



## Distribución de Pareto: un modelo de probabilidad en investigación social

Pareto distribution: a probability model in social research

José Moral de la Rubia

### Resumen

Este artículo metodológico tiene como objetivo exponer la distribución de Pareto de una forma clara y ejemplificada para su comprensión entre investigadores sociales. Esta distribución continua con forma de J invertida, asimetría y cola pesada a la derecha constituye un buen modelo de probabilidad para diversas variables sociales, como riqueza e ingreso, y de conductas que son altamente frecuentes en pocas personas y poco frecuentes en la mayoría. Se desarrolla la distribución tipo I que tiene un parámetro de escala  $x_m$  y otro de forma  $\alpha$ . Se inicia con una breve reseña histórica. Se muestran las funciones de densidad, distribución acumulativa, de cola, de momentos y característica. Se sigue con las medidas descriptivas, estimadores por los métodos de momentos y máxima verosimilitud, su relación con otras distribuciones y las pruebas de bondad de ajuste. Se aplica todo este material con dos ejemplos: uno de cálculo de probabilidades y medidas descriptivas, y otro de estimación de parámetros y comprobación del ajuste. Finalmente, se hacen sugerencias para su uso.

**Palabras clave:** distribución de probabilidad continua; estimación de parámetros; medidas descriptivas; índice de cola de Pareto; índice de concentración de Gini.

### Abstract

This methodological article aims to expose the Pareto distribution in a clear and exemplified way for its understanding among social researchers. This continuous distribution with an inverted J shape, skewness, and a heavy tail to the right constitutes a good probability model for various social variables, such as wealth and income, and for behaviors that are highly frequent in a few people and infrequent in the majority. The type-I distribution is developed, which has a scale parameter  $x_m$  and another of shape  $\alpha$ . It begins with a brief historical review. The density, cumulative distribution, tail, moment, and characteristic functions are shown. It continues with the descriptive measures, estimators by the methods of moments and maximum likelihood, its relationship with other distributions, and the goodness-of-fit tests. All this material is applied with two examples: one of calculating probabilities and descriptive measures and the other of parameter estimation and fit testing. Finally, suggestions for its use are made.

**Keywords:** continuous probability distribution; parameter estimation; descriptive measures; Pareto tail index; Gini concentration index.

## 1. Introducción<sup>1</sup>

La distribución de Pareto es una de las distribuciones continuas no normales que puede ser de gran utilidad en el área de las Ciencias Sociales y de la Conducta (Barnoy y Reich, 2022; Feng et al., 2020). Cabe señalar que la exposición explicativa de esta distribución suele ser muy teórica y confusa, sobre todo por ser presentada bajo distintas parametrizaciones (Sarabia, Jorda y Prieto, 2019) y muy enfocada al área de la economía (Barczy, Nedényi y Sütő, 2023), de ahí que el objetivo de este artículo es presentar esta distribución de probabilidad de una forma clara, comprensible y ejemplificada para distintos tipos de investigadores sociales, como psicólogos, sociólogos y trabajadores sociales.

La distribución fue desarrollada entre 1896 y 1897 por el ingeniero civil, economista y sociólogo italiano Vilfredo Federico Damaso Pareto (1848–1923) durante su curso de economía política impartido en la universidad de Lausana (UNIL), Suiza. Surge para describir el reparto de la tierra, riqueza e ingresos en una sociedad. De estos estudios, se infiere el principio 20-80 que enuncia que el 80% de la riqueza se concentra en el 20% de la población. Cabe señalar que tanto la distribución como el principio de desigualdad hoy en día se aplican como modelos de probabilidad para estas mismas variables (Charpentier y Flachaire, 2022; Feng et al., 2020), aunque la desigualdad suele ser menor, con un máximo de 30-70 (McCarthy y Winer, 2019). Además, la distribución y la regla de desigualdad sirven como modelos en otros ámbitos distintos a la economía, como en psicología (Campbell y Brauer, 2021; Rajeev, 2022), ingeniería (Chen et al., 2019; Sudharson y Prabha, 2019), climatología (Le Gall, et al., 2022) y física (Rácz et al., 2023; Xiong et al., 2019). Las reclamaciones a las aseguradoras por accidentes y enfermedades siguen una distribución de Pareto, lo que lleva a estas compañías a castigar excesivamente las primas de seguros de determinados sectores de la población, como adultos mayores (Diawara et al., 2021; Zhang, Wu y Yao, 2022).

## 2. Caracterización de la distribución

### 2.1. Parámetros y soporte

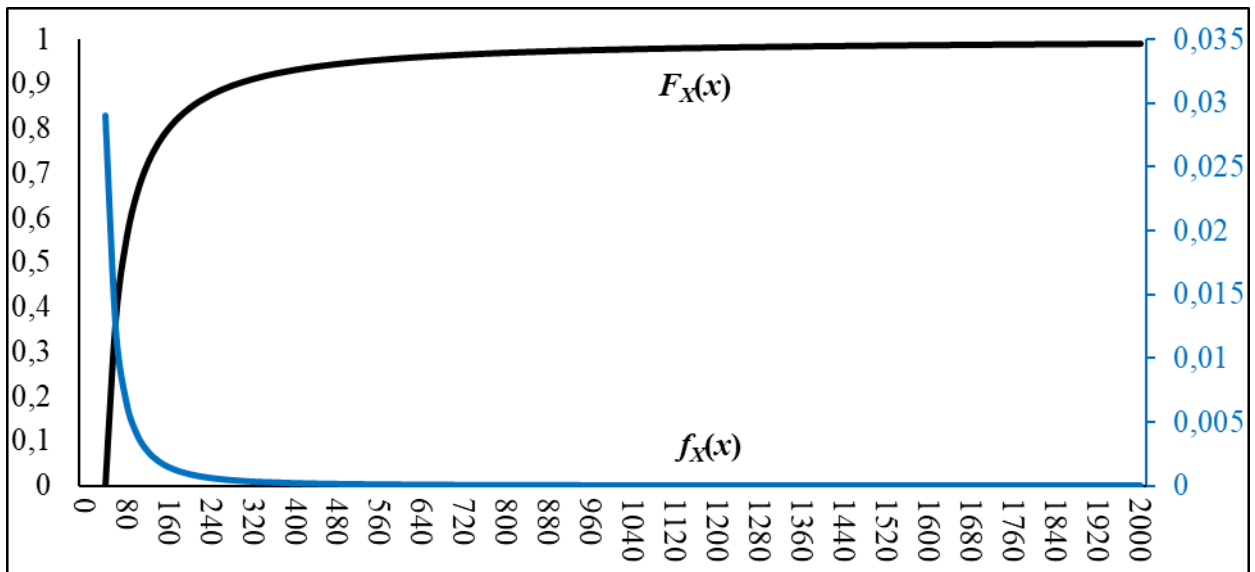
Esta distribución continua de la familia de distribuciones de potencia está determinada por dos parámetros en su presentación más sencilla o tipo I (Ahmad y Almetwally, 2020). Un parámetro es de escala, se puede denotar por  $x_m$ , aunque su notación es muy variable, y corresponde al pico o moda y al valor mínimo de la distribución (Chattamvelli y Shanmugam, 2021). El otro parámetro es de forma, se denota por  $\alpha$  (Fedotenkov, 2020) y se suele denominar índice de cola de Pareto (Andria, 2022). El espacio parametral de ambos abarca el intervalo  $(0, \infty)$ . El soporte de la distribución corresponde al intervalo  $[x_m, \infty)$ . El perfil de su función de densidad describe una curva con un descenso rápido desde su valor máximo hasta hacerse asíntota al eje de abscisas, por lo que presenta asimetría positiva y leptocurtosis con cola larga a la derecha (Figura 1). El valor de  $\alpha$  que modela con exactitud la ley 20-80 es  $\log_4(5) = \ln(5)/\ln(4) \approx 1.161$ , el cual es un valor demasiado extremo para la distribución de la riqueza que, usualmente, es mayor que 1.5 (Yang y Zhou, 2022).

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

$$x_m \text{ y } \alpha \in (0, \infty) \text{ y } x \in [x_m, \infty)$$

<sup>1</sup> Esta es la versión en español del artículo: Moral de la Rubia, J. (2025). Pareto distribution: a probability model in social research. *Open Journal of Social Sciences*, 13(1), 86-121. <https://doi.org/10.4236/jss.2025.131007>

**Figura 1. Función de densidad  $f_X(x)$  y probabilidad acumulativa  $F_X(x)$  del ingreso por hora trabajada de una distribución de Pareto con parámetro de escala  $x_m = 40$  y forma  $\alpha = \log_4(5)$ .**



Fuente: elaboración propia.

