Journal of Mathematics Research; Vol. 8, No. 4; August 2016
ISSN 1916-9795 E-ISSN 1916-9809
Published by Canadian Center of Science and Education

A Double-indexed Functional Hill Process and Applications

Modou Ngom[1] & Gane Samb Lo[2]

[1] LERSTAD, Université Gaston Berger de Saint-Louis, SENEGAL

[2] LERSTAD, Université Gaston Berger de Saint-Louis, SENEGAL and LSTA, Université Pierre et Marie Curie, France (affiliated), gane-samb.

Correspondence: Gane Samb Lo, LERSTAD, Université Gaston Berger de Saint-Louis, SENEGAL and LSTA, Université Pierre et Marie Curie, France (affiliated). E-mail: lo@ugb.edu.sn

Received: April 23, 2016 Accepted: May 30, 2016 Online Published: July 26, 2016

doi:10.5539/jmr.v8n4p144 URL: http://dx.doi.org/10.5539/jmr.v8n4p144

Abstract

Let $X_{1,n} \leq .... \leq X_{n,n}$ be the order statistics associated with a sample $X_1, ..., X_n$ whose pertaining distribution function (df) is F. We are concerned with the functional asymptotic behaviour of the sequence of stochastic processes

$$T_n(f, s) = \sum_{j=1}^{j=k} f(j)\left(\log X_{n-j+1,n} - \log X_{n-j,n}\right)^s, \tag{0.1}$$

indexed by some classes $\mathcal{F}$ of functions $f : \mathbb{N}^* \longmapsto \mathbb{R}_+$ and $s \in]0, +\infty[$ and where $k = k(n)$ satisfies

$$1 \leq k \leq n, k/n \to 0 \text{ as } n \to \infty.$$

We show that this is a stochastic process whose margins generate estimators of the extreme value index when F is in the extreme domain of attraction. We focus in this paper on its finite-dimension asymptotic law and provide a class of new estimators of the extreme value index whose performances are compared to analogous ones. The results are next particularized for one explicit class $\mathcal{F}$..

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.

2000 Mathematics Subject Classification. Primary 62EG32, 60F05. Secondary 62F12, 62G20.

1. Introduction

1.1 General Introduction

In this paper, we are concerned with the statistical estimation of the univariate extreme value index of a df F, when it is available. But rather than doing this by one statistic, we are going to use a stochastic process whose margins generate estimators of the extreme value index (SPMEEXI). To precise this notion, let $X_1, X_2, ...$ be a sequence of independent copies (s.i.c) of a real random variable (rv) $X > 1$ with df $F(x) = \mathbb{P}(X \leq x)$. F is said to be in the extreme value domain of attraction of a nondegenerate df M whenever there exist real and nonrandom sequences $(a_n > 0)_{n\geq1}$ and $(b_n)_{n\geq1}$ such that for any continuity point x of M,

$$\lim_{n\to\infty} P(\frac{X_{n,n} - b_n}{a_n} \leq x) = \lim_{n\to\infty} F^n(a_n x + b_n) = M(x). \tag{1.1}$$

It is known that M is necessarily of the family of the Generalized Extreme Value (GEV) df :

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

parameterized by $\gamma \in \mathbb{R}$. The parameter γ is called the extreme value index. There exists a great number of estimators of γ, going back to first of all of them, the Hill's one defined by

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

where for each n, $k = k(n)$ is an integer such that

$$1 \leq k \leq n,\ k \to \infty,\ k/n \to 0 \text{ as } n \to \infty.$$

A modern and large account of univariate Extreme Value Theory can be found in Beirlant, Goegebeur and Teugels (2004), Galambos (1985), de Haan (1970) and de Haan and Feireira (2006), Embrechts *et al.* (1997) and Resnick (1987). One may estimate γ by one statistic only. This is widely done in the literature. But one also may use a stochastic process of statistics $\{T_n(f), f \in \mathcal{F}\}$ indexed by $\mathcal{F}$, such that for any fixed $f \in \mathcal{F}$, there exists a sequence of nonrandom and positive real coefficients $(a_n(f))_{n\geq 1}$ such that $T_n^*(f) = T_n(f)/a_n(f)$ is an asymptotic estimator of γ. We name such families *Stochastic Processess with Margins Estimating of the EXtreme value Index (SPMEEXI's)*. Up to our knowledge, the first was introduced in Lo (1997) (see also Lo (2012)) as follows

$$T_n(p) = k^{-1} \sum_{h=1}^{p} \sum_{(s_1......s_h)\in\mathcal{P}(p,h)} \sum_{i_1=\ell+1}^{i_0} ... \sum_{i_h=\ell+1}^{i_{h-1}} i_h \prod_{i=i_1}^{i_h} \frac{(\log X_{n-i+1,n} - \log X_{n-i,n})^{s_i}}{s!},$$

for $1 \leq \ell < k < n$, $p \geq 1$, $i_0 = k$, where $\mathcal{P}(p,h)$ is the set of all ordered partitions of $p > 0$ into positive integers, $1 \leq h \leq p$:

$$\mathcal{P}(p,h) = \{(s_1...s_h), \forall i,\ 1 \leq i \leq h, s_i > 0; s_1 + ... + s_h = p\}.$$

Further Lo *et al.* (2006, 2009) introduced continuous and functional forms described in (1.2) below. Meanwhile, without denoting it such that, Segers (2002) and others considered the Pickands process $\{P_n(s), \sqrt{k/n} \leq s \leq 1\}$, with

$$P_n(s) = \log \frac{X_{n-[k/s],n} - X_{n-[k/1],n}}{X_{n-[k/s^2],n} - X_{n-[k/s],n}}, \sqrt{k/n} \leq s \leq 1.$$

Groeneboom (2003), proposed a family of kernel estimators, indexed by kernels. This family is surely a SPMEEXI although the authors did not consider a stochastic process view in the kernels K.

The main interest of *SPMEEXI*'s is first to have in hands an infinite class of estimators and especially, as shown in Segers (2002), to have the possibility to build discrete and continuous combinations of the margins as new and powerful estimators.

Number of the estimators of the extreme value index are either functions of consecutive log-spacings $\log X_{n-j+1,n} - \log X_{n-j,n}$ for $1 \leq j \leq k$, $1 \leq k \leq n$, or are functions of log-spacings from a threshold $\log t$: $\log X_{n-j+1,n} - \log t$, $1 \leq j \leq k$, $1 \leq k \leq n$. In the last case, the threshold is usually taken as $t = X_{n-k,n}$. This simple remark teases the idea that taking functions of the log-spacings in place of the simple ones may lead to more general estimators. Dekkers *et al.* (1989) successfully experimented to the so-called moment estimator by using the power functions $h(x) = x^p$, $x \in \mathbb{R}_+$, $p = 1, 2$. Some available SPMEEX's are functions of these log-spacings as we will see soon. Here, In this paper, we aim at presenting a more general functional form in the following

$$T_n(f, s) = \sum_{j=1}^{k} f(j) \left(\log X_{n-j+1,n} - \log X_{n-j,n}\right)^s, \tag{1.2}$$

indexed by some classes $\mathcal{F}$ of functions $f : \mathbb{N}^* = \mathbb{N}\backslash\{0\} \longmapsto \mathbb{R}_+$, and by $s > 0$. We have two generalizations. First, for $s = 1$, we get

$$T_n(f, 1)/k = \sum_{j=1}^{k} f(j) \left(\log X_{n-j+1,n} - \log X_{n-j,n}\right)/k,$$

which is the functional generalization of the Diop and Lo statistics (Diop and Lo (2006, 2009) for $f(j) = j^\tau$, for $0 < \tau$ and Deme *et al.* (2012). Secondly, if f is the identity function and $s = 1$, we see that $T_n(Identity, 1)/k$ is Hill's statistic.

On the other hand, when utilizing the threshold method, we have, with the same properties of the parameters, the following statistic process :

$$S_n(f, s) = \sum_{j=1}^{k} f(j) \left(\log X_{n-j+1,n} - \log X_{n-k,n}\right)^s, . \tag{1.3}$$

This leads the couple of statistics

$$(M_{1,n}, M_{1,n}) = (S_n(\mathbf{1}, 1)/k, M_n(\mathbf{1}, 2)/k)$$

where $\mathbf{1}$ is the constant function $\mathbf{1}(x) = 1$. From this couple of statistics Dekkers *et al.* (1989) deduced the following estimator of the extreme value index

$$D_n = M_{1,n}1 + (1 - (M_{2,n}/M_{1,n}^2))^{-1}/2.$$

Our objective is to show that these two stochastic processes (1.2) and (1.3) are SPMEEXI's. In this paper, we focus on the stochastic process $T_n(f, s)$ which uses sums of independant random variables. As to $S_n(f, s)'s$, to the contrary, it uses sums of dependent random variables. Its study will be done in coming up papers.

1.2 Motivations and Scope of the Paper

As announced, we focus on the stochastic process (1.2) here. We have been able to establish its finite-dimension asymptotic distribution. As already noticed in earlier works in Lo *et al.* (2006, 2009) and in Lo, Diop and Deme (2012), the limiting law may be Gaussian or non-Gaussian. In both cases, statistical tests may be implemented. In case of non-Gaussian asymptotic limits, the limiting distribution is represented through an infinite series of standard exponential random variables. Its law may be approximated through monte-Carlo methods, as showed in Fall *et al.* (2011)

Then we prove that it is a SPMEEXI in the sense of convergence in probability. Both for asymptotic distribution and convergence in probability, the used conditions are expressed with respect to an infinite series of standard exopnential randoms variables and through the auxiliary functions a and p in the representations of df's in the extreme domain of attraction that will be recalled in the just next subsection. The conditions are next notably simplified by supposing that the *df* F is differentiable in the neighborhood of its upper endpoint.

To show how work the results for specific classes of functions f, we adapt them for $f_\tau(j) = j^\tau$, $\tau > 0$. It is interesting to see that although we have the existence of the asymptotic laws for any $\tau > 0$ and $s \geq 1$, we don't have an estimation of γ in the region $\tau < s - 1$, when $s > 1$.

One advantage of using SPMEEXI's is that we may consider the best estimators, in some sense to be precised, among all margins. We show in Theorem 3 that $T_n(f_\tau, s)$ is asymptotically Gaussian for $\tau \geq s - 1/2$. When we restrict ourselves in that domain, we are able to establish that the minimum asymptotic variance is reached for $\tau = s$. Then we construc the best estimator $T_n^{(\tau)} = T_n(f_\tau, \tau)$, that is for $\tau = s$. This is very important since the Hill estimator is $T_n^{(1)}$ itself and, as a consequence, the Hill estimator is an element of a set of best estimators indexed by τ. In fact, it is the best of all, that is $T_n^{(1)}$ has less asymptotic variance than $T_n^{(\tau)}, \tau > 1$.

