

STRONG CONVERGENCE BOUNDS OF THE HILL-TYPE ESTIMATOR UNDER SECOND-ORDER REGULARLY VARYING CONDITIONS

ZUOXIANG PENG AND SARALEES NADARAJAH

Received 22 April 2005; Revised 7 July 2005; Accepted 10 July 2005

Bounds on strong convergences of the Hill-type estimator are established under second-order regularly varying conditions

Copyright © 2006 Z. Peng and S. Nadarajah. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

1. Introduction

Suppose X_1, X_2, \dots are independent and identically distributed (iid) random variables with common distribution function (df) F . Let $M_n = \max\{X_1, \dots, X_n\}$ denote the maximum of the first n random variables and let $w(F) = \sup\{x : F(x) < 1\}$ denote the upper end point of F . The extreme value theory seeks norming constants $a_n > 0$, $b_n \in \mathfrak{R}$ and a nondegenerate df G such that the df of a normalized version of M_n converges to G , that is,

$$\Pr\left(\frac{M_n - b_n}{a_n} \leq x\right) = F^n(a_n x + b_n) \longrightarrow G(x) \quad (1.1)$$

as $n \rightarrow \infty$. If this holds for suitable choices of a_n and b_n then it is said that G is an extreme value df and F is in the domain of attraction of G , written as $F \in D(G)$. For suitable constants $a > 0$ and $b \in \mathfrak{R}$, one can write

$$G(ax + b) = G_\gamma(x) = \exp\{- (1 + \gamma x)^{-1/\gamma}\} \quad (1.2)$$

for all $1 + \gamma x > 0$ and $\gamma \in \mathfrak{R}$. For $\gamma > 0$, (1.1) is equivalent to

$$\lim_{t \rightarrow \infty} \frac{U(tx)}{U(t)} = x^\gamma, \quad (1.3)$$

where $U(t) = (1/(1 - F))^{-}(t) = \inf\{t \in \mathfrak{R} : 1/(1 - F(x)) \geq t\}$, that is, $U(t)$ is a regularly varying function at infinity with index γ .

2 Strong convergence bounds of the Hill-type estimator

The distribution given by (1.2) is known as the extreme value distribution. Its practical applications have been wide-ranging: fire protection and insurance problems, model for the extremely high temperatures, prediction of the high return levels of wind speeds relevant for the design of civil engineering structures, model for the extreme occurrences in Germany's stock index, prediction of the behavior of solar proton peak fluxes, model for the failure strengths of load-sharing systems and window glasses, model for the magnitude of future earthquakes, analysis of the corrosion failures of lead-sheathed cables at the Kennedy space center, prediction of the occurrence of geomagnetic storms, and estimation of the occurrence probability of giant freak waves in the sea area around Japan.

Each of the above problems requires estimation of the extremal index γ in (1.2). Several estimators for γ have been proposed in the extreme value theory literature. One of the first estimators of γ is due to Pickands [11]. Peng [10] proposed a general Pickands type estimator. Another kind of extremal index is the moment estimator proposed by Dekkers et al. [6], which generalizes the Hill type estimator for positive γ (Hill [8]).

However, there has been little work on trying to study the convergence properties of the estimators for γ . The question is: what is the penultimate form of the limit in (1.1)? Addressing this question is important because it will enable one to improve the modeling in each of the problems above. The convergence properties of the Pickands estimator such as consistency, asymptotic normality and the strong convergence rate have been discussed by Dekkers and De Haan [5], De Haan [1] and Pan [9]. Dekkers et al. [6] considered the weak consistency, strong consistency and the asymptotic normality of the Hill type estimator under different conditions. The aim of this paper is to consider the strong convergence rate of the Hill type estimator for γ under the second-order regularly varying conditions.

The Hill type estimator for $P(X_i > 0) = 1, i \geq 1$ is defined by

$$H_n = \frac{1}{k(n)} \sum_{i=1}^{k(n)} \log X_{n-i+1,n} - \log X_{n-k(n),n}, \quad (1.4)$$

where $X_{1,n} \leq X_{2,n} \leq \dots \leq X_{n,n}$ are order statistics of X_1, X_2, \dots, X_n , and $k(n)$ are positive integers satisfying $k(n) \rightarrow \infty, k(n)/n \rightarrow 0$ as $n \rightarrow \infty$. If Y_1, Y_2, \dots, Y_n are i.i.d random variables with common distribution function $\Pr(Y_1 \leq x) = 1 - 1/x, x \geq 1$ and if $Y_{1,n} \leq Y_{2,n} \leq \dots \leq Y_{n,n}$ are the order statistics of Y_1, Y_2, \dots, Y_n then $(U(Y_1), U(Y_2), \dots) \stackrel{d}{=} (X_1, X_2, \dots)$ and thus one may simplify H_n as

$$H_n = \frac{1}{k(n)} \sum_{i=1}^{k(n)} \log U(Y_{n-i+1,n}) - \log U(Y_{n-k(n),n}). \quad (1.5)$$

