1. Pareto Distribution

Welcome to the series of E-learning modules on Pareto Distribution. Here, we will discuss about the distribution function, mean, variance, raw moments, skewness and kurtosis, moment generating function, median, mode, method of obtaining conditional distribution, uses, applications, and examples of the distribution.

By the end of this session, you will be able to:

- Explain the Pareto distribution
- Explain the distribution function
- Explain the mean and variance
- Explain the raw moments, skewness and kurtosis
- Explain the moment generating function
- Explain the median and mode
- Explain the uses, applications and examples

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.

The Pareto distribution is a skewed, heavy-tailed distribution that is sometimes used to model the distribution of incomes and other financial variables. As the distribution is heavy-tailed, it can have large values, which makes it a tool for analyzing extreme values.

A continuous random variable X is said to have Pareto distribution with parameters alpha and b if its probability density function is given by,
f of x is equal to alpha into b power alpha divided by x power alpha plus 1, where x is greater than or equal to b and alpha is greater than b.
Here, alpha is called shape parameter and b is called scale parameter.
Since f(x) is a probability density function, we can write
Integral from b to infinity, f of x dx is equal to 1 implies alpha into b power alpha into integral from b to infinity x power minus alpha minus 1 dx is equal to 1
Implies integral from b to infinity dx is equal to 1 divided by alpha into b power alpha
Let us define incomplete gamma integral,
Integral from beta to infinity e power minus x into x power alpha minus 1 dx
Is equal to gamma alpha, beta,
Is equal to alpha minus 1 factorial into e power minus beta into summation over k is equal to zero to alpha minus 1 beta power k divided by k factorial

2. Distribution Function, Mean and Variance

Let us obtain the distribution function of the Pareto distribution.
Distribution function is given by,
F of x is equal to integral from b to x f of x dx
Is equal to integral from b to x, alpha into b power alpha divided by x power alpha plus 1 dx
Is equal to alpha into b power alpha into integral from b to x, x power minus alpha minus 1 dx
Is equal to alpha into b power alpha into x power minus alpha divided by minus alpha, ranges from b to x
Is equal to minus b power alpha into x power minus alpha minus b power minus alpha
Is equal to 1 minus b power alpha divided by x power alpha, for x greater than or equal to b

Now, let us find mean of Pareto distribution.
Mean of the distribution is given by mu 1 dash is equal to expectation of X is equal to integral from b to infinity x into f of x dx
Is equal to alpha into b power alpha into integral from b to infinity x into x power minus alpha minus 1 dx
Is equal to alpha into b power alpha into 1 divided by alpha minus 1 into b power alpha minus 1
Is equal to alpha b divided by alpha minus 1

Variance of the Pareto distribution is given by mu 2 is equal to mu 2 dash minus mu 1 dash whole square. Hence, first let us find mu 2 dash is equal to expectation of X square
Is equal to integral from b to infinity x square into f of x dx
Is equal to alpha into b power alpha into integral from b to infinity x power minus of alpha minus 2 minus 1 dx
Is equal to alpha into b power alpha into 1 divided by alpha minus 2 into b power alpha minus 2
Is equal to alpha into b square divided by alpha minus 2

Therefore, variance
mu 2 is equal to mu 2 dash minus mu 1 dash square
Is equal to alpha into b square divided by alpha minus 2 minus alpha into b divided by alpha minus 1 the whole square
Is equal to alpha into b square into 1 divided by alpha minus 2 minus alpha divided by alpha minus 1 the whole square
By taking common denominator we get, alpha into b square into alpha minus 1 whole square minus alpha into alpha minus 2 whole divided by alpha minus 1 square into alpha minus 2
On simplification, we get,
Alpha into b square divided by alpha minus 1 square into alpha minus 2, for alpha greater than 2. If alpha is 1 or 2, the variance is infinity.

3. Raw Moments, Skewness and Kurtosis

Now, let us obtain r^{th} raw moment of Pareto distribution.
r^{th} raw moment mu r dash of Pareto distribution is given by,
Mu r dash is equal to expectation of x power r
Is equal to integral from b to infinity x power r into f of x dx
Is equal to alpha into b power alpha into integral from b to infinity x power r into x power minus alpha minus 1 dx
It can be written as, alpha into b power alpha integral from b to infinity x power minus of alpha minus r, minus 1 dx
Is equal to alpha into b power alpha into 1 divided by alpha minus r into b power alpha minus r
On simplifying we get, alpha into b power r divided by alpha minus r.