## 2.2. Funciones

Función de densidad:

$$f_X(x) = \alpha \frac{x_m^\alpha}{x^{\alpha+1}}, \text{ para valores de } x \geq x_m$$

Función de distribución acumulativa:

$$F_X(x) = P(X \leq x) = \int_{x_m}^x f_X(x) dx = 1 - \left(\frac{x_m}{x}\right)^\alpha, \text{ para valores de } x \geq x_m$$

Función de distribución acumulativa complementaria o de cola:

$$\bar{F}_X(x) = P(X > x) = \int_{x_m}^{\infty} f_X(x) dx = \left(\frac{x_m}{x}\right)^\alpha, \text{ para valores de } x \geq x_m$$

Función cuantil:

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

Función generadora de momentos: Para un intervalo no trivial de  $t$  en torno a 0 no está definida, pero sí para valores de  $t \leq 0$ .

$$M_X(t) = E(e^{tx}) = \int_{x_m}^{+\infty} e^{tx} f_X(x) dx = \begin{cases} \alpha(-x_m t)^\alpha \Gamma(-\alpha, -x_m t) & t < 0 \\ 1 & t = 0 \\ \text{Indefinida} & t > 0 \end{cases}$$

Función gamma incompleta superior:

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

Función característica:

$$i = \sqrt{-1} \text{ y } t \in \mathbb{R}$$

$$C_X(t): \mathbb{R} \rightarrow \mathbb{C}; C_X(t) = E(e^{itx}) = \int_{x_m}^{\infty} e^{itx} f_X(x) dx = \alpha(-ix_m t)^{\alpha} \Gamma(-\alpha, -ix_m t)$$

### 2.3. Medidas descriptivas de tendencia central, variación y forma

Esperanza matemática o media aritmética:

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

Media geométrica:

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

Media armónica:

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

Mediana:

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

Moda:

$$Mo(X) = \{x | \max[f_X(x)]\} = x_m$$

Momento no central de orden  $k$ :

$$E(X^k) = \int_{x_m}^{\infty} x^k f_X(x) dx = \frac{\alpha}{\alpha-k} x_m^k, \alpha > k$$

Varianza:

$$Var(X) = \sigma^2(X) = E[(X - E(X))^2] = \int_{x_m}^{\infty} (x - E(X))^2 f_X(x) dx = \frac{\alpha x_m^2}{(\alpha-1)^2(\alpha-2)}, \alpha > 2$$

Desviación estándar:

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

Entropía: Al estar el logaritmo con base neperiana, la entropía viene en nats.

$$H(X) = E(-\ln[f_X(x)]) = - \int_{x_m}^{\infty} f_X(x) \ln[f_X(x)] dx = \ln\left[\frac{x_m}{\alpha} e^{1+\frac{1}{\alpha}}\right] = \ln(x_m) - \ln(\alpha) + 1 + 1/\alpha$$

Asimetría y exceso de curtosis basadas en momentos centrales con la notación original de Karl Pearson:

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

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

### 3. Estimación de parámetros

Sea una muestra aleatoria de tamaño  $n$  de una variable cuantitativa continua  $X$  que sigue una distribución de Pareto tipo I con parámetros  $x_m$  y  $\alpha$ . Por ejemplo, los datos muestrales podrían ser registros del salario mensual de trabajadores elegidos al azar en una empresa grande.

$$\{x_1, x_2, \dots, x_n\} \subseteq X \sim \text{Pareto}(x_m, \alpha)$$

#### 3.1. Estimador de $\alpha$ por el método de momentos

El estimador del parámetro de forma  $\alpha$  por el método de momentos, se obtiene a partir de la esperanza matemática o media aritmética de  $X$ , si se conoce el valor del parámetro de localización  $x_m$ , que es el pico y valor mínimo de la distribución. Dicha esperanza matemática corresponde al cociente entre la media muestral (numerador) y la diferencia entre la media muestral y el parámetro conocido  $x_m$  (denominador). Es válido cuando  $\alpha > 2$ .

$$E(X) = \frac{\alpha}{\alpha - 1} x_m$$

$$\alpha E(X) - E(X) - \alpha x_m = 0$$

$$\alpha(E(X) - x_m) = E(X)$$

$$\alpha = \frac{E(X)}{E(X) - x_m} = \frac{\mu(X)}{\mu(X) - x_m}$$

$$\hat{\alpha} = \frac{\hat{E}(X)}{\hat{E}(X) - x_m} = \frac{\hat{\mu}(X)}{\hat{\mu}(X) - x_m} = \frac{\bar{x}}{\bar{x} - x_m}, \text{ donde } \hat{\mu}(X) = \bar{x} = \frac{\sum_{i=1}^n x_i}{n}$$

El estimador puntual converge a una distribución normal, al tener X una distribución con momentos finitos (Rao, 1973), lo que permite obtener un estimador por intervalo asintótico.

$$\lim_{n \rightarrow \infty} \hat{\alpha}_n = \frac{\bar{x}}{\bar{x} - x_m} \xrightarrow{d} N\left(\alpha, \frac{\alpha(\alpha - 1)^2}{n(\alpha - 2)}\right)$$

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

### 3.2. Estimadores de $X_m$ y $\alpha$ por el método de máxima verosimilitud

El estimador por el método de máxima verosimilitud del parámetro de escala  $x_m$  es el valor mínimo muestral y el del parámetro de forma  $\alpha$  es el inverso de la media aritmética de los logaritmos de los cocientes entre cada valor  $x_i$  y el mínimo muestral de los valores (Siudem, Nowak y Gagolewski, 2022).

$$\hat{x}_m = \min(x_1, x_2, \dots, x_n) = \min(\{x_i\}_{i=1}^n)$$

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

La información de Fisher para  $n$  datos de estos dos parámetros viene dada por la siguiente matriz cuadrada  $2 \times 2$ :

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

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

La varianza de un parámetro  $\theta$ ,  $Var(\theta)$ , siempre es mayor o igual que su cuota inferior de Cramer y Rao,  $CICR(\theta)$  o inverso de su información de Fisher para  $n$  datos,  $1/I(\theta)$  (Xu, Sedory y Singh, 2022).

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

Los estimadores obtenidos por el método de máxima verosimilitud tienen como propiedades asintóticas la insesgadez, consistencia, eficiencia y normalidad, lo que los hacen muy útiles (Song, Ryoung-Park y Yoon, 2022).

$$\hat{\theta}_n \text{ es un estimador asintóticamente insesgado de } \theta, \text{ si } \lim_{n \rightarrow \infty} E(\hat{\theta}_n) = \theta$$

$$\hat{\theta}_n \text{ es un estimador asintóticamente consistente de } \theta, \text{ si } \lim_{n \rightarrow \infty} P(|\hat{\theta}_n - \theta| > \varepsilon) = 0$$

$\hat{\theta}_n$  es un estimador asintóticamente eficiente de  $\theta$ , si  $\lim_{n \rightarrow \infty} \frac{CICR(\hat{\theta})}{Var(\hat{\theta}_n)} = 1$

$\hat{\theta}_n$  como estimador  $\theta$  converge a la normalidad, si  $\lim_{n \rightarrow \infty} \frac{\sqrt{n}(\hat{\theta}_n - \theta)}{DE(\hat{\theta})} \xrightarrow{d} N(0, 1)$

Desde estas propiedades asintóticas, se puede definir un error asintótico para los estimadores por máxima verosimilitud de  $x_n$  y  $\alpha$ ; así como un intervalo de confianza asintótico.

$$EEA(\hat{\theta}) = \sigma_{\hat{\theta}} = \sqrt{1/I(\hat{\theta})}$$

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

$$EEA(\hat{x}_m) = \sqrt{\frac{1}{n\alpha/x_m^2}} = \sqrt{\frac{x_m^2}{n\alpha}} = \frac{x_m}{\sqrt{n\alpha}}; \widehat{EEA}(\hat{x}_m) = \frac{\hat{x}_m}{\sqrt{n\hat{\alpha}}}$$

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

$$EEA(\hat{\alpha}) = \sqrt{\frac{1}{n/\alpha^2}} = \sqrt{\frac{\alpha^2}{n}} = \frac{\alpha}{\sqrt{n}}; \widehat{EEA}(\hat{\alpha}) = \frac{\hat{\alpha}}{\sqrt{n}}$$

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

El estimador de  $\alpha$  por el método de máxima verosimilitud es más eficiente que el del método de momentos, pero no es insesgado. Ante esta salvedad, se puede definir un estimador corregido de sesgo que tiene una varianza menor que el estimador sesgado (Rytgaard, 1990) y se puede usar para lograr una estimación asintótica más eficiente. Como las anteriores definiciones, el uso de estas fórmulas asintóticas requiere una muestra grande, a saber, un tamaño mayor que 30 y de preferencia de al menos 100 (Mateus y Caeiro, 2022).

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

$$\bar{V}ar(\hat{\alpha}_c) = \hat{\alpha}_c^2/n - 2 < Var(\hat{\alpha})$$

$$\widehat{EE}(\hat{\alpha}_c) = \sqrt{\hat{\alpha}_c^2/(n-2)} = \hat{\alpha}_c/\sqrt{n-2}$$

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

### 3.3. Distribuciones exactas

Sea  $X$  una variable aleatoria con distribución de Pareto  $(x_m, \alpha)$ . La distribución exacta de la suma de los logaritmos de cada uno de los  $n$  datos muestrales de  $X$  divididos por el valor mínimo muestral sigue una distribución exponencial de parámetro de tasa igual que  $\alpha$  (Rytgaard, 1990).

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

La distribución exacta del estimador por máxima verosimilitud del parámetro de escala  $x_m$  es una distribución Pareto con parámetro de escala  $x_m$  y parámetro de forma  $n\alpha$  y la distribución exacta del estimador por máxima verosimilitud de parámetro de forma  $\alpha$  es una distribución gamma inversa con parámetro de forma  $n - 1$  y parámetro de escala  $n\alpha$  (Qian, Chen y He, 2021).

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

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

### 4. Relación con otras distribuciones e índice de concentración de Gini

Existe una forma generalizada de la distribución de Pareto de cuatro parámetros: de localización  $\mu$ , escala  $\sigma$  y dos de forma (índice de cola de Pareto  $\alpha$  e índice de desigualdad  $\gamma$ ). Su función de distribución acumulativa corresponde a la siguiente expresión (Arnold, 2015):

Soporte:  $x \geq \mu$

Espacio parametral:  $\mu \in (-\infty, \infty)$  y  $\sigma, \alpha$  y  $\gamma \in (0, \infty)$ .

