University of Nevada, Reno

ESTIMATION OF TAIL INDICES OF HEAVY-TAILED DISTRIBUTIONS

A Thesis Submitted in Partial Fulfillment
of the Requirements for the Degree of Master of Science in
Mathematics

by
Md. Rashidul Hasan
Dr. Javier Rojo / Thesis Advisor
August 2016

© 2016 Md. Rashidul Hasan

ALL RIGHTS RESERVED

THE GRADUATE SCHOOL

We recommend that the thesis
prepared under our supervision by

MD RASHIDUL HASAN

Entitled

Estimation Of Tail Indices Of Heavy-Tailed Distributions

be accepted in partial fulfillment of the
requirements for the degree of

MASTER OF SCIENCE

Javier Rojo, Ph.D., Advisor

Raul Rojas, Ph.D., Committee Member

Luis Rodolfo Garcia Carrillo, Ph.D., Graduate School Representative

David W. Zeh, Ph.D., Dean, Graduate School

August, 2016

Abstract

Heavy-tailed distributions have found many real life applications in analyzing extreme values. Therefore extreme value theory has received much attention in recent years. Several estimators for the tail index of heavy-tailed distributions have been proposed in the literature and their properties have been established. These estimators suffer from some drawbacks. Most of the estimators including Hill and Pickands, make use of intermediate upper order statistics, $X_{(k(n))}, X_{(k(n)+1)}, ..., X_{(n)}$. Choosing the suitable value of k is painstaking work. The estimated values can be highly sensitive to the choice of k and the choice of slowly varying function. Here, a simple and very useful estimator is proposed and its operating characteristics are examined in terms of its bias and mean-squared-error properties. To compare with other estimators, we introduce several slowly varying functions and generate some synthetic data. Then we compute our new estimator using the generated data and compare with other estimators. In addition, using the new estimator we develop some block estimators and examine the behavior of these block estimators too. These block estimators show excellent results for every survival function. The estimators are illustrated on real data-sets.

Acknowledgements

I would like to express my gratitude to my advisor Professor Javier Rojo for his encouragement, support, and all the help throughout the thesis work. I would also like to thank National Science Foundation (grant CMMI 1436844) for supporting my research, all of the faculty, graduate students, and staff of the department of Mathematics and Statistics at UNR.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

The concept of a heavy-tailed probability distribution plays an important role in applications in numerous fields such as finance, internet traffic, environmental science, civil engineering, and insurance. For example, in finance, Li and Wei (2011) and LeBaron and Samanta (2005) provide an overview of the topic and analyze the data of stock market close indexes; in insurance, Resnick (2007) and Embrechts et al. (2013) use heavy-tailed distribution for insurance claims; in economics, Huisman et al. (2001) and Koedijk et al. (1990) estimate tail index for extreme returns of exchange rates; in internet traffic, Rezaul and Grout (2007) and Markovich (2005) use heavy tails for internet traffic; in geology, Caers et al. (1999) use heavy-tailed distributions for large magnitudes of earthquakes; in hydrology, Gardes and Girard (2010), Meerschaert and Scheffler (2003), Willems et al. (2007) and El Adlouni et al. (2008) use heavy-tailed distributions for practical application of hydrologic extremes (e.g. river flows, rainfall or discharge), and in environmental science, Phalitnonkiat et al. (2016) use heavy tails for extreme ozone events. Different types of heavy-tailed distributions are commonly used in these applications. Because of the heavy-tailed nature of these distributions, not all moments exist and this represents a technical difficulty, see, e.g., Rojo (1996). These difficulties motivated our interest in studying estimators of heavy-tailed indices. Parametric and semiparametric estimators have been proposed, (see, e.g., Hill (1975); Pickands III (1975)) to estimate the tail indices of heavy-tailed distributions. A heavy-tailed distribution is defined by

$$\bar{F}(x) = P(X > x) = x^{-\alpha}h(x) \tag{1.1}$$

where $\alpha > 0$ is the "tail index" and for $x > 0, h(x)$ is a slowly varying function defined as

$$\lim_{x\to\infty} \frac{h(tx}{h(x)} = 1 \quad for \quad all \quad t > 0 \tag{1.2}$$

Rojo (2013) presented a classification scheme for distribution function using the theory of function of regular variation. Let the order statistics be $X_{1,n} \leq X_{2,n} \leq ...... \leq X_{n,n}$. where $X_{1,n} = min(X_1, ..., X_n)$ and $X_{n,n} = max(X_1, ...X_n)$ and X_k is the k^{th} order statistic. Rojo (1996) provided a more precise characterization of various classes of distributions as a theorem.

Theorem 1.0.1. *Let $S_n = X_{n,n} - X_{n-1,n}$ be the n^{th} extreme spacing in a random variable of size n and let $F(x)$ be a distribution function then*

F is short-tailed $\Leftrightarrow S_n = X_{n,n} - X_{n-1,n} \xrightarrow{a.s.} 0$*,*

F is medium-tailed $\Leftrightarrow S_n = X_{n,n} - X_{n-1,n} \xrightarrow{a.s.} Exp(\theta)$*,*

F is long-tailed $\Leftrightarrow S_n = X_{n,n} - X_{n-1,n} \xrightarrow{a.s.} \infty$*,*

Where $Exp(\theta)$ denote the exponential distribution with parameter θ. i.e. mean $1/\theta$.

Let $X_1, X_2, ...$ be independent and identically distributed random variables from a distribution F. Then we have $P(max(X_1, ..., X_n)) = P(X_1 \leq x, X_2 \leq x, ..., X_n \leq x) = F^n(x)$. Suppose that there are sequences $a_n > 0$, and b_n with $n = 1, 2, ...$ such that

$$\lim_{n\to\infty} F^n(a_n x + b_n) = G(x) \tag{1.3}$$

where G is a nondegenerate distribution function. The class of distribution functions F satisfying (1.3) is called the maximum domain of attraction or simply domain of attraction of G. The distribution functions G that can occur as a limit are called extreme value distributions. We can find alternative characterizations from (1.3). Fisher and Tippett (1928) and Gnedenko (1943) provided the following theorem that define the class of possible limit distributions in (1.3).

Theorem 1.0.2. *The class of extreme value distributions is $G_\gamma(ax+b)$ with $a>0$, b real, where*

$$G_\gamma(x) = exp(-(1+\gamma x)^{-1/\gamma}) \quad , 1+\gamma x > 0 \tag{1.4}$$

with γ real and where for $\gamma = 0$ the right-hand is interpreted as $exp(-e^{-x})$. The parameter γ is called the extreme value index. If relation (1.3) holds with $G = G_\gamma$ for some $\gamma \in \mathbb{R}$, we say that the distribution function F is in the domain of attraction of G_γ and denote it as $D(G_\gamma)$. The parametrization in theorem (1.0.2) is due to von Mises (1936) and Jenkinson (1955). Moreover, the limit distribution functions have a form of a simple explicit one-parameter family apart from the scale and location parameters. We can now define some sub classes according to $\gamma > 0, \gamma = 0$, and $\gamma < 0$ respectively (De Haan and Ferreira (2007)):

(**a**) For $\gamma > 0$ clearly $G_\gamma(x) < 1$ for all x, i.e., the right endpoint of the support of the distribution is infinity. Moreover, as $x \longrightarrow \infty$, $1 - G_\gamma(x) \sim \gamma^{-1/\gamma}x^{-1/\gamma}$, i.e., the distribution has a heavy right tail; for example, moments of order greater than or equal to $1/\gamma$ do not exist.

(**b**) For $\gamma = 0$ the right endpoint of the support of the distribution equals infinity. The distribution, however, is rather light-tailed: $1 - G_0(x) \sim e^{-x}$ as $x \longrightarrow \infty$, and all moments exist.

(**c**) For $\gamma < 0$ the right endpoint of the support of the distribution is $-1/\gamma$ so it has a short tail, verifying $1 - G_\gamma(-\gamma^{-1} - x) \sim (-\gamma x)^{-1/\gamma}$, as $x \downarrow 0$.

Another alternative characterization of the limit distribution in theorem (1.0.2) is possible as follows (see De Haan and Ferreira (2007)):

(**a**) For γ >0 use $G_\gamma((x-1))/\gamma$ and get with $\alpha = 1/\gamma > 0,$

$$\Phi_\alpha(x) = \begin{cases} 0 & , x \leq 0, \\ exp(-x^{-\alpha}) & , x > 0. \end{cases}$$

This is often called the frechet class of distributions (Frechet(1927)).

(**b**) The distribution function with $\gamma = 0$,

$$G_0(x) = exp(-e^{-x}),$$

for real x , is called the Gumbel distribution.

(**c**) for $\gamma < 0$ use $G_\gamma(-(1+x)/\gamma)$ and get with $\alpha = -1/\gamma > 0,$

$$\Psi_\alpha(x) = \begin{cases} exp(-(-x)^\alpha) & , x < 0, \\ 1 & , x \geq 0. \end{cases}$$

This class is sometimes called the reverse-Weibull class of distributions.

In analysis of heavy-tailed distribution the index, α is always positive. The smaller the α the slower the decay of $\bar{F}(x)$ and this generates more extreme values (Meerschaert and Scheffler (2003)). When this phenomenon is discussed for the distribution of univariate random variables, we find that regularly varying functions are the appropriate mathematical concept for the development of analysis of heavy-tailed indices. We also need to define regularly varying functions.

Definition 1.0.1 (De Haan and Ferreira (2007))**.** A Lebesgue measurable function $h : \mathbb{R}^+ \longrightarrow \mathbb{R}$ that is eventually positive is regularly varying (at infinity) if for some

$\alpha \in \mathbb{R}$,

$$\lim_{t\to\infty} \frac{h(tx)}{h(t)} = x^{\alpha}, \quad x > 0.$$

The number α in the above definition is called the index of regular variation (RV). A function satisfying this equation with $\alpha = 0$ is called slowly varying .

We mostly face three problems which appear naturally in dealing with heavy-tailed distributions explained in detail by Smith (1987). One of them is estimating an index of regular variation. The other two are estimating an endpoint and estimating tail probabilities.

This thesis is organized as follows. Different types of estimators for estimating the tail index of heavy-tailed distributions are discussed in chapter 2. In chapter 3 we review the applications of heavy-tailed distributions and the performance of the different estimators of underlying distribution. The analyses of some data-sets taken from real life events are presented. A new estimator is presented and discussed in chapter 4 for large and small values of α. Using the definition of regularly varying functions, we define some distribution function and generate data-sets from these distribution functions to compute and study in new estimator. Block estimators are also introduced using new estimator and applied to those functions. In addition, we compare our estimates with the Hill's, Pickands and De Haan and Resnick's estimators. In chapter 5, we choose data-sets from S&P 500 and Danish fire insurance claims and apply our estimator to compare with other estimators. After analyzing and applying our estimator we give our conclusions in chapter 6.

Chapter 2

Different estimators

There are parametric and non-parametric approaches to the problem of the tail-index estimation. In this chapter we describes numerous estimators proposed in the literature.

2.1 Hill's estimator

One of the most popular estimators was proposed by Hill (1975). Let X_1,....,X_n be i.i.d. (independent and identically distributed) and define the order statistics as $X_{1,n} \leq ..... \leq X_{n,n}$ where $X_{1,n}$=min(X_1......X_n) and $X_{n,n}$=max(X_1.......X_n). Then $X_{k,n}$ is the k^{th} upper order statistic. Consider the order statistics $X_{n-k,n}$ with $n \longrightarrow \infty, k = k(n) \longrightarrow \infty$, and $k(n)/n \longrightarrow 0$. These are called intermediate order statistics (for details see De Haan and Ferreira (2007)). Hill's estimator is defined by

$$\hat{\alpha} = \hat{\alpha}_{k,n} = (\frac{1}{k}\sum_{j=0}^{k-1} lnX_{n-j,n} - lnX_{n-k,n})^{-1} \tag{2.1}$$

where $k = k(n) \longrightarrow \infty$ defines as an increasing sequence of upper order statistics. Below we summarize the main properties of Hill's estimator in a theorem. Before stating the theorem, we define weak and strong consistency. Consider a sequence of i.i.d. random variables from a distribution with a parameter α. Let $\hat{\alpha}$ be an estimator of α. The estimator $\hat{\alpha}$ is weakly consistent as an estimator of α if $\hat{\alpha} \xrightarrow{p} \alpha$ or,

$$\lim_{n\to\infty} P(|\hat{\alpha} - \alpha| \leq \epsilon) = 1 \quad \forall \epsilon > 0$$

and we define strong consistency if $\hat{\alpha} \xrightarrow{a.s.} \alpha$ or,

$$P(\lim_{n\to\infty} |\hat{\alpha} - \alpha| \leq \epsilon) = 1 \quad \forall \epsilon > 0$$

An i.i.d sequence (X_n) has representation as a linear process (Embrechts et al. (2013)), i.e.

$$X_n = \sum_{j=-\infty}^{\infty} \psi_j Z_{n-j}, \quad n \in \mathbb{Z} \tag{2.2}$$

where ψ_j are real and Z_n are i.i.d.sequence. Strictly stationary (Embrechts et al. (2013)) is a sequence of random variables (X_n) if its finite-dimensional distributions are invariant under shifts of time, i.e. $(X_{t_1}, ..., X_{t_m}) \stackrel{d}{=} (X_{t_1+h}, ..., X_{t_m+h})$ for any choice of indices $t_1 < .... < t_m$ and integers h. A strictly stationary sequence is weakly dependent if x_t and x_{t+h} are "almost independent" as h increases. Next we mention the theorem

Theorem 2.1.1 (Properties of the Hill estimator (see De Haan and Resnick (1998); Rodionov (2014); Resnick (2007); Embrechts et al. (2013))). *Let $\hat{\alpha} = \hat{\alpha}_{k,n}$ be the Hill's estimator*

***(a)** (Weak consistency) Assume that one of the following conditions is satisfied*

$-(X_n)$ *is iid*

$-(X_n)$ *is weakly dependent*

$-(X_n)$ is a linear process

If $k \longrightarrow \infty$, $\frac{k}{n} \longrightarrow 0$ for $n \longrightarrow \infty$, then $\hat{\alpha} \xrightarrow{p} \alpha$.

***(b)** (Strong consistency) If $\frac{k}{n} \longrightarrow 0$, $\frac{k}{ln(ln(n))} \longrightarrow \infty$ for $n \longrightarrow \infty$, and (X_n) iid sequence, then $\hat{\alpha} \xrightarrow{a.s.} \alpha$.*

***(c)** (Asymptotic normality) If (X_n) is an iid sequence and if $k \longrightarrow \infty$, $\frac{k}{n} \longrightarrow 0$ for $n \longrightarrow \infty$, then*

$$\sqrt{k}(\hat{\alpha}-\alpha) \xrightarrow{d} N(0, \alpha^2). \tag{2.3}$$

Thus the Hill's estimator has an asymptotic variance of α^2. If the value of k increases, the asymptotic variance α^2/k of $\hat{\alpha}$ decreases but Hill's estimator can exhibit considerable bias in this case. Choosing the right upper order statistic of $k(n)$ requires a lot of experience and skill (see Beirlant et al. (1996)). Since Hill's estimator is scale invariant, it is widely used in estimating thickness of heavy tails. On the other hand, this estimator is not shift invariant. That means that a shift location of the data can affect the accuracy of the estimation. Embrechts et al. (2013) indicated that this the situation arises when the slowly varying function is far away from a constant in the tail, this estimator produces so-called "Hill horror plot". To illustrate Hill's horror plot, we generated 40000 observations from the survival function $\bar{F} = \frac{1}{xlnx}$ with $\alpha = 1$. Figure 2.1 shows the histograms of estimates of the Hill's estimator. The histogram for $k = 500$ is close to 1 with small bias. Hill's estimator misleads us for large values of k. When $k = 15000$ the histogram of Hill's estimator is far away from the value 1.

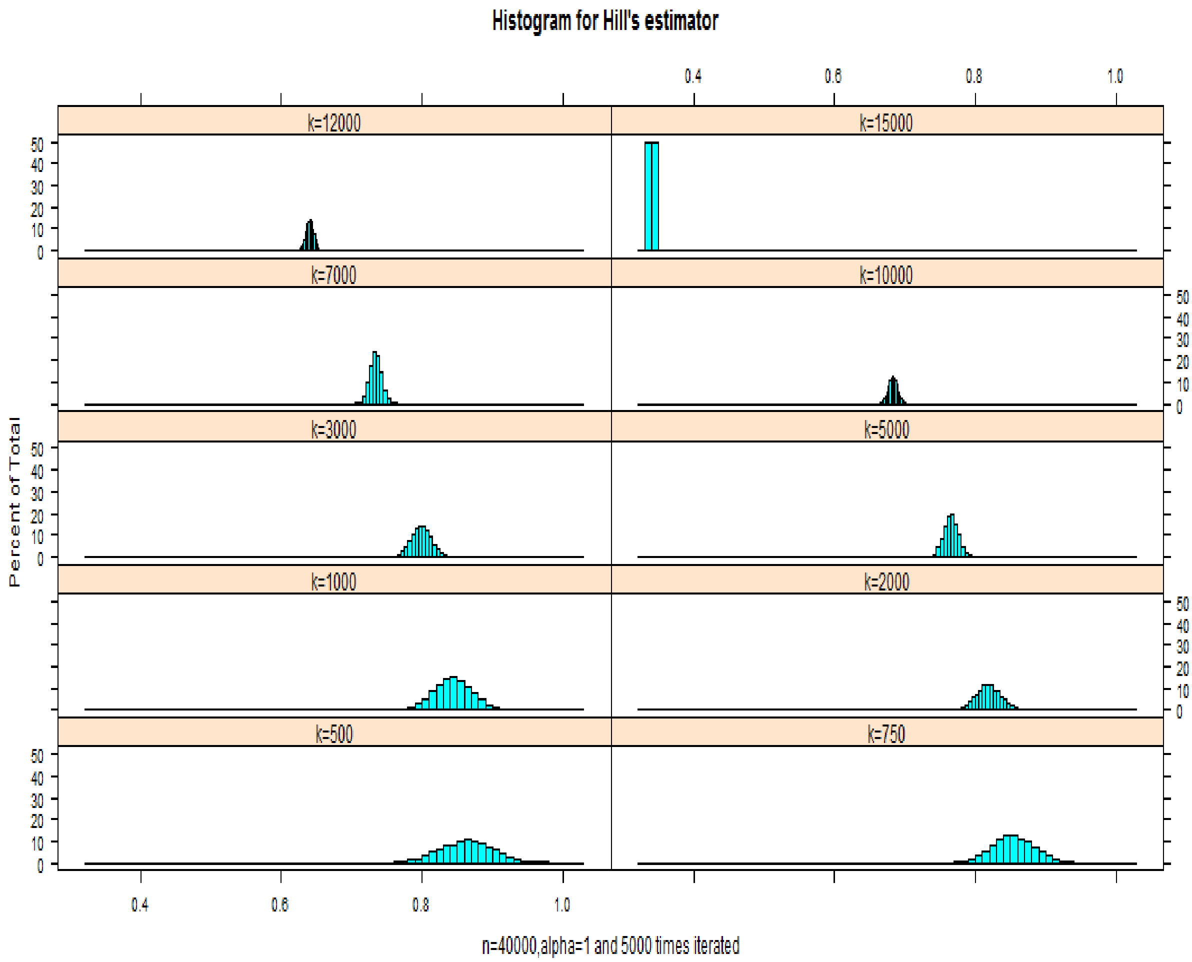

Figure 2.1: Hill horror plot for different values of k

2.2 Deckers-Enmahl-de Haan's estimator (DEdH)

The Hill's estimator is specially designed for $\alpha > 0$. In developing Hill estimator for the general case of all $\alpha \in \mathbb{R}$, Dekkers et al. (1989) proposed a new estimator. This estimator for $\hat{\alpha}$ is also referred to as a moment estimator of α (Embrechts et al. (2013)). Let $X_1, X_2, ...$ be i.i.d. random variables with distribution function F and for $k = k(n) \longrightarrow \infty, \frac{k}{n} \longrightarrow 0$, $n \longrightarrow \infty$, the moment estimator is defined (see De Haan and Ferreira (2007)) by

$$\hat{\alpha}_M = (M_n^{(1)} + 1 - \frac{1}{2}(1 - \frac{(M_n^{(1)})^2}{M_n^{(2)}})^{-1})^{-1} \tag{2.4}$$

where for $j = 1, 2,$

$$M_n^{(j)} = \frac{1}{k}\sum_{i=0}^{k-1}(logX_{n-i,n} - logX_{n-k,n})^j. \tag{2.5}$$

Note that $M_n^{(1)}$ is the reciprocal of Hill's estimator. For $\alpha \in \mathbb{R}$, $\hat{\alpha}_M$ is consistent for α. The moment estimator is not location invariant but it is scale invariant.

2.3 Shifted Hill's estimator

Since Hill's estimator is not shift invariant, Aban and Meerschaert (2001) modified the Hill's estimator and proposed a shifted Hill's estimator as follows

$$\hat{\alpha} = [\frac{1}{r}\sum_{i=1}^{r} ln(X_i - \hat{s}) - ln(X_{r+1} - \hat{s})]^{-1} \tag{2.6}$$

where $X_r > X_{r+1}$ and the optimal shift $\hat{s}$ satisfies the equation

$$\hat{\alpha}(X_{r+1} - \hat{s})^{-1} = (\hat{\alpha} + 1)r^{-1} \sum_{i=1}^{r} (X_i - \hat{s})^{-1} \tag{2.7}$$

Their research investigated the behavior of the shifted Hill's estimator for data from stable distributions with the values of $0 < \alpha \leq 2$ and $-1 \leq \beta \leq 1$. They concluded that the original Hill's estimator gives very good results for smaller value of α but performs poorly for stable distributions with $1.5 < \alpha \leq 2$. Moreover, the shifted Hill's estimator does better when α lies between 1.5 and 2. However, although the shifted Hill's estimator is shift and scale invariant, the resulting estimator has a relatively wide sampling distribution (Bianchi and Meerschaert (2000)).

2.4 Pickands estimator

Another simple estimator was proposed by Pickands III (1975). We now define the Pickands estimator for $\hat{\gamma}(= 1/\alpha)$

$$\hat{\gamma}_{k,n} = \frac{1}{ln2} ln \frac{X_{n-k,n} - X_{n-2k,n}}{X_{n-2k,n} - X_{n-4k,n}} \tag{2.8}$$

where $k \longrightarrow \infty$, $\frac{k}{n} \longrightarrow 0$ for $n \longrightarrow \infty$, then

$$\hat{\gamma} \xrightarrow{P} \gamma, n \longrightarrow \infty.$$

Below we state a theorem to discuss properties of Pickands estimator.

