Received: 12 September 2019 | Revised: 1 September 2020 | Accepted: 12 October 2020

DOI: 10.1002/cmm4.1133

RESEARCH ARTICLE

WILEY

A new class of estimators for the shape parameter of a Pareto model

Ayana Mateus[1,2] | **Frederico Caeiro**[1,2]

[1]Faculdade de Ciências e Tecnologia (FCT), Universidade Nova de Lisboa, Caparica, Portugal

[2]Centro de Matemática e Aplicações (CMA), Universidade Nova de Lisboa, Caparica, Portugal

Correspondence
Ayana Mateus, Faculdade de Ciências e Tecnologia, 2829-516 Caparica, Portugal.
Email: amf@fct.unl.pt

Funding information
Fundação para a Ciência e a Tecnologia, Grant/Award Number: Research partially supported by National Funds thr

Abstract

The Pareto model, first used in socioeconomic problems, has successfully been applied in many other areas such as astronomy, biology, bibliometrics, demography, insurance, or risk management. Although there are several variants of this distribution, the current study focuses on the classic Pareto distribution, also known as the Pareto type I distribution. We propose a new class of estimators for the Pareto shape parameter, obtained through a modification of the probability weighted moment method, called the log generalized probability weighted moments method. In addition to the asymptotic distribution, Monte Carlo simulations were performed to analyze the finite sample behavior of the proposed new estimators. A comparison with the most used estimators, such as the moment, the maximum likelihood, the least squares, and the probability weighted moments estimators was also performed. In addition, the estimators were used to construct asymptotic confidence intervals. To illustrate an application of the different estimation methods to a real data set from a clinical trial complete the article. Results indicate an overall good performance of the new proposed class.

KEYWORDS

asymptotic distribution, confidence interval, least squares, maximum likelihood, moments, Pareto distribution

1 | INTRODUCTION

The Pareto distribution was initially introduced by Vilfredo Pareto in 1897[1] to describe the income distribution among individuals, showing that the number of taxpayers with an income greater than x could be approximated by b/x^a, for some vector of positive components (a, b). More recently, this distribution has been extensively used in many other areas such as astronomy,[2] bibliometrics,[3] biology,[4] demography,[5] insurance,[6] and risk management.[7] Although there are several variants of this model, the current study applies the classic Pareto distribution, also known as Pareto type I[8-11] or Power-Law,[12,13] with distribution function (d.f.) given by

$$F(x) = 1 - \left(\frac{x}{c}\right)^{-a}, \quad x > c, \quad c > 0, \quad a > 0, \tag{1}$$

and with a quantile function

$$Q(p) = F^{\leftarrow}(p) = c(1-p)^{-1/a}, \quad 0 < p < 1, \quad c > 0, \quad a > 0, \tag{2}$$

where c and a are the scale and shape parameters respectively, and p denotes the lower tail probability. In the following we consider both parameters a and c unknown and the notation $X \sim P(a, c)$ whenever X has the d.f. in (1). The a parameter, also known as the tail index or Pareto index, measures the heaviness of the right tail. Smaller values of the tail index corresponds to heavier tails. The scale parameter corresponds to the left endpoint of the support of X. This distribution is of great importance since in many models the upper tail is asymptotically Paretian. More precisely, the following relation holds

$$F(x) = 1 - x^{-a}L(x), \quad x \to \infty, \tag{3}$$

for some slowly variation function L,[14] that is, a function L which satisfies

$$\lim_{t\to\infty} \frac{L(tx)}{L(t)} = 1, \quad \forall x > 0.$$

Distributions such as the generalized Pareto, Student's t or the Burr (type XII) are examples of models with a Pareto-type tail. Thus

$$\lim_{t\to\infty} \frac{1 - F(tx)}{1 - F(t)} = x^{-a}, \quad \forall x > 0,$$

and this class of Pareto-type models are in the maximum domain of attraction of the Fréchet distribution. Under the semiparametric framework in (3) the so-called extreme value index parameter, $\xi = 1/a$, is commonly used. Details on the estimation of ξ for models with a Pareto-type tail can be found in References 15-18, among others. Also, notice that the Pareto distribution is a special case of the generalized Pareto distribution (GPD), with d.f.

$$F(x) = 1 - \left(1 + \xi\left(\frac{x - \mu}{\sigma}\right)\right)^{-1/\xi}, \quad x > \mu, \quad \sigma > 0, \quad \xi \in \mathbf{R}.$$

Estimation of shape and scale parameters in (1) has already been extensively addressed in the literature: Quandt[19] compared the maximum likelihood estimator with the moments estimator, a least squares estimator, and four percentile estimators. Lu and Tau, Caeiro et al.,[20] and Kim et al.[21] studied several least squares estimators and Brazauskas and Serfling[22] and Vandewalle et al.[23] introduced robust estimators of the shape parameter. Bayesian estimators can be found in Arnold and Press,[24] Rasheed and Al-Gazi,[25] and Han[26]. Caeiro and Gomes[27] and Munir et al.[28] considered probability weighted moment estimators. Dalpatadu and Singh[29] proposed estimators based on minimization of a distance function. Modified percentile and maximum likelihood estimators were studied in Bhatti et al..[30,31] The estimation of a parametric function of the scale and shape parameters was studied in Barranco-Chamorro and Jiménez-Gamero[5].

The maximum likelihood estimators have a simple closed form and are often used in applications. However, their optimal efficiency is only valid asymptotically, meaning that for a small or moderate sample size other estimation methods could be more efficient in terms of bias or mean squared error. In this article we introduce a new class of estimators for the shape parameter a. These new estimators, called log generalized probability weighted moment estimators, are obtained through a modification of the probability weighted moments method.

The present article is organized as follows. In Section 2, we briefly describe the most used estimation methods for the parameters of a Pareto distribution, namely the moment, the maximum likelihood, the least squares, and the probability weighted moments. Next, we introduce a new class of estimators and derive the asymptotic nondegenerated distribution of all estimators under study. In Section 3 we have performed a Monte Carlo simulation study to compare the finite sample behavior of the new class of estimators with the most used estimators. In Section 4, we derive asymptotic confidence intervals for the shape parameter based on the different methods of estimation considered in this study and compare their performance via their coverage probabilities. To illustrate the applicability of our results a real data application is provided in Section 5. Section 6 concludes our article with some conclusions about the performance of the different estimation methods.

2 | ESTIMATORS UNDER STUDY

In this section we review the most common estimation methods and introduce the new method of estimation. In the following, we shall assume that $X_1, X_2, \ldots, X_n$ is a sample of size n of independent and identically distributed random variables, with a common Pareto distribution, given in (1). The corresponding sample of nondecreasing order statistics is denoted by $X_{1:n}, X_{2:n}, \dots, X_{n:n}$.

2.1 | Maximum likelihood estimators

The density function of a $P(a, c)$ random variable X is

$$f(x) = \frac{ac^a}{x^{a+1}}, \quad x > c > 0, \quad a > 0. \tag{4}$$

Hence, the likelihood function is given by

$$L(a, c|X_1, \dots, X_n) = \prod_{i=1}^{n} \frac{ac^a}{X_i^{a+1}}. \tag{5}$$

The maximum likelihood (ML) estimators of the parameters a and c are obtained by maximization of the log-likelihood function, in (5), and are given by,

$$\hat{a}^{ML} = \frac{1}{\frac{1}{n}\sum_{i=1}^{n} \ln X_i - \ln X_{1:n}}, \quad \text{and} \quad \hat{c}^{ML} = X_{1:n}. \tag{6}$$

The ML estimators are consistent and are asymptotically the most efficient estimators. However, for small sample sizes, they can be outperformed by other estimators in terms of bias or root mean squared error. Quandt[19] concluded that for small sample sizes the ML estimators and quantile-type estimators were reported to be the best compared with the moment and least squares estimators. In 2007, Lu and Tao[32] proposed a weighted least square estimator and found its performance to be better or close to the ML estimators when the sample size is smaller than 100.

