REVSTAT – Statistical Journal
Volume 2, Number 1, June 2004

ESTIMATING PARETO TAIL INDEX BASED ON SAMPLE MEANS

Authors: Alena Fialová, Jana Jurečková
– Department of Statistics, Charles University in Prague,
Czech Republic (fialova@karlin.mff.cuni.cz, jurecko@karlin.mff.cuni.cz)

Jan Picek
– Department of Applied Mathematics, Technical University in Liberec,
Czech Republic (jan.picek@vslib.cz)

Received: October 2003 Revised: October 2003 Accepted: January 2004

Abstract:

- We propose an estimator of the Pareto tail index m of a distribution, that competes well with the Hill, Pickands and moment estimators. Unlike the above estimators, that are based only on the extreme observations, the proposed estimator uses all observations; its idea rests in the tail behavior of the sample mean $\bar{X}_n$, having a simple structure under heavy-tailed F. The observations, partitioned into N independent samples of sizes n, lead to N sample means whose empirical distribution function is the main estimation tool. The estimator is strongly consistent and asymptotically normal as $N \to \infty$, while n remains fixed. Its behavior is illustrated in a simulation study.

Key-Words:

- *domain of attraction; Pareto index; strong embedding of empirical process; tail behavior.*

AMS Subject Classification:

- 62G05, 62G30.

1. INTRODUCTION

Let $X_1, ..., X_n$ be independent nonnegative random variables, identically distributed with distribution function F. The exact shape of F is generally unknown, but we assume that F is absolutely continuous with density f and nondegenerate right tail of the Pareto type satisfying

$$\lim_{x\to\infty} \frac{-\log\bigl(1-F(x)\bigr)}{m\log x} = 1 \tag{1.1}$$

for some $m > 0$. Then, by the von Mises condition (see Embrechts *et al.* [4]),

$$1 - F(x) = x^{-m} L(x) \tag{1.2}$$

where $L(x)$ is a function, slowly varying at ∞, and hence F belongs to the domain of attraction of the Fréchet distribution with the distribution function $\Phi_m(x) = \exp\{-x^{-m}\}\,,\ x > 0$.

Among the estimators of the Pareto index m or its reciprocal $\gamma = \frac{1}{m}$, proposed in the literature, the Hill [9], Pickands [15] and moment estimators [3] are the most well-known. Either of these estimators is based only on the fraction of the observations, namely on k_n largest ones, where $k_n \to \infty$ and $k_n/n \to 0$ as $n \to \infty$. The consistency and asymptotic normality of these estimators was proved under various regularity conditions on k_n and on F, some of them not easy to verify. The problem of the estimating was considered by many other authors, e.g. Smith [16], Beirlant *et al.* [2], Feuerverger *et al.* [8], Gomes and Martins [7].

We propose another estimator of the Pareto index m, that competes well with the above estimators; the regularity conditions, required for its strong consistency and asymptotic normality, are apparently more transparent and less restrictive. The proposed estimator uses all observations, unlike the estimators mentioned above. The idea of the estimator is based on the tail behavior of the sample mean $\bar{X}_n$, that has a simple structure under heavy-tailed F, satisfying (1.1). The estimator is strongly consistent and asymptotically normal and it was also discussed by the same authors in [5, 6].

The tail behavior of the sample mean is described is Section 2. The estimator is defined in Section 3, along with the formulation of its consistency and asymptotic normality. Its behavior is illustrated in a simulation study in Section 4. The proofs of the main results are postponed to Section 5. In Section 6 we propose a test of a one-sided hypothesis on m, that can be used as a preliminary test before the estimation.

2. TAIL-BEHAVIOR OF THE SAMPLE MEAN

Let $X_1, ..., X_n$ be a random sample from a distribution with an absolutely continuous distribution function F and density f, positive on interval (K_f, ∞), $K_f \geq 0$. Let $\bar{X}_n = \frac{1}{n}\sum_{i=1}^n X_i$.

For a heavy-tailed F, symmetric around 0, Jurečková [10] showed that the tail behavior of $\bar{X}_n$ coincides with that of F. The following lemma demonstrates a similar behavior of the sample mean also for heavy-tailed F concentrated only on the positive half-axis.

Lemma 2.1. *Let $X_1, ..., X_n$ be a random sample from the distribution with absolutely continuous d.f. F and density f such that*

(i) $f(x) = 0$ *for* $x < 0$ *and* $0 < f(x) < \infty$ *for* $x \geq K_f \geq 0$.

(ii) F *satisfies (1.1) for some* m, $0 < m < \infty$.

Then, for any fixed n,

$$\lim_{a\to\infty} \frac{-\log \mathbb{P}_m(\bar{X}_n > a)}{-\log\big(1 - F(a)\big)} = \frac{-\log\big(1 - F_{\bar{X}_n}(a)\big)}{-\log\big(1 - F(a)\big)} = 1 . \tag{2.1}$$

Proof: Let $0 \leq X_{n:1} \leq ... \leq X_{n:n}$ be the order statistics corresponding to $X_1, ..., X_n$. Then

$$\mathbb{P}(\bar{X}_n > a) = \mathbb{P}\left(\sum_{i=1}^n X_i > na\right) \geq \mathbb{P}(X_{n:n} > na) \geq 1 - F(na)$$

and

$$\mathbb{P}(\bar{X}_n > a) \leq \mathbb{P}(X_{n:n} > a) = 1 - \big(F(a)\big)^n \leq n\big(1 - F(a)\big) ,$$

hence

$$\varliminf_{a\to\infty} \frac{-\log \mathbb{P}(\bar{X}_n > a)}{-\log\big(1 - F(a)\big)} \geq \lim_{a\to\infty} \frac{-\log\big(n(1 - F(a))\big)}{-\log\big(1 - F(a)\big)} = 1$$

and

$$\varlimsup_{a\to\infty} \frac{-\log \mathbb{P}(\bar{X}_n > a)}{-\log\big(1 - F(a)\big)} \leq \lim_{a\to\infty} \frac{-\log\big(1 - F(na)\big)}{-\log\big(1 - F(a)\big)} = 1 ,$$

what implies (2.1). □

Notice that (2.1) and (1.1) imply

$$\lim_{a\to\infty} \frac{-\log \mathbb{P}_m(\bar{X}_n > a)}{m \log a} \;=\; 1 \;, \tag{2.2}$$

hence

$$m \;=\; \lim_{a\to\infty} m_n(a) \;, \tag{2.3}$$

where

$$m_n(a) \;=\; \frac{-\log \mathbb{P}_m(\bar{X}_n > a)}{\log a} \;=\; \frac{-\log\bigl(1 - F_{\bar{X}_n}(a)\bigr)}{\log a} \tag{2.4}$$

with $F_{\bar{X}_n}$ being the distribution function of $\bar{X}_n$. There are two possibilities how to estimate m with the aid of formula (2.4): First, we can estimate the unknown $F_{\bar{X}_n}$ in (2.4) by the empirical distribution function, based on N realizations of $\bar{X}_n$ (nonparametric approach). Second, the distribution function can be modelled by the by some parametric model whose parameters are then estimated. The "perturbed Pareto distribution", considered recently by Feuerverger and Hall [8], is a possible parametric model. Both approaches lead to the asymptotically normal estimators, that are generally biased, unless the distribution has exactly Pareto tails. The parametric model enables to reduce the bias, provided it is correct, e.g. using efficient estimators of its parameters. The bias in the nonparametric approach is expressed by means on the unknown slowly varying function; it can be still reduced if the slowly varying function can be further parametrized.

In the present paper, we shall develop the nonparametric approach, replacing $F_{\bar{X}_n}$ by the empirical distribution function. In this way we obtain a consistent estimator of m under $N \to \infty$, while n remains fixed. Because we need to estimate the limit of (2.4) as $a \to \infty$, the argument a_N of the empirical distribution function should be sufficiently large, but some observations should be still greater than a_N.

The estimator and its properties are described in the next section.

3. ESTIMATOR OF THE TAIL INDEX BASED ON SAMPLE MEANS

Let us partition the set of observations into N non-overlapping samples of the same sizes n (a modification to different sample sizes is possible), denoted as $(X_1^{(1)}, ..., X_n^{(1)})$, ..., $(X_1^{(N)}, ..., X_n^{(N)})$. Then the vector $(\bar{X}_n^{(1)}, ..., \bar{X}_n^{(N)})$ of the corresponding sample means is a random sample from a distribution with distribution function $F_{\bar{X}_n}(x) = \mathbb{P}(\bar{X}_n \le x)$ (unknown).

Denote $\widehat{F}^{(N)}_{\bar{X}_n}(x) = \frac{1}{N}\sum_{j=1}^{N} I[\bar{X}_n^{(j)} \leq x]$ the empirical distribution function, based on $(\bar{X}_n^{(1)}, ..., \bar{X}_n^{(N)})$.

The argument a_N of the empirical distribution function should be sufficiently large, but such that there are still some observations behind a_N. If we know that F is not lighter than the Pareto distribution with index m_0 for some fixed m_0, $0 < m_0 < \infty$, hence we know that $0 < m \leq m_0$, then a possible choice of a_N is as in (3.1) below. This situation is considered in the present paper. We can either have such information from the experience or from the character of the experiment.

