# Maximum lq–likelihood estimator of the heavy–tailed distribution parameter

Mohammed Ridha Kouider[1,*], Nesrine Idiou [2], Samia Toumi [3] and Fatah Benatia[4]

[1,3,4] Mohamed Khider University of Biskra, Faculty of Exact Science and Natural and Life Science, Algeria

[2] Salah Boubnider University of Constantine 3, Faculty of Process Engineering, Algeria

**SUMMARY**

Studying the extreme value theory (EVT) involves multiple main objectives, among them the estimation of the tail index parameter. Some estimation methods are used to estimate the tail index parameter like maximum likelihood estimation (MLE). Additionally, the Hill estimator is one type of maximum likelihood estimator, which is a more robust with a large sample than a small sample. This research proposes the construction of an alternative estimator for the parameter of the heavy–tailed distribution using the maximum lq–likelihood estimation (MLqE) approach in order to adapt the ML and Hill estimator with the small sample. Furthermore, the maximum lq–likelihood estimator asymptotic normality is established. Moreover, several simulation studies in order to compare the MLq estimator with the ML estimators are provided. In the excesses over high suitable threshold values the number of the largest observation $k$ will lead to an efficient estimate of the Hill estimator. For this, selection of $k$ in the Hill estimator was investigated using the method of the quantile type 8 which is effective with the hydrology data. The performance of the Hill estimator and the lq–Hill estimator is subsequently compared by employing real relies with the distribution of hydrology data.

ARTICLE TYPE
**Preliminary communication**

ARTICLE INFO
Received: September 7, 2024
Accepted: November 5, 2024
DOI: 10.62366/crebss.2024.2.003
JEL: C13, C46

**KEYWORDS**

*excesses over threshold, extreme value index, heavy–tailed distribution, maximum lq–likelihood estimator*

## 1. Introduction

Considering $X_1, X_2, \ldots, X_n$ of independently and identically distributed (iid) random variables (rv) defined over some probability space $(\Omega; A; P)$, with cumulative distribution function (cdf) $F$. We are interested in the probability that the maximum is not beyond a certain threshold $x$. This probability is given by

$$P\left(\max\left(X_1, X_2, \ldots, X_n\right) \leq x\right) = F^n\left(x\right). \tag{1}$$

*Corresponding author

As it is well known, when we are interested in the central part of a sample, the central limit theorem (CLT) giving the asymptotic law of the sum of observations which says that the sampling distribution of the mean will always be normally distributed as $n \to +\infty$ . On the other hand, if we want to study the extreme values of this sample, the CLT presents only little of interest. Instead, we use a result establishing the asymptotic distribution of the maximum of the sample. This result is stated under EVT as demonstrated particularly by Gnedenko (1943). The EVT gives the conditions under which there exist sequences of normalizing constants $a_n > 0$ and $b_n > 0$ such that

$$\lim_{n \to +\infty} F^n (a_n x + b_n) = G_\gamma (x) . \tag{2}$$

$G_\gamma (x)$ is so–called the extreme value distribution, defined by

$$G_\gamma (x) = \begin{cases} \exp\left( - (1 + \gamma x)^{-1/\gamma} \right), & \text{if } \gamma \neq 0 \\ \exp(-\exp(-x)), & \text{if } \gamma = 0 \end{cases} \tag{3}$$

where $G_\gamma (x)$ is a well–defined non–degenerate cdf. This law depends only on the parameter $\gamma \in \mathbb{R}$ called the extreme value index or the tail index, or the shape parameter. According to the sign of $\gamma$, there are three domains of attraction that are defined from $G_\gamma (x)$ depending on the tail index; among them is domain attraction of Fréchet. Also, it is referred to as heavy–tailed distribution. It contains laws whose survival function decreases as a power function like distributions of Pareto, Student, Cauchy, etc. So the heavy tailed limit distribution is the Fréchet distribution (Balkema and de Haan, 1974), which is defined by

$$G_{\gamma > 0} (x) = e^{-x^{-1/\gamma}}. \tag{4}$$

As it is known and associated with EVT, the characterization of the domains of attraction makes extensive use of the notion of functions with regular variations which we define below.

Let $X_1, X_2, \ldots, X_n$ be a iid sequence of a non–negative rv $X$ over some probability space $(\Omega; A; P)$, with cdf $F$. We assume that the distribution tail $\overline{F} = 1 - F$ is regularly varying at infinity, with index $(-1/\gamma)$, notation: $\overline{F} \in RV_{(-1/\gamma)}$. That is

$$\lim_{t \to +\infty} \frac{\overline{F}(tx)}{\overline{F}(t)} = x^{-1/\gamma}, \text{ for any } x > 0. \tag{5}$$

A distribution function $F$ belongs to the domain attraction of Fréchet if and only if $\overline{F} \in RV_{(-1/\gamma)}$. Hence, the tail behaves approximately as a power function $x^{-1/\gamma}$. This implies that the distribution for the maximum has a one–to–one relationship with the shape parameter. Then, we will take heavy tailed to mean sub–exponential (definition below), but several definitions exist in the literature. For the convenience of the reader, we also sketch other common definitions and, where possible, relate them to one another. Sub-exponential distributions exhibit one of the general properties expected of heavy–tailed distributions on the level of aggregate losses, namely that the tail of the maximum determines the tail of the sum. All of the distributions considered here are sub–exponential.

Let $F$ be a cdf with support in $[0; +\infty[$. Then $F$ is sub–exponential if, for all $n \geq 2$,

$$\lim_{x \to +\infty} \frac{\overline{F}^n (x)}{\overline{F}(x)} = n, \text{ for any } x > 0. \tag{6}$$

Then $\overline{F}^n(x) \approx n\overline{F}(x)$ as $x \to +\infty$. Subexponentiality implies another property that is sometimes taken as the definition of heavy tail, i.e. the tail decays more slowly than any exponential function. With the notation as above, the precise formulation is that for all $t > 0$,

$$\lim_{t\to+\infty} e^{tx}\overline{F}(x) := \infty. \tag{7}$$

An important subclass of sub–exponential distributions consists of regularly varying functions. For a regularly varying with tail index $\gamma > 0$ , all moments of the associated rv higher than $\gamma > 0$ will be unbounded (Embrechts et al., 1997).

The parameter of interest is $\gamma > 0$ is the tail index of $F$. Now for $\overline{F} \in RV_{(-1/\gamma)}$ we take as $\overline{F}(x) = x^{-1/\gamma}$. Then we can check for $F$ is sub–exponential that

$$\lim_{t\to+\infty}\left(\lim_{x\to+\infty}\frac{1}{\overline{F}^n(t)}\int_t^\infty \frac{\overline{F}^n(x)}{x}dx\right) = \frac{1}{\overline{F}(t)}\int_t^\infty \frac{\overline{F}(x)}{x}dx := \gamma. \tag{8}$$

Consider $X_{1,n} \le X_{2,n} \le \ldots \le X_{n,n}$ the order statistics of $X_1, X_2, \ldots, X_n$. Let's replace the distribution $\overline{F}$ by its empirical version $\overline{F}_n$ and $t$ by $X_{n-k,n}$. Thus, we find the Hill estimator (Hill, 1975) defined by:

$$\widehat{\gamma}^H_{X_{n-k,n}} := \frac{1}{\overline{F}_n(X_{n-k,n})}\int_{X_{n-k,n}}^\infty \overline{F}_n(x)\frac{dx}{x}. \tag{9}$$

The Hill estimator can only be used for distributions belonging to the Fréchet domain. The Hill estimator, which is a type of ML estimator, is the most common estimators for the tail index of heavy–tailed distributions; is probably the most studied estimator in the literature. As agreed that, Hill estimator and ML estimator are goods with the large sample is susceptible to be biased. But if we use very small $k$, both estimators have a large variance.

In this article, we investigate a new class of parametric estimators based on the q–entropy function proposed by Havrda and Charvát (1967). It has been of considerable interest in different domains of application like physics, finance and biomedical sciences. As well, Altun and Smola (2006) have seen that the classical maximum entropy is dual of MLE. Ditto, Ferrari and Yang (2010) proposed a new parametric estimation method based on the q–entropy function, the MLqE where $q$ is called the distortion parameter. Also, they have proven to be a very useful method when estimating high–dimensional parameters and small tail probabilities. This is important in many applications where the number of available observations is not great. They have shown that MLqE becomes the MLE with $q = 1$.

Since for large sample the ML and Hill estimators are at least as precise as any other estimators. However, for a moderate or small sample size the MLq estimator can offer dramatic improvement in mean squared error at the expense of a slight increase in bias.

This paper has been organized as follows. In Section (1) it is presented as an introduction the asymptotic distribution of the maximum of the sample under the EVT and especially the heavy–tailed distribution or the Fréchet distribution. We also gave a definition about sub–exponential. As for later, we will focus on presenting the new estimate about the tail index of heavy–tailed distributions, i.e. MLq estimator. Next, with Section (2), we present the basic asymptotic normality of MLq estimator with their consistent for exponential families which we introduce in the same section. In Section (3), we present a simulation study with Pareto distribution to compare the MLqE with MLE. Also, the real data are utilized to illustrate the usefulness of the Fréchet distribution as the distribution of hydrology data. Finally, concluding notes are provided in Section (4).

