Chapter

Asymptotic Normality of Hill's Estimator under Weak Dependence

Boualam Karima and Berkoun Youcef

Abstract

This note is devoted to the asymptotic normality of Hill's estimator when data are weakly dependent in the sense of Doukhan. The primary results on this setting rely on the observations being strong mixing. This assumption is often the key tool for establishing the asymptotic behavior of this estimator. A number of attempts have been made to relax the assumption of stationarity and mixing. Relaxing this condition, and assuming the weak dependence, we extend the results obtained by Rootzen and Starica. This approach requires less restrictive conditions than the previous results.

Keywords: tail index, Hill's estimator, regularly varying function, linear process, weak dependence

1. Introduction

Extreme value theory (EVT) is a branch of statistics which focus on modeling and measuring extremes events occurring with small probability. Rare events can have severe consequences for human and economic society. The protection against these events is therefore of particular interest. EVT have been extensively applied in various many fields including hydrology, finance, insurance and telecommunications. Unlike most traditional statistical analysis that deal with the center of the underlying distribution, EVT enables us to restrict attention to the behavior of the tails of the distribution which is strongly connected to limiting distribution of extremes values, i.e., maximum or minimum of a sample.

Let $X_1, X_2, \ldots, X_n$ be *i.i.d* random variables with a common distribution F and let $X_{(n)} \leq \ldots \leq X_{(1)}$ the order statistics pertaining to $X_1, X_2, \ldots, X_n$, where $X_{(n)} = \text{Min}(X_i)$ and $X_{(1)} = \text{Max}(X_i)$. Suppose that there exist two normalizing constants a_n, b_n, $(b_n > 0)$ and a nondegenerate distribution function H such that $F^n(b_n x + a_n) \to H(x)$, for every continuity point x of H, then H belongs to one of the three types

- Type I $(\beta > 0)$: $\Phi_\beta(x) = \begin{cases} \exp\left(-x^{-\beta}\right) & \text{if } x > 0 \\ 0 & \text{if } x \leq 0 \end{cases}$

This class is often called the Frechet class of distributions (fat tailed distribution).

- Type II ($\beta < 0$): $\Psi_\beta(x) = \begin{cases} \exp\left(-\left(-x^\beta\right)\right) & \text{if } x < 0 \\ 1 & \text{if } x \geq 0 \end{cases}$

This class is called Weibull class of distributions (short tailed distributions).

- Type III ($\beta = 0$): $\Lambda_0(x) = \exp\left(-e^{-x}\right), \quad \forall x \in \mathbb{R},$

called the Gumbel type (moderate tail).

This result is known as Fisher-Tippett theorem (see [14]) or the extreme value theorem.

These three family of distributions can be nested into a single representation called the generalized extreme value distribution (GEV) and is given by

$$H_\gamma(x) = \exp\left(-(1+\gamma x)^{-\frac{1}{\gamma}}\right), \;\; 1+\gamma x > 0$$

This representation is useful in practice since it nets three types of limiting distributions behavior in one framework.

For a positive $\gamma = \frac{1}{\beta}$, we recover the Frechet distribution with, negative $\gamma = \frac{1}{\beta}$ corresponds to the Weibull type, and the limit case $\gamma \to 0$ describes the Gumbel family. The shape parameter γ governs the tail behavior of the distribution. The extreme value theorem remains true if condition of independence of the rv's is replaced by the requirement that the form a stationary sequences satisfying a weak dependence condition called distributional mixing condition (e.g., Leadbetter et al. [20]).