It will be interesting to found out whether this minimim variance can be improved for other functional classes.

Even when we have a minimum asymptotic variance estimator, it is not sure that the performance is better for finite samples. This is why simulation studies mean reveal a best combination between bias and asymptotic variance. At finite sample size, the performance of an estimator is measured both by the bias and the variance and we don't know how the random value of the estimator is far from the exact value. We will see in the simulation Section 3 that the boundary case $\tau = s - 1/2$ gives performances similar to the optimal case.

Before we present the theoritical results and their consequences, we feel obliged to present a brief reminder of basic univariate extreme value theory and some related notation on which the statements of the results will rely on.

1.3 Basics of Extreme Value Theory

Let us make this reminder by continuing the lines of (1.1) above. If (1.1) 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-kwown that the three possible nondegenerate limits in (1.1), called extreme value *df*, are the following :

The Gumbel *df* of parameter $\gamma = 0$,

$$\Lambda(x) = \exp(-\exp(-x)),\ x \in \mathbb{R}, \tag{1.4}$$

or the Fréchet *df* of parameter $\gamma > 0$,

$$\phi_\gamma(x) = \exp(-x^{-\gamma})\mathbb{I}_{[0,+\infty[}(x),\ x \in \mathbb{R} \tag{1.5}$$

or the Weibull *df* of parameter $\gamma < 0$,

$$\psi_\gamma(x) = \exp(-(x)^{-\gamma})\mathbb{I}_{]-\infty,0]}(x) + (1 - 1_{]-\infty,0]}(x)),\quad x \in \mathbb{R}, \tag{1.6}$$

where I_A denotes the indicator function of the set A. Now put $D(\phi) = \cup_{\gamma>0} D(\phi_\gamma)$, $D(\psi) = \cup_{\gamma>0} D(\psi_\gamma)$, and $\Gamma = D(\phi) \cup D(\psi) \cup D(\Lambda)$.

In fact the limiting distribution function M is defined by an equivalence class of the binary relation $\mathcal{R}$ on the set of *df* $\mathcal{D}$ on F defined as follows :

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

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

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

$$a_n/d_n \to a \text{ and } (b_n - d_n)/c_n \to b \text{ as } n \to \infty. \tag{1.7}$$

Theses facts allow to parameterize the class of extremal distribution functions. For this purpose, suppose that (1.1) holds for the three *df*'s given in (1.4), (1.5) and (1.6). We may take sequences $(a_n > 0)_{n \geq 1}$ and $(b_n)_{n \geq 1}$ such that the limits in (1.7) are $a = \gamma = 1/\alpha$ and $b = 1$ (in the case of Fréchet extremal domain), and $a = -\gamma = -1/\alpha$ and $b = -1$ (in the case of Weibull extremal domain). Finally, one may interprets $(1 + \gamma x)^{-1/\gamma} = exp(-x)$ for $\gamma = 0$ (in the case of Gumbel extremal domain). This leads to the following parameterized extremal distribution function

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

called the Generalized Extreme Value (GEV) distribution of parameter $\gamma \in R$.

Now we give the usual representations of $df's$ lying in the extremal domain in terms of the quantile function of $G(x) = F(e^x), x \geq 1$, that is $G^{-1}(1-u) = \log F^{-1}(1-u), 0 \leq u \leq 1$.

Theorem 1. *We have :*

1. *Karamata's representation (KARARE)*

   (a) If $F \in D(\phi_{1/\gamma})$, $\gamma > 0$, then

$$G^{-1}(1-u) = \log c + \log(1 + p(u)) - \gamma \log u + (\int_u^1 b(t) t^{-1} dt), \; 0 < u < 1, \tag{1.8}$$

   where $\sup(|p(u)|, |b(u)|) \to 0$ *as* $u \to 0$ *and* c *is a positive constant and* $G^{-1}(1-u) = \inf\{x, G(x) \geq u\}, 0 \leq u \leq 1$, *is the generalized inverse of* G *with* $G^{-1}(0) = G^{-1}(0+)$.

   (b) If $F \in D(\psi_{1/\gamma})$, $\gamma > 0$, then $y_0(G) = \sup\{x, G(x) < 1\} < +\infty$ *and*

$$y_0 - G^{-1}(1-u) = c(1 + p(u)) u^\gamma \exp\Big(\int_u^1 b(t) t^{-1} dt\Big), \; 0 < u < 1, \tag{1.9}$$

   where c, $p(\cdot)$ and $b(\cdot)$ are as in (1.8)

2. *Representation of de Haan (Theorem 2.4.1 in de Haan (1970)),*

   If $G \in D(\Lambda)$, then

$$G^{-1}(1-u) = d - a(u) + \int_u^1 a(t) t^{-1} dt, \; 0 < u < 1, \tag{1.10}$$

   where d is a constant and $a(\cdot)$ admits this KARARE :

$$a(u) = c(1 + p(u)) \exp(\int_u^1 b(t) t^{-1} dt), \; 0 < u < 1, \tag{1.11}$$

   c, $p(\cdot)$ anf $b(\cdot)$ being defined as in (1.8). We warn the reader to not confuse this function $a(.)$ with the function $a_n(.,.)$ which will be defined later.

Finally, we shall also use the uniform representation of $Y_1 = \log X_1, Y_2 = \log X_2, ...$ by $G^{-1}(1-U_1), G^{-1}(1-U_2), ...$ where $U_1, U_2, ...$ are independent and uniform random variables on $(0, 1)$ and where G is the df of Y, in the sense of equality in distribution (denoted by $=_d$)

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

and hence

$$\{\{Y_{1,n}, Y_{2,n}, ...Y_{n,n}\}, n \geq 1\} \tag{1.12}$$

$$=_d \left\{\{G^{-1}(1 - U_{n,n}), G^{-1}(1 - U_{n-1,n}), ..., G^{-1}(1 - U_{1,n})\}, n \geq 1\right\}.$$

In connexion with this, we shall use the following Malmquist representation (see Shorack and Wellner (1986, p. 336) :

$$\{\log(\frac{U_{j+1,n}}{U_{j,n}})^j, j = 1, ..., n\} =_d \{E_{1,n}, ..., E_{n,n}\},$$

where $E_{1,n}, ..., E_{n,n}$ is an array of independent standard exponential random variables. We write E_i instead of $E_{i,n}$ for simplicity sake. Some conditions will be expressed in terms of these exponential random variables. We are now in position to state our first results for finite distribution asymptotic normality.

2. Our Results

We need the following conditions. First define for $n \geq 1$, f and s fixed,

$$B_n(f, s) = \max\{f(j)j^{-s}/\sigma_n(f, s), 1 \leq j \leq k\},$$

$$a_n(f, s) = \Gamma(s + 1)\sum_{j=1}^{k} f(j)\, j^{-s}$$

and

$$\sigma_n^2(f, s) = \sum_{j=1}^{k} f^2(j)\, j^{-2s}.$$

We will use the two main conditions of f and s fixed :

$$\sum_{j=1}^{\infty} f(j)^2 j^{-2s} < \infty \tag{K1}$$

and

$$\sum_{j=1}^{\infty} f(j)^2 j^{-2s} = +\infty \text{ and } B_n(f, s) \to 0 \text{ as } n \to \infty. \tag{K2}$$

Further, any df in $D(G_\gamma)$ is associated with a couple of functions (p, b) as given in the representations (1.8), (1.9) and (1.11). Define then the following notation for $\lambda > 1$,

$$b_n(\lambda) = \sup\{|b(t)|, 0 \leq t \leq \lambda k/n\}$$

and

$$p_n(\lambda) = \sup\{|p(t)|, 0 \leq t \leq \lambda k/n\}$$

We will require below that, for some $\lambda > 1$,

$$b_n(\lambda)\log k \to 0, \text{ as } n \to +\infty. \tag{CR1}$$

From now, all the limits below are meant as $n \to \infty$ unless the contrary is specified.

Here are our fundamental results. First, we have marginal estimations of the extreme value index as expected. The conditions of the results are given in very general forms that allow further, specific hypotheses as particular cases. As well, although we focus here on finite-distribution limits, the conditions are stated in a way that will permit to handle uniform studies further.

Theorem 2. *Let $F \in D(G_\gamma)$, $0 \le \gamma < +\infty$.*

(A) Case $0 < \gamma < +\infty$.

1) Let $(K1)$ hold. If $a_n(f,s) \to \infty$ and for an arbitrary $\lambda > 1$

$$a_n^{-1}(f,s)\left[\sum_{j=1}^{k} f(j)\, s\left\{\frac{\gamma}{j}E_j + \left(p_n(\lambda) + \frac{E_j}{j}b_n(\lambda)\right)\right\}^{s-1} \times \left(p_n(\lambda) + \frac{E_j}{j}b_n(\lambda)\right)\right] \to_{\mathbb{P}} 0, \tag{H0a}$$

then

$$(T_n(f,s)/a_n(f,s))^{1/s} \to_{\mathbb{P}} \gamma,$$

where $\to_{\mathbb{P}}$ stands for convergence in probability.

2) Let $(K2)$ hold. If $a_n^{-1}(f,s)\sigma_n^{-1}(f,s) \to 0$ and for an arbitrary $\lambda > 1$

$$\sigma_n^{-1}(f,s)\left[\sum_{j=1}^{k} f(j)\, s\left\{\frac{\gamma}{j}E_j + \left(p_n(\lambda) + \frac{E_j}{j}b_n(\lambda)\right)\right\}^{s-1} \times \left(p_n(\lambda) + \frac{E_j}{j}b_n(\lambda)\right)\right] \to_{\mathbb{P}} 0, \tag{H1a}$$

then

$$(T_n(f,s)/a_n(f,s))^{1/s} \to_{\mathbb{P}} \gamma.$$

(B) case $\gamma = 0$.

