December 7, 2012

ESTIMATION OF THE TAIL INDEX FOR LATTICE-VALUED SEQUENCES

MUNEYA MATSUI, THOMAS MIKOSCH, AND LALEH TAFAKORI

Abstract. If one applies the Hill, Pickands or Dekkers-Einmahl-de Haan estimators of the tail index of a distribution to data which are rounded off one often observes that these estimators oscillate strongly as a function of the number k of order statistics involved. We study this phenomenon in the case of a Pareto distribution. We provide formulas for the expected value and variance of the Hill estimator and give bounds on k when the central limit theorem is still applicable. We illustrate the theory by using simulated and real-life data.

Keywords: tail index, Hill estimator, Pickands estimator, Dekkers-Einmahl-de Haan estimator, discretized Pareto random variable, central limit theorem, consistency.

2010 Mathematics Subject Classification: Primary 62G32, Secondary 62G20; 62G30

1. Introduction

Numerous real-life data (X_n) have power-like tails in the sense that for some $\alpha > 0$ and large x,

$$P(X_n > x) \approx x^{-\alpha}\,.$$

One finds such data sets in insurance (e.g. claim sizes for fire, storm, motor insurance; see Embrechts et al. [4]), telecommunications (e.g. file sizes, throughput rates, transmission durations; see Resnick [14]), finance (e.g. log-returns of speculative prices; see [4] and Mikosch [12]), seismological studies (e.g. magnitudes of earthquake aftershocks; see Kagan and Vere-Jones [10]). Power law tails are also observed in the context of *Zipf's law* which is empirically observed for the distributional tail of the sizes of large city populations in a given country, the distribution of words in a given national literature, and in other demographic, linguistic, financial and economic applications; see e.g. Gabaix and Ioannides [6]. Power law tails are also the basis for defining the notion of *integral dimension* of the attractor of a dynamical system; see Falconer [5] for its definition and Takens [15] for its estimation; the latter estimator is closely related to the Hill estimator used in extreme value theory.

A convenient way of describing power law behavior is the notion of *regular variation.* Recall that the distribution function F of a positive random variable X has a *regularly varying tail* if it can be written in the form

$$\overline{F}(x) = P(X > x) = \frac{L(x)}{x^{\alpha}}\,, \quad x > 0\,, \tag{1.1}$$

where $\alpha \geq 0$ is the *tail index* and L is a *slowly varying function*, i.e., for every $c > 0$, $\lim_{x\to\infty} L(cx)/L(x) = 1$. The theory of regularly varying functions is well studied; see e.g. the encyclopedic treatment in Bingham et al. [1]. The function L is an infinite-dimensional nuisance parameter which makes the statistical estimation of the parameter α a very difficult task. The appearance of a slowly varying function L in the tail $\overline{F}(x)$ is due to limit theory for sums and

Thomas Mikosch's research is partly supported by the Danish Natural Science Research Council (FNU) Grants 09-072331 "Point process modelling and statistical inference" and 10-084172 "Heavy tail phenomena: Modeling and estimation". Muneya Matsui's research is partly supported by the JSPS Grant-in-Aid for Research Activity start-up Grant Number 23800065. Parts of this paper were written when Laleh Tafakori and Muneya Matsui were visiting the Department of Mathematics of the University of Copenhagen. They would like to thank for hospitality of the host institution.

partial maxima of iid random variables. Then (1.1) is a necessary domain of attraction property; see e.g. Embrechts et al. [4], Chapters 3 and 4.

For real-life data, the tail index α has to be estimated. There exists a well developed statistical theory for this purpose; see de Haan and Ferreira [8] for a complete theory in the case of iid data, see also the discussion in Chapter 6 of [4]. The estimation of α for stationary sequences is even more challenging due to clusters of exceedances of high thresholds; see e.g. Drees and Rootzén [3]. There exists a multitude of estimators of α; see the literature mentioned above. Our main focus is on the most popular *Hill estimator*; see Hill [9]. Writing

$$X_{(1)} \leq \cdots \leq X_{(n)}$$

for the order statistics of the observations $X_1, \ldots, X_n$, the Hill estimator of α^{-1} is given by

$$\widehat{\alpha}_k^{-1} = \frac{1}{k-1} \sum_{i=1}^{k-1} \log \frac{X_{(n-i+1)}}{X_{(n-k+1)}} .$$

In order to achieve desirable statistical properties such as consistency and asymptotic normality one needs the conditions $k = k_n \to \infty$ and $k_n/n \to 0$, i.e. one needs to consider a whole family of estimators $(k, \widehat{\alpha}_k^{-1})$, $k = 2, 3, \ldots, n-1$, defining the so-called *Hill plot*. The rationale of the Hill estimator is easily explained when assuming a *Pareto tail*:

$$\overline{F}(x) = u^{\alpha} x^{-\alpha} , \quad x > u , \tag{1.2}$$

for some positive value u. For a given data set, the slowly varying function L in (1.1) is in general unknown. The introduction of the model (1.2) is based on the belief that the model (1.1) can be approximated in some sense by (1.2) if the threshold u is "sufficiently high". A theoretical basis for this belief is the Pickands-Balkema-de Haan theorem (cf. Theorem 3.4.5 in [4]) which states that the generalized Pareto distribution (GPD) appears as limit distribution of the normalized sample excesses above high thresholds. In heavy-tail situations, the GPD and the Pareto distribution with parameter $\alpha > 0$ belong to the same location-scale family. The Hill estimator $\widehat{\alpha}_k$ can be derived as the maximum likelihood estimator of α based on the k largest order statistics in an iid sample with tail (1.2). If $k_n \to \infty$ and $k_n/n \to 0$, the Hill estimator is consistent under the more general condition (1.1) for some $\alpha > 0$, i.e. $\widehat{\alpha}_k \stackrel{P}{\to} \alpha$. If one assumes the exact Pareto tail (1.2) one also has asymptotic normality $\sqrt{k}\,(\widehat{\alpha}_k - \alpha) \stackrel{d}{\to} Z$ for a normal $N(0, \alpha^2)$ random variable Z for any $k_n \to \infty$, in particular for $k_n = n$. This is simply due to the fact that

$$\widehat{\alpha}_k^{-1} = \frac{1}{k-1} \sum_{i=1}^{k-1} \log \frac{X_{(n-i+1)}}{X_{(n-k+1)}} \stackrel{d}{=} \alpha^{-1} \frac{\Gamma_{k-1}}{k-1} , \quad 2 \leq n , \quad 3 \leq k \leq n-1 , \tag{1.3}$$

where $\Gamma_i = E_1 + \cdots + E_i$, $i \geq 1$, for an iid standard exponential sequence (E_i); see the second display on p. 192 in [4]. Then $\sqrt{k}(\widehat{\alpha}_k^{-1} - \alpha^{-1}) \stackrel{d}{\to} Y$ for a normal $N(0, \alpha^{-2})$ random variable Y, and the relation $\sqrt{k}\,(\widehat{\alpha}_k - \alpha) \stackrel{d}{\to} Z$ for a normal $N(0, \alpha^2)$ random variable Z follows by an application of the Δ-method for any sequence $k_n \to \infty$ such that $k_n \leq n$.

Real-life data are always discrete, i.e. for any practical purposes one would only collect data which are concentrated on a lattice with equidistant grid size. Any data saved on computers or other electronic devices are of this kind and therefore the number of digits is limited. More importantly, various data sets are rather imprecise due to the lack of information or measurement error. A typical example are seismological data: the 3-dimensional coordinates (latitude, longitude and depth) of the epicenter of an earthquake can often only be determined up to tens or even hundreds of kilometers. Another example are the longest life spans of humans: due to the rareness of the event that a person survives the age of 100 years such extreme life spans are usually registered in mortality or demographic tables only as belonging to certain intervals, e.g. $(101,110]$ or $(101,120]$.

It is the aim of this note to discuss the influence of discretization effects (such as round-off, imprecise data) on the estimation of α^{-1}. We focus on the Hill estimator but we also touch on the Pickands and Dekkers-Einmahl-de Haan estimators and show simulation evidence that these estimators may suffer from the discreteness of the data. To illustrate the effect of discreteness of data on the Hill estimator we consider a data set of word counts from the English language. The data are available at www.wordfrequency.inf, where one also finds a description how the data were cleaned. This data set is often used to support the evidence on Zipf's law as regards the distribution of words in a given language. The data have Pareto-like tails with an estimated index close to 1; see Figure 1, top left. The data consists of the 10 000 largest counts between 3 000 and 23 million. In the remaining 3 graphs in Figure 1 we show the Hill plots of the same data when the last one, two or three digits in the count data are replaced by zeros. This means that the counting units are tens, hundreds, thousands, respectively. While the effect for units of tens is hardly visible, we see significant changes in the Hill plot for units of hundreds and thousands. We do not show the plots for counting units of ten thousands. Then about 60% of the data turns into zero and the Hill plot oscillates even more wildly than for smaller counting units.

We do not aim at a general distribution with regularly varying tail but we choose a Pareto distribution as a toy model. Throughout we consider a Pareto distributed iid sequence with representation

$$X_i = U_i^{-1/\alpha}\,, \quad i \in \mathbb{N}\,, \tag{1.4}$$

for an iid $U(0,1)$ sequence (U_i) and some positive α. It is easy to see that

$$\overline{F}(x) = P(U_i^{-1/\alpha} > x) = x^{-\alpha}\,, \quad x \geq 1\,, \tag{1.5}$$

and hence the order statistics satisfy the relation $X_{(i)} = U_{(n-i+1)}^{-1/\alpha}$ for $i \leq n$.

In Figure 2, top left, we show a Hill plot based on a sample of size 10 000 with $\alpha = 1$. The plot nicely shows the trade-off between bias and variance depending on the chosen values k: too small values of k lead to a large variance while too large k lead to a larger bias. In the remaining graphs of Figure 2 we show the Hill plots for the iid sample $10^{-l}[10^l U_i^{-1/\alpha}]$ for $l = 0,1,2$, where $[x]$ denotes the integer part of any real number x. (Due to the scale invariance of the Hill estimator, these Hill plots coincide with those based on $([10^l U_i^{-1/\alpha}])$.) This transformation of $U_i^{-1/\alpha}$ turns all digits but the first l ones behind the comma into zeros. In this sense, we obtain a discretization of the random variables $U_i^{-1/\alpha}$ by rounding off. The tail of the transformed random variable is given by

$$\begin{aligned} \overline{F}(x) &= 1 - P([10^l U_i^{-1/\alpha}] \leq [10^l x]) \\ &= 1 - P(10^l U_i^{-1/\alpha} < [10^l x] + 1) \\ &= \frac{(10^l)^\alpha}{([10^l x] + 1)^\alpha}\,. \end{aligned} \tag{1.6}$$

Since $[y] \in (y-1, y]$ for $y \in \mathbb{R}$ one immediately concludes that

$$P(10^{-l}[10^l U_i^{-1/\alpha}] > x) \sim P(U_i^{-1/\alpha} > x) = x^{-\alpha}\,, \quad x \to \infty\,.$$

Therefore the standard theory (see Mason [11] or Theorem 3.2.2 in de Haan and Ferreira [8]) yields that the Hill estimator is consistent:

$$\widehat{\alpha}_k \stackrel{P}{\to} \alpha\,, \quad \text{if } k \to \infty \text{ and } k/n \to 0$$

and even strongly consistent (i.e. $\widehat{\alpha}_k \stackrel{\text{a.s.}}{\to} \alpha$) if $k/n \to 0$ and $k/\log\log n \to \infty$; see Deheuvels et al. [2].

