1. Examples and Applications of Log Normal Distribution

Welcome to the series of E-learning modules on Log Normal Distribution and its properties.

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

- Explain the examples of Log Normal Distribution and its applications
- Explain the probability distribution function and cumulative distribution function
- Explain median and mode
- Explain r^{th} raw moment, mean, variance, skewness and kurtosis
- Explain the relation between Pareto and Log Normal Distributions

Log normal is also written as log-normal or lognormal. The distribution is occasionally referred as Galton distribution or Galton's distribution, after Francis Galton, and other names such as McAlister, Gibrat and Cobb–Douglas have also been associated.
A variable might be modeled as log-normal if it is the multiplicative product of many independent random variables, each of which is positive.

For example, in finance, the variable could represent the compound return from a sequence of many trades (each expressed as its return + 1); or a long-term discount factor can be derived from the product of short-term discount factors. In wireless communication, the attenuation caused by shadowing or slow fading from random objects is often assumed to be log-normally distributed.

Examples of variates, which have approximately log normal distributions are:

- The size of silver particles in a photographic emulsion
- The survival time of bacteria in disinfectants
- The weight and blood pressure of humans
- Number of words written in sentences by George Bernard Shaw, etc.

Following are the different fields, where log normal distribution can be used:

- In Biology, variables whose logarithms tends to have a normal distribution include:
  - Measures of size of living tissue (length, skin area, weight)
  - The *length* of *inert* appendages (hair, claws, nails, teeth) of biological specimens, *in the direction of growth*
- Certain Physiological Measurements, such as blood pressure of adult humans (after separation on male/female subpopulations)
- In Hydrology, the log-normal distribution is used to analyze extreme values of such variables as monthly and annual maximum values of daily rainfall and river discharge volumes
- In Economics, there is evidence that the income of 97%–99% of the population is distributed log-normally
- In Finance, in particular the Black–Scholes model, changes in the *logarithm* of

exchange rates, price indices, and stock market indices are assumed normal (these variables behave like compound interest, not like simple interest, and so are multiplicative)

- In Reliability Analysis, the lognormal distribution is often used to model times to repair a maintainable system
- In Wireless Communication, the local-mean power expressed in logarithmic values has a normal (i.e. Gaussian) distribution
- It has been proposed that coefficients of friction and wear may be treated as having a lognormal distribution

2. PDF and CDF

Now, let us define log normal distribution.
A positive random variable X is said to have a log normal distribution if log X to the base e is normally distributed.
Let Y is equal to log X to the base e follows normal distribution with parameters mu and sigma square. For x greater than zero,
FX of x is equal to probability that X is less than or equal to x
Taking log on both the sides, we get probability that log X to the base e is less than or equal to log x to the base e
Is equal to probability of Y is less than or equal to log x to the base e, since logX is monotonic increasing function.
Is equal to 1 divided by sigma into root 2 into phi into integral from minus infinity to log x e power minus y minus mu the whole square divided by 2 into sigma square dy
Substituting y is equal to log u, above integral can be written as, 1 divided by sigma into root 2 into phi into integral from minus infinity to x, e power minus log u minus mu the whole square divided by 2 into sigma square du divided by u

For x less than or equal to zero, FX of x is equal to Probability of X less than or equal to x is equal to zero because X is a positive random variable.
Let us define
fX of u is equal to 1 divided by u into sigma into root of 2 into phi into e power minus log u minus mu the whole square divided by 2 into sigma square, if u is greater than zero and is equal to zero if u is less than or equal to zero.
Then, FX of x is equal to integral from minus infinity to x fX of u du for x and hence f of x is a probability density function of X.

Now, let us consider some remarks.

- If X follows Normal distribution with parameters mu and sigma square, then Y is equal to e power X is called a log normal random variable since logarithm log Y is equal to X is a normal random variable.
- Log normal distribution arises in problems of economics, biology, geology, and reliability theory. In particular, it arises in the study of dimensions of a particle under pulverization.
- Moment generating function of log normal distribution does not exist on the domain R but it exists on the half interval minus infinity to zero.
- If X_1, X_2, up to X_n is a set of independently identically distributed random variables such that mean of each log X_i is mu and variance is sigma square, then the product X_1 into X_2, into up to X_n is asymptotically distributed according to logarithmic normal distribution and with mean mu and variance n into sigma square.

