Simple tail index estimation for dependent and heterogeneous data with missing values

Ivana Ilić[a] and Vladica M. Veličković[b,c]

[a]*University of Nis*
[b]*Synergus AB*
[c]*UMIT*

Abstract. Financial returns are known to be nonnormal and tend to have fat-tailed distribution. Also, the dependence of large values in a stochastic process is an important topic in risk, insurance and finance. In the presence of missing values, we deal with the asymptotic properties of a simple "median" estimator of the tail index based on random variables with the heavy-tailed distribution function and certain dependence among the extremes. Weak consistency and asymptotic normality of the proposed estimator are established. The estimator is a special case of a well-known estimator defined in Bacro and Brito [*Statistics & Decisions* **3** (1993) 133–143]. The advantage of the estimator is its robustness against deviations and compared to Hill's, it is less affected by the fluctuations related to the maximum of the sample or by the presence of outliers. Several examples are analyzed in order to support the proofs.

1 Introduction and motivation

The last three decades have witnessed rapid expansion of the study of heavy-tailedness phenomena in economics and finance. Following the initial work by Mandelbrot (1960, 1963) numerous studies have documented that time series encountered in many fields in economics and finance are typically thick-tailed and can be well approximated using distributions with the tails exhibiting the property:

$$1 - F(x) = x^{-\alpha}L(x), \qquad x > 0, \tag{1.1}$$

where $\alpha > 0$ denotes the index of regular variation and L is slowly varying at infinity, that is:

$$\lim_{t \to \infty} \frac{L(tx)}{L(t)} \to 1 \qquad \text{for } x > 0. \tag{1.2}$$

One of the most important parameters for measuring the thickness of the distribution tail, attracting the great attention among the statisticians over the past

20 years is the parameter α, known as Pareto index or the index of regular variation. Since it is a well-known fact that financial returns tend to have empirical distributions that exhibit fatter tails than the normal distribution (see Jansen and de Vries (1991), Koedijk and Kool (1994) and Loretan and Phillips (1994)), the approximate evaluation of this parameter is of great importance. Therefore, the tail thickness has been rigorously studied and different estimators of α have been suggested and analyzed in the case of complete data.

Suppose $\{X_t\} = \{X_t : 1 \le t \le n\}$ is a sample of rv's equally distributed with the heavy-tailed distribution function F and $X_{(1)} > X_{(2)} > \cdots > X_{(n)}$ are the order statistics. Probably the most popular estimator of $\gamma = 1/\alpha$, based on the extreme portion of the sample and obtained as the arithmetic mean of certain logarithmic differences is by all means the Hill's estimator $H_{\alpha,n}$, defined as follows (Hill (1975)):

$$H_{k,n} = \frac{1}{k} \sum_{i=1}^{k} \ln X_{(i)} - \ln X_{(k+1)}, \qquad (1.3)$$

where $k = k_n$ is a sequence of positive integers satisfying

$$1 \le k_n < n, \qquad \lim_{n \to \infty} k_n = \infty \quad \text{and} \quad \lim_{n \to \infty} k_n/n = 0. \qquad (1.4)$$

Hill's estimator has been well studied when $\{X_t\}$ are i.i.d. (independent identically distributed) (see Davis and Resnick (1984) and Haeusler and Teugels (1985)). Other estimators of α in i.i.d. settings were also studied in the case of complete samples: Bacro and Brito (1995) and Csörgó and Viharos (1998). See also Hsing (1991) and Hill (2010) and the extensive list of citations therein for the estimation of the tail index using a full sample of possibly dependent data. The results of Hsing (1991) were extended on the incomplete sample of possibly dependent variables, see Mladenovic and Piterbarg (2008) and Ilic and Mladenovic (2008). Also see Ilic (2012) for the proof of the asymptotic normality of the Hill's estimator under the assumption of the incompleteness of the sample consisting of heterogeneous and dependent data.

It is well known that $H_{k,n}$ is a strong consistent estimator of γ, asymptotically normally distributed (see, for instance, Haeusler and Teugels (1985)).

