Journal of Mathematics and Modeling in Finance (JMMF)
Vol. 6, No. 2, Summer & Fall 2026
Research paper

Closed-Form Estimation for the Pareto Distribution Based on Logarithmic Moments

Zahra Karimiezmareh[1], **Ghazal Elahi**[2]

[1] Department of Statistics, Faculty of Statistics, Mathematics and Computer, Allameh Tabataba'i University, Tehran, Iran

zahra.karimiez@atu.ac.ir

[2] Department of Statistics, Faculty of Statistics, Mathematics and Computer, Allameh Tabataba'i University, Tehran, Iran

ghazal_elahi@atu.ac.ir

Abstract:
The Pareto distribution is a cornerstone for modeling heavy-tailed phenomena in fields like economics, finance, and risk management. While maximum likelihood (ML) estimation is prevalent, its estimators lack closed-form expressions, requiring iterative numerical methods. This paper introduces closed-form estimators for the Pareto distribution using the method of logarithmic moments. The proposed estimators are computationally simple and eliminate convergence issues associated with ML. We derive their large-sample properties, establishing consistency. For comparative purposes, L-moments estimators for the Pareto distribution are also considered. A comprehensive simulation study demonstrates that the proposed logarithmic moments estimators perform competitively with ML for heavy-tailed distributions. Moreover, an empirical application to real-world fire insurance claims data confirms the practical advantages of the proposed method, where it achieves a better fit compared to ML. The combination of closed-form simplicity, competitive performance in heavy-tailed settings, and computational efficiency makes the proposed estimators a powerful and practical alternative for rapid data analysis, pedagogical purposes, and applications with limited computational resources.

Keywords: Pareto Distribution, Logarithmic Moments, Closed-Form Estimators, Asymptotic Properties.
Classification: 62Exx; 62P99.

1 Introduction

The Pareto distribution was first introduced by the Italian economist [14] in his work "Cours d'Économie Politique" to model wealth distribution in society. Since its inception, this distribution has expanded beyond its original economic applications to become a fundamental tool in modeling heavy-tailed phenomena across various disciplines. In actuarial science, it serves to model extreme insurance claims and operational risk losses. In hydrology, it helps predict extreme flood levels and

[1]Corresponding author
Received: 14/01/2026 Accepted: 07/06/2026
https://doi.org/10.22054/jmmf.2026.91047.1253

rainfall amounts. The distribution has also proven valuable in telecommunications for modeling network traffic, in geology for characterizing earthquake magnitudes, and in finance for estimating value at risk. Furthermore, recent applications include modeling social network phenomena, file size distributions on the internet, and protein family abundances in bioinformatics. The enduring relevance of the Pareto distribution across these varied domains stems from its mathematical tractability and its effectiveness in characterizing the ubiquitous 80-20 rule observed in many natural and man-made systems [8]. There is an authoritative reference providing a detailed treatment of its statistical properties and applications for comprehensive coverage of this distribution [2].

Definition 1.1. A continuous random variable X is said to follow a Pareto distribution with parameters $\alpha > 0$ and $\beta > 0$, denoted as $X \sim Par(\alpha, \beta)$, if its probability density function (PDF) and cumulative distribution function (CDF) are given by:

$$f(x; \alpha, \beta) = \frac{\alpha \beta^{\alpha}}{x^{\alpha+1}}, \quad x \geq \beta, \tag{1}$$

$$F(x; \alpha, \beta) = 1 - (\frac{\beta}{x})^{\alpha}, \quad x \geq \beta. \tag{2}$$

Recent years have witnessed a renewed and sustained interest in developing advanced estimation methods for the Pareto distribution, reflecting its undiminished relevance across various scientific domains. [9] tackled the complex challenge of Bayesian stress-strength reliability estimation with fully unknown parameters, necessitating sophisticated computational algorithms like acceptance-rejection sampling due to the intractable form of the posterior distribution. Concurrently, [15] conducted a comprehensive comparison of Bayesian and frequentist paradigms, demonstrating the competitive performance of the Reference Intrinsic Bayesian approach. The critical role of the estimation method itself was further underscored by [13], who revealed that the choice between ML and moment-based estimation profoundly impacts the performance of goodness-of-fit tests. In the realm of specialized applications, [6] advanced the estimation of differential entropy for Pareto data in the presence of outliers, while [11] developed a robust Bayesian framework using mixed Gibbs sampling that outperforms MLE for small censored samples. This collective research effort highlights two dominant themes: the persistent complexity of obtaining tractable estimators, especially in Bayesian settings, and the clear advantage of innovative methods over traditional MLE in specific, challenging scenarios. It is within this context that our proposed method of logarithmic moments emerges, aiming to bridge the gap by offering a simple, closed-form, and frequentist alternative that circumvents the computational burdens evident in these recent studies.

Although the method of moments is a standard technique for parameter estimation and distribution summarization, it possesses notable limitations. Higher-order moments often lack intuitive interpretability regarding distributional shape, and their

sample estimates—particularly for small samples—can be highly unstable and biased. Consequently, parameter estimates derived from this method frequently exhibit inferior accuracy compared to alternatives like ML. To address these shortcomings, [7] introduced L-moments as a robust alternative. Unlike conventional moments, L-moments are defined as linear combinations of order statistics (L-statistics). This foundation grants them with several superior properties: they are more robust to outliers in the data, less prone to estimation bias, and capable of characterizing a wider range of probability distributions. Empirical evidence further suggests that estimators based on L-moments often achieve greater accuracy in small samples and converge more rapidly to their asymptotic normal distribution compared to both traditional method-of-moments and, in some cases, ML. Thus, L-moments provide a unified and powerful framework for distribution characterization, parameter estimation, and goodness-of-fit testing.

Building on the theoretical appeal of L-moments, recent research has demonstrated their practical utility in overcoming estimation challenges across various distributions. This has led to a growing adoption of moment-based methods in applied statistics, particularly when seeking a balance between computational efficiency and estimation accuracy. [10] addressed the critical issue of computational speed in both univariate and bivariate Weibull distributions. Recognizing that the ML is often unavailable in closed-form, the author developed a novel, closed-form estimator. This method is shown to be not only statistically efficient and asymptotically normal but also dramatically faster—approximately 11 times for the univariate case and 85 times for the bivariate case—making it highly suitable for real-time processing and state-space models. In a different application, [16] leveraged L-moments for modeling urban wind speed distributions. Their study compared the method of L-moments against the method of moments and ML for fitting both standard and mixture Weibull distributions. A key finding was that while ML was generally the most accurate, the L-moments method outperformed it at specific locations. Furthermore, the mixture model fitted via L-moments was particularly effective at capturing complex, bimodal patterns in the data, showcasing the method's robustness in environmental modeling. Most directly related to this work, [12] introduced a breakthrough for the Fréchet distribution by proposing the first closed-form estimators using the method of logarithmic moments. This approach, which shares the philosophical goal of deriving tractable estimators, yielded estimators applicable for all parameter values and established their large-sample properties, providing a compelling alternative to the traditional ML.