1) Let $(K1)$ hold. If $a_n(f,s) \to +\infty$ and for an arbitrary $\lambda > 1$,

$$a_n^{-1}(f,s)\left[\sum_{j=1}^{k} f(j)\, s\left\{j^{-1}E_j + \left(p_n(\lambda) + \frac{E_j}{j}p_n(\lambda) \vee b_n(\lambda)\log k\right)\right\}^{s-1} \times \left(p_n(\lambda) + \frac{E_j}{j}p_n(\lambda) \vee b_n(\lambda)\log k\right)\right] \to_{\mathbb{P}} 0, \tag{H0b}$$

then

$$\left(\frac{T_n(f,s)}{a_n(f,s)}\right)^{1/s} / a(k/n) \to_{\mathbb{P}} 1.$$

2) Let $(K2)$ hold. If $a_n^{-1}(f,s)\sigma_n^{-1}(f,s) \to 0$ and for an arbitrary $\lambda > 1$, and

$$\sigma_n^{-1}(f,s)\left[\sum_{j=1}^{k} f(j)\, s\left\{j^{-1}E_j + \left(p_n(\lambda) + \frac{E_j}{j}p_n(\lambda) \vee b_n(\lambda)\log k\right)\right\}^{s-1} \times \left(p_n(\lambda) + \frac{E_j}{j}p_n(\lambda) \vee b_n(\lambda)\log k\right)\right] \to_{\mathbb{P}} 0, \tag{H1b}$$

then

$$\left(\frac{T_n(f,s)}{a_n(f,s)}\right)^{1/s} / a(k/n) \to_{\mathbb{P}} 1.$$

Theorem 3. *Let* $F \in D(G_\gamma)$*,* $0 \leq \gamma < +\infty$*.*

(A) Case $0 < \gamma < +\infty$*.*

1) If $(K1)$ *and*

$$\sum_{j=1}^{k} f(j)\, s \left\{ \frac{\gamma}{j} E_j + \left(p_n(\lambda) + \frac{E_j}{j} b_n(\lambda) \right) \right\}^{s-1} \left(p_n(\lambda) + \frac{E_j}{j} b_n(\lambda) \right) \to_{\mathbb{P}} 0, \tag{H2a}$$

then

$$T_n(f,s) - \gamma^s a_n(f,s) \to \gamma^s \left\{ \Gamma(2s+1) - \Gamma(s+1)^2 \right\}^{1/2} \mathcal{L}(f,s),$$

where

$$\mathcal{L}(f,s) = \sum_{j=1}^{\infty} f(j)\, j^{-s} F_j^{(s)},$$

and the F_j^s*'s are independent and centred random variables with variance one.*

2) If $(K2)$ *and* $(H2a)$ *hold for an arbitrary* $\lambda > 1$*, then*

$$\sigma_n^{-1}(f,s)\,(T_n(f,s) - \gamma^s a_n(f,s)) \to \mathcal{N}(0, \gamma^{2s} \left\{ \Gamma(2s+1) - \Gamma(s+1)^2 \right\}).$$

(B) case $\gamma = 0$*.*

1) If $(K1)$ *and for an arbitrary* $\lambda > 1$*,*

$$\left[\sum_{j=1}^{k} f(j)\, s \left\{ j^{-1} E_j + \left(p_n(\lambda) + \frac{E_j}{j} p_n(\lambda) \vee b_n(\lambda) \log k \right) \right\}^{s-1} \right. \tag{H2b}$$

$$\left. \times \left(p_n(\lambda) + \frac{E_j}{j} p_n(\lambda) \vee b_n(\lambda) \log k \right) \right] \to_{\mathbb{P}} 0,$$

then

$$\frac{T_n(f,s)}{a^s\left(\frac{k}{n}\right)} - a_n(f,s) \to \left\{ \Gamma(2s+1) - \Gamma(s+1)^2 \right\}^{1/2} \mathcal{L}(f,s)),$$

where

$$\mathcal{L}(f,s) = \sum_{j=1}^{\infty} f(j)\, j^{-s} F_j^{(s)}.$$

2) If $(K2)$ *and* $(H2b)$ *hold for an arbitrary* $\lambda > 1$*, then*

$$\sigma_n^{-1}(f,s) \left[\frac{T_n(f,s)}{a^s(k/n)} - a_n(f,s) \right] \to \mathcal{N}\left(0, \left\{ \Gamma(2s+1) - \Gamma(s+1)^2 \right\}\right).$$

2.1 Remarks and Applications.

2.1.1 General Remarks on the Conditions

The conditions $(H0a)$, $(H0b)$, $(H1a)$, $(H1b)$, $(H2a)$ and $(H2b)$ hold if we show that the expectations of the *rv*'s of their right members tend to zero, for an arbitrary $\lambda > 1$, simply by the use of Markov's inequality. These expectations include intergrals $I(a,b,s) = \int_0^\infty (a+bx)^s e^{-x} dx$ and $J(a,b,c,s) = \int_0^\infty (a+cx)^s (a+bx) e^{-x} dx$ for real numbers a, b, c and $s \geq 1$ computed in (4.1) and (4.2) of Section (4), ***for integer values of*** s, as

$$I(a,b,s) = \int_0^\infty (a+bx)^s e^{-x} dx = s! \sum_{h=0}^{s} b^h a^{s-h}/(s-h)!.$$

and

$$\begin{aligned} J(a,b,c,s) &= s!a \sum_{h=0}^{s-1} c^h a^{s-h}/(s-h)! + s!b \sum_{h=0}^{s-1} c^h I(a,c,s-h)/(s-h)! \\ &+ s!c^s I(a,b,1). \end{aligned}$$

Then the conditions $(H0a)$, $(H0b)$, $(H2a)$, $(H2b)$, $(H1a)$ and $(H1b)$ respectively hold when hold these ones

(HE0a) $a_n^{-1}(f,s)\sum_{j=1}^{k} f(j)I_n(1,j,s) \to 0$;
(HE0b) $a_n^{-1}(f,s)\sum_{j=1}^{k} f(j)I_n(2,j,s) \to 0$;
(HE1a) $\sigma_n^{-1}(f,s)\sum_{j=1}^{k} f(j)I_n(1,j,s) \to 0$;
(HE1b) $\sigma_n^{-1}(f,s)\sum_{j=1}^{k} f(j)I_n(2,j,) \to 0$;
(HE2a) $\sum_{j=1}^{k} f(j)I_n(1,j,s) \to 0$;
(HE2b) $\sum_{j=1}^{k} f(j)I_n(2,j,s) \to 0$;

with

$$I_n(1,j,s) = sJ(p_n(\lambda), b_n(\lambda)/j, (\gamma + b_n(\lambda))/j, s-1)$$

and

$$I_n(2,j,s) = sJ(p_n(\lambda), (p_n(\lambda) \vee b_n(\lambda)\log k)/j, (1 + (p_n(\lambda) \vee b_n(\lambda)\log k))/j, s-1)$$

2.2 Weakening the Conditions for s Interger.

When the distribution function G admits an ultimate derivative at $x_0(G) = \sup\{x, G(x) < 1\}$, and this is the case for the usual df's, one may take $p(u) = 0$, as pointed out in (2014). In that case, the conditions $(HE0x), (HE1x)$ and $(HE2x)$, for $x = a$ or $x = b$, are much simpler. We then have $I_n(1,j,s) = ss!b_n(\lambda)(\gamma + b_n(\lambda))^{s-1}j^{-s}$ and $I_n(2,j,s) = ss!(b_n(\lambda)\log k)(\gamma + b_n(\lambda)\log k)^{s-1}j^{-s}$. We get these simpler conditions :

(HE0a) $a_n^{-1}(f,s)\left(ss!b_n(\lambda)(\gamma + b_n(\lambda))^{s-1}\sum_{j=1}^{k} f(j)j^{-s}\right) \to 0$;
(HE0b) $a_n^{-1}(f,s)\left(ss!(b_n(\lambda)\log k)^{s-1}\sum_{j=1}^{k} f(j)j^{-s)}\right) \to 0$;
(HE1a) $\sigma_n^{-1}(f,s)\left(ss!b_n(\lambda)(\gamma + b_n(\lambda))^{s-1}\sum_{j=1}^{k} f(j)j^{-s}\right) \to 0$;
(HE1b) $\sigma_n^{-1}(f,s)\left(ss!(b_n(\lambda)\log k)^{s-1}\sum_{j=1}^{k} f(j)j^{-s)}\right) \to 0$;
(HE2a) $\left(ss!b_n(\lambda)(\gamma + b_n(\lambda))^{s-1}\sum_{j=1}^{k} f(j)j^{-s)}\right) \to 0$;
(HE2b) $\left(ss!(b_n(\lambda)\log k)(\gamma + b_n(\lambda)\log k)^{s-1}\sum_{j=1}^{k} f(j)j^{-s)}\right) \to 0$.

It is interesting to remark that all these conditions automaticaly hold whenever $b_n(\lambda) \to 0$, and/or $(CR1)$ holds. Indeed, we remark, by the Cauchy-Scharwz's inequality, that :

$$\sigma_n^{-1}(f,s)\sum_{j=1}^{k} f(j)j^{-s} \leq \sigma_n^{-1}(f,s)\left(\sum_{j=1}^{k} f^2(j)j^{-2s}\right)^{1/2} = 1.$$

Then for $\gamma > 0$, the corresponding conditions always hold since $b_n(\lambda) \to 0$ and for $\gamma = 0$, the corresponding conditions hold with $(CR1)$. This surely leads to powerfull results. It also happens that for the usual cases, we know the values of $b_n(\lambda)$, based on $b(u) = (G^{-1}(1-u))' + \gamma$ for $\gamma > 0$ and $b(u) = us'(u)/s(u)$ with $s(u) = u(G^{-1}(1-u))'$ (see for instance Fall and Lo (2011) or Segers (2002)).

2.3 the Special Case of Diop-Lo

Now it is time to see how the preceeding results work for the particular case the functions class $f_\tau(j) = j^\tau, \tau > 0$. This special study should be a model of how to apply the results for other specific classes. Here, we will replace f by τ in all the notation meaning that $f = f_\tau$. We summarize the holding conditions depending on $\tau > 0$ and $s \geq 1$, in the following table

We may see the details as follows. First $\sum f(j)^2 j^{-2s} = \sum j^{-2(s-\tau)}$ is finite if and only if $2(s-\tau) > 1$. This gives the cases (I) and (II). For (III) in Table 1, we have

$$\sigma_n^2(\tau,s) = \sum_{j=1}^{k} j^{-2(s-\tau)} = \sum_{j=1}^{k} j^{-1} \sim (\log k),$$

by (4.3) in Section (4). Since for $1 \leq j \leq k$, $f(j)j^{-s}/\sigma_n(\tau,s) = j^{-1/2}/\sigma_n(\tau,s) \leq 1/\sigma_n(\tau,s)$, we have $B_n(\tau,s) \leq \sigma_n^{-1}(\tau,s) \to 0$ and then $(K2)$ holds. For (IV),

$$\sigma_n^2(\tau,s) = \sum_{j=1}^{k} j^{-2(s-\tau)} \sim k^{(2(\tau-s)+1))}/(2(\tau-s)+1)),$$

Table 1. Checking the conditions for the Diop-Lo class

(I)	(II)	(III)	(IV)
$\tau < s-1$	$s-1 \leq \tau < s-1/2$	$\tau = s-1/2$	$\tau > s-1/2$
$(K1)$	$(K1)$	$(K2)$	$(K2)$
a_n bounded	$a_n \to \infty$	$a_n \sim 2s!k^{1/2}$	$a_n \sim s!\frac{k^{\tau-s+1}}{\tau-s+1}$
σ_n bounded	σ_n bounded	$\sigma_n \sim (\log k)^{1/2}$	$\sigma_n \sim \frac{k^{(\tau-s+1/2)}}{\sqrt{2(\tau-s)+1}}$

by (4.4). Next $1 \leq j \leq k$, $f(j)j^{-s}/\sigma_n(\tau,s) = j^{(\tau-s)}/\sigma_n(\tau,s) = C_0(j/k)^{(\tau-s)}k^{-1/2}$. Since $(\tau - s) > 0, B_n(\tau,s) \leq C_0k^{-1/2} \to 0$. Then $(K2)$ also holds. The lines above also explain the fourth row of the table. The third row is immediate since $a_n = \sum_{j=1}^{k} j^{-(s-\tau)} \to \infty$ for $(s-\tau) \leq 1$ and remains bounded for $(s-\tau) > 1$.