The investigation of the strong convergence rate of H_n requires knowing the convergence rate of (1.3). For this, we need to define second-order regularly varying functions. Firstly, a measurable real function $g(t)$ defined on $(0, \infty)$ is said to be a general regularly varying function with auxiliary function $a(t)$ if there exists a measurable function

$a(t) \rightarrow 0$ (as $t \rightarrow \infty$) with constant sign near infinity such that

$$\lim_{t \rightarrow \infty} \frac{g(tx) - g(t)}{a(t)} = S(x), \tag{1.6}$$

where $S(x)$ is not zero for some $x > 0$. It is known that $S(x)$ must be of the form $c\{x^\rho - 1\}/\rho$ (see, e.g., Resnick [12]) where $\rho \in \mathfrak{R}$ is referred to as the index of regular variation.

Now, suppose that there exists a regularly varying function $A(t) \rightarrow 0$ (as $t \rightarrow \infty$) such that

$$\lim_{t \rightarrow \infty} \frac{U(tx)/U(t) - x^\gamma}{A(t)} = H(x) \tag{1.7}$$

for all $x > 0$, where $H(x)$ is not a multiplier of x^γ . Then, $H(x)$ must be of the form $x^\rho\{x^\rho - 1\}/\rho$ for some $\rho \leq 0$, where ρ is the regularly varying index of $A(t)$, and (1.7) is locally uniformly convergent (De Haan [1]). We say that $U(t)$ satisfies second-order regularly varying conditions. It is easy to check that (1.7) is equivalent to

$$\lim_{t \rightarrow \infty} \frac{\log U(tx) - \log U(t) - \gamma \log x}{A(t)} = \frac{x^\rho - 1}{\rho} \tag{1.8}$$

for all $x > 0$, which is also locally uniform convergent.

2. Main results

We need the following four technical lemmas.

LEMMA 2.1. *If $k(n)$ satisfies $k(n)/n \rightarrow 0$ and $k(n)/(\log \log n) \uparrow \infty$ then*

$$\lim_{n \rightarrow \infty} \frac{k(n)}{n} Y_{n-k(n),n} = 1 \tag{2.1}$$

almost surely.

Proof. The result follows from Wellner [13] by noting that $1/Y_i$ are uniformly distributed on $(0, 1)$. □

LEMMA 2.2. *If $k(n) \sim \alpha_n \uparrow \infty$, $\log \log n = o(k(n))$ and $k(n)/n \sim \beta_n \downarrow 0$ then*

$$\limsup_{n \rightarrow \infty} \pm \{S_n(k(n)) - \mu_n(k(n))\} / \sqrt{2k(n) \log \log n} = \sqrt{2} \tag{2.2}$$

almost surely, where

$$S_n(k(n)) = \sum_{i=1}^{k(n)} \log Y_{n-i+1,n}, \tag{2.3}$$

$$\mu_n(k(n)) = k(n)(\log n - \log k(n) + 1).$$

Proof. The result follows from Deheuvels and Mason [4] after noting that $\log Y_i$ are i.i.d standard exponential variables. □

4 Strong convergence bounds of the Hill-type estimator

LEMMA 2.3. *If $k(n) \uparrow \infty$, $k(n)/n \sim \beta_n \downarrow 0$ and $\log \log n = o(k(n))$ then*

$$\limsup_{n \rightarrow \infty} \pm \frac{n/Y_{n-k(n)+1,n} - k(n)}{\sqrt{2k(n)\log \log n}} = 1 \quad (2.4)$$

almost surely.

Proof. The result follows from Deheuvels [2, 3] after noting that $1/Y_i$ are uniformly distributed on $(0, 1)$. \square

LEMMA 2.4. *If (1.8) holds then for arbitrary $\epsilon > 0$ there exists $t_0 > 0$ such that*

$$\left| \frac{\log U(tx) - \log U(t) - \gamma \log x}{A(t)} - \frac{x^\rho - 1}{\rho} \right| \leq \epsilon x^{\rho+\epsilon} \quad (2.5)$$

for all $x > 1$ and $t > t_0$.

