

Research Article

Improved Shape Parameter Estimation for the Three-Parameter Log-Logistic Distribution

Ayana Mateus^{1,2} and Frederico Caeiro^{1,2}

¹NOVA School of Science and Technology (FCT NOVA), NOVA University Lisbon, Caparica, Portugal

²Center for Mathematics and Applications (CMA), FCT NOVA and Department of Mathematics, Portugal

Correspondence should be addressed to Ayana Mateus; amf@fct.unl.pt

Received 9 September 2021; Revised 19 October 2021; Accepted 23 November 2021; Published 13 February 2022

Academic Editor: Qichun Zhang

Copyright © 2022 Ayana Mateus and Frederico Caeiro. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The log-logistic distribution is widely used in different fields of study such as survival analysis, hydrology, insurance, and economics. Recently, Ahsanullah and Alzaatreh studied the best linear unbiased estimators for the location and the scale parameters of the three-parameter log-logistic model. The same authors also propose a shift-invariant Hill estimator for the unknown shape parameter. In this work, we propose a new estimation method for the shape parameter. We derive its nondegenerate asymptotic behaviour and analyse its finite sample performance through a Monte Carlo simulation study. To have precise estimates, we present a method for selecting the threshold. To illustrate the improvement achieved, efficiency comparisons are also provided.

1. Introduction

Over the last decades, the need to solve problems in a diversity of applied areas, such as finance, hydrology, insurance, or survival analysis, gave rise to many new statistical distributions in the literature. A significant effort has been made toward the generalization of some classic distributions. A common technique consists in adding a parameter to a well-known distribution. It is often observed that the introduction of such an extra parameter brings more flexibility to the class of distribution functions. Here, we are interested in the three-parameter or shifted log-logistic distribution which is obtained from the classic log-logistic model with the addition of a location parameter. Thus, a random variable X follows a three-parameter log-logistic distribution if its probability density function and distribution function (d.f.) are, respectively, given by

$$f(x | \alpha, \mu, \sigma) = \frac{\alpha((x - \mu)/\sigma)^{\alpha-1}}{(1 + ((x - \mu)/\sigma)^\alpha)^2}, \quad x > \mu, \quad (1)$$

$$F(x | \alpha, \mu, \sigma) = 1 - \frac{1}{1 + ((x - \mu)/\sigma)^\alpha} = \frac{1}{1 + ((x - \mu)/\sigma)^{-\alpha}}, \quad x > \mu. \quad (2)$$

The corresponding quantile function is

$$Q_X(p | \alpha, \mu, \sigma) = \mu + \sigma \left(\frac{p}{1-p} \right)^{1/\alpha}, \quad 0 \leq p < 1. \quad (3)$$

The constants $\alpha > 0$, $\mu \in \mathbf{R}$, and $\sigma > 0$ are the shape, location, and scale parameters, respectively. Once there is no restriction between the values of the parameters, they can vary freely in the parameter space. If $\alpha = 1$, X has a location and scale beta prime distribution. When $\alpha > 1$, the probability density function f is unimodal with mode at the value $\mu + \sigma((\alpha - 1)/(\alpha + 1))^{1/\alpha}$. Since the log-logistic model is heavy tailed, with a tail index equal to α , the i -th moment of X is finite only when $i < \alpha$. In the following, the notation $X \sim \text{llogist}(\alpha, \mu, \sigma)$ will be used whenever the random variable X has the distribution function in (2). If $X \sim \text{llogist}(\alpha, \mu, \sigma)$, then $(X - \mu)/\sigma$, with $\sigma > 0$, has a standard log-logistic distribution, $\text{llogist}(\alpha, 0, 1)$, with d.f. given by

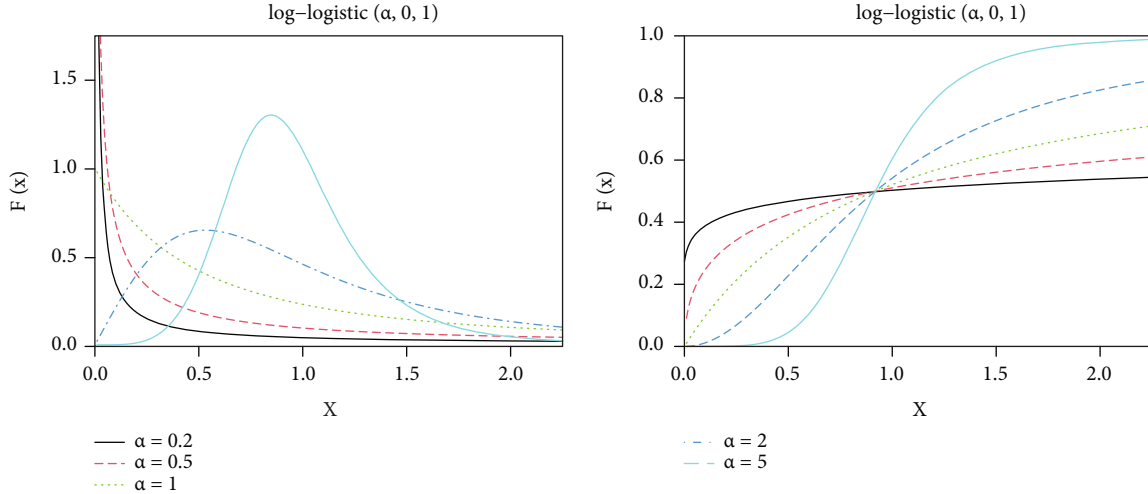


FIGURE 1: Plots of the probability density function (a) and cumulative distribution function (a) and for some shape parameter values of the standard log-logistic distribution.

$$F(x | \alpha, 0, 1) = 1 - \frac{1}{1 + x^\alpha}, \quad x > 0. \quad (4)$$

Figure 1 illustrates the probability density function and cumulative distribution function, for selected values of the shape parameter of the standard log-logistic distribution. As noted in [1], the standard log-logistic model is characterized by the relation

$$xf(x | \alpha, 0, 1) = \alpha F(x | \alpha, 0, 1)(1 - F(x | \alpha, 0, 1)), \quad \alpha > 0. \quad (5)$$

The three-parameter log-logistic distribution is also known as a Pareto type III distribution (see Arnold [2]). Also, the two-parameter log-logistic distribution is a member of Burr's type XII family of distributions [3] and is known as the Fisk [4] distribution in the economic literature.

The log-logistic distribution in (2) is closely related with the logistic distribution. More precisely, if $X \sim \text{llogist}(\alpha, \mu, \sigma)$, then

$$W = \alpha \ln \left(\frac{X - \mu}{\sigma} \right), \quad (6)$$

is a logistic random variable with probability density function $f(w) = e^w / (1 + e^w)^2$, $w \in \mathbf{R}$.

The half-logistic model, that is, the absolute value of the standard logistic model is another important model in the scientific literature. Extensions of the log-logistic or half-logistic models can be found in Cordeiro et al. [5, 6], Alizadeh et al. [7], Mohammad [8], Lemonte [9], and Shakhatareh [10] among others.

Considerable attention has been paid to the estimation of the model parameters of the log-logistic distribution. Although more attention has been paid to the two-parameter case, several estimation procedures for the three-parameter log-logistic model are already available in the literature. Balakrishnan et al. [1] derived the best linear unbiased estimators (BLUE)

