

On the Generalized Hill Process for Small Parameters and Applications

Gane Samb Lo^{1,2}, El Hadji Deme², and Aliou Diop²

¹ *LSTA, UPMC, France**

² *LERSTAD, Université Gaston-Berger de Saint-Louis, Sénégal*

gane-samb.lo@ugb.edu.sn, ehdeme@ufrsat.org, aliou.diop@ugb.edu.sn

Received 25 November 2011

Accepted 9 September 2012

Let X_1, X_2, \dots be a sequence of independent copies (s.i.c) of a real random variable (r.v.) $X \geq 1$, with distribution function *df* $F(x) = \mathbb{P}(X \leq x)$ and let $X_{1,n} \leq X_{2,n} \leq \dots \leq X_{n,n}$ be the order statistics based on the $n \geq 1$ first of these observations. The following continuous generalized Hill process

$$T_n(\tau) = k^{-\tau} \sum_{j=1}^{j=k} j^\tau (\log X_{n-j+1,n} - \log X_{n-j,n}),$$

$\tau > 0$, $1 \leq k \leq n$, has been introduced as a continuous family of estimators of the extreme value index, and largely studied for statistical purposes with asymptotic normality results restricted to $\tau > 1/2$. We extend those results to $0 < \tau \leq 1/2$ and show that asymptotic normality is still valid for $\tau = 1/2$. For $0 < \tau < 1/2$, we get non Gaussian asymptotic laws which are closely related to the Riemann function $\zeta(s) = \sum_{n=1}^{\infty} n^{-s}$, $s > 1$.

Keywords: Extreme values theory; Asymptotic distribution; Functional Gaussian and nongaussian laws; Uniform entropy numbers; Asymptotic tightness, Stochastic process of estimators of extremal index; Slowly and regularly varying functions.

2010 Mathematics Subject Classifications: Primary 62E20, 62F12, 60F05
Secondary 60B10, 60F17.

1. Introduction

The aim of this note is to settle an asymptotic theory for some functional forms of Hill's estimators. Precisely, let X_1, X_2, \dots be a sequence of independent copies (s.i.c) of a real random variable (r.v.) X with distribution function (*df*) $F(x) = \mathbb{P}(X \leq x)$, $x \in \mathbb{R}$. Since we are only concerned with the upper tail of F in this paper, we assume without loss of generality that $X \geq 1$ and define a s.i.c. of the r.v. $Y = \log X$ denoted Y_1, Y_2, \dots with *df* $G(x) = P(Y \leq x) = F(e^x)$, $x \geq 0$. Finally $Y_{1,n} = \log X_{1,n} \leq \dots \leq Y_{n,n} = \log X_{n,n}$ are their respective order statistics. The following generalized continuous Hill

*ganesamblo@ufrsat.org, www.lsta.upmc.fr

process

$$T_n(\tau) = k^{-\tau} \sum_{j=1}^{j=k} j^\tau (\log X_{n-j+1,n} - \log X_{n-j,n}), \quad (1.1)$$

has been introduced and studied in [6] and [5], for a continuous parameter $\tau > 0$, and as throughout this paper,

$$1 \leq k = k(n) \leq n, \quad k \rightarrow +\infty, \quad k/n \rightarrow 0 \quad \text{as } n \rightarrow +\infty.$$

But some margins of (1.1) appeared before that, since $T_n(1)$ is the Hill estimator introduced in [8] in 1975. Also, De Haan and Resnick [4] proposed $DR_n = T_n(0)/(\log k)$ as a simple estimator of the extreme value index. It has been shown in [10] that for $F \in D(\Lambda)$, DR_n – when appropriately centered and normalized – converges in law to a multiple of a Gumbel law.

However, we are only concerned in this paper with values $\tau > 0$. In this case, the process (1.1) provides a family of estimators of the Univariate Extreme Value Index as shown in [5] in a sense to be defined shortly, after a brief recall on Extreme Value Theory.

Since this work arises in the just mentioned theory, it would be appropriate to make some reminders before going any further. The reader is referred to de Haan [3], de Haan and Ferreira [2], Resnick [12] and Galambos [7] for a modern and large account of the Extreme Value Theory. However, the least to say is that the *df* F is said to be attracted to a non degenerated *df* M iff the maximum $X_{n,n} = \max(X_1, \dots, X_n)$, when appropriately centered and normalized by two sequences of real numbers $(a_n > 0)_{n \geq 0}$ and $(b_n)_{n \geq 0}$, converges to M , in the sense that

$$\lim_{n \rightarrow +\infty} \mathbb{P}(X_{n,n} \leq a_n x + b_n) = \lim_{n \rightarrow +\infty} F^n(a_n x + b_n) = M(x), \quad (1.2)$$

for continuity points x of M . If (1.2) holds, it is said that F is attracted to M or F belongs to the domain of attraction of M , written $F \in D(M)$. It is well-known that the three nondegenerate possible limits in (1.2), called extremal *df*'s, are the following.

The Gumbel *df*

$$\Lambda(x) = \exp(-\exp(-x)), \quad x \in \mathbb{R}, \quad (1.3)$$

the Fréchet *df* with parameter $\gamma > 0$

$$\phi_\gamma(x) = \exp(-x^{-\gamma}) \mathbb{I}_{[0, +\infty[}(x), \quad x \in \mathbb{R} \quad (1.4)$$

and the Weibull *df* with parameter $\beta > 0$

$$\psi_\beta(x) = \exp(-(x)^\beta) \mathbb{I}_{]-\infty, 0]}(x) + (1 - 1_{]-\infty, 0]}(x)), x \in \mathbb{R}, \quad (1.5)$$

where \mathbb{I}_A denotes the indicator function of the set A .

Actually the limiting df M is defined by an equivalence class of the binary relation \mathcal{R} on the set \mathcal{D} of df 's on \mathbb{R} , defined as follows

$$\forall (M_1, M_2) \in \mathcal{D}^2, (M_1 \mathcal{R} M_2) \Leftrightarrow \exists (a, b) \in \mathbb{R}_+ \setminus \{0\} \times \mathbb{R}, \forall (x \in \mathbb{R}) :$$

$$M_2(x) = M_1(ax + b).$$

One easily checks that if $F^n(a_n x + b_n) \rightarrow M_1(x)$, then $F^n(c_n x + d_n) \rightarrow M_1(ax + b) = M_2(x)$ whenever

$$a_n/d_n \rightarrow a \text{ and } (b_n - d_n)/c_n \rightarrow b \text{ as } n \rightarrow \infty. \quad (1.6)$$

These facts allow to parameterize the class of extremal df 's. For this purpose, suppose that (1.2) holds for the three df 's given in (1.3), (1.4) and (1.5). If we take sequences $(c_n > 0)_{n \geq 1}$ and $(d_n)_{n \geq 1}$ such that the limits in (1.6) are $a = \gamma = 1/\alpha$ and $b = 1$ (in the case of Fréchet extremal domain), and $a = -\beta = -1/\alpha$ and $b = -1$ (in the case of Weibull extremal domain), and finally, if we interpret $(1 + \gamma x)^{-1/\gamma}$ as $\exp(-x)$ for $\gamma = 0$ (in the case of Gumbel extremal domain), we are entitled to write the three extremal df 's in the parameterized shape

$$G_\gamma(x) = \exp(-(1 + \gamma x)^{-1/\gamma}), \quad 1 + \gamma x \geq 0, \quad (1.7)$$

called the Generalized Extreme Value (GEV) df with parameter $\gamma \in \mathbb{R}$.

