Journal of Applied Mathematics and Physics, 2023, 11, 4000-4015
https://www.scirp.org/journal/jamp
ISSN Online: 2327-4379
ISSN Print: 2327-4352

Distributed Trimmed Hill Estimator

Tao Guo

School of Mathematics and Statistics, Southwest University, Chongqing, China
Email: xndxgt@163.com

How to cite this paper: Guo, T. (2023) Distributed Trimmed Hill Estimator. *Journal of Applied Mathematics and Physics*, **11**, 4000-4015.
https://doi.org/10.4236/jamp.2023.1112256

Received: October 27, 2023
Accepted: December 24, 2023
Published: December 27, 2023

Abstract

Proceeded from trimmed Hill estimators and distributed inference, a new distributed version of trimmed Hill estimator for heavy tail index is proposed. Considering the case where the number of observations involved in each machine can be either the same or different and either fixed or varying to the total sample size, its consistency and asymptotic normality are discussed. Simulation studies are particularized to show the new estimator performs almost in line with the trimmed Hill estimator.

Keywords

Extreme Value Index, Distributed Trimmed Hill Estimator

1. Introduction

Let $\{X_1, X_2, \cdots, X_n\}$ be independent and identically distributed (i.i.d.) random variables drawn from *F*, a distributed function which belongs to the max-domain of attraction of an extreme value distribution G_γ with extreme value index $\gamma \in \mathbb{R}$. It is well-known that if $F \in D(G_\gamma)$, mathematically, there exist constants $a_n > 0$ and $b_n \in \mathbb{R}$ such that

$$\lim_{n\to\infty} F^n(a_n x + b_n) = G_\gamma(x) := \exp\left\{-(1+\gamma x)^{-1/\gamma}\right\}, \quad (1)$$

for all $1+\gamma x > 0$. For a tail index $\gamma > 0$, the limit relation (1) is equivalent to

$$\lim_{t\to\infty} \frac{U(tx)}{U(t)} = x^\gamma, \quad x > 0 \quad (2)$$

where $U = \{1/(1-F)\}^{\leftarrow}$ is the left-continuous inverse function of $1/(1-F)$ and a regular varying function with γ, see [1].

The estimation of tail index for heavy-tailed distributions may be the one of the most studied problems in the extreme value theory. Since the numerous works of this aspect such as the Hill estimator, the Pickands estimator and the maximum likelihood estimator have already been explored referring to [1] [2] [3]

[4] and [5] for detailed discussions and reviews.

This extreme value analysis often rely on high order statistics. However, in many applications, one may face the challenges when it quickly run out of data since the observations can be corrupted and this contamination can lead to severe bias in the estimation of the tail index. Considering the problem, plenty of researchers did a lot of work. Based on the classic Hill eatimator of γ:

$$\hat{\gamma}_k(n) := \frac{1}{k}\sum_{i=1}^{k}\log\left(\frac{X_{(n-i+1,n)}}{X_{(n-k,n)}}\right), \quad 1 \le k \le n-1,$$

where $X_{1,n} \le \cdots \le X_{n,n}$ are the associated order statistics of $\{X_1, X_2, \cdots, X_n\}$ i.i.d. random variables with unknown distribution $F \in D(G_\gamma)$ with $\gamma > 0$, [6] trimmed a certain number of the largest order statistics in order to obtain a robust estimator of γ and (among other robust estimators) defined a trimmed version of the Hill estimator:

$$\hat{\gamma}_{k_0,k}^{trim}(n) := \sum_{i=k_0+1}^{k} c_{k_0,k}(i)\log\left(\frac{X_{(n-i+1,n)}}{X_{(n-k,n)}}\right), \quad 0 \le k_0 < k < n.$$

[7] then chose the weights $c_{k_0,k}(i)$ so that the estimator is asymptotically optimal where,

$$c_{k_0,k}(i) = \begin{cases} \dfrac{k_0+1}{k-k_0}, & i = k_0+1 \\ \dfrac{1}{k-k_0}, & i = k_0+2, \cdots, k \end{cases}$$

and they also found the method for the trimming parameter which yields the trimmed Hill estimator that can adapt to the unknown level of contamination in the extremes. While removing the lower order statistics from the classical Hill estimator, [8] derived an alternative estimator of the tail index and it was shown to have lower variance than the classic Hill estimator. A number of reseachers also considered trimming but of the models rather than the data, see [9] and [10]. Moreover, the random censoring for heavy-tailed distribution was discussed in [11] [12] [13] and [14]. Contrary to the above, here we assume to have non-truncated heavy-tailed model and only the top order statistics are contaminated in the associated data.

The rapid emergence of massive datasets in various fields becomes more and more challenging to traditional statistical methods. Account of that, distributed inference theory which refers to analyzing data stored in distributed machines has been proposed. It is developed to deal with large-scale statistical optimization problems and requires a divide-and-conquer algorithm which estimates a desired quantity or parameter on each machine and transmits the results to a central machine often by simple averaging. With the conditions of [15] [16] and [17] [18] reported on a first attempt in distributed inference for extreme value index and proposed a distributed Hill estimator and establish its asymptotic theories.

In this paper, considering the massive datasets contaminated of the top order statistics we apply the method of distributed inference and then derive a new estimator of the extreme value index for heavy-tailed distributions. The new estimator can be used for the situation when large datasets are distributedly stored and cannot be combined into one oracle sample and the top order statistics are corrupted.

We assume that the i.i.d. observations $\{X_1, X_2, \cdots, X_n\}$ are stored in k machines with m observations each, *i.e.*, $n = mk$ and the operation mechanism of each machine is independent. Let $M_j^{(1)} \geq \cdots \geq M_j^{(m)}$ denote the order statistics within the machine j. Suppose we have identified that the top d_0 order statistics have been corrupted in each machine, we use the top d_j exceedance radios $M_j^{(i)} \big/ M_j^{(d_j)}$ and cut the same level of top d_0 exceedance radios for $i = 1, \cdots, d_j$ and $0 < d_0 < d_j < m$ to build the estimator in each machine. Then we take the average of the estimators from all machines and the distributed trimmed Hill estimator is defined as:

$$\hat{\gamma}_{DH}^{trim} = \frac{1}{k} \sum_{i=1}^{k} \left(\frac{d_0 + 1}{d_j - d_0} \log \frac{M_j^{(d_0+1)}}{M_j^{(d_j+1)}} + \frac{1}{d_j - d_0} \sum_{i=d_0+2}^{d_j} \log \frac{M_j^{(i)}}{M_j^{(d_j+1)}} \right). \tag{3}$$

To derive the asymptotic normality of distributed trimmed Hill estimator $\hat{\gamma}_{DH}^{trim}$, we need impose the following condition on the sequences k and m,

$$m(n) \to \infty, \ k(n) \to \infty \ \text{and} \ m/\log k \to \infty \ \text{as} \ n \to \infty. \tag{4}$$

And we need the second order regular varying condition as follows: there exists an positive or negative function A with $\lim_{t\to\infty} A(t) = 0$ and a real number $\rho \leq 0$, satisfying $|A(t)| \in RV_\rho$, such that

$$\lim_{t\to\infty} \frac{\frac{U(tx)}{U(t)} - x^\gamma}{A(t)} = x^\gamma \frac{x^\rho - 1}{\rho} \tag{5}$$

for all $x > 0$ (see e.g. [1], Corollary 2.3.4).

By (5), we have that there exists a function $A_0(t)$ such that $A_0(t) \sim A(t)$ as $t \to \infty$, and for all $\varepsilon > 0$, $\delta > 0$, there is a $t_0(\varepsilon, \delta) > 0$ such that for $tx \geq t_0$, $t \geq t_0$,

$$\left| \frac{\log U(tx) - \log U(t) - \gamma \log x}{A_0(t)} - \frac{x^\rho - 1}{\rho} \right| \leq \varepsilon x^\rho x^{\pm\delta}, \ \text{where} \ x^{\pm\delta} = \max\left\{x^\delta, x^{-\delta}\right\}. \tag{6}$$

By the adoption of (5) on the function L, we also have,

$$\left| \frac{\log L(tx) - \log L(t)}{A_0(t)} - \frac{x^\rho - 1}{\rho} \right| \leq \varepsilon x^\rho x^{\pm\delta}, \ \text{where} \ x^{\pm\delta} = \max\left\{x^\delta, x^{-\delta}\right\}, \tag{7}$$

for more details on A_0, see page 48 in [1].