### 1.1. Adaptive ML and Hill estimators for small sample

In the domain of heavy–tailed distributions, the statistic of EVT translates into a semi-parametric estimation problem. Indeed, if $F$ belongs to the Fréchet domain, then $\overline{F}$ is of the form $x^{-1/\gamma}\ell(x)$ with $\ell$ a slowly varying function i.e, $\lim_{t\to+\infty} \ell(tx)/\ell(x) := 1$. This paper concentrates on the distributions that have a regularly varying tail,

$$\frac{\overline{F}(x)}{x^{-1/\gamma}\ell(x)} := 1, \text{ as } x \to +\infty, \gamma > 0 \tag{10}$$

Note that $G_{\gamma>0}(x)$ satisfies (10). Here $(1/\gamma)$ is the index of regular variation, or the tail index. Then $\overline{F}$ has a parametric part $x^{-1/\gamma}$ depending only on $\gamma$, and a non–parametric part $\ell$. For a real $t > 0$ it's clear that we deduce

$$\overline{F}(x) := \overline{F}(t)\left(\frac{x}{t}\right)^{-1/\gamma} \tag{11}$$

Let $X_1, X_2, \ldots, X_n$ be a sequence of iid rv from distribution function (df) $F$ and let $X_{1,n} \leq X_{2,n} \leq \ldots \leq X_{n,n}$ denote the order statistics correspondence. We denoting the number of absolute excesses over $t$ by $k$ for rather the largest observations $(X_{n-k,n}, \ldots, X_{n,n})$ where in the asymptotic setting $k = k_n$ an intermediate sequence, that is, $k_n \to \infty$ and $k_n/n \to 0$ as $n \to \infty$. Then, see Haan and Ferreira (2006) Lemma 3.4.1, the joint distribution of $(X_{n-k,n}, \ldots, X_{n,n})$ for $k = 1, \ldots, n-1$ be the df given by

$$F_t(x) = P(X \leq x | X > t) = \frac{F(x) - F(t)}{1 - F(t)} \text{ for } x > t \tag{12}$$

Such that $\overline{F}_t(x) = 1 - F_t(x)$ we can rewrite (12) for $x > t$ as

$$\overline{F}_t(x) = \frac{\overline{F}(x)}{\overline{F}(t)} \tag{13}$$

Since $\overline{F}(x)$ is given by (11), also we can rewrite $\overline{F}_t(x)$ given in (13) for $x > t$ by

$$\overline{F}_t(x) = \left(\frac{x}{t}\right)^{-1/\gamma} \tag{14}$$

It's easy to checked that under (5) that $\overline{F}_t \in RV_{(-1/\gamma)}$.

The parameter of the $\overline{F}_t(x)$ can be estimated using standard methods such as the MLE. Another estimation method is the MLqE which based on q–order entropy. The q–order entropy which is provided by Havrda and Charvát (1967), has the function

$$L_q(u) = \begin{cases} \frac{u^{1-q}-1}{1-q} & \text{for } q < 1 \\ \log u & \text{for } q = 1 \end{cases} \tag{15}$$

where $u$ is probability density function(pdf) and $q$ called the distortion parameter. The MLqE, which is proposed by Ferrari and Yang (2010), use the $L_q(u)$ function instead of the log–likelihood function as in the MLE. Observed that when $q = 1$, we have the MLq estimator approaches as the ML estimator approaches. The MLqE method reduces the effect of extreme observations on parameter estimates using $q$. The choice of $q$ is another difficult problem in

MLqE estimation. In this research, we take $q = 1 - \frac{1}{k}$ as given by Ferrari and Yang (2010) and we note that $q \to 1$ as $k \to \infty$. Then, the MLqE of $\gamma > 0$ is given by

$$\widehat{\gamma} = \arg\max \sum_{i=1}^{k} L_q\left(f_t\left(x\right)\right) \tag{16}$$

with

$$L_q\left(f_t\left(x\right)\right) = \begin{cases} \frac{f_t(x)^{1-q}-1}{1-q} & \text{for } 0 < q < 1 \\ \log u & \text{for } q = 1 \end{cases} \tag{17}$$

where $f_t\left(x\right)$ is the pdf of $\overline{F}_t\left(x\right)$ given in (14). Thus defense for $\gamma > 0$ by

$$f_t\left(x\right) = \frac{1}{\gamma}\frac{1}{t}\left(\frac{x}{t}\right)^{-1/\gamma-1} \tag{18}$$

As is known, MLq estimator of $\gamma > 0$ can be obtained by maximizing $\sum_{i=1}^{k} L_q\left(f_t\left(x\right)\right)$ with respect to the parameter $\gamma > 0$. Then, for $0 < q \leq 1$ we get

$$\frac{\partial}{\partial\gamma}L_q\left(f_t\left(x\right)\right) := \frac{\partial}{\partial\gamma}\log\left(f_t\left(x\right)\right)f_t\left(x\right)^{1-q} \tag{19}$$

Then, the equations from (19) are then given in term of the partial derivative respect to $\gamma$ by:

$$\frac{\partial}{\partial\gamma}\sum_{i=1}^{k} L_q\left(f_t\left(X_i\right)\right) := \sum_{i=1}^{k}\frac{1}{\gamma^2}\left(\log\left(X_i\right) - \log\left(t\right) - \gamma\right)f_t\left(X_i\right)^{1-q} = 0 \tag{20}$$

Next, we can define the estimator of $\gamma > 0$ by

$$\widehat{\gamma}_t^{MLq} = \frac{\sum\limits_{i=1}^{k} w_i\left(\log\left(X_i\right) - \log\left(t\right)\right)}{\sum\limits_{i=1}^{k} w_i} \text{ with } w_i = f_t\left(X_i\right)^{1-q} \tag{21}$$

There more, if we take $t = X_{n-k,n}$ we find a new Hill estimator called lq–Hill estimator which is defined by

$$\widehat{\gamma}_{X_{n-k,n}}^{HLq} = \frac{\sum\limits_{i=1}^{k} w_{n-i+1,n}\left(\log\left(X_{n-i+1,n}\right) - \log\left(X_{n-k,n}\right)\right)}{\sum\limits_{i=1}^{k} w_{n-i+1,n}} \text{ with } w_{n-i+1,n} = f_t\left(X_{n-i+1,n}\right)^{1-q} \tag{22}$$

It is clear that if $q = 1$ gives us the classic the MLE. We get ML estimator of $\gamma > 0$ which is given by:

$$\widehat{\gamma}_t^{ML} = \frac{1}{k}\sum_{i=1}^{k}\log\left(X_i\right) - \log\left(t\right) \tag{23}$$

And by their equivalents in the order statistic, we find the formula of the Hill estimator with $t = X_{n-k,n}$ by

$$\widehat{\gamma}_{X_{n-k,n}}^{H} = \frac{1}{k}\sum_{i=1}^{k}\log\left(X_{n-i+1,n}\right) - \log\left(X_{n-k,n}\right) \tag{24}$$

Also, the Hill estimator was found to be very sensitive to the choice of index $k$. Because choosing the optimal value for the $k$ index will lead to an effective estimate of Hill estimator, the choice of $k$ is another difficult problem and there are many researches on this. However, in this article, we will present in sub–section (3.2) a method that allows calculating the number $k$ based on the method of the quantile type 8 with the hydrology data. It is important to mention that Hill estimator is a consistent for the tail index and asymptotically normal with mean $\gamma$ and variance $\gamma^2/k$. Hence, if one uses a very small $k$, the estimator has large variance, however, for very large $k$, the estimator is likely to be biased i.e asymptotically normal with mean 0 and variance $\gamma^2$. This is why it is good with large sample also, ML estimator is effected with large sample. In this research, we have written this tow estimator with the MLqE. We can show that for $q = 1 - \frac{1}{k}$ we rewrite the estimators $\widehat{\gamma}_t^{MLq}$ and $\widehat{\gamma}_{X_{n-k,n}}^{HLq}$ as

$$\widehat{\gamma}_t^{MLq} = \frac{\sum\limits_{i=1}^{k} f_t\left(X_i\right)^{1/k}\left(\log\left(X_i\right) - \log\left(t\right)\right)}{\sum\limits_{i=1}^{k} f_t\left(X_i\right)^{1/k}} \tag{25}$$

and

$$\widehat{\gamma}_{X_{n-k,n}}^{HLq} = \frac{\sum\limits_{i=1}^{k} f_t\left(X_{n-i+1,n}\right)^{1/k}\left(\log\left(X_{n-i+1,n}\right) - \log\left(X_{n-k,n}\right)\right)}{\sum\limits_{i=1}^{k} f_t\left(X_{n-i+1,n}\right)^{1/k}} \tag{26}$$

Recall that when $q$ is chosen correctly for small samples, the MLqE can trade bias for accuracy successfully. This leads to a significant decrease in the mean squared error. There are more for large sample if $(k \to \infty)$ $q = 1 - \frac{1}{k}$ or $q \to 1$ we focus on a necessary and sufficient condition, to ensure a proper asymptotic normality and efficiency of MLqE is established. This is what will be discussed in the next section.