Remark 3.1. Another possibility would be a preliminary test estimation, when we first apply a preliminary test of the hypothesis $\mathbf{H}$: $0 < m \leq m_0$. In Section 5 we shall briefly describe one possible test of $\mathbf{H}$ based on the sample means. Other tests of $\mathbf{H}$ were recently proposed and numerically illustrated by Picek and Jurečková [14], Jurečková and Picek [12]; the test on the tail of errors in linear model was proposed by Jurečková [11]. A preliminary test estimator will be a subject of the next study.

Choose the sequence $\{a_N\}_{N=1}^{\infty}$, $a_N \to \infty$ as $N \to \infty$, in the following way:

$$a_N = N^{\frac{1-\delta}{m_0}}, \qquad \text{with a fixed } \delta \in (0,1) \tag{3.1}$$

and consider the sequence of random functions

$$\widehat{m}_N(a) = \tilde{m}_N(a)\, I\Big[0 < \widehat{F}^{(N)}_{\bar{X}_n}(a) < 1\Big] + m_0\, I\Big[\widehat{F}^{(N)}_{\bar{X}_n}(a) = 0 \text{ or } 1\Big], \quad a > 0, \tag{3.2}$$

where

$$\tilde{m}_N(a) = \frac{-\log\Big(1 - \widehat{F}^{(N)}_{\bar{X}_n}(a)\Big)}{\log a}, \qquad a > 0. \tag{3.3}$$

We propose $\widehat{m}_N = \widehat{m}_N(a_N)$ as an estimator of the parameter m; more precisely,

$$\widehat{m}_N = \tilde{m}_N(a_N)\, I\Big[0 < \widehat{F}^{(N)}_{\bar{X}_n}(a_N) < 1\Big] + m_0\, I\Big[\widehat{F}^{(N)}_{\bar{X}_n}(a_N) = 0 \text{ or } 1\Big] \tag{3.4}$$

with $\tilde{m}_N(a)$ defined in (3.3) and a_N defined in (3.1) with a fixed choice of δ, $0 < \delta < 1$.

We must first show that the estimator $\widehat{m}_N$ is well defined. It follows from the following lemma:

Lemma 3.1. *Let F satisfy the conditions of Lemma 2.1 with $0<m\le m_0$, $m_0>0$ fixed. Let $\{a_N\}$ be the sequence defined in (3.1). Then $a_N\to\infty$ and*

$$\mathbb{P}_m\Big(0<\hat F^{(N)}_{\bar X_n}(a_m)<1\Big)\;\to\;1 \qquad \text{as}\quad N\to\infty\,. \tag{3.5}$$

Proof: If F is heavy-tailed with Pareto index m, satisfying (1.1), then, by Lemma 2.1,

$$\lim_{a\to\infty}\frac{1-F(a)}{m\log a}\;=\;\lim_{a\to\infty}\frac{1-F_{\bar X_n}(a)}{m\log a}\;=\;1 \tag{3.6}$$

and both F and $F_{\bar X_n}$ belong to the domain of attraction of the Fréchet distribution Φ_m with the distribution function $\Phi_m(x)=\exp\{-x^{-m}\},\ x>0$. Let $\bar{\mathcal{X}}^{(N)}_n=\max_{1\le j\le N}\bar X^{(j)}_n$ denote the maximum of $\bar X^{(1)}_n,\ldots,\bar X^{(N)}_n$. Then

$$\mathbb{P}_m\left(\frac{\bar{\mathcal{X}}^{(N)}_n}{\xi_N}\le x\right)\;\to\;\Phi_m(x) \qquad \text{as}\quad N\to\infty \tag{3.7}$$

with ξ_N satisfying $N\big[1-F_{\bar X_n}(\xi_N)\big]=1,\ N=1,2,\ldots$; then we conclude from (3.10) that $\xi_N=N^{\frac1m}L^*_2(N)$ with some slowly varying function L^*_2 and, by (3.7),

$$\mathbb{P}_m\Big(\bar{\mathcal{X}}^{(N)}_n\le a_N\Big)\;=\;\mathbb{P}_m\left(\frac{\bar{\mathcal{X}}^{(N)}_n}{\xi_N}\le\frac{a_N}{\xi_N}\right)\;\to\;0 \qquad \text{as}\quad N\to\infty\,. \tag{3.8}$$

It means that at least one $\bar X^{(j)}_n$ lies above a_N with probability tending to 1, and thus

$$\lim_{N\to\infty}\mathbb{P}_m\Big(\hat F^{(N)}_{\bar X_n}(a_N)<1\Big)\;=\;1\,.$$

On the other hand, we obtain from (3.10),

$$\begin{aligned}\mathbb{P}_m\left(\min_{1\le j\le N}\bar X^{(j)}_n\ge a_N\right)\;&=\;\Big(1-F_{\bar X_n}(a_N)\Big)^N\\ &=\;a_N^{-mN}\big(L^*(a_N)\big)^N\\ &=\;N^{-\frac{m(1-\delta)}{m_0}N}\big(L^*(a_N)\big)^N\;\to\;0 \qquad \text{as}\quad N\to\infty\,,\end{aligned}$$

and hence there is at least one $\bar X^{(j)}_n$ below a_N with probability tending to one. This completes the proof of (3.5). □

The first main property of estimator $\widehat m_N$ is its strong consistency with respect to the asymptotics $N\to\infty$:

Theorem 3.1. *Let $\{X_1, X_2, ...\}$ be a sequence of random variables, identically distributed according to distribution function F of the Pareto type (1.1), satisfying the conditions (i) and (ii) of Lemma 2.1 with $0 < m \le m_0 < \infty$. Let $\widehat{m}_N$ be the estimator of m defined in (3.4). Then*

$$\widehat{m}_N \to m \quad \text{with probability } 1, \qquad \text{as } N \to \infty \, . \tag{3.9}$$

The second main result is the asymptotic normality of $\widehat{m}_N$. The problem of estimating m is semiparametric in its nature, involving an unknown slowly varying function. If distribution function F is of the type (1.1) with index m, then Lemma 2.1 implies that $F_{\bar{X}_n}$ also satisfies (1.1) with the same m; hence, by the von Mises condition, it has the form

$$1 - F_{\bar{X}_n}(x) \;=\; x^{-m} L^*(x) \;, \tag{3.10}$$

where $L^*(x)$ is a function, slowly varying at ∞. The presence of L^* can cause a bias in the asymptotic distribution of $\widehat{m}_N$, generally not asymptotically negligible, unless we impose some more restrictive condition on F. We shall see (Lemma 5.1) that $(\widehat{m}_N - m_n(a_N))$, with $m_n(\cdot)$ defined in (2.4), is asymptotically normal and unbiased, while the bias of $(\widehat{m}_N - m)$ is due to the term $(m_n(a_N) - m)$, that tends to 0, but generally not fast enough to eliminate function L^*.

Theorem 3.2. *Under the conditions of Theorem 3.1, the sequence*

$$N^{\frac{1}{2}} \, \log a_N \left(\frac{1 - F_{\bar{X}_n}(a_N)}{F_{\bar{X}_n}(a_N)} \right)^{\frac{1}{2}} \left(\widehat{m}_N - m + \frac{\log L^*(a_N)}{\log a_N} \right) \tag{3.11}$$

is asymptotically normally distributed as $N \to \infty$, where L^* is the function, defined in (3.10).

Remark 3.2. The order of the coefficient by $\left(\widehat{m}_N(a_N) - m + \frac{\log L^*(a_N)}{\log a_N} \right)$ in (3.11) can be alternatively expressed as

$$\begin{aligned} N^{\frac{1}{2}} \, \log a_N \left(\frac{1 - F_{\bar{X}_n}(a_N)}{F_{\bar{X}_n}(a_N)} \right)^{\frac{1}{2}} &\sim \tfrac{1-\delta}{m_0} \big(L^*(a_N) \big)^{\frac{1}{2}} \cdot N^{\frac{1}{2}\left(1 - \frac{m}{m_0}(1-\delta)\right)} \, \log N \\ &\ge \tfrac{1-\delta}{m_0} \big(L^*(a_N) \big)^{\frac{1}{2}} \cdot N^{\frac{\delta m}{2 m_0}} \quad \Big(\to \infty \ \text{ as } \ N \to \infty \Big) \, , \end{aligned} \tag{3.12}$$

where $b_N \sim c_N$ means that $\lim_{N \to \infty} \frac{b_N}{c_N} \to 1$.