2. Main Results

In the homogeneous case where $d_1 = \cdots = d_k = d$ is a fixed integer, the following theorem shows the asymptotic normality of the distributed trimmed Hill estimator.

Theorem 2.1. *Suppose* $F \in D(G_\gamma)$ *with* $\gamma > 0$ *and* (4) *and* (5) *hold. Let* $d_1 = \cdots = d_k = d$, *where* $d > d_0 > 0$ *is a fixed integer. If* $\left[k(d-d_0)\right]^{\frac{1}{2}} A(m/d) = O(1)$, *as* $n \to \infty$, *then*

$$\left[k(d-d_0)\right]^{\frac{1}{2}} \left\{\hat{\gamma}_{DH}^{trim} - A(m/d) B(m,d,d_0,\rho)\right\} \xrightarrow{d} N(0,\gamma^2),$$

where

$$B(m,d,d_0,\rho) = \frac{1}{\rho(d-d_0)} \left(\frac{m}{d}\right)^{-\rho} \frac{\Gamma(m+1)\Gamma(d-\rho+1)}{\Gamma(d+1)\Gamma(m-\rho+1)} \cdot \left\{ d_0 \left(\frac{d}{d-\rho}\right)^{d-d_0} - \sum_{i=1}^{d_0} \left(\frac{d}{d-\rho}\right)^{d-i+1} + \frac{d\rho}{1-\rho} \right\}$$

with $\rho < 0$ and $B(m,d,d_0,\rho) = 1$ with $\rho = 0$.

In the heterogeneous case where $\{d_j\}_{j=1}^k$ are uniformly bounded positive integer series, *i.e.*, $\sup_{j\in\mathbb{N}} d_j = d_{\max} < \infty$, the following theorem shows the asymptotic normality of the distributed trimmed Hill estimator.

Theorem 2.2. *Suppose* $F \in D(G_\gamma)$ *with* $\gamma > 0$ *and* (4) *and* (5) *hold. Let* $\{d_j\}_{j=1}^k$ *be uniformly bounded positive integer series, i.e.,* $\sup_{j\in\mathbb{N}} d_j = d_{\max} < \infty$ *and* $0 < d_0 < \inf_{j\in\mathbb{N}} d_j = d_{\min}$. *If* $\left[k(\bar{d}-d_0)\right]^{\frac{1}{2}} A(m/\bar{d}) = O(1)$, *as* $n \to \infty$, *then*

$$\left[k(\bar{d}-d_0)\right]^{\frac{1}{2}} \left\{\hat{\gamma}_{DH}^{trim} - \gamma - A(m/\bar{d}) k^{-1} \sum_{j=1}^{k} (\bar{d}/d_j)^\rho B(m,d_j,d_0,\rho)\right\} \xrightarrow{d} N(0,\gamma^2),$$

where $\bar{d} = k^{-1}\sum_{j=1}^k d_j$.

In the homogeneous case where $d_1 = \cdots = d_k = d$ and $d = d(m)$ is an intermediate sequence, *i.e.*, $d = d(m) \to \infty$, $d/m \to 0$ as $n \to \infty$, the following theorem shows the asymptotic normality of the distributed trimmed Hill estimator.

Theorem 2.3. *Suppose* $F \in D(G_\gamma)$ *with* $\gamma > 0$, *and* (4) *and* (5) *hold. Let* $d_1 = \cdots = d_k = d$, *where* $d = d(m) \to \infty$ *and* $d/m \to 0$ *as* $n \to \infty$. *If* $\left[k(d-d_0)\right]^{1/2} A(m/d) = O(1)$, *as* $n \to \infty$, *then*

$$\left[k(d-d_0)\right]^{\frac{1}{2}} \left\{\hat{\gamma}_{DH}^{trim} - \gamma - A(m/d) H(d,m,\rho)\right\} \xrightarrow{d} N(0,\gamma^2),$$

where

$$H(d,m,\rho) = \frac{1}{1-\rho}\left(\frac{m}{d}\right)^{-\rho} \frac{\Gamma(m+1)\Gamma(d-\rho+1)}{\Gamma(d+1)\Gamma(m-\rho+1)} \frac{d}{d-d_0}.$$

3. Simulation Studies

In this section, we study the finite sample performance of the distributed trimmed Hill estimator $\hat{\gamma}_{DH}^{trim}$ and compare it with [7]'s estimator, *i.e.*, the trimmed Hill estimator on the following three distributions which all belong to the max-domain of attraction of an extreme value distribution for varying parameters with three sub-cases for each distribution.

We obtain the mean value and mean squared error (MSE) for $r = 2000$ Monte Carlo simulations of all considered estimators of heavy-tailed models with sample size $n = 10{,}000$. We assume the contamination occurs in the top d_0 order statictics in each machine and vary the level of d in the distributed trimmed Hill estimator to verify the theoretical results on the property we give in Section2 and to compare the finite sample performance of the distributed trimmed Hill estimator with that of the trimmed Hill estimator for different values of d. The sample $\{X_1, X_2, \cdots, X_n\}$ contains $n = 10{,}000$ observations stored in k machines with m observations each. We fix $k = 20$ and $m = 500$ and vary d from 30 to 100 with $d_0 = 8$.The results are presented in **Figures 1-3**.

- The Fréchet distribution with distribution function

$$F(x) = \exp\left(-x^{-\alpha}\right), \quad \alpha > 0,$$

which implies $\gamma = 1/\alpha$ and $\rho = -1$. We consider the three parameters $\alpha = 1$, 0.5 and 2.

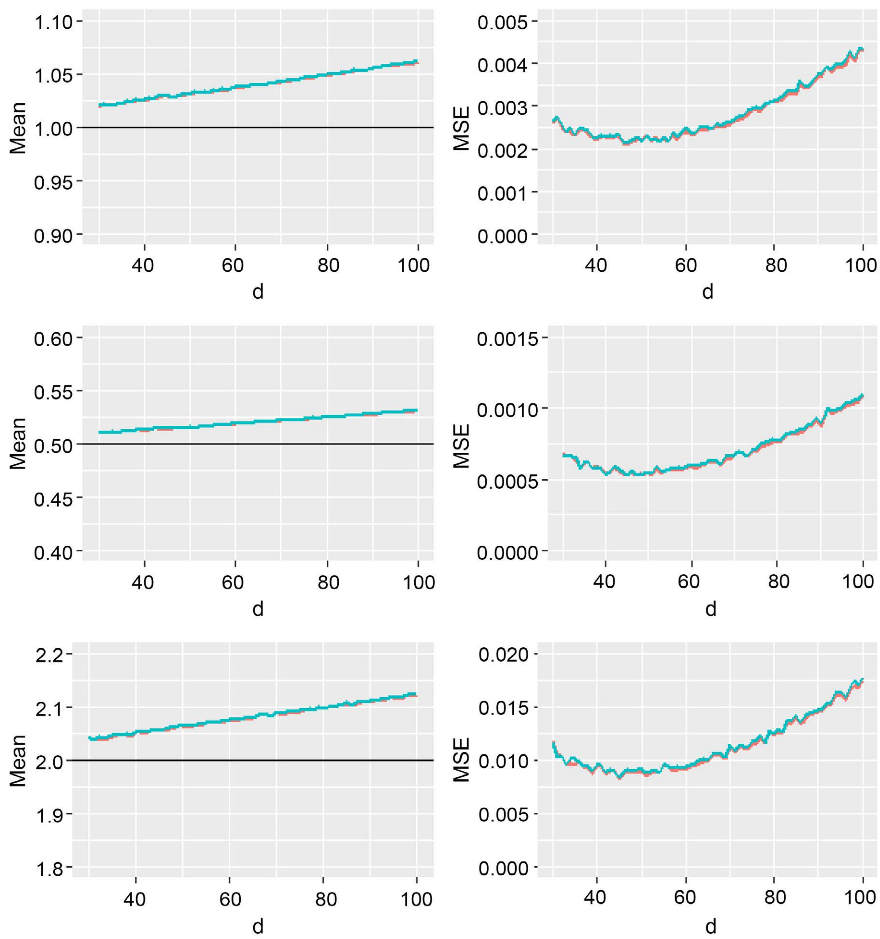

Figure 1. Fréchet distribution, parameters $\alpha = 1$, 0.5 and 2. Diagnostics of trimmed Hill estimator (coral) and distributed trimmed Hill estimator(skyblue) as a function of d.