## 2. Main result

In this section, we discuss the basic asymptotic properties of the MLq estimator when the degree of distortion depends on the amount of information available in the sample. Such properties will be used later on to derive our main results. In the reminder of the paper, our analysis focuses on the distributions belonging to the exponential family. In particular, we consider pdf of $\overline{F}_t\left(x\right)$ given on (22) in the form

$$f_t\left(x\right) = \exp\left\{\frac{1}{\gamma}b\left(x\right) - c\left(x\right) - A\left(\gamma\right)\right\} \tag{27}$$

For

$$b\left(x\right) = \log\left(\frac{x}{t}\right) \text{ and } c\left(x\right) = \log\left(x\right),\, A\left(\gamma\right) = \log\left(\gamma\right)$$

In addition, we define $\psi_\gamma\left(x\right) = \frac{1}{\gamma}b\left(x\right) - c\left(x\right) - A\left(\gamma\right)$. So that $f_t\left(x\right)$ and $\log f_t\left(x\right)$ with its derivative function respect to $\gamma$ expressed respectively as

$$f_t\left(x\right) = \exp\psi_\gamma\left(x\right) \text{ and } \frac{\partial}{\partial\gamma}f_t\left(x\right) = \left(-\frac{1}{\gamma^2}b\left(x\right) - \frac{1}{\gamma}\right)\exp\psi_\gamma\left(x\right), \tag{29}$$

and

$$\log f_t(x) = \psi_\gamma(x) \text{ and } \frac{\partial}{\partial\gamma}\log f_t(x) = \left(-\frac{1}{\gamma^2}b(x) - \frac{1}{\gamma}\right). \tag{30}$$

Until, for $\gamma > 0$ we can check that

$$A(\gamma) = \log \int_t^{+\infty} \exp\left(\frac{1}{\gamma}b(x) - c(x)\right)dx = \log\gamma, \tag{31}$$

is the cumulative generating function (or log normalize) and differentiating $k$ times gives

$$\frac{\partial^k}{\partial^k\gamma}A(\gamma) = \frac{(-1)^{k-1}(k-1)!}{\gamma^k}, \tag{32}$$

Throughout the course of the discussion the true parameter will be denoted by $\gamma > 0$ . Next, we explore consistency, which is a basic requirement for a good estimator. Let, for $0 < q \leq 1$

$$\varphi_k(\gamma) = \frac{1}{k}\sum_{i=1}^{k}\frac{\partial}{\partial\gamma}L_q(f_t(X_i)) := \frac{1}{k}\sum_{i=1}^{k}\left(-\frac{1}{\gamma^2}b(x) - \frac{1}{\gamma}\right)e^{(1-q)\left(\frac{1}{\gamma}b(x)-c(x)-A(\gamma)\right)} \tag{33}$$

The MLqE is found by setting $\varphi_k(\gamma) = 0$ and solving for $\gamma$. Since for $q = 1$ and $\gamma \neq 0$ in the above expression gives the usual MLE equation

$$\frac{1}{k}\sum_{i=1}^{k}\left(-\frac{1}{\widehat{\gamma}^{MLq}}b(x) - 1\right) = 0 \tag{34}$$

**Theorem 2.1.** *Let $X_1, X_2, \ldots, X_n$ be a sequence of iid rv from df $F$ and let $X_{1,n} \leq X_{2,n} \leq \ldots \leq X_{n,n}$ denote the order statistics correspondence. Considering the largest observations $(X_{n-k,n}, \ldots, X_{n,n})$ for $k = 1, \ldots, n-1$ from the df $\overline{F}_t(x)$ and pdf $f_t(x)$ as (27). Then, for any MLq estimator of $\widehat{\gamma} = \arg\max\sum_{i=1}^{k}L_q(f_t(x))$ as $k \to \infty$ we have that*

$$\widehat{\gamma} \xrightarrow{P} \gamma.$$

*with $L_q(f_t(x))$ is given in (15) and $q \to 1$ as $k \to \infty$.*

*Proof.* Define, $\varphi(\gamma) = E_\gamma\left[\frac{\partial}{\partial\gamma}\log f_t(x)\right]$. Then we can rewrite $\varphi(\gamma) = E_\gamma\left[-\frac{1}{\gamma^2}b(x) - \frac{1}{\gamma}\right]$. Now, we want to show that for all $\varphi_k - \varphi \to 0$ as $k \to \infty$ where

$$\varphi_k - \varphi = \frac{1}{k}\sum_{i=1}^{k}\left(-\frac{1}{\gamma^2}b(x) - \frac{1}{\gamma}\right)e^{(1-q)\psi_\gamma(x)} - E_\gamma\left[-\frac{1}{\gamma^2}b(x) - \frac{1}{\gamma}\right] \tag{35}$$

Then we have

$$|\varphi_k - \varphi| = \left|\frac{1}{k}\sum_{i=1}^{k}\left(-\frac{1}{\gamma^2}b(x) - \frac{1}{\gamma}\right)\left(e^{(1-q)\psi_\gamma(x)} - 1\right) + \frac{1}{k}\sum_{i=1}^{k}\left(-\frac{1}{\gamma^2}b(x) - \frac{1}{\gamma}\right) - E_\gamma\left[-\frac{1}{\gamma^2}b(x) - \frac{1}{\gamma}\right]\right|$$

thus

$$\leq \left|\frac{1}{k}\sum_{i=1}^{k}\left(-\frac{1}{\gamma^2}b(x) - \frac{1}{\gamma}\right)\left(e^{(1-q)\psi_\gamma(x)} - 1\right)\right| + \left|\frac{1}{k}\sum_{i=1}^{k}\left(-\frac{1}{\gamma^2}b(x) - \frac{1}{\gamma}\right) - E_\gamma\left[-\frac{1}{\gamma^2}b(x) - \frac{1}{\gamma}\right]\right| \tag{36}$$

By the law of large numbers we find

$$\left|\frac{1}{k}\sum_{i=1}^{k}\left(-\frac{1}{\gamma^2}b\left(x\right)-\frac{1}{\gamma}\right)-E_\gamma\left[-\frac{1}{\gamma^2}b\left(x\right)-\frac{1}{\gamma}\right]\right|\to 0 \tag{37}$$

Then the inequality (35) becomes

$$|\varphi_k-\varphi|=\left|\frac{1}{k}\sum_{i=1}^{k}\left(-\frac{1}{\gamma^2}b\left(x\right)-\frac{1}{\gamma}\right)\left(e^{(1-q)\psi_\gamma(x)}-1\right)\right| \tag{38}$$

By the Hölder's inequality, we can rewrite (38) as follows

$$|\varphi_k-\varphi|\leq\sqrt{\frac{1}{k}\sum_{i=1}^{k}\left(e^{(1-q)\psi_\gamma(x)}-1\right)^2}\sqrt{\frac{1}{k}\sum_{i=1}^{k}\left(-\frac{1}{\gamma^2}b\left(x\right)-\frac{1}{\gamma}\right)^2} \tag{39}$$

And under Jensen's inequality we have

$$|\varphi_k-\varphi|\leq\frac{1}{4}\frac{1}{k}\sum_{i=1}^{k}\left(e^{(1-q)\psi_\gamma(x)}-1\right)^2\frac{1}{k}\sum_{i=1}^{k}\left(-\frac{1}{\gamma^2}b\left(x\right)-\frac{1}{\gamma}\right)^2 \tag{40}$$

For the left side from the previous inequality and by the basic fact that $(1+u)^2\leq e^{2u}$ for any real number $u$ we rewrite

$$\frac{1}{\gamma^2 k}\sum_{i=1}^{k}\left(-\frac{1}{\gamma}b\left(x\right)-1\right)^2\leq\frac{1}{\gamma^2 k}\sum_{i=1}^{k}e^{\frac{2}{\gamma}b(x)}=\frac{1}{\gamma^2 k}\sum_{i=1}^{k}\left(\frac{x}{t}\right)^{-\frac{2}{\gamma}} \tag{41}$$

Then we have

$$\frac{1}{\gamma^2 k}\sum_{i=1}^{k}e^{\frac{2}{\gamma}b(x)}=\frac{1}{\gamma^2 k}\sum_{i=1}^{k}\overline{F}_t\left(x\right)^2<\infty \tag{42}$$

And for the right side of the inequality that under the law of large numbers we get

$$\frac{1}{k}\sum_{i=1}^{k}\left(e^{(1-q)\psi_\gamma(x)}-1\right)^2\to E_\gamma\left(e^{(1-q)\psi_\gamma(x)}-1\right)^2 \tag{43}$$

With $q\to 1$ as $k\to\infty$ we have

$$E_\gamma\left(e^{(1-q)\psi_\gamma(x)}-1\right)^2\to 0$$

Finally we obtained that $\varphi_k\to\varphi$ as $k\to\infty$. □

As defined in theorem (2.1), the MLqE is a consistent estimator of $\gamma$. Although for a fixed $q\neq 1$. The MLqE is clearly asymptotically biased; a clear improvement is obtained by letting the distortion parameter depends on the sample size. To obtain the asymptotic normality of MLqE, we shall discuss the reduction in terms of variance achieved by considering a slightly different target parameter for each $\gamma>0$ and $k\geq 1$.