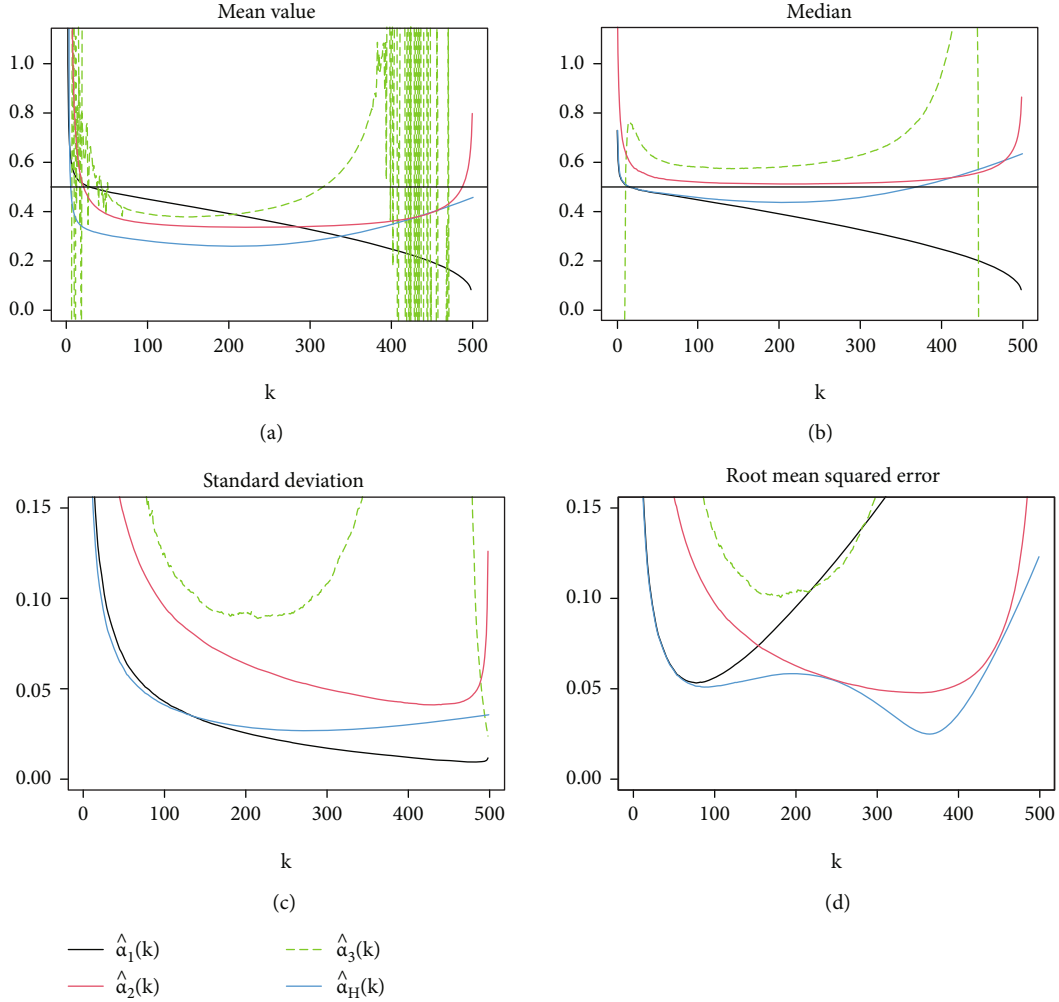
for the location and scale parameters of a three-parameter log-logistic model, with a known shape parameter. In practice, it is unrealistic to assume that the shape parameter α is known, and it should be estimated. More recently, Ahsanullah and Alzaatreh [11] considered again the BLUE for the location and scale parameters of the log-logistic model. The authors propose the estimation of the shape parameter with the reciprocal of a Hill-type estimator applied to a tail sample fraction, shifted by the sample minimum. Moreover, Ahsanullah and Alzaatreh [11] proposed a sample fraction of 10%, if the sample size is greater than 100. However, such a simple suggestion has no theoretical or empirical support. The main objective of this paper is to improve the estimation of the shape parameter α of the log-logistic model. The motivation comes from the fact that the Hill estimator is biased and the sample fraction proposed in [11] is nonoptimal.

The paper is organized as follows. Section 2 describes the estimator proposed [11] and presents two alternative estimators and their asymptotic properties. It is shown that the estimators are asymptotically normal distributed and the choice of the sample fraction is discussed. In Section 3, we introduce a simple threshold selection method. The results of a Monte Carlo simulation that evaluates the mean value, the median, the standard deviation, and the root mean squared error of the estimators under study are reported in Section 4. Finally, Section 5 concludes the paper.

2. Estimator of the Shape Parameter and Their Properties

In what follows, we consider the estimation of the shape parameter of the log-logistic model in (2). We shall assume that (X_1, X_2, \dots, X_n) is a sample of size n of independent and identically distributed random variables, with a common d.f. F , given in (2). All three parameters are assumed unknown. The corresponding sample of nondecreasing order statistics is denoted by $(X_{1:n}, X_{2:n}, \dots, X_{n:n})$.

- (i) Generate 5000 samples of size n from a log-logistic distribution, with n taking values between 50 and 2000, with step 50
- (ii) Let $\hat{\alpha}_j(k, i)$ denote the estimates based on the i -th sample. For each sample size, compute $\hat{\alpha}_j(k, i)$, $k = 1, 2, \dots, n - 2$, $i = 1, \dots, 5000$
- (iii) Compute the empirical root mean squared error as a function of k
- (iv) Obtain the level $k_j(n)$ that minimizes the empirical root mean squared error
- (v) Finally, perform a power regression with $k_j(n)$ as the response variable and the sample size as the predictor variable. The regression coefficients are the vector values (a_1, a_2) .

 ALGORITHM 1: Method to determine the constants (a_1, a_2) .

 FIGURE 2: Simulated mean values (a), median (b), standard deviation (c), and RMSE (d) of the estimators of α , versus k for samples of size $n = 500$ drawn from the log-logistic distribution with $(\alpha, \mu, \sigma) = (0.5, 1, 1)$.

2.1. Hill-Type Estimator. Ahsanullah and Alzaatreh [11] noted that the log-logistic distribution has a Pareto-type tail behaviour. More precisely, we can write

$$1 - F(x | \alpha, 0, 1) = x^{-\alpha} l(x), \quad (7)$$

with α the tail (or Pareto) index and $l(x) = (1 + x^{-\alpha})^{-1}$, $x > 0$, a slowly varying function at infinity. Moreover, since $l(x)$ admits the Taylor expansion $l(x) = 1 - x^{-\alpha} + o(-x^{-\alpha})$, as $x \rightarrow \infty$, the standard log-logistic model belongs to Hall's subclass (see equation (1) of [12]) of Pareto-type models with survival function,

$$1 - F(x) = \left(\frac{x}{c}\right)^{-(1/\xi)} \left(1 + \frac{\beta}{\rho} \left(\frac{x}{c}\right)^{\rho/\xi} + o\left(x^{\rho/\xi}\right)\right), \quad (8)$$

with $\xi > 0$, $c > 0$ is a first-order scale parameter, and $\rho < 0$ and β are second-order tail parameters. The second-order parameter ρ quantifies the deviation of the model to the Pareto distribution. For the standard log-logistic model, we have $\xi = \alpha^{-1}$, $\rho = -1$, and $c = \beta = 1$. Therefore, F is in the max domain of attraction of the extreme value distribution with a positive shape parameter. This means that there exist

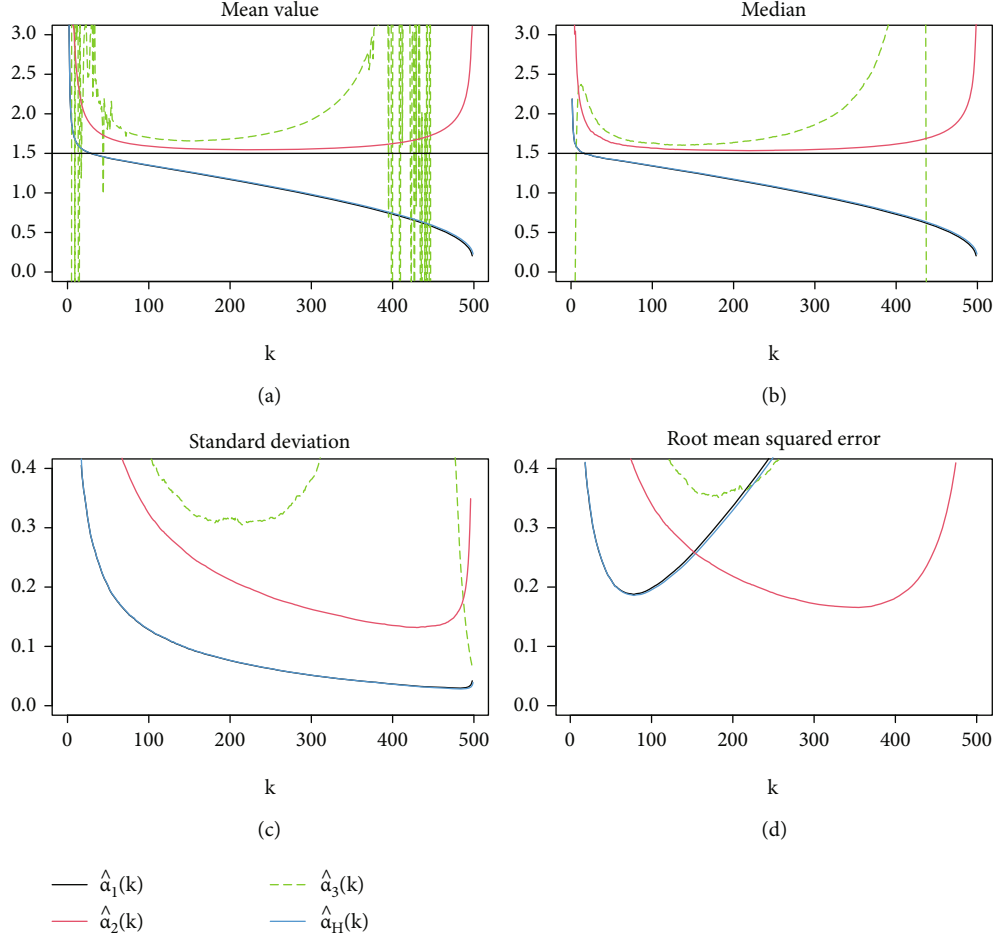


FIGURE 3: Simulated mean values (a), median (b), standard deviation (c), and RMSE (d) of the estimators of α , versus k for samples of size $n = 500$ drawn from the log-logistic distribution with $(\alpha, \mu, \sigma) = (1.5, 0, 1)$.