The problem of estimating the tail index has received much attention and a variety of estimators have been proposed in the literature in the context of *i.i.d* observations, see Hill [16], Pickands [23], Dekkers and De Haan [12]. We focus on the popular Hill estimator (available only for $\beta > 0$) based on the k- upper order statistics and defined as follows.

$$H_{k,n} = \frac{1}{k}\sum_{i=1}^{k} \log \frac{X_{(i)}}{X_{(k+1)}} \tag{1}$$

where $k = k_n$ and $(k_n)_n$ is an intermediate sequence that is, $k_n \to \infty$, $\frac{k_n}{n} \to 0$ as $n \to \infty$ The asymptotic behavior of this estimator has been extensively investigated in the *i.i.d* setup. Mason [22] proved weak consistency of $H_{k,n}$ for any intermediate sequence k_n and Deheuvels et al. [11] derived strong consistency under the condition that $\frac{k}{log(log(n))} \to \infty$, as $n \to \infty$. Under varying conditions on the sequence k_n and the second-order behavior of F, asymptotic normality of $H_{k,n}$ was discussed among others in Hall [15], Davis and Resnick [7], Csörgo and Mason [5], De Haan and Peng [10], De Haan and Resnick [9].

Hill estimator can still be used for dependent data. In this context, we give below the asymptotic behavior of this estimator.

Hsing [19] and Rootzen et al. [26] established the consistency and the asymptotic normality of $H_{k,n}$ under some general conditions for strictly stationary strong mixing sequences. Brito and Freitas [3] also gave a simplified sufficient condition for consistency, appropriate for applications.

Resnick and Starica [24, 25] proves the weak consistency of Hill's estimator for certain class of stationary sequences with heavy tailed observations which can be approximated by m-dependent sequences. Using this result, they also proved consistency and asymptotic normality of this estimator for an infinite order moving average and autoregressive sequences with regularly varying marginal distribution. However, Ling and Peng [21] extend their results to an ARMA model with *i.i.d* residuals, based on the estimated residuals, this method can achieve a smaller asymptotic variance than applying hill's estimator to the original data.

Hill [17] proved that $H_{k,n}$ still asymptotically normal for dependent, heterogeneous processes with extremes that form mixingale sequences and for near-Epoch-dependent process.

Hill [18] extends the results of Resnick and Starica [24] and Ling and Peng [21] to a wide range of filtered time series satisfying β- mixing condition. Without using the strong mixing condition, Zhang and McCormick [28], established the asymptotic normality of Hill's estimator, for shot noise sequence provided some mild conditions on the impulse response function.

As mentioned above, the asymptotic normality of Hill's estimator has so far been proved for dependent data under various mixing conditions, but not for weak dependent which is the aim of this note. This notion of weak dependence is more general than the classical frameworks of mixing, associated sequences and Markovian models. This type of dependence covers a broad range of time series models.

In order to establish the asymptotic behavior of Hill's estimator in this setting, we first extend the result of Rootzen et al. [26] to random variables, which fulfill the weak dependence condition. Secondly, we derive the asymptotic normality of the Hill estimator when the observations are generated by a linear process satisfying the η- weak dependence condition (see Boualam and Berkoun [2]). This result extends the work of Resnick and Starica [24].

The novelty of using weak dependence instead of mixing dependence lies in the fact that conditions ensuring the normality of the Hill estimator are weaker than the existing conditions.

To make the chapter self-contained, we present definitions and some important results that we need in the sequel.

2. Definitions and auxiliary results

2.1 Regularly varying functions

We start with some background theory on regular variation.

2.1.1 Regularly varying

A positive measurable function $1-F$ is called regularly varying function at infinity with index $-\beta, \beta>0$ (written $1-F\in RV_\beta$) if

$$\lim_{t\to+\infty}\frac{1-F(tx)}{1-F(t)}=x^{-\beta},\quad x>0$$

Recall that F belongs to the domain of attraction of $H_{\frac{1}{\beta}}, \beta>0$ if and only if $1-F(x)\in RV_\beta$.

To establish the asymptotic normality of $H_{k,n}$, a second order regular variation is imposed on the survival function distribution $1-F$.

