3.3 MEASURES OF CENTRAL TENDENCY

3.3.1 Measures of Central Tendency:

Definition: - Given a series of observations measured on a quantitative variable, there is a general tendency among the values to cluster around a central value. Such clustering is called central tendency and measures these tendency are called measures of central tendency or averages.

Average: - Average is a value which is the representative of a set of data.

3.3.2 Functions or objects of an average:

[i] It facilitates quick understanding of complex data.

[ii] It facilitates comparison

[iii] To know about the universe from the sample

[iv] To get the single value that describes the characteristic of the entire group.

3.3.3 Requisites of a good average:

[i] It should be easy to understand

[ii] It should be easy to calculate.

[iii] It should be based on all the observations of the data.

[iv] It should not be affected by the extreme values.

[v] It should be strictly defined, so that it has one and only one interpretation.

[vi] It should be capable of further algebraic treatment.

[vii] It's definition should be in the form of a mathematical formula.

[viii] It should have sampling stability.

3.3.4 Types of averages (or) Measures of central tendency:-

The following are the important types of averages.

[1] Arithmetic mean.

[2] Median

[3] Mode

The first one is called **'mathematical average'** where as other two are called **'measures of location'** or **'measures of position'** or **'positional averages'**

3.3.5 [1] Arithmetic Mean:-

The most popular and widely used measure of representing the entire data by one value is what a layman call an 'average' and what the statisticians called as 'arithmetic mean'.

[A] Simple arithmetic mean:

Case-(i) Calculation of simple arithmetic mean –ungrouped data:

The process of computing mean in case of ungrouped data (i.e. where frequencies are not given) is very simple. Add together the various values of the variable and divide the total by the no of items.

Direct method:

If $x_1, x_2, x_3, ..., x_n$ are 'n' individual observed values of a variable X, then the A.M. is denoted by $\bar{x}$ and is defined as

$$\bar{x} = \frac{x_1 + x_2 + x_3 + ... + x_n}{n}$$

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

Short cut method:

Under this method the formula for calculating mean is

$$\bar{x} = A + \frac{\sum_{i=1}^{n} d_i}{n}$$

Where,

A= assumed mean, n = Number of observations

d_i = deviations of items taken from the assumed mean,

Note:

Any value whether existing in the data or not can be taken as the assumed mean and the final answer would be the same. However, it's better to take assumed mean nearer to the actual mean for lesser calculations.

Case-(ii) Calculation of A.M -Discrete series (or) ungrouped frequency distribution:

Direct method:

If $x_1, x_2, x_3, ..., x_n$ are 'n' individual observed values of a variable X occurred with frequencies $f_1, f_2, f_3, ..., f_n$ then the arithmetic mean is defined as

$$\bar{x} = \frac{\sum_{i=1}^{n} f_i x_i}{\sum_{i=1}^{n} f_i} \quad \text{or} \quad \bar{x} = \frac{\sum_{i=1}^{n} f_i x_i}{N}$$

Short cut method:

Under this method the formula for calculating mean is

$$\bar{x} = A + \left[\frac{\sum_{i=1}^{n} f_i d_i}{\sum_{i=1}^{n} f_i}\right] \quad \text{or} \quad \bar{x} = A + \left[\frac{\sum_{i=1}^{n} f_i d_i}{N}\right]$$

Where,

A= assumed mean; d_i = deviations of items taken from the assumed mean, n = number of observations f_i = frequency of the i^{th} observation.

Case-(iii) Continuous series (or) Grouped frequency distribution:

Direct method:

If $x_1, x_2, x_3, ..., x_n$ are 'n' mid points of the classes and $f_1, f_2, f_3, ..., f_n$ be the corresponding frequencies then the arithmetic mean is defined as

$$\bar{x}=\frac{\sum_{i=1}^{n} f_i x_i}{\sum_{i=1}^{n} f_i} \quad \text{or} \quad \bar{x}=\frac{\sum_{i=1}^{n} f_i x_i}{N}$$

Where, x_i is the mid- point of the i^{th} class.

Short cut method:

Under this method the formula for calculating mean is

$$\bar{x}= A+\left[\frac{\sum_{i=1}^{n} f_i d_i}{\sum_{i=1}^{n} f_i}\right]$$

Where, A=assumed mean $d_i = x_i - A$

n = number of observations

f_i = frequency of the ith observation.

Step deviation method:

Under this method the formula for calculating mean is

$$\bar{x}= A+\left[\frac{\sum_{i=1}^{n} f_i d_i}{\sum_{i=1}^{n} f_i}\right]\times h$$

Where, $d_i = \frac{x_i - A}{h}$

A =assumed mean

h = class width

n = number of observations

f_i = frequency of the i^{th} observation.

3.3.5 (i) (a)Properties of arithmetic mean:

(i) The sum of the deviations of the observation taken from their arithmetic mean is zero.

$$\text{i.e. } \sum_{i=1}^{n}\left(x_i - \bar{x}\right) = 0$$

(ii) The sum of the squares of the deviations of the observations taken from arithmetic mean is minimum.

$$\text{i.e. } \sum_{i=1}^{n}\left(x_i - \bar{x}\right)^2 \text{ is minimum}$$

$$\text{i.e. } \sum_{i=1}^{n}\left(x_i - \bar{x}\right)^2 < \sum_{i=1}^{n}\left(x_i - A\right)^2 \text{, Where A is any arbitrary value.}$$

(iii) If a series of n observations consists of two components having n_1 and n_2 observations ($n_1 + n_2 = n$), and means $\bar{x}_1$ and $\bar{x}_2$ respectively then the Combined mean $\bar{x}_C$ of n observations is given by

$$\text{Combined mean } \bar{x}_C = \frac{n_1\bar{x}_1 + n_2\bar{x}_2}{n_1 + n_2}$$

(iv) Let the values of X-series be $x_1, x_2, x_3, \ldots, x_n$ and values of Y-series be $y_1, y_2, y_3, \ldots, y_n$ and define $z_i = x_i \pm y_i$ then $\bar{z} = \bar{x} \pm \bar{y}$.

3.3.5 (i) (b) Merits of Arithmetic Mean:

[i] Arithmetic mean is most popular among averages used in statistical analysis.

[ii] It is very simple to under -stand and easy to calculate.

[iii] The calculation of A.M. is based on all the observations in the series.

[iv] The A.M. is responsible for further algebraic treatment.

[v] It is strictly defined.

[vi] It provides a good means of comparison.

[vii] It has more sampling stability.

3.3.5 (i) (c) Demerits of Arithmetic Mean:

[i] The A.M. is affected by the extreme values in a series.

[ii] In case of a missing observation in a series it is not possible to calculate the A.M.

[iii] In case frequency distribution with open end classes the calculation of A.M. is theoretically impossible.

[iv] The arithmetic mean is an unsuitable average for qualitative data.

3.3.6 Weighted arithmetic mean:

In calculating simple arithmetic mean, it is assumed that all the items in the series carry equal importance, but in practice, there are many cases where relative importance should be given to different items. Hence the limitation of not giving equal importance in case of simple arithmetic mean can be eliminated by giving relative importance i.e. weights by computing weighted arithmetic mean.

Formula:

If $w_1, w_2, w_3, \ldots\ldots, w_n$ are weights of n observations in a series $x_1, x_2, x_3, \ldots, x_n$ then the weighted mean is calculated as

$$\bar{x}_w = \frac{\sum_{i=1}^{n} f_i x_i}{\sum_{i=1}^{n} f_i}$$

Note: If the weights of all the observations are equal i.e. $w_1 = w_2 = w_3 = \ldots = w_n = w$. Then the weighted A.M is equal to simple A.M i.e. $\bar{x}_w = \bar{x}$.

3.3.7 [2] MEDIAN

The median is the middle most or central value of the observations made on a variable when the values are arranged either in ascending order or descending order.

As distinct from the arithmetic mean which is calculated from each and every item in the series, the median is what is called ‘positional average’. The term position refers to the place of value in a series. The place of the median in a series is such that an equal number of items lie on either side of it, i.e. it splits the observations into two halves.

Definition: Median is the value which divides it into two equal parts.

3.3.8. (i) (a) Calculation of median – Individual Data:

Step-1: Arrange the data in ascending order of magnitude.

Step-2: Case-(i) If the number of observations is odd then median is the $\left(\frac{n}{2}+1\right)^{th}$ observation in the arranged order.