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

El parámetro de localización  $\mu$  no es la esperanza matemática o media aritmética de la distribución y el parámetro de escala  $\sigma$  no es la desviación estándar. La media depende de los parámetros  $\sigma, \alpha$  y  $\gamma$  y unas funciones gamma y requiere que  $\gamma < \alpha$ .

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

En la distribución de Pareto tipo I,  $\mu = \sigma = x_m$  (mínimo y moda) y  $\gamma = 1$ .

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

En la distribución de Pareto tipo II,  $\gamma = 1$ . Cuando el parámetro  $\mu$  de la distribución de Pareto tipo II es igual que 0, recibe el nombre de distribución de Lomax (1954).

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



En la distribución de Pareto tipo III,  $\alpha = 1$ .

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

Sea  $X$  una variable aleatoria con distribución de Pareto con parámetros  $x_m$  y  $\alpha$ . El producto del parámetro de escala  $x_m$  y la exponencial con base el número de Euler ( $e = 2.718281\dots$ ) yponente  $Y$  sigue una distribución exponencial de parámetro de tasa o inverso de escala  $\lambda = \alpha$ . A la inversa, sea  $Y$  una variable aleatoria con distribución exponencial de parámetro de tasa o inverso de escala  $\lambda$ . El logaritmo neperiano del cociente entre la variable aleatoria  $Y$  y el valor  $x_m$  sigue una distribución de Pareto tipo I de parámetro de escala  $x_m$  y parámetro de forma  $\alpha = \lambda$ . Esto implica que las funciones de distribución acumulativas de  $X$  e  $Y$  son iguales e intercambiables cuando se calculan las probabilidades acumuladas.

$$X \sim \text{Pareto}(x_m, \alpha); x_m e^X = Y \sim \text{Exponencial}(\lambda = \alpha)$$

$$Y \sim \text{Exp}(\lambda); \ln\left(\frac{Y}{x_m}\right) = X \sim \text{Pareto}(x_m, \alpha = \lambda)$$

$$P(X \leq x) = F_X(x|x_m, \alpha) = F_Y[y = \ln(x/x_m) | \lambda = \alpha] = 1 - (x_m/x)^\alpha = 1 - e^{-\alpha y}$$

Cabe señalar que entre las distribuciones exponencial y de Pareto hay una relación análoga a la existente entre las distribuciones normal y lognormal. Las distribuciones de Pareto y lognormal se pueden aplicar a los mismos datos, aunque una probablemente tenga mejor ajuste que la otra, y ambas son una transformación exponencial de una variable más ampliamente conocida y usada, la primera de una distribución exponencial y la segunda de una distribución normal (Feng et al., 2020).

La distribución zeta es el equivalente discreto de la distribución de Pareto. Si el soporte de la distribución zeta está acotado en un intervalo de números naturales constituye la denominada distribución de Zipf (Arnold, 2015).

$$X \sim Z(s)$$

$$s \in (1, \infty), x \in \mathbb{N}$$

$$P(X = x) = f_X(x) = \frac{1}{x^s \zeta(s)} = \frac{1}{x^s \sum_{n=1}^{\infty} 1/n^s}$$

$$P(X \leq x) = F_X(x) = \frac{\sum_{n=1}^x 1/n^s}{\zeta(s)} = \frac{\sum_{n=1}^x 1/n^s}{\sum_{n=1}^{\infty} 1/n^s}$$

$$E(X) = \frac{\zeta(s-1)}{\zeta(s)} = \frac{\sum_{n=1}^{\infty} 1/n^{s-1}}{\sum_{n=1}^{\infty} 1/n^s}; s > 2$$

$$E(X^k) = \frac{\zeta(s-k)}{\zeta(s)} = \frac{\sum_{n=1}^{\infty} 1/n^{s-k}}{\sum_{n=1}^{\infty} 1/n^s}; s > k + 1$$

$$\text{Var}(X) = E(X^2) - E(X)^2 = \frac{\zeta(s)\zeta(s-2) - [\zeta(s-1)]^2}{[\zeta(s)]^2}; s > 3$$

$$Mo(X) = 1$$

El índice de cola de Pareto (1897) está relacionado con el índice de concentración de Gini (Safari *et al.*, 2019). Corrado Gini en 1936 creó este índice para resumir en un solo número la información contenida en la curva de Lorenz (Mojiri y Ahmadi, 2022). La curva de Lorenz (1905) se obtiene a partir de un diagrama bidimensional que representa el número de personas en el eje horizontal de abscisas ( $x$ ) y el ingreso o la riqueza acumulada en el eje vertical o de ordenadas,  $F(x)$ . Dentro de esta gráfica se dispone una recta con una pendiente de 45 grados que se denomina línea de igualdad. En la medida que la curva definida por los pares coordenados  $(x, F(x))$  se separa de esta recta, hay más desigualdad. El índice de Gini es la proporción del área entre la línea de igualdad y la curva de Lorenz y el área del triángulo recto inferior. Un valor de 0 indica total igualdad, y aparece cuando todos los individuos tienen el mismo ingreso; un valor de 1 refleja la máxima diferencia, y surge cuando una persona recibe todo el ingreso y las demás personas serían mano de obra no remunerada o sin ingresos (Sitthiyot y Holasut, 2021). En países con más igualdad de ingresos, como los escandinavos, el índice Gini es menor que 0.3. En los países con más desigualdad, como en algunos del continente africano (Sudáfrica, Namibia, Surinam, Zambia Eswatini, Botsuana, Angola y Zimbabue) y americano (Bélice, Brasil, Colombia y Panamá), el índice Gini es mayor que 0.5 (World Bank, 2022). Cabe señalar que los países con sistemas de economía comunista no poseen índices de Gini bajos y las políticas socialdemócratas que implican aumento del gasto y deuda públicos lo incrementan en lugar de disminuirlo (Tokhirov, 2021). Así, se ha denominado a las políticas populistas de izquierdas políticas de igualación en la pobreza con la emergencia de una oligarquía política (Aránguez-Sánchez, 2019; Landoni y Villegas, 2022).

El valor en la curva de Lorenz en el punto  $x$  para una variable aleatoria continua  $X$  con función de densidad  $f_x(x)$  se obtiene mediante la siguiente proporción:

$$L(F(x)) = \frac{\int_{-\infty}^x xf(x) dx}{\int_{-\infty}^{\infty} xf(x) dx} = \frac{\int_{-\infty}^x xf(x) dx}{E(X)}$$

Aplicado a la distribución de Pareto, la expresión anterior queda del siguiente modo:

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

$$F(x) = 1 - (x_m/x)^\alpha$$

$$L(F(x)) = 1 - (1 - F(x))^{1-\frac{1}{\alpha}} = 1 - \left[1 - \left(1 - \left(\frac{x_m}{x}\right)^\alpha\right)\right]^{1-\frac{1}{\alpha}} = 1 - \left(\frac{x_m}{x}\right)^{\alpha-1}$$

$$L(F(x)) = 0, 0 < \alpha < 1$$

El índice de Gini se puede expresar como uno menos el doble de la integral de 0 a 1 de la curva de Lorenz. Aplicado a la distribución de Pareto, queda como el inverso del doble del parámetro de cola menos uno,  $1/(2\alpha-1)$ , siempre que  $\alpha > 1$ . Cuando  $\alpha < 1$ , el índice de Gini toma su valor mínimo que es 1. En la Tabla 1, se muestra la correspondencia entre valores de los índices de cola de Pareto y concentración de Gini y su interpretación.

$$G(X) = 1 - 2 \int_0^1 L(F(x)) dx = 1 - 2 \int_0^1 \left(1 - \left(\frac{x_m}{x}\right)^{\alpha-1}\right) dx = \frac{1}{2\alpha-1}, \alpha \geq 1$$

$$G(X) = 1, 0 < \alpha < 1$$

$$\alpha = 0.5(1 + 1/G(X))$$

**Tabla 1. Correspondencia entre los índices de Gini y Pareto y su interpretación**

G	$\alpha$	Interpretación
0.05	10.5	Muy poca desigualdad
0.1	5.5	
0.15	3.833	Poca desigualdad
0.2	3	
0.25	2.5	
0.3	2.167	
0.35	1.929	Desigualdad
0.4	1.75	
0.45	1.61	Bastante desigualdad
0.5	1.5	
0.55	1.409	Mucha desigualdad
0.6	1.333	
0.65	1.269	
0.7	1.214	
0.75	1.167	Muchísima desigualdad
0.8	1.125	
0.85	1.088	
0.9	1.056	
0.95	1.026	
1	1	

Nota. G = índice de concentración de Gini y  $\alpha$  = índice de cola de Pareto.

Fuente: elaboración propia

## 5. Generación de muestras aleatorias paretianas y contraste de bondad de ajuste

Sea  $U$  una variable aleatoria con distribución uniforme estándar, la transformación  $x_m(1-U)^{-1/\alpha}$  sigue una distribución de Pareto con parámetro de escala  $x_m$  y de forma  $\alpha$ . Este procedimiento permite obtener una muestra aleatoria con distribución de Pareto a partir de una muestra aleatoria de una variable uniforme estándar y se denomina muestreo por transformada inversa (Gass y Fu, 2016).

