

Penultimate Approximation for Hill's Estimator

S. Cheng*(Peking University) L. de Haan**(Erasmus University Rotterdam)

Summary We prove that the probability distribution of Hill's estimator can be better approximated by a series of appropriate gamma distributions than by the limiting normal distribution.

1. Introduction

An often used estimator for the tail index of a probability distribution is Hill's estimator. Let X_1, X_2, \dots be i.i.d random variables with distribution function F and suppose that the function $1 - F$ is regularity varying at infinity with index $1/\gamma$, $\gamma > 0$. Hill's estimator[6]

$$H_{k,n} := \frac{1}{k} \sum_{i=0}^{k-1} \log X_{n-i,n} - \log X_{n-k,n},$$

where $X_{1,n} \leq \dots \leq X_{n,n}$ are the order statistics of X_1, \dots, X_n is a consistent estimator for γ if one chooses $k = k(n)$ in such a way that $k(n) \rightarrow \infty$, $k(n)/n \rightarrow 0$ ($n \rightarrow \infty$) (cf.[7]). For asymptotic normality one needs a further expansion of $1 - F$: suppose that for some function c not changing sign with $\lim_{t \rightarrow \infty} c(t) = 0$,

$$\lim_{t \rightarrow \infty} \frac{1}{c(t)} \left\{ \frac{1 - F(tx)}{1 - F(t)} - x^{-1/\gamma} \right\}$$

exists and is not zero. In that case we have, with a possibly different function c ,

$$\lim_{t \rightarrow \infty} \frac{1}{c(t)} \left\{ \frac{1 - F(tx)}{1 - F(t)} - x^{-1/\gamma} \right\} = x^{-1/\gamma} \cdot \frac{x^{\rho'} - 1}{\rho'} \quad (1.1)$$

for all $x > 0$ with ρ' a non-positive second order parameter([2]). Condition (1.1) is fulfilled e.g. when $1 - F(x) = c_1 x^{-1/\gamma} + c_2 x^{-\gamma + \rho'} (1 + o(1))$ ($x \rightarrow \infty$) with $\rho' < 0$. Under these conditions

$$\sqrt{k} \left\{ \frac{H_{k,n}}{\gamma} - 1 \right\}$$

is asymptotically standard normal for sequences $k = k(n)$ that do not increase too quickly, more precisely: if

*Supported in part by the Netherlands Organization for Scientific Research.

**Research carried out for a large part when working at Delft University of Technology.

$$\lim_{n \rightarrow \infty} \sqrt{k}a(n/k) = 0, \quad (1.2)$$

with U the inverse of the function $1/(1 - F)$ and $a(t) = c(U(t))$. Cheng and Pan [1] proved that under condition (1.1),

$$\lim_{n \rightarrow \infty} \sqrt{k}[P(H_n \leq x) - \Phi(x)] = \left[\frac{1 - x^2}{3} - \frac{C}{\gamma(1 - \rho)} \right] \gamma^2 \phi(x) \quad (1.3)$$

uniformly for all x where $\rho = \rho' \gamma$, Φ is the standard normal distribution function and ϕ its density, provided

$$\lim_{n \rightarrow \infty} ka(n/k) = C. \quad (1.4)$$

Note that (1.4) implies (1.2).

We shall prove a rate of convergence result for $H_{k,n}$ when the approximating distribution is not the normal one but a standardized version (mean zero, variance one) of the gamma distribution function with k degrees of freedom.

Theorem 1. *Suppose (1.1), (1.2) and*

$$\lim_{n \rightarrow \infty} \frac{\log n}{k} = 0 \quad (1.5)$$

hold. Then

$$\lim_{n \rightarrow \infty} \frac{1}{\sqrt{k_n}a(n/k)} [P\{\sqrt{k}(\frac{H_{k,n}}{\gamma} - 1) \leq x\} - \Gamma_k(k + x\sqrt{k})] = -\frac{\gamma\phi(x)}{1 - \rho} \quad (1.6)$$

uniformly for all x where Γ_k is the distribution function of a gamma distribution with k degrees of freedom.