The Pareto distribution is the cornerstone model for analyzing income and wealth data due to its power-law tail and inherent scale-invariance, which directly captures economic inequality. Its self-similar nature, where moments of the distribution themselves follow Pareto laws, makes it theoretically coherent for this domain. However, accurate parameter estimation is crucial, as the shape parameter α is a direct measure of inequality. Existing estimation methods, including moments and

ML, often rely on complex order statistics or iterative procedures, limiting their accessibility for applied researchers. This paper addresses this gap by proposing a new closed-form estimator based on logarithmic moments. This novel approach provides computationally simple estimators, intuitively derived from the properties of the log-transformed data, and is highly accessible for practical use in economics and social sciences. We derive the asymptotic properties of the proposed logarithmic moments estimators and demonstrate their competitive performance against ML method via simulation.

The remainder of this paper is organized as follows. Section 2 reviews the classical method of moments and its limitations for the Pareto distribution. Section 3 briefly presents ML estimation. In Section 4, we provide an overview of the L-moments method, which is included for completeness. The main contribution of this work is presented in Section 5, where we introduce a new closed-form estimator for the Pareto distribution based on logarithmic moments. Section 6 presents simulation results comparing the proposed estimator with existing methods. Finally, Section 7 provides an empirical application to real-world insurance data, demonstrating the practical advantages of the proposed estimator.

2 Moment Estimation

Let $X_1, X_2, ..., X_n$ be a random sample from a $Par(\alpha, \beta)$. The r^{th} theoretical moment of the Pareto distribution exists only if $\alpha > r$ and is given by

$$E[X^r] = \frac{\alpha\beta^r}{\alpha - r}. \tag{3}$$

To estimate the parameters α and β, we typically use the first two moments:

$$E(X) = \frac{\alpha\beta}{\alpha - 1}, \quad \alpha > 1, \tag{4}$$

$$E(X^2) = \frac{\alpha\beta^2}{\alpha - 2}, \quad \alpha > 2. \tag{5}$$

Equating the population and sample moments yields the following system of equations:

$$\frac{\alpha\beta}{\alpha - 1} = \overline{X}, \tag{6}$$

$$\frac{\alpha\beta^2}{\alpha - 2} = \frac{1}{n}\sum_{i=1}^{n} X_i^2. \tag{7}$$

Solving this system for α and β, we obtain the moment estimators:

$$\tilde{\alpha} = 1 + \sqrt{1 + (\frac{\overline{X}}{S})^2}, \tag{8}$$

$$\tilde{\beta} = \frac{(\tilde{\alpha} - 1)\overline{X}}{\tilde{\alpha}}. \tag{9}$$

However, the practical utility of these estimators is severely limited by a critical constraint: they require $\alpha > 2$ for existence, a condition often violated by real-world heavy-tailed data, where the tail index α is frequently found to be between 1 and 2. Furthermore, even when $\alpha > 2$, these estimators are known to be inefficient and can be highly sensitive to outliers in the upper tail, leading to substantial instability. It is precisely these limitations that provide a critical benchmark for evaluation. The persistent shortcomings of the method of moments in estimating Pareto parameters underscore the necessity for more robust and universally applicable estimation techniques.

3 ML Estimation

For a random sample $X_1, X_2, ..., X_n$ from a $Par(\alpha, \beta)$ distribution, the log-likelihood function is given by:

$$l(\alpha, \beta) = n \log \alpha + n\alpha \log \beta - (\alpha + 1) \sum_{i=1}^{n} \log x_i. \tag{10}$$

Maximizing this log-likelihood function yields the following well-known ML estimators:

$$\hat{\alpha} = \frac{n}{\sum_{i=1}^{n} \log \frac{x_{i:n}}{x_{1:n}}}, \tag{11}$$

$$\hat{\beta} = x_{1:n}, \tag{12}$$

where $x_{1:n} \leq x_{2:n} \leq ... \leq x_{n:n}$ denote the order statistics in the sample.

Despite its optimal asymptotic properties, the ML estimator for the Pareto distribution presents several practical limitations that the proposed method effectively addresses. The ML estimate of the threshold parameter $\hat{\beta} = x_{(1)}$ is highly sensitive to the smallest observation, making it vulnerable to outliers in the lower tail and potentially yielding unstable estimates. Furthermore, while the shape parameter estimator is conditionally closed-form, the overall solution depends on an order statistic, complicating the theory of its finite-sample distribution. In contrast, the proposed estimators are derived simultaneously in a fully closed-form, utilize information from the entire dataset for both parameters, and offer computational simplicity, making them a more practical alternative for real-world applications.

3.1 Remarks on the Fisher Information Matrix

For the Pareto distribution, the support of the distribution depends on the threshold parameter β, since $x \geq \beta$. This violates the standard regularity conditions required for the Cramér–Rao lower bound (CRLB). In particular, the Fisher information matrix cannot be derived in the usual way because the order of differentiation and integration does not commute.

Therefore, the usual asymptotic variance expressions based on the inverse Fisher information matrix are not valid for the Pareto distribution when β is unknown. For this reason, we do not present an asymptotic variance comparison based on the CRLB. Instead, we rely on our simulation study in Section 6 to evaluate the finite-sample performance of the proposed estimators and compare them with ML estimators.

It is worth noting that if the scale parameter β is known, the standard regularity conditions hold for the shape parameter α alone. However, this case is not the focus of the present work, as in most practical applications, both parameters are unknown.

4 L-moments Estimation

Before presenting our main contribution in Section 5, it is helpful to briefly review the L-moments method for the Pareto distribution.

The L-moments of a real-valued random variable X with CDF $F(x)$ and quantile function $x(F)$ are defined as linear combinations of expected order statistics. Consider a sample of size n with order statistics $X_{1:n} \leq X_{2:n} \leq ... \leq X_{n:n}$. The r^{th} population L-moment, denoted μ_r, is formally defined as:

$$\mu_r = \frac{1}{r}\sum_{k=0}^{r-1}(-1)^k \binom{r-1}{k} E(X_{r-k:r}), \quad r = 1, 2, 3, ... \tag{13}$$

The "L" in L-moments highlights that each μ_r is a linear function of the expectations of these order statistics. The expected value of the j^{th} order statistic $X_{j:r}$ in a sample of size r can be expressed as:

$$E(X_{j:r}) = \frac{r!}{(j-1)!(r-j)!}\int x[F(x)]^{j-1}[1-F(x)]^{r-j}dF(x). \tag{14}$$

The corresponding r^{th} sample L-moment is

$$l_r = \binom{n}{r}^{-1}\sum_{1\leq i_1<\cdots<i_r\leq n}\cdots\sum\frac{1}{r}\sum_{k=0}^{r-1}(-1)^k\binom{r-1}{k}X_{i_{r-k}:n}, \quad r = 1, 2, \ldots, n. \tag{15}$$

L-moments estimators are derived by equating the theoretical L-moments of the probability distribution to their corresponding sample L-moments calculated from the observed data.

The first two sample L-moments are defined as follows:

$$l_1 = \frac{1}{n}\sum_{i=1}^{n} X_{i:n} = \overline{X}, \tag{16}$$

$$l_2 = \frac{2}{n(n-1)}\sum_{i=1}^{n}(i-1)X_{i:n} - \overline{X}. \tag{17}$$

These statistics serve as robust alternatives to conventional sample moments, with l_1 representing the sample mean and l_2 measuring the scale based on linear combinations of order statistics.