Case-(ii) If the number of observations is even then the median is the mean of $\left(\frac{n}{2}\right)^{th}$ and $\left(\frac{n}{2}+1\right)^{th}$ observations in the arranged order.

3.3.8. (i) (b) Calculation of median-Discrete series:

Step-i: Arrange the data in ascending order of magnitude.

Step-ii: find out the cumulative frequency (c.f.)

Step-iii: Apply the formula: Median = size of $\frac{N+1}{2}$

Step-iv: Now look at the cumulative frequency column and find that total which is either equal to $\frac{N+1}{2}$ or next higher to that and determine the value of the variable corresponding to it. That gives the value of median.

3.3.8 (i) (c) Computation of median-Continuous series:

The median of a continuous series can be calculated.

$$\text{Median} = l + \frac{h}{f_k}\left(\frac{N}{2} - F_{k-1}\right)$$

Where,

l =lower limit of the median class.

f_k = frequency corresponding to the median class

$\frac{N}{2} = \frac{\text{Total frequency}}{2}$

F_{k-1} = cumulative frequency of the class preceding to the median class.

3.3.8 (i) (d) Mathematical property of median:

The sum of the deviations of the items from median, ignoring signs is the least.

$$\text{i.e } \sum_{i=1}^{n} |x_i - \text{median}| \text{ is least.}$$

3.3.8 (i) (e) Merits:

[i] The median is useful in case of frequency distribution with open-end classes.

[ii] The median is recommended if distribution has unequal classes.

[iii] Extreme values do not affect the median as strongly as they affect the mean.

[iv] It is the most appropriate average in dealing with qualitative data.

[v] The value of median can be determined graphically.

[vi] It is easy to calculate and understand.

3.3.8 (i) (f) Demerits:

[i] For calculating median it is necessary to arrange the data, where as other averages do not need arrangement.

[ii] Since, it is a positional average its value is not determined by all the observations in the series.

[iii] Median is not capable for further algebraic calculations.

[iv] The sampling stability of the median is less as compared to mean.

3.3.8 (i) (g) Determining the median graphically:

Median can be determined graphically by applying any one of the following methods.

Method-1:

Step-1: Draw two Ogives- one by less than method and other by more than method.

Step-2: From the point where these both curves intersect each other draw a perpendicular on the X-axis.

Step -3: The point where this perpendicular touches the X-axis gives the value of median.

Method-2:

Step-1: Draw only one ogive by less than method or more than method by taking variable on the X-axis and frequency on the Y-axis.

Step-2: Determine the median value by the formula median = size of $\frac{N}{2}$ th item.

Step-3: Locate this value on the Y-axis and from it draw a perpendicular on the ogive

Step -4: The point where this perpendicular touches the X-axis gives the value of median.

Note: The other partition values like quartiles, deciles, etc can be also determined graphically by this method no. II

3.3.9. [3] MODE:

The mode or the modal value is that value in which a series of observations occurs with the greatest frequency.

For example: The mode of the series 3, 5, 8, 5, 4, 5, 9, 3 would be 5.

In certain cases there may not be a mode or there may be more than one mode.

[i] 40, 44,57,78,84 (no mode)

[ii] 3, 4, 5, 5, 4, 2, 1 (modes 4 and 5)

[iii] 8, 8, 8, 8, 8 (no mode)

A series of data which having one mode is called 'unimodal' and a series of data which having two modes are called 'bimodal'. It may also have several modes and be called 'multimodal'.

3.3.10 (3) (a) Calculation of mode – discrete series:

(a) Simple inspection method:

In a discrete series the value of the variable against which the frequency is the largest, would be the modal value.

Age	5	7	10	12	15	18
No. of Boys	4	6	9	7	5	3

From the above data we can clearly say that mode is 10 because 10 has occurred maximum number of times i.e. 9.

(b) Grouping and Analysis table method:

This method is practically applied when the below problems occurs.

When the difference between the maximum frequency and the frequency preceding it or succeeding it is very small.

Process:

In order to find mode, a grouping table and an analysis table are to be prepared in the following manner.

(c) Grouping Table:

A grouping table consists of 6 columns.

(i) Arrange the values in ascending order and write down their corresponding frequencies in the column-1.

In column-2 the frequencies are grouped into two's and added.

(ii) In column-3 the frequencies are grouped into two's, leaving the first frequency and added.

(iii) In column-4 the frequencies are grouped into three's, and added.

(iv) In column-5 the frequencies are grouped into three's, leaving the first frequency and added.

(v) In column-6 the frequencies are grouped into three's, leaving the first and second frequencies and added.

(vi) Now in each these columns mark the highest total with a circle.

(d) Analysis Table:

After preparing a grouping table, prepare an analysis table. While preparing this table take the column numbers as rows and the values of the variable as columns. Now for each column number see the highest total in the grouping table (Which is marked with a circle) and mark the corresponding values of the variable to which the frequencies are related by using bars in the relevant boxes. Now the value of the variable (class) which gets highest number of bars is the modal value (modal class).

3.3.10 (3) (b) Calculation of mode – continuous series:

In a continuous series, to find out the mode we need one step more than those used for discrete series. As explained in the discrete series, modal class is determined by inspection or by preparing grouping and analysis tables. Then we apply the following formula.

$$\text{Mode}\left(M_O\right) = l + \left[\frac{f_1 - f_0}{2f_1 - f_0 - f_2}\right] \times h$$

Where,

l = lower limit of the modal class.

f_1 = Corresponding frequency of the modal class.

f_0 = Preceding frequency of the modal class.

f_2 = Succeeding frequency of the modal class.

h = width of the class.

Note:

[i] While applying the above formula for calculating mode, it is necessary to see that the class intervals are uniform through- out. If they are unequal they should first made equal on the assumption that the frequencies are equally distributed through -out.

[ii] In case of bimodal distribution the mode can't be found.

Finding mode in case of bimodal distribution:

In a bimodal distribution the value of mode can't be determined by the help of the above formula. In this case the mode can be determined by using the empirical relation given below.

Mode = 3Median - 2Mean

And the mode which is obtained by using the above relation is called **'Empirical mode'**

3.3.10 (3) (c) Locating mode graphically:

In a frequency distribution the value of mode can be determined graphically using following steps.

[i] Draw a histogram of the given data.

[ii] Draw two lines diagonally in the inside of the modal class bar, starting from each upper corner of the bar to the upper corner of the adjacent bars.

[iii] Draw a perpendicular line from the intersection of the two diagonal lines to the X-axis which gives the modal value.

Note:

The graphical method of determining mode can be used only where the data is unimodal.

3.3.10 (3) (d) Merits:

(i) It is easy to calculate and simple to understand.

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

(iii) The value of mode can be determined graphically.

(iv) Its value can be determined in case of open-end class interval.

(v) The mode is the most representative of the distribution.

3.3.10 (3) (e) Demerits:

(i) It is not suitable for further mathematical treatments.

(ii) The value of mode can- not always be determined.

(iii) The value of mode is not based on each and every items of the series.

(iv) The mode is strictly defined.

(v) It is difficult to calculate when one of the observations is zero or the sum of the observations is zero.

3.3.10 (3) (f) Empirical relation between Mean, Median and Mode:

The relationship between mean, median and mode depends upon the nature of the distribution. A distribution may be symmetrical or asymmetrical.

In symmetrical distribution the mean, median and mode are equal

i.e. **Mean(AM) = Median(M) = Mode(Mo)**

In a highly asymmetrical distribution it is not possible to find a relationship among the averages. But in a moderately asymmetric distribution the difference between the mean and mode is three times the difference between the mean and median.

$$\textbf{Mean} - \textbf{Mode} \cong 3(\textbf{Mean} - \textbf{Median})$$

$$\textbf{Mode} \cong 3\ \textbf{median} - 2\ \textbf{mean}$$

3.3.11 Partition Values

3.3.11.1 Quartiles:

Quartile is that value which divides the total distribution into four equal parts. So there are three quartiles, i.e. Q_1, Q_2 and Q_3 . Q_1, Q_2 and Q_3 are termed as first quartile, second quartile and third quartile or lower quartile, middle quartile and upper quartile respectively. Q_1 (quartile one) covers the first 25% items of the series and it divides the first half of the series into two equal parts. Q_2 (quartile two) is the median or middle value of the series and Q_3 (quartile three) covers 75% items of the series.

3.3.11.2 Calculation of Quartiles:

The calculation of quartiles is done exactly in the same manner as it is in case of the calculation of median.

