Hindawi Publishing Corporation
Journal of Inequalities and Applications
Volume 2010, Article ID 209156, 8 pages
doi:10.1155/2010/209156

Research Article

Strong Convergence Bound of the Pareto Index Estimator under Right Censoring

Bao Tao[1] and Zuoxiang Peng[2]

[1] *College of Mathematics and Statistics, Chongqing Technology and Business University, Chongqing 400067, China*
[2] *School of Mathematics and Statistics, Southwest University, Chongqing 400715, China*

Correspondence should be addressed to Bao Tao, taobao@ctbu.edu.cn

Received 25 October 2009; Revised 14 March 2010; Accepted 11 April 2010

Academic Editor: Siegfried Carl

Let $\{X_n, n \geq 1\}$ be a sequence of positive independent and identically distributed random variables with common Pareto-type distribution function $F(x) = 1 - x^{-1/\gamma} l_F(x)$ as $\gamma > 0$, where $l_F(x)$ represents a slowly varying function at infinity. In this note we study the strong convergence bound of a kind of right censored Pareto index estimator under second-order regularly varying conditions.

1. Introduction

A distribution F is said to be of Pareto-type if there exists a positive constant γ such that

$$1 - F(x) = x^{-1/\gamma} l_F(x), \tag{1.1}$$

where $l_F(x)$ is a slowly varying function at infinity, that is,

$$\lim_{t \to \infty} \frac{l_F(tx)}{l_F(t)} = 1 \quad \forall x > 0. \tag{1.2}$$

The parameter γ is called the Pareto index.

Estimating the Pareto index γ is very important in theoretical analysis and practical applications of extreme value theory; for example, Embrechts et al. [1], Reiss and Thomas [2] and references therein. For recent work on estimating extreme value index, see Beirlant and

Guillou [3], Fraga Alves [4, 5], Gomes et al. [6], Gomes and Henriques Rodrigues [7], and Li et al. [8, 9].

Suppose that $\{X_n, n \geq 1\}$ is a sequence of positive independent and identically distributed (i.i.d.) random variables with common distribution function F, and let $X_{1,n} \leq X_{2,n} \leq \cdots \leq X_{n,n}$ denote the order statistics of $X_1, X_2, \ldots, X_n$. By using maximum likelihood method, Hill [10] introduced the following well-known estimator of γ, that is,

$$H_{k,n} = \frac{1}{k}\sum_{j=1}^{k} \log X_{n-j+1,n} - \log X_{n-k,n}. \tag{1.3}$$

Mason [11] proved weak consistency of $H_{k,n}$ for any sequence $k = k(n) \to \infty$, $k(n)/n \to 0(n \to \infty)$ and Deheuvels et al. [12] proved its strong consistency for any sequence $k(n)$ with $k(n)/n \to 0, k(n)/\log\log n \to \infty(n \to \infty)$. The strong convergence bounds of the Hill estimator $H_{k,n}$ have been considered by Peng and Nadarajah [13] under second-order regularly varying conditions.

Recently Beirlant and Guillou [3] proposed a new kind of Hill estimator in case of the sample being right censored. In actuarial setting, most insurance policies limit the claim size and reinsure the large claim size exceeding the given level. So the observed claim size series are right censored. Suppose that $X_{1,n} \leq X_{2,n} \leq \cdots \leq X_{N,n}$ are the first N $(= N(n))$ ascending order statistics of $X_1, X_2, \ldots, X_n$, where N is an integer random variable. Define

$$\gamma(k,n) = \frac{1}{k-n+N}\left\{\sum_{j=n-N+1}^{k} \log\frac{X_{n-j+1,n}}{X_{n-k,n}} + (n-N)\log\frac{X_{N,n}}{X_{n-k,n}}\right\} \tag{1.4}$$

as the estimator of γ, where $k = n-N+1, \ldots, n-1$. This estimator reduces to the Hill estimator in the absence of censoring. Beirlant and Guillou [3] proved weak and strong consistency, and asymptotic normality of $\gamma(k,n)$. In this paper, we consider the strong convergence bound of this Pareto index estimator $\gamma(k,n)$ under second-order regularly varying conditions.