More recently, Bacro and Brito (1993) proposed and well studied a simple estimator defined as:

$$B_p = -\frac{1}{\ln p} \ln \frac{X_{(\lceil pk_n \rceil)}}{X_{(\lceil k_n \rceil)}}, \qquad (1.5)$$

where $0 < p < 1$ is a fixed constant and $\lceil x \rceil$ is a ceiling function and denotes the smallest integer greater than or equal to x and k_n is a nondecreasing sequence of positive real numbers satisfying (1.4). Also, this estimator was proposed in Fraga Alves (1992, 1995) and the same paper has also been referred in Bacro and Brito (1998). For the results on strong asymptotic behavior of B_p see Bacro and Brito

(1993). General conditions which ensure the asymptotic normality of this estimator are given in Bacro and Brito (1995) and in the same paper, the weak limiting behavior of the B_p has been analyzed.

In this paper, we investigate an asymptotic behavior of the median estimator $B_{2,n}$:

$$B_{2,n} = \frac{1}{\ln 2} \ln \frac{X_{(k_n)}}{X_{(2k_n)}}, \tag{1.6}$$

where k_n is a sequence of positive integers satisfying (1.4).

Note that as well as Hill's, this estimator presents an important statistical measure, i.e. Hill's estimator may be pointed as a mathematical expectation of certain sequence of logarithmic differences and $B_{2,n}$ as their median. On the other hand, it is well known that maximum likelihood estimators are usually sensitive to few particular observations, known as outliers. These observations do not fit the model imposed by the majority of the data points. However, in extreme value statistics, modeling these rare events with low frequency but high impact receives great attention. So, the idea is to investigate the outlier-resistant estimator under certain additional conditions attractive for the researchers in modeling financial, insurance and economic extreme events. The main motivation for studying $B_{2,n}$ is its "robustness" against deviations of the slowly varying function L from the constant (see Csörgo and Viharos (1998)). It doesn't include the top values of the sample, so one can expect that $B_{2,n}$ is less affected by the fluctuations related to the maximum of the sample or by the presence of outliers. The proposed procedure provides estimation of α which is still reliable and reasonably efficient under small deviations from the assumed parametric model.

Robust methods for extreme values have been discussed in recent literature. For example, Vandewalle et al. (2007) consider a robust estimator of the tail index by combining a refinement of the Pareto approximation for the conditional distribution of relative excesses over a large threshold with an integrated squared error approach on partial density component estimation. Also, Hubert et al. (2013) introduce new robust GLM estimator for alpha, which determines high quantiles, based on a robust M-estimator proposed by Cantoni and Ronchetti (2001). On the other hand, the formulation of an estimator robust to outliers, may seem as a contradiction to estimation of models with the main interest laying in the extreme values-values that are associated exactly to outliers. Robust methods can help us identify deviating structure, influential observations and guarantee good statistical properties over a whole set of underlying distributions. Therefore, they can considerably enhance the data analysis. In this sense there is no "obvious" contradiction between robustness and EVT. Overall, we find that robust statistical methods can improve the data analysis process of the skilled analyst and provide useful additional information.

Under the assumptions of extremal dependence structure and the incompleteness of the sample, we track the weak limiting behavior and provide general conditions which ensure the asymptotic normality of the median estimator $B_{2,n}$. In

next sections, we define more accurately the model of the incompleteness and the notion of extremal dependence.

2 A sample with missing observations

If the condition (1.1) is satisfied one can easily prove that (see Leadbetter et al. (1983), Theorems 1.5.1 and 1.7.3):

$$\lim_{x \to \infty} \frac{1 - F(x)}{1 - F(x - 0)} = 1, \tag{2.1}$$

$$1 - F\left(F^{-1}\left(1 - \frac{1}{t}\right)\right) \sim \frac{1}{t} \qquad \text{as } t \to \infty, \tag{2.2}$$

where $F^{-1}(y) := \inf\{x : F(x) \geq y\}$ is the left continuous inverse of the function F. Since any inference of α should be made with the tail portion of the empirical distribution of the sample, without loss of generality, we may assume that F has support on $(0, \infty)$.

Now, assume that only observations at certain points are available. Denote the observed random variables among $\{X_1, \ldots, X_n\}$ with $\widetilde{X}_1, \ldots, \widetilde{X}_{S_n}$. Here the random variable S_n presents the number of observed rv's among first n terms of the sequence $\{X_t\}_{t \in N}$. We shall assume that observed random variables are determined by a general point process and only conditions on S_n will be imposed. This model was considered in Mladenovic and Piterbarg (2008), where consistency of Hill's estimator was proved. Finally, through out this paper we used specific model which defines a structure of the incompleteness of data, a common problem in modern statistical analysis.