In order to compare the two results note that for $ka(n/k) \rightarrow 0$ and $\log n = o(k)$, the rate in (1.6) is better than in (1.3), for $ka(n/k) \rightarrow C \in (0, \infty)$ the two rates are of the same order and for $ka(n/k) \rightarrow \infty$ but $\sqrt{k}a(n/k) \rightarrow 0$, only (1.6) is valid. Approximation by the gamma distribution is attractive since the density of the density has a known functional form like the normal distribution, moreover the distribution is asymmetric and this is what seems to happen to Hill's estimator in simulations and applications.

2. Lemmas

First note (cf.[2]) that (1.1) implies

$$\lim_{t \rightarrow \infty} \frac{\log U(tx) - \log U(t) - \gamma \log x}{\gamma^2 a(t)} = \frac{x^\rho - 1}{\rho} \quad (2.1)$$

for $x > 0$. It follows that a is regularity varying at infinity with index $\rho \leq 0$ and therefore for every $\epsilon > 0$, there exists a $t_0 > 0$ such that

$$(1 - \epsilon) \min(x^{\rho - \epsilon}, x^{\rho + \epsilon}) \leq a(tx)/a(t) \leq (1 + \epsilon) \max(x^{\rho - \epsilon}, x^{\rho + \epsilon}) \quad (2.2)$$

holds for all $t \geq t_0$, $tx \geq t_0$ and $x > 0$. Moreover, we need three lemmas.

Lemma 1. ([3]) Suppose (1.1) holds. Then there exists a function \hat{a} with $\hat{a}(t)/a(t) \rightarrow 1$ ($t \rightarrow \infty$), such that for every $\epsilon > 0$, there exists a t_0 with: uniformly for $x \geq 1$ and $t \geq t_0$

$$\left| \frac{\log U(tx) - \log U(t) - \gamma \log x}{\hat{a}(t)} - \gamma^2 \frac{x^\rho - 1}{\rho} \right| \leq \epsilon x^{\rho+\epsilon}. \quad (2.3)$$

Lemma 2. Suppose $k = o(n)$ and (1.5) hold. Then there exists a sequence $\{t_n\}$ with $t_n \rightarrow 0$ such that for $n \rightarrow \infty$

$$P(|kY_{n-k,n}/n - 1| \geq t_n) = o(\sqrt{k}\hat{a}(n/k)) \quad (2.4)$$

where $\{Y_{i,n}\}_{i=1}^n$ are the n -th order statistics from the distribution function $1 - x^{-1}$, $x \geq 1$.

Proof. Define $I_i := I_{\{Y_i > n(1-t_n)/k\}}$ with Y_1, Y_2, \dots i.i.d. with the distribution function $1 - x^{-1}$, $x \geq 1$. Then $\sum_{i=1}^n I_i \in \text{Bin}(n, r_n^{-1})$ with $r_n := n(1 - t_n)/k$. Hoeffding's inequality (cf.[8], page 440) gives

$$\begin{aligned} P\left(\frac{kY_{n-k,n}}{n} \leq 1 - t_n\right) &= P\left(\sum_{i=1}^n I_i \leq k\right) \\ &= P\left(\frac{\sum_{i=1}^n I_i - nr_n^{-1}}{\sqrt{n}} \leq -\frac{kt_n}{\sqrt{n(1-t_n)}}\right) \\ &\leq \exp\left\{-\left[\frac{kt_n}{\sqrt{n(1-t_n)}}\right]^2 / \left[\frac{2k}{n(1-t_n)}\left(1 - \frac{k}{n(1-t_n)}\right)\right]\right\} \\ &= \exp\left\{-\frac{kt_n^2[1+o(1)]}{2}\right\}. \end{aligned}$$

If $\rho = 0$, we take $t_n = [(2 \log n)/k]^{1/2}$ and by (2.2) get for some $K > 0$ and $\epsilon \in (0, 1)$,

$$\frac{\exp\{-kt_n^2[1+o(1)]/2\}}{\sqrt{k}\hat{a}(n/k)} \leq \frac{K}{n^{1-\epsilon}k^{1/2+\epsilon}} \rightarrow 0.$$

