

Pareto distribution

From Wikipedia, the free encyclopedia

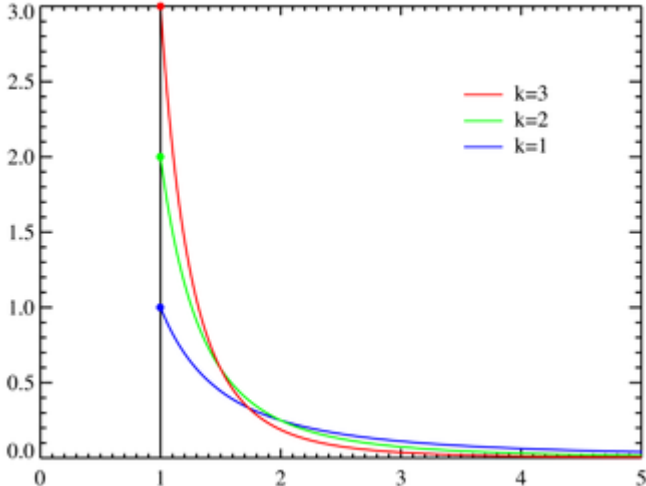
The **Pareto distribution**, named after the Italian economist Vilfredo Pareto, is a power law probability distribution that coincides with social, scientific, geophysical, actuarial, and many other types of observable phenomena. Outside the field of economics it is sometimes referred to as the **Bradford distribution**.

Contents

- 1 Properties
 - 1.1 Definition
 - 1.2 Density function
 - 1.3 Moments and characteristic function
 - 1.4 Degenerate case
 - 1.5 Conditional distributions
 - 1.6 A characterization theorem
- 2 Applications
- 3 Relation to other distributions
 - 3.1 Relation to the log-normal distribution
 - 3.2 Relation to the exponential distribution
 - 3.3 Relation to Zipf's law
 - 3.4 Relation to the "Pareto principle"

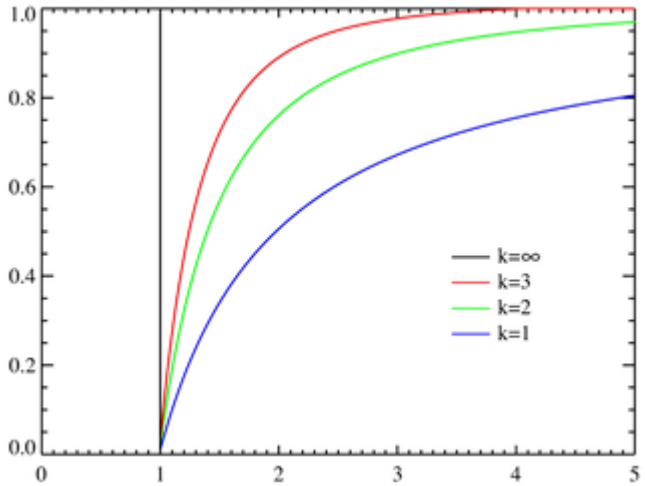
Pareto

Probability density function



Pareto probability density functions for various α with $x_m = 1$. The horizontal axis is the x parameter. As $\alpha \rightarrow \infty$ the distribution approaches $\delta(x - x_m)$ where δ is the Dirac delta function.

Cumulative distribution function



Pareto cumulative distribution functions for various α with $x_m = 1$. The horizontal axis is the x parameter.

parameters:	$x_m > 0$ scale (real) $\alpha > 0$ shape (real)
support:	$x \in [x_m; +\infty)$
pdf:	$\frac{\alpha x_m^\alpha}{x^{\alpha+1}}$ for $x > x_m$
cdf:	$1 - \left(\frac{x_m}{x}\right)^\alpha$