It is worth mentioning that the case $\tau < s-1$ is not possible for $s = 1$. This unveils a new case comparatively with former studies of Deme *et al.* (2012) for $s = 1$.

Now, based on these facts, we are able to get the following estimations (results $(CR1)$ for $\gamma = 0$) :

1. For $s-1 \leq \tau < s-1/2$, $a_n \to \infty$. Hence we have the estimation

$$\left(\frac{T_n(\tau,s)}{a_n(\tau,s)}\right)^{1/s} / a(k/n) \to_{\mathbb{P}} 1.$$

   For $\gamma > 0$, $a(k/n) \to \gamma$. For $\gamma = 0$, $\left(\frac{T_n(\tau,s)}{a_n(\tau,s)}\right)^{1/s} \to \gamma = 0$ at the rate of $a_n(k/n)$.

2. $0 < \tau < s-1$. We do not have an estimation of γ.

For testing the hypothesis $F \in D(G_\gamma)$, $\gamma \geq 0$, we derive the following laws by the delta-method under $(CR1)$, especially for $\gamma = 0$.

Let $s-1 \leq \tau < s-1/2$. For $\gamma > 0$,

$$a_n(\tau,s)\left\{\left(\frac{T_n(\tau,s)}{a_n(\tau,s)}\right) - \gamma^s\right\}$$

$$\to \gamma^s \left\{\Gamma(2s+1) - \Gamma(s+1)^2\right\}^{1/2} \mathcal{L}(\tau,s). \tag{2.1}$$

For $\gamma = 0$

$$a_n(\tau,s)\left\{a(k/n)^{-s}\left(\frac{T_n(\tau,s)}{a_n(\tau,s)}\right) - 1\right\} \to \left\{\Gamma(2s+1) - \Gamma(s+1)^2\right\}^{1/2} \mathcal{L}(\tau,s).$$

Let $\tau \geq s-1/2$. In this case $a_n(\tau,s)\sigma_n^{-1}(\tau,s) \to +\infty$. This enables the delta-method application to the limit in Theorem 3, case (A-2), that is

$$\frac{a_n(\tau,s)}{\sigma_n(\tau,s)}\left\{\left(\frac{T_n(\tau,s)}{a_n(\tau,s)}\right) - \gamma^s\right\} \to \mathcal{N}(0, \gamma^{2s}\left\{\Gamma(2s+1) - \Gamma(s+1)^2\right\}). \tag{2.2}$$

We derive

$$\frac{a_n(\tau,s)}{\sigma_n(\tau,s)}\left\{\left(\frac{T_n(\tau,s)}{a_n(\tau,s)}\right)^{1/s} - \gamma\right\} \to \mathcal{N}(0, (\gamma/s)\left\{\Gamma(2s+1) - \Gamma(s+1)^2\right\})$$

For $\gamma = 0$

$$\frac{a_n(\tau,s)}{\sigma_n(\tau,s)}\left\{a(k/n)^{-s}\left(\frac{T_n(\tau,s)}{a_n(\tau,s)}\right) - 1\right\} \to \mathcal{N}(0, \left\{\Gamma(2s+1) - \Gamma(s+1)^2\right\}).$$

For the new case $\tau < s-1$, we have for $\gamma > 0$,

$$T_n(\tau,s) - A(\tau,s)\gamma^s \to \gamma^s \left\{\Gamma(2s+1) - \Gamma(s+1)^2\right\}^{1/2} \mathcal{L}(\tau,s),$$

where $A(\tau, s) = \sum_{j=1}^{\infty} j^{-(s-\tau)} < \infty$, for $\gamma = 0$,

$$T_n(\tau, s)/a^s(k/n) - A(\tau, s) \to \left\{\Gamma(2s+1) - \Gamma(s+1)^2\right\}^{1/2} \mathcal{L}(\tau, s).$$

These two limiting laws also allow statistical tests based on Monte-Carlo methods as in Deme *et al.* (2012).

2.4 Best Performance Estimators

In pratical situations, we have to select a particular function f from a particular class $\mathcal{F}$ of function f. A natural question is to select a couple (f, s) for which the estimator is the best in some sense. Here we consider the class of Diop-Lo, $f(j) = j^{\tau}$ and we are interested in finding the best performance of the estimator $\left(\frac{T_n(\tau,s)}{a_n(\tau,s)}\right)^{1/s}$ of $\gamma > 0$. We place ourselves in the normality domain, that is $\tau > s - 1/2$. Straigthforward computations from (2.2) yield

$$V_n(\tau, s)\left[\left(\frac{T_n(\tau, s)}{a_n(\tau, s)}\right)^{1/s} - \gamma\right] \rightsquigarrow \boldsymbol{N}(0, 1)$$

with

$$V_n(\tau, s) = \frac{a_n(\tau, s)}{\sigma_n(\tau, s)} \times \frac{s\gamma^{-1}}{\sqrt{\Gamma(2s+1) - \Gamma(s+1)^2}}.$$

So, finding the best performance is achieved for minimum value for the asymptotic variance $V_n(\tau, s)^{-2}$. We then have to find the greatest value of variance $V_n(\tau, s)$. But maximizing this function both in s and τ might be tricky. However, for a fixed $s \geq 1$, we may find that the maximum value of $V_n(\tau, s)$ for $\tau \in [s - 1/2, +\infty[$. First, we have to isolate the boundary point $\tau = s - 1/2$. We prove in Subsection 4.4 below that the maximum value of $V_n(s, \tau)$ is reached when $\tau = s$. Using the formulae in Table 1, we see that, for $\tau \geq s - 1/2$, we have

$$V_n(\tau, \tau) = \tau!\,\sqrt{(k)} \times \frac{s\gamma^{-1}}{\sqrt{\Gamma(2\tau+1) - \Gamma(\tau+1)^2}} \tag{2.3}$$

and for $\tau = s - 1/2$, we get

$$V_n(\tau, \tau + 1/2) = 2s!\,\sqrt{(k/\log k)} \times \frac{s\gamma^{-1}}{\sqrt{\Gamma(2(\tau+1)) - \Gamma(\tau+3/2)^2}}. \tag{2.4}$$

We get as best estimator with least asymptotic variance

$$T_n(\tau)^{(\tau)} = \left(\frac{T_n(\tau, \tau)}{a_n(\tau, \tau)}\right)^{1/s}$$

for the normality zone $\tau \geq s - 1/2$. Its asymptotic variance (2.3) increases when τ decreases. This means that the Hill estimator is the best with respect to this sense. Now, let us move to the non-Gaussian zone, that is $0 \leq s - 1 \leq \tau < s - 1/2$, corresponding to the column II in the Table 1. We may easily derive from (2.1) that the asymptotic variance is of equivalent to

$$\gamma\frac{\Gamma(2s+1) - \Gamma(s+1)^2}{s(\tau - s + 1)},$$

which is still dominated by $V_n(\tau, \tau)^{-1}$. To sum up, we say that the Hill estimator has best asymptotic variance for all margins.

However for finite sample, we do not know how far the centered and normalized statistic is from the limiting Gaussian variable or the non-Gaussan limiting random. Here we are obliged to back on simulation studies. Let us consider

$$T_n^{(\tau)} = \left(\frac{T_n(\tau, \tau)}{a_n(\tau, \tau)}\right)^{1/\tau} \tag{2.5}$$

and

$$T_n^{(\tau+1/2)} = \left(\frac{T_n(\tau, \tau + 1/2)}{a_n(\tau, \tau + 1/2)}\right)^{1/(\tau+1/2)}$$

We get that these two estimators generally behave better than the Hill's and the Dekkers *et al.*'s ones. At least, they have equivalent performances. But absolutely, they seem to be more stable in a sense to precised later. This must result in lesser biases that constitute a compensation of their poorer performance regarding the asymptotic variance point of view. A full report of simulation studies are given in Section 3.

But we should keep in mind that the results presented here, go far beyond the Diop-Lo family for which the Hill's estimator demonstates to be the least asymptotic variance estimator.

Further, researches will be conducted on other functions families in order to possibly find out estimators with asymptotic variances better that $1/V_n(1,1)$.

2.5 Proofs

We will prove both theorems together. For both cases $\gamma > 0$ and $\gamma = 0$, we will arrive at the final statement based on the hypotheses $(K1)$ or $(K2)$, $(H1a)$ or $(H1a),(H2a)$ in the case $\gamma > 0$ (A) , and on $(K1)$ or $(K2)$, $(H0b)$ or $(H1b),(H2b)$ in the case $\gamma = 0$ (B). In each case, an analysis will give the corresponding parts (1) and (2) for the two theorems. We begin with the A case.

Case (A) : Here, $F \in D(G_\gamma)$ with $0 < \gamma < +\infty$.