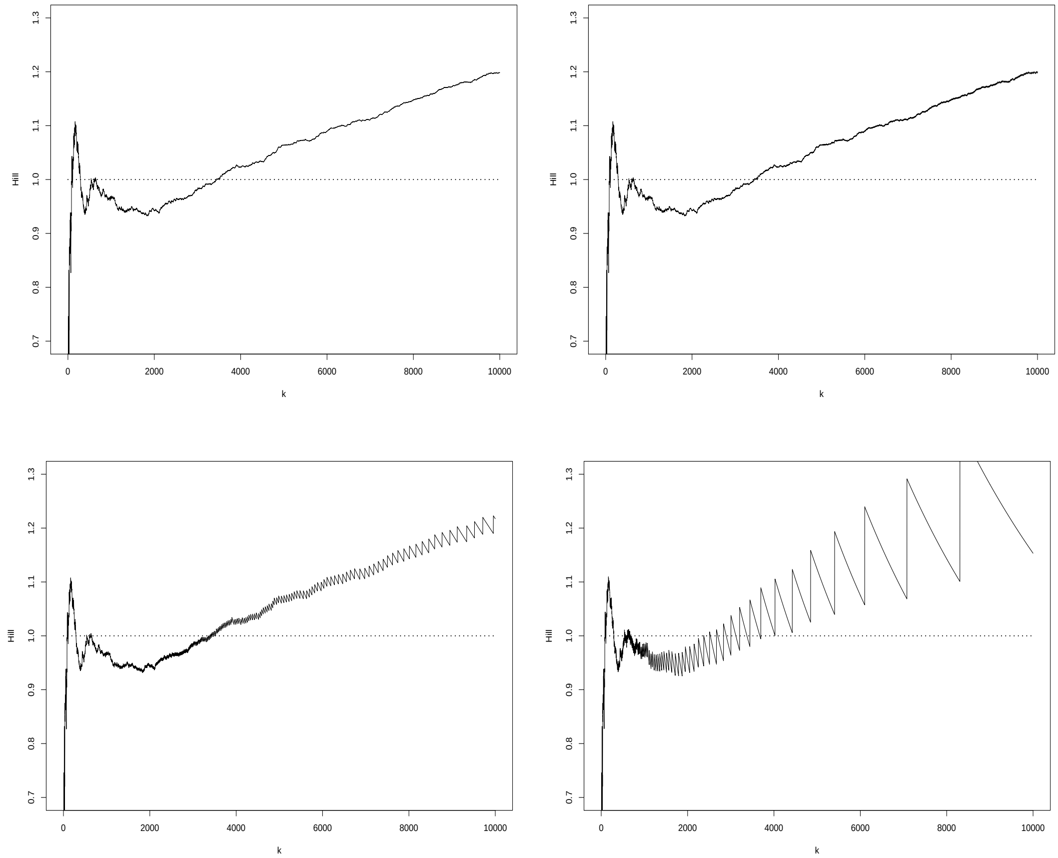

Figure 1. Hill plot $(k, \widehat{\alpha}_k^{-1})$, $k = 2, \ldots, n-1$, based on the counts of 10 000 words which are used most frequently in the English language. Notice that the Hill plot yields a reliable estimator only for small k, up to 1000 say. *Top left.* The estimated tail index of the count data is close to one. *Top right.* The last digit in the counts is replaced by zero. *Bottom left.* The last two digits are replaced by zeros. *Bottom right.* The last three digits are replaced by zeros.

Standard results about asymptotic normality of the Hill estimator are not available in this case since such a theory requires that a second order condition on $\overline{F}$ must be satisfied. According to de Haan and Ferreira [8], Theorem 3.2.5, asymptotic normality of $\widehat{\alpha}_k$ can be achieved if the following second order condition holds as $x \to \infty$ for $t > 0$

$$\frac{\overline{F}(tx)}{\overline{F}(x)} - t^{-\alpha} \sim b(t)a(x)\,,$$

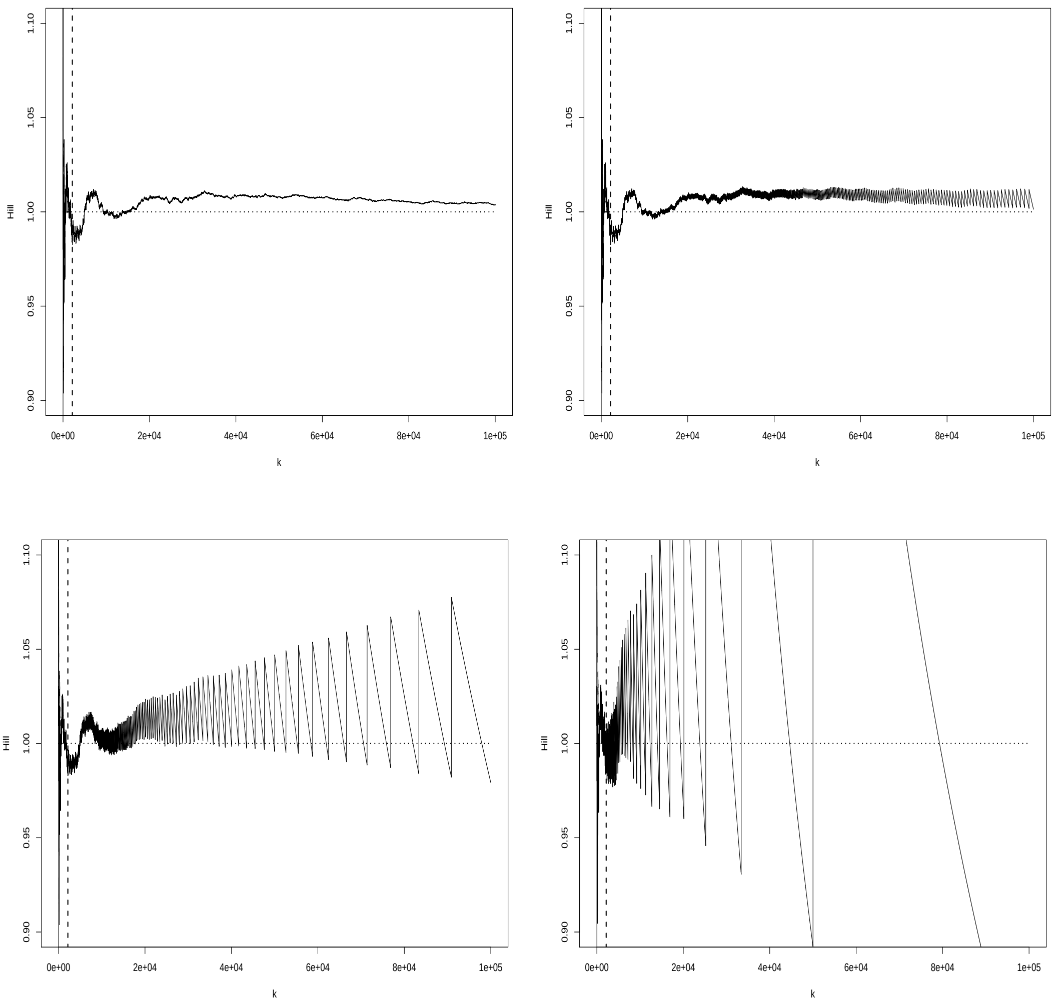

FIGURE 2. Hill plot $(k, \widehat{\alpha}_k^{-1})$, $k = 2, \ldots, n-1$, for a sample of size $n = 10\,000$. Top left. The data have a Pareto distribution (1.5) with $\alpha = 1$. In the other figures all digits but the first l behind the comma are set equal to zero. Top right. $l = 2$. Bottom left. $l = 1$. Bottom right. $l = 0$. The vertical line at $k = n^{2/3}$ is an upper limit for those k for which the central limit theorem is still valid; see Corollary 2.4.

where $|a(x)|$ is regularly varying with a non-positive index and $b(t)$ is a positive function of t. We observe that ($\{x\}$ denotes the fractional part of x)

$$\begin{aligned}(1.7)\qquad \frac{\overline{F}(tx)}{\overline{F}(x)} - t^{-\alpha} &= \frac{([10^l x]+1)^\alpha}{([10^l tx]+1)^\alpha} - t^{-\alpha} \\ &\sim t^{-\alpha}(10^l x)^{-1}\alpha\big(-\{10^l x\} + (1-t^{-1}) + t^{-1}\{10^l tx\}\big)\,.\end{aligned}$$

The right-hand side exhibits very erratic behavior. For irrational $10^l t$, the sequence $(\{10^l tx\})_{x=1,2,\ldots}$ is uniformly distributed in the number theoretical sense; see Weyl [16]. In particular, it visits any interval $(a,b)\subset(0,1)$ infinitely often. Then the sequence

$$-\{10^l x\}+(1-t^{-1})+t^{-1}\{10^l tx\} \quad = \quad 1+t^{-1}(-1+\{10^l tx\})\,,\quad x=1,2,\ldots\,,$$

is uniformly distributed on $(1-t^{-1},1)$. If x assumes the integers $1,2,\ldots$ and $10^l t$ is an integer, the right-hand side in (1.7) vanishes. Hence $|\overline{F}(tx)/\overline{F}(x)-t^{-\alpha}|$ is not a regularly varying function as required. However, asymptotic normality of $\widehat{\alpha}_k^{-1}$ can still be derived from the corresponding results for $(U_i^{-1/\alpha})$ if $k_n=o(n^{2/(2+\alpha)})$; see Lemma 2.4 below. We show a similar result for the Pickands and Dekkers-Einmahl-de Haan estimators. We have simulation evidence showing that these estimators fail for k_n of a magnitude larger than $n^{2/(2+\alpha)}$.

Our paper is organized as follows. In Section 2 we give some theoretical explanation for the erratic behavior of the mentioned tail index estimators in the presence of discretized data. We calculate the expectation and variance of the Hill estimator $\widehat{\alpha}_k^{-1}$ for discretized Pareto random variables and provide bounds for the deviation of this estimator from the pure Pareto case.

2. Basic properties of the Hill estimator for an integer-valued sequence

Throughout we assume that the iid sequence (X_i) is given by (1.4). Recall that for an iid uniform $U(0,1)$ distributed sequence (U_i), the ith order statistic $U_{(i)}$ has a $\beta(i,n-i+1)$ density (see e.g. [4], Proposition 4.1.2) given by

$$\beta(i,n-i+1)(x)=\frac{n!}{(n-i)!(i-1)!}x^{i-1}(1-x)^{(n-i+1)-1}\,,\quad x\in(0,1)\,. \tag{2.1}$$

From (1.3) it follows that the Hill estimator $\widehat{\alpha}_k^{-1}$ is an unbiased estimator of α^{-1}. The situation changes in the case of round-off effects:

Lemma 2.1. *Consider the sequence $X_i^{(l)}=10^{-l}[10^l U_i^{-1/\alpha}]$, $i=1,2,\ldots$, for a fixed integer $l\ge 0$. Then the following relation holds for the Hill estimator $\widehat{\alpha}_{k,l}^{-1}$ based on $X_1^{(l)},\ldots,X_n^{(l)}$, $n\ge 3$, $2\le k\le n-1$:*

$$E\widehat{\alpha}_{k,l}^{-1}$$

$$= \sum_{s=10^l+1}^{\infty}\log\frac{s}{s-1}\,\frac{1}{k-1}\sum_{i=1}^{k-1}\int_0^{(10^l/s)^\alpha}\big(\beta(i,n-i+1)(x)-\beta(k,n-k+1)(x)\big)\,dx\,, \tag{2.2}$$

$$= \frac{n}{k-1}\sum_{s=10^l+1}^{\infty}\log\frac{s}{s-1}(10^l/s)^\alpha\int_{(10^l/s)^\alpha}^{1}\beta(k-1,n-k+1)(x)\,dx\,. \tag{2.3}$$

