscr{E}(X_1^* - b_n^*)_+ I\big(X_j^* - b_n^* > \zeta/\sqrt{m}\big)S_n,$$

$$\mathscr{E}(X_j^* - b_n^*)_+ I\big(X_1^* - b_n^* > \zeta/\sqrt{m}\big)$$
$$= \mathscr{E}(X_j^* - b_n^*)_+ I\big(X_1^* - b_n^* > \zeta/\sqrt{m}\big)S_n,$$

$$P\big[X_1^* - b_n^* > \zeta/\sqrt{m},\, X_j^* - b_n^* > \zeta/\sqrt{m}\big]$$
$$= \mathscr{E}I\big(X_1^* - b_n^* > \zeta/\sqrt{m},\, X_j^* - b_n^* > \zeta/\sqrt{m}\big)S_n,$$

where S_n is defined by (4.7), and mimic the steps leading to (a). The modification is quite straightforward and is omitted. □

The following result follows simply from Theorems 3.5 and 4.4.

THEOREM 4.5. *Let m_n be positive integers satisfying* (1.4) *and* $\{X_j\}$ *be the finite moving average sequence under discussion in this section. Then* (3.3)–(3.5) *hold with*

$$\chi = 2\sum_{j=2}^{l}\sum_{r=j}^{l}\rho_1(a_r, a_{r-j+1}),$$

$$\psi = \sum_{j=2}^{l}\sum_{r=j}^{l}\{\rho_2(a_r, a_{r-j+1}) + \rho_2(a_{r-j+1}, a_r)\},$$

$$\omega = 2\sum_{j=2}^{l}\sum_{r=j}^{l}\rho_3(a_r, a_{r-j+1}),$$

where ρ_1, ρ_2 *and* ρ_3 *are defined by* (4.3), (4.4) *and* (4.5), *respectively.*

APPENDIX

PROOF OF LEMMA 4.1. The proof is based on the simple fact that for any nonnegative r.v.'s X, Y,

$$EXY = \int_{x=0}^{\infty}\int_{y=0}^{\infty} P[X > x, Y > y]\,dy\,dx. \tag{A.1}$$

Thus

$$\begin{aligned}
&E((\kappa Z)^* - b_n^*)_+((\lambda Z)^* - b_n^*)_+ \\
&\quad = \int_{x=0}^{\infty}\int_{y=0}^{\infty} P\big[Z > (\kappa^{-1}e^x \vee \lambda^{-1}e^y)b_n\big]\,dy\,dx \\
&\quad = \int_{x=0}^{\infty}\int_{y=0}^{(x+\log(\lambda/\kappa))_+} P\big[Z > \kappa^{-1}e^x b_n\big]\,dy\,dx \\
&\qquad + \int_{y=0}^{\infty}\int_{x=0}^{(y-\log(\lambda/\kappa))_+} P\big[Z > \lambda^{-1}e^y b_n\big]\,dx\,dy \\
&\quad = \int_{x=0}^{\infty}\left(x + \log\frac{\lambda}{\kappa}\right)_+ P\big[Z > \kappa^{-1}e^x b_n\big]\,dx \\
&\qquad + \int_{y=0}^{\infty}\left(y - \log\frac{\lambda}{\kappa}\right)_+ P\big[Z > \lambda^{-1}e^y b_n\big]\,dy.
\end{aligned}$$

Assume henceforth that $\kappa \le \lambda$. The above becomes

$$\int_{x=0}^{\infty}\left(x+\log\frac{\lambda}{\kappa}\right)P\left[Z>\kappa^{-1}e^{x}b_n\right]dx + \int_{y=\log(\lambda/\kappa)}^{\infty}\left(y-\log\frac{\lambda}{\kappa}\right)P\left[Z>\lambda^{-1}e^{y}b_n\right]dy.$$