4. NUMERICAL ILLUSTRATION

The performance of the estimation procedure for different choices of m and δ is illustrated on the simulated random samples: The replications ($N=200$ and $N=2000$) of samples of sizes $n=5$ were simulated 1000 times from the following distributions:

Pareto $$F(x) = 1-\left(\frac{1}{1+x}\right)^m, \quad x \ge 0\ ;$$

Burr $$F(x) = 1-\left(\frac{1}{1+x^m}\right)^\alpha, \quad x \ge 0\ ;$$

Generalized Pareto $$F(x) = \begin{cases} 1-\left(1+\frac{x}{m\beta}\right)^{-m} & \text{if } x\ge 0,\ 0<m<\infty,\ \beta>0, \\ 1-\left(1+\frac{x}{m\beta}\right)^{-m} & \text{if } 0\le x\le -m\beta,\ m<0,\ \beta>0, \\ 1-e^{-x/\beta} & \text{if } m=\infty,\ \beta>0, \\ 0 & \text{otherwise}\ ; \end{cases}$$

Inverse normal $$F(x) = \begin{cases} 2\left(1-\Phi\left(\frac{1}{\sqrt{x}}\right)\right) & x>0, \\ 0 & x\le 0\ . \end{cases}$$

For each distribution we proceeded as follows:

(1) we generated the independent observations $X_1,\ ...,\ X_n, X_{n+1},\ ...,\ X_{2n},\ ...,\ X_{Nn}$;

(2) computed sample means $\overline{X}_n^{(1)}, ..., \overline{X}_n^{(N)}$

(3) and found the empirical distribution function based on $\overline{X}_n^{(1)}, ..., \overline{X}_n^{(N)}$;

(4) for $a_N = N^{\frac{1-\delta}{m_0}}$ we calculated

$$\widehat{m}_N = \tilde{m}_N(a_N)\, I\left[0<\widehat{F}_{\bar{X}_n}^{(N)}(a_N)<1\right] + m_0\, I\left[\widehat{F}_{\bar{X}_n}^{(N)}(a_N)=0 \text{ or } 1\right]\ ;$$

(5) Step (4) was repeated for various values m_0, δ;

(6) For a comparison, the Hill estimator

$$H(k) \; = \; \frac{1}{k} \sum_{i=1}^{k} \log X_{(Nn-i+1:Nn)} - \log X_{(Nn-k:Nn)} \; ,$$

the Pickands estimator

$$P(k) \; = \; \frac{1}{\log 2} \; \log \left(\frac{X_{Nn-k+1:Nn} - X_{Nn-2k+1:Nn}}{X_{Nn-2k+1:Nn} - X_{Nn-4k+1:Nn}} \right) \; ,$$

the moment estimator

$$M(k) \; = \; 1 + M(k)^{(1)} + \frac{1}{2} \left(\frac{\left(M(k)^{(1)}\right)^2}{M(k)^{(2)}} - 1 \right)^{-1} ,$$

where

$$M(k)^{(j)} \; = \; \frac{1}{k} \sum_{i=1}^{k} \Big(\log X_{(Nn-i+1:Nn)} - \log X_{(Nn-k:Nn)} \Big)^j \; ,$$

and Gomes and Martins [7] estimator

$$GM(k) \; = \; \frac{1}{k} \sum_{i=1}^{k} U_i \; - \left(\frac{1}{k} \sum_{i=1}^{k} i \, U_i \right) \frac{\sum_{i=1}^{k} (2\,i - k - 1)\, U_i}{\sum_{i=1}^{k} i\,(2\,i - k - 1)\, U_i} \; ,$$

where

$$U_i \; = \; i \left[\log \frac{X_{Nn-i+1:Nn}}{X_{Nn-i:Nn}} \right] \; ,$$

were computed for $k = 1, ..., Nn - 1$.

(7) steps (1)–(6) were repeated 1 000 times.

(8) Selected sample quantiles of estimates $\left(\widehat{m}_N^1, ..., \widehat{m}_N^{1000}\right)$ and selected sample statistics of pertaining estimates were computed and tabulated.

Selected sample quantiles for different distributions of the errors are summarized in Table 1 and Table 2. Fig. 1 and 2 show the behaviour of the tail index estimator with regard to δ and m_0 in 1000 simulated samples ($N = 2000$) of Pareto with $m = 1$.

For a comparison, the Hill, Pickands, moment and Gomes and Martins [7] estimators were computed. The question is the choice of k, respectively δ for our procedures. To compare we followed the standard approach of minimizing the mean squared error (MSE); the Table 3 give the selected sample statistics of estimators of m for various distribution shapes of errors.

Table 1: Sample quantiles of the estimation of Pareto index under different distributions for some values m_0 and δ $(N=200)$.

sample distr.	m_0	δ	min	5%	25%	50% quantile	75%	95%	max
Pareto	0.75	0.1	0.279	0.321	0.354	0.379	0.397	0.443	0.580
$m=0.5$		0.5	0.205	0.242	0.263	0.278	0.297	0.327	0.355
	2	0.1	0.141	0.165	0.187	0.200	0.214	0.239	0.274
		0.5	0.043	0.063	0.080	0.092	0.105	0.123	0.164
Pareto	1.5	0.1	0.694	0.815	0.885	0.942	1.013	1.160	1.449
$m=1$		0.5	0.548	0.663	0.721	0.763	0.808	0.884	1.038
	3	0.1	0.489	0.617	0.670	0.707	0.747	0.824	0.938
		0.5	0.253	0.311	0.356	0.388	0.428	0.479	0.551
Pareto	3.5	0.1	3.083	3.380	3.500	3.500	3.889	3.889	3.889
$m=3$		0.5	3.042	3.422	3.717	3.958	4.253	4.874	6.084
	5	0.1	3.252	3.677	4.102	4.404	4.829	5.556	5.556
		0.5	2.272	2.774	2.991	3.185	3.344	3.711	4.254
Burr	0.75	0.1	0.283	0.327	0.354	0.379	0.407	0.443	0.615
$\alpha=1$		0.5	0.220	0.267	0.293	0.310	0.327	0.360	0.416
$m=0.5$	2	0.1	0.200	0.232	0.258	0.274	0.295	0.321	0.359
		0.5	0.164	0.202	0.238	0.259	0.280	0.314	0.373
Burr	1.5	0.1	0.694	0.837	0.912	0.975	1.055	1.164	1.667
$\alpha=1$		0.5	0.654	0.752	0.820	0.870	0.926	1.020	1.250
$m=1$	3	0.1	0.617	0.737	0.812	0.860	0.911	0.983	1.173
		0.5	0.541	0.667	0.741	0.796	0.843	0.930	1.066
Burr	3.5	0.1	3.083	3.500	3.500	3.500	3.889	3.889	3.889
$\alpha=1$		0.5	3.500	3.958	4.429	4.633	5.168	6.084	7.000
$m=3$	5	0.1	3.375	3.868	4.404	4.829	5.000	5.556	5.556
		0.5	3.290	3.847	4.166	4.443	4.767	5.310	6.327
General.	0.75	0.1	0.315	0.370	0.407	0.430	0.456	0.527	0.661
Pareto		0.5	0.270	0.323	0.355	0.376	0.398	0.435	0.528
$m=0.5$	2	0.1	0.243	0.287	0.317	0.330	0.354	0.390	0.434
β=1		0.5	0.192	0.232	0.264	0.291	0.314	0.349	0.418
General.	1.5	0.1	0.694	0.837	0.912	0.975	1.055	1.164	1.667
Pareto		0.5	0.617	0.737	0.812	0.860	0.911	0.983	1.173
$m=1$	3	0.1	0.654	0.752	0.820	0.870	0.926	1.020	1.250
$\beta=1$		0.5	0.541	0.667	0.741	0.796	0.843	0.930	1.066
General.	3.5	0.1	1.901	2.129	2.363	2.574	2.708	3.083	3.889
Pareto		0.5	1.798	2.025	2.180	2.314	2.468	2.716	3.252
$m=3$	5	0.1	1.635	1.971	2.126	2.266	2.421	2.646	3.042
$\beta=1$		0.5	1.753	2.007	2.151	2.304	2.437	2.655	3.038
Inverse	0.75	0.1	0.304	0.354	0.388	0.407	0.443	0.488	0.661
normal		0.5	0.207	0.243	0.270	0.287	0.308	0.335	0.390
	2	0.1	0.242	0.297	0.323	0.341	0.365	0.398	0.456
		0.5	0.132	0.168	0.192	0.217	0.238	0.269	0.337

Table 2: Sample quantiles of the estimation of Pareto index under different distributions for some values m_0 and δ ($N = 2000$).