The parameter γ is called extreme value index. Originally, Hill (1975) [8] introduced the so-called Hill estimator $T_n(1)$ of the parameter γ of the Pareto df $\{1 - Cx^{-1/\gamma}\} \mathbb{I}_{(x \geq 0)}$. So did Pickands [11] as well. In the same way, the Generalized Hill process $T_n(\tau)$ provides an infinite family of estimators for the extremal index value in the sense that for any $\tau > 0$, for $F \in D(G_\gamma)$, $0 < \gamma \leq +\infty$,

$$\tau T_n(\tau) \rightarrow \gamma \text{ as } n \rightarrow +\infty.$$

This motivated Diop and Lo [6] to find out asymptotic normality results for this family in order to construct statistical tests. Although their results seem to be valuable, they restricted themselves to values $\tau > 1/2$. Actually, this restriction is imposed by the so-called Hungarian approximation methods based on M. Csörgö *et al.* [1] results that they used. It is then necessary to call for another approach to get around this obstacle.

In this paper, we will complete the work of [6] by giving the asymptotic law of $T_n(\tau)$ for $0 < \tau \leq 1/2$ under the hypothesis $F \in D(G_\gamma)$, $0 \leq \gamma < +\infty$. Our results are also complements to the related ones for $T_n(0)$ in connection with those in [11], [4] and [10]. We point out for once that $D(G_0) = D(\Gamma)$ for $\gamma = +\infty$.

We restrict ourselves here to the cases $0 \leq \gamma < +\infty$, since the case $\gamma < 0$, may be studied through the transform $F(x_0(F) - 1/\cdot) \in D(G_{-\gamma})$ for estimating γ . This leads to replace $X_{n-j+1,n}$ by $x_0(F) - 1/1/X_{j,n}$ in (1.1). However, a direct investigation of (1.1) for $\gamma < 0$ is possible. This requires the theory of sums of dependent random variables while this paper uses results on sums of independent random variables, as it will be seen shortly. We consequently consider a special handling of this case in a distinct paper.

Our achievement is the full description of the asymptotic behavior of $T_n(\tau)$, under the hypothesis $F \in D(G_\gamma)$, $0 \leq \gamma < +\infty$, for $0 < \tau \leq 1/2$, which is still Gaussian for $\tau = 1/2$ and non Gaussian for $\tau < 1/2$. In the latter case, the limiting laws are described in connection with the Riemann function.

Our best results concern $\tau \in]0, 1/2[$, in which case the usual conditions, (C2) is useless and (C3) becomes very weak. The asymptotic results become very general (see Subsection 4.2).

Before we state our results, we need to introduce some representations used in the theorems and to describe the conditions we shall require, at the appropriate time, in the next section. In Section 3, we will state our results and their proofs. In Section 4, statistical illustrations of the results and simulation study results are provided while Section 5 includes the key tools of this paper, which ends by a general conclusion in Section 6.

2. Representations Tools

Throughout the paper, we use the usual representation of Y_1, Y_2, \dots by $G^{-1}(1 - U_1), G^{-1}(1 - U_2), \dots$ where U_1, U_2, \dots are independent and uniform random variables on $(0, 1)$ in the sense of equality in distribution (denoted by $=_d$)

$$\{Y_j, n \geq 1\} =_d \{G^{-1}(1 - U_j), n \geq 1\},$$

and hence

$$\begin{aligned} & \{ \{Y_{1,n}, Y_{2,n}, \dots, Y_{n,n}\}, n \geq 1 \} \\ & =_d \{ \{G^{-1}(1 - U_{n,n}), G^{-1}(1 - U_{n-1,n}), \dots, G^{-1}(1 - U_{1,n})\}, n \geq 1 \}. \end{aligned} \quad (2.1)$$

In connexion with this, we also use the following Malmquist representation (see [13], p. 336)

$$\left\{ \log \left(\frac{U_{j+1,n}}{U_{j,n}} \right)^j, j = 1, \dots, n \right\} =_d \{E_1, \dots, E_n\}, \quad (2.2)$$

where E_1, \dots, E_n are independent standard exponential random variables.

Now we recall the classical representations of df 's attracted to some nondegenerated extremal df . For each df F in the extremal domain, an appropriate representation is given for $G^{-1}(1 - u) = \log F^{-1}(1 - u)$, $u \in (0, 1)$.

Theorem 2.1. *We have:*

(1) *Karamata's representation (KARARE)*

(a) *If $F \in D(G_\gamma)$, $\gamma > 0$, then there exist two measurable functions $p(u)$ and $b(u)$ of $u \in (0, 1)$ such that $\sup(|p(u)|, |b(u)|) \rightarrow 0$ as $u \rightarrow 0$ and a positive constant c so that*

$$G^{-1}(1 - u) = \log c + \log(1 + p(u)) - \gamma \log u + \left(\int_u^1 b(t)t^{-1} dt \right), \quad 0 < u < 1, \quad (2.3)$$

where $G^{-1}(u) = \inf\{x, G(x) \geq u\}$, $0 < u \leq 1$ is the generalized inverse of G with $G^{-1}(0) = G^{-1}(0+)$.

(b) *If $F \in D(G_\gamma)$, $\gamma < 0$, $\gamma > 0$, then $y_0(G) = \sup\{x, G(x) < 1\} < +\infty$ and there exist two measurable functions $p(u)$ and $b(u)$ for $u \in (0, 1)$ and a positive constant c as defined in (2.3) such that*

$$y_0 - G^{-1}(1 - u) = c(1 + p(u))u^{-\gamma} \exp \left(\int_u^1 b(t)t^{-1} dt \right), \quad 0 < u < 1. \quad (2.4)$$

(2) Representation of de Haan (Theorem 2.4.1 in [3]):

If $G \in D(\Lambda)$, then there exist two measurable functions $p(u)$ and $b(u)$ of $u \in (0, 1)$ and a positive constant c as defined in (2.3) such that for

$$s(u) = c(1 + p(u)) \exp \left(\int_u^1 b(t)t^{-1} dt \right), \quad 0 < u < 1, \quad (2.5)$$

we have for some constant $d \in \mathcal{R}$,

$$G^{-1}(1 - u) = d - s(u) + \int_u^1 s(t)t^{-1} dt, \quad 0 < u < 1. \quad (2.6)$$