sequences of normalizing constants $a_n > 0$ and b_n such that $(X_{n:n} - b_n)/a_n$ converges in distribution to a nondegenerate random variable with distribution function

$$\text{EV}_\xi(x) = \begin{cases} \exp\left(-(1 + \xi x)^{-1/\xi}\right), & 1 + \xi x > 0, \quad \text{if } \xi \neq 0, \\ \exp(-\exp(-x)), & x \in \mathbb{R}, \quad \text{if } \xi = 0, \end{cases} \quad (9)$$

with a positive shape parameter ξ . This parameter is the so-called extreme value index. In the statistic literature, one can find several estimators for ξ or, equivalently, for the shape parameter α . For a general overview of the available estimators, we refer to [13, 14]. Whenever working with Pareto-type models, the Hill estimator [15] is frequently used to estimate the extreme value index. This estimator is defined as the average of the log excesses over the threshold $u = X_{n-k:n} > 0$,

$$H(k) := \frac{1}{k} \sum_{i=1}^k \ln X_{n-i+1:n} - \ln X_{n-k:n}, \quad k = 1, 2, \dots, n-1, \quad (10)$$

where k represents the number of upper order statistics used in the estimation. For the strict Pareto model, with d.f. $F(y) = 1 - (y/c)^{-\alpha}$, $y > 0$, ($\alpha > 0$, $c > 0$), the Hill estimator is consistent and asymptotically unbiased for the estimation of α^{-1} . Moreover, if $k = n - 1$, $H(k)$ is the maximum likelihood estimator of the reciprocal of the shape parameter α of the strict Pareto distribution. If F differs from the strict Pareto model, $H(k)$ is consistent if k is intermediate, i.e., if $k = k_n \in [1, n - 1]$ is a sequence of positive integers satisfying

$$k \longrightarrow \infty, \quad \frac{k}{n} \longrightarrow 0, \quad \text{as } n \longrightarrow \infty. \quad (11)$$

Also, the variance of $H(k)$ decreases and the absolute bias increases, as k increases. Therefore, the choice of k leads to a trade-off between the bias and the variance of the estimator.

Many of the estimators that have been suggested to estimate a positive extreme value index, including the one in (10), are only scale-invariant. A change in the location can modify the asymptotic behaviour of the tail and the bias of a location-variant estimator (for more information, see the papers [16, 17]). The properties of the Hill estimator and the fact that the d.f. in (2) has a location parameter, lead Ahsanullah and Alzaatreh [11] to propose the estimation

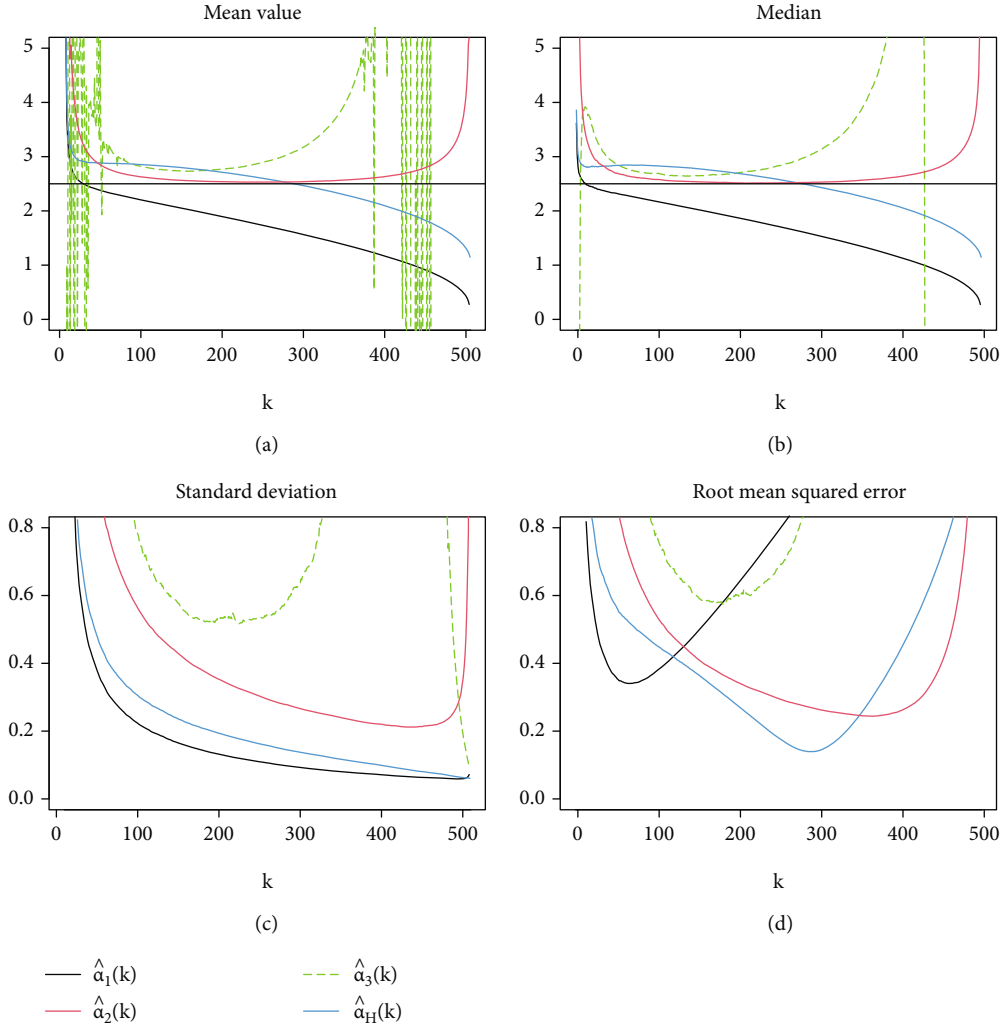


FIGURE 4: Simulated mean values (a), median (b), standard deviation (c), and RMSE (d) of the estimators of α , versus k for samples of size $n = 500$ drawn from the log-logistic distribution with $(\alpha, \mu, \sigma) = (2.5, 2, 3)$.

of the shape parameter α with the following location-invariant Hill-type estimator,

$$\hat{\alpha}_1(k) = \frac{1}{H^*(k)}, \quad H^*(k) = \frac{1}{k} \sum_{i=1}^k \ln \frac{X_{n-i+1:n} - X_{1:n}}{X_{n-k:n} - X_{1:n}}, \quad 1 \leq k \leq n-2. \quad (12)$$

The sample values are thus shifted by the sample minimum. In the following, we always add the notation $*$ to any statistic based on shifted data. Note that the estimator $H^*(k)$ in (12) is a member of the class of estimators in references [16, 18, 19]. Regarding the choice of the parameter k , Ahsanullah and Alzaatreh [11] proposed $k = \lfloor n/10 \rfloor$, if $n > 100$, where $\lfloor x \rfloor$ denotes the integer part of x . In order to improve the choice of the threshold k , it is crucial to gain information about the asymptotic behaviour of $\hat{\alpha}_1$. We begin with the following proposition that provides the distributional representation of $\hat{\alpha}_1$. The proof can be found in the appendix.