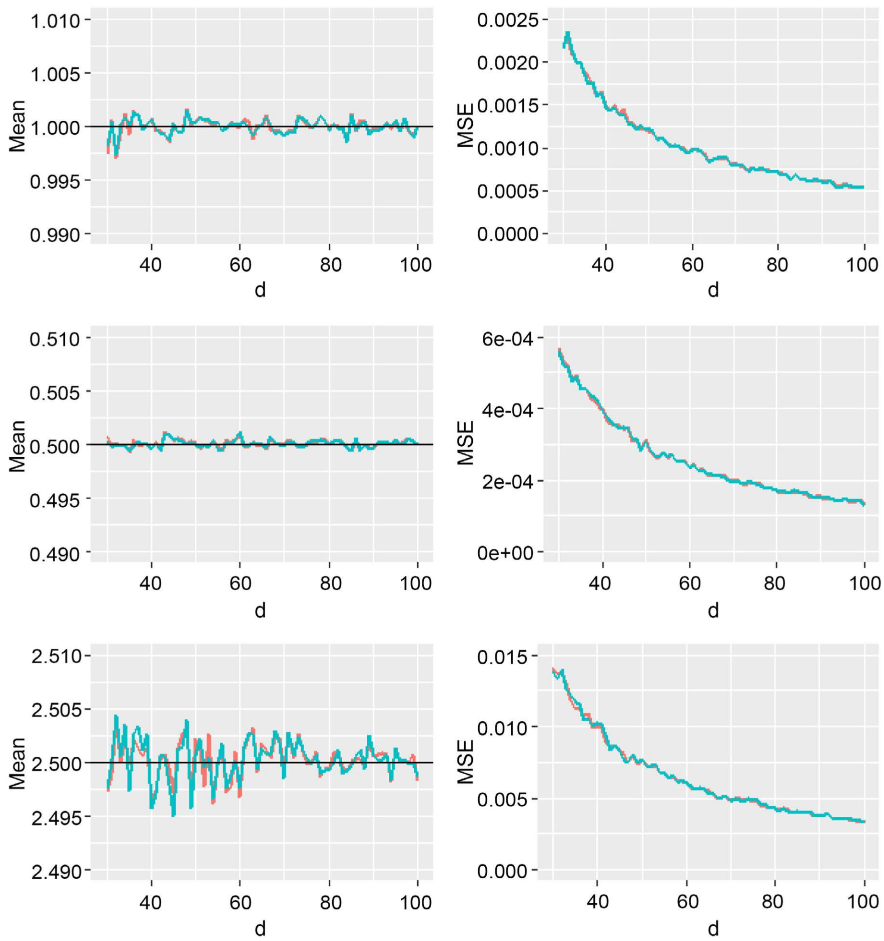

Figure 2. Pareto (σ, ξ) distribution, sets of parameters $\sigma = 1$, $\xi = 1$; $\sigma = 2$, $\xi = 0.5$ and $\sigma = 1$, $\xi = 2.5$. Diagnostics of trimmed Hill estimator(coral) and distributed trimmed Hill estimator (skyblue) as a function of d.

- The Pareto(σ, ξ) distribution with distribution function

$$F(x) = 1 - \left(\frac{x}{\sigma}\right)^{-1/\xi}, \quad x \geq \sigma, \quad \xi > 0,$$

which implies $\gamma = \xi$ and $\rho = -\xi$. We consider the three sets of parameters $\sigma = 1$, $\xi = 1$; $\sigma = 2$, $\xi = 0.5$ and $\sigma = 1$, $\xi = 2.5$.

- The Burr(τ, λ) distribution with distribution function

$$F(x) = 1 - \left(1 + x^{\tau}\right)^{-\lambda}, \quad x > 0, \quad \tau, \lambda > 0,$$

which implies $\gamma = \frac{1}{\tau\lambda}$ and $\rho = -\frac{1}{\lambda}$.We consider the three sets of parameters $\tau = 2$, $\lambda = 0.5$; $\tau = 3$, $\lambda = 0.5$ and $\tau = 3$, $\lambda = 1$.

For the Fréchet distribution, Figure 1 shows that as d increases, the MSE increases for the estimators with different α. For the Pareto distribution in Figure 2, the bias between the estimators and the true value is virtually zero for all levels of d. For the Burr distribution in Figure 3, we observe a trade off for the estimators with different sets of parameters: as d increases, the MSE increases when λ is low while the MSE decreases when λ takes a larger value.

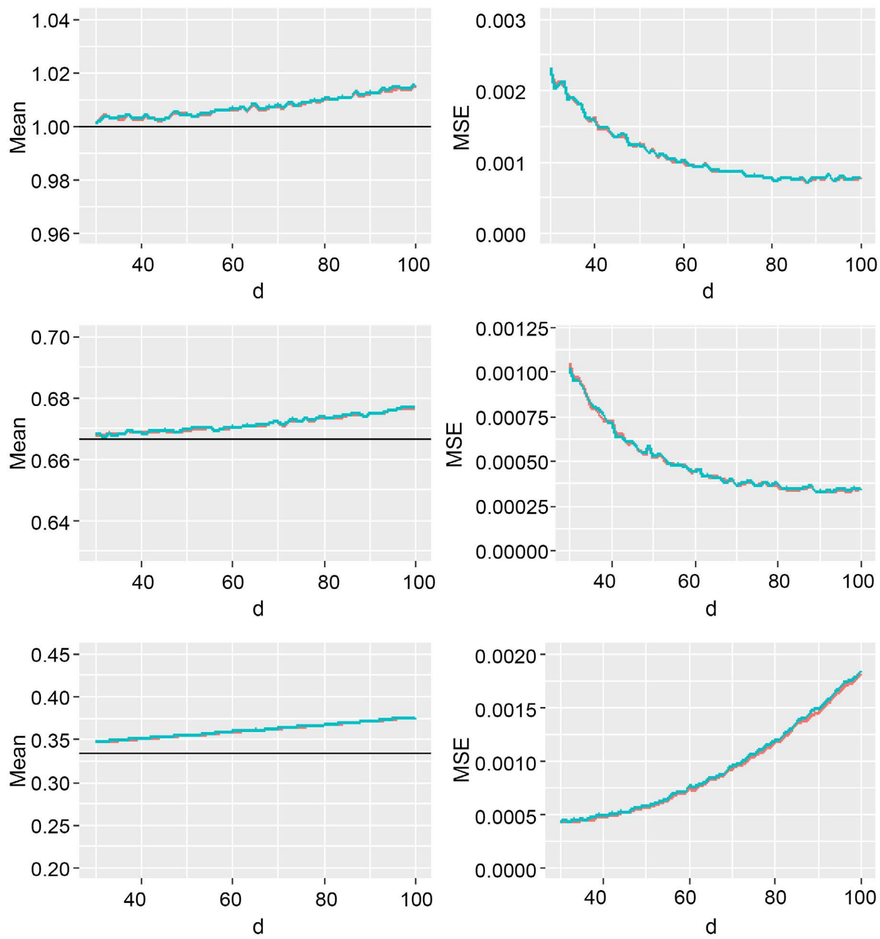

Figure 3. Burr (τ, λ) distribution, sets of parameters $\tau = 2$, $\lambda = 0.5$; $\tau = 3$, $\lambda = 0.5$ and $\tau = 3$, $\lambda = 1$. Diagnostics of trimmed Hill estimator (coral) and distributed trimmed Hill estimator (skyblue) as a function of *d*.

Figures 1-3 show that the difference in MSE between the distributed trimmed Hill estimator and the trimmed Hill estimator is not sizeable. Consequently, we can infer that when dealing with the estimation problem of extreme value index with massive and corrupted datasets the new estimator we derive performs well.