Theorem 4.1. *Let $X_1, X_2, ..., X_n$ be a random sample from a $Par(\alpha, \beta)$ distribution. Then, the method of L-moments estimators for the parameters α and β are given by the following closed-form expressions:*

$$\hat{\alpha}_{lm} = \frac{1}{2} + \frac{\overline{X}}{\frac{4}{n(n-1)} \sum_{i=1}^{n} (i-1) X_{i:n} - \overline{X}},$$

$$\hat{\beta}_{lm} = \frac{\overline{X}(\hat{\alpha}_{lm} - 1)}{\hat{\alpha}_{lm}}.$$

Proof. The proof is based on equating the first two population L-moments to their sample counterparts. The first two population L-moments are defined as:

$$\mu_1 = E(X_{1:1}), \tag{18}$$

$$\mu_2 = \frac{1}{2}[E(X_{2:2}) - E(X_{1:2})]. \tag{19}$$

Since the density of the first order statistic equals the population density, $f_{1:1}(x) = f(x)$, the first population L-moment μ_1 equals the first moment of the Pareto distribution:

$$\mu_1 = \frac{\alpha\beta}{\alpha - 1}.$$

For the second population L-moment, we use $\mu_2 = \frac{1}{2}[E(X_{2:2}) - E(X_{1:2})]$. The marginal density of the second order statistic in a sample of size 2 is $f_{2:2}(x) = 2F(x)f(x)$. Thus,

$$E(X_{2:2}) = \int_{\beta}^{\infty} x \ . \ 2 \ [1 - (\frac{\beta}{x})^{\alpha}] \ \frac{\alpha\beta^{\alpha}}{x^{\alpha+1}} dx = \frac{2\alpha^2\beta}{(\alpha - 1)(2\alpha - 1)}.$$

Also $E(X_{1:2})$ can be found using $f_{1:2}(x) = 2[1 - F(x)]f(x)$, so

$$E(X_{1:2}) = \int_{\beta}^{\infty} x \ . \ 2 \ (\frac{\beta}{x})^{\alpha} \ \frac{\alpha\beta^{\alpha}}{x^{\alpha+1}} dx = \frac{2\alpha\beta}{2\alpha - 1}.$$

Therefore,

$$\mu_2 = \frac{\alpha\beta}{(\alpha - 1)(2\alpha - 1)}.$$

Equating sample and population L-moments, $l_1 = \mu_1$ and $l_2 = \mu_2$, gives the system:

$$\overline{X} = \frac{\alpha\beta}{\alpha - 1},$$

$$\frac{2}{n(n-1)} \sum_{i=1}^{n} (i-1) X_{i:n} - \bar{X} = \frac{\alpha\beta}{(\alpha - 1)(2\alpha - 1)}.$$

Solving for parameters α and β, we express β from the first equation as $\beta = \frac{\overline{X}(\alpha-1)}{\alpha}$. Substituting into the second equation:

$$\hat{\alpha}_{lm} = \frac{1}{2} + \frac{\overline{X}}{\frac{4}{n(n-1)} \sum_{i=1}^{n} (i-1) X_{i:n} - \overline{X}}.$$

□

5 Proposed Estimators

The proposed approach is a direct application of the classical method of moments, but applied to the transformed variable $Z = \log X$ rather than to the original variable X. Recall from Section 2 that applying the method of moments directly to X leads to several well-known problems: the estimators require $\alpha > 2$ for existence, and their finite-sample performance is often poor. By working on the logarithmic scale, we obtain a shifted exponential distribution for Z, for which the moment equations become simple and yield closed-form solutions. Furthermore, the proposed logarithmic moment estimators exist for all $\alpha > 0$, without the restrictive condition $\alpha > 2$ required by classical moments.

The main advantages of working on the log-scale are:

- **Simplicity of the transformed distribution:** The variable $Z = \log X$ follows a shifted exponential distribution, which is mathematically tractable and well-understood.
- **Linear moment equations:** The moment equations for the shifted exponential distribution are linear in the parameters, yielding closed-form estimators without iterative numerical methods.
- **No restriction on α:** Classical moment estimators require $\alpha > 2$ for the second moment to exist. In contrast, the proposed logarithmic estimators exist for all $\alpha > 0$, covering the entire range of practical interest, including very heavy-tailed cases ($\alpha < 2$) where classical moments fail.

The core idea of our estimation approach is to transform the random variable to a logarithmic scale, where the parameter estimation becomes more tractable. This method, known as the method of logarithmic moments (LM), leverages the fact that the distribution of the log-transformed variable often yields simpler moment equations.

For a Pareto-distributed variable $X \sim Par(\alpha, \beta)$, we consider the logarithmic transformation $Z = \log X$. The transformed variable Z follows a shifted exponential distribution with probability density function:

$$f_Z(z) = \alpha e^{-\alpha(z - \log \beta)}, \quad z > \log \beta. \tag{20}$$

The first two theoretical moments of Z are:

$$\mathrm{E}(Z) = \log\beta + \frac{1}{\alpha}, \tag{21}$$

$$E(Z^2) = \frac{2}{\alpha^2} + \frac{2\log\beta}{\alpha} + \log^2\beta. \tag{22}$$

By equating these theoretical moments to their sample counterparts, we obtain:

$$\overline{Z} = \log\beta + \frac{1}{\alpha}, \tag{23}$$

$$\frac{1}{n}\sum_{i=1}^{n} Z_i^2 = \overline{Z}^2 + \frac{1}{\alpha^2}\cdot \tag{24}$$

Solving this system yields the closed-form estimators:

$$\hat{\alpha} = \frac{1}{\sqrt{\frac{1}{n}\sum_{i=1}^{n} Z_i^2 - \overline{Z}^2}} = \frac{1}{S_z}, \tag{25}$$

$$\hat{\beta} = e^{\overline{Z} - \frac{1}{\hat{\alpha}}} = e^{\overline{Z} - S_z}. \tag{26}$$

This approach provides computationally efficient estimators while maintaining statistical robustness.

5.1 Large Sample Properties

This section establishes the strong consistency and asymptotic normality of the proposed estimators, along with their asymptotic variances.

Theorem 5.1. *The estimators $\hat{\alpha} = \frac{1}{S_z}$ and $\hat{\beta} = e^{\overline{Z} - S_z}$ are strongly consistent for α and β, respectively.*

Proof. Let $X_1, X_2, ..., X_n$ be a random sample from the Pareto distribution with parameters α and β. Define $Z_i = \log X_i$, and let $\overline{Z}$ and S_z^2 denote the sample mean and variance of the Z_i's. From the distributional properties of Z, we have:

$$E(\log X) = E(Z) = \log\beta + \frac{1}{\alpha},$$

$$E(\log^2 X) = E(Z^2) = \frac{2}{\alpha^2} + \frac{2\log\beta}{\alpha} + \log^2\beta.$$

Let $\overline{U} = \frac{1}{n}\sum_{i=1}^{n} \log X_i$ denote the sample mean of $\log X_i$, and let $\overline{V} = \frac{1}{n}\sum_{i=1}^{n} \log^2 X_i$ denote the mean of $\log^2 X_i$. By the strong law of large numbers,

$$\begin{pmatrix} \overline{U} \\ \overline{V} \end{pmatrix} \to \begin{pmatrix} \log\beta + \frac{1}{\alpha} \\ \frac{2}{\alpha^2} + \frac{2\log\beta}{\alpha} + \log^2\beta \end{pmatrix},$$

almost surely as $n \to \infty$. Define the functions

$$h(u,v) = \frac{1}{\sqrt{v-u^2}},$$
$$g(u,v) = \exp\left(u - \frac{1}{h(u,v)}\right).$$