Proposition 1. *Assume that k is an intermediate sequence of integers satisfying (11). Then, the following distributional representation*

$$\hat{\alpha}_1(k) \stackrel{d}{=} \alpha \left(1 - \frac{Z_k}{\sqrt{k}} + O_p \left(\frac{\sqrt{k}}{n} \right) - \frac{k}{2n} (1 + o_p(1)) \right), \quad (13)$$

is valid, where Z_k is asymptotically standard normal. Moreover, if $\sqrt{k}(k/2n) \rightarrow c$ then

$$\sqrt{k}(\hat{\alpha}_1(k) - \alpha) \xrightarrow{d} N(-\alpha c, \alpha^2). \quad (14)$$

A typical approach in the literature is to choose the threshold through the minimization of the asymptotic mean squared error (AMSE) of $\hat{\alpha}_1(k)$. Because such a choice depends on asymptotic arguments, it may only be reliable when the sample size becomes large. Alternative methods for selecting the threshold can be found in references [20–23]. From (13), it follows that the AMSE of

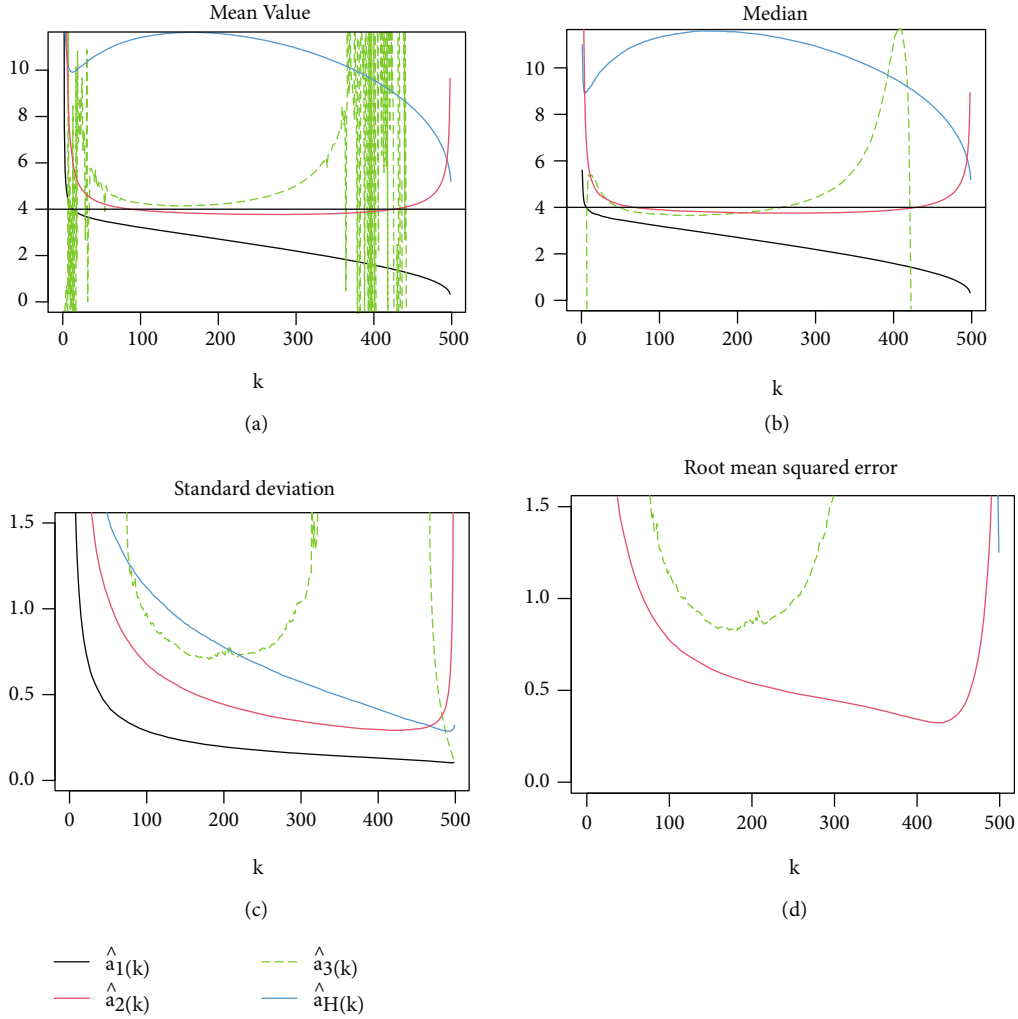


FIGURE 5: Simulated mean values (a), median (b), standard deviation (c), and RMSE (d) of the estimators of α , versus k for samples of size $n = 500$ drawn from the log-logistic distribution with $(\alpha, \mu, \sigma) = (4, 2, 0.5)$.

$\hat{\alpha}_1(k)$ is given by

$$\text{AMSE}(\hat{\alpha}_1(k)) = \alpha^2 \left(\frac{1}{k} + \frac{k^2}{4n^2} \right). \quad (15)$$

Then, the level k_0 that minimizes the AMSE in Equation (15) is asymptotically equivalent to

$$k_0^{(1)} = \text{argmin}[\text{AMSE}(\hat{\alpha}_1(k))] = (2n^2)^{1/3}. \quad (16)$$

Furthermore, the level $k_0^{(1)}$ leads to a biased estimate of the α ($\sqrt{k_0^{(1)}}(k_0^{(1)}/2n) \rightarrow 1/\sqrt{2}$).

2.2. Alternative Estimators. Despite its wide use to estimate the extreme value index, the Hill estimator is difficult to apply in real data problems due to the substantial bias. This problem motivated several researchers to reduce the bias of such an estimator and to construct alternative estimators. Reduced bias estimators of the extreme value

index usually have a stable sample path, close to the target value, which makes them less sensitive to the choice of k . We mention the first reduced bias estimators of the extreme value index in references [24–27] (see also the papers [28, 29] for a general overview on bias reduction).

Bias reduction of tail parameter estimators typically requires the estimation of tail second-order parameters. However, in the present paper, we are assuming the parametric log-logistic distribution, a model within the class of heavy tailed models in (8), with $(\rho, \beta) = (-1, 1)$. Therefore, in this paper, we shall consider two alternative estimators, which provide a reduction of the bias whenever $(\rho, \beta) = (-1, 1)$. We shall consider the following alternative estimators:

$$\hat{\alpha}_2(k) = \frac{1}{G_1^*(k)}, \quad \hat{\alpha}_3(k) = \frac{1}{WH^*(k)}, \quad k = 1, 2, \dots, n-2, \quad (17)$$

TABLE 1: Simulated values of the optimal sample fraction, mean values, and RMSE at the simulated optimal threshold, computed for the $\text{llogist}(\alpha, \mu, \sigma)$ model.

n	20	50	100	200	500	1000	2000	5000
$(\alpha, \mu, \sigma) = (0.5, 1, 1)$								
$\hat{k}_0^{(1)}/n$	0.4500	0.3600	0.2700	0.2200	0.1540	0.1110	0.0980	0.0754
$\hat{k}_0^{(2)}/n$	0.8000	0.7800	0.7600	0.7400	0.7040	0.6410	0.5940	0.5456
$E(\hat{\alpha}_1(\hat{k}_0^{(1)}))$	0.3962	0.4172	0.4412	0.4500	0.4645	0.4741	0.4765	0.4816
$E(\hat{\alpha}_2(\hat{k}_0^{(2)}))$	0.6951	0.5928	0.5584	0.5404	0.5266	0.5174	0.5129	0.5090
$\text{RMSE}(\hat{\alpha}_1(\hat{k}_0^{(1)}))$	0.1677	0.1259	0.1002	0.0815	0.0620	0.0501	0.0408	0.0305
$\text{RMSE}(\hat{\alpha}_2(\hat{k}_0^{(2)}))$	0.3532	0.1812	0.1207	0.0852	0.0557	0.0406	0.0302	0.0202
$(\alpha, \mu, \sigma) = (1.5, 0, 1)$								
$\hat{k}_0^{(1)}/n$	0.4000	0.3000	0.2700	0.2150	0.1540	0.1110	0.0980	0.0744
$\hat{k}_0^{(2)}/n$	0.8000	0.7800	0.7600	0.7400	0.7040	0.6410	0.5940	0.5456
$E(\hat{\alpha}_1(\hat{k}_0^{(1)}))$	1.1794	1.2904	1.3065	1.3449	1.3896	1.4203	1.4282	1.4453
$E(\hat{\alpha}_2(\hat{k}_0^{(2)}))$	2.0714	1.7706	1.6712	1.6188	1.5785	1.5513	1.5382	1.5268
$\text{RMSE}(\hat{\alpha}_1(\hat{k}_0^{(1)}))$	0.5302	0.3913	0.3081	0.2489	0.1877	0.1512	0.1230	0.0918
$\text{RMSE}(\hat{\alpha}_2(\hat{k}_0^{(2)}))$	1.0482	0.5306	0.3559	0.2525	0.1656	0.1210	0.0902	0.0603
$(\alpha, \mu, \sigma) = (2.5, 2, 3)$								
$\hat{k}_0^{(1)}/n$	0.3500	0.2800	0.2200	0.1750	0.1340	0.1090	0.0870	0.0656
$\hat{k}_0^{(2)}/n$	0.8000	0.7800	0.7800	0.7550	0.7340	0.6860	0.6290	0.5844
$E(\hat{\alpha}_1(\hat{k}_0^{(1)}))$	1.8626	2.0328	2.1485	2.2263	2.2932	2.3333	2.3710	2.4038
$E(\hat{\alpha}_2(\hat{k}_0^{(2)}))$	3.1386	2.7714	2.6780	2.6219	2.5901	2.5635	2.5455	2.5348
$\text{RMSE}(\hat{\alpha}_1(\hat{k}_0^{(1)}))$	0.9756	0.7189	0.5638	0.4548	0.3403	0.2706	0.2183	0.1624
$\text{RMSE}(\hat{\alpha}_2(\hat{k}_0^{(2)}))$	1.4794	0.7467	0.5109	0.3682	0.2438	0.1810	0.1363	0.0920
$(\alpha, \mu, \sigma) = (4, 2, 0.5)$								
$\hat{k}_0^{(1)}/n$	0.3000	0.2200	0.1600	0.1400	0.0900	0.0760	0.0600	0.0430
$\hat{k}_0^{(2)}/n$	0.8000	0.8400	0.8500	0.8600	0.8540	0.8320	0.8045	0.7684
$E(\hat{\alpha}_1(\hat{k}_0^{(1)}))$	2.6513	3.0091	3.2582	3.3449	3.5425	3.6160	3.6949	3.7747
$E(\hat{\alpha}_2(\hat{k}_0^{(2)}))$	4.1997	3.9457	3.9195	3.9602	4.0043	4.0102	4.0082	4.0095
$\text{RMSE}(\hat{\alpha}_1(\hat{k}_0^{(1)}))$	1.7873	1.3664	1.1050	0.9007	0.6932	0.5593	0.4547	0.3429
$\text{RMSE}(\hat{\alpha}_2(\hat{k}_0^{(2)}))$	1.8306	0.9646	0.6696	0.4872	0.3233	0.2392	0.1802	0.1243