4. Proof

Recall that $1/(1-F)$, $\{X_1, X_2, \cdots, X_m\} \overset{d}{=} \{U(Z_1), U(Z_2), \cdots, U(Z_m)\}$, where $\{Z_1, Z_2, \cdots, Z_n\}$ is a random sample of Z with the distribution function $1-1/z$, $z \ge 1$. For each machine *j*, let $Z_j^{(1)} \ge \cdots \ge Z_j^{(m)}$ denote the order statistics of the m Pareto (1) distributed variables corresponding to the m observations in this machine. Notting that $\{M_j^{(1)}, M_j^{(2)}, \cdots, M_j^{(m)}\} \overset{d}{=} \{U(Z_j^{(1)}), U(Z_j^{(2)}), \cdots, U(Z_j^{(m)})\}$, we have

$$\hat{\gamma}_{DHj}^{trim} \overset{d}{=} \frac{d_0+1}{d_j-d_0} \log \frac{U\left(Z_j^{(d_0+1)}\right)}{U\left(Z_j^{(d_j+1)}\right)} + \frac{1}{d_j-d_0} \sum_{i=d_0+2}^{d_j} \log \frac{U\left(Z_j^{(i)}\right)}{U\left(Z_j^{(d_j+1)}\right)},$$

and then

$$\hat{\gamma}_{DH}^{trim} \overset{d}{=} \frac{1}{k}\sum_{j=1}^{k}\left(\frac{d_0+1}{d_j-d_0}\log\frac{U\left(Z_j^{(d_0+1)}\right)}{U\left(Z_j^{(d_j+1)}\right)}+\frac{1}{d_j-d_0}\sum_{i=d_0+2}^{d_j}\log\frac{U\left(Z_j^{(i)}\right)}{U\left(Z_j^{(d_j+1)}\right)}\right).$$

(2) implies that $U \in RV_\gamma$, then $U(x)=x^\gamma \cdot L(x)$, where $L(x)$ is slowly varying function. Hence $U\left(Z_j^{(i)}\right)=\left(Z_j^{(i)}\right)^\gamma \cdot L\left(Z_j^{(i)}\right)$ and

$$\hat{\gamma}_{DH}^{trim} \overset{d}{=} \frac{\gamma}{k}\sum_{j=1}^{k}\left(\frac{d_0+1}{d_j-d_0}\log\frac{Z_j^{(d_0+1)}}{Z_j^{(d_j+1)}}+\frac{1}{d_j-d_0}\sum_{i=d_0+2}^{d_j}\log\frac{Z_j^{(i)}}{Z_j^{(d_j+1)}}\right)$$
$$+\frac{1}{k}\sum_{j=1}^{k}\left(\frac{d_0+1}{d_j-d_0}\log\frac{L\left(Z_j^{(d_0+1)}\right)}{L\left(Z_j^{(d_j+1)}\right)}+\frac{1}{d_j-d_0}\sum_{i=d_0+2}^{d_j}\log\frac{L\left(Z_j^{(i)}\right)}{L\left(Z_j^{(d_j+1)}\right)}\right).$$

Lemma 4.1. *Suppose* $F \in D(G_\gamma)$ *with* $\gamma > 0$, *define*

$$\gamma_j^*(d_0) := \frac{d_0+1}{d_j-d_0}\log\frac{Z_j^{(d_0+1)}}{Z_j^{(d_j+1)}}+\frac{1}{d_j-d_0}\sum_{i=d_0+2}^{d_j}\log\frac{Z_j^{(i)}}{Z_j^{(d_j+1)}},$$

for $j=1,\cdots,k$, and $\gamma_{DH}^{trim*} := \frac{\gamma}{k}\sum_{j=1}^{k}\gamma_j^*(d_0)$. Under the assumption of (4), $\gamma_{DH}^{trim*} \overset{P}{\to} \gamma$.

Proof. Note that $\{\log Z_i, i=1,2,\cdots,m\}$ forms a random sample from the standard exponential distribution. In the machine *j*, for any $i \in \{d_0+1,\cdots,d_j\}$ by Rényi's representation we have

$\log Z_j^{(i)} - \log Z_j^{(d_j+1)} \overset{d}{=} \sum_{q=1}^{d_j-i+1} E_{j,q}/(d_j-q+1) \overset{d}{=} Y_j^{(i)}$ with $E_{j,1},\cdots,E_{j,d_j}$ i.i.d. standard exponential, where $Y_j^{(1)} \ge Y_j^{(2)} \ge \cdots \ge Y_i^{(d)}$ are the order statistics of Exp (1) corresponding to the d_j obeservation.

The joint distribution of $\gamma_j^*(d_0)$, $0 \le d_0 \le d_j-1$, can be expressed as follows:

$$\begin{aligned}
&\left\{\gamma_j^*(d_0)\right\}_{d_0=0}^{d_j-1}\\
&\overset{d}{=}\left\{\frac{d_0+1}{d_j-d_0}Y_j^{(d_0+1)}+\frac{1}{d_j-d_0}\sum_{i=d_0+2}^{d_j}Y_j^{(i)}\right\}_{d_0=0}^{d_j-1}\\
&\overset{d}{=}\left\{\frac{d_0+1}{d_j-d_0}\sum_{q=1}^{d_j-d_0}\frac{E_{j,q}}{d_j-q+1}+\frac{1}{d_j-d_0}\sum_{i=d_0+2}^{d_j}\sum_{q=1}^{d_j-i+1}\frac{E_{j,q}}{d_j-q+1}\right\}_{d_0=0}^{d_j-1}\\
&=\left\{\frac{d_0+1}{d_j-d_0}\sum_{q=1}^{d_j-d_0}\frac{E_{j,q}}{d_j-q+1}+\frac{1}{d_j-d_0}\sum_{q=1}^{d_j-d_0-1}\sum_{i=d_0+2}^{d_j-q+1}\frac{E_{j,q}}{d_j-q+1}\right\}_{d_0=0}^{d_j-1}\\
&=\left\{\sum_{q=1}^{d_j-d_0-1}\left(\frac{d_0+1}{d_j-d_0}+\sum_{i=d_0+2}^{d_j-q+1}\frac{1}{d_j-d_0}\right)\frac{E_{j,q}}{d_j-q+1}+\frac{E_{j,d_j-d_0}}{d_j-d_0}\right\}_{d_0=0}^{d_j-1}\\
&=\left\{\sum_{q=1}^{d_j-d_0}\frac{E_{j,q}}{d_j-d_0}\right\}_{d_0=0}^{d_j-1}
\end{aligned}\tag{8}$$

By (8) it implies that $E\gamma_j^*(d_0)=1$ for $d_0=0,\cdots,d_j-1$, and by WLLN and

(4) it follows that $\gamma_{DH}^{trim*} \xrightarrow{P} \gamma$ as $n \to \infty$.

Lemma 4.2. *Under the condition of Theorem* 3 *and define*

$$R_{d_0,d} := \frac{1}{k}\sum_{j=1}^{k}\left(\frac{d_0+1}{d-d_0}\log\frac{L\left(Z_j^{(d_0+1)}\right)}{L\left(Z_j^{(d_j+1)}\right)} + \frac{1}{d-d_0}\sum_{i=d_0+2}^{d}\log\frac{L\left(Z_j^{(i)}\right)}{L\left(Z_j^{(d_j+1)}\right)}\right),$$

if $h(d) = o(d)$, as $n \to \infty$, we have

$$\left[k(d-d_0)\right]^{\frac{1}{2}} \max_{0\le d_0<h(d)}\left|R_{d_0,d} - \frac{A_0(m/d)}{1-\rho}\frac{d}{d-d_0}\right| \xrightarrow{P} 0.$$

Proof. Note that

$$\begin{aligned}&\left[k(d-d_0)\right]^{\frac{1}{2}} \max_{0\le d_0<h(d)}\left|R_{d_0,d} - \frac{A_0(m/d)}{1-\rho}\frac{d}{d-d_0}\right| \\ &\le \left[k(d-d_0)\right]^{\frac{1}{2}} \max_{0\le d_0<h(d)}\left|R_{d_0,d} - S_{d_0,d}\right| + \left[k(d-d_0)\right]^{\frac{1}{2}} \max_{0\le d_0<h(d)}\left|S_{d_0,d} - \frac{A_0(m/d)}{1-\rho}\frac{d}{d-d_0}\right|,\end{aligned} \tag{9}$$

where

$$S_{d_0,d} = \frac{1}{k}\sum_{j=1}^{k}\left\{\frac{A_0\left(Z_j^{(d+1)}\right)}{d-d_0}\left[(d_0+1)\frac{\left(Z_j^{(d_0+1)}/Z_j^{(d+1)}\right)^{\rho}-1}{\rho} + \sum_{i=d_0+2}^{d}\frac{\left(Z_j^{(i)}/Z_j^{(d+1)}\right)^{\rho}-1}{\rho}\right]\right\}.$$