Proof of (2.2). By the scale invariance of the Hill estimator,

$$E\widehat{\alpha}_{k,l}^{-1}=\frac{1}{k-1}\sum_{i=1}^{k-1}E\log[10^l U_{(i)}^{-1/\alpha}]-E\log[10^l U_{(k)}^{-1/\alpha}]\,,$$

where

$$\begin{aligned} E\log[10^l U_{(i)}^{-1/\alpha}] &= \sum_{s=10^l}^{\infty}\log s\,P((10^l/(s+1))^\alpha\le U_{(i)}\le((10^l/s)^\alpha)\\ &= \sum_{s=10^l}^{\infty}\log s\int_{(10^l/(s+1))^\alpha}^{(10^l/s)^\alpha}\beta(i,n-i+1)(x)\,dx \\ &= \sum_{s=10^l+1}^{\infty}\log(s/(s-1))\int_0^{(10^l/s)^\alpha}\beta(i,n-i+1)(x)\,dx+\log 10^l\,. \qquad (2.4)\end{aligned}$$

In the last step we used Abel's formula. Then (2.2) follows. □

Proof of (2.3). For an iid standard exponential sequence (E_i) write $\Gamma_i = E_1 + \cdots + E_i$, $i \ge 1$. Then it is well known (e.g. p. 189 in [4]) that $(U_{(i)})_{i=1,\ldots,k} \stackrel{d}{=} \Gamma_{n+1}^{-1}(\Gamma_i)_{i=1,\ldots,k}$. If we now condition on $\Gamma_k/\Gamma_{n+1} = u$ on the right-hand side we see that

$$((U_{(1)},\ldots,U_{(k-1)}) \mid U_{(k)} = u) \stackrel{d}{=} u\Gamma_k^{-1}(\Gamma_1,\ldots,\Gamma_{k-1})\,.$$

Hence the left-hand side has the same distribution as $u\,(U^{(k)}_{(1)},\ldots,U^{(k)}_{(k-1)})$, where $U^{(k)}_{(1)} \le \cdots \le U^{(k)}_{(k-1)}$ are the order statistics of an iid $U(0,1)$ sample $U^{(k)}_1,\ldots,U^{(k)}_{k-1}$. Thus, for almost every $u \in (0,1)$,

$$\begin{aligned} & E(\widehat{\alpha}^{-1}_{k,l} \mid U_{(k)} = u) + \log([10^l u^{-1/\alpha}]) \\ = & \frac{1}{k-1}\sum_{i=1}^{k-1} E(\log([10^l U_{(i)}^{-1/\alpha}]) \mid U_{(k)} = u) \\ = & \frac{1}{k-1}\sum_{i=1}^{k-1} E\Big(\log([10^l (u\,U_i)^{-1/\alpha}])\Big) \\ = & E\Big(\log([10^l (uU_1)^{-1/\alpha}])\Big)\,. \end{aligned}$$

The right-hand side can be written as follows

$$\begin{aligned} & \sum_{s+1>10^l u^{-1/\alpha}}^{\infty} \log s\; P((10^l/(s+1))^\alpha \le u\,U_1 \le (10^l/s)^\alpha) \\ = & \sum_{s=10^l}^{\infty} \log s\; I_{\{s+1>10^l u^{-1/\alpha}\}}\Big(10^{l\alpha}u^{-1}(s^{-\alpha}-(s+1)^{-\alpha})I_{\{(10^l/s)^\alpha \le u\}} \\ & \qquad +\big(1-(10^l/(s+1))^\alpha u^{-1}\big)I_{\{(10^l/(s+1))^\alpha < u \le (10^l/s)^\alpha\}}\Big) \\ = & \sum_{s=10^l}^{\infty} \log s\,\Big(10^{l\alpha}u^{-1}(s^{-\alpha}-(s+1)^{-\alpha})I_{\{u\ge(10^l/s)^\alpha\}} \\ & \qquad +\big(1-(10^l/(s+1))^\alpha u^{-1}\big)I_{\{(10^l/(s+1))^\alpha < u \le (10^l/s)^\alpha\}}\Big) \\ = & \sum_{s=10^l}^{\infty} \log s\,\Big(10^{l\alpha}u^{-1}s^{-\alpha}I_{\{u\ge(10^l/s)^\alpha\}} + I_{\{u\le(10^l/s)^\alpha\}} \\ & \qquad -10^{l\alpha}u^{-1}(s+1)^{-\alpha}I_{\{u>(10^l/(s+1))^\alpha\}} - I_{\{u\le(10^l/(s+1))^\alpha\}}\Big). \end{aligned}$$

Take expectations with respect to the distribution of $U_{(k)}$ and recall (2.4) to obtain

$$\begin{aligned} E\widehat{\alpha}^{-1}_{k,l} &= \sum_{s=10^l+1}^{\infty} \log\frac{s}{s-1}\Big(10^{l\alpha}s^{-\alpha}EU_{(k)}^{-1}I_{\{U_{(k)}\ge(10^l/s)^\alpha\}} + P(U_{(k)} \le (10^l/s)^\alpha)\Big) + \log 10^l \\ & \quad -E\log[10^l U_{(k)}^{-1/\alpha}] \\ &= \sum_{s=10^l+1}^{\infty} \log\frac{s}{s-1} 10^{l\alpha}s^{-\alpha}\int_{(10^l/s)^\alpha}^{1} u^{-1}\beta(k,n-k+1)(u)\,du\,. \end{aligned}$$

□

Figure 3 exhibits $E\widehat{\alpha}^{-1}_{k,l}$ as a function of k for $\alpha = 1$, $l = 0,1,2$, and sample size $n = 10\,000$. Evidently, the erratic behavior of the Hill estimators $\widehat{\alpha}^{-1}_{k,l}$ is also inherited by its mean value function.

It shows significant deviations from the value α^{-1}, in particular for large k. This fact is a clear warning against using the maximum likelihood estimator $\widehat{\alpha}_n$ of α. Figures 2 and 3 give some convincing evidence that the Hill estimator based on a relatively small number k of upper order statistics (in agreement with the conditions $k = k_n \to \infty$ and $k_n/n \to 0$) provides some reasonable approximations of α^{-1}. On the other hand, the estimator $\widehat{\alpha}_n^{-1}$ yields the best results (see the top left graph in Figure 2) only if X_i has an exact Pareto distribution and $\widehat{\alpha}_{n,l}^{-1}$ is extremely unreliable for the discretized Pareto random variables $X_i^{(l)}$. The variance of the Hill estimator $\widehat{\alpha}_{k,l}^{-1}$ for round-off data is different from the pure Pareto case and very complicated; see Appendix A. Based on the latter result, in Figure 4 we show the graphs of the variance as a function of l and k. As expected, the variance of the Hill estimator oscillates as a function of k. The size of the oscillations decreases as the round-off error becomes smaller and the frequency is high for large l. In the case $l = 2$, the oscillations are rather tiny. In this case, the numerical calculations take an enormous time, and we decided to restrict ourselves to $k \in [9\,000, 10\,000]$; for $k \le 9\,000$ one rarely observes an oscillation.

In the next result we measure the deviation of $E\widehat{\alpha}_{k,l}^{-1}$ from $E\widehat{\alpha}_k^{-1} = \alpha^{-1}$.

Proposition 2.2. *Under the conditions of Lemma 2.1, for $2 \le k \le n-1$, $n \ge 3$, $l \ge 0$,*

$$-10^{-l} E\big(U_{(k)}^{1/\alpha}\big) \le E\widehat{\alpha}_{k,l}^{-1} - \alpha^{-1} \le 10^{-l}(1+10^{-l}) E\big(U_{(k)}^{1/\alpha}\big)\,. \tag{2.5}$$

Moreover, writing $D_k^{(1)} = \widehat{\alpha}_k^{-1} - \widehat{\alpha}_{k,l}^{-1}$, we have for any $p > 0$,

$$E|D_k^{(1)}|^p \le 10^{-lp}(1+10^{-l})^p E\big(U_{(k)}^{p/\alpha}\big)\,. \tag{2.6}$$

The left and right hand sides in (2.5) and (2.6) converge to zero as $n \to \infty$ if $k = k_n \to \infty$ and $k/n \to 0$ or k is fixed and $l \to \infty$.

Proof. We will exploit the fact that

$$D_k^{(1)} \stackrel{d}{=} \widetilde{D}_k^{(1)} = \frac{1}{k-1}\sum_{i=1}^{k-1} \log \frac{U_{(i)}^{-1/\alpha}}{U_{(k)}^{-1/\alpha}} - \frac{1}{k-1}\sum_{i=1}^{k-1} \log \frac{[10^l U_{(i)}^{-1/\alpha}]}{[10^l U_{(k)}^{-1/\alpha}]}\,.$$

We will frequently use the inequality for $i \le k$:

$$\begin{aligned} 0 &\le \log \frac{10^l U_{(i)}^{-1/\alpha}}{[10^l U_{(i)}^{-1/\alpha}]} \le \log\Big(1 + \frac{1}{[10^l U_{(i)}^{-1/\alpha}]}\Big) \le \frac{1}{[10^l U_{(i)}^{-1/\alpha}]} \\ &\le \frac{10^l U_{(i)}^{-1/\alpha}}{[10^l U_{(i)}^{-1/\alpha}]} 10^{-l} U_{(i)}^{1/\alpha} \le (1+10^{-l}) 10^{-l} U_{(i)}^{1/\alpha}. \end{aligned} \tag{2.7}$$

Then straightforward calculation yields

$$\begin{aligned} E\widehat{\alpha}_{k,l}^{-1} - \alpha^{-1} &\le E\Big(\log \frac{10^l U_{(k)}^{-1/\alpha}}{[10^l U_{(k)}^{-1/\alpha}]}\Big) \\ &\le 10^{-l}(1+10^{-l}) E\big(U_{(k)}^{1/\alpha}\big) \to 0\,, \quad n \to \infty\,, \end{aligned}$$

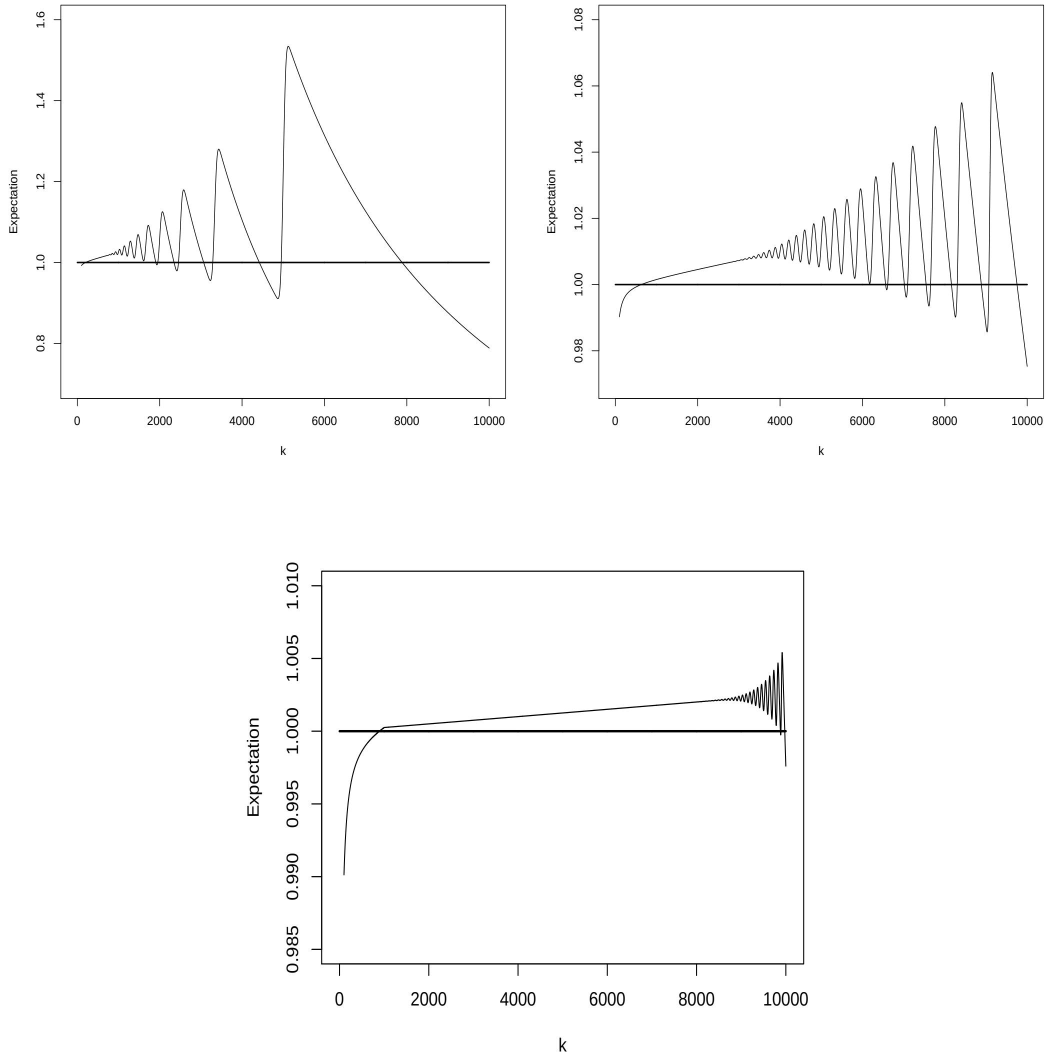

FIGURE 3. The mean value function of the Hill estimator $\widehat{\alpha}_{k,l}^{-1}$ for a sample of size $n = 10\,000$ of discretized Pareto distributed random variables $X_1^{(l)}, \ldots, X_n^{(l)}$ with parameter $\alpha = 1$. Top left: $l = 0$. Top right: $l = 1$. Bottom: $l = 2$.