These functions are continuous at the point

$$\left(\log\beta + \frac{1}{\alpha}, \frac{2}{\alpha^2} + \frac{2\log\beta}{\alpha} + \log^2\beta\right).$$

By the continuous mapping theorem, the estimators $\hat{\alpha}$ and $\hat{\beta}$ converge almost surely to α and β, respectively, as $n \to \infty$. □

Theorem 5.2. *The estimators $\hat{\alpha}$ and $\hat{\beta}$ given in (25) and (26) are asymptotically normally distributed as*

$$\sqrt{n}(\hat{\alpha}-\alpha, \hat{\beta}-\beta) \xrightarrow{d} N(\mathbf{0}, \mathbf{D}), \quad as \quad n \to \infty,$$

where

$$D = \begin{pmatrix} 2\alpha^2 & \beta \\ \beta & \dfrac{\beta^2}{\alpha^2} \end{pmatrix}.$$

Proof. We first calculate the raw moments of Z.

$$E(Z) = \int_{\log\beta}^{\infty} z.\alpha e^{-\alpha(z-\log\beta)}dz = \log\beta + \frac{1}{\alpha},$$
$$E(Z^2) = \int_{\log\beta}^{\infty} z^2.\alpha e^{-\alpha(z-\log\beta)}dz = \frac{2}{\alpha^2} + \frac{2\log\beta}{\alpha} + \log^2\beta,$$
$$E(Z^3) = \int_{\log\beta}^{\infty} z^3.\alpha e^{-\alpha(z-\log\beta)}dz = \log^3\beta + \frac{3\log^2\beta}{\alpha} + \frac{6\log\beta}{\alpha^2} + \frac{6}{\alpha^3},$$
$$E(Z^4) = \int_{\log\beta}^{\infty} z^4.\alpha e^{-\alpha(z-\log\beta)}dz = \log^4\beta + \frac{4\log^3\beta}{\alpha} + \frac{12\log^2\beta}{\alpha^2} + \frac{24\log\beta}{\alpha^3} + \frac{24}{\alpha^4}.$$

Therefore,

$$\mathrm{Var}(\sqrt{n}\,\overline{U}) = \mathrm{Var}(Z) = E(Z^2) - E^2(Z) = \frac{1}{\alpha^2},$$
$$\mathrm{Var}(\sqrt{n}\,\overline{V}) = E(Z^4) - E^2(Z^2) = \frac{16\log\beta}{\alpha^3} + \frac{20}{\alpha^4} + \frac{4log^2\beta}{\alpha^2},$$
$$\mathrm{Cov}(\sqrt{n}\,\overline{U}, \sqrt{n}\,\overline{V}) = E(Z^3) - E(Z)E(Z^2) = \frac{2\log\beta}{\alpha^2} + \frac{4}{\alpha^3}.$$

By the Central Limit Theorem, as $n \to \infty$,

$$\sqrt{n}\left((\overline{U},\overline{V}) - (u_0,v_0)\right) \xrightarrow{d} N(\mathbf{0}, \mathbf{\Sigma}),$$

where $(u_0, v_0) = \left(\log\beta + \frac{1}{\alpha} \quad , \quad \frac{2}{\alpha^2} + \frac{2\log\beta}{\alpha} + \log^2\beta\right)$ and

$$\mathbf{\Sigma} = \begin{pmatrix} \mathrm{Var}(U) & \mathrm{Cov}(U,V) \\ \mathrm{Cov}(U,V) & \mathrm{Var}(V) \end{pmatrix} = \begin{pmatrix} \frac{1}{\alpha^2} & \frac{2\log\beta}{\alpha^2} + \frac{4}{\alpha^3} \\ \frac{2\log\beta}{\alpha^2} + \frac{4}{\alpha^3} & \frac{16\log\beta}{\alpha^3} + \frac{20}{\alpha^4} + \frac{4log^2\beta}{\alpha^2} \end{pmatrix}.$$

Using the delta method, as $n \to \infty$,

$$\sqrt{n}(\hat{\alpha} - \alpha, \hat{\beta} - \beta) \xrightarrow{d} N(\mathbf{0}, \mathbf{W\Sigma W}^T),$$

where

$$\mathbf{W} = \begin{pmatrix} \frac{\partial h}{\partial u} & \frac{\partial h}{\partial v} \\ \frac{\partial g}{\partial u} & \frac{\partial g}{\partial v} \end{pmatrix}.$$

The partial derivatives in the matrix $\mathbf{W}$ are evaluated at the point (u_0, v_0). Evaluating the partial derivatives at this point yields the matrix

$$\mathbf{W} = \begin{pmatrix} \alpha^3\log\beta + \alpha^2 & -\frac{1}{2}\alpha^3 \\ 2\beta + \alpha\beta\log\beta & -\frac{1}{2}\alpha\beta \end{pmatrix}. \tag{27}$$

Thus,

$$\mathbf{W\Sigma W}^T = \begin{pmatrix} 2\alpha^2 & \beta \\ \beta & \frac{\beta^2}{\alpha^2} \end{pmatrix}. \tag{28}$$

□

6 Simulation

6.1 Simulation Design

To comprehensively evaluate the performance of the proposed LM estimator against the classical ML estimator for the Pareto distribution, we conducted an extensive Monte Carlo simulation study. The simulation was designed to examine various realistic scenarios encountered in practical applications of Pareto modeling.
The simulation study considered the following factors:

- **Shape parameter (α)**: We examined four distinct values: $\alpha = 0.05, 0.10, 0.80, 1.20$. This range covers extremely heavy-tailed distributions ($\alpha = 0.05, 0.10$) to moderately heavy-tailed distributions ($\alpha = 0.80, 1.20$), representing common scenarios in financial, environmental, and risk management applications.

- **Scale parameter (β)**: Three different scale parameters were considered: $\beta = 1, 2, 5$. This variation tests the performance of the estimators across different location shifts of the distribution.
- **Sample sizes**: Four sample sizes were examined: $n = 20, 50, 100, 200$. This range represents scenarios from small datasets to moderately large samples.
- **Performance metrics**: For each scenario, we computed multiple statistical measures to assess estimator performance:
  - Bias: $\text{Bias}(\hat{\theta}) = E(\hat{\theta}) - \theta$
  - Root Mean Square Error (RMSE): $\sqrt{E[(\hat{\theta} - \theta)^2]}$
  - Mean Absolute Error (MAE): $E[|\hat{\theta} - \theta|]$
  - Variance: $\text{Var}(\hat{\theta})$
- **Replications**: Each scenario was replicated 5,000 times to ensure stable estimates of the performance metrics.

6.2 Computational Implementation

The simulation was implemented in R, utilizing the following workflow:

(i) For each combination of (α, β, n), generate random samples from the Pareto distribution using the inverse transform method: $X = \beta(1 - U)^{-1/\alpha}$, where $U \sim \text{Uniform}(0, 1)$.

(ii) For each sample, compute both ML and LM estimates:

- ML: $\hat{\alpha}_{\text{ML}} = n/\sum_{i=1}^{n} \log(X_{i:n}/X_{1:n})$, $\hat{\beta}_{\text{ML}} = X_{1:n}$
- LM: $\hat{\alpha}_{\text{LM}} = 1/S_z$, $\hat{\beta}_{\text{LM}} = e^{\bar{Z} - S_z}$, where $Z_i = \log X_i$

(iii) Compute performance metrics across all replications for each estimator.

