Scientific Research Publishing

Open Journal of Social Sciences, 2025, 13(1), 86-121
https://www.scirp.org/journal/jss
ISSN Online: 2327-5960
ISSN Print: 2327-5952

Pareto Distribution: A Probability Model in Social Research

José Moral de la Rubia

School of Psychology, Universidad Autónoma de Nuevo León, Monterrey, Mexico
Email: jose.morald@uanl.edu.mx

How to cite this paper: Moral de la Rubia, J. (2025). Pareto Distribution: A Probability Model in Social Research. *Open Journal of Social Sciences, 13,* 86-121.
https://doi.org/10.4236/jss.2025.131007

Received: December 9, 2024
Accepted: January 11, 2024
Published: January 14, 2024

Abstract

This methodological article aims to present the type I Pareto distribution in a clear and illustrative manner for better understanding among social researchers. It also provides R scripts for practical application. This continuous distribution, with its inverted J shape, skewness towards the right side, and heavy right tail, serves as an effective probability model for various social variables, such as wealth and income, as well as behaviors that are highly frequent in a few individuals and infrequent in the majority. The type I distribution, which has a scale parameter x_m and a shape parameter α, is introduced, beginning with a brief historical overview. The density, cumulative distribution, tail, moment, and characteristic functions are presented. The article proceeds with descriptive measures, estimators based on the method of moments and maximum likelihood, its relationship with other distributions, and goodness-of-fit tests. This material is applied through two examples: one involving probability and descriptive measure calculations, and the other focused on parameter estimation and fit testing using the Kolmogorov-Smirnov and Anderson-Darling tests. Additionally, scripts were developed to perform the corresponding calculations in R, a freely available software. Simulated data were used in two examples illustrating the application of the distribution. Finally, suggestions for its use are provided.

Keywords

Continuous Probability Distribution, Parameter Estimation, Descriptive Measures, Pareto Tail Index, Gini Concentration Index

1. Introduction

The Pareto distribution is a non-normal continuous distribution that has significant applications in the social and behavioral sciences (Barnoy & Reich, 2022; Feng, Deng, Chen, Perc, & Kurths, 2020). However, its explanation is often theoretical

and confusing, primarily due to its presentation under varying parameterizations that are frequently interchanged (Sarabia, Jorda, & Prieto, 2019). These include the classic two-parameter form, which is the most widely used (type I: scale parameter x_m and shape parameter α); the three-parameter forms (type II: location parameter μ, x_m, and α; and type III: μ, x_m, and shape parameter γ); and the four-parameter form (type IV: μ, x_m, α, and γ). Additionally, the distribution is predominantly developed with an economic focus (Barczy, Nedényi, & Sütő, 2023). The purpose of this article is, therefore, to present the type I Pareto distribution in a clear, comprehensible, and illustrative manner for social researchers, including psychologists, sociologists, and social workers.

The type I Pareto (1896; 1897) distribution was developed between 1896 and 1897 by the Italian civil engineer, economist, and sociologist Vilfredo Federico Damaso Pareto (1848-1923) during his political economy course at the University of Lausanne (UNIL) in Switzerland. It was initially conceived to describe the distribution of land, wealth, and income within a society. From these studies, the 80/20 principle was derived, positing that 80% of wealth is concentrated among 20% of the population. Notably, both the distribution and the inequality principle are now applied as probabilistic models for similar variables (Charpentier & Flachaire, 2022; Feng et al., 2020), although contemporary inequality levels tend to be lower, with a maximum ratio of 70/30 (McCarthy & Winer, 2019).

Furthermore, the Pareto distribution and inequality rule have found applications beyond economics, serving as models in fields such as psychology (Campbell & Brauer, 2021; Rajeev, 2022), virtual education (Valkanas & Diamandis, 2022), engineering (Chen, Zhang, Wang, Jiang, & Liu, 2019; Sudharson et al., 2022; Sudharson & Prabha, 2019), climatology (Le Gall, Favre, Naveau, & Prieur, 2022), and physics (Rácz et al., 2023; Cheng et al., 2023). For instance, claims to insurers for accidents and illnesses often follow a Pareto distribution, prompting these companies to impose disproportionately high premiums on certain sectors of the population, such as older adults (Diawara, Kane, Dembele, & Lo, 2021; Zhang, Wu, & Yao, 2022).

2. Characterization of the Distribution

2.1. Parameters and Support

This continuous distribution, belonging to the family of power distributions, is defined by two parameters in its simplest or type I form (Ahmad & Almetwally, 2020). The first parameter, a scale parameter often denoted as x_m (though its notation varies widely), represents both the peak or mode and the minimum value of the distribution (Chattamvelli & Shanmugam, 2021). The second parameter, a shape parameter denoted as α (Fedotenkov, 2020), is commonly referred to as the Pareto tail index (Andria, 2022). The parameter space for both spans the interval $(0, \infty)$, while the support of the distribution is defined over the interval $[x_m, \infty)$.

The profile of its density function depicts a curve that sharply decreases from its maximum value and becomes asymptotic to the horizontal axis, exhibiting

positive skewness and leptokurtosis with a long right tail (Figure 1). The value of α that precisely models the 80/20 law is $\log_4(5) = \ln(5)/\ln(4) \approx 1.161$, which is considered too extreme for wealth distributions, where α is typically greater than 1.5 (Yang & Zhou, 2022).

$$X \sim \text{Pareto}\left(x_m, \alpha\right)$$

$$x_m \text{ and } \alpha \in \left(0, \infty\right) \text{ and } x \in \left[x_m, \infty\right)$$

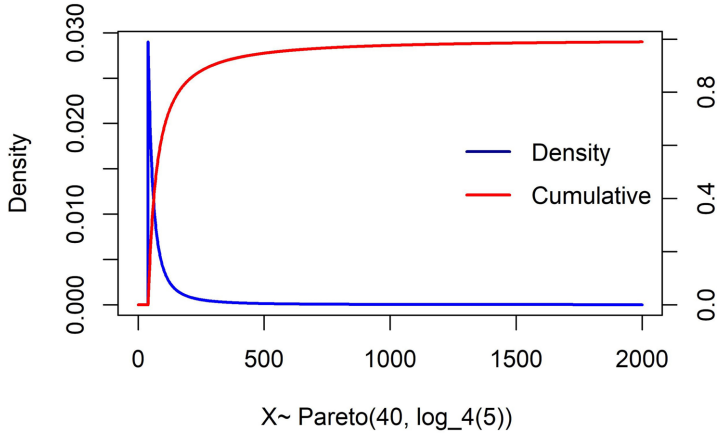

Figure 1. Density function $f_X(x)$ and cumulative distribution function $F_X(x)$ of hourly income modeled by a type I Pareto distribution with scale parameter $x_m = 40$ and shape $\alpha = \log_4(5)$.

2.2. Functions

2.2.1. Density Function

The definite integral of the density function allows the calculation of the probability that a value x of the continuous random variable X falls within a given integration interval, which must lie within the domain of the distribution. It is denoted by $f_X(x)$. The analytical expression of this function for the type I Pareto distribution is shown in Equation 1.

$$f_X\left(x\right) = \alpha \frac{x_m^{\alpha}}{x^{\alpha+1}}, x \geq x_m \tag{1}$$

2.2.2. Cumulative Distribution Function

This function provides the probability that a continuous random variable X takes a value between the lower limit of the distribution and a specified value x, where x belongs to the domain of the distribution. it is denoted by $F_X(x)$. The analytical expression of this function for the type I Pareto distribution is presented in Equation 2.

$$F_X\left(x\right) = P\left(X \leq x\right) = \int_{x_m}^{x} f_X\left(x\right) \mathrm{d}x = 1 - \left(\frac{x_m}{x}\right)^{\alpha}, x \geq x_m \tag{2}$$

2.2.3. Complementary Cumulative Distribution Function

Also known as the tail function, it is denoted by $\overline{F}_X\left(x\right)$. This function provides

the probability that a continuous random variable X takes a value between a specified value x and the upper limit of the distribution, where x belongs to the domain of the distribution. The analytical expression for this function in the case of the type I Pareto distribution is presented in Equation 3.

$$\bar{F}_X(x) = P(X > x) = \int_{x_m}^{\infty} f_X(x)\,\mathrm{d}x = \left(\frac{x_m}{x}\right)^{\alpha}, x \geq x_m \tag{3}$$

2.2.4. Quantile Function

It is the inverse of the cumulative distribution function. It is denoted by $Q_X(x)$ or $F_X^{-1}(x)$. It provides the value of X that accumulates the probability corresponding to the given argument, where the argument takes values in the interval (0, 1). The analytical expression for this function in the case of the type I Pareto distribution is shown in Equation 4.

$$Q_X(p) = \frac{x_m}{\sqrt[\alpha]{1-p}} = x_m (1-p)^{-\frac{1}{\alpha}}, p \in (0,1) \tag{4}$$

2.2.5. Moment Generating Function

The k-th derivative of this function, evaluated at the point 0, gives the k-th order moments of the distribution, where the k-th order moment is the mathematical expectation of the variable raised to the k-th power. It is denoted by $M_X(t)$. The moment generating function of the type I Pareto distribution is not defined for a non-trivial interval of t around 0, but it is defined for values of $t \leq 0$. Its analytical expression is shown in Equation 5.

$$M_X(t) = E\left(\mathrm{e}^{tx}\right) = \int_{x_m}^{+\infty} \mathrm{e}^{tx} f_X(x)\,\mathrm{d}x = \begin{cases} \alpha(-x_m t)^{\alpha}\,\Gamma(-\alpha, -x_m t) & t < 0 \\ 1 & t = 0 \\ \text{Undefined} & t > 0 \end{cases} \tag{5}$$

Upper incomplete gamma function

$$\Gamma(-\alpha, -x_m t) = \int_{-x_m t}^{\infty} u^{-\alpha-1} \mathrm{e}^{-u}\,\mathrm{d}u$$

2.2.6. Characteristic Function

The characteristic function corresponds to the Fourier transform (with a sign inversion) applied to the analytical expression of the density function. It allows for the analysis of the behavior of the moments of a distribution. It is denoted by $C_X(t)$. The analytical expression of this function for the type I Pareto distribution is shown in Equation 6, where $i = \sqrt{-1}$ and $t \in \mathrm{R}$.

$$\begin{aligned} &C_X(t): \mathbb{R} \to \mathbb{C}; \\ &C_X(t) = E\left(\mathrm{e}^{itx}\right) = \int_{x_m}^{\infty} \mathrm{e}^{itx} f_X(x)\,\mathrm{d}x = \alpha(-ix_m t)^{\alpha}\,\Gamma(-\alpha, -ix_m t) \end{aligned} \tag{6}$$

2.3. Descriptive Measures

2.3.1. Measures of Central Tendency

Mathematical expectation or arithmetic mean: This measure corresponds to the

first (non-central) moment. For the type I Pareto distribution, it can be calculated using its two parameters, as shown in Equation 7. Its computation requires the shape parameter α to be greater than 1.

$$\mu(X) = E(X) = \int_{x_m}^{\infty} x f_X(x)\,\mathrm{d}x = \frac{\alpha}{\alpha - 1} x_m,\ \alpha > 1 \tag{7}$$

Geometric mean: It is the antilogarithm of the mathematical expectation of the logarithm of the values, typically using the natural base. For the type I Pareto distribution, it can be calculated from its two parameters, as shown in Equation 8.

$$\mu_g(X) = G(X) = \mathrm{e}^{E(\ln(X))} = x_m \sqrt[\alpha]{\mathrm{e}} = x_m \mathrm{e}^{1/\alpha} \tag{8}$$

Harmonic Mean: It is the reciprocal of the mathematical expectation of the reciprocals of the values. For the type I Pareto distribution, it can be calculated using its two parameters, as shown in Equation 9.

$$\mu_h(X) = H(X) = \frac{1}{E(1/X)} = x_m \left(1 + \frac{1}{\alpha}\right) \tag{9}$$

Median: It is the 0.5 quantile. For the type I Pareto distribution, it can be calculated using its two parameters, as shown in the Equation 10.

$$Mdn(X) = Q_X(p = 0.5) = \frac{x_m}{\sqrt[\alpha]{1/2}} = x_m \sqrt[\alpha]{2} \tag{10}$$

Mode: It corresponds to the peak of the distribution. The type I Pareto distribution is unimodal, with its single peak occurring at the scale parameter x_m (Equation 11).

$$Mo(X) = \left\{x \mid \max\left[f_X(x)\right]\right\} = x_m \tag{11}$$

Non-central k-th moment: It is the mathematical expectation of the values raised to the k-th power. For the type I Pareto distribution, it can be calculated using its two parameters, as shown in Equation 12. For its calculation, the shape parameter α must be greater than the power k to which the values are raised.

$$E\left(X^k\right) = \int_{x_m}^{\infty} x^k f_X(x)\,\mathrm{d}x = \frac{\alpha}{\alpha - k} x_m^k,\ \alpha > k \tag{12}$$

2.3.2. Measures of Variation

Variance: It corresponds to the second central moment or the expected value of the squared differences from the arithmetic mean. For the type I Pareto distribution, it can be calculated using its two parameters, as shown in in Equation 13. Its computation requires the shape parameter α to be greater than 2.

$$\begin{aligned} Var(X) = \sigma^2(X) &= E\left[\left(X - E(X)\right)^2\right] = \int_{x_m}^{\infty} \left(x - E(X)\right)^2 f_X(x)\,\mathrm{d}x \\ &= \frac{\alpha x_m^2}{(\alpha - 1)^2 (\alpha - 2)},\ \alpha > 2 \end{aligned} \tag{13}$$

Standard deviation: It is the square root of the variance (Equation 14).

$$DE(X)=\sigma(X)=\sqrt{Var(X)}=\frac{x_m}{\alpha-1}\sqrt{\frac{\alpha}{\alpha-2}},\ \alpha>2 \tag{14}$$

Entropy: It corresponds to the mathematical expectation of the additive inverses of the logarithms of the densities, or the mathematical expectation of the information of the values. When the natural logarithm is used, entropy is expressed in natural units of information (nats), which is the most common. If base 10 is used, it is expressed in decimal units of information (dits); whereas, if base 2 is used, it is expressed in binary units of information (bits). For the type I Pareto distribution, it can be calculated using its two parameters, as shown in Equation 15.