If $\rho < 0$, we take $t_n = [(-3\rho \log n)/k]^{1/2}$ and by (2.2) get for some $K > 0$ and $\epsilon \in (0, -\rho/2)$,

$$\frac{\exp\{-kt_n^2[1+o(1)]/2\}}{\sqrt{k}\hat{a}(n/k)} \leq \frac{K}{n^{-\rho/2-\epsilon}k^{1/2-\rho+\epsilon}} \rightarrow 0.$$

Hence $P(kY_{n-k,n}/n \leq 1 - t_n) = o(\sqrt{k}\hat{a}(n/k))$. similarly we may show $P(kY_{n-k,n}/n \geq 1 + t_n) = o(\sqrt{k}\hat{a}(n/k))$, completing the proof of (2.5). ■

Lemma 3. For any sequence $f_k \rightarrow 0$, we have

$$\lim_{k \rightarrow \infty} \frac{\Gamma_k(k + (x + f_k)\sqrt{k}) - \Gamma_k(k + x\sqrt{k})}{f_k} = \phi(x) \quad (2.6)$$

uniformly for all x .

Proof. By Cramer's theorem (cf. [4] Theorem 1 of XVI.4)

$$\Gamma_k(k + x\sqrt{k}) = \Phi(x) + \frac{(1 - x^2)\phi(x)}{3\sqrt{k}} + o\left(\frac{1}{\sqrt{k}}\right)$$

uniformly for all x . Hence

$$\begin{aligned} & \frac{\Gamma_k(k + (x + f_k)\sqrt{k}) - \Gamma_k(k + x\sqrt{k})}{f_k} \\ &= \frac{\Phi(x + f_k) - \Phi(x)}{f_k} + \frac{\phi(x + f_k)[1 - (x + f_k)^2] - \phi(x)(1 - x^2)}{3\sqrt{k}f_k} + o\left(\frac{1}{\sqrt{k}}\right) \end{aligned}$$

uniformly in x . The first term goes to $\phi(x)$ uniformly and the other terms go to zero uniformly, so that (2.6) holds uniformly. ■

3. Proof of Theorem 1 We use the representation $\{X_i\}_{i=1}^\infty \stackrel{d}{=} \{U(Y_i)\}_{i=1}^\infty$ and may assume

$$H_{k,n} := \frac{1}{k} \sum_{i=0}^{k-1} \log U(Y_{n-i,n}) - \log U(Y_{n-k,n}).$$

Write

$$\begin{aligned} & P\left\{\sqrt{k}\left(\frac{H_{k,n}}{\gamma} - 1\right) \leq x\right\} - G_k(x) \\ &= [P\left\{\sqrt{k}\left(\frac{H_{k,n}}{\gamma} - 1\right) \leq x \mid \left|\frac{kY_{n-k,n}}{n} - 1\right| \leq t_n\right\} - G_k(x)] P\left(\left|\frac{kY_{n-k,n}}{n} - 1\right| \leq t_n\right) \\ & \quad - G_k(x) P\left(\left|\frac{kY_{n-k,n}}{n} - 1\right| > t_n\right) + P\left(\sqrt{k}\left(\frac{H_{k,n}}{\gamma} - 1\right) \leq x, \left|\frac{kY_{n-k,n}}{n} - 1\right| > t_n\right). \end{aligned}$$

According to Lemma 2, $P(|kY_{n-k,n}/n - 1| \leq t_n) \rightarrow 1$ and the two last terms are $o(\sqrt{k}\hat{\alpha}(n/k))$, so we only have to deal with the first term and in fact, we only have to find the limit of

$$\frac{P\left\{\sqrt{k}\left(H_{k,n}/\gamma - 1\right) \leq x \mid \left|kY_{n-k,n}/n - 1\right| \leq t_n\right\} - G_k(x)}{\sqrt{k}\hat{\alpha}(n/k)}.$$