sample distr.	m_0	δ	min	5%	25%	50% quantile	75%	95%	max
Pareto	0.75	0.1	0.361	0.388	0.402	0.411	0.424	0.440	0.476
$m = 0.5$		0.5	0.313	0.327	0.336	0.342	0.349	0.358	0.374
	2	0.1	0.249	0.261	0.269	0.274	0.279	0.288	0.302
		0.5	0.137	0.146	0.151	0.156	0.161	0.168	0.179
Pareto	1.5	0.1	0.881	0.921	0.961	0.989	1.021	1.088	1.185
$m = 1$		0.5	0.825	0.857	0.880	0.899	0.915	0.942	0.978
	3	0.1	0.786	0.826	0.848	0.865	0.881	0.907	0.933
		0.5	0.540	0.562	0.577	0.590	0.604	0.623	0.649
Pareto	3.5	0.1	3.180	3.500	3.500	3.500	3.889	3.889	3.889
$m = 3$		0.5	3.899	4.112	4.288	4.447	4.570	4.879	5.350
	5	0.1	3.739	4.036	4.379	4.542	4.753	5.556	5.556
		0.5	3.648	3.780	3.877	3.968	4.051	4.202	4.488
Burr	0.75	0.1	0.361	0.390	0.402	0.411	0.424	0.440	0.480
$\alpha = 1$		0.5	0.322	0.336	0.345	0.351	0.358	0.368	0.383
$m = 0.5$	2	0.1	0.276	0.293	0.301	0.307	0.313	0.322	0.338
		0.5	0.237	0.248	0.257	0.263	0.270	0.279	0.294
Burr	1.5	0.1	0.881	0.921	0.961	0.989	1.021	1.088	1.185
$\alpha = 1$		0.5	0.843	0.891	0.915	0.933	0.953	0.980	1.028
$m = 1$	3	0.1	0.833	0.873	0.898	0.917	0.933	0.960	0.990
		0.5	0.735	0.783	0.805	0.820	0.838	0.863	0.897
Burr	3.5	0.1	3.180	3.500	3.500	3.500	3.889	3.889	3.889
$\alpha = 1$		0.5	4.112	4.391	4.570	4.712	4.976	5.208	6.362
$m = 3$	5	0.1	3.873	4.133	4.542	4.753	5.049	5.556	5.556
		0.5	4.171	4.391	4.528	4.635	4.752	4.907	5.285
General.	0.75	0.1	0.394	0.424	0.437	0.450	0.464	0.485	0.529
Pareto		0.5	0.373	0.392	0.402	0.410	0.418	0.429	0.444
$m = 0.5$	2	0.1	0.335	0.356	0.366	0.373	0.380	0.391	0.409
β=1		0.5	0.285	0.298	0.307	0.314	0.322	0.333	0.353
General.	1.5	0.1	0.881	0.921	0.961	0.989	1.021	1.088	1.185
Pareto		0.5	0.833	0.873	0.898	0.917	0.933	0.960	0.990
$m = 1$	3	0.1	0.843	0.891	0.915	0.933	0.953	0.980	1.028
$\beta = 1$		0.5	0.735	0.783	0.805	0.820	0.838	0.863	0.897
General.	3.5	0.1	2.410	2.539	2.662	2.765	2.893	3.065	3.889
Pareto		0.5	2.267	2.390	2.471	2.527	2.588	2.696	2.878
$m = 3$	5	0.1	2.178	2.270	2.326	2.379	2.430	2.510	2.671
$\beta = 1$		0.5	2.028	2.161	2.216	2.260	2.301	2.363	2.466
Inverse	0.75	0.1	0.377	0.411	0.426	0.437	0.450	0.468	0.523
normal		0.5	0.306	0.323	0.332	0.338	0.345	0.355	0.377
	2	0.1	0.354	0.369	0.380	0.386	0.394	0.406	0.420
		0.5	0.230	0.242	0.251	0.257	0.264	0.272	0.288

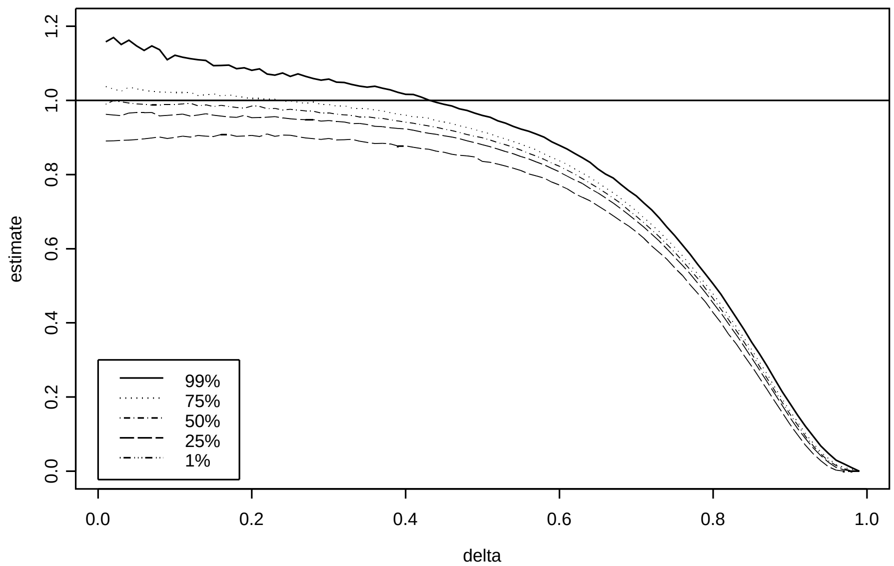

Figure 1: Dependence of tail index estimator in 1000 simulated samples of Pareto ($m = 1$) on the parameter δ for $m_0 = 1.5$.
Plotted are the median and the 1, 25, 75 and 99 percentiles.

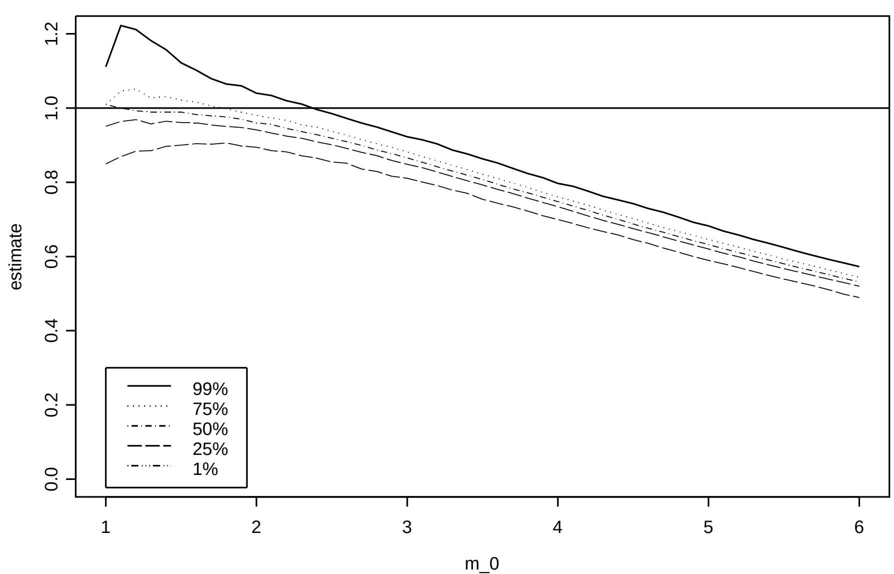

Figure 2: Dependance of tail index estimator in 1000 simulated samples of Pareto ($m = 1$) on the value m_0 for $\delta = 0.1$.
Plotted are the median and the 1, 25, 75 and 99 percentiles.

The Table 3 shows that the estimator based on the sample means (FJP) can be considered as comparable with the most popular estimators of the tail index. The regularity conditions, required for its strong consistency and asymptotic normality, are apparently more transparent and less restrictive.

Table 3: Sample statistics of the estimates of the Pareto index under different distributions for minimal MSE and $N = 200$ and $n = 5$

sample	method	fraction	MSE	mean	median	var	MAD
Pareto $m = 1$	Hill	$k = 998$	0.0010	1.0003	0.9984	0.0010	0.0321
	Moment	$k = 998$	0.0023	1.0053	1.0033	0.0022	0.0454
	Pickands	$k = 985$	0.0221	1.0177	0.9967	0.0218	0.1349
	Gomes	$k = 997$	0.0044	1.0016	0.9968	0.0044	0.0655
	FJP	$\delta = 0.15$	0.0123	0.9542	0.9371	0.0102	0.0900
Burr $\alpha = 1$ $m = 1$	Hill	$k = 112$	0.0098	0.9517	0.9489	0.0075	0.0885
	Moment	$k = 257$	0.0101	0.9478	0.9383	0.0074	0.0802
	Pickands	$k = 985$	0.0221	1.0177	0.9967	0.0218	0.1349
	Gomes	$k = 998$	0.0012	1.0007	0.9989	0.0012	0.0345
	FJP	$\delta = 0.22$	0.0111	0.9574	0.9402	0.0093	0.0981
General. Pareto $m = 0.5$ β=1	Hill	$k = 310$	0.0010	0.4847	0.4841	0.0007	0.0261
	Moment	$k = 367$	0.0010	0.4880	0.4863	0.0009	0.0283
	Pickands	$k = 993$	0.0020	0.5030	0.4997	0.0020	0.0429
	Gomes	$k = 482$	0.0025	0.5227	0.5210	0.0020	0.0440
	FJP	$\delta = 0.01$	0.0084	0.4177	0.4123	0.0016	0.0395
General. Pareto $m = 1$ $\beta = 1$	Hill	$k = 112$	0.0098	0.9517	0.9489	0.0075	0.0885
	Moment	$k = 257$	0.0101	0.9478	0.9383	0.0074	0.0802
	Pickands	$k = 985$	0.0221	1.0177	0.9967	0.0218	0.1349
	Gomes	$k = 998$	0.0012	1.0007	0.9989	0.0012	0.0345
	FJP	$\delta = 0.22$	0.0111	0.9574	0.9402	0.0093	0.0981
General. Pareto $m = 3$ $\beta = 1$	Hill	$k = 23$	0.5527	2.4329	2.3598	0.2314	0.4397
	Moment	$k = 257$	0.5037	2.5140	2.4248	0.2678	0.4368
	Pickands	$k = 890$	16.1112	3.6237	3.0364	15.7379	1.1255
	Gomes	$k = 102$	0.4795	2.4276	2.4020	0.1520	0.3966
	FJP	$\delta = 0.01$	0.2869	2.5618	2.5565	0.0949	0.2841
Inverse normal	Hill	$k = 360$	0.0008	0.4894	0.4888	0.0007	0.0250
	Moment	$k = 472$	0.0008	0.4889	0.4881	0.0007	0.0258
	Pickands	$k = 893$	0.0026	0.5142	0.5111	0.0024	0.0467
	Gomes	$k = 588$	0.0021	0.5202	0.5184	0.0017	0.0407
	FJP	$\delta = 0.01$	0.0127	0.3937	0.3890	0.0014	0.0347