Theorem 2.4.1 (Properties of the Pickands estimator (Embrechts et al. (2013)))**.** *. Let $\hat{\gamma} = \hat{\gamma}_{k,n}$ be the Pickands estimator* (**a**) *(Weak consistency) If* $k \longrightarrow \infty$, $\frac{k}{n} \longrightarrow 0$ *for* $n \longrightarrow \infty$, *then*

$$\hat{\gamma} \xrightarrow{p} \gamma, \quad n \longrightarrow \infty.$$

(**b**) *(Strong consistency) If* $\frac{k}{n} \longrightarrow 0, \frac{k}{lnlnn} \longrightarrow \infty$ *for* $n \longrightarrow \infty$, *and* (X_n) *is an i.i.d. sequence, then*

$$\hat{\gamma} \xrightarrow{a.s.} \gamma, \quad n \longrightarrow \infty.$$

(**c**) *(Asymptotic normality) If* $k \longrightarrow \infty$, $\frac{k}{n} \longrightarrow 0$ *for* $n \longrightarrow \infty$ *and* (X_n) *is an i.i.d. sequence, then*

$$\sqrt{k}(\hat{\gamma}-\gamma) \xrightarrow{d} N(0, \nu(\gamma)), \quad n \longrightarrow \infty. \tag{2.9}$$

where

$$\nu(\gamma) = \frac{\gamma^2(2^{2\gamma+1}+1)}{(2(2^\gamma - 1)ln2)^2}$$

Pickands estimator is weakly (or strongly) consistent under the asymptotic condition in for γ. This estimator is invariant under shift and scale transformations. Since asymptotic variance is approximately $2(\frac{\gamma}{2})^2$, it has larger asymptotic variance with the increases of values of k.

2.5 QQ-estimator

A visual way for assessing goodness of fit and estimating location and scale parameter is the qq-plot. Kratz and Resnick (1996) applied this idea to a data-set generated from Pareto distribution (a distribution with heavy tail). However, for estimating a heavy-tailed index, the qq-plot technique is used for estimating location and scale parameters. The method of estimating the index parameter is motivated by the following observation: If $U_{1,n} \leq U_{2,n} \leq ....U_{n,n}$ are the order statistics of n i.i.d observations which are uniformly distributed on [0,1], then by symmetry

$$E(U_{i+1,n} - U_{i,n}) = \frac{1}{n+1}$$

and hence

$$EU_{i,n} = \frac{i}{n+1}.$$

where $U_{i,n}$ is close to its mean $i/(n+1)$ and the plot of $(1/(n+1), U_{i,n}), 1 \leq i \leq n$ may be roughly linear. Now we define the left-continuous inverse function

Definition 2.5.1. Suppose $G : \mathbb{R} \mapsto (a,b)$ is a nondecreasing function on $\mathbb{R}$ with range (a,b), where $-\infty \leq a < b \leq \infty$. With the convention that the infimum of an empty set is $+\infty$, we define the (left-continuous) inverse(see Resnick (2007)) $G^{-1} : (a,b) \mapsto \mathbb{R}$ of G as

$$G^{-1}(y) = inf\{s : G(s) \geq y\}$$

Now suppose that $X_{1,n} \leq X_{2,n} \leq .....X_{n,n}$ are the order statistics from an i.i.d sample of size n. If this random sample is from a specific continuous distribution G, then the plot of $\{(i/(n+1), G(X_{i,n})), 1 \leq i \leq n\}$ is approximately linear and hence the plot of $\{G^{-1}(i/(n+1), X_{i,n}), 1 \leq i \leq n\}$ is also approximately linear. Here $G^{-1}(i/(n+1))$ is a theoretical quantile and $X_{i,n}$ is the corresponding quantile of the empirical distribution function and hence the name $qq-plot$. Let the data come from a location-scale family

$$G_{\mu,\sigma}(x) = G_{0,1}(\frac{x-\mu}{\sigma}) \tag{2.10}$$

where μ, σ are unknown. The plot of $\{(G^{-1}_{\mu,\sigma}(i/(n+1), X_{i,n}), 1 \leq i \leq n\}$ should follow a line through 0 of slope 1 and since

$$G^{-1}_{\mu,\sigma}(y) = \sigma G^{-1}_{0,1}(y) + \mu$$

the plot of

$$\{(G^{-1}_{0,1}(i/(n+1), X_{i,n}), 1 \leq i \leq n\}$$

is approximately a line of slope σ and intercept μ.

This gives us visual assessment of the goodness of fit of the location-scale family and provides estimates of μ, σ. To use this technique in heavy tails suppose we have $Z_{1,n} \leq .... \leq Z_{n,n}$ the order statistics from a Pareto family indexed by its shape parameter $\alpha > 0$ that is, $F_{\alpha}(x) = 1 - x^{-\alpha}, x \geq 1$. Then we get for $y > 0$

$$G_{0,\alpha(y)} = P[logZ_1 > y] \ = \ e^{-\alpha y}$$

and the plot of

$$\{(G^{-1}_{0,1}(\frac{i}{n+1}), logZ_{i,n}), 1 \leq i \leq n\} = \{(-log(1 - \frac{i}{n+1}, logZ_{i,n}), 1 \leq i \leq n\}$$

is also approximately a line with intercept 0 and slope α^{-1}. If $\{(x_i, y_i), 1 \leq i \leq n\}$ are n points in the plane, then the slope of the least squares line through these points is

$$SL(\{(x_i, y_i), 1 \leq i \leq n\}) = \frac{\bar{S}_{xy} - \bar{x}\bar{y}}{\bar{S}_{xx} - \bar{x}^2}$$

where as usual

$$S_{xy} = \sum_{i=1}^{n} x_i y_i, \quad S_{xx} = \sum_{i=1}^{n} x_i^2$$

and "bar" indicates the average. Thus for the Pareto example, if we set

$$x_i = -log(1-\frac{i}{n+1}), \quad y_i = logZ_{n,i},$$

then an estimator of α^{-1} is

$$\widehat{\alpha^{-1}} = \frac{\sum_{i=1}^{n} -log(\frac{i}{n+1})\{nlogZ_{n-i+1,n} - \sum_{j=1}^{n} logZ_{n-j+1,n}\}}{n\sum_{i=1}^{n}(-log(\frac{i}{n+1}))^2 - (\sum_{i=1}^{n} -log(\frac{i}{n+1}))^2} \tag{2.11}$$

which is called the $QQ\, estimator$. The following theorem describes the properties of the QQ-estimator.

Theorem 2.5.1 (Properties of the QQ-estimator (Resnick (2007)))**.** *(**a**) (Weak consistency) Suppose* $k = k(n) \longrightarrow \infty$ *in such a way that as* $n \longrightarrow \infty$ *we have* $k/n \longrightarrow 0$*. Suppose* $Z_1, ...Z_n$ *are a random sample from* F*, a distribution with regularly varying tail satisfying*

$$1 - F(x) \sim x^{-\alpha}L(x), \quad (x \longrightarrow \infty)$$

where L is a slowly varying function. Then the $qq-estimator$ $\widehat{\alpha^{-1}}$ *is weakly consistent for* $1/\alpha$*:*

$$\widehat{\alpha^{-1}} \xrightarrow{P} \alpha^{-1}$$

as $n \longrightarrow \infty$*.*

*(**b**) (Asymptotic normality) If* $k \longrightarrow \infty, k/n \longrightarrow 0, \sqrt{k}A(n/k) \longrightarrow 0$ *then*

$$\sqrt{k}(\widehat{\alpha^{-1}} - \alpha^{-1}) \xrightarrow{d} N(0, 2\alpha^{-2}) \tag{2.12}$$

where A(.) is a positive or negative (not changing sign) function with $\lim_{t\to\infty} A(t) = 0$ and it satisfies the condition of regularly varying function. Asymptotic variance of $\sqrt{k}(\widehat{\alpha^{-1}} - \alpha^{-1})$ is $2\alpha^{-2}$. In contrast, from (2.3) we can write asymptotic variance of $\sqrt{k}(\widehat{\alpha^{-1}} - \alpha^{-1})$ for Hill's estimator is α^{-2}. In section (2.1) we have seen that Hill's estimator can produce considerable bias, if the value of k increases. Thus we can conclude that the QQ-estimator shows bias with the increases of the valus of k.

2.6 Davydov-Paulauskas-Rackauska's estimator

In some cases, we have only the largest value or several largest observations are available for analysis. Particularly, when the data are divided into several blocks but only the largest observations within each block can be used to investigate, for example, in meteorology data, only the highest and lowest temperatures of each data are forcasted. Davydov et al. (2000) proposed a new estimator where they divided observations into several blocks and the estimator of the tail index is constructed from the ratios of the first largest and the second largest terms within the blocks. This Davydov-Paulauskas-Rackauska's estimator is widely known as DPR approach. The steps are as follows. Divide the sample $X_1, ..., X_n$ into k blocks (or groups), $V_1, ..., V_k$, and each block contains $m = m_n = [n/k]$ observations, where $[x]$ denotes the integer part of $x > 0$. To be more specific, $V_i = X_{(i-1)m+1}, ..., X_{im}$ for $i \leq i \leq k_n$. Let $X_{m,1}^{(i)} \geq ..... \geq X_{m,m}^{(i)}$ denote the order statistics of the m observations in the i-th block. Set

$$S_k = \sum_{i=1}^{k} \frac{X_{m,2}^{(i)}}{X_{m,1}^{(i)}}$$

and define

$$\hat{\gamma}_{DPR} = S_k^{-1}(k - S_k)$$

an estimator of γ. Under the condition

$$1 - F(x) = cx^{-1/\gamma} + dx^{-\beta} + o(x^{-\beta}) \quad as \quad x \longrightarrow \infty,$$

where $0 < \gamma^{-1} < \beta \leq \infty$, it is proved that

$$\sqrt{k}(\hat{\gamma}_{DPR} - \gamma) \xrightarrow{d} N(0, \frac{\gamma^2(1+\gamma)^2}{(1+2\gamma)})$$

where $k = o(n^{2(\beta\gamma-1)/(2\beta\gamma-1)})$.

2.7 Maximum likelihood estimator

Smith (1987) proposed a maximum likelihood estimator (MLE) for $\gamma \in \mathbb{R}$. If $X_1, X_2, ..., X_n$ is and i.i.d. random sample from a probability distribution given by the probability density function (PDF) $f(x|\gamma)$, then the MLE of γ maximizes the function

$$L(\gamma) = \prod_{i=1}^{\infty} f(x_i|\gamma)$$

is called likelihood function. The likelihood equations are given by (De Haan and Ferreira (2007)).

$$\sum_{i=0}^{k-1} \frac{1}{\gamma^2} log(1 + \frac{\gamma}{\sigma}(X_{n,n-i} - X_{n,n-k})) - (\frac{1}{\gamma} + 1)\frac{\frac{1}{\sigma}(X_{n,n-i} - X_{n,n-k})}{1 + \frac{\gamma}{\sigma}(X_{n,n-i} - X_{n,n-k})} = 0 \quad (2.13)$$

and

$$\sum_{i=0}^{k-1}(\frac{1}{\gamma} + 1)\frac{\frac{1}{\sigma}(X_{n,n-i} - X_{n,n-k})}{1 + \frac{\gamma}{\sigma}(X_{n,n-i} - X_{n,n-k})} = 0 \quad (2.14)$$

For the case, $\gamma \neq 0$ the maximum likelihood method can be simplified by solving the following two equations:

$$\gamma = k^{-1} \sum_{i=0}^{k-1} log(1 + \frac{\gamma}{\sigma}(x_{n,n-i} - x_{n,n-k})) \quad (2.15)$$

$$\frac{1}{1+\gamma} = k^{-1}\sum_{i=0}^{k-1}\frac{1}{1+\frac{\gamma}{\sigma}(X_{n,n-i}-X_{n,n-k})} \tag{2.16}$$

where the maximum likelihood estimator of σ is a scale parameter. If $k \longrightarrow \infty$, $\frac{k}{n} \longrightarrow 0$ for $n \longrightarrow \infty$, then MLE is consistent and asymptotically normal. Note that the maximum likelihood estimator of γ is shift and scale invariant. Maximizing the likelihood function is not straightforward and numerical methods need to be used. For example in the case of Cauchy distribution, sometimes no solutions are obtained by this method (De Haan and Ferreira (2007)).

2.8 Probability weighted moment estimator

The probability weighted moment estimator (PWM) was first introduced by Hosking and Wallis in 1987 (AghaKouchak and Nasrollahi (2010); De Haan and Ferreira (2007)) as follows:

$$\hat{\gamma} = \frac{\sum_{i=0}^{k-1}(x_{n,n-i}-x_{n,n-k}) - 4\sum_{i=0}^{k-1}\frac{i}{k}(x_{n,n-i}-x_{n,n-k})}{\sum_{i=0}^{k-1}(x_{n,n-i}-x_{n,n-k}) - 2\sum_{i=0}^{k-1}\frac{i}{k}(x_{n,n-i}-x_{n,n-k})} \tag{2.17}$$

If $k \longrightarrow \infty$, $\frac{k}{n} \longrightarrow 0$ for $n \longrightarrow \infty$, then The probability weighted estimator is consistent and asymptotically normal. PWM is also scale and shift invariant.

2.9 De Haan and Resnick's estimator

De Haan and Resnick (1980) proposed an estimator for the index of a stable distribution. Where a stable distribution with parameter $\alpha, \sigma, \beta,$ and μ is given by

$$ln\phi(t) = \begin{cases} -\sigma^{\alpha}|t|^{\alpha}\{1 - i\beta sin(t)tan\frac{\pi\alpha}{2} + i\mu t\} & ,\alpha \neq 1, \\ -\sigma|t|\{1 + i\beta sin(t)\frac{2}{\pi}ln|t| + i\mu t\} & ,\alpha = 1. \end{cases}$$

Suppose $X_1, X_2, ...$ are independent and identically distributed random variables with order statistics of $X_{1,n} \leq ...... \leq X_{(n,n)}$ then

$$\hat{\alpha} = \frac{logk}{logX_{(n,n)} - logX_{(n-k+1,n)}} \tag{2.18}$$

where it is weakly consistent if $k(n) \longrightarrow \infty, k(n)/n \longrightarrow 0, n \longrightarrow \infty$. it is not asymptotically efficient (Basu and Sen (1982)).

In the next chapter we show some real data that follows the heavy-tailed distributions. The estimators that we described above have been applied to these data-sets.

Chapter 3

Applications of Heavy-tailed Distributions

Heavy-tailed distributions continue to attract attention from a substantial number of interesting and important applications. In the previous chapter, we discussed many methods to estimate the tail index. However, different methods need to choose for different real life problems. This chapter depicts various applications of estimators.

3.1 Application to Internet traffic

The applications of heavy-tailed distributions to network traffic have been illustrated in several works. When files are transferred from a web-server, the server sends large number of small files. But the number of very large files remains significant. Define the term ”self-similar” as viewing the same or behaving the same when looked at from different degrees of magnification or different scales on a dimension. We also define the term ”bursty” as transmission of high-bandwidth over a short period. If a self-similar process exhibits remarkable bursts at a wide range of time scales, then all the values at any time show long-range dependence (LRD) (Long-range dependence means statistically significant correlations across large time scales). Thus traffic burstiness follows a heavy-tailed distribution (see Rezaul and Grout (2007)). Almost all modern telecommunication system such as the Ethernet, wireless local area networks (WLAN) or wireless wide area networks (WWAN) exhibit this phenomenon of LRD (see Lee and Fapojuwo (2005); Pacheco (2006)). Heavy-tailed distributions are being used in characterizing the internet traffic more precisely from a number of several sources (for example video, audio, web requests, email, chat, game, etc) show the properties of self-similarity and LRD. A major feature of LRD is that it has high variability of the traffic at mini time scales also present in the wide time scales. Rezaul and Grout (2007) analyzed six traffic traces like EPA, NASA-Jul95, NASA-Aug95, ClarkNet, Saskatchewan and Calgary, all publicly available and each of the traces of sample size 10000. They compared various methods for estimating the tail index of heavy-tailed internet traffic. They came up with the results that moment (DEDh) estimator and qq-plot were not good enough to provide an acceptable measured index because of unstable regions observed in the graph. They concluded that Hill estimator is a good in estimating the index when there are infinite variances (i.e. when $\alpha < 2$). Pacheco (2006) tried to estimate the heavy tail index for WWW traffic. He collected the data of traces from Boston University over the period of January-February 1995 which are publicly available. It was concluded that the QQ estimator is more robust and less volatile than the Hill estimator and there should have more accurate estimates of the tail-index.

3.2 Application to earthquake

Some natural phenomena show heavy-tailed distributions. Predicting the highest sample value for such phenomena is a difficult problem. Moreover the modeling heavy-tailed distributions in geology depends on the number of extreme data. Hill and the moment estimator are widely used in estimating tail index for this kind of problem. Caers et al. (1999) collected data from the Harvard Central Moment Tensor (CMT) catalog of global earthquakes. They considered a listed large earthquake magnitudes above 6.5m on the Richter scale (seismic moments 0.56×10^{19}Nm and above). They estimated the magnitude of a large earthquake with sample size of 715 that happens once in 2000 times using Hill and moment estimators. They investigated that the graph of the moment estimator was stable between the earthquake magnitudes of 7.2 and 7.7 and Hill estimator showed high variability. They identified that finding the value of k was the problem for these estimators. To over come this problem, they chose k above a series of thresholds. These thresholds were selected by computing extreme qunatile, q_ϵ, defined as $P(X > q_\epsilon) = \epsilon$ and in the case of $\gamma > 0$, extreme quantile is given by

$$\hat{q}_{\epsilon,k} = X_{n-k} + H_{k,n} X_{n-k} \frac{(\frac{k+1}{\epsilon n})^{\hat{\gamma}} - 1}{\hat{\gamma}}$$

where $H_{k,n}$ is the Hill's estimator and $\hat{\gamma}$ is the estimates of either Hill or moment estimator, in there case they chose moment estimator. After choosing the value of k above a threshold they re-sampled k data and calculated the MSE, bias, and variance of the estimators. They did this process repeatedly for every threshold and selected the threshold that gave the minimum MSE.

3.3 Applications to extreme flood data

Hydrological events (e.g. floods) cause severe damage. The information about Extreme rainfall statistics is often used to assess when a flood is likely. Rainfall data are always considered as auto-correlated. To overcome this problem of correlation only the largest values from consecutive rainfall events that exceed a high threshold are used for the risk assessment of extreme floods. Typically the data shows long right tails and then it is of interest to estimate the tail index. However, there are no fixed models governing which type of distribution is most suitable for a particular event. AghaKouchak and Nasrollahi (2010) took daily long-term rain gauge data from 6 different places across the world including USA, Australia, France, and Netherlands. In their study, they considered data stretching over a period of 38-71 years. A sample of approximately 14000 to 25000 data points can be chosen from each station. They used these data to Pickands, Hill, Probability weighted moment (PWM) and Maximum likelihood estimator. It was shown that for the small value of k the semi-parametric Pickands estimator approach had large variability while the other estimators showed their consistency. In conclusion, they illustrated that different estimators may give different outcomes in considerable differences in tail index estimators. Anderson and Meerschaert (1998) used a data-set of 72 years of river flow from October 1912 to September 1983 of the Salt River near Roosevelt, Arizona. The average monthly flow rate in cubic feet per second of those years was calculated and the graph of sharp spikes of the data followed the characteristic of heavy-tailed distributions. Although the data has heavy tails, it could not be concluded clearly that the proportion of the data was large enough for estimating the tail index by applying the method of Hill and Maximum likelihood estimators.

3.4 Application to air pollution due to ozone

Chronic and acute exposure to surface ozone can increase the risk of death from cardiovascular and respiratory causes. In addition, high levels of surface ozone damage agriculture. Secondary pollutant ozone is generated through the oxidation of volatile organic carbons, carbon monoxide and methane in the presence of NO_X and sunlight. The NAAQS (National Ambient Air Quality Standard) for ozone was set, in 1997, to a maximum daily 8-h average ozone (MDA8 O_3) of 84ppb (Parts per billion (ppb) is the number of units of mass of a contaminant per 1000 million units of total mass). Phalitnonkiat et al. (2016) characterized ozone extremes over the continental U.S. They analyzed the ozone measured at 25 sites throughout the continental U.S. using data from the CASNET (www.epa.gov/casnet) observational network. They used Maximum likelihood estimator (MLE) and Hill estimator to estimate the tail index. They had to chose proper value of a threshold (that is a value of k) because if the threshold is so big, we will not get robust result. On the other hand, if the threshold is too small the MLE method is supposed to underestimate the shape parameter. However, they used a technique to generate some data as CASNET data. First, the data were used to Hill's estimator. If the estimates were greater than 0.3, they took it as their estimate, otherwise, MLE method was used to the data They generated 500 data-sets as close to the CASNET data by using this technique. Their technique of estimating was then used to measured ozone data whether the data follow long-tailed distribution in some places or light-tailed distribution in some other places.

We discuss more examples of applications of heavy-tailed distributions as presented in De Haan and Ferreira (2007). In these examples we consider $\gamma = 1/\alpha$ as the tail index of heavy-tailed distributions.

3.5 Application to extreme sea levels due to windstorms

The first example deals with sea level data in the Netherlands. About 40% of the Netherlands is below sea level. Storm surges cause the seawater level to go up along the coastal area. The coastal area has to be protected against sea by dikes. The problem consists in determining the height of the dike such that the probability of a flood in a given year is 40^{-4}. More than 100 years of storm data in the town of Delfzijl in the northeast of the Netherlands has been collected. The data corresponds to winter storms during the years 1882-1991. Within this period, 1877 severe wind storms have been recorded and the high-tide water level at Delfzijl forms approximately a set of independent observations. So we can assume that they are independent and identically distributed. Then a $(1-\alpha)100\%$ approximating confidence interval is given by

$$\hat{\gamma} - Z_{\alpha/2}\sqrt{\frac{var_{\hat{\gamma}}}{k}} < \gamma < \hat{\gamma} + Z_{\alpha/2}\sqrt{\frac{var_{\hat{\gamma}}}{k}},$$

where $\mathrm{var}_{\hat{\gamma}}$ is the respective asymptotic variance with γ replaced by its estimate and $Z_{\alpha/2}$ is the $1-\alpha/2$ quantile of the standard normal distribution. They apply Pickands, Moment and Probability weighted moment estimator with 95% asymptotic confidence intervals for some values of k. The confidence intervals of all these three estimators contain the value zero which does not contradict the hypothesis that the extreme value index is zero.