On particular, we consider $\widehat{\gamma}^{MLq}$ the value such that

$$E\left[\frac{\partial L_q\left(f_t\left(x\right)\right)}{\partial\widehat{\gamma}^{MLq}}\right]=0 \tag{44}$$

First, under the formula (27) can be represent $E\left[\frac{\partial L_q(f_t(x))}{\partial \widehat{\gamma}^{MLq}}\right]$ as,

$$E\left[\frac{\partial L_q\left(f_t\left(x\right)\right)}{\partial \widehat{\gamma}^{MLq}}\right] = \int_t^{\infty} \frac{\partial}{\partial \widehat{\gamma}^{MLq}} L_q\left(f_t\left(x\right)\right) f_t\left(x\right) dx \tag{45}$$

then

$$E\left[\frac{\partial L_q\left(f_t\left(x\right)\right)}{\partial \widehat{\gamma}^{MLq}}\right] = \frac{1}{\gamma} t^{\frac{1}{\gamma}} \left(\widehat{\gamma}^{MLq}\right)^{q-3} t^{\frac{1-q}{\widehat{\gamma}^{MLq}}} \int_t^{\infty} \left(\log\left(X_i\right) - \log\left(t\right) - \widehat{\gamma}^{MLq}\right) e^{-\left[\widehat{\alpha}(1-q)+\alpha\right]\log x} dx \tag{46}$$

with

$$\alpha = 1 + \frac{1}{\gamma} \text{ and } \widehat{\alpha} = 1 + \frac{1}{\widehat{\gamma}^{MLq}} \tag{47}$$

After performing simple arithmetic operations, we find

$$E\left[\frac{\partial L_q\left(f_t\left(x\right)\right)}{\partial \widehat{\gamma}^{MLq}}\right] = \frac{1}{\gamma} t^{\frac{1}{\gamma}} \left(\widehat{\gamma}^{MLq}\right)^{q-3} t^{\frac{1-q}{\widehat{\gamma}^{MLq}}} \left(\frac{1}{\widehat{\theta}^2} - \frac{1}{\widehat{\theta}}\right) e^{-\widehat{\theta}\log x} \tag{48}$$

where $\widehat{\theta} = \widehat{\alpha}\left(1-q\right) + \alpha - 1$

**Proposition 2.1.** *By solving the equation* $E\left[\frac{\partial L_q(f_t(x))}{\partial \widehat{\gamma}^{MLq}}\right] = 0$, *we find these results shown as follows*

*1)* $\widehat{\alpha}\left(1-q\right) + \alpha - 1 = \frac{1}{\widehat{\gamma}^{MLq}}$.

*2)* $\alpha = \widehat{\alpha} q$.

*3)* $\widehat{\gamma}^{MLq} = \frac{q}{\alpha - q}$ *or* $\widehat{\gamma}^{MLq} = \frac{1}{\widehat{\alpha}-1}$. *And each of the previous values satisfies* $E\left[\frac{\partial L_q(f_t(x))}{\partial \widehat{\gamma}^{MLq}}\right] = 0$.

In particular, $\widehat{\gamma}^{MLq}$ arbitrarily close to the true parameter $\gamma$ depending on the value of the distortion parameter .Since, for $q = 1$ we can get that $\alpha = \widehat{\alpha}$ and $\widehat{\gamma}^{MLq} = \gamma$.

**Theorem 2.2.** *Let* $X_1, X_2, \ldots, X_n$ *be a sequence of iid rv from df* $F$ *and let* $X_{1,n} \leq X_{2,n} \leq \ldots \leq X_{n,n}$ *denote the order statistics correspondence. Considering the largest observations* $\left(X_{n-k,n}, \ldots, X_{n,n}\right)$ *for* $k = 1, \ldots, n-1$ *from the df* $\overline{F}_t\left(x\right)$ *and pdf* $f_t\left(x\right)$ *as (27). Then, for* $\widehat{\gamma}^{MLq}$ *the MLq estimator of* $\widehat{\gamma} = \arg\max \sum_{i=1}^{k} L_q\left(f_t\left(x\right)\right)$ *if* $q \to 1$ *as* $k \to \infty$ *are satisfied, then we have as* $k \to \infty$ *that*

$$\sqrt{k}\left(\widehat{\gamma}^{MLq} - \widetilde{\gamma}\right) \to N\left(0; \frac{E\left[\left(\frac{\partial L_q(f_t(x))}{\partial \widetilde{\gamma}}\right)^2\right]}{\left(E\left[\left(\frac{\partial^2 L_q(f_t(x))}{\partial^2 \widetilde{\gamma}}\right)\right]\right)^2}\right).$$

*Where* $\widetilde{\gamma} = \frac{q}{\alpha - q}$, $\alpha = 1 + \frac{1}{\gamma}$ *with* $L_q\left(f_t\left(x\right)\right)$ *is given in (17) and* $N$ *is the normal distribution.*

*Proof.* Under (33) for $0 < q \leq 1$ we have

$$\varphi_k\left(\gamma\right) = \frac{1}{k}\sum_{i=1}^{k} \frac{\partial}{\partial \gamma} L_q\left(f_t\left(X_i\right)\right) := \frac{1}{k}\sum_{i=1}^{k}\left(-\frac{1}{\gamma^2} b\left(x\right) - \frac{1}{\gamma}\right) e^{(1-q)\left(\frac{1}{\gamma}b(x) - c(x) - A(\gamma)\right)} \tag{49}$$

Then by Taylor's theorem we have

$$\frac{\partial \varphi_k\left(\widehat{\gamma}^{MLq}\right)}{\partial \widetilde{\gamma}} \approx \frac{\partial \varphi_k\left(\widetilde{\gamma}\right)}{\partial \widetilde{\gamma}} + \left(\widehat{\gamma}^{MLq} - \widetilde{\gamma}\right)\frac{\partial^2 \varphi_k\left(\widetilde{\gamma}\right)}{\partial^2 \widetilde{\gamma}} \tag{50}$$

Since $\frac{\partial \varphi_k\left(\widehat{\gamma}^{MLq}\right)}{\partial \widetilde{\gamma}} = 0$. Then

$$\left(\widehat{\gamma}^{MLq} - \widetilde{\gamma}\right) = \frac{\frac{1}{\sqrt{k}} \frac{\partial \varphi_k(\widetilde{\gamma})}{\partial \widetilde{\gamma}}}{-\frac{1}{k} \frac{\partial^2 \varphi_k(\widetilde{\gamma})}{\partial^2 \widetilde{\gamma}}} \tag{51}$$

Since $E\left[\frac{\partial L_q(f_t(x))}{\partial \widetilde{\gamma}}\right] = 0$ and $S_i = \frac{\partial}{\partial \widetilde{\gamma}} L_q\left(f_t\left(x\right)\right)$. Then

$$\frac{1}{\sqrt{k}} \sum_{i=1}^{k} S_i = \sqrt{k}\left(\overline{S} - 0\right). \tag{52}$$

By the central limit theorem, we find that

$$\sqrt{k}\left(\overline{S} - 0\right) \to N\left(0; \sigma_a\right) \text{ with } \sigma_a = E\left[\left(\frac{\partial L_q\left(f_t\left(x\right)\right)}{\partial \widetilde{\gamma}}\right)^2\right] \tag{53}$$

Let $R_i = -\frac{\partial^2 L_q(f_t(x))}{\partial^2 \widetilde{\gamma}}$ , then $E\left[\overline{R}\right] = -E\left[\frac{\partial^2 L_q(f_t(x))}{\partial^2 \widetilde{\gamma}}\right]$ and

$$-\frac{1}{k} \sum_{i=1}^{k} \frac{\partial^2 \varphi_k\left(\widetilde{\gamma}\right)}{\partial^2 \widetilde{\gamma}} = \frac{1}{k} \sum_{i=1}^{k} R_i = \overline{R} \tag{54}$$

Hence, under the law of large numbers, we have $\overline{R} \to E\left[\frac{\partial^2 L_q(f_t(x))}{\partial^2 \widetilde{\gamma}}\right]$. Then, apply Slutsky's theorem we get

$$\sqrt{k}\left(\widehat{\gamma}^{MLq} - \widetilde{\gamma}\right) \to N\left(0; \frac{E\left[\left(\frac{\partial L_q(f_t(x))}{\partial \widetilde{\gamma}}\right)^2\right]}{\left(E\left[\left(\frac{\partial^2 L_q(f_t(x))}{\partial^2 \widetilde{\gamma}}\right)\right]\right)^2}\right). \tag{55}$$

□

**Corollary 2.1.** *Under the assumptions of theorem (2.2), we have*

$$\sqrt{k}\left(\widehat{\gamma}^{MLq} - \widetilde{\gamma}\right) \to N\left(0; \sigma^2\right),$$

*where N is the normal distribution, with* $\widetilde{\gamma} = \frac{q}{\alpha - q}$ $\widetilde{\alpha} = 1 + \frac{1}{\widetilde{\gamma}}$ *with* $q = 1 - \frac{1}{k}$, *and* $\sigma^2$ *equal*

$$\sigma^2 = \frac{1}{\left(q\widetilde{\alpha} - 1\right)} \frac{.}{1 + \left(\left(\widetilde{\gamma} + 1\right)\left(1 - q\right)\right)^2} \left(\widetilde{\alpha}\left(2 - q\right) - 1\right)^3 \left(\left(1 - q\right)\left(1 - 2\widetilde{\gamma}\right) + \widetilde{\gamma}\right)^2.$$

**Corollary 2.2.** *Under the assumptions of corollary (2.1), for* $q = 1$ *we have*

$$\sqrt{k}\left(\widehat{\gamma}^{MLq} - \widetilde{\gamma}\right) \to N\left(0; \gamma^2\right),$$

*where N is the normal distribution.*

## 3. Numerical examples

### 3.1. Simulation study

We checked that the upper tail of a distribution given by (14) is based on the exceedances observations $(X_{n-k,n}, \ldots, X_{n,n})$ for $k = 1, \ldots, n-1$ over a threshold $t > 0$. Then one basic statistical model is that of Pareto distribution function which is defined as

$$H_\gamma(x) = 1 - \left(\frac{x}{t}\right)^{-\frac{1}{\gamma}}, \text{ with } x \geq t \tag{56}$$

or $\frac{x}{t} \geq 1$ with $0 < \gamma < 1$ is called the tail index. When $t = 1$ the cdf $H_\gamma$ in (56) becomes a continuous power law distribution, also known as heavy-tailed distributions, and it's easy to check that $\overline{H} \in RV_{(-1/\gamma)}$. In this section, we perform a simulation study by using selected $H_\gamma$ the Pareto distribution. We generate data sets from the Pareto distribution using the following parameter values $\gamma = 0.7$ with $t = 1$ when the sample size $n = 15$ under the following steps:

1) We use inverse of cumulative distribution function method of $H_{0.7}$ with $t = 1$ to generate the data.