Para comprobar si los datos muestrales se ajustan a la distribución de Pareto, Chu, Dickin y Nadarajah (2019) recomiendan la prueba de Kolmogorov y Smirnov con base en un estudio de simulación. La forma más sencilla para aplicar esta prueba inferencial es transformar los datos con supuesta distribución de Pareto a valores con una distribución exponencial a través del logaritmo neperiano del cociente entre cada dato y el valor mínimo muestral  $y$ , a continuación, aplicar la prueba de Kolmogorov y Smirnov adaptada a muestras exponencialmente distribuidas (Stephens, 1974). El contraste inferencial se puede complementar con una comprobación gráfica a través del diagrama cuantil-cuantil y el histograma. En el primero, se busca una alineación de puntos en una recta con pendiente de 45 grados y, en el segundo, un perfil con forma de J invertida (Bhoj y Chandra, 2021).

Hipótesis estadísticas:  $H_0: X \sim \text{Pareto}(x_m, \alpha) \equiv Y = \ln[X/\min(X)] \sim \text{Exponencial}(\lambda = \alpha)$  y  $H_1: X \not\sim \text{Pareto}(x_m, \alpha)$ .

Supuestos: Muestra aleatoria de tamaño  $n$  de la variable cuantitativa continua  $X$ .

Estadístico de contraste: Se ordenan los datos en sentido ascendente, se les asignan rangos u órdenes y se transforman para que sigan una distribución exponencial bajo el supuesto de que la hipótesis nula es cierta.

$X_{(1)}$	$X_{(2)}$	...	$X_{(i)}$	...	$X_{(n-1)}$	$X_{(n)}$
1	2	...	$i$	...	$n-1$	$n$
$\ln[X_{(1)}/X_{(1)}]$	$\ln[X_{(2)}/X_{(1)}]$		$\ln[X_{(i)}/X_{(1)}]$		$\ln[X_{(n-1)}/X_{(1)}]$	$\ln[X_{(n)}/X_{(1)}]$

Se calcula la probabilidad acumulada para cada valor transformado y los estadísticos  $D^+$ ,  $D^-$  y  $D$ .

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

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

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

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

$$D = \max(D^+, D^-)$$

Se aplica la corrección de estimación muestral al estadístico  $D$  (Stephens, 1974).

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

Se toma la decisión desde el estadístico  $D$  transformado. Si  $D_c \leq D_\alpha$ , se mantiene  $H_0$  y si  $D_c > D_\alpha$ , se rechaza  $H_0$  con un nivel de significación de  $\alpha$ . Los valores críticos  $D_\alpha$  dependen del nivel de significación  $\alpha$  (Stephens, 1974), el cual usualmente se fija en 0.05 ( $D_{0.05} = 1.094$ ). Con muestras pequeñas ( $n < 20$ ), se puede subir a 0.10 ( $D_{0.10} = 0.990$ ) y, con muestras grandes ( $n > 500$ ), se puede bajar a 0.01 ( $D_{0.01} = 1.308$ ).

Otra opción para el contraste inferencial es la prueba de Anderson y Darling (1952). Al igual que con el test de Kolmogorov y Smirnov se transforman los datos para que sigan una distribución exponencial y se aplica una corrección al estadístico de contraste cuando los parámetros son estimados desde la muestra (Stephens, 1986).

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

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

Como en la prueba de Kolmogorov y Smirnov, se toma la decisión desde el estadístico  $A^2$  corregido. Si  $A_c^2 \leq A_\alpha^2$  (valor crítico), se mantiene  $H_0$  y si  $A_c^2 > A_\alpha^2$ , se rechaza  $H_0$  con un nivel de significación de  $\alpha$ . Los valores críticos  $A_\alpha^2$  dependen del tipo de distribución (exponencial), método de estimación (estimador de máxima verosimilitud) y nivel de significación  $\alpha$  (Stephens, 1986), el

cual usualmente se fija en 0.05 ( $D_{0.05} = 1.321$ ). Con muestras pequeñas ( $n < 20$ ), el nivel de significación  $\alpha$  se puede subir a 0.10 ( $D_{0.10} = 1.062$ ) y, con muestras grandes ( $n > 500$ ), se puede bajar a 0.01 ( $D_{0.01} = 1.959$ ).

Existe una modificación de la prueba de Anderson y Darling para distribuciones con asimetría positiva y casos muy atípicos en la cola derecha, la cual fue desarrollada por Sinclair, Spurr y Ahmad (1990). La distribución de Pareto corresponde a este tipo de distribución, por lo que la prueba se aplica a la muestra original en caso de usarse. Su estadístico de contraste se denota por  $AU_n^2$ .

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

$$F_X(x_{(i)}) = 1 - (x_i/\hat{x}_m)^2 = 1 - (x_i/\min(x))^2$$

Cálculo del valor crítico para la toma de decisión:

$${}_pAU_n^2 = 1 - \frac{1}{1 + e^{G(p)}}$$

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

$$t = \ln \frac{p}{1-p}, u = \frac{1}{1 + 0.3/\sqrt{n}}$$

Si  $AU_n^2 \leq {}_pAU_n^2$ , se mantiene  $H_0: X \sim \text{Pareto}(x_m, \alpha)$  con el nivel de significación de  $p$  y  $AU_n^2 > {}_pAU_n^2$ , se rechaza. Cuando el tamaño muestral tiende a infinito, el valor de  $u$  se hace unitario y el valor crítico  ${}_pAU_n^2$  alcanza su valor asintótico:  ${}_pAU_n^2 = 0.356$  con  $p = 0.1$ ,  ${}_pAU_n^2 = 0.432$  con  $p = 0.05$  y  ${}_pAU_n^2$  es 0.610 con  $p = 0.01$ .

## 6. Ejemplos de cálculo con la distribución Pareto

### 6.1. Cálculo de probabilidades y medidas descriptivas

**Instrucciones:** Sea  $X$  una variable aleatoria con distribución de Pareto de parámetros  $x_m = 500$  y  $\alpha = 4.1$  que mide el salario quincenal en dólares en una empresa. Calcule la probabilidad de tener un ingreso menor o igual que 592, menor que 1000, entre 800 y 1400, mayor o igual que 600 y mayor que 1500. Compute la esperanza matemática o media aritmética  $\mu(X)$ , media geométrica  $\mu_g(X)$ , media armónica  $\mu_h(X)$ , mediana  $Mdn(X)$  y moda  $Mo(X)$  como medidas de tendencia central; la varianza  $Var(X)$ , desviación estándar  $DE(X)$  y entropía  $H(X)$  como medidas de variación; así como las medidas de la forma basadas en momentos centrales estandarizados de asimetría  $\sqrt{\beta_1(X)}$  y exceso de curtosis  $\beta_2(X)$ . Obtenga el índice de concentración de Gini  $G(X)$ . Finalmente, represente su función de densidad  $f_X(x)$  y distribución acumulativa  $F_X(x)$ .

$$X \sim \text{Pareto}(x_m = 500, \alpha = 4.1)$$

Probabilidad de tener un ingreso menor o igual que 592:

$$P(X \leq 592) = \int_{500}^{592} f_X(x) dx = 1 - (x_m/x)^\alpha = 1 - (500/592)^{4.1} = 0.4997$$

Probabilidad de tener un ingreso menor que 1000:

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

Probabilidad de tener un ingreso entre 800 y 1400:

$$P(800 \leq X \leq 1400) = P(X \leq 1400) - P(X \leq 800) = 0.9853 - 0.8544 = 1309$$

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

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

Probabilidad de tener un ingreso mayor o igual que 600:

$$P(X \geq 600) = (500/600)^{4.1} = 0.4735$$

Probabilidad de tener un ingreso mayor que 1500:

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

Medidas de tendencia central:

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

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

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

$$Mdn(X) = Q_X(p = 0.5) = \sqrt[4]{2} \times x_m = 2^{1/4.1} \times 500 = 592.10$$

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

Medidas de variación:

$$Var(X) = \mu_2(X) = E[(X - E(X))^2] = x_m^2 \frac{\alpha}{(\alpha - 1)^2(\alpha - 2)} = 500^2 \frac{4.1}{3.1^2 \times 2.1} = 50790.35$$

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

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

Asimetría y exceso de curtosis basadas en momentos centrales:

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

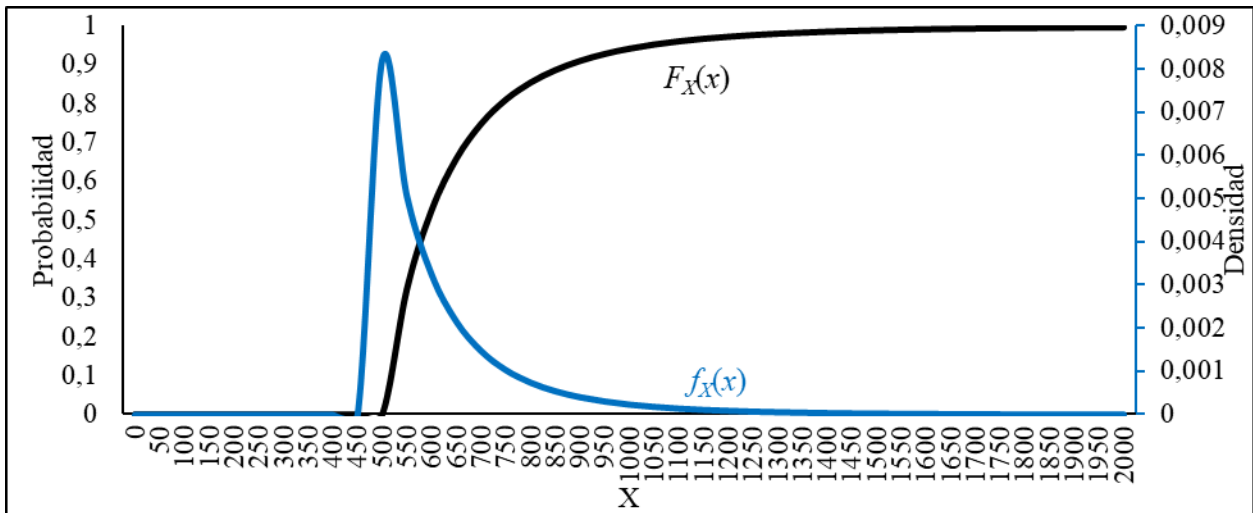
$$\beta_2(X) - 3 = \frac{\mu_4(X)}{\mu_2^2(X)} - 3 = \frac{6(\alpha^3 + \alpha^2 - 6\alpha - 2)}{\alpha(\alpha - 3)(\alpha - 4)} = \frac{6(4.1^3 + 4.1^2 - 6 \times 4.1 - 2)}{4.1 \times 1.1 \times 0.1} = 786.67$$

Índice de Gini:

$$G(X) = 1/(2\alpha - 1) = 1/(2 \times 4.1 - 1) = 0.1389$$

El índice de Gini está próximo a 0, lo que indica bastante igualdad en los salarios dentro de la empresa. En la Figura 2, se representa la función de densidad  $f_X(x)$  y distribución acumulativa  $F_X(x)$  de la variable aleatoria X de salario quincenal (en dólares) que sigue una distribución Pareto( $x_m = 500$ ,  $\alpha = 4.1$ ).