3.6 Application to S&P 500 data

Another example discusses the S&P500 daily price quotes stands for Standard & Poor's 500, and it is an American stock market index of the 500 largest companies chosen for market size, liquidity and industry grouping, among other factors. The data set consists of 5835 daily price quotes of the S&P500 total return index taken over the period from 01/01/1980 to 14/05/2002. Let p_t be the daily price quotes and r_t be the daily continuous compounded returns where r_t is calculated by taking the logarithmic first differences of the price series that is $r_t = log(p_t/p_{t-1})$. Since, due to positive growth of the economy, stock returns generally give a positive mean, we need to be concerned only about the loss returns. Assume that the observations are independent and identically distributed. Making a decision on the risk of a big investment while one wants to avoid a loss larger than a certain amount. It is of interest to know the probability of the occurrence of such a loss. Economists typically suppose that this data can be understood by a heavy-tailed distribution. Suppose F is the distribution function of the log-loss returns and x is a critical (large) amount. The problem is then the estimation of $1 - F(x)$ which leads us to a tail probability estimator under the extreme value theory approach. Hence the underlying function is in the domain of attraction of some G_γ with positive index γ. That implies that the right endpoint of the underlying function is infinite. De Haan and Ferreira (2007) used Hill, Pickands, and Moment estimators to estimate the tail index of loss log-returns. They observed that for small values of k there is large variability in the estimates. The Pickands estimator has the large variance for positive γ, when compared to the others. On the other hand, Hill and moment estimators give positive estimates for γ.

3.7 Application to exchange rate returns

Another example from economic theory where the estimation of tail indices is an important problem is the exchange rate returns. Some analyses are related to the extreme cases. Exchange rate return problems lend themselves to be analyzed through the use of extreme value theory. Hols and De Vries (1991) applied extreme value theory to foreign exchange rate data. They analyzed the data consisting of weekly Thursday quotations of the Canadian-United States dollar spot exchange rate. They take a total of 475 observations starting from the last week of 1973 and ending the fourth week of 1983. The investors focus on the maximum returns. Let $X_i; i = 1, .., n$ be an i.i.d. a sequence of the data-set and $x_p, x_p > 0$ be a high return level. The subscript p of x_p refers to the probability of returning to a high-level exchange rate. If $X_i > x_p$, define $Y_i = X_i - x_p$. The return level x_p corresponding to a given low probability p is estimated by (Hols and De Vries (1991))

$$\hat{x_p} = \frac{(mr/pn)^{\hat{\gamma}} - 1}{1 - 2^{-\hat{\gamma}}}(X_{n-r} - X_{n-2r}) + X_{n-r}$$

where n is the number of observations, m is the time period considered, $r = k/2$, and k is the lowest descending-order statistic $X_{(k)}$ used in computing $\hat{\gamma}$. It turns out that $\hat{\gamma}$ is also a consistent estimator of γ. They used Pickands, Hill and De Haan and Resnick estimators to estimate γ. For example, they observed that a weekly return of x_p over a period of 100 weeks estimates a probability of maximum returns is 0.1 of Hill's estimator while the others two estimators have notable differences.

Other examples of real applications where estimating tail indices is a concern include: McElroy (2016) gives an example of OctExt Ethernet trace known as counting of packets arriving in one-second intervals; Hill and Shneyerov (2013) describe first-price auctions and apply it to British Columbia Timber Sales data; Chernozhukov and Fernández-Val (2011) represent the results with extremely low percentiles of live infants' birthweights in the range between 250 and 1500 grams; Gardes and Girard (2010) illustrate an application to the estimation of return levels of extreme rainfalls. As all of these example, demonstrate, the problem of estimating the tail index of heavy-tailed distributions is of actual importance in many important application.

In the next chapter, a new estimator is proposed and its performance, using simulations, is compared to the estimators already available in the literature.

Chapter 4

New Estimator

In this chapter a new estimator is proposed which is easy to compute and exhibits good operating charcteristic in terms of mean squared error and bias when compare to other estimators. Let X_1,....,X_n be an i.i.d and consider the order statistics as $X_{1,n} \leq ..... \leq X_{n,n}$ of n number of observations, where $X_{1,n}$=min(X_1......X_n) and $X_{n,n}$=max(X_1.......X_n). Define the new estimator as

$$\hat{\alpha} = \frac{ln(n)}{lnX_{n,n}} \tag{4.1}$$

One striking feature of this estimator is that it is based on only one order statistic where estimators such as Hill, Pickands, De Haan and Resnick, Dekkers-Einmahl and De Haan (DedH), and some other estimators that are based on a proper selection of a tunning parameter k, and the use of k of the largest observations in the sample. The value of k, besides heaving to be selected, goes to infinity as $n \to \infty$. Choosing a proper value for k is not always so easy. There are several proposed ways of selecting k. Haan and Peng (1998) compared few estimators like Hill, Pickands and DEdH. For a fixed sequence of $k = k(n) \longrightarrow \infty$ they selected the k value such that the estimators are asymptotically normal as well as minimum asymptotic mean squared error occurred. Gomes and Guillou (2015) discussed the challenges of threshold selection. The estimator proposed by (4.1) uses $k = 1$ and this may suggest that it cannot be competitive in terms of its asymptotic properties. However, as the simulation work will suggest, this estimator does very well when compared to others. By blocking the data into independent blocks, computing (4.1) within each block, and then averaging the estimators thus computed, improves its performance in terms of mean squared error. The estimator (4.1) is not shift nor scale invariant but its behavior is somewhat robust when shifting and rescaling the data.

4.1 Simulation study and results

Simulation studies were carried out for various slowly varying function $h(x)$. Data was generated from distributions with survival functions $\hat{F}(x) = x^{-\alpha}h(x)$ for various values of α and a wide selection of function $h(x)$. To be able to generate a random sample from these distributions, it became necessary to develop code to implement numerical approximations to the quantile functions of such distribution functions. The various estimators of interest were then computed from each sample and, based on 5000 simulations, their biases and MSE were calculated. These results were then used to compare the various estimators in various settings consisting of different slowly varying functions and various values of α. In addition, and to somewhat assuage the potential concern that (4.1) makes use of only the largest observation in the sample, our simulation studies also considered the estimator obtained by averaging the within-blocks estimators of α computed using (4.1). These results are reported and discussed in this chapter. As it turns out these ”block” estimators exhibit good MSE behavior. However, as it can be the case with those estimators that need to worry about choosing "k", a similar issue arises with "block" estimator. Namely, what is the "best" number of blocks? We do not attempt to answer this question here and leave it for future work. The tables below report the standard deviation, Bias, and Mean squared error for the new estimator, Hill, Pickands and De Haan and Resnick, and block estimators. A sample size of 40000 observations has been generated from the survival function, $\bar{F} = \frac{\frac{ln(2+\frac{1}{1+x})}{ln(3)}}{(1+x)^{\alpha}}$ for values of $\alpha = 0.5, 0.75, 1, 2, 3, 5, 7,\ and\, 10$. In the appendix we present different tables for different slowly varying functions. Table 4.1 represents the standard deviation(Sd),

Table 4.1: Standard deviation, Bias and Mean squared error of estimators from survival function of $\bar{F} = \frac{\frac{ln(2+\frac{1}{1+x})}{ln(3)}}{(1+x)^{\alpha}}$ for true value of $\alpha = 0.5$

α =0.5, k=5000, n=40000, 5000 iterations, $\bar{F} = \frac{\frac{ln(2+\frac{1}{1+x})}{ln(3)}}{(1+x)^{\alpha}}$, x>0			
Statistics	Sd	Bias	MSE
Hill's estimator	0.007113	0.002589	5.73E-05
Pickands estimator	0.010439	0.019172	0.000477
New estimator	0.055703	0.000778	0.003103
De Haan & Resnick	0.062178	-0.02314	0.004401
n=40,1000blks	0.006029	-0.0181	0.000364
n=100,400blks	0.007312	-0.00401	6.95E-05
n=200, 200blks	0.009012	0.001137	8.25E-05
n=500, 80blks	0.012021	0.003775	0.000159
n=1000, 40blks	0.014873	0.005243	0.000249
n=2000, 20blks	0.019107	0.005458	0.000395
n=4000, 10blks	0.024639	0.00566	0.000639
n=5000, 8blks	0.027273	0.00612	0.000781

bias, and Mean squared error (MSE) of estimates of true value $\alpha = 0.5$ for different estimators. This table reports that Hill's estimator produces the smallest sd and MSE when compared with Pickands, De Haan and Resnick, and new estimators. The block estimators also provide smaller Sd, bias, and MSE compare to others. The block of size 400 with the sample size of n = 100 per block gives an MSE of

6.95E-05 which is very close to the MSE of Hill's estimator. However, the new estimator gives the smallest bias among all the biases in this table. Figure 4.1 shows the histograms of estimates of $\alpha = 0.5$ of all different estimators. Because of having the smallest bias, the histogram of new estimator is more consistent and symmetric with value 0.5 than the others. This figure also depicts the negative biases for the block estimators where using blocks of size 40 and 100, and De Haan and Resnick estimators. Table 4.2 reports the standard deviation(Sd), bias, and Mean squared

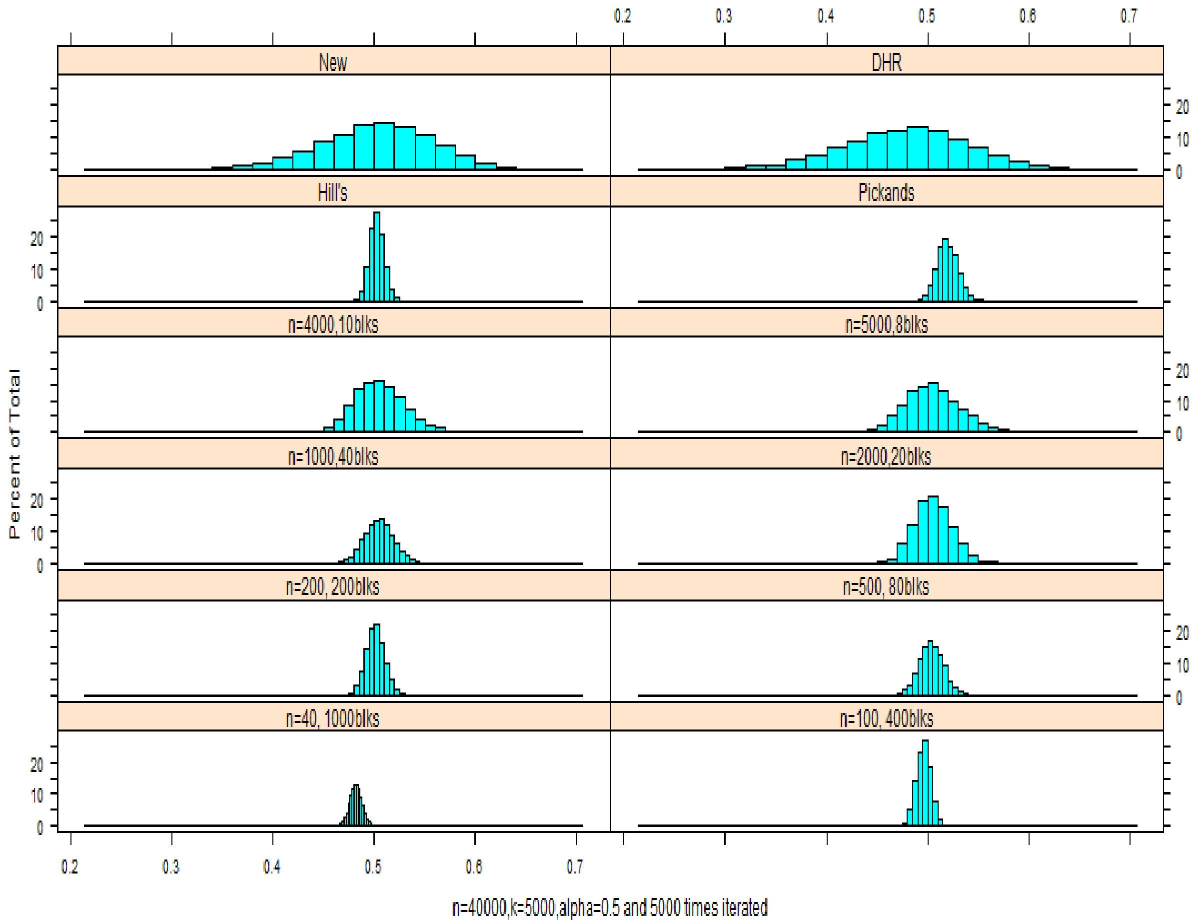

Figure 4.1: Data generated from $\bar{F} = \frac{\frac{ln(2+\frac{1}{1+x})}{ln(3)}}{(1+x)^{\alpha}}$, $x > 0$; Histograms of estimators for $\alpha = 0.5$

error (MSE) of estimates of true value $\alpha = 0.75$ for different estimators. This table also shows that Hill's estimator gives the smallest Sd and MSE when compared to Pickands, De Haan and Resnick, and new estimators. Block estimators also provide smaller Sd, bias and MSE when compared to others. The block of size 1000 with the sample size of n = 40 per block gives the smallest Sd of 0.008521 and an MSE of 0.000292, very close to smallest MSE of estimates of Hill's estimator. However, the new estimator presents the smallest bias of 0.002759 among all the biases in this table. Figure 4.2 shows the histograms of estimates of $\alpha = 0.75$ for all different estimators. Because of having the smallest bias, the histogram of the new estimator is more consistent and symmetric about the 0.75 than the others. This figure 4.2 also exhibits the negative biases in histograms for the Hill and De Haan and Resnick estimators. In addition, the histogram of estimates of Hill's estimator seems more compact.

Table 4.2: Standard deviation, Bias and Mean squared error of estimators from survival function of $\bar{F} = \frac{\frac{ln(2+\frac{1}{1+x})}{ln(3)}}{(1+x)^{\alpha}}$ for true value of $\alpha = 0.75$

α =0.75, k=5000, n=40000, 5000 iterations, $\bar{F} = \frac{\frac{ln(2+\frac{1}{1+x})}{ln(3)}}{(1+x)^{\alpha}}$, x>0			
Statistics	Sd	Bias	MSE
Hill's estimator	0.010601	-0.00577	0.000146
Pickands est.	0.020396	0.04105	0.002101
New estimator	0.084171	0.002759	0.007091
De Haan & Resnick	0.093569	-0.03445	0.00994
n=40,1000blks	0.008521	0.014823	0.000292
n=100,400blks	0.010543	0.015223	0.000343
n=200, 200blks	0.012757	0.014324	0.000368
n=500, 80blks	0.017258	0.013646	0.000484
n=1000, 40blks	0.022162	0.012511	0.000648
n=2000, 20blks	0.028024	0.011411	0.000915
n=4000, 10blks	0.037148	0.01037	0.001487
n=5000, 8blks	0.03972	0.009676	0.001671

Table 4.3 demonstrates the standard deviation(Sd), bias, and Mean squared error

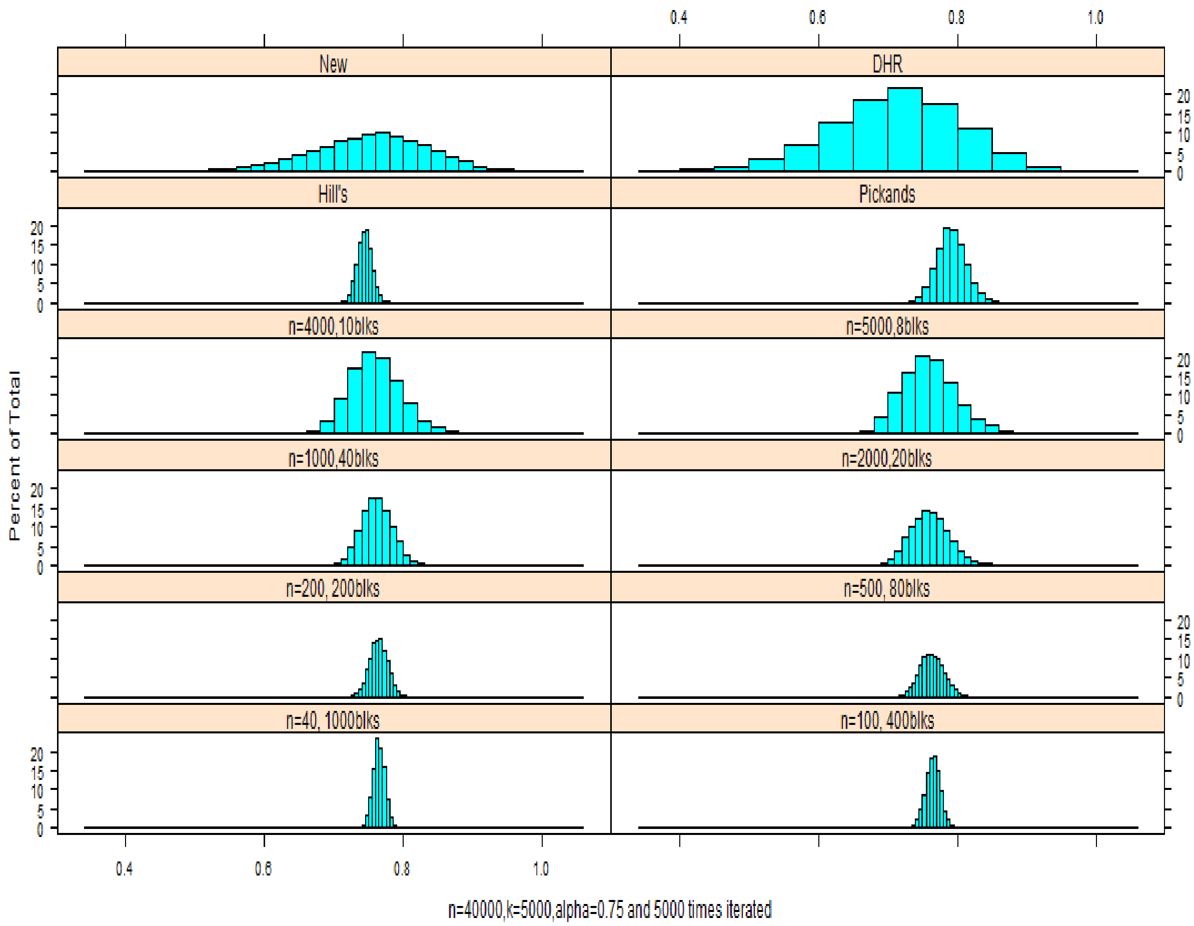

Figure 4.2: Data generated from $\bar{F} = \frac{\frac{ln(2+\frac{1}{1+x})}{ln(3)}}{(1+x)^{\alpha}}$, $x > 0$; Histograms of estimators for $\alpha = 0.75$

Table 4.3: Standard deviation, Bias and Mean squared error of estimators from survival function of $\bar{F} = \frac{\frac{ln(2+\frac{1}{1+x})}{ln(3)}}{(1+x)^{\alpha}}$ for true value of $\alpha = 1$

α =1, k=5000, n=40000, 5000 iterations, $\bar{F} = \frac{\frac{ln(2+\frac{1}{1+x})}{ln(3)}}{(1+x)^{\alpha}}$, x>0			
Statistics	Sd	Bias	MSE
Hill's estimator	0.013138	-0.04086	0.001842
Pickands est.	0.034381	0.06115	0.004921
New estimator	0.110314	0.003318	0.012178
De Haan & Resnick	0.120952	-0.05364	0.017504
n=40,1000blks	0.011046	0.031437	0.00111
n=100,400blks	0.013899	0.025313	0.000834
n=200, 200blks	0.017199	0.0217	0.000767
n=500, 80blks	0.023255	0.018583	0.000886
n=1000, 40blks	0.029559	0.015996	0.001129
n=2000, 20blks	0.037682	0.015476	0.001659
n=4000, 10blks	0.049095	0.014037	0.002607
n=5000, 8blks	0.054032	0.013938	0.003113

(MSE) of estimates of true value $\alpha = 1$ for different estimators. This table also shows that Hill's estimator gives the smaller Sd and MSE when compared with Pickands, De Haan and Resnick, and new estimators. Block estimators also give smaller Sd, bias and MSE compared to others. The block of size 40 with the sample size of n = 1000 per block and block size of 200 with the sample of n= 200 per block give the smallest Sd of 0.011046 and an MSE of 0.000767 respectively compared to others. However, the new estimator illustrates the smallest bias of 0.003318 among all the biases in this table. Figure 4.3 shows the histograms of the estimates for $\alpha = 1$ for all different estimators. Because of having the smallest bias, the histogram of the new estimator is more consistent and symmetric about 1 than the others. This figure 4.3 also depicts the negative biases in histograms for the Hill and De Haan and Resnick estimators. In addition, the histograms of estimates for the block estimators with block sizes of 40, 100, and 200 seem more compact about 1.