In the first term of (9), we have that

$$\begin{aligned}&\left[k(d-d_0)\right]^{\frac{1}{2}} \max_{0\le d_0<h(d)}\left|R_{d_0,d} - S_{d_0,d}\right| \\ &= \left[k(d-d_0)\right]^{\frac{1}{2}} \max_{0\le d_0<h(d)}\frac{dA_0(m/d)}{d-d_0}\left(\frac{d-d_0}{dA_0(m/d)}\left|R_{d_0,d} - S_{d_0,d}\right|\right) \\ &\le \frac{\left[k(d-d_0)\right]^{\frac{1}{2}} A_0(m/d)}{1-\frac{h(d)}{d}} \max_{0\le d_0<h(d)}\left(\frac{d-d_0}{dA_0(m/d)}\left|R_{d_0,d} - S_{d_0,d}\right|\right).\end{aligned} \tag{10}$$

By the assumption of $\left[k(d-d_0)\right]^{1/2} A(m/d) = O(1)$ as $n \to \infty$ and $A_0(t) \sim A(t)$ as $t \to \infty$, we can get that as $n \to \infty$, $\left[k(d-d_0)\right]^{1/2} A_0(m/d) = O(1)$. By (7) choose $\rho + \delta < 0$ and we can get that

$$\begin{aligned}&\max_{0\le d_0<h(d)}\left(\frac{d-d_0}{dA_0(m/d)}\left|R_{d_0,d} - S_{d_0,d}\right|\right) \\ &\le \max_{0\le d_0<h(d)}\frac{d_0\varepsilon}{dk}\sum_{j=1}^{k}\frac{A_0\left(Z_j^{(d+1)}\right)}{A_0(m/d)}\left(\frac{Z_j^{(d_0+1)}}{Z_j^{(d+1)}}\right)^{\rho+\delta} \\ &\quad + \max_{0\le d_0<h(d)}\frac{\varepsilon}{dk}\sum_{j=1}^{k}\frac{A_0\left(Z_j^{(d+1)}\right)}{A_0(m/d)}\sum_{i=d_0+1}^{d}\left(\frac{Z_j^{(i)}}{Z_j^{(d+1)}}\right)^{\rho+\delta}.\end{aligned} \tag{11}$$

Since $Z_j^{(d_0+1)}/Z_j^{(d+1)}$ is the $(d_0+1)^{th}$ order statistic from the standard Pareto distribution, [19] implies that $d_0 Z_j^{(d_0+1)}/(d+1)Z_j^{(d+1)} \xrightarrow{P} 1$. Recall that $A \in RV_\rho$

and we have $Z_j^{(d+1)}/(m/d)\xrightarrow{P}1$ as $n\to\infty$. Combining with $A_0(t)\sim A(t)$ as $t\to\infty$, we can get that as $n\to\infty$, $A_0\left(Z_j^{(d+1)}\right)/A_0(m/d)\xrightarrow{P}1$ and then

$$\max_{0\le d_0<h(d)}\frac{d_0\varepsilon}{dk}\sum_{j=1}^{k}\frac{A_0\left(Z_j^{(d+1)}\right)}{A_0(m/d)}\left(\frac{d+1}{d_0}\right)^{\rho+\delta}\left(\frac{d_0Z_j^{(d_0+1)}}{(d+1)Z_j^{(d+1)}}\right)^{\rho+\delta}\xrightarrow{P}0. \tag{12}$$

Similarly, as $n\to\infty$,

$$\max_{0\le d_0<h(d)}\frac{\varepsilon}{dk}\sum_{j=1}^{k}\frac{A_0\left(Z_j^{(d+1)}\right)}{A_0(m/d)}\sum_{i=d_0+1}^{d}\left(\frac{d+1}{i-1}\right)^{\rho+\delta}\left(\frac{(i-1)Z_j^{(i)}}{(d+1)Z_j^{(d+1)}}\right)^{\rho+\delta}\xrightarrow{P}0. \tag{13}$$

Combining with the (10) (11) (12) and (13), we can get that as $n\to\infty$,

$$\left[k(d-d_0)\right]^{\frac{1}{2}}\max_{0\le d_0<h(d)}\left|R_{d_0,d}-S_{d_0,d}\right|\xrightarrow{P}0. \tag{14}$$

In the second term of (9), we have that

$$\begin{aligned}&\left[k(d-d_0)\right]^{\frac{1}{2}}\max_{0\le d_0<h(d)}\left|S_{d_0,d}-\frac{A_0(m/d)}{1-\rho}\frac{d}{d-d_0}\right|\\&\le\frac{\left[k(d-d_0)\right]^{\frac{1}{2}}A_0(m/d)}{1-\frac{h(d)}{d}}\max_{0\le d_0<h(d)}\left|\frac{d-d_0}{d}\frac{S_{d_0,d}}{A_0(m/d)}-\frac{1}{1-\rho}\right|.\end{aligned} \tag{15}$$

It is known that

$$\begin{aligned}&\max_{0\le d_0<h(d)}\left|\frac{d-d_0}{d}\frac{S_{d_0,d}}{A_0(m/d)}-\frac{1}{1-\rho}\right|\\&\le\max_{0\le d_0<h(d)}\left|\frac{d_0}{dk}\sum_{j=1}^{k}\frac{A_0\left(Z_j^{(d+1)}\right)}{A_0(m/d)}\frac{\left(Z_j^{(d_0+1)}/Z_j^{(d+1)}\right)^{\rho}-1}{\rho}\right|\\&\quad+\max_{0\le d_0<h(d)}\left|\frac{1}{dk}\sum_{j=1}^{k}\frac{A_0\left(Z_j^{(d+1)}\right)}{A_0(m/d)}\sum_{i=d_0+1}^{d}\frac{\left(Z_j^{(i)}/Z_j^{(d+1)}\right)^{\rho}-1}{\rho}-\frac{1}{1-\rho}\right|,\end{aligned} \tag{16}$$

by WLLN for triangular array and $Z_j^{(d+1)}$ is independent with $Z_j^{(i)}/Z_j^{(d+1)}$ for $i=d_0+1,\cdots,d$ and $j=1,\cdots,k$, we have that $n\to\infty$,

$$\max_{0\le d_0<h(d)}\left|\frac{d_0}{dk}\sum_{j=1}^{k}\frac{A_0\left(Z_j^{(d+1)}\right)}{A_0(m/d)}\frac{\left(Z_j^{(d_0+1)}/Z_j^{(d+1)}\right)^{\rho}-1}{\rho}\right|\xrightarrow{P}0 \tag{17}$$

and

$$\max_{0\le d_0<h(d)}\left|\frac{1}{dk}\sum_{j=1}^{k}\frac{A_0\left(Z_j^{(d+1)}\right)}{A_0(m/d)}\sum_{i=d_0+1}^{d}\frac{\left(Z_j^{(i)}/Z_j^{(d+1)}\right)^{\rho}-1}{\rho}-\frac{1}{1-\rho}\right|\xrightarrow{P}0. \tag{18}$$

Combining with (17) and (18), it illustrates that $n\to\infty$,

$$\max_{0\le d_0<h(d)}\left|\frac{d-d_0}{d}\frac{S_{d_0,d}}{A_0(m/d)}-\frac{1}{1-\rho}\right|\xrightarrow{P}0$$

and finally we can get that $n \to \infty$,

$$\left[k(d-d_0)\right]^{\frac{1}{2}} \max_{0\le d_0<h(d)} \left| S_{d_0,d} - \frac{A_0(m/d)}{1-\rho}\frac{d}{d-d_0} \right| \xrightarrow{P} 0$$

which yield the Lemma.