2.1.2 Second-order regular condition

A function $1-F$ is said to be of second-order regular variation with parameter $\rho\le 0$, if there exists a function $g(t)$ having constant sign with $\lim_{t\to+\infty}g(t)=0$ and a constant $c\ne 0$ such that

$$\lim_{t\to+\infty} \frac{\frac{1-F(tx)}{1-F(t)} - x^{-\beta}}{g(t)} = cx^{-\beta}\int_1^x \mu^{\rho-1}d\mu, \quad x>0 \tag{2}$$

Then it is written as $1 - F \in 2RV(-\beta, \rho)$ and $g(t)$ is referred as the auxiliary function of $1 - F$. The convergence in (2) is uniform in x on compact intervals of $(0, +\infty)$.

Under this assumption, de Haan and Peng derived the asymptotic expansion

$$H_{k,n} = \gamma + \frac{\gamma}{\sqrt{k}} Z_k + \frac{g((n/k)}{1-\rho}(1+o_p(1))$$

where $Z_k = \sqrt{k}\left(\frac{\sum_{i=1}^k E_i - 1}{k}\right)$ and E_i is a sequence of i.i.d standard exponential random variables. Hence, choosing k such that $\sqrt{k}g(n/k) = \lambda \neq 0$ leads to asymptotic normality of $\sqrt{k}(H_{k,n} - \gamma)$ with mean $\frac{\lambda}{1-\rho}$ and variance γ^2.

2.2 Strong mixing condition and weak dependence

Several ways of modeling dependence have already been proposed. One of the most popular is the notion of strong mixing introduced by Rosenblatt [27].

2.2.1 Strong mixing

The sequence $(X_n)_n$ is called strongly mixing with mixing coefficient

$$\alpha_{n,l} = \sup\left\{|P(A\cap B) - P(A)P(B)| : A \in \mathcal{F}_{1,p}, B \in \mathcal{F}_{p+l,n}, 1 \le p \le n-l\right\}$$

if

$$\lim_{n\to\infty}\alpha_{n,l} = 0$$

$\mathcal{F}_{i,j}$ is the $\sigma-$field generated by $\left(X_p : i \le p \le j\right)$

It turns out certain classes of processes are not mixing. Inspired by such problems, and in order to generalize mixing and other dependence, Doukhan and Louhichi introduced a new weak dependence condition.

Recall that random variables U, V with values in a measurable space χ are independent if for some rich enough class $\mathcal{F}$ of numerical functions on χ

$$Cov(f(U), g(V)) = 0, \quad \forall f, g \in \mathcal{F}$$

Weakening this assumption leads to definition of weak dependence condition. More precisely, assume that, for convenient functions f and g, $\operatorname{cov}\left(f\left(\text{"past"}\right), g\left(\text{"future"}\right)\right)$ converge to zero as the distance between the "past" and the "future" converge to infinity. Here "past" and "future" refer to the values of the process of interest. This makes explicit the asymptotic dependence between past and future.

Now we describe the notion of weak dependence (in the sense of Doukhan and Louhichi) considered here (see [13]).

2.2.2 Weak dependence

A process $(X_n)_n$ is called $(\varepsilon, \pounds_n, \Psi)$-weakly dependent if there exists a function $\Psi : \mathbb{R}_+^2 \to \mathbb{R}_+$ and a sequence $\varepsilon = (\varepsilon_l)_l \in \mathbb{N}$ decreasing to zero at infinity, such that for any $(h, k) \in \pounds_u \times \pounds_v$ and $(u, v) \in \mathbb{N}^2$

$$Cov\left(h\left(X_{i_1}, \ldots, X_{i_u}\right), k\left(X_{j_1}, \ldots, X_{j_v}\right)\right) \le \Psi(Lip(h), Lip(k), u, v)\varepsilon_l$$

For any $(i_1, \ldots, i_u)$ and $\left(j_1, \ldots, j_v\right)$ with $i_1 < \ldots . < i_u \le i_u + l \le j_1 < \ldots . < j_v$.