If $kY_{n-k,n}/n \geq 1 - t_n$ with $t_n \rightarrow 0$, we have by Lemma 1 eventually for $0 \leq i \leq k - 1$

$$\begin{aligned} & \left| \frac{\log U(Y_{n-i,n}) - \log U(Y_{n-k,n}) - \gamma \log(Y_{n-i,n}/Y_{n-k,n})}{\hat{\alpha}(Y_{n-k,n})} - \gamma^2 \frac{(Y_{n-i,n}/Y_{n-k,n})^\rho - 1}{\rho} \right| \\ & \leq \epsilon \left(\frac{Y_{n-i,n}}{Y_{n-k,n}}\right)^{\rho+\epsilon}. \end{aligned}$$

Moreover, if $|kY_{n-k,n}/n - 1| \leq t_n$ with $t_n \rightarrow 0$ and n is so large that $(1 - t_n)^{\rho - \epsilon} \leq 1 + \epsilon$, $(1 - t_n)^\epsilon \geq 1 - \epsilon$ and $(1 + t_n)^{\rho - \epsilon} \geq 1 - \epsilon$, then we have from (2.2) that

$$\begin{aligned} (1 - \epsilon)^2 \hat{a}(n/k) &\leq (1 - \epsilon) \min\{(1 + t_n)^{\rho - \epsilon}, (1 - t_n)^\epsilon\} \hat{a}(n/k) \\ &\leq \hat{a}(Y_{n-k,n}) \leq (1 + \epsilon)(1 - t_n)^{\rho - \epsilon} \hat{a}(n/k) \leq (1 + \epsilon)^2 \hat{a}(n/k) \end{aligned}$$

Hence for $|kY_{n-k,n}/n - 1| \leq t_n$ with $t_n \rightarrow 0$, we get by addition

$$\begin{aligned} &\sqrt{k} \left(\frac{H_{k,n}}{\gamma} - 1 \right) - \sqrt{k} \left\{ \frac{1}{k} \sum_{i=0}^{k-1} \log \frac{Y_{n-i,n}}{Y_{n-k,n}} - 1 \right\} \\ &\geq \frac{\hat{a}(Y_{n-k,n})}{\sqrt{k}} \left[\gamma \sqrt{k} \sum_{i=0}^{k-1} \frac{(Y_{n-i,n}/Y_{n-k,n})^\rho - 1}{\rho} - \frac{\epsilon}{\gamma} \sum_{i=0}^{k-1} \left(\frac{Y_{n-i,n}}{Y_{n-k,n}} \right)^{\rho + \epsilon} \right] \geq R_k \\ &:= \frac{\hat{a}(n/k)}{\sqrt{k}} \left[\gamma (1 - \epsilon)^2 \sum_{i=0}^{k-1} \frac{(Y_{n-i,n}/Y_{n-k,n})^\rho - 1}{\rho} - \frac{\epsilon(1 + \epsilon)^2}{\gamma} \sum_{i=0}^{k-1} \left(\frac{Y_{n-i,n}}{Y_{n-k,n}} \right)^{\rho + \epsilon} \right] \end{aligned}$$

and a similar lower inequality. So we have for sufficiently large n ,

$$\begin{aligned} &P \left\{ \sqrt{k} \left(\frac{H_{k,n}}{\gamma} - 1 \right) \leq x \mid \left| \frac{kY_{n-k,n}}{n} - 1 \right| \leq t_n \right\} \tag{3.1} \\ &\leq P \left(\sqrt{k} \left\{ \frac{1}{k} \sum_{i=0}^{k-1} \left[\log \frac{Y_{n-i,n}}{Y_{n-k,n}} - 1 \right] \right\} + R_k \leq x \mid \left| \frac{kY_{n-k,n}}{n} - 1 \right| \leq t_n \right) \\ &= P \left(\sqrt{k} \left\{ \frac{1}{k} \sum_{i=0}^{k-1} \left[\log \frac{Y_{n-i,n}}{Y_{n-k,n}} - 1 \right] \right\} + R_k \leq x \right) \end{aligned}$$

since $\{Y_{n-i,n}/Y_{n-k,n}\}_{i=0}^{k-1}$ and $Y_{n-k,n}$ are independent. Note by the lack of memory property of the exponential distribution