It is important to remark at once that any *df* F in the domain of attraction is associated with a couple of functions (p, b) linked to $G(x) = F(e^x)$ used in each appropriate representation. Our conditions will rely on them. In order to state them, we set first for $0 < \tau \leq 1/2$

$$a_n(\tau) = k^{-\tau} \sum_{j=1}^k j^{\tau-1} \quad (2.7)$$

and

$$\sigma_n^2(\tau) = k^{-2\tau} \sum_{j=1}^k j^{2(\tau-1)}. \quad (2.8)$$

From there, our conditions are the following, where λ is an arbitrary real number greater than one. The first is

$$g_{1,n}(p, \lambda) (\sigma_n(\tau) k^\tau)^{-1} \sum_{j=1}^k j^\tau \rightarrow 0 \quad \text{as } n \rightarrow +\infty \quad (C1)$$

for

$$0 \leq g_{1,n}(p, \lambda) = \sup_{0 \leq u \leq \lambda k/n} |p(u)|.$$

The second is

$$g_{2,n}(p, \lambda) (\sigma_n(\tau) k^\tau)^{-1} \sum_{j=1}^k j^{\tau-1} \rightarrow 0 \quad \text{as } n \rightarrow +\infty \quad (C2)$$

for

$$g_{2,n}(b, \lambda) = \sup_{0 \leq u \leq \lambda k/n} |b(u)|.$$

The third is

$$d_n(p, b, \lambda) (\sigma_n(\tau) k^\tau)^{-1} \sum_{j=1}^k j^{\tau-1} \rightarrow 0 \quad \text{as } n \rightarrow +\infty \quad (C3)$$

for $d_n(p, b, \lambda) = \max(g_{1,n}(p, \lambda), g_{2,n}(b, \lambda) \log k) \rightarrow 0$ as $n \rightarrow \infty$.

We finally point out that all the limits in the sequel are meant as $n \rightarrow +\infty$ unless the contrary is specified. We are now able to state our results in the following section.

3. The results

Theorem 3.1. Let $F \in D(\Lambda) = D(G_0)$. Then, if (C3) holds,

$$(s(k/n)\sigma_n(\tau))^{-1}(T_n(\tau) - a_n(\tau)s(k/n)) \rightarrow \mathcal{N}(0, 1) \text{ for } \tau = 1/2$$

and

$$(s(k/n)\sigma_n(\tau))^{-1}(T_n(\tau) - a_n(\tau)s(k/n)) \rightarrow \mathcal{L}(\tau) \text{ for } 0 < \tau < 1/2,$$

and where

$$\mathcal{L}(\tau) = \zeta(2(1-\tau))^{-1/2} \sum_{j=1}^{\infty} j^{-(1-\tau)}(E_j - 1)$$

is a centered and normed random variable and

$$\zeta(s) = \sum_{j=1}^{\infty} j^{-s}, \quad s > 1,$$

is the Riemann function.

Let $F \in D(G_\gamma)$, $0 < \gamma < +\infty$. Then if (C1) and (C2) hold,

$$(a_n(\tau)/\sigma_n(\tau))(T_n(\tau)/a_n(\tau) - \gamma) \rightarrow \mathcal{N}(0, \gamma^2)$$

for $\tau = 1/2$ and

$$(a_n(\tau)/\sigma_n(\tau))(T_n(\tau)/a_n(\tau) - \gamma) \rightarrow \gamma\mathcal{L}(\tau).$$

Remark 3.1. We should remark that the methods used here apply for $0 < \tau < 1$. Moreover, these results do not extend to parameters $\tau \in]1/2, 1[$ since for such ones, $k^\tau \sigma_n(\tau) \rightarrow 0$, while our results are based on the hypothesis $k^\tau \sigma_n(\tau) \rightarrow C \in]0, +\infty[$.

Proof. Let us use the representations (2.1). We have, for any $n \geq 1$,

$$\{\log X_{n-j+1,n} = Y_{n-j+1,n}, 1 \leq j \leq n\} =_d \{G^{-1}\{1 - U_{j,n}\}, 1 \leq j \leq n\}.$$

Our proofs will rely on Lemma 5.1 below which establishes that

$$\left\{ k^{-\tau} \sum_{j=1}^b j^\tau \log \left(\frac{U_{j+1,n}}{U_{j,n}} \right) - a_n(\tau) \right\} / \sigma_n(\tau) =_d (k^{-\tau} \sigma_n(\tau))^{-1} \sum_{j=1}^k j^{\tau-1} (E_j - 1)$$

converges in distribution to $\mathcal{N}(0, 1)$ random variable for $\tau = 1/2$ and to the finite random variable

$$\mathcal{L}(\tau) = A(2(1-\tau))^{-1/2} \sum_{j=1}^{\infty} j^{-(1-\tau)}(E_j - 1)$$

for $0 < \tau < 1/2$, where the independent standard exponential random variables E_1, E_2, \dots are defined in (2.2). We begin by letting $F \in D(\Lambda) = D(G_{1/\gamma})$ for $\gamma = +\infty$. By (2.4), we get

$$T_n(\tau) = k^{-\tau} \sum_{j=1}^k j^\tau (s(U_{j,n}) - s(U_{j+1,n})) + k^{-\tau} \sum_{j=1}^k j^\tau \int_{U_{j,n}}^{U_{j+1,n}} s(t)/t dt \equiv S_n(1) + S_n(2). \quad (3.1)$$

By using (2.5), we have for $U_{1,n} \leq v, u \leq U_{k,n}$,

$$\frac{s(u)}{s(v)} = \frac{1+p(u)}{1+p(v)} \exp\left(-\int_u^v t^{-1}b(t)dt\right).$$

By putting

$$g_{1,n,0}(p) = \sup\{|p(u)|, 0 \leq u \leq U_{k+1,n}\} \text{ and } g_{2,n,0}(p) = \sup\{|b(u)|, 0 \leq u \leq U_{k+1,n}\}, \quad (3.2)$$

and by remarking that $\log(U_{k+1,n}/U_{1,n}) = O_p(\log k)$, we get

$$\frac{s(u)}{s(v)} = (1 + O(g_{1,n,0})) \exp(-O_p(g_{2,n,0} \log k)),$$

whenever

$$g_{2,n,0} \log k \rightarrow_p 0, \quad (3.3)$$

and next

$$\sup_{U_{1,n} \leq u, v \leq U_{k,n}} |s(u)/s(v) - 1| = O_p(\max(g_{1,n,0}, g_{2,n,0} \log k)) \quad (3.4)$$

and finally,

$$\sup_{U_{1,n} \leq u, v \leq U_{k,n}} \left| \frac{s(u) - s(v)}{s(k/n)} \right| = O_p(\max(g_{1,n,0}, g_{2,n,0} \log k)). \quad (3.5)$$

We have then to prove (3.3). But since $nk^{-1}U_{k+1,n} \rightarrow 1$, for a fixed value $\lambda > 1$ such that (C3) holds, we may find for any $\varepsilon > 0$, an integer N_0 such that for any $n \geq N_0$,

$$\mathbb{P}(g_{2,n,0} \log k \geq g_{2,n} \log k) \equiv \mathbb{P}(A_n) < \varepsilon.$$

Now fix $\delta > 0$. By using the sets decomposition

$$(|g_{2,n,0} \log k| > \delta) = (|g_{2,n,0} \log k| > \delta) \cap A_n + (|g_{2,n,0} \log k| > \delta) \cap A_n^c,$$

we have, for $n \geq N_0$

$$\mathbb{P}(|g_{2,n,0} \log k| > \delta) \leq \mathbb{P}(|g_{2,n} \log k| > \delta) + \varepsilon.$$

By letting first $n \rightarrow \infty$ and next $\varepsilon \downarrow 0$, and by taking (C3) into account, we see that (3.3) holds and so does (3.5). By putting together (3.1) and (3.5), and by using the preceding techniques, we get for a picked value $\lambda > 1$ like in (C3), for any $\varepsilon > 0$, a value N_0 such that for any $n \geq N_0$ and for any