In case of Individual and Discrete Series:

$$Q_i = \text{Size of} \left(\frac{i(N+1)}{4}\right)^{th} \text{item of the series;} \qquad i = 1, 2, 3$$

In case of continuous series:

$$Q_i = \text{Size of} \left(\frac{iN}{4}\right)^{th} \text{item of the series;} \qquad i = 1, 2, 3$$

Interpolation formula for continuous series:

$$Q_i = l + \frac{h}{f_k}\left(i\frac{N}{4} - F_{k-1}\right); \qquad i = 1, 2, 3$$

Where,

l=lower limit of the quartile class.

f_k = frequency corresponding to the quartile class

$\frac{N}{4}$, N = total frequency

F_{k-1} = cumulative frequency of the class preceding to the quartile class.

3.3.11.3 Deciles:

Deciles are those values which divide the series into ten equal parts. There are nine deciles i.e. D_1, D_2, D_3,.......,D_9 in a series and 5th decile is same as median and 2nd quartile, because those values divide the series in two equal parts.

3.3.11.4 Calculation of Deciles:

The calculation of deciles is done exactly in the same manner as it is in case of calculation of median.

In case of Individual and Discrete Series:

$$D_i = \text{Size of } \left(\frac{i(N+1)}{10}\right)^{th} \text{ item of the series;} \qquad i = 1, 2, 3, \ldots\ldots, 9$$

In case of continuous series:

$$D_i = \text{Size of } \left(\frac{iN}{10}\right)^{th} \text{ item of the series;} \qquad i = 1, 2, 3, \ldots\ldots\ldots, 9$$

Interpolation formula for continuous series:

$$D_i = l + \frac{h}{f_k}\left(\frac{iN}{10} - F_k\right); \quad i = 1, 2, 3, \ldots\ldots, 9$$

Where,

l = Lower limit of i^{th} decile class

F_k = Cumulative frequency preceding the i^{th} decile class

f_k = Frequency of i^{th} decile class.

3.3.11.5 Percentiles:

Percentiles are the values which divides the series into hundred equal parts. There are 99 percentiles i.e. P_1, P_2, P_3, ……..,P_{99} in a series, The value of 50th percentile = the value of 5^{th} decile = Value of median. The 50^{th} percentile divides the series into two equal parts. Similarly the value of $Q_1 = P_{25}$ and value of $Q_3 = P_{75}$; $Q_2 = P_{50}$

3.3.11.6 Calculation of Percentiles:

The calculation of percentiles is done exactly in the same manner as it is in the case of the calculation of median. The series must have to organize either in ascending or descending order, then necessary formula be put in order to get the percentile value.

In case of Individual and Discrete Series:

$$P_i = \text{Size of } \left(\frac{i(N+1)}{100}\right)^{th} \text{ item of the series, } i = 1, 2, 3, \ldots\ldots, 99$$

In case of continuous series:

$$P_i = \text{Size of } \left(\frac{iN}{100}\right)^{th} \text{ item of the series, } i = 1, 2, 3, \ldots\ldots, 99$$

Interpolation formula for continuous series:

$$P_i = l + \frac{h}{f_K}\left(\frac{iN}{100} - F_{k-1}\right), \qquad i = 1, 2, 3, \ldots\ldots, 99$$

Where,

l = Lower limit of i^{th} percentile class

F_{k-1} = Cumulative frequency preceding the i^{th} percentile class

f_k = Frequency of i^{th} percentile class.

3.3.11.7 Advantages of Quartiles, Deciles and Percentiles:

[i] These averages can be directly determined in case of open end class intervals without knowing the lower limit of lowest class and upper limit of the largest class.

[ii] These averages can be calculated easily in absence of some data in a series.

[iii] These averages are helpful in the calculation of measures of dispersion.

[iv] These averages are not affected very much by the extreme items.

[v] These averages can be located graphically.

3.3.11.8 Disadvantages of Quartiles, Deciles and Percentiles:

[i] These averages are not easily understood by a common man. These are not well defined and easy to calculate.

[ii] These averages are not based on all the observations of a series.

[iii] These averages cannot be computed if items are not given in ascending

or descending order.

[iv] These averages are affected very much by the fluctuation of sampling.

[v] The computation of these averages is not so easy in case of continuous series as the formula of interpolation is to be used.

3.3.12 Geometric mean:

Geometric mean is defined as the nth root of the product of n items (or) values.

3.3.12.1 Calculation of G.M.- Individual series:

If $x_1, x_2, x_3, ..., x_n$ be n observations studied on a variable X, then the G.M of the observations is defined as

$$\text{G.M.} = \left(x_1 \times x_2 \times x_3 \times ..... \times x_n\right)^{\frac{1}{n}}$$

Applying log both sides

$$\log \text{G.M.} = \frac{1}{n}\log\left(x_1.x_2....x_n\right)$$

$$= \frac{1}{n}\left(\log x_i + \log x_2 + ...+ \log x_n\right)$$

$$= \frac{1}{n}\sum_{i=1}^{n}\log x_i$$

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

3.3.12.2 Calculation of G.M.- Discrete series:

If $x_1, x_2, x_3, ..., x_n$ be n observations of a variable X with frequencies $f_1, f_2, f_3, ..., f_n$ respectively then the G.M is defined as

$$\text{G.M.} = \left(x_1^{f_1} \times x_2^{f_2} \times x_3^{f_3} \times .... \times x_n^{f_1}\right)^{\frac{1}{N}} ....(*)$$

Where, $N = \sum_{i=1}^{n} f_i$ i.e. total frequency

Applying log on both sides in (*) we get,

$$\log G.M.=\frac{1}{N}\log\left(x_1^{f_1}\times x_2^{f_2}\times x_3^{f_3}\times....\times x_n^{f_1}\right)$$

$$\log G.M.=\frac{1}{N}\left(f_1\log x_1\times f_2\log x_2\times f_3\log x_3\times....\times f_n\log x_n\right)$$

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

3.3.12.3 Calculation of G.M-Continuous series:

In continuous series the G.M is calculated by replacing the value of x_i by the mid points of the class's i.e. m_i.

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

Where, m_i is the mid value of the i^{th} class interval.

3.3.12.4 Properties of G.M.:

[i] If G_1 and G_2 are geometric means of two components having n_1 and n_2 observations and G is the geometric mean of the combined series of n (n_1+n_2) values then

$$G = G_1^{W_1}\times G_2^{W_2}$$

Where, $w_1=\frac{n_1}{n_1+n_2}$ & $w_2=\frac{n_2}{n_1+n_2}$

3.3.12.5 Uses of G.M:

Geometrical Mean is especially useful in the following cases.

[i] The G.M is used to find the average percentage increase in sales, production, or other economic or business series.

For example, from 1992 to 1994 prices increased by 5%, 10% and 18% respectively, then the average annual income is not 11% which is calculated by A.M but it is 10.9 calculated by G.M.

[ii] It is most appropriate average when dealing with ratios, percentages and rate of increase between two periods.

[iii] It is applied when increase or decrease in time is proportional e.g. growth of population is proportional to the time, increase in bacterial population is proportional to the time and rate of interest.

[iv] G.M is theoretically considered to be best average in the construction of Index numbers.

3.3.12.6 Weighted Geometric Mean:

Like weighted Arithmetic mean, we can also find weighted geometric mean and the formula is given by

$$\text{G.M.} = \left(x_1^{w_1} \times x_2^{w_2} \times x_3^{w_3} \times \ldots\ldots \times x_n^{w_n} \right)^{\frac{1}{N}}$$

Where,

$N = \sum_{i=1}^{n} w_i$ i.e. total weight

Applying log both sides we get,

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

3.3.12.7 Merits of G. M.:-

[i] It is based on all the observations.

[ii] It is rigidly defined.

[iii] It is capable for further mathematical treatment, such as combined G.M. of two sets.

[iv] It is not unduly affected by extreme observations.

3.3.12.8 Demerits of G. M.:-

[i] The serious drawback of G.M., is, it is zero if any of the observations is zero.

[ii] It is not simple to understand and calculate.

[iii] It may be imaginary if some observations are negative. Therefore it is calculated only for the data containing positive values.

[iv] It is not applicable to qualitative data.

[v] It cannot be determined graphically.

[vi] It cannot be computed if frequency distribution included open and class.

[vii] It may not be an actual observation in the data.

3.3.13 Harmonic mean:

The harmonic mean (H.M.) is defined as the reciprocal of the arithmetic mean of the reciprocal of the individual observations.