Table 4.4 reports the standard deviation(Sd), bias, and Mean squared error (MSE)

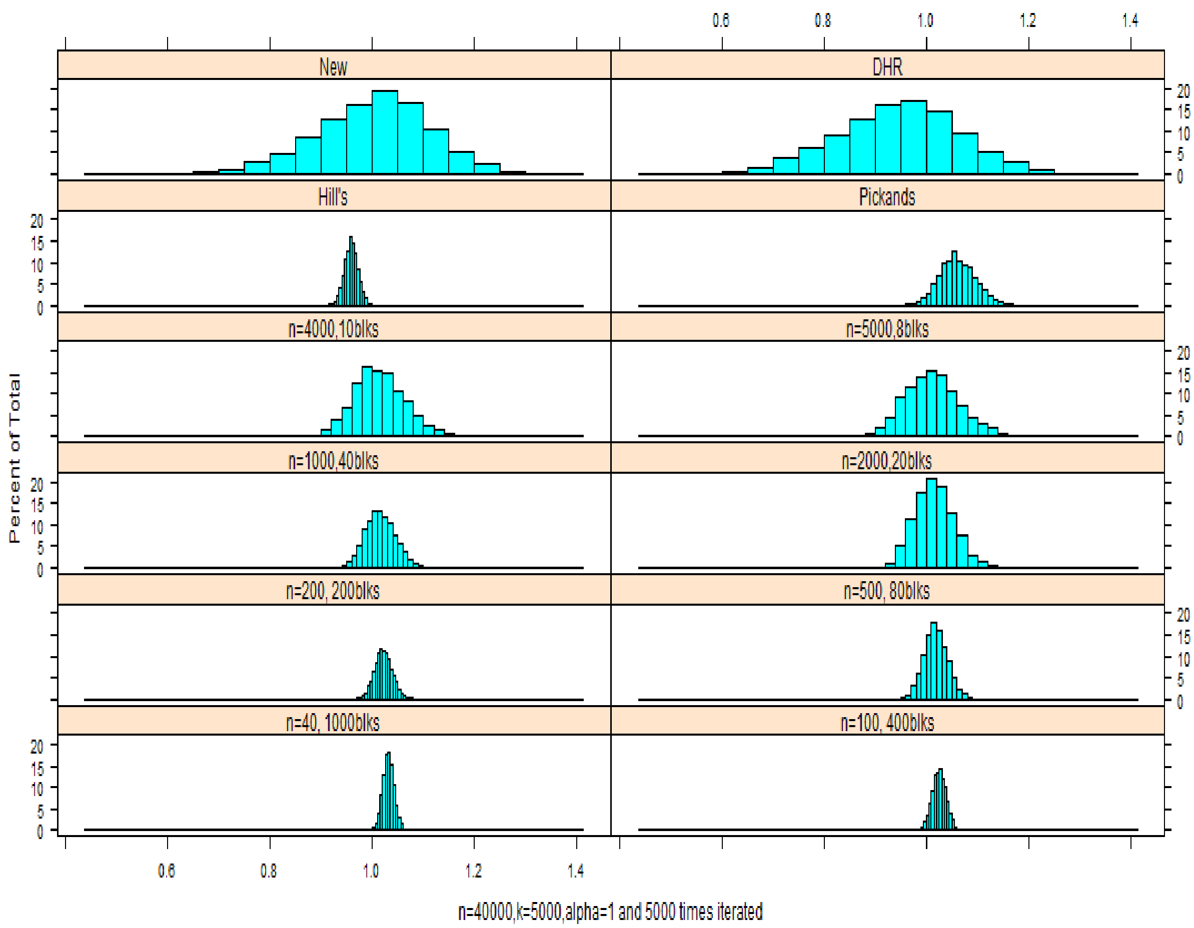

Figure 4.3: Data generated from $\bar{F} = \frac{\frac{ln(2+\frac{1}{1+x})}{ln(3)}}{(1+x)^{\alpha}}$, $x > 0$; Histograms of estimators for $\alpha = 1$

Table 4.4: Standard deviation, Bias and Mean squared error of estimators from survival function of $\bar{F} = \frac{\frac{ln(2+\frac{1}{1+x})}{ln(3)}}{(1+x)^{\alpha}}$ for true value of $\alpha = 2$

α =2, k=5000, n=40000, 5000 iterations, $\bar{F} = \frac{\frac{ln(2+\frac{1}{1+x})}{ln(3)}}{(1+x)^{\alpha}}$, x>0			
Statistics	Sd	Bias	MSE
Hill's estimator	0.020396	-0.45767	0.209874
Pickands est.	0.126099	0.113702	0.028826
New estimator	0.222226	0.003731	0.049388
De Haan & Resnick	0.210319	-0.24825	0.105852
n=40,1000blks	0.021332	0.070638	0.005445
n=100,400blks	0.027574	0.053193	0.00359
n=200, 200blks	0.034152	0.044891	0.003181
n=500, 80blks	0.045916	0.037892	0.003544
n=1000, 40blks	0.058849	0.032965	0.004549
n=2000, 20blks	0.075107	0.031264	0.006617
n=4000, 10blks	0.098603	0.027363	0.010469
n=5000, 8blks	0.104399	0.025454	0.011545

of estimates of true value $\alpha = 2$ for different estimators. This table represents that Pickands estimator gives the smallest MSE when compared to Hill, De Haan and Resnick, and the new estimators. Block estimators also provide smaller Sd, bias and MSE when compared to others. The block of size 200 with the sample size of n = 200 per block gives the smallest MSE of 0.003181. However, the new estimator presents the smallest bias of 0.003731 among all the biases in this table. Figure 4.4 shows the histograms of estimates of $\alpha = 2$ for all the estimators. Because of having the smallest bias, the histogram of new estimator is more consistent and symmetric about 2 than the others. This figure 4.4 also exhibits the negative biases in histograms for the Hill and De Haan and Resnick estimators. In addition, the histogram of estimates of Hill's estimator seems more compact and far away from the value 2.

Table 4.5 demonstrates the standard deviation(Sd), bias, and Mean squared error (MSE) of estimates of true value $\alpha = 3$ for different estimators. This table also

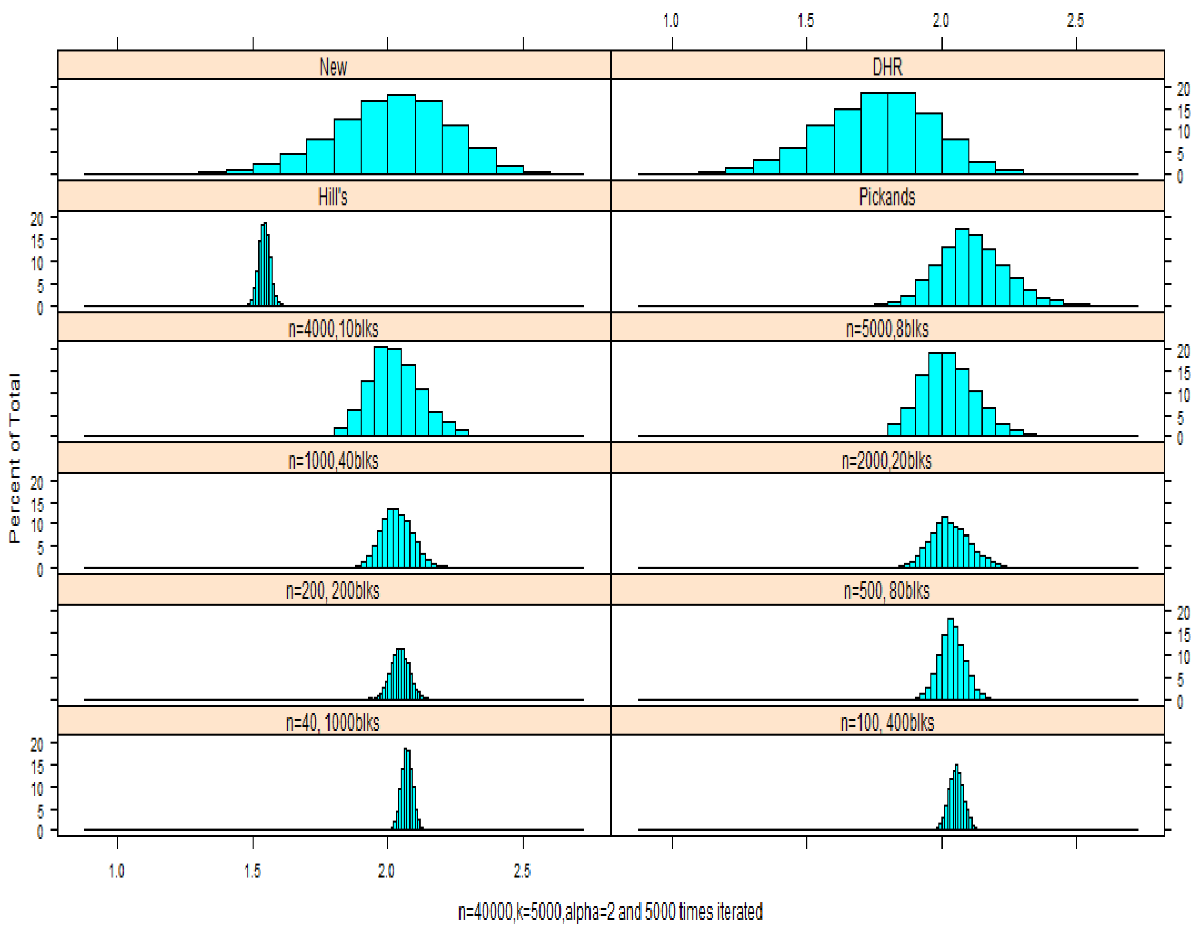

Figure 4.4: Data generated from $\bar{F} = \frac{\frac{ln(2+\frac{1}{1+x})}{ln(3)}}{(1+x)^{\alpha}}$, $x > 0$; Histograms of estimators for $\alpha = 2$

Table 4.5: Standard deviation, Bias and Mean squared error of estimators from survival function of $\bar{F} = \frac{\frac{ln(2+\frac{1}{1+x})}{ln(3)}}{(1+x)^{\alpha}}$ for true value of $\alpha = 3$

α =3, k=5000, n=40000, 5000 iterations, $\bar{F} = \frac{\frac{ln(2+\frac{1}{1+x})}{ln(3)}}{(1+x)^{\alpha}}$, x>0			
Statistics	Sd	Bias	MSE
Hill's estimator	0.02367	-1.13965	1.299361
Pickands est.	0.272029	0.153411	0.09752
New estimator	0.348411	0.025289	0.122005
De Haan & Resnick	0.262863	-0.64734	0.48813
n=40,1000blks	0.032448	0.104566	0.011987
n=100,400blks	0.041915	0.079456	0.00807
n=200, 200blks	0.052019	0.067761	0.007297
n=500, 80blks	0.070413	0.059512	0.008499
n=1000, 40blks	0.087785	0.052399	0.01045
n=2000, 20blks	0.114869	0.048307	0.015526
n=4000, 10blks	0.147367	0.037942	0.023152
n=5000, 8blks	0.156951	0.035241	0.025871

provides that Pickands estimator gives the smaller MSE when compared with Hill, De Haan and Resnick, and new estimators. Block estimators also give smaller Sd, bias and MSE when compared to others. The block size of 200 with the sample of n= 200 per block gives the smallest MSE of 0.007297 when compared to others. However, the new estimator illustrates the smallest bias of 0.025289 among all the biases in this table. Figure 4.5 shows the histograms of estimates of $\alpha = 3$ of all different estimators. Because of having the smallest bias, the histogram of the new estimator is more consistent and symmetric about the 3 than the others. This figure 4.5 also depicts the negative biases in histograms for the Hill and De Haan and Resnick estimators. In addition, the histograms of estimates of the block estimator of sizes 40, 100, 200, and 500 seem more compact about 3 and the histogram of estimates of Hill's estimator seems more compact and far away from the value 3.

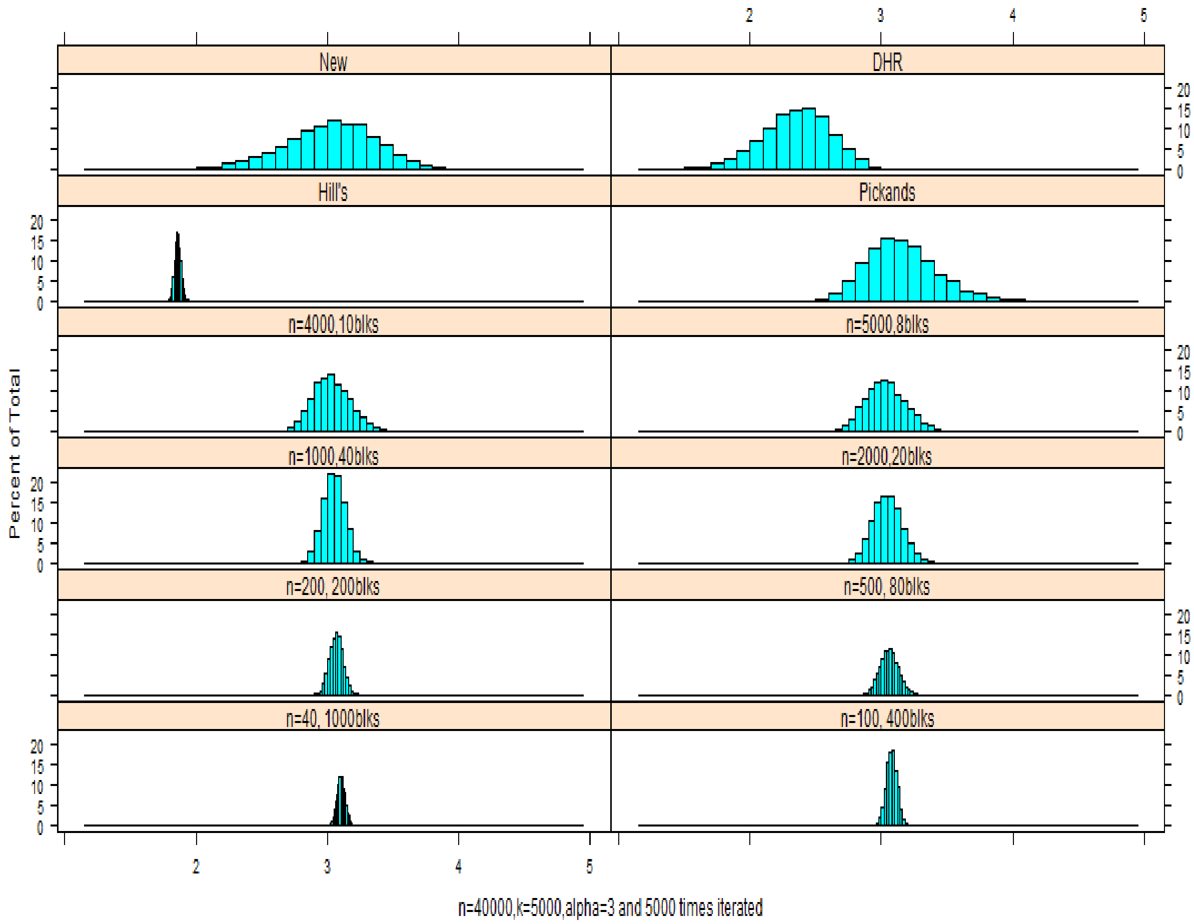

Figure 4.5: Data generated from $\bar{F} = \frac{\frac{ln(2+\frac{1}{1+x})}{ln(3)}}{(1+x)^{\alpha}}$, $x > 0$; Histograms of estimators for $\alpha = 3$

Table 4.6 reports the standard deviation(Sd), bias, and Mean squared error (MSE) of estimates of true value $\alpha = 5$ for different estimators. This table shows that the new estimator gives the smallest MSE when compared with Hill, De Haan and Resnick, and Pickands estimators. Block estimators also provide smaller Sd, bias and MSE compare to others. The block estimator of size 200 with the sample size of n = 200 per block gives the smallest MSE of 0.019744 and block estimator of size 5000 with the sample size of n = 8 per block presents the smallest bias of 0.067669 among all the biases in this table. Figure 4.6 shows the histograms of estimates of

Table 4.6: Standard deviation, Bias and Mean squared error of estimators from survival function of $\bar{F} = \frac{\frac{ln(2+\frac{1}{1+x})}{ln(3)}}{(1+x)^{\alpha}}$ for true value of $\alpha = 5$

α =5, k=5000, n=40000, 5000 iterations, $\bar{F} = \frac{\frac{ln(2+\frac{1}{1+x})}{ln(3)}}{(1+x)^{\alpha}}$, x>0			
Statistics	Sd	Bias	MSE
Hill's estimator	0.027067	-2.81462	7.922835
Pickands est.	0.742035	0.235809	0.606112
New estimator	0.691814	0.315214	0.577871
De Haan & Resnick	0.302054	-1.88153	3.631368
n=40,1000blks	0.054503	0.174551	0.033438
n=100,400blks	0.068549	0.132766	0.022325
n=200, 200blks	0.085114	0.11181	0.019744
n=500, 80blks	0.116197	0.093426	0.022227
n=1000, 40blks	0.146292	0.085168	0.028651
n=2000, 20blks	0.190137	0.077852	0.042206
n=4000, 10blks	0.242265	0.070843	0.063699
n=5000, 8blks	0.267464	0.067669	0.076101

$\alpha = 5$ of all different estimators. Because of having the smallest bias, the histogram of block estimator of size 5000 is more consistent and symmetric about the 5 than the others. This figure 4.6 also exhibits the negative biases in histograms for the Hill and De Haan and Resnick estimators. In addition, the histogram of estimates of Hill's estimator seems more compact and far away from the value 5.

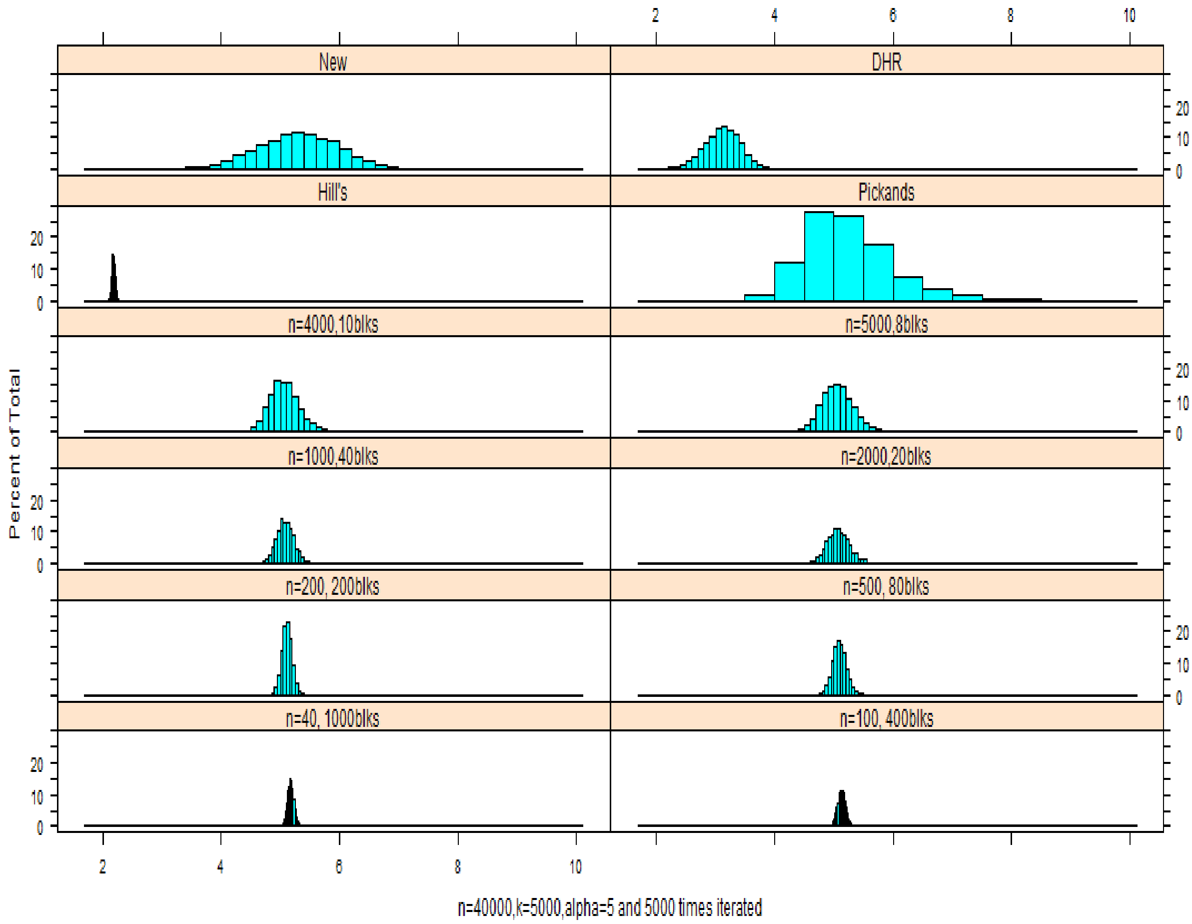

Figure 4.6: Data generated from $\bar{F} = \frac{\frac{ln(2+\frac{1}{1+x})}{ln(3)}}{(1+x)^{\alpha}}$, $x > 0$; Histograms of estimators for $\alpha = 5$

Table 4.7: Standard deviation, Bias and Mean squared error of estimators from survival function of $\bar{F} = \frac{\frac{ln(2+\frac{1}{1+x})}{ln(3)}}{(1+x)^{\alpha}}$ for true value of $\alpha = 7$

α =7, k=5000, n=40000, 5000 iterations, $\bar{F} = \frac{\frac{ln(2+\frac{1}{1+x})}{ln(3)}}{(1+x)^{\alpha}}$, x>0			
Statistics	Sd	Bias	MSE
Hill's estimator	0.029157	-4.64994	21.62282
Pickands est.	1.523826	0.373497	2.461082
New estimator	1.338066	1.311522	3.510151
De Haan & Resnick	0.312749	-3.44856	11.99033
n=40,1000blks	0.077657	0.246041	0.066566
n=100,400blks	0.095979	0.188146	0.044609
n=200, 200blks	0.118683	0.157819	0.03899
n=500, 80blks	0.163394	0.134822	0.044869
n=1000, 40blks	0.2074	0.116975	0.056689
n=2000, 20blks	0.261257	0.104575	0.079178
n=4000, 10blks	0.346386	0.104312	0.130841
n=5000, 8blks	0.374323	0.109244	0.152024

Table 4.7 shows the standard deviation(Sd), bias, and Mean squared error (MSE) of estimates of true value $\alpha = 7$ for different estimators. This table also shows that Pickands estimator gives the smallest MSE when compared with Hill, De Haan and Resnick, and new estimators. Block estimators also give smaller Sd, bias and MSE compare to others. The block size of 200 with the sample of n= 200 per block gives the smallest MSE of 0.03899 compared to other estimators. Moreover, the block estimator with block of size 4000 with the sample size of n = 10 per block presents the smallest bias of 0.104312 among all the biases in this table. Figure 4.7 shows the histograms of estimates of $\alpha = 7$ of all different estimators. The histogram of the new estimator is more symmetric about the 7 than the others. This figure 4.7 also depicts the negative biases in histograms for the Hill and De Haan and Resnick estimators. In addition, the histogram of estimates of Hill's estimator seems more compact and far away from the value 7.