Proof of Theorem 2.1. When $\rho<0$ and by Lemma S.2 in [18], we have that $\lim_{n\to\infty}\left\{Z^{(d+1)}>t_0\right\}=1$, for any $t_0>1$.Then by applying (6) twice with $t=m/d$ and $x=dZ_j^{(j)}/m, i=d_0+1,\cdots,d+1$ and $x=dZ_j^{(d+1)}/m$ we get that as $n\to\infty$,

$$\begin{aligned}&\frac{\log U\left(Z_j^{(i)}\right)-\log U\left(Z_j^{(d+1)}\right)-\gamma\left(\log Z_j^{(i)}-\log Z_j^{(d+1)}\right)}{A_0(m/d)}\\&=\frac{\left(dZ_j^{(j)}/m\right)^{\rho}-1}{\rho}-\frac{\left(dZ_j^{(d+1)}/m\right)^{\rho}-1}{\rho}+o_p(1)\left\{\left(dZ_j^{(j)}/m\right)^{\rho\pm\delta}+\left(dZ_j^{(d+1)}/m\right)^{\rho\pm\delta}\right\}.\end{aligned} \tag{19}$$

Here, the $o_p(1)$ term is uniform for all $1\le j\le k$, $d_0+1\le i\le d+1$ and all $k\in\mathbb{N}$. We obtain that

$$\left[k(d-d_0)\right]^{\frac{1}{2}}\left(\hat{\gamma}_{DH}^{trim}-\gamma\right):=I_1+I_2+I_3,$$

where

$$I_1=\gamma\left\{\left[k(d-d_0)\right]^{\frac{1}{2}}\left[\frac{1}{k}\sum_{j=1}^{k}\left(\frac{d_0+1}{d-d_0}\log\frac{Z_j^{(d_0+1)}}{Z_j^{(d+1)}}+\sum_{i=d_0+2}^{d}\frac{1}{d-d_0}\log\frac{Z_j^{(i)}}{Z_j^{(d+1)}}\right)-1\right]\right\},$$

$$\begin{aligned}I_2&=\left[k(d-d_0)\right]^{\frac{1}{2}}\frac{A_0(m/d)}{\rho}\frac{1}{k}\sum_{j=1}^{k}\left\{\left(\frac{dZ_j^{(d+1)}}{m}\right)^{\rho}\frac{d_0+1}{d-d_0}\left(\left(\frac{Z_j^{(d_0+1)}}{Z_j^{(d+1)}}\right)^{\rho}-1\right)\right\}\\&\quad+\left[k(d-d_0)\right]^{\frac{1}{2}}\frac{A_0(m/d)}{\rho}\frac{1}{k}\sum_{j=1}^{k}\left\{\left(\frac{dZ_j^{(d+1)}}{m}\right)^{\rho}\sum_{i=d_0+2}^{d}\frac{1}{d-d_0}\left(\left(\frac{Z_j^{(i)}}{Z_j^{(d+1)}}\right)^{\rho}-1\right)\right\},\end{aligned}$$

$$\begin{aligned}I_3&=o_p(1)\left[k(d-d_0)\right]^{\frac{1}{2}}A_0\left(\frac{m}{d}\right)\frac{1}{k}\sum_{j=1}^{k}\left\{\left(\frac{dZ_j^{(d+1)}}{m}\right)^{\rho\pm\delta}\frac{d_0+1}{d-d_0}\left(\left(\frac{Z_j^{(d_0+1)}}{Z_j^{(d+1)}}\right)^{\rho+\delta}+1\right)\right\}\\&\quad+o_p(1)\left[k(d-d_0)\right]^{\frac{1}{2}}A_0\left(\frac{m}{d}\right)\frac{1}{k}\sum_{j=1}^{k}\left\{\left(\frac{dZ_j^{(d+1)}}{m}\right)^{\rho\pm\delta}\sum_{i=d_0+2}^{d}\frac{1}{d-d_0}\left(\left(\frac{Z_j^{(i)}}{Z_j^{(d+1)}}\right)^{\rho+\delta}+1\right)\right\}.\end{aligned}$$

By Lemma 4.1. and the central limit theorem, we have that $I_1\xrightarrow{d}N\left(0,\gamma^2\right)$, as $n\to\infty$.

$$\begin{aligned}I_2&=\left[k(d-d_0)\right]^{\frac{1}{2}}\frac{A_0(m/d)}{\rho}E\left\{\left(\frac{dZ_1^{(d+1)}}{m}\right)^{\rho}\right\}\\&\quad\cdot\frac{1}{k(d-d_0)}\sum_{j=1}^{k}\frac{\left(\frac{dZ_j^{(d+1)}}{m}\right)^{\rho}}{E\left\{\left(\frac{dZ_1^{(d+1)}}{m}\right)^{\rho}\right\}}\left\{d_0\left(\left(\frac{Z_j^{(d_0+1)}}{Z_j^{(d+1)}}\right)^{\rho}-1\right)+\sum_{i=d_0+1}^{d}\left(\left(\frac{Z_j^{(i)}}{Z_j^{(d+1)}}\right)^{\rho}-1\right)\right\}.\end{aligned}$$

By WLLN for triangular array and $Z_j^{(d+1)}$ is independent with $Z_j^{(i)}/Z_j^{(d+1)}$ for $i=d_0+1,\cdots,d$ and $j=1,\cdots,k$, we have that

$$\frac{1}{k}\sum_{j=1}^{k}\frac{\left(\frac{dZ_j^{(d+1)}}{m}\right)^{\rho}}{E\left\{\left(\frac{dZ_1^{(d+1)}}{m}\right)^{\rho}\right\}}\left\{\frac{d_0+1}{d-d_0}\left(\left(\frac{Z_j^{(d_0+1)}}{Z_j^{(d+1)}}\right)^{\rho}-1\right)+\sum_{i=d_0+2}^{d}\frac{1}{d-d_0}\left(\left(\frac{Z_j^{(i)}}{Z_j^{(d+1)}}\right)^{\rho}-1\right)\right\}$$

$$\xrightarrow{P} E\left\{\frac{d_0+1}{d-d_0}\left(\left(\frac{Z_j^{(d_0+1)}}{Z_j^{(d+1)}}\right)^{\rho}-1\right)+\sum_{i=d_0+2}^{d}\frac{1}{d-d_0}\left(\left(\frac{Z_j^{(i)}}{Z_j^{(d+1)}}\right)^{\rho}-1\right)\right\}$$

$$=\frac{1}{d-d_0}\left\{d_0\left(\frac{d}{d-\rho}\right)^{d-d_0}-\sum_{i=1}^{d_0}\left(\frac{d}{d-\rho}\right)^{d-i+1}+\frac{d\rho}{1-\rho}\right\}$$

as $n\to\infty$, where the second equality follows from a direct calculation. By the Stirling's formula, it follows that,

$$E\left\{\left(\frac{dZ_1^{(d+1)}}{m}\right)^{\rho}\right\}=\left(\frac{m}{d}\right)^{-\rho}\frac{\Gamma(m+1)\Gamma(d-\rho+1)}{\Gamma(d+1)\Gamma(m-\rho+1)}\to\frac{d^{\rho}\Gamma(d-\rho+1)}{\Gamma(d+1)}$$

as $m\to\infty$. Hence, combing with $\left[k(d-d_0)\right]^{1/2}A(m/d)=O(1)$ as $n\to\infty$, we can replace A_0 by A and obtain that as $n\to\infty$,

$$I_2=\left[k(d-d_0)\right]^{\frac{1}{2}}A_0(m/d)B(m,d,d_0,\rho)\left\{1+o_p(1)\right\}. \tag{20}$$

Similarly, as for I_3, we obtain that $I_3\xrightarrow{P}0$ as $n\to\infty$. Combining with $I_1\xrightarrow{d}N(0,\gamma^2)$ and (20) as $n\to\infty$ the statement in Theorem 2.1 follows.