5. PROOFS OF THEOREMS 3.1 AND 3.2

5.1. Asymptotic normality

We shall start with the asymptotic normality of $\widehat{m}_N$; and first prove that the sequence

$$N^{\frac{1}{2}} \log a_N \left(\frac{1-F_{\bar{X}_n}(a_N)}{F_{\bar{X}_n}(a_N)}\right)^{\frac{1}{2}} \Big(\widehat{m}_N - m_n(a_N)\Big) ,$$

with $m_n(\cdot)$ given in (2.4), has asymptotically standard normal distribution:

Lemma 5.1. *Let $\{X_1, X_2, ...\}$ be a sequence of independent random variables, identically distributed with distribution function F of the Pareto type satisfying the conditions (i) and (ii) of Lemma 2.1 with $0 < m \le m_0 < \infty$. Put $a_N = N^{\frac{1-\delta}{m_0}}$, $0 < \delta < 1$ and*

$$(5.1)\qquad \begin{aligned} \widehat{m}_N &= \tilde{m}_N(a_N)\, I\Big[0 < \hat{F}^{(N)}_{\bar{X}_n}(a_N) < 1\Big] + m_0\, I\Big[\hat{F}^{(N)}_{\bar{X}_n}(a_N) = 0 \text{ or } 1\Big] ,\\ \tilde{m}_N(a) &= \frac{-\log\Big(1 - \hat{F}^{(N)}_{\bar{X}_n}(a)\Big)}{\log a} , \qquad a > 0 . \end{aligned}$$

Then the sequence

$$(5.2)\qquad N^{\frac{1}{2}} \log a_N \left(\frac{1-F_{\bar{X}_n}(a_N)}{F_{\bar{X}_n}(a_N)}\right)^{\frac{1}{2}} \Big(\widehat{m}_N - m_n(a_N)\Big)$$

with $m_n(x)$ defined in (2.4), is asymptotically normally distributed $\mathcal{N}(0,1)$, as $N \to \infty$ and for any fixed n.

Proof: By the Hungarian embedding theorems (see, e.g., [13]), there exists a sequence of Brownian bridges $\{\mathcal{B}_N\}$, $\mathcal{B}_N$ dependent on $\bar{X}^{(1)}_n, ..., \bar{X}^{(N)}_n$, such that

$$(5.3)\qquad \sup_{a\in\mathbb{R}} \left| \sqrt{N}\Big[1 - \hat{F}^{(N)}_{\bar{X}_n}(a) - \big(1 - F_{\bar{X}_n}(a)\big)\Big] + \mathcal{B}_N\big(F_{\bar{X}_n}(a)\big) \right| = \mathcal{O}\Big(N^{-\frac{1}{2}} \log N\Big) \text{ a.s.}$$

as $N \to \infty$.

Because $B_N(F_{\bar{X}_n}(a))$ is normally distributed $\mathcal{N}\big(0, F_{\bar{X}_n}(a)(1 - F_{\bar{X}_n}(a))\big)$, then

$$\mathbb{P}_m\left(\mathcal{B}_N\big(F_{\bar{X}_n}(a)\big) > C\Big[F_{\bar{X}_n}(a)\Big(1 - F_{\bar{X}_n}(a)\Big)\Big]^{\frac{1}{2}}\right) = 1 - \Phi(C) ,$$

holds for all $a \in \mathbb{R}$ and all $C > 0$, where Φ is the standard normal distribution function; hence $\forall \varepsilon > 0 \ \exists C > 0$ such that, for all $a \in \mathbb{R}$,

$$\mathbb{P}_m \left(\mathcal{B}_N\big(F_{\bar{X}_n}(a)\big) > C\Big[F_{\bar{X}_n}(a)\left(1 - F_{\bar{X}_n}(a)\right)\Big]^{\frac{1}{2}} \right) < \varepsilon \ . \tag{5.4}$$

Let us first consider the first term of $\widehat{m}_N$, i.e.

$$\tilde{m}_N(a_N)\, I\Big[0 < \hat{F}^{(N)}_{\bar{X}_n}(a_N) < 1\Big] = \frac{-\log\Big(1 - \hat{F}^{(N)}_{\bar{X}_n}(a_N)\Big)}{\log a_N}\, I\Big[0 < \hat{F}^{(N)}_{\bar{X}_n}(a_N) < 1\Big] \ .$$

We can write

$$\begin{aligned}
&\sqrt{N}\, \log a_N \left(\frac{1 - F_{\bar{X}_n}(a_N)}{F_{\bar{X}_n}(a_N)}\right)^{\frac{1}{2}} \Big(\tilde{m}_N(a_N) - m_n(a_N)\Big)\, I\Big[0 < \hat{F}^{(N)}_{\bar{X}_n}(a_N) < 1\Big] = \\
&\quad = \sqrt{N}\, \log a_N \left(\frac{1 - F_{\bar{X}_n}(a_N)}{F_{\bar{X}_n}(a_N)}\right)^{\frac{1}{2}} \\
&\quad\quad \cdot \left[\frac{-\log\Big(1 - \hat{F}^{(N)}_{\bar{X}_n}(a_N)\Big)}{\log a_N} - \frac{-\log\Big(1 - F_{\bar{X}_n}(a_N)\Big)}{\log a_N}\right] I\Big[0 < \hat{F}^{(N)}_{\bar{X}_n}(a_N) < 1\Big] \\
&\quad = \sqrt{N} \left(\frac{1 - F_{\bar{X}_n}(a_N)}{F_{\bar{X}_n}(a_N)}\right)^{\frac{1}{2}} \\
&\quad\quad \cdot \left(- \log \left[\frac{1 - \hat{F}^{(N)}_{\bar{X}_n}(a_N)}{1 - F_{\bar{X}_n}(a_N)} - 1 + 1\right]\right) I\Big[0 < \hat{F}^{(N)}_{\bar{X}_n}(a_N) < 1\Big] \ .
\end{aligned} \tag{5.5}$$

An expansion of $\log(1 + x)$ or $\log(1 - x)$, $x > 0$, gives

$$\begin{aligned}
&- \log \left[\frac{1 - \hat{F}^{(N)}_{\bar{X}_n}(a_N)}{1 - F_{\bar{X}_n}(a_N)} - 1 + 1\right] = \\
&\quad = 1 - \frac{1 - \hat{F}^{(N)}_{\bar{X}_n}(a_N)}{1 - F_{\bar{X}_n}(a_N)} + \mathcal{O}\left(\left[1 - \frac{1 - \hat{F}^{(N)}_{\bar{X}_n}(a_N)}{1 - F_{\bar{X}_n}(a_N)}\right]^2\right);
\end{aligned} \tag{5.6}$$

further we obtain from (5.3)

$$\begin{aligned}
&\sqrt{N} \left(\frac{1 - F_{\bar{X}_n}(a_N)}{F_{\bar{X}_n}(a_N)}\right)^{\frac{1}{2}} \left[1 - \frac{1 - \hat{F}^{(N)}_{\bar{X}_n}(a_N)}{1 - F_{\bar{X}_n}(a_N)}\right] = \\
&\quad = \frac{\mathcal{B}_N\big(F_{\bar{X}_n}(a_N)\big)}{\Big[F_{\bar{X}_n}(a_N)\Big(1 - F_{\bar{X}_n}(a_N)\Big)\Big]^{\frac{1}{2}}} \Big(1 + \mathbf{o}_p(1)\Big)
\end{aligned} \tag{5.7}$$

and

$$\sqrt{N}\left(\frac{1-F_{\bar{X}_n}(a_N)}{F_{\bar{X}_n}(a_N)}\right)^{\frac{1}{2}}\left[1-\frac{1-\hat{F}^{(N)}_{\bar{X}_n}(a_N)}{1-F_{\bar{X}_n}(a_N)}\right]^2 =$$

$$(5.8)\qquad = N^{-\frac{1}{2}\left(1-\frac{m}{m_0}\right)-\frac{\delta m}{2m_0}}\left(L^*(a_N)\right)^{-\frac{1}{2}}\,\frac{\Big(\mathcal{B}_N\left(F_{\bar{X}_n}(a_N)\right)\Big)^2}{F_{\bar{X}_n}(a_N)\Big(1-F_{\bar{X}_n}(a_N)\Big)}$$

$$= \mathbf{o}_p\left(N^{-\delta/2}\right) .$$

It follows from (5.6), (5.7), (5.8) that

$$\sqrt{N}\,\log a_N\left(\frac{1-F_{\bar{X}_n}(a_N)}{F_{\bar{X}_n}(a_N)}\right)^{\frac{1}{2}}\Big(\tilde{m}_N(a_N)-m_n(a_N)\Big)\,I\Big[0<\hat{F}^{(N)}_{\bar{X}_n}(a_N)<1\Big] =$$

$$= \left\{\frac{\mathcal{B}_N\big(F_{\bar{X}_n}(a_N)\big)}{\Big(\mathrm{var}\;\mathcal{B}_N\big(F_{\bar{X}_n}(a_N)\big)\Big)^{\frac{1}{2}}}\Big(1+\mathbf{o}_p(1)\Big)+\mathcal{O}_p\big(N^{-\delta/2}\big)\right\} I\Big[0<\hat{F}^{(N)}_{\bar{X}_n}(a_N)<1\Big] ,$$

hence

$$\lim_{N\to\infty}\mathbb{P}_m\left(\sqrt{N}\,\log a_N\left(\frac{1-F_{\bar{X}_n}(a_N)}{F_{\bar{X}_n}(a_N)}\right)^{\frac{1}{2}}\Big(\tilde{m}_N(a_N)-m_n(a_N)\Big)\le y\right) =$$

$$(5.9)\qquad = \lim_{N\to\infty}\mathbb{P}_m\left(\sqrt{N}\,\log a_N\left(\frac{1-F_{\bar{X}_n}(a_N)}{F_{\bar{X}_n}(a_N)}\right)^{\frac{1}{2}}\Big(\tilde{m}_N(a_N)-m_n(a_N)\Big)\le y\,,\; 0<\hat{F}^{(N)}_{\bar{X}_n}(a_N)<1\right) = \Phi(y) .$$