Proof. follows from Drees [7]. \square

THEOREM 2.5. *If (1.7) holds with $k(n)$ and $A(n/k(n))$ satisfying $k(n)/n \sim \beta_n \downarrow 0$, $\sqrt{k(n)/(2\log \log n)} A(n/k(n)) \rightarrow \beta \in [0, \infty)$ and $k(n)/(\log n)^\delta \rightarrow \infty$ for some $\delta > 0$ then*

$$\limsup_{n \rightarrow \infty} \pm \sqrt{\frac{k(n)}{2\log \log n}} (H_n - \gamma) \leq (\sqrt{2} + 1)\gamma \pm \frac{\beta}{1 - \rho} \quad (2.6)$$

almost surely.

Proof. We prove the case for $\rho < 0$. The proof for $\rho = 0$ is similar. One can write

$$\begin{aligned} H_n - \gamma &= \frac{1}{k(n)} \sum_{i=1}^{k(n)} B_i(n) A(Y_{n-k(n),n}) + \frac{\gamma}{k(n)} \sum_{i=1}^{k(n)} (\log Y_{n-i+1,n} - \log Y_{n-k(n),n} - 1) \\ &\quad + \frac{A(Y_{n-k(n),n})}{\rho k(n)} \sum_{i=1}^{k(n)} \left[\left(\frac{Y_{n-i+1,n}}{Y_{n-k(n),n}} \right)^\rho - 1 \right], \end{aligned} \quad (2.7)$$

where

$$\begin{aligned} B_i(n) &= \frac{\log U(Y_{n-i+1,n}) - \log U(Y_{n-k(n),n}) - \gamma \log (Y_{n-i+1,n}/Y_{n-k(n),n})}{A(Y_{n-k(n),n})} \\ &\quad - \frac{(Y_{n-i+1,n}/Y_{n-k(n),n})^\rho - 1}{\rho} \end{aligned} \quad (2.8)$$

for $i = 1, 2, \dots, k(n)$. By Lemmas 2.1 and 2.4,

$$\left| \sum_{i=1}^{k(n)} B_i(n) \right| \leq \epsilon \sum_{i=1}^{k(n)} \left(\frac{Y_{n-i+1,n}}{Y_{n-k(n),n}} \right)^{\rho+\epsilon} \quad (2.9)$$

for all sufficiently large n . Note that $\Pr(Y_i^{\rho+\epsilon} \leq x) = x^{-1/(\rho+\epsilon)}$ for $0 \leq x \leq 1$ and $i = 1, 2, \dots, n$. By [5, Lemma 2.3(i)]

$$\lim_{n \rightarrow \infty} \frac{1}{k(n)} \sum_{i=1}^{k(n)} \left(\frac{Y_{n-i+1,n}}{Y_{n-k(n),n}} \right)^{\rho+\epsilon} = \frac{1}{1-\rho-\epsilon} \quad (2.10)$$

for almost surely. By Lemma 2.1 and since $A(t) \in Rv(\rho)$,

$$\lim_{n \rightarrow \infty} \frac{A(Y_{n-k(n),n})}{A(n/k(n))} = 1 \quad (2.11)$$

almost surely. Hence

$$\limsup_{n \rightarrow \infty} \sqrt{\frac{k(n)}{2 \log \log n}} |A(Y_{n-k(n),n})| \frac{1}{k(n)} \left| \sum_{i=1}^{k(n)} B_i(n) \right| \leq \frac{\epsilon \beta}{1-\rho-\epsilon} \quad (2.12)$$

almost surely. Letting $\epsilon \rightarrow 0$,

$$\lim_{n \rightarrow \infty} \sqrt{\frac{k(n)}{2 \log \log n}} \frac{A(Y_{n-k(n),n})}{k(n)} \sum_{i=1}^{k(n)} B_i(n) = 0 \quad (2.13)$$

almost surely. Similarly,

$$\lim_{n \rightarrow \infty} \sqrt{\frac{k(n)}{2 \log \log n}} \frac{A(Y_{n-k(n),n})}{k(n)} \sum_{i=1}^{k(n)} \left[\left(\frac{Y_{n-i+1,n}}{Y_{n-k(n),n}} \right)^\rho - 1 \right] = \frac{\beta \rho}{1-\rho} \quad (2.14)$$

almost surely. By Lemmas 2.2 and 2.3,

$$\begin{aligned} \limsup_{n \rightarrow \infty} \pm \sqrt{\frac{k(n)}{2 \log \log n}} \frac{\gamma}{k(n)} \sum_{i=1}^{k(n)} (\log Y_{n-i+1,n} - \log Y_{n-k(n),n} - 1) \\ \leq \limsup \pm \frac{\gamma}{\sqrt{2k(n) \log \log n}} \left(\sum_{i=1}^{k(n)} \log Y_{n-i+1,n} - \mu_n(k(n)) \right) \\ + \limsup_{n \rightarrow \infty} \pm \frac{\gamma}{\sqrt{2k(n) \log \log n}} (\mu_n(k(n)) - k(n) - k(n) \log Y_{n-k(n),n}) \leq (\sqrt{2} + 1) \gamma \end{aligned} \quad (2.15)$$

almost surely. The result of the theorem follows by combining (2.13)–(2.15).