$$\begin{aligned} H(X)&=E\left(-\ln\left[f_X(x)\right]\right)=-\int_{x_m}^{\infty}f_X(x)\ln\left[f_X(x)\right]\mathrm{d}x=\ln\left[\frac{x_m}{\alpha}\mathrm{e}^{1+\frac{1}{\alpha}}\right] \\ &=\ln(x_m)-\ln(\alpha)+1+1/\alpha \end{aligned} \tag{15}$$

2.3.2. Measures of Shape

Skewness: It corresponds to the third standardized central moment. Karl Pearson's original notation is used: $\sqrt{\beta_1}$ (Pearson, 1895). For the type I Pareto distribution, it can be calculated using the shape parameter α, as shown in Equation 16. Its computation requires this parameter to be greater than 3.

$$\sqrt{\beta_1(X)}=\frac{E\left[\left(X-E(X)\right)^3\right]}{\left(E\left[\left(X-E(X)\right)^2\right]\right)^{3/2}}=\frac{2(\alpha+1)}{\alpha-3}\sqrt{\frac{\alpha-2}{\alpha}},\ \alpha>3 \tag{16}$$

Excess Kurtosis: It corresponds to the standardized fourth central moment minus the kurtosis value of the normal distribution, which is 3. Karl Pearson's original notation is used: β_2-3 (Pearson, 1905). For the type I Pareto distribution, it can be calculated using the shape parameter α, as shown in Equation 17. Its computation requires this parameter to be greater than 4.

$$\beta_2(X)-3=\frac{E\left[\left(X-E(X)\right)^4\right]}{\left(E\left[\left(X-E(X)\right)^2\right]\right)^2}=\frac{6\left(\alpha^3+\alpha^2-6\alpha-2\right)}{\alpha(\alpha-3)(\alpha-4)},\ \alpha>4 \tag{17}$$

3. Parameter Estimation

Let X be a random sample of size n from a continuous quantitative variable that follows a type I Pareto distribution with parameters x_m and α. For example, the sample data could consist of records of the monthly salaries of randomly chosen workers in a large firm. Refer to Equation 18 for its notation.

$$x=\{x_1,x_2,\cdots,x_n\}\subseteq X\sim\text{Pareto}(x_m,\alpha) \tag{18}$$

3.1. Estimator of α by the Method of Moments

The estimator of the shape parameter α by the method of moments is obtained from the mathematical expectation or arithmetic mean of X, if the value of the location

parameter x_m, which is the peak and minimum value of the distribution, is known. This mathematical expectation corresponds to the quotient of the sample mean (numerator) and the difference between the sample mean and the known parameter x_m (denominator). It is valid when α > 2. For more details, refer to Equation 19.

$$\begin{aligned} &E(X)=\frac{\alpha}{\alpha-1}x_m \\ &\alpha E(X)-E(X)-\alpha x_m=0 \\ &\alpha\left(E(X)-x_m\right)=E(X) \\ &\alpha=\frac{E(X)}{E(X)-x_m}=\frac{\mu(X)}{\mu(X)-x_m} \\ &\hat{\alpha}=\frac{\hat{E}(X)}{\hat{E}(X)-x_m}=\frac{\hat{\mu}(X)}{\hat{\mu}(X)-x_m}=\frac{\overline{x}}{\overline{x}-x_m},\ \text{where } \hat{\mu}(X)=\overline{x}=\frac{\sum_{i=1}^{n}x_i}{n} \end{aligned} \tag{19}$$

The point estimator converges to a normal distribution, since X has a distribution with finite moments (Rao, 1973), which allows for the computation of an asymptotic interval estimator (Equation 20).

$$\begin{aligned} &\hat{\alpha}_n=\frac{\overline{x}}{\overline{x}-x_m}\xrightarrow[d]{}N\left(\alpha,\frac{\alpha(\alpha-1)^2}{n(\alpha-2)}\right) \\ &P\left(\hat{\alpha}-z_{1-\frac{\alpha}{2}}\times\sqrt{\frac{\hat{\alpha}(\hat{\alpha}-1)^2}{n(\hat{\alpha}-2)}}\le\alpha\le\hat{\alpha}+z_{1-\frac{\alpha}{2}}\times\sqrt{\frac{\hat{\alpha}(\hat{\alpha}-1)^2}{n(\hat{\alpha}-2)}}\right) \end{aligned} \tag{20}$$

3.2. Estimators of x_m and α by the Maximum Likelihood Method

The estimator by the maximum likelihood method of the scale parameter x_m is the minimum sample value, and the estimator of the shape parameter α is the inverse of the arithmetic mean of the logarithms of the ratios between each value x_i and the minimum sample value (Siudem, Nowak, & Gagolewski, 2022). See Equation 21 and Equation 22 for these estimators, respectively.

$$\hat{x}_m=\min\left(x_1,x_2,\cdots,x_n\right)=\min\left(\left\{x_i\right\}_{i=1}^{n}\right) \tag{21}$$

$$\hat{\alpha}=\frac{n}{\sum_{i=1}^{n}\ln\left(\frac{x_i}{\hat{x}_m}\right)}=\frac{n}{\sum_{i=1}^{n}\ln\left(\frac{x_i}{\min\left(\left\{x_i\right\}_{i=1}^{n}\right)}\right)} \tag{22}$$

The Fisher information for n data points for these two parameters is given by the 2 × 2 matrix in Equation 23.

$$\begin{aligned} &I\left(x_m,\alpha\,|\,x\right) \\ &=n\begin{pmatrix} -E\left[\frac{\delta}{\delta x_m\delta x_m}\ln f_X\left(x\,|\,x_m,\alpha\right)\right] & -E\left[\frac{\delta}{\delta x_m\delta\alpha}\ln f_X\left(x\,|\,x_m,\alpha\right)\right] \\ -E\left[\frac{\delta}{\delta\alpha\delta x_m}\ln f_X\left(x\,|\,x_m,\alpha\right)\right] & -E\left[\frac{\delta}{\delta\alpha\delta\alpha}\ln f_X\left(x\,|\,x_m,\alpha\right)\right] \end{pmatrix} \end{aligned}$$

$$= n\begin{pmatrix} \alpha/x_m^2 & -1/x_m \\ -1/x_m & 1/\alpha^2 \end{pmatrix} = \begin{pmatrix} n\alpha/x_m^2 & -n/x_m \\ -n/x_m & n/\alpha^2 \end{pmatrix} \tag{23}$$

The variance of a parameter θ, *Var*(θ), is always greater than or equal to its Cramér-Rao lower bound CRLB(θ) or the inverse of its Fisher information for *n* data, $1/I(\theta)$ (Xu, Sedory, & Singh, 2022). Refer to Equation 24.

$$Var\left(\hat{\theta}\right) = \sigma_{\hat{\theta}}^2 \geq CICR\left(\hat{\theta}\right) = 1/I\left(\theta\right) \tag{24}$$

The estimators obtained by the maximum likelihood method have asymptotic properties of unbiasedness (Equation 25), consistency (Equation 26), efficiency (Equation 27), and normality (Equation 28), which make them very useful (Song, Roung-Park, & Yoon, 2022).

$$\hat{\theta}_n \text{ is an asymptotically unbiased estimator of } \theta \text{ if } \lim_{n\to\infty} E\left(\hat{\theta}_n\right) = \theta \tag{25}$$

$$\hat{\theta}_n \text{ is an asymptotically consistent estimator of } \theta \text{ if } \lim_{n\to\infty} P\left(\left|\hat{\theta}_n - \theta\right| > \varepsilon\right) = 0 \tag{26}$$

$$\hat{\theta}_n \text{ is an asymptotically efficient estimator of } \theta \text{ if } \lim_{n\to\infty} \frac{CICR\left(\hat{\theta}\right)}{Var\left(\hat{\theta}_n\right)} = 1 \tag{27}$$

$$\hat{\theta}_n \text{ as an estimator } \theta \text{ converges to normality if } \lim_{n\to\infty} \frac{\sqrt{n}\left(\hat{\theta}_n - \theta\right)}{DE\left(\hat{\theta}\right)} \xrightarrow[d]{} N\left(0,1\right) \tag{28}$$

Based on these asymptotic properties, asymptotic standard errors (*aes*) can be defined for the maximum likelihood estimators of x_m (Equation 29) and α (Equation 31), along with asymptotic confidence intervals (Equation 30 for the estimator of x_m and Equation 32 for the estimator of α).

$$ase\left(\hat{\theta}\right) = \sigma_{\hat{\theta}} = \sqrt{1/I\left(\theta\right)}$$
$$ase\left(\hat{x}_m\right) = \sigma_{\hat{x}_m} = \sqrt{\frac{1}{n\alpha/x_m^2}} = \sqrt{\frac{x_m^2}{n\alpha}} = \frac{x_m}{\sqrt{n\alpha}}; \widehat{EEA}\left(\hat{x}_m\right) = \frac{\hat{x}_m}{\sqrt{n\hat{\alpha}}} \tag{29}$$

$$P\left(\hat{\theta} - z_{1-\frac{\alpha}{n}}\sigma_{\hat{\theta}} \leq \theta \leq \hat{\theta} + z_{1-\frac{\alpha}{n}}\sigma_{\hat{\theta}}\right) = 1-\alpha$$
$$P\left(\hat{x}_m - z_{1-\frac{\alpha}{n}}\frac{\hat{x}_m}{\sqrt{n\hat{\alpha}}} \leq x_m \leq \hat{x}_m + z_{1-\frac{\alpha}{n}}\frac{\hat{x}_m}{\sqrt{n\hat{\alpha}}}\right) = 1-\alpha \tag{30}$$

$$ase\left(\hat{\theta}\right) = \sigma_{\hat{\theta}} = \sqrt{1/I\left(\theta\right)}$$
$$ase\left(\hat{\alpha}\right) = \sigma_{\hat{\alpha}} = \sqrt{\frac{1}{n/\alpha^2}} = \sqrt{\frac{\alpha^2}{n}} = \frac{\alpha}{\sqrt{n}}; \widehat{EEA}\left(\hat{\alpha}\right) = \frac{\hat{\alpha}}{\sqrt{n}} \tag{31}$$

$$P\left(\hat{\theta}-z_{1-\frac{\alpha}{n}}\sigma_{\hat{\theta}}\le\theta\le\hat{\theta}+z_{1-\frac{\alpha}{n}}\sigma_{\hat{\theta}}\right)=1-\alpha$$
$$P\left(\hat{\alpha}-z_{1-\frac{\alpha}{n}}\frac{\hat{\alpha}}{\sqrt{n}}\le\alpha\le\hat{\alpha}+z_{1-\frac{\alpha}{n}}\frac{\hat{\alpha}}{\sqrt{n}}\right)=1-\alpha \quad (32)$$

The estimator of α by the maximum likelihood method is more efficient than that of the method of moments, but it is biased. Given this limitation, a bias-corrected estimator can be defined that has a lower variance than the biased estimator (Rytgaard, 1990) and can be used to achieve more efficient asymptotic estimation (Equation 33). As with the previous definitions, the use of these asymptotic formulas requires a large sample, specifically one larger than 30 and preferably at least 100 (Mateus & Caeiro, 2022). See the variance, standard error, and asymptotic confidence interval of the bias-corrected estimator in Equations 34, 35, and 36, respectively.

$$\hat{\alpha}_c=\frac{n-2}{\sum_{i=1}^{n}\ln\left(\frac{x_i}{\min\left(\{x_i\}_{i=1}^{n}\right)}\right)} \quad (33)$$

$$\hat{\sigma}^2\left(\hat{\alpha}_c\right)=\hat{\alpha}_c^2/\left(n-2\right)<\hat{\sigma}^2\left(\hat{\alpha}\right) \quad (34)$$

$$\hat{\sigma}\left(\hat{\alpha}_c\right)=\sqrt{\hat{\alpha}^2/\left(n-2\right)}=\hat{\alpha}_c/\sqrt{n-2} \quad (35)$$

$$P\left(\hat{\alpha}_c-z_{1-\frac{\alpha}{n}}\frac{\hat{\alpha}_c}{\sqrt{n-2}}\le\alpha\le\hat{\alpha}_c+z_{1-\frac{\alpha}{n}}\frac{\hat{\alpha}_c}{\sqrt{n-2}}\right)=1-\alpha \quad (36)$$

3.3. Exact Distributions of the Parameters

Let X be a random variable following a type I Pareto distribution with scale parameter x_m and shape parameter α. The exact distribution of the sum of the logarithms of each of the n sample data points of X, divided by the minimum sample value, follows an exponential distribution with a rate parameter equal to α (Rytgaard, 1990). See Equation 37.

$$\sum_{i=1}^{n}\ln\left(\frac{x_i}{\min\left(\{x_i\}\right)_{i=1}^{n}}\right)=\sum_{i=1}^{n}\ln\left(\frac{x_i}{\hat{x}_m}\right)\sim\text{Exponential}\left(\lambda=\alpha\right) \quad (37)$$

The exact distribution of the maximum likelihood estimator of the scale parameter x_m is a type I Pareto distribution with a scale parameter x_m and a shape parameter $n\times\alpha$, while the exact distribution of the maximum likelihood estimator of the shape parameter α is an inverse gamma distribution with a shape parameter n - 1 and a scale parameter $n\times\alpha$ (Qian, Chen, & He, 2021). See Equation 38.

$$\hat{x}_m=\min\left(\{x_i\}_{i=1}^{n}\right)\sim\text{Pareto}\left(x_m,n\times\alpha\right)$$

$$\hat{\alpha} = \frac{n}{\sum_{i=1}^{n} \ln\left(\frac{x_i}{\min\left(\{x_i\}_{i=1}^{n}\right)}\right)} \sim \text{Inv-Gamma}\left(\alpha = n-1, \beta = n \times \alpha\right) \tag{38}$$

4. Generalized Form of the Pareto Distribution

There is a generalized form of the four-parameter Pareto distribution with location parameter μ, scale parameter σ, and two shape parameters: the Pareto tail index α and the inequality index γ. Its cumulative distribution function is provided in Equation 39 (Arnold, 2015).