Assumption A. The sequence $X_1, X_2, \ldots$ does not depend on S_n and

$$\frac{S_n}{n} \xrightarrow{p} \nu > 0 \qquad \text{as } n \to +\infty.$$

Suppose β_n is a sequence of real numbers such that

$$\lim_{n \to \infty} \beta_n = \infty \quad \text{and} \quad \lim_{n \to \infty} \frac{\beta_n}{n} = 0. \tag{2.3}$$

Let

$$M_n = \left[\frac{S_n}{\beta_n}\right] \quad \text{and} \quad P\{M_n = k\} = p_k, \tag{2.4}$$

where the floor function $[\cdot]$ denotes the largest previous integer.

Remark 2.1. Let ξ_k be the indicator of the event that random variable X_k is observed. Assume that $(\xi_n)_{n \in N}$ is an i.i.d. sequence, independent of $(X_n)_{n \in N}$ satisfying that $P\{\xi_k = 1\} = p$ and $P\{\xi_k = 0\} = 1 - p$, where $0 < p < 1$. Then $S_n \in B(n, p)$, i.e. $S_n = \xi_1 + \xi_2 + \cdots + \xi_n$ is a binomial random variable with parameters n and p. In this case, the Assumption A is satisfied.

We are interested in the estimation of α, using some portion of the sample. Let $\widetilde{X}_{(1)} \geq \widetilde{X}_{(2)} \geq \cdots \geq \widetilde{X}_{(S_n)}$ be the order statistics from S_n observed variables.

Median estimator is now defined as:

$$\widetilde{B}_{2,S_n} = \frac{1}{\ln 2}(\ln \widetilde{X}_{(M_n)} - \ln \widetilde{X}_{(2M_n)}).$$

3 Extremal dependence

Using the above model, we prove the consistency and asymptotic normality of $B_{2,n}$ estimator by using processes with extremes that are NED (near epoch dependent) on some arbitrary mixing functional. This property has substantial practical advantages because it only requires computation of conditional expectation and it is easy to verify. It indicates a possible dependence among the extremes. This assumption is particularly convenient since it covers wide variety of processes and it is suitable for applications. We do not require stationarity in general and our results cover, for example, processes satisfying Hsing's mixing condition, nonlinear distributed lag processes, strong mixing GARCH, explosive GARCH, and much more. See Hill (2010) who expedites the theory by analyzing the processes whose extremal support is NED and proves a broad class of processes have this property.

In order to demonstrate the asymptotic properties of $\widetilde{B}_{2,S_n}$ we appeal to the concept in Hill (2010, Theorem 1). He imposes new tail dependence properties on $\{I(X_t > b_{k_n}e^u)\}$, ($b_{k_n} := F^{-1}(1 - \frac{k_n}{n})$ and I is a notation for an indicator function) that cover and generalize Hsing's mixing condition (see Hsing, 1991).

Let $\Im_t := \sigma(\varepsilon_\tau : \tau \leq t)$ be a σ-field induced by some α-mixing base ε_t. Let q_n be an arbitrary sequence of integers satisfying $1 \leq q_n < n$, and $q_n \to \infty$. For example, we may assume $\varepsilon_t = I(X_t > b_{k_n}e^u)$ and impose a mixing condition on ε_t as in Hsing (1991).

L_p-*Extremal-Near Epoch Dependence (L_p-E-NED) property.* $\{X_t\}$ is L_p-E-NED on $\{\Im_t\}$, $p > 0$, with size $\lambda > 0$ if

$$\left\| I(X_t > b_{k_n}e^u) - P(X_t > b_{k_n}e^u | \Im_{t-q_n}^{t+q_n}) \right\|_p \leq f_{nt}(u) \times \psi_{q_n},$$

where $\psi_{q_n} = o(q_n^{-\lambda})$, $f_{nt} : R_+ \to R_+$ is Lebesgue measurable, $\sup_{1 \leq t \leq n} \sup_{u \geq 0} f_{nt}(u) = O((k_n/n)^{1/p})$ and $\Im_s^{s+t} := \sigma(\widetilde{\varepsilon}_\tau : s \leq \tau \leq s+t)$, where $\widetilde{\varepsilon}_s$ is some functional of $\widetilde{X}_s$. See Hill (2010, p. 1402).