where

$$G_1^*(k) = \frac{M^{(2)*}(k) - (M^{(1)*}(k))^2}{M^{(1)*}(k)} = MR^*(k) - H^*(k), \quad k = 1, 2, \dots, n-2, \quad (18)$$

$$M_n^{(j)*}(k) := \frac{1}{k} \sum_{i=1}^k \left(\ln \frac{X_{n-i+1:n} - X_{1:n}}{X_{n-k:n} - X_{1:n}} \right)^j, \quad j > 0, \left(M_n^{(1)*}(k) \equiv H^*(k) \right), \quad (19)$$

are the moments of order j of the shifted log-excesses. The statistic WH^* is defined as

$$WH^*(k) = \frac{1}{k} \sum_{i=1}^k 4 \left(1 - \frac{i}{k+1} \right) \left(\ln \frac{X_{n-i+1:n} - X_{1:n}}{X_{n-k:n} - X_{1:n}} \right), \quad k = 1, 2, \dots, n-2. \quad (20)$$

Both statistics in (18) and (20) are data-shifted versions of statistics already considered in the literature. The nonshifted version of (18) was first introduced in [30], a particular case of a generalized Jackknife statistic based on the Hill and the moment ratio (MR) [31] estimators. Moreover, both $H(k)$ and $MR(k)$ are members of Lehmer mean-of-order- p class of extreme value index estimators studied in [32, 33]. The nonshifted version of the estimator $G_1^*(k)$ was also independently introduced in [34] (see also [35]). The statistic in (20) is a location invariant modification of the class of weighted Hill estimators introduced in [36]. Next, we provide the asymptotic representation of the alternative estimators in (17). We do not provide proofs, since the arguments required to obtain the results are similar to the ones provided in the proof of Proposition 1 and in references [30, 36].

Proposition 2. *Assume the conditions of Proposition 1. Then, the following distributional representation holds:*

$$\hat{\alpha}_2(k) \stackrel{d}{=} \alpha \left(1 - \frac{\sqrt{5}Z_k}{\sqrt{k}} + O_p \left(\frac{\sqrt{k}}{n} \right) + o_p \left(\frac{k}{2n} \right) (1 + o_p(1)) \right), \quad (21)$$

$$\hat{\alpha}_3(k) \stackrel{d}{=} \alpha \left(1 - \frac{2Z_k}{\sqrt{k}} + O_p \left(\frac{\sqrt{k}}{n} \right) + o_p \left(\frac{k}{2n} \right) (1 + o_p(1)) \right), \quad (22)$$

where Z_k is an asymptotically standard normal variable. Moreover, if $\sqrt{k}(k/2n) \rightarrow c$, then

$$\sqrt{k}(\hat{\alpha}_2(k) - \alpha) \xrightarrow{d} N(0, 5\alpha^2), \quad \sqrt{k}(\hat{\alpha}_3(k) - \alpha) \xrightarrow{d} N(0, 4\alpha^2). \quad (23)$$

Regarding the asymptotic variances, note that we have $V(\hat{\alpha}_1(k)) < V(\hat{\alpha}_3(k)) < V(\hat{\alpha}_2(k))$. The absolute asymptotic bias of either $\hat{\alpha}_2(k)$ or $\hat{\alpha}_3(k)$ is null, while the absolute asymptotic bias of $\hat{\alpha}_1(k)$ is $|\text{Bias}(\hat{\alpha}_1(k))| = k/2n$. Moreover,

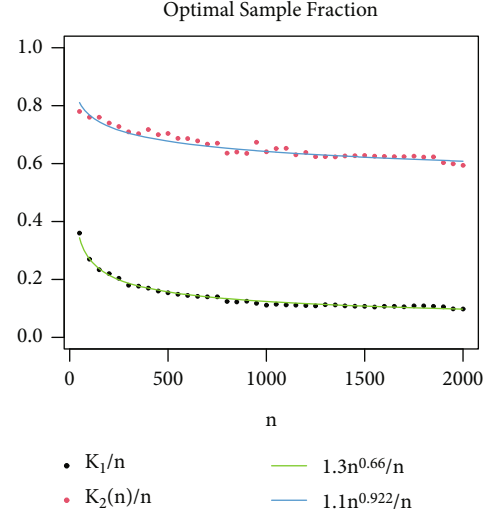


FIGURE 6: Simulated mean values of the optimal sample fraction for $\hat{\alpha}_1(k)$ and $\hat{\alpha}_2(k)$ and the corresponding regression curves.

the results in (21) and (22) do not allow to evaluate the asymptotic optimal value of k . To find such an optimal value, it is necessary to obtain higher-order terms of the asymptotic expansion of bias. The major drawback of such an approach is that it can lead to inaccurate results for small sample sizes. For that reason, we shall address the choice of k in the next section.

3. A Threshold Selection Method

In practice, the threshold is fundamental to obtain accurate estimates, and it must be chosen before applying any of the aforementioned estimators of the parameter α . The naive method of selecting the top 10% of the data is extremely simple but has some drawbacks. Since such a threshold is not an intermediate sequence, it is not well supported by the theory. To have precise estimates for small sample sizes, we now propose a method for selecting the threshold that combines theoretical and empirical results. Hall [12] noted that the optimal performance of the Hill estimator is achieved if $k = O(\delta_1 n^{\delta_2})$ for some $\delta_1 > 0$ and $\delta_2 \in [0, 1]$. Such a result also holds for other estimators of α , with a similar limit distribution. In the following, let $\hat{\alpha}_j(k)$ denote any of the aforementioned estimators of the shape parameter. We propose the following threshold selection procedure:

$$k_j = k_j(n) = \lfloor a_1 n^{a_2} \rfloor, \quad (24)$$

with $a_1 > 0$ and $a_2 \in (0, 1)$ as suitable real numbers, chosen independently for each estimator. Additionally, it is worth noting that, at least for the estimator $\hat{\alpha}_1(k)$, the constants a_1 and a_2 are asymptotically independent of any parameter of the log-logistic model in (2). We propose Algorithm 1 to obtain empirically both constants (a_1, a_2) in (24).

TABLE 2: Simulated mean value and RMSE of the estimators $\hat{\alpha}_j^*$, $j = 1, 2, 3$.