Support: $x \geq \mu$

Parameter space: $\mu \in (-\infty, \infty)$ and σ, α and $\gamma \in (0, \infty)$.

$$F(X) = 1 - \frac{1}{\left[1 + \sqrt[\gamma]{\frac{x-\mu}{\sigma}}\right]^{\alpha}} \tag{39}$$

The location parameter μ is not the mathematical expectation or arithmetic mean of the distribution, and the scale parameter σ is not the standard deviation. The mean depends on the parameters σ, α, and γ, as well as certain gamma functions, and requires that γ < α (Equation 40).

$$E(X) = \frac{\sigma \times \Gamma(\alpha - \gamma) \times \Gamma(1+\gamma)}{\Gamma(\alpha)}, \gamma < \alpha \tag{40}$$

When μ = σ = x_m (minimum and mode) and γ = 1, this generalized four-parameter form reduces to the type I Pareto distribution (Equation 41).

$$\begin{aligned} F(X) &= 1 - \frac{1}{\left[1 + \left(\frac{x-\mu}{\sigma}\right)^{1/\gamma}\right]^{\alpha}} = 1 - \frac{1}{\left[1 + \left(\frac{x-x_m}{x_m}\right)^{1/1}\right]^{\alpha}} \\ &= 1 - \frac{1}{\left(1 + \frac{x}{x_m} - 1\right)^{\alpha}} = 1 - \left(x_m / x\right)^{\alpha} \end{aligned} \tag{41}$$

In the type II Pareto distribution, γ = 1 (Equation 42). When the parameter μ of the type II Pareto distribution is equal to 0, it is referred to as the Lomax (1954) distribution.

$$F(X) = 1 - \frac{1}{\left[1 + \left(\frac{x-\mu}{\sigma}\right)^{1/\gamma}\right]^{\alpha}} = 1 - \frac{1}{\left[1 + \left(\frac{x-\mu}{\sigma}\right)^{1/1}\right]^{\alpha}} = 1 - \frac{1}{\left[1 + \frac{x-\mu}{\sigma}\right]^{\alpha}} \tag{42}$$

In the type III Pareto distribution, α = 1 (Equation 43).

$$F(X) = 1 - \frac{1}{\left[1 + \left(\frac{x-\mu}{\sigma}\right)^{1/\gamma}\right]^{\alpha}} = 1 - \frac{1}{\left[1 + \left(\frac{x-\mu}{\sigma}\right)^{1/\gamma}\right]^{1}}$$

$$= 1 - \frac{1}{1 + \left(\frac{x-\mu}{\sigma}\right)^{1/\gamma}} = 1 - \frac{1}{1 + \sqrt[\gamma]{\frac{x-\mu}{\sigma}}} \tag{43}$$

5. Relationship between Type I Pareto Distribution and Other Distributions

Let X be a random variable following a type I Pareto distribution with parameters x_m and α. The product of the scale parameter x_m and an exponential distribution based on the number e and exponent X (random variable Y) follows an exponential distribution with a rate parameter or inverse scale $\lambda = \alpha$. Conversely, let Y be a random variable with an exponential distribution with rate parameter or inverse scale λ. The natural logarithm of the quotient of the random variable Y and the value x_m (random variable X) follows a type I Pareto distribution with a scale parameter x_m and a shape parameter $\alpha = \lambda$. This implies that the cumulative distribution functions of X and Y are equal and interchangeable when calculating cumulative probabilities. Refer to Equation 44.

$$\begin{aligned} &X \sim \text{Pareto}\left(x_m, \alpha\right); x_m \mathrm{e}^X = Y \sim \text{Exponential}\left(\lambda = \alpha\right) \\ &Y \sim \text{Exp}\left(\lambda\right); \ln\left(Y/x_m\right) = X \sim \text{Pareto}\left(x_m, \alpha = \lambda\right) \\ &P\left(X \le x\right) = F_X\left(x \mid x_m, \alpha\right) = F_Y\left[y = \ln\left(\frac{x}{x_m}\right) \mid \lambda = \alpha\right] = 1 - \left(\frac{x_m}{x}\right)^{\alpha} = 1 - \mathrm{e}^{-\alpha y} \end{aligned} \tag{44}$$

It should be noted that there is a relationship between the exponential and Pareto distributions that is analogous to the one between the normal and lognormal distributions. The Pareto and lognormal distributions can be applied to the same data, although one may provide a better fit than the other. Both are exponential transformations of a more widely known and used variable: the former from an exponential distribution and the latter from a normal distribution (Feng et al., 2020).

The zeta distribution is the discrete equivalent of the Pareto distribution. If the support of the zeta distribution is bounded within an interval of natural numbers, it becomes the so-called Zipf distribution (Arnold, 2015).

6. Relationship between Type I Pareto Distribution and Gini Concentration Index

The Pareto (1897) tail index is related to the Gini concentration index (Safari, Masseran, Ibrahim, & Hussain, 2019). Corrado Gini (1936), created this index to summarize in a single number the information contained in the Lorenz curve (Mojiri & Ahmadi, 2022). The Lorenz (1905) curve is obtained from a two-dimensional diagram that represents the number of people on the horizontal axis (x) and the accumulated income or wealth on the vertical axis, $F_X(x)$. A straight line with a 45-degree slope, known as the line of equality, is included in this graph. As the curve, defined by the coordinate pairs (x, $F_X(x)$), diverges from this line, inequality increases. The Gini index is the ratio of the area between the line of equality and the Lorenz curve

to the area of the lower right triangle. A value of 0 indicates total equality, which occurs when all individuals have the same income. A value of 1 reflects maximum inequality, occurring when one person receives all the income and the others have no income or are unpaid labor (Sitthiyot & Holasut, 2021).

In countries with more income equality, such as the Scandinavian countries, the Gini index is less than 0.3. In countries with greater inequality, such as some African nations (South Africa, Namibia, Suriname, Zambia, Eswatini, Botswana, Angola, and Zimbabwe) and countries in the Americas (Belize, Brazil, Colombia, and Panama), the Gini index is higher than 0.5 (World Bank, 2022). It should be noted that countries with communist economic systems do not have low Gini indices, and social democratic policies involving increased public spending and debt tend to increase rather than decrease inequality (Tokhirov, 2021). Thus, left-wing populist policies have been described as policies of poverty equalization, leading to the emergence of a wealthy political oligarchy (Benczes, 2022; Landoni & Villegas, 2022).

The value on the Lorenz curve at point x for a continuous random variable X with density function $f_X(x)$ is obtained by the following proportion given in Equation 45.

$$L\left(F\left(x\right)\right)=\frac{\int_{-\infty}^{x}xf\left(x\right)\mathrm{d}x}{\int_{-\infty}^{\infty}xf\left(x\right)\mathrm{d}x}=\frac{\int_{-\infty}^{x}xf\left(x\right)\mathrm{d}x}{E\left(X\right)} \tag{45}$$

When applied to the type I Pareto distribution, the above formula is transformed into the expression shown in Equation 46.

$$\begin{aligned}&X\sim\text{Pareto}\left(x_n,\alpha\right)\text{ and }F\left(x\right)=1-\left(x_m/x\right)^{\alpha}\\&L\left(F\left(x\right)\right)=\begin{cases}0 & 0<\alpha<1\\1-\left(1-F\left(x\right)\right)^{1-\frac{1}{\alpha}}=1-\left(\dfrac{x_m}{x}\right)^{\alpha-1} & \alpha\geq 1\end{cases}\end{aligned} \tag{46}$$

The Gini index can be expressed as one minus twice the integral from 0 to 1 of the Lorenz curve. When applied to the type I Pareto distribution, it is the inverse of twice the tail parameter minus one, $1/(2\alpha-1)$, whenever $\alpha > 1$. When $\alpha < 1$, the Gini index takes its minimum value, which is 1 (Equation 47). **Table 1** shows the correspondence between values of the Pareto tail and Gini concentration indices and their interpretation.

$$\begin{aligned}&G\left(X\right)=\begin{cases}1 & 0<\alpha<1\\1-2\int_0^1 L\left(F\left(x\right)\right)\mathrm{d}x=\dfrac{1}{2\alpha-1} & \alpha\geq 1\end{cases}\\&\alpha=0.5\left(1+1/G\left(X\right)\right)\end{aligned} \tag{47}$$

Table 1. Correspondence between Gini and Pareto indexes and their interpretation.

G	α	Interpretation
0.05	10.5	Very little inequality
0.1	5.5	
0.15	3.833	Little inequality

Continued

0.2	3	
0.25	2.5	
0.3	2.167	
0.35	1.929	Moderate inequality
0.4	1.75	
0.45	1.61	Fairly significant inequality
0.5	1.5	
0.55	1.409	A lot of inequality
0.6	1.333	
0.65	1.269	
0.7	1.214	
0.75	1.167	Very much inequality
0.8	1.125	
0.85	1.088	
0.9	1.056	
0.95	1.026	
1	1	

Note. G = Gini concentration index and α = Pareto tail index.

7. Generation of Type I Pareto Random Samples and Goodness-of-Fit Testing

Let U be a random variable with a standard uniform distribution: $U \subseteq$ U [0, 1]. The transformation $x_m(1 - U)^{-1/\alpha}$ follows a type I Pareto distribution with scale parameter x_m and shape parameter α. This procedure allows for obtaining a random sample with a type I Pareto distribution from a random sample of a standard uniform variable and is called inverse transform sampling (Ross, 2022).

To test whether the sample data fit the type I Pareto distribution, Chu, Dickin and Nadarajah (2019) recommend the Kolmogorov-Smirnov test based on a simulation study. The simplest way to apply this inferential test is to transform the data, assuming a type I Pareto distribution, into values with an exponential distribution by taking the natural logarithm of the ratio of each data point to the minimum sample value. Then, the Kolmogorov-Smirnov test, adapted for exponentially distributed samples, is applied (Stephens, 1974). The inferential test can be supplemented with a graphical assessment using a quantile-quantile plot and a histogram with an overlaid density curve. In the former, an alignment of points along a straight line with a 45-degree slope is sought, and in the latter, an inverted J-shaped profile is expected (Bhoj & Chandra, 2021).

Statistical hypotheses: H_0: $X \sim$ Pareto(x_m, α) ≡ $Y = \ln[X/\min(X)] \sim$ Exponential(λ = α) y H_1: $X \nsim$ Pareto(x_m, α).

Assumptions: Random sample of size n of a continuous quantitative variable X.

Test statistic: The data are sorted in ascending order, assigned ranks or orders, and transformed to follow an exponential distribution under the assumption that the null hypothesis is true.

$x_{(1)} \leq$	$x_{(2)} \leq$	$\ldots \leq$	$x_{(i)} \leq$	$\ldots \leq$	$x_{(n-1)} \leq$	$x_{(n)}$
1	2	…	i	…	n-1	n
$\ln[x_{(1)}/x_{(1)}]$	$\ln[x_{(2)}/x_{(1)}]$	…	$\ln[x_{(i)}/x_{(1)}]$	…	$\ln[x_{(n-1)}/x_{(1)}]$	$\ln[x_{(n)}/x_{(1)}]$

The cumulative probability (theorical cumulative relative frequency) under the type I Pareto distribution model (Equation 48) and the (empirical) cumulative relative frequency (Equation 49) are calculated for each transformed value, along with the D^+, D^-, and D statistics (Equation 50).

$$y_i = \ln\left[x_i / x_{(1)}\right] \in Y \sim \text{Exponetial}\left(\hat{\lambda} = \frac{n}{\sum_{i=1}^{n} \ln\left(x_i\right)}\right) \tag{48}$$

$$F_Y\left(y_{(i)}\right) = 1 - \mathrm{e}^{-\hat{\lambda} y_{(i)}}$$

$$F_n\left(x_{(i)}\right) = (i)/n \tag{49}$$

$$\begin{aligned} D^+ &= \max\left[F_n\left(x_{(i)}\right) - F_Y\left(y_{(i)}\right)\right] = \max\left[(i)/n - F_Y\left(y_{(i)}\right)\right] \\ D^- &= \max\left[F_Y\left(y_{(i)}\right) - F_n\left(x_{(i-1)}\right)\right] = \max\left[(i-1)/n - F_Y\left(y_{(i)}\right)\right] \\ D &= \max\left(D^+, D^-\right) \end{aligned} \tag{50}$$

The sample estimation correction to the D statistic, as given by Stephens (1974), is applied. See Equation 51.

$$D_c = \left(D - 0.2/n\right)\left(\sqrt{n} + 0.26 + 0.5/\sqrt{n}\right) \tag{51}$$

The decision is based on the transformed D statistic (Equation 48). If $D_c \leq D_\alpha$, H_0 holds, and if $D_c > D_\alpha$, H_0 is rejected at a significance level of α. The critical D_α values depend on the significance level α (Stephens, 1974), which is typically set at .05 ($D_{0.05}$ = 1.094). For small samples (n < 20), it may be increased to .10 ($D_{.10}$ = .990), and for large samples (n > 500), it may be reduced to .01 ($D_{.01}$ = 1.308).

Another option for inferential testing is the Anderson & Darling (1952) test. The formula for its test statistic is shown in Equation 52. Similar to the Kolmogorov-Smirnov test, the data are transformed to follow an exponential distribution (Equation 48), and Stephens' (1986) correction is applied to the test statistic when the parameters are estimated from the sample data (Equation 53).

$$A^2 = -n - \sum_{i=1}^{n} \frac{2i-1}{n}\left[\ln\left(F_Y\left(y_i\right)\right) + \ln\left(1 - F_Y\left(y_{n+1-i}\right)\right)\right] \tag{52}$$

$$A_c^2 = A\left(1 + \frac{0.6}{n}\right) \tag{53}$$

As in the Kolmogorov-Smirnov test, the decision is based on the corrected A^2 statistic, denoted as A_c^2. If $A_c^2 \leq A_\alpha^2$ (critical value), H_0 is retained; if $A_c^2 > A_\alpha^2$, H_0 is rejected at a significance level of α. The critical values A_α^2 depend on the type of distribution (in this case, exponential), the estimation method (in this case, maximum likelihood estimator), and the significance level α, which is typically set at .05 ($A_{.05}$ = 1.321). For small samples (n < 20), the significance level α can be

increased to .10 ($A_{.1}$ = 1.062), while for large samples (n > 500), it can be reduced to .01 ($A_{.01}$ = 1.959). Refer to Stephens (1986).