$\pounds_n$ denotes the class of real Lipschitz functions, bounded by 1 and defined on $\mathbb{R}^n$ $(n \in \mathbb{N}^*)$. $Lipf$ denotes the Lipschitz modulus of continuity of function f, that is

$$Lip(f) = \sup_{x \neq y} \frac{|f(x) - f(y)|}{\|x - y\|_1}$$

with $\|x - y\|_1 = \sum_{i=1}^{n} |x_i - y_i|$.

Specific functions Ψ yield variants of weak dependence appropriate to describe various examples of models:

- η-weakly dependence for which $\Psi(Lip(h), Lip(k), u, v) = uLip(h) + vLip(k)$
- λ-weakly dependence for which $\Psi(Lip(h), Lip(k), u, v) = uLip(h) + vLip(k) + uvLip(h)Lip(k)$
- κ-weakly dependence for which $\Psi(Lip(h), Lip(k), u, v) = uvLip(h)Lip(k)$
- ζ-weakly dependence for which $\Psi(Lip(h), Lip(k), u, v) = min(u, v)Lip(h)Lip(k)$

Several class of processes satisfy the weak dependence assumption, as the Bernoulli shift, a Gaussian or an associated process, linear process, $GARCH(p, q)$ and $ARCH(\infty)$ processes (more examples and details can be found in the Dedecker et al. [8]).

The coefficients of weak dependence have some hereditary properties. If the sequence $(X_t)_t$ is κ, λ or θ weakly dependent, then for a Lipschitz function h, the sequence $(h(X_t))_t$ is also weakly dependent.

Mixing conditions refer to $\sigma-$algebras rather than to random variables. The main inconvenience of mixing coefficients is the difficulty of checking them. The weak dependence in the sense of Doukhan is measured in terms of covariance which is much easier to compute than mixing coefficients.

3. Asymptotic normality of Hill's estimator under strong mixing condition

In order to proof the asymptotic normality of Hill estimator, we use the approach of Rootzen described in the following.

Let $(Y_n)_n$ be a sequence of stationary strong mixing random variables with mixing coefficients α_{n, l_n} tending to zero at infinity and $l_n = o(n)$. Suppose that the common distribution function F of Y_n is such that

$$\lim_{t \to +\infty} \frac{1 - F(t + x)}{1 - F(t)} = e^{-\beta x}, \quad x \ge 0 \tag{3}$$

i.e., $1 - F(x)$ decays approximately in an exponential manner $e^{-\beta x}$ as $x \to \infty$ or (by log transformations) as an approximate Inverse power law in the sense of regular variation.

Rootzen et al. [26] considered the estimator

$$\beta_n^* = \frac{1}{k}\sum_{i=1}^{k} Y_{(i)} - Y_{(k)} \tag{4}$$

Where $Y_{(1)}, Y_{(2)}, \ldots, Y_{(n)}$ are the order statistics pertaining to a sample $Y_1, Y_2, \ldots, Y_n$.

Under certain conditions, they proved that

$$\sqrt{\frac{k_n}{\lambda_n}}\left(\beta_n^* - \beta_n\right) \xrightarrow{d} N(0,1) \tag{5}$$

where $\beta_n = \frac{n}{k_n} E(Y_1 - u_n)_+$ and $\lambda_n = \frac{n}{k_n r_n} \mathrm{var} \sum_{i=1}^{r_n} \left\{ \left(Y_j - u_n\right) 1_{\{Y_j - u_n \geq 0\}} - \frac{1}{\beta} 1_{\{Y_j - u_n \geq 0\}} \right\}$.

The sequences (u_n), (r_n) are chosen such that $\lim_{n\to\infty} p_n \left(\alpha_{n, l_n} + \frac{l_n}{n}\right) = 0$, $\lim_{n\to\infty} \frac{n(1-F(u_n))}{k_n} = 1$ and

$$r_n = \left[\frac{n}{p_n}\right], \quad \text{with } p_n \to \infty, \quad \frac{r_n}{n} \to 0 \text{ as } n \to \infty \tag{6}$$

Note that if we replace $Y = \log X$ in (13), we find the expression of the Hill estimator.