2.2 | Least squares estimators

There are several estimators based on a regression framework (see References 19,21,32). We shall consider the estimators obtained through a regression approach based on the quantile function. By taking the natural logarithm of the quantile function in (2), we have a linear relation between the $\ln Q(p)$ and the parameters β_0 and β_1,

$$\ln Q(p) = \beta_0 - \beta_1 \ln(1 - p), \quad \text{with} \quad \beta_0 = \ln c \text{ and } \beta_1 = a^{-1}. \tag{7}$$

The quantile function can be easily estimated through the empirical estimator defined by $\hat{Q}(p) = X_{i:n}$, for $p \in]\frac{i-1}{n}, \frac{i}{n}]$ and the least squares estimators (LS) of the parameters β_0 and β_1 are obtained by the minimization of the sum of squares,

$$\sum_{i=1}^{n} \left\{ \ln X_{i:n} - \beta_0 - \beta_1 \left(-\ln \left(1 - \frac{i}{n+1} \right) \right) \right\}^2$$

and are given by

$$\hat{\beta}_1^{LS} = \frac{\sum_{i=1}^{n} -\ln\left(1 - \frac{i}{n+1}\right) \left\{ n \ln X_{i:n} - \sum_{j=1}^{n} \ln X_{j:n} \right\}}{n \sum_{i=1}^{n} \left(\ln\left(1 - \frac{i}{n+1}\right)\right)^2 - \left(\sum_{i=1}^{n} -\ln\left(1 - \frac{i}{n+1}\right)\right)^2} \quad \text{and} \quad \hat{\beta}_0^{LS} = \frac{1}{n} \sum_{i=1}^{n} \ln X_{i:n} + \frac{\hat{\beta}_1^{LS}}{n} \sum_{i=1}^{n} -\ln\left(1 - \frac{i}{n+1}\right). \tag{8}$$

The estimators of a and c follows straightforwardly from the relation in (7).

2.3 | Moment estimators

The noncentral moments of order k of the Pareto distribution are given by

$$E(X^k) = \frac{ac^k}{a-k}, \quad \text{if } a > k. \tag{9}$$

The method of moments based on the two first moments is not very popular in applications. This is because the second moment only exist for $a > 2$ and because other moment-based estimators have appeared in the literature. The most used moment estimators are the ones proposed by Quandt.[19] By equating the first noncentral moment, given in (9), to the sample mean it follows that

$$\hat{a} = \frac{\overline{X}}{\overline{X} - \hat{c}},$$

where $\hat{c}$ is the estimator for c and $\overline{X}$ denotes the arithmetic sample mean. To extend the interval of values where the estimators based on moments are valid, Quandt proposed the use of the sample minimum which has a Pareto distribution with shape and scale parameters given by (an, c), respectively. By equating the moment of the sample minimum,

$$E(X_{1:n}) = \frac{anc}{an-1}$$

to the minimum sample value, Quandt derived the moments (M) estimators

$$\hat{a}^M = \frac{n\overline{X} - X_{1:n}}{n(\overline{X} - X_{1:n})}, \quad \text{and} \quad \hat{c}^M = \left(1 - \frac{1}{n\hat{a}^M}\right) X_{1:n}. \tag{10}$$

Remark 1. Since the first noncentral moment only exists for $a > 1$, the moment estimators in (10) are only consistent for $a > 1$.

2.4 | Probability weighted moment estimators

The method of probability weighted moment (PWM) was introduced in Greenwood et al.[33] as a useful tool to estimate distribution parameters from models that have a well-defined quantile function. It is nowadays a well-established estimation method in the field of hydrology.[34] The PWMs of a continuous random variable X with distribution function F are defined as

$$M_{k,r,s} = E(X^k (F(X))^r (1 - F(X))^s), \tag{11}$$

where k, r, and s are real numbers. The PWM estimators are obtained by equating $M_{k,r,s}$ with their corresponding sample moments, and then by solving those equations for the unknown parameters. This method is a generalization of the classic method of moments. Indeed, when $r = s = 0$, $M_{k,0,0}$ are the usual noncentral moments of order k.

Greenwood et al.[33] and Hosking et al.[35] advise the use of $M_{1,r,s} = E(X(F(X))^r(1 - F(X))^s)$, because the relations between parameters and moments have usually a much simpler form. Another advantage is that if the mean value, $E(X) = M_{1,0,0}$ exists, then $M_{1,r,s}$ exists for any real $r > 0$ and $s > 0$. It is also known that the empirical estimate of $M_{1,r,s}$ is usually less sensitive to outliers and has good properties when the sample size is small. Although the constants k, r, and s in (11) can take any real value, for convenience, many authors have considered $k = 1$ and the first nonnegative integer values for r and s. We shall refer to this approach as the classic PWM method. Also, when r and s are nonnegative integers it is more convenient to work with the PWMs

$$\alpha_r = M_{1,0,r} = E(X(1 - F(X))^r), \quad r = 0, 1, \dots, \tag{12}$$

or

$$\beta_r = M_{1,r,0} = E(X(F(X))^r) \quad r = 0, 1, \dots. \tag{13}$$

Notice that for nonnegative integers r and s, $F(X)^r(1-F(X))^s$ can be written as a linear combination of powers of $F(X)$ or $1-F(X)$. As a consequence we can relate α_r and β_r through the equations

$$\alpha_r = \sum_{j=0}^{r}(-1)^j \binom{r}{j} \beta_j \quad \text{and} \quad \beta_r = \sum_{j=0}^{r}(-1)^j \binom{r}{j} \alpha_j,$$

where $\binom{r}{j}$ denotes the binomial coefficient. Therefore, it is equivalent to work with α_r or β_r given that the values for r are chosen as small as possible. For nonnegative integer values of r, the unbiased estimators of the PWMs α_r and β_r in (12) and (13) are respectively,[36]

$$\hat{\alpha}_r = \frac{1}{n}\sum_{i=1}^{n-r}\frac{\binom{n-i}{r}}{\binom{n-1}{r}}X_{i:n} \quad \text{and} \quad \hat{\beta}_r = \frac{1}{n}\sum_{i=r+1}^{n}\frac{\binom{i-1}{r}}{\binom{n-1}{r}}X_{i:n}. \tag{14}$$

Landwehr et al.[37] noticed empirically that moderated biased estimators of the PWMs could provide more efficient estimates of upper quantiles. The biased estimators of the PWMs α_r and β_r are respectively,

$$\tilde{\alpha}_r = \frac{1}{n}\sum_{i=1}^{n}(1-p_{i:n})^r X_{i:n} \quad \text{and} \quad \tilde{\beta}_r = \frac{1}{n}\sum_{i=1}^{n} p_{i:n}^r X_{i:n}, \tag{15}$$

where $p_{i:n}$ are the so-called plotting positions, that is, empirical estimates of $F(X_{i:n})$. The most common choices for the plotting positions are

$$p_{i:n} = \frac{i-b}{n}, \quad 0 \le b \le 1 \tag{16}$$

and

$$p_{i:n} = \frac{i-b}{n+1-2b}, \quad -0.5 \le b \le 0.5, \tag{17}$$

where b is a continuity correction factor.

Landwehr et al.[37] confirmed with a Monte Carlo simulation study that, for small sample sizes, the PWM estimators compare favorably to other estimation methods.

For the Pareto distribution under study, the PWMs in (11) have the following analytical expression

$$M_{k,r,s} = c^k\, \mathrm{B}\left(s+1-\frac{k}{a}, r+1\right), \quad s-\frac{k}{a} > -1, \quad r > -1,$$

where B stands for the complete beta function.

Caeiro and Gomes[27] and Caeiro et al.[38] considered the PWMs, α_s in (12) and obtained the correspondents PWMs for the Pareto distribution:

$$\alpha_s = M_{1,0,s} = \frac{c}{\left(s+1-\frac{1}{a}\right)}, \quad a > \frac{1}{1+s}.$$

By considering the two PWMs $\alpha_0 = \frac{c}{1-1/a}$ and $\alpha_1 = \frac{c}{2-1/a}$ and equating them with their corresponding sample moments,

$$\hat{\alpha}_0 = \overline{X} \quad \text{and} \quad \hat{\alpha}_1 = \frac{1}{n}\sum_{i=1}^{n}\frac{n-i}{n-1}X_{i:n}, \tag{18}$$

the corresponding PWM estimators for the shape and scale parameters of the Pareto distribution are given by

$$\hat{a}^{PWM} = \frac{\hat{\alpha}_0-\hat{\alpha}_1}{\hat{\alpha}_0-2\hat{\alpha}_1} \quad \text{and} \quad \hat{c}^{PWM} = \frac{\hat{\alpha}_0\hat{\alpha}_1}{\hat{\alpha}_0-\hat{\alpha}_1} \quad a > 1. \tag{19}$$

A comparison of the PWM estimators in (19) with other estimation methods can be found in Reference 28.