	20	50	100	200	500	1000	2000	5000
$(\alpha, \mu, \sigma) = (0.5, 1, 1)$								
$E(\hat{\alpha}_0^*)$	0.8577	0.5804	0.5199	0.4955	0.4815	0.4774	0.4760	0.4748
$E(\hat{\alpha}_1^*)$	0.3962	0.4250	0.4412	0.4526	0.4639	0.4704	0.4765	0.4826
$E(\hat{\alpha}_2^*)$	0.7148	0.5962	0.5584	0.5389	0.5240	0.5174	0.5139	0.5100
RMSE($\hat{\alpha}_0^*$)	1.3470	0.3452	0.1803	0.1135	0.0698	0.0512	0.0408	0.0325
RMSE($\hat{\alpha}_1^*$)	0.1677	0.1262	0.1002	0.0818	0.0621	0.0503	0.0408	0.0306
RMSE($\hat{\alpha}_2^*$)	0.3650	0.1818	0.1207	0.0854	0.0560	0.0406	0.0303	0.0203
$(\alpha, \mu, \sigma) = (1.5, 0, 1)$								
$E(\hat{\alpha}_0^*)$	2.5008	1.7205	1.5499	1.4807	1.4415	1.4304	1.4268	1.4237
$E(\hat{\alpha}_1^*)$	1.1078	1.2414	1.3065	1.3491	1.3879	1.4093	1.4282	1.4474
$E(\hat{\alpha}_2^*)$	2.1729	1.7834	1.6712	1.6139	1.5706	1.5513	1.5411	1.5298
RMSE($\hat{\alpha}_0^*$)	3.8964	1.0149	0.5355	0.3391	0.2097	0.1541	0.1231	0.0980
RMSE($\hat{\alpha}_1^*$)	0.5365	0.3924	0.3081	0.2493	0.1880	0.1520	0.1230	0.0922
RMSE($\hat{\alpha}_2^*$)	1.1515	0.5347	0.3559	0.2531	0.1665	0.1210	0.0903	0.0607
$(\alpha, \mu, \sigma) = (2.5, 2, 3)$								
$E(\hat{\alpha}_0^*)$	3.8239	2.7151	2.4844	2.3968	2.3540	2.3476	2.3501	2.3533
$E(\hat{\alpha}_1^*)$	1.6140	1.9042	2.0571	2.1622	2.2573	2.3097	2.3527	2.3950
$E(\hat{\alpha}_2^*)$	3.3060	2.7909	2.6644	2.6051	2.5627	2.5461	2.5393	2.5302
RMSE($\hat{\alpha}_0^*$)	5.8773	1.5590	0.8494	0.5547	0.3584	0.2719	0.2217	0.1785
RMSE($\hat{\alpha}_1^*$)	1.0385	0.7497	0.5832	0.4643	0.3451	0.2760	0.2211	0.1635
RMSE($\hat{\alpha}_2^*$)	1.6688	0.7516	0.5123	0.3717	0.2496	0.1835	0.1373	0.0923
$(\alpha, \mu, \sigma) = (4, 2, 0.5)$								
$E(\hat{\alpha}_0^*)$	5.2055	3.8334	3.5848	3.5134	3.5073	3.5357	3.5697	3.6099
$E(\hat{\alpha}_1^*)$	2.0964	2.5999	2.8959	3.1196	3.3372	3.4675	3.5746	3.6860
$E(\hat{\alpha}_2^*)$	4.4281	3.8722	3.7920	3.7822	3.7981	3.8234	3.8520	3.8817
RMSE($\hat{\alpha}_0^*$)	7.8812	2.1558	1.2880	0.9334	0.6941	0.5778	0.5004	0.4228
RMSE($\hat{\alpha}_1^*$)	2.0392	1.5397	1.2340	0.9990	0.7595	0.6148	0.4974	0.3714
RMSE($\hat{\alpha}_2^*$)	2.0809	0.9698	0.7181	0.5600	0.4120	0.3239	0.2544	0.1847

Remark 3. The constants in the first step of Algorithm 1 can be changed to modify the precision of the estimates of (a_1, a_2) .

4. Numeric Results

In this section, we use Monte Carlo simulation to study and compare the finite-sample properties of the estimators of the shape parameter α , in (12) and (17). For comparison purposes, we also include the reciprocal of the Hill estimator, $\hat{\alpha}_H = 1/H(k)$, with $H(k)$ in (10). All the computations were done using R [37] software. We generated 5000 samples of size n from the log-logistic model with the same set of

parameters considered in paper [11], that is,

$$\begin{aligned}
 (\alpha, \mu, \sigma) &= (0.5, 1, 1), \\
 (\alpha, \mu, \sigma) &= (1.5, 0, 1), \\
 (\alpha, \mu, \sigma) &= (2.5, 2, 3), \\
 (\alpha, \mu, \sigma) &= (4, 2, 0.5).
 \end{aligned}
 \tag{25}$$

For the sample size n , we considered values between 50 and 5000. Let $\hat{\alpha} \cdot (k)$, with $\cdot \in \{1, 2, 3, H\}$, denote any of the aforementioned estimators under study. For each sample, the estimates of α are first computed for every k . We then

calculate the mean value (E), the median, the standard deviation (SD), and the root mean squared error (RMSE) for each k . We have further computed the simulated optimum level:

$$\widehat{k}_0^{(\bullet)} = \arg \min_k \text{RMSE}[\widehat{\alpha}^\bullet(k)], \quad (26)$$

and the simulated previous characteristics at the optimal level in (26). Furthermore, since the exact minimizer of the RMSE is not known, we also study the estimators with the level proposed in Section 3.

4.1. Results for Every Value of k . In Figures 2–5, we present the Monte Carlo estimates of the mean value (top left), the median (top right), the SD (bottom left), and the RMSE, with respect to k , for the four different estimators. The horizontal solid line, at the top plots, indicates the true value of the parameter α . A good performance is assessed by the flatness of the curves of the mean value/median in a large continuous region of values of k , close to the true value of α and by a small RMSE in such a region. A small RMSE without a stability region, of the estimates of the mean value/median, may not be useful in practical applications.

In all of the figures, estimator $\widehat{\alpha}_1(k)$ shows a decreasing negative bias and the smallest standard deviation, as a function of k . The performance of the location-variant estimator $\widehat{\alpha}_H(k)$ is affected by the values of the location and shape parameters. It can perform equally, worse, or better than the location invariant estimator $\widehat{\alpha}_1(k)$. Also, in all of the figures, $\widehat{\alpha}_2(k)$ and $\widehat{\alpha}_3(k)$ provide a region with a stable sample path of the mean value and median of the estimates, as expected. However, the mean value plots suggest that $\widehat{\alpha}_3(k)$ does not provide good estimates if k is very small or very large. This problem occurs only in a few generated samples. Based on the simulation results, $\widehat{\alpha}_2(k)$ provides globally a good bias and RMSE patterns as a function of k . Therefore it constitutes an important estimation procedure for the shape parameter α .

4.2. Results for the Optimal Level of k . We now compare the performance of the estimators at their simulated optimal level. The exact optimal value of k provides a benchmark of the best possible performance obtainable with each estimator of α . Notice that in practice, such an optimal level may not be achieved. For the sake of simplicity, we excluded the estimators $\widehat{\alpha}_H(k)$ and $\widehat{\alpha}_3(k)$, the first due to the mixed performance, the second because it is less efficient than $\widehat{\alpha}_2(k)$. Thus, we only provide results for the estimator $\widehat{\alpha}_1(k)$, suggested in [11], and $\widehat{\alpha}_2(k)$, the estimator that provided a good performance for the four sets of the model parameters. The simulated values of the optimal sample fraction (the optimal level divided by the sample size), mean value, and RMSE, both computed at the simulated optimal threshold, are given in Table 1. Values associated to a smaller absolute bias and smaller RMSE are presented in bold.

Notice that the $\widehat{\alpha}_2(k)$ has always an optimal level much higher than $\widehat{\alpha}_1(k)$. At their corresponding optimal level, $\widehat{\alpha}_1$ gives better results for small sample sizes ($n < 500$ if $\alpha = 0.5$ or 1.5 , $n < 100$ if $\alpha = 2.5$, and $n < 50$ if $\alpha = 4$). For large values of n , $\widehat{\alpha}_1$ is less biased and provides a smaller RMSE.