2) Then, we use the pdf from the Pareto distribution of $H_{0.7}$ with $t = 1$.

3) Calculate the estimators $\widehat{\gamma}_{t=1}^{MLq}$ and $\widehat{\gamma}_{t=1}^{ML}$ as

$$\widehat{\gamma}_{t=1}^{MLq} = \frac{1}{r}\sum_{i=1}^{n}\left(h_{0.7}(X_i)\right)^{1/n}\log(X_i) \text{ and } \widehat{\gamma}_{t=1}^{ML} = \frac{1}{n}\sum_{i=1}^{n}\log(X_i), \tag{57}$$

with $r = \sum_{i=1}^{n}\left(h_{0.7}(X_i)\right)^{1/n}$ and $h_{0.7}$ is pdf associated to $H_{0.7}$.

The data are collected from the threshold $t = 1$ in generate data sets from $H_{0.7}$ for a random value $(X_1, X_2, \ldots, X_{15})$ of 15 random are listed in an increasing order in Table 1.

Table 1. *Simulated data ($n = 15$) from Pareto distribution $H_{0.7}$ with $t = 1$*

| | | | | |
|---|---|---|---|---|
| 1.00 | 1.06 | 1.06 | 1.11 | 1.20 |
| 1.26 | 1.57 | 1.87 | 2.06 | 2.33 |
| 2.54 | 2.62 | 4.66 | 7.08 | 40.03 |

We focus on this first part of simulation to compare between two estimators, that is $\widehat{\gamma}_{t=1}^{MLq}$ and $\widehat{\gamma}_{t=1}^{ML}$ . Then , after doing the calculations we find $\widehat{\gamma}_{t=1}^{MLq} = 0.7$ and $\widehat{\gamma}_{t=1}^{ML} = 0.825$ . And we can explicitly note that $\widehat{\gamma}_{t=1}^{MLq} = 0.7$ is robust then $\widehat{\gamma}_{t=1}^{ML} = 0.825$ with this 15 generate data.

In the next parts we generate data sets from Pareto df $H_{0.7}$ with $t = 1$ with various samples "the small sample where for $n = \{8; 10; 25\}$, the moderate sample for $n = \{50; 100\}$ , the largest sample for $n = \{200; 500\}$ "and all results are calculated by averaging over 1000 simulation runs. Also the RMSE of $\gamma$ for $\widehat{\gamma}_{t=1}^{MLq}$ and $\widehat{\gamma}_{t=1}^{ML}$ are calculated in each sample respectively as

$$\text{RMSE} = \frac{1}{1000}\sqrt{\sum_{i=1}^{1000}\left(\gamma - \widehat{\gamma}_{t=1}^{MLq}\right)^2} \text{ and RMSE} = \frac{1}{1000}\sqrt{\sum_{i=1}^{1000}\left(\gamma - \widehat{\gamma}_{t=1}^{ML}\right)^2} \tag{58}$$

By using the MLqE method, the results are summarized in Table 2.

Table 2. *The RMSE results in estimating the tail index with ML and MLq estimators*

| $\gamma = 0.7$ | $n$ | $r$ | Estimator | $\widehat{\gamma}_{t=1}$ | RMSE($\gamma$) |
|---|---|---|---|---|---|
| Small sample | 8 | 6.55 | ML | 0.922 | 0.2311998 |
| | | | MLq | 0.701 | 0.0011944 |
| | 10 | 8.61 | ML | 0.831 | 0.1437046 |
| | | | MLq | 0.705 | 0.0059549 |
| | 25 | 23.6 | ML | 0.742 | 0.0487582 |
| | | | MLq | 0.696 | 0.0047946 |
| Moderate sample | 50 | 48.6 | ML | 0.732 | 0.0374020 |
| | | | MLq | 0.706 | 0.0071408 |
| | 100 | 98.6 | ML | 0.715 | 0.0177394 |
| | | | MLq | 0.702 | 0.0023871 |
| Large sample | 200 | 199 | ML | 0.71 | 0.0118678 |
| | | | MLq | 0.703 | 0.0035780 |
| | 500 | 499 | ML | 0.702 | 0.0023871 |
| | | | MLq | 0.699 | 0.0011961 |

We note that the RMSE of $\widehat{\gamma}_{t=1}^{MLq}$ are lower than $\widehat{\gamma}_{t=1}^{ML}$. Since $r = \sum_{i=1}^{n} \left(h_{0.7}\left(X_i\right)\right)^{1/n}$ is also estimated. We notice that $\widehat{\gamma}_{t=1}^{MLq}$ is almost equal to the estimate for large samples. As well, we can find theoretically that when $n \to \infty$ we have $r = n$. Precisely, under the inequalities (21) and (23) and via the Hölder's inequality we obtain that,

$$\widehat{\gamma}_t^{MLq} \leq \sqrt{\sum_{i=1}^{n} \left(h_t\left(X_i\right)\right)^{1/n} \widehat{\gamma}_t^{ML}} \tag{59}$$

Since $\sqrt{x} \leq x$ for any $0 \leq x \leq 1$. Then, when $n \to \infty$ we have $\widehat{\gamma}_t^{MLq} \leq \widehat{\gamma}_t^{ML}$ implies that $\sqrt{k}\left(\widehat{\gamma}_t^{MLq} - \gamma\right) \leq \sqrt{k}\left(\widehat{\gamma}_t^{ML} - \gamma\right)$ for $0 < \gamma < 1$. These results indicate that MLqE gives the tailed behavior estimator reliable.

## 3.2. Illustrative example

In this part, we use real data set related to minimum monthly flows of water ($m3/s$) on the Piracicaba River, located in São Paulo state, Brazil. The data obtained from the Water Resources Department and Director of the Energy Agency for Water Resources in the state São Paulo from May to September from 1960 to 2014. Ramos et al. (2018) were proposed methodology can be used successfully to analyze the minimum flow of water during. They demonstrated that the data from the study on the Piracicaba River follows the Fréchet distribution. However, in our study, we will only use data related to the month of August. The minimum water flows in August ($m3/s$) consists of 41 data, listed in an increasing order in Table 3.

Table 3. *The minimum water flows pf Piracicaba River in August from 1960 to 2014*

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| 6.80 | 7.11 | 7.72 | 8.24 | 8.52 | 8.78 | 8.93 | 9.43 | 9.52 |
| 9.69 | 9.80 | 10.00 | 10.00 | 10.40 | 10.96 | 11.22 | 11.63 | 11.84 |
| 12.01 | 12.46 | 12.46 | 12.98 | 13.10 | 13.74 | 14.29 | 15.23 | 15.51 |
| 16.00 | 16.58 | 17.10 | 18.26 | 18.39 | 20.01 | 26.99 | 28.06 | 28.97 |
| 30.20 | 37.86 | 51.43 | 53.72 | 58.98 | | | | |

In this study we will focus on estimation the tail index by the Hill estimator and the lq–Hill estimator. Hence, the number of order statistics $k$ is important for estimating the Hill's estimators. Therefore, we use the method of the quantile type 8 or $\widehat{Q}_8(p)$ for selecting $k$ which was efficient method for an estimate of the shape parameter of hydrology data. Let $X_1, X_2, \ldots, X_n$ are iid rv from a cdf and $X_{1,n} > X_{2,n} > \cdots > X_{n,n}$ the order statistics correspondence, Boonradsamee et al. (2021) used a new method for selecting $k$ in Hill's estimator by the quantile type 8 from the stable region of Hill plot which depends on the order statistics for heavy tailed distributions. They said that they could choose the optimal $k$ of the stable region of Hill plot by $Q_8(p)$ method with $p = 0.75$. The quantile type 8 is the theoretical of being median unbiased, and we can write as

$$Q_8(p) = X_j + \alpha\left(X_{j+1} - X_j\right) \text{ with } \frac{j-m}{n} \le p \le \frac{j-m+1}{n} \text{ for } m = \frac{p+1}{3} \tag{60}$$

where $0 < p < 1$ with $\alpha = \mathbf{1}\left(\frac{j}{n} < p\right)$ and $j = [np + m]$ (Here, $[u]$ denotes the largest integer not greater than $u$). Hyndman and Fan (1996) investigated their motivation and some of their properties. And they presented the estimator of $Q_8(p)$ which is given in (60) as $\widehat{Q}_8(p)$ which is the median position and by using an approximation to the incomplete beta function ratio $\beta(k, n-k+1)$, for $X_{1,n} > X_{2,n} > \cdots > X_{n,n}$ we get

$$\widehat{Q}_8(p) \approx \frac{p_k}{\beta(X_k)} \text{ with } p_k = \frac{3k-1}{3n+1} \tag{61}$$

where $k$ represents the rank of $X_{k,n}$. Hence, the sample quantiles are not defined for all $p$ including hinges and other letter values. Therefore, we select the plotting position in a quantile-quantile plotting which $X_k$ is plotted against $Q(p)$ the inverse of empirical distribution function where $Q(p_k) = X_k$ where $0 \le p_k \le 1$ with $p_k$ is the empirical distribution function. However, if we take the order statistics as $X_{1,n} < X_{2,n} < \cdots < X_{n,n}$ it is clear that from (61) with $s = n - k$ we can take that

$$s = n - \frac{\widetilde{p}+1}{3} \text{ with } \widetilde{p} = (3n+1)\, p_k \text{ and } \frac{\widetilde{p}+1}{3} = k \tag{62}$$

where $n-k$ represents the rank of $X_{n-k,n}$. Hence, in this study with the order statistics $X_{1,n} < X_{2,n} < \cdots < X_{n,n}$ if we choose $p_k = 1 - p_s = 0.7$ with $p_s = 0.3$ and under (62) we find $k = 29$. For $n = 41$ we have $k/n = 0.7 = p_k$ and we find $m = 0.5$ with $j = k$ and $\alpha = 0$, thus under (60) we get $\widehat{Q}_8(q_s) = X_{12,41}$. Then, $k = 29$ is the number of the largest observations $(X_{n-k+1,n}, \ldots, X_{n,n})$ given $X_j > X_{n-k,n}$ for $j = 1, \ldots, k$.