using dominated convergence and $U_{(k)} \stackrel{\text{a.s.}}{\to} 0$ as $k = k_n \to \infty$ and $k/n \to 0$. Similarly,

$$
\begin{aligned}
E\widehat{\alpha}_{k,l}^{-1} - \alpha^{-1} &\geq E\Big(\frac{1}{k-1} \sum_{i=1}^{k-1} \log \frac{[10^l U_{(i)}^{-1/\alpha}]}{10^l U_{(i)}^{-1/\alpha}}\Big) \\
&\geq -\frac{1}{k-1} \sum_{i=1}^{k-1} E\Big(\frac{U_{(i)}^{1/\alpha}}{10^l}\Big) \geq -10^{-l} E\big(U_{(k)}^{1/\alpha}\big) \to 0\,, \quad n \to \infty\,.
\end{aligned}
$$

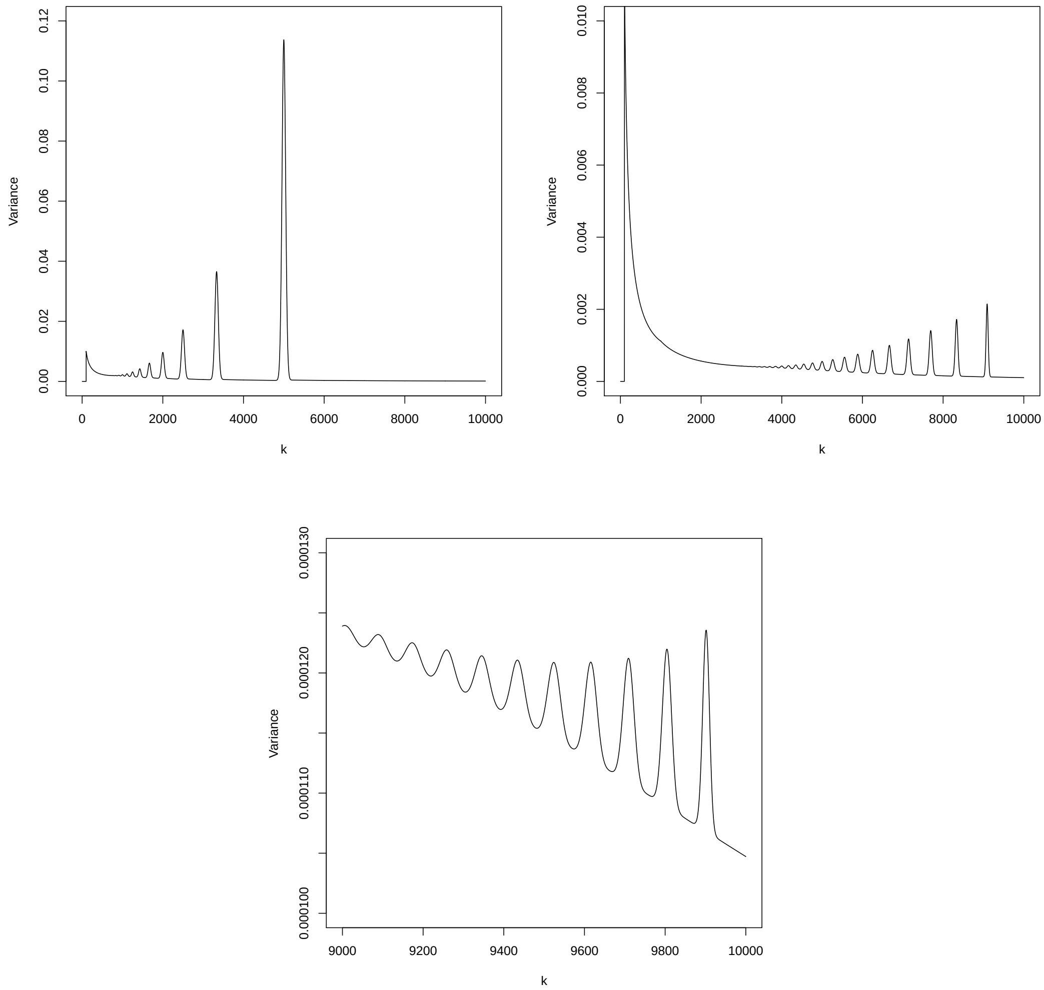

Figure 4. The variance function of the Hill estimator $\widehat{\alpha}_{k,l}^{-1}$ for a sample of size $n = 10\,000$ of discretized Pareto distributed random variables $X_1^{(l)}, \ldots, X_n^{(l)}$ with parameter $\alpha = 1$. Top left: $l = 0$. Top right: $l = 1$. Bottom: $l = 2$ and for $9\,000 \le k \le 10\,000$.

This proves (2.5).

Next we observe that

$$
\begin{aligned}
-\log \frac{10^l U_{(k)}^{-1/\alpha}}{[10^l U_{(k)}^{-1/\alpha}]} \quad &\le \quad \widetilde{D}_k^{(1)} = \frac{1}{k-1} \sum_{i=1}^{k-1} \log \frac{10^l U_{(i)}^{-1/\alpha}}{[10^l U_{(i)}^{-1/\alpha}]} - \log \frac{10^l U_{(k)}^{-1/\alpha}}{[10^l U_{(k)}^{-1/\alpha}]} \\
&\le \quad \frac{1}{k-1} \sum_{i=1}^{k-1} \log \frac{10^l U_{(i)}^{-1/\alpha}}{[10^l U_{(i)}^{-1/\alpha}]}
\end{aligned}
$$

and apply (2.7). We conclude that (2.6) holds for any $p > 0$. □

Before we proceed further we recall the following benchmark result. Its proof is an immediate consequence of the representation (1.3).

Lemma 2.3. *Let (X_i) be an iid sequence with common Pareto distribution defined in* (1.5). *Assume that $k = k_n \to \infty$ and $k_n \le n$. Then*

$$\sqrt{k}\,(\widehat{\alpha}_k^{-1} - \alpha^{-1}) \stackrel{d}{\to} Y \quad \text{and} \quad \sqrt{k}\,(\widehat{\alpha}_k - \alpha) \stackrel{d}{\to} Z,$$

where Y and Z are normally distributed with mean zero and variances α^{-2} and α^2, respectively.

A combination of Proposition 2.2 and Lemma 2.3 yields the following result.

Corollary 2.4. *Assume the conditions of Lemma 2.1 and that $k = k_n \to \infty$ and $k = o(n^{2/(\alpha+2)})$. Then for a normal $N(0, \alpha^{-2})$ distributed random variable Y,*

$$\begin{aligned} \sqrt{k}\,(\widehat{\alpha}_{k,l}^{-1} - \alpha^{-1}) &\stackrel{d}{\to} Y\,, \\ E\Big(\big|\sqrt{k}\,(\widehat{\alpha}_{k,l}^{-1} - \alpha^{-1})\big|^p\Big) &\to E\big(|Y|^p\big)\,. \end{aligned}$$

Proof. Recall the representation (1.3) in law of the Hill estimator $\widehat{\alpha}_k^{-1}$. Then the central limit theorem $\sqrt{k}(\widehat{\alpha}_k^{-1} - \alpha^{-1}) \stackrel{d}{\to} Y$ and the relation $E\Big(\big|\sqrt{k}(\widehat{\alpha}_k^{-1} - \alpha^{-1})\big|^p\Big) \to E\big(|Y|^p\big)$ follow for any $p > 0$; see Theorem 4.2 in Gut [7]. In view of (2.6) the desired results follow if $E|\sqrt{k}D_k^{(1)}|^p = o(1)$. Recalling the density (2.1), an application of Stirling's formula yields

$$E\big(U_{(k)}^{p/\alpha}\big) = \frac{\Gamma(n+1)}{\Gamma((p/\alpha)+n+1)}\frac{\Gamma((p/\alpha)+k)}{\Gamma(k)} \sim (k/n)^{p/\alpha}\,.$$

Thus $E|\sqrt{k}D_k^{(1)}|^p = o(1)$ holds if $k = o(n^{2/(\alpha+2)})$. □

We could not prove whether Corollary 2.4 is optimal in the sense that it does not hold for $k \ge c n^{2/(\alpha+2)}$. However, simulations indicate that the Hill estimator $\widehat{\alpha}_{k,l}^{-1}$ is very unreliable for such k-values. In the next section, we make an excursion to two other classical estimators of the extreme value index. Also in these cases, $k = n^{2/(\alpha+2)}$ appears as a borderline case for central limit behavior of the corresponding estimators.

3. An excursion to the moment and Pickands estimators

There exists a wide range of estimators of the tail index of a distribution with regularly varying tail and, more generally, of the extreme value index of a distribution; cf. de Haan and Ferreira [8], Embrechts et al. [4]. We consider two classical estimators, the *moment estimator* (or *Dekkers-Einmahl-de Haan (DEdH)* estimator) and the *Pickands estimator.* In what follows, we again assume that (X_i) is an iid Pareto sequence with representation (1.4) and $X_i^{(l)} = 10^{-l}[10^l U_i^{-1/\alpha}]$, $l \ge 0$.

3.1. Asymptotic normality for the DEdH estimator for discretized Pareto variables.

The DEdH estimator of α^{-1} is given by the relation

$$\widehat{\alpha}_k^{-1} = M_k^{(1)} + 1 - \frac{1}{2}\Big(1 - \frac{(M_k^{(1)})^2}{M_k^{(2)}}\Big)^{-1}\,,$$

where

$$M_k^{(j)} = \frac{1}{k-1}\sum_{i=1}^{k-1}\Big(\log\frac{X_{(n-i+1)}}{X_{(n-k+1)}}\Big)^j\,, \quad j = 1, 2.$$

If we replace (X_i) by $(X_i^{(l)})$ we adapt the notation correspondingly by adding the subscript l, e.g. $\widehat{\alpha}_{k,l}^{-1}$, $M_{k,l}^{(j)}$.

Lemma 3.1. *Assume that $k = k_n \to \infty$ and $k \le n$. Then*

$$\sqrt{k}\big(\widehat{\alpha}_k^{-1} - \alpha^{-1}\big) \stackrel{d}{\to} Y \tag{3.1}$$

for a normal $N(0, \alpha^{-2} + 1)$ distributed random variable Y.

The fact that the DEdH estimator is asymptotically normal is well known under second order assumptions on the tail and resulting restrictions on k; cf. de Haan and Ferreira [8], Theorem 3.5.4. The latter result is not directly applicable since the Pareto case is a degenerate one. We give a short proof of (3.1).