- 4 Pareto, Lorenz, and Gini
- 5 Parameter estimation
- 6 Graphical representation
- 7 Generating a random sample from Pareto distribution
- 8 Bounded Pareto distribution
- 9 Generating bounded Pareto random variables
- 10 Generalized Pareto distribution
- 11 Generating generalized Pareto random variables
- 12 Symmetric Pareto distribution
- 13 See also
- 14 Notes
- 15 References
- 16 External links

mean:	$\frac{\alpha x_m}{\alpha - 1}$ for $\alpha > 1$
median:	$x_m \sqrt[\alpha]{2}$
mode:	x_m
variance:	$\frac{x_m^2 \alpha}{(\alpha - 1)^2 (\alpha - 2)}$ for $\alpha > 2$
skewness:	$\frac{2(1 + \alpha)}{\alpha - 3} \sqrt{\frac{\alpha - 2}{\alpha}}$ for $\alpha > 3$
ex.kurtosis:	$\frac{6(\alpha^3 + \alpha^2 - 6\alpha - 2)}{\alpha(\alpha - 3)(\alpha - 4)}$ for $\alpha > 4$
entropy:	$\ln \left(\frac{\alpha}{x_m} \right) - \frac{1}{\alpha} - 1$
mgf:	$\alpha (-x_m t)^{\alpha} \Gamma(-\alpha, -x_m t)$ for $t < 0$
cf:	$\alpha (-i x_m t)^{\alpha} \Gamma(-\alpha, -i x_m t)$
Fisher information:	$\begin{pmatrix} \frac{\alpha}{x_m^2} & -\frac{1}{x_m} \\ -\frac{1}{x_m} & \frac{1}{\alpha^2} \end{pmatrix}$

Properties

Definition

If X is a random variable with a Pareto (Type I) distribution,^[1] then the probability that X is greater than some number x is given by

$$\Pr(X > x) = \begin{cases} \left(\frac{x_m}{x}\right)^{\alpha} & \text{for } x \geq x_m, \\ 1 & \text{for } x < x_m. \end{cases}$$

where x_m is the (necessarily positive) minimum possible value of X , and α is a positive parameter. The family of Pareto distributions is parameterized by two quantities, x_m and α . When this distribution is used to model the distribution of wealth, then the parameter α is called the Pareto index.

It follows from the above that therefore the cumulative distribution function of a Pareto random variable with parameters α and x_m is

$$F_X(x) = \begin{cases} 1 - \left(\frac{x_m}{x}\right)^\alpha & \text{for } x \geq x_m, \\ 0 & \text{for } x < x_m. \end{cases}$$

Density function

It follows (by differentiation) that the probability density function is

$$f_X(x) = \begin{cases} \alpha \frac{x_m^\alpha}{x^{\alpha+1}} & \text{for } x > x_m, \\ 0 & \text{for } x < x_m. \end{cases}$$

Moments and characteristic function

- The expected value of a random variable following a Pareto distribution with $\alpha > 1$ is

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

(if $\alpha \leq 1$, the expected value does not exist).

- The variance is

$$\text{var}(X) = \left(\frac{x_m}{\alpha - 1}\right)^2 \frac{\alpha}{\alpha - 2}.$$

(If $\alpha \leq 2$, the variance does not exist).

- The raw moments are

$$\mu'_n = \frac{\alpha x_m^n}{\alpha - n},$$

but the n th moment exists only for $n < \alpha$.

- The moment generating function is only defined for non-positive values $t \leq 0$ as

$$M(t, \alpha, x_m) = E(e^{tX}) = \alpha(-x_m t)^\alpha \Gamma(-\alpha, -x_m t) \text{ and } M(0, \alpha, x_m) = 1.$$

- The characteristic function is given by

$$\varphi(t; \alpha, x_m) = \alpha(-ix_m t)^\alpha \Gamma(-\alpha, -ix_m t),$$

where $\Gamma(a, x)$ is the incomplete gamma function.

Degenerate case

The Dirac delta function is a limiting case of the Pareto density:

$$\lim_{\alpha \rightarrow \infty} f(x; \alpha, x_m) = \delta(x - x_m).$$

Conditional distributions

The conditional probability distribution of a Pareto-distributed random variable, given the event that it is greater than or equal to a particular number x_1 exceeding x_m , is a Pareto distribution with the same Pareto index α but with minimum x_1 instead of x_m .