$\delta > 0$,

$$\mathbb{P} \left(\left| \frac{S_n(1)}{\sigma_n(\tau)s(k/n)} \right| > \delta \right) \leq \mathbb{P} \left(d_n(p, b, \lambda) (k^\tau \sigma_n(\tau))^{-1} \sum_{j=1}^k j^\tau > \delta \right) + \varepsilon,$$

where $d_n(p, b, \lambda) = \max(g_{1,n}, g_{2,n} \log k)$. Again, by letting first $n \uparrow \infty$ and next $\varepsilon \downarrow 0$, and by using (C3), we get that $S_n(1)(\sigma_n(\tau)s(k/n))^{-1} \rightarrow_P 0$. We now treat $S_n(2)$ and remark that

$$\begin{aligned} \frac{S_n(2)}{\sigma_n(\tau)s(k/n)} &= (k^\tau \sigma_n(\tau))^{-1} \sum_{j=1}^k j^\tau \int_{U_{j,n}}^{U_{j+1,n}} t^{-1} \left\{ \frac{s(t)}{s(k/n)} \right\} dt \\ &= (k^\tau \sigma_n(\tau))^{-1} \sum_{j=1}^k j^\tau \int_{U_{j,n}}^{U_{j+1,n}} t^{-1} dt \\ &+ (k^\tau \sigma_n(\tau))^{-1} \sum_{j=1}^k j^\tau \int_{U_{j,n}}^{U_{j+1,n}} \{s(t)/s(k/n) - 1\} / t dt = S_n(2, 1) + S_n(2, 2). \end{aligned}$$

We have, by (3.4) and the Malmquist representation (2.2),

$$\begin{aligned} |S_n(2, 2)| &\leq O_p(1) d_n(p, b, \lambda) (k^\tau \sigma_n(\tau))^{-1} \sum_{j=1}^k j^{\tau-1} E_j \leq O_p(1) \times \\ &\left\{ d_{n,0}(p, b, \lambda) (k^\tau \sigma_n(\tau))^{-1} \sum_{j=1}^k j^{\tau-1} (E_j - 1) + d_{n,0}(p, b, \lambda) (k^\tau \sigma_n(\tau))^{-1} \sum_{j=1}^k j^{\tau-1} \right\}. \end{aligned}$$

where $d_{n,0}(p, b, \lambda) = \max(g_{1,n,0}, g_{2,n,0} \log k)$. The first term tends to zero since $d_{n,0}(p, b, \lambda) \rightarrow_P 0$ by (3.3), and $(k^\tau \sigma_n(\tau))^{-1} \sum_{j=1}^k j^{\tau-1} (E_j - 1)$ converges in distribution to a finite random variable by Lemma 5.1 in Section 5. We handle the second term also by the previous techniques. We choose a value $\lambda > 1$ like in (C3). For any $\varepsilon > 0$, we get value N_0 such that for any $n \geq N_0$ and for any $\delta > 0$,

$$\mathbb{P} \left(\left| d_{n,0}(p, b, \lambda) (k^\tau \sigma_n(\tau))^{-1} \sum_{j=1}^k j^{\tau-1} \right| > \delta \right) \leq \varepsilon + \mathbb{P} \left(\left| d_n(p, b, \lambda) (k^\tau \sigma_n(\tau))^{-1} \sum_{j=1}^k j^{\tau-1} \right| > \delta \right).$$

By letting first $n \uparrow \infty$ and next $\varepsilon \downarrow 0$, we see by (C3) that this term goes to 0 in probability. Finally, by the Malmquist representation (2.2), one arrives at

$$S_n(2, 1) = (k^\tau \sigma_n(\tau))^{-1} \sum_{j=1}^k j^{\tau-1} E_j + o_P(1).$$

And this leads to

$$S_n(2, 1) - a_n(\alpha)/\sigma_n(\tau) = (k^\tau \sigma_n(\tau))^{-1} \sum_{j=1}^k j^{\tau-1} (E_j - 1) \equiv V_n(\tau)$$

and finally, by summing up what precedes,

$$(\sigma_n(\tau)s(k/n))^{-1} (T_n(\tau) - a_n(\tau)) = V_n(\tau) + o_P(1).$$

This concludes the proof since, by Lemma 5.1 in Section 5, $V_n(\tau)$ converges in probability to a $\mathcal{N}(0, 1)$ random variable for $\tau = 1/2$ and to $\mathcal{L}(\tau)$ for $\tau < 1/2$.

Now let $F \in D(G_\gamma)$, $0 < \gamma < +\infty$. We will have similar proofs but we do not explicit the convergences in distribution as we did in the previous case. We have by (2.3) and the usual representations,

$$\begin{aligned} T_n(\tau) &= k^{-\tau} \sum_{j=1}^k j^\tau \{ \log(1 + p(U_{j+1,n})) - \log(1 + p(U_{j,n})) \} \\ &+ \gamma k^{-\tau} \sum_{j=1}^k j^\tau \log(U_{j+1,n}/U_{j,n}) + k^{-\tau} \sum_{j=1}^k j^{-\tau} \int_{U_{j,n}}^{U_{j+1,n}} b(t)/t dt \\ &\equiv S_n(1) + S_n(2) + S_n(3). \end{aligned}$$

We have, for large values of k ,

$$\left| \frac{S_n(1)}{\sigma_n(\tau)} \right| \leq 2g_{1,n,0}(f)(\sigma_n(\tau)k^\tau)^{-1} \sum_{j=1}^k j^\tau,$$

where $g_{1,n,0}$ is defined in (3.2), which tends to zero in probability by (C1) and (3.4). Next

$$\begin{aligned} \left| \frac{S_n(3)}{\sigma_n(\tau)} \right| &\leq g_{2,n,0}(b)(\sigma_n(\tau)k^\tau)^{-1} \sum_{j=1}^k j^\tau \log\left(\frac{U_{j+1,n}}{U_{j,n}}\right) \\ &= g_{2,n,0}(b)(\sigma_n(\tau)k^\tau)^{-1} \sum_{j=1}^k j^{\tau-1}(E_j - 1) + g_{2,n,0}(b)(\sigma_n(\tau)k^\tau)^{-1} \sum_{j=1}^k j^{\tau-1}, \end{aligned}$$

where $g_{2,n,0}$ defined in (3.2). Then $S_n(3)/\sigma_n(\tau) \rightarrow 0$ by (C2), Lemma 5.1 and the methods described above. Finally, always by Lemma 5.1, we get

$$\{(S_n(3) - \gamma a_n(\tau))/\sigma_n(\tau) = \gamma(\sigma_n(\tau)k^\tau)^{-1} \sum_{j=1}^k j^{\tau-1}(E_j - 1)$$

and next

$$\frac{\{T_n(\tau) - \gamma a_n(\tau)\}}{\sigma_n(\tau)} = \gamma(\sigma_n(\tau)k^\tau)^{-1} \sum_{j=1}^k j^{\tau-1}(E_j - 1) + o_P(1) = \gamma W_n(\tau) + o_P(1)$$

and this converges in distribution to $\mathcal{N}(0, \gamma^2)$ random variable for $\tau = 1/2$ and to $\gamma \mathcal{L}(\tau)$ for $\tau < 1/2$. The proof is completed with the remark that $a_n(\tau) \rightarrow \tau^{-1}$. \square