Proof. We have

$$\Delta_k = \sqrt{k}\,(\widehat{\alpha}_k^{-1} - \alpha^{-1}) = \sqrt{k}\,(M_k^{(1)} - \alpha^{-1}) + \sqrt{k}\,\frac{M_k^{(2)} - 2(M_k^{(1)})^2}{2(M_k^{(2)} - (M_k^{(1)})^2)}.$$

The same argument as on pp. 191-192 in [4] and the law of large numbers show that

$$(M_k^{(1)}, M_k^{(2)}) \stackrel{d}{=} \frac{1}{k-1}\Big(\alpha^{-1}\sum_{i=1}^{k-1} E_i, \alpha^{-2}\sum_{i=1}^{k-1} E_i^2\Big) \stackrel{P}{\to} (\alpha^{-1}, 2\alpha^{-2}), \tag{3.2}$$

where (E_i) is iid standard exponential. In what follows, we will use the representation (3.2) of $(M_k^{(1)}, M_k^{(2)})$ via the exponential random variables. Then

$$\begin{aligned}
&\Delta_k - \Big[k^{-1/2}\alpha^{-1}\sum_{i=1}^{k}(E_i - 1) + o_P(1)\Big] \\
&= k^{1/2} 0.5(1 + o_P(1))\Big(\frac{1}{k-1}\sum_{i=1}^{k-1}(E_i^2 - 2) + 2(1 - (\frac{1}{k-1}\sum_{i=1}^{k-1} E_i)^2)\Big) \\
&= k^{1/2} 0.5(1 + o_P(1))\Big(\frac{1}{k-1}\sum_{i=1}^{k-1}(E_i^2 - 2) + 2\frac{1}{k-1}\sum_{i=1}^{k-1}(1 - E_i)\frac{1}{k-1}\sum_{i=1}^{k-1}(1 + E_i)\Big) \\
&= k^{-1/2} 0.5(1 + o_P(1))\sum_{i=1}^{k-1}(E_i^2 - 4E_i + 2).
\end{aligned}$$

By the central limit theorem,

$$\Delta_k \stackrel{d}{\to} N(0, \mathrm{var}((\alpha^{-1} - 2)E_i + 0.5E_i^2)).$$

Straightforward calculation yields $\mathrm{var}((\alpha^{-1} - 2)E_i + 0.5E_i^2) = \alpha^{-2} + 1$. This concludes the proof. □

Corollary 3.2. *Assume the conditions of Lemma 2.1 and that $k = k_n \to \infty$ and $k = o(n^{2/(\alpha+2)})$. Then for a normal $N(0, \alpha^{-2} + 1)$ distributed random variable Y,*

$$\sqrt{k}(\widehat{\alpha}_{k,l}^{-1} - \alpha^{-1}) \stackrel{d}{\to} Y. \tag{3.3}$$

It is again not obvious whether (3.3) is valid for $k \ge n^{2/(\alpha+2)}$. Due to the presence of ratios of random variables in the definition of the DEdH estimator it seems difficult to get explicit bound for the moments $E\big(|\sqrt{k}(\widehat{\alpha}_{k,l}^{-1} - \alpha^{-1})|^p\big)$, $p > 0$. This is in contrast to the Hill estimator of α^{-1}.

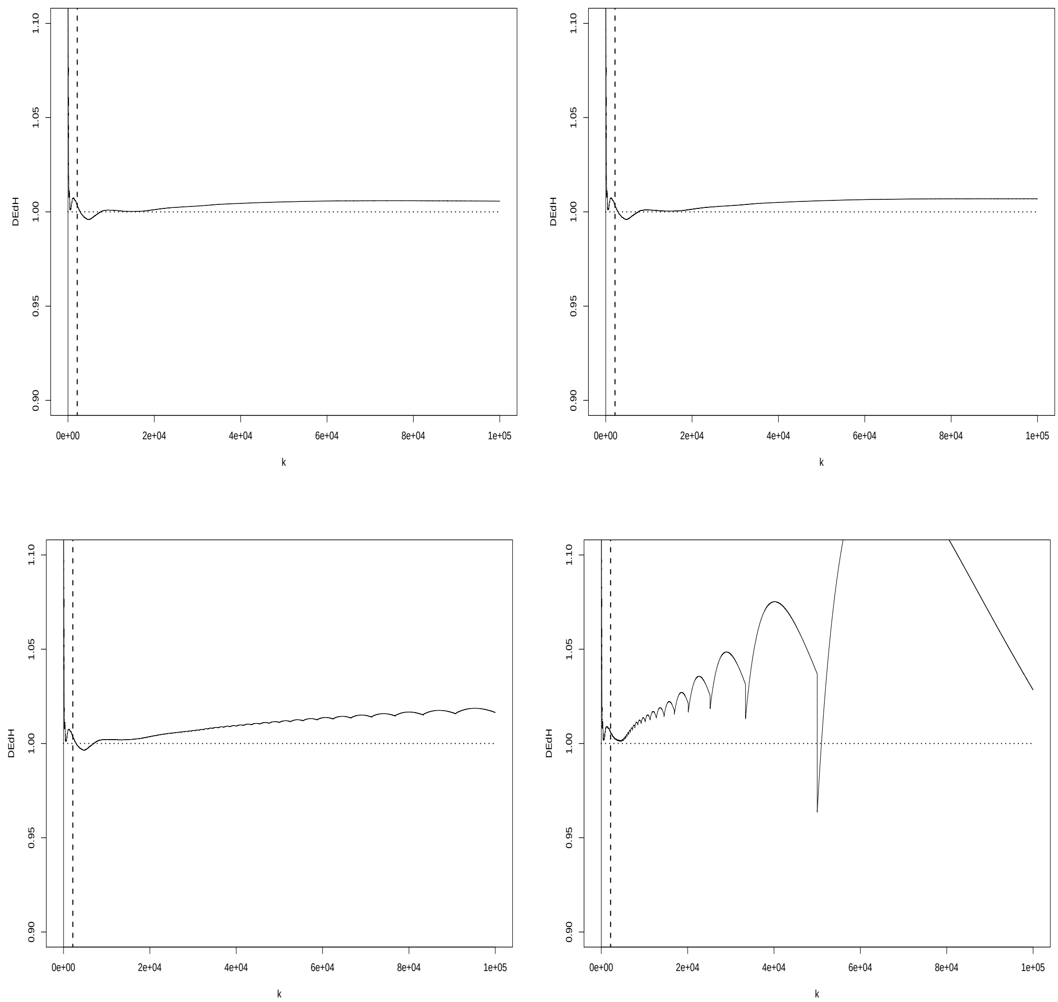

FIGURE 5. DEdH plot $(k, \widehat{\alpha}_k^{-1})$, $k = 2, \dots, n-1$, for a sample of size $n = 10\,000$. Top left. The data have a Pareto distribution (1.5) with $\alpha = 1$. In the other figures all digits but the first l behind the comma are set equal to zero. Top right. $l = 2$. Bottom left. $l = 1$. Bottom right. $l = 0$. The vertical line shows the value $k = n^{2/3}$ which is an upper bound for those k for which the central limit theorem is still valid; see Corollary 3.2.

In Figure 5 we show the DEdH estimator for the same sample as in Figure 2 and the corresponding plots for the discretized data. The vertical line shows the value $k = n^{2/(\alpha+2)}$ which is an upper bound for those k for which the central limit theorem is still valid.

Proof. Observe that

$$\begin{aligned} D_k^{(2)} &= M_k^{(2)} - M_{k,l}^{(2)} \\ &\stackrel{d}{=} \frac{1}{k-1}\sum_{i=1}^{k-1}\Big\{\Big(\log\frac{10^l U_{(i)}^{-1/\alpha}}{10^l U_{(k)}^{-1/\alpha}}\Big)^2 - \Big(\log\frac{[10^l U_{(i)}^{-1/\alpha}]}{[10^l U_{(k)}^{-1/\alpha}]}\Big)^2\Big\} \\ &= \widetilde{D}_k^{(2)}\,. \end{aligned}$$

Therefore, using (2.7),

$$\begin{aligned} |\widetilde{D}_k^{(2)}| &= \frac{1}{k-1}\Big|\sum_{i=1}^{k-1}\Big(\log\frac{10^l U_{(i)}^{-1/\alpha}}{10^l U_{(k)}^{-1/\alpha}} + \log\frac{[10^l U_{(i)}^{-1/\alpha}]}{[10^l U_{(k)}^{-1/\alpha}]}\Big)\Big(\log\frac{10^l U_{(i)}^{-1/\alpha}}{[10^l U_{(i)}^{-1/\alpha}]} - \log\frac{10^l U_{(k)}^{-1/\alpha}}{[10^l U_{(k)}^{-1/\alpha}]}\Big)\Big| \\ &\le 10^{-l}(1+10^{-l})U_{(k)}^{1/\alpha}\big(M_l^{(1)} + M_{k,l}^{(1)}\big)\,. \end{aligned}$$

The Hill estimators $M_l^{(1)}$, $M_{k,l}^{(1)}$ are consistent estimators of α^{-1}. Hence $\sqrt{k}D_k^{(2)} \stackrel{P}{\to} 0$ if $k = o(n^{2/(\alpha+2)})$. Under the latter condition, we also know from Proposition 2.2 that $\sqrt{k}D_k^{(1)} \stackrel{P}{\to} 0$. Then an application of the Δ-method and Lemma 3.1 imply that $M_k^{(i)}$ can be replaced by $M_{k,l}^{(i)}$ for $i = 1, 2$ in the definition of $\widehat{\alpha}_k^{-1}$, leading to the central limit theorem (3.3). □

3.2. **Asymptotic normality of Pickands's estimator for discretized Pareto variables.** The Pickands estimator of the extreme value index α^{-1} is defined as

$$\widehat{\alpha}_k^{-1} = \frac{1}{\log 2}\log\frac{X_{(n-k+1)} - X_{(n-2k+1)}}{X_{(n-2k+1)} - X_{(n-4k+1)}}\,, \quad k \ge 1\,.$$

If we replace (X_i) by $(X_i^{(l)})$ we write $\widehat{\alpha}_{k,l}^{-1}$, $l \ge 0$. We will give the asymptotic results analogous to the DEdH estimator and start with the asymptotic normality for $\widehat{\alpha}_k^{-1}$.

Lemma 3.3. *Assume that $k = k_n \to \infty$ and $k/n \to 0$. Then*

$$\sqrt{k}\big(\widehat{\alpha}_k^{-1} - \alpha^{-1}\big) \stackrel{d}{\to} Y \tag{3.4}$$

for a normal $N\big(0, \alpha^{-2}(2^{1+2/\alpha}+1)/(4(\log 2)^2(2^{1/\alpha}-1)^2)\big)$ distributed random variable Y.

The fact that the Pickands estimator is asymptotically normal is well known under second order assumptions on the tail and resulting restrictions on k; cf. de Haan and Ferreira [8], Theorem 3.3.5. The latter is not directly applicable since the Pareto case is a degenerate one. The proof of (3.4) can be given by direct calculation, using

$$\widehat{\alpha}_k^{-1} \stackrel{d}{=} \frac{1}{\log 2}\log\frac{U_{(k)}^{-1/\alpha} - U_{(2k)}^{-1/\alpha}}{U_{(2k)}^{-1/\alpha} - U_{(4k)}^{-1/\alpha}}\,,$$

a result of Smirnov [13] (see Lemma 3.3.2 of de Haan and Ferreira [8]): if $k = k_n \to \infty$, $k/n \to 0$, then

$$\sqrt{k}\Big(\frac{\sqrt{2}U_{(2k)}}{2U_{(k)}} - \sqrt{2}, \frac{U_{(4k)}}{U_{(2k)}} - 2\Big) \stackrel{d}{\to} \mathbf{Y}\,,$$

for a bivariate standard normal vector $\mathbf{Y}$, and applying the Δ-method. We omit further details and refer to the argument in Theorem 3.3.5 of [8] which simplifies in the Pareto setting. A consequence is the following result.