4.3. Results for the Proposed Level k . In this section, we study the distributional properties of the estimators, based on the tail data defined by the threshold proposed in (26). Again, we restrict ourselves to the estimators $\widehat{\alpha}_1(k)$ and $\widehat{\alpha}_2(k)$ in (12) and (17), respectively. We applied algorithm 3 to samples of the log-logistic model with parameters $(\alpha, \mu, \sigma) = (0.5, 1, 1)$. The algorithm provided the thresholds

$$k_1 = \lceil 1.3n^{0.66} \rceil \text{ for } \widehat{\alpha}_1(k) \text{ and } k_2 = \lceil 1.1n^{0.922} \rceil \text{ for } \widehat{\alpha}_2(k). \quad (27)$$

Figure 6 presents the empirical optimal sample fraction and the corresponding regression curve for both estimators. The overall agreement between empirical and fitted values is quite good.

Next, we provide numerical results of the finite sample performance of the estimators with the thresholds proposed in [11] and (27). More precisely, the following estimators are considered:

$$\begin{aligned} \widehat{\alpha}_0^* &= \widehat{\alpha}_1(\lfloor 0.1n \rfloor), \\ \widehat{\alpha}_1^* &= \widehat{\alpha}_1(\lceil 1.3n^{0.66} \rceil), \\ \widehat{\alpha}_2^* &= \widehat{\alpha}_2(\lceil 1.1n^{0.922} \rceil). \end{aligned} \quad (28)$$

From the simulation results obtained in Section 4.1, for every value of k , we obtained the mean value and the root mean squared error of the estimators under study. Results are presented in Table 2.

When looking at the empirical bias, $\widehat{\alpha}_0^*$ achieves the smallest absolute bias if $\alpha < 4$ and the sample size has a moderate size (usually between 50 and 500). Otherwise, $\widehat{\alpha}_0^*$ has usually the smallest absolute bias. In terms of RMSE, the threshold approach in (27) compared favourably to the threshold proposed in [11]. If the sample size is small, $\widehat{\alpha}_1^*$ provides the smallest RMSE. Otherwise, $\widehat{\alpha}_2^*$ has the smallest RMSE. For all sets of parameters, the values of RMSE of α_1^* and α_2^* are very close to the corresponding optimal level in Table 1.

5. Conclusion

It is known that the Hill estimator can be seriously biased. To deal with this problem, we introduced two reduced bias estimators for the shape parameter of the log-logistic model and established its asymptotic normality. Additionally, we improved earlier guidelines for the choice of the threshold. Also, the presented simulation study shows improvements in the estimation of the shape parameter when compared to the classical estimation method in Ahsanullah and Alzaatreh [11]. These improvements are more pronounced if the sample size is large. We conclude by noting that the approach taken in this paper could be applicable to other bias-reduced estimators. Therefore, further research concerning alternative estimators of the shape parameter will be taken in the future.

Appendix

A.1. Proofs

A.1. Preliminary Results. Here, we provide several lemmas that are useful in the derivation of the asymptotic properties of the estimators. Let $E_{1:n}, E_{2:n}, \dots, E_{n:n}$ be the order statistics from n mutually independent and identically distributed exponential random variables $E_i, 1 \leq i \leq n$, with a common distribution function $F_E(x) = 1 - e^{-x}, x > 0$.

$$\lim_{t \rightarrow \infty} \frac{Q(1 - 1/tx)/Q(1 - 1/t) - x^\xi}{A(t)} = x^\xi \frac{x^\rho - 1}{\rho}, \quad (\text{A.3})$$

for all $x > 1$, where $\xi > 0$ and ρ are, respectively, positive and negative real numbers and the function $A(t)$ satisfies for any $x > 0$,

$$\lim_{t \rightarrow \infty} \frac{A(tx)}{A(t)} = x^\rho. \quad (\text{A.4})$$

Lemma 1 (see Balakrishnan and Basu [38]). *Considering the convention that $E_{0:n} \equiv 0$, we have*

$$E_{j:n} - E_{i:n} \stackrel{d}{=} E_{j-i:n-i}, \quad 1 \leq i \leq j \leq n. \quad (\text{A.1})$$

Lemma 2 (see Girard [39]). *Suppose k is an intermediate sequence, i.e., (11) holds. Then,*

$$\frac{E_{n-i+1:n}}{\ln(n/i)} \xrightarrow{p} 1, \quad i = 1, \dots, k. \quad (\text{A.2})$$

Lemma 3 (see Santos et al. [19]). *Assume that the quantile function Q satisfies the following second order regular variation condition*

Then, for any intermediate sequence k , such as $k = k(n) \rightarrow \infty$ and $k/n \rightarrow 0$,

$$\frac{1}{k} \sum_{i=1}^k \ln \frac{X_{n-i+1:n} - X_{[np]+1:n}}{X_{n-k:n} - X_{[np]+1:n}} - \frac{1}{k} \sum_{i=1}^k \ln \frac{X_{n-i+1:n} - \chi_p}{X_{n-k:n} - \chi_p} = o_p\left(\frac{1}{Q(1 - k/n)}\right), \quad (\text{A.5})$$

where χ_p denotes the quantile of order p ($0 \leq p < 1$) for the random variable X .

A.2. Proof of Proposition 1. First notice that condition (A.3) holds for the quantile function in (3) and $X_{1:n} \xrightarrow{p} \chi_0 = Q_X(0 | \alpha, \mu, \sigma) = \mu$. Then, using Lemma 6,

$$H^*(k) = \frac{1}{k} \sum_{i=1}^k \ln \frac{X_{n-i+1:n} - X_{1:n}}{X_{n-k:n} - X_{1:n}} \stackrel{d}{=} \frac{1}{k} \sum_{i=1}^k \ln \frac{X_{n-i+1:n} - \mu}{X_{n-k:n} - \mu}. \quad (\text{A.6})$$

Let us now define the random variable $X^* = (X - \mu)/\sigma, i = 1, \dots, n$. From the inverse probability integral transform,

we have that $X^* \stackrel{d}{=} Q_{X^*}(1 - e^{-E}) = (e^E(1 - e^{-E}))^{1/\alpha} \sim \text{llogist}(\alpha, 0, 1)$ and $X_{i:n}^* \stackrel{d}{=} Q_{X^*}(1 - e^{-E_{i:n}}), 1 \leq i \leq n$. We then have

$$\begin{aligned} H^*(k) &\stackrel{d}{=} \frac{1}{k} \sum_{i=1}^k \ln \frac{X_{n-i+1:n}^*}{X_{n-k:n}^*} \stackrel{d}{=} \frac{1}{k} \sum_{i=1}^k \ln \frac{Q_{X^*}(1 - e^{-E_{n-i+1:n}})}{Q_{X^*}(1 - e^{-E_{n-k:n}})} \\ &= \frac{1}{k\alpha} \sum_{i=1}^k \ln \frac{e^{E_{n-i+1:n}}(1 - e^{-E_{n-i+1:n}})}{e^{E_{n-k:n}}(1 - e^{-E_{n-k:n}})} \\ &= \frac{1}{\alpha} \left(\frac{1}{k} \sum_{i=1}^k (E_{n-i+1:n} - E_{n-k:n}) + \frac{1}{k} \sum_{i=1}^k \ln \frac{1 - e^{-E_{n-i+1:n}}}{1 - e^{-E_{n-k:n}}} \right). \end{aligned} \quad (\text{A.7})$$

Using Lemma 4 and the Taylor approximation $\ln(1 + x) \sim x$, as $x \rightarrow 0$

$$\begin{aligned} H^*(k) &\stackrel{d}{=} \frac{1}{\alpha} \left(\frac{1}{k} \sum_{i=1}^k E_{k-i+1:k} + \frac{1}{k} \sum_{i=1}^k (e^{-E_{n-k:n}} - e^{-E_{n-i+1:n}}) \right) \\ &= \frac{1}{\alpha} \left(\frac{1}{k} \sum_{i=1}^k E_i + \frac{1}{k} \sum_{i=1}^k e^{-E_{n-k:n}} (1 - e^{-E_{n-i+1:n} + E_{n-k:n}}) \right) \\ &\stackrel{d}{=} \frac{1}{\alpha} \left(\frac{1}{k} \sum_{i=1}^k E_i + e^{-E_{n-k:n}} \frac{1}{k} \sum_{i=1}^k (1 - e^{-E_{k-i+1:k}}) \right) \\ &= \frac{1}{\alpha} \left(\frac{1}{k} \sum_{i=1}^k E_i + e^{-E_{n-k:n}} \left(1 - \frac{1}{k} \sum_{i=1}^k e^{-E_i} \right) \right). \end{aligned} \quad (\text{A.8})$$

From the weak law of large numbers, $(1/k) \sum_{i=1}^k e^{-E_i} \xrightarrow{p} 1/2$. Also using Lemmas 4 and 5, we have

$$H^*(k) \stackrel{d}{=} \frac{1}{\alpha} \left(1 + \frac{Z_k}{\sqrt{k}} + \frac{1}{2} e^{-\ln(n/k)} (1 + o_p(1)) \right), \quad (\text{A.9})$$

where $Z_k = \sqrt{k}(1/k) \sum_{i=1}^k (E_i - 1)$ is an asymptotic standard normal random variable by the central limit theorem. After an application of the Taylor series, (13) and (14) follow straightforwardly.