2.5 | The new class of extended PWM estimators

Although the theoretical PWMs in (11) is defined for any real exponents k, r, and s, first applications have considered nonnegative integers exponents. Rasmussen[39] studied PWMs with real exponents and called the method generalized PWM (GPWM) to distinguish them from the classic PWM method. He concluded that in most cases the GPWM method outperforms the classic PWM method. For the sake of simplicity it is advisable to restrict the class of GPWMs by setting $(k,r,s)=(1,0,s),\ s\in\mathbf{R}$ or $(k,r,s)=(1,r,0),\ r\in\mathbf{R}$. Such restriction allows us to work with much simpler analytical formulas for the GPWMs. The GPWM estimators are thus the ones in (15) for any real value of r.

Another modification of the PWM method was introduced by Caeiro and Prata Gomes.[40] The authors worked under a framework of Pareto-type tails and considered a different class of PWMs, given by

$$M'_{g,r,s} = E(g(X)(F(X))^r(1-F(X))^s) \tag{20}$$

with $g(x)=\ln(x)$. Such class of PWMs was called Log PWM (LPWM) and, for the Pareto model, has the advantage of extending the domain of validity of the estimators to the complete parameter space. Caeiro and Mateus[41] considered the LPWMs in (20) with $r=0$ and derived the corresponding LPWMs for the Pareto model,

$$l'_s = M'_{\ln,0,s} = \frac{\ln(c)}{1+s} + \frac{1}{a(1+s)^2}, \quad s>-1. \tag{21}$$

If we consider the LPWMs l'_0 and l'_1, the corresponding LPWM estimators of the shape and scale parameters of the Pareto distribution in (1) are respectively

$$\hat{a}^{LPWM} = \frac{1}{2\hat{l}_0 - 4\hat{l}_1} \quad \text{and} \quad \hat{c}^{LPWM} = \exp(4\hat{l}_1 - \hat{l}_0), \tag{22}$$

where $\hat{l}_s$, $s=0,1$ are the unbiased empirical estimator of l_s given by

$$\hat{l}'_s = \frac{1}{n}\sum_{i=1}^{n-s} \frac{\binom{n-i}{s}}{\binom{n-1}{s}} \ln X_{i:n}. \tag{23}$$

Recently, Chen et al.[42] proposed a wider class of GPWM (called extended version of GPWM) by considering the PWMs in (20) with $g(\cdot)$ a suitable measurable function and r and s being real values.

In the following we propose a new class of estimators for the shape parameter of a Pareto model based on a trivial characterization of this distribution. This characterization will allow the estimation of the shape parameter a, separately of c, rather than estimating simultaneous both parameters (a,c). Since this approach only provides an estimator for the a parameter, we consider the scale parameter c estimated by maximum likelihood. It is know that left truncation in a Pareto distribution is equivalent to a rescaling (see section 3.9.2 of Arnold[11]). More precisely,

$$P(X \le x | X > t) = 1 - \left(\frac{x}{t}\right)^{-a}, \quad x > t \ge c.$$

As a consequence, the relative excess X/t given $X>t$ has a Pareto distribution with the same shape parameter as X and scale parameter equal to 1 $(X/t|X>t)\sim P(a,1)$. To use all values from the random sample, we choose $t=X_{1:n}$. Notice that the sample minimum is consistent for c ($X_{1:n} \overset{P}{\to} c$, where $\overset{P}{\to}$ represents the convergence in probability). Next we explain how one can base the estimation of the shape parameter on the sample values of $X^* = \frac{X}{X_{1:n}}|(X > X_{1:n}) \sim P(a,1)$, which are independent of the scale parameter c. We shall assume without loss of generality that the random sample $(X_1, X_2, \dots, X_n)$ has $n-1$ values greater than $X_{1:n}$, that is, we assume $X_{1:n} < X_{2:n}$. Then for a real value s, we shall define the log generalized probability weighted moment (LGPWM) as

$$l^*_s = E(\ln(X^*)(1-F(X^*)^s)) = \frac{1}{a(1+s)^2}, \quad s>-1. \tag{24}$$

The corresponding (biased) estimator of the moment in (24) is,

$$\tilde{\alpha}_s^* = \frac{1}{n-1}\sum_{i=2}^{n}(1-p_{i-1:n-1})^s \ln X_{i-1:n-1}^*, \quad s > -1,$$

where $p_{i:n}$ is the aforementioned empirical estimate of $F(X_{i:n})$.

Solving the equation $l_s = \tilde{\alpha}_s^*$ in order of the parameter a, we obtain the class of log generalized probability weighted moment estimator for the shape parameter of the Pareto distribution,

$$\hat{a}^{LGPWM} = \frac{1}{(1+s)^2\tilde{\alpha}_s^*} = \frac{1}{\frac{(1+s)^2}{n-1}\sum_{i=2}^{n}\left(1-\frac{i-1}{n-1}\right)^s \ln X_{i-1:n-1}^*}, \quad s > -1. \tag{25}$$

This class of estimators depends on a tuning parameter $s > -1$. When $s = 0$ we obtain an estimator asymptotically equivalent to the ML estimator for the shape parameter. For the estimation of the scale parameter we can consider the maximum likelihood estimator in (6) or the bias corrected estimator

$$\hat{c}^{LGPWM} = \left(1 - \frac{1}{n\hat{a}^{LGPWM}}\right)X_{1:n}. \tag{26}$$

2.6 | Limiting distribution of the shape parameter estimators

To understand the behavior of the estimators of the shape parameter and to compare the new class of the LGPWM estimators with other estimators for the same parameter, it is important to know their sampling distribution. Since it is difficult to obtain the finite sample distribution of several of the estimators here considered, we shall derive their asymptotic or limiting nondegenerate distribution. For small sample sizes, the asymptotic results must be used carefully because they may be inaccurate.[43] The ML, LS, M, and PWM estimators of the parameter a in (6), (8), (10), and (19), respectively, are already studied in the literature and their corresponding limit behavior are presented without proof.

Proposition 1. *Suppose that $(X_1, X_2, \ldots, X_n)$ is an i.i.d. sample from Pareto population with d.f. in (1). Then*

$$\sqrt{n}(\hat{a}^{ML} - a) \underset{n\to\infty}{\overset{d}{\to}} N(0, a^2), \tag{27}$$

$$\sqrt{n}(\hat{a}^{LS} - a) \underset{n\to\infty}{\overset{d}{\to}} N(0, 2a^2), \tag{28}$$

$$\sqrt{n}(\hat{a}^{M} - a) \underset{n\to\infty}{\overset{d}{\to}} N\left(0, \frac{a(a-1)^2}{a-2}\right), \quad \text{if } a > 2, \tag{29}$$

and

$$\sqrt{n}(\hat{a}^{PWM} - a) \underset{n\to\infty}{\overset{d}{\to}} N\left(0, \frac{a(a-1)(2a-1)^2}{(a-2)(3a-2)}\right), \quad \text{if } a > 2, \tag{30}$$

where $N(\mu, \sigma^2)$ denotes a normal random variable with mean value μ and variance σ^2 and $\overset{d}{\to}$ means convergence in distribution.

Remark 2. Since the M and the PWM estimators are based on moments of the random variable X, the asymptotic distribution is only valid if $a > 2$.

Before providing the main result related to the asymptotic behavior of the LGPWM shape estimator in (25) we first need to establish a few lemmas.

Lemma 1 (Kleiber and Kotz,[9] section 3). *Let $X \sim P(a, c)$. Then*

$$a \ln\left(\frac{X}{c}\right) \overset{d}{=} E,$$

where E denotes a standard exponential random variable with d.f. $F_E(x) = 1 - e^{-x}$, $x > 0$ and $\overset{d}{=}$ stands for equally distributed.