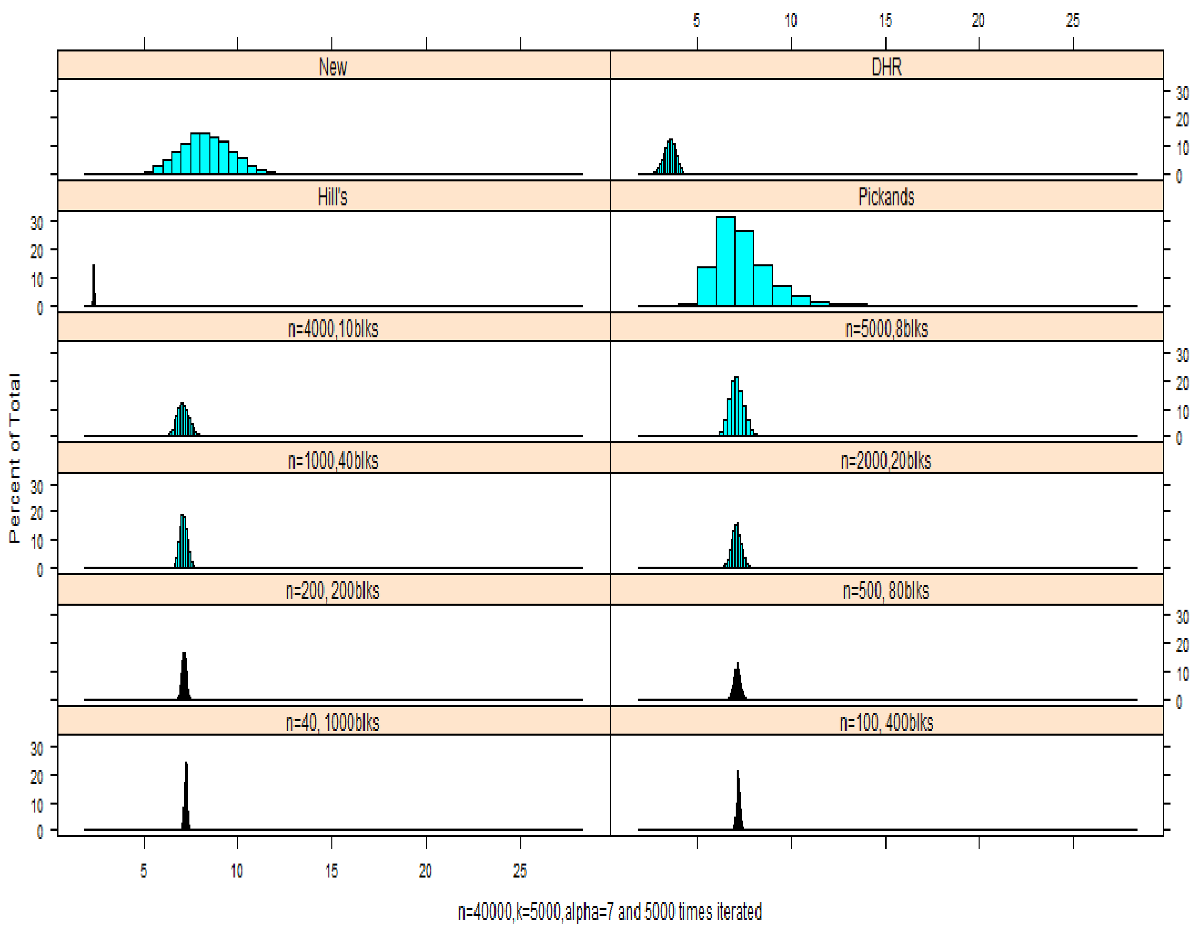

Figure 4.7: Data generated from $\bar{F} = \frac{\frac{ln(2+\frac{1}{1+x})}{ln(3)}}{(1+x)^{\alpha}}$, $x > 0$; Histograms of estimators for $\alpha = 7$

Table 4.8: Standard deviation, Bias and Mean squared error of estimators from survival function of $\bar{F} = \frac{\frac{ln(2+\frac{1}{1+x})}{ln(3)}}{(1+x)^{\alpha}}$ for true value of $\alpha = 10$

α =10, k=5000, n=40000, 5000 iterations, $\bar{F} = \frac{\frac{ln(2+\frac{1}{1+x})}{ln(3)}}{(1+x)^{\alpha}}$, x>0			
Statistics	Sd	Bias	MSE
Hill's estimator	0.03081	-7.51323	56.44964
Pickands est.	4.625555	1.20609	22.84613
New estimator	4.488374	6.808022	66.49064
De Haan & Resnick	0.320895	-6.05393	36.75299
n=40,1000blks	0.108546	0.349651	0.134036
n=100,400blks	0.139193	0.262576	0.088317
n=200, 200blks	0.171123	0.220592	0.077938
n=500, 80blks	0.229732	0.18872	0.088382
n=1000, 40blks	0.289404	0.172265	0.113413
n=2000, 20blks	0.382975	0.153199	0.170111
n=4000, 10blks	0.487449	0.144817	0.258531
n=5000, 8blks	0.527803	0.131782	0.295887

Table 4.8 reports the standard deviation(Sd), bias, and Mean squared error (MSE) of estimates of true value $\alpha = 10$ for different estimators. This table represents that block estimators provide smaller Sd, bias and MSE compare to others. The block estimator with block size of 200 with the sample size of n = 200 per block gives the smallest MSE of 0.077938 and block estimator with block size of 5000 with the sample size of n = 8 per block presents the smallest bias of 0.131782 among all the biases in this table. Examining the standard deviation, bias, mean squared error and histograms in the above tables and figures, the new estimator has always smallest- or very close to smallest bias for any α. It has the smallest MSE or very close to having the smallest MSE. Blocks estimator produces always smallest MSE. we add more tables for different slowly varying functions in the appendix. In the next chapter we apply our new estimator to real life data and show how good it behaves.

Chapter 5

Illustrating the New Estimator

This chapter considers several examples of real life data and illustrate the proposed estimators.

5.1 Application to S&P 500 data

Recall the S&P 500 total returns used by De Haan and Ferreira (2007) discussed in chapter 2. They paid particular attention to loss-returns from 01/01/1980 to 14/05/2002. We use the same data obtained from the website http://home.isa.utl.pt/~anafh/logretS&P.txt. Table 5.1 is a short summary of the data-set.

Table 5.1: Summary of S&P 500 data from 01/01/1980 to 14/05/2002

$X_{1,n}$	1st Quantile	Median	Mean	3rd Quantile	$X_{n-1,n}$	$X_{n,n}$
0.0000048	0.002178	0.004972	0.007096	0.009478	0.0863	0.2282

The various estimators for $\gamma = 1/\alpha$, including Hill, Pickands, moment (Deckers-Einmahl-De Haan or DEdH) and new estimators, are computed and discussed. For the new estimator, we get 0.1875015 and it is shown in figure 5.1. The diagram of

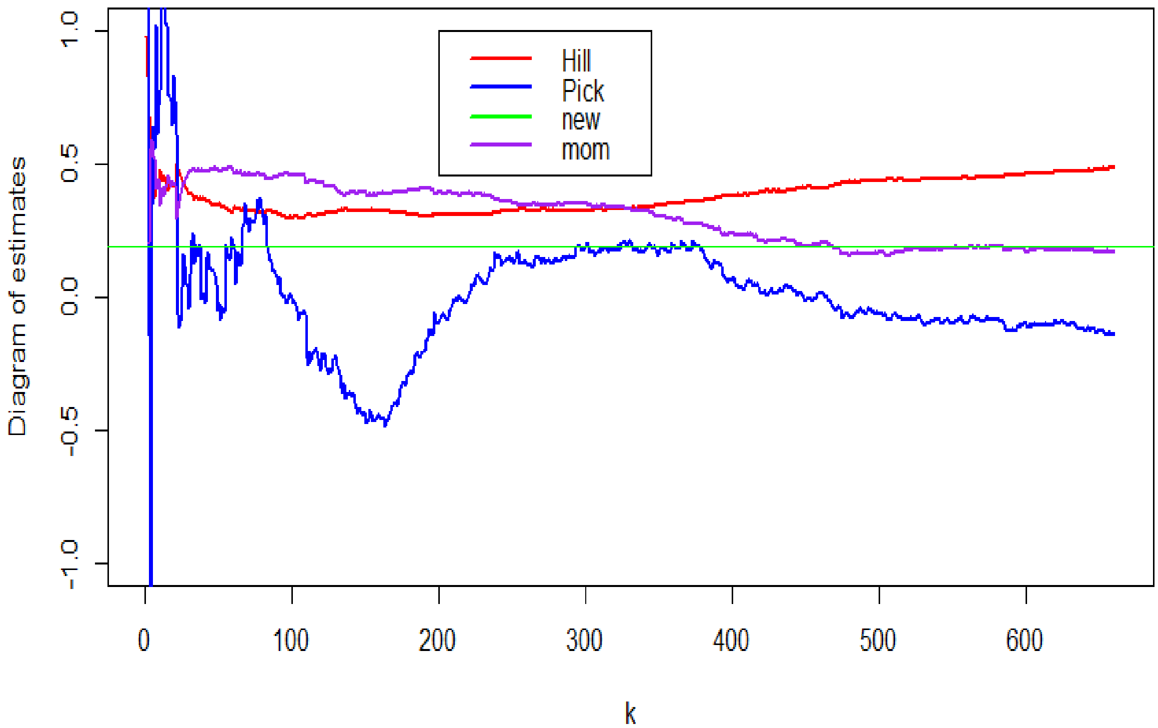

Figure 5.1: S&P 500 data from 01/01/1980 to 14/05/2002; diagram of estimates of γ for various values of k

estimates shows that Hill, Pickands, and moment estimators are highly variable for small values of k. Pickands estimator does not have positive values always. Hill and moment estimators depict almost same value around $k = 300$. However, the Pickands estimator has high variability for all values of k. Table 5.2 reports estimates of different estimators for different values of k. Hill's estimator shows some variability for the smaller value of k, for example, when $k = 25$ it produces 0.442504 while 0.297989 if $k = 100$. The moment and the new estimators give approximately same value after $k = 500$ and for other larger values of k.

In addition we take the data from 01/03/1950 to 07/08/2016 using the website

Table 5.2: S&P 500 data from 01/01/1980 to 14/05/2002: estimates of γ for different values of k

Estimators	$k = 25$	$k = 50$	$k = 100$	$k = 500$	$k = 600$
Hill's	0.4425044	0.3492949	0.2979289	0.4343314	0.4634224
Pickands	0.03423893	-0.02098013	-0.003257236	-0.06545672	-0.1013936
Moment	0.4074895	0.4684189	0.458767	0.1674885	0.1705947
New	0.1875015	0.1875015	0.1875015	0.1875015	0.1875015

https://finance.yahoo.com/q/hp?s=%5EGSPC+Historical+Prices. There is 16737 daily prices and for loss log-returns we compromise with 8853 observations. Table 5.3 is the short summary of this data-set.

Table 5.3: Summary of S&P 500 data from 01/03/1950 to 07/08/2016

$X_{1,n}$	1st Quantile	Median	Mean	3rd Quantile	$X_{n-1,n}$	$X_{n,n}$
5.140e-06	2.109e-03	4.638e-03	6.481e-03	8.618e-03	0.1024	.1096

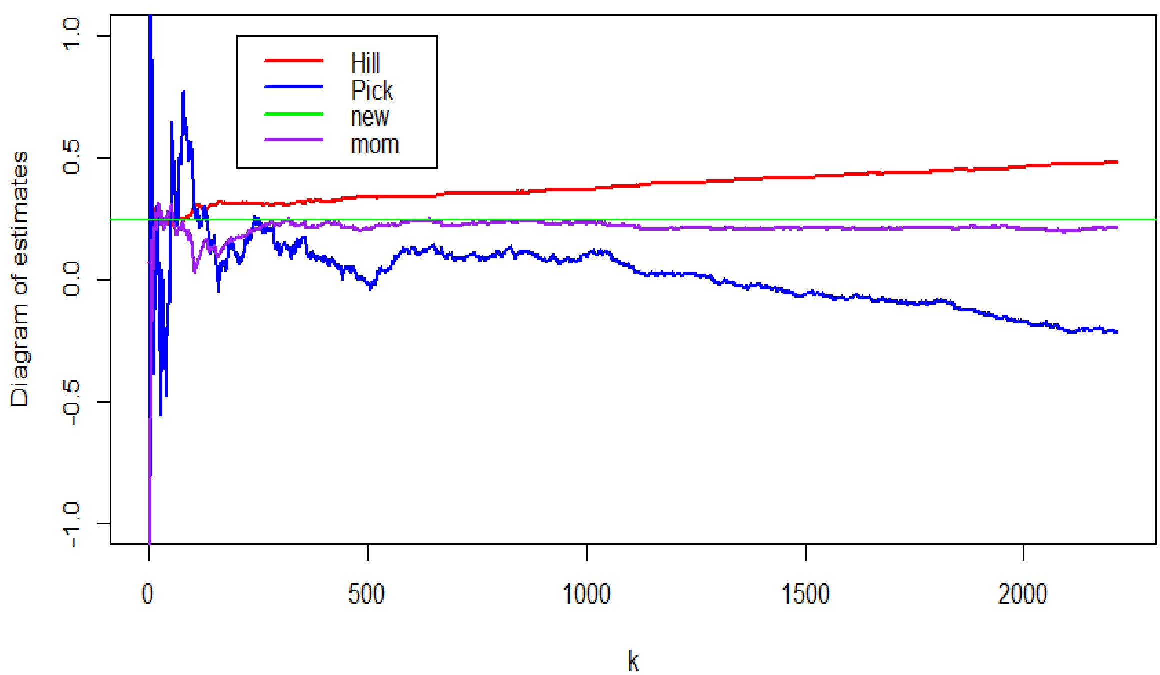

Figure 5.2: S&P 500 data from 01/03/1950 to 07/08/2016; diagram of estimates of γ

The various estimator illustrated this data-set. Figure 5.2 shows that Hill, Pickands,

Table 5.4: S&P 500 data from 01/03/1950 to 07/08/2016; estimates for different k

Estimators	$k = 50$	$k = 75$	$k = 100$	$k = 500$
Hill	0.2213214	0.2466857	0.2703108	0.338164
Pickands	0.1334873	0.4709856	0.4841695	-0.01410704
Moment	0.2941624	0.2138563	0.1487995	0.2027242
New	0.2432933	0.2432933	0.2432933	0.2432933

and moment estimators behave erratically for the various values of k and for small values of k, we observe large variability. For larger values of k, Pickands estimator seems to deviate more and more from the other estimators. After certain values of k, the moment estimator seems more stable than other estimators and new estimator give very close values to those of the moment estimator. Table 5.4 illustrates of estimates for different values of k. We observe negative estimates from the Pickands estimator at $k = 500$ and so on.

5.2 Application to Danish Data

This data-set consists of fire insurance claims over a specific time period. The observations are assumed to be i.i.d. and it is of interest to know the next year's premium volume needed to cover the future losses with very high probability. Embrechts et al. (2013) considered the Danish fire insurance claims data in millions of Danish kroner (1985 prices). This data-set is now available in "$R - Packages$". There is a total of n = 2492 observations. Table 5.5 provides a short summary of the data-set.

At $k = 500$ all the values of γ are very close to each other. Table 5.6 represents the estimates of γ for different values of k. The estimates of Hill, moment and

Table 5.5: Danish fire insurance claims data

$X_{1,n}$	1st Quantile	Median	Mean	3rd Quantile	$X_{n-1,n}$	$X_{n,n}$
0.3134	1.1570	1.6340	3.0630	2.6450	152.4132	263.3000

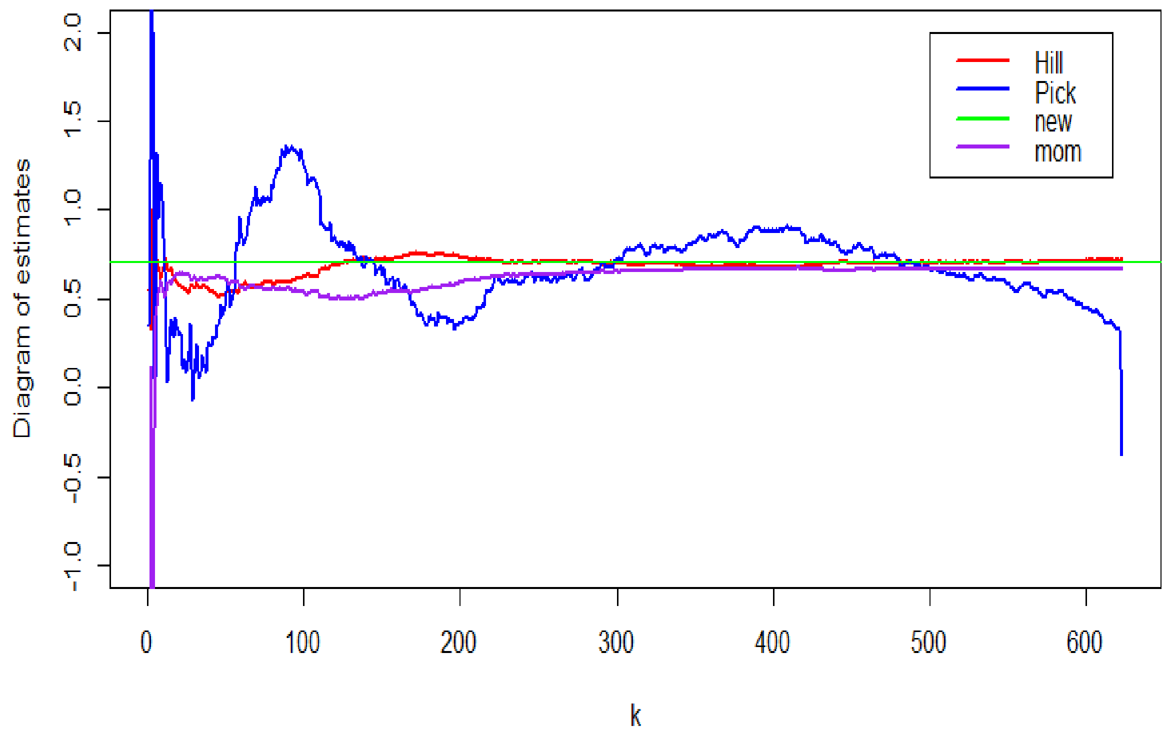

Figure 5.3: Danish Data insurance claims caused by fire; diagram of estimates of γ

Table 5.6: Estimates of γ for different values of k

Estimators	$k = 100$	$k = 200$	$k = 350$	$k = 500$	$k = 623$
Hill's	0.6246393	0.734206	0.685226	0.7038363	0.7183668
Pickand's	1.256662	0.3691794	0.8164744	0.6645386	-0.3788121
Moment	0.537924	0.5945406	0.6664806	0.6654947	0.6670751
New	0.7125967	0.7125967	0.7125967	0.7125967	0.7125967

new estimators give smaller differences $k = 100, 200, 350, 500$, and 623 whether the Pickands estimator shows notable differences for various values of k. Figure 5.3 indicates that estimates of Hill and moment estimators are always very close to our estimator. But Pickands estimator exhibits high variability for different values of k. Resnick (2007) also used this danish data in his book. He found $\gamma = 0.71$ for different estimators except Pickands estimator. Our estimate, using (4.1), is 0.7125967.

Chapter 6

Conclusions

The problem of estimating the tail index of a heavy-tailed distribution has attracted a lot of attention in the last 30 years. In this work, a new simple to compute estimator has been proposed. The estimator avoids the issue of having to select the number of order statistics to be used for estimation by considering only the behavior of the largest order statistic in sharp contrast to other estimators that depend on finding the "right" k when k is the number of order statistics. The behavior of the other estimator can be to be highly sensitive to the choice of k. Almost every estimator including Hill and Pickands need intermediate upper order statistics, $k(n)$. Failure of choosing a perfect value of k can produce large variability in the estimates. To compare with other estimators, we introduced several slowly varying functions $h(.)$ and generated data from the survival function $\hat{F} = x^{-\alpha}h(x)$ for various values of α. We then use these data to compute the various estimators and then analyzed. Furthermore, using the new estimator we developed some block estimators and compared with other estimators. These block estimators showed excellent estimates for every $\alpha > 0$.

Bibliography

Aban, I. B. and Meerschaert, M. M. (2001). Shifted hill's estimator for heavy tails. *Communications in Statistics-Simulation and Computation*, 30(4):949--962.

AghaKouchak, A. and Nasrollahi, N. (2010). Semi-parametric and parametric inference of extreme value models for rainfall data. *Water resources management*, 24(6):1229--1249.

Anderson, P. L. and Meerschaert, M. M. (1998). Modeling river flows with heavy tails. *Water Resources Research*, 34(9):2271--2280.

Basu, S. K. and Sen, P. K. (1982). *Asymptotically efficient estimator for the index of a stable distribution*. Department of Biostatistics, University of North Carolina at Chapel Hill.

Beirlant, J., Vynckier, P., and Teugels, J. L. (1996). Tail index estimation, Pareto quantile plots regression diagnostics. *Journal of the American Statistical Association*, 91(436):1659--1667.

Bianchi, K. and Meerschaert, M. (2000). Scale and shift invariant estimators for the heavy tail index alpha. *AMS 2000 Subject Classification: 62F12, 60F05*.

Caers, J., Beirlant, J., and Maes, M. A. (1999). Statistics for modeling heavy tailed distributions in geology: Part i. methodology. *Mathematical geology*, 31(4):391--410.

Chernozhukov, V. and Fernández-Val, I. (2011). Inference for extremal conditional quantile models, with an application to market and birthweight risks. *The Review of Economic Studies*, 78(2):559--589.

Davydov, Y., Paulauskas, V., and Račkauskas, A. (2000). More on p-stable convex sets in banach spaces. *Journal of Theoretical Probability*, 13(1):39--64.

De Haan, L. and Ferreira, A. (2007). *Extreme value theory: an introduction*. Springer Science & Business Media.

De Haan, L. and Resnick, S. (1998). On asymptotic normality of the Hill estimator. *Stochastic Models*, 14(4):849--866.

De Haan, L. and Resnick, S. I. (1980). A simple asymptotic estimate for the index of a stable distribution. *Journal of the Royal Statistical Society. Series B (Methodological)*, pages 83--87.

Dekkers, A. L., Einmahl, J. H., and De Haan, L. (1989). A moment estimator for the index of an extreme-value distribution. *The Annals of Statistics*, pages 1833--1855.

El Adlouni, S., Bobée, B., and Ouarda, T. (2008). On the tails of extreme event distributions in hydrology. *Journal of Hydrology*, 355(1):16--33.

Embrechts, P., Klüppelberg, C., and Mikosch, T. (2013). *Modelling extremal events: for insurance and finance*, volume 33. Springer Science & Business Media.

Fisher, R. A. and Tippett, L. H. C. (1928). Limiting forms of the frequency distribution of the largest or smallest member of a sample. In *Mathematical Proceedings of the Cambridge Philosophical Society*, volume 24, pages 180--190. Cambridge Univ Press.

Gardes, L. and Girard, S. (2010). Conditional extremes from heavy-tailed distributions: An application to the estimation of extreme rainfall return levels. *Extremes*, 13(2):177--204.

Gnedenko, B. (1943). Sur la distribution limite du terme maximum d'une serie aleatoire. *Annals of Mathematics*, pages 423--453.

Gomes, M. I. and Guillou, A. (2015). Extreme value theory and statistics of univariate extremes: A review. *International Statistical Review*, 83(2):263--292.

Haan, D. and Peng, L. (1998). Comparison of tail index estimators. *Statistica Neerlandica*, 52(1):60--70.

Hill, B. M. (1975). A simple general approach to inference about the tail of a distribution. *The Annals of Statistics*, 3(5):1163--1174.

Hill, J. B. and Shneyerov, A. (2013). Are there common values in first-price auctions? a tail-index nonparametric test. *Journal of Econometrics*, 174(2):144--164.