2. Main Results

Denote $U(x) = (1/(1-F))^{\leftarrow}(x) = \inf\{t : F(t) \geq 1 - 1/x\}$ $(x > 1)$; then (1.1) is equivalent to

$$\lim_{t\to\infty}\frac{U(tx)}{U(t)} = x^{\gamma}. \tag{2.1}$$

In order to investigate the strong convergence bound of $\gamma(k,n)$, we require knowing the convergence rate of (2.1). For this reason, we need the following second-order regular condition.

Suppose that there exists a measurable function $A(t)$ satisfying $\lim_{t\to\infty}A(t) = 0$, and a function $H(x) \neq cx^{\gamma}$ such that for all $x > 0$,

$$\lim_{t\to\infty}\frac{1}{A(t)}\left\{\frac{U(tx)}{U(t)} - x^{\gamma}\right\} = H(x). \tag{2.2}$$

Then $H(x)$ must be of the form $x^{\gamma}((x^{\rho}-1)/\rho)$ for some $\rho \le 0$ ($x^{\gamma}((x^{\rho}-1)/\rho) =: x^{\gamma}\log x$ as $\rho = 0$), and ρ is the regularly varying index of $A(t)$, that is, $A(t) \in RV_{\rho}$; for example, de Haan and Stadtmüller [14]. (2.2) holds if and only if for all $x > 0$,

$$\lim_{t\to\infty}\frac{\log U(tx)-\log U(t)-\gamma\log x}{A(t)}=\frac{x^{\rho}-1}{\rho}. \tag{2.3}$$

In order to obtain the strong convergence bound of $\gamma(k,n)$, we give the following two results firstly.

Theorem 2.1. *Suppose that* (2.2) *holds, and further assume that* $k/n \sim p_n \downarrow 0$, $k/(\log n)^{\delta} \to \infty$ *for some* $\delta > 0$, $\sqrt{k/(2\log\log n)}A(n/k) \to \beta \in [0,\infty)$, *and* $(n-N)/\sqrt{2k\log\log n} \to 0$ *a.s. as* $n \to \infty$, *where* $0 < N < n$. *Then*

$$\limsup_{n\to\infty}\frac{1}{\sqrt{2k\log\log n}}\left|\sum_{j=n-N+1}^{k}\left(\log X_{n-j+1,n}-\log X_{n-k,n}-\gamma\right)\right| \le \left(\sqrt{2}+1\right)\gamma+\frac{\beta}{1-\rho} \quad a.s. \tag{2.4}$$

Theorem 2.2. *Suppose that* (2.2) *holds, and further assume that* $k/n \sim p_n \downarrow 0$, $(n-N)/\log\log n \uparrow \infty$, $\sqrt{k/(2\log\log n)}A(n/k) \to \beta \in [0,\infty)$, *and* $(n-N)(2k\log\log n)^{-1/2}\log[(n-N)/k] \to 0$ *a.s. as* $n \to \infty$, *where* $0 < N < n$. *Then*

$$\limsup_{n\to\infty}\frac{n-N}{\sqrt{2k\log\log n}}\left|\log X_{N,n}-\log X_{n-k,n}\right| = 0 \quad a.s. \tag{2.5}$$

By using Theorems 2.1 and 2.2, we can deduce the following theorem easily.