Assumption B. $\{X_t\}$ is L_2-E-NED on $\{\Im_t\}$, with coefficients ψ_{q_n} of size $1/2$ and constants $f_{nt}(u)$ where $f_{nt} : R_+ \to R_+$ is Lebesgue measurable, $\sup_{1 \leq t \leq n} \int_0^\infty f_{nt}(u)\,du = O((k_n/n)^{1/2})$.

Remark 3.1. Note that, if the sequence $\{X_t\}$ is α-mixing, then so is $\{\widetilde{X}_t\}$ because

$$\sup_{s \in Z} \sup_{A \in \mathfrak{F}^s_{-\infty}, B \in \mathfrak{F}^{+\infty}_{s+d}} \left| P(A \cap B) - P(A)(B) \right|$$

$$\leq \sup_{t \in Z} \sup_{A \in \mathfrak{I}^t_{-\infty}, B \in \mathfrak{I}^{+\infty}_{t+d}} \left| P(A \cap B) - P(A)(B) \right|.$$

4 Main results

Lemma 4.1. *Let $k_n = k$ be fixed. Suppose that Assumptions* A, B *and* (1.1) *hold. Then $\widetilde{B}_{2,S_n} \xrightarrow{P} \gamma$, as $n \to \infty$.*

Lemma 4.2. *Assume that Assumptions* A *and* B *hold. Then: $E(\widetilde{B}_{2,S_n}) \to \gamma$ and $\mathrm{Var}(\widetilde{B}_{2,S_n}) \sim \gamma^2 \frac{1}{\ln^2 2} E(\frac{1}{M_n})$, as $n \to \infty$.*

Assumption C. There exists positive measurable function g on $(0, \infty)$ such that for any $\lambda > 0$,

$$L(\lambda x)/L(x) - 1 = O(g(x)),$$

as $x \to \infty$. Also, there exist $D > 0$, $z_0 < \infty$ and $\tau \leq 0$, such that $g(\lambda z)/g(z) \leq D\lambda^\tau$, for some $\lambda \geq 1$ and $z \geq z_0$. We require k_n and g to satisfy $k_n^{1/2} g(F^{-1}(1 - \frac{k_n}{n})) \to 0$, as $n \to \infty$.

Theorem 4.3. *Let Assumptions* A, B *and* C *hold. Then*

$$\ln 2 E(M_n^{1/2})(\widetilde{B}_{2,S_n} - \gamma) \xrightarrow{d} N(0, \gamma^2),$$

as $n \to \infty$.

Example 4.4. Let us consider a model with missing observations such that the number of observed variables in the sample is described with the random variable S_n with the probability distribution given by:

$$P\left\{S_n = \frac{n}{2}\right\} = \frac{1}{n}, \qquad P\left\{S_n = \frac{n}{2} - 1\right\} = 1 - \frac{1}{n},$$

$$E(S_n) = \frac{n}{2} \cdot \frac{1}{n} + \left(\frac{n}{2} - 1\right) \cdot \left(1 - \frac{1}{n}\right) = \frac{n^2 - 2n + 2}{2n}. \tag{4.1}$$

We conclude straightforward that the Assumption A holds. Let us consider the finite moving average processes

$$X_j = \sum_{i=0}^m c_i Z_{i+j}, \qquad j = 1, 2, \dots,$$

where $c_i \in R^+$, $i = 0, \ldots, m$ and Z_k are independent identically distributed random variables with $P\{Z_k > z\} = z^{-\alpha}L(x)$ and $L(x)$ is slowly varying function at infinity defined as:

$$L(x) = B\big(1 + \mathrm{O}(\log x)^{-\beta}\big), \tag{4.2}$$

as $x \to \infty$ and $B, \beta > 0$. Since this kind of processes is covered with the L_2-E-NED property (see for example Hill (2010) for more details) we have that the Assumption B holds. Finally, $L(x)$ satisfies the Assumption C with $g(x) = (\log x)^{-\beta}$ and $k_n = o((\log n)^{2\beta})$. For more details, please see Haeusler and Teugels (1985). So the conclusion of Theorem 4.3 holds for this type of processes, where certain portion of the data is missing.