3.3.13.1 Calculation of H.M. -Individual series:

If $x_1, x_2, x_3, ..., x_n$ be 'n' observations of a variable X then harmonic mean is defined as

$$\text{H. M.} = \frac{n}{\frac{1}{x_1} + \frac{1}{x_2} + ... + \frac{1}{x_n}}$$

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

3.3.13.2 Calculation of H.M. -Discrete series:

If $x_1, x_2, x_3, ..., x_n$ be 'n' observations occurs with frequencies $f_1, f_2, f_3, ..., f_n$ respectively then H.M is defined as

$$\text{H.M.} = \frac{\sum_{i=1}^{n} f_i}{\frac{f_1}{x_1} + \frac{f_2}{x_2} + ... + \frac{f_n}{x_n}} \quad ....(*)$$

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

3.3.13.3 Calculation of H.M. – Continuous series:

In continuous series H.M can be calculated by replacing mid values (m_i) in place of x_i's in the equation. Hence H.M is given by

$$\Rightarrow \text{H.M.} = \frac{\sum_{i=1}^{n} f_i}{\sum_{i=1}^{n} \frac{f_i}{m_i}}$$

Where, m_i is the mid- value of the i^{th} class interval

3.3.13.4 Uses of harmonic mean:

[i] H.M. is used in finding averages involving speed, time, price and ratios.

[ii] It is useful for computing the average rate of increase of profits of a concern or average speed at which a journey has been performed or the average price at which an article has been sold.

[iii] The rate usually indicates the relation between two different types of measuring units that can be expressed reciprocally.

[iv] The H.M. is used for the problems about work, time and rate, where the amount of work is held constant and the average rate is required, or in problems about total cost, number of persons and per capita cost is called for or in problems of similar nature involving rates.

[v] The arithmetic mean (A.M.), the geometric mean (G.M.) and the harmonic mean (H.M.) of a series of n observations are connected by the relation A.M. ≥ G.M. ≥ H.M.

3.3.13.5 Merits of H. M.:-

[i] It is based on all observations.

[ii] It is rigidly defined.

[iii] It is capable of further mathematical treatment.

3.3.13.6 Demerits of H.M.:-

[i] If any of the observation is zero, H.M. cannot be defined.

[ii] It is not simple to compute and easy to understand as compared to A.M.

[iii] It is not applicable to qualitative data.

[iv] It cannot be computed for frequency distribution with open end class.

3.3.13.7 Relation between A.M, G.M, and H.M:

The relation between A.M. G.M. and H.M. is given by

$\text{A.M.} \geq \text{G.M.} \geq \text{H.M.}$

Note: The equality condition holds true only if all the items are equal in the distribution.

Prove that if a and b are two positive numbers then $\text{A.M.} \geq \text{G.M.} \geq \text{H.M.}$

Solution:

Let a and b are two positive numbers then

The Arithmetic mean of a and b is $= \frac{a + b}{2}$

The Geometric mean of a and b is $= \sqrt{ab}$

The harmonic men of a and b is $= \frac{2ab}{a + b}$

Let us assume $\text{A.M.} \geq \text{G.M.}$

$\Rightarrow \frac{a + b}{2} \geq \sqrt{ab}$

$\Rightarrow a + b \geq 2\sqrt{ab}$

$\Rightarrow (a + b)^2 \geq 4ab$

$\Rightarrow (a-b)^2 \geq 0$

Which is always true.

$\text{A.M.} \geq \text{G.M.}$ ………………………… (1)

let us assume $\text{G.M.} \geq \text{H.M.}$

$\Rightarrow \sqrt{ab} \geq \frac{2ab}{a+b}$

$\Rightarrow a+b \geq 2\sqrt{ab}$

$\Rightarrow (a+b)^2 \geq 4ab$

$\Rightarrow (a-b)^2 \geq 0$

Which is always true.

$\therefore$ G.M. $\geq$ H.M. .... (2)

From (1) and (2) we get A.M. $\geq$ G.M. $\geq$ H.M.

3.3.13.8 Numerical Examples:-

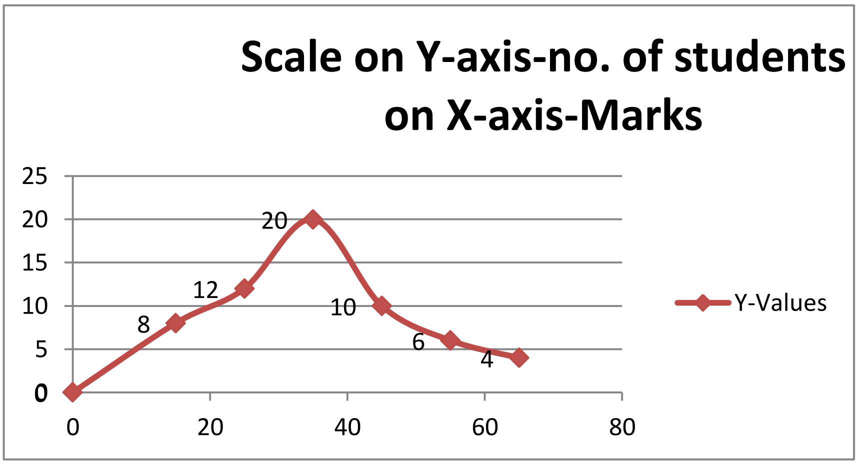
[1] From the following data of the marks obtained by 60 students of a class. Draw frequency polygon.

Marks	10-20	20-30	30-40	40-50	50-60	60-70
No. of students	8	12	20	10	6	4

Solution:

First convert class into mid-values then draw frequency curve

Marks	10-20	20-30	30-40	40-50	50-60	60-70
No. of students	8	12	20	10	6	4
Mid-values	15	25	35	45	55	65



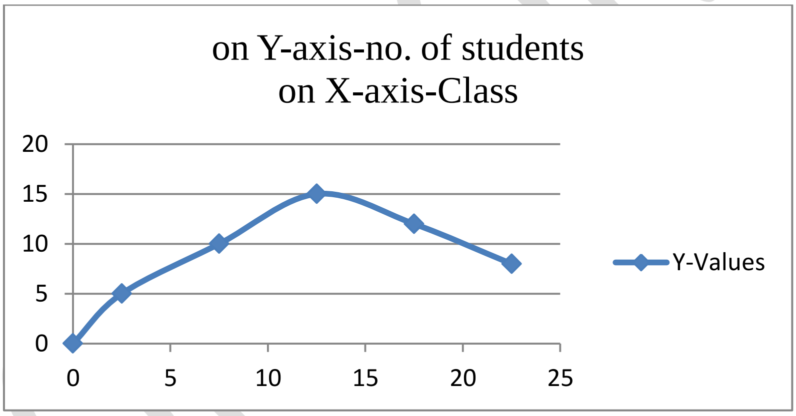
[2] From the following data of the class. Draw frequency curve.

class	0-5	5-10	10-15	15-20	20-25
No. of students	5	10	15	12	8

Solution:

First convert class into mid-values then draw frequency curve

class	2.5	7.5	12.5	17.5	22.5
No. of students	5	10	15	12	8



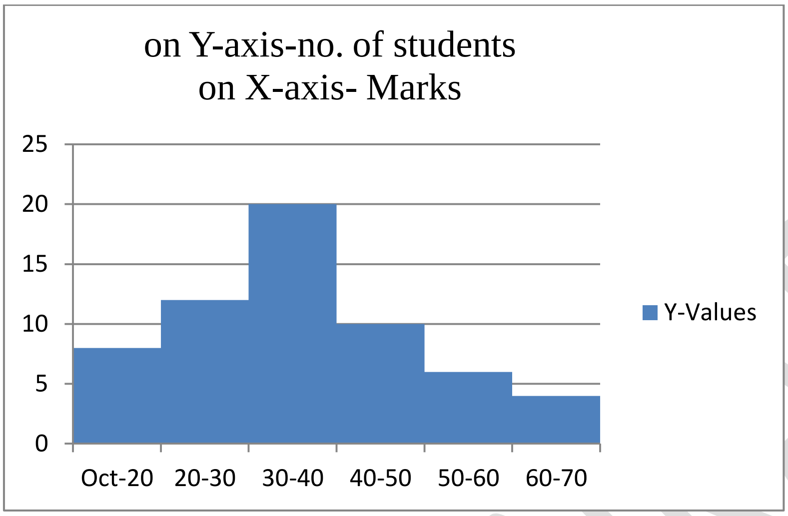
[3] From the following data of the marks obtained by 60 students of a class. Draw histogram

Marks	10-20	20-30	30-40	40-50	50-60	60-70
No. of students	8	12	20	10	6	4

Solution:

Marks	10-20	20-30	30-40	40-50	50-60	60-70

No. of students	8	12	20	10	6	4



[4] The frequency distribution of marks obtained by 100 students in F. Y. B. Com. is given below:

Marks	Less than 5	5-10	10-15	15-20	20-25	Total
No. of Students	10	24	30	20	16	100

Answer the following questions:

[i] State the type of classification

[ii] Identify the open end class

[iii] Find the class mark of 3^{rd} class

[iv] Find the class boundaries of 4^{th} class

[v] Find class width of 4^{th} class

[vi] Find the number of students getting marks less than 10

[vii] Find the percentage of students getting 10-15 marks.

