1. Introduction

Welcome to the series of e-learning modules on Geometric Mean. In this module we are going to cover the definition and calculation of geometric mean for the various statistical series, followed by its application, advantages and limitations.

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

- The concept of geometric mean
- Calculation of geometric mean for statistical series
- Application of geometric mean
- Properties of geometric mean

First, We shall have a look at the definition of a mean.
Means are mathematical formulations used to characterize the central tendency of a set of numbers. Most people are familiar with the "arithmetic mean", which is also commonly called an average. Geometric mean has specific utility in the field of science, finance, and statistics.

Geometric mean is the ‘nth’ root of the product of ‘n’ items of a series. For example if there are two or three items in a series then the square root or the cube root of the products of the item in the series is taken.

A geometric mean is denoted by the alphabet ‘G’ or as G.M.
Symbolically it is shown as G is equal to nth root of the variables X one into X two into X three and so on to X n

Let us take an example of a series having three items 4, 6 and 9 and find their geometric mean.
In this case the geometric mean G is equal to the cube root of the products of the variable 4,6 and 9 which is equal to cube root of 216 which is equal to 6

Whenever the number of items in the series are two or three it is easy to take out the roots and calculate the geometric mean, but when there are more items in the series then it is excessively difficult finding the roots of the items.

To simplify we use the logarithms for calculation. The average of the logarithmic values of a data set, converted back to a base 10 number.
That is the Log of the geometric mean is equal to the Log of the variables X one plus X two plus X three plus and on so to plus X n, divided by the number of the data N.
Or Log of the geometric mean is equal to the sigma of Log X divided by N.
Thus by removing the Log we get that the geometric mean is equal to the Antilog of sigma Log X divided by N.

In statistics we classify the data using various statistical series. Statistical series are classified into – Individual series, discrete series and continuous series.

2. Calculation of Geometric Mean for Individual or Ungrouped Data

We have seen that averages can be calculated differently for various types of series. Similarly the geometric mean can be calculated for three different types of series.
In the section that follows we shall discuss the calculation of geometric mean for these three types of series.
First we shall look at the computation of geometric mean for individual or ungrouped data.
An individual or ungrouped data is a data where all the values are shown separately and is called an Individual series.

Suppose an individual series contains n items as X one, X two, X three and so on to X N, then the geometric mean G is defined as Geometric mean is equal to 1 by n to the power of the variables.

Taking logarithm on both sides of the equation,
We get, Log of Geometric mean is equal to Log 1 divided by n multiplied by Log of x one plus Log of x two plus Log of x three plus so on to Log of x n
Which is equal to the total value of the logarithm of X variable divided by n

Thus the Geometric mean is equal to Antilog of the total value of the logarithm of X variables divided by n.

Here is a look at the steps in the calculation of geometric mean for individual or ungrouped data.

- Take the logarithms of the variable X
- Obtain the total of the logarithms Sigma of Log X
- Divide Sigma of log X by N
- Take the antilog of the value obtained in the previous step
- This gives the value of the geometric mean

Let us take a few examples to understand the computation of geometric mean in an individual series
In the first example we are given that the annual rate of growth for a factory for 5 years as 7 percent, 8 percent, 4 percent, 6 percent and 10 percent respectively. We have to calculate the average rate of growth per annum for this period.

Now, If the growth for the company is 100 in the beginning then in the five years, the growth is given as 100 plus 7, 100 plus 8, 100 plus 4, 100 plus 6 and 100 plus 10 respectively;
We create a table where, the first column of the table we take the year and in the second column we have taken 100 as the beginning growth and added the percentage increase in the growth, the third column is the logarithm value of the variables and the total of the log of the values is equal to 10.14654

Table 1

Year	Growth (X)	Log (X)
1	100+7 = 107	2.0293
2	100+8 = 108	2.0334
3	100+4 = 104	2.0170
4	100+6 = 106	2.0253
5	100+10 = 110	2.0413
		$\Sigma \log x = 10.14654$

We shall now calculate the geometric mean using the formula for Geometric mean.
Substituting the values we get,
Geometric mean is equal to the antilog of the total of logarithm of the X variable that is 10.14654 in this case divided by 5 the total number of the data
which is equal to antilog of 2.0293 which is equal to 106.98
Thus the average rate growth is 106.98 minus 100
Which is equal to 6.98, And therefore we can say that the average rate of growth percentage per annum is 6.98 percent.