**Figura 2. Función de densidad  $f_X(x)$  y distribución acumulativa  $F_X(x)$  de  $X \sim \text{Pareto}(500, 4.1)$ .**



Fuente: elaboración propia.

## 6.2. Generación de muestra aleatoria, estimación de parámetros y bondad de ajuste

**Instrucciones:** Genere una muestra aleatoria de 40 datos desde la variable  $X \sim \text{Pareto}(500, 4.1)$ . Estime puntualmente los parámetros  $x_m$  y  $\alpha$ , usando los estimadores de máxima verosimilitud, y por intervalos con un nivel de confianza al 95% mediante el error y distribución asintóticos. Finalmente, compruebe la bondad de ajuste con la gráfica de los cuantiles teóricos versus los cuantiles empíricos y el histograma, así como con las pruebas de Kolmogorov y Smirnov, y de Anderson y Darling.

En la primera columna de la Tabla 2, se dispone una muestra aleatoria de 40 datos  $u_i$  extraída de una distribución uniforme estándar  $U \sim U[0, 1]$ , para lo cual se usó el generador de números aleatorios de Excel versión 2021. A través de la función cuantil, los datos  $u_i$  se transforman en una muestra aleatoria con un tamaño de 40 de una variable X, la cual sigue una distribución de Pareto con parámetros poblacionales:  $x_m = 500$  (de escala) y  $\alpha = 4.1$  (de forma). Este proceso generativo de datos constituye el denominado muestreo por transformada inversa.

$$x_i = x_m(1 - u_i)^{-\frac{1}{\alpha}} = 500(1 - u_i)^{-\frac{1}{4.1}}$$

$$x_1 = 500(1 - 0.226)^{-\frac{1}{4.1}} = 532.239$$

**Tabla 2. Generación de una muestra aleatoria y contraste del ajuste**

Generación de X		Prueba de Kolmogorov y Smirnov						Gráfica C-C	
$u_i$	$x_i$	$(i)$	$x_{(i)}$	$y_{(i)}$	$F_Y(y_{(i)})$	$D^+$	$D^-$	$p_{(i)}$	$x_{t(i)}$
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
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
$\Sigma$			10.108						

Nota. Generación de X:  $u_i$  = muestra en su orden aleatorio extraída de una distribución uniforme estándar y  $x_i = 500(1-u_i)^{-1/4.1}$  = datos transformados para seguir una distribución de Pareto(500, 4.1) mediante muestreo por transformada inversa. Transformación logarítmica de X y prueba de Kolmogorov y Smirnov: (i) = orden del dato  $x_i$  dentro de la muestra de 40 elementos,  $x_{(i)}$  = datos  $x_i$  ordenados en sentido ascendente o cuantiles empíricos,  $y_{(i)} = \ln(x_{(i)}/502.595)$  = transformación logarítmica de los datos  $x_i$  divididos por el mínimo muestral,  $F_Y(y_{(i)})$  = función de distribución acumulativa de la variable Y que sigue una distribución exponencial con parámetro de tasa  $\lambda = 3.957$ ,  $D^+ = (i)/40 - F_Y(y_{(i)})$  = diferencia entre las funciones de distribución acumulativas empírica y teórica y  $D^- = F_Y(y_{(i)}) - ((i)-1)/40$  = diferencia entre la función de distribución empírica teórica y la empírica con regazo de 1. Gráfica Cuantil-Cuantil:  $p_{(i)} = ((i)-1/3)/(n+1/3)$  = orden del cuantil teórico y  $x_{(i)} = Q_X[p_{(i)}]$  = cuantiles teóricos calculados mediante la función cuantil de una distribución de Pareto, dando a los parámetros sus valores estimados.  $\Sigma$  = suma por columna, Max = valor máximo por columna.

Fuente: elaboración propia.

A continuación, se estiman puntualmente los dos parámetros mediante los estimadores del método de máxima verosimilitud y por intervalo con los errores asintóticos.

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

$$\widehat{EEA}(\hat{x}_m) = \frac{\hat{x}_m}{\sqrt{n\hat{\alpha}}} = \frac{502.595}{\sqrt{40 \times 3.957}} = 39.948$$

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

$$P(502.595 - 1.96 \times 39.948 \leq x_m \leq 502.595 + 1.96 \times 39.948) = 0.95$$

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

$$\hat{\alpha} = \frac{n}{\sum_{i=1}^n \ln\left(\frac{x_i}{\hat{x}_m}\right)} = \frac{n}{\sum_{i=1}^n \ln\left(\frac{x_i}{\min(\{x_i\}_{i=1}^n)}\right)} = \frac{40}{\sum_{i=1}^{40} \ln\left(\frac{x_i}{502.595}\right)} = \frac{40}{10.108} = 3.957$$

$$\widehat{EEA}(\hat{\alpha}) = \hat{\alpha}/\sqrt{n} = 3.957/\sqrt{40} = 0.626$$

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

$$P(3.957 - 1.96 \times 0.626 \leq \alpha \leq 3.957 + 1.96 \times 0.626) = 0.95$$

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

También se acude a la fórmula corregida de sesgo en la estimación del parámetro de forma  $\alpha$ . En un principio, la estimación puntual de  $\alpha$  parece menos exacta, pero la estimación por intervalo es más eficiente y adecuada.

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

$$\widehat{EE}(\hat{\alpha}_c) = \hat{\alpha}_c/\sqrt{n-2} = 3.759/\sqrt{38} = 0.610$$

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

$$P(3.759 - 1.96 \times 0.610 \leq \alpha \leq 3.759 + 1.96 \times 0.610) = 0.95$$

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

El contraste de la bondad de ajuste se inicia con las pruebas inferenciales. Los 40 datos con distribución de Pareto  $x_i$  se transforman para que sigan una distribución exponencial  $y_i$  y sobre esta muestra se aplica la prueba de Kolmogorov y Smirnov con el procedimiento de Stephens (1974). En la Tabla 2, se encuentra la transformación de los datos y los cálculos para obtener los estadísticos  $D^+$  (diferencia máxima entre la función de distribución acumulativa empírica y la teórica) y  $D^-$  (diferencia máxima entre la función de distribución acumulativa teórica y la empírica con rezago de 1) y  $D$  (máximo de  $D^+$  y  $D^-$ ). Una vez transformado  $D$ , se observa que es menor que el valor crítico para un nivel de significación del 5%, por lo que se mantiene la hipótesis nula de bondad de ajuste. Los datos muestrales transformados  $y_i$  siguen una distribución exponencial y los datos sin transformar  $x_i$  siguen una distribución de Pareto.

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

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

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

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

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

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

Se transforma el estadístico de contraste D (Stephens, 1974)

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

$$D_c = 0.275 < D_{\alpha=0.05} = 1.094, \text{ se mantiene } H_0.$$

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

Los datos  $y_i$  también se usan en la aplicación de la prueba de Anderson y Darling (1952). Se ordenan en sentido ascendente. Para el cálculo de estadístico  $A^2$ , se requiere eliminar el primero, cuyo valor es 0, ya que le corresponde un valor de probabilidad acumulada de 0 y el logaritmo de esta probabilidad no es un número real. Para el cálculo de las probabilidades acumuladas, se estima el parámetro de tasa  $\lambda$  a través de su estimador de máxima verosimilitud que es el inverso de la media muestral de los 39 datos restantes (Tabla 3). El estadístico  $A^2$  corregido es menor que el valor crítico para un nivel de significación del 5% (Stephens, 1986), por lo que se mantiene la hipótesis nula de que los datos  $y_i = \ln(x_i/502.595)$  siguen una distribución exponencial y, por consiguiente, los datos  $x_i = 502.595 \times e^{y_i}$  una distribución de Pareto.