Lemma 2. *Let $(Y_{1:n}, Y_{2:n}, \dots, Y_{n:n})$ denote the ascending order statistics from an i.i.d. sample drawn from a population distributed as $Y \sim P(1,1)$. Then, the following result holds*

$$\frac{Y_{j:n}}{Y_{i:n}} \stackrel{d}{=} Y_{j-i:n-i}, \quad 1 \le i < j \le n. \tag{31}$$

Proof. Note that (31) can be rewritten as $\ln Y_{j:n} - \ln Y_{i:n} \stackrel{d}{=} \ln Y_{j-i:n-i}$ $1 \le i < j \le n$. Using Lemma 1, we obtain $E \stackrel{d}{=} \ln(Y)$. Then, since ln is an increasing continuous function on the interval $(0, \infty)$, we have $E_{i:n} \stackrel{d}{=} \ln(Y_{i:n})$, $1 \le i \le n$, where $E_{i:n}$ denotes the ith ascending order statistic from a sample of size n from the standard exponential distribution. From the properties of the spacings of the exponential order statistics (see Reference [44, section 3]) we know that $E_{j:n} - E_{i:n} \stackrel{d}{=} E_{j-i:n-i}$ and the remaining of the proof follows straightforwardly. ■

Lemma 3. *Let $(X_{1:n}, X_{2:n}, \dots, X_{n:n})$ denote the ascending order statistics from a sample of size n from a Pareto distribution in (1). Then,*

$$\ln \frac{X_{i:n}}{X_{1:n}} \stackrel{d}{=} \frac{1}{a} E_{i-1:n-1}.$$

Proof. The quotient $\ln \frac{X_{i:n}}{X_{1:n}}$ can be rewritten as a quotient of two standard Pareto and using Lemmas 1 and 2, we have

$$\ln \frac{X_{i:n}}{X_{1:n}} = \ln \left(\frac{(X_{i:n}/c)^a}{(X_{1:n}/c)^a} \right)^{\frac{1}{a}} \stackrel{d}{=} \ln (Y_{i-1:n-1})^{\frac{1}{a}} \stackrel{d}{=} \frac{1}{a} E_{i-1:n-1}.$$

■

Next, we derive the main result, the asymptotic nondegenerate limit behavior of the LGPWM shape estimator in (25).

Proposition 2. *For a sample of size n, from a Pareto distribution in (1) and for $\hat{a}^{LGPWM}$ defined in (25), we have,*

$$\sqrt{n}(\hat{a}^{LGPWM} - a) \underset{n\to\infty}{\overset{d}{\to}} N\left(0, \frac{(1+s)^2 a^2}{1+2s}\right), \quad s > -0.5. \tag{32}$$

Proof. Let us first consider the linear function of $X^*_{1:n-1}, X^*_{2:n-1}, \dots, X^*_{n-1:n-1}$

$$T = \frac{1}{n-1} \sum_{i=2}^{n} \left(1 - \frac{i-1}{n-1}\right)^s \ln \frac{X_{i:n}}{X_{1:n}} = \frac{1}{n-1} \sum_{i=1}^{n-1} \left(1 - \frac{i}{n-1}\right)^s \ln X^*_{i:n-1},$$

with $X^*_{i:n-1} = \frac{X_{i+1:n}}{X_{1:n}}$, $i = 1, \dots, n-1$. Consequently, using the result in Lemma 3, we can assure that T has the same distribution of

$$\frac{1}{n-1} \sum_{i=1}^{n-1} \frac{1}{a} \left(1 - \frac{i}{n-1}\right)^s E_{i:n-1}.$$

Using the asymptotic results for linear function of order statistics,[45] we get,

$$\sqrt{n}\,(T - \mu_T) \underset{n\to\infty}{\overset{d}{\to}} N(0, \sigma_T^2),$$

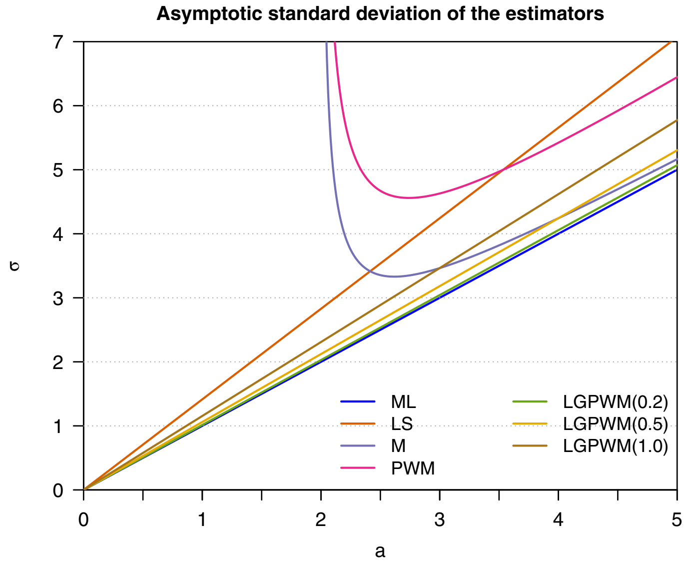
where

$$\mu_T = \frac{1}{a} \int_0^1 (1-x)^s \ln(1-x) dx = \frac{1}{a(1+s)^2} \quad \text{and} \quad \sigma_T^2 = \frac{2}{a^2} \int_0^1 \left(\int_x^1 x(1-x)^{s-1}(1-y)^s dy \right) dx = \frac{1}{a^2(1+s)^2(1+2s)}.$$

Applying the delta method to $\frac{1}{(1+s)^2 T}$, the result in (32) follows straightforwardly. ■

Remark 3. All estimators of the shape parameter studied in this article are asymptotically unbiased and normally distributed. In Figure 1 we plot the asymptotic standard deviation of the estimators under consideration. As expected, $\hat{a}^{ML}$

FIGURE 1 Asymptotic standard deviation of the estimators of the shape parameter



has the smallest asymptotic standard deviation. Furthermore, the asymptotic standard deviation of $\hat{a}^{LGPWM}$ takes the minimum value a when $s = 0$.

3 | FINITE SAMPLE BEHAVIOR OF THE ESTIMATORS

In this section, we have implemented a Monte Carlo simulation experiment in R software environment,[46] with 50,000 samples of sizes $n = 10$, 15, 20, 30, 40, 50, 75, 100, 150, 200, 300, and 500 from the Pareto distribution with parameters $(a, c) = (0.5, 1)$, $(a, c) = (1, 1)$ and $(a, c) = (2, 1)$. The simulation was designed to evaluate the performance of the estimators for the shape parameter, a. The shape parameter was estimated by the PWM, M, ML, LS, and LGPWM methods presented in Section 2. Note that the M and PWM estimators are not consistent when $a = 0.5$. In the LGPWM method, we have considered for the parameter s values between −0.4 and 0.3, discretized in small steps of length 0.1 and the plotting positions in (16), $p_{i:n} = (i - 0.5)/n$. All simulated estimates obtained by the LGPWM method using s negative values as input, result in larger (positive) bias and root mean squared error as opposed to results with $s = 0$. This can be confirmed in Figures 2–5 where we present the simulated mean values and the root mean square error (RMSE) of the LGPWM shape estimator with $s = -0.4, -0.3, -0.2, -0.1$, and 0. Due to this fact we did not consider negative values for the parameter s in the comparison. In Tables 1,2, and 3 we present the simulated mean values and the RMSE, up to four decimal places, of the estimators under study. In the LGPWM estimator we have considered $s = 0.1, 0.2, 0.3$. The "best" mean value and RMSE, for each sample size, are written in bold.

Based on the simulation results, we have reached the following conclusions:

- The M and PWM estimators must be used with care since they always provide a numerical estimate, even when they are not consistent (see Table 1).
- The simulated bias is usually positive. Almost all estimators, except the LS, overestimate the true value of the shape parameter. Also, as expected, the simulated bias falls to zero, as the sample size increases (except when the estimators are inconsistent).
- LS estimator proved to be the best estimator for small sample sizes with regard to the minimum absolute bias.
- Regarding minimum absolute bias for the three Pareto models here used, the LGPWM shape estimator, with $s = 0.2$, was always the best estimator for moderate and large sample sizes.
- The LS shape estimator achieves a minimum RMSE for small sample sizes. For moderate sample sizes the best results are achieved with $s = 0.1$ and for large sample sizes the best results are achieved with the ML.