Theorem 2.3. *Suppose that* (2.2) *holds and assume that* $k/n \sim p_n \downarrow 0$, $\sqrt{k/(2\log\log n)}\ A(n/k) \to \beta \in [0,\infty)$, $k/(\log n)^{\delta} \to \infty$ *for some* $\delta > 0$, $(n-N)(2k\log\log n)^{-1/2}\log[(n-N)/k] \to 0$, *and* $(n-N)(\log\log n)^{-1} \uparrow \infty$ *a.s. as* $n \to \infty$ *where* $0 < N < n$; *then*

$$\limsup_{n\to\infty}\sqrt{\frac{k}{2\log\log n}}\left|\gamma(k,n)-\gamma\right| \le \left(\sqrt{2}+1\right)\gamma+\frac{\beta}{1-\rho} \quad a.s. \tag{2.6}$$

3. Proofs

Suppose that $Y_1, Y_2, \ldots, Y_n$ are independent and identically distributed random variables with common distribution function $P(Y_1 \le x) = 1 - 1/x$ $(x \ge 1)$. Let $Y_{1,n} \le Y_{2,n} \le \cdots \le Y_{n,n}$ denote the order statistics of $Y_1, Y_2, \ldots, Y_n$. It is easy to see that $(U(Y_1), U(Y_2), \ldots) \stackrel{d}{=} (X_1, X_2, \ldots)$. For the sake of simplicity, define $\gamma(k,n)$ as

$$\gamma(k,n)=\frac{1}{k-n+N}\left\{\sum_{j=n-N+1}^{k}\log\frac{U(Y_{n-j+1,n})}{U(Y_{n-k,n})}+(n-N)\log\frac{U(Y_{N,n})}{U(Y_{n-k,n})}\right\}. \tag{3.1}$$

The following auxiliary results are necessary for the proofs of the main results. The first two results are correct due to Wellner [15].

Lemma 3.1. *If* $k/n \to 0$ *and* $k/\log\log n \to \infty$, *then*

$$\lim_{n\to\infty} \frac{k}{n} Y_{n-k,n} = 1 \quad a.s. \tag{3.2}$$

Lemma 3.2. *Suppose that* $k \to \infty$, $k/n \sim p_n \downarrow 0$, *and* $np_n/\log\log n \to \infty$. *Then*

$$\limsup_{n\to\infty} \frac{k\left|\log n - \log k - \log Y_{n-k,n}\right|}{\sqrt{2k \log\log n}} = 1 \quad a.s. \tag{3.3}$$

Proof. Applying Lemma 3.1, we find

$$k\left(\log n - \log k - \log Y_{n-k,n}\right) = \left(\frac{n}{Y_{n-k,n}} - k\right)(1 + o(1)) \quad \text{a.s.} \tag{3.4}$$

Notice that $1/Y_i$ is uniformly distributed on $(0,1)$; the result follows from Wellner [15]. □

Lemma 3.3. *Let* $k \to \infty$, $k/n \sim p_n \downarrow 0$ *and* $k/\log\log n \to \infty$. *Then*

$$\limsup_{n\to\infty} \frac{\left|\sum_{i=1}^{k} \log Y_{n-i+1,n} - k(\log n - \log k + 1)\right|}{\sqrt{2k \log\log n}} = \sqrt{2} \quad a.s. \tag{3.5}$$

Proof. The result follows from Deheuvels and Mason [16]. □

The following bound of (2.3) is from Drees [17]; for example, Theorem B.2.18 of de Haan and Ferreira [18].

Lemma 3.4. *If* (2.2) *holds, then for every* $\epsilon > 0$, *there exists* $t_0 > 0$ *such that for* $t \ge t_0$ *and* $x \ge 1$,

$$\left|\frac{\log U(tx) - \log U(t) - \gamma \log x}{A(t)} - \frac{x^{\rho} - 1}{\rho}\right| \le \epsilon x^{\rho+\epsilon}. \tag{3.6}$$

Lemma 3.5. *Assume that* $k/n \to 0$, $k/(\log n)^{\delta} \to \infty$ *for some* $\delta > 0$ *and* $(n-N)/k \to 0$ *a.s. as* $n \to \infty$ *where* $0 < N < n$. *Suppose* $\xi_{1,n} \le \xi_{2,n} \le \cdots \le \xi_{n,n}$ *are order statistics from parent* ξ *with distribution function* $F(x) = x^{\alpha}(0 < x < 1)$ *for some* $\alpha > 0$. *Then*

$$\lim_{n\to\infty} \frac{1}{k-n+N} \sum_{j=n-N+1}^{k} \frac{\xi_{j,n}}{\xi_{k+1,n}} = \frac{\alpha}{\alpha+1} \quad a.s. \tag{3.7}$$

Proof. The proof is similar to the proof of Lemma 2.3(i) in Dekkers et al. [19]; for example, Lemma 1 of Beirlant and Guillou [3]. □

Based on the above lemmas, we prove Theorems 2.1 and 2.2.