Hols, M. C. and De Vries, C. G. (1991). The limiting distribution of extremal exchange rate returns. *Journal of Applied Econometrics*, 6(3):287--302.

Huisman, R., Koedijk, K. G., Kool, C. J. M., and Palm, F. (2001). Tail-index estimates in small samples. *Journal of Business & Economic Statistics*, 19(2):208--216.

Jenkinson, A. F. (1955). The frequency distribution of the annual maximum (or minimum) values of meteorological elements. *Quarterly Journal of the Royal Meteorological Society*, 81(348):158--171.

Koedijk, K. G., Schafgans, M. M., and De Vries, C. G. (1990). The tail index of exchange rate returns. *Journal of International Economics*, 29(1-2):93--108.

Kratz, M. and Resnick, S. I. (1996). The qq-estimator and heavy tails. *Stochastic Models*, 12(4):699--724.

LeBaron, B. and Samanta, R. (2005). Extreme value theory and fat tails in equity markets. *Available at SSRN 873656*.

Lee, I. W. and Fapojuwo, A. O. (2005). Estimating heavy-tails in long-range dependent wireless traffic. In *2005 IEEE 61st Vehicular Technology Conference*, volume 4, pages 2132--2136. IEEE.

Li, X. and Wei, X. (2011). A shift and scale invariant estimator for index of extreme value distribution. In *Consumer Electronics, Communications and Networks (CECNet), 2011 International Conference on*, pages 994--997. IEEE.

Markovich, N. M. (2005). On-line estimation of the tail index for heavy-tailed distributions with application to www-traffic. In *Next Generation Internet Networks, 2005*, pages 388--395. IEEE.

McElroy, T. (2016). On the measurement and treatment of extremes in time series. *Extremes*, pages 1--24.

Meerschaert, M. M. and Scheffler, H.-P. (2003). Nonparametric methods for heavy tailed vector data: A survey with applications from finance and hydrology. *Recent advances and trends in nonparametric statistics*, pages 265--279.

Pacheco, J. R. (2006). Estimating the heavy-tail index for www traces. In *2006 3rd International Conference on Electrical and Electronics Engineering*, pages 1--4. IEEE.

Phalitnonkiat, P., Sun, W., Grigoriu, M. D., Hess, P., and Samorodnitsky, G. (2016). Extreme ozone events: Tail behavior of the surface ozone distribution over the U.S. *Atmospheric Environment*, 128:134--146.

Pickands III, J. (1975). Statistical inference using extreme order statistics. *The Annals of Statistics*, pages 119--131.

Resnick, S. I. (2007). *Heavy-tail phenomena: probabilistic and statistical modeling*. Springer Science & Business Media.

Rezaul, K. M. and Grout, V. (2007). A comparison of methods for estimating the tail index of heavy-tailed internet traffic. In *Innovative Algorithms and Techniques in Automation, Industrial Electronics and Telecommunications*, pages 219--222. Springer.

Rodionov, I. (2014). Properties of Hill's estimator of extreme value index for impure samples. *Moscow University Mathematics Bulletin*, 69(1):1--4.

Rojo, J. (1996). On tail categorization of probability laws. *Journal of the American Statistical Association*, 91(433):378--384.

Rojo, J. (2013). Heavy-tailed densities. *Wiley Interdisciplinary Reviews: Computational Statistics*, 5(1):30--40.

Smith, R. L. (1987). Estimating tails of probability distributions. *The Annals of Statistics*, pages 1174--1207.

von Mises, R. (1936). La distribution de la plus grande de n valeurs, vol. 2 of selected papers of richard von mises. *Providence, RI: American Mathematical Society, pp. 271r*, 294.

Willems, P., Guillou, A., and Beirlant, J. (2007). Bias correction in hydrologic gpd based extreme value analysis by means of a slowly varying function. *Journal of Hydrology*, 338(3):221--236.

Appendices

Appendix A

Table A.1: Standard deviation, Bias and Mean squared error of estimators from survival function of $\bar{F} = \frac{1}{ln(x+e)(1+x)^{\alpha}}$ for true value of $\alpha = 0.5$

k=5000, n=40000, 5000 iterations, $\bar{F} = \frac{1}{ln(x+e)(1+x)^{\alpha}}$, x>0			
Statistics	Sd	Bias	MSE
Hill's estimator	0.010994	0.235537	0.055599
Pickands est.	0.023084	0.342519	0.117852
New estimator	0.081357	0.143215	0.027128
De Haan & Resnick	0.087052	0.099787	0.017534
n=40,1000blks	0.005461	-0.07378	0.005473
n=100,400blks	0.006679	-0.05892	0.003516
n=200, 200blks	0.007957	-0.05379	0.002956
n=500, 80blks	0.010875	-0.05035	0.002654
n=1000, 40blks	0.013395	-0.04963	0.002643
n=2000, 20blks	0.017155	-0.04911	0.002706
n=4000, 10blks	0.022014	-0.04863	0.002849
n=5000, 8blks	0.024435	-0.0484	0.00294

Table A.2: Standard deviation, Bias and Mean squared error of estimators from survival function of $\bar{F} = \frac{1}{ln(x+e)(1+x)^{\alpha}}$ for true value of $\alpha = 0.75$

k=5000, n=40000, 5000 iterations, $\bar{F} = \frac{1}{ln(x+e)(1+x)^{\alpha}}$, x>0			
Statistics	Sd	Bias	MSE
Hill's estimator	0.013757	0.226717	0.05159
Pickands est.	0.038054	0.377664	0.144078
New estimator	0.111138	0.177332	0.043796
De Haan & Resnick	0.121037	0.121924	0.029512
n=40,1000blks	0.007148	-0.10319	0.0107
n=100,400blks	0.009044	-0.10133	0.010349
n=200, 200blks	0.01142	-0.10179	0.010492
n=500, 80blks	0.014946	-0.10174	0.010575
n=1000, 40blks	0.019507	-0.1017	0.010723
n=2000, 20blks	0.024918	-0.10101	0.010825
n=4000, 10blks	0.032316	-0.10039	0.011122
n=5000, 8blks	0.035857	-0.09942	0.011171

Table A.3: Standard deviation, Bias and Mean squared error of estimators from survival function of $\bar{F} = \frac{1}{ln(x+e)(1+x)^{\alpha}}$ for true value of $\alpha = 1$

k=5000, n=40000, 5000 iterations, $\bar{F} = \frac{1}{ln(x+e)(1+x)^{\alpha}}$, x>0			
Statistics	Sd	Bias	MSE
Hill's estimator	0.016026	0.170823	0.029437
Pickands est.	0.056851	0.402279	0.16506
New estimator	0.141344	0.199081	0.059607
De Haan & Resnick	0.152493	0.122122	0.038163
n=40,1000blks	0.008892	-0.16633	0.027746
n=100,400blks	0.011623	-0.16995	0.029019
n=200, 200blks	0.014191	-0.17084	0.029388
n=500, 80blks	0.019954	-0.17054	0.029483
n=1000, 40blks	0.025507	-0.16841	0.029011
n=2000, 20blks	0.032626	-0.16565	0.028503
n=4000, 10blks	0.04248	-0.16305	0.028389
n=5000, 8blks	0.048586	-0.16091	0.028251

Table A.4: Standard deviation, Bias and Mean squared error of estimators from survival function of $\bar{F} = \frac{1}{ln(x+e)(1+x)^{\alpha}}$ for true value of $\alpha = 2$

k=5000, n=40000, 5000 iterations, $\bar{F} = \frac{1}{ln(x+e)(1+x)^{\alpha}}$, x>0			
Statistics	Sd	Bias	MSE
Hill's estimator	0.021783	-0.33599	0.113362
Pickands est.	0.166826	0.452093	0.232213
New estimator	0.250077	0.240052	0.120151
De Haan & Resnick	0.231038	-0.06563	0.057675
n=40,1000blks	0.01652681	-0.5627745	0.3169882
n=100,400blks	0.02162468	-0.5552731	0.3087957
n=200, 200blks	0.02750165	-0.5416408	0.294131
n=500, 80blks	0.03818421	-0.519484	0.2713213
n=1000, 40blks	0.05008286	-0.5037726	0.2562946
n=2000, 20blks	0.06353328	-0.4857156	0.2399553
n=4000, 10blks	0.08477473	-0.4709861	0.2290132
n=5000, 8blks	0.09270431	-0.4673176	0.2269781

Table A.5: Standard deviation, Bias and Mean squared error of estimators from survival function of $\bar{F} = \frac{1}{ln(x+e)(1+x)^{\alpha}}$ for true value of $\alpha = 3$

k=5000, n=40000, 5000 iterations, $\bar{F} = \frac{1}{ln(x+e)(1+x)^{\alpha}}$, x>0			
Statistics	Sd	Bias	MSE
Hill's estimator	0.023708	-1.06771	1.140568
Pickands est.	0.325472	0.485223	0.341352
New estimator	0.371984	0.278946	0.216155
De Haan & Resnick	0.270768	-0.49536	0.318683
n=40,1000blks	0.02374269	-1.067933	1.141045
n=100,400blks	0.03200677	-1.0271	1.055959
n=200, 200blks	0.04078602	-0.9900388	0.9818401
n=500, 80blks	0.05734418	-0.9392235	0.8854285
n=1000, 40blks	0.07361633	-0.9021961	0.819376
n=2000, 20blks	0.09583906	-0.8655588	0.7583753
n=4000, 10blks	0.1271034	-0.8325749	0.709333
n=5000, 8blks	0.1401381	-0.8210087	0.69369

Table A.6: Standard deviation, Bias and Mean squared error of estimators from survival function of $\bar{F} = \frac{1}{ln(x+e)(1+x)^{\alpha}}$ for true value of $\alpha = 5$

k=5000, n=40000, 5000 iterations, $\bar{F} = \frac{1}{ln(x+e)(1+x)^{\alpha}}$, x>0			
Statistics	Sd	Bias	MSE
Hill's estimator	0.027323	-2.78106	7.735041
Pickands est.	0.858417	0.591218	1.086272
New estimator	0.729499	0.589058	0.879051
De Haan & Resnick	0.304541	-1.79464	3.313451
n=40,1000blks	0.03837284	-2.238089	5.010516
n=100,400blks	0.05135158	-2.116474	4.482099
n=200, 200blks	0.06640241	-2.01634	4.070036
n=500, 80blks	0.09294263	-1.887974	3.573082
n=1000, 40blks	0.1210395	-1.801709	3.260804
n=2000, 20blks	0.1581419	-1.719591	2.981997
n=4000, 10blks	0.2136436	-1.63975	2.734413
n=5000, 8blks	0.2314971	-1.621067	2.681439

Table A.7: Standard deviation, Bias and Mean squared error of estimators from survival function of $\bar{F} = \frac{1}{ln(x+e)(1+x)^{\alpha}}$ for true value of $\alpha = 7$

k=5000, n=40000, 5000 iterations, $\bar{F} = \frac{1}{ln(x+e)(1+x)^{\alpha}}$, x>0			
Statistics	Sd	Bias	MSE
Hill's estimator	0.029345	-4.63127	21.44952
Pickands est.	1.68253	0.782163	3.442121
New estimator	1.41195	1.692344	4.857233
De Haan & Resnick	0.312316	-3.38685	11.56828
n=40,1000blks	0.0517698	-3.535281	12.50089
n=100,400blks	0.07165377	-3.318277	11.01609
n=200, 200blks	0.09112558	-3.154205	9.957311
n=500, 80blks	0.1276948	-2.943618	8.681191
n=1000, 40blks	0.1670837	-2.799035	7.86251
n=2000, 20blks	0.2160819	-2.668606	7.168138
n=4000, 10blks	0.2865966	-2.539821	6.532812
n=5000, 8blks	0.3201474	-2.504886	6.37693

Table A.8: Standard deviation, Bias and Mean squared error of estimators from survival function of $\bar{F} = \frac{1}{ln(x+e)(1+x)^{\alpha}}$ for true value of $\alpha = 10$

k=5000, n=40000, 5000 iterations, $\bar{F} = \frac{1}{ln(x+e)(1+x)^{\alpha}}$, x>0			
Statistics	Sd	Bias	MSE
Hill's estimator	0.030269	-7.5038	56.3079
Pickands est.	4.396818	1.563076	21.77135
New estimator	4.982569	7.568754	82.10706
De Haan & Resnick	0.32495	-6.02426	36.39722
n=40,1000blks	0.07054307	-5.633296	31.73899
n=100,400blks	0.0998428	-5.263666	27.71615
n=200, 200blks	0.1290942	-4.990105	24.91781
n=500, 80blks	0.1787395	-4.64266	
n=1000, 40blks	0.2390361	-4.412663	19.52872
n=2000, 20blks	0.3116778	-4.191211	17.66337
n=4000, 10blks	0.4078714	-3.987638	16.06758
n=5000, 8blks	0.4457378	-3.932987	15.66703

Table A.9: Standard deviation, Bias and Mean squared error of estimators from survival function of $\bar{F} = \frac{1}{1+\frac{ln(1+x)}{1+x}(1+x)^{\alpha}}$ for true value of $\alpha = 0.5$

α =0.5, k=5000, n=40000, 5000 iterations, $\bar{F} = \frac{1}{1+\frac{ln(1+x)}{1+x}(1+x)^{\alpha}}$, x>0			
Statistics	Sd	Bias	MSE
Hill's estimator	0.006586	-0.02346	0.000594
Pickands est.	0.007206	-0.0819	0.00676
New estimator	0.051365	-0.02049	0.003057
De Haan & Resnick	0.061165	-0.02793	0.004521
n=40,1000blks	0.006586	-0.00036	4.35E-05
n=100,400blks	0.007939	0.019521	0.000444
n=200, 200blks	0.009344	0.026135	0.00077
n=500, 80blks	0.012243	0.029903	0.001044
n=1000, 40blks	0.01553	0.03023	0.001155
n=2000, 20blks	0.019727	0.030373	0.001312
n=4000, 10blks	0.025051	0.02937	0.00149
n=5000, 8blks	0.027313	0.029243	0.001601

Table A.10: Standard deviation, Bias and Mean squared error of estimators from survival function of $\bar{F} = \frac{1}{1+\frac{ln(1+x)}{1+x}(1+x)^{\alpha}}$ for true value of $\alpha = 0.75$

α =0.75, k=5000, n=40000, 5000 iterations, $\bar{F} = \frac{1}{1+\frac{ln(1+x)}{1+x}(1+x)^{\alpha}}$, x>0			
Statistics	Sd	Bias	MSE
Hill's estimator	0.008721	-0.08456	0.007226
Pickands est.	0.012477	-0.15562	0.024373
New estimator	0.075232	-0.03156	0.006655
De Haan & Resnick	0.08656	-0.05605	0.010633
n=40,1000blks	0.008942	0.071237	0.005155
n=100,400blks	0.011018	0.072149	0.005327
n=200, 200blks	0.013251	0.068927	0.004926
n=500, 80blks	0.017577	0.063607	0.004355
n=1000, 40blks	0.021984	0.059149	0.003982
n=2000, 20blks	0.028705	0.054869	0.003834
n=4000, 10blks	0.036497	0.05108	0.003941
n=5000, 8blks	0.039727	0.04981	0.004059

Table A.11: Standard deviation, Bias and Mean squared error of estimators from survival function of $\bar{F} = \frac{1}{1+\frac{ln(1+x)}{1+x}(1+x)^{\alpha}}$ for true value of $\alpha = 1$

α =1, k=5000, n=40000, 5000 iterations, $\bar{F} = \frac{1}{1+\frac{ln(1+x)}{1+x}(1+x)^{\alpha}}$, x>0			
Statistics	Sd	Bias	MSE
Hill's estimator	0.011253	-0.15968	0.025623
Pickands est.	0.020474	-0.19849	0.039816
New estimator	0.100656	-0.03897	0.011649
De Haan & Resnick	0.111098	-0.09014	0.020465
n=40,1000blks	0.011679	0.130055	0.017051
n=100,400blks	0.014335	0.115154	0.013466
n=200, 200blks	0.01704	0.104157	0.011139
n=500, 80blks	0.023213	0.090941	0.008809
n=1000, 40blks	0.029665	0.082252	0.007645
n=2000, 20blks	0.037672	0.075289	0.007087
n=4000, 10blks	0.049421	0.069133	0.007221
n=5000, 8blks	0.05324	0.067562	0.007398

Table A.12: Standard deviation, Bias and Mean squared error of estimators from survival function of $\bar{F} = \frac{1}{1+\frac{ln(1+x)}{1+x}(1+x)^{\alpha}}$ for true value of $\alpha = 2$

α =2, k=5000, n=40000, 5000 iterations, $\bar{F} = \frac{1}{1+\frac{ln(1+x)}{1+x}(1+x)^{\alpha}}$, x>0			
Statistics	Sd	Bias	MSE
Hill's estimator	0.018248	-0.60209	0.36285
Pickands est.	0.079347	-0.31388	0.104817
New estimator	0.210662	-0.06858	0.049073
De Haan & Resnick	0.192793	-0.33988	0.152679
n=40,1000blks	0.021931	0.315808	0.100216
n=100,400blks	0.027474	0.251327	0.06392
n=200, 200blks	0.034068	0.218109	0.048732
n=500, 80blks	0.046445	0.186127	0.0368
n=1000, 40blks	0.057956	0.166794	0.031178
n=2000, 20blks	0.076631	0.153346	0.029386
n=4000, 10blks	0.098801	0.136828	0.028482
n=5000, 8blks	0.105259	0.134619	0.0292

Table A.13: Standard deviation, Bias and Mean squared error of estimators from survival function of $\bar{F} = \frac{1}{1+\frac{ln(1+x)}{1+x}(1+x)^{\alpha}}$ for true value of $\alpha = 3$

α =3, k=5000, n=40000, 5000 iterations, $\bar{F} = \frac{1}{1+\frac{ln(1+x)}{1+x}(1+x)^{\alpha}}$, x>0			
Statistics	Sd	Bias	MSE
Hill's estimator	0.021765	-1.26121	1.591112
Pickands est.	0.183006	-0.41594	0.206489
New estimator	0.333592	-0.06865	0.115975
De Haan & Resnick	0.241259	-0.76838	0.648598
n=40,1000blks	0.032569	0.478211	0.229747
n=100,400blks	0.041565	0.379528	0.145769
n=200, 200blks	0.050074	0.327911	0.110032
n=500, 80blks	0.06967	0.278682	0.082517
n=1000, 40blks	0.087526	0.252171	0.071249
n=2000, 20blks	0.114353	0.228517	0.065294
n=4000, 10blks	0.145752	0.208908	0.064882
n=5000, 8blks	0.161761	0.207645	0.069278

Table A.14: Standard deviation, Bias and Mean squared error of estimators from survival function of $\bar{F} = \frac{1}{1+\frac{ln(1+x)}{1+x}(1+x)^{\alpha}}$ for true value of $\alpha = 5$

α =5, k=5000, n=40000, 5000 iterations, $\bar{F} = \frac{1}{1+\frac{ln(1+x)}{1+x}(1+x)^{\alpha}}$, x>0			
Statistics	Sd	Bias	MSE
Hill's estimator	0.026435	-2.89553	8.384786
Pickands est.	0.531507	-0.56371	0.60021
New estimator	0.682345	0.216261	0.51227
De Haan & Resnick	0.284906	-2.005	4.101179
n=40,1000blks	0.054581	0.798526	0.640623
n=100,400blks	0.069315	0.632003	0.404231
n=200, 200blks	0.084805	0.547269	0.306694
n=500, 80blks	0.115384	0.46658	0.231008
n=1000, 40blks	0.148596	0.413841	0.193341
n=2000, 20blks	0.188726	0.382444	0.181874
n=4000, 10blks	0.244937	0.349109	0.181859
n=5000, 8blks	0.2634297	0.3391247	0.1843869

Table A.15: Standard deviation, Bias and Mean squared error of estimators from survival function of $\bar{F} = \frac{1}{1+\frac{ln(1+x)}{1+x}(1+x)^{\alpha}}$ for true value of $\alpha = 7$

α =7, k=5000, n=40000, 5000 iterations, $\bar{F} = \frac{1}{1+\frac{ln(1+x)}{1+x}(1+x)^{\alpha}}$, x>0			
Statistics	Sd	Bias	MSE
Hill's estimator	0.028084	-4.70667	22.15349
Pickands est.	1.138575	-0.59435	1.649343
New estimator	1.37734	1.352836	3.72685
De Haan & Resnick	0.305566	-3.53167	12.56608
n=40,1000blks	0.075633	1.119706	1.259461
n=100,400blks	0.0967	0.886991	0.796102
n=200, 200blks	0.120252	0.768041	0.604345
n=500, 80blks	0.158841	0.652483	0.450959
n=1000, 40blks	0.205835	0.590059	0.39053
n=2000, 20blks	0.265209	0.534025	0.355504
n=4000, 10blks	0.345147	0.494266	0.363402
n=5000, 8blks	0.369953	0.479875	0.367117

Table A.16: Standard deviation, Bias and Mean squared error of estimators from survival function of $\bar{F} = \frac{1}{1+\frac{ln(1+x)}{1+x}(1+x)^{\alpha}}$ for true value of $\alpha = 10$

α =10, k=5000, n=40000, 5000 iterations, $\bar{F} = \frac{1}{1+\frac{ln(1+x)}{1+x}(1+x)^{\alpha}}$, x>0			
Statistics	Sd	Bias	MSE
Hill's estimator	0.030897	-7.54991	57.00207
Pickands est.	2.830777	-0.39161	8.165054
New estimator	4.847947	7.156302	74.71054
De Haan & Resnick	0.313934	-6.12436	37.6063
n=40,1000blks	0.108553	1.596537	2.560713
n=100,400blks	0.13886	1.268649	1.628748
n=200, 200blks	0.171073	1.095257	1.228848
n=500, 80blks	0.228082	0.931123	0.919002
n=1000, 40blks	0.286844	0.836115	0.781351
n=2000, 20blks	0.374523	0.752812	0.706965
n=4000, 10blks	0.489446	0.699708	0.729101
n=5000, 8blks	0.522073	0.679745	0.734559

Table A.17: Standard deviation, Bias and Mean squared error of estimators from survival function of $\bar{F} = \frac{1}{[ln(eln(x+e))](1+x)^{\alpha}}$ for true value of $\alpha = 0.5$

