1.1. Meaning

One of the most important objects of statistical analysis is to get one single value that describes the characteristics of the whole data. Such a value is called the **'central value'** or an **'average'**. A central tendency is, thus, a single figure that represents the whole mass of data. For example, if a comparison is to be made between the marks in economics obtained by the students of class XI in a school and marks obtained by the students of XI class of another school in the same town, it may be difficult to arrive at any conclusion. But if we say that average marks obtained by students of a school is 90 and that of another school is 80, we can easily compare the performance of students of the two schools.

Individual observations in a distribution generally have characteristics of showing a tendency to concentrate somewhere in the centre of the distribution which is called 'average' or 'central value. For example, suppose we give a test in economics to a class of 40 students out of a maximum marks of 100. We shall find that most of the students would have secured marks around a central figure, say 60. Here this figure of 60 marks can be considered as the average or central tendency. An average is called as a measure of central tendency because it is a typical value around which other figures lie or a value which occupies the central position. In this way, when we talk of average cost of production, average revenue, average productivity, average size of holdings, per capita income, etc., we mean central tendencies. Measures of central tendency thus, refer to all those methods of statistical analysis which are used to calculate the average of a set of data. Averages occupy an important place in statistical methods. That is why, **Bowley** has rightly remarked, statistics as the **Science of averages**.

In brief, *an average is a single value within the range of data that is representative of all the items in the series.* It is also known as **'measures of location'** or **'summary number'**.

average productivity of labour etc.

1.3. Essentials of an Ideal Average.

An ideal average must possess the following characteristics:

1. **Simple to Calculate and Easy to Understand.** It should be easy to understand. It should be simple to calculate.

2. **Rigidly Defined.** It should be rigidly defined. It means whatever may be the method of calculation, answer should be the same. So, it should have a definite and fixed value.

3. **Based on all Observations.** It should be based on all observations. For example, there are five values of a series, namely 20, 30, 40, 50 and 60. The average of these values is 40 $\left(\frac{20+30+40+50+60}{5}\right)$. The average thus, obtained is based on all observations.

4. **Capable of Algebraic Treatment.** It should be capable of algebraic treatment. For example, the sum of all the values of a series should be equal to the product of the average and the number. In algebraic form:

$$N \times \overline{X} = \Sigma\left(X^1 + X^2 + X^3 + ... X^n\right)$$

$N \times \overline{X}$ denotes the product of average and the number of observations whereas Σ $(X^1 + X^2 + X^3 + ... X^n)$ denotes sum of all the values of observations.

Similarly, if averages of two or more than two groups are given differently, combined mean should also be calculated.

5. **No Undue Effect of Extreme Values.** It should not be unduly affected by extreme values. Although each and every value should influence the value of the average but it should not effect it unduly. One or two very small or very large values should not reduce or increase the value of the average unduly.

6. **Least Affected by Fluctuations of Sample.** It should be least affected by fluctuations of sampling. If we take two or more independent random samples of the same size from a given population and compute the average for each, then the values so obtained from different samples should not differ much from one another.

7. **Expressed in Absolute Number.** A good average should always be expressed in a absolute number. Average measured in terms of percentage or according to any relative method is not considered a good measure of central tendency.

2.1. Arithmetic Mean

2.1.1. Meaning of Arithmetic Mean. Arithmetic mean is a most popular and useful measure of central tendency. Generally, when we talk of 'average', it signifies arithmetic mean. *Arithmetic mean is defined as the sum of values of a group of items divided by the number of items.* It is very easy to calculate. For example, marks obtained by 5 students in Statistics are 10, 12, 14, 16, and 18. To find out arithmetic mean, we shall just add these values and then divide the total by 5, i.e.,

$\frac{10+12+14+16+18}{5}$ The value so obtained is 14 which is arithmetic mean. Arithmetic mean may be of two types :

1. Simple arithmetic mean
2. Weighted arithmetic mean

2.1.4. Mathematical Properties

The following are some mathematical properties of the arithmetic mean:

1. **The sum of the deviations of the values from mean is zero.** Symbolically,

$$\Sigma\left(X - \overline{X}\right) = 0.$$

This property becomes clear from the following example:

From the above example, it is clear that, when the sum of the deviations from the actual mean *i.e.*, 40 is taken, it comes out to be zero.