4. Remarks, Applications and Simulations

4.1. An asymptotically Gaussian estimator

Now let us illustrate some statistical applications of our results. Let $F \in D(G_\gamma)$. Recall that for $\tau > 1/2$, $\tau T_n(\tau)$ has been proved to be asymptotic normal estimators of γ . Our present results extend those obtained to $\tau = 1/2$ and show that this latter value is a critical one since $\tau T_n(\tau)$ is still

an estimator of γ for $\tau < 1/2$ but has a non-Gaussian asymptotic behavior. Now let us particularize the results for each case and illustrate the conditions (C1) and (C2). For $\tau = 1/2$, we have under the (C1) and (C2),

$$2k^{1/2}(\log k)^{-1/2}(T_n(1/2)/a_n(1/2) - \gamma) \rightarrow N(0, \gamma^2),$$

Thus $T_n(1/2)/a_n(1/2)$ is a new estimator of the extreme value index which is asymptotically normal. Since $a_n(1/2) \rightarrow 1/2$, we rediscover the general result $\tau T_n(\tau) \rightarrow \gamma$. By (S4) below, $a_n(1/2) = (1/2)(1 + O(k^{-1/2}))$ so that we also have

$$2k^{1/2}(\log k)^{-1/2}(2T_n(1/2) - \gamma) \rightarrow N(0, \gamma^2).$$

From there we may form another family of estimators $T_n(1/2, a) = (2a + (1 - a)/a_n(1/2))T_n(1/2)$, for $0 < a < 1$, which are asymptotically normal since

$$2k^{1/2}(\log k)^{-1/2}(T_n(1/2, a) - \gamma) \rightarrow N(0, \gamma^2). \tag{4.1}$$

This parameter a , that may be randomly picked from $(0,1)$, may be usefull in applications. The condition (C1) turns to

$$g_{1,n}(p, \lambda)(\log k)^{-1/2} \sum_{j=1}^k j^{1/2} \sim g_{1,n}(p, \lambda)k^{-1/2}(\log k)^{-1/2} \rightarrow 0.$$

But in many real situations, we may suppose that the df admits a derivative so that we may take $p(u) = 0$, which renders (C1) useless. Indeed, suppose that $F(x)$ admits a derivative $F'(x)$ for large values of x . A condition for F to belong to $D(G_{1/\gamma})$, $0 < \gamma < +\infty$, is

$$\lim_{x \rightarrow \infty} \frac{1 - F(x)}{xF'(x)} = \frac{1}{\gamma}. \tag{4.2}$$

(see [2], p. 17, Theorem 1.1.11). This implies that

$$b(u) = u(G^{-1}(1 - u))' + \gamma = uG'(G^{-1}(1 - u))^{-1} + \gamma \rightarrow 0 \text{ as } n \rightarrow \infty$$

and next, for some constant $c > 0$ and $u_0 \in (0, 1)$,

$$G^{-1}(1 - u) = c - \gamma \log u + \int_u^{u_0} b(t)t^{-1} dt.$$

For such df 's, (C1) is effectively useless. The condition (C2) is more consistent. However, it is somewhat related to the extreme second order condition (see [2], p. 43) and some popular models used in the literature. For example, in the Hall one, that is $1 - F(x) = C_1x^{-1/\gamma}(1 + C_2x^{-\beta} + o(x^{-\beta}))$, where β , C_1 and C_2 are positive constants, we have for constants C_i , $i = 1, 2, \dots$, $F^{-1}(1 - u) =$

$C_3 u^{-\gamma}(1 + C_4 u^{\beta\gamma} + o(u^{\beta\gamma}))$ and then

$$G^{-1}(1 - u) = \log C_3 - \gamma \log u + O(u^{\beta\gamma}),$$

which may be written, with $b_1(t) = u^{\beta\gamma}$ and $c_1(u) = O(1)$ as $u \rightarrow 0$,

$$G^{-1}(1 - u) = C_4 - \gamma \log u + c_1(u) \int_u^1 (b_1(t)/t) dt.$$

Suppose now that $c_1(u) = D$ or that $c_1(u) \int_u^1 (b_1(t)/t) dt$ is obtained from $\int_u^1 c(t)(b_1(t)/t) dt$ by the Integral Mean Theorem with a bounded function $c(\cdot)$. Then condition (C2) turns to

$$k^{\beta\gamma+1/2}(\log k)^{-1/2}/n^{\beta\gamma} \rightarrow 0$$

and (C3) to

$$k^{\beta\gamma+1/2}(\log k)^{1/2}/n^{\beta\gamma} \rightarrow 0.$$

We conclude that for $\tau = 1/2$, $T_n(1/2)/a_n(1/2)$ is an asymptotic normal estimator of γ under reasonable assumptions based on the regularity on the df F . Simulation studies suggest its good performances as well.

4.2. A class of non asymptotically Gaussian estimators

For $0 < \tau < 1/2$, we also have under (C1) and (C2),

$$\frac{k^{\tau\gamma}}{\tau \zeta(2(1 - \tau))} (T_n(\tau)/a_n(\tau) - \gamma) \rightarrow \mathcal{L}(\tau). \quad (4.3)$$

Here again, we get a family of estimators of γ . However, the limiting law is not Gaussian. This seems to be new. Remark here that we cannot simplify (4.3) to

$$\frac{k^{\tau\gamma}}{\tau \zeta(2(1 - \tau))} (\tau T_n(\tau) - \gamma) \rightarrow \mathcal{L}(\tau)$$

like for $\tau = 1/2$ because of $\sigma_n(\tau)$ being finite and because of (S3) below.

This results make a connexion of Extreme Value Theory and Number Theory since they deeply depend of the Riemann function closely related to the prime numbers since, for $s > 1$,

$$\zeta(s) = \prod_{p=2}^{\infty} (1 - p^{-s})^{-1} = \frac{s}{s-1} + \frac{1}{s} \int_1^{\infty} \frac{x - [x]}{x^{s+1}} dx,$$

where the product extends to the prime numbers $p \geq 2$ and $[x]$ denotes the integer part of x . The limiting law $\mathcal{L}(\tau)$ is characterized by its characteristic function

$$\psi_{\infty}(t) = \exp \left(\sum_{n=2}^{\infty} \frac{(it)^n}{n} \zeta(n(1 - \tau)) \zeta(2(1 - \tau))^{-n/2} \right),$$

calculated and justified in (5.1). This random variable has all its moments finite and is related to the Riemann function. Indeed put $\psi_{\infty}(t) = \exp(\phi(t))$ and let $\phi^{(r)}$ denote the r^{th} derivative function

of ϕ . We easily see that

$$\psi_{\infty}^{(1)}(t) = \phi^{(1)}(t) \exp(\phi(t)), \text{ with } \phi^{(1)}(t) = \sum_{n=2}^{\infty} i(it)^{n-1} \zeta(n(1-\tau)) \zeta(2(1-\tau))^{-n/2},$$

and

$$\psi_{\infty}^{(2)}(t) = (\phi^{(1)}(t))^2 + \phi^{(2)}(t) \exp(\phi(t)),$$

with

$$\phi^{(2)}(t) = \sum_{n=2}^{\infty} -(n-1)(it)^{n-2} \zeta(n(1-\tau)) \zeta(2(1-\tau))^{-n/2}.$$