For this we will take the exceedances $(X_{13,41}, \ldots, X_{41,41})$ over a threshold $t = X_{12,41} = 10$. There is $k = 29$ observations of $C-$sample with

$$C_0 = X_{12,41}; C_1 = X_{13,41}; C_2 = X_{14,41}; \ldots; C_{29} = X_{41,41}$$

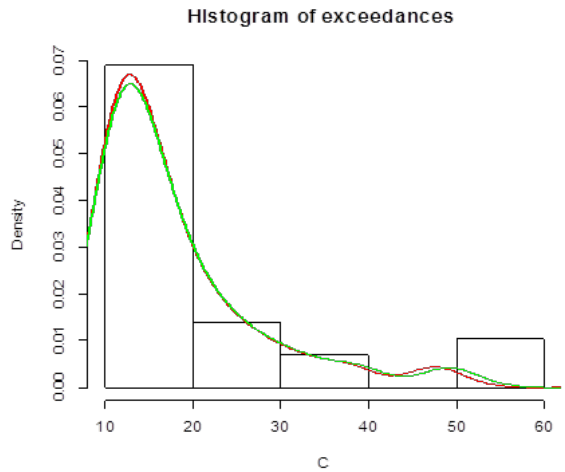


Figure 1. *Histogram of C–sample with the pdf for* $\widehat{\gamma}_{C_0}^{HLq}$ *and* $\widehat{\gamma}_{C_0}^{H}$

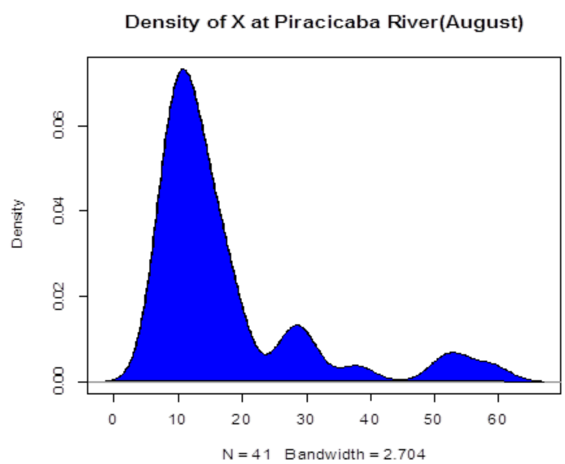


Figure 2. *Pdf of X-sample from the Piracicaba River (m3/s) in August from 1960 to 2014*

The histogram of $C-$sample for $i = 1, \ldots, k$ shows a moderate deviation to the right (Figure 1). Therefore, heavy–tailed distributions are plausible to modeling this data. We estimate the unknown tail index of the heavy–tailed distribution for $C-$sample. We use the lq–Hill estimator and Hill estimator which are given in (22) and (24) respectively. The results are given by $\widehat{\gamma}_{C_0}^{HLq} = 0.583649$ and $\widehat{\gamma}_{C_0}^{H} = 0.5961899$, respectively. Also, the fitted densities obtained from the lq–Hill estimator and Hill estimator are shown in Figure 1. The red color is the density obtained from $\widehat{\gamma}_{C_0}^{HLq}$ and the green color is the density obtained from $\widehat{\gamma}_{C_0}^{H}$. We got them via inverse cdf of $C-$sample which is given in (14) as the following form: $C_i = t \times (p_i)^{\gamma_t^*}$ where $t = 10$ and $p_i$ is the empirical cdf according the $C-$sample for $i = 1, \ldots, 29$.

We notice that the value of $\widehat{\gamma}_{C_0}^{HLq}$ and $\widehat{\gamma}_{C_0}^{H}$ are very close to each other. Hence from Figure 1, we observed that the best fit is obtained from the MLq method is closer to pdf of $X-$sample from the Piracicaba River (m3/s) in August (Figure 2) .

## 4. Concluding notes

Heavy–tailed distributions would be a good alternative to the distributions that are used in economics, reliability, survival analysis and so on. The parameters of this distribution have been estimated using MLE method. Recently, ML and Hill estimators are the most estimators used to estimate the parameter of the tail behavior. For too large sample, ML and Hill estimators are likely to be good and robust methods should be used to estimate the shape parameter. However, ML and Hill estimators cannot be compatible with the small sample. In this paper, we have used the MLq estimation method to estimate the shape parameter of the heavy–tailed distribution for small sample. We have carried out to adapt the ML and Hill estimators in the case of small sample, $\widehat{\gamma}_t^{MLq}$ and $\widehat{\gamma}_{X_{n-k,n}}^{HLq}$ as in (25) and (26), respectively. Since $\widehat{\gamma}_t^{MLq}$ and $\widehat{\gamma}_{X_{n-k,n}}^{HLq}$ are both result from the MLqE, we establish the asymptotic normality of the MLq estimator of Heavy tailed distributions parameter $\gamma > 0$. According to corollary (2.1) for very large $k$, the MLq estimator is likely to be biased. And if $q = 1$, the MLqE becomes MLE and the normality asymptotic of MLq estimator becomes as ML estimator. Since the MLq estimator depends on the order statistical index $k$ for heavy tailed distributions, we used an approach for selecting $k$ which is defined in (62) by using type 8 quantile estimator $Q_8(0.7)$ in (60) from the stable region of Hill plot, which will be a more flexible alternative for use in Hill's estimator. This approach will be an approximation of the extreme value index at the tail end of the distribution of hydrology data.

## Appendix

First, we calculate $E\left[\left(\frac{\partial L_q(f_t(x))}{\partial\widetilde{\gamma}}\right)^2\right]$. Let's consider

$$E\left[\left(\frac{\partial L_q\left(f_t\left(x\right)\right)}{\partial\widetilde{\gamma}}\right)^2\right] = \frac{1}{\gamma\widetilde{\gamma}^6} t^{\frac{1}{\gamma}+\frac{2}{\widetilde{\gamma}}} \widetilde{\gamma}^{2q} t^{\frac{-2q}{\widetilde{\gamma}}} \int_t^{\infty} \left(\log\left(X_i\right) - \log\left(t\right) - \widetilde{\gamma}\right)^2 e^{-\left[2\widetilde{\alpha}(1-q)+\alpha\right]\log x} dx \quad (66)$$

with $\widetilde{\alpha} = 1 + \frac{1}{\widetilde{\gamma}}$. Putting $\log\left(x\right) = u$ , then the previous equality (66) becomes

$$E\left[\left(\frac{\partial L_q\left(f_t\left(x\right)\right)}{\partial\widetilde{\gamma}}\right)^2\right] = \beta \int_{\log t}^{\infty} \left(u^2 - 2u\left(\log t + \widetilde{\gamma}\right) + \left(\log t + \widetilde{\gamma}\right)^2\right) e^{-\left[2\widetilde{\alpha}(1-q)+\alpha-1\right]u} du \quad (67)$$

where $\beta = \frac{1}{\gamma\widetilde{\gamma}^6} t^{\frac{1}{\gamma}+\frac{2}{\widetilde{\gamma}}} \widetilde{\gamma}^{2q} t^{\frac{-2q}{\widetilde{\gamma}}}$ is constant with $\alpha = 1 + \frac{1}{\gamma}$. Then, for $\widetilde{\gamma} = q/\left(\alpha - q\right)$ we have $q\widetilde{\alpha} = \alpha$ with $\widetilde{\gamma} = 1/\left(\widetilde{\alpha} - 1\right)$. For $\widetilde{\theta} = \widetilde{\alpha}\left(q-1\right) + \alpha - 1 = \widetilde{\alpha} - 1$then $2\widetilde{\alpha}\left(1-q\right) + \alpha - 1 = \widetilde{\alpha}\left(2-q\right) - 1$.