Now, we provide an analogue of Theorem 2.5 when $U(tx)/U(t)$ converges to x^γ with faster speed. Specifically, suppose there exists a regularly varying function $A(t) \rightarrow 0$

6 Strong convergence bounds of the Hill-type estimator

(as $t \rightarrow \infty$) with index $\rho \leq 0$ such that

$$\lim_{t \rightarrow \infty} \frac{U(tx)/U(t) - x^\gamma}{A(t)} = 0, \quad (2.16)$$

where the convergence is locally uniform for $x > 0$. It is easy to check that (2.16) is equivalent to

$$\lim_{t \rightarrow \infty} \frac{\log U(tx) - \log U(t) - \gamma \log x}{A(t)} = 0, \quad (2.17)$$

which is also locally uniformly convergent for all $x > 0$. Under this assumption, the following result holds. Its proof is similar to that of Theorem 2.5. \square

THEOREM 2.6. *If (2.16) holds with $k(n)$ and $A(n/k(n))$ satisfying $k(n)/n \sim \beta_n \downarrow 0$ and $\sqrt{k(n)/(2 \log \log n)} A(n/k(n)) \rightarrow \beta \in [0, \infty)$ as $n \rightarrow \infty$ then*

$$\limsup_{n \rightarrow \infty} \pm \sqrt{\frac{k(n)}{2 \log \log n}} (H_n - \gamma) \leq (\sqrt{2} + 1)\gamma \quad (2.18)$$

almost surely.

Acknowledgment

The authors would like to thank the Editor-in-Chief and the referee for carefully reading the paper and for their great help in improving the paper.

References

- [1] L. De Haan, *Extreme value statistics*, Extreme Value Theory and Applications (J. Galambos, ed.), Kluwer Academic, Massachusetts, 1994, pp. 93–122.
- [2] P. Deheuvels, *Strong laws for the k th order statistic when $k \leq c \log_2 n$* , Probability Theory and Related Fields **72** (1986), no. 1, 133–154.
- [3] ———, *Strong laws for the k th order statistic when $k \leq c \log_2 n$. II*, Extreme Value Theory (Oberwolfach, 1987), Lecture Notes in Statistics, vol. 51, Springer, New York, 1989, pp. 21–35.
- [4] P. Deheuvels and D. M. Mason, *The asymptotic behavior of sums of exponential extreme values*, Bulletin des Sciences Mathématiques, Series 2 **112** (1988), no. 2, 211–233.
- [5] A. L. M. Dekkers and L. de Haan, *On the estimation of the extreme-value index and large quantile estimation*, The Annals of Statistics **17** (1989), no. 4, 1795–1832.
- [6] 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 **17** (1989), no. 4, 1833–1855.
- [7] H. Drees, *On smooth statistical tail functionals*, Scandinavian Journal of Statistics **25** (1998), no. 1, 187–210.
- [8] B. M. Hill, *A simple general approach to inference about the tail of a distribution*, The Annals of Statistics **3** (1975), no. 5, 1163–1174.
- [9] J. Pan, *Rate of strong convergence of Pickands' estimator*, Acta Scientiarum Naturalium Universitatis Pekinensis **31** (1995), no. 3, 291–296.
- [10] Z. Peng, *An extension of a Pickands-type estimator*, Acta Mathematica Sinica **40** (1997), no. 5, 759–762 (Chinese).
- [11] J. Pickands III, *Statistical inference using extreme order statistics*, The Annals of Statistics **3** (1975), no. 1, 119–131.

- [12] S. I. Resnick, *Extreme Values, Regular Variation, and Point Processes*, Applied Probability. A Series of the Applied Probability Trust, vol. 4, Springer, New York, 1987.
- [13] J. A. Wellner, *Limit theorems for the ratio of the empirical distribution function to the true distribution function*, *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete* **45** (1978), no. 1, 73–88.

Zuoxiang Peng: Department of Mathematics, Southwest Normal University, Chongqing 400715, China

E-mail address: pzx@swnu.edu.cn

Saralees Nadarajah: Department of Statistics, University of Nebraska–Lincoln, Lincoln, NE 68583, USA

E-mail address: snadaraj@unlserve.unl.edu