Solution:

[i] Type of classification is

Answer: Exclusive

[ii] The open end class

Answer: Less than 5

[iii] Find the class mark of 3rd class

Mid-value = $\frac{\text{Upper limit} + \text{Lower limit}}{2}$

Mid-value = $\frac{15+10}{2}$

Mid-value = $\frac{25}{2}$

Mid-value =12.5

OR

Mid-value = $\frac{\text{Upper boundary} + \text{Lower boundary}}{2}$

In Exclusive classification boundary and limit are same.

[iv] Class boundaries of 4th class are

15-20

Answer: In Exclusive classification boundary and limit are same.

[v] Find class width of 5th class

Class width = (Upper limit of the succeeding class) – (Upper limit of the considering class)

Class width = 25 – 20

Class width = 5

[vi] Number of students getting marks less than 10 is

Answer: 10 + 24 = 34

[vii] Find the percentage of students getting 10-15 marks.

$= \frac{\text{No. of Students from 10-15}}{\text{Total no. of students}} \times 100$

$= \frac{30}{100} \times 100$

=30%

30% students getting 10-15 marks

[5] The following table gives the monthly income of 10 employees in an office. Income (in Rs): 1780, 1760, 1690, 1750, 1840, 1920, 1100, 1810, 1050, 1950. Calculate the A.M.

Solution:-

Calculation of arithmetic mean

Employee	**Income(in Rs) x**	**d=(x-1800)**
1	1780	20
2	1760	-40
3	1690	-110
4	1750	-50
5	1840	+40
6	1920	+120
7	1100	-700
8	1810	+10
9	1050	-750
10	1950	+150
n=10	$\sum x = 16,650$	$\sum d = -1350$

(i)Direct method:

$$\text{Mean} = \bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i = \frac{16,650}{10} = 1665$$

Thus the average income is Rs.1665.

(ii) Shortcut method:

$$\text{Mean } \bar{x} = A + \frac{1}{n}\sum_{i=1}^{n} d_i$$

Since A=1800

$\bar{x} = 1800 - \frac{1350}{10}$

$= 1800-135$

$\bar{x} = 1665.$

Thus the average income is Rs.1665.

[6] From the following data of the marks obtained by 60 students of a class. Calculate arithmetic mean

Marks	20	30	40	50	60	70
No. of students	8	12	20	10	6	4

Solution: Calculation of Arithmetic mean

Marks (x)	No. of students(f)	$f_i x_i$	d = (x-40)	$f_i d_i$
20	8	160	-20	-160
30	12	360	-10	-120
40	20	800	0	0
50	10	500	10	100
60	6	360	20	120
70	4	280	30	120
	N=60	$\sum fx =$ 2,460		$\sum fd = 60$

(i) **Direct method**:

Here,

N= total frequency = 60

$$\bar{x} = \frac{\sum_{i=1}^{n} f_i x_i}{\sum_{i=1}^{n} f_i}$$

$$\bar{x} = \frac{2460}{60} = 41$$

Hence the average marks are 41.

(ii) **Short cut method:**

$$\text{Mean} = A + \frac{\sum_{i=1}^{n} f_i d_i}{\sum_{i=1}^{n} f_i}$$

Since A = 40, $\bar{x} = 40 + \frac{60}{60} = 40 + 1 = 41.$

Hence the average marks are 41.

Calculation of Arithmetic mean

Marks	No. of students (f)	Mid-points (x_i)	$f_i\ x_i$	$d_i = \frac{x_i - A}{h}$ $d_i = \frac{x_i - 35}{10}$	Fidi
10-20	8	15	120	-2	-16
20-30	12	25	300	-1	-12
30-40	20	35	700	0	00
40-50	10	45	450	1	10
50-60	6	55	330	2	12
60-70	4	65	260	3	12
	N=60		$\sum f_i m_i =$ 2,160	$\sum \mathbf{d_i} = 3$	$\sum fd =$ 6

We know,

$$\bar{x} = A + \left[\frac{\sum_{i=1}^{n} f_i d_i}{\sum_{i=1}^{n} f_i} \right] \times h\ ;$$

$$\bar{x} = 35 + \frac{6}{60} \times 10$$

$$\bar{x} = 35 + \frac{60}{60}$$

$$\bar{x} = 35 + 1$$

$$\bar{x} = 36$$

[7] The mean monthly salary paid to all employees in a certain company was Rs. 4000. The mean monthly salaries paid to the male and female employees were Rs. 4200 and Rs. 3200 respectively. Obtain the percentage of male and female employees in the company.

Solution: Let n_1 be the number of male employees and n_2 the number of female employees in the company.

$$\bar{x}_C = 4000, \quad \bar{x}_1 = 4200, \quad \bar{x}_2 = 3200$$

$$\bar{x}_C = \frac{n_1\bar{x}_1 + n_2\bar{x}_2}{n_1 + n_2}$$

$$4000 = \frac{4200n_1 + 3200n_2}{n_1 + n_2}$$

$$4000(n_1 + n_2) = 4200n_1 + 3200n_2$$

$$4000n_1 + 4000n_2 = 4200n_1 + 3200n_2$$

$$4000n_1 - 4200n_1 = 3200n_2 - 4000n_2$$

$$-200n_1 = -800n_2$$

$$200n_1 = 800n_2$$

$$\frac{n_1}{n_2} = \frac{800}{200}$$

$$\frac{n_1}{n_2} = \frac{4}{1}$$

Male and female employees in the company are 4 and 1

Percentage of male

5 - 4

100-?

$$\text{Male percentage} = \frac{100 \times 4}{5}$$

$$\text{Male percentage} = 80\%$$

Percentage of female

5 - 1

100-?

$$\text{Female percentage} = \frac{100 \times 1}{5}$$

$$\text{Female percentage} = 20\%$$

[8] The means of two samples of sizes 50 and 100 with their means are 70 and 40 respectively. Obtain the combined mean.

Solution: Let n_1 be the size of first sample and n_2 be the size of second sample.

$$\overline{x}_c = ?,\ \overline{x}_1 = 70, \overline{x}_2 = 40, n_1 = 50, n_2 = 100$$

$$\overline{x}_c = \frac{50 \times 70 + 100 \times 40}{50 + 100}$$

$$\overline{x}_c = \frac{3500 + 4000}{150}$$

$$\overline{x}_c = \frac{7500}{150}$$

$$\overline{x}_c = 50$$

[9] The following table gives the monthly income of 10 employees in an office. Income (in Rs): 1780, 1760, 1690, 1750, 1840, 1920, 1100, 1810, 1050. Calculate the median.

Solution:

First arrange the data in ascending order

1050, 1100, 1690, 1750, 1760, 1780, 1810, 1840, 1920, 1950

Here, n = 9 i.e. odd then,

$$\text{Median} = \frac{\left(\frac{n+1}{2}\right)^{th} \text{observation}}{2}$$

$$\text{Median} = \frac{\left(\frac{9+1}{2}\right)^{th} \text{observation}}{2}$$

$$\text{Median} = \frac{\left(\frac{10}{2}\right)^{th} \text{observation}}{2}$$

$$\text{Median} = \frac{(5)^{th} \text{observation}}{2}$$

$$\text{Median} = \frac{1760}{2}$$

Median = 880

[10] The following table gives the monthly income of 10 employees in an office. Income (in Rs): 1780, 1760, 1690, 1750, 1840, 1920, 1100, 1810, 1050, 1950. Calculate the median.