Example 4.5. Volatility of financial time series plays a key role in both risk management and option pricing. The Generalized Autoregressive Conditional Heteroscedasticity (GARCH) model introduced by Bollerslev (1986) is designed to capture the volatility clustering feature by modeling the dynamic of volatility. When considering the volatility dynamic with only one lagged period, the GARCH(1, 1) model has become a valuable tool in practice due to its simplicity and intuitive interpretation. Under very general conditions on the noise sequence (Z_t) the GARCH(1, 1) process has Pareto-like marginal distribution with the specific slowly varying function defined as:

$$L(x) = C_1\big(1 + C_2 x^{-\beta} + \mathrm{o}\big(x^{-\beta}\big)\big), \tag{4.3}$$

as $x \to \infty$, $C_2 \neq 0$, $C_1, \beta > 0$ and it represents very attractive tool to model the heavier-then-normal tails of the financial data. It is defined, specifying σ_t, as follows:

$$\sigma_t^2 = \alpha_0 + \beta_1 \sigma_{t-1}^2 + \alpha_1 X_{t-1}^2 = \alpha_0 + \sigma_{t-1}^2\big(\beta_1 + \alpha_1 Z_{t-1}^2\big),$$

$t \in Z$ and the parameters α_0, α_1 and β_1 are nonnegative. The GARCH(1, 1) process is L_2-E-NED if certain conditions on this parameters are satisfied (see Hill (2005)). Finally, for $g(x) = x^{-\beta}$ and $k_n = o(n^{\frac{2\beta}{2\beta+\alpha}})$ the Assumption C holds (see Bacro and Brito (1995)).

We may conclude that in this particular case:

$$\ln 2E\big(M_n^{1/2}\big)(\widetilde{B}_{2,S_n} - \gamma) \xrightarrow{d} N\big(0, \gamma^2\big).$$

5 Proofs

In order to prove Lemma 4.1 and Lemma 4.2 and to study further asymptotic properties of $\widetilde{B}_{2,S_n}$ it is more convenient to formulate the model (1.1) in the terms of quantile function. Let us denote by $F^{-1}(t)$ the left continuous inverse of F, i.e. $F^{-1}(t) = \inf\{x : F(x) \geq t\}$. According to De Haan (1970), the model (1.1) is

equivalent to $F^{-1}(1 - \frac{1}{t}) = t^{\frac{1}{\alpha}}\widetilde{L}(t)$, where $\widetilde{L}$ satisfies (1.2). Moreover, we may choose $X_i = F^{-1}(\frac{\Gamma_{n-i+1}}{\Gamma_{n+1}})$, where $\Gamma_i = E_1 + E_2 + \cdots + E_i$ and $\{E_t\}$, $1 \le t \le n$ is the sequence of identically independent exponential random variables with mean one. The above is obtained by Renyi's representation for the order statistics from the exponential sample.

Proof of Lemma 4.1. According to the above, we may write:

$$\frac{\widetilde{X}_{(M_n)}}{\widetilde{X}_{(2M_n)}} = \frac{F^{-1}(\frac{\Gamma_{S_n-M_n+1}}{\Gamma_{S_n+1}})}{F^{-1}(\frac{\Gamma_{S_n-2M_n+1}}{\Gamma_{S_n+1}})}$$

$$= \left(\frac{\Gamma_{S_n+1} - \Gamma_{S_n-2M_n+1}}{\Gamma_{S_n+1} - \Gamma_{S_n-M_n+1}}\right)^{\gamma} \frac{L(\frac{\Gamma_{S_n+1}}{\Gamma_{S_n+1}-\Gamma_{S_n-M_n+1}})}{L(\frac{\Gamma_{S_n+1}}{\Gamma_{S_n+1}-\Gamma_{S_n-2M_n+1}})}.$$

Since the Assumption A holds, we may apply the law of large numbers. Because of the fact that $k_n = k$ is fixed, we obtain:

$$A = \frac{\Gamma_{S_n+1} - \Gamma_{S_n-2M_n+1}}{\Gamma_{S_n+1} - \Gamma_{S_n-M_n+1}} \to 2 \qquad \text{a.s.,} \tag{5.1}$$

$$\frac{\Gamma_{S_n+1}}{\Gamma_{S_n+1} - \Gamma_{S_n-M_n+1}} \to \infty \qquad \text{a.s.} \tag{5.2}$$

and

$$\frac{\Gamma_{S_n+1}}{\Gamma_{S_n+1} - \Gamma_{S_n-2M_n+1}} \to \infty \qquad \text{a.s.,} \tag{5.3}$$

as $n \to \infty$.