2. If some specific value is added to or subtracted from different items in a series, the mean value of the items would increase or decrease by the same specific value respectively.

Likewise, if diferent items of a series are multiplied or divided by any specific value, the mean value of the items would be the multipled or dvided by the same specific value.

3. **The sum of the squared deviations of the values from arithmetic mean is minimum.** It is less than the sum of the squared deviations from any value. Observe the following example:

5. **If, we have the arithmetic mean and number of items of two or more than two related groups, we can calculate the combined arithmetic mean of these groups.** Combined average can be computed by applying the following formula:

$$\overline{X}_{1,2} = \frac{N_1\overline{X}_1 + N_2\overline{X}_2}{N_1 + N_2}$$

Here

$\overline{X}_{1,2}$ = Combined mean of the first and second groups

N_1 = Number of items in the first group

N_2 = Number of items in the second group

$\overline{X}_1$ = Arithmetic mean of the first group

$\overline{X}_2$ = Arithmetiuc mean of the second group.

If we have to find out combined mean of the three groups of a series, the above formula can be extended as follows:

$$\overline{X}_{1,2,3} = \frac{N_1\overline{X}_1 + N_2\overline{X}_2 + N_3\overline{X}_3}{N_1 + N_2 + N_3}$$

2.1.5. Merits and Demerits of Arithmetic Mean

Merits. Arithmetic mean is widely used because of the following merits:

1. It is rigidly defined. Therefore, its value is always definite.
2. It is easy to calculate and simple to understand.
3. It is based on all observations. In other words, it is affected by each and every item of the series.
4. It is capable of further algebraic treatment. So it is widely used in further statistical analysis.
5. It provides a good basis to compare two or more than two series.
6. It is a calculated value and not based on the position in the series like median and mode,

Demerits. Following are the main drawbacks of arithmetic mean:

1. It is highly affected by extreme values. For example, mean of 3, 5, 499 is 169. In this example, none of the value is represented by the mean.
2. It cannot be determined in open end classes. For such classes, we cannot determine middle values.
3. Sometimes arithmetic mean gives absurd and ridiculous results. For instance, average number of children born per married couple in India is 2.3, average number of patients admitted in a hospital is 18.7 per day.
4. It cannot be determined graphically.
5. It does not throw any light on the composition of the series. The arithmetic mean of different series may be the same even when their composition is altogether different.

   For example, three students of school A may obtain marks in English as 60, 70 and 80, while 3 students of school B obtain 95, 30 and 85 marks. The average marks in English of students in both the schools is 70. It appears that the students of both the schools have the same level of performance. But this actually is not the case. In such cases, arithmetic mean gives misleading conclusions.

2.2. Median

2.2.1 Meaning. *Median is the middle value of the series when arranged either in ascending order or in descending order.* It is the value which divides the arranged series into two equal parts. One part comprises all values greater than the median value and the other part comprises all values smaller than the median value. The number of observations smaller than median is equal to the number greater than it. Median is, thus, an average of position. The word **'position'** here implies the place of a value in the series. Its value lies in the middle. Median is particularly useful when the items are not capable of exact measurement. For example, when we talk of average intelligence, average health of people, we are talking about median as an average.

"The median is that value of the variable which divides the group into two equal parts, one part comprising all values greater and the other values less than the median." – L.R. Connor

The concept of median can be better explained with the help of

2.2.4. Merits and Demerits of Median

Merits:

(i) It is very easy to understand and easy to locate.

(ii) It is not influenced by the extreme values because it is a positional value.

(iii) The value of median can also be determined graphically. Mean cannot be determined graphic

(iv) It is most appropriate average in dealing with qualitative data, that is, where ranks are give where items are not counted or measured.

(v) It is especially useful in case of open end classes since only the position is to be known.

(vi) It can also be calculated even in case of some missing values.

(vii) It is rigidly defined. This means that it has a definite value.

Demerits:

(i) It requires arranging of data in ascending or descending order whereas other averages do require any such arrangement. So, lot of time and labour is wasted.