3.1 Hill's estimator in case of infinite order moving average process

Resnick and Starica [24] generalize the Hill estimator for more general settings with possibly dependent data especially for infinite moving average model and AR(p) process.

For a sequence $(X_n)_n$ of random variables generated by a strong mixing linear process, with common distribution F satisfying the following von Mises condition

$$\lim_{t\to+\infty} \frac{tf(t)}{1 - F(t)} = \beta \tag{7}$$

Resnick and Starica [24] have adopted the approach of Rootzen applied to $(Y_n)_n = \left(logX_n\right)_n$ for proving the normality of Hill's estimator. It is well known that if (15) holds then $1 - F \in RV_{-\beta}$.

Let $(X_t)_t$ be a strictly stationary linear process defined by

$$X_t = \sum_{i\geq 0} c_i \varepsilon_{t-i} \tag{8}$$

$(\varepsilon_t)_t$ is an *i.i.d* sequence of random variables with marginal distribution satisfying

$$\overline{G}(x) = 1 - G(x) = x^{-\beta} l(x), \quad x > 0, \quad \beta > 0 \tag{9}$$

l is a slowly varying function at infinity and $(c_i)_i$ is a sequence of real numbers satisfying certain mild summability conditions.

Throughout this paper, assume that:

$$\sum_{j=0}^{\infty} \left|c_j\right|^{\delta} < \infty, \text{ for some } \ 0 < \delta < 1 \wedge \beta \tag{10}$$

then (Cline [4]) $\sum_{j=0}^{\infty} c_j \varepsilon_j < \infty$ which implies that $\sum_{j=0}^{\infty} |c_j||\varepsilon_j| < \infty$ (Datta and McCormick [6]). Next, assume that

$$\sum_{j=1}^{\infty}\sum_{k=0}^{\infty} |c_{j+k}|^{\beta} \wedge |c_k|^{\beta} \log\left(\frac{|c_{j+k}| \vee |c_k|}{|c_{j+k}| \wedge |c_k|}\right) < \infty \tag{11}$$

then $X_t = \sum_{i \geq 0} c_i \varepsilon_{t-i}$ is regularly varying.

As a direct consequence of the lemma 2.1 for Resnick and Starica [24] we have

$$\lambda_n \to \lambda = \frac{1}{\beta^2}\left(1 + 2\frac{\sum_{j=1}^{\infty}\sum_{k=0}^{\infty} |c_k|^{\beta} \wedge |c_{k+j}|^{\beta}}{\sum_{j=0}^{\infty} |c_k|^{\beta}}\right), \text{as} \, n \to \infty \tag{12}$$

Note that λ is finite and depends only on the coefficients c_j.

Theorem 3.1 *(Resnick and Starica [24]) Let $(X_t)_t$ be a strongly mixing linear process and assume that conditions* (7), (9), (10) *and* (11) *hold. If the intermediate sequence k is such that*

$$\liminf_{n\to\infty} \frac{n}{k^{\frac{3}{2}}} > 0 \qquad or \qquad \limsup_{n\to\infty} \frac{n}{k^{\frac{3}{2}}} < \infty \tag{13}$$

and $1 - F \in 2RV(-\beta, \rho)$ with the auxiliary function g satisfying:

$$\sqrt{n}g\left(b\left(\frac{n}{k}\right)\right) \to 0, \qquad as\, n \to \infty \;\; where\, b\, is\; the\; quantile\; function \tag{14}$$

Then

$$\sqrt{k}\left(H_{k,n} - \frac{1}{\beta}\right) \to N(0, \lambda)$$

Note that the second order condition imposed on F implies condition (11) required by Rootzen (see Rootzen et al. [26], Appendix. p44). Condition (13) on the intermediate sequence k allows us to prove the existence of sequence $(r_n)_n$ previously defined.