A characterization theorem

Suppose X_i , $i = 1, 2, 3, \dots$ are independent identically distributed random variables whose probability distribution is supported on the interval $[x_m, \infty)$ for some $x_m > 0$. Suppose that for all n , the two random variables $\min\{X_1, \dots, X_n\}$ and $(X_1 + \dots + X_n)/\min\{X_1, \dots, X_n\}$ are independent. Then the common distribution is a Pareto distribution.

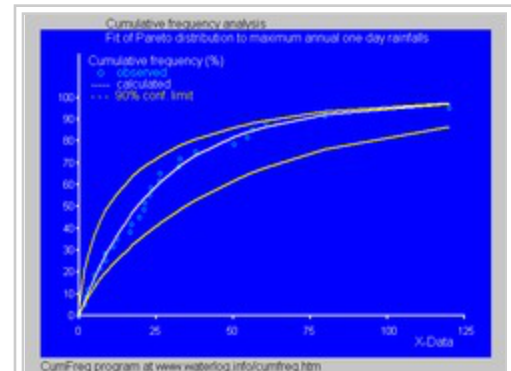
Applications

Pareto originally used this distribution to describe the allocation of wealth among individuals since it seemed to show rather well the way that a larger portion of the wealth of any society is owned by a smaller percentage of the people in that society. He also used it to describe distribution of income.^[2] This idea is sometimes expressed more simply as the Pareto principle or the "80-20 rule" which says that 20% of the population controls 80% of the wealth.^[3] The probability density function (PDF) graph on the right shows that the "probability" or fraction of the population that owns a small amount of wealth per person is rather high, and then decreases steadily as wealth increases. This distribution is not limited to describing wealth or income, but to many situations in which an equilibrium is found in the distribution of the "small" to the "large". The following examples are sometimes seen as approximately Pareto-distributed:

- The sizes of human settlements (few cities, many hamlets/villages)
[citation needed]
- File size distribution of Internet traffic which uses the TCP protocol (many smaller files, few larger ones)^{*[citation needed]*}
- Hard disk drive error rates^[4]
- Clusters of Bose–Einstein condensate near absolute zero^{*[citation needed]*}
- The values of oil reserves in oil fields (a few large fields, many small fields)
[citation needed]

- The length distribution in jobs assigned supercomputers (a few large ones, many small ones)^[citation needed]
- The standardized price returns on individual stocks^[citation needed]

- Sizes of sand particles
- Sizes of meteorites
- Numbers of species per genus (There is subjectivity involved: The tendency to divide a genus into two or more increases with the number of species in it)^[citation needed]
- Areas burnt in forest fires
- Severity of large casualty losses for certain lines of business such as general liability, commercial auto, and workers compensation.^{[6][7]}
- In hydrology the Pareto distribution is applied to extreme events such as annually maximum one-day rainfalls and river discharges.



Fitted cumulative Pareto distribution to maximum one-day rainfalls using CumFreq ^[5]

The blue picture illustrates an example of fitting the Pareto distribution to ranked annually maximum one-day rainfalls showing also the 90% confidence belt based on the binomial distribution. The rainfall data are represented by plotting positions as part of the cumulative frequency analysis.

Relation to other distributions

Relation to the log-normal distribution

Note that the Pareto distribution and log-normal distribution are alternative distributions for describing the same types of quantities. One of the connections between the two is that they are both the distributions of the exponential of random variables distributed according to other common distributions, respectively the exponential distribution and normal distribution. (Both of these latter two distributions are "basic" in the sense that the logarithms of their density functions are linear and quadratic, respectively, functions of the observed values.)^[citation needed]

Relation to the exponential distribution

The Pareto distribution is related to the exponential distribution as follows. If X is Pareto-distributed with minimum x_m and index α , then

$$Y = \log \left(\frac{X}{x_m} \right).$$

is exponentially distributed with intensity α . Equivalently, if Y is exponentially distributed with intensity α , then

$$x_m e^Y$$

is Pareto-distributed with minimum x_m and index α .