α =0.5, k=5000, n=40000, 5000 iterations, $\bar{F} = \frac{1}{[ln(eln(x+e))](1+x)^{\alpha}}$, x>0			
Statistics	Sd	Bias	MSE
Hill's estimator	0.008629	0.085606	0.007403
Pickands est.	0.01495	0.147321	0.021927
New estimator	0.064108	0.048502	0.006461
De Haan & Resnick	0.069043	0.01266	0.004926
n=40,1000blks	0.005748	-0.04814	0.00235
n=100,400blks	0.007109	-0.03261	0.001114
n=200, 200blks	0.008583	-0.02691	0.000798
n=500, 80blks	0.011308	-0.02336	0.000673
n=1000, 40blks	0.014341	-0.02172	0.000677
n=2000, 20blks	0.017987	-0.02111	0.000769
n=4000, 10blks	0.023493	-0.02006	0.000954
n=5000, 8blks	0.02584	-0.0192	0.001036

Table A.18: Standard deviation, Bias and Mean squared error of estimators from survival function of $\bar{F} = \frac{1}{[ln(eln(x+e))](1+x)^{\alpha}}$ for true value of $\alpha = 0.75$

α =0.75, k=5000, n=40000, 5000 iterations, $\bar{F} = \frac{1}{[ln(eln(x+e))](1+x)^{\alpha}}$, x>0			
Statistics	Sd	Bias	MSE
Hill's estimator	0.011753	0.086942	0.007697
Pickands est.	0.026871	0.179486	0.032937
New estimator	0.094875	0.063586	0.013043
De Haan & Resnick	0.103158	0.014159	0.01084
n=40,1000blks	0.0079	-0.04621	0.002197
n=100,400blks	0.009889	-0.0433	0.001973
n=200, 200blks	0.012231	-0.04231	0.00194
n=500, 80blks	0.01658	-0.04167	0.002011
n=1000, 40blks	0.020787	-0.0413	0.002138
n=2000, 20blks	0.027331	-0.04057	0.002392
n=4000, 10blks	0.035626	-0.04085	0.002938
n=5000, 8blks	0.0079	-0.04621	0.002197

Table A.19: Standard deviation, Bias and Mean squared error of estimators from survival function of $\bar{F} = \frac{1}{[ln(eln(x+e))](1+x)^{\alpha}}$ for true value of $\alpha = 1$

α =1, k=5000, n=40000, 5000 iterations, $\bar{F} = \frac{1}{[ln(eln(x+e))](1+x)^{\alpha}}$, x>0			
Statistics	Sd	Bias	MSE
Hill's estimator	0.014544	0.04762	0.002479
Pickands est.	0.043264	0.203956	0.04347
New estimator	0.12087	0.078083	0.020703
De Haan & Resnick	0.130321	0.007825	0.017041
n=40,1000blks	0.010168	-0.06908	0.004876
n=100,400blks	0.012878	-0.07061	0.005151
n=200, 200blks	0.016211	-0.07055	0.005241
n=500, 80blks	0.022055	-0.06885	0.005227
n=1000, 40blks	0.0276	-0.06705	0.005258
n=2000, 20blks	0.035689	-0.06673	0.005726
n=4000, 10blks	0.047118	-0.06495	0.006438
n=5000, 8blks	0.051391	-0.06375	0.006705

Table A.20: Standard deviation, Bias and Mean squared error of estimators from survival function of $\bar{F} = \frac{1}{[ln(eln(x+e))](1+x)^{\alpha}}$ for true value of $\alpha = 2$

α =2, k=5000, n=40000, 5000 iterations, $\bar{F} = \frac{1}{[ln(eln(x+e))](1+x)^{\alpha}}$, x>0			
Statistics	Sd	Bias	MSE
Hill's estimator	0.02107	-0.40425	0.163863
Pickands est.	0.139308	0.258892	0.086428
New estimator	0.236854	0.111384	0.068495
De Haan & Resnick	0.220103	-0.17033	0.077448
n=40,1000blks	0.019859	-0.23902	0.057524
n=100,400blks	0.025972	-0.23071	0.0539
n=200, 200blks	0.03265	-0.22103	0.049919
n=500, 80blks	0.043452	-0.20684	0.04467
n=1000, 40blks	0.055744	-0.19567	0.041393
n=2000, 20blks	0.072781	-0.18428	0.039255
n=4000, 10blks	0.095399	-0.1754	0.039863
n=5000, 8blks	0.104102	-0.17241	0.040562

Table A.21: Standard deviation, Bias and Mean squared error of estimators from survival function of $\bar{F} = \frac{1}{[ln(eln(x+e))](1+x)^{\alpha}}$ for true value of $\alpha = 3$

α =3, k=5000, n=40000, 5000 iterations, $\bar{F} = \frac{1}{[ln(eln(x+e))](1+x)^{\alpha}}$, x>0			
Statistics	Sd	Bias	MSE
Hill's estimator	0.023936	-1.10804	1.228317
Pickands est.	0.291622	0.29743	0.173491
New estimator	0.360398	0.146641	0.151365
De Haan & Resnick	0.265906	-0.58036	0.407507
n=40,1000blks	0.0298965	-0.4562098	0.209021
n=100,400blks	0.03858009	-0.4206543	0.1784381
n=200, 200blks	0.04778821	-0.3943127	0.1577657
n=500, 80blks	0.06563101	-0.3600863	0.1339687
n=1000, 40blks	0.0849514	-0.3400228	0.1228308
n=2000, 20blks	0.1098889	-0.3216738	0.1155472
n=4000, 10blks	0.142615	-0.3042849	0.1129243
n=5000, 8blks	0.1555065	-0.2944732	0.1108919

Table A.22: Standard deviation, Bias and Mean squared error of estimators from survival function of $\bar{F} = \frac{1}{[ln(eln(x+e))](1+x)^{\alpha}}$ for true value of $\alpha = 5$

α =5, k=5000, n=40000, 5000 iterations, $\bar{F} = \frac{1}{[ln(eln(x+e))](1+x)^{\alpha}}$, x>0			
Statistics	Sd	Bias	MSE
Hill's estimator	0.027411	-2.79928	7.836692
Pickands est.	0.797712	0.391582	0.789553
New estimator	0.720624	0.458934	0.729815
De Haan & Resnick	0.305581	-1.84153	3.4846
n=40,1000blks	0.0505054	-0.9422479	0.8903813
n=100,400blks	0.06519905	-0.8428541	0.7146531
n=200, 200blks	0.08171109	-0.7761008	0.6090078
n=500, 80blks	0.1132071	-0.6949482	0.4957663
n=1000, 40blks	0.1443464	-0.6480802	0.4408396
n=2000, 20blks	0.1847917	-0.6053936	0.4006425
n=4000, 10blks	0.2382278	-0.5719583	0.3838774
n=5000, 8blks	0.2580852	-0.5562211	0.3759766

Table A.23: Standard deviation, Bias and Mean squared error of estimators from survival function of $\bar{F} = \frac{1}{[ln(eln(x+e))](1+x)^{\alpha}}$ for true value of $\alpha = 7$

α =7, k=5000, n=40000, 5000 iterations, $\bar{F} = \frac{1}{[ln(eln(x+e))](1+x)^{\alpha}}$, x>0			
Statistics	Sd	Bias	MSE
Hill's estimator	0.029089	-4.64006	21.53102
Pickands est.	1.640103	0.584292	3.030797
New estimator	1.420973	1.520372	4.330292
De Haan & Resnick	0.322061	-3.42105	11.8073
n=40,1000blks	0.06979401	-1.472174	2.172167
n=100,400blks	0.09160463	-1.299717	1.697654
n=200, 200blks	0.1158134	-1.185764	1.419445
n=500, 80blks	0.1534725	-1.061476	1.150279
n=1000, 40blks	0.1997038	-0.9848263	1.009756
n=2000, 20blks	0.2536211	-0.9140467	0.8997921
n=4000, 10blks	0.3330588	-0.8618532	0.853697
n=5000, 8blks	0.3598507	-0.8417578	0.8380228

Table A.24: Standard deviation, Bias and Mean squared error of estimators from survival function of $\bar{F} = \frac{1}{[ln(eln(x+e))](1+x)^{\alpha}}$ for true value of $\alpha = 10$

α =10, k=5000, n=40000, 5000 iterations, $\bar{F} = \frac{1}{[ln(eln(x+e))](1+x)^{\alpha}}$, x>0			
Statistics	Sd	Bias	MSE
Hill's estimator	0.030145	-7.50869	56.38135
Pickands est.	4.290617	1.453424	20.51815
New estimator	4.844116	7.366777	77.73017
De Haan & Resnick	0.32177	-6.03277	36.49778
n=40,1000blks	0.1025723	-2.317968	5.383496
n=100,400blks	0.1305565	-2.018449	4.091178
n=200, 200blks	0.163355	-1.838046	3.405091
n=500, 80blks	0.2222236	-1.62985	2.705786
n=1000, 40blks	0.2807136	-1.506659	2.348807
n=2000, 20blks	0.3620385	-1.39833	2.086372
n=4000, 10blks	0.4740665	-1.304796	1.927186
n=5000, 8blks	0.5234977	-1.293076	1.94604

Table A.25: Standard deviation, Bias and Mean squared error of estimators from survival function of $\bar{F} = \frac{1}{[ln(x+e)]x^{\alpha}}$ for true value of $\alpha = 0.5$

α =0.5, k=5000, n=40000, 5000 iterations, $\bar{F} = \frac{1}{[ln(x+e)]x^{\alpha}}$, x>1			
Statistics	Sd	Bias	MSE
Hill's estimator	0.012131	0.27728	0.077031
Pickands est.	0.023041	0.340003	0.116133
New estimator	0.080741	0.144267	0.027331
De Haan & Resnick	0.088254	0.106618	0.019155
n=40,1000blks	0.004552	-0.00913	0.000104
n=100,400blks	0.005854	-0.02521	0.00067
n=200, 200blks	0.007363	-0.03243	0.001106
n=500, 80blks	0.010041	-0.03875	0.001602
n=1000, 40blks	0.012811	-0.04174	0.001907
n=2000, 20blks	0.016919	-0.04372	0.002198
n=4000, 10blks	0.021997	-0.04483	0.002494
n=5000, 8blks	0.0239	-0.04529	0.002622

Table A.26: Standard deviation, Bias and Mean squared error of estimators from survival function of $\bar{F} = \frac{1}{[ln(x+e)]x^{\alpha}}$ for true value of $\alpha = 0.75$

α =0.75, k=5000, n=40000, 5000 iterations, $\bar{F} = \frac{1}{[ln(x+e)]x^{\alpha}}$, x>1			
Statistics	Sd	Bias	MSE
Hill's estimator	0.016711	0.342449	0.11755
Pickands est.	0.037696	0.375664	0.142544
New estimator	0.111782	0.172272	0.04217
De Haan & Resnick	0.127174	0.136569	0.034821
n=40,1000blks	0.006701	-0.0682	0.004697
n=100,400blks	0.008554	-0.08576	0.007427
n=200, 200blks	0.010726	-0.09299	0.008762
n=500, 80blks	0.01464	-0.09744	0.00971
n=1000, 40blks	0.019019	-0.09942	0.010246
n=2000, 20blks	0.024623	-0.09972	0.010551
n=4000, 10blks	0.032633	-0.09935	0.010935
n=5000, 8blks	0.035224	-0.09857	0.010957

Table A.27: Standard deviation, Bias and Mean squared error of estimators from survival function of $\bar{F} = \frac{1}{[ln(x+e)]x^{\alpha}}$ for true value of $\alpha = 1$

α =1, k=5000, n=40000, 5000 iterations, $\bar{F} = \frac{1}{[ln(x+e)]x^{\alpha}}$, x>1			
Statistics	Sd	Bias	MSE
Hill's estimator	0.020665	0.392722	0.154658
Pickands est.	0.056355	0.400845	0.163852
New estimator	0.142757	0.196509	0.058991
De Haan & Resnick	0.166668	0.167111	0.055699
n=40,1000blks	0.008665	-0.14559	0.021271
n=100,400blks	0.011189	-0.16203	0.02638
n=200, 200blks	0.014329	-0.16693	0.02807
n=500, 80blks	0.020141	-0.16793	0.028605
n=1000, 40blks	0.025528	-0.16726	0.028626
n=2000, 20blks	0.032969	-0.16559	0.028508
n=4000, 10blks	0.043139	-0.1634	0.028559
n=5000, 8blks	0.046001	-0.16216	0.028411

Table A.28: Standard deviation, Bias and Mean squared error of estimators from survival function of $\bar{F} = \frac{1}{[ln(x+e)]x^{\alpha}}$ for true value of $\alpha = 2$

α =2, k=5000, n=40000, 5000 iterations, $\bar{F} = \frac{1}{[ln(x+e)]x^{\alpha}}$, x>1			
Statistics	Sd	Bias	MSE
Hill's estimator	0.036505	0.509444	0.260865
Pickands est.	0.168345	0.469127	0.248414
New estimator	0.247542	0.239765	0.118752
De Haan & Resnick	0.305499	0.245727	0.153693
n=40,1000blks	0.016348	-0.55804	0.311677
n=100,400blks	0.021225	-0.5533	0.306592
n=200, 200blks	0.02766	-0.54133	0.293808
n=500, 80blks	0.038442	-0.5196	0.271457
n=1000, 40blks	0.050147	-0.50326	0.255782
n=2000, 20blks	0.065396	-0.48561	0.240096
n=4000, 10blks	0.087055	-0.46903	0.227561
n=5000, 8blks	0.091935	-0.46254	0.222395

Table A.29: Standard deviation, Bias and Mean squared error of estimators from survival function of $\bar{F} = \frac{1}{[ln(x+e)]x^{\alpha}}$ for true value of $\alpha = 3$

α =3, k=5000, n=40000, 5000 iterations, $\bar{F} = \frac{1}{[ln(x+e)]x^{\alpha}}$, x>1			
Statistics	Sd	Bias	MSE
Hill's estimator	0.051723	0.568673	0.326064
Pickands est.	0.326943	0.515786	0.372905
New estimator	0.350503	0.238217	0.179575
De Haan & Resnick	0.441273	0.281489	0.273919
n=40,1000blks	0.023609	-1.06528	1.135372
n=100,400blks	0.031954	-1.02776	1.057312
n=200, 200blks	0.041426	-0.98909	0.980012
n=500, 80blks	0.056948	-0.93775	0.882619
n=1000, 40blks	0.074129	-0.90047	0.81634
n=2000, 20blks	0.098038	-0.86445	0.756875
n=4000, 10blks	0.125465	-0.83387	0.711079
n=5000, 8blks	0.137217	-0.81876	0.689187

Table A.30: Standard deviation, Bias and Mean squared error of estimators from survival function of $\bar{F} = \frac{1}{[ln(x+e)]x^{\alpha}}$ for true value of $\alpha = 5$

α =5, k=5000, n=40000, 5000 iterations, $\bar{F} = \frac{1}{[ln(x+e)]x^{\alpha}}$, x>1			
Statistics	Sd	Bias	MSE
Hill's estimator	0.079981	0.624678	0.396619
Pickands est.	0.875537	0.626787	1.159273
New estimator	0.540183	0.157048	0.316404
De Haan & Resnick	0.690214	0.261594	0.544732
n=40,1000blks	0.037684	-2.23701	5.00562
n=100,400blks	0.051482	-2.11461	4.474228
n=200, 200blks	0.066303	-2.01677	4.071755
n=500, 80blks	0.09187	-1.89077	3.583431
n=1000, 40blks	0.121693	-1.80204	3.26217
n=2000, 20blks	0.159884	-1.71937	2.981789
n=4000, 10blks	0.209772	-1.64211	2.740516
n=5000, 8blks	0.229469	-1.62431	2.691022

Table A.31: Standard deviation, Bias and Mean squared error of estimators from survival function of $\bar{F} = \frac{1}{[ln(x+e)]x^{\alpha}}$ for true value of $\alpha = 7$

α =7, k=5000, n=40000, 5000 iterations, $\bar{F} = \frac{1}{[ln(x+e)]x^{\alpha}}$, x>1			
Statistics	Sd	Bias	MSE
Hill's estimator	0.109423	0.65086	0.43559
Pickands est.	1.777492	0.856183	3.891896
New estimator	0.737192	0.05094	0.545938
De Haan & Resnick	0.946328	0.214042	0.941172
n=40,1000blks	0.051551	-3.53564	12.50343
n=100,400blks	0.070529	-3.31842	11.01688
n=200, 200blks	0.090747	-3.15139	9.939471
n=500, 80blks	0.125775	-2.94305	8.677349
n=1000, 40blks	0.171286	-2.79672	7.850976
n=2000, 20blks	0.221802	-2.66019	7.125787
n=4000, 10blks	0.291136	-2.53824	6.527385
n=5000, 8blks	0.316129	-2.50613	6.380602

Table A.32: Standard deviation, Bias and Mean squared error of estimators from survival function of $\bar{F} = \frac{1}{[ln(x+e)]x^{\alpha}}$ for true value of $\alpha = 10$

α =10, k=5000, n=40000, 5000 iterations, $\bar{F} = \frac{1}{[ln(x+e)]x^{\alpha}}$, x>1			
Statistics	Sd	Bias	MSE
Hill's estimator	0.151528	0.6708	0.472929
Pickands est.	11.50347	1.64519	135.01
New estimator	1.013694	-0.12121	1.042062
De Haan & Resnick	1.304817	0.126747	1.718271
n=40,1000blks	0.071302	-5.63425	31.74988
n=100,400blks	0.099829	-5.26544	27.7348
n=200, 200blks	0.125856	-4.98871	24.90308
n=500, 80blks	0.179173	-4.64577	21.61523
n=1000, 40blks	0.233685	-4.4055	19.46302
n=2000, 20blks	0.316822	-4.1962	17.70841
n=4000, 10blks	0.411496	-3.98954	16.08576
n=5000, 8blks	0.450539	-3.92876	15.63811

Table A.33: Standard deviation, Bias and Mean squared error of estimators from survival function of $\bar{F} = \frac{1}{(1+x)^\alpha}$ for true value of $\alpha = 0.5$

α =0.5, k=5000, n=40000, 5000 iterations, $\bar{F} = \frac{1}{(1+x)^\alpha}$, x>0			
Statistics	Sd	Bias	MSE
Hill's estimator	0.006974	-0.0025	5.49E-05
Pickands est.	0.0097	0.000333	9.42E-05
New estimator	0.050825	-0.02055	0.003005
De haan & Resnick	0.061517	-0.02424	0.004371
n=40,1000blks	0.006259707	0.03857669	0.001527337
n=100,400blks	0.007506953	0.04308462	0.001912628
n=200, 200blks	0.009009118	0.04225627	0.00186674
n=500, 80blks	0.01184758	0.03998551	0.001739178
n=1000, 40blks	0.01493691	0.037245	0.001610257
n=2000, 20blks	0.01917761	0.03554999	0.001631508
n=4000, 10blks	0.02477839	0.03293315	0.001698438
n=5000, 8blks	0.02744176	0.03264338	0.00181849

Table A.34: Standard deviation, Bias and Mean squared error of estimators from survival function of $\bar{F} = \frac{1}{(1+x)^\alpha}$ for true value of $\alpha = 0.75$

α =0.75, k=5000, n=40000, 5000 iterations, $\bar{F} = \frac{1}{(1+x)^\alpha}$, x>0			
Statistics	Sd	Bias	MSE
Hill's estimator	0.010045	-0.02016	0.000507
Pickands est.	0.018491	0.000349	0.000342
New estimator	0.077173	-0.02964	0.006833
De haan & Resnick	0.092515	-0.03813	0.010011
n=40,1000blks	0.008565643	0.1038257	0.01085314
n=100,400blks	0.01075557	0.08824107	0.007902146
n=200, 200blks	0.01302044	0.07859902	0.006347304
n=500, 80blks	0.01718599	0.0686574	0.005009138
n=1000, 40blks	0.02205111	0.06217935	0.004352426
n=2000, 20blks	0.02794153	0.05615302	0.003933734
n=4000, 10blks	0.0366867	0.05201068	0.004050755
n=5000, 8blks	0.03988517	0.04988283	0.004078805

Table A.35: Standard deviation, Bias and Mean squared error of estimators from survival function of $\bar{F} = \frac{1}{(1+x)^{\alpha}}$ for true value of $\alpha = 1$

α =1, k=5000, n=40000, 5000 iterations, $\bar{F} = \frac{1}{(1+x)^{\alpha}}$, x>0			
Statistics	Sd	Bias	MSE
Hill's estimator	0.01292	-0.06338	0.004183
Pickands est.	0.031312	0.000842	0.000981
New estimator	0.099743	-0.03799	0.01139
De haan & Resnick	0.11744	-0.05821	0.017178
n=40,1000blks	0.01103639	0.1532378	0.02360359
n=100,400blks	0.01387641	0.1240727	0.01558655
n=200, 200blks	0.01729714	0.1081451	0.01199448
n=500, 80blks	0.02329814	0.09252761	0.009104054
n=1000, 40blks	0.02965933	0.08305087	0.007776946
n=2000, 20blks	0.03893367	0.07540556	0.007201526
n=4000, 10blks	0.04949827	0.07005244	0.007356932
n=5000, 8blks	0.05320117	0.06822817	0.007484882

Table A.36: Standard deviation, Bias and Mean squared error of estimators from survival function of $\bar{F} = \frac{1}{(1+x)^{\alpha}}$ for true value of $\alpha = 2$

α =2, k=5000, n=40000, 5000 iterations, $\bar{F} = \frac{1}{(1+x)^{\alpha}}$, x>0			
Statistics	Sd	Bias	MSE
Hill's estimator	0.019969	-0.48847	0.239002
Pickands est.	0.108844	0.005403	0.011874
New estimator	0.207261	-0.07998	0.049345
De haan & Resnick	0.209288	-0.26456	0.113785
n=40,1000blks	0.02194616	0.3191198	0.102319
n=100,400blks	0.02804897	0.2529105	0.06475031
n=200, 200blks	0.03373862	0.2190757	0.04913224
n=500, 80blks	0.04607082	0.186597	0.03694055
n=1000, 40blks	0.05823441	0.1675705	0.03147044
n=2000, 20blks	0.07402977	0.1531882	0.02894594
n=4000, 10blks	0.09777919	0.1390228	0.0288862
n=5000, 8blks	0.1079021	0.1347941	0.02980998

Table A.37: Standard deviation, Bias and Mean squared error of estimators from survival function of $\bar{F} = \frac{1}{(1+x)^\alpha}$ for true value of $\alpha = 3$

α =3, k=5000, n=40000, 5000 iterations, $\bar{F} = \frac{1}{(1+x)^\alpha}$, x>0			
Statistics	Sd	Bias	MSE
Hill's estimator	0.023343	-1.1652	1.358225
Pickands est.	0.249798	0.029087	0.063233
New estimator	0.321102	-0.09951	0.112987
De haan & Resnick	0.258124	-0.66871	0.513789
n=40,1000blks	0.03289679	0.4793042	0.2308145
n=100,400blks	0.04193313	0.3801758	0.1462916
n=200, 200blks	0.05096263	0.3297897	0.1113579
n=500, 80blks	0.06820186	0.2772793	0.08153436
n=1000, 40blks	0.08821962	0.2514543	0.07101043
n=2000, 20blks	0.1122283	0.2282023	0.06466897
n=4000, 10blks	0.1471962	0.2129455	0.06700818
n=5000, 8blks	0.1576334	0.2006923	0.06512071