Proof of Theorem 2.1. We only prove the case for $\rho < 0$. For the case for $\rho = 0$, the proof is similar. Clearly,

$$\begin{aligned}\sum_{j=n-N+1}^{k} &[\log U(Y_{n-j+1,n}) - \log U(Y_{n-k,n}) - \gamma] \\ &= \sum_{j=n-N+1}^{k} D_j(n)A(Y_{n-k,n}) + \frac{A(Y_{n-k,n})}{\rho}\sum_{j=n-N+1}^{k}\left[\left(\frac{Y_{n-j+1,n}}{Y_{n-k,n}}\right)^{\rho} - 1\right] \\ &\quad + \gamma\sum_{j=n-N+1}^{k}(\log Y_{n-j+1,n} - \log Y_{n-k,n} - 1),\end{aligned} \tag{3.8}$$

where

$$\begin{aligned}D_j(n) = &\frac{\log U(Y_{n-j+1,n}) - \log U(Y_{n-k,n}) - \gamma\log(Y_{n-j+1,n}/Y_{n-k,n})}{A(Y_{n-k,n})} \\ &- \frac{(Y_{n-j+1,n}/Y_{n-k,n})^{\rho} - 1}{\rho}\end{aligned} \tag{3.9}$$

for $j = n - N + 1, \ldots, k$. By using Lemmas 3.1 and 3.4, for sufficiently large n, we have

$$\left|\sum_{j=n-N+1}^{k} D_j(n)\right| \le \epsilon\sum_{j=n-N+1}^{k}\left(\frac{Y_{n-j+1,n}}{Y_{n-k,n}}\right)^{\rho+\epsilon}. \tag{3.10}$$

Noting that $P(Y_i^{\rho+\varepsilon} \le x) = x^{-1/(\rho+\varepsilon)}$ for $i = 1, \ldots, n$ and by using Lemma 3.5 we find

$$\lim_{n\to\infty}\frac{1}{k}\sum_{j=n-N+1}^{k}\left(\frac{Y_{n-j+1,n}}{Y_{n-k,n}}\right)^{\rho+\epsilon} = \frac{1}{1-\rho-\epsilon} \quad \text{a.s.} \tag{3.11}$$

Since $A(t) \in RV_{\rho}$ and $\sqrt{k/2\log\log n}A(n/k) \to \beta$, by (3.10), and (3.11), and Lemma 3.1, it follows that

$$\lim_{n\to\infty}\frac{1}{\sqrt{2k\log\log n}}\left|\sum_{j=n-N+1}^{k} D_j(n)A(Y_{n-k,n})\right| = 0 \quad \text{a.s.} \tag{3.12}$$

by letting $\epsilon \to 0$. Similarly,

$$\lim_{n\to\infty}\frac{1}{\sqrt{2k\log\log n}}\frac{A(Y_{n-k,n})}{\rho}\sum_{j=n-N+1}^{k}\left[\left(\frac{Y_{n-j+1,n}}{Y_{n-k,n}}\right)^{\rho} - 1\right] = \frac{\beta}{1-\rho} \quad \text{a.s.} \tag{3.13}$$