By using (1.12), we have

$$T_n(f,s) = \sum_{j=1}^{k} f(j)(G^{-1}(1-U_{j,n}) - G^{-1}(1-U_{j+1,n}))^s$$

By (1.8), we also have

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

For $1 \leq j \leq k$,

$$\begin{aligned} G^{-1}(1-U_{j,n}) - G^{-1}(1-U_{j+1,n}) &= \log\left[\frac{1+p(U_{j,n})}{1+p(U_{j+1,n})}\right] \\ +\gamma \log\left(\frac{U_{j+1,n}}{U_{j,n}}\right) &+ \int_{U_{j,n}}^{U_{j+1,n}} b(t)t^{-1}dt. \end{aligned}$$

Put $p_n = \sup\{|p(t)|, 0 \leq t \leq U_{k+1,n}\}$ and $b_n = \sup\{|b(t)|, 0 \leq t \leq U_{k+1,n}\}$. Both b_n and p_n tend to zero in probability as $n \to +\infty$, since $U_{k+1,n} \to 0$ when $(n, k/n) \to (+\infty, 0)$. We then get

$$T_n(f,s) = \sum_{j=1}^{k} f(j)\left\{O_{\mathbb{P}}(p_n) + \frac{\gamma}{j}E_j + \frac{E_j}{j}O_{\mathbb{P}}(b_n)\right\}^s$$

Put

$$A_{n,j} = \left(O_{\mathbb{P}}(p_n) + \frac{\gamma}{j}E_j + \frac{E_j}{j}O_{\mathbb{P}}(b_n)\right)^s.$$

By the mean value Theorem, we have for $s \geq 1$

$$\begin{aligned} A_{n,j} - \left(\frac{\gamma}{j}E_j\right)^s &= s\left\{\frac{\gamma}{j}E_j + \theta_{n,j}\left(O_{\mathbb{P}}(p_n) + \frac{E_j}{j}O_{\mathbb{P}}(b_n)\right)\right\}^{s-1} \\ &\times \left(O_{\mathbb{P}}(p_n) + \frac{E_j}{j}O_{\mathbb{P}}(b_n)\right) \end{aligned}$$

where for $\left|\theta_{n,j}\right| \leq 1$. Put

$$\zeta_{n,j}(s) = s\left\{\frac{\gamma}{j}E_j + \theta_{n,j}\left(O_{\mathbb{P}}(p_n) + \frac{E_j}{j}O_{\mathbb{P}}(b_n)\right)\right\}^{s-1}\left(O_{\mathbb{P}}(p_n) + \frac{E_j}{j}O_{\mathbb{P}}(b_n)\right) \tag{2.6}$$

We have

$$T_n(f,s) = \gamma^s \sum_{j=1}^{k} f(j) j^{-s} E_j^s + \sum_{j=1}^{k} f(j)\zeta_{n,j}(s).$$

Recall that $\mathbb{E}\left(E_j^s\right) = \Gamma(s+1)$ and $\mathbb{V}(E_j^s) = \Gamma(2s+1) - \Gamma(s+1)^2$ and denote

$$\begin{aligned} V_n(f,s) &= \sum_{j=1}^{k} f(j)\, j^{-s}(E_j^s - s!) \\ &= \left\{\Gamma(2s+1) - \Gamma(s+1)^2\right\}^{\frac{1}{2}} \sum_{j=1}^{k} f(j)\, j^{-s} F_j^{(s)}(s), \end{aligned}$$

and for a fixed $s \geq 1$,

$$F_j^{(s)}(s) = (E_j^s - s!)/\left\{\Gamma(2s+1) - \Gamma(s+1)^2\right\}^{\frac{1}{2}}$$

is a sequence of independent mean zero random variables with variance one. We have

$$\begin{aligned} T_n(f,s) - \gamma^s a_n(f,s) &= \gamma^s \left\{\Gamma(2s+1) - \Gamma(s+1)^2\right\} \sum_{j=1}^{k} f(j)\, j^{-s} F_j^{(s)}(s) \\ &+ \sum_{j=1}^{k} f(j)\zeta_{n,j}(s). \end{aligned} \tag{2.7}$$

By (2.6),

$$\begin{aligned} \left|\zeta_{n,j}(s)\right| &= s\left\{\frac{\gamma}{j}E_j + \theta_{n,j}\left(|O_{\mathbb{P}}(p_n)| + \frac{E_j}{j}|O_{\mathbb{P}}(b_n)|\right)\right\}^{s-1} \\ &\times \left(|O_{\mathbb{P}}(p_n)| + \frac{E_j}{j}|O_{\mathbb{P}}(b_n)|\right) \end{aligned}$$

When $(K1)$ holds, Kolmogorov's Theorem on sums of centered random variables ensures that

$$\left\{\Gamma(2s+1) - \Gamma(s+1)^2\right\}^{1/2} \sum_{j=1}^{k} f(j)\, j^{-s} F_j^{(s)}(s)$$

converges to the *rv*

$$\left\{\Gamma(2s+1) - \Gamma(s+1)^2\right\}^{1/2} \sum_{j=1}^{+\infty} f(j)\, j^{-s} F_j^{(s)}(s) = \left\{\Gamma(2s+1) - \Gamma(s+1)^2\right\}^{1/2} \mathcal{L}(f,s)$$

which is centered and has variance one, as completely described in Lemma (1).

Now $(H0a)$ and Lemma (2) ensure that the second term of (2.7) tends to zero in probability. Then if $a_n(f,s) \to \infty$, we get that $T_n(f,s)/a_n(f,s) \to \gamma^s$. This proves Part (A)(1) of Theorem (2) for $\gamma > 0$. Further if $(H2a)$ holds, Lemma (2) and the first point yields that $T_n(f,s) - \gamma^s a_n(f,s)$ asymptotically behaves as $L(f,s) = \left\{\Gamma(2s+1) - \Gamma(s+1)^2\right\}^{1/2} \sum_{j=1}^{+\infty} f(j)\, j^{-s} F_j^{(s)}(s)$ since the second term of (2.7) is zero at infinity. This proves Part (A)(1) of Theorem (3) for $\gamma > 0$.

Now suppose that $(K1)$ does not hold and that $(K2)$ and $(H2a)$ both hold. Also $(H2a)$ implies via Lemma (1) that the second term, when divided by $\sigma_n(f,s)$, tends to zero in probability. Next by Lemma (2),

$$\sigma_n^{-1}(f,s)\left\{\Gamma(2s+1)-\Gamma(s+1)^2\right\}^{1/2}\sum_{j=1}^{k} f(j)\, j^{-s}F_j^{(s)}(s)$$

asymptotically behaves as a $\mathcal{N}(0,1)$ rv under $(K2)$. It follows under these circumtances that

$$(a_n(f,s)/\sigma_n(f,s))(T_n(f,s)/a_n(f,s)-\gamma^s)\to N(0,1).$$

This ends the proof of Part (A)(2) of Theorem (2). Further, whenever $a_n(f,s)/\sigma_n(f,s)\to\infty$,

$$T_n(f,s)/a_n(f,s)\to\gamma^s,$$

which establishes Part (A) (2) of Theorem (2).

Case B : $F\in D(G_0)$, $\gamma=0$. Use representations (1.10) and (1.11) to get for $1\le j\le U_{k+1,n}$,

$$G^{-1}(1-U_{j,n})-G^{-1}(1-U_{j+1,n})=a\left(U_{j+1,n}\right)-a\left(U_{j,n}\right)+\int_{U_{j,n}}^{U_{j+1,n}} a(t)t^{-1}dt.$$

Remark that for $U_{1,2}<u,v<U_{k,n}$

$$\frac{v}{u}<\frac{U_{k,n}}{U_{1,n}}=\frac{U_{k,n}}{U_{k-1,n}}\times\frac{U_{k-1,n}}{U_{k-2,n}}\times\frac{U_{k-2,n}}{U_{k-3,n}}\times\cdots\times\frac{U_{2,n}}{U_{1,n}}$$

and

$$0\le\log\left(\frac{u}{v}\right)=\sum_{j=1}^{k-1}\log\left(\frac{U_{j+1,n}}{U_{j,n}}\right)=\sum_{j=1}^{k-1} j^{-1}E_j=\sum_{j=1}^{k-1} j^{-1}\left(E_j-1\right)+\sum_{j=1}^{k-1} j^{-1}.$$

By Kolmogorov's theorem for partial sums of independent and mean zero random variables, $\sum_{j=1}^{k-1} j^{-1}\left(E_j-1\right)$ converges in law to a finite rv E. We have $\sum_{j=1}^{k-1} j^{-1}\sim\log k$ and this ensures that for $1\le u\le U_{k+1,n}$, $(1+p(u))/(1+p(U_{k+1,n}))-1=O_P(p_n)$,

$$\exp\left(\int_u^1 b(t)t^{-1}dt\right)/\exp\left(\int_{U_{k+1,n}}^1 b(t)t^{-1}dt\right)-1=O_P(b_n\log k)$$

both uniformly in $1\le u\le U_{k+1,n}$ and finally

$$a(U_{j+1,n})/a(U_{j,n})=(1+O_P(p_n))\exp(O_P(b_n)E_j/j),$$

But for $1\le j\le k$, $\left|j^{-1}E_j\right|\le\sum_{j=1}^{k-1} j^{-1}E_j=O_p(\log k)$. Then if $b_n\log k\to_{\mathbb{P}}0$,

$$\begin{aligned}a(U_{j+1,n})/a(U_{j,n})&=(1+O_P(p_n))(1+b_nE_j/j)\\&=1+O_P(p_n)+O_P(p_nb_nE_j/j)+O_P(b_nE_j/j).\end{aligned}$$

Finally, since $a(k/n)/a(U_{k+1,n})=1+O_P(1)$, it follows that

$$a(u)/a(k/n)=1+O_P((p_n\vee b_n)\log k)$$

uniformly in $1\le u\le U_{k+1,n}$. This finally leads to for $1\le j\le U_{k+1,n}$,

$$\begin{aligned}B_{j,n}(s)&=\left(G^{-1}(1-U_{j,n})-G^{-1}(1-U_{j+1,n})\right)/a(k/n)\\&=(1+O_P((p_n\vee b_n)\log k))\times\left\{O_P(p_n)+O_P(p_nb_nE_j/j)+O_P(b_nE_j/j)\right\}\end{aligned}$$

$$+\{1+O_P((p_n\vee b_n)\log k)\}\,E_j/j$$

$$=E_j/j+R_{j,n}(s),$$

where

$$R_{j,n}(s)=(1+O_P((p_n\vee b_n)\log k))\times\left\{O_P(p_n)+O_P(p_nb_nE_j/j)+O_P(b_nE_j/j)\right\}$$

$$+O_P((p_n\vee b_n)\log k)E_j/j.$$

We can easily show that

$$R_{j,n}(s)=O_P(p_n)+O_P((p_n\vee b_n)\log k)E_j/j$$

and remark that $E_j/j=O_P(\log k)$ uniformly in $j\in\{1,...,k\}$. This yields

$$\frac{T_n\,(f,s)}{a^s\,(k/n)}=\sum_{j=1}^{k}f\,(j)\left\{O_P(p_n)+O_P((p_n\vee b_n)\log k)E_j/j\right\}^s.$$

We have by the same methods used above

$$\frac{T_n\,(f,s)}{a^s(k/n)}=\sum_{j=1}^{k}f\,(j)\,j^{-s}E_j^s+\sum_{j=1}^{k}f\,(j)\,\xi_{n,j}(s)$$

$$\begin{aligned}\xi_{n,j}(s) &= s\left\{E_jj^{-1}+\theta_{n,j}\left(O_{\mathbb{P}}(p_n)+\frac{E_j}{j}O_{\mathbb{P}}(p_n\vee b_n\log k)\right)\right\}^{s-1}\\ &\times\ \left(O_{\mathbb{P}}(p_n)+\frac{E_j}{j}O_{\mathbb{P}}(p_n\vee b_n\log k)\right)\end{aligned}$$

And further

$$\frac{T_n\,(f,s)}{a^s\,(k/n)}-a_n\,(f,s)=\sum_{j=1}^{k}f\,(j)\,j^{-s}\left(E_j^s-s!\right)+\sum_{j=1}^{k}f\,(j)\,\xi_{n,j}(s).$$

$$\frac{T_n\,(f,s)}{a^s\,(k/n)}=\left\{\Gamma\,(2s+1)-\Gamma\,(s+1)^2\right\}^{1/2}\sum_{j=1}^{k}f\,(j)\,j^{-s}F_j^{(s)}\,(s)+\sum_{j=1}^{k}f\,(j)\,\xi_{n,j}(s).\tag{2.8}$$

When we compare Formulas (2.7) and (2.8), $(H0a)-(H2a)-(H1a)$ with $(H0b)-(H2b)-(H1b)$, we use Lemmas (1) and (2) and reconduct almost the same conclusion already done for the $\gamma>0$ case to prove the parts (B) of Theorems (2) and (3).

