Measures of Central Tendency

Central tendency: In a representative sample, the values of a series of data have a tendency to cluster around a certain point usually at the center of the series. This tendency of clustering the values around the center of the series is called central tendency and its numerical measures are called the measures of central tendency.

Objectives:

There are two main objectives of the study of the averages:

i) To get one single value that describes the characteristics of the entire data.

ii) To facilitate comparison.

For example, the figure of average sales for December may be compared with the sales figure of previous months or with the sales figure of another competitive firm.

Characteristics of a good average:

i) It should be easy to understand.
ii) It should be simple to compute.
iii) It should be based on all the observations.
iv) It should be rigidly defined.
v) It should be capable of further algebraic treatment.
vi) It should have sampling stability.
vii) It should not be unduly affected by the presence of extreme values.

The following are the important measure of central tendency which are generally used in business.

1. Arithmetic Mean (A.M.) or Mean
2. Median
3. Mode
4. Geometric Mean (G.M.)
5. Harmonic Mean (H.M.)

Arithmetic Mean (A.M.) or Mean: Arithmetic mean of a set of data observations is the sum of all observations divided by the number of observations. It is denoted by A.M. or $\bar{x}$.

Calculation of Arithmetic Mean (A.M.):

For ungrouped or discrete series:

Suppose $x_1, x_2, \ldots\ldots, x_n$ are n numbers ungrouped observations. So arithmetic mean is given by-

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n} = \frac{x_1 + x_2 + \ldots\ldots + x_n}{n}$$

Most. Salma Yeasmin Jannaty (SYN)
Lecturer (Statistics), MNS Department,
BRAC University.
E-mail: salma.yeasmin@bracu.ac.bd

Example: The monthly income (in $) of 10 employees in a firm is as follows:
4487, 4493, 4502, 4446, 4475, 4492, 4572, 4516, 4468, 4489
Find the average monthly income.

For grouped or continuous series:

Suppose $x_1, x_2, \ldots\ldots, x_n$ are n numbers observations and their corresponding frequencies are $f_1, f_2, \ldots\ldots, f_n$. So arithmetic mean is given by-

$$\bar{x} = \frac{f_1x_1 + f_2x_2 + \ldots\ldots + f_nx_n}{f_1 + f_2 + \ldots\ldots + f_n} = \frac{\sum_{i=1}^{n} f_i x_i}{N}$$

Where, $i = 1, 2, \ldots\ldots, n$ and $N = f_1 + f_2 + \ldots\ldots + f_n$

Example:

The following are the figures of profits earned by 1,400 companies during 2007-08:

Profits (Taka lakhs)	No. of companies
200-400	500
400-600	300
600-800	280
800-1000	120
1000-1200	100
1200-1400	80
1400-1600	20

Calculate the average profits of all the companies.

(i) By Direct method:

Solution:

Calculation of Average Profits

Profits Tk Lakhs	**Mid-value** x_i	**No. of companies** f_i	$f_i x_i$
200-400	300	500	1,50,000
400-600	500	300	1,50,000
600-800	700	280	1,96,000
800-1000	900	120	1,08,000
1000-1200	1100	100	1,10,000
1200-1400	1300	80	1,04,000
1400-1600	1500	20	30,000
Total		N=1,400	**8,48,000**

Most. Salma Yeasmin Jannaty (SYN)
Lecturer (Statistics), MNS Department,
BRAC University.
E-mail: salma.yeasmin@bracu.ac.bd

We know,

$$Arithmetic\,Mean, \bar{x} = \frac{\sum_{i=1}^{n} fx}{N}$$

$$= \frac{8,48,000}{1,400} = 605.71$$

Thus, the average profit is 605.71 lakh Taka.

Properties of Arithmetic Mean:

1. Every set of interval or ratio level data has a mean.
2. All the values are included in computing the mean.
3. The mean is unique.
4. The sum of the deviations of each value from their mean is zero. Symbolically, $\sum_{i=1}^{n}(x_i - \bar{x}) = 0$

Merits of A.M.:

Arithmetic mean is the most important and useful measure in statistics.

1. It is widely used and understood by the common people.
2. It is easy to calculate and suitable for further mathematical treatment.
3. It takes all the observation into account.
4. It is less affected by sampling fluctuations.

Demerits of A.M.:

1. It is badly affected by extreme value.
2. We cannot use arithmetic mean when we have qualitative data.
3. Sometimes it gives wrong information.

Uses: The principle uses of averages are-

1. A typical value or average value is useful for describing a group.
2. Average is useful as a basis for comparing groups of people or two or more series.
3. It is useful to estimate an average condition for a large group people from the average of a sample.

Median:

It is the middle most value in a set of figures when they are arranged in ascending or descending order of magnitude.

Most. Salma Yeasmin Jannaty (SYN)
Lecturer (Statistics), MNS Department,
BRAC University.
E-mail: salma.yeasmin@bracu.ac.bd

For ungrouped data(Discrete data):
When n is odd:
Let $x_1, x_2, ..., x_n$ are n numbers observations. When n is odd then median is given by-

Median = $\frac{n+1}{2}$ th observation in a series arranged in ascending or descending order.