Solution:

First arrange the data in ascending order

1050, 1100, 1690, 1750, 1760, 1780, 1810, 1840, 1920, 1950

Here, n = 10, even then,

$$\text{Median} = \frac{\left(\frac{n}{2}\right)^{th} \text{observation} + \left(\frac{n}{2}+1\right)^{th} \text{observation}}{2}$$

$$\text{Median} = \frac{\left(\frac{10}{2}\right)^{th} \text{observation} + \left(\frac{10}{2}+1\right)^{th} \text{observation}}{2}$$

$$\text{Median} = \frac{(5)^{th} \text{observation} + (5+1)^{th} \text{observation}}{2}$$

$$\text{Median} = \frac{(5)^{th} \text{observation} + (6)^{th} \text{observation}}{2}$$

$$\text{Median} = \frac{1760 + 1780}{2}$$

$$\text{Median} = \frac{3540}{2}$$

Median=1770

[11] From the following data of the marks obtained by 60 students of a class. Calculate median

Marks	20	30	40	50	60	70
No. of students	8	12	20	10	6	4

Solution:

Marks (x_i)	20	30	40	50	60	70
No. of students (f_i)	8	12	20	10	6	4
Cumulative frequency (c. f.)	8	20	40	50	56	60

$$\text{Median} = \text{size of } \frac{N+1}{2}$$

$$\text{Median} = \text{size of } \frac{60+1}{2}$$

$$\text{Median} = \text{size of } \frac{61}{2}$$

Median = size of 30.5

Which is just greater than 30.5 is 40.Value of the variable corresponding to 40 is 40. Hence, value of median is 40.

[12] From the following data of the marks obtained by 60 students of a class. Calculate median

Marks	10-20	20-30	30-40	40-50	50-60	60-70
No. of students	8	12	20	10	6	4

Solution:

Marks	10-20	20-30	30-40	40-50	50-60	60-70
No. of students	8	12	20	10	6	4
Cumulative frequency (c. f.)	8	20	40	50	56	60

$$\text{Median} = \text{size of } \frac{N}{2}$$

$$\text{Median} = \text{size of } \frac{60}{2}$$

$$\text{Median} = \text{size of } 30$$

Cumulative frequency which is, just greater than 30 is 40.

Median class of the variable corresponding to 40 is 30-40. Hence,

$$\text{Median} = l + \frac{h}{f_k}\left(\frac{N}{2} - F_{k-1}\right)$$

Where,

l =lower limit of the median class = 30

f_k = frequency corresponding to the median class = 20

$$\frac{N}{2} = \frac{\text{total frequency}}{2}$$

F_{k-1} = cumulative frequency of the class preceding to the median class=20

h = 10

$$\text{Median} = l + \frac{h}{f_k}\left(\frac{N}{2} - F_{k-1}\right)$$

$$\text{Median} = 30 + \frac{10}{20}(30 - 20)$$

$$\text{Median} = 30 + \frac{10}{20}(10)$$

$$\text{Median} = 30 + \frac{100}{20}$$

$$\text{Median} = 30 + 5$$

$$\text{Median} = 35$$

[13] The following table gives the monthly income of 10 employees in an office. Income (in Rs): 1780, 1760,1760, 1690, 1750, 1840, 1920, 1100, 1810, 1050, 1760, 1950. Calculate the mode.

Solution:

It can be clearly seen that observation 1760 has occurred maximum number of times i.e. 3

Hence, mode = Rs. 1760

[14] From the following data of the marks obtained by 60 students of a class. Calculate mode

Marks	20	30	40	50	60	70
No. of students	8	12	20	10	6	4

It can be clearly seen that observation (marks) 40 has occurred maximum number of times it means, it has maximum frequency i. e. 20

Hence, mode = 40 marks

[15] From the following data of the marks obtained by 60 students of a class. Calculate mode

Marks	10-20	20-30	30-40	40-50	50-60	60-70
No. of students	8	12	20	10	6	4

It can be clearly seen that observation (marks) 40 has occurred maximum number of times means, it has maximum frequency i. e. 20

Hence, corresponding to 20, variable class is 30-40, it is called modal class

Where,

l = lower limit of the modal class = 30

f_1 =frequency of the modal class = 20

f_0 = frequency of the class preceding to the modal class = 12

f_2 = frequency of the class succeeding to the modal class = 10

h = size of the class = 10

$$\text{Mode}\ (M_o) = l + \frac{f_1 - f_0}{2f_1 - f_0 - f_2} \times h$$

$$\text{Mode}\ (M_o) = 30 + \frac{20-12}{2 \times 20-12-10} \times 10$$

$$\text{Mode}\ (M_o) = 30 + \frac{8}{40-22} \times 10$$

$$\text{Mode}\ (M_o) = 30 + \frac{80}{18}$$

$$\text{Mode}\ (M_o) = 30 + 4.44$$

$$\text{Mode}\ (M_o) = 34.44$$

[16] The following is the frequency distribution of the number of tablets necessary to cure the persons affected by fever. The frequency against the class 20-24 is missing.

No. of tablets	No. of persons cured
4-8	11
8-12	13
12-16	16
16-20	14
20-24	X
24-28	9
28-32	17
32-36	6
36-40	4

However, it is known for the above data the arithmetic mean of the number of tablets required to cure fever is 19.92. Find the missing frequency and hence obtain the mode for the above data.

Solution:

No. of tablets	No. of persons cured f_i	Mid-points (x_i)	$\sum d_i = \frac{x_i - A}{h}$ $\sum d_i = \frac{x_i - 22}{4}$	$f_i d_i$
4-8	11	6	-4	-44
8-12	13	10	-3	-39
12-16	16	14	-2	-32
16-20	14	18	-1	-14
20-24	X	22	0	00
24-28	9	26	1	09
28-32	17	30	2	34
32-36	6	34	3	18
36-40	4	38	4	16

Here, A = 22;   N= 90 + x

$$\overline{x} = A + \left[\frac{\sum_{i=1}^{n} f_i d_i}{\sum_{i=1}^{n} f_i}\right] \times h \ ;$$

$$19.92 = 22 + \frac{(-52)}{90+x} \times 4$$

$$19.92 = 22 - \frac{208}{90+x}$$

$$19.92 - 22 = -\frac{208}{90+x}$$

$$-2.08 = -\frac{208}{90+x}$$

$$2.08 = \frac{208}{90 + x}$$

$$90 + x = \frac{208}{2.08}$$

$$90 + x = 100$$

$$x = 10$$

Thus missing frequency is 10

Now, Calculation of mode:

It can be clearly seen that observation 17 has occurred maximum number of times, means, it has maximum frequency i. e. 17

Hence, corresponding to 17, variable class is 28-32, it is called modal class

Where,

l = lower limit of the modal class = 28

f_1 =frequency of the modal class = 17

f_0 =frequency of the class preceding to the modal class = 9

f_2 = frequency of the class succeeding to the modal class = 6

h = size of the class = 4

$$\text{Mode } (M_o) = l + \frac{f_1 - f_0}{2f_1 - f_0 - f_2} \times h$$

$$\text{Mode } (M_o) = 28 + \frac{17-9}{2 \times 17-9-6} \times 4$$

$$\text{Mode } (M_o) = 28 + \frac{8}{34-15} \times 4$$

$$\text{Mode } (M_o) = 28 + \frac{32}{19}$$

$$\text{Mode } (M_o) = 28 + 1.68$$

$$\text{Mode } (M_o) = 29.68 \text{ Rs.}$$

[17] Find the geometric mean of 2,8,20,62,54

Solution: G.M.= $(2.8.20.62.54)^{1/5}$= $(1071360)^{1/5}$=16.0689

Alternative Method: Using logarithms computations will be as follows:

X_i	2	8	20	62	54
Log x_i	0.3010	0.9031	1.3010	1.7924	1.7324

G= Antilog = = 16.068

[18] A xyz co. bank gave interest at the rate of 10% p. c. p. a. on affixed deposit for the first year. In the second year and third year the rate of interest was 12% and 15% respectively. If the amount is compounded yearly, find the average rate of interest.

Solution: In this case geometric mean is appropriate. Let x = rate of interest

Y = current year amount + previous year amount = +1

y is relative change in amount. In other words y indicates per unit change in amount due to interest.

X	10	12	15
Y	1.10	1.12	1.15

Average rate of interest = (G.M of y -1) x 100%

$$= (1.1\text{x } 1.12\text{x } 1.15)^{1/3} -1) \text{ x } 100\%$$

$$= (1.1231-1) \text{ x } 100 = 12.31\%$$

Therefore average rate of interest is 12.31%

[19] A variable takes values $a, ar, ar^2,...,ar^{n-1}$. Find geometric mean

Solution: By definition

$$G = a \times ar \times ar^2 \times ,..., \times ar^{n-1}$$

$$G = \left[a^n \ r^{1+2+3..+(n-1)1/n} \right]$$

$$G = \left[a^n r^{n(n-1)/2} \right]^{\frac{1}{n}}$$

$$G = ar^{(n-1)/2}$$

Note: G.M. of all the four terms in geometric progression is same as that of the first term and the last term.