(iv) Perform comparative analysis between the two estimators.

6.3 Main Findings

The comprehensive simulation results comparing the LM estimator with the ML estimator for the Pareto distribution are summarized in Tables 1-12. Tables 1-3 correspond to $\alpha = 0.05$ with $\beta = 1, 2, 5$; Tables 4-6 correspond to $\alpha = 0.10$ with $\beta = 1, 2, 5$; Tables 7-9 correspond to $\alpha = 0.80$ with $\beta = 1, 2, 5$; and Tables 10-12 correspond to $\alpha = 1.20$ with $\beta = 1, 2, 5$. The key findings are as follows:

(i) **Sample size effect**: As expected, both estimators show improved performance with increasing sample size across all scenarios.

(ii) **Performance for extremely heavy-tailed distributions ($\alpha = 0.05$ and $\alpha = 0.10$)**: For estimating the shape parameter α, the proposed LM estimator is highly competitive with ML. The differences in bias, RMSE, and MAE between the two estimators are very small across all sample sizes and scale parameters. Given the closed-form simplicity and computational efficiency of the LM estimator, it is recommended for such extremely heavy-tailed cases.

(iii) **Performance for moderately heavy-tailed distributions ($\alpha = 0.80$ and $\alpha = 1.20$)**: For these cases, the ML estimator generally outperforms the LM estimator for both α and β. However, both estimators improve with increasing sample size, and the LM estimator remains a viable alternative due to its computational advantages.

(iv) **Performance for the scale parameter β**: The ML estimator consistently provides better estimates for β across all scenarios. This is expected since $\hat{\beta}_{\text{ML}} = X_{(1)}$ is the natural and efficient estimator for the threshold parameter.

(v) **Advantages of the LM estimator**: The main advantages of the proposed LM estimator are its closed-form simplicity (no iterative numerical procedures required), computational efficiency (can be computed instantly), and existence for all $\alpha > 0$ without the restrictive condition $\alpha > 2$ required by the classical method of moments.

Table 1: Comparing estimators with assumed values $\alpha = 0.05$, $\beta = 1$

Scenario	n	Method	Bias	RMSE	MAE	Variance
α	20	ML	0.00541	0.01421	0.01054	0.00017
		LM	0.00662	0.01852	0.01347	0.00030
	50	ML	0.00201	0.00783	0.00602	0.00006
		LM	0.00284	0.01063	0.00821	0.00010
	100	ML	0.00102	0.00533	0.00416	0.00003
		LM	0.00142	0.00732	0.00578	0.00005
	200	ML	0.00044	0.00363	0.00289	0.00001
		LM	0.00056	0.00503	0.00400	0.00003
β	20	ML	8.37758	78.01623	8.37758	6017.55214
		LM	702.80650	10122.03544	703.37740	101982060.94765
	50	ML	0.64471	1.29845	0.64471	1.27058
		LM	29.29322	833.48523	29.85345	693978.33320
	100	ML	0.24929	0.39489	0.24929	0.09381
		LM	3.86222	11.74362	4.40981	123.02036
	200	ML	0.10888	0.16187	0.10888	0.01435
		LM	1.33558	3.83574	1.84244	12.93172

Table 2: Comparing estimators with assumed values $\alpha = 0.05$, $\beta = 2$

Scenario	n	Method	Bias	RMSE	MAE	Variance
α	20	ML	0.00574	0.01487	0.01083	0.00019
		LM	0.00710	0.01901	0.01378	0.00031
	50	ML	0.00204	0.00791	0.00611	0.00006
		LM	0.00295	0.01089	0.00833	0.00011
	100	ML	0.00100	0.00530	0.00417	0.00003
		LM	0.00141	0.00722	0.00568	0.00005
	200	ML	0.00052	0.00358	0.00286	0.00001
		LM	0.00079	0.00498	0.00392	0.00002
β	20	ML	17.44739	190.34623	17.44739	35934.46336
		LM	2306.37786	61812.39936	2307.50261	3816216578.88087
	50	ML	1.37419	2.89575	1.37419	6.49827
		LM	34.73044	182.21227	35.83480	32001.50682
	100	ML	0.48852	0.82038	0.48852	0.43446
		LM	8.23036	28.44072	9.31881	741.28408
	200	ML	0.21925	0.32345	0.21925	0.05656
		LM	3.04622	8.48485	4.01868	62.72584

Table 3: Comparing estimators with assumed values $\alpha = 0.05$, $\beta = 5$

Scenario	n	Method	Bias	RMSE	MAE	Variance
α	20	ML	0.00578	0.01486	0.01088	0.00019
		LM	0.00717	0.01914	0.01391	0.00031
	50	ML	0.00206	0.00786	0.00604	0.00006
		LM	0.00290	0.01056	0.00811	0.00010
	100	ML	0.00099	0.00529	0.00414	0.00003
		LM	0.00145	0.00725	0.00566	0.00005
	200	ML	0.00050	0.00368	0.00290	0.00001
		LM	0.00070	0.00508	0.00400	0.00003
β	20	ML	32.66829	328.03837	32.66829	106563.26479
		LM	4128.86476	85777.09070	4131.72069	7342130190.94824
	50	ML	3.31869	7.19167	3.31869	40.71452
		LM	94.83789	655.55372	97.62656	420840.61682
	100	ML	1.25470	2.21422	1.25470	3.32918
		LM	22.63877	88.02074	25.35821	7236.58460
	200	ML	0.55079	0.81810	0.55079	0.36599
		LM	7.01421	20.75794	9.45365	381.76937

Table 4: Comparing estimators with assumed values $\alpha = 0.10$, $\beta = 1$

Scenario	n	Method	Bias	RMSE	MAE	Variance
α	20	ML	0.01082	0.02896	0.02138	0.00072
		LM	0.01367	0.03724	0.02713	0.00120
	50	ML	0.00401	0.01551	0.01193	0.00022
		LM	0.00551	0.02059	0.01592	0.00039
	100	ML	0.00176	0.01045	0.00823	0.00011
		LM	0.00277	0.01437	0.01122	0.00020
	200	ML	0.00098	0.00727	0.00571	0.00005
		LM	0.00141	0.01011	0.00800	0.00010
β	20	ML	1.03746	2.86151	1.03746	7.11332
		LM	7.69243	78.27504	8.13617	6069.02262
	50	ML	0.24608	0.40018	0.24608	0.09961
		LM	1.39291	3.95964	1.81807	13.74126
	100	ML	0.10934	0.16273	0.10934	0.01453
		LM	0.64873	1.67174	1.02525	2.37434
	200	ML	0.05269	0.07650	0.05269	0.00308
		LM	0.29946	0.92654	0.62938	0.76896