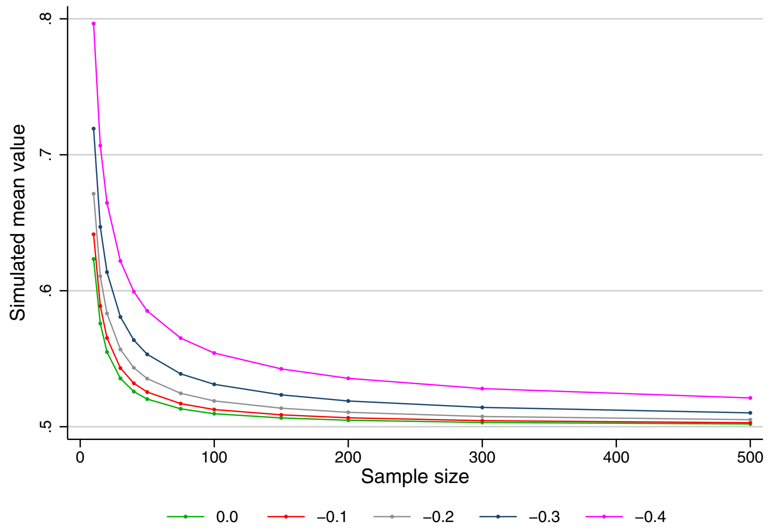


FIGURE 2 Simulated mean value of the LGPWM estimator with $s = -0.4, -0.3, -0.2, -0.1$, and 0 from a $P(0.5, 1)$ model

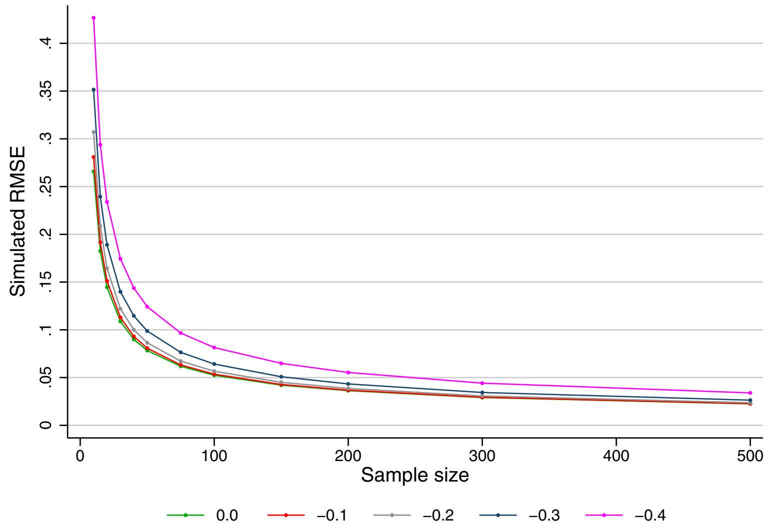


FIGURE 3 Simulated RMSE of the LGPWM estimator with $s = -0.4, -0.3, -0.2, -0.1$, and 0 from a $P(0.5, 1)$ model

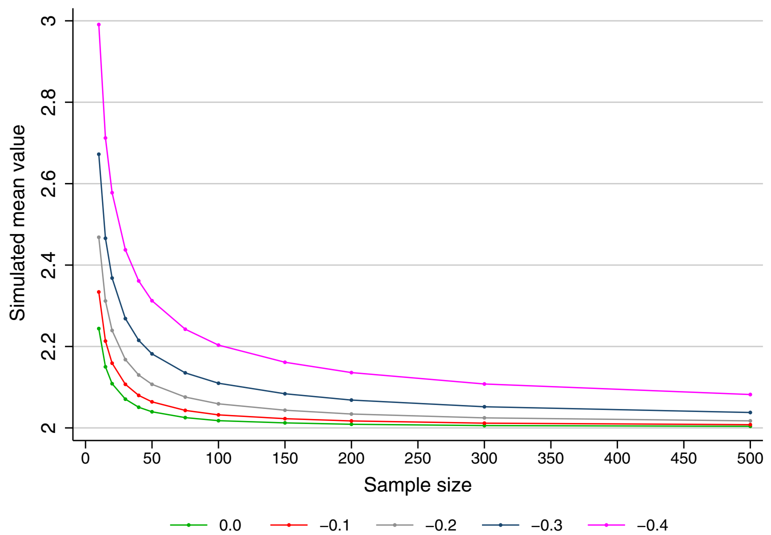


FIGURE 4 Simulated mean value of the LGPWM estimator with $s = -0.4, -0.3, -0.2, -0.1$, and 0 from a $P(2, 1)$ model

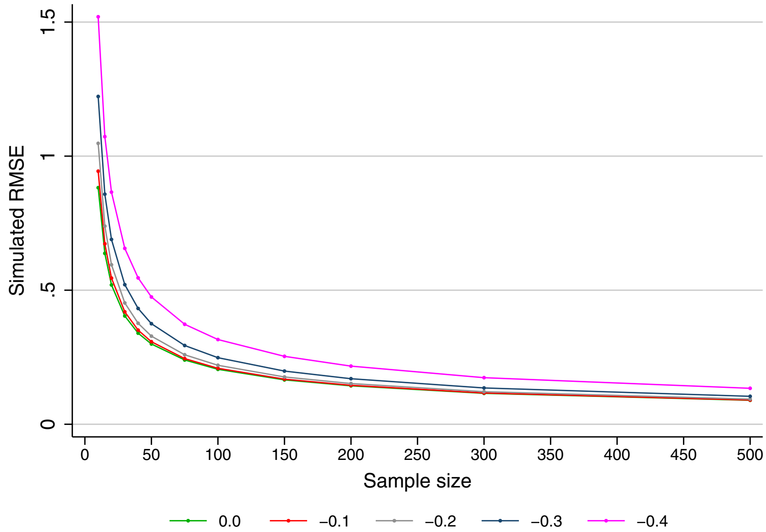


FIGURE 5 Simulated RMSE of the LGPWM estimator for with $s = -0.4, -0.3, -0.2, -0.1$, and 0 from a $P(2, 1)$ model

TABLE 1 Simulated mean value/RMSE of the shape estimators, for the Pareto model with $(a, c) = (0.5, 1)$

$\hat{a}$ (a = 0.5)				
n	PWM	M	ML	LS
10	1.1614/0.6903	1.0774/0.5899	0.6243/0.2660	**0.4893**/**0.2219**
15	1.1126/0.6266	1.0479/0.5527	0.5762/0.1829	**0.4788**/**0.1717**
20	1.0887/0.5974	1.0348/0.5373	0.5550/0.1450	**0.4761**/0.1467
30	1.0633/0.5677	1.0223/0.5233	0.5351/0.1088	**0.4753**/0.1190
40	1.0499/0.5527	1.0164/0.5170	0.5257/0.0900	0.4761/0.1035
50	1.0417/0.5436	1.0130/0.5133	0.5204/0.0785	0.4776/0.0928
75	1.0302/0.5312	1.0086/0.5087	0.5134/0.0621	0.4805/0.0769
100	1.0240/0.5247	1.0064/0.5065	0.5100/0.0527	0.4827/0.0673
150	1.0173/0.5177	1.0042/0.5043	0.5067/0.0421	0.4859/0.0554
200	1.0138/0.5140	1.0032/0.5032	0.5051/0.0363	0.4879/0.0486
300	1.0099/0.5100	1.0021/0.5021	0.5034/**0.0293**	0.4904/0.0400
500	1.0065/0.5066	1.0013/0.5013	0.5021/**0.0227**	0.4930/0.0313
n	LGPWM(0.1)	LGPWM(0.2)	LGPWM(0.3)	
10	0.6139/0.2582	0.6089/0.2553	0.6077/0.2559	
15	0.5691/0.1785	0.5660/0.1773	0.5656/0.1783	
20	0.5496/0.1420	0.5473/**0.1415**	0.5472/0.1427	
30	0.5314/**0.1073**	0.5300/0.1074	0.5300/0.1086	
40	0.5227/**0.0891**	**0.5217**/0.0894	0.5218/0.0904	
50	0.5180/**0.0779**	**0.5172**/0.0782	0.5173/0.0792	
75	0.5118/**0.0618**	**0.5113**/0.0622	0.5114/0.0630	
100	0.5087/**0.0525**	**0.5084**/0.0529	0.5085/0.0536	
150	0.5058/**0.0420**	**0.5055**/0.0424	0.5056/0.0430	
200	0.5044/**0.0363**	**0.5042**/0.0366	0.5043/0.0371	
300	0.5029/0.0294	**0.5028**/0.0297	0.5029/0.0301	
500	0.5018/0.0227	**0.5018**/0.0230	0.5018/0.0233	

TABLE 2 Simulated mean value/RMSE of the shape estimators, for the Pareto model with $(a, c) = (1, 1)$