Now, from (5.1), (5.2) and (5.3) we have that:

$$\frac{L(\frac{\Gamma_{S_n+1}}{\Gamma_{S_n+1}-\Gamma_{S_n-M_n+1}})}{L(\frac{\Gamma_{S_n+1}}{\Gamma_{S_n+1}-\Gamma_{S_n-2M_n+1}})} = \frac{L(A\frac{\Gamma_{S_n+1}}{\Gamma_{S_n+1}-\Gamma_{S_n-2M_n+1}})}{L(\frac{\Gamma_{S_n+1}}{\Gamma_{S_n+1}-\Gamma_{S_n-2M_n+1}})} \to 1 \qquad \text{a.s.} \tag{5.4}$$

We conclude from the above that $\widetilde{B}_{2,S_n} \overset{p}{\to} \gamma$. □

Proof of Lemma 4.2. Since (1.1) holds, we may write, according to the above and to David (1981):

$$\ln\frac{\widetilde{X}_{(M_n)}}{\widetilde{X}_{(2M_n)}} = E^*_{M_n} + E^*_{M_n+1} + \cdots + E^*_{2M_n-1},$$

where $E^*_i = \widetilde{E}_i - \widetilde{E}_{i-1}$ may be considered as independent exponentially distributed rv's with the corresponding mathematical expectations $\frac{1}{\alpha}E(\frac{1}{i})$, $i = M_n, \ldots, 2M_n - 1$. Note that we assumed that Assumption B holds, meaning that,

without loss of generality, $\widetilde{X}_{(M_n)}$ and $\widetilde{X}_{(2M_n)}$ may be considered independent. Now we have:

$$
E\left(\ln \frac{\widetilde{X}_{(M_n)}}{\widetilde{X}_{(2M_n)}}\right) = E\left(\sum_{i=M_n}^{2M_n-1} E_i^*\right)
$$

$$
= \sum_{k=1}^{n} E\left(\sum_{i=M_n}^{2M_n-1} E_i^* \Big| M_n = k\right) P\{M_n = k\}
$$

$$
= \sum_{k=1}^{n} E\left(\sum_{i=k}^{2k-1} E_i^*\right) P\{M_n = k\}
$$

$$
= \sum_{k=1}^{n} \sum_{i=k}^{2k-1} E(E_i^*) p_k
$$

$$
= \gamma \sum_{k=1}^{n} p_k \left(\frac{1}{k} + \frac{1}{k+1} + \cdots + \frac{1}{2k-1}\right)
$$

$$
\sim \gamma \sum_{k=1}^{n} p_k \ln 2
$$

$$
= \gamma \ln 2 \sum_{k=1}^{n} p_k = \gamma \ln 2.
$$

Further, without loss of generality and under the assumption of the independency of E_i^*, $M_n \le i \le 2M_n - 1$, we have that:

$$
\mathrm{Var}\left(\ln \frac{\widetilde{X}_{(M_n)}}{\widetilde{X}_{(2M_n)}}\right) = \mathrm{Var}\left(\sum_{i=M_n}^{2M_n-1} E_i^*\right)
$$

$$
= E\left(\sum_{i=M_n}^{2M_n-1} E_i^*\right)^2
$$

$$
= \sum_{k=1}^{n} E\left\{\left(\sum_{i=M_n}^{2M_n-1} E_i^*\right)^2 \Big| M_n = k\right\} P\{M_n = k\}
$$

$$
= \sum_{k=1}^{n} E\left(\sum_{i=k}^{2k-1} E_i^*\right)^2 P\{M_n = k\}
$$

$$
= \sum_{k=1}^{n} \sum_{i=k}^{2k-1} E(E_i^*)^2 p_k
$$

$$= \sum_{k=1}^{n} p_k \gamma^2 \left(\frac{1}{k^2} + \frac{1}{(k+1)^2} + \cdots + \frac{1}{(2k-1)^2} \right)$$

$$\sim \gamma^2 \sum_{k=1}^{n} p_k \frac{1}{k} = \gamma^2 E\left(\frac{1}{M_n} \right).$$