Table 5: Comparing estimators with assumed values $\alpha = 0.10$, $\beta = 2$

Scenario	n	Method	Bias	RMSE	MAE	Variance
α	20	ML	0.01162	0.03021	0.02193	0.00078
		LM	0.01400	0.03775	0.02752	0.00123
	50	ML	0.00414	0.01587	0.01232	0.00023
		LM	0.00561	0.02146	0.01651	0.00043
	100	ML	0.00223	0.01060	0.00833	0.00011
		LM	0.00307	0.01448	0.01131	0.00020
	200	ML	0.00091	0.00710	0.00563	0.00005
		LM	0.00140	0.01012	0.00803	0.00010
β	20	ML	2.00805	6.32895	2.00805	36.03060
		LM	12.84226	66.25968	13.75800	4226.26710
	50	ML	0.50730	0.82968	0.50730	0.43110
		LM	2.96781	8.34968	3.84749	60.92145
	100	ML	0.22307	0.32509	0.22307	0.05593
		LM	1.24200	3.31488	2.00299	9.44775
	200	ML	0.10505	0.15340	0.10505	0.01250
		LM	0.63185	1.92266	1.30597	3.29807

Table 6: Comparing estimators with assumed values $\alpha = 0.10$, $\beta = 5$

Scenario	n	Method	Bias	RMSE	MAE	Variance
α	20	ML	0.01129	0.02947	0.02174	0.00074
		LM	0.01414	0.03824	0.02794	0.00126
	50	ML	0.00465	0.01595	0.01219	0.00023
		LM	0.00627	0.02113	0.01624	0.00041
	100	ML	0.00195	0.01062	0.00826	0.00011
		LM	0.00280	0.01468	0.01145	0.00021
	200	ML	0.00095	0.00727	0.00575	0.00005
		LM	0.00138	0.01016	0.00802	0.00010
β	20	ML	4.87259	14.36997	4.87259	182.79056
		LM	30.45670	193.21878	32.68798	36413.16767
	50	ML	1.25342	2.08490	1.25342	2.77630
		LM	7.10976	18.33073	9.22001	285.52412
	100	ML	0.56727	0.84774	0.56727	0.39694
		LM	3.27121	8.78485	5.21594	66.48604
	200	ML	0.26701	0.38807	0.26701	0.07932
		LM	1.52268	4.72661	3.20505	20.02628

Table 7: Comparing estimators with assumed values $\alpha = 0.80$, $\beta = 1$

Scenario	n	Method	Bias	RMSE	MAE	Variance
α	20	ML	0.08902	0.23197	0.17044	0.04589
		LM	0.11129	0.29731	0.21897	0.07602
	50	ML	0.03165	0.12515	0.09701	0.01466
		LM	0.04547	0.16835	0.13005	0.02628
	100	ML	0.01752	0.08407	0.06562	0.00676
		LM	0.02470	0.11871	0.09303	0.01349
	200	ML	0.00795	0.05862	0.04619	0.00337
		LM	0.01109	0.08128	0.06478	0.00648
β	20	ML	0.06725	0.09797	0.06725	0.00508
		LM	0.08370	0.25683	0.19916	0.05897
	50	ML	0.02555	0.03668	0.02555	0.00069
		LM	0.03941	0.16758	0.13182	0.02653
	100	ML	0.01282	0.01809	0.01282	0.00016
		LM	0.01865	0.12120	0.09697	0.01434
	200	ML	0.00636	0.00901	0.00636	0.00004
		LM	0.00906	0.08689	0.06945	0.00747

Table 8: Comparing estimators with assumed values $\alpha = 0.80$, $\beta = 2$

Scenario	n	Method	Bias	RMSE	MAE	Variance
α	20	ML	0.08421	0.22962	0.16869	0.04564
		LM	0.10491	0.29869	0.21809	0.07823
	50	ML	0.03495	0.12570	0.09729	0.01458
		LM	0.04726	0.16989	0.13100	0.02663
	100	ML	0.01795	0.08492	0.06590	0.00689
		LM	0.02429	0.11705	0.09138	0.01311
	200	ML	0.00803	0.05753	0.04570	0.00325
		LM	0.01212	0.08024	0.06335	0.00629
β	20	ML	0.13481	0.19803	0.13481	0.02105
		LM	0.16081	0.52866	0.40575	0.25367
	50	ML	0.05209	0.07509	0.05209	0.00293
		LM	0.07455	0.33705	0.26613	0.10806
	100	ML	0.02555	0.03608	0.02555	0.00065
		LM	0.03646	0.24088	0.19165	0.05670
	200	ML	0.01306	0.01852	0.01306	0.00017
		LM	0.02140	0.17000	0.13574	0.02845

Table 9: Comparing estimators with assumed values $\alpha = 0.80$, $\beta = 5$

Scenario	n	Method	Bias	RMSE	MAE	Variance
α	20	ML	0.08852	0.23451	0.17127	0.04717
		LM	0.11444	0.30418	0.22227	0.07945
	50	ML	0.03033	0.12483	0.09600	0.01466
		LM	0.04310	0.16873	0.13038	0.02662
	100	ML	0.01772	0.08503	0.06628	0.00692
		LM	0.02425	0.11782	0.09206	0.01330
	200	ML	0.00762	0.05803	0.04613	0.00331
		LM	0.01109	0.08030	0.06327	0.00633
β	20	ML	0.32425	0.47322	0.32425	0.11883
		LM	0.42349	1.28276	1.00427	1.46643
	50	ML	0.13264	0.18926	0.13264	0.01823
		LM	0.19200	0.84054	0.66475	0.66978
	100	ML	0.06287	0.08920	0.06287	0.00401
		LM	0.08887	0.59280	0.47040	0.34359
	200	ML	0.03138	0.04428	0.03138	0.00098
		LM	0.04777	0.42756	0.34059	0.18056

Table 10: Comparing estimators with assumed values $\alpha = 1.2$, $\beta = 1$

Scenario	n	Method	Bias	RMSE	MAE	Variance
α	20	ML	0.13665	0.35296	0.25901	0.10593
		LM	0.17750	0.46122	0.33427	0.18125
	50	ML	0.05044	0.18999	0.14540	0.03356
		LM	0.06506	0.25247	0.19391	0.05952
	100	ML	0.02398	0.12626	0.09826	0.01537
		LM	0.03265	0.17281	0.13556	0.02880
	200	ML	0.01167	0.08565	0.06770	0.00720
		LM	0.01654	0.12099	0.09609	0.01437
β	20	ML	0.04369	0.06344	0.04369	0.00212
		LM	0.05300	0.16853	0.13159	0.02560
	50	ML	0.01689	0.02415	0.01689	0.00030
		LM	0.01941	0.10971	0.08654	0.01166
	100	ML	0.00834	0.01192	0.00834	0.00007
		LM	0.00998	0.07918	0.06266	0.00617
	200	ML	0.00426	0.00606	0.00426	0.00002
		LM	0.00523	0.05773	0.04587	0.00331

Table 11: Comparing estimators with assumed values $\alpha = 1.2$, $\beta = 2$

Scenario	n	Method	Bias	RMSE	MAE	Variance
α	20	ML	0.13037	0.34916	0.25558	0.10494
		LM	0.16148	0.44684	0.32284	0.17362
	50	ML	0.04637	0.18514	0.14406	0.03213
		LM	0.06661	0.25496	0.19680	0.06058
	100	ML	0.02509	0.13005	0.10187	0.01629
		LM	0.03408	0.17731	0.13840	0.03028
	200	ML	0.01336	0.08793	0.06956	0.00756
		LM	0.01794	0.12264	0.09639	0.01472
β	20	ML	0.08915	0.12959	0.08915	0.00885
		LM	0.09901	0.33713	0.26264	0.10387
	50	ML	0.03499	0.04947	0.03499	0.00122
		LM	0.04530	0.22624	0.17879	0.04914
	100	ML	0.01681	0.02396	0.01681	0.00029
		LM	0.02023	0.16071	0.12761	0.02542
	200	ML	0.00833	0.01189	0.00833	0.00007
		LM	0.00989	0.11501	0.09143	0.01313