By substituting r is equal to 1, we get,
Mean mu 1 dash is equal to alpha into b divided by alpha minus 1
Substituting r is equal to 2, we get, mu 2 dash is equal to alpha into b square divided by alpha minus 2.
Hence, variance mu 2 is equal to mu 2 dash minus mu 1 dash square
Is equal to alpha into b square divided by alpha minus 2 minus alpha into b divided by alpha minus 1 the whole square
On simplification we get,
Alpha into b square divided by alpha minus 1 square into alpha minus 2, for alpha greater than 2. If alpha is 1 or 2, the variance is infinity.

By substituting r is equal to 3, we get mu 3 dash is equal alpha into b cube divided by alpha minus 3 and
r is equal to 4, we get, mu 4 dash is equal to alpha into b power 4 divided by alpha minus 4
Hence, Mu 3 is equal to mu 3 dash minus 3 into mu 2 dash into mu 1 dash plus 2 into mu 1 dash cube
Is equal to alpha into b cube divided by alpha minus 3 minus 3 into alpha into b square divided by alpha minus 2 into alpha into beta divided by alpha minus 1 plus 2 into alpha into b divided by alpha minus 1 the whole cube
On simplifying the above expression, we get
2 into alpha into alpha plus 1 into b cube divided by alpha minus 1 whole cube into alpha minus 2 into alpha minus 3, for alpha greater than 3.

Mu 4 is equal to mu 4 dash minus 4 into mu 3 dash into mu 1 dash plus 6 into mu 2 dash into mu 1 dash square minus 3 into mu 1 dash power 4
Is equal to alpha into b power 4 divided by alpha minus 4 minus 4 into alpha into b cube divided by alpha minus 3 into alpha into b divided by alpha minus 1 plus 6 into alpha into b square divided by alpha minus 2 into alpha into b divided by alpha minus 1 the whole square minus 3 into alpha into b divided by alpha minus 1 whole power 4.
On simplifying the above expression, we get, 3 into alpha into 3 alpha cube plus alpha plus 3

into b power 4 divided by alpha minus 1 power 4 into alpha minus 2 into alpha minus 3 into alpha minus 4, for alpha greater than 4.

Coefficient of skewness is given by, beta 1 is equal to mu 3 square divided by mu 2 cube
Is equal to 2 into alpha into alpha plus 1 into b cube divided by alpha minus 1 cube into alpha minus 2 into alpha minus 3 the whole square divided by alpha into b square divided by alpha minus 1 square into alpha minus 2 the whole cube
On simplifying, we get, 4 into alpha plus 1 square into alpha minus 2 divided by alpha into alpha minus 3 square.
For alpha greater than 3, Pareto distribution is positively skewed.

Now, let us find the kurtosis of the distribution. Coefficient of kurtosis is given by,
Beta 2 is equal to mu 4 divided by mu 2 square
Is equal to 3 into alpha into 3 alpha cube plus alpha plus 3 into b power 4 divided by alpha minus 1 power 4 into alpha minus 2 into alpha minus 3 into alpha minus 4 whole divided by alpha into b square divided by alpha minus 1 square into alpha minus 2 the whole square
Is equal to 3 into 3 alpha cube plus alpha plus 3 into alpha minus 2 divided by alpha into alpha minus 3 into alpha minus 4.
For alpha greater than 4, the Pareto distribution has leptokurtic curve.

4. MGF, Median and Mode

Moment generating function of Pareto distribution is given by,
Mx of t is equal to expectation of e power t into X
Is equal to integral from b to infinity e power t into x f of x dx
Is equal to alpha into b power alpha into integral from b to infinity e power t into x into x power minus alpha minus 1 dx
Is equal alpha into b power alpha into integral from b to infinity e power minus x into minus t into x power minus alpha minus 1 dx
Using, the incomplete gamma function, we get,
alpha into b power alpha into gamma minus alpha, minus b into t divided by minus t whole power minus alpha
Is equal to alpha into minus t into b whole power alpha into gamma minus alpha, minus b into t.