Data Availability

No data were used to support this article.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work is funded by national funds through the FCT-Fundação para a Ciência e a Tecnologia, I.P., under the scope of the projects UIDB/00297/2020 and UIDP/00297/2020 (Center for Mathematics and Applications).

References

- [1] N. Balakrishnan, H. Malik, and S. Puthenpura, "Best linear unbiased estimation of location and scale parameters of the log-logistic distribution," *Communications in Statistics - Theory and Methods*, vol. 16, no. 12, pp. 3477–3495, 1987.
- [2] B. C. Arnold, "Pareto distributions," *Monographs on Statistics and Applied Probability Chapman and Hall/CRC*, Chapman and Hall/CRC, 2015.
- [3] I. W. Burr, "Cumulative frequency functions," *The Annals of Mathematical Statistics*, vol. 13, no. 2, pp. 215–232, 1942.
- [4] P. R. Fisk, "The graduation of income distributions," *Econometrica*, vol. 29, no. 2, pp. 171–185, 1961.
- [5] G. M. Cordeiro, M. Alizadeh, and E. M. M. Ortega, "The exponentiated half-logistic family of distributions: properties and applications," *Journal of Probability and Statistics*, vol. 2014, Article ID 864396, 21 pages, 2014.
- [6] G. M. Cordeiro, M. Alizadeh, and P. R. D. Marinho, "The type I half-logistic family of distributions," *Journal of Statistical Computation and Simulation*, vol. 86, no. 4, pp. 707–728, 2015.
- [7] M. Alizadeh, A. Nematollahi, E. Altun, and M. Rasekhi, "A study on a new type 1 half-logistic family of distributions and its applications," *Statistics, Optimization and Information Computing*, vol. 8, no. 4, pp. 934–949, 2020.
- [8] G. S. Mohammad, "A new two-parameter modified half-logistic distribution: properties and applications," *Statistics, Optimization & Information Computing*, 2021.
- [9] A. J. Lemonte, "The beta log-logistic distribution," *Brazilian Journal of Probability and Statistics*, vol. 28, no. 3, pp. 313–332, 2014.
- [10] M. K. Shakhatareh, "A new three-parameter extension of the log-logistic distribution with applications to survival data," *Communications in Statistics-Theory and Methods*, vol. 47, no. 21, pp. 5205–5226, 2017.
- [11] M. Ahsanullah and A. Alzaatreh, "Parameter estimation for the log-logistic distribution based on order statistics," *REVSTAT*, vol. 16, no. 4, pp. 429–443, 2018.
- [12] P. Hall, "On some simple estimates of an exponent of regular variation," *Journal of the Royal Statistical Society: Series B*, vol. 44, no. 1, pp. 37–42, 1982.
- [13] J. Beirlant, K. Herrmann, and J. Teugels, *Estimation of the extreme value index*, JohnWiley & Sons, Inc., 2016.
- [14] I. Fedotenkov, "A review of more than one hundred Pareto-tail index estimators," *Statistica*, vol. 80, no. 3, pp. 245–299, 2020.
- [15] B. M. Hill, "A simple general approach to inference about the tail of a distribution," *The Annals of Statistics*, vol. 3, no. 5, pp. 1163–1174, 1975.
- [16] M. I. Gomes and O. Oliveira, "How can non-invariant statistics work in our benefit in the semi-parametric estimation of parameters of rare events," *Communications in Statistics - Simulation and Computation*, vol. 32, no. 4, pp. 1005–1028, 2003.
- [17] F. Caeiro and M. I. Gomes, "Minimum-variance reduced-bias tail index and high quantile estimation," *Revstat*, vol. 6, no. 1, pp. 1–20, 2008.
- [18] M. Fraga-Alves, "A location invariant Hill-type estimator," *Extremes*, vol. 4, no. 3, pp. 199–217, 2001.
- [19] P. Araújo Santos, M. Fraga Alves, and M. I. Gomes, "Peaks over random threshold methodology for tail index and high quantile estimation," *Revstat*, vol. 4, no. 3, pp. 227–247, 2006.
- [20] C. Caballero-Megido, J. Hillier, D. Wyncoll, L. Boshier, and B. Gouldby, "Technical note: comparison of methods for threshold selection for extreme sea levels," *Journal of Flood Risk Management*, vol. 11, no. 2, pp. 127–140, 2017.
- [21] F. Caeiro and M. Gomes, "Threshold selection in extreme value analysis," in *Extreme Value Modeling and Risk Analysis*, D. K. Dey and J. Yan, Eds., pp. 69–86, Chapman and Hall/CRC, 2015.
- [22] C. Scarrott and A. MacDonald, "A review of extreme value threshold estimation and uncertainty quantification," *Revstat*, vol. 10, no. 1, pp. 33–60, 2012.
- [23] B. Bader, J. Yan, and X. Zhang, "Automated threshold selection for extreme value analysis via ordered goodness-of-fit tests with adjustment for false discovery rate," *The Annals of Applied Statistics*, vol. 12, no. 1, 2018.
- [24] H. Drees, "Refined pickands estimators with bias correction," *Communications in Statistics - Theory and Methods*, vol. 25, no. 4, pp. 837–851, 1996.
- [25] L. Peng, "Asymptotically unbiased estimators for the extreme-value index," *Statistics & Probability Letters*, vol. 38, no. 2, pp. 107–115, 1998.
- [26] J. Beirlant, G. Dierckx, Y. Goegebeur, and G. Matthys, "Tail index estimation and an exponential regression model," *Extremes*, vol. 2, no. 2, pp. 177–200, 1999.
- [27] A. Feuerverger and P. Hall, "Estimating a tail exponent by modelling departure from a Pareto distribution," *The Annals of Statistics*, vol. 27, no. 2, pp. 760–781, 1999.
- [28] J. Beirlant, F. Caeiro, and M. I. Gomes, "An overview and open research topics in statistics of univariate extremes," *Revstat*, vol. 10, no. 1, pp. 1–31, 2012.
- [29] M. I. Gomes and A. Guillou, "Extreme value theory and statistics of univariate extremes: a review," *International Statistical Review*, vol. 83, no. 2, pp. 263–292, 2015.
- [30] M. I. Gomes, M. J. Martins, and M. Neves, "Alternatives to a semi-parametric estimator of parameters of rare events—the jackknife methodology," *Extremes*, vol. 3, no. 3, pp. 207–229, 2000.
- [31] J. Danielsson, D. W. Jansen, and C. G. de Vries, "The method of moments ratio estimator for the tail shape parameter," *Communications in Statistics - Theory and Methods*, vol. 25, no. 4, pp. 711–720, 1996.
- [32] H. Penalva, M. I. Gomes, F. Caeiro, and M. M. Neves, "A couple of non reduced bias generalized means in extreme value theory: an asymptotic comparison," *Revstat*, vol. 18, no. 3, pp. 281–298, 2020.
- [33] H. Penalva, M. I. Gomes, F. Caeiro, and M. M. Neves, "Lehmer's mean-of-order-p extreme value index estimation: a simulation study and applications," *Journal of Applied Statistics*, vol. 47, no. 13–15, pp. 2825–2845, 2020.
- [34] P. Jordanova, Z. Fabián, P. Hermann et al., "Weak properties and robustness of t-Hill estimators," *Extremes*, vol. 19, no. 4, pp. 591–626, 2016.
- [35] M. Stehlik, J. Kiseľák, M. Vaičiulis et al., "Priority statement and some properties of t-lgHill estimator," *Extremes*, vol. 23, no. 3, pp. 493–499, 2020.
- [36] F. Caeiro, A. Mateus, and L. Soltane, "A class of weighted Hill estimators," *Computational and Mathematical Methods*, 2021.

- [37] R Core Team, *R: a language and environment for statistical computing*, R Foundation for Statistical Computing, Vienna, Austria, 2020.
- [38] N. Balakrishnan and A. P. Basu, *Exponential Distribution: Theory, Methods and Applications*, CRC Press, 1995.
- [39] S. Girard, "A Hill type estimator of the Weibull tail-coefficient," *Communications in Statistics - Theory and Methods*, vol. 33, no. 2, pp. 205–234, 2004.