Table 12: Comparing estimators with assumed values $\alpha = 1.2$, $\beta = 5$

Scenario	n	Method	Bias	RMSE	MAE	Variance
α	20	ML	0.12921	0.34545	0.25689	0.10266
		LM	0.16484	0.44328	0.32806	0.16935
	50	ML	0.04861	0.18836	0.14511	0.03312
		LM	0.06775	0.25058	0.19328	0.05821
	100	ML	0.02304	0.12645	0.09939	0.01546
		LM	0.03481	0.17481	0.13661	0.02935
	200	ML	0.01011	0.08681	0.06853	0.00744
		LM	0.01775	0.11951	0.09489	0.01397
β	20	ML	0.21446	0.31143	0.21446	0.05101
		LM	0.24847	0.83099	0.64440	0.62893
	50	ML	0.08526	0.12211	0.08526	0.00764
		LM	0.11358	0.55840	0.44066	0.29896
	100	ML	0.04278	0.06118	0.04278	0.00191
		LM	0.06087	0.39951	0.31703	0.15594
	200	ML	0.02091	0.02966	0.02091	0.00044
		LM	0.03667	0.28125	0.22410	0.07777

6.4 Robustness to Extreme Observations

To evaluate the sensitivity of both estimators to extreme observations, we conducted an additional simulation study. We generated samples from the Pareto distribution and artificially contaminated them by replacing the largest observation $X_{(n)}$ with $X_{(n)} \times c$, where $c = 10, 50, 100$ represents the contamination factor. We then compared the bias and RMSE of both estimators under contamination. The results are presented in Tables 13 through 15.

Table 13: Robustness analysis for $\alpha = 0.03$ (extremely heavy-tailed) with $\beta = 1$

n	c	Bias		RMSE	
		ML	LM	ML	LM
50	10	0.0012	0.0015	0.0046	0.0062
50	50	0.0011	0.0013	0.0046	0.0062
50	100	0.0012	0.0014	0.0047	0.0062
100	10	0.0006	0.0007	0.0032	0.0043
100	50	0.0006	0.0007	0.0032	0.0043
100	100	0.0006	0.0007	0.0032	0.0043

Table 14: Robustness analysis for $\alpha = 0.10$ (extremely heavy-tailed) with $\beta = 1$

n	c	Bias		RMSE	
		ML	LM	ML	LM
50	10	0.0035	0.0036	0.0152	0.0200
50	50	0.0031	0.0022	0.0150	0.0195
50	100	0.0033	0.0021	0.0154	0.0193
100	10	0.0017	0.0015	0.0105	0.0142
100	50	0.0016	0.0010	0.0105	0.0139
100	100	0.0015	0.0008	0.0106	0.0140

Table 15: Robustness analysis for $\alpha = 0.80$ (moderately heavy-tailed) with $\beta = 1$

n	c	Bias		RMSE	
		ML	LM	ML	LM
50	10	0.0006	-0.0766	0.1109	0.1441
50	50	-0.0203	-0.1543	0.1071	0.1829
50	100	-0.0265	-0.1828	0.1088	0.2033
100	10	0.0005	-0.0503	0.0803	0.1104
100	50	-0.0094	-0.1001	0.0793	0.1315
100	100	-0.0139	-0.1207	0.0795	0.1456

Based on the results presented in Tables 13, 14, and 15, we observe that for extremely heavy-tailed distributions ($\alpha = 0.03$ and $\alpha = 0.10$), the differences between the two estimators are very small, and the LM estimator remains competitive with ML. For moderately heavy-tailed distributions ($\alpha = 0.80$), the ML estimator shows better performance.

6.5 Summary

The simulation results demonstrate that the proposed LM estimator is a practical alternative to ML, particularly for extremely heavy-tailed distributions ($\alpha = 0.03$ and $\alpha = 0.10$), where it performs competitively for the shape parameter α with negligible differences in bias and RMSE. For moderately heavy-tailed distributions ($\alpha = 0.80$ and $\alpha = 1.20$), ML generally outperforms LM for both α and β. The robustness analysis also shows that for extremely heavy-tailed cases the two estimators have similar performance under contamination, while for $\alpha = 0.80$ ML is more resilient to extreme observations.

Despite these considerations, the proposed LM estimator offers several practical advantages. Its closed-form nature eliminates the need for iterative numerical procedures, making it computationally efficient and free from convergence issues.

Furthermore, unlike the classical method of moments which requires $\alpha > 2$, the LM estimator exists for all $\alpha > 0$, covering the entire range of practical interest. Therefore, the LM estimator is recommended for applications where computational simplicity and closed-form solutions are prioritized, particularly when dealing with extremely heavy-tailed data where its performance for α is competitive with ML.

7 Illustrative Example: Norwegian Fire Insurance Claims

The Pareto distribution has found extensive applications in various fields. In economics, it models the distribution of income and wealth, capturing the well-known Pareto principle where a small proportion of the population holds a large fraction of total resources. In finance, it is used for modeling extreme asset returns and calculating value-at-risk (VaR). In risk management and actuarial science, the distribution is widely employed for modeling large insurance claims and operational risk losses. To illustrate the practical utility of the proposed LM estimator in one of these domains, we present an empirical application to insurance data.

To demonstrate the practical application of the proposed LM estimator for heavy-tailed data, we analyze the well-known Norwegian fire insurance claims dataset. This dataset, which has become a benchmark in the actuarial literature for studying extreme losses, contains 9,181 fire insurance claims from a Norwegian insurance company over the period 1972 to 1992, with claim amounts expressed in thousands of Norwegian Kroner (NOK) and a deductible of 500 NOK [3]. The data are publicly available through the `ReIns` package in R.

The dataset has been extensively used in recent studies of heavy-tailed phenomena and risk assessment. [5] analyzed severity and tail risk of these claims for the years 1981 through 1992, while [1] employed the dataset for financial risk analysis and value-at-risk assessment, demonstrating its continued relevance in actuarial research. [4] also utilized this data for studying heavy-tailed distributions and rating in insurance.

We focus on claims from the year 1981, following the methodology of [5]. Table 16 presents the summary statistics for the 1981 claims. The data consist of 429 claims ranging from 500 (the deductible) to 77,839 thousand NOK. The mean claim amount is 1,847 thousand NOK with a standard deviation of 4,485 thousand NOK, indicating substantial variability and a heavy-tailed pattern. The Pareto distribution is particularly suitable for this data because insurance claims exceeding a high threshold typically follow a heavy-tailed distribution, capturing the phenomenon where a small proportion of claims account for a large proportion of total losses.

Table 16: Summary statistics for Norwegian fire insurance claims (1981)

Statistic	Value
Number of claims	429
Minimum (1000 NOK)	500
Maximum (1000 NOK)	77,839
Mean (1000 NOK)	1,847
Median (1000 NOK)	766
Standard deviation (1000 NOK)	4,485

We fit the Pareto distribution to the 1981 claims using both the proposed LM estimator and the traditional ML estimator. Table 17 presents the estimated parameters and goodness-of-fit measures.