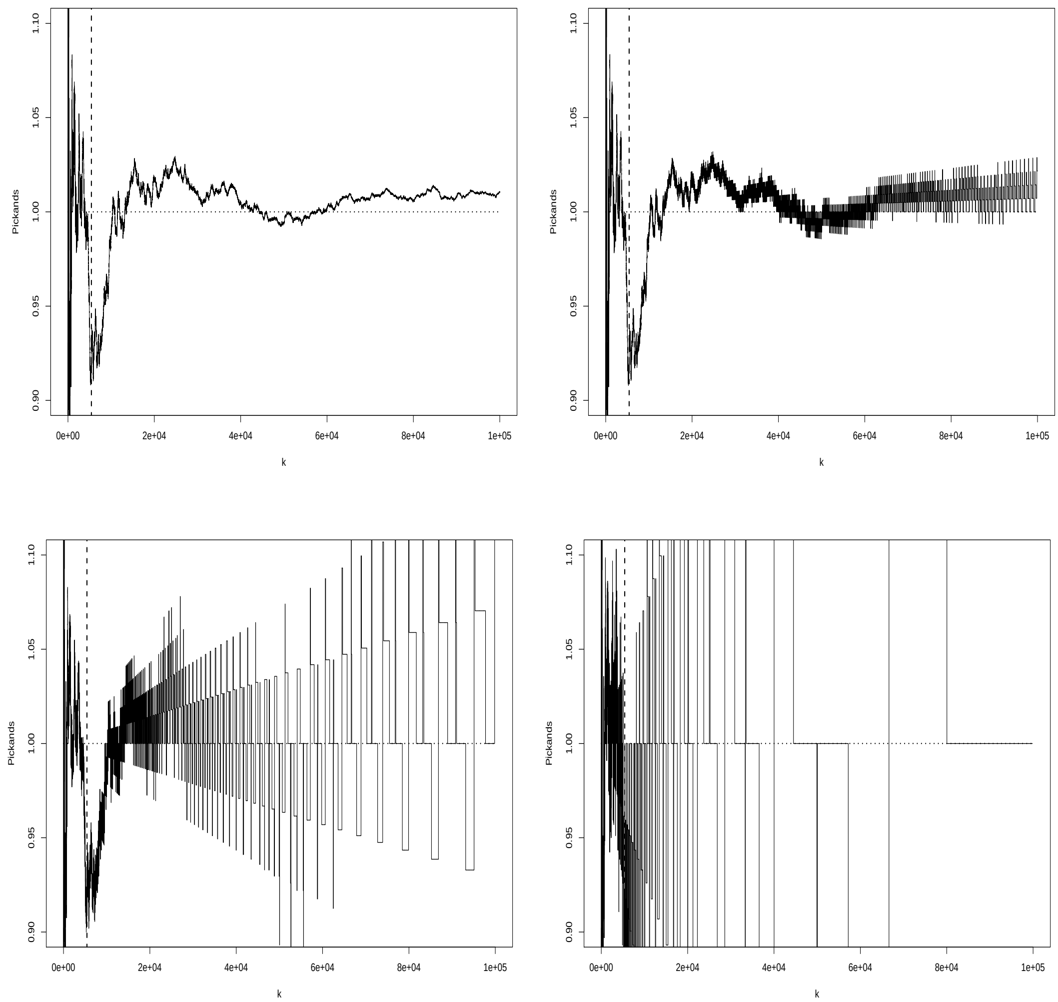

FIGURE 6. Pickands plot $(k, \widehat{\alpha}_k^{-1})$, $k = 2, \ldots, n-1$, for a sample of size $n = 4 * 10\,000$. Top left. The data have a Pareto distribution (1.5) with $\alpha = 1$. In the other figures all digits but the first l behind the comma are set equal to zero. Top right. $l = 2$. Bottom left. $l = 1$. Bottom right. $l = 0$. The vertical line shows the value $k = n^{2/3}$ which is an upper bound for those k for which the central limit theorem is still valid; see Corollary 3.4.

Corollary 3.4. *Assume the conditions of Lemma 2.1 and that $k = k_n \to \infty$ and $k = o(n^{2/(\alpha+2)})$. Then for a normal $N\big(0, \alpha^{-2}(2^{1+2/\alpha}+1)/(4(\log 2)^2(2^{1/\alpha}-1)^2)\big)$ distributed random variable Y,*

$$\sqrt{k}\,(\widehat{\alpha}_{k,l}^{-1} - \alpha^{-1}) \quad \stackrel{d}{\to} \quad Y\,.$$

In Figure 6 we illustrate the behavior of the Pickands estimator for the same sample as in Figure 2 and for the corresponding discretized data.

Proof. We observe that

$$\begin{aligned} D_k^{(3)} &= \log 2\big(\widehat{\alpha}_k^{-1} - \widehat{\alpha}_{k,l}^{-1}\big) \\ &\stackrel{d}{=} \log \frac{10^l U_{(k)}^{-1/\alpha} - 10^l U_{(2k)}^{-1/\alpha}}{[10^l U_{(k)}^{-1/\alpha}] - [10^l U_{(2k)}^{-1/\alpha}]} - \log \frac{10^l U_{(2k)}^{-1/\alpha} - 10^l U_{(4k)}^{-1/\alpha}}{[10^l U_{(2k)}^{-1/\alpha}] - [10^l U_{(4k)}^{-1/\alpha}]} \\ &= I_1 + I_2\,. \end{aligned}$$

By virtue of (2.7) we have

$$\begin{aligned} I_1 &= \log \frac{10^l U_{(k)}^{-1/\alpha}}{[10^l U_{(k)}^{-1/\alpha}]} + \log \frac{1 - 10^l U_{(2k)}^{-1/\alpha}/10^l U_{(k)}^{-1/\alpha}}{1 - [10^l U_{(2k)}^{-1/\alpha}]/[10^l U_{(k)}^{-1/\alpha}]} \\ &\le (1+10^{-l})10^{-l} U_{(k)}^{1/\alpha} + \log \frac{1 - 10^l U_{(2k)}^{-1/\alpha}/10^l U_{(k)}^{-1/\alpha}}{1 - [10^l U_{(2k)}^{-1/\alpha}]/[10^l U_{(k)}^{-1/\alpha}]} \\ &= I_{11} + I_{12}\,. \end{aligned}$$

We analyze I_{12} by using the inequality $x/(1+x) \le \log(1+x) \le x$, $|x| < 1$. Observing that $U_{(k)}^{1/\alpha}/U_{(2k)}^{1/\alpha} \stackrel{P}{\to} 2^{-1/\alpha}$, we have for large n with probability 1,

$$\begin{aligned} |I_{12}| &\le O_P\Big(\Big|\frac{[10^l U_{(2k)}^{-1/\alpha}]}{[10^l U_{(k)}^{-1/\alpha}]} - \frac{10^l U_{(2k)}^{-1/\alpha}}{10^l U_{(k)}^{-1/\alpha}}\Big|\Big) \\ &= O_P\Big(U_{(k)}^{2/\alpha}\Big(10^l U_{(2k)}^{-1/\alpha} 10^l U_{(k)}^{-1/\alpha} - [10^l U_{(2k)}^{-1/\alpha}]\,[10^l U_{(k)}^{-1/\alpha}]\Big)\Big) \\ &= O_P\Big(U_{(k)}^{2/\alpha}\Big(10^l U_{(2k)}^{-1/\alpha}(10^l U_{(k)}^{-1/\alpha} - [10^l U_{(k)}^{-1/\alpha}]) + [10^l U_{(k)}^{-1/\alpha}](10^l U_{(2k)}^{-1/\alpha} - [10^l U_{(2k)}^{-1/\alpha}])\Big)\Big) \\ &= O_P\Big(U_{(k)}^{2/\alpha}\Big(U_{(2k)}^{-1/\alpha} + U_{(k)}^{-1/\alpha}\Big)\Big) = O_P(U_{(k)}^{1/\alpha})\,. \end{aligned}$$

Using these bounds and $k_n = o(n^{2/(\alpha+2)})$, we have $\sqrt{k} I_i \stackrel{P}{\to} 0$, $i = 1,2$, hence $\sqrt{k} D_k^{(3)} \stackrel{P}{\to} 0$. Finally, an application of Lemma 3.3 yields the result. □

4. Concluding remarks

The estimation of the tail index α is a complicated statistical problem. The results and graphs above show that the estimation also depends on round-off effects which often are neglected, e.g. by assuming that the data have a Lebesgue density.

There exist numerous applied papers where power law behavior of the tails of the data has been postulated (e.g. in the literature on Zipf's law or on fractal dimensions of real-life data). The tail index α is often estimated by ordinary least squares (OLS) based on a plot of $-\log \overline{F}_n(x)$ (F_n is the empirical distribution function) against $\log x$, where x is chosen from the whole range of the data or from a "far-out" x-region where the plot is "roughly linear" . The round-off effect leads to undesirable oscillations of the log-log plot and, in turn, yields unreliable estimates of α.

For the Hill estimator $\alpha_{k,l}^{-1}$ of Pareto variables one can calculate the expectation and variance explicitly; numerical calculations and simulations show that these moments and the estimator itself may oscillate strongly, depending on the size of the round-off error. The results of this paper indicate that the region of k-values where the Hill and related estimators are asymptotically normal is rather small and strongly depends on the size of the round-off error described by the parameter l. On the positive side, even under round-off effects the classical estimators are reliable (i.e. satisfy the usual asymptotic properties) in these k-regions and, in contrast to the estimation of α based on OLS, a body of standard theory is applicable.

A referee of this paper pointed out that rounding of data could be considered as a special case of interval censoring which can be handled in the framework of maximum likelihood. For example, if we assume that the data come from a particular distribution (the generalized Pareto distribution would be a natural candidate in the extreme value context) an interval censored likelihood approach would be possible. We did not explore this method.

Acknowledgment. We would like to thank both referees for careful reading of the manuscript and for their constructive comments.

Appendix A. Expression for second moment

Direct calculation of the second moment of the Hill estimator $\widehat{\alpha}_{k,l}^{-1}$ is complicated, involving infinite series of incomplete beta functions. As in Lemma 2.1, we chose a proof based on a conditioning argument. The result is rather complex, but involves only two incomplete beta functions in each term of the infinite series below. This formula can be evaluated numerically.