A modification of the Anderson-Darling test was developed by Sinclair, Spurr and Ahmad (1990) for distributions with positive skewness and highly atypical cases in the right tail. The type I Pareto distribution falls into this category; hence, this modification can be applied. The Sinclair-Spurr-Ahmad statistic is calculated directly from the original sample data, without transformation (Equation 54), and is denoted as AU_n^2.

$$\begin{aligned} AU_n^2 &= \frac{n}{2} - 2\sum_{i=1}^{n} F_X\left(x_{(i)}\right) - \sum_{i=1}^{n}\left(2 - \frac{2i-1}{n}\right)\ln\left(1 - F_X\left(x_{(i)}\right)\right) \\ F_X\left(x_{(i)}\right) &= 1 - \left(x_i / \hat{x}_m\right)^{\alpha} = 1 - \left(x_i / \min(x)\right)^{\alpha} \end{aligned} \tag{54}$$

The critical value for decision making is calculated with the formula shown in Equation 55.

$$\begin{aligned} {}_pAU_n^2 &= 1 - \frac{1}{1 + e^{G(p)}} \\ G(p) &= 0.1170 - 0.03791 \times t + 0.06318 \times u + 0.09878 \times t \times u \\ &\quad + 0.009184 \times t^2 \times u - 0.00009742 \times t^4 \times u \\ t &= \ln\frac{p}{1-p}, u = \frac{1}{1 + 0.3/\sqrt{n}} \end{aligned} \tag{55}$$

If $AU_n^2 \le {}_pAU_n^2$, H_0 holds: $X \sim$ Pareto (x_m, α) at a significance level of p. If $AU_n^2 > {}_pAU_n^2$, H_0 is rejected. As the sample size approaches infinity, the value of u becomes unitary, and the critical value ${}_pAU_n^2$ reaches its asymptotic value: ${}_pAU_n^2$ = 0.356 for p = .1, ${}_pAU_n^2$ = 0.432 for p = .05, and ${}_pAU_n^2$ = 0.610 for p = .01.

8. Example Calculations Using the Type I Pareto Distribution

Next, simulated data are used in two examples illustrating the application of the distribution. In the first example, probabilities, descriptive statistics, and the representation of the distribution are calculated for a variable X that follows a type I Pareto distribution, characterized by the parameters x_m = 500 and α = 4.1, representing the biweekly salary (in dollars) within a company. In the second example, the data are generated using the inverse sampling transform in Excel to follow a type I Pareto distribution (x_m =500, α = 4.1). Based on these generated data, the distribution parameters are estimated both pointwise and using confidence intervals. Additionally, the goodness of fit is assessed using statistical tests and graphical methods.

8.1 Calculation of Probabilities and Descriptive Measures

Let X be a variable with a type I Pareto distribution characterized by parameters x_m = 500 and α = 4.1, representing the biweekly salary in dollars within a company. The task is to calculate the probability of having an income less than or equal to

592, less than 1000, between 800 and 1400, greater than or equal to 600, and greater than 1500. Additionally, the following measures of central tendency are required: the mathematical expectation or arithmetic mean $\mu(X)$, geometric mean $\mu_g(X)$, harmonic mean $\mu_h(X)$, median $Mdn(X)$, and mode $Mo(X)$; measures of variation such as variance $\sigma^2(X)$, standard deviation $\sigma(X)$, and entropy $H(X)$; as well as shape measures based on standardized central moments, including skewness $\sqrt{\beta_1(X)}$ and excess kurtosis $\beta_2(X)$. Furthermore, the Gini concentration index $G(X)$ is to be calculated. Finally, it is considered illustrative to plot its density function $f_X(x)$ and cumulative distribution $F_X(x)$.

$$X \sim \text{Pareto}\left(x_n = 500, \alpha = 4.1\right)$$

Equation 2 is applied for the calculation of the probability of having an income less than or equal to 592.

$$P\left(X \le 592\right) = \int_{500}^{592} f_X\left(x\right)\mathrm{d}x = 1 - \left(x_m / x\right)^{\alpha} = 1 - \left(500/592\right)^{4.1} = 0.4997$$

Equation 2 is also used to calculate the probability of having an income less than 1000. Since it is a continuous distribution, including or excluding the specific point does not affect the value of the cumulative probability.

$$P\left(X < 1000\right) = \int_{500}^{1000} f_X\left(x\right)\mathrm{d}x = 1 - \left(500/1000\right)^{4.1} = 0.9417$$

Equation 2 is used to calculate the probability of having an income between 800 and 1400 by taking the difference between two cumulative probabilities.

$$P\left(800 \le X \le 1400\right) = P\left(X \le 1400\right) - P\left(X \le 800\right) = 0.9853 - 0.8544 = 1309$$

$$P\left(X \le 1400\right) = 1 - \left(500/1400\right)^{4.1} = 0.9853$$

$$P\left(X \le 800\right) = 1 - \left(500/800\right)^{4.1} = 0.8544$$

Equation 3 is used to calculate the probability of having an income greater than or equal to 600.

$$P\left(X \ge 600\right) = \left(500/600\right)^{4.1} = 0.4735$$

Equation 3 is also used to calculate the probability of having an income greater than 1500. Since it is a continuous distribution, including or excluding the specific point does not affect the complementary or right-tailed probability.

$$P\left(X > 1500\right) = \left(500/1500\right)^{4.1} = 0.0111$$

Measures of central tendency. The arithmetic mean is calculated using Equation 7, the geometric mean using Equation 8, the harmonic mean using Equation 9, the median using Equation 10, and the mode using Equation 11.

$$\mu\left(X\right) = E\left(X\right) = \frac{\alpha}{\alpha - 1} x_m = \frac{4.1}{3.1} \times 500 = 661.29$$

$$\mu_g\left(X\right) = \mathrm{e}^{E\left(\ln\left(X\right)\right)} = x_m \mathrm{e}^{1/\alpha} = 500 \times \mathrm{e}^{1/4.1} = 638.11$$

$$\mu_h\left(X\right) = H\left(X\right) = \frac{1}{E\left(1/X\right)} = \left(1 + \frac{1}{\alpha}\right) x_m = \left(1 + \frac{1}{4.1}\right) \times 500 = 621.95$$

$$Mdn(X) = Q_X(p = 0.5) = \sqrt[\alpha]{2} \times x_m = 2^{1/4.1} \times 500 = 592.10$$

$$Mo(X) = \{x \mid \max(f_X(x))\} = x_m = 500$$

Measures of variation. Variance is calculated using Equation 13, standard deviation using Equation 14, and entropy using Equation 15.

$$\sigma^2(X) = \mu_2(X) = E\left[\left(X - E(X)\right)^2\right] = x_m^2 \frac{\alpha}{(\alpha-1)^2(\alpha-2)}$$

$$= 500^2 \frac{4.1}{3.1^2 \times 2.1} = 50790.35$$

$$\sigma(X) = \mu_2^{1/2}(X) = \sqrt{E\left[\left(X - E(X)\right)^2\right]} = \frac{x_m}{\alpha-1}\sqrt{\frac{\alpha}{\alpha-2}} = \frac{500}{3.1}\sqrt{\frac{4.1}{2.1}} = 225.37$$

$$H(X) = E\left[-\ln(f_X(x))\right] = \ln\left[\frac{x_m}{\alpha} e^{1+\frac{1}{\alpha}}\right] = \ln\left[\frac{500}{4.1} e^{1+\frac{1}{4.1}}\right] = 6.05 \text{ nats}$$

Skewness and Excess Kurtosis Based on Central Moments. These two shape statistics are calculated using Equations 16 and 17, respectively.

$$\sqrt{\beta_1(X)} = \frac{\mu_3(X)}{\mu_2^{3/2}(X)} = \frac{2(\alpha+1)}{\alpha-3}\sqrt{\frac{\alpha-2}{\alpha}} = \frac{2 \times 5.1}{1.1}\sqrt{\frac{2.1}{4.1}} = 6.64$$

$$\beta_2(X) - 3 = \frac{\mu_4(X)}{\mu_2^2(X)} - 3 = \frac{6\left(\alpha^3 + \alpha^2 - 6\alpha - 2\right)}{\alpha(\alpha-3)(\alpha-4)} = \frac{6\left(4.1^3 + 4.1^2 - 6 \times 4.1 - 2\right)}{4.1 \times 1.1 \times 0.1} = 786.67$$

The Gini Index is obtained using Equation 47.

$$G(X) = 1/(2\alpha - 1) = 1/(2 \times 4.1 - 1) = 0.14$$

The Gini index is close to 0, indicating fairly equal wages within the firm. Figure 2 displays the density function $f_X(x)$ and the cumulative distribution $F_X(x)$ of the random variable X, representing biweekly wages (in dollars). The wages follow a type I Pareto distribution with parameters $x_m = 500$ and $\alpha = 4.1$.

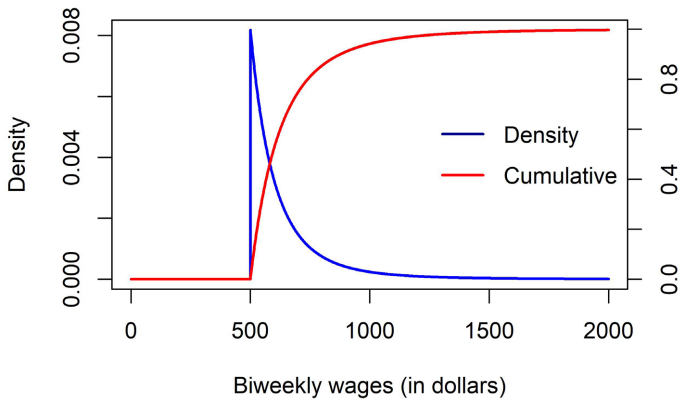

Figure 2. Density function $f_X(x)$ and cumulative distribution function $F_X(x)$ for $X \sim$ Pareto($x_m = 500$, $\alpha = 4.1$).

The R script to calculate these probabilities, descriptive statistics, and the plot

of the density and cumulative distribution functions of the type I Pareto distribution (x_m = 500 and α = 4.1) is shown below:

```
# Load required libraries (packages)
library(cascsim)
library(modeest)
cat ("Probability calculations (rounding to four decimal places)", "\n")
cat ("Cumulative probability less than or equal to 592 =", round(ppareto(592, xm = 500, alpha = 4.1), 4), "\n")
cat ("Cumulative probability less than 1000 =", round(ppareto(1000, xm = 500, alpha = 4.1), 4), "\n")
cat ("Probability in the interval [800, 1400] =", round(ppareto(1400, xm = 500, alpha = 4.1) - ppareto(800, xm = 500, alpha = 4.1), 4), "\n")
cat ("Probability greater than or equal to 600 =", round(1 - ppareto(600, xm = 500, alpha = 4.1), 4), "\n")
cat ("Probability greater than 1500 =", round(1 - ppareto(1500, xm = 500, alpha = 4.1), 4), "\n")
# Descriptive measures for type I Pareto distribution (rounding to three decimal places).
x_m <- 500 # Scale parameter
alpha <- 4.1 # Shape parameter
cat ("Measures of central tendency", "\n")
cat ("Mathematical expectation of X: μ(X) =", round(alpha / (alpha - 1) * x_m, 3), "\n")
cat ("Geometric mean of X: μg(X) =", round(x_m * exp(1 / alpha), 3), "\n")
cat ("Harmonic mean of X: μh(X) =", round((1 + 1 / alpha) * x_m, 3), "\n")
cat ("Median de X: Mdn(X) =", round(qpareto(0.5, xm = 500, alpha = 4.1), 3), "\n")
cat ("Mode of X: Mdn(X) =", round(qpareto(0.5, xm = 500, alpha = 4.1), 3), "\n")
cat ("Measures of variation ", "\n")
cat ("Variance of X: σ^2(X) =", round(x_m^2 * alpha /(( alpha -1)^2 * (alpha - 2)), 3), "\n")
cat ("Measures of variation ", "\n")
cat ("Standard deviation of X: σ(X) =", round(sqrt(x_m^2 * alpha /(( alpha -1)^2 * (alpha - 2))), 3), "\n")
cat ("Entropy of X: H(X) =", round(log(x_m / alpha * exp(1+ 1/ alpha)), 3), "nats", "\n")
cat ("Shape measures", "\n")
cat("Skewness of X:  √b1(X) =", round(2 * (alpha + 1) / (alpha - 3) * sqrt((alpha - 2) / alpha), 3), "\n")
cat ("Kurtosis of X: b2(X) =", round(6* (alpha^3 + alpha^2 - 6 * alpha - 2) / (alpha * (alpha - 3) * (alpha - 4)), 3), "\n")
cat ("Gini Index of X: G(X) =", round(1 / (2 * alpha - 1), 3), "\n")
# Plot of density and cumulative distribution functions.
x <- seq(0, 2000, length.out = 2000)
```

```
dpareto <- function (x, x_m, alpha) {
ifelse (x < x_m, 0, alpha * x_m^alpha / (x^(alpha + 1)))}
ppareto2 <- function (x, x_m, alpha) {ifelse(x < x_m, 0, 1 - (x_m / x)^alpha)}
density <- dpareto(x, x_m, alpha)
cumulative <- ppareto(x, x_m, alpha)
par (mar = c(4, 4, 1, 2) + 0.1)
plot (x, density, type = "l", col = "blue", lwd = 2, xlab = "Biweekly wages (in dollars)",
ylab = "Density", ylim = c(0, max(density)))
par (new = TRUE)
plot (x, cumulative, type = "l", col = "red", lwd = 2, axes = FALSE, xlab = "", ylab = "")
axis (4)
mtext ("Cumulative Probability", side = 4, line = 3, col = "red")
legend ("right", legend = c("Density", "Cumulative"),
col = c ("darkblue", "red"), lwd = 2, bty = "n", y.intersp = 1.5)
```

8.2. Random Sample Generation, Parameter Estimation, and Goodness of Fit

In this second example, the aim is to generate a random sample of 40 observations from the variable $X \sim$ Pareto (500, 4.1). Based on this sample, the goal is to estimate the parameters x_m and α using maximum likelihood estimators, and to construct 95% confidence intervals for these parameters based on their asymptotic distributions and errors. Finally, the goodness-of-fit will be assessed through a plot of theoretical versus empirical quantiles, a histogram with an overlaid density curve, and the Kolmogorov-Smirnov and Anderson-Darling tests.