Table 17: Parameter estimates and goodness-of-fit for Norwegian fire insurance claims (1981)

Estimator	$\hat{\alpha}$	$\hat{\beta}$	RMSE	KS p-value
ML	1.107	500.00	2048.84	0.059
LM	1.145	515.35	1357.98	0.179

The results demonstrate that the proposed LM estimator produces reliable parameter estimates for this real-world heavy-tailed data. Notably, the LM estimator achieves a substantially lower RMSE (1357.98) compared to the ML estimator (2048.84), indicating a better fit to the data. This represents an improvement of approximately 33.7% in RMSE. The Kolmogorov-Smirnov (KS) test further supports this finding. The KS test compares the empirical distribution of the data with the fitted Pareto distribution; a higher p-value indicates a better fit. The LM estimator yields a p-value of 0.179, which is well above 0.05, confirming that the Pareto distribution provides an adequate fit when estimated by the proposed LM method.

These findings demonstrate that the proposed LM estimator outperforms ML for this extremely heavy-tailed data, achieving a substantially better fit. Given its closed-form computational simplicity and superior goodness-of-fit, the LM estimator offers a practical and efficient alternative for rapid risk assessment, real-time claims analysis, and applications where computational resources are limited.

Conclusion

This paper has introduced a LM estimator for the Pareto distribution, providing closed-form solutions that overcome the computational limitations of traditional ML estimation. The proposed estimator is simple to implement and eliminates the convergence issues inherent in iterative numerical methods.

Theoretical analysis establishes the consistency of the proposed estimator, ensuring its reliability for large-sample applications. Through an extensive simulation study across a wide range of parameter values and sample sizes, we have shown that for extremely heavy-tailed distributions, the differences between the LM and ML estimators are very small, and the two methods perform very similarly. In such challenging scenarios, which are common in finance, insurance, hydrology, and telecommunications, the LM estimator competes closely with ML while offering the distinct advantage of closed-form computation.

The practical relevance of the proposed estimator is further demonstrated through an empirical application to Norwegian fire insurance claims data, a well-known benchmark dataset in the actuarial literature. For the 1981 claims, the LM estimator achieves a substantially lower RMSE compared to the ML estimator, indicating a better fit to the data. The Kolmogorov-Smirnov test also confirms that the Pareto distribution provides an adequate fit when estimated by the LM method, while the ML fit is marginal. These results demonstrate that the LM estimator not only competes with ML but can also provide a better fit for real-world heavy-tailed data.

Given the negligible differences between the two estimators in extremely heavy-tailed cases, and the superior performance of LM in the empirical example, the simplicity and computational efficiency of the LM estimator make it a preferable choice for such applications. For moderately heavy-tailed distributions and for estimating the scale parameter, ML generally provides better performance. Nevertheless, the LM estimator remains a viable alternative due to its computational advantages. Robustness analysis further confirms that in extremely heavy-tailed cases, both estimators exhibit similar behavior under data contamination.

The key advantages of the proposed LM estimator are its closed-form nature, instant computation, and existence for all admissible parameter values, thus avoiding the restrictive conditions imposed by classical moment estimators. These features, combined with its strong performance in both simulations and real-data applications, make it particularly valuable for rapid risk assessment, real-time data analysis, pedagogical purposes, and applications with limited computational resources, especially when dealing with extremely heavy-tailed phenomena.

Bibliography

[1] A. Aljadani, *Extreme PORT for Norwegian fire financial claims: Empirical assessment and financial VAR analysis*, Alexandria Engineering Journal, 108: 852-862 (2024), Elsevier.

[2] B. C. Arnold, *Pareto Distributions*, (1983), Fairland, MD: International Co-operative Publishing House.

[3] J. Beirlant, Y. Goegebeur, J. Segers, and J. L. Teugels, *Statistics of extremes: theory and applications*, (2006) John Wiley & Sons.

[4] J. Beirlant, G. Matthys, and G. Dierckx, *Heavy-tailed distributions and rating*, ASTIN Bulletin: The Journal of the IAA, 31(1): 37-58, (2001) Cambridge University Press.

[5] V. Brazauskas, and A. Kleefeld, *Modeling severity and measuring tail risk of Norwegian fire claims*, North American Actuarial Journal, 20(1): 1-16, (2016) Taylor & Francis.

[6] A. S. HASSAN, E. A. ELSHERPIENY, AND R. E. MOHAMED, *Classical and Bayesian estimation of entropy for Pareto distribution in presence of outliers with application*, SANKHYA A, 85(1): 707-740, (2023) SPRINGER.

[7] J. RM. HOSKING, *L-moments: analysis and estimation of distributions using linear combinations of order statistics*, JOURNAL OF THE ROYAL STATISTICAL SOCIETY SERIES B: STATISTICAL METHODOLOGY, 52(1): 105-124, (1990) OXFORD UNIVERSITY PRESS.

[8] N. L. JOHNSON, S. KOTZ, AND N. BALAKRISHNAN, *Continuous Univariate Distributions* (Vol. 1, 2nd ed., Chapter 20), (1994) JOHN WILEY & SONS.

[9] Z. KARIMI EZMAREH, AND G. YARI, *Bayesian Estimation of the Stress-Strength Reliability Based on Generalized Order Statistics for Pareto Distribution*, JOURNAL OF PROBABILITY AND STATISTICS, 1: 8648261, (2023) WILEY ONLINE LIBRARY.

[10] H. M. KIM, Y. H. JANG, AND B. C. ARNOLD, AND J. ZHAO, *New efficient estimators for the Weibull distribution*, COMMUNICATIONS IN STATISTICS-THEORY AND METHODS, 53(13): 4576-4601, (2024) TAYLOR & FRANCIS.

[11] F. LI, S. WEI AND M. ZHAO, *Bayesian estimation of a new Pareto-type distribution based on mixed Gibbs sampling algorithm*, MATHEMATICS, 12(1): 18, (2023) MDPI.

[12] V. NAWA, S. NADARAJAH, *Logarithmic method of moments estimators for the Fréchet distribution*, JOURNAL OF COMPUTATIONAL AND APPLIED MATHEMATICS, 457:116293, (2025) ELSEVIER.

[13] L. NDWANDWE, J. S. ALLISON, L. SANTANA, AND I. J. H. VISAGIE, *Testing for the Pareto type I distribution: A comparative study*, sc Metron, 81(2): 215-256, (2023) SPRINGER.

[14] V. PARETO, *Cours d'économie politique*, (1964) LIBRAIRIE DROZ, 1.

[15] J. SHARPE, M. A. JUÁREZ, *Estimation of the Pareto and related distributions–A reference-intrinsic approach*, COMMUNICATIONS IN STATISTICS-THEORY AND METHODS, 52(3): 523-542, (2023) TAYLOR & FRANCIS.

[16] W. WANG, Y. GAO, AND N. IKEGAYA, *Approximating wind speed probability distributions around a building by mixture weibull distribution with the methods of moments and L-moments*, JOURNAL OF WIND ENGINEERING AND INDUSTRIAL AERODYNAMICS, 257:106001, (2025) ELSEVIER.

How to Cite: Zahra Karimiezmareh[1], Ghazal Elahi[2], *Closed-Form Estimation for the Pareto Distribution Based on Logarithmic Moments*, Journal of Mathematics and Modeling in Finance (JMMF), Vol. 6, No. 2, Pages:159–182, (2026).