In the second example we are given the daily income of ten families of a particular place of which we need to calculate the geometric mean. The values given are: 85, 70, 15, 75, 500, 8, 45, 250, 40 and 36.

So first we create a table with the number of families in the first column and the income the X variable in the second column the third column represents the logarithm value of the X variable the total of which is 17.6373.

Table 2

Families	Income (X)	Log (X)
1	85	1.9294
2	70	1.8451
3	15	1.1761
4	75	1.8751
5	500	2.6990
6	8	0.9031
7	45	1.6532
8	250	2.3979
9	40	1.6021
10	36	1.5563
		$\Sigma \log x = 17.6373$

Applying the formulae we get that the geometric mean is equal to the antilog of total logarithmic value of the variable divided by the number of the data that is 17.6373 divided by 10 which is equal to antilog of 1.7637 which is equal to 58.03.

3. Calculation of Geometric Mean for Discrete Frequency Distribution

Next we shall look at the computation of the second series of geometric mean for Discrete frequency distribution. A discrete frequency distribution is a statistical series which represents the variables in the data and the number of time each variable has occurred. Here while calculating the geometric mean we need to consider both the variable, X and frequency F of the data.

Suppose there is a discrete frequency distribution with X one, X two and so on to X N, with data having frequencies F one, F two and so on to F N respectively then the geometric mean is the root of the product of the frequency with the variable in this case the frequencies are F one, F two and so on to F N and the variables are X one, X two and so on to X N.

Taking logarithms on both sides of the equation we get,
Log of the geometric mean is equal to one by N of frequency and its product of the log of the variable. The logarithm of the geometric mean is equal to one by n into the sigma or total of the product of the frequency with the logarithm value of the X variable divided by N.

Thus by removing the log from the left side of the equation we get geometric mean is equal to antilog of the sigma or total of the product of the frequency with the logarithm value of the X variable divided by N

Steps:

- Find the logarithms of the variable X
- Multiply these logarithms with the respective frequencies
- Obtain the total of summation of F logx
- Divide Summation of F Log X by N, which is the total of frequency
- Take the antilog of the value so obtained
- This gives the value of the geometric mean

Let us take a few examples to understand the computation of geometric mean in a discrete frequency distribution series.
In the first example, we are required to calculate the geometric mean from the data given in a table consisting of wages and the number of people receiving the corresponding wages.

Table 3

Wages	20	30	40	50	60	70	80
No. of Person	5	2	3	10	3	2	5

First step find the logarithms of the X variable as shown in the table the log of the variable wages 20 is equal to 1.3010, 30 is 1.4771 and so on

Table 4

Wages	Log
20	1.3010
30	1.4771
40	1.6021
50	1.6990
60	1.7782
70	1.8451
80	1.9031

Multiply the log with the frequency that is when we multiply 1.3010 the log value of the variable with the frequency 5 we get 6.5051 similarly the values for the other variables are calculated. We also find the total of the product that is 49.7954

Table 5

Wages	Log	Frequency	F Log X
20	1.3010	5	6.5051
30	1.4771	2	2.9542
40	1.6021	3	4.8062
50	1.6990	10	16.9897
60	1.7782	3	5.3345
70	1.8451	2	3.6902
80	1.9031	5	9.5154
		N=30	Σf logX= 49.7954

Calculate the geometric mean by applying the formulae we get that the geometric mean is equal to antilog of the sigma or total of the product of the frequency with the logarithm value of the X variable divided by N
Which is equal to antilog of 49.7954 divided by 30) which is equal to antilog of 1.65985 which is equal to 45.69 as the geometric mean.