Letting $x = y - \log(\lambda/\kappa)$ in the second term in the last expression, we get

$$E((\kappa Z)^* - b_n^*)_+((\lambda Z)^* - b_n^*)_+$$

$$\text{(A.2)}\qquad = 2\int_{x=0}^{\infty} xP\left[Z>\kappa^{-1}e^{x}b_n\right]dx + \left(\log\frac{\lambda}{\kappa}\right)\int_{x=0}^{\infty}P\left[Z>\kappa^{-1}e^{x}b_n\right]dx.$$

It follows from (4.1) that $P[Z > b_n] \sim (m/n)(\sum_{i=1}^{l} a_i^{\alpha})^{-1}$. Thus

$$2\int_{x=0}^{\infty} xP\left[Z>\kappa^{-1}e^{x}b_n\right]dx$$

$$\text{(A.3)}\qquad = \int_{x=0}^{\infty}P\left[Z>\kappa^{-1}e^{\sqrt{x}}b_n\right]dx$$

$$\sim \frac{m}{n}\left(\sum a_i^{\alpha}\right)^{-1}\int_{x=0}^{\infty}\left(\kappa^{-1}e^{\sqrt{x}}\right)^{-\alpha}dx = \frac{2}{\alpha^2}\frac{m}{n}\frac{\kappa^{\alpha}}{\sum a_i^{\alpha}}.$$

Similarly,

$$\text{(A.4)}\qquad \int_{x=0}^{\infty}P\left[Z>\kappa^{-1}e^{x}b_n\right]dx \sim \frac{1}{\alpha}\frac{m}{n}\frac{\kappa^{\alpha}}{\sum a_i^{\alpha}}.$$

Condition (a) now follows from combining (A.2), (A.3) and (A.4). Conditions (b) and (c) are proved by basically the same line of techniques as above, and therefore the proofs are not presented. □

Acknowledgments. The author wishes to thank the referees and an Associate Editor for their helpful comments and criticisms.

REFERENCES

BEIRLANT, J. and TEUGELS, J. L. (1989). Asymptotic normality of Hill's estimator. *Extreme Value Theory. Proceedings, Oberwolfach 1987. Lecture Notes in Statist.* **51** 148–155. Springer, Berlin.

BINGHAM, N. H., GOLDIE, C. M. and TEUGELS, J. L. (1987). *Regular Variation*. Cambridge Univ. Press.

CSÖRGŐ, S., DEHEUVELS, P. and MASON, D. M. (1985). Kernel estimates of the tail index of a distribution. *Ann. Statist.* **13** 1050–1077.

DE HAAN, L. (1970). *On Regular Variation and Its Application to the Weak Convergence of Sample Extremes*. Math. Centre Tract **32**. Mathematisch Centrum, Amsterdam.

DAVIS, R. A. and RESNICK, S. I. (1984). Tail estimates motivated by extreme value theory. *Ann. Statist.* **12** 1467–1487.

FELLER, W. (1971). *An Introduction to Probability Theory and Its Applications* **2**, 2nd ed. Wiley, New York.

GOLDIE, C. M. and SMITH, R. L. (1987). Slow variation with remainder: Theory and applications. *Quart. J. Math. Oxford Ser.* (2) **38** 45–71.

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

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

HALL, P. and WELSH, A. H. (1985). Adaptive estimates of parameters of regular variation. *Ann. Statist.* **13** 331–341.

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

IBRAGIMOV, I. A. and LINNIK, Y. V. (1969). *Independent and Stationary Sequences of Random Variables*. Wolters–Noordhoff, Groningen.

LEADBETTER, M. R., LINDGREN, G. and ROOTZÉN, H. (1983). *Extremes and Related Properties of Random Sequences and Processes*. Springer, New York.

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

SENETA, E. (1976). *Regularly Varying Functions. Lecture Notes in Math.* **508**. Springer, New York.

SMITH, R. L. (1982). Uniform rates of convergence in extreme-value theory. *Adv. in Appl. Probab.* **14** 600–622.

SMITH, R. L. (1987). Estimating tails of probability distributions. *Ann. Statist.* **15** 1174–1207.

DEPARTMENT OF STATISTICS
TEXAS A & M UNIVERSITY
COLLEGE STATION, TEXAS 77843-3143