3. Simulation Studies

Nowadays, simulation studies are very sophisticated and may be very difficult to follow. Here, we want give a serious comparison of our estimators with several analoges while keeping the study reasonably simple. Let us begin to explain the stakes before proceeding any further. The estimators of the extremal index generally use a number, say k like in this text, of the greatest observations : X_{n-j+1}, $1\leq j\leq k$. For almost all such estimators, we have a small bias and a great variance for large values of k, and the contrary happens for small values of k. This leads to the sake of an optimal value of k keeping both the bias and the variance at a low level. A related method consists in considering a range of values $kv(j)=kmin+j(kmin-kmin)/ksize$, $1\leq j\leq ksize$ over which the observed values of the statistic are stable and well approximate the index. This second method seems preferable when comparing two estimators with respect to the bias.

So we fix a sample size n and consider the range of values as described above where $kmin$ and $kmax$ are suitably chosen. Thus checking the curves of two statistics over the interval $[mink, maxk]$ is a good tool for comparing their performances. Next, for each j, $1\leq kmin\leq j\leq kmax$, we compute the mean square error of the estimated values of γ for values of k in a neighborhood of $kv(j)$, that is for $k\in[kv(j)-kstab, kv(j)+kstab]$, $1\leq j\leq ksize$, where $kstab$ is also suitably fixed.

The mimimun of these MES's certainly corresponds to the most stable zone and may be taken as the best estimation when it is low enough.

Here, we compare our class of estimators, represented the optimal estimator (2.4) and the boundary form (2.5) for $s \in [1, 5]$, with the estimators of Hill and Dekkers *et al.*. The estimators (2.4) and (2.5) for $s \in [1, 5]$ fall in the asymptotic normality area $s - 1/2 \leq \tau$. In a larger study, we will include the Pickands' statistic and consider the nongaussian asymptotic area.

The study will only cover the heavy tail case, that $\gamma > 0$. The case $\gamma = 0$ will be part of a large simulation paper. And we consider a pure Pareto law (I), and two perturbed ones laws (II) and (III):

$$\text{(I) } F^{-1}(1-u) = u^{-\gamma}$$
$$\text{(II) } F^{-1}(1-u) = u^{-\gamma}(1-u^{\beta})$$
$$\text{(III) } F^{-1}(1-u) = u^{-\gamma}(1-(-1/\log u)^{\beta}).$$

3.1 Simulations for $\gamma > 0$

Let $f(j) = j^{\tau}$, $\tau > 0$. Our results say that $\left(\frac{T_n(\tau,s)}{a_n(\tau,s)}\right)^{1/s}$ is an estimator of $\gamma \geq 0$ if $s - 1 \leq \tau$.

Theoritically, we then have in hands an infinite class of estimators. We should be able to find values of s and τ leading a lowest stable bias and hope that this bias will be lower than of the other analogues, or to be at their order at least. We know from Subsection 4.4 that our class of estimators for $f = f_{\tau}$ has miminal asymptotic variance for $s = \tau$. But it is not sure that this corresponds to the best performance for finite samples. And following the remark of Deme *et al.* (2012) who noticed that the boundary case, that discriminates the Gaussian and the non Gaussian asymptotic laws, behaves well, we include it also here, that is the case $s - 1/2 = \tau$.

We fix the following values : $n = 1000$, $kmin = 105$, $kmax = 375$, $ksize = 100$, $kstab = 5$. And we fix the number of replications to $B = 1000$.

How to read our results? For each $j \in [1, ksize]$, we compue the mean square error $MSE(j)$ when k spans $[kv(j) - kstab, kv(j) + kstab]$, that is

$$MSE(j)^2 = \frac{1}{2kstab+1} \sum_{k \in [kv(j)-kstab, kv(j)+kstab]} (T_n(k) - \gamma)^2.$$

Next we take the mimimum and the maximum values of these MSE(j)'s denoted as *Min* and *Max*. The difference (*df*) Diff=*Max-Min* is reported as well as the middle term $Mid = (Min + Max)/2$.

We classify the estimators with respect to both the values of *Mid* and *Diff*. If the values *Diff* are of the same order for two estimators, we will prefer the one with the minimum value of *Mid*. That mean that this latter estimator is more stable and then, is better.

Since the Hill estimator and the Dekkers *et al.* do not depend on parameters, we have conducted a series of simulations and the performances are of the order of values given in Table 2. Next, we conduct simulations on the performances of the boundary Double Hill and the optimal double Hill statistics for $s = 1, ..., s = 5$ in Tables 3, 4, 5, 6, 7, 8.

Table 2. Performances of Hill and Dekkers *et al.* estimators

Values	Model I		Model II		Model III	
	Hill	Dekkers	Hill	Dekkers	Hill	Dekkers
Min	1.529110^{-5}	7.683510^{-3}	3.575110^{-3}	9.324510^{-5}	1.042810^{-1}	5.575110^{-2}
Max	2.033810^{-3}	1.515810^{-2}	3.724210^{-2}	1.648410^{-2}	8.442510^{-1}	5.456810^{-1}
Diff	2.018510^{-3}	7.474510^{-3}	3.366710^{-2}	1.639110^{-2}	7.399610^{-1}	4.901110^{-1}
Mid	1.024510^{-3}	1.142010^{-2}	2.040810^{-2}	8.288710^{-3}	4.742710^{-1}	3.008010^{-1}

Table 3. Performances of the Boundary Double Hill statistics for s=1,...,5 with Model I

Values	Model I Double Hill				
	s=1	s=2	s=3	s=4	s=5
Min	6.203010^{-4}	1.111310^{-3}	2.598810^{-2}	5.237910^{-3}	6.655810^{-3}
Max	1.903710^{-3}	2.181410^{-3}	3.014210^{-2}	6.032410^{-3}	8.443310^{-3}
Diff	1.283410^{-3}	1.070110^{-3}	4.153710^{-3}	7.945410^{-4}	1.787410^{-3}
Mid	1.262010^{-3}	1.646310^{-3}	2.806510^{-2}	5.635210^{-3}	7.549610^{-3}

Table 4. Performances of the Optimal Double Hill statistics for s=1,...,5 with Model I

Values	Model I Optimal Double Hill				
	s=1	s=2	s=3	s=4	s=5
Min	4.712810^{-3}	1.404210^{-4}	9.436910^{-7}	5.456610^{-3}	3.849510^{-3}
Max	9.851410^{-3}	5.374710^{-3}	4.117010^{-3}	1.080210^{-2}	1.606810^{-2}
Diff	5.138610^{-3}	5.234210^{-3}	4.116010^{-3}	5.345410^{-3}	1.221810^{-2}
Mid	7.282110^{-3}	2.757510^{-3}	2.058910^{-3}	8.129310^{-3}	9.958910^{-3}

Table 5. Performances of the Boundary Double Hill statistics for s=1,...,5 with Model II

Values	Model II Double Hill				
	s=1	s=2	s=3	s=4	s=5
Min	1.825810^{-3}	1.322610^{-3}	2.981110^{-2}	1.562110^{-5}	8.124610^{-6}
Max	1.009710^{-2}	9.692610^{-3}	3.399510^{-2}	5.227810^{-3}	6.245110^{-3}
Diff	8.272010^{-3}	8.369910^{-3}	4.184010^{-3}	5.212210^{-3}	6.236910^{-3}
Mid	5.961810^{-3}	5.507610^{-3}	3.190310^{-2}	2.621710^{-3}	3.126610^{-3}

Table 6. Performances of the Optimal Double Hill statistics for s=1,...,5 with Model II

Values	Model II Optimal Double Hill				
	s=1	s=2	s=3	s=4	s=5
Min	1.843210^{-3}	1.345110^{-3}	1.495410^{-3}	1.562110^{-5}	8.124610^{-6}
Max	1.009710^{-2}	9.692610^{-3}	3.399510^{-2}	7.997210^{-3}	6.171210^{-3}
Diff	8.254610^{-3}	8.347410^{-3}	3.249910^{-2}	7.981610^{-3}	6.163010^{-3}
Mid	5.970510^{-3}	5.518810^{-3}	1.774510^{-2}	4.006410^{-3}	3.089610^{-3}

Table 7. Performances of the Boundary Double Hill statistics for s=1,...,5 with Model III

Values	Model III Double Hill				
	s=1	s=2	s=3	s=4	s=5
Min	3.351710^{-2}	3.144610^{-2}	3.144610^{-2}	2.493710^{-2}	2.550910^{-2}
Max	1.879410^{-1}	2.711010^{-1}	2.711010^{-1}	6.190110^{-1}	9.538710^{-1}
Diff	1.544210^{-1}	2.396510^{-1}	2.396510^{-1}	5.940710^{-1}	9.283610^{-1}
Mid	1.107210^{-1}	1.511210^{-1}	1.512710^{-1}	3.219710^{-1}	4.896910^{-1}

Table 8. Performances of the Optimal Double Hill statistics for s=1,...,5 with Model III

Values	Model III Optimal Double Hill				
	s=1	s=2	s=3	s=4	s=5
Min	3.197910^{-2}	3.182110^{-2}	3.182110^{-2}	2.536110^{-2}	2.606010^{-2}
Max	1.879210^{-1}	2.711010^{-1}	2.711010^{-1}	6.190110^{-1}	9.538710^{-1}
Diff	1.559410^{-1}	2.392810^{-1}	2.392810^{-1}	5.936410^{-1}	9.278110^{-1}
Mid	1.099510^{-1}	1.514610^{-1}	1.514610^{-1}	3.221810^{-1}	4.899610^{-1}

Now, we are able to draw a number of conclusions and remarks based on the Tables 2, 3, 4, 5, 6, 7, 8.

(1.) Model I : Compared to Hill's statistic and Dekkers *et al.*'s estimator, our Double Hill and Optimal Double Hill estimors are generally more stable and mostly better with regard to the middle value for $s = 1$, $s = 2$, $s = 4$.

(2.) Model II : we have simular results.

(3.) Model III : For $s \leq 3$. We also get similar results.