□

Proof of Theorem 3.2: By (3.10), $1-F_{\bar{X}_n}(x)=x^{-m}L^*(x)$ where L^* is slowly varying at ∞. Moreover, $\widehat{m}_N-m=(\widehat{m}_N-m_n(a_N))+(m_n(a_N)-m)$, while

$$\sqrt{N}\,\log a_N\left(\frac{1-F_{\bar{X}_n}(a_N)}{F_{\bar{X}_n}(a_N)}\right)^{\frac{1}{2}}\Big(\widehat{m}_N-m_n(a_N)\Big)$$

has asymptotically the standard normal distribution by Lemma 5.1. By (2.2), $m_n(a_N)-m\to 0$ as $N\to\infty$; more precisely,

$$\lim_{N\to\infty}\Big(m_n(a_N)-m\Big) = \lim_{N\to\infty}\left[\frac{-\log\Big(1-F_{\bar{X}_n}(a_N)\Big)}{\log a_N}-m\right]$$

$$(5.10)\qquad = \lim_{N\to\infty}\left[\frac{m\log a_N-\log\big(L^*(a_N)\big)}{\log a_N}-m\right]$$

$$= \lim_{N\to\infty}\frac{-\log\big(L^*(a_N)\big)}{\log a_N} = 0 ,$$

but the term

$$\sqrt{N}\left(\frac{1-F_{\bar{X}_n}(a_N)}{F_{\bar{X}_n}(a_N)}\right)^{\frac{1}{2}}\big|\log\big(L^*(a_N)\big)\big|$$

converges to 0 in only special cases and hence will create a bias; this, together with Lemma 5.1, implies the theorem. □

5.2. Strong consistency

We shall prove Theorem 3.1 with the aid of the following lemma.

Lemma 5.2. *Under the assumptions of Theorem 3.1,*

$$\Big(\tilde{m}_N(a_N)-m\Big)\, I\Big[0<\widehat{F}^{(N)}_{\bar{X}_n}(x)<1\Big] \;\to\; 0 \tag{5.11}$$

with probability 1, as $N\to\infty$.

Proof: Because $\tilde{m}_N(a_N)-m = [\tilde{m}_N(a_N)-m_n(a_N)]+[m_n(a_N)-m]$ and because of (2.4), it suffices to prove that

$$\Big(\tilde{m}_N(a_N)-m_n(a_N)\Big)\, I\Big[0<\widehat{F}^{(N)}_{\bar{X}_n}(x)<1\Big] \;\to\; 0 \tag{5.12}$$

with probability 1, as $N\to\infty$. Using (5.5) and (5.6), we obtain

$$\begin{aligned}
&\Big(\tilde{m}_N(a_N)-m_n(a_N)\Big)\, I\Big[0<\hat{F}^{(N)}_{\bar{X}_n}(a_N)<1\Big] \;=\\
&\quad=\;(\log a_N)^{-1}\left\{\left[1-\frac{1-\hat{F}^{(N)}_{\bar{X}_n}(a_N)}{1-F_{\bar{X}_n}(a_N)}\right]+\mathcal{O}\left(\left[1-\frac{1-\hat{F}^{(N)}_{\bar{X}_n}(a_N)}{1-F_{\bar{X}_n}(a_N)}\right]^2\right)\right\} \qquad (5.13)\\
&\quad=\;A_N^{(1)}+A_N^{(2)}
\end{aligned}$$

and, using again the strong embedding of empirical processes,

$$\begin{aligned}
A_N^{(1)} \;&=\; (\log a_N)^{-1}\Big(1-F_{\bar{X}_n}(a_N)\Big)^{-1}\\
&\qquad\cdot\;\Big[N^{-\frac{1}{2}}\,\mathcal{B}_N\big(F_{\bar{X}_n}(a_N)\big)+\mathcal{O}_{\text{a.s.}}\big(N^{-1}\log N\big)\Big]\\
&=\;(\log a_N)^{-1}N^{\frac{1}{2}\left(\frac{m}{m_0}-1\right)-\frac{m\delta}{2m_0}}\big(L^*(a_N)\big)^{-\frac{1}{2}}\,\frac{\mathcal{B}_N\big(F_{\bar{X}_n}(a_N)\big)}{\Big(F_{\bar{X}_n}(a_N)\Big(1-F_{\bar{X}_n}(a_N)\Big)\Big)^{\frac{1}{2}}} \qquad (5.14)\\
&\qquad+\;N^{\frac{m}{m_0}-1-\delta\frac{m}{m_0}}\big(L^*(a_N)\big)^{-1}\,\mathcal{O}_{\text{a.s.}}(1)\;.
\end{aligned}$$

The second term on the right-hand side of (5.14) converges to 0 almost surely as $N \to \infty$. The first term is normally distributed, hence, because $m \le m_0$, it holds for any $\varepsilon > 0$,

$$\sum_{N=1}^{\infty} \mathbb{P}_m \left((\log a_N)^{-1} N^{\frac{1}{2}\left(\frac{m}{m_0}-1\right)-\frac{m\delta}{2m_0}} \left(L^*(a_N)\right)^{-\frac{1}{2}} \frac{\left|\mathcal{B}_N\left(F_{\bar{X}_n}(a_N)\right)\right|}{\left(F_{\bar{X}_n}(a_N)\left(1-F_{\bar{X}_n}(a_N)\right)\right)^{\frac{1}{2}}} > \varepsilon \right) \le$$

$$\le\; 2 \sum_{N=1}^{\infty} \left[1 - \Phi\left(\varepsilon \,\frac{1-\delta}{m_0}\, N^{\frac{m\delta}{2m_0}} \log N \left(L_1^*(N)\right)^{\frac{1}{2}} \right) \right] . \tag{5.15}$$

Using the inequality $1-\Phi(x) \le \frac{1}{x\sqrt{2\pi}}\, \mathrm{e}^{-\frac{x^2}{2}},\ x>0$ in (5.15), we obtain

$$\sum_{N=1}^{\infty} \frac{1}{N^{\frac{m\delta}{2m_0}} \log N} \exp\left\{ -\frac{\varepsilon^2}{2} \left(\frac{1-\delta}{m_0}\right)^2 (\log N)^2\, N^{\frac{m\delta}{m_0}} L_1^*(N) \right\} \le$$

$$\le\; K_1 \sum_{N=1}^{\infty} \exp\left\{-K_2\, N^{\kappa}\right\} \;<\; \infty$$

where $K_1, K_2, \kappa > 0$ are constants, and by the Borel–Cantelli lemma we conclude that the first term on the right-hand side of (5.14) also converges to 0 almost surely as $N\to\infty$. Similarly we prove that $A_N^{(2)} = \mathbf{o}(1)$ *a.s.* as $N\to\infty$. This proves (5.12) and, in turn, (5.11). ☐

Proof of Theorem 3.1: For any $\varepsilon > 0$, it holds

$$\sum_{N=1}^{\infty} \mathbb{P}_m\left(\left|\widehat{m}_N(a_N) - m\right| > \varepsilon\right) \le$$

$$\le \sum_{N=1}^{\infty} \mathbb{P}_m\left(\left|\tilde{m}_N(a_N) - m\right| I\left[0<\hat{F}_{\bar{X}_n}^{(N)}(a_N)<1\right] > \frac{\varepsilon}{2}\right) + \sum_{N=1}^{\infty} \mathbb{P}_m\left((m_0-m)\, I\left[\hat{F}_{\bar{X}_n}^{(N)}(a_N)=0 \text{ or } 1\right] > \frac{\varepsilon}{2}\right) . \tag{5.16}$$

The convergence of the first series on the right-hand side of (5.16) follows from Lemma 5.2. The sum of the second series is bounded from above by

$$\sum_{N=1}^{\infty} \mathbb{P}_m\left((m_0-m)\, I\left[\hat{F}_{\bar{X}_n}^{(N)}(a_N)=0\right] > \frac{\varepsilon}{4}\right) + \sum_{N=1}^{\infty} \mathbb{P}_m\left((m_0-m)\, I\left[\hat{F}_{\bar{X}_n}^{(N)}(a_N)=1\right] > \frac{\varepsilon}{4}\right) \le$$

$$\le \sum_{N=1}^{\infty}\left(1 - F_{\bar{X}_n}(a_N)\right)^N + \sum_{N=1}^{\infty}\left(F_{\bar{X}_n}(a_N)\right)^N \tag{5.17}$$

$$= \sum_{N=1}^{\infty}\left(a_N^{-m} L^*(a_N)\right)^N + \sum_{N=1}^{\infty}\left(1 - a_N^{-m} L^*(a_N)\right)^N .$$

Because $a_N^{-\eta} < L^*(a_N) < a_N^{\eta}$ for $N > N_\eta$ and $\forall\, \eta > 0$, we conclude

$$(5.18) \qquad \left(a_N^{-m} L^*(a_N)\right)^N < \left(a_N^{-m+\eta}\right)^N \le \left(\frac{1}{N^{\frac{1-\delta}{m_0}(m-\eta)}}\right)^N$$

hence the first series on the right-hand side of (5.17) converges for sufficiently small η. Similarly,

$$(5.19) \qquad \begin{aligned} \left(1 - a_N^{-m} L^*(a_N)\right)^N &< \left(1 - a_N^{-m-\eta}\right)^N \\ &\le \left(1 - \frac{1}{N^{\frac{1-\delta}{m_0}(m+\eta)}}\right)^N \\ &\le \left[\exp\left\{-\left(N^{\frac{1-\delta}{m_0}(m+\eta)}\right)^{-1}\right\}\right]^N , \end{aligned}$$

what implies the convergence of the second series on the right-hand side of (5.17).
□

6. TEST ON THE PARETO INDEX

We shall now briefly describe one possible test of the hypothesis on the Pareto index, based on the sample means. For other tests we refer to [12] and [14].