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

$$F_Y(y_{(i)}) = \hat{\lambda} e^{-\hat{\lambda} y_{(i)}} = 3.858 e^{-3.858 y_{(i)}}$$

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

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

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

$$A_c^2 = 0.081 < A_{0.05^2} = 1.321, \text{ se mantiene } H_0: Y = \ln(X/x_m) \sim \text{Exp}(\alpha); X \sim \text{Pareto}(x_m, \alpha)$$

**Tabla 3. Cálculos para obtener el estadístico  $A^2$  de Anderson y Darling desde los valores  $y_i$**

$i$	$y_{(i)}$	$(2i-1)/39$	$F_Y(y_{(i)})$	$\ln[F_Y(y_{(i)})]$	$y_{(n+1-i)}$	$F_Y[y_{(n+1-i)}]$	$\ln[1-F_Y(y_{(n+1-i)})]$	$A_i$
	0	0	0					
1	0.008	0.026	0.031	-3.481	1.095	0.985	-4.224	-0.198
2	0.014	0.077	0.054	-2.920	0.858	0.964	-3.312	-0.479
3	0.022	0.128	0.081	-2.514	0.757	0.946	-2.919	-0.697
4	0.027	0.179	0.100	-2.300	0.623	0.910	-2.405	-0.844
5	0.033	0.231	0.120	-2.123	0.614	0.906	-2.369	-1.037
6	0.043	0.282	0.153	-1.880	0.522	0.867	-2.015	-1.099
7	0.053	0.333	0.184	-1.694	0.471	0.837	-1.817	-1.170
8	0.057	0.385	0.198	-1.618	0.416	0.799	-1.604	-1.239
9	0.073	0.436	0.246	-1.402	0.377	0.766	-1.453	-1.244
10	0.074	0.487	0.247	-1.398	0.362	0.753	-1.396	-1.361
11	0.091	0.538	0.295	-1.220	0.337	0.727	-1.300	-1.357
12	0.093	0.590	0.301	-1.200	0.311	0.698	-1.198	-1.415
13	0.102	0.641	0.326	-1.122	0.308	0.695	-1.188	-1.481
14	0.110	0.692	0.345	-1.063	0.271	0.649	-1.046	-1.460
15	0.124	0.744	0.381	-0.965	0.252	0.621	-0.971	-1.439
16	0.128	0.795	0.390	-0.942	0.232	0.592	-0.896	-1.460
17	0.131	0.846	0.396	-0.927	0.213	0.561	-0.824	-1.481
18	0.159	0.897	0.458	-0.782	0.205	0.547	-0.792	-1.412
19	0.174	0.949	0.490	-0.714	0.192	0.524	-0.742	-1.381
20	0.176	1.000	0.493	-0.708	0.176	0.493	-0.679	-1.387
21	0.192	1.051	0.524	-0.647	0.174	0.490	-0.673	-1.387
22	0.205	1.103	0.547	-0.603	0.159	0.458	-0.612	-1.340
23	0.213	1.154	0.561	-0.578	0.131	0.396	-0.504	-1.248
24	0.232	1.205	0.592	-0.525	0.128	0.390	-0.494	-1.228
25	0.252	1.256	0.621	-0.476	0.124	0.381	-0.480	-1.201
26	0.271	1.308	0.649	-0.433	0.110	0.345	-0.424	-1.120
27	0.308	1.359	0.695	-0.364	0.102	0.326	-0.394	-1.030

28	0.311	1.410	0.698	-0.359	0.093	0.301	-0.358	-1.012
29	0.337	1.462	0.727	-0.318	0.091	0.295	-0.350	-0.977
30	0.362	1.513	0.753	-0.284	0.074	0.247	-0.284	-0.860
31	0.377	1.564	0.766	-0.266	0.073	0.246	-0.283	-0.859
32	0.416	1.615	0.799	-0.224	0.057	0.198	-0.221	-0.720
33	0.471	1.667	0.837	-0.177	0.053	0.184	-0.203	-0.634
34	0.522	1.718	0.867	-0.143	0.043	0.153	-0.166	-0.530
35	0.614	1.769	0.906	-0.098	0.033	0.120	-0.127	-0.399
36	0.623	1.821	0.910	-0.095	0.027	0.100	-0.106	-0.365
37	0.757	1.872	0.946	-0.055	0.022	0.081	-0.084	-0.262
38	0.858	1.923	0.964	-0.037	0.014	0.054	-0.055	-0.178
39	1.095	1.974	0.985	-0.015	0.008	0.031	-0.031	-0.091
$\Sigma$	10.108							-39.08

Nota.  $i$  = orden o rango del dato  $y = \ln[x_i/\min(x)]$  cuando se orden en sentido ascendente,  $y_{(i)}$  = valor  $y$  en el orden  $i$ ,  $(2i-1)/39$  = primer factor de  $A_i$ ,  $F_Y(y_{(i)})$  = probabilidad acumulada por el valor  $y_{(i)}$  bajo un modelo de distribución exponencial de parámetro de tasa  $\lambda = 3.858$ ,  $\ln[F_Y(y_{(i)})]$  = logaritmo de la probabilidad acumulada por el valor  $y_{(i)}$ ,  $y_{(n+1-i)}$  = valor  $y$  en el orden  $n+1-i$ ,  $F_Y[y_{(n+1-i)}]$  = probabilidad acumulada por el valor  $y_{(n+1-i)}$  bajo el mismo modelo de probabilidad,  $\ln[1-F_Y(y_{(n+1-i)})]$  = logaritmo neperiano del complemento de la probabilidad acumulada por el valor  $y_{(n+1-i)}$ ,  $A_i = ((2i-1)/39)(\ln[F_Y(y_{(i)})]+\ln[1-F_Y(y_{(n+1-i)})])$ ,  $\Sigma$  = suma por columna.

Fuente: elaboración propia.

La prueba de Anderson y Darling modificada por Sinclair, Spurr y Ahmad (1990) para una muestra con asimetría positiva se aplica con los datos originales ordenados en sentido ascendente  $x_{(i)}$ . Se usa la función de distribución acumulativa de la distribución de Pareto,  $F_X(x_{(i)}) = 1-(x_{(i)}/502.595)^{3.858}$ ; no obstante, las probabilidades son exactamente las mismas que las dadas por la función de distribución acumulativa de una distribución exponencial con parámetro de tasa  $\lambda = 3.858$ . El valor del estadístico de contraste  $AU_n^2$  es menor que el valor crítico para un tamaño de muestra de 40 y un nivel de significación del 5%, por lo que se mantiene la hipótesis nula de ajuste a una distribución de Pareto. Véase a continuación los cálculos, parte de los cuales se hallan en la Tabla 4.

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

$$AU_{40}^2 = \frac{40}{2} - 2 \times 19.636 - (-19.306) = 0.035$$

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

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

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

$$G(p = 0.05) = 0.1170 - 0.03791 \times (-2.944) + 0.06318 \times 0.955 + 0.09878 \times (-2.944) \times 0.955 + 0.009184 \times (-2.944)^2 \times 0.955 - 0.00009742 \times (-2.944)^4 \times 0.955 = 0.080$$

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

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

$$AU_{40}^2 = 0.035 < {}_{0.05}AU_{40}^2 = 0.520, \text{ se mantiene } H_0: X \sim \text{Pareto}(x_m, \alpha)$$

**Tabla 4. Cálculos para obtener el estadístico de contraste de la prueba de Anderson y Darling modificada por Sinclair Spurr y Ahmad (1990)**

<i>i</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	0.031	1.925	-0.031	-0.060
3	509.868	0.054	1.875	-0.055	-0.104
4	513.713	0.081	1.825	-0.084	-0.154
5	516.552	0.100	1.775	-0.106	-0.188
6	519.474	0.120	1.725	-0.127	-0.220
7	524.641	0.153	1.675	-0.166	-0.277
8	529.753	0.184	1.625	-0.203	-0.330
9	532.239	0.198	1.575	-0.221	-0.348
10	540.797	0.246	1.525	-0.283	-0.431
11	540.979	0.247	1.475	-0.284	-0.419
12	550.305	0.295	1.425	-0.350	-0.499
13	551.504	0.301	1.375	-0.358	-0.493
14	556.651	0.326	1.325	-0.394	-0.522
15	560.951	0.345	1.275	-0.424	-0.540
16	569.140	0.381	1.225	-0.480	-0.588
17	571.286	0.390	1.175	-0.494	-0.581
18	572.739	0.396	1.125	-0.504	-0.567
19	588.961	0.458	1.075	-0.612	-0.658
20	598.325	0.490	1.025	-0.673	-0.690

21	599.242	0.493	0.975	-0.679	-0.662
22	609.166	0.524	0.925	-0.742	-0.686
23	617.102	0.547	0.875	-0.792	-0.693
24	622.201	0.561	0.825	-0.824	-0.680
25	633.900	0.592	0.775	-0.896	-0.694
26	646.362	0.621	0.725	-0.971	-0.704
27	659.182	0.649	0.675	-1.046	-0.706
28	683.834	0.695	0.625	-1.188	-0.743
29	685.652	0.698	0.575	-1.198	-0.689
30	703.919	0.727	0.525	-1.300	-0.682
31	721.765	0.753	0.475	-1.396	-0.663
32	732.467	0.766	0.425	-1.453	-0.618
33	761.718	0.799	0.375	-1.604	-0.602
34	804.876	0.837	0.325	-1.817	-0.590
35	847.360	0.867	0.275	-2.015	-0.554
36	928.627	0.906	0.225	-2.369	-0.533
37	937.437	0.910	0.175	-2.405	-0.421
38	1071.113	0.946	0.125	-2.919	-0.365
39	1185.755	0.964	0.075	-3.312	-0.248
40	1502.032	0.985	0.025	-4.224	-0.106
$\Sigma$		19.636			-19.306

Nota.  $i$  = orden o rango del dato muestral de  $X$  cuando los 40 datos se ordenan en sentido ascendente,  $x_{(i)}$  = dato muestral de  $X$  en el orden o rango  $i$ ,  $F_X(x_{(i)})$  = probabilidad acumulada de  $x_{(i)}$  bajo una distribución de Pareto( $x_m = 502.595$ ,  $\alpha = 3.858$ ),  $2 - [(2i-1)/40]$  = primer factor de  $AU_i$ ,  $\ln[1 - F_X(x_{(i)})]$  = logaritmo neperiano del probabilidad a la cola derecha de  $x_{(i)}$  o segundo factor de  $AU_i$ ,  $AU_i = (2 - [(2i-1)/40]) \times \ln[1 - F_X(x_{(i)})]$ ,  $\Sigma$  = suma por columna.