When $\rho=0$, (19) is equivalent to

$$\frac{\log U\left(Z_j^{(i)}\right)-\log U\left(Z_j^{(d+1)}\right)-\gamma\left(\log Z_j^{(i)}-\log Z_j^{(d+1)}\right)}{A_0(m/d)}$$
$$=\log\left(Z_j^{(i)}\right)-\log\left(Z_j^{(d+1)}\right)+o_p(1)\left\{\left(dZ_j^{(i)}/m\right)^{\pm\delta}+\left(dZ_j^{(d+1)}/m\right)^{\pm\delta}\right\},$$

as $n\to\infty$, where $o_p(1)$ term is uniform for all $1\le j\le k$, all $k\in\mathbb{N}$ and $d_0+1\le i\le d+1$. Similarly, we obtain that

$$\left[k(d-d_0)\right]^{\frac{1}{2}}\left(\hat{\gamma}_{DH}^{trim}-\gamma\right):=I_1+I_2+I_3,$$

where

$$I_1=\gamma\left\{\left[k(d-d_0)\right]^{\frac{1}{2}}\left[\frac{1}{k}\sum_{j=1}^{k}\left(\frac{d_0+1}{d-d_0}\log\frac{Z_j^{(d_0+1)}}{Z_j^{(d+1)}}+\sum_{i=d_0+2}^{d}\frac{1}{d-d_0}\log\frac{Z_j^{(i)}}{Z_j^{(d+1)}}\right)-1\right]\right\},$$

$$I_2=\left[k(d-d_0)\right]^{\frac{1}{2}}A_0(m/d)\frac{1}{k}\sum_{j=1}^{k}\left\{\frac{d_0+1}{d-d_0}\log\frac{Z_j^{(d_0+1)}}{Z_j^{(d+1)}}+\frac{1}{d-d_0}\sum_{i=d_0+2}^{d}\log\frac{Z_j^{(i)}}{Z_j^{(d+1)}}\right\},$$

$$I_3 = o_p(1)\left[k(d-d_0)\right]^{\frac{1}{2}} A_0(m/d)\frac{1}{k}\sum_{j=1}^{k}\left\{\left(\frac{dZ_j^{(d+1)}}{m}\right)^{\pm\delta}\frac{d_0+1}{d-d_0}\left(\left(\frac{Z_j^{(d_0+1)}}{Z_j^{(d+1)}}\right)^{\delta}+1\right)\right\}$$
$$+o_p(1)\left[k(d-d_0)\right]^{\frac{1}{2}} A_0(m/d)\frac{1}{k}\sum_{j=1}^{k}\left\{\left(\frac{dZ_j^{(d+1)}}{m}\right)^{\pm\delta}\sum_{i=d_0+2}^{d}\frac{1}{d-d_0}\left(\left(\frac{Z_j^{(i)}}{Z_j^{(d+1)}}\right)^{\delta}+1\right)\right\}.$$

We can show that $I_1 \xrightarrow{d} N(0,\gamma^2)$, $I_2 = \left[k(d-d_0)\right]^{\frac{1}{2}} A_0(m/d)\{1+o_p(1)\}$ and $I_3 \xrightarrow{P} 0$ as $n\to\infty$, similar to the proof above, the statement in Theorem 2.1 follows.

Proof of Theorem 2. 2. We only show the proof for $\rho<0$ and the proof for $\rho=0$ is similar.

By Lemma S.2 in [18], we have $\lim_{n\to\infty}\left\{Z^{(d+1)}>t_0\right\}=1$, for any $t_0>1$. Then by applying (6) twice with $t=m/\bar{d}$ and $x=\bar{d}Z_j^{(i)}/m, i=d_0+1,\cdots,d+1$ and $x=\bar{d}Z_j^{(d_j+1)}/m$ and using the same method as shown in the proof of Theorem 2.1, we obtain that

$$\left[k(\bar{d}-d_0)\right]^{\frac{1}{2}}\left(\hat{\gamma}_{DH}^{trim}-\gamma\right) := I_1+I_2+I_3,$$

where

$$I_1=\gamma\left\{\left[k(\bar{d}-d_0)\right]^{\frac{1}{2}}\left[\frac{1}{k}\sum_{j=1}^{k}\left(\frac{d_0+1}{d_j-d_0}\log\frac{Z_j^{(d_0+1)}}{Z_j^{(d_j+1)}}+\sum_{i=d_0+2}^{d_j}\frac{1}{d_j-d_0}\log\frac{Z_j^{(i)}}{Z_j^{(d_j+1)}}\right)-1\right]\right\},$$

$$I_2=\left[k(\bar{d}-d_0)\right]^{\frac{1}{2}}\frac{A_0(m/\bar{d})}{\rho}\frac{1}{k}\sum_{j=1}^{k}\left\{\left(\frac{\bar{d}}{d_j}\right)^{\rho}\left(\frac{d_jZ_j^{(d_j+1)}}{m}\right)^{\rho}\frac{d_0+1}{d_j-d_0}\left(\left(\frac{Z_j^{(d_0+1)}}{Z_j^{(d_j+1)}}\right)^{\rho}-1\right)\right\}$$
$$+\left[k(\bar{d}-d_0)\right]^{\frac{1}{2}}\frac{A_0(m/\bar{d})}{\rho}\frac{1}{k}\sum_{j=1}^{k}\left\{\left(\frac{\bar{d}}{d_j}\right)^{\rho}\left(\frac{d_jZ_j^{(d_j+1)}}{m}\right)^{\rho}\sum_{i=d_0+2}^{d_j}\frac{1}{d_j-d_0}\left(\left(\frac{Z_j^{(i)}}{Z_j^{(d_j+1)}}\right)^{\rho}-1\right)\right\},$$

$$I_3=o_p(1)\left[k(\bar{d}-d_0)\right]^{\frac{1}{2}}A_0(m/\bar{d})\frac{1}{k}\sum_{j=1}^{k}\left\{\left(\frac{\bar{d}}{d_j}\right)^{\rho\pm\delta}\left(\frac{d_jZ_j^{(d_j+1)}}{m}\right)^{\rho\pm\delta}\frac{d_0+1}{d_j-d_0}\left(\left(\frac{Z_j^{(d_0+1)}}{Z_j^{(d_j+1)}}\right)^{\rho+\delta}+1\right)\right\}$$
$$+o_p(1)\left[k(\bar{d}-d_0)\right]^{\frac{1}{2}}A_0(m/\bar{d})\frac{1}{k}\sum_{j=1}^{k}\left\{\left(\frac{\bar{d}Z_j^{(d_j+1)}}{m}\right)^{\rho\pm\delta}\sum_{i=d_0+2}^{d_j}\frac{1}{d_j-d_0}\left(\left(\frac{Z_j^{(i)}}{Z_j^{(d_j+1)}}\right)^{\rho+\delta}+1\right)\right\}.$$

as $n\to\infty$.

By Lemma 4.1. and the central limit theorem, we have that $I_1\xrightarrow{d}N(0,\gamma^2)$ as $n\to\infty$.

As for I_2, by WLLN for triangular array and for each *j*, $Z_j^{(d+1)}$ is independent with $Z_j^{(i)}/Z_j^{(d+1)}$, where $i=d_0+1,\cdots,d_j$, similar to the proof of Theorem 2.1, we have that

$$I_2 = \left[k\left(\bar{d}-d_0\right)\right]^{\frac{1}{2}} A_0\left(m/\bar{d}\right)\frac{1}{k}\sum_{j=1}^{k}\left(\frac{\bar{d}}{d_j}\right)^{\rho} B\left(m,d_j,d_0,\rho\right)\left\{1+o_p\left(1\right)\right\}, \tag{21}$$

as $n \to \infty$.

Similarly, as for I_3, we obtain that $I_3 \overset{P}{\to} 0$ as $n \to \infty$. Combining with $I_1 \overset{d}{\to} N\left(0,\gamma^2\right)$ and (21) as $n \to \infty$ the statement in Theorem 2.2 follows.

Proof of Theorem 2.3. We only show the proof for $\rho < 0$ and the proof for $\rho = 0$ is similar.