In the first column of Table 2, a random sample of 40 u_i values is drawn from a standard uniform distribution $U \sim$ U [0, 1], generated by means of the 2021 version of the Excel random number generator. Using the quantile function, the u_i values are transformed into a random sample of size 40 for a variable X that follows a type I Pareto distribution with population parameters: x_m = 500 (scale) and α = 4.1 (shape). This data generation process is known as inverse transform sampling, as illustrated in Equation 56. The first value of X obtained by applying Equation 56, starting at u_1 = 0.226, is 532.239.

$$\begin{aligned} x_i &= x_m \left(1-u_i\right)^{-\frac{1}{\alpha}} = 500\left(1-u_i\right)^{-\frac{1}{4.1}} \\ x_1 &= 500\left(1-0.226\right)^{-\frac{1}{4.1}} = 532.239 \end{aligned} \tag{56}$$

Table 2. Random sample generation and fit testing to the generating distribution.

Generation of X			Kolmogorov-Smirnov test					Q-Q plot	
u_i	x_i	(i)	$x_{(i)}$	$y_{(i)}$	$F_Y(y_{(i)})$	$D+$	$D-$	$p_{(i)}$	$x_{t(i)}$
0.226	532.239	1	502.595	0	0	0.025	0	0.017	504.716
0.678	659.182	2	506.685	0.008	0.032	0.018	0.007	0.041	507.983

Continued

0.754	703.919	3	509.868	0.014	0.055	0.020	0.005	0.066	511.358
0.791	732.467	4	513.713	0.022	0.083	0.017	0.008	0.091	514.847
0.427	572.739	5	516.552	0.027	0.103	0.022	0.003	0.116	518.457
0.989	1502.032	6	519.474	0.033	0.123	0.027	−0.002	0.140	522.196
0.956	1071.113	7	524.641	0.043	0.156	0.019	0.006	0.165	526.073
0.356	556.651	8	529.753	0.053	0.188	0.012	0.013	0.190	530.097
0.651	646.362	9	532.239	0.057	0.203	0.022	0.003	0.215	534.278
0.592	622.201	10	540.797	0.073	0.252	−0.002	0.027	0.240	538.628
0.275	540.797	11	540.979	0.074	0.253	0.022	0.003	0.264	543.159
0.885	847.360	12	550.305	0.091	0.302	−0.002	0.027	0.289	547.886
0.921	928.627	13	551.504	0.093	0.308	0.017	0.008	0.314	552.824
0.524	599.242	14	556.651	0.102	0.333	0.017	0.008	0.339	557.991
0.145	519.474	15	560.951	0.110	0.353	0.022	0.003	0.364	563.406
0.726	685.652	16	569.140	0.124	0.389	0.011	0.014	0.388	569.093
0.053	506.685	17	571.286	0.128	0.398	0.027	−0.002	0.413	575.075
0.325	550.305	18	572.739	0.131	0.404	0.046	−0.021	0.438	581.384
0.421	571.286	19	588.961	0.159	0.466	0.009	0.016	0.463	588.050
0.578	617.102	20	598.325	0.174	0.498	0.002	0.023	0.488	595.114
0.489	588.961	21	599.242	0.176	0.501	0.024	0.001	0.512	602.620
0.021	502.595	22	609.166	0.192	0.533	0.017	0.008	0.537	610.619
0.924	937.437	23	617.102	0.205	0.556	0.019	0.006	0.562	619.175
0.331	551.504	24	622.201	0.213	0.570	0.030	−0.005	0.587	628.359
0.622	633.900	25	633.900	0.232	0.601	0.024	0.001	0.612	638.261
0.723	683.834	26	646.362	0.252	0.630	0.020	0.005	0.636	648.989
0.211	529.753	27	659.182	0.271	0.658	0.017	0.008	0.661	660.674
0.412	569.140	28	683.834	0.308	0.704	−0.004	0.029	0.686	673.482
0.077	509.868	29	685.652	0.311	0.707	0.018	0.007	0.711	687.625
0.276	540.979	30	703.919	0.337	0.736	0.014	0.011	0.736	703.373
0.521	598.325	31	721.765	0.362	0.761	0.014	0.011	0.760	721.090
0.822	761.718	32	732.467	0.377	0.775	0.025	0.000	0.785	741.265
0.125	516.552	33	761.718	0.416	0.807	0.018	0.007	0.810	764.590
0.971	1185.755	34	804.876	0.471	0.845	0.005	0.020	0.835	792.076
0.858	804.876	35	847.360	0.522	0.873	0.002	0.023	0.860	825.283
0.105	513.713	36	928.627	0.614	0.912	−0.012	0.037	0.884	866.783
0.376	560.951	37	937.437	0.623	0.915	0.010	0.015	0.909	921.249
0.555	609.166	38	1071.113	0.757	0.950	0.000	0.025	0.934	998.449
0.179	524.641	39	1185.755	0.858	0.967	0.008	0.017	0.959	1124.364
0.778	721.765	40	1502.032	1.095	0.987	0.013	0.012	0.983	1417.313
Σ				10.108					
max						0.046	0.037		

Note. Generation of X: u_i = sample data in its random order (i = 1, 2, …, 40) drawn from a standard uniform distribution U [0, 1] and x_i = 500(1 - u_i)-1/4.1 = data transformed to follow a type I Pareto distribution (with scale parameter x_m = 500 and shape parameter α = 4.1) via inverse transform sampling. Logarithmic transformation of X and Kolmogorov- Smirnov test: (i) = the order of data xi within the 40-item sample, $x_{(i)}$ = data x_i sorted in ascending order (empirical quantiles), $y_{(i)}$ = ln($x_{(i)}$/502.595) = the logarithmic transformation of x_i normalized by the sample minimum, $F_Y(y_{(i)})$ = the cumulative distribution function of the Y variable following an exponential distribution with rate parameter λ = 3.957, D^+ = (i)/40 - $F_Y(y_{(i)})$ = the difference between the empirical and theoretical cumulative distribution functions and D^- = $F_Y(y_{(i)})$ - ((i)-1)/40 = the difference between the theoretical distribution function and empirical cumulative distribution function with lap of 1. Quantile-Quantile plot: $p_{(i)}$ = ((i)-1/3)/(n+1/3) = the order of the theoretical quantile and $x_{t(i)}$ = $Q_X[p_{(i)}]$ = theoretical quantiles calculated using the quantile function of a type I Pareto distribution with estimated parameters. Σ = sum per column, *max* = maximum value per column.

Next, the two parameters are estimated pointwise using maximum likelihood estimation (Equation 21 for thse estimate of x_m and Equation 22 for the estimate of α), and their corresponding standard errors (Equation 29 for $\hat{x}_m$ and Equation 31 for $\hat{\alpha}$) and 95% asymptotic confidence intervals (Equation 30 for $\hat{x}_m$ and Equation 32 for $\hat{\alpha}$) are computed.

Estimation of the scale parameter (x_m):

$$\hat{x}_m = \min\left(\{x_i\}_{i=1}^{40}\right) = 502.595$$

$$ase\left(\hat{x}_m\right) = \hat{\sigma}\left(\hat{x}_m\right) = \frac{\hat{x}_m}{\sqrt{n\times\hat{\alpha}}} = \frac{502.595}{\sqrt{40\times 3.957}} = 39.948$$

$$P\left(\hat{x}_m - z_{1-\frac{\alpha}{n}}\frac{\hat{x}_m}{\sqrt{n\times\hat{\alpha}}} \le x_m \le \hat{x}_m + z_{1-\frac{\alpha}{n}}\frac{\hat{x}_m}{\sqrt{n\times\hat{\alpha}}}\right) = 1-\alpha$$

$$P\left(502.595 - 1.96\times 39.948 \le x_m \le 502.595 + 1.96\times 39.948\right) = 0.95$$

$$P\left(x_m \in \left[424.299, 580.891\right]\right) = 0.95$$

Estimation of the shape parameter (α):

$$\hat{\alpha} = \frac{n}{\sum_{i=1}^{n}\ln\left(\frac{x_i}{\hat{x}_m}\right)} = \frac{n}{\sum_{i=1}^{n}\ln\left(\frac{x_i}{\min\left(\{x_i\}_{i=1}^{n}\right)}\right)}$$

$$= \frac{40}{\sum_{i=1}^{40}\ln\left(\frac{x_i}{502.595}\right)} = \frac{40}{10.108} = 3.957$$

$$ase\left(\hat{\alpha}\right) = \hat{\sigma}\left(\hat{\alpha}\right) = \hat{\alpha}/\sqrt{n} = 3.957/\sqrt{40} = 0.626$$

$$P\left(\hat{\alpha} - z_{1-\frac{\alpha}{n}}\frac{\hat{\alpha}}{\sqrt{n}} \le \alpha \le \hat{\alpha} + z_{1-\frac{\alpha}{n}}\frac{\hat{\alpha}}{\sqrt{n}}\right) = 1-\alpha$$

$$P\left(3.957 - 1.96\times 0.626 \le \alpha \le 3.957 + 1.96\times 0.626\right) = 0.95$$

$$P\left(\alpha \in \left[2.731, 5.184\right]\right) = 0.95$$

The bias-corrected formula is also used to estimate the shape parameter α. (Equation 33). Initially, the point estimate of α appears less accurate than the unbiased one; however, the confidence interval is more efficient and appropriate. The standard error is calculated using Equation 35, and the asymptotic confidence interval is determined using Equation 36.

$$\hat{\alpha}_c = \frac{n-2}{\sum_{i=1}^{n}\ln\left(\frac{x_i}{\hat{x}_m}\right)} = \frac{38}{\sum_{i=1}^{40}\ln\left(\frac{x_i}{502.595}\right)} = \frac{38}{10.108} = 3.759$$

$$ase\left(\hat{\alpha}_c\right) = \hat{\sigma}\left(\hat{\alpha}_c\right) = \hat{\alpha}_c/\sqrt{n-2} = 3.759/\sqrt{38} = 0.610$$

$$P\left(\hat{\alpha}_c - z_{1-\frac{\alpha}{n}}\frac{\hat{\alpha}_c}{\sqrt{n-2}} \le \alpha \le \hat{\alpha}_c + z_{1-\frac{\alpha}{n}}\frac{\hat{\alpha}_c}{\sqrt{n-2}}\right) = 1-\alpha$$

$$P\left(3.759-1.96\times 0.610\le \alpha \le 3.759+1.96\times 0.610\right)=0.95$$

$$P\left(\alpha \in \left[2.564,4.955\right]\right)=0.95$$

The following is the R script for point estimates and 95% asymptotic confidence intervals for the two parameters of the type I Pareto distribution.

```
# Vector of scores.
x <- c(532.2385845, 659.1817235, 703.9185616, 732.4668655, 572.7389011, 1502.031724, 1071.11279, 556.6512741, 646.3622704, 622.2013372, 540.796541, 847.3602594, 928.6272967, 599.2422898, 519.4737837, 685.6524203, 506.685297, 550.3047107, 571.2856056, 617.1023659, 588.9614346, 502.594959, 937.4374491, 551.5044233, 633.9000483, 683.8337772, 529.7526944, 569.1404291, 509.8675755, 540.9786305, 598.3247283, 761.718308, 516.5523976, 1185.754777, 804.8759389, 513.7129142, 560.9510722, 609.1662801, 524.6406256, 721.7654816)
# Calculation of statistics.
n <- length(x)
x_m_hat <- min(x)
y <- log(x / x_m_hat)
alpha_hat <- n / sum(y)
ase_x_m_hat <- x_m_hat /sqrt(n * alpha_hat)
LL_x_m <- x_m_hat - qnorm(0.975) * ase_x_m_hat
UL_x_m <- x_m_hat + qnorm(0.975) * ase_x_m_hat
ase_alpha_hat <- alpha_hat / sqrt(n)
LL_alpha_hat <- alpha_hat - qnorm(0.975) * ase_alpha_hat
UL_alpha_hat <- alpha_hat + qnorm(0.975) * ase_alpha_hat
alpha_c <- (n - 2) / sum(y)
ase_alpha_c <- alpha_c / sqrt(n - 2)
LL_alpha_c <- alpha_c - qnorm(0.975) * ase_alpha_c
UL_alpha_c <- alpha_c + qnorm(0.975) * ase_alpha_c
# Display results (rounding to three decimal places).
cat("sample size: n =", n, "\n")
cat("Estimate of the scale parameter: x_m_hat =", round(x_m_hat, 3), "\n")
cat("Asymptotic standard error for x_m_hat: ase(x_m_hat) =", round(ase_x_m_hat, 3), "\n")
cat("95% asymptotic confidence interval for x_m: 95% CI", "[", round(LL_x_m, 3), ",", round(UL_x_m, 3), "]", "\n")
cat("Estimate of shape parameter: alpha_hat =", round(alpha_hat, 3), "\n")
cat("Asymptotic standard error for alpha_hat: ase(alpha_hat) =", round(ase_alpha_hat, 3), "\n")
cat("95% asymptotic confidence interval for alpha: 95% CI", "[", round(LL_alpha_hat, 3), ",", round(UL_alpha_hat, 3), "]", "\n")
cat("Biased-corrected estimate of shape parameter: alpha_bc =", round(alpha_c, 3), "\n")
cat("Asymptotic standard error for alpha_bc: ase(alpha_bc) =",
```

```
round(ase_alpha_c, 3), "\n")
cat("95% asymptotic confidence interval for alpha_bc: 95% CI", "[", round(LL_alpha_c, 3), ",", round(UL_alpha_c, 3), "]", "\n")
```

The goodness-of-fit test begins with inferential tests. The 40 data points, x_i, following a type I Pareto distribution, are transformed using Equation 48 to follow an exponential distribution, y_i. The Kolmogorov-Smirnov test statistic is then calculated by applying Equation 50 to the transformed sample. Subsequently, Stephens' correction is applied (Equation 51). Table 2 presents the data transformation and the calculations to derive the statistics D^+ (the maximum difference between the empirical and theoretical cumulative distribution functions), D^- (the maximum difference between the theoretical cumulative distribution function and the empirical cumulative distribution function with a lag of 1), and D (the maximum of D^+ and D^-). After correcting D, it is found to be less than the critical value at a 5% significance level, supporting the null hypothesis of goodness-of-fit. Thus, the transformed sample data, y_i, follow an exponential distribution, and the original data, x_i, follow a type I Pareto distribution.