In the above example,

$$\text{G. M. of } \left(a \times ar \times ar^2 \times ... \times ar^{n-1}\right) = \text{G. M. of } \left(a \times ar^{n-1}\right)$$

$$= \text{G. M. of } \left(ar^{(n-1)/2}\right)$$

[20] A population of a certain city is given below. Obtain the average increase in population between the period 1991-1994.

Year	1991	1992	1993	1994
Population (in Lakhs)	15.22	15.01	15.97	16.45

Solution: Let x = population y =   y is relative change or per unit change in population.

Year	X	Y
1992	15.61	15.6115.22 = 1.0256
1993	15.97	15.97 15.61 = 1.0231
1994	16.45	16.45 15.97 = 1.0300

Average percent increase in x $= \left(\text{G.M. of y -1}\right)$ 100

$= \left(1.0256\ 1.0231\ 1.03\right)^{\frac{1}{2}}$ -1) 100

$= \left[\left(1.0808\right)^{\frac{1}{3}} -1\right] 100$

$= \left(1.0262-1\right)100$

=2.62%

Therefore average increase in population is 2.62% per year

Use of Geometric Mean:

Arithmetic mean is an important and widely used average, whereas geometric mean is not much in use. However, in certain situations geometric mean is more appropriate. The following are situations where G.M. is preferred.

[i] Average change in percent.

[ii] Average of bank interest rates.

[iii] Average of depreciation in the cost of a certain machine.

[iv] Average of population growth.

[v] Average rate of returns on share.

In general, G.M. is appropriate if the values are rations or percentages. Similarly if the values are approximately in geometric progression then also F.M. is proper average to find rate of growth. Due to some mathematical properties of G.M., it is popularly used in the construction of index numbers.

[21] Find the Q_1 and Q_3 of the following:

(a) 4, 5, 6, 7, 8, 9, 12, 13, 15, 10, 20

(b) 100, 500, 1000, 800, 600, 400, 7000 and 1200

Answer:

(a) Values of the variable are in ascending order:

i.e. 4, 5, 6, 7, 8, 9, 10, 12, 13, 15, 20, So N = 11 (No. of Values)

$$Q_1 = \text{Size of} \left[\frac{(N+1)}{4}\right]^{th} \text{item of the series}$$

$$Q_1 = \text{Size of} \left[\frac{(11+1)}{4}\right]^{th} \text{item of the series}$$

$$Q_1 = \text{Size of} \left[\frac{12}{4}\right]^{th} \text{item of the series}$$

Q_1 = size of 3rd item is 6

Now,

$$Q_3 = \text{Size of} \left[\frac{3(N+1)}{4}\right]^{th} \text{item of the series}$$

$$Q_3 = \text{Size of} \left[\frac{3(11+1)}{4}\right]^{th} \text{item of the series}$$

$$Q_3 = \text{Size of} \left[\frac{3 \times 12}{4}\right]^{th} \text{item of the series}$$

Q_3 = size of 9^{th} item is 13

$\therefore$ Required Q_1 and Q_3 are 6 and 13 respectively

(b) The values of the variable in ascending order are:

100, 400, 500, 600, 700, 800, 1000, 1200,   N = 8

$$Q_1 = \text{Size of } \left[\frac{(N+1)}{4}\right]^{th} \text{ item of the series}$$

$$Q_1 = \text{Size of } \left[\frac{(8+1)}{4}\right]^{th} \text{ item of the series}$$

$$Q_1 = \text{Size of } [2.25]^{th} \text{ item of the series}$$

$$Q_1 = \text{ size of } \{\text{Second item} + 0.25(\text{Third item - Second item})\}$$

$$Q_1 = \text{ size of } \{400 + 0.25(500 - 400)\}$$

$$Q_1 = 425$$

Now,

$$Q_3 = \text{Size of } \left[\frac{3(N+1)}{4}\right]^{th} \text{ item of the series}$$

$$Q_3 = \text{Size of } \left[\frac{3(8+1)}{4}\right]^{th} \text{ item of the series}$$

$$Q_3 = \text{Size of } \left[\frac{3\times 9}{4}\right]^{th} \text{ item of the series}$$

$$Q_3 = \text{Size of } [6.75]^{th} \text{ item of the series}$$

$$Q_3 = \text{ size of } \{\text{Sixth item} + 0.75(\text{Seventh item - Sixth item})\}$$

$$Q_3 = \text{ size of } \{800 + 0.75(1000 - 800)\}$$

$$Q_3 = 800+150$$

$$Q_3 = 950$$

$\therefore$ Required Q_1 and Q_3 are 425 and 950 respectively

Objective questions on measures of central tendency

Fill in the blanks

[1] The most frequently occurring value of a data set is called …

Answer: **Mode**

[2] The interquartile range is…..

Answer:**The difference between the third quartile and the first quartile**

[3] Sum of all observation divided by number of observation is known as…

Answer: **Mean**

[4] ………. quartile is also known as upper quartiles

Answer: **Third**

[5] ………. quartile is also known as lower quartiles

Answer: **First**

[6] The class having maximum frequency is called ………

Answer: **Modal class**

[7] Mode is found graphically by…………..

Answer: **Histogram**

[8] The series of observation which contains two modes is called

Answer: **Bimodal**

[9] The algebraic sum of the deviation of observations from their arithmetic mean is _________.

Answer: **Zero**

[10] The middle most value of set of observations is _________.

Answer: **Median**

[11] The value which occurs with the maximum frequency is called__________.

Answer: **Mode**

[12] The relationship between mean, median and mode for a symmetrical distribution is__________.

Answer: **Mean = Median = Mode**

[13] The relationship between mean, median and mode for a moderately asymmetrical distribution is__________.

Answer: $\mathbf{Mean - Mode \cong 3(Mean - Median)}$

[14] ________ is the reciprocal of the arithmetic mean of the reciprocal of the observations.

Answer: **Harmonic mean**

Choose the correct alternative from the following.

[1] The most frequently occurring value of a data set is called ..

(a) range (b) mode

(c) mean (d)median

Answer: (**b) mode**

[2] The interquartile range is….

(a) the 50th percentile

(b) another name for the variance

(c) the difference between the largest and smallest values

(d) difference between the third quartile and the first quartile

Answer: **(d)** difference between the third quartile and the first quartile

[3] Sum of all observation divided by number of observation is known as

(a) median (b) mean

(c) mode (d) none of these

Answer: (**b) mean**

[4] While computing mean of grouped data, we assume that the frequencies a re:

(a) evenly distributed over all the classes

(b) centred at the class-marks of the classes

(c) centred at the upper limits of the classes

(d) centred at the lower limits of the classes

Answer:**(b) centred at the class-marks of the classes**

[5] The mean of the following data is : 10, 20, 30, 40, 50 :

(a) 15 (b) 25

(c) 35 (d) 30

Answer: **(d) 30**

[6] Which of the following is not a measure of central tendency?

(a) Mean (b) Median

(c) Range (d) Mode

Answer: **(c) Range**

[7] Which of the following provides a measure of central location for the data?

(a) standard deviation (b) mean

(c) variance (d) range

Answer: (**b) mean**

[8] Which of the following measures of central tendency can have more than one value in a single sample?

(a) mean (b) median

(c) mode (d) none of the above

Answer: **(d) none of the above**

[9] What would happen to the mean if all of the scores were converted by

Subtracting10 points from each score?

(a) The mean would be unchanged.

(b) The mean would increase by 10 points.

(c) The mean would decrease by 10 points.

(d) The mean would decrease by an amount equal to 10 points divided by the number of participants.