Let us find the distribution function.
Cumulative distribution function of log normal distribution is given by
FX of x is equal to integral from minus infinity to x fX of x dx
Is equal to half into erfc of minus log x to the base e minus mu divided by sigma into square root of 2
Is equal to phi of log x to the base e minus mu divided by sigma, where erfc is the

complementary error function and is given by
Erfc of x is equal to 2 divided by square root of phi into integral from zero to x e power minus t square dt and phi is the cumulative distribution function of the standard normal distribution.

3. Median and Mode

If M is the median of the distribution, then it divides the whole distribution into two equal parts. Or if we consider the cumulative distribution function, FX of (x) is equal to half.
That is,
Phi of log e to the base e minus mu divided by sigma is equal to half.
Since phi is standard normal variate, we know that phi of zero is equal to half.
That is, log x to the base e minus mu whole divided by sigma is equal to zero
Implies, log x to the base e is equal to mu or
X is equal to e power mu.
Therefore, median of log normal distribution is given by
M is equal to e power mu.

Mode is the value of x for which the probability density function attains its maximum. Hence, it is a solution of f dash of x is equal to zero and f double dash of x is less than zero.
Probability density function of log normal distribution is given by,
f of x is equal to 1 divided by x into sigma into square root of 2 into phi into e power minus log x to the base e minus mu the whole square divided by 2 into sigma square.
Instead of differentiating the probability density function f of (x), we can work with the logarithm. Hence,
Log f of x to the base is equal to log c to the base e minus log x to the base e minus log x to the base e minus mu the whole square divided by 2 into sigma square, where c is equal to 1 divided by square root of 2 into phi.

By differentiating with respect to x, we get,
f dash of x divided by f of x is equal to minus 1 divided by x minus log x to the base e minus mu divided by sigma square into 1 divided by x
Taking f of x on the other sided, and taking common term outside, we get,
f dash of x is equal to minus f of x divided by x into 1 minus log x to the base e minus mu divided by sigma square.
f dash of x is equal to zero implies,
1 minus log x to the base e minus mu divided by sigma square is equal to zero.
Implies log x to the base e minus mu is equal to sigma square.
Implies log x to the base e is equal to mu plus sigma square.
Implies x is equal to e power mu plus sigma square.

Now, let us obtain the 2^{nd} derivative f double dash of x.
f double dash of x is equal to d by d x of minus f of x divided by x into 1 minus log x to the base e minus mu divided by sigma square.
Is equal to f of x into 1 divided by x square minus 1 divided by sigma square into 1 minus log x to the base e minus mu whole divided by x square minus 1 divided by x minus log x to the base e minus mu divided by x into sigma square whole multiplied by f dash of x.
Observe that the above function is less than zero at x is equal to e power mu plus sigma square
Hence, mode of the log normal distribution is given by x is equal to e power mu plus sigma square.

4. r^{th} Raw Moment, Mean, Variance, Skewness and Kurtosis

Now, let us obtain rth raw moment about origin of log normal distribution.
If Y follows normal mu sigma square, then X is equal to e power y is called a log normal random variable. rth raw moment about origin is given by,
Mu r dash is equal to expectation of X power r
Is equal to expectation of e power y the whole power r
Is equal to expectation of e power Y into r, which is same as M Y of r, that is moment generating function of Y, the normal distribution with parameters mu and sigma square. We know that the moment generating function of normal distribution is given by
M Y of r is equal to e power r into mu plus r square into sigma square divided by 2
Hence, rth raw moment is given by
Mu r dash is equal to e power r into mu plus r square into sigma square divided by 2

By putting r is equal to 1, we get mean of the log normal distribution as
Mu 1 dash is equal to e power mu plus sigma square divided by 2.
By putting r is equal to 2, we get,
Mu 2 dash is equal to e power 2 into mu plus 2 into sigma square
Variance of the distribution is given by,
Mu 2 is equal to mu 2 dash minus mu 1 dash the whole square.
Is equal to e power 2 into mu plus 2 into sigma square minus e power mu plus sigma square divided by 2 the whole square.
By taking the common terms outside, and simplifying we get,
e power 2 into mu plus sigma square into e power sigma square minus 1.