$\hat{a}$ ($\boldsymbol{a = 1}$)				
n	PWM	M	ML	LS
10	1.5693/0.7456	1.4326/0.5710	1.2485/0.5320	**0.9786**/**0.4439**
15	1.4655/0.5848	1.3532/0.4442	1.1524/0.3658	**0.9576**/**0.3433**
20	1.4115/0.5059	1.3132/0.3833	1.1100/0.2900	**0.9522**/0.2935
30	1.3508/0.4204	1.2697/0.3209	1.0702/0.2177	**0.9505**/0.2380
40	1.3164/0.3743	1.2455/0.2875	1.0513/0.1801	0.9522/0.2070
50	1.2944/0.3446	1.2301/0.2664	1.0407/0.1570	0.9553/0.1857
75	1.2604/0.3004	1.2064/0.2352	1.0269/0.1241	0.9610/0.1538
100	1.2405/0.2751	1.1926/0.2173	1.0200/0.1053	0.9654/0.1347
150	1.2172/0.2454	1.1763/0.1965	1.0134/0.0841	0.9718/0.1108
200	1.2033/0.2281	1.1666/0.1844	1.0101/0.0725	0.9758/0.0972
300	1.1859/0.2068	1.1543/0.1694	1.0068/**0.0587**	0.9809/0.0799
500	1.1677/0.1847	1.1414/0.1537	1.0043/**0.0454**	0.9860/0.0626
n	LGPWM(0.1)	LGPWM(0.2)	LGPWM(0.3)	
10	1.2278/0.5163	1.2178/0.5105	1.2154/0.5118	
15	1.1383/0.3569	1.1321/0.3546	1.1311/0.3567	
20	1.0992/0.2841	1.0947/**0.2831**	1.0943/0.2853	
30	1.0627/**0.2146**	1.0599/0.2149	1.0600/0.2172	
40	1.0455/**0.1782**	**1.0435**/0.1788	1.0437/0.1809	
50	1.0359/**0.1557**	**1.0343**/0.1565	1.0346/0.1584	
75	1.0236/**0.1236**	**1.0226**/0.1244	1.0229/0.1261	
100	1.0175/**0.1050**	**1.0168**/0.1058	1.0171/0.1073	
150	1.0115/**0.0841**	**1.0111**/0.0848	1.0113/0.0860	
200	1.0087/**0.0725**	**1.0084**/0.0732	1.0086/0.0741	
300	1.0058/0.0588	**1.0056**/0.0593	1.0057/0.0601	
500	1.0037/0.0455	**1.0035**/0.0459	1.0036/0.0465	

- Regarding the different estimation methods, we conclude that there is no overall winner. But the results presented here allow us to conclude that the LGPWM estimators provide a consistently better fit than those from other considered estimation methods.
- A data-driven heuristic for the choice of s would help practitioners to use the new LGPWM estimator. One possibility is to consider a goodness of fit statistic such as the one suggested in Reference 19. However, this topic is outside the scope of the current article and should be addressed in future research.

4 | CONFIDENCE INTERVALS

Asymptotic confidence intervals (CI) for the shape parameter a can be obtained from the asymptotic distribution theory of Section 2.6. In the construction of the CI based on the M and PWM estimators we replaced the unknown asymptotic variance by its estimate. One advantage of this procedure is that the expressions of the extremes of the CI can be obtained with less tedious algebra. Based on the result in Propositions 1 and 2, we have the following two sided $(1 - \alpha) \times 100\%$ asymptotic confidence interval for the a parameter:

TABLE 3 Simulated mean value/RMSE of the shape estimators, for the Pareto model with $(a, c) = (2, 1)$

$\hat{a}$ ($\boldsymbol{a = 2}$)				
n	PWM	M	ML	LS
10	2.6152/1.1825	2.4360/0.9408	2.4971/1.0641	**1.9572**/**0.8877**
15	2.4397/0.8751	2.2991/0.6874	2.3048/0.7316	**1.9151**/**0.6867**
20	2.3513/0.7266	2.2331/**0.5659**	2.2200/0.5799	**1.9045**/0.5870
30	2.2563/0.5697	2.1653/0.4439	2.1405/0.4354	**1.9011**/0.4761
40	2.2052/0.4868	2.1299/0.3787	2.1026/0.3602	1.9044/0.4140
50	2.1745/0.4336	2.1091/0.3372	2.0815/0.3140	1.9106/0.3713
75	2.1288/0.3569	2.0789/0.2776	2.0538/0.2482	1.9220/0.3076
100	2.1043/0.3133	2.0630/0.2429	2.0401/0.2107	1.9308/0.2693
150	2.0780/0.2619	2.0463/0.2025	2.0267/0.1683	1.9436/0.2215
200	2.0634/0.2330	2.0373/0.1799	2.0203/0.1451	1.9516/0.1944
300	2.0472/0.1971	2.0273/0.1516	2.0136/**0.1174**	1.9617/0.1599
500	2.0322/0.1609	2.0185/0.1233	2.0086/**0.0908**	1.9719/0.1252
n	LGPWM(0.1)	LGPWM(0.2)	LGPWM(0.3)	
10	2.4555/1.0326	2.4356/1.0210	2.4308/1.0235	
15	2.2766/0.7139	2.2641/0.7092	2.2623/0.7134	
20	2.1983/0.5681	2.1894/0.5662	2.1886/0.5707	
30	2.1254/**0.4293**	2.1198/0.4297	2.1200/0.4343	
40	2.0910/**0.3564**	**2.0870**/0.3575	2.0874/0.3617	
50	2.0719/**0.3115**	**2.0687**/0.3129	2.0692/0.3169	
75	2.0471/**0.2471**	**2.0452**/0.2488	2.0458/0.2521	
100	2.0349/**0.2101**	**2.0336**/0.2117	2.0341/0.2145	
150	2.0231/**0.1682**	**2.0222**/0.1697	2.0226/0.1720	
200	2.0174/**0.1451**	**2.0168**/0.1463	2.0172/0.1483	
300	2.0116/0.1176	**2.0112**/0.1187	2.0114/0.1203	
500	2.0073/0.0910	**2.0071**/0.0918	2.0073/0.0930	

- Asymptotic CI based on the ML estimator

$$\left(\frac{\hat{a}^{ML}\sqrt{n}}{\sqrt{n} + z_{1-\alpha/2}}, \frac{\hat{a}^{ML}\sqrt{n}}{\sqrt{n} - z_{1-\alpha/2}} \right). \tag{33}$$

- Asymptotic CI based on the LS estimator

$$\left(\frac{\hat{a}^{LS}}{\sqrt{n} + z_{1-\alpha/2}\sqrt{2}/\sqrt{n}}, \frac{\hat{a}^{LS}\sqrt{n}}{\sqrt{n} - z_{1-\alpha/2}\sqrt{2}} \right). \tag{34}$$

- Asymptotic CI based on the M estimator

$$\left(\hat{a}^M - z_{1-\alpha/2}\frac{\sqrt{\hat{a}^M}(\hat{a}^M - 1)}{\sqrt{n}\sqrt{\hat{a}^M - 2}}, \ \hat{a}^M + z_{1-\alpha/2}\frac{\sqrt{\hat{a}^M}(\hat{a}^M - 1)}{\sqrt{n}\sqrt{\hat{a}^M - 2}} \right), \quad \hat{a}^M > 2. \tag{35}$$

TABLE 4 Simulated coverage probabilities for the asymptotic confidence intervals

	ML	LS	M	PWM	LGPWM(0.1)	LGPWM(0.2)	LGPWM(0.3)
10	0.829	0.957	0.945	0.940	0.841	0.848	0.851
15	0.869	0.968	0.959	0.953	0.879	0.884	0.885
20	0.888	0.972	0.968	0.960	0.897	0.901	0.902
30	0.908	0.974	0.976	0.970	0.913	0.916	0.917
40	0.920	0.974	0.979	0.975	0.925	0.926	0.927
50	0.925	0.973	0.980	0.978	0.929	0.931	0.931
75	0.933	0.969	0.982	0.983	0.935	0.936	0.936
100	0.938	0.967	0.977	0.984	0.940	0.941	0.941
150	0.942	0.964	0.968	0.984	0.943	0.944	0.944
200	0.944	0.960	0.964	0.982	0.944	0.944	0.945
300	0.946	0.957	0.962	0.971	0.947	0.946	0.946
500	0.946	0.953	0.958	0.965	0.947	0.947	0.947

- Asymptotic CI based on the PWM estimator

$$\left(\hat{a}^{PWM} - z_{1-\alpha/2}\delta_n,\ \hat{a}^{PWM} + z_{1-\alpha/2}\delta_n\right), \quad \delta_n = \frac{\sqrt{\hat{a}^{PWM}(\hat{a}^{PWM}-1)}(2\hat{a}^{PWM}-1)}{\sqrt{n}\sqrt{(\hat{a}^{PWM}-2)(3\hat{a}^{PWM}-2)}}, \quad \hat{a}^{PWM} > 2. \tag{36}$$

- Asymptotic CI based on the LGPWM estimator

$$\left(\frac{\hat{a}^{LGPWM}\sqrt{n}}{\sqrt{n} + z_{1-\alpha/2}(1+s)/\sqrt{1+2s}},\ \frac{\hat{a}^{LGPWM}\sqrt{n}}{\sqrt{n} - z_{1-\alpha/2}(1+s)/\sqrt{1+2s}}\right), \tag{37}$$

where z_α denotes the α quantile from the standard normal distribution.