Lemma A.1. *Let $X_i^{(l)} = 10^{-l}[10^l U_i^{-1/\alpha}]$, $i = 1, 2, \dots,$ for a fixed integer $l \geq 0$. Then the second moment for the Hill estimator $\widehat{\alpha}_{k,l}^{-1}$ from the sample $(X_i^{(l)})_i$, $n \geq 3, 2 \leq k \leq n-1$:*

$$
\begin{aligned}
&E(\widehat{\alpha}_{k,l}^{-1})^2 \\
= \; &\sum_{s=10^l}^{\infty} \Bigg[\frac{n}{(k-1)^2} \{(\log(s+1))^2 - (\log s)^2\} (10^l/(s+1))^\alpha (1 - J_1) \\
&\quad - \frac{2n}{k-1} \Big(\log \frac{s+1}{s} \sum_{t=s}^{\infty} \log \frac{t+1}{t} (10^l/(t+1))^\alpha - \log(s+1) \log \frac{s+1}{s} (10^l/(s+1))^\alpha \Big) J_1 \\
&\quad + \frac{2n(k-2)}{(k-1)^2} \Big(\log \frac{s+1}{s} \sum_{t=s+1}^{\infty} \log \frac{t+1}{t} (10^l/(t+1))^\alpha - \log s \log(s+1) (10^l/(s+1))^\alpha \\
&\qquad\qquad + (\log(s+1))^2 (10^l/(s+2))^\alpha \Big) J_1 \\
&\quad - \frac{2n(k-2)}{(k-1)^2} \Big\{ (\log(s+1))^2 (10^l/(s+2))^\alpha - (\log s)^2 (10^l/(s+1))^\alpha \Big\} J_1 \Bigg] \\
&+ \sum_{s=10^l}^{\infty} \Bigg[\frac{2n(n-1)}{(k-1)^2} \Big\{ \log \frac{s+1}{s} \sum_{t=s+1}^{\infty} \log \frac{t+1}{t} (10^l/(t+1))^\alpha - \log s \log(s+1) (10^l/(s+1))^\alpha \\
&\qquad\qquad + (\log(s+1))^2 (10^l/(s+2))^\alpha \Big\} (10^l/(s+1))^\alpha (1 - J_2) \\
&\quad + \frac{n(n-1)}{(k-1)^2} (\log(s+1))^2 \big\{ (10^l/(s+1))^\alpha - (10^l/(s+2))^\alpha \big\}^2 (1 - J_2) \\
&\quad + \frac{n(n-1)}{(k-1)^2} \Big\{ (\log(s+1))^2 (10^l/(s+2))^{2\alpha} - (\log s)^2 (10^l/(s+1))^{2\alpha} \Big\} J_2 \Bigg] \\
&- \frac{2n}{k-1} \log 10^l \sum_{s=10^l}^{\infty} \log \frac{s+1}{s} (10^l/(s+1))^\alpha \\
&+ \frac{2n(k-2)}{(k-1)^2} \log 10^l \Big(\sum_{s=10^l+1}^{\infty} \log \frac{s+1}{s} (10^l/(s+1))^\alpha + \log(10^l+1)(10^l/(10^l+1))^\alpha \Big)
\end{aligned}
$$

$$-\frac{2n(k-2)}{(k-1)^2}(\log 10^l)^2(10^l/(10^l+1))^\alpha + \frac{n(n-1)}{(k-1)^2}(\log 10^l)^2(10^l/(10^l+1))^{2\alpha},$$

where the quantities J_i are given by

$$J_i := \int_0^{(10/(s+1))^\alpha} \beta(k-i, n-k+1)(u)du, \qquad i = 1, 2.$$

Proof. The conditional second moment of $\alpha_{k,l}^{-1}$ given $\{U_{(k)} = u\}$ is calculated as

$$\begin{aligned}
E[(\widehat{\alpha}_{k,l}^{-1})^2 \mid U_{(k)} = u] &= E\Big[\Big(\frac{1}{k-1}\sum_{j=1}^{k-1}\log[10^l V_j^{-1/\alpha}] - \log[10^l u^{-1/\alpha}]\Big)^2\Big] \\
&= E\Big[\frac{1}{(k-1)^2}\sum_{i=1}^{k-1}\sum_{j=1}^{k-1}\log[10^l V_j^{-1/\alpha}]\log[10^l V_i^{-1/\alpha}] \\
&\qquad - \frac{2}{k-1}\sum_{j=1}^{k-1}\log[10^l V_j^{-1/\alpha}]\log[10 u^{-1/\alpha}] + (\log[10^l u^{-1/\alpha}])^2\Big] \\
&= \frac{1}{k-1}E\Big[(\log[10^l V_1^{-1/\alpha}])^2\Big] + \frac{k-2}{k-1}\big(E\big[\log[10^l V_1^{-1/\alpha}]\big]\big)^2 \\
&\qquad - 2E\big[\log[10^l V_1^{-1/\alpha}]\big]\log[10^l u^{-1/\alpha}] + (\log[10^l u^{-1/\alpha}])^2 \\
&=: \frac{1}{k-1}I_A + \frac{k-2}{k-1}I_B - 2I_C + I_D,
\end{aligned} \tag{A.1}$$

where V_i, $i = 1, \dots, k-1$, are iid $U(0, u)$ distributed random variables. We start by observing that

$$\begin{aligned}
E(I_A) &= E\Big[\sum_{s=10^l}^\infty (\log s)^2 P((10^l/(s+1))^\alpha \le V_1 \le (10^l/s)^\alpha)\Big] \\
&= E\Big[\sum_{s=10^l}^\infty \int_{(10^l/(s+1))^\alpha}^{(10^l/s)^\alpha} \frac{I_{\{0\le x\le U_{(k)}\}}}{U_{(k)}}dx\Big] \\
&= \sum_{s=10^l}^\infty (\log s)^2 \frac{n}{k-1}\int_0^1 du\beta(k-1, n-k+1)(u)\Big[\big\{(10^l/s)^\alpha - (10^l/(s+1))^\alpha\big\}I_{\{(10^l/s)^\alpha \le u\}} \\
&\qquad + \{u - (10/(s+1))^\alpha\}I_{\{(10^l/(s+1))^\alpha \le u \le (10^l/s)^\alpha)\}}\Big] \\
&= \sum_{s=10^l}^\infty (\log s)^2\Big[\frac{n}{k-1}\Big\{(10^l/s)^\alpha \int_{(10^l/s)^\alpha}^1 \beta(k-1, n-k+1)(u)du \\
&\qquad - (10^l/(s+1))^\alpha \int_{(10^l/(s+1))^\alpha}^1 \beta(k-1, n-k+1)(u)du\Big\} \\
&\qquad + \int_0^{(10^l/s)^\alpha} \beta(k, n-k+1)(u)du - \int_0^{(10^l/(s+1))^\alpha} \beta(k, n-k+1)(u)du\Big] \\
&= \sum_{s=10^l}^\infty \big\{(\log(s+1))^2 - (\log s)^2\big\}\Big\{\frac{n}{k-1}(10^l/(s+1))^\alpha \int_{(10^l/(s+1))^\alpha}^1 \beta(k-1, n-k+1)(u)du \\
&\qquad + \int_0^{(10^l/(s+1))^\alpha} \beta(k, n-k+1)(u)du\Big\} + (\log 10^l)^2.
\end{aligned}$$

Similar to the calculation of $E(I_A)$, we have

$$E(I_D) = \sum_{s=10^l}^{\infty} \{(\log(s+1))^2 - (\log s)^2\} \int_0^{(10^l/(s+1))^\alpha} \beta(k, n-k+1)(u)du + (\log 10^l)^2.$$

As for the expression I_C, we need more calculations. First, we see that

$$\begin{aligned}
E(I_C) &= E\Bigg[\sum_{s=10^l}^{\infty} \log s \int_{(10^l/(s+1))^\alpha}^{(10/s)^\alpha} dx \frac{I_{\{0\le x\le U_{(k)}\}}}{U_{(k)}} \log[10^l U_{(k)}^{-1/\alpha}]\Bigg] \\
&= \sum_{s=10^l}^{\infty} \log s \frac{n}{k-1} \int_0^1 du \log[10^l u^{-1/\alpha}] \beta(k-1, n-k+1)(u) \Big\{(10^l/s)^\alpha I_{\{u\ge(10^l/s)^\alpha\}} \\
&= \sum_{s=10^l}^{\infty} \log \frac{s+1}{s} (10^l/(s+1))^\alpha \frac{n}{k-1} \int_{(10^l/(s+1))^\alpha}^1 \log[10^l u^{-1/\alpha}] \beta(k-1, n-k+1)(u)du \\
&\quad + \sum_{s=10^l}^{\infty} \{(\log(s+1))^2 - (\log s)^2\} \int_0^{(10^l/(s+1))^\alpha} \beta(k, n-k+1)(u)du + (\log 10^l)^2.
\end{aligned}$$

The first integral is further calculated as

$$\begin{aligned}
&\sum_{s=10^l}^{\infty} \log \frac{s+1}{s} (10^l/(s+1))^\alpha \frac{n}{k-1} \sum_{t=10^l}^{s} \log t \int_{(10^l/(t+1))^\alpha}^{(10^l/t)^\alpha} \beta(k-1, n-k+1)(u)du \\
&= \sum_{t=10^l}^{\infty} \log t \sum_{s=t}^{\infty} \log \frac{s+1}{s} (10^l/(s+1))^\alpha \frac{n}{k-1} \int_{(10^l/(t+1))^\alpha}^{(10^l/t)^\alpha} \beta(k-1, n-k+1)(u)du \\
&= \sum_{t=10^l}^{\infty} \Big\{ \log(t+1) \sum_{s=t+1}^{\infty} \log \frac{s+1}{s} (10^l/(s+1))^\alpha - \log t \sum_{s=t}^{\infty} \log \frac{s+1}{s} (10^l/(s+1))^\alpha \Big\} \\
&\times \frac{n}{k-1} \int_0^{(10^l/(t+1))^\alpha} \beta(k-1, n-k+1)(u)du + \frac{n}{k-1} \log 10^l \sum_{s=10^l}^{\infty} \log \frac{s+1}{s} (10^l/(s+1))^\alpha \\
&= \sum_{t=10^l}^{\infty} \Big\{ \log \frac{t+1}{t} \sum_{s=t}^{\infty} \log \frac{s+1}{s} (10^l/(s+1))^\alpha - \log(t+1) \log \frac{t+1}{t} (10^l/(t+1))^\alpha \Big\} \\
&\times \frac{n}{k-1} \int_0^{(10^l/(t+1))^\alpha} \beta(k-1, n-k+1)(u)du + \frac{n}{k-1} \log 10^l \sum_{s=10^l}^{\infty} \log \frac{s+1}{s} (10^l/(s+1))^\alpha
\end{aligned}$$

where in the last step, we change the summation and related arguments. Hence, we obtain

$$\begin{aligned}
E(I_C) &= \sum_{t=10^l}^{\infty} \Big\{ \log \frac{t+1}{t} \sum_{s=t}^{\infty} \log \frac{s+1}{s} (10^l/(s+1))^\alpha - \log(t+1) \log \frac{t+1}{t} (10^l/(t+1))^\alpha \Big\} \\
&\quad \times \frac{n}{k-1} \int_0^{(10^l/(t+1))^\alpha} \beta(k-1, n-k+1)(u)du + \frac{n}{k-1} \log 10^l \sum_{s=10^l}^{\infty} \log \frac{s+1}{s} (10^l/(s+1))^\alpha \\
&\quad + \sum_{s=10^l}^{\infty} \Big\{(\log(s+1))^2 - (\log s)^2\Big\} \int_0^{(10^l/(s+1))^\alpha} \beta(k, n-k+1)(u)du + (\log 10^l)^2.
\end{aligned}$$

As for I_B, by symmetry, we write

$$\begin{aligned} E(I_B) &= \Big(2\sum_{s<t}+\sum_{s=t}\Big)\log s\log t\ E\Big[\int_{(10^l/t)^\alpha}^{(10^l/(t+1))^\alpha}\frac{I_{\{0\le x\le U_{(k)}\}}}{U_{(k)}}dx\int_{(10^l/s)^\alpha}^{(10^l/(s+1))^\alpha}\frac{I_{\{0\le y\le U_{(k)}\}}}{U_{(k)}}dy\Big] \\ &=: 2I_{B1}+I_{B2}. \end{aligned}$$