In the second example we are required to calculate the geometric mean from the discrete frequency distribution table given.

Table 6

Age	5	10	15	20	25	30	35	40
No. of Person	5	2	3	10	3	2	5	14

Finding the logarithms of the age as the x variable we get the log of five as 0.6990, for 10 as 1.0000 and so on for the other variables as shown in the table.

Table 7

Age	Log
5	0.6990
10	1.0000
15	1.1761
20	1.3010
25	1.3979
30	1.4771
35	1.5441
40	1.60206

In step 2 we multiply the logarithms with the frequency. That ism in this case we multiply the log value 0.6990 with 12, which is the frequency and get 8.3876 and so on for the other variables then the total of the product is taken that is 125.5482 in this case along with the total of the frequency N which is equal to 100.

Table 8

Age	Log	Frequency	Σ F logx
5	0.6990	12	8.3876
10	1.0000	15	15.0000
15	1.1761	13	15.2892
20	1.3010	18	23.4185
25	1.3979	11	15.3773
30	1.4771	9	13.2941
35	1.5441	8	12.3525
40	1.60206	14	22.4288
		N =100	Σ f logx = 125.5482

Calculation of GM by applying the formula we get that the geometric mean is equal to the antilog of the of the sigma or total of the product of the frequency with the logarithm value of the X variable divided by N which is equal to the antilog of 125.5482 divided by 100 which is equal to antilog 1.25548 which gives the geometric mean equal to 18.009

4. Calculation of Geometric Mean for Continuous Frequency Distribution

We shall now look at the computation of geometric mean for continuous frequency distribution.
A continuous frequency distribution data does not assume exact value, but is given only within a certain range and the number of data points in that range is mentioned.
Here while calculating the geometric mean we need to consider the mid value (m) of the range variable(x) and frequency (f) of the data.

We now need to look at the steps involved in the computation of geometric mean for a continuous frequency distribution.

- Find the mid values of the classes
- Take the logarithms of the mid value (m)
- Multiply these logarithms with the respective frequencies
- Obtain the total of summation of F log M
- Divide summation of F log M by N (total of frequency)
- Take the antilog of the value so obtained
- This gives the value of the geometric mean

Therefore Geometric Mean is equal to the sigma or total of the product of the frequency with the logarithm value of the X variable divided by N.

Let us take a few examples to understand the computation of geometric mean in a continuous frequency distribution series.
In the first example we shall compute the geometric mean for the data given in the table as shown. The table contains marks in the series with class intervals of 4, as 4 to 6, 8 to 12, 12 to 16, 16 to 20 and so on and the respective frequencies are given as 6, 10, 18, 30, 15 and so on.
Table 9

Marks	Frequency	Marks	Frequency
4-8	6	24-28	12
8-12	10	28-32	10
12-16	18	32-36	6
16-20	30	36-40	2
20-24	15		

In the solution, first let us find the mid value of the given classes and then calculate the logarithms.
The mid value in this example is calculated by taking the class interval, say for the interval 4-8 we add

the upper limit and the lower limit that is 4 plus 8 which is 12 divided by 2 to take the average, which is 6, the mid-point.
Similarly we calculate the mid values of all the class intervals as shown in the table.
Then the logarithmic value of the mid-point values are taken, in this case the logarithmic value of 6 is 0.7782 which is further multiplied with the frequency 6 which is equal to 4.66891.
The same method is followed for the other values as represented in the table.

Table 10

Class Intervals	Mid points	Frequency	Logm	flogm
4-8	6	6	0.7782	4.66891
4-8	10	10	1.0000	10.00000
12-16	14	18	1.1461	20.63030
16-20	18	30	1.2553	37.65818
20-24	22	15	1.3424	20.13634
24-28	26	12	1.4150	16.97968
28-32	30	10	1.4771	14.77121
32-36	34	6	1.5315	9.18887
36-40	38	2	1.5798	3.15957
		N =109		Σf log m =137.19306

Now, the calculation of geometric mean is done using the formulae that is the geometric mean is equal to the Antilog of the sigma or total of the product of the frequency with the logarithm value of the X variable divided by N.
In this case the geometric mean is equal to antilog of 137.19306 divided by 109 which is equal to antilog of 1.25865 thus the geometric mean is equal to 18.14.

