Faculty Science

Everest Shiwach

Department: Botany

B Sc III -Paper I (Plant Resource Utilization, Palynology, Plant Pathology and Biostatistics)

Unit- IV Topic- Mean, Median and Mode

Average

An average is a single value within the range of data that is used to represent all the values in the series. This tendency of distribution is known as Central tendency.

Types of measure of Central tendency-

1. Arithmetic mean or simple mean

2. Geometric mean

3. Harmonic mean

4. Median

5. Mode

1. Arithmetic mean- It is the most popular and important measure of Central tendency. In daily life it is called as an average, in statistics it is called as arithmetic mean or simple mean.

It is defined as the sum of items divided by the number of items

If the values of N items are $x_1, x_2, x_3, \ldots\ldots, x_N$ be the value of variate x, the arithmetic mean usually denoted by

$$\bar{x} = \frac{x_1 + x_2 + \cdots + x_N}{N} = \frac{\sum_{i=1}^{N} x}{N}$$

If the variate x takes the values $x_1, x_2, x_3, \ldots\ldots, x_N$ with frequencies $f_1, f_2, f_3, \ldots\ldots, f_N$ respectively

$$\bar{x} = \frac{\sum f_1x_1 + f_2x_2 + ... + f_Nx_N}{f_1 + f_2 + f_3 + ... + f_N}$$

Example 1. Calculate the mean of the marks obtained by ten students of a class in Biochemistry subject. The marks obtained are: 24, 22, 30, 45, 57, 10, 15, 45, 17, 35

Solution:

Let $\bar{x}$ be the average marks obtained. Sum of all the observations

$\sum x$ = 24+22+30+45+57+10+15+45+17+35=300

Number of students n=10

Mean $\bar{x} = \frac{\sum x}{n} = \frac{300}{10}$ =30.

Example 2. The following table gives the number of marks obtained in Botany by 50 students. Calculate the average number of marks obtained by each student.

Marks obtained (x)	Number of students (f)
10	10
20	5
30	14
40	21

Solution:

Computation of the average number of marks obtained by the students.

Marks obtained (x)	Number of students (f)	fx
10	10	100
20	5	100
30	14	420
40	21	840
	$\sum f = 50$	$\sum fx = 1460$

$$\bar{x} = \frac{\sum fx}{\sum f}$$

where,

$\bar{x}$ = mean

$\sum fx$ = sum of the products of the size of items and corresponding frequencies
$\sum f$ = total frequencies

$$\bar{x} = \frac{\sum fx}{\sum f} = \frac{1460}{50} = 29.2$$

The average number of marks obtained by each student is approximately 29.

Merits-1. For any distribution the sum of deviations from their Arithmetic mean is always zero.

2. It is simple and easy to calculate.

3. It includes all the values.

4.It cannot be calculated if a certain item is missing.

Demerits-1. It is affected by the unusually large or small data values.

2. It may not be actually present in the series.

3. It cannot be determined by inspection nor can be represented graphically.

2. Geometric mean- The geometric mean for *N* positive items is defined as the positive Nth root of the products of *N* items.

$$GM = \sqrt[N]{x_1 \cdot x_2 \cdot x_3 \cdots \cdot x_N}$$

The geometric mean of any series is always less than or equal to arithmetic mean i.e AM ≥ GM

Merits-1. It is based on all the observations.

2. It is not much affected by extreme value.

3. It is rigidly defined.

4. It is capable of further algebraic manipulation.

5. It is particularly useful in dealing with ratios, rates and percentage.

Demerits-1. It is difficult to calculate.

2. It cannot be used when any one value is zero or negative.

3. It may be a value which is not present in the actual data.

3. Harmonic mean- It is defined as reciprocal of the arithmetic mean of the reciprocals of the given values. If $x_1, x_2, x_3, \ldots\ldots, x_N$ be the observations then

HM of N observations is defined as

$$H = \frac{N}{\frac{1}{x_1} + \frac{1}{x_2} + \ldots\ldots\ldots + \frac{1}{x_N}} = \frac{N}{\sum_{i=1}^{N} \frac{1}{x_i}}$$

Merits-1. It is based on all the observations of a series.

2. It gives largest weightage to smaller items.

3.It is useful to study the rate of respiration i.e averages of rates or times

4. It is not much affected by sampling fluctuations.

Demerits-1. It is not easy to calculate and understand.

2. It cannot be calculated if negative and positive values are given in a series.

3. It cannot be calculated if one value is zero.

Average of Position

4. Median- The value of a middle most of observation, when the data are arranged in ascending or descending order of magnitudes, is called the median of the data.

We arrange the data in increasing or decreasing order. Let N be the total number of observations

i). If N is odd then median is the value of $\left(\frac{N+1}{2}\right)^{th}$ observations.

ii). If N is even then the average of $\left(\frac{N}{2}\right)^{th}$ and $\left\{\left(\frac{N}{2}\right)+1\right\}^{th}$ observations is the median.

Example 3. Case I- When the number of observations is odd

Find out the median of the following items:

5, 7, 12, 10, 8, 7, 20

Solution: To calculate median, first we arrange the items in ascending order as

5, 7, 7, 8, 10, 12, 20

If *M* represents the median, and *N* the number of items,

$$M = \text{size of } \left(\frac{N+1}{2}\right)^{th} \text{ term}$$

$$= \text{size of } \left(\frac{7+1}{2}\right)^{th} \text{ term}$$

= size of 4^{th} term = 8 Ans.

Case II- When the number of observations is even

Example: Find out the median of the following items:

2,4,1, 5, 8, 7, 5, 10

Solution: To calculate median, first we arrange the items in ascending order as

1, 2, 4, 5, 5, 7, 8, 10

If *M* represents the median, and *N* the number of items,

$$M = \text{size of average of } \frac{1}{2}\left[\left(\frac{N}{2}\right)^{th} + \left\{\left(\frac{N}{2}\right)+1\right\}^{th}\right] \text{ term}$$

= size of average of $(4)^{th}$ and $(5)^{th}$ term

$$= \frac{5+5}{2}$$

= 5 Ans.

Merits-1. It is easy to understand and to calculate.

2. It is well defined.

3. Extreme values are eliminated.

4. The position of the median is based on all observations but required only the value of middle observation.

5. It can be located graphically.

6. It can be useful in open end distribution.

Demerits-1. It is not a calculated figure obtained by mathematical formula.

2.The data has to be arranged in order of magnitude.

3. It may not be representative when the distribution is irregular.

4. It is not capable of further algebraic treatment.

5. Mode- Mode of a frequency distribution is defined as that value of variable for which frequency is maximum.

Example: Find out the mode of the following items:

2, 4,1, 5, 8, 4, 3, 2, 4

Solution: Since mode is the most frequent term. If we observe the given data we can see that the most frequent term is 4.

Merits-1. It is easy to understand. Sometime it is found only by inspection.
2. It is least affected by extreme values.
3. It can also be located graphically.
4. It is assumed to be the best representative of the distribution because its frequency is maximum.

Demerits-1. It is not well defined.
2. Arithmetic explanation of mode is not possible.
3. Sometimes it is indefinite.
4. It becomes difficult in multi modal distribution.
5. It is not based on all the observations of a series.

Relationship between Mean, Median and Mode
There is empirical relationship between Mean, Median and Mode of a series of items. If distribution of item values be symmetrical then the Mean, Median and Mode coincides, otherwise the distance between the Mode and the Median is usually twice the distance between the Median and the Mean:
Mode – Median = 2(Median – Mean)
or Mode – Mean = 3(Median – Mean)
or **Mode = 3 Median – 2 Mean**