An analytical expression of I_{B1} is derived as

$$\begin{aligned} E(I_{B1}) &= \sum_{s=10^l+1}^{\infty}\log s\{(10^l/s)^\alpha-(10^l/(s+1))^\alpha\}\sum_{t=10^l}^{s-1}\log t E\int_{(10^l/(t+1))^\alpha}^{(10^l/t)^\alpha}\frac{I_{\{0\le x\le U_{(k)}\}}}{U_{(k)}^2}dx \\ &= \sum_{s=10^l+1}^{\infty}\log s\{(10^l/s)^\alpha-(10^l/(s+1))^\alpha\}\sum_{t=10^l}^{s-1}\log t\frac{n(n-1)}{(k-1)(k-2)} \\ &\quad\times\int_0^1 du\beta(k-2,n-k+1)(u)\Big\{(10^l/t)^\alpha I_{\{(10/t)^\alpha\le u\}} \\ &\qquad -(10^l/(t+1))^\alpha I_{\{(10/(t+1))^\alpha\le u\}}+uI_{\{(10^l/(t+1))^\alpha\le u\le(10^l/t)^\alpha)\}}\Big\} \\ &= \sum_{t=10^l}^{\infty}\log t\sum_{s=t+1}^{\infty}\log s\Big\{(10^l/s)^\alpha-(10^l/(s+1))^\alpha\Big\} \\ &\quad\times\Bigg[\frac{n(n-1)}{(k-1)(k-2)}\Big\{(10^l/t)^\alpha\int_{(10^l/t)^\alpha}^1\beta(k-2,n-k+1)(u)du \\ &\qquad -(10^l/(t+1))^\alpha\int_{(10^l/(t+1))^\alpha}^1\beta(k-2,n-k+1)(u)du\Big\} \\ &\qquad +\frac{n}{k-1}\Big\{\int_0^{(10^l/t)^\alpha}\beta(k-1,n-k+1)(u)du \\ &\qquad\qquad -\int_0^{(10^l/(t+1))^\alpha}\beta(k-1,n-k+1)(u)du\Big\}\Bigg] \\ &= \sum_{t=10^l}^{\infty}\Big[\log(t+1)\sum_{s=t+1}^{\infty}\log s\{(10^l/s)^\alpha-(10^l/(s+1))^\alpha\} \\ &\quad -\log t\sum_{s=t+1}^{\infty}\log s\{(10^l/s)^\alpha-(10^l/(s+1))^\alpha\}\Big] \\ &\quad\times\Bigg\{\frac{n(n-1)}{(k-1)(k-2)}(10^l/(t+1))^\alpha\int_{(10^l/(t+1))^\alpha}^1\beta(k-2,n-k+1)(u)du \\ &\qquad +\frac{n}{k-1}\int_0^{(10^l/(t+1))^\alpha}\beta(k-1,n-k+1)(u)du\Bigg\} \\ &\quad +\frac{n}{k-1}\log 10^l\Big\{\sum_{s=10^l+1}^{\infty}\log\frac{s+1}{s}(10^l/(s+1))^\alpha+\log(10^l+1)(10^l/(10^l+1))^\alpha\Big\} \\ &= \sum_{t=10^l}^{\infty}\Big\{\log\frac{t+1}{t}\sum_{s=t+1}^{\infty}\log\frac{s+1}{s}(10^l/(s+1))^\alpha+(\log(t+1))^2(10^l/(t+2))^\alpha \\ &\qquad -\log t\log(t+1)(10^l/(t+1))^\alpha\Big\} \end{aligned}$$

$$
\begin{aligned}
&\times\Big\{\frac{n(n-1)}{(k-1)(k-2)}(10^l/(t+1))^\alpha \int_{(10^l/(t+1))^\alpha}^1 \beta(k-2,n-k+1)(u)du \\
&\quad +\frac{n}{k-1}\int_0^{(10^l/(t+1))^\alpha} \beta(k-1,n-k+1)(u)du\Big\} \\
&+\frac{n}{k-1}\log 10^l\Big\{\sum_{s=10^l+1}^{\infty}\log\frac{s+1}{s}(10^l/(s+1))^\alpha+\log(10^l+1)(10^l/(10^l+1))^\alpha\Big\}.
\end{aligned}
$$

We consider the expression for $E(I_{B2})$,

$$
\begin{aligned}
E(I_{B2}) &= \sum_{s=10^l}^{\infty}(\log s)^2 E\int_{(10/(s+1))^\alpha}^{(10^l/s)^\alpha}dx\int_{(10/(s+1))^\alpha}^{(10^l/s)^\alpha}dy\frac{I_{\{0\le x\le U_{(k)}\}}I_{\{0\le y\le U_{(k)}\}}}{(U_{(k)})^2} \\
&= \sum_{s=10^l}^{\infty}(\log s)^2\frac{n(n-1)}{(k-1)(k-2)}\int_0^1 du\beta(k-2,n-k+1)(u) \\
&\quad\times\Big(\big\{(10^l/s)^\alpha-(10/(s+1))^\alpha\big\}I_{\{(10^l/s)^\alpha\le u\}} \\
&\qquad +\big\{u-(10/(s+1))^\alpha\big\}I_{\{(10^l/(s+1))^\alpha\le u\le(10^l/s)^\alpha)\}}\Big)^2 \\
&= \sum_{s=10^l}^{\infty}(\log s)^2\frac{n(n-1)}{(k-1)(k-2)}\int_0^1 du\beta(k-2,n-k+1)(u) \\
&\quad\times\Big(\big\{(10^l/s)^\alpha-(10/(s+1))^\alpha\big\}^2 I_{\{(10^l/s)^\alpha\le u\}} \\
&\qquad +\big\{u-(10/(s+1))^\alpha\big\}^2 I_{\{(10^l/(s+1))^\alpha\le u\le(10^l/s)^\alpha)\}}\Big) \\
&= \sum_{s=10^l}^{\infty}(\log s)^2\frac{n(n-1)}{(k-1)(k-2)}\big\{(10^l/s)^\alpha-(10/(s+1))^\alpha\big\}^2 \\
&\qquad\times\int_{(10^l/s)^\alpha}^1\beta(k-2,n-k+1)(u)du \\
&\quad +\sum_{s=10^l}^{\infty}(\log s)^2\frac{n(n-1)}{(k-1)(k-2)}\int_0^1\beta(k-2,n-k+1)(u)du \\
&\qquad\times\big\{u^2-2u(10^l/(s+1))^\alpha+(10^l/(s+1))^{2\alpha}\big\}\big(1_{\{u\le(10^l/s)^\alpha\}}-1_{\{u\le(10^l/(s+1))^\alpha\}}\big) \\
&= \sum_{s=10^l}^{\infty}(\log s)^2\frac{n(n-1)}{(k-1)(k-2)}\big\{(10^l/s)^\alpha-(10/(s+1))^\alpha\big\}^2 \\
&\qquad\times\Big(1-\int_0^{(10^l/s)^\alpha}\beta(k-2,n-k+1)(u)du\Big) \\
&\quad +\sum_{s=10^l}^{\infty}(\log s)^2\int_{(10^l/(s+1))^{2\alpha}}^{(10^l/s)^\alpha}\beta(k,n-k+1)(u)du \\
&\quad -2\sum_{s=10^l}^{\infty}(\log s)^2(10^l/(s+1))^\alpha\frac{n}{k-1}\int_{(10^l/(s+1))^{2\alpha}}^{(10^l/s)^\alpha}\beta(k-1,n-k+1)(u)du
\end{aligned}
$$

$$
\begin{aligned}
&\quad + \sum_{s=10^l}^{\infty} (\log s)^2 (10^l/(s+1))^{2\alpha} \frac{n(n-1)}{(k-1)(k-2)} \int_{(10^l/(s+1))^\alpha}^{(10^l/s)^\alpha} \beta(k-2, n-k+1)(u)du \\
&= \sum_{s=10^l}^{\infty} (\log(s+1))^2 \frac{n(n-1)}{(k-1)(k-2)} \big\{(10^l/(s+1))^\alpha - (10/(s+2))^\alpha\big\}^2 \\
&\qquad \times \Big(1 - \int_0^{(10^l/(s+1))^\alpha} \beta(k-2, n-k+1)(u)du\Big) \\
&\quad + \sum_{s=10^l}^{\infty} \big\{(\log(s+1))^2 - (\log s)^2\big\} \int_0^{(10^l/(s+1))^\alpha} \beta(k, n-k+1)(u)du + (\log 10^l)^2 \\
&\quad -2 \sum_{s=10^l}^{\infty} \Big\{(\log(s+1))^2 (10^l/(s+2))^\alpha - (\log s)^2 (10^l/(s+1))^\alpha\Big\} \\
&\qquad \times \frac{n}{k-1} \int_0^{(10^l/(s+1))^\alpha} \beta(k-1, n-k+1)(u)du \\
&\quad + \sum_{s=10^l}^{\infty} \Big\{(\log(s+1))^2 (10^l/(s+2))^{2\alpha} - (\log s)^2 (10^l/(s+1))^{2\alpha}\Big\} \\
&\qquad \times \frac{n(n-1)}{(k-1)(k-2)} \int_0^{(10^l/(s+1))^\alpha} \beta(k-2, n-k+1)(u)du \\
&\quad -2(\log 10^l)^2 (10^l/(10^l+1))^\alpha \frac{n}{k-1} + (\log 10^l)^2 (10^l/(10^l+1))^{2\alpha} \frac{n(n-1)}{(k-1)(k-2)}.
\end{aligned}
$$

Substituting these integrals into (A.1) and putting together the coefficients of the two kinds of incomplete beta functions, we obtain the desired result. □

References

[1] Bingham, N.H., Goldie, C.M. and Teugels, J.L. (1987) *Regular Variation.* Cambridge University Press, Cambridge (UK).

[2] Deheuvels, P., Häusler, E. and Mason, D.M. (1988) Almost sure convergence of the Hill estimator. *Math. Proc. Cambridge Philos. Soc.* **104**, 371–381.

[3] Drees, H. and Rootzén, H. (2010) Limit theorems for empirical processes of cluster functionals. *Ann. Statist.* **38**, 2145–2186.

[4] Embrechts, P., Klüppelberg, C. and Mikosch, T. (1997) *Modelling Extremal Events for Insurance and Finance.* Springer, Berlin.

[5] Falconer, K.J. (2003) *Fractal Geometry: Mathematical Foundations and Applications.* John Wiley & Sons, Chichester.

[6] Gabaix, X. and Ioannides, Y.M. (2005) The evolution of city size distributions. In: *Handbook of Regional and Urban Economics,* Chapter 53, Elsevier, Amsterdam.

[7] Gut, A. (1988) *Stopped Random Walks.* Springer, New York.

[8] Haan, L. de and Ferreira, A. (2006) *Extreme Value Theory. An Introduction.* Springer Series in Operations Research and Financial Engineering. Springer, New York.

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

[10] Kagan, Y.Y. and Vere-Jones, D. (1996) Problems in the modelling and statistical analysis of earthquakes. In: Heyde, C.C., Prohorov Yu.V., Pyke, R. and Rachev, S.T. *Athens Conference on Applied Probability and Time Series Analysis (Lecture Notes Statist. 114)*, Springer, New York, pp. 398–425.

[11] Mason, D.M. (1982) Laws of large numbers for sums of extreme values. *Ann. Probab.* **10**, 756–764.

[12] Mikosch, T. (2003) Modelling dependence and tails of financial time series. In: B. Finkenstädt and H. Rootzén (Eds.) *Extreme Values in Finance, Telecommunications and the Environment.* Chapman and Hall (2003), 185–286.

[13] Smirnov, N.V. (1949) Limit distributions for the terms of a variational series. In Russian: *Trudy Mat. Inst. Steklov.* **25** (1949). Translation: *Transl. Amer. Math. Soc.* **11**, (1952) 82–143.

[14] Resnick, S.I. (2007) *Heavy-Tail Phenomena: Probabilistic and Statistical Modeling.* Springer, New York.

[15] Takens, F. (1985) On the numerical determination of the dimension of an attractor. In: *Dynamical Systems and Bifurcations (Groningen, 1984). Lecture Notes in Math.* **1125**, Springer, Berlin, pp. 99–106,

[16] Weyl, H. (1916) Über die Gleichverteilung von Zahlen mod. Eins. *Math. Ann.* **77**, 313–352.

M. Matsui, Nanzan University, Department of Business Administration, 18 Yamazato-cho Showa-ku Nagoya, 466-8673, Japan
E-mail address: mmuneya@nanzan-u.ac.jp

T. Mikosch, University of Copenhagen, Department of Mathematics, Universitetsparken 5, DK-2100 Copenhagen, Denmark
E-mail address: mikosch@math.ku.dk

L. Tafakori, University of Shiraz, Department of Statistics, College of Sciences, Shiraz, 71454, Iran
E-mail address: tafakori@shirazu.ac.ir