We conclude that:

$$E(\widetilde{B}_{2,S_n}) \to \gamma \tag{5.5}$$

and

$$\mathrm{Var}(\widetilde{B}_{2,S_n}) \sim \gamma^2 \frac{1}{\ln^2 2} E\left(\frac{1}{M_n} \right), \tag{5.6}$$

as $n \to \infty$. $\qquad \square$

Before we give the prove of the Theorem 4.3, let us formulate the *Lindeberg condition* in case of a sequence $\{E_i^*\}$, $i = M_n, \ldots, 2M_n - 1$. Please see Rychlik (1979) for the definition of this condition in the general case and the prove of central limit theorem for sums of a random number of independent random variables.

Random Lindeberg condition for the sequence $\{E_i^*\}$. A sequence $\{E_i^*\}$, $i = M_n, \ldots, 2M_n - 1$ of independent random variables with the corresponding mathematical expectations $m_i = E(E_i^*)$ and $\sigma_i^2 = \mathrm{Var}(E_i^*)$ is said to satisfy the *Random Lindeberg condition* if, for every $\varepsilon > 0$:

$$\frac{1}{D_{M_n}^2} E \left\{ \sum_{i=M_n}^{2M_n-1} \int_{|x-m_i| \geq \varepsilon \sqrt{D_{M_n}^2}} (x - m_i)^2 \, dF_i(x) \right\} \to 0, \tag{5.7}$$

as $n \to \infty$, where

$$D_{M_n}^2 = \mathrm{Var}\left(\sum_{i=M_n}^{2M_n-1} E_i^* \right)$$

and F_i are distribution functions of E_i^*, for $i = M_n, \ldots, 2M_n - 1$.

Proof of Theorem 4.3. By using similar arguments as in the proof of Lemma 4.2, we conclude that:

$$E\left(\sum_{i=M_n}^{2M_n-1} \frac{\gamma^3}{i^3} \right) = \gamma^3 E\left(\sum_{i=M_n}^{2M_n-1} \frac{1}{i^3} \right) = \gamma^3 \sum_{j=1}^{\infty} p_j \sum_{i=j}^{2j-1} \frac{1}{i^3}. \tag{5.8}$$

The last sum $\sum_{i=j}^{2j-1} \frac{1}{i^3}$ may be easily transformed and then approximated into $\frac{1}{j^2} \int_0^1 \frac{1}{(1+x)^3} \, dx$. Therefore, we obtain:

$$E\left(\sum_{i=M_n}^{2M_n-1} \frac{\gamma^3}{i^3} \right) \sim \frac{3\gamma^3}{8} \sum_{j=1}^{\infty} p_j \frac{1}{j^2} = \frac{3\gamma^3}{8} E\left(\frac{1}{M_n^2} \right). \tag{5.9}$$

Finally, by using (5.6), one can see that:

$$\frac{1}{D^2_{M_n}} E\left\{ \sum_{i=M_n}^{2M_n-1} \int_{|x|\geq\varepsilon\sqrt{D^2_{M_n}}} x^2\, dF_i(x) \right\} \sim \frac{3\gamma\, E(\frac{1}{M_n^2})}{4E(\frac{1}{M_n})} \to 0,$$

when $n \to \infty$.

Now, by following all the steps from above and by applying the regularity Assumption C (see Bacro and Brito (1995), Corollary 2), we conclude straightforward that:

$$\ln 2E(M_n^{1/2})(\widetilde{B}_{2,S_n} - \gamma) \xrightarrow{d} N(0, \gamma^2). \qquad \square$$

References

Bacro, J. N. and Brito, M. (1993). Strong limiting behaviour of a simple tail Pareto-index estimator. *Statistics & Decisions* **3**, 133–143.

Bacro, J. N. and Brito, M. (1995). Weak limiting behaviour of a simple tail Pareto-index estimator. *Journal of Statistical Planning and Inference* **45**, 7–19. MR1342083

Bacro, J. N. and Brito, M. (1998). A tail bootstrap procedure for estimating the tail Pareto-index. *Journal of Statistical Planning and Inference* **71**, 245–260.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics* **31**, 307–327.