Next we compare the performance of the previous asymptotic CIs via their coverage probabilities. Since CIs based on the M and PWM estimators are only valid if $a > 2$, we conducted a Monte Carlo simulation study with 50,000 samples from the P(3,1) model. We considered the same set of sample sizes as in the previous section and the values $s = 0.1$, 0.2, and 0.3 for the CI in (37). The simulated coverage probabilities are presented in Table 4. We conclude that the CIs based on the LS, M, and PWM estimators are conservative. The largest coverage probability is provided by the interval in (36), except when $n < 50$. The CIs based on the remaining estimation methods fail to reach a desired 95% coverage probability for sample sizes smaller than 300 although their coverage probability converges to the 95% nominal level as the sample size increases. Since the associated pointwise estimators do not provide a substantial bias, the behavior for $n < 300$ can be explained by the poor normal approximation based on the results in (27) and (32). Therefore, to have a coverage probability close to the nominal level, we advise the use of the CIs based on the LS method if $n < 300$ and on the ML or LGPWM methods if $n \geq 300$.

5 | REAL DATA ANALYSIS

To illustrate the use of the different estimation methods, presented in Section 2, we now analyze the fit of a Pareto model to a real data set. The data used for the empirical analysis come from a clinical trial on endogenous creatinine clearance of 34 male patients. The data set also includes three covariates: body weight in kilograms, serum creatinine concentration in milligrams/deciliter, and age in years and is available in Reference 47 and in R package heavy.[48] We shall consider only the sample of body weights. The data are the following: 71, 69, 85, 100, 59, 73, 63, 81, 74, 87, 79, 93, 60, 70, 83, 70, 73, 85,

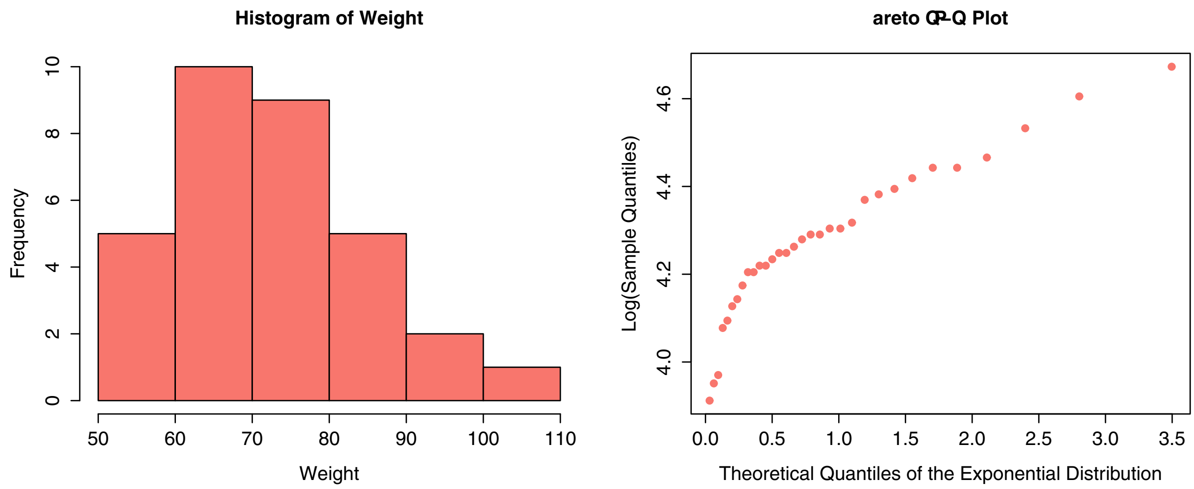


FIGURE 6 Histogram and Q-Q plot

TABLE 5 Estimates of the shape and scale parameters and value of the Cramér–von Mises criterion

	PWM	M	LS	ML	LGPWM(0.1)	LGPWM(0.2)	LGPWM(0.3)
$\hat{a}$	7.289	6.855	6.526	6.972	6.540	6.436	6.359
$\hat{c}$	67.270	66.593	66.998	67.000	67.000	67.000	67.000
W_n^2	0.057	0.404	0.024	0.066	0.021	0.056	0.094

68, 65, 53, 50, 74, 67, 80, 67, 68, 72.2, NA, NA, 107, 75, 62, 52 (NA means not available). Notice that two male patients had no weight recorded and were excluded from the sample.

In order to analyze the fit of the Pareto model to the data we used the Cramér–von Mises criterion

$$W_n^2 = \sum_{i=1}^{n}\left(F(x_{i:n}|\hat{a},\hat{c}) - \frac{2i-1}{2n}\right)^2 + \frac{1}{2n},$$

where $F(\cdot|\hat{a},\hat{c})$ is the Pareto distribution function, in (1), with scale and shape parameters estimated by $\hat{a}$, and $\hat{c}$, respectively. Smaller values of W_n^2 corresponds to a better fit of the Pareto model.

The histogram and the Pareto quantile-quantile (Q-Q) plot of these observations, in Figure 6, indicate that the underlying distribution has a Pareto tail. For Pareto-type models, the Pareto Q-Q Plot should be linear above a certain threshold, which in this case appears to be close to the value 65 kg. Thus, to model the tail, we shall only consider the 24 largest values, above the threshold 65. Recent overviews of data-driven heuristic choices of the tail threshold can be found in References 49,50. The parameter estimates of the Pareto model and the empirical value of the Cramér–von Mises criterion are shown in Table 5. It is evident that the Pareto model with parameters estimated by LGPWM method with $s = 0.1$ provides the best fit. Overall, the LS ML, and LGPWM methods, with $s = 0.1$ and 0.2 provides a good fit. Also, notice that the scale PWM estimate is invalid, since it is greater than the sample minimum.

6 | CONCLUSION

In this article, we propose a new class of estimators for the shape parameter of a Pareto model, called the log generalized probability weighted moments estimator and present some of its asymptotic properties. A finite sample comparison with other estimation methods showed that the new class of estimators is very competitive for pointwise estimation. The confidence interval provided by the LGPWM method evidence a coverage probability very close to the nominal level when $n > 300$. The usefulness of the new estimation method was illustrated with a real data application. Regarding the different estimation methods, we find the LGPWM estimator has an overall good performance and is also able to outperform the remaining estimation methods.

ACKNOWLEDGMENTS

The authors thank the anonymous reviewers for their careful reading and constructive comments and suggestions. This work is funded by national funds through the FCT—Fundação para a Ciência e a Tecnologia, I.P., under the scope of the project UIDB/00297/2020 (Center for Mathematics and Applications)

CONFLICT OF INTEREST

The authors declare no potential conflict of interests.

ORCID

Ayana Mateus https://orcid.org/0000-0002-5630-3321
Frederico Caeiro https://orcid.org/0000-0001-8628-7281