(4.) As a general conclusion, we say that our Double Hill estimator, for both cases of of boundary and minimum variance, behaves like these Hill's and Dekkers *et al.*'s estimator, are more stable and sloghtly better.

(5.) All these estimators present poorer performances for the very perturbated model III. But unlikely to the Hill's and Dekkers *et al.*'s estimators, we have in hand a SPMEEXI and hopefully we will be able to get better new estimators for model III through suitable combinations in future works. This is important since the Hill's and Dekkers *et al.*'s are the most used in the applications.

(6.) We may find theses patterns on Figure 1 and Figure 2.

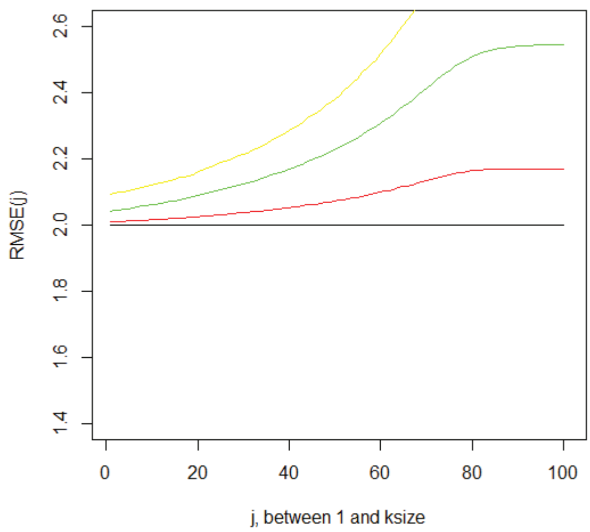

Figure 1. Curves of the mean values of the RMSE's at 100 values of k(j) (j=1,...,100) computed on eleven points around k(j) for the statistics : Hill [blue], Dekkers *et al.* [green], Boundary Double Hill [red] and Optimal Double Hill [yellow] for s=1

4. Technical Results

4.1 Technical Lemmas

We begin by this simple lemma where we suppose that we are given a sequence of independent and identically mean zero $rv's$ $F_1, F_2, ..., F_k$ with variance unity. Denote

$$A(f, s) = \sum_{j=1}^{\infty} f(j)^2 j^{-2s}$$

and

$$\sigma_n^2(f, s) = \sum_{j=1}^{k} f(j)^2 j^{-2s}$$

Lemma 1. *Let*

$$V_n(f, s) = \sum_{j=1}^{k} f(j) j^{-s}(E_j^s - s!) =: \left\{\Gamma(2s+1) - \Gamma(s+1)^2\right\}^{1/2} \sum_{j=1}^{k} f(j) j^{-s} F_j^{(s)},$$

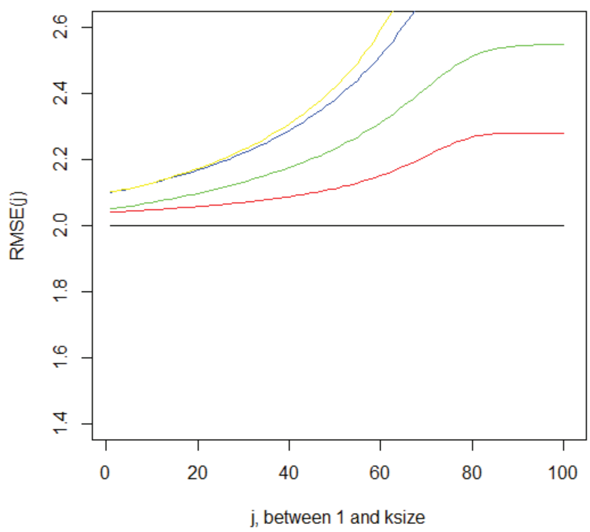

Figure 2. Curves of the mean values of the RMSE's at 100 values of k(j) (j=1,...,100) computed on eleven points around k(j) for the statistics : Hill [blue], Dekkers *et al.* [green], Boundary Double Hill [red] and Optimal Double Hill [yellow] for s=2

wher the $F_j^{(s)}$ are independent centered random variables with variance one.

(1) If $(K1)$: $A(f,s) < \infty$, then

$$V_n(f,s) \rightsquigarrow \left\{\Gamma(2s+1) - \Gamma(s+1)^2\right\}^{1/2} \mathcal{L}(f,s).$$

(2) If $(K2)$: $B_n(f,s) = \max\{f(j)j^{-s}/\sigma_n(f,s), 1 \leq j \leq k\} \to 0$, then

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

Proof. Put

$$V_n^*(f,s) = \sigma_n(f,s)^{-1}V_n(f,s).$$

and suppose that $(K1)$ holds. Then Kolmogorov's Theorem for sums of zero mean indenpendent rv's applies. Since the series $\sum_{j\geq 1} Var(f(j)j^{-s}F_j^{(s)})$ is finite, we have :

$$V_n(f,s) \to \left\{\Gamma(2s+1) - \Gamma(s+1)^2\right\}^{-1/2} \sum_{j=1}^{\infty} f(j)j^{-s}F_j^{(s)} = \left\{\Gamma(2s+1) - \Gamma(s+1)^2\right\}^{1/2} \mathcal{L}(f,s),$$

Now suppose that $(K2)$ holds. Let us evaluate the moment generating function of $V_n^{**}(f,s) = \sigma_n(f,s)^{-1}\sum_{j=1}^{k} f(j)j^{-s}F_j^{(s)}$:

$$\phi_{V_n^{**}(f,s)}(t) = \prod_{j=1}^{k} \phi_{F_j^{(s)}}(tf(j)j^{-s}\sigma_n(f,s)^{-1}). \tag{4.1}$$

We use the expansion common characteristic function $\phi(u) = \phi_{F_j^{(s)}}(u) = 1 - u^2/2 + u^2\varepsilon(v)$, where $\varepsilon(v) \to 0$, uniformly in $|u| \leq v \to 0$. Remind that for $1 \leq j \leq k$, $0 \leq f(j)j^{-s}\sigma_n(f,s)^{-1} \leq B_n(f,s) \to 0$ and then

$$\phi_{V_n^{**}(f,s)}(t) = \prod_{j=1}^{k}(1 - t^2f(j)^2j^{-2s}\sigma_n^{-2}(f,s)/2 + t^2f(j)^2j^{-2s}\sigma_n^2(f,s)\varepsilon(B_n))$$

$$= \exp(\sum_{j=1}^{k} \log(1 - t^2f(j)^2j^{-2s}\sigma_n^2(f,s)/2 + t^2f(j)^2j^{-2s}\sigma_n^2(f,s)\varepsilon(B_n))).$$

By using a first order expansion the logarithmic function in the neighborhood of unity, we have

$$\phi_{V_n^*(f)}(t) = \exp(\sum_{j=1}^{k} -t^2 f(j)^2 j^{-2s}\sigma_n(f)^{-2}/2 + t^2 f(j)^2 j^{-2s}\sigma_n^{-2}(f)\varepsilon(B_n)),$$

where the function $\varepsilon(B_n)$ may change from one line to an other, but always tends to zero. Hence

$$\phi_{V_n^{**}(f,s)}(t) = \exp(\sum_{j=1}^{k} -t^2/2 + t^2\varepsilon(B_n)) \rightarrow \exp(-t^2/2).$$

and

$$V_n^{**}(f) \rightarrow \mathcal{N}(0,1).$$

and then

$$V_n^{*}(f) \rightarrow \mathcal{N}(0, \left\{\Gamma(2s+1) - \Gamma(s+1)^2\right\}).$$

□

Lemma 2. *If any of* $(H1a)$, $(H1b)$, $(H2a)$ *and* $(H2b)$ *holds with an abitrary* $\lambda > 1$, *then their analogues where* $b_n(\lambda)$ *is replaced with* b_n *and* $p_n(\lambda)$ *is replaced with* p_n *also hold.*

Proof. We have to prove this only for one case. The others are similarly done. We begin to recall the Balkema result, that is $\sqrt{n}((n/k)U_{k,n} - 1) \rightarrow_d N(0,1)$ which entails that $\sqrt{n}((n/k)U_{k+1,n} - 1) \rightarrow_d N(0,1)$ and next $(n/k)U_{k+1,n} \rightarrow_{\mathbb{P}} 1$. Then for any $\varepsilon > 0$, for any $\lambda > 1$, we have for large values of n, say $n \geq n_1$,

$$\mathbb{P}(U_{k+1,n} > \lambda n/k) \leq \varepsilon/3.$$

Next by the definition of $V_n = O_p(p_n)$ and $W_n = O_p(b_n)$, there exists C_0 such for large values of n, say $n \geq n_2$,

$$\mathbb{P}(|V_n| > C_0 p_n) \leq \varepsilon/3, P(|W_n| > C_0 b_n) \leq \varepsilon/3.$$

Then for $n \geq \max(n_1, n_2)$. Recall that $b_n(\lambda) = \{|b(t)|, t \leq \lambda k/n\}$, $p_n(\lambda) = \{|p(t)|, t \leq \lambda k/n\}$, $b_n = \{|b(t)|, t \leq U_{k+1,n}\}$ and $p_n = \{|p(t)|, t \leq U_{k+1,n}\}$. We have

$$\mathbb{P}(|V_n| \leq C_0 p_n, |W_n| \leq C_0 b_n, b_n \leq b_n(\lambda), p_n \leq p_n(\lambda)) \geq 1 - \varepsilon.$$

And next, since $s \geq 1$,

$$\mathbb{P}\left|s\left\{\frac{\gamma}{j}E_j + \theta_{n,j}\left(O_{\mathbb{P}}(p_n) + \frac{E_j}{j}O_{\mathbb{P}}(b_n)\right)\right\}^{s-1}\left(O_{\mathbb{P}}(p_n) + \frac{E_j}{j}O_{\mathbb{P}}(b_n)\right)\right|$$

$$\leq s\left|\frac{\gamma}{j}E_j + \left(C_0 p_n(\lambda) + \frac{E_j}{j}C_0 b_n(\lambda)\right)\right|^{s-1}\left(C_0 p_n(\lambda) + \frac{E_j}{j}C_0 b_n(\lambda)\right)\Bigg) \geq 1 - \varepsilon$$

Suppose that for an arbitrary $\lambda > 1$,

$$c_n(\lambda) = s\left|\frac{\gamma}{j}E_j + \left(C_0 b_n(\lambda) + \frac{E_j}{j}C_0 b_n(\lambda)\right)\right|^{s-1}\left(C_0 p_n(\lambda) + \frac{E_j}{j}C_0 b_n(\lambda)\right) \rightarrow_{\mathbb{P}} 0$$

and put

$$c_n = \left|s\left\{\frac{\gamma}{j}E_j + \theta_{n,j}\left(O_{\mathbb{P}}(p_n) + \frac{E_j}{j}O_{\mathbb{P}}(b_n)\right)\right\}^{s-1}\left(O_{\mathbb{P}}(p_n) + \frac{E_j}{j}O_{\mathbb{P}}(b_n)\right)\right|$$

, we have for any $\eta > 0$, and a fixed $\lambda > 1$ and for large values of n. This gives

$$\mathbb{P}(c_n > \eta) = P((c_n > \eta) \cap (c_n < c_n(\lambda))) + \mathbb{P}((c_n > \eta) \cap (c_n \geq c_n(\lambda)))$$

$$\leq (c_n(\lambda) > \eta) + \varepsilon.$$

Now letting $n \rightarrow +\infty$, we have

$$\limsup_{n\rightarrow\infty} \mathbb{P}(c_n > \eta) \leq \varepsilon.$$

By letting $\varepsilon \downarrow 0$, one achieves the proof, that is $c_n \rightarrow_{\mathbb{P}} 0$.

