STATISTICS

(**Statistics** is the branch of Mathematics dealing with the collection, organization, and interpretation of numerical data)

1. In statistics, a **central tendency** (or, more commonly, a **measure of central tendency**) is a central value or a typical value for a probability distribution. It is occasionally called an average or just the **center** of the distribution. The most common measures of central tendency are

    a. The arithmetic mean or mean

    b. The median and

    c. The mode.

2. $Class\ mark\ (or)\ mid-value = \dfrac{\text{Lower class limit + Upper class limit}}{2}$

3. Mean of grouped data:

The mean (or average) of observations is the sum of the values of all the observations divided by the total number of observations.

It can be found by three methods:

    I.   **THE DIRECT METHOD**

$$\overline{x} = \frac{f_1 x_1 + f_2 x_2 + \ldots\ldots + f_n x_n}{f_1 + f_2 \ldots\ldots f_n} = \frac{\sum f_i x_i}{\sum f_i}$$

Where,

$\overline{x}$ = mean,

$\sum$ = Sum operation,

f_i = frequency of the observations of a data,

x_i = Midvalue of the observations of a data,

$f_i x_i$ = Product of midvalue of the observations and their corresponding frequencies of a data,

i = varies from 1 to n

STATISTICS

II. THE ASSUMED MEAN METHOD

$$\overline{x} = a + \frac{\sum f_i d_i}{\sum f_i}$$

Where,

$\overline{x}$ = mean,

$\sum$ = Sum operation,

f_i = frequency of the observations of a data,

d_i = <u>deviations</u> = (midvalues − assumed mean)
       = $(x_i − a)$,

$f_i x_i$ = Product of midvalue of the observations and their corresponding frequencies

i = varies from 1 to n

III. THE STEP DEVIATION METHOD

$$\overline{x} = a + \left(\frac{\sum f_i u_i}{\sum f_i}\right) \times h$$

Where,

$\overline{x}$ = mean,

$\sum$ = sum operation,

a = assumed mean,

f_i = frequency of the observations of a data,

h = class size,

u_i = ratio of deviations and a common number = $\frac{x_i − a}{h} = \frac{d_i}{h}$,

i = varies from 1 to n

STATISTICS

4. Mode of grouped data:

A mode is that value among the observations which occurs most often, that is, the value of the observation having the maximum frequency.

A class with the maximum frequency, called the **modal class** and the mode is a value inside the modal class.

$$l + \left(\frac{f_1 - f_0}{2f_1 - f_0 - f_2} \right) \times h$$

Where,

l= lower limit of the modal class,

f_1= frequency of the modal class,

f_0= frequency of the class, preceding the modal class,

f_2= frequency of the class succeeding the modal class,

h= class size (assuming all class sizes to be equal)

5. The cumulative frequency of a class is the frequency obtained by adding the frequencies of all the classes preceding the given class.

6. Median of grouped data:

A median is a value that is positioned exactly in the middle of the set of observations arranged in an ascending or descending order. Hence, it separates the higher half of a data from the lower half.

A class whose cumulative frequency is greater than (and nearest to) $\frac{n}{2}$, i.e., half of the total frequency of a data, is called the **median class** and the median is a value inside the median class.

$$l + \left(\frac{\frac{n}{2} - cf}{f} \right) \times h$$

STATISTICS

Where,

l = lower limit of the modal class,

$n = \sum f\mathbf{i}$ = Total frequency,

cf = Cumulative frequency of class preceding the median class,

f = Frequency of the median class,

h = class size (assuming all class sizes to be equal)

7. Cumulative frequency distribution can be represented graphically as a cumulative frequency curve (an ogive) of a less than type and of the more than type.

8. The median of a grouped data can be obtained graphically as the x-coordinate of the point of intersection of the two ogives ('Less than' & 'More than') for a data.