So

$$\phi^{(1)}(0) = 0 \text{ and } \phi^{(2)}(0) = - \left[(n-1)(it)^{n-2} \zeta(n(1-\tau)) \zeta(2(1-\tau))^{-n/2} \right]_{n=2} = -1$$

and thus

$$\begin{aligned} \mathbb{E}X &= 0, \\ \mathbb{E}X^2 &= 1. \end{aligned}$$

One may repeat the same methods to get

$$\begin{aligned} \mathbb{E}X^3 &= 2\zeta(3(1-\tau))/\zeta(2(1-\tau))^{3/2}, \\ \mathbb{E}X^4 &= 3 + 6\zeta(4(1-\tau))/\zeta(2(1-\tau))^2 \end{aligned}$$

and higher moments. For such values of τ , the condition (C2) and (C3) become

$$g_{2,n}(p, \lambda) \rightarrow 0$$

and

$$\max(g_{1,n}(p, \lambda), g_{2,n}(b, \lambda)) \rightarrow 0.$$

We get here the interesting thing that (C2) always holds for any $df F \in D(G_{\gamma})$, $0 < \gamma < +\infty$. And this is a quite general and rare result in EVT. For $\gamma = 0$, the condition (C3) is also very weak since $g_{1,n}$ and $g_{2,n}$ are only requested to tend to zero faster than $\log k$. We get then very general asymptotic laws that sound as a compensation of the lack of normality!

Since these limits seem new in the extreme value index estimation, it may be reasonable that we give some comments on how to use them. Although we cannot give a simple expression of the df of $\mathcal{L}(\tau)$, $F_{\mathcal{L}(\tau)}(u) = \mathbb{P}(\mathcal{L}(\tau) \leq u)$, we may use computer-based methods to compute approximations of its values. We may fix a great value of k and then consider a large sample, of size B , of the random variable

$$\mathcal{L}(\tau, k) = \zeta(2(1-\tau), m)^{-1/2} \sum_{j=1}^k j^{\tau-1} (E_j - 1), \quad (4.4)$$

where $\zeta(s, m) = \sum_{j=1}^m j^{-s}$, with a sufficiently large value of m . We surely get the empirical distribution function (edf) of this sample as a fair uniform approximation of $F_{\mathcal{L}(\tau)}$. Fix $m = 10000$,

$B = 10000$, we may observe in the right graph in Figure 3, that the df 's of $F_{L(\tau)}$ for $\tau \in \{0.1, 0.2, 0.25, 0.3, 0.25, 0.3, 0.35, 0.40, 0.48, 0.48\}$ are very close one another. When we consider a τ uniformly drawn from $]0, 1/2[$, we obtain a significantly different df as illustrated in the left graph in Figure 3, where we simply add this last df to those already in the first graph.

From there, we are able to compute the probabilities and the quantiles of such limiting laws, as illustrated in the Table 1 in the Appendix. Statistical tests may then be based on these results.

4.3. Simulations

We wish to show the performance of these class of statistics as estimators of γ and compare them to some other ones available in the literature, such as Hill's one $T_n(1)$ (see [8]), the Pickands estimator (see [11])

$$P_n(k) = (\log 2)^{-1} \log \left\{ \frac{(X_{n-k,n} - X_{n-2k,n})}{(X_{n-2k,n} - X_{n-4k,n})} \right\},$$

and the Lo estimator (see [9]) $L_n(k)$ with

$$L_n^2(k) = k^{-1} \sum_{j=1}^k \sum_{i=j}^k jg(i, j)(\log X_{n-j+1,n} - \log X_{n-j,n})(\log X_{n-i+1,n} - \log X_{n-i,n}),$$

where $g(i, j) = 1/2$ for $i = j$ and $g(i, j) = 1$ otherwise. We may see that the performance of these estimators largely depend on the nature of the tail $1 - F$. In the simple case where $1 - F(x) = x^{-1/\gamma}$, all these estimators perform pretty well for $n = 300$, $k \sim n^{0.75}$ for $\gamma = 0.5$ and $T_n(1/2)/a_n(1/2)$ has particularly good results. They also present good performance for the model $1 - F(x) = x^{-1/\gamma}(1 + O(x^{-\eta}))$ with $\eta = 100$. The results are less sound for small η ($\eta = 10$ for example) for sample sizes around 300. We check the performances in Table 2 in the Appendix, where we report the estimated values for γ for the sample case and the MSE, and in Figure 1 all in the Appendix. The asymptotic normality of (4.1) is illustrated in the graphs of Figure 2 as good.

The simulation results for the asymptotic normality of (4.1) are also acceptable since the empirical p -value is high, of order 90%.

For values $\tau \in]0, 1/2[$, we illustrate the results for $\tau = 0.3$. In the graph in right in Figure 4, we realize that $T_n(0.3)/a_n(0.3)$ is a fair estimate of γ in the simple model described above.

As said previously Figure 3 successfully compares the law of $\frac{k^{\tau\gamma}}{\tau\zeta(2(1-\tau))}(T_n(\tau)/a_n(\tau) - \gamma)$ in (4.3) and that of (4.4).

5. Lemmas and Tools

We begin by this simple lemma where we suppose that we are given a sequence of independent and uniformly distributed random variables U_1, U_2, \dots as in (2.1).

Lemma 5.1. *Let*

$$V_n(\tau) = k^{-\tau} \sum_{j=1}^b j^\tau \log \left(\frac{U_{j+1,n}}{U_{j,n}} \right).$$

Then for $\tau = 1/2$,

$$\sigma_n^{-1}(\tau)(V_n(\tau) - a_n(\tau)) \rightsquigarrow \mathcal{N}(0, 1)$$

and for $0 < \tau < 1/2$,

$$\sigma_n^{-1}(\tau)(V_n(\tau) - a_n(\tau)) \rightsquigarrow \mathcal{L}(\tau),$$

where $a_n(\tau)$ and $\sigma_n(\tau)$ are defined in (2.7) and (2.8), and

$$\mathcal{L}(\tau) = A(2(1 - \tau))^{-1/2} \sum_{j=1}^{\infty} j^{-(1-\tau)}(E_j - 1),$$

is a centred and reduced random variable having all moments finite.