(ii) It is not based on all observations of the series as at is a positional average.

(iii) In case of median, algebraic treatment is not possible. For example, median cannot be used determine combined median of two or more than two groups of a series.

(iv) The value of median is affected by changes in the sample.

(v) It is very difficult to calculate, if the number of items is very small or large.

2.3 Partition Values

When a series is divided into more than two parts, the dividing values are called partitio values. As discussed earlier, median divides a statistical series into two equal parts. Lik median, there are other partition values also such as quartiles, deciles and percentiles etc. Partitic values which divide series into four equal parts, are called as **quartiles. Deciles** divide the series in ten equal parts. Similarly, those values which divide a series into hundred equal parts are known as **percentiles.**

Quartiles

There are four quartiles in each statistical series—Q_1, Q_2, Q_3 and Q_4. Second quartile (Q_2) is median and fourth quartile (Q_4) is the last value of the series. First quartile (Q_1) or lower quartile is that value of the variable which divides the first half of a series into two equal parts. Whereas, the third quartile or upper quartile (Q_3) divides the latter half of the series into two equal parts. The first quartile (Q_1) or lower quartile has 25% of the items of the distribution below it and 75% of the items are greater than it. The second Quartile (Q_2) has 50% items below it and 50% of the observations above it. The third Quartile (Q_3) or upper quartile has 75 per cent of the items below it and 25% of the items above it. In actual practice, only first and third quartiles are used within which 50% of the central items lie. **It should be noted here that Q_1 is always less than Q_2 and Q_3 and median lies between Q_1 and Q_3.**

Calculation of Partition

2.4 Mode

2.4.1. Meaning.

The English word *'mode'* has been taken from a French word 'La Mode' which means fashion or custom. It is like mean and median a central value of the distribution. It is most oftenly said to be the value which occurs most frequently in a series. In simple words, it is the most common value found in a series. For example, if most of the workers in a factory get Rs. 70 as daily wages, then the value of mode will be 70. Wages as high Rs. 100 and as low Rs. 40 are much less frequented than the modal wages. Mode is very useful in expressing most common size of shoes or readymade garments. If we say that mode is 40 number shirt, this means maximum number of people wear shirts of size no. 40.

2.4.3. Advantages and Disadvantages of Mode

Advantages:

(i) It is very easy to understand and simple to calculate. It can be located even by inspection.

(ii) It is not affected by the extreme values.

(iii) It is precise. It is an actual item of the series except in the case of continuous series.

(iv) It can be located even graphically.

(v) It is a true representative of frequency distribution, since it is the value which occurs most frequently.

(vi) It is also useful in showing qualitative facts. For example, when we talk of likings and dislikings of the people, we mean mode.

(vii) It is the most suitable average in certain circumstances. For example, when an investigation is to be conducted concerning living standard, the modal wage will tell us the maximum number of persons getting a particular wage.

Disadvantages:

(i) It is not suited to algebraic treatment.

(ii) In some cases, mode can be ill-defined. A clearly defined mode does not always exist.

(iii) There can be bi-modal frequency distribution.

(iv) The value of mode is not based on all items of the series.

(v) In the computation of mode, grouping of the data is necessary whereas ,mean can be calculated even when the sum and the number of items are given.

5. Limitations of Measures of Central Tendency

In spite of their importance, measures of central tendency have many limitations which are mentioned below:

1. Measures of central tendency do not describe the true structure of the distribution. Two or more than two distributions may have same mean but different structure.
2. Sometimes, average is such value which is not in the distribution hence, is not a true representative. For example, mean of 200, 10, 6 and 4 is 55 which is not true representative.
3. Sometimes, average gives absurd results. For example, we find average number of members per family as 2.3, 1.3 etc.

(ii) It is not based on all values.

4. **What are the essentials of an Ideal average?**

Ans. **Essentials of an Ideal Average :**

(i) It should be simple to calculate and easy to understand.

(ii) It should be rigidly defined.

(iii) It should be based on all values.

(iv) It should be capable of algebraic treatment.

(v) It should not be affected by extreme values.

(vi) It should be least affected by fluctuations of sampling.