3.2 Hill's estimator in case of AR(p)process

Similar result to 3.1 where obtained by Rootzen et al. [26] for $AR(p)$ process. Consider a stationary, pth-order autoregression $(X_t)_t$ satisfying

$$X_t = \sum_{i=1}^{p} \phi_i X_{t-i} + \varepsilon_t, \qquad t \in \mathbb{N} \tag{15}$$

We assume that the common distribution of $i.i.d$ sequence ε_i satisfy condition (9). Under mild conditions the process (15) has a causal representation of the form (8); if these conditions are not verified then the procedure of applying the Hill estimator directly to an autoregressive process is first to estimate the autoregressive coefficients and then estimating β using estimated residuals.

We assume that we have a sequence $\hat{\phi}^{(n)} = \left(\hat{\phi}_1^{(n)}, \ldots, \hat{\phi}_p^{(n)}\right), n \geq 1,$ of consistent estimators for the coefficients of the autoregression such that $d(n)\left(\hat{\phi}^{(n)} - \phi\right) \to S$ where S is nondegenerate random vector and $d(n) \to \infty$. So that $\varepsilon_t - \hat{\varepsilon}_t^{(n)} = \sum_{i=1}^{p}\left(\hat{\phi}_i^{(n)} - \phi_i\right)X_{t-i}$.

Applying the Hill estimator to the estimated residuals $\left|\hat{\varepsilon}_1^{(n)}\right|, \left|\hat{\varepsilon}_2^{(n)}\right|, \ldots, \left|\hat{\varepsilon}_n^{(n)}\right|$, Resnick and Starica [24] obtained that, if the distribution $\overline{G}_{|\varepsilon|} \in 2RV(-\beta, \rho)$ and the sequence k is chosen to satisfy the condition (13) and $\frac{\sqrt{k} b\left(\frac{n}{\sqrt{k}}\right)}{b\left(\frac{n}{k}\right)} = o(d(n))$, as $n \to \infty$ then,

$$\sqrt{k}\left(\frac{1}{k}\sum_{i=1}^{k} \log \frac{\left|\hat{\varepsilon}_{(i)}^{(n)}\right|}{\left|\hat{\varepsilon}_{(k+1)}^{(n)}\right|} - \frac{1}{\beta}\right) \xrightarrow{d} N\left(0, \frac{1}{\beta^2}\right)$$

For AR(p) process, the approach used is quite different than the previous one. Instead, of working with the original observations, the authors used the estimated residuals in order to get the asymptotic normality of Hill's estimator. This method achieves a smaller variance of the Hill estimator than the first one.

4. Asymptotic normality of Hill's estimator under weak dependence

Following the approach of Rootzen et al. [26], we investigate the asymptotic normality of the Hill estimator when the observations are drawn from a causal weakly dependent process in Doukhan sense. In order to check the asymptotic normality of the Hill estimator, we first extend the normality asymptotic of β_n^* defined by (13) for η-weakly dependent random variables. Therefore, applying this to the process $(Y_t)_t$ where $Y_t = \log X_t$, we obtain the desired result.

Let $(Y_n)_n$ be a stationary sequence of random variables η-weakly dependent. We suppose that for each sequences $(p_n)_n$ and $(r_n)_n$ the condition (6) is satisfied and $(l_n)_n$ is such that

$$\lim_{n\to\infty} \frac{l_n}{n} = 0, \ \lim_{n\to\infty} \frac{l_n}{r_n} = 0 \ \text{ and } \ \lim_{n\to\infty} \frac{n l_n^{\frac{1}{2}-\mu}}{r_n} = 0, \ \ \mu > \frac{1}{2} \tag{16}$$

To establish the asymptotic normality of Hill's estimator, we need to show that under suitable conditions and even if the function *logarithm* does not satisfy the conditions of proposition 2.1 of [8], the sequence $(Y_t)_t = (logX_t)_t$ is η-weakly dependent and possess the hereditary property.