Relation to Zipf's law

Pareto distributions are continuous probability distributions. Zipf's law, also sometimes called the zeta distribution, may be thought of as a discrete counterpart of the Pareto distribution.

Relation to the "Pareto principle"

The "80-20 law", according to which 20% of all people receive 80% of all income, and 20% of the most affluent 20% receive 80% of that 80%, and so on, holds precisely when the Pareto index is $\alpha = \log_4 5$. Moreover, the following have been shown^[8] to be mathematically equivalent:

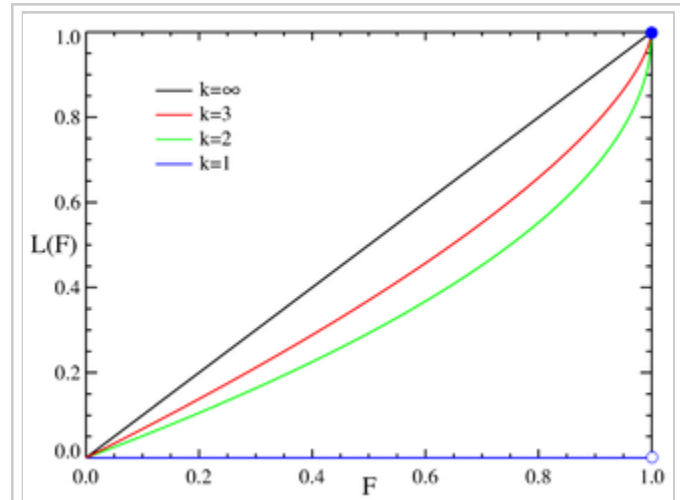
- Income is distributed according to a Pareto distribution with index $\alpha > 1$.
- There is some number $0 \leq p \leq 1/2$ such that $100p\%$ of all people receive $100(1 - p)\%$ of all income, and similarly for every real (not necessarily integer) $n > 0$, $100p^n\%$ of all people receive $100(1 - p)^n\%$ of all income.

This does not apply only to income, but also to wealth, or to anything else that can be modeled by this distribution.

This excludes Pareto distributions in which $0 < \alpha \leq 1$, which, as noted above, have infinite expected value, and so cannot reasonably model income distribution.

Pareto, Lorenz, and Gini

The Lorenz curve is often used to characterize income and wealth distributions. For any distribution, the Lorenz curve $L(F)$ is written in terms of the PDF f or the CDF F as



Lorenz curves for a number of Pareto distributions. The case $\alpha = \infty$ corresponds to perfectly equal distribution ($G = 0$) and the $\alpha = 1$ line corresponds to complete inequality ($G = 1$)

$$L(F) = \frac{\int_{x_m}^{x(F)} x f(x) dx}{\int_{x_m}^{\infty} x f(x) dx} = \frac{\int_0^F x(F') dF'}{\int_0^1 x(F') dF'}$$

where $x(F)$ is the inverse of the CDF. For the Pareto distribution,

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

and the Lorenz curve is calculated to be

$$L(F) = 1 - (1 - F)^{1-1/\alpha},$$

where α must be greater than or equal to unity, since the denominator in the expression for $L(F)$ is just the mean value of x . Examples of the Lorenz curve for a number of Pareto distributions are shown in the graph on the right.

The Gini coefficient is a measure of the deviation of the Lorenz curve from the equidistribution line which is a line connecting $[0, 0]$ and $[1, 1]$, which is shown in black ($\alpha = \infty$) in the Lorenz plot on the right. Specifically, the Gini coefficient is twice the area between the Lorenz curve and the equidistribution line. The Gini coefficient for the Pareto distribution is then calculated to be

$$G = 1 - 2 \int_0^1 L(F) dF = \frac{1}{2\alpha - 1}$$

(see Aaberge 2005).