4.2 Integral Computations

Let $b \geq 1$, we get by comparing the area under the curve $x \longmapsto x^{-b}$ from j to $k-1$ and those of the rectangles based on the intervals $[h, h+1]$, $h = j, .., k-2$, we get

$$\sum_{h=j+1}^{k-1} h^{-b} \leq \int_j^{k-1} x^{-b} dx \leq \sum_{h=j}^{k-2} h^{-b},$$

that is

$$\int_j^{k-1} x^{-b} dx + (k-1)^{-b} \leq \sum_{h=j}^{k-1} h^{-b} \leq \int_j^{k-1} x^{-b} dx + j^{-b}. \tag{4.2}$$

For $b = 1$, we get

$$\frac{1}{k-1} \leq (\sum_{h=j}^{k-1} \frac{1}{h}) - \log((k-1)/j) \leq \frac{1}{j}. \tag{4.3}$$

For $b = 2$,, we have

$$\frac{1}{j} - \frac{1}{k-1} + \frac{1}{(k-1)^2} \leq \sum_{h=j}^{k-1} h^{-2} \leq \frac{1}{j} - \frac{1}{k-1} + \frac{1}{j^2},$$

that is

$$\frac{1}{(k-1)^2} \leq \sum_{h=j}^{k-1} h^{-2} - \frac{1}{j}(1 - \frac{j}{k-1}) \leq \frac{1}{j^2}.$$

As well, we have for $b > 0$,

$$\sum_{h=j}^{k-2} h^b \leq \int_j^{k-1} x^b dx \leq \sum_{h=j+1}^{k-1} h^b$$

and then

$$\frac{1}{b+1}((k-1)^{b+1} - j^{b+1}) + j^b \leq \sum_{h=j}^{k-1} h^b \leq \frac{1}{b+1}((k-1)^{b+1} - j^{b+1}) + (k-1)^b. \tag{4.4}$$

Hence for j fixed and $k \to \infty$, we get $\sum_{h=j}^{k-1} h^b \sim (k-1)^{b+1}/(b+1)$.

4.3 Computation of J(a,b,c,s)

Recall

$$J(a, b, c, s) = \int_0^{\infty} (a+cx)^s (a+bx) e^{-x} dx \text{ and } I(a, b, s) = \int_0^{\infty} (a+bx)^s e^{-x} dx.$$

4.3.1 Computation of I(a,b,s).

We have by integration by parts

$$I(a, b, s) = \int_0^{\infty} (a+bx)^s e^{-x} dx = [-e^{-x}(a+bx)^s]_0^{\infty} + bs \int_0^{\infty} (a+bx)^{s-1} e^{-x} dx,$$

that is, for any $s \geq 1$,

$$I(a, b, s) = a^s + sbI(a, b, s-1).$$

By induction for $s \geq 1$,this leads to

$$I(a, b, s) = s! \sum_{h=0}^{s} b^h a^{s-h}/(s-h)!.$$

4.3.2 Computation of J(a,b,c,s)

We have integration by parts

$$J(a,b,c,s) = a^{s+1} + bI(a,c,s) + csJ(a,b,c,s-1).$$

We get by induction for $\ell \geq 1$,

$$\begin{aligned} J(a,b,c,s) &= s!a\sum_{h=0}^{\ell} c^h a^{s-h}/(s-h)! + s!b\sum_{h=0}^{\ell} c^h I(a,c,s-h)/(s-h)! \\ &+ s!c^{\ell+1} J(a,b,c,s-\ell-1)/(s-\ell-1)! \end{aligned}$$

For $\ell + 1 = s$, we arrive at

$$\begin{aligned} J(a,b,c,s) &= s!a\sum_{h=0}^{s-1} c^h a^{s-h}/(s-h)! + s!b\sum_{h=0}^{s-1} c^h I(a,c,s-h)/(s-h)! \\ &+ s!c^s J(a,b,c,0). \end{aligned}$$

Since $J(a,b,c,0) = I(a,b,1)$, we finally get

$$\begin{aligned} J(a,b,c,s) &= s!a\sum_{h=0}^{s-1} c^h a^{s-h}/(s-h)! + s!b\sum_{h=0}^{s-1} c^h I(a,c,s-h)/(s-h)! \\ &+ s!c^s I(a,b,1) \end{aligned}$$

4.4 Minimization of the Asymptotic Variance

For $s \geq 1$ fixed, We have to maximize

$$V_n(\tau,s) = \frac{d_n(\tau,s)}{\sigma_n(\tau,s)} \times \frac{s\gamma^{-1}}{s!\sqrt{\Gamma(2s+1)-\Gamma(s+1)^2}}.$$

with respect of $\tau > s - 1/2$ for s, where $d_n(\tau,s) = \sum_{j=1}^{k} j^{\tau-s}$. We denote $q(s) = \frac{s\gamma^{-1}}{s!\sqrt{\Gamma(2s+1)-\Gamma(s+1)^2}}$. Let us find critical points. It is easy to see that

$$\frac{\partial V_n(\tau,s)}{\partial \tau} = \frac{d_n'(\tau,s)\sigma_n(\tau,s) - d_n(\tau,s)\sigma_n'(\tau,s)}{(\sigma_n(\tau,s))^2} \times q(s).$$

A zero-point of $\frac{\partial V_n(\tau,s)}{\partial \tau}$ obviously, is a solution of the ordinary differential equation.

$$\frac{d_n'(\tau,s)}{d_n(\tau,s)} = \frac{\sigma_n'(\tau,s)}{\sigma_n(\tau,s)}.$$

Its general solution is given by

$$\log d_n(\tau,s) = \log \sigma_n(\tau,s) + C(s). \tag{4.5}$$

By taking the particular value of $\tau = s$, we find that $C(s) = (1/2)\log k$,and (4.5) becomes

$$\sum_{j=1}^{k} j^{\tau-s} = \left(k\sum_{j=1}^{k} j^{2(\tau-s)}\right)^{1/2}.$$

This is the equality form of Cauchy-Schwarz's inequality with respect to the usual scalar product in $\mathbb{R}^k$. Then there exists a constant $\lambda(s)$ such that $j^{\tau-s} = \lambda(s)$ for $1 \leq j \leq k$. The only solution is $\tau = s$. Now to show that τ is the global maximum point, it suffices to notice that

$$\frac{1}{s!q(s)}\frac{\partial^2 V_n(\tau,s)}{\partial^2 \tau} = \frac{1}{k\sqrt{k}}\left(\left(\sum_{j=1}^{k}\log j\right)^2 - k\sum_{j=1}^{k}(\log j)^2\right) < 0.$$

$$\frac{1}{s!q(s)\ \sqrt{(}k)} \frac{\partial^2 V_n(\tau, s)}{\partial^2 \tau} = -\left(\left(k^{-1} \sum_{j=1}^{k} (\log j)^2_(\, k^{-1}) \sum_{j=1}^{k} \log j\right)^2\right) < 0.$$

since the left member is the opposite of a the empirical variance of log j, $1 \leq j \leq k$. We conclude that the point $\tau = s$ is the unique local miximum point. Then the global maximum is reached at $\tau = s$.

References

Beirlant, J., Goegebeur, Y., & Teugels, J.(2004). *Statistics of Extremes Theory and Applications*. Wiley. (MR2108013)

de Haan, L. (1970). *On regular variation and its application to the weak convergence of sample extremes*. Mathematical Centre Tracts, 32, Amsterdam. (MR0286156)

de Haan, L., & Feireira A. (2006). *Extreme value theory: An introduction.* Springer. (MR2234156)

Dekkers, A. L. M., Einmahl, J. H. J., & Haan, L. D. (1989). A moment estimator for the index of an extreme value distribution. *Ann. Statist., 17*(4), 1833-1855.(MR1026315)

Dème, E., LO, G. S., & Diop, A. (2012). On the generalized Hill process for small parameters and applications. *Journal of Statistical Theory and Applications*, *11*(4), 397-418. http://dx.doi.org/10.2991/jsta.2013.12.1.3.(MR3191797)

Diop, A., & Lo, G. S. (2006). Generalized Hill's Estimator. *Far East J. Theor. Statist., 20*(2), 129-149. (MR2294728)

Diop, A., & Lo, G. S. (2009). Ratio of Generalized Hill's Estimator and its asymptotic normality theory. Math. Method. *Statist., 18*(2), 117-133. (MR2537361)

Embrechts, P., Kűppelberg C., & Mikosh T. (1997). *Modelling extremal events for insurance and Finance*. Springer Verlag.

Fall, A. M., LO, G. S., Ndiaye, C. H., & Adekpedjou, A. (2014). Supermartingale argument for characterizing the Functional Hill process weak law for small parameters. Submitted. available at : http://arxiv.org/pdf/1306.5462

Galambos, J. (1985). *The Asymptotic theory of Extreme Order Statistics*. Wiley, Nex-York. (MR0489334)

Groeneboom, Lopuhaä, H. P., & Wolf, P. P.(2003). Kernel-Type Estimator for the extreme Values index. *Ann. Statist., 31*(6), 1956-1995.(MR2036396)

Lo, G. S.(1991). Caractérisation empirique des extrêmes et questions liées. Thèse d'Etat. Université de Dakar.

Lo, G. S, & Fall, A. M. (2011). Another look at Second order condition in Extreme Value Theory. *Afrik. Statist., 6*, 346-370.(zbl:1258.62058)

Lo, G. S. (2012). On a discrete Hill's statistical process based on sum-product statistics and its finite-dimensional asymptotic theory. available at : http://arxiv.org/abs/1203.0685

Resnick, S. I. (1987). *Extreme Values, Regular Variation and Point Processes*. Springer-Verbag, New-York. (M-R0900810)

Segers, J. (2002). Generalized Pickands Estimators for the Extreme Value Index J. *Statist. Plann. Inference, 128*(2), 381-396. (MR2102765)

Shorack, G. R., & Wellner J. A. (1986). *Empirical Processes with Applications to Statistics*. wiley-Interscience, New-York. (MR0838963)

Copyrights