REFERENCES

1. Pareto V. Cours d'Economie politique. Lausanne, vol. 2; 1897.
2. Zaninetti L, Ferraro M. On the truncated Pareto distribution with applications. *Open Phys*. 2008;6(1):1-6.
3. Glänzel W. On some new bibliometric applications of statistics related to the h-index. *Scientometrics*. 2008;77(1):187-196.
4. Purdom E, Holmes SP. Error distribution for gene expression data. *Stat Appl Genet Molecul Biol*. 2005;4(1):16.
5. Barranco-Chamorro I, Jiménez-Gamero M. Asymptotic results in partially non-regular log-exponential distributions. *J Stat Comput Simulat*. 2012;82(3):445-461.
6. Wang S. Insurance pricing and increased limits ratemaking by proportional hazards transforms. *Insur Math Econom*. 1995;17(1):43-54.
7. Packová V, Brebera D. Loss distributions in insurance risk management. *Recent Advances on Economics and Business Administration*; 2015:17-22.
8. Johnson NL, Kotz S, Balakrishnan N. *Continuous Univariate Distributions*. New York, NY: Springer; 1994.
9. Kleiber C, Kotz S. *Statistical Size Distributions in Economics and Actuarial Sciences. 470 of Wiley Series in Probability and Statistics*. Hoboken, NJ: John Wiley & Sons; 2003.
10. Arnold BC. Pareto and generalized pareto distributions. *Modeling Income Distributions and Lorenz Curves*. Vol 5. New York, NY: Springer; 2008:119-145.
11. Arnold BC. *Pareto distributions. Monographs on Statistics and Applied Probability*. Boca Raton, FL: Chapman and Hall/CRC Press; 2015.
12. Newman ME. Power laws, Pareto distributions and Zipf's law. *Contemp Phys*. 2005;46(5):323-351.
13. Gabaix X. Power laws in economics: an introduction. *J Econom Perspect*. 2016;30(1):185-206.
14. Bingham NH, Goldie CM, Teugels JL. *Regular Variation*. Cambridge, MA: Cambridge University Press; 1989.
15. Beirlant J, Teugels JL, Vynckier P. *Practical Analysis of Extreme Values*. Leuven: Leuven University Press; 1996.
16. Beirlant J, Goegebeur Y, Segers J, Teugels JL. *Statistics of Extremes: Theory and Applications*. Hoboken, NJ: John Wiley & Sons; 2006.
17. Beirlant J, Caeiro F, Gomes MI. An overview and open research topics in statistics of univariate extremes. *Revstat Stat J*. 2012;10(1):1-31.
18. Gomes MI, Guillou A. Extreme value theory and statistics of univariate extremes: a review. *Int Stat Rev*. 2015;83(2):263-292.
19. Quandt RE. Old and new methods of estimation and the Pareto distribution. *Metrika*. 1966;10(1):55-82.
20. Caeiro F, Martins AP, Sequeira IJ. Finite sample behaviour of classical and quantile regression estimators for the Pareto distribution. Paper presented at: Proceedings of the AIP Conference; vol. 1648; 2015:540007; AIP Publishing.
21. Kim JH, Ahn S, Ahn S. Parameter estimation of the Pareto distribution using a pivotal quantity. *J Korean Stat Soc*. 2017;46(3):438-450.
22. Brazauskas V, Serfling R. Robust estimation of tail parameters for two-parameter Pareto and exponential models via generalized quantile statistics. *Extremes*. 2000;3(3):231-249.
23. Vandewalle B, Beirlant J, Christmann A, Hubert M. A robust estimator for the tail index of Pareto-type distributions. *Comput Stat Data Anal*. 2007;51(12):6252-6268.
24. Arnold BC, Press SJ. Bayesian estimation and prediction for Pareto data. *J Am Stat Assoc*. 1989;84(408):1079-1084.
25. Rasheed HA, Al-Gazi NAA. Bayes estimators for the shape parameter of Pareto type I distribution under generalized square error loss function. *Math Theory Model*. 2014;4(11):20-32.
26. Han M. The E-Bayesian estimation and its E-MSE of Pareto distribution parameter under different loss functions. *J Stat Comput Simulat*. 2020;90(10):1834-1848.
27. Caeiro F, Gomes MI. Semi-parametric tail inference through probability-weighted moments. *J Stat Plann Infer*. 2011;141(2):937-950.
28. Munir R, Saleem M, Aslam M, Ali S. Comparison of different methods of parameters estimation for Pareto Model. *Casp J Appl Sci Res*. 2013;2(1):45-56.
29. Dalpatadu RJ, Singh AK. Estimation of the parameter of the pareto distribution using a minimization technique. *Int J Appl Sci Technol*. 2015;5(4):8-10.
30. Bhatti SH, Hussain S, Ahmad T, Aslam M, Aftab M, Raza MA. Efficient estimation of Pareto model: some modified percentile estimators. *PloS One*. 2018;13(5):e0196456.
31. Bhatti SH, Hussain S, Ahmad T, Aftab M, Ali Raza M, Tahir M. Efficient estimation of Pareto model using modified maximum likelihood estimators. *Scientia Iranica*. 2019;26(1):605-614.
32. Lu HL, Tao SH. The estimation of Pareto distribution by a weighted least square method. *Qual Quant*. 2007;41(6):913-926.

33. Greenwood JA, Landwehr JM, Matalas NC, Wallis JR. Probability weighted moments: definition and relation to parameters of several distributions expressable in inverse form. *Water Resour Res*. 1979;15(5):1049-1054.
34. Hamed K, Rao AR. *Flood Frequency Analysis*. Boca Raton, FL: CRC Press; 2019.
35. Hosking JRM, Wallis JR, Wood EF. Estimation of the generalized extreme-value distribution by the method of probability-weighted moments. *Technometrics*. 1985;27(3):251-261.
36. Landwehr JM, Matalas N, Wallis J. Probability weighted moments compared with some traditional techniques in estimating Gumbel parameters and quantiles. *Water Resour Res*. 1979;15(5):1055-1064.
37. Landwehr JM, Matalas N, Wallis J. Estimation of parameters and quantiles of Wakeby distributions. known lower bounds. *Water Resour Res*. 1979;15(6):1361-1372.
38. Caeiro F, Gomes MI, Vandewalle B. Semi-parametric probability-weighted moments estimation revisited. *Methodol Comput Appl Probab*. 2014;16(1):1-29.
39. Rasmussen PF. Generalized probability weighted moments: application to the generalized Pareto distribution. *Water Resour Res*. 2001;37(6):1745-1751.
40. Caeiro F, Prata GD. A log probability weighted moment estimator of extreme quantiles. In: Kitsos C, Oliveira T, Rigas A, Gulati S, eds. *Theory and Practice of Risk Assessment. 136 of Springer Proceedings in Mathematics & Statistics*. Cham, Switzerland: Springer; 2015:293-303.
41. Caeiro F, Mateus A. Log probability weighted moments method for Pareto distribution. In: Skiadas CH, ed. *Proceedings of the 17th ASMDA Conference*; Springer Proceedings in Mathematics & Statistics: Springer; 2017:211-218.
42. Chen H, Cheng W, Zhao J, Zhao X. Parameter estimation for generalized Pareto distribution by generalized probability weighted moment-equations. *Commun Stat Simulat Comput*. 2017;46(10):7761-7776.
43. Hosking JR, Wallis JR. Parameter and quantile estimation for the generalized Pareto distribution. *Technometrics*. 1987;29(3):339-349.
44. Balakrishnan K, Basu AP. *Exponential Distribution: Theory, Methods and Applications*. Boca Raton, FL: CRC Press; 1996.
45. Arnold BC, Balakrishnan N, Nagaraja HN. *A First Course in Order Statistics*. Vol 54. Philadelphia, PA: SIAM; 1992.
46. R Development Core Team R, *R: A Language and Environment for Statistical Computing*. Austria: R foundation for statistical computing Vienna; 2018. http://www.r-project.org/index.html.
47. Liu C, Rubin DB. ML estimation of the t distribution using EM and its extensions, ECM and ECME. *Statistica Sinica*. 1995;19-39.
48. Osorio F. Heavy: robust estimation using heavy-tailed distributions. R package version 0.38.196; 2019.
49. Scarrott C, MacDonald A. A review of extreme value threshold estimation and uncertainty quantification. *Revstat Stat J*. 2012;10(1):33-60.
50. Caeiro F, Gomes M. Threshold selection in extreme value analysis. *Extreme Value Modeling and Risk Analysis*. Hoboken, NJ: Chapman and Hall/CRC Press; 2015:69-86.

AUTHOR BIOGRAPHIES

Ayana Mateus is an Auxiliary Professor at the Mathematics Department of the NOVA School of Science and Technology (FCT NOVA) and is a member of the Center for Mathematics and Applications. She has a master's degree in Operation Research by Instituto Superior Técnico, University of Lisbon (1997) and a PhD degree in Statistics from FCT NOVA (2006). Her current research interests include Statistics of Extremes, Nonparametric Statistics and Computational Statistical Methods.

Frederico Caeiro is an Auxiliary Professor at the Mathematics Department of the NOVA School of Science and Technology (FCT NOVA) and is a member of the Mathematics and Application Research Center (Portugal). He has an MSc degree in Probability and Statistics (2001) and a PhD degree in Statistics (2006) from the University of Lisbon. His current research interests include Statistics of Extremes, Extreme Value Theory, Nonparametric Statistics and Computational Statistical Methods.

How to cite this article: Mateus A, Caeiro F. A new class of estimators for the shape parameter of a Pareto model. *Comp and Math Methods*. 2021;3:e1133. https://doi.org/10.1002/cmm4.1133