Parameter estimation

The likelihood function for the Pareto distribution parameters α and x_m , given a sample $x = (x_1, x_2, \dots, x_n)$, is

$$L(\alpha, x_m) = \prod_{i=1}^n \alpha \frac{x_m^\alpha}{x_i^{\alpha+1}} = \alpha^n x_m^{n\alpha} \prod_{i=1}^n \frac{1}{x_i^{\alpha+1}}.$$

Therefore, the logarithmic likelihood function is

$$\ell(\alpha, x_m) = n \ln \alpha + n\alpha \ln x_m - (\alpha + 1) \sum_{i=1}^n \ln x_i.$$

It can be seen that $\ell(\alpha, x_m)$ is monotonically increasing with x_m , that is, the greater the value of x_m , the greater the value of the likelihood function. Hence, since $x \geq x_m$, we conclude that

$$\hat{x}_m = \min_i x_i.$$

To find the estimator for α , we compute the corresponding partial derivative and determine where it is zero:

$$\frac{\partial \ell}{\partial \alpha} = \frac{n}{\alpha} + n \ln x_m - \sum_{i=1}^n \ln x_i = 0.$$

Thus the maximum likelihood estimator for α is:

$$\hat{\alpha} = \frac{n}{\sum_i (\ln x_i - \ln \hat{x}_m)}.$$

The expected statistical error is:

$$\sigma = \frac{\hat{\alpha}}{\sqrt{n}}.^{[9]}$$

Graphical representation

The characteristic curved 'long tail' distribution when plotted on a linear scale, masks the underlying simplicity of the function when plotted on a log-log graph, which then takes the form of a straight line with negative gradient.^[*citation needed*]

Generating a random sample from Pareto distribution

Random samples can be generated using inverse transform sampling. Given a random variate U drawn from the uniform distribution on the unit interval $(0, 1)$, the variate T given by

$$T = \frac{x_m}{U^{1/\alpha}}$$

is Pareto-distributed.^[*citation needed*]

Bounded Pareto distribution

The **bounded Pareto distribution** has three parameters α , L and H . As in the standard Pareto distribution α determines the shape. L denotes the minimal value, and H denotes the maximal value. (The Variance in the table on the right should be interpreted as 2nd Moment).

The probability density function is

$$\frac{\alpha L^\alpha x^{-\alpha-1}}{1 - \left(\frac{L}{H}\right)^\alpha}$$

where $L \leq x \leq H$, and $\alpha > 0$.

Bounded Pareto	
parameters:	$L > 0$ location (real) $H > L$ location (real) $\alpha > 0$ shape (real)
support:	$L \leq x \leq H$
pdf:	$\frac{\alpha L^\alpha x^{-\alpha-1}}{1 - \left(\frac{L}{H}\right)^\alpha}$
cdf:	$\frac{1 - L^\alpha x^{-\alpha}}{1 - \left(\frac{L}{H}\right)^\alpha}$
mean:	$\frac{L^\alpha}{1 - \left(\frac{L}{H}\right)^\alpha} \cdot \left(\frac{\alpha}{\alpha-1}\right) \cdot \left(\frac{1}{L^{\alpha-1}} - \frac{1}{H^{\alpha-1}}\right), \alpha > 1$
median:	$L \left(1 - \frac{1}{2} \left(1 - \left(\frac{L}{H}\right)^\alpha\right)\right)^{-\frac{1}{\alpha}}$
mode:	
variance:	$\frac{L^\alpha}{1 - \left(\frac{L}{H}\right)^\alpha} \cdot \left(\frac{\alpha}{\alpha-2}\right) \cdot \left(\frac{1}{L^{\alpha-2}} - \frac{1}{H^{\alpha-2}}\right), \alpha > 2$
skewness:	
ex.kurtosis:	
entropy:	
mgf:	

cf:

Generating bounded Pareto random variables

If U is uniformly distributed on $(0, 1)$, then

$$\left(- \left(\frac{UH^\alpha - UL^\alpha - H^\alpha}{H^\alpha L^\alpha} \right) \right)^{-1/\alpha}$$

is bounded Pareto-distributed.^[*citation needed*]

Generalized Pareto distribution

The family of **generalized Pareto distributions (GPD)** has three parameters μ, σ and ξ .