Consider this decomposition

$$D\left(t\right) = \int_{\log t}^{\infty} \left(u^2 - 2u\left(\log t + \widetilde{\gamma}\right) + \left(\log t + \widetilde{\gamma}\right)^2\right) e^{-\left[\widetilde{\alpha}(2-q)-1\right]u} du = I_1 - I_2 + I_3 \quad (68)$$

where

$$I_1 = \int_{\log t}^{\infty} u^2 e^{-\left[\widetilde{\alpha}(2-q)-1\right]u} du,$$

$$I_2 = \int_{\log t}^{\infty} 2u\left(\log t + \widetilde{\gamma}\right) e^{-\left[\widetilde{\alpha}(2-q)-1\right]u} du$$

And

$$I_3 = \int_{\log t}^{\infty} \left(\log t + \widetilde{\gamma}\right)^2 e^{-\left[\widetilde{\alpha}(2-q)-1\right]u} du$$

Beginning with $I_1 = \int_{\log t}^{\infty} u^2 e^{-\left[\widetilde{\alpha}(2-q)-1\right]u} du$, under integral by parts we have

$$I_1 = \int_{\log t}^{\infty} u^2 e^{-\left[\widetilde{\alpha}(2-q)-1\right]u} du = \frac{-1}{\widetilde{\alpha}\left(2-q\right)-1}\left(\left[u^2 e^{-\left[\widetilde{\alpha}(2-q)-1\right]u}\right]_{\log t}^{\infty} - 2\int_{\log t}^{\infty} u e^{-\left[\widetilde{\alpha}(2-q)-1\right]u} du\right) \tag{69}$$

Also with integral by parts we have

$$\int_{\log t}^{\infty} u e^{-\left[\widetilde{\alpha}(2-q)-1\right]u} du = \frac{-1}{\widetilde{\alpha}\left(2-q\right)-1}\left(\left[u e^{-\left[\widetilde{\alpha}(2-q)-1\right]u} + \frac{e^{-\left[\widetilde{\alpha}(2-q)-1\right]u}}{\widetilde{\alpha}\left(2-q\right)-1}\right]_{\log t}^{\infty}\right) \tag{70}$$

Substituting (26) into $I_1$ and since $\widetilde{\alpha}\left(2-q\right)-1>0$ we get

$$I_1 = \frac{e^{-\left[\widetilde{\alpha}(2-q)-1\right]\log t}}{\widetilde{\alpha}\left(2-q\right)-1}\left(\left(\log t\right)^2 + \frac{2}{\widetilde{\alpha}\left(2-q\right)-1}\left(\log t + \frac{1}{\widetilde{\alpha}\left(2-q\right)-1}\right)\right) \tag{71}$$

Then we go to $I_2 = \int_{\log t}^{\infty} 2u\left(\log t + \widetilde{\gamma}\right) e^{-\left[\widetilde{\alpha}(2-q)-1\right]u} du$, . By (26) and with $\widetilde{\alpha}\left(2-q\right)-1>0$ we have

$$I_2 = \frac{2\left(\log t + \widetilde{\gamma}\right) e^{-\left[\widetilde{\alpha}(2-q)-1\right]\log t}}{\widetilde{\alpha}\left(2-q\right)-1}\left(\log t + \frac{1}{\widetilde{\alpha}\left(2-q\right)-1}\right) \tag{72}$$

And for $I_3 = \int_{\log t}^{\infty} \left(\log t + \widetilde{\gamma}\right)^2 e^{-\left[\widetilde{\alpha}(2-q)-1\right]u} du$, under-performing simple arithmetic operations, we get

$$I_3 = \left(\log t + \widetilde{\gamma}\right)^2\left[\frac{-e^{-\left[\widetilde{\alpha}(2-q)-1\right]u}}{\widetilde{\alpha}\left(2-q\right)-1}\right]_{\log t}^{\infty} = \frac{\left(\log t + \widetilde{\gamma}\right)^2 e^{-\left[\widetilde{\alpha}(2-q)-1\right]\log t}}{\widetilde{\alpha}\left(2-q\right)-1} \tag{73}$$

Under (71)-(73) and $D\left(t\right) = I_1 - I_2 + I_3$ we have

$$\left(\widetilde{\alpha}\left(2-q\right)-1\right) e^{\left[\widetilde{\alpha}(2-q)-1\right]\log t} D\left(t\right) = \frac{1+\left(\widetilde{\alpha}\widetilde{\gamma}\left(1-q\right)\right)^2}{\left(\widetilde{\alpha}\left(2-q\right)-1\right)^2} \tag{74}$$

Then it gives us that

$$E\left[\left(\frac{\partial L_q\left(f_t\left(x\right)\right)}{\partial\widetilde{\gamma}}\right)^2\right] = \beta\frac{e^{-\left[\widetilde{\alpha}(2-q)-1\right]\log t}}{\widetilde{\alpha}\left(2-q\right)-1}\left(\frac{1+\left(\widetilde{\alpha}\widetilde{\gamma}\left(1-q\right)\right)^2}{\left(\widetilde{\alpha}\left(2-q\right)-1\right)^2}\right) \tag{75}$$

Since

$$\beta = \frac{1}{\gamma\widetilde{\gamma}^6} t^{\frac{1}{\gamma}+\frac{2}{\widetilde{\gamma}}} \widetilde{\gamma}^{2q} t^{\frac{-2q}{\widetilde{\gamma}}}$$

Next, for $\widetilde{\alpha}\widetilde{\gamma} = 1 + \widetilde{\gamma}$ we can find that

$$E\left[\left(\frac{\partial L_q\left(f_t\left(x\right)\right)}{\partial \widetilde{\gamma}}\right)^2\right] = \frac{1}{\gamma} t^{\frac{1}{\gamma}} \frac{e^{2(q-3)\log \widetilde{\gamma} + \left(\frac{2}{\widetilde{\gamma}}(1-q) - (\widetilde{\alpha}(2-q)-1)\right)\log t}}{\widetilde{\alpha}\left(2-q\right)-1} \left(\frac{1+\left(\left(1+\widetilde{\gamma}\right)\left(1-q\right)\right)^2}{\left(\widetilde{\alpha}\left(2-q\right)-1\right)^2}\right)$$

Then

$$E\left[\left(\frac{\partial L_q\left(f_t\left(x\right)\right)}{\partial \widetilde{\gamma}}\right)^2\right] = \frac{1}{\gamma} t^{\frac{1}{\gamma}} e^{2(q-3)\log \widetilde{\gamma} - (q(\widetilde{\alpha}-2)+1)\log t} \left(\frac{1+\left(\left(1+\widetilde{\gamma}\right)\left(1-q\right)\right)^2}{\left(\widetilde{\alpha}\left(2-q\right)-1\right)^3}\right) \tag{76}$$

Next, we have to calculate $E\left[\left(\frac{\partial^2 L_q(f_t(x))}{\partial^2 \widetilde{\gamma}}\right)\right]$. Then, consider

$$E\left[\left(\frac{\partial^2 L_q\left(f_t\left(x\right)\right)}{\partial^2 \widetilde{\gamma}}\right)\right] = \frac{\partial}{\partial \widetilde{\gamma}} \int_t^\infty \frac{\partial L_q\left(f_t\left(x\right)\right)}{\partial \widetilde{\gamma}} f_t\left(x\right) dx \tag{77}$$

Referring to equality (48), we conclude that

$$\begin{aligned} E\left[\left(\frac{\partial^2 L_q\left(f_t\left(x\right)\right)}{\partial^2 \widetilde{\gamma}}\right)\right] &= \frac{\partial}{\partial \widetilde{\gamma}} \left(\frac{1}{\gamma} t^{\frac{1}{\gamma}} \widetilde{\gamma}^{q-3} t^{\frac{1}{\widetilde{\gamma}}(1-q)} \left(\frac{1}{\widetilde{\theta}^2} - \frac{1}{\widetilde{\theta}}\right) e^{-\widetilde{\theta}\log t}\right) \\ &= \frac{\partial}{\partial \widetilde{\gamma}} \left(\frac{1}{\gamma} t^{\frac{1}{\gamma}} \left(\frac{1}{\widetilde{\theta}^2} - \frac{1}{\widetilde{\theta}}\right) e^{\left(-\widetilde{\theta} + \frac{1}{\widetilde{\gamma}}(1-q)\right)\log t + (q-3)\log \widetilde{\gamma}}\right) \end{aligned}$$

where $\widetilde{\theta} = \widetilde{\alpha}\left(1-q\right) + \alpha - 1$ with $\widetilde{\alpha} = 1 + 1/\widetilde{\gamma}$ implies $\partial\widetilde{\theta}/\partial\widetilde{\gamma} = -\left(1-q\right)/\widetilde{\gamma}^2$ and $\partial\widetilde{\alpha}/\partial\widetilde{\gamma} = -1/\widetilde{\gamma}^2$. Since

$$\frac{\partial e^{\left(-\widetilde{\theta} + \frac{1}{\widetilde{\gamma}}(1-q)\right)\log t + (q-3)\log \widetilde{\gamma}}}{\partial \widetilde{\gamma}} = \frac{(q-3)}{\widetilde{\gamma}} e^{\left(-\widetilde{\theta} + \frac{1}{\widetilde{\gamma}}(1-q)\right)\log t + (q-3)\log \widetilde{\gamma}}$$

Therefore, we have

$$E\left[\left(\frac{\partial^2 L_q\left(f_t\left(x\right)\right)}{\partial^2 \widetilde{\gamma}}\right)\right] = \frac{1}{\gamma} t^{\frac{1}{\gamma}} e^{\left(-\widetilde{\theta} + \frac{1}{\widetilde{\gamma}}(1-q)\right)\log t + (q-3)\log \widetilde{\gamma}} \left(\frac{1-q}{\widetilde{\gamma}^2\widetilde{\theta}^2} - \frac{2\left(1-q\right)}{\widetilde{\gamma}^2\widetilde{\theta}^3} + \frac{q-2}{\widetilde{\theta}} - \frac{q-3}{\widetilde{\gamma}\widetilde{\theta}^2}\right)$$

Under $\widetilde{\theta} = \widetilde{\alpha}\left(1-q\right) + \alpha - 1 = 1/\widetilde{\gamma}$. Then, we get

$$E\left[\left(\frac{\partial^2 L_q\left(f_t\left(x\right)\right)}{\partial^2 \widetilde{\gamma}}\right)\right] = \frac{1}{\gamma} t^{\frac{1}{\gamma}} e^{-\frac{1}{\widetilde{\gamma}}\log t + (q-3)\log \widetilde{\gamma}} \left(\left(1-q\right)\left(1-2\widetilde{\gamma}\right) + \widetilde{\gamma}\right) \tag{78}$$