Lemma 4.1 *(Boualam and Berkoun [2]) Let $(X_t)_t$ be a stationary sequence of positive random variables η -weakly dependent. Suppose that there exists a constant $C > 0$, such that $\|X_1\|_p \le C$, with $p > 1$ then $(Y_t)_t$ where $Y_t = logX_t$ is also η -weakly dependent with $\eta_Y(r) = \mathcal{O}\left(\eta^{\frac{p}{p-1}}(r)\right)$.*

Let $(X_t)_t$ be a causal linear process given by (8) where

$$c_k = \mathcal{O}(|k|^{-\mu}), \quad \text{with } \ \mu > 1/2 \tag{17}$$

then $(X_t)_t$ is η-weak dependent with $\eta_{l_n} = \mathcal{O}\left(\frac{1}{l_n^{\mu-1/2}}\right)$ (see Bardet et al. [1]).

Now, we extend theorem 4.3 of Rootzen et al. [26] obtained for strong mixing sequences to η-weakly dependent random variables.

Theorem 4.1 *(Boualam and Berkoun [2]) Let $(Y_n)_n$ be a stationary sequence of η-weakly dependent random variables. If condition $\lim_{n\to\infty} p_n\left(\alpha_{n,l_n} + \frac{l_n}{n}\right) = 0$ of theorem 4.3 of Rootzen et al. [26] is replaced by (16), then*

DOI: http://dx.doi.org/10.5772/intechopen.84555

$$\sqrt{\frac{k_n}{\lambda_n}}\left(\beta_n^* - \beta_n\right) \xrightarrow{d} N(0,1) \tag{18}$$

The above results allows us to state our main result which extend the result obtained by Resnick and Starica [24] for strong mixing to weak dependent sequences.

Theorem 4.2 *(Boualam and Berkoun [2]) Let* $(X_t)_t$ *be a linear process given by* (16) *with common distribution F, satisfying assumptions* (7), (10), (11), (13), (14), (16) *and* (17) *then*

$$\sqrt{k}\left(H_{k,n} - \frac{1}{\beta}\right) \to N(0,\lambda)$$

5. Conclusion

In a primary work, Hsing showed the asymptotic normality of Hill's estimator in a weak dependent setting under suitable mixing and stationary conditions. Similar results have derived for data with several types of dependence or some specific structures. These conditions have been considerably weakened in Hill. We extend the results obtained by Rootzen and Resnick and Starica. The contribution of this note is threefold. First, the weak dependence in the sense of Doukhan is more general than the framework of mixing and several class of processes possesses this type of dependence. It is important to stress that this dependence allows us to prove the asymptotic normality of the Hill estimator without requiring the assumption that the linear process enjoys the strong mixing property. Consequently, the conditions ensuring the asymptotic normality are weakened with our approach. Second, mixing is hard to verify and requires some regularity conditions. However, using weak dependence which focus on covariances is much easier to compute and this assumption is more often checked by several process. Third, our work can be extended to linear process with dependent innovations (under mild conditions, linear process with dependent innovations is η-weak dependent).

Author details

Boualam Karima and Berkoun Youcef*
Laboratoire de Mathématiques Pures et Appliquées, Faculté des Sciences, Université Mouloud Mammeri, Tizi Ouzou, Algeria

*Address all correspondence to: youberk@yahoo.com

References

[1] Bardet JM, Doukhan P, León JF. A functional limit theorem for η-weakly dependent processes and its applications. Statistical Inference for Stochastic Processes. 2008;**11**(3): 265-280

[2] Boualam K, Berkoun Y. Hill's estimator under weak dependence. Communications in Statistics - Theory and Methods. 2017;**46**(18):9218-9229

[3] Brito M, Freitas ACM. Consistent estimation of the tail index for dependent data. Statistics & Probability Letters. 2010;**80**:1835-1843