The cumulative distribution function is

Generalized Pareto

parameters:	$\mu \in (-\infty, \infty)$ location (real) $\sigma \in (0, \infty)$ scale (real) $\xi \in (-\infty, \infty)$ shape (real)
support:	$x \geq \mu \ (\xi \geq 0)$ $\mu \leq x \leq \mu - \sigma/\xi \ (\xi < 0)$
pdf:	$\frac{1}{\sigma} (1 + \xi z)^{-(1/\xi+1)}$ where $z = \frac{x - \mu}{\sigma}$
cdf:	$1 - (1 + \xi z)^{-1/\xi}$
mean:	$\mu + \frac{\sigma}{1 - \xi} \ (\xi < 1)$
median:	$\mu + \frac{\sigma(2^\xi - 1)}{\xi}$
mode:	
variance:	$\frac{\sigma^2}{(1 - \xi)^2(1 - 2\xi)} \ (\xi < 1/2)$
skewness:	
ex.kurtosis:	
entropy:	
mgf:	
cf:	

$$F_{(\xi, \mu, \sigma)}(x) = \begin{cases} 1 - \left(1 + \frac{\xi(x-\mu)}{\sigma}\right)^{-1/\xi} & \text{for } \xi \neq 0, \\ 1 - \exp\left(-\frac{x-\mu}{\sigma}\right) & \text{for } \xi = 0. \end{cases}$$

for $x \geq \mu$ when $\xi \geq 0$, and $x \leq \mu - \sigma/\xi$ when $\xi < 0$, where $\mu \in \mathbb{R}$ is the location parameter, $\sigma > 0$ the scale parameter and $\xi \in \mathbb{R}$ the shape parameter. Note that some references give the "shape parameter" as $\kappa = -\xi$.

The probability density function is:

$$f_{(\xi, \mu, \sigma)}(x) = \frac{1}{\sigma} \left(1 + \frac{\xi(x-\mu)}{\sigma}\right)^{(-\frac{1}{\xi}-1)}.$$

or

$$f_{(\xi, \mu, \sigma)}(x) = \frac{\sigma^{\frac{1}{\xi}}}{(\sigma + \xi(x-\mu))^{\frac{1}{\xi}+1}}.$$

again, for $x \geq \mu$, and $x \leq \mu - \sigma/\xi$ when $\xi < 0$.

Generating generalized Pareto random variables

If U is uniformly distributed on $(0, 1]$, then

$$X = \mu + \frac{\sigma(U^{-\xi} - 1)}{\xi} \sim \text{GPD}(\mu, \sigma, \xi).$$

In Matlab Statistics Toolbox, you can easily use "gprnd" command to generate generalized Pareto random numbers.

With GNU R you can use the packages POT or evd with the "rgpd" command (see for exact usage: <http://rss.acs.unt.edu/Rdoc/library/POT/html/simGPD.html>)

Symmetric Pareto distribution

The symmetric Pareto distribution can be defined by the probability density function:^[10]

$$f(x; \alpha, x_m) = \begin{cases} (\alpha x_m^\alpha / 2) |x|^{-\alpha-1} & \text{for } |x| > x_m \\ 0 & \text{otherwise.} \end{cases}$$

The density is zero for x between $-x_m$ and x_m . It has a similar shape to a Pareto distribution for $x > x_m$ while looking like an inverted Pareto distribution for $x < x_m$

See also

- Cumulative frequency analysis
- Pareto analysis
- Pareto efficiency
- Pareto interpolation
- Pareto principle
- The Long Tail
- Traffic generation model