Because the problem is of semiparametric nature, we should first think over a proper formulation of the hypothesis. Following [12], we shall consider the hypothesis

$$(6.1) \qquad \mathbf{H}_{m_0} : \qquad x^{m_0}\left(1 - F(x)\right) \ge 1 \quad \forall\, x > x_0$$

with a hypothetical $m_0 > 0$ and with some $x_0 \ge 0$. Such hypothesis and hence the test are nonparametric; the test is based on splitting the set of observations into N subsamples of sizes n and on the empirical distribution function of the means of the subsamples; the asymptotics is for $N \to \infty$ and fixed n (eventually small), and the asymptotic null distribution of the test criterion is normal. The proposed test is consistent against exponentially tailed alternatives, as well as against heavy tailed alternatives with index $m > m_0$. The test is asymptotically unbiased for the broad family of distributions represented by $\mathbf{H}_{m_0}$ and its alternative. Such test may be used as a supplement to the usual tests of the Gumbel hypothesis $m = \infty$ against $m < \infty$, namely in the situation that the latter tests reject the hypothesis of exponentiality.

Similarly as in the estimation, we partition the set of observations into N non-overlapping samples of the same sizes n, denoted as

$$(6.2) \qquad \left(X_1^{(1)},...,X_n^{(1)}\right), \quad ..., \quad \left(X_1^{(N)},...,X_n^{(N)}\right)$$

and denote $\bar{X}_n^{(1)},...,\bar{X}_n^{(N)}$ the respective sample means. Let $F_{\bar{X}_n}(x)=\mathbb{P}_m(\bar{X}_n\leq x)$ be the common distribution function of the sample means and let $\widehat{F}_{\bar{X}_n}^{(N)}(x)=\frac{1}{N}\sum_{j=1}^N I[\bar{X}_n^{(j)}\leq x]$ be the corresponding empirical distribution function. Let

$$(6.3) \qquad a_N = N^{(1-\delta)/m_0}\,, \qquad 0<\delta<1\,.$$

We propose the test rejecting $\mathbf{H}_{m_0}$ if

$$\begin{aligned} &\text{either} \quad \hat{F}_{\bar{X}_n}^{(N)}(a_n)=1\\ &\text{or} \quad \hat{F}_{\bar{X}_n}^{(N)}(a_n)<1 \quad \text{and simultaneously}\\ (6.4) \qquad &N^{\delta/2}\left[-\log\left(1-\hat{F}_{\bar{X}_n}^{(N)}(a_n)\right)-(1-\delta)\log N\right] \geq \Phi^{-1}(1-\alpha) \end{aligned}$$

where Φ is the standard normal distribution function. If F satisfies (6.1) as an equality, as the Pareto distribution, then α is the asymptotic probability of the error of the first kind; for any other distribution satisfying (6.1), the asymptotic probability of the error of the first kind it is $\leq\alpha$.

The asymptotic null distribution of the test is described in the following theorem:

Theorem 6.1. *Let $X_1,X_2,...$ be independent observations, identically distributed according to absolutely continuous distribution function F satisfying (6.1). Let $\hat{F}_{\bar{X}_n}^{(N)}$ be the empirical distribution function of the means of samples (6.2). Then*

$$(6.5) \qquad \lim_{N\to\infty}\mathbb{P}_{m_0}\left(\hat{F}_{\bar{X}_n}^{(N)}(a_N)<1\right) = 1$$

with a_N defined in (6.3), and

$$(6.6) \qquad \lim_{N\to\infty}\mathbb{P}_{m_0}\left\{N^{\delta/2}\left[-\log\left(1-\hat{F}_{\bar{X}_n}^{(N)}(a_N)\right)-(1-\delta)\log N\right]\geq\tau_\alpha\,,\ \hat{F}_{\bar{X}_n}^{(N)}(a_N)<1\right\}\leq\alpha\,,$$

where $\tau_\alpha=\Phi^{-1}(1-\alpha)$, $0<\alpha<1$, and Φ is the standard normal distribution function.

Moreover, if there exists x_0 *such that*

$$x^{m_0}\Big(1 - F(x)\Big) = 1 \qquad \textit{for} \quad x > x_0 \ , \tag{6.7}$$

then

$$\lim_{N\to\infty} \mathbb{P}_{m_0}\left\{N^{\delta/2}\Big[-\log\Big(1-\hat{F}^{(N)}_{\bar{X}_n}(a_N)\Big) - (1-\delta)\log N\Big] \ge \tau_\alpha\,,\ \hat{F}^{(N)}_{\bar{X}_n}(a_N) < 1\right\} = \alpha\,. \tag{6.8}$$

Proof: First, (6.5) follows from Lemma 3.1. Further, (6.8) follows from the proof of Lemma 5.1, namely from (5.9), where we insert the pertinent expressions for $m_n(a)$ and a_N according to (2.4), (5.1) and (6.3), respectively.

If F satisfies (6.1), then the right tail of $1-\hat{F}^{(N)}_{\bar{X}_n}$ is ultimately not smaller stochastically than that of F satisfying (6.7); this implies (6.6). □

The following Corollary shows the set of alternatives against which is the test asymptotically unbiased:

Corollary 6.1.

(i) Under the conditions of Theorem 6.1, the test with the critical region (6.4) is asymptotically unbiased for the hypothesis $\mathbf{H}_{m_0}$ *against the alternative*

$$x^{m_0}\Big(1 - F(x)\Big) < 1 \qquad \textit{for} \quad x > x_0 \ . \tag{6.9}$$

(ii) The test attains the asymptotic power 1 against the alternative that F *is of type (1.1) with index* $m > m_0$*, including* $m = \infty$*.*

Proof: Under $\mathbf{H}_{m_0}$, (6.6) holds by Theorem 6.1, hence the asymptotic size of the test is equal to α for the whole hypothesis $\mathbf{H}_{m_0}$.

Under (6.9), $1-\hat{F}^{(N)}_{\bar{X}_n}$ is ultimately stochastically smaller than under Pareto with index m_0, hence

$$\lim_{N\to\infty} \mathbb{P}_{m_0}\left\{N^{\delta/2}\Big[-\log\Big(1-\hat{F}^{(N)}_{\bar{X}_n}(a_N)\Big) - (1-\delta)\log N\Big] \ge \tau_\alpha\,,\ \hat{F}^{(N)}_{\bar{X}_n}(a_N) < 1\right\} \ \ge\ \alpha\,.$$

This proves the asymptotic unbiasedness.

Let now F satisfy (1.1) with index $m > m_0$. Then $F_{\bar{X}_n}$ also satisfies (1.1) and it implies that, given $\varepsilon > 0$, there exists N_0 such that, for $N > N_0$,

$$N^{-\frac{m}{m_0}(1+\varepsilon)(1-\delta)} \;=\; a_N^{-m(1+\varepsilon)} \;\le\; 1 - F_{\bar{X}_n}(a_N) \;\le\; a_N^{-m(1-\varepsilon)} \;=\; N^{-\frac{m}{m_0}(1-\varepsilon)(1-\delta)}\ .$$

If $1-\hat{F}^{(N)}_{\bar{X}_n}(a_N)=0$, we reject $\mathbf{H}_{m_0}$. If $1-\hat{F}^{(N)}_{\bar{X}_n}(a_N)>0$, then

$$\mathbb{P}_{m_0}\left\{N^{\delta/2}\left[-\log\left(1-\hat{F}^{(N)}_{\bar{X}_n}(a_N)\right)-(1-\delta)\log N\right]\geq\tau_\alpha\right\} = \tag{6.10}$$

$$= \mathbb{P}_m\left\{N^{\delta/2}\left[-\log\left(\frac{1-\hat{F}^{(N)}_{\bar{X}_n}(a_N)}{1-F_{\bar{X}_n}(a_N)}\right)-\log\left(1-F_{\bar{X}_n}(a_N)\right)-(1-\delta)\log N\right]\geq\tau_\alpha\right\}$$

$$\to 1 \qquad \text{as} \quad N\to\infty\ ,$$

because the first term of the argument on the right-hand side of (6.10) is stochastically bounded under index m (cf. the proof of Lemma 3.1), while the second term tends to infinity for $m>m_0$. Hence, we reject $\mathbf{H}_{m_0}$ with probability tending to 1. The case $m=\infty$ corresponds to the exponential tail. □

The performance of the test procedure for different choices of m_0 is illustrated again on the simulated random samples. The replications ($N=200$) of samples of sizes $n=5$ were simulated 1000 times. Fig. 3–5 show the number of rejection of the null hypothesis $\mathbf{H}_{m_0}$ as a function of m_0 for Pareto ($m=1$), Burr ($m=2$) and generalized Pareto ($m=0.5$) distributions with $\delta=0.1,\ 0.5$ on the level $\alpha=0.01,\ 0.05$.