Cantoni, E. and Ronchetti, E. (2001). Robust inference for generalized linear models. *Journal of the American Statistical Association* **96**, 1022–1030.

Chörgó, S. and Viharos, L. (1998). Estimating the tail index. In *Asymptotic Methods in Probability and Statistics*, 833–881. Amsterdam: North Holand.

David, H. A. (1981). Order Statistics. New York: Wiley.

Davis, R. A. and Resnick, S. T. (1984). Tail estimates motivated by extreme value theory. *The Annals of Statistics* **12**, 1467–1487. MR0760700

De Haan, L. (1970). On Regular Variation and its Application to the Weak Convergence of Sample Extremes. *Mathematical Centre Tract* **32**.

Fraga Alves, M. I. (1992, 1995). Estimation of the tail parameter in the domain of attraction of an extremal distribution. *Journal of Statistical Planning and Inference* **45**, 143–173.

Haeusler, E. and Teugels, J. L. (1985). On asymptotic normality of Hill's estimator for the exponent of regular variation. *The Annals of Statistics* **13**, 743–756.

Hill, B. M. (1975). A simple general approach to inference about the tail of a distribution. *The Annals of Statistics* **3**, 1163–1174. MR0378204

Hill, J. B. (2005). On tail index estimation using dependent, heterogenous data. Working paper.

Hill, J. B. (2010). On tail index estimation for dependent, heterogeneus data. *Econometric Theory* **26**, 1398–1436.

Hsing, T. (1991). On tail index estimation using dependent data. *The Annals of Statistics* **19**, 1547–1569.

Hubert, M., Dierckx, G. and Vanpaemel, D. (2013). Detecting influential data points for the Hill estimator in Pareto-type distributions. *Computational Statistics & Data Analysis* **65**, 13–28.

Ilic, I. (2012). On tail index estimation using a sample with missing observations. *Statistics & Probability Letters* **82**, 949–958.

Ilic, I. and Mladenović, P. (2008). Incomplete samples and tail estimation for stationary sequences. *Novi Sad Journal of Mathematics* **38**, 97–104.

Jansen, D. and de Vries, C. (1991). On the frequency of large stock returns: Putting boom and busts into perspective. *Review of Economics and Statistics* **73**, 18–24.

Koedijk, K. and Kool, C. (1994). Tail estimates and the EMS target zone. *Review of International Economics* **2**, 153–165.

Leadbetter, M. R., Lindgren, G. and Rootzen, H. (1983). *Extremes and Related Properties of Random Sequences and Processes*. New York: Springer-Verlag.

Loretan, M. and Phillips, P. (1994). Testing the covariance structure of heavy-tailed time series. *Journal of Empirical Finance* **1**, 211–248.

Mandelbrot, B. (1960). The Pareto–Levy law and the distribntion of income. *International Economic Review* **1**, 79–106.

Mandelbrot, B. (1963). The variation of certain speculative prices. *Journal of Business* **36**, 394–419.

Mladenović, P. and Piterbarg, V. (2008). On estimation of the exponent of regular variation using a sample with missing observations. *Statistics & Probability Letters* **78**, 327–335.

Rychlik, Z. (1979). Martingale random central limit theorems. *Acta Mathematica Academiae Scientarum Hungaricae Tomus* **34**, 129–139. MR0546728

Vandewalle, B., Beirlant, J., Christmann, A. and Hubert, M. (2007). A robust estimator for the tail index of Pareto-type distributions. *Computational Statistics & Data Analysis* **51**, 6252–6268.

Department of Mathematics and Informatics
Faculty of Medical Sciences
University of Nis
Bulevar Dr. Zorana Djindjića 81
18000 Nis
Serbia
E-mail: ivana@medfak.ni.ac.rs

Department for Health Economics and Evidence Synthesis
Synergus AB
Kevinge Strand 20
182 57 Danderyd (Stockholm)
Sweden

Institute of Public Health, Medical Decision Making and
Health Technology Assessment
Department of Public Health, Health Services Reseaech and
Health Technology Assessment
UMIT - University for Health Sciences, Medical Informatics
and Technology
Eduard-WallnÃŭfer-Zentrum 1
6060 Hall i.T.
Austria
E-mail: vladica.velickovic@synergus.com