Fuente: elaboración propia

Para obtener la gráfica cuantil-cuantil, se ordenan los datos generados en sentido ascendente  $x_{(i)}$  y se asigna a cada dato  $x_{(i)}$  su correspondiente orden ( $i$ ). Con el orden del dato se calcula el orden del cuantil teórico  $p_{(i)} = ((i+1/3)/(n+1/3))$ , usando la fórmula correspondiente a la mediana de los estadísticos de orden de una distribución beta,  $Mdn = (\alpha-1/3)/(\alpha+\beta-2/3)$ , de parámetros de forma:  $\alpha = i$  y  $\beta = n+1-i$ . Debe considerarse que el estadístico de orden  $i$  de una distribución uniforme estándar sigue dicha distribución beta. Para calcular los cuantiles teóricos, se usa la función cuantil dando como valores para los parámetros las estimaciones obtenidas (Tabla 2).

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

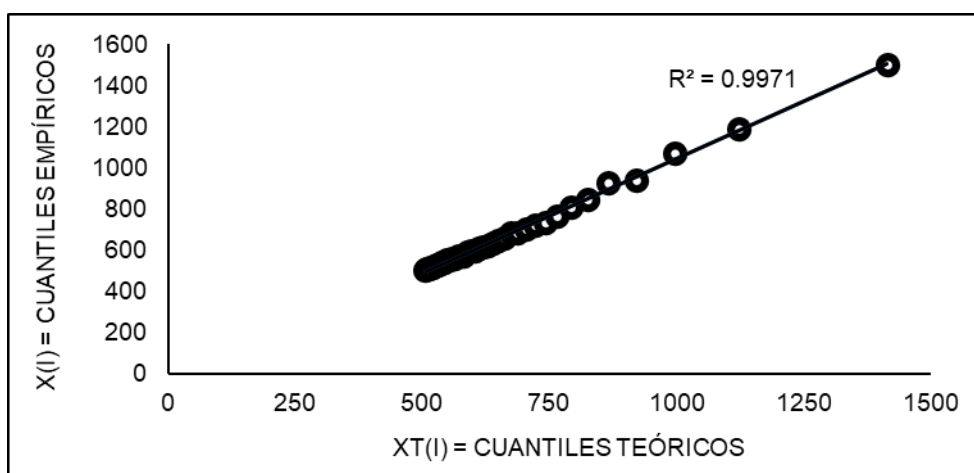
$$p_{(1)} = \frac{(i) - 1/3}{n + 1/3} = \frac{1 - 1/3}{40 + 1/3} = 0.017$$

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

$$(x_{\tau_{(1)}} \quad x_{(1)}) = (502.595 \quad 504.716)$$

En la Figura 3, aparecen los pares coordenados de cuantiles teóricos  $x_{(i)}$  y empíricos  $x_{(i)}$ . Los primeros se disponen en el eje horizontal o de abscisas porque predicen los cuantiles empíricos que se colocan en el eje vertical o de ordenadas. En la gráfica, los puntos coordenados se alinean a 45 grados, lo que indica que los datos empíricos siguen el modelo teórico. Esta alineación se puede cuantificar por el coeficiente de correlación lineal; a su vez, el cuadrado de la correlación informa de la varianza compartida. En este ejemplo, la correlación lineal es de 0.999 y la varianza compartida es del 99.7%. Adicionalmente, el histograma muestra el perfil de J invertido esperado para una muestra que sigue una distribución de Pareto (Figura 4), con lo que se puede concluir que los datos muestrales siguen un modelo de distribución de Pareto.

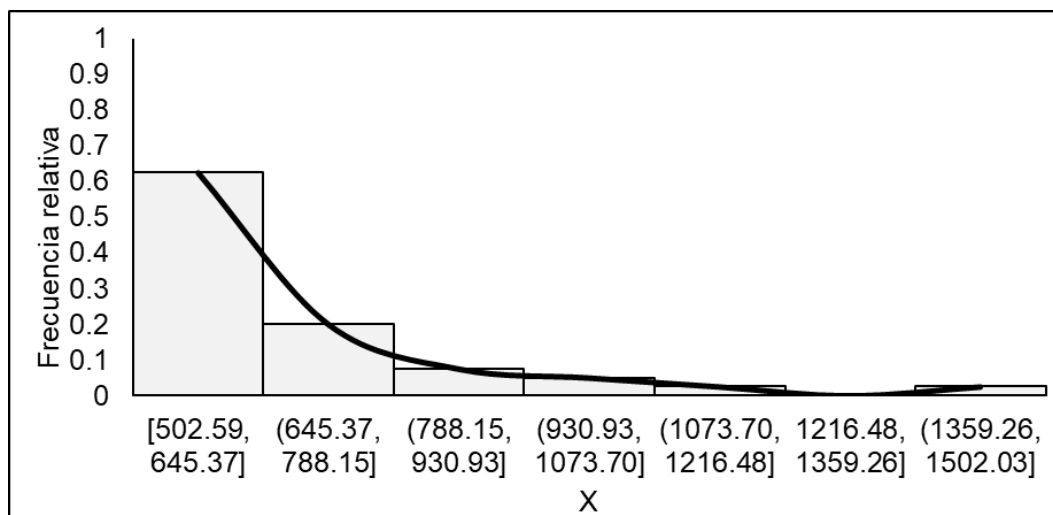
**Figura 3. Gráfica cuantil-cuantil.**



Fuente: elaboración propia.



**Figura 4. Histograma de frecuencias relativas con el polígono de frecuencias superpuesto.**



Fuente: elaboración propia.

## 7. Conclusiones

La presentación ejemplificada de la distribución de Pareto con parámetro de escala  $x_m$  y forma  $\alpha$  permite constatar que el cálculo de probabilidades es muy simple, al igual que el cómputo de las medidas descriptivas. Como en otras distribuciones, los mejores estimadores de sus parámetros son los definidos mediante el método de máxima verosimilitud (Warsono et al., 2019), los cuales son sencillos y rápidos de calcular. El estimador de  $x_m$  es el mínimo muestral y el estimador de  $\alpha$  es el inverso del promedio de los datos muestrales transformados logarítmicamente (Siudem et al., 2022). Naturalmente, se requiere contrastar el ajuste de los datos empíricos al modelo de probabilidad combinando una estrategia gráfica a través del histograma (forma de J invertida) y diagrama cuantil-cuantil (alineación a 45 grados) y una estrategia inferencial mediante las pruebas de Kolmogorov y Smirnov (Chu et al., 2019), Anderson y Darling (Stephens, 1986) o Anderson y Darling modificada (Sinclair et al., 1990). Con las dos primeras se transforman los datos para que sigan una distribución exponencial y con la tercera prueba no se requiere transformación.

Dentro de las Ciencias Sociales, esta distribución es un buen modelo de probabilidad para la distribución del ingreso, renta, recursos adquiridos acumulables en una empresa, región o país y reclamos de seguros (Feng et al., 2020), así como para la frecuencia promedio de un comportamiento raro en la mayoría de las personas, pero muy usual en unas pocas, como conductas compulsivas, adicciones comportamentales o a sustancias y parafilias (Rajeev, 2022).

Cabe señalar que se puede aplicar a los mismos datos de variables continuas que la distribución lognormal, como datos epidemiológicos (Beare y Toda, 2020), por lo que hay que determinar qué modelo de probabilidad tiene mejor ajuste (Charpentier y Flachaire, 2022; Feng et al., 2020). Cuando los datos son discretos en un conjunto acotado de números naturales, la opción es la distribución de Zipf, que es una variante de la distribución zeta. En cuanto al trasfondo social de lo expuesto, el buen ajuste de los datos a un modelo de probabilidad paretiano revela reglas de embudo y desigualdad muy definidas en reparto de recursos. A su vez, puede motivar políticas de segregación de un sector social responsable del disparo del gasto o la deuda, como ocurre hacia los adultos mayores con los seguros de salud (Granados-Martínez y Nava-Bolaños, 2019).