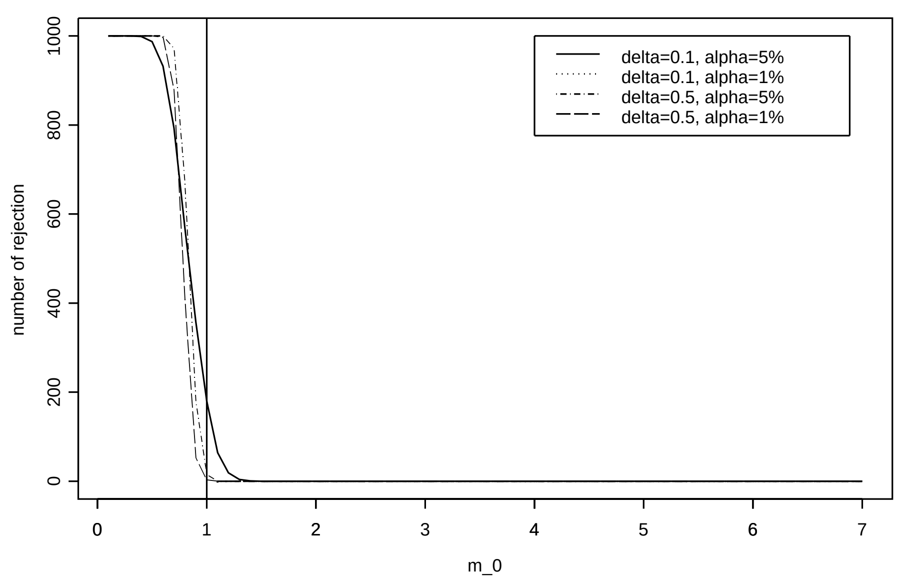

Figure 3: The number of rejection of $\mathbf{H}_{m_0}$ as a function of m_0 for Pareto distribution with $m=1$.

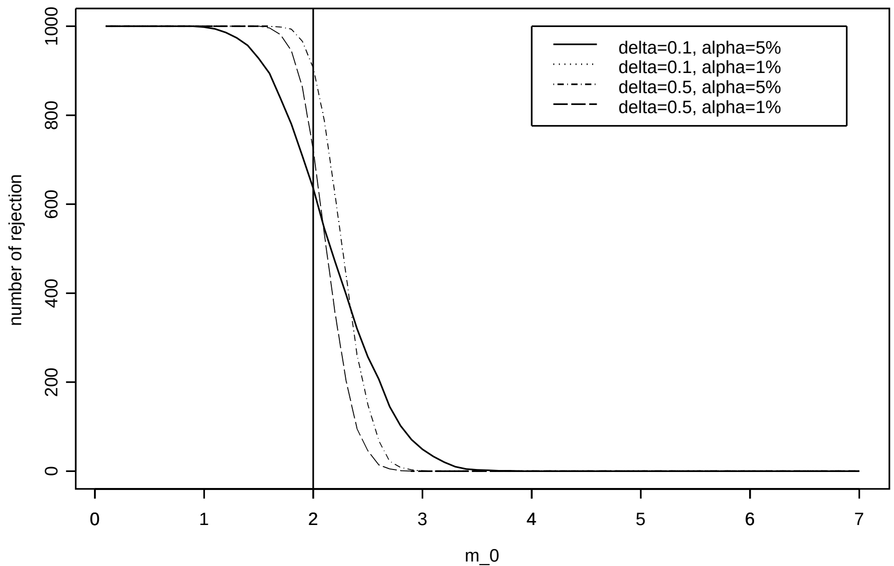

Figure 4: The number of rejection of $\mathbf{H}_{m_0}$ as a function of m_0 for Burr distribution with $m = 2$.

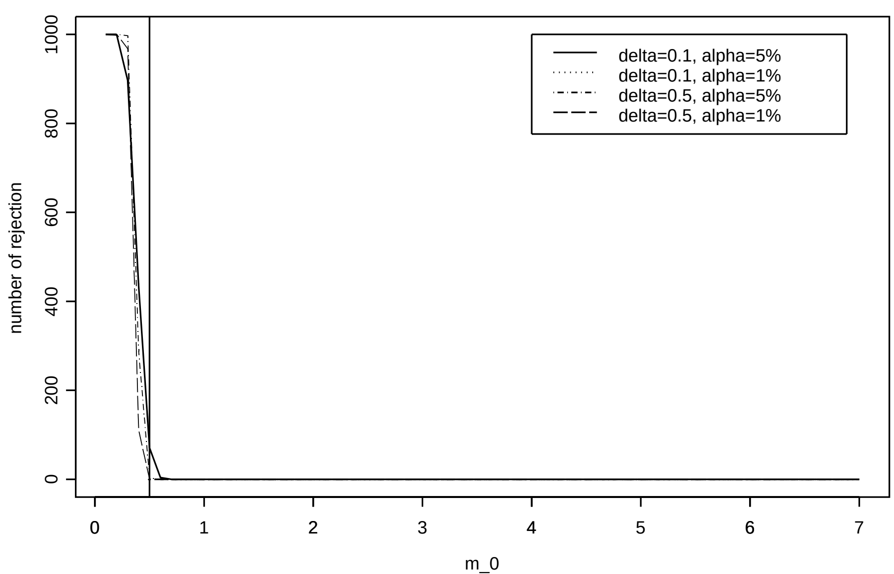

Figure 5: The number of rejection of $\mathbf{H}_{m_0}$ as a function of m_0 for Generalized Pareto distribution with $m = 0.5$.

ACKNOWLEDGMENTS

The authors thank two referees for their valuable comments and suggestions, which led to a better readability of the text and to further considerations.

The research of the first and the second authors was supported by the Research Project CEZ: MSM 113200008 J13/98:113200008 "Mathematical Methods in Stochastics", the third author was supported by postdoctoral grant from the Portuguese Fundação para a Ciência e a Tecnologia (FCT) at the University of Lisbon (the Centro de Estatística e Aplicações). The second author was also supported by the Czech Republic Grant 201/02/0621 and the third by GAČR 205/02/0871.

REFERENCES

[1] Beirlant, J.; Teugels, J.L. and Vynckier, P. (1996). *Practical Analysis of Extreme Values*, Leuven University Press.

[2] Beirlant, J.; Dierckx, G.; Goegebeur, Y. and Matthys, G. (1999). Tail index estimation and an exponential regression model, *Extremes* **2**, 177–200.

[3] Dekkers, A.L.M.; Einmahl, J.H.J. and de Haan, L. (1989). A moment estimator for the index of the extreme-value distribution, *Ann. Statist.*, **17**, 1833–1855.

[4] Embrechts, P.; Klüppelberg, C. and Mikosch, T. (1997). *Modelling of Extremal Events in Insurance and Finance*, Springer Verlag.

[5] Fialová, A. (2001). *Estimators and tests of Pareto index* (in Czech), PhD Thesis, Charles University, Prague.

[6] Fialová, A.; Jurečková, J. and Picek, J. (2002). *New type of estimator of Pareto index*, Preprint #28, Dept. of Statistics, Charles University, Prague.

[7] Gomes, M.I. and Martins, M.J. (2002). Asymptotically unbiased estimators of the tail index based on external estimation of the second order parameter, *Extremes*, **5**(1), 5–31.

[8] Feuerverger, A. and Hall, P. (1999). Estimating a tail exponent by modelling departure from a Pareto distribution, *Ann. Statist.*, **27**, 760–781.

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

[10] Jurečková, J. (1981). Tail-behavior of location estimators, *Ann. Statist.*, **9**, 578–585.

[11] Jurečková, J. (2000). Tests of tails based on extreme regression quantiles, *Statist. & Probab. Letters* **49**, 53–61.

[12] Jurečková, J. and Picek, J. (2001). A class of tests on the tail index, *Extremes*, **4**(2), 165–183.

[13] Komlós, J.; Major, P. and Tusnády, G. (1975). An approximation of partial sums of independent R.V.'s and the sample DF, *Y. Wahrscheinlichkeitstheorie verw. Geb.*, **32**, 111–131.

[14] Picek, J. and Jurečková, J. (2001). *A class of tests on the tail index using the modified extreme regression quantiles.* In "ROBUST'2000" (J. Antoch, G. Dohnal, Eds.), Union of Czech Mathematicians and Physicists, Prague, 217–226.

[15] Pickands, J. (1975). Statistical inference using extreme order statistics, *Ann. Statist.*, **3**, 119–131.

[16] Smith, R.L. (1987). Estimating tails of probability distribution, *Ann. Statist.*, **15**(3), 1174–1207.