In the next example here, the table contains two columns, one containing the class interval of marks and the other with the number of students in each class. From this data we have to calculate the geometric mean.
The class interval given for this is 10, so the ranges in the marks table will be 0 to 10, 10 to 20, 20 to 30, 30 to 40 and so on. The corresponding values of number of students are given as 5, 10, 25, 30, 20 and 10 respectively.

Table 11

Marks	No. of students
0-10	5
10-20	10
20-30	25
30-40	30
40-50	20
50-60	10

The first step here is the calculation of mid value and logarithms.
Here we shall find the mid-value of the classes and then calculate the logarithms of the the mid-values.
In this example when we calculate we take the class interval 0 to10 and then add the upper limit and the

lower limit that is 0 plus 10 which is equal to 10, divided by 2 to take the average. We get 5. Similarly we calculate the mid values of all the class intervals as shown in the table, then the logarithmic value of these mid-point values are taken. In this case the logarithmic value of 5 is 0.6990 which is further multiplied with the frequency 5, which is equal to 145.23227. The same method is followed for the other values as represented in the table.

Table 12

Class Intervals	Mid points	Frequency	Logm	flogm
0-10	5	5	0.6990	3.49485
10-20	15	10	1.0000	10.00000
20-30	25	25	1.3979	34.94850
30-40	35	30	1.5441	46.32204
40-50	45	20	1.6532	33.06425
50-60	55	10	1.7404	17.40363
		N =109		Σ flogm =145.23327

Calculation of geometric mean is done using the formulae that is the geometric mean is equal to the Antilog of the sigma or total of the product of the frequency with the logarithm value of the X variable divided by N.
In this case the geometric mean is equal to antilog of 145.23227 divided by 100 which is equal to antilog of 1.45233.
Thus, the geometric mean is equal to 28.34.

We will now look at the formula for computing Compound Interest.
In business and economic problems very often we are faced with questions pertaining to percentage rates change over time. The most generally useful interpretation of this term is the constant percentage rate of change which if applied each year would take us from the first to the last figure.

The correct answer can be obtained through the use of geometric mean or what amounts to same thing through the use of compound interest calculations.
In this discussion the relationship between the average and compound interest calculations is indicated.
The compound interest formula is expressed as
P-N is equal to P-O multiplied by 1 plus r whole power n, where, P-N is the Amount accumulated at the end of 'n' periods
P-O is the original principal
R is the rate of interest expressed as a decimal and
N is the number of compounding periods.

It follows from the above formula that rate of interest is equal to n root of the amount accumulated at the end of n period divided by original principal minus one.

If interest is compounded at different rates in each time period, and if these successive rates are denoted by r one, r two and so on to r-n, then the amount accumulated at the end of n periods with an original principal of P-O is calculated as
P-N is equal to P-O into one plus r one into 1 plus r two plus and so on to 1 plus r-n

5. Application of Geometric Mean

We shall now look at the application of Geometric Mean.

- The geometric mean is used to find the average percent increase in sales, production, population or other economic or business series
- The geometric mean is a useful average for measuring the growth of the population, because population increases in geometric progression
- The geometric mean is considered to be the best average in the construction of index numbers
- The geometric mean is the most suitable when large weights have to be given to small items and small weights to large items for situations we usually face in economic and social fields

Here is an example to illustrate the use of geometric mean.
In a garment factory, the overhead expenses went up by 32% as compared to the previous year, this increased by 40% in the next year and by 50% in the following year. Calculate the average rate of increase in the overhead expenses over the three years.