$$\sum_{i=0}^{k-1} \log \frac{Y_{n-i,n}}{Y_{n-k,n}} \stackrel{d}{=} \sum_{i=0}^{k-1} E_i$$

with E_0, E_1, \dots, E_{k-1} i.i.d. standard exponential. A similar simplification may be applied to $\sum_{i=0}^{k-1} (Y_{n-i,n}/Y_{n-k,n})^\rho$ and $\sum_{i=0}^{k-1} (Y_{n-i,n}/Y_{n-k,n})^{\rho + \epsilon}$. Write

$$Q_k := \frac{1}{\sqrt{k}} \sum_{i=0}^{k-1} \left\{ \gamma (1 - \epsilon)^2 \left(\frac{e^{\rho E_i} - 1}{\rho} - \frac{1}{1 - \rho} \right) - \frac{\epsilon(1 + \epsilon)^2}{\gamma} \left[e^{(\rho - \epsilon) E_i} - \frac{1}{1 - \rho + \epsilon} \right] \right\},$$

which is asymptotically normal, and

$$\hat{\Gamma}_k := \sqrt{k} \left(\frac{1}{k} \sum_{i=0}^{k-1} E_i - 1 \right) + \sqrt{k} \hat{a}(n/k) \left[\frac{\gamma(1 - \epsilon)^2}{1 - \rho} - \frac{\epsilon(1 + \epsilon)^2}{\gamma(1 - \rho + \epsilon)} \right].$$

Then we have from (3.1)

$$\begin{aligned}
& P\left\{\sqrt{k}\left(\frac{H_{k,n}}{\gamma} - 1\right) \leq x \mid \left|\frac{kY_{n-k,n}}{n} - 1\right| \leq t_n\right\} \\
& \leq P(\hat{\Gamma}_k + \hat{a}(n/k)Q_k \leq x) \\
& \leq P(\hat{\Gamma}_k \leq x + \epsilon\sqrt{k}\hat{a}(n/k)) + P(Q_k \leq -\epsilon\sqrt{k}).
\end{aligned} \tag{3.2}$$

We consider the two terms separately. Noting that $\sqrt{k}(\sum_{i=0}^{k-1} E_i/k - 1)$ has distribution $\Gamma_k(k + \sqrt{k}x)$, we have from Lemma 3

$$\begin{aligned}
& \frac{P(\hat{\Gamma}_k \leq x + \epsilon\sqrt{k}\hat{a}(n/k)) - \Gamma_k(k + x\sqrt{k})}{\sqrt{k}\hat{a}(n/k)} \\
& = \frac{\Gamma_k(k + (x + \mu_\epsilon\sqrt{k}\hat{a}(n/k))\sqrt{k}) - \Gamma_k(k + x\sqrt{k})}{\sqrt{k}\hat{a}(n/k)} \\
& \rightarrow \mu_\epsilon\phi(x)
\end{aligned}$$

uniformly for all x where

$$\mu_\epsilon := \frac{-\gamma(1-\epsilon)^2}{1-\rho} + \frac{\epsilon(1+\epsilon)^2}{\gamma(1-\rho-\epsilon)} + \epsilon.$$

For the second term we use the exponential inequality (cf.[8], page 851) for independent random variables

$$P(Q_k \leq -\epsilon\sqrt{k}) \leq \exp\{-Ck\}$$

with $C > 0$ and the right hand side is $o(\sqrt{k}\hat{a}(n/k))$ under the stated condition (1.6). It follows from (3.2) that

$$\limsup_{n \rightarrow \infty} \frac{P\left\{\sqrt{k}(H_{k,n}/\gamma - 1) \leq x \mid \left|\frac{kY_{n-k,n}}{n} - 1\right| \leq t_n\right\} - \Gamma_k(k + x\sqrt{k})}{\sqrt{k}\hat{a}(n/k)} \leq \mu_\epsilon\phi(x).$$