Putting r is equal to 3 in mu r dash, we get,
Mu 3 dash is equal to e power 3 into mu plus 3 square into sigma square divided by 2
Is equal to e power 3 into mu plus 9 into sigma square divided by 2.

Third central moment 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 e power 3 into mu plus 9 into sigma square divided by 2 minus 3 into e power 2 into mu plus 2 into sigma square into e power mu plus sigma square divided by 2 plus 2 into e power mu plus sigma square divided by 2 the whole cube.
By taking common terms outside we get, e power 3 into mu plus 3 into sigma square divided by 2 into e power 3 into sigma square minus 3 into e power sigma square plus 2.

Putting r is equal to 4 into mu r dash, we get
Mu 4 dash is equal to e power 4 into mu plus 4 square into sigma square divided by 2
Is equal to e power 4 into mu plus 8 into sigma square.
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 whole power 4
Is equal to e power 4 into mu plus 8 into sigma square minus 4 into e power 3 into mu plus 9 into sigma square divided by 2 into e power mu plus sigma square divided by 2 plus 6 into e

power 2 into mu plus 2 into sigma square into e power mu plus sigma square divided by 2 whole square minus 3 into e power mu plus sigma square divided by 2 the whole power 4.
By taking common terms outside we get,
E power 4 into mu plus 2 into sigma square into e power 6 into sigma square minus 4 into e power 3 into sigma square plus 6 into e power sigma square minus 3.

Coefficient of skewness is given by,
Beta 1 is equal to mu 3 square divided by mu 2 cube
Is equal to e power 3 into mu plus 3 into sigma square divided by 2 into e power 3 into sigma square minus 3 into e power sigma square plus 2 whole square divided by e power 2 into mu plus sigma square into e power sigma square minus 1 the whole cube.
Is equal to e power 3 into sigma square minus 3 into e power sigma square plus 2 whole square divided by e power sigma square minus 1 the whole cube, which is greater than zero.
Therefore, log normal distribution is positively skewed.

Coefficient of kurtosis is given by,
Beta 2 is equal to mu 4 divided by mu 2 square
Is equal to e power 4 into mu plus 2 into sigma square into e power 6 into sigma square minus 4 into e power 3 into sigma square plus 6 into e power sigma square minus 3 divided by e power 2 into mu plus sigma square into e power sigma square minus 1 the whole square
Is equal to e power 6 into sigma square minus 4 into e power 3 into sigma square plus 6 into e power sigma square minus 3 divided by e power sigma square minus 1 the whole square
On simplification, we get, e power 4 into sigma square plus 2 into e power 3 into sigma square plus 3 into e power 2 into sigma square minus 3, which is always greater than 3.
Hence, log normal distribution has leptokurtic curve.

Consider the following remarks.
We know that variance of the distribution is,
Mu 2 is equal to e power 2 into mu plus sigma square into e power sigma square minus 1
Hence, the standard deviation, SD is equal to e power mu plus sigma square divided by 2 into e power sigma square minus 1 power half.
Also, mean is given by, e power mu plus sigma square divided by 2.
Hence coefficient of variation is given by,
C.V. is equal to standard deviation divided by mean
Is equal to e power mu plus sigma square divided by 2 into e power sigma square minus 1 power half divided by e power mu plus sigma square divided by 2 is equal to e power sigma square minus 1 power half

Given mean mu 1 dash and variance mu 2 of log normal distribution, we can find mu and sigma square as follows.
Mu 1 dash is equal to e power mu plus sigma square divided by 2 and mu 2 is equal to e power 2 into mu plus sigma square into e power sigma square minus 1.
Taking logarithms and simplifying the above equations, we get,
Mu is equal to log mu 1 dash to the base e minus half log 1 plus mu 2 divided by mu 1 dash square to the base e
And sigma square is equal to log 1 plus mu 2 divided by mu 1 dash square to the base e

5. Relation between Pareto and Log Normal Distributions

Now, let us observe the relation between Pareto distribution and log normal distribution.
Note that the Pareto distribution and log-normal distributions are alternative distributions for describing the same types of quantities. One of the connections between the two is that they are 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.)

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

- The examples of log normal distribution and its applications
- The pdf and cumulative distribution function
- The median and mode
- The r^{th} raw moment, mean, variance, skewness and kurtosis
- The relation between Pareto and log normal distributions