[4] Cline D. Estimation and linear prediction for regression, autoregression and ARMA with infinite variance data [thesis]. Fort Collins, CO, USA: Deptartment of Statistics, Colorado State University; 1983

[5] Csörgo S, Mason D. Central limit theorems for sums of extreme values. Mathematical Proceedings of the Cambridge Philosophical Society. 1985; **98**:547-558

[6] Datta S, McCormick WP. Inference for the tail parameters of a linear process with heavy tailed innovations. Annals of the Institute of Statistical Mathematics. 1998;**50**(2):337-359

[7] Davis R, Resnick S. Tail estimates motivated by extreme value theory. The Annals of Statistics. 1984;**12**:1467-1487

[8] Dedecker J, Doukhan P, Lang G, León JR, Louhichi S, Prieur C. Weak Dependence: With Examples and Applications. New York: Springer-Verlag; 2007

[9] De Haan L, Resnick SI. On asymptotic normality of the hill estimator. Communications in Statistics. Stochastic Models. 1997;**4**:849-866

[10] De Haan L, Peng L. Comparison of tail index estimators. Statistica Neerlandica. 1998;**52**(1):60-70

[11] Deheuvels P, Haeusler E, Mason DM. Almost sure convergence of the Hill estimator. Mathematical Proceedings of the Cambridge Philosophical Society. 1988;**104**:2, 371-381

[12] Dekkers ALM, De Haan L. On the estimation of the extreme-value index and large quantile estimation. The Annals of Statistics. 1989;**17**:1795-1832

[13] Doukhan P, Louhichi S. A new weak dependence condition and applications to moment inequalities. Stochastic Processes and their Applications. 1999; **84**:313-342

[14] Fisher RA, Tippett LHC. Limiting forms of the frequency distribution of the largest or smallest member of a sample. Proceedings of the Cambridge Philosophical Society. 1928;**24**:180-190

[15] Hall P. On simple estimates of an exponent of regular variation. Journal of the Royal Statistical Society, Series B. 1982;**44**(1):37-42

[16] Hill BM. A simple general approach to inference about the tail of a distribution. The Annals of Statistics. 1975;**3**(5):1163-1174

[17] Hill JB. On tail index estimation for dependent, heterogeneous data. Econometric Theory. 2010;**26**: 1398-1436

[18] Hill JB. Tail index estimation for a filtered dependent time series. Statistica Sinica. 2015;**25**(2):609-629

[19] Hsing T. On tail index estimation using dependent data. The Annals of Statistics. 1991;**19**:1547-1569

[20] Leadbetter MR, Lindgren G, Rootzen H. Extremes and Related Properties of Random Sequences and Processes. New York: Springer Verlag; 1983

[21] Ling S, Peng L. Hill's estimator for the tail index of an ARMA model. Journal of Statistical Planning and Inference. 2004;**123**(2):279-293

[22] Mason DM. Laws of large numbers for sums of extreme values. The Annals of Probability. 1982;**10**(3):754-764

[23] Pickands J. Statistical inference using extreme order statistics. The Annals of Statistics. 1975;**3**:119-131

[24] Resnick S, Starica C. Asymptotic behavior of Hill's estimator for autoregressive data. Communications in Statistics. Stochastic Models. 1997;**13**: 703-723

[25] Resnick S, Starica C. Tail index estimation for dependent data. Annals of Applied Probability. 1998;**8**:1156-1183

[26] Rootzen H, Leadbetter M, De Haan L. Tail and quantile estimation for strongly mixing stationary sequences. Techical report. No. 292. Chapel Hill: Center for Stochastic Processes, Departement of Statistics, University of North Carolina; 1990

[27] Rosenblatt M. A central limit theorem and a strong mixing condition. Proceedings of the National Academy of Sciences of the USA. 1956;**42**(1):43-47

[28] Zhang C, McCormick WP. Asymptotic properties of the tail distribution and Hill's estimator for shot noise sequence. Extremes: Statistical Theory and Applications in Science. 2012;**15**(4):407-435