Now, let us obtain median of the distribution.
If M is the median of the Pareto distribution, then
Integral from b to M f of x dx is equal to half is equal to integral from M to infinity f of x dx
We can consider any one of the two integrals. Let us consider the first integral and solve for M
Integral from b to M f of x dx is equal to half
Substituting for f of x we get,
Alpha into b power alpha into integral from b to M x power minus alpha minus 1 dx is equal half

On integrating the function we get,
Alpha into b power alpha into x power minus alpha divided by minus alpha, ranges from b to M is equal to half
On substituting the limits, we get,
Minus b power alpha into M power minus alpha minus b power minus alpha is equal to half
Or 1 minus b power alpha into M power minus alpha is equal to half
That is b power alpha into M power minus alpha is equal to half
Implies M power alpha is equal to 2 into b power alpha
By taking alphath root, we get
M is equal to b into alphath root of 2

Mode is that value of x for which f of (x) is maximum. Observe that Pareto distribution is strictly a decreasing function. Hence, it attains its maximum at the very first value taken by the variable. That is If X has Pareto distribution with parameters alpha and b, then f of (x) attains its maximum at x is equal to b. Therefore, b is mode of the distribution.

5. Uses, Applications and Examples

Let us define some particular Pareto distribution.
The classical Pareto distribution has a survival function of the form
f of x is equal to x divided by b whole power minus alpha, where x is greater than or equal b, and both alpha and b are positive.
The probability density function of basic Pareto distribution is f of x is equal to alpha divided by x power alpha plus 1 for x greater than or equal to 1.

Feller defines a Pareto variable by transformation Y is equal to X power minus 1 minus 1of a beta random variable of first kind with parameters m and n.
That is, W is equal to alpha plus b into Y into gamma, has a Feller-Pareto distribution with parameters (alpha, b, m, n, and gamma)

Let us see how to get the conditional distribution of a Pareto Variate.
The conditional probability distribution of a Pareto Distributed random variable given the event that it is greater than or equal to a particular number n, exceeding b, is a Pareto distribution with the same Pareto index alpha but minimum n instead of b, that is in range, x is greater than or equal to n instead of b.

Applications
Pareto originally used this distribution to describe the allocation of wealth among individuals since it seemed 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. This idea is sometimes expressed more simply as the Pareto principle or the "80, 20 rule", which says that 20 per cent of the population controls 80 per cent of the wealth.

However, the 80, 20 rule corresponds to a particular value of alpha. In fact, Pareto's data on British income taxes in his *Cours d'économie politique* indicates that about 30 per cent of the population had about 70 per cent of the income.

The probability density function graph tells 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. (Note that the Pareto distribution is not realistic for wealth for the lower end. In fact, net worth may even be negative.) 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".

Following are the examples which are seen as approximately Pareto-distributed:

- The sizes of human settlements (few cities, many hamlets/villages)
- File size distribution of Internet traffic which uses the TCP protocol (many smaller files, few larger ones)
- Hard disk drive error rates
- Clusters of Bose-Elistein condensate near absolute zero

- The values of oil reserves in oil fields (a few large fields, many small fields)
- The length distribution in jobs assigned supercomputers (a few large ones, many small ones)
- The standardized price returns on individual stocks
- 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)
- Areas burnt in forest fires
- Severity of large casualty losses for certain lines of business such as general liability, commercial auto, and workers compensation

Consider the following illustration on Pareto distribution.
Suppose that the income of a certain population has the Pareto distribution with shape parameter 3 and scale parameter 1000. Find the proportion of the population with incomes between 2000 and 4000.

Let us solve this problem as follows.
Let X denote the income of a certain population. Hence, X has Pareto distribution with shape parameter (alpha) is equal to 3 and scale parameter (b) is equal to1000. Also find the median income.

Hence, we can write probability density function as follows:
f of x is equal to alpha into b power alpha divided by x power alpha plus 1 is equal to 3 into 1000 cube divided by x power 4, where x is greater than or equal to 1000
To find proportion of the population having income between 2000 and 4000, we need to find
Probability of income between 2000 and 4000
Is equal to probability that 2000 less than X less than 4000
Is equal to integral from 2000 to 4000 f of x dx
Is equal to 3 into 1000 cube into integral from 2000 to 4000 x power minus 4 dx
Is equal to 3 into 1000 cube into x power minus 3 divided by minus 3, ranges from 2000 to 4000
Is equal to 1000 cube into 2000 power minus 3 minus 4000 thousand power minus 3
On simplifying the above figures, we get
Zero point 1, zero 9, 4
We know that median of the distribution is given by,
M is equal to b into alphath root of 2
Is equal to 1000 into cube root of 2 is equal to 125

Here's a summary of our learning in this session, where we understood:

- The Pareto distribution – definition through pdf
- The distribution function
- The mean and variance
- The raw moments, skewness and kurtosis
- The moment generating function
- The median and mode
- The uses, applications and examples.