When n is even:
When n is even then median is given by-

Median=Average of $\frac{n}{2}$ the and $(\frac{n}{2}+1)$th item in the series.

For Grouped data (Continuous Data):
For grouped frequency distribution median is given by-

Median= $L + \frac{N/2 - c.f}{f_m} \times c$

Median class= size of $\frac{N}{2}$ th item

Where, L = lower limit of the median class
N =total number of observations
$c.f$ =cumulative frequency of the class just preceding the median class.
f_m =frequency of the median class.
c =class interval

Merits:
1. Median is rigidly defined
2. It is not affected by the extreme value.
3. It can be calculated from frequency distribution with open end class interval.

Demerits:
1. It is not based on all the observations.
2. It is not easy for algebraic treatment.
3. It is affected much by sampling fluctuation.

Mode:
Mode is the value which possesses the maximum frequency i.e. it is that item which is repeated largest number of time.

It is the value of the variable which occurs most frequently.

Most. Salma Yeasmin Jannaty (SYN)
Lecturer (Statistics), MNS Department,
BRAC University.
E-mail: salma.yeasmin@bracu.ac.bd

Calculation of Mode:

For ungrouped data (Discrete data):

Let $x_1, x_2, ..., x_n$ are n numbers observations.

Mode=The value in the series x_i which has maximum frequency or which occurs most frequently.

For Grouped data (Continuous Data):

Suppose $x_1, x_2, ..., x_n$ are n numbers observations and their corresponding frequencies are $f_1, f_2, ..., f_n$. So Mode is given by-

$$\text{Mode}= L + \frac{\Delta_1}{\Delta_1 + \Delta_2} \times c$$

Where, L = the lower limit of the modal class

Δ_1 = The difference between the frequency of the modal class and pre-modal class.

Δ_2 = The difference between the frequency of the modal class and post-modal class.

c =class interval.

Merits:

1. It is useful for non-quantitative data.
2. It is not affected by extreme values.
3. It can be calculated from frequency distribution with open class.

Demerits:

1. Mode is not clearly defined in case of bimodal or multimodal distribution.
2. It is not based on all the observations.
3. It is affected to a great extent by sampling fluctuation.

Relation between A.M (mean), Median and Mode:

Mode=3 Median-2 Mean

Geometric Mean:

The geometric mean is useful in finding the average change of percentages, ratios, indexes or growth rate. It has wide application in business and economics because we are often interested in finding the percentage changes in sales, salaries or economic figures such as GDP.

The n th root of the product of n observation is called geometric mean.

i.e. Geometric mean=(product of n observation)$^{1/n}$

To define G.M. here is some conditions:

i) none of the $x_i = 0$

ii) Product will not be negative.

iii) Odd number of values should not be negative.

Most. Salma Yeasmin Jannaty (SYN)
Lecturer (Statistics), MNS Department,
BRAC University.
E-mail: salma.yeasmin@bracu.ac.bd

Calculations of Geometric Mean:

For ungrouped data: Suppose $x_1, x_2, ..., x_n$ are the n numbers non-zero positive observations. So G.M. is given by-

$$\text{G.M.} = (x_1.x_2......x_n)^{1/n}$$

$$\text{Log G.M.} = \frac{\sum_{i=1}^{n} \log x_i}{n}$$

$$\text{G.M.} = \text{antilog}\left(\frac{\sum_{i=1}^{n} \log x_i}{n}\right)$$

Merits:

1. It is determinate. i.e. any data have only one G.M.
2. It gives more weights to small values than to large values.
3. It takes all the values into account.

Demerits:

1. If any value of the series is zero then G.M. cannot be determined.
2. If product x, $\prod x_i = -ve$, then geometric mean cannot be defined.
3. It is somewhat troublesome to calculate.
4. General people do not have the concept of logarithm. So, they cannot understand easily this measure.

Harmonic Mean: Harmonic mean is the reciprocal of the average of the reciprocal of the observations in a series. It is the ratio of the number of observations and the sum of reciprocal of observations.

i.e. Harmonic Mean(H.M)= (Number of observations/Sum of the reciprocal of the observations)

Conditions: None of the observations can be zero.

Calculations of H.M.:

For ungrouped data: Suppose $x_1, x_2, ..., x_n$ are n numbers non-zero positive observations. Then Harmonic Mean is given by-

$$\text{H.M.} = \frac{n}{\sum_{i=1}^{n} \frac{1}{x_i}}$$

Merits:

1. it is rigidly defined.
2. it takes all the observations into account.
3. It is not affected much by small observations.

Most. Salma Yeasmin Jannaty (SYN)
Lecturer (Statistics), MNS Department,
BRAC University.
E-mail: salma.yeasmin@bracu.ac.bd

Demerits:

1. it is not easy to understand.
2. it is impossible to calculate if the extreme classes of the frequency distribution are open.

Most. Salma Yeasmin Jannaty (SYN)
Lecturer (Statistics), MNS Department,
BRAC University.
E-mail: salma.yeasmin@bracu.ac.bd