## 8. Bibliografía

- Ahmad, H. A. H. y Almetwally, E. M. (2020). Marshall-Olkin generalized Pareto distribution: Bayesian and non-Bayesian estimation. *Pakistan Journal of Statistics y Operation Research*, 16(1), 21–33. <http://dx.doi.org/10.18187/pjsor.v16i1.2935>
- Anderson, T. W. y Darling, D. A. (1952). Asymptotic theory of certain goodness-of-fit criteria based on stochastic processes. *Annals of Mathematical Statistics*, 23(2), 193–212. <http://dx.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 108048. <https://doi.org/10.1016/j.asoc.2021.108048>
- Aránguez-Sánchez, T. (2019). Luces y sombras del populismo. *Astrolabio: Revista Internacional de Filosofía*, 23, 39–53.
- Arnold, B. C. (2015). *Pareto distribution* (2nd ed.). New York: John Wiley y Sons, Ltd. <https://doi.org/10.1201/b18141>
- Barczy, M., Nedényi, F. K. y Sütő, L. (2023). Probability equivalent level of Value at Risk and higher-order Expected Shortfalls. *Insurance: Mathematics and Economics*, 108(1), 107–128. <https://doi.org/10.1016/j.insmatheco.2022.11.004>
- Barnoy, A. y Reich, Z. (2022). Trusting others: A pareto distribution of source and message credibility among news reporters. *Communication Research*, 49(2), 196–220. <https://doi.org/10.1177/0093650206298071>
- Beare, B. K. y Toda, A. A. (2020). On the emergence of a power law in the distribution of COVID-19 cases. *Physica D: Nonlinear Phenomena*, 412, article 132649. <https://doi.org/10.1016/j.physd.2020.132649>
- Bhoj, D. S. y Chandra, G. (2021). Ranked set sampling with lowest order statistics for Pareto distribution. *Communications in Statistics-Simulation and Computation*, 50, article 1904143. <https://doi.org/10.1080/03610918.2021.1904143>
- Campbell, M. R. y Brauer, M. (2021). Is discrimination widespread? Testing assumptions about bias on a university campus. *Journal of Experimental Psychology: General*, 150(4), 756–777. <https://doi.org/10.1037/xge0000983>
- Charpentier, A. y Flachaire, E. (2022). Pareto models for top incomes and wealth. *The Journal of Economic Inequality*, 20(1), 1–25. <https://doi.org/10.1007/s10888-021-09514-6>
- Chattamvelli, R. y Shanmugam, R. (2021). Pareto Distribution. In *Continuous Distributions in Engineering and the Applied Sciences-Part II* (pp. 179-188). Cham: Springer International Publishing. [https://doi.org/10.1007/978-3-031-02435-1\\_3](https://doi.org/10.1007/978-3-031-02435-1_3)
- Chen, B. Y., Zhang, K. Y., Wang, L. P., Jiang, S. y Liu, G. L. (2019). Generalized extreme value-pareto distribution function and its applications in ocean engineering. *China Ocean Engineering*, 33(2), 127–136. <https://doi.org/10.1007/s13344-019-0013-9>
- Chu, J., Dickin, O. y Nadarajah, S. (2019). A review of goodness of fit tests for Pareto distributions. *Journal of Computational and Applied Mathematics*, 361(17), 13–41. <https://doi.org/10.1016/j.cam.2019.04.018>
- Diawara, D., Kane, L., Dembele, S. y 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(3), 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(3), 245–299. <https://doi.org/10.6092/issn.1973-2201/9533>
- Feng, M., Deng, L. J., Chen, F., Perc, M. y 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(2237), article 2020.0019. <https://doi.org/10.1098/rspa.2020.0019>

- Gass, S. I. y Fu, M. C. (2016). Inverse transform method. In *Encyclopedia of operations research and management science* (pp. 815–815). Boston, MA: Springer. [https://doi.org/10.1007/978-1-4419-1153-7\\_200343](https://doi.org/10.1007/978-1-4419-1153-7_200343)
- Gini, C. (1936). On the measure of concentration with special reference to income and statistics. *Colorado College Publication, General Series*, 208, 73–79.
- Granados-Martínez, A. y Nava-Bolaños, I. (2019). Gastos catastróficos por motivos de salud y hogares con personas mayores en México. *Papeles de Población*, 25(99), 113–141. <https://doi.org/10.22185/24487147.2019.99.05>
- Landoni, J. S. y Villegas, L. (2022). *Pagan los pobres: Consecuencias negativas de políticas públicas con buenas (y malas) intenciones*. Buenos Aires: Editorial Galerna.
- Le Gall, P., Favre, A. C., Naveau, P. y Prieur, C. (2022). Improved regional frequency analysis of rainfall data. *Weather and Climate Extremes*, 36, article 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(268), 847–852. <http://dx.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(70), 209–219. <https://doi.org/10.2307/2276207>
- Mateus, A. y Caeiro, F. (2022). Confidence intervals for the shape parameter of a Pareto distribution. *AIP Conference Proceedings*, 2425, 320003. <https://doi.org/10.1063/5.0081541>
- McCarthy, D. M. y Winer, R. S. (2019). The Pareto rule in marketing revisited: is it 80/20 or 70/20?. *Marketing Letters*, 30(1), 139–150. <https://doi.org/10.1007/s11002-019-09490-y>
- Mojiri, A. y آحمادی, 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). *Payesh*, 21(3), 227–236. <https://doi.org/10.52547/payesh.21.3.227>
- Pareto, V. F. D. (1896). *Cours d'economie politique* (Vol. 1). Lausanne, Francia: F. Rouge éditeur.
- Pareto, V. F. D. (1897). *Cours d'economie politique* (Vol. 2). Lausanne, Francia: F. Rouge éditeur.
- Qian, W., Chen, W. y 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. y Filip, R. (2023). Quantifying optical rogue waves. *arXiv*, article 2303.04615. <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(2), 44–59.
- Rao, C. R. (1973). *Linear statistical inference and its applications*. New York: John Wiley and Sons. <https://doi.org/10.1002/9780470316436>
- Rytgaard, M. (1990). Estimation in the Pareto distribution. *ASTIN Bulletin: The Journal of the IAA*, 20(2), 201–216. <https://doi.org/10.2143/AST.20.2.2005443>
- Safari, M. A. M., Masseran, N., Ibrahim, K. y 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(3), 431–439. <https://doi.org/10.1016/j.physa.2018.11.029>
- Sarabia, J. M., Jorda, V. y 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 121277. <https://doi.org/10.1016/j.physa.2019.121277>
- Sinclair, C. D., Spurr, B. D. y Ahmad, M. I. (1990). Modified Anderson-Darling test. *Communications in Statistics - Theory and Methods*, 19(10), 3677–3686. <https://doi.org/10.1080/03610929008830405>

- Sitthiyot, T. y Holasut, K. (2021). A simple method for estimating the Lorenz curve. *Humanities and Social Sciences Communications*, 8(1), 1-9. <https://doi.org/10.1057/s41599-021-00948-x>
- Siudem, G., Nowak, P. y Gagolewski, M. (2022). Luck, reason, and the price-Pareto type-2 distribution. *Physica A: Statistical Mechanics and its Applications*, 606, article 128059. <https://doi.org/10.1016/j.physa.2022.128059>
- Song, I., Ryoung-Park, S. y Yoon, S. (2022). *Probability and random variables: theory and applications*. Cham, Switzerland: Springer International Publishing. [https://doi.org/10.1007/978-3-030-97679-8\\_3](https://doi.org/10.1007/978-3-030-97679-8_3)
- Stephens, M. A. (1974). EDF statistics for goodness of fit and some comparisons. *Journal of the American Statistical Association*, 69(347), 730–737. <https://doi.org/10.2307/2286009>
- Stephens, M. A. (1986). Test based on EDF statistics. En R. B. D'Agostino y M. A. Stephens (Eds.), *Goodness-of-fit techniques* (pp. 97–193) New York: Marcel Dekker, Inc. <https://doi.org/10.1201/9780203753064-4>
- Sudharson, D. y Prabha, D. (2019). A novel machine learning approach for software reliability growth modelling with pareto distribution function. *Soft Computing*, 23(18), 8379–8387. <https://doi.org/10.1007/s00500-019-04047-7>
- Tokhirov, A. (2021). Remittances and inequality: the post-communist region. *Prague Economic Papers*, 30(4), 426–448. <https://doi.org/10.18267/j.pep.776>
- Warsono, Gustavia, E., Kurniasari, D., Amanto y Antonio, Y. (2019). On the comparison of the methods of parameter estimation for Pareto distribution. *Journal of Physics: Conference Series*, 1338(1) article 012042. <https://dx.doi.org/10.1088/1742-6596/1338/1/012042>
- World Bank (2022). *Gini index*. Recuperado de: <https://data.worldbank.org/indicator/>
- Xiong, B., Fu, H., Xu, F. y Jin, Y. (2019). Data-driven discovery of partial differential equations for multiple-physics electromagnetic problem. *arXiv*, article 1910.13531. <https://doi.org/10.48550/arXiv.1910.13531>
- Xu, T., Sedory, S. A. y Singh, S. (2022). Lowering the Cramer-Rao lower bounds of variance in randomized response sampling. *Communications in Statistics-Simulation and Computation*, 51(7), 4112–4126. <https://doi.org/10.1080/03610918.2020.1737874>
- Yang, X. y Zhou, P. (2022). Wealth inequality and social mobility: A simulation-based modelling approach. *Journal of Economic Behavior y Organization*, 196, 307–329. <https://doi.org/10.1016/j.jebo.2022.02.012>
- Zhang, Y., Wu, Y. y Yao, H. (2022). Optimal health insurance with constraints under utility of health, wealth and income. *Journal of Industrial and Management Optimization*, 18(3), 1519–1540. <https://doi.org/10.3934/jimo.2021031>

**Autor.**

**José Moral de la Rubia**

Facultad de Psicología de la Universidad Autónoma de Nuevo León, México.

Doctor en Filosofía y Ciencias de la Educación. Sección de Psicología por la Universidad de Alcalá Henares, Madrid. Psicólogo Especialista en Psicología Clínica vía PIR, Madrid. Licenciatura en Filosofía y Ciencias de la Educación. Sección de Psicología por la Universidad Pontificia Comillas, Madrid.

E-mail: jose.morald@uanl.edu.mx

**Citado.**

MORAL DE LA RUBIA, José (2025). Distribución de Pareto: un modelo de probabilidad en investigación social. *Revista Latinoamericana de Metodología de la Investigación Social – ReLMIS*, N°29, Año 15, pp. 22-50.

**Plazos.**

Recibido: 01/04/2023. Aceptado: 07/11/2023.