From Lemmas 3.2, and 3.3, and the conditions provided by Theorem 2.1, we have

$$\begin{aligned}
&\limsup_{n\to\infty}\frac{\gamma}{\sqrt{2k\log\log n}}\left|\sum_{j=n-N+1}^{k}\left(\log Y_{n-j+1,n}-\log Y_{n-k,n}-1\right)\right| \\
&\quad\le \limsup_{n\to\infty}\frac{\gamma}{\sqrt{2k\log\log n}}\left|\sum_{j=n-N+1}^{k}\log Y_{n-j+1,n}-(k-n+N)\left[\log n-\log(k-n+N)+1\right]\right| \\
&\qquad+\limsup_{n\to\infty}\frac{\gamma}{\sqrt{2k\log\log n}}(k-n+N)\left|\log n-\log k-\log Y_{n-k,n}\right| \\
&\qquad+\limsup_{n\to\infty}\frac{\gamma}{\sqrt{2k\log\log n}}(k-n+N)\left|\log k-\log(k-n+N)\right| \\
&\quad\le\left(\sqrt{2}+1\right)\gamma \quad \text{a.s.}
\end{aligned} \tag{3.14}$$

Combining (3.12), (3.13) with (3.14), we complete the proof. □

Proof of Theorem 2.2. We only prove the case of $\rho<0$. Clearly,

$$\begin{aligned}
&\log U(Y_{N,n})-\log U(Y_{n-k,n}) \\
&\quad=G(n)A(Y_{n-k,n})+A(Y_{n-k,n})\frac{(Y_{N,n}/Y_{n-k,n})^{\rho}-1}{\rho}+\gamma\left(\log Y_{N,n}-\log Y_{n-k,n}\right),
\end{aligned} \tag{3.15}$$

where

$$G(n)=\frac{\log U(Y_{N,n})-\log U(Y_{n-k,n})-\gamma\log(Y_{N,n}/Y_{n-k,n})}{A(Y_{n-k,n})}-\frac{(Y_{N,n}/Y_{n-k,n})^{\rho}-1}{\rho}. \tag{3.16}$$

By using Lemmas 3.1 and 3.4, for sufficiently large n, we have

$$|G(n)|\le\epsilon\left(\frac{Y_{N,n}}{Y_{n-k,n}}\right)^{\rho+\epsilon}. \tag{3.17}$$

Since $A(t)\in RV_{\rho}$, by using Lemma 3.1 and the conditions provided by this theorem, we have

$$\limsup_{n\to\infty}\frac{n-N}{\sqrt{2k\log\log n}}\left|G(n)A(Y_{n-k,n})\right|=0 \quad \text{a.s.} \tag{3.18}$$

Similarly,

$$\limsup_{n\to\infty}\frac{n-N}{\sqrt{2k\log\log n}}\left|A(Y_{n-k,n})\frac{(Y_{N,n}/Y_{n-k,n})^{\rho}-1}{\rho}\right|=0 \quad \text{a.s.} \tag{3.19}$$

By using Lemma 3.2 and the conditions of Theorem 2.2, it follows that

$$\begin{aligned}
&\limsup_{n\to\infty}\frac{n-N}{\sqrt{2k\log\log n}}\left|\gamma(\log Y_{N,n}-\log Y_{n-k,n})\right| \\
&\quad\le \limsup_{n\to\infty}\frac{n-N}{\sqrt{2k\log\log n}}\gamma\left|\log n-\log k-\log Y_{n-k,n}\right| \\
&\qquad+\limsup_{n\to\infty}\frac{n-N}{\sqrt{2k\log\log n}}\gamma\left|\log n-\log(n-N)-\log Y_{N,n}\right| \\
&\qquad+\limsup_{n\to\infty}\frac{n-N}{\sqrt{2k\log\log n}}\gamma\left|\log k-\log(n-N)\right|=0 \quad \text{a.s.}
\end{aligned} \tag{3.20}$$

Combining (3.18), (3.19) with (3.20), we complete the proof. □

Acknowledgments

The authors would like to sincerely thank the Editor-in-Chief and the anonymous referees for their valuable comments and suggestions on a previous draft, which resulted in the present version of the paper. This paper was partially supported by Chongqing Municipal Education Commission Projects (No. KJ100726, KJ080725) and Chongqing NSF (No. 2009BB8221).