By Lemma S.2 in [18], we have $\lim_{n\to\infty}\left\{Z^{(d+1)} > t_0\right\} = 1$, for any $t_0 > 1$. Then by applying (6) twice with $t = m/d$ and $x = dZ_j^{(i)}/m, i = d_0+1,\cdots,d+1$ and $x = dZ_j^{(d+1)}/m$ and using the same method in the prood of Theorem 2.1, we obtain that

$$\left[k\left(d-d_0\right)\right]^{\frac{1}{2}}\left(\hat{\gamma}_{DH}^{trim}-\gamma\right) := I_1 + I_2 + I_3,$$

where

$$I_1 = \gamma\left\{\left[k\left(d-d_0\right)\right]^{\frac{1}{2}}\left[\frac{1}{k}\sum_{j=1}^{k}\left(\frac{d_0+1}{d-d_0}\log\frac{Z_j^{(d_0+1)}}{Z_j^{(d+1)}} + \sum_{i=d_0+2}^{d}\frac{1}{d-d_0}\log\frac{Z_j^{(i)}}{Z_j^{(d+1)}}\right)-1\right]\right\},$$

$$I_2 = \left[k\left(d-d_0\right)\right]^{\frac{1}{2}}\frac{A_0\left(m/d\right)}{\rho}\frac{1}{k}\sum_{j=1}^{k}\left\{\left(\frac{dZ_j^{(d+1)}}{m}\right)^{\rho}\frac{d_0+1}{d-d_0}\left(\left(\frac{Z_j^{(d_0+1)}}{Z_j^{(d+1)}}\right)^{\rho}-1\right)\right\}$$
$$+\left[k\left(d-d_0\right)\right]^{\frac{1}{2}}\frac{A_0\left(m/d\right)}{\rho}\frac{1}{k}\sum_{j=1}^{k}\left\{\left(\frac{dZ_j^{(d+1)}}{m}\right)^{\rho}\sum_{i=d_0+2}^{d}\frac{1}{d-d_0}\left(\left(\frac{Z_j^{(i)}}{Z_j^{(d+1)}}\right)^{\rho}-1\right)\right\},$$

$$I_3 = o_p\left(1\right)\left[k\left(d-d_0\right)\right]^{\frac{1}{2}} A_0\left(\frac{m}{d}\right)\frac{1}{k}\sum_{j=1}^{k}\left\{\left(\frac{dZ_j^{(d+1)}}{m}\right)^{\rho\pm\delta}\frac{d_0+1}{d-d_0}\left(\left(\frac{Z_j^{(d_0+1)}}{Z_j^{(d+1)}}\right)^{\rho+\delta}+1\right)\right\}$$
$$+o_p\left(1\right)\left[k\left(d-d_0\right)\right]^{\frac{1}{2}} A_0\left(\frac{m}{d}\right)\frac{1}{k}\sum_{j=1}^{k}\left\{\left(\frac{dZ_j^{(d+1)}}{m}\right)^{\rho\pm\delta}\sum_{i=d_0+2}^{d}\frac{1}{d-d_0}\left(\left(\frac{Z_j^{(i)}}{Z_j^{(d+1)}}\right)^{\rho+\delta}+1\right)\right\}.$$

By Lemma 4.1. and the central limit theorem, we have that $I_1 \overset{d}{\to} N\left(0,\gamma^2\right)$ as $n \to \infty$.

By WLLN for triangular array and $Z_1^{(d+1)}$ is independent with $Z_j^{(i)}/Z_j^{(d+1)}$ for $i = d_0+1,\cdots,d$ and $j = 1,\cdots,k$, we have that

$$I_2 = \left[k\left(d-d_0\right)\right]^{\frac{1}{2}} A_0\left(m/d\right) E\left\{\left(\frac{dZ_1^{(d+1)}}{m}\right)^{\rho}\right\}\frac{d}{d-d_0}\frac{1}{1-\rho}\left\{1+o_p\left(1\right)\right\}. \tag{22}$$

By the Stirling's formula, it follows that

$$E\left\{\left(\frac{dZ_1^{(d+1)}}{m}\right)^{\rho}\right\} = \left(\frac{m}{d}\right)^{-\rho}\frac{\Gamma\left(m+1\right)\Gamma\left(d-\rho+1\right)}{\Gamma\left(d+1\right)\Gamma\left(m-\rho+1\right)} \sim 1$$

as $m \to \infty$.

Similarly, as for I_3, we obtain that $I_3 \xrightarrow{P} 0$ as $n \to \infty$. Combining with $I_1 \xrightarrow{d} N\left(0, \gamma^2\right)$ and (22) as $n \to \infty$ the statement in Theorem 2.3 follows.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

[1] de Haan, L. and Ferreira, A. (2006) Extreme Value: An Introduction. Springer, New York.

[2] Dekkers, A.L.M. and de Haan, L. (1989) On the Estimation of the Extreme-Value Index and Large Quantile Estimation. *The Annals of Statistics*, **17**, 1795-1832. https://doi.org/10.1214/aos/1176347396

[3] Dekkers, A.L.M., Einmahl, J.H.J. and de Haan, L. (1989) A Moment Estimator for the Index of an Extreme-Value Distribution. *The Annals of Statistics*, **17**, 1833-1855. https://doi.org/10.1214/aos/1176347397

[4] Beirlant, J., Goegebeur, Y., Teugels, J. and Segers, J. (2004) Statistics of Extremes: Theory and Applications. John Wiley & Sons, Ltd., New York. https://doi.org/10.1002/0470012382

[5] Li, J., Peng, Z. and Nadarajah, S. (2010) Asymptotic Normality of Location Invariant Heavy Tail Index Estimator. *Extremes*, **13**, 269-290. https://doi.org/10.1007/s10687-009-0088-4

[6] Brazauskas, V. and Serfling, R. (2001) Robust Estimation of Tail Parameters for Two-parameter Pareto and Exponential Models via Generalized Quantile Statistics. *Extremes*, **3**, 231-249.

[7] Bhattacharya, S., Kallitsis, M. and Stoev, S. (2019) Data-Adaptive Trimming of the Hill Estimator and Detection of Outliers in the Extremes of Heavy-Tailed Data. *Electronic Journal of Statistics*, **13**, 1872-1925. https://doi.org/10.1214/19-EJS1561

[8] Bladt, M., Albrecher, H. and Beirlant, J. (2020). Threshold Selection and Trimming in Extremes. *Extremes*, **23**, 629-665. https://doi.org/10.1007/s10687-020-00385-0

[9] Aban, I.B., Meerschaert, M.M. and Panorska, A.K. (2006) Parameter Estimation for the Truncated Pareto Distribution. *Journal of the American Statistical Association*, **101**, 270-277. https://doi.org/10.1198/016214505000000411

[10] Beirlant, J., Alves, I.F., Gomes, I. and Meerschaert, M.M. (2014) Extreme Value Statistics for Truncated Pareto-Type Distributions. John Wiley & Sons, Ltd., New York. https://doi.org/10.1002/9781118445112.stat04418

[11] Beirlant, J., Guillou, A. and Dierckx, G. (2007) Estimation of the Extreme Value Index and Extreme Quantiles under Random Censoring. *Extremes*, **10**, 151-174. https://doi.org/10.1007/s10687-007-0039-x

[12] Einmahl, J.H., Fils-Villetard, A. and Guillou, A. (2008) Statistics of Extremes under Random Censoring. *Bernoulli*, **14**, 207-227. https://doi.org/10.3150/07-BEJ104

[13] Gomes, M.I. and Neves, M.M. (2011) Estimation of the Extreme Value Index for Randomly Censored Data. *In Biometrical Letters*, **48**, 1-22.

[14] Bladt, M., Albrecher, H. and Beirlant, J. (2021) Trimmed Extreme Value Estimators

for Censored Heavy-Tailed Data. *Electronic Journal of Statistics*, **15**, 3112-3136. https://doi.org/10.1214/21-EJS1857

[15] Battey, H., Fan, J., Liu, H., Lu, J. and Zhu, Z. (2015) Distributed Estimation and Inference with Statistical Guanrantees. arXiv: 1509.05457. https://doi.org/10.48550/arXiv.1509.05457

[16] Li, R., Lin, D.K.J. and Li, B. (2013) Statistical Inference in Massive Data Sets. *Applied Stochastic Models in Business and Industry*, **29**, 399-409. https://doi.org/10.1002/asmb.1927

[17] Volgushev, S., Chao, S.-K. and Cheng, G. (2019). Distributed Inference for Quantile Regression Processes. *The Annals of Statistics*, **47**, 1634-1662. https://doi.org/10.1214/18-AOS1730

[18] Chen, L., Li, D. and Zhou, C. (2021). Distributed Inference for Extreme Value Index. *Biometrika*, **109**, 257-264. https://doi.org/10.1093/biomet/asab001

[19] Wellner, R.L. (1978) Limit Theorems for the Ratio of the Empirical Distribution Function to the True Distribution Fuction. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, **45**, 73-88. https://doi.org/10.1007/BF00635964