Letting $\epsilon \rightarrow 0$, we get

$$\limsup_{n \rightarrow \infty} \frac{P\left\{\sqrt{k}(H_{k,n}/\gamma - 1) \leq x \mid \left|\frac{kY_{n-k,n}}{n} - 1\right| \leq t_n\right\} - \Gamma_k(k + x\sqrt{k})}{\sqrt{k}\hat{a}(n/k)} \leq \frac{-\gamma\phi(x)}{1-\rho}$$

uniformly for all x . Similarly, we may prove

$$\liminf_{n \rightarrow \infty} \frac{P\left\{\sqrt{k}(H_{k,n}/\gamma - 1) \leq x \mid \left|\frac{kY_{n-k,n}}{n} - 1\right| \leq t_n\right\} - \Gamma_k(k + x\sqrt{k})}{\sqrt{k}\hat{a}(n/k)} \geq \frac{-\gamma\phi(x)}{1-\rho}$$

uniformly for all x , completing the proof of the theorem. \blacksquare

4. Remarks

Under some condition like (1.4) used by Cheng and Pan, we may also get approximation by gamma distributions. In fact, instead of Lemma 3, we have

Lemma 4. *Suppose (1.2) holds and there exists a $\eta > 0$ such that*

$$\lim_{n \rightarrow \infty} k^{\eta+1/2} a(n/k) = C \in [0, \infty]. \quad (4.1)$$

Then there exists a sequence $\{t_n\}$ such that for $n \rightarrow \infty$, $t_n \rightarrow 0$ and

$$P(|kY_{n-k,n}/n - 1| \geq t_n) = \begin{cases} o(\sqrt{k}\hat{a}(n/k)) & \text{for } C \in (0, \infty] \\ o(k^{-\eta}) & \text{for } C = 0. \end{cases} \quad (4.2)$$

Proof. As in the proof of Lemma 2, Hoeffding's inequality gives

$$P\left(\frac{kY_{n-k,n}}{n} \leq 1 - t_n\right) \leq \exp\left\{-\frac{kt_n^2[1 + o(1)]}{2}\right\}.$$

Take

$$t_n = \begin{cases} \{-2[\log(\sqrt{k}\hat{a}(n/k))]/\sqrt{k}\}^{1/2} & \text{if } C = 0; \\ [4\eta \log k/k]^{1/2} & \text{if } C \in (0, 1); \\ [2\eta \log k/k]^{1/2} & \text{if } C = \infty. \end{cases}$$

Then (4.2) holds. ■

Following the steps in proving Theorem 1 and using Lemma 4, we can get the following.

Theorem 2. *Suppose (1.1), (1.2) and (4.1) with some $\eta > 0$ hold. If $C \in (0, \infty]$ in (4.1), then (1.6) holds and if $C = 0$ in (4.1), then*

$$\lim_{n \rightarrow \infty} k^\eta \left[P\left\{ \sqrt{k} \left(\frac{H_{k,n}}{\gamma} - 1 \right) \leq x \right\} - \Gamma_k(k + x\sqrt{k}) \right] = -\frac{\gamma\phi(x)}{1-\rho}. \quad (4.3)$$

REFERENCES

- [1] S.Cheng and J.Pan(1998). Asymptotic Expansions of Estimators for the Tail Index with Applications. Scand. J. Statist. 25: 717-728.
- [2] L.de Haan and U.Stadtmeuler(1996). Generalized Regular Variation of Second Order. J. Australian Math. Soc., Series A(61): 381-395.
- [3] H. Drees(1998). Smooth Statistical Tail Functionals. Scand. J. Statist. 25: 187-210.
- [4] W.Feller(1971). An Introduction on Probability and its Applications, Volume II. Wiley, New York.
- [5] P.Hall(1982). On Some Simple Estimates of an Exponent of Regular Variation. Roy. Statistical Soc., Ser. B(44):37-42.
- [6] B.M.Hill(1975). A Simple General Approach to Inference about the Tail of a Distribution. Ann. Statist. 3: 1163-1174.
- [7] D.M.Mason(1982). Laws of Large Numbers for Sums of Extreme Values. Ann. Statist. 10: 754-764.
- [8] G.R.Shorack and J.Wellner(1986). Empirical Processes with Applications to Statistics. John Wiley & Sons. New York.