References

[1] P. Embrechts, C. Klüppelberg, and T. Mikosch, *Modelling Extremal Events: For Insurance and Finance*, vol. 33 of *Applications of Mathematics*, Springer, Berlin, Germany, 1997.

[2] R.-D. Reiss and M. Thomas, *Statistical Analysis of Extreme Values: From Insurance, Finance, Hydrology and Other Fields*, Birkhäuser, Basel, Switzerland, 1997.

[3] J. Beirlant and A. Guillou, "Pareto index estimation under moderate right censoring," *Scandinavian Actuarial Journal*, vol. 101, no. 2, pp. 111–125, 2001.

[4] M. I. Fraga Alves, "A location invariant Hill-type estimator," *Extremes*, vol. 4, no. 3, pp. 199–217, 2001.

[5] M. I. Fraga Alves, M. I. Gomes, L. de Haan, and C. Neves, "Mixed moment estimator and location invariant alternatives," *Extremes*, vol. 12, no. 2, pp. 149–185, 2009.

[6] M. I. Gomes, L. de Haan, and L. H. Rodrigues, "Tail index estimation for heavy-tailed models: accommodation of bias in weighted log-excesses," *Journal of the Royal Statistical Society. Series B*, vol. 70, no. 1, pp. 31–52, 2008.

[7] M. I. Gomes and L. Henriques Rodrigues, "Tail index estimation for heavy tails: accommodation of bias in the excesses over a high threshold," *Extremes*, vol. 11, no. 3, pp. 303–328, 2008.

[8] J. Li, Z. Peng, and S. Nadarajah, "A class of unbiased location invariant Hill-type estimators for heavy tailed distributions," *Electronic Journal of Statistics*, vol. 2, pp. 829–847, 2008.

[9] J. Li, Z. Peng, and S. Nadarajah, "Asymptotic normality of location invariant heavy tail index estimator," to appear in *Extremes*.

[10] B. M. Hill, "A simple general approach to inference about the tail of a distribution," *The Annals of Statistics*, vol. 3, no. 5, pp. 1163–1174, 1975.
[11] D. M. Mason, "Laws of large numbers for sums of extreme values," *The Annals of Probability*, vol. 10, no. 3, pp. 754–764, 1982.
[12] P. Deheuvels, E. Haeusler, and D. M. Mason, "Almost sure convergence of the Hill estimator," *Mathematical Proceedings of the Cambridge Philosophical Society*, vol. 104, no. 2, pp. 371–381, 1988.
[13] Z. Peng and S. Nadarajah, "Strong convergence bounds of the Hill-type estimator under second-order regularly varying conditions," *Journal of Inequalities and Applications*, vol. 2006, Article ID 95124, 7 pages, 2006.
[14] L. de Haan and U. Stadtmüller, "Generalized regular variation of second order," *Journal of the Australian Mathematical Society. Series A*, vol. 61, no. 3, pp. 381–395, 1996.
[15] J. A. Wellner, "Limit theorems for the ratio of the empirical distribution function to the true distribution function," *Probability Theory and Related Fields*, vol. 45, no. 1, pp. 73–88, 1978.
[16] P. Deheuvels and D. M. Mason, "The asymptotic behavior of sums of exponential extreme values," *Bulletin des Sciences Mathématiques. 2e Série*, vol. 112, no. 2, pp. 211–233, 1988.
[17] H. Drees, "On smooth statistical tail functionals," *Scandinavian Journal of Statistics*, vol. 25, no. 1, pp. 187–210, 1998.
[18] L. de Haan and A. Ferreira, *Extreme Value Theory: An Introduction*, Springer Series in Operations Research and Financial Engineering, Springer, New York, NY, USA, 2006.
[19] A. L. M. Dekkers, J. H. J. Einmahl, and L. de Haan, "A moment estimator for the index of an extreme-value distribution," *The Annals of Statistics*, vol. 17, no. 4, pp. 1833–1855, 1989.