Now, we have to count the variance of $\sqrt{k}\left(\widehat{\gamma}^{MLq} - \widetilde{\gamma}\right)/\sigma \to N\left(0;1\right)$ as $k \to \infty$. With

$$\sigma^2 = \frac{E\left[\left(\frac{\partial L_q(f_t(x))}{\partial \widetilde{\gamma}}\right)^2\right]}{\left(E\left[\left(\frac{\partial^2 L_q(f_t(x))}{\partial^2 \widetilde{\gamma}}\right)\right]\right)^2} \tag{79}$$

And under (76)-(78) we get

$$\sigma^2 = \frac{1}{\gamma} t^{-\frac{1}{\gamma}} e^{2\frac{1}{\widetilde{\gamma}}q\log t - (q(\widetilde{\alpha}-2)+1)\log t} \left(\frac{\left(\frac{1+((1+\widetilde{\gamma})(1-q))^2}{(\widetilde{\alpha}(2-q)-1)^3}\right)}{\left(\left(1-q\right)\left(1-2\widetilde{\gamma}\right) + \widetilde{\gamma}\right)^2}\right) \tag{80}$$

It's easy to find that

$$-\left(q\left(\widetilde{\alpha}-2\right)+1\right)\log t+\frac{2}{\widetilde{\gamma}}q\log t=\frac{1}{\gamma}\log t$$

Since $q\widetilde{\alpha}=\alpha$ we find, $\gamma=1/\left(q\widetilde{\alpha}-1\right)$. Then the variance becomes

$$\sigma^2=\frac{1}{\left(q\widetilde{\alpha}-1\right)}\frac{1+\left(\left(\widetilde{\gamma}+1\right)\left(1-q\right)\right)^2}{\left(\widetilde{\alpha}\left(2-q\right)-1\right)^3\left(\left(1-q\right)\left(1-2\widetilde{\gamma}\right)+\widetilde{\gamma}\right)^2}. \tag{81}$$

with $q=1$ then $\widetilde{\alpha}=\alpha$ and $\widetilde{\gamma}=\gamma$ the variance $\sigma^2$ write as

$$\sigma^2=\frac{1}{\left(\alpha-1\right)}\frac{1}{\left(\alpha-1\right)^3\gamma^2} \tag{82}$$

Since $(\alpha-1)=1/\gamma$ we have $\sigma^2=\gamma^2$.

## References


Altun, Y. & Smola, A. (2006). *Unifying divergence minimization and statistical inference via convex duality*. In: Lugosi, G. & Simon, H. U. (Eds) Learning Theory, Conference Proceedings of the 19th Annual Conference on Learning Theory – COLT 2006, 139–153. Springer: Berlin. doi: 10.1007/11776420_13

Balkema, A. A. & de Haan, L. (1974). Residual life time at great age. *The Annals of Probability*, 2(5):792–804. doi: 10.1214/aop/1176996548

Boonradsamee, J., Bodhisuwan, W. & Jaroengeratikun, U. (2021). *An estimation approach of classical tail index and an application to hydrology assessment*. Proceedings of the World Congress on Engineering – WCE 2021, 19–24. https://www.iaeng.org/publication-WCE2021/WCE2021_pp19-24.pdf

Davison, A. C. & Smith, R. L. (1990). Models for exceedances over high thresholds. *Journal of the Royal Statistical Society*, 52(3):393–425. doi: 10.1111/j.2517-6161.1990.tb01796.x

Drees, H., Ferreira, A. & den Haan, L. (2004). On maximum likelihood estimation of the extreme value index. *The Annals of Applied Probability*, 14(3):1179–1201. doi: 10.1214/105051604000000279

Embrechts, P., Klüppelberg, C. & Mikosch, T. (1997). *Modelling extremal events for insurance and finance*. Springer: New York. doi: 10.1007/978-3-642-33483-2

Ferrari, D. & Yang, Y. (2010). Maximum lq–likelihood estimation. *The Annals of Statistics*, 38(2):753–783. doi: 10.1214/09-AOS687

Fisher, R. A. & Tippett, L. H. C. (1928). Limiting forms of the frequency distribution in the largest particle size and smallest member of a sample. *Mathematical Proceedings of the Cambridge Philosophical Society*, 24(2):180–190. doi: 10.1017/S0305004100015681

Gnedenko, B. (1943). Sur la distribution limite du terme maximum d'une série aléatoire. *Annals of Mathematics*, 44(3):423–453.doi: 10.2307/1968974

Haan, L. & Ferreira, A. (2006). *Extreme value theory: An introduction*. Springer: New York. doi: 10.1007/0-387-34471-3

Havrda, J. & Charvát, F. (1967). Quantification method of classification processes: Concept of structural $a-$entropy. *Kybernetika*, 1(3):30–35. https://www.kybernetika.cz/content-1967/1/30/paper.pdf

Hill, B. M. (1975). A simple general approach to inference about the tail of a distribution, *The Annals of Statistics*, 3(5):1163–1174. doi: 10.1214/aos/1176343247

Hyndman, R. J. & Fan, Y. (1996). Sample quantiles in statistical packages. *The American Statistician*, 50(4):361–365.doi: 10.2307/2684934

Kouider, M. R. (2019). On maximum likelihood estimates for the shape parameter of the generalized Pareto distribution. *Science Journal of Applied Mathematics and Statistics*, 7(5):89–94. doi: 10.11648/j.sjams.20190705.15

Kouider, M. R. & Benatia, F. (2023). Modified bisection algorithm in estimating the extreme value index. *International Conference of Young Mathematicians*, The Institute of Mathematics of the National Academy of Sciences of Ukraine. https://www.imath.kiev.ua-young/youngconf2023/Abstracts_2023/PS/Kouider-Benatia.pdf

Kouider, M. R., Idiou, N. & Benatia, F. (2023a). Modified bisection algorithm in estimating the extreme value index under random censoring. *TWMS Journal of Applied and Engineering Mathematics*, 13(4):1408–1422. https://jaem.isikun.edu.tr/web/images/articles-vol.13.no.4/13.pdf

Kouider, M. R., Idiou, N. & Benatia, F. (2023b). Adaptive estimators of the general Pareto distribution parameters under random censorship and application. *Journal of Science and Arts*, 23(2):395–412. doi: 10.46939/j.sci.arts-23.2-a07

Kouider, M. R., Idiou, N. & Benatia, F. (2024). Adaptive hazard estimator of the peak over threshold model under random censoring with application. *Studies in Engineering and Exact Sciences*, 5(2):1–24. doi: 10.54021/seesv5n2-243

Kouider, M. R., Kheireddine, S. & Benatia, F. (2024). The adaptive mean estimation of heavy tailed distribution under random censoring. *Studies in Engineering and Exact Sciences*, 5(2):1–31. doi: 10.54021/seesv5n2-164

Pickands III, J. (1975). Statistical inference using extreme order statistics. *The Annals of Statistics*, 3(1):119–131. doi: 10.1214/aos/1176343003

Ramos, P. L., Louzada, F., Ramos, E. & Dey, S. (2018). The Fréchet distribution: Estimation and application an overview. Repository arXiv. doi: 10.48550/arXiv.1801.05327

Smith, R. L. (1987). Estimating tails of probability distributions. *The annals of Statistics*, 15(3):1174–1207. doi: 10.1214/aos/1176350499

# Procjenitelj najveće lq–vjerodostojnosti parametra distribucije s teškim repom

**SAŽETAK**
Proučavanje teorije ekstremnih vrijednosti (EVT) uključuje nekoliko glavnih ciljeva, među kojima je procjena parametra repnog indeksa. Različite metode procjene koriste se za procjenu ovog parametra, poput metode najveće vjerodostojnosti (MLE). Dodatno, Hillov procjenitelj je jedan tip procjenitelja najveće vjerodostojnosti koji je robusniji kod velikih uzoraka nego kod malih. Ovo istraživanje predlaže konstrukciju alternativnog procjenitelja za parametar distribucije s "teškim repom" koristeći pristup najveće lq–vjerodostojnosti (MLqE), kako bi se prilagodili MLE i Hillov procjenitelj za male uzorke. Nadalje, utvrđena je asimptotska normalnost procjenitelja najveće lq–vjerodostojnosti. Osim toga, provedene su simulacijske studije kako bi se usporedio MLq procjenitelj s MLE procjeniteljem. U slučajevima prekoračenja visokih razina pragova, prikladnih vrijednosti, broj najvećih opažanja $k$ vodi do efikasne procjene pomoću Hillovog procjenitelja. U tu svrhu, izbor $k$ kod Hillovog procjenitelja istražen je metodom kvantila tipa 8, koja se pokazala učinkovitom u analizi podataka iz hidrologije. Učinkovitost Hillovog procjenitelja i lq–Hillovog procjenitelja zatim je uspoređena primjenom stvarnih podataka s distribucijom hidroloških vrijednosti.

VRSTA ČLANKA
**Prethodno priopćenje**

INFORMACIJE O ČLANKU
Primljeno: 7. rujna 2024.
Prihvaćeno: 5. studenog 2024.
DOI: 10.62366/crebss.2024.2.003
JEL: C13, C46

**KLJUČNE RIJEČI**
prekoračenja iznad razine praga, indeks ekstremnih vrijednosti, distribucija s teškim repom, procjenitelj najveće lq–vjerodostojnosti