Statistical hypothesis

$$H_0 : X \sim \text{Pareto}\left(x_m, \alpha\right) \equiv Y = \ln\left(\frac{X}{\min\left(X\right)}\right) \sim \text{Exp}\left(\lambda = \alpha\right) \wedge H_1 : X \not\sim \text{Pareto}\left(x_m, \alpha\right)$$

Data transformation

$$y_i = \ln\left[x_i / 502.595\right] \in Y \sim \text{Exponetial}\left(\hat{\lambda} = n \Big/ \sum_{i=1}^{40} \ln\left(x_i\right) = 3.957\right)$$

$$F_Y\left(y_{(i)}\right) = 1 - \text{e}^{-\hat{\lambda} \times y_{(i)}} = 1 - \text{e}^{-3.957 \times y_{(i)}}$$

Test statistic

$$D^+ = \max\left[\left(i\right)/n - F_Y\left(y_{(i)}\right)\right] = 0.046$$

$$D^- = \max\left[F_Y\left(y_{(i)}\right) - \left(\left(i\right)-1\right)/n\right] = 0.037$$

$$D = \max\left(D^+, D^-\right) = \max\left(0.046, 0.037\right) = 0.046$$

Test statistic with Stephens' correction

$$D_c = \left(D - \frac{0.2}{n}\right)\left(\sqrt{n} + 0.26 + \frac{0.5}{\sqrt{n}}\right) = \left(0.046 - \frac{0.2}{40}\right)\left(\sqrt{40} + 0.26 + \frac{0.5}{\sqrt{40}}\right) = 0.275$$

Decision on the null statistical hypothesis of the type I Pareto distribution

$$D_c = 0.275 < D_{\alpha=0.05} = 1.094, H_0 \text{ is not rejected.}$$

$$H_0 : Y \sim \text{Exp}\left(\lambda\right) \Rightarrow X = 502.595\text{e}^Y \sim \text{Pareto}\left(x_m, \alpha\right)$$

The y_i data are also used in the application of the Anderson-Darling test. These data are arranged in ascending order. For the calculation of the statistic A^2 (Equation 52), the first value, which is 0, must be excluded because it corresponds to a cumulative probability of 0, and the logarithm of this probability is undefined. To compute the cumulative probabilities, the rate parameter λ is required, and it is

estimated using its maximum likelihood estimator, which is the reciprocal of the sample mean of the remaining 39 data points (Table 3). The corrected A^2 statistic, calculated using Equation 53, is less than the critical value at a 5% significance level (Stephens, 1986). Consequently, the null hypothesis holds, indicating that the transformed data $y_i = \ln(x_i/502.595)$ follow an exponential distribution and, therefore, the original data $x_i = 502.595 \times e^{y_i}$ follow a type I Pareto distribution.

Rate parameter estimate

$$\hat{\lambda} = \frac{39}{\sum_{i=1}^{39} y_{(i)}} = \frac{39}{10.108} = 3.858$$

Expected cumulative probability following an exponential distribution with a rate parameter of 3.858.

$$F_Y\left(y_{(i)} \mid \hat{\lambda} = 3.858\right) = \hat{\lambda} \times e^{-\hat{\lambda} y_{(i)}} = 3.858 \times e^{-3.858 \hat{y}_{(i)}}$$

$$\sum_{i=2}^{40} A_i = \sum_{i=2}^{40} \frac{2i-1}{39}\left[\ln\left(F_Y\left(y_{(i)}\right)\right) + \ln\left(1 - F_Y\left(y_{(n+1-i)}\right)\right)\right] = -39.080$$

Anderson-Darling test statistic

$$n' = \sum_{i=2}^{40} i = 39$$

$$A^2 = -n' - \sum_{i=1}^{39} A_i = -39 + 39.080 = 0.080$$

Anderson-Darling test statistic with Stephens' correction

$$A_c^2 = A^2\left(1 + \frac{0.6}{n'}\right) = 0.080\left(1 + \frac{0.6}{39}\right) = 0.081$$

Table 3. Calculations for deriving the Anderson-Darling A^2 statistic from the y_i values.

i	$y_{(i)}$	$(2i-1)/39$	$F_Y(y_{(i)})$	Ln $[F_Y(y_{(i)})]$	$y_{(n+1-i)}$	$F_Y[y_{(n+1-i)}]$	Ln $[1 - F_Y(y_{(n+1-i)})]$	A_i
1	0	0	0					
2	0.008	0.026	0.031	−3.481	1.095	0.985	−4.224	−0.198
3	0.014	0.077	0.054	−2.920	0.858	0.964	−3.312	−0.479
4	0.022	0.128	0.081	−2.514	0.757	0.946	−2.919	−0.697
5	0.027	0.179	0.100	−2.300	0.623	0.910	−2.405	−0.844
6	0.033	0.231	0.120	−2.123	0.614	0.906	−2.369	−1.037
7	0.043	0.282	0.153	−1.880	0.522	0.867	−2.015	−1.099
8	0.053	0.333	0.184	−1.694	0.471	0.837	−1.817	−1.170
9	0.057	0.385	0.198	−1.618	0.416	0.799	−1.604	−1.239
10	0.073	0.436	0.246	−1.402	0.377	0.766	−1.453	−1.244
11	0.074	0.487	0.247	−1.398	0.362	0.753	−1.396	−1.361
12	0.091	0.538	0.295	−1.220	0.337	0.727	−1.300	−1.357
13	0.093	0.590	0.301	−1.200	0.311	0.698	−1.198	−1.415
14	0.102	0.641	0.326	−1.122	0.308	0.695	−1.188	−1.481
15	0.110	0.692	0.345	−1.063	0.271	0.649	−1.046	−1.460
16	0.124	0.744	0.381	−0.965	0.252	0.621	−0.971	−1.439
17	0.128	0.795	0.390	−0.942	0.232	0.592	−0.896	−1.460
18	0.131	0.846	0.396	−0.927	0.213	0.561	−0.824	−1.481
19	0.159	0.897	0.458	−0.782	0.205	0.547	−0.792	−1.412
20	0.174	0.949	0.490	−0.714	0.192	0.524	−0.742	−1.381

Continued

21	0.176	1.000	0.493	−0.708	0.176	0.493	−0.679	−1.387
22	0.192	1.051	0.524	−0.647	0.174	0.490	−0.673	−1.387
23	0.205	1.103	0.547	−0.603	0.159	0.458	−0.612	−1.340
24	0.213	1.154	0.561	−0.578	0.131	0.396	−0.504	−1.248
25	0.232	1.205	0.592	−0.525	0.128	0.390	−0.494	−1.228
26	0.252	1.256	0.621	−0.476	0.124	0.381	−0.480	−1.201
27	0.271	1.308	0.649	−0.433	0.110	0.345	−0.424	−1.120
28	0.308	1.359	0.695	−0.364	0.102	0.326	−0.394	−1.030
29	0.311	1.410	0.698	−0.359	0.093	0.301	−0.358	−1.012
30	0.337	1.462	0.727	−0.318	0.091	0.295	−0.350	−0.977
31	0.362	1.513	0.753	−0.284	0.074	0.247	−0.284	−0.860
32	0.377	1.564	0.766	−0.266	0.073	0.246	−0.283	−0.859
33	0.416	1.615	0.799	−0.224	0.057	0.198	−0.221	−0.720
34	0.471	1.667	0.837	−0.177	0.053	0.184	−0.203	−0.634
35	0.522	1.718	0.867	−0.143	0.043	0.153	−0.166	−0.530
36	0.614	1.769	0.906	−0.098	0.033	0.120	−0.127	−0.399
37	0.623	1.821	0.910	−0.095	0.027	0.100	−0.106	−0.365
38	0.757	1.872	0.946	−0.055	0.022	0.081	−0.084	−0.262
39	0.858	1.923	0.964	−0.037	0.014	0.054	−0.055	−0.178
40	1.095	1.974	0.985	−0.015	0.008	0.031	−0.031	−0.091
Σ	10.108							

Note. i = order or rank of the data $y = \ln[x_i/\min(x)]$ where the data are ordered in ascending order, $y_{(i)}$ = y-value at order i, $(2i - 1)/39$ = first factor of A_i, $F_Y(y_{(i)})$ = cumulative probability for the $y_{(i)}$ value, assuming an exponential distribution with a rate parameter λ = 3.858, $\ln[F_Y(y_{(i)})]$ = natural logarithm of the cumulative probability for the value $y_{(i)}$, $y_{(n+1-i)}$ = y-value at order $n + 1 - i$, $F_{Y[y(n+1-i)]}$ = cumulative probability for the value $y_{(n+1-i)}$, following the same probability distribution; $\ln[1 - F_Y(y_{(n+1-i)})]$ = natural logarithm of the complement of the cumulative probability for the value $y_{(n+1-i)}$, $A_i = ((2i - 1)/39) \times (\ln[F_Y(y_{(i)})]+\ln[1-F_Y(y_{(n+1-i)})])$, Σ = sum by column.

The Anderson-Darling test, modified by Sinclair et al. (1990) for distributions with positive skewness, is applied to the original data sorted in ascending order $x_{(i)}$. Refer to Equation 54. The cumulative distribution function of the type I Pareto distribution, $F_X(x_{(i)}) = 1 - (x_{(i)}/502.595)^{3.957}$, is used for these calculations; however, the probabilities are identical to those given by the cumulative distribution function of an exponential distribution with rate parameter λ = 3.858. The test statistic AU_n^2 is less than the critical value for a sample size of 40 and a significance level of 5%, so the null hypothesis of fit to a type I Pareto distribution is not rejected. The detailed calculations are provided below, with part of the results summarized in Table 4.

Test statistic

$$AU_n^2 = \frac{n}{2} - 2\sum_{i=1}^{n} F_X\left(x_{(i)}\right) - \sum_{i=1}^{n}\left(2 - \frac{2i-1}{n}\right)\ln\left(1 - F_X\left(x_{(i)}\right)\right)$$

$$AU_n^2 = \frac{n}{2} - 2\sum_{i=1}^{n}\left(\frac{\hat{x}_m}{x_{(i)}}\right)^{\hat{\alpha}} - \sum_{i=1}^{n}\left(2 - \frac{2i-1}{n}\right)\ln\left(1 - \left(\frac{\hat{x}_m}{x_{(i)}}\right)^{\hat{\alpha}}\right)$$

$$AU_{40}^2 = \frac{40}{2} - 2\times 19.885 - \left(-19.802\right) = 20 - 19.969 = 0.031$$

Critical value:

$$t = \ln\left(\frac{p}{1-p}\right) = \ln\left(\frac{0.05}{0.95}\right) = -2.944$$

$$u = \frac{1}{1+0.3/\sqrt{n}} = \frac{1}{1+0.3/\sqrt{40}} = \frac{1}{1.0480} = 0.955$$

$$G(p) = 0.1170 - 0.03791t + 0.06318u + 0.09878tu + 0.009184t^2u - 0.00009742t^4u$$

$$\begin{aligned} G(p=0.05) &= 0.1170 - 0.03791\times(-2.944) + 0.06318\times 0.955 \\ &\quad + 0.09878\times(-2.944)\times 0.955 + 0.009184\times(-2.944)^2 \\ &\quad \times 0.955 - 0.00009742\times(-2.944)^4\times 0.955 \\ &= 0.080 \end{aligned}$$

$${}_pAU_n^2 = 1 - 1/\left[1+e^{G(p)}\right]$$

$${}_{0.05}AU_{40}^2 = 1 - 1/\left[1+e^{0.080}\right] = 0.520$$

$$AU_{40}^2 = 0.031 < {}_{0.05}AU_{40}^2 = 0.520,\ H_0 \text{ is not rejected}: X \sim \text{Pareto}(x_m, \alpha)$$

Table 4. Calculations for deriving the test statistic of the Anderson-Darling test modified by Sinclair et al. (1990).

i	$x_{(i)}$	$F_X(x_{(i)})$	$2 - [(2i-1)/40]$	$\ln[1 - F_X(x_{(i)})]$	AU_i
1	502.595	0	1.975	0	0
2	506.685	0.032	1.925	−0.032	−0.062
3	509.868	0.055	1.875	−0.057	−0.107
4	513.713	0.083	1.825	−0.087	−0.158
5	516.552	0.103	1.775	−0.108	−0.192
6	519.474	0.123	1.725	−0.131	−0.225
7	524.641	0.156	1.675	−0.170	−0.285
8	529.753	0.188	1.625	−0.208	−0.338
9	532.239	0.203	1.575	−0.227	−0.357
10	540.797	0.252	1.525	−0.290	−0.442
11	540.979	0.253	1.475	−0.291	−0.430
12	550.305	0.302	1.425	−0.359	−0.511
13	551.504	0.308	1.375	−0.367	−0.505
14	556.651	0.333	1.325	−0.404	−0.536
15	560.951	0.353	1.275	−0.435	−0.554
16	569.140	0.389	1.225	−0.492	−0.603
17	571.286	0.398	1.175	−0.507	−0.596
18	572.739	0.404	1.125	−0.517	−0.582
19	588.961	0.466	1.075	−0.628	−0.675
20	598.325	0.498	1.025	−0.690	−0.707
21	599.242	0.501	0.975	−0.696	−0.679
22	609.166	0.533	0.925	−0.761	−0.704
23	617.102	0.556	0.875	−0.812	−0.711
24	622.201	0.570	0.825	−0.845	−0.697
25	633.900	0.601	0.775	−0.919	−0.712
26	646.362	0.630	0.725	−0.996	−0.722

Continued

27	659.182	0.658	0.675	−1.073	−0.724
28	683.834	0.704	0.625	−1.219	−0.762
29	685.652	0.707	0.575	−1.229	−0.707
30	703.919	0.736	0.525	−1.333	−0.700
31	721.765	0.761	0.475	−1.432	−0.680
32	732.467	0.775	0.425	−1.490	−0.633
33	761.718	0.807	0.375	−1.645	−0.617
34	804.876	0.845	0.325	−1.863	−0.606
35	847.360	0.873	0.275	−2.067	−0.568
36	928.627	0.912	0.225	−2.429	−0.547
37	937.437	0.915	0.175	−2.467	−0.432
38	1071.113	0.950	0.125	−2.994	−0.374
39	1185.755	0.967	0.075	−3.397	−0.255
40	1502.032	0.987	0.025	−4.332	−0.108
Σ		19.885			−19.802

Note. i = order or rank of the sample data of X, where the 40 data points are sorted in ascending order, $x_{(i)}$ = sample data of X at order or rank i, $F_X(x_{(i)})$ = cumulative probability of $x_{(i)}$ under a Pareto distribution(x_m = 502.595, α = 3.957), 2 − [(2i − 1)/40] = first factor of AU_i, ln[1 − $F_X(x_{(i)})$] = natural logarithm of the probability to the right tail of $x_{(i)}$ or second factor of AU_i, AU_i = (2 − [(2i − 1)/40]) × ln[1 − $F_X(x_{(i)})$], Σ = sum per column.