Notes

1. ^ See Arnold (1983).
2. ^ Pareto, Vilfredo, *Cours d'Économie Politique: Nouvelle édition par G.-H. Bousquet et G. Busino*, Librairie Droz, Geneva, 1964, pages 299-345.
3. ^ For a two-quantile population, where 18% of the population owns 82% of the wealth, the Theil index takes the value 1.
4. ^ Schroeder, Bianca; Damouras, Sotirios; Gill, Phillipa (2010-02-24), "Understanding latent sector error and how to protect against them" (http://www.usenix.org/event/fast10/tech/full_papers/schroeder.pdf) , *8th Usenix Conference on File and Storage Technologies (FAST 2010)*, http://www.usenix.org/event/fast10/tech/full_papers/schroeder.pdf, retrieved 2010-09-10, "We experimented with 5 different distributions (Geometric, Weibull, Rayleigh, Pareto, and Lognormal), that are commonly used in the context of system reliability, and evaluated their fit through the total squared differences between the actual and hypothesized frequencies (χ^2 statistic). We found consistently across all models that the geometric distribution is a poor fit, while the Pareto distribution provides the best fit."
5. ^ "Cumfreq, a free computer program for cumulative frequency analysis" (<http://www.waterlog.info/cumfreq.htm>) . <http://www.waterlog.info/cumfreq.htm>.
6. ^ Kleiber and Kotz (2003): page 94.
7. ^ Seal, H. (1980). Survival probabilities based on Pareto claim distributions. *ASTIN Bulletin*, 11, 61-71.
8. ^ Michael Hardy (2010) "Pareto's Law", *Mathematical Intelligencer*, 32 (3), 38-43. doi: 10.1007/s00283-010-9159-2 (<http://dx.doi.org/10.1007%2Fs00283-010-9159-2>)
9. ^ M. E. J. Newman (2005). "Power laws, Pareto distributions and Zipf's law" (<http://arxiv.org/abs/cond-mat/0412004v3>) . *Contemporary Physics* **46**: 323-351. doi:10.1080/00107510500052444 (<http://dx.doi.org/10.1080%2F00107510500052444>) . <http://arxiv.org/abs/cond-mat/0412004v3>.
10. ^ "Do Financial Returns Have Finite or Infinite Variance? A Paradox and an

Explanation" (<http://people.orie.cornell.edu/~gennady/techreports/RetTailParadoxExplFinal.pdf>) . pp. 7-8. <http://people.orie.cornell.edu/~gennady/techreports/RetTailParadoxExplFinal.pdf>.

References

- Barry C. Arnold (1983). *Pareto Distributions*, International Co-operative Publishing House, Burtonsville, Maryland. ISBN 0-899974-012-1.
- Christian Kleiber and Samuel Kotz (2003). *Statistical Size Distributions in Economics and Actuarial Sciences*, New York:Wiley. xi+332 pp. ISBN 0-471-15064-9.
- Lorenz, M. O. (1905). "Methods of measuring the concentration of wealth". Publications of the American Statistical Association. 9: 209–219.

External links

- The Pareto, Zipf and other power laws / William J. Reed - PDF (http://linkage.rockefeller.edu/wli/zipf/reed01_el.pdf)
- Gini's Nuclear Family / Rolf Aabergé. - In: International Conference to Honor Two Eminent Social Scientists (<http://www.unisi.it/eventi/GiniLorenz05/>) , May, 2005 - PDF (http://www.unisi.it/eventi/GiniLorenz05/25%20may%20paper/PAPER_Aaberge.pdf)
- syntraf1.c (<http://www.csee.usf.edu/~christen/tools/syntraf1.c>) is a C program to generate synthetic packet traffic with bounded Pareto burst size and exponential interburst time.

Retrieved from "http://en.wikipedia.org/wiki/Pareto_distribution"

Categories: Actuarial science | Continuous distributions | Power laws | Socioeconomics

- This page was last modified on 3 March 2011 at 05:19.
- Text is available under the Creative Commons Attribution-ShareAlike License; additional terms may apply. See Terms of Use for details. Wikipedia® is a registered trademark of the Wikimedia Foundation, Inc., a non-profit organization.