Proof. By using the Malmquist representation (2.2), we have

$$V_n(k) = k^{-\tau} \sum j^{\tau-1} E_j.$$

It follows that $\mathbb{E}(V_n(k)) = a_n(k)$ and $\mathbb{V}\text{ar}(V_n(k)) = \sigma_n^2(k)$. Put

$$V_n^*(\tau) = \sigma_n(k)^{-1}(V_n(k) - a_n(k)).$$

Then

$$V_n^*(\tau) = (k^\tau \sigma_n(\tau))^{-1} \sum_{j=1}^k j^{\tau-1} (E_j - 1).$$

For $0 < \tau < 1/2$, we have by (S1) below, $k^\tau \sigma_n(\tau) \rightarrow A(2(1 - \tau))^{-1/2}$ and

$$V_n^*(\tau) \rightarrow \zeta(2(1 - \tau))^{-1/2} \sum_{j=1}^{\infty} j^{\tau-1} (E_j - 1) = L(\tau).$$

Now, we have to prove that $\mathcal{L}(\tau)$ is a well-defined random variable with all finite moments. The characteristic function of $V_n^*(\tau)$ is

$$\psi_{V_n^*(\tau)}(t) = \exp(-\zeta(2(1 - \tau))^{-1/2} \sum_{j=1}^k j^{\tau-1} (it)) \prod_{j=1}^k (1 - it j^{-(1-\tau)} \zeta(2(1 - \tau))^{-1/2})^{-1}.$$

By using the development of $\log(1 - \cdot)$ and by the Lebesgue's Theorem, one readily proves that

$$\begin{aligned} \psi_{V_n^*(\tau)}(t) &= \exp\left(\sum_{j=1}^k \sum_{n=2}^{\infty} \frac{(it)^n}{n} j^{-n(1-\tau)} \zeta(2(1 - \tau))^{-n/2}\right) \\ &\rightarrow \psi_{\infty}(t) = \exp\left(\sum_{n=2}^{\infty} \frac{(it)^n}{n} \zeta(n(1 - \tau)) \zeta(2(1 - \tau))^{-n/2}\right). \end{aligned} \quad (5.1)$$

Recall as in [14], p. 506, that for $s > 1$,

$$\zeta(s) = \frac{s}{s-1} + \frac{1}{s} \int_1^{\infty} \frac{x - [x]}{x^{s+1}} dx,$$

where $[x]$ denotes the integer part of x . This leads to

$$s(s-1)^{-1} \leq \zeta(s) \leq s(1 + (s-1)^{-1}).$$

By using this, we see that the absolute value of the general term of the series in (5.1) is dominated, for large values of n , as follows

$$\left| \frac{(it)^n}{n} \zeta(n(1-\tau)) \zeta(2(1-\tau))^{-n/2} \right| \leq 2(1-2\tau) \left(\frac{|t|}{2(1-\tau)} \right)^n.$$

This shows that $\psi_\infty(t)$ is well defined and characterizes the df of $\mathcal{L}(\tau)$. In Section 4, we anticipated by remarking that $\mathcal{L}(\tau)$ is centered and has all its moment finite and indicated the way to get them.

It remains to handle the case $\tau = 1/2$. Let us evaluate the moment generating function of $V_n^*(1/2)$:

$$\psi_{V_n^*(1/2)}(t) = \prod_{j=1}^k \psi_{(E_{j-1})}(ti^{\tau-1}(k^{1/2}\sigma_n(1/2))^{-1}),$$

But we have $\sigma_n^2(1/2) = S(k, 1) \sim k^{-1} \log k$ and next $k^{1/2}\sigma_n(1/2) = (\log k)^{1/2} \rightarrow \infty$. It follows that for a fixed, for k large enough,

$$\left| ti^{-1/2}(k^{1/2}\sigma_n(1/2))^{-1} \right| \leq 2t(\log k)^{-1/2} \tag{5.2}$$

uniformly in $i \geq 1$. At this step, we use the expansion of $\psi_{(E_{j-1})}$ in the neighbourhood of zero:

$$\psi_{(E_{j-1})}(u) = 1 + u^2/2 + u^4 g(u),$$

where there exists u_0 such that

$$0 \leq u \leq u_0 \Rightarrow |g(u)| \leq 1.$$

By using the uniform bound in (5.2), for a value of k such that $2t(\log k)^{1/2} \leq u_0$, we get

$$\begin{aligned} \psi_{(E_{j-1})}(t j^{-1/2}(k^{1/2}\sigma_n(1/2))^{-1}) &= 1 + \frac{1}{2}(t j^{-1/2}(k^{1/2}\sigma_n(1/2))^{-1})^2 \\ &\quad + (t j^{-1/2}(k^{1/2}\sigma_n(1/2))^{-1})^4 g_{0,j,n}(t), \end{aligned}$$

where $|g_{0,j,n}(t)| \leq 1$ for all $1 \leq j \leq k$. By the uniform boundedness of the error term, we have

$$\begin{aligned} \log \psi_{(E_{j-1})}(t j^{-1/2}(k^{1/2}\sigma_n(1/2))^{-1}) &= \\ \frac{1}{2}(t j^{-1/2}(k^{1/2}\sigma_n(1/2))^{-1})^2 &+ (t j^{-1/2}(k^{1/2}\sigma_n(1/2))^{-1})^4 g_{0,j,n}(t) \\ (t j^{-1/2}(k^{1/2}\sigma_n(1/2))^{-1})^4 &g_{1,j,n}(t), \end{aligned}$$

where always $|g_{1,j,n}(t)| \leq 1$ for all $1 \leq j \leq k$. Finally

$$\begin{aligned} \psi_{V_n^*(1/2)}(t) &= \exp \left(\sum_{j=1}^k \log \psi_{(E_{j-1})}(ti^{\tau-1}(k^{1/2}\sigma_n(1/2))^{-1}) \right) \\ &= \exp \left(t^2/2 + t^4 g_2 \sum_{j=1}^k j^{-2} \right), \end{aligned}$$

where for large values of k ,

$$|g_2| \leq (\log k)^{-2}.$$

Hence

$$\psi_{V_n^*(1/2)}(t) \rightarrow \exp(t^2/2)$$

and

$$V_n^*(1/2) \rightarrow \mathcal{N}(0, 1),$$

which achieves the proofs. □

Lemma 5.2. Define for $\tau > 0$, $S(k, \tau) = k^{-\tau} \sum_{j=1}^k j^{-\tau}$. Then:

$$S(k, \tau) \sim \zeta(\tau)k^{-\tau}, \quad \text{for } \tau > 1, \quad (\text{S1})$$

$$S(k, \tau) = 2(1 + O(k^{-1/2})), \quad \text{for } \tau = 1/2, \quad (\text{S2})$$

$$S(k, \tau) = k^{1-2\tau}/(1-\tau)(1 + O(k^{-1-\tau})), \quad \text{for } \tau < 1, \text{ and } (\text{S3})$$

$$S(k, \tau) = (\log k)k^{-1}(1 + O(1/\log k)), \quad \text{for } \tau = 1. \quad (\text{S4})$$

Proof. These formulas are readily obtained by comparing the $\sum_{j=1}^k j^{-\tau}$ and $\int_1^k x^{-\tau} dx$, as in classical analysis. □

6. Conclusion

The family $\{T_n(\tau), 0 < \tau \leq 1/2\}$ is only studied for F belonging to $D(G_\gamma)$, $\gamma \geq 0$. This includes the negative case under the appropriate transform. However the remaining case, that is the Weibull domain, presents a radically different approach including sums of dependent random variables and deserves a separated study. Furthermore, we still have to develop a Bayesian approach by considering a random choice of the parameter τ leading to an optimal choice among a large class of admissible laws.

Appendix

We give in this appendix all the tables and figures of the paper.

Table 1.

	0.1	0.15	0.2	0.25	0.3	0.35	0.4	0.45	0.48	unif
$F(-1.96)$	0.004	0.0035	0.006	0.007	0.009	0.0137	0.016	0.0166	0.018	0.067
$F(0)$	0.5572	0.557	0.549	0.544	0.538	0.54	0.524	0.501	0.524	0.537
$F(1.96)$	0.9530	0.9558	0.9545	0.9623	0.961	0.965	0.971	0.9842	0.966	0.913

Table 2.