We know that, in average ratios and percentages geometric mean is more appropriate.
Hence, applying geometric mean here:
If the growth of the company is 100 in the beginning then in the three years, the growth is given as follows;
That is in the first column of the table we take the year and in the second column we have taken 100 as the beginning growth and added the percentage increase in the growth, the third column is the logarithm value of the variables and the total of the logarithm of the values which is equal to 6.4428.

Table 13

% Rise	Expenses at the end of the year taking preceding year as 100	LogX
32	100+32=132	2.1206
40	100+40=140	2.1461
50	100+50=150	2.1761
		Σ logx = 6.4428

We will now calculate the geometric mean using the formula.
That is the Geometric mean is equal to the antilog of the total of logarithm of the X variable that is 6.4428 in this case divided by 3 the total number of the data which is equal to anti log of 2.1476 which is equal to 140.5.
Therefore the average rate of increase in overhead expenses is equal to 140.5 minus 100, which is equal to 40.5%.

Combined geometric mean:
As we calculate the combined arithmetic mean we can also calculate the geometric mean for items having a number of sets in them.
If the geometric mean of N items is 6 and then these N items are divided into sets first containing N

one items and second containing N two items having G one and G two as their respective geometric mean.
Then, the logarithm value is taken on both sides of the equation that is the log of the geometric mean is equal to the product of the number of items and the log of the geometric mean divided by the total number of items

To understand the above equation let us take an example
In this example we have two sets N one with 5 items having a geometric mean 20 and N two with 10 items having a geometric mean 35.28. Find the respective geometric mean.

By applying the formula and substituting the given values that is
N one as 5, N two as 10, G one is equal to 20 and G two is 35.28, we find the GM of the entire series 'N'.

N is equal to the total of N one plus N two. Substituting the values in the equation and simplifying, we get the log of the geometric mean as 1.465 and when we find the geometric mean we get the value as 29.17 which is the antilog of 1.465.

Corrected Geometric Mean is used when there is a mistake in the observations while calculating the geometric mean.
We use the following formula to apply the correction:
The correct geometric mean is equal to the product of the geometric mean of the observation and the correct observation divided by the wrong observation to the power of one by N.

Here is an example.
The geometric mean of 10 observations was calculated as 28.6. It was later discovered that one of the observations was recorded as 23.4 instead of 32.4.
We have to apply the correction and re-calculate the correct geometric mean.

So, we know that n is equal to 10, Calculated GM is 28.6, Correct observation is 32.4 and Wrong observation recorded was 23.4

Substituting the values in the equation we get

Correct Geometric mean is equal to the product of the geometric mean 28.6 and the correct value 32.4 divided by the wrong observation 23.4 to the power of one by 10.
Taking log on both sides of the equation and calculating we will get the logarithmic value of geometric mean as 1.47053 thus the geometric mean is equal to antilog of 1.47053 which is equal to 29.54.
Properties of geometric mean:
There are two important mathematical properties of geometric mean:
Property 1: The product of the value of the series will remain unchanged when the value of geometric mean is substituted for each value.
Property 2: The sum of the deviations of the logarithms of the original observations above or below the logarithm of the geometric mean is equal. This also means that the value of the geometric mean is such as to balance the ratio deviations of the observations from it.

Here, to illustrate the first property, we shall take an example.
Here we have to find the geometric mean for series 2,4 and 8.
Using the formula for GM,

GM is equal to cube root of 2 into 4 into 8, Which is equal to cube root of 64 which is 4.
When substituting the mean value in the series we get the same value series.
That is 4 into 4 into 4 which will be 64.

To illustrate the second property, here is an example.
If 4 by 2 into 4 by 4 is equal to 2 which can also be represented as 8/4.
This property is useful as it is adapted in calculating average ratios, rates of change and logarithmically distributed series.

Here’s a summary of our learning in this session:
In this session

- We have understood the concept of geometric mean
- The calculation of geometric mean for various statistical series
- Properties and Application of geometric mean