To obtain the quantile-quantile plot, the data x_i (i = 1, 2, …, 40) are sorted in ascending order and each value is assigned its corresponding order (i), representing the empirical quantiles: $x_{(i)}$. Using the orders (i), the theoretical quantiles are computed based on the median of the sampling distribution of the i-th order statistic from a standard uniform continuous distribution: Mdn = (α − 1/3)/(α + β − 2/3). This distribution corresponds to a beta distribution with parameters α = (i) and β = n + 1 − (i). Consequently, the theoretical quantile's cumulative probability (order of quantile) is given by $p_{(i)}$ = ((i) + 1/3)/(n + 1/3). The theoretical quantiles are then obtained by evaluating the quantile function of a type I Pareto distribution at $p_{(i)}$. The values for its two parameters are taken from the estimates previously calculated using the maximum likelihood estimators (Table 2).

First coordinate pair of the theoretical quantile (on the abscissa axis) and the empirical quantile (on the ordinate axis) plot.

$$x_{(1)} = 502.595$$

$$p_{(1)} = \frac{(i)-1/3}{n+1/3} = \frac{1-1/3}{40+1/3} = 0.017$$

$$x_{t_{(1)}} = \hat{x}_m\left(1-p_{(i)}\right)^{-\frac{1}{\alpha}} = 502.595\left(1-0.017\right)^{-\frac{1}{3.957}} = 504.716$$

$$\left(x_{t_{(1)}} \quad x_{(1)}\right) = \left(502.595 \quad 504.716\right)$$

Figure 3 illustrates the coordinate pairs of theoretical quantiles $x_t(i)$ and empirical quantiles $x_{(i)}$. The theoretical quantiles are plotted on the horizontal (abscissa) axis as they predict the empirical quantiles, which are positioned on the vertical

(ordinate) axis. In the plot, the coordinate points align closely along a 45-degree line, indicating that the empirical data closely adhere to the theoretical model. This alignment can be quantified using the linear correlation coefficient; the square of this coefficient represents the proportion of shared variance. In this example, the linear correlation is 0.999, corresponding to a shared variance of 99.7%. Furthermore, the histogram with an overlaid density curve (Figure 4) exhibits the characteristic inverted J-shaped profile of a sample drawn from a type I Pareto distribution, supporting the conclusion that the sample data follow a type I Pareto distribution pattern.

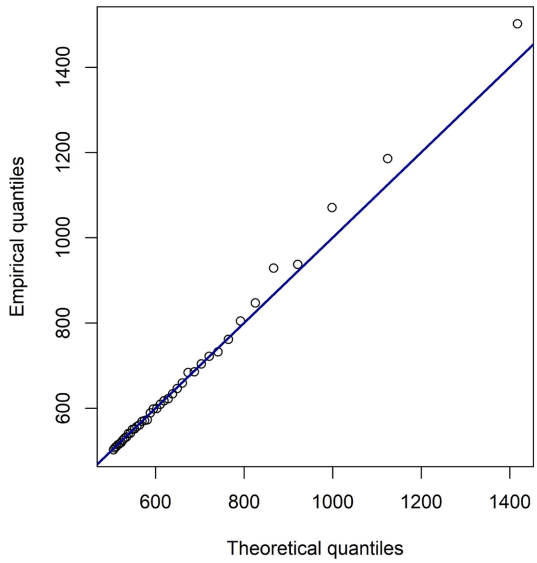

Figure 3. Quabtile-Quantile plot.

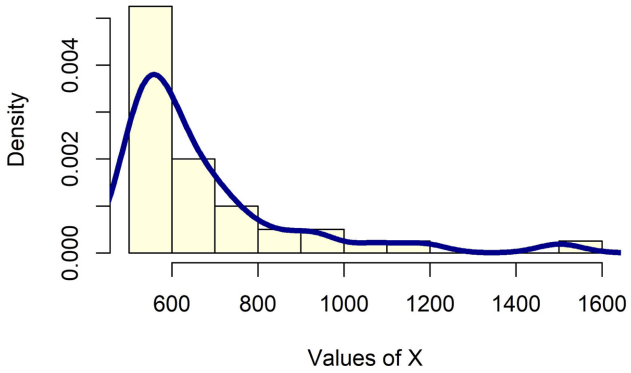

Figure 4. Histogram with an overlaid density curve.

The R script for testing the fit of sample data to a type I Pareto distribution using the Kolmogorov-Smirnov and Anderson-Darling tests, and for displaying the q-q plot and histogram with an overlaid density curve, is shown.

```
# Sample data.
x <- c(532.2385845, 659.1817235, 703.9185616, 732.4668655, 572.7389011,
```

```
1502.031724, 1071.11279, 556.6512741, 646.3622704, 622.2013372, 540.796541,
847.3602594, 928.6272967, 599.2422898, 519.4737837, 685.6524203, 506.685297,
550.3047107, 571.2856056, 617.1023659, 588.9614346, 502.594959, 937.4374491,
551.5044233, 633.9000483, 683.8337772, 529.7526944, 569.1404291, 509.8675755,
540.9786305, 598.3247283, 761.718308, 516.5523976, 1185.754777, 804.8759389,
513.7129142, 560.9510722, 609.1662801, 524.6406256, 721.7654816)

# Load library
Library (EnvStats)

# Kolmogorov-Smirnov test.
x_sorted <- sort(x)
n <- length(x_sorted) # Sample size
i <- 1:n
F_empirical <- i / n
x_m <- min(x_sorted)
y <- log(x_sorted / x_m)
lambda_hat <- n / sum(y)
F_theoretical <- 1 - exp(-lambda_hat * y)
D_plus <- max(F_empirical - F_theoretical)
F_empirical_prev <- c(0, F_empirical[-n]) # Add 0 at the beginning for F_n(x_(i-1))
D_minus <- max(F_theoretical - F_empirical_prev)
D <- max(D_plus, D_minus)
# Stephens' correction for D (D_c).
Dc <- (D - 0.2 / n) * (sqrt(n) + 0.26 + 0.5 / sqrt(n))
alpha <- 0.05 # Significance level
D_alpha <- 1.094 # Critical value for α = 0.05 (Stephens, 1974)
if (Dc <= D_alpha) {decision <- "H_0 holds: The data follow the theoretical distribution."
} else {decision <- "H_0 is rejected: The data do not follow the theoretical distribution."}
# Display results (rounding to three decimal places).
cat("D⁺:", round(D_plus, 3), "\n")
cat("D⁻:", round(D_minus, 3), "\n")
cat("D:", round(D, 3), "\n")
cat("D_c:", round(Dc, 3), "\n")
cat("Significance level: α =", alpha, "\n")
cat("Decision:", decision, "\n")

# Anderson-Darling test.
y_prime <- y[-1]
n_prime <- length(y_prime)
```

```
i_prime <- 1: n_prime
lambda_hat_prime <- n_prime / sum(y_prime)
FY_i <- 1 - exp(-lambda_hat_prime * y_prime)
log_FY_i <- log(FY_i)
FY_complement_log <- log(1 - FY_i[n_prime:i_prime])
A <- -n_prime - sum((2 * i_prime - 1) / n * (log_FY_i + FY_complement_log))
# Apply Stephens' correction.
Ac <- A * (1 + 0.6 / n_prime)
alpha <- 0.05 # Significance level
A_alpha <- 1.321 # Critical value for α = 0.05 (Stephens, 1974)
if (Ac <= A_alpha) {
decision <- "H_0 holds: The data follow the theoretical distribution."
} else {decision <- "H_0 is rejected: The data do not follow the theoretical distribution."}
# Display results.
cat("A:", A, "\n")
cat("Ac:", Ac, "\n")
cat("Significance level: α =", alpha, "\n")
cat("Decision:", decision, "\n")

# Anderson-Darling test modified by Sinclair et al. (1990).
alpha_hat <- lambda_hat
FX <- 1 - (x_m / x_sorted)^alpha_hat
sum(FX)
sum((2 - (2 * i - 1) / n) * log(1 - FX))
AU <- n / 2 - 2 * sum(FX) - sum((2 - (2 * i - 1) / n) * log(1 - FX))
p <- 0.05 # Significance level
t <- log(p / (1 - p))
u <- 1 / (1 + 0.3 / sqrt(n))
G_p <- 0.1170 - 0.03791 * t + 0.06318 * u + 0.09878 * t * u + 0.009184 * t^2 * u - 0.00009742 * t^4 * u
AU_np <- 1 - 1 / (1 + exp(G_p))
if (AU <= AU_np) {
decision <- "H_0 holds: The data follow the theoretical distribution."
} else {decision <- "H_0 is rejected: The data do not follow the theoretical distribution."}
# Display results (rounding to three decimal places).
cat("Estimate of scale parameter: x_m =", round(x_m, 3), "\n")
cat("Estimate of shape parameter: α = :", round(alpha_hat, 3), "\n")
cat("Test statistic: AU =:", round(AU, 3), "\n")
cat("Significance level: p =:", p, "\n")
cat("Critical value (AU_n,p):", round(AU_np, 3), "\n")
cat("Decision:", decision, "\n")
```

```
# Q-Q plot.
u <- (i - 1/3) / (n + 1/3)
q <- eqpareto(x, p = u, method = "mle", plot.pos.con = 1/3)
plot(q$quantiles, x_sorted, pch = 1, xlab = " Theoretical quantiles ", ylab = "Empirical quantiles")
abline(0, 1, col = "darkblue", lwd = 2)

# Histogram with an overlaid density curve.
hist(x, breaks = "fd", col = "lightyellow", border = "black", freq = FALSE, xlab = "Values of X", ylab = "Density", main = "")
lines(density(x), col="darkblue", lwd=4)
```

R provides a command to generate random samples that follow a Type I Pareto distribution. The EnvStats package must be loaded, and a seed can be set to ensure the results are stable (e.g., 123). It is only necessary to specify the sample size (n), the value of the scale parameter (x_m) in "location" (since in the Type I Pareto distribution, the scale and location parameters coincide), and the value of the shape parameter (α) in "shape".

```
Library (EnvStats)
set.seed (123)
rpareto (n = 40, location = 500, shape = log (5)/log(4))
```

9. Strengths and Weaknesses of Type I Pareto Distribution

The Pareto distribution is a valuable tool in social research for analyzing extreme inequality, but its simplicity and specific focus make it less versatile than other models for more balanced or nuanced distributions (Charpentier & Flachaire, 2022). The following are some of its most distinctive features:

Flexibility with inequality measures: The Pareto model is directly linked to measures of inequality, such as the Gini coefficient. It can succinctly model disparities, making it a preferred choice over less interpretable models like the exponential distribution.

Simplicity in application: With only two parameters, the type I Pareto distribution is relatively simple compared to multi-parameter models such as generalized Pareto or Weibull distributions. This simplicity can be advantageous for interpretability but might limit its ability to model more complex phenomena.

Fit for empirical data: While the type I Pareto distribution fits well for data with pronounced inequality, it might not perform as effectively for datasets with moderate or low inequality. In such cases, models like the log-normal or gamma distributions might provide better fits.

Possible combination with other distributions: The type I Pareto distribution typically focuses on the upper tail of a distribution (e.g., the wealthiest individuals). For comprehensive population modeling, hybrid models (e.g., combining Pareto for the tail and log-normal or beta for the body distribution) may be more

accurate (Akinsete, Famoye, & Lee, 2008; Martins, Liska, Beijo, Menezes, & Cirillo, 2020). However, this feature can be seen as a limitation of the Pareto distribution that needs to be combined.

Specifically, the Type I Pareto distribution has several limitations that should be considered when applying it to real world data. First, it is primarily suited to modelling the upper tails of distributions and may not accurately represent the entire data set, especially for moderate or low values. Second, its reliance on a power-law relationship imposes strict assumptions about the proportion of large values, which may oversimplify complex phenomena. In addition, the distribution assumes that the scale parameter (x_m) represents a minimum threshold, which may not always match the data structure. The Type I Pareto distribution also struggles to capture multimodal distributions or data sets with varying degrees of inequality across different ranges. Finally, while its simplicity makes it attractive in terms of interpretability, this can limit its flexibility in capturing more nuanced patterns, necessitating the use of hybrid or alternative models in certain contexts.

The type I Pareto distribution fails in contexts where the data are light-tailed, multimodal, highly dynamic or dominated by intermediate values, or where structural factors introduce complexity that cannot be accounted for by a simple power-law relationship. For example, in a society with a large and stable middle class, income distributions often peak around the median, which the Pareto model cannot effectively capture due to its emphasis on extremes.

10. Conclusion

The illustrated presentation of the type I Pareto distribution with scale parameter x_m and shape parameter α demonstrates that calculating probabilities is straightforward, as is computing descriptive statistics. As with other distributions, the most effective estimators for its parameters are those derived using the maximum likelihood method (Martín, Parra, Pizarro, & Sanjuán, 2022), which are both simple and computationally efficient. The estimator for x_m is the sample minimum, while the estimator for α is the inverse of the mean of the log-transformed sample data (Siudem et al., 2022). Naturally, it is necessary to assess the fit of the empirical data to the probability model by employing both a graphical approach—using a histogram with an overlaid density curve (characterized by an inverted J-shape) and a quantile-quantile plot (showing a 45-degree alignment)—and an inferential approach, involving the Kolmogorov-Smirnov (Chu et al., 2019), Anderson-Darling (Stephens, 1986), or modified Anderson-Darling (Sinclair et al., 1990) tests. In the first two tests, the data are transformed to follow an exponential distribution, whereas the third test does not require such a transformation.

All these calculations can be performed using the scripts developed and presented in this article, which can be adapted to other score vectors. Additionally, Excel is also a valuable tool for this purpose. It should be noted that R is not only a free-access program, but it is currently the most comprehensive among statistical software, and it is gaining popularity among researchers in psychology and

social sciences (Navarro, 2024).