Table A.38: Standard deviation, Bias and Mean squared error of estimators from survival function of $\bar{F} = \frac{1}{(1+x)^\alpha}$ for true value of $\alpha = 5$

α =5, k=5000, n=40000, 5000 iterations, $\bar{F} = \frac{1}{(1+x)^\alpha}$, x>0			
Statistics	Sd	Bias	MSE
Hill's estimator	0.026751	-2.82905	8.00424
Pickands est.	0.692809	0.073695	0.48532
New estimator	0.630792	0.101117	0.408044
De haan & Resnick	0.293935	-1.89963	3.694981
0.05407554	0.7998903	0.642748	0.05407554
0.0698187	0.6320741	0.4043914	0.0698187
0.0860035	0.5482112	0.3079307	0.0860035
0.115717	0.4649748	0.2295893	0.115717
0.1467542	0.4178196	0.1961057	0.1467542
0.1887935	0.3813704	0.1810792	0.1887935
0.2421337	0.3462732	0.1785221	0.2421337
0.2656874	0.3365799	0.1838617	0.2656874

Table A.39: Standard deviation, Bias and Mean squared error of estimators from survival function of $\bar{F} = \frac{1}{(1+x)^{\alpha}}$ for true value of $\alpha = 7$

α =7, k=5000, n=40000, 5000 iterations, $\bar{F} = \frac{1}{(1+x)^{\alpha}}$, x>0			
Statistics	Sd	Bias	MSE
Hill's estimator	0.028682	-4.65818	21.69946
Pickands est.	1.477972	0.233258	2.238373
New estimator	1.24739	0.963239	2.483501
De haan & Resnick	0.314162	-3.46154	12.08095
n=40,1000blks	0.07746162	1.117881	1.255656
n=100,400blks	0.09745315	0.8854813	0.7935723
n=200, 200blks	0.118176	0.7651034	0.5993459
n=500, 80blks	0.1615618	0.6500484	0.4486599
n=1000, 40blks	0.2045863	0.5866075	0.3859555
n=2000, 20blks	0.2614978	0.5250804	0.3440769
n=4000, 10blks	0.344094	0.4841648	0.3527926
n=5000, 8blks	0.3690909	0.4780904	0.3647713

Table A.40: Standard deviation, Bias and Mean squared error of estimators from survival function of $\bar{F} = \frac{1}{(1+x)^{\alpha}}$ for true value of $\alpha = 10$

α =10, k=5000, n=40000, 5000 iterations, $\bar{F} = \frac{1}{(1+x)^{\alpha}}$, x>0			
Statistics	Sd	Bias	MSE
Hill's estimator	0.030792	-7.51904	56.53689
Pickands est.	5.955733	0.907357	36.28696
New estimator	3.970494	5.785645	49.23536
De haan & Resnick	0.318109	-6.0654	36.8902
n=40,1000blks	0.1106577	1.599295	2.569986
n=100,400blks	0.1375319	1.263387	1.615057
n=200, 200blks	0.1697337	1.090715	1.218463
n=500, 80blks	0.2355912	0.9316792	0.9235183
n=1000, 40blks	0.2903284	0.8328911	0.7779813
n=2000, 20blks	0.3763366	0.7631112	0.7239395
n=4000, 10blks	0.4992994	0.6903523	0.7258363
n=5000, 8blks	0.5477182	0.678252	0.7599611

Table A.41: Standard deviation, Bias and Mean squared error of estimators from survival function of $\bar{F} = \frac{1}{(1+x-\theta)^{\alpha}}$ for true value of $\alpha = 0.5$

α =0.5, k=5000, n=40000, 5000 iterations, $\bar{F} = \frac{1}{(1+x-\theta)^{\alpha}}$, $x > \theta$			
Statistics	Sd	Bias	MSE
Hill's estimator	0.007571	0.017876	0.000377
Pickands est.	0.009965	0.000184	9.93E-05
New estimator	0.051361	-0.01994	0.003035
De haan & Resnick	0.063014	-0.02019	0.004378
n=40,1000blks	0.003437583	0.2553995	0.06524071
n=100,400blks	0.004982032	0.1634803	0.02675063
n=200, 200blks	0.006563532	0.120951	0.01467221
n=500, 80blks	0.00959648	0.08492821	0.007304875
n=1000, 40blks	0.0129373	0.06688331	0.004640717
n=2000, 20blks	0.01742018	0.05529185	0.003360591
n=4000, 10blks	0.02282441	0.04585495	0.002623527
n=5000, 8blks	0.02529751	0.04359693	0.002540529

Table A.42: Standard deviation, Bias and Mean squared error of estimators from survival function of $\bar{F} = \frac{1}{(1+x-\theta)^{\alpha}}$ for true value of $\alpha = 0.75$

α =0.75, k=5000, n=40000, 5000 iterations, $\bar{F} = \frac{1}{(1+x-\theta)^{\alpha}}$, $x > \theta$			
Statistics	Sd	Bias	MSE
Hill's estimator	0.013645	0.130848	0.017307
Pickands est.	0.01855	0.00019	0.000344
New estimator	0.075324	-0.02864	0.006493
De haan & Resnick	0.097172	-0.01008	0.009542
n=40,1000blks	0.006797103	0.2051315	0.04212512
n=100,400blks	0.009151039	0.1335256	0.0179128
n=200, 200blks	0.01188914	0.1029798	0.01074617
n=500, 80blks	0.01679767	0.0788601	0.006501021
n=1000, 40blks	0.02129532	0.06722411	0.00497248
n=2000, 20blks	0.02758686	0.05880392	0.004218783
n=4000, 10blks	0.03643546	0.05321159	0.00415875
n=5000, 8blks	0.04007629	0.05251352	0.004363458

Table A.43: Standard deviation, Bias and Mean squared error of estimators from survival function of $\bar{F} = \frac{1}{(1+x-\theta)^{\alpha}}$ for true value of $\alpha = 1$

α =1, k=5000, n=40000, 5000 iterations, $\bar{F} = \frac{1}{(1+x-\theta)^{\alpha}}$, $x > \theta$			
Statistics	Sd	Bias	MSE
Hill's estimator	0.023891	0.392972	0.154997
Pickands est.	0.030346	0.000764	0.000921
New estimator	0.100247	-0.03912	0.011578
De haan & Resnick	0.14051	0.028451	0.020549
n=40,1000blks	0.009993234	0.2000776	0.04013087
n=100,400blks	0.01335903	0.1403838	0.01988605
n=200, 200blks	0.0167828	0.1160613	0.01375183
n=500, 80blks	0.02320665	0.09576368	0.009709124
n=1000, 40blks	0.02967404	0.08422288	0.007973866
n=2000, 20blks	0.03741284	0.07615468	0.007198976
n=4000, 10blks	0.0490683	0.06938247	0.007221144
n=5000, 8blks	0.05355237	0.06712313	0.007372797

Table A.44: Standard deviation, Bias and Mean squared error of estimators from survival function of $\bar{F} = \frac{1}{(1+x-\theta)^{\alpha}}$ for true value of $\alpha = 2$

α =2, k=5000, n=40000, 5000 iterations, $\bar{F} = \frac{1}{(1+x-\theta)^{\alpha}}$, $x > \theta$			
Statistics	Sd	Bias	MSE
Hill's estimator	0.092715	2.985625	8.922549
Pickands est.	0.110944	0.006172	0.012344
New estimator	0.196705	-0.08832	0.046486
De haan & Resnick	0.45329	0.646902	0.623913
n=40,1000blks	0.02183738	0.3216499	0.1039354
n=100,400blks	0.02722008	0.2538086	0.06515958
n=200, 200blks	0.03424072	0.2187795	0.04903667
n=500, 80blks	0.04585899	0.1859612	0.03668421
n=1000, 40blks	0.05967194	0.1671103	0.03148586
n=2000, 20blks	0.07473086	0.1528975	0.02896122
n=4000, 10blks	0.09923653	0.1391613	0.02921179
n=5000, 8blks	0.1052021	0.1364321	0.029679

Table A.45: Standard deviation, Bias and Mean squared error of estimators from survival function of $\bar{F} = \frac{1}{(1+x-\theta)^{\alpha}}$ for true value of $\alpha = 3$

α =3, k=5000, n=40000, 5000 iterations, $\bar{F} = \frac{1}{(1+x-\theta)^{\alpha}}, x > \theta$			
Statistics	Sd	Bias	MSE
Hill's estimator	0.19193	7.392638	54.68792
Pickands est.	0.245621	0.015068	0.060545
New estimator	0.242429	-0.25291	0.122722
De haan & Resnick	1.033	2.24955	6.127349
n=40,1000blks	0.03260829	0.4796574	0.2311344
n=100,400blks	0.0415323	0.3793168	0.1456058
n=200, 200blks	0.05112637	0.3276319	0.1099561
n=500, 80blks	0.06882232	0.2780589	0.08205232
n=1000, 40blks	0.08815056	0.2505875	0.07056308
n=2000, 20blks	0.1137798	0.2276353	0.0647611
n=4000, 10blks	0.1484035	0.2092291	0.06579602
n=5000, 8blks	0.1623122	0.2053091	0.06849181

Table A.46: Standard deviation, Bias and Mean squared error of estimators from survival function of $\bar{F} = \frac{1}{(1+x-\theta)^{\alpha}}$ for true value of $\alpha = 5$

α =5, k=5000, n=40000, 5000 iterations, $\bar{F} = \frac{1}{(1+x-\theta)^{\alpha}}, x > \theta$			
Statistics	Sd	Bias	MSE
Hill's estimator	0.404037	18.7133	350.351
Pickands est.	0.703154	0.097947	0.50392
New estimator	0.193586	-1.20623	1.492457
De haan & Resnick	2.709914	8.520775	79.94578
n=40,1000blks	0.0538579	0.7985147	0.6405258
n=100,400blks	0.07017635	0.6307157	0.402726
n=200, 200blks	0.08515138	0.5466608	0.3060873
n=500, 80blks	0.1149072	0.465747	0.2301213
n=1000, 40blks	0.145221	0.4193986	0.1969801
n=2000, 20blks	0.1885228	0.3758026	0.1767613
n=4000, 10blks	0.2468368	0.3499019	0.1833476
n=5000, 8blks	0.2692236	0.3335713	0.1837367

Table A.47: Standard deviation, Bias and Mean squared error of estimators from survival function of $\bar{F} = \frac{1}{(1+x-\theta)^{\alpha}}$ for true value of $\alpha = 7$

α =7, k=5000, n=40000, 5000 iterations, $\bar{F} = \frac{1}{(1+x-\theta)^{\alpha}}, x > \theta$			
Statistics	Sd	Bias	MSE
Hill's estimator	0.619676	31.38167	985.1933
Pickands est.	1.496374	0.247897	2.300139
New estimator	0.134542	-2.73267	7.485574
De haan & Resnick	4.69563	17.40603	325.0145
n=40,1000blks	0.07764942	1.120586	1.261741
n=100,400blks	0.09849588	0.8842685	0.7916303
n=200, 200blks	0.1169323	0.7659327	0.6003233
n=500, 80blks	0.1618841	0.6513114	0.4504077
n=1000, 40blks	0.2038136	0.586217	0.385182
n=2000, 20blks	0.2658311	0.5339154	0.3557177
n=4000, 10blks	0.3365086	0.4937448	0.3569993
n=5000, 8blks	0.372606	0.478281	0.3675601

Table A.48: Standard deviation, Bias and Mean squared error of estimators from survival function of $\bar{F} = \frac{1}{(1+x-\theta)^{\alpha}}$ for true value of $\alpha = 10$

α =10, k=5000, n=40000, 5000 iterations, $\bar{F} = \frac{1}{(1+x-\theta)^{\alpha}}, x > \theta$			
Statistics	Sd	Bias	MSE
Hill's estimator	0.94593	51.30462	2633.058
Pickands est.	4.00982	0.881264	16.85206
New estimator	0.081024	-5.41025	29.27734
De haan & Resnick	7.869851	33.81811	1205.587
n=40,1000blks	0.1094805	1.600368	2.57316
n=100,400blks	0.1378932	1.26112	1.609435
n=200, 200blks	0.1731845	1.094652	1.22825
n=500, 80blks	0.2339506	0.9281397	0.9161652
n=1000, 40blks	0.2958544	0.8393098	0.7919533
n=2000, 20blks	0.3786369	0.7598368	0.7206893
n=4000, 10blks	0.490598	0.6940294	0.7223151
n=5000, 8blks	0.5368291	0.6822754	0.7536276

Table A.49: Standard deviation, Bias and Mean squared error of estimators from survival function of $\bar{F} = \frac{1}{[1+c(x-\theta)]^{\alpha}}, c = 2, \theta = 5, x > \theta$ for true value of $\alpha = 0.5$

α =0.5, k=5000, n=40000, 5000 iterations, $\bar{F} = \frac{1}{[1+c(x-\theta)]^{\alpha}}, c = 2, \theta = 5, x > \theta$			
Statistics	Sd	Bias	MSE
Hill's estimator	0.007654	0.022909	0.000583
Pickands est.	0.009842	0.000105	9.68E-05
New estimator	0.055509	-0.00462	0.003102
De haan & Resnick	0.064123	-0.01989	0.004506
n=40,1000blks	0.003194655	0.1008403	0.01017897
n=100,400blks	0.004660769	0.03406787	0.001182338
n=200, 200blks	0.006286719	0.004806984	6.262203e-05
n=500, 80blks	0.009203174	-0.01747658	0.0003901122
n=1000, 40blks	0.01267552	-0.0267696	0.000877248
n=2000, 20blks	0.01658092	-0.03242022	0.001325943
n=4000, 10blks	0.02198588	-0.03511162	0.001716108
n=5000, 8blks	0.02473935	-0.03481688	0.001824128

Table A.50: Standard deviation, Bias and Mean squared error of estimators from survival function of $\bar{F} = \frac{1}{[1+c(x-\theta)]^{\alpha}}, c = 2, \theta = 5, x > \theta$ for true value of $\alpha = 0.75$

α =0.75, k=5000, n=40000, 5000 iterations, $\bar{F} = \frac{1}{[1+c(x-\theta)]^{\alpha}}, c = 2, \theta = 5, x > \theta$			
Statistics	Sd	Bias	MSE
Hill's estimator	0.014856	0.164975	0.027437
Pickands est.	0.018758	0.000199	0.000352
New estimator	0.084433	0.006174	0.007166
De haan & Resnick	0.100294	-0.00613	0.010094
n=40,1000blks	0.00647512	0.03637488	0.001365051
n=100,400blks	0.008923533	-0.007926859	0.0001424486
n=200, 200blks	0.01163372	-0.02229308	0.0006322978
n=500, 80blks	0.01652112	-0.0297116	0.001155672
n=1000, 40blks	0.02185479	-0.03144547	0.001466354
n=2000, 20blks	0.02740294	-0.03050318	0.001681215
n=4000, 10blks	0.03602945	-0.02932349	0.002157728
n=5000, 8blks	0.03947938	-0.02912784	0.002406741

Table A.51: Standard deviation, Bias and Mean squared error of estimators from survival function of $\bar{F} = \frac{1}{[1+c(x-\theta)]^{\alpha}}$, $c = 2, \theta = 5, x > \theta$ for true value of $\alpha = 1$

α =1, k=5000, n=40000, 5000 iterations, $\bar{F} = \frac{1}{[1+c(x-\theta)]^{\alpha}}$, $c = 2, \theta = 5, x > \theta$			
Statistics	Sd	Bias	MSE
Hill's estimator	0.02605	0.492741	0.243472
Pickands est.	0.030742	0.001931	0.000949
New estimator	0.115158	0.023654	0.013818
De haan & Resnick	0.146313	0.041483	0.023124
n=40,1000blks	0.009696335	0.02227045	0.0005899732
n=100,400blks	0.01325967	-0.006463801	0.0002175645
n=200, 200blks	0.01655129	-0.01296875	0.0004420789
n=500, 80blks	0.02309595	-0.01548508	0.0007731041
n=1000, 40blks	0.02959901	-0.01553582	0.001117288
n=2000, 20blks	0.03811381	-0.01469597	0.001668343
n=4000, 10blks	0.04939047	-0.01212413	0.002585925
n=5000, 8blks	0.05395253	-0.01314631	0.003083118

Table A.52: Standard deviation, Bias and Mean squared error of estimators from survival function of $\bar{F} = \frac{1}{[1+c(x-\theta)]^{\alpha}}$, $c = 2, \theta = 5, x > \theta$ for true value of $\alpha = 2$

α =2, k=5000, n=40000, 5000 iterations, $\bar{F} = \frac{1}{[1+c(x-\theta)]^{\alpha}}$, $c = 2, \theta = 5, x > \theta$			
Statistics	Sd	Bias	MSE
Hill's estimator	0.114203	3.785928	14.34629
Pickands est.	0.111127	0.007023	0.012396
New estimator	0.254889	0.181591	0.097931
De haan & Resnick	0.505937	0.799374	0.89492
n=40,1000blks	0.02168271	0.1349181	0.01867294
n=100,400blks	0.02815573	0.1028053	0.01136151
n=200, 200blks	0.03479732	0.08755143	0.008875864
n=500, 80blks	0.04650336	0.075078	0.007798835
n=1000, 40blks	0.0584126	0.06706987	0.007909717
n=2000, 20blks	0.07447703	0.05912982	0.009042054
n=4000, 10blks	0.09825269	0.05857767	0.013083
n=5000, 8blks	0.1079379	0.05420849	0.01458682

Table A.53: Standard deviation, Bias and Mean squared error of estimators from survival function of $\bar{F} = \frac{1}{[1+c(x-\theta)]^{\alpha}}$, $c = 2, \theta = 5, x > \theta$ for true value of $\alpha = 3$

α =3, k=5000, n=40000, 5000 iterations, $\bar{F} = \frac{1}{[1+c(x-\theta)]^{\alpha}}$, $c = 2, \theta = 5, x > \theta$			
Statistics	Sd	Bias	MSE
Hill's estimator	0.234065	9.437967	89.13
Pickands est.	0.244323	0.01697	0.05997
New estimator	0.331579	0.307235	0.204316
De haan & Resnick	1.203472	2.816332	9.379783
n=40,1000blks	0.03219889	0.2922663	0.08645614
n=100,400blks	0.04275406	0.2276833	0.05366722
n=200, 200blks	0.05030808	0.1962209	0.04103303
n=500, 80blks	0.06916485	0.1686867	0.03323802
n=1000, 40blks	0.08938557	0.1510256	0.03079693
n=2000, 20blks	0.1145911	0.1379939	0.03217081
n=4000, 10blks	0.1479548	0.1273676	0.03810875
n=5000, 8blks	0.1590826	0.1231564	0.04046971

Table A.54: Standard deviation, Bias and Mean squared error of estimators from survival function of $\bar{F} = \frac{1}{[1+c(x-\theta)]^{\alpha}}$, $c = 2, \theta = 5, x > \theta$ for true value of $\alpha = 5$

α =5, k=5000, n=40000, 5000 iterations, $\bar{F} = \frac{1}{[1+c(x-\theta)]^{\alpha}}$, $c = 2, \theta = 5, x > \theta$			
Statistics	Sd	Bias	MSE
Hill's estimator	0.503987	24.00829	576.6521
Pickands est.	0.701981	0.091224	0.501
New estimator	0.269985	-0.2129	0.118202
De haan & Resnick	3.270214	10.9159	129.849
n=40,1000blks	0.05549214	0.6125445	0.3782895
n=100,400blks	0.06847292	0.4817731	0.2367929
n=200, 200blks	0.0851942	0.4152388	0.1796798
n=500, 80blks	0.1146755	0.3524383	0.1373606
n=1000, 40blks	0.147393	0.3194035	0.1237389
n=2000, 20blks	0.1873472	0.2918233	0.1202528
n=4000, 10blks	0.2444365	0.2645508	0.1297244
n=5000, 8blks	0.2664808	0.2565561	0.1368188

Table A.55: Standard deviation, Bias and Mean squared error of estimators from survival function of $\bar{F} = \frac{1}{[1+c(x-\theta)]^{\alpha}}$, $c = 2, \theta = 5, x > \theta$ for true value of $\alpha = 7$

α =7, k=5000, n=40000, 5000 iterations, $\bar{F} = \frac{1}{[1+c(x-\theta)]^{\alpha}}$, $c = 2, \theta = 5, x > \theta$			
Statistics	Sd	Bias	MSE
Hill's estimator	0.76656	40.29976	1624.658
Pickands est.	1.486896	0.277177	2.287244
New estimator	0.185814	-1.54511	2.42187
De haan & Resnick	5.798514	22.52088	540.8062
n=40,1000blks	0.07694896	0.9318923	0.8743433
n=100,400blks	0.09617021	0.7374939	0.5531442
n=200, 200blks	0.1213006	0.6342712	0.4170108
n=500, 80blks	0.1626008	0.536206	0.3139506
n=1000, 40blks	0.2023034	0.4827879	0.2740026
n=2000, 20blks	0.2629798	0.4391793	0.262023
n=4000, 10blks	0.3388937	0.3884696	0.2657346
n=5000, 8blks	0.3645092	0.3975266	0.2908678

Table A.56: Standard deviation, Bias and Mean squared error of estimators from survival function of $\bar{F} = \frac{1}{[1+c(x-\theta)]^{\alpha}}$, $c = 2, \theta = 5, x > \theta$ for true value of $\alpha = 10$

α =10, k=5000, n=40000, 5000 iterations, $\bar{F} = \frac{1}{[1+c(x-\theta)]^{\alpha}}$, $c = 2, \theta = 5, x > \theta$			
Statistics	Sd	Bias	MSE
Hill's estimator	1.16069	65.88528	4342.218
Pickands est.	17.66205	1.064788	313.0194
New estimator	0.112666	-4.10511	16.86459
De haan & Resnick	9.553479	43.26414	1963.037
n=40,1000blks	0.1093612	1.409759	1.999377
n=100,400blks	0.1365757	1.113008	1.257435
n=200, 200blks	0.1705809	0.9601782	0.9510343
n=500, 80blks	0.2319675	0.8178184	0.7226251
n=1000, 40blks	0.2962669	0.7358197	0.6291871
n=2000, 20blks	0.3775196	0.6690641	0.5901392
n=4000, 10blks	0.4931979	0.6223421	0.6305052
n=5000, 8blks	0.5352447	0.58932	0.6337277