	$T_n^*(1/2)$	Hill Est.	Pickands Est.	Lo Est.
Estimated γ	0.5047	0.5008	0.4992	0.4733
MSE	0.0051	0.0025	0.045	0.009

Fig. 1.

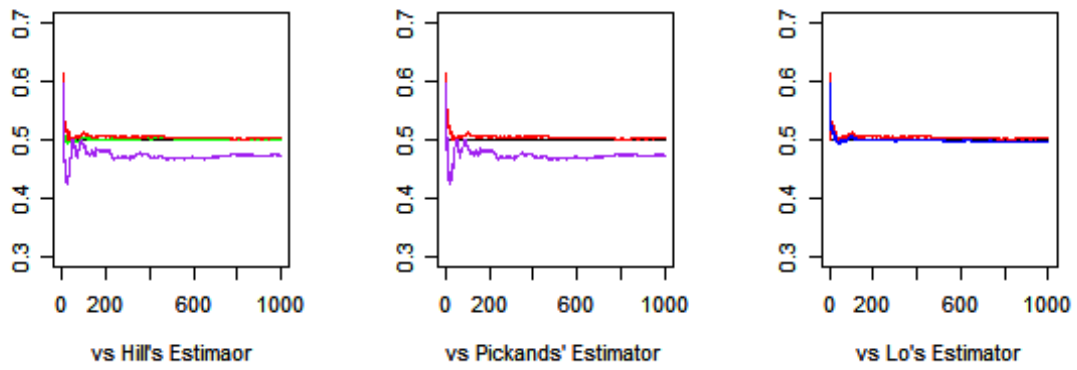


Fig. 2.

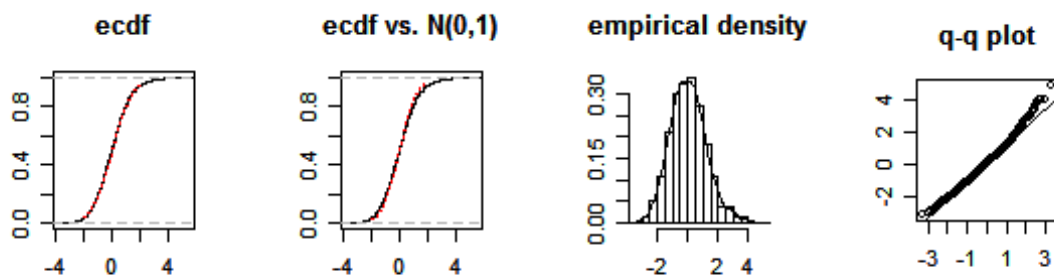


Fig. 3.

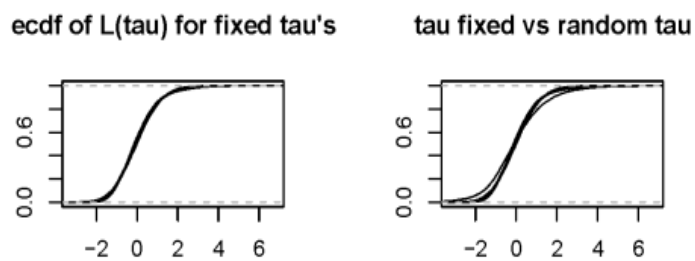
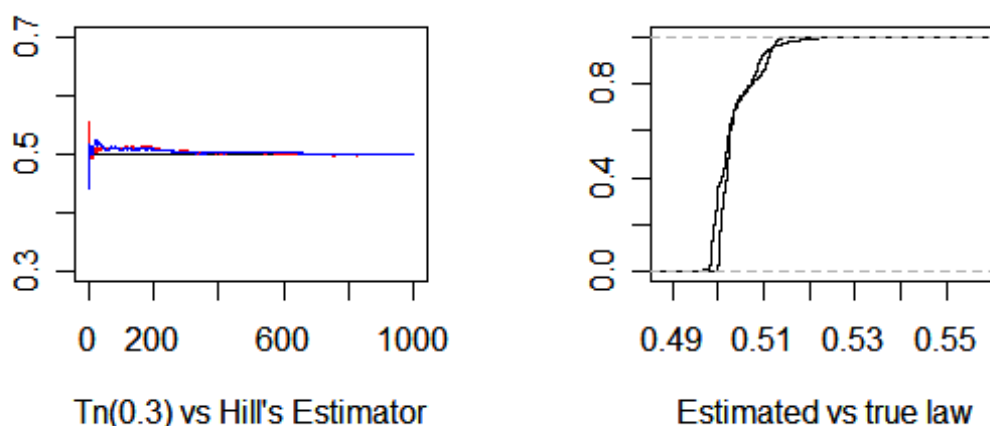


Fig. 4.



References

- [1] Csörgő, M., Csörgő, S., Horváth, L. and Mason, M. (1986). Weighted empirical and quantile processes. *Ann. Probab.* **14**, 31–85.
- [2] de Haan, L. and Ferreira A. (2006). *Extreme value theory: An introduction*. Springer.
- [3] de Haan, L. (1970). *On regular variation and its application to the weak convergence of sample extremes*. Mathematical Centre Tracts, 32, Amsterdam.
- [4] De Haan, L. and Resnick, S.I. (1980). A simple asymptotic estimate for the index of a stable law. *J. Roy. Statist. Soc.*, B, 83–87.
- [5] Diop Aliou and Lo G.S. (1990). Generalized Hill's Estimator. *Far East J. Theor. Statist.*, **20** (2), pp. 129–149.
- [6] Diop, A. and G.S. Lo. (2009). Ratio of Generalized Hill's Estimator and its asymptotic normality theory. *Math. Method. Statist.*, **18** (2), pp. 117-133.
- [7] Galambos, J. (1985). *The Asymptotic theory of Extreme Order Statistics*. Wiley, New York.
- [8] Hill, B.M. (1975). A simple general approach to the inference about the tail index of a distribution. *Ann. Statist.* **3**, 1163–1174.
- [9] Lo, G.S. (1992). Sur la caractérisation empirique des extrêmes. *C. R. Math. Rep. Acad. Sci. Canada*, **XIV**, n° 2, 3, 89–94.
- [10] Lo, G.S. (1986). *Sur quelques estimateurs de l'index d'une loi de Pareto : Estimation de Deheuvels-Csörgő-Mason, de De Haan-Resnick et lois limites pour des sommes de valeurs extrêmes pour une variable dans le domaine de Gumbel*. Thèse de Doctotat. Université Paris-6. France.
- [11] Pickands III, J. (1975). Statistical inference using extreme order statistics. *Ann. Statist.*, **3**, 119–131.
- [12] Resnick, S.I. (1987). *Extreme Values, Regular Variation and Point Processes*. Springer-Verbag, New York.
- [13] Shorack G.R. and Wellner J.A. (1986). *Empirical Processes with Applications to Statistics*. Wiley-Interscience, New York.
- [14] Valiron, G. (1990). *Cours d'Analyse Mathématique : Théorie des fonctions*. 3ième Edition. Masson, Paris.