Within the social sciences, this distribution serves as an effective probability model for various phenomena, including the distribution of income, rent, accumulated resources within a company, region, or country, and insurance claims (Feng et al., 2020). It is also applicable to the average frequency of behaviors that are rare among most individuals but highly prevalent in a few, such as compulsive behaviors, behavioral or substance addictions, and paraphilias (Rajeev, 2022). Notably, it can be applied to the same continuous variables as the lognormal distribution, such as epidemiological data (Beare & Toda, 2020), necessitating an evaluation to determine which probability model provides the best fit (Charpentier & Flachaire, 2022; Feng et al., 2020). For discrete data defined on a bounded set of natural numbers, the Zipf distribution—a variant of the zeta distribution—is the appropriate choice. From a social perspective, a good fit to a Paretian probability model highlights the presence of well-defined funnel rules and inequality in resource allocation. This, in turn, may inspire policies of segregation targeting specific social sectors deemed responsible for disproportionate expenditure or debt, such as health insurance for the elderly (Rodríguez-Abreu, 2021).

Acknowledgements

The author expresses his gratitude to the reviewers and editors for the suggestions and corrections received for the improvement of the manuscript.

Conflicts of Interest

The author declares no conflicts of interest regarding the publication of this paper.

References

Ahmad, H. A. H., & Almetwally, E. M. (2020). Marshall-Olkin Generalized Pareto Distribution: Bayesian and Non Bayesian Estimation. *Pakistan Journal of Statistics and Operation Research, 16,* 21-33. https://doi.org/10.18187/pjsor.v16i1.2935

Akinsete, A., Famoye, F., & Lee, C. (2008). The Beta-Pareto Distribution. *Statistics, 42,* 547-563. https://doi.org/10.1080/02331880801983876

Anderson, T. W., & Darling, D. A. (1952). Asymptotic Theory of Certain "Goodness of Fit" Criteria Based on Stochastic Processes. *The Annals of Mathematical Statistics, 23,* 193-212. https://doi.org/10.1214/aoms/1177729437

Andria, J. (2022). A Computational Proposal for a Robust Estimation of the Pareto Tail Index: An Application to Emerging Markets. *Applied Soft Computing, 114,* Article ID: 108048. https://doi.org/10.1016/j.asoc.2021.108048

Arnold, B. C. (2015). *Pareto Distribution* (2nd ed.). John Wiley y Sons, Ltd. https://doi.org/10.1201/b18141

Barczy, M., K. Nedényi, F., & Sütő, L. (2023). Probability Equivalent Level of Value at Risk and Higher-Order Expected Shortfalls. *Insurance: Mathematics and Economics, 108,* 107-128. https://doi.org/10.1016/j.insmatheco.2022.11.004

Barnoy, A., & Reich, Z. (2022). Trusting Others: A Pareto Distribution of Source and Message Credibility among News Reporters. *Communication Research, 49,* 196-220.

Beare, B. K., & Toda, A. A. (2020). On the Emergence of a Power Law in the Distribution of COVID-19 Cases. *Physica D: Nonlinear Phenomena, 412,* Article ID: 132649. https://doi.org/10.1016/j.physd.2020.132649

Benczes, I. (2022). Taking Back Control over the Economy: From Economic Populism to the Economic Consequences of Populism. *European Policy Analysis, 8,* 109-123. https://doi.org/10.1002/epa2.1134

Bhoj, D. S., & Chandra, G. (2021). Ranked Set Sampling with Lowest Order Statistics for Pareto Distribution. *Communications in Statistics-Simulation and Computation, 52,* 2327-2335. https://doi.org/10.1080/03610918.2021.1904143

Campbell, M. R., & Brauer, M. (2021). Is Discrimination Widespread? Testing Assumptions about Bias on a University Campus. *Journal of Experimental Psychology: General, 150,* 756-777. https://doi.org/10.1037/xge0000983

Charpentier, A., & Flachaire, E. (2022). Pareto Models for Top Incomes and Wealth. *The Journal of Economic Inequality, 20,* 1-25. https://doi.org/10.1007/s10888-021-09514-6

Chattamvelli, R., & Shanmugam, R. (2021). Pareto Distribution. In *Continuous Distributions in Engineering and the Applied Sciences-Part II* (pp. 179-188). Springer International Publishing. https://doi.org/10.1007/978-3-031-02435-1_3

Chen, B., Zhang, K., Wang, L., Jiang, S., & Liu, G. (2019). Generalized Extreme Value-Pareto Distribution Function and Its Applications in Ocean Engineering. *China Ocean Engineering, 33,* 127-136. https://doi.org/10.1007/s13344-019-0013-9

Cheng, W., Fu, H., Wang, L., Dong, C., Jin, Y., Jiang, M. et al. (2023). Data-Driven, Multi-moment Fluid Modeling of Landau Damping. *Computer Physics Communications, 282,* Article ID: 108538. https://doi.org/10.1016/j.cpc.2022.108538

Chu, J., Dickin, O., & Nadarajah, S. (2019). A Review of Goodness of Fit Tests for Pareto Distributions. *Journal of Computational and Applied Mathematics, 361,* 13-41. https://doi.org/10.1016/j.cam.2019.04.018

Diawara, D., Kane, L., Dembele, S., & Lo, G. S. (2021). Applying of the Extreme Value Theory for Determining Extreme Claims in the Automobile Insurance Sector: Case of a China Car Insurance. *Afrika Statistika, 16,* 2883-2909. https://doi.org/10.16929/as/2021.2883.188

Fedotenkov, I. (2020). A Review of More than One Hundred Pareto-Tail Index Estimators. *Statistica, 80,* 245-299. https://doi.org/10.6092/issn.1973-2201/9533

Feng, M., Deng, L., Chen, F., Perc, M., & Kurths, J. (2020). The Accumulative Law and Its Probability Model: An Extension of the Pareto Distribution and the Log-Normal Distribution. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences, 476,* Article ID: 20200019. https://doi.org/10.1098/rspa.2020.0019

Gini, C. (1936). On the Measure of Concentration with Special Reference to Income and Statistics. *Colorado College Publication, General Series, 208,* 73-79.

Landoni, J. S., & Villegas, L. (2022). *Pagan los Pobres: Consecuencias Negativas de Políticas Públicas con Buenas (y Malas) Intenciones [The Poor Pay: Negative Consequences of Public Policies with Good (and Bad) Intentions]*. Editorial Galerna.

Le Gall, P., Favre, A., Naveau, P., & Prieur, C. (2022). Improved Regional Frequency Analysis of Rainfall Data. *Weather and Climate Extremes, 36,* Article ID: 100456. https://doi.org/10.1016/j.wace.2022.100456

Lomax, K. S. (1954). Business Failures: Another Example of the Analysis of Failure Data. *Journal of the American Statistical Association, 49,* 847-852. https://doi.org/10.1080/01621459.1954.10501239

Lorenz, M. O. (1905). Methods of Measuring the Concentration of Wealth. *Publications of*

the American Statistical Association, 9, 209-219. https://doi.org/10.2307/2276207

Martín, J., Parra, M. I., Pizarro, M. M., & Sanjuán, E. L. (2022). Baseline Methods for the Parameter Estimation of the Generalized Pareto Distribution. *Entropy, 24,* Article No. 178. https://doi.org/10.3390/e24020178

Martins, A. L. A., Liska, G. R., Beijo, L. A., Menezes, F. S. d., & Cirillo, M. Â. (2020). Generalized Pareto Distribution Applied to the Analysis of Maximum Rainfall Events in Uruguaiana, RS, Brazil. *SN Applied Sciences, 2,* Article No. 1479. https://doi.org/10.1007/s42452-020-03199-8

Mateus, A., & Caeiro, F. (2022). Confidence Intervals for the Shape Parameter of a Pareto Distribution. *AIP Conference Proceedings, 2425,* Article ID: 320003. https://doi.org/10.1063/5.0081541

McCarthy, D. M., & Winer, R. S. (2019). The Pareto Rule in Marketing Revisited: Is It 80/20 or 70/20? *Marketing Letters, 30,* 139-150. https://doi.org/10.1007/s11002-019-09490-y

Mojiri, A., & Ahmadi, K. (2022). Inequality in the Distribution of Resources in Health Care System by Using the Gini Coefficient and Lorenz Curve (A Case Study of Sistan and Baluchestan Province over a Five-Year Period). *Health Monitor Journal of the Iranian Institute for Health Sciences Research, 21,* 227-236. https://doi.org/10.52547/payesh.21.3.227

Navarro, D (2024). *Learning Statistics with R—A Tutorial for Psychology Students and Other Beginners*. LibreTexts Libraries. Statistics. https://stats.libretexts.org/Bookshelves/Applied_Statistics/Learning_Statistics_with_R_-_A_tutorial_for_Psychology_Students_and_other_Beginners_(Navarro)

Pareto, V. F. D. (1896). *Cours d'Economie Politique* (Vol. 1). F. Rouge éditeur.

Pareto, V. F. D. (1897). *Cours d'Economie Politique* (Vol. 2). F. Rouge éditeur.

Pearson, K. (1895). Contributions to the Mathematical Theory of Evolution. II. Skew Variation in Homogeneous Material. *Philosophical Transactions of the Royal Society of London A, 186,* 343-414. https://doi.org/10.1098/rsta.1895.0010

Pearson, K. (1905). "Das fehlergesetz und seine verallgemeiner-ungen durch fechner und pearson." A rejoinder. *Biometrika, 4,* 169-212. https://doi.org/10.1093/biomet/4.1-2.169

Qian, W., Chen, W., & He, X. (2021). Parameter Estimation for the Pareto Distribution Based on Ranked Set Sampling. *Statistical Papers, 62,* 395-417. https://doi.org/10.1007/s00362-019-01102-1

Rácz, E., Spasibko, K., Manceau, M., Ruppert, L., Chekhova, M. V., & Filip, R. (2023). *Quantifying Optical Rogue Waves*. https://doi.org/10.48550/arXiv.2303.04615

Rajeev, C. D. S. (2022). Pareto Principle and Compulsive Buying Disorder—An Analysis. *Journal of Educational and Social Research, 8,* 44-59.

Rao, C. R. (1973). *Linear Statistical Inference and Its Applications*. Wiley. https://doi.org/10.1002/9780470316436

Rodríguez Abreu, M. (2021). Gasto de bolsillo y gastos catastróficos en salud en hogares mexicanos. *Carta Económica Regional, 34,* 59-83. https://doi.org/10.32870/cer.v0i128.7825

Ross, S. M. (2022). *Simulation* (6th ed.). Academic Press.

Rytgaard, M. (1990). Estimation in the Pareto Distribution. *ASTIN Bulletin, 20,* 201-216. https://doi.org/10.2143/ast.20.2.2005443

Safari, M. A. M., Masseran, N., Ibrahim, K., & Hussain, S. I. (2019). A Robust and Efficient Estimator for the Tail Index of Inverse Pareto Distribution. *Physica A: Statistical Mechanics and its Applications, 517,* 431-439.

https://doi.org/10.1016/j.physa.2018.11.029

Sarabia, J. M., Jordá, V., & Prieto, F. (2019). On a New Pareto-Type Distribution with Applications in the Study of Income Inequality and Risk Analysis. *Physica A: Statistical Mechanics and Its Applications, 527,* Article ID: 121277. https://doi.org/10.1016/j.physa.2019.121277

Sinclair, C. D., Spurr, B. D., & Ahmad, M. I. (1990). Modified Anderson Darling Test. *Communications in Statistics—Theory and Methods, 19,* 3677-3686. https://doi.org/10.1080/03610929008830405

Sitthiyot, T., & Holasut, K. (2021). A Simple Method for Estimating the Lorenz Curve. *Humanities and Social Sciences Communications, 8,* Article No. 268. https://doi.org/10.1057/s41599-021-00948-x

Siudem, G., Nowak, P., & Gagolewski, M. (2022). Power Laws, the Price Model, and the Pareto Type-2 Distribution. *Physica A: Statistical Mechanics and its Applications, 606,* Article ID: 128059. https://doi.org/10.1016/j.physa.2022.128059

Song, I., Ryoung-Park, S., & Yoon, S. (2022). *Probability and Random Variables: Theory and Applications*. Springer International Publishing.

Stephens, M. A. (1974). EDF Statistics for Goodness of Fit and Some Comparisons. *Journal of the American Statistical Association, 69,* 730-737. https://doi.org/10.2307/2286009

Stephens, M. A. (1986). Tests Based on EDF Statistics. In R. B. D'Agostino, & M. A. Stephens (Eds.), *Goodness-of-Fit Techniques* (pp. 97-193) Marcel Dekker, Inc. https://doi.org/10.1201/9780203753064-4

Sudharson, D., & Prabha, D. (2019). Retracted Article: A Novel Machine Learning Approach for Software Reliability Growth Modelling with Pareto Distribution Function. *Soft Computing, 23,* 8379-8387. https://doi.org/10.1007/s00500-019-04047-7

Sudharson, D., Divya, P., Ratheeshkumar, M., Saravanan, A., Nithiyashree, V. K., & Srinithi, J. (2022). A PD ANN Machine Learning Framework for Reliability Optimization in Application Software. In *2022 Smart Technologies, Communication and Robotics (STCR)* (pp. 1-4). Institute of Electrical and Electronics Engineers. https://doi.org/10.1109/stcr55312.2022.10009626

Tokhirov, A., Harmáček, J., & Syrovátka, M. (2021). Remittances and Inequality: The Post-Communist Region. *Prague Economic Papers, 30,* 426-448. https://doi.org/10.18267/j.pep.776

Valkanas, K., & Diamandis, P. (2022). Pareto Distribution in Virtual Education: Challenges and Opportunities. *Canadian Medical Education Journal, 13,* 102-104. https://doi.org/10.36834/cmej.73511

World Bank (2022). *Gini Index*. https://data.worldbank.org/indicator/

Xu, T., Sedory, S. A., & Singh, S. (2022). Lowering the Cramer-Rao Lower Bounds of Variance in Randomized Response Sampling. *Communications in Statistics—Simulation and Computation, 51,* 4112-4126. https://doi.org/10.1080/03610918.2020.1737874

Yang, X., & Zhou, P. (2022). Wealth Inequality and Social Mobility: A Simulation-Based Modelling Approach. *Journal of Economic Behavior & Organization, 196,* 307-329. https://doi.org/10.1016/j.jebo.2022.02.012

Zhang, Y., Wu, Y., & Yao, H. (2022). Optimal Health Insurance with Constraints under Utility of Health, Wealth and Income. *Journal of Industrial and Management Optimization, 18,* 1519-1540. https://doi.org/10.3934/jimo.2021031