Answer: **(c) The mean would decrease by 10 points.**

[10] In the assumed mean method, if A is the assumed mean, than deviation d_i is

(a) $x_i + A$ (b) $x_i – A$

(c) $A – x_i$ (d) none of these

Answer: **(b) $x_i – A$**

[11] When the given data arranged in increasing or decreasing order of their values, the middle most observation is called

(a) median (b) mean

(c) mode (d) none of these

Answer: **(a) median**

[12] Median is found graphically by

(a) frequency polygon (b) ogive curve

(c) simple bar diagram (d) histogram

Answer: (**b) ogive curve**

[13] Measures of central tendency are known as

(a) standard deviation (b) sum

(c) average (d) minimum

Answer: (**c) average**

[14] _________ quartile is also known as upper quartiles

(a) first (b) second

(c) third (d) forth

Answer: **(c) third**

[15] In case of open-ended classes, an appropriate measure of dispersion to be used is

(a) Range **(b) Quartile Deviation**

(c) Mean Deviation (d) Standard Deviation

Answer: **(b) Quartile Deviation**

[16] The median of a sample will always equal the

(a) mode (b) mean

(c) 50th percentile (d) all of the above

Answer: **(c) 50th percentile**

[17] The class having maximum frequency is called __________

(a) median class (b) median class

(c) modal class (d) none of these

Answer: **(c) modal class**

[18]__________ quartile is also known as lower quartiles

(a) first (b) second

(c) third (d) forth

Answer: **(a) first**

[19] The first quartile

(a) contains at least one third of the data elements

(b) is the same as the 25th percentile

(c) is the same as the 50th percentile

(d) is the same as the 75th percentile

Answer: **(b) is the same as the 25th percentile**

[20] Mode is found graphically by ______________

(a) frequency polygon (b) ogive curve

(c) simple bar diagram (d) histogram

Answer: (**d) histogram**

[21] Mode is:

(a) Middle most value (b) least frequent value

(c) most frequent value (d) none of these

Answer: **(c) most frequent value**

[22]The median is a measure of

(a) relative dispersion (b) absolute dispersion

(c) central location (d) relative location

Answer: **(c) central location**

[23] The 75th percentile is referred to as the

(a) first quartile (b) second quartile

(c) third quartile (d) fourth quartile

Answer: **(c) third quartile**

[24] If the class interval 10-20 then mid-value or mid -point of the class is

(a) 10 (b) 15 (c) 20 (d) 30

Answer: **(b) 15**

[25] The series of observation which contains two modes is called

(a) unimodal (b) bimodal

(c) tri-modal (d) none of these

Answer: **(b) bimodal**

[26] A researcher has collected the following sample data. 5, 12, 6, 8, 5, 6, 7, 5, 12, 4, 1 .The median is

(a) 5 (b) 6

(c) 7 (d) 8

Answer: (**b) 6**

[27] A researcher has collected the following sample data. 5, 12, 6, 8, 5, 6, 7, 5, 12, 4. The mode is

(a) 5 (b) 6

(c) 7 (d) NOTA

Answer: **(a) 5**

[28]Find out the mean and median of 1, 2, 3, 6, 8.

(a) 4, 6 (b) 4, 3

(c) 4, 2 (d) 4, 8

Answer: (**b) 4, 3**

[29] Find the mode of 3, 5, 6, 6, 5, 3, 5, 3, 6, 5, 3, 5, 7, 6, 5, 7, 5.

(a) 3 (b) 5

(c) 7 (d) 6

Answer:(**b) 5**

[30] Find the median and mode for the set of numbers 2, 2, 3, 6, 6, 6, 7, 8, 9.

(a) 3, 3 (b) 7, 7

(c) 8, 8 (**d) 6,6**

Answer:(**d)6,6**

[31] For what value of x, the mode of the following data is 27 ? 25, 26, 27, 23, 27, 26, 24, x, 27, 26, 25, 25.

(a) 24 (b) 25

(c) 26 (d) 27

Answer:(**d)27**

[32] Find the mode, median and mean of the following data.

15, 17, 16, 7, 10, 12, 14, 16, 19, 12, 16.

(a) 16, 15, 14 (b) 15, 16, 17

(c) 16, 17, 18 (d) 14, 16, 18

Answer: **(a) 16, 15, 14**

[33] Find the mode from the following distributions.

Marks	10	12	15	20	25	35	45	50	60
No. of Students	4	6	10	14	20	19	10	6	3

(a) 20 (b) 35

(c) 25 (d) 45

Answer: (**c) 25**

[34] The values of all items are taken into consideration in the calculation of-

(a) Mean (b) Median

(c) Mode (d) None

Answer: **(a) Mean**

[35] The average which is not effected by the extreme values is

(a) Mean (b) Median

(c) G.M (d) None

Answer**: (b) Median**

EXERCISE

THEORY QUESTIONS:

[1] What is central tendency?

[2] State two objectives of the measures of central tendency?

[3] Which average is called average of position?

[4] Where mode is considered as ill-defined?

[5] What do you mean by mode?

[6] Suggest appropriate averages for the following data

(i) Average size of shoes
(ii) Average number of children
(iii) Average height of students
(iv) Average speed of motor car
(v) Average changes in the price level

NUMERICAL EXAMPLES:

[1] Calculate AM of the following data.

(i) 4, 3, 2, 5, 3, 4, 5, 1, 7, 3, 2, 1 [3.33]

(ii) 30,70,10,75,500,8,42,250,40,36 [106.1]

(iii) 35, 46, 27, 38, 52, 44,50, 37, 41, 50[42]

[2] Find the A.M of first n natural numbers.

[3] Find the A.M of first n even numbers.

[4] Find the A.M of first n odd numbers.

[5] Find the A.M of first 10 even numbers.

[6] Find the A.M of first 100 odd numbers.

[7] Find A.M, G.M, H.M, median and mode of following

X	5	6	7
f	1	4	3

[8] Find A.M, G.M, H.M, median and mode of following data.

Marks	20	30	40	50	60	70
No. of Students	8	12	20	10	6	4

[9] Find A.M, G.M, H.M, median and mode of following

CI	0-10	10-20	20-30	30-40	40-50	50-60
Frequency	6	14	16	27	22	15

[10] Find A.M, G.M, H.M, median and mode of following data.

CI	20-25	25-30	30-35	35-40	40-45	45-50	50-55
Frequency	10	12	8	20	11	4	5

[11] Find A.M, G.M, H.M, median and mode of following data.

CI	10-20	20-40	40-70	70-120	120-200
Frequency	4	10	26	8	2

[12] Find A.M, G.M, H.M, median and mode of following data.

CI	1-7	8-14	15-21	22-28	29-35
Frequency	3	17	12	11	7

[13] In an examination marks secured by three students A, B, C along with the respective weights of the subjects are given below. Determine the best performance

Students	Math (wt-4)	Science (wt-3)	English (wt-2)	History (wt-1)
A	90	80	60	70
B	85	75	65	75
C	95	70	55	80

[14] Find A.M, G.M, H.M, median and mode of following data

Wages in rupees	Less than 10	Less than 20	Less than 30	Less than 40	Less than 50
No. of workers	5	17	20	22	25

[15] Find the missing frequencies from the data given below if mean is 60.

Marks	50	55	60	65	70	Total
No. of Students	?	20	25	?	10	100

[16] Find the missing frequencies from the data given below if mean is 60.

Marks	60-62	63-65	66-68	69-71	72-74
No. of Students	15	54	?	81	24

[17] In a class of 60 students 10 have failed with an average mark of 15. If the total marks of all the students were 1800, find the average marks of those who have passed?

[18] The average marks of HSC students are 68.4 and that of +2 students is 71.2. if the combined average of all these students be 70, find the ratio of the number of students in the HSC and +2.

[19] The mean of 10 observations is 10 and the sum of first four observations is 10. Find the 5^{th} observation.

[20] The average weight of group of 25 boys was calculated to be 50kg. It was later discovered that one weight was misread as 34 instead of 43. Calculate the correct average.

[21] The numbers 3,5,7 and 4 have frequencies x, (x+2), (x-2), (x+1) respectively. If the arithmetic mean is 4.424. Find the value of x.

[22] The mean height of 25 male workers in a factory is 61cms and the mean height of 35 female workers is 58 cms. Find the combined mean height of 60 workers in the factory.

[23] The average marks secured by 36 students were 50. But it was later discovered that an item 64 was misread as 46. Find the correct mean of marks.

[24] The mean of 50 iems was 80. It was later discovered that two items were wrongly taken as 23 and 24 in place of 32 and 42 respectively. Ascertain the correct mean.

[25] The mean of 100 items was 46. Later on it was discovered that two items 16 and 43 were misread as 61 and 34. It was also found that the number of items 80 and not 100. Find the correct mean.

[26] The mean age of a group of 100 persons was found to be 32. Later it was discovered that one age 57 was misread as 27. Find the correct mean.

[27] Find the G.M of

(i) 3, 6, 24 and 48

(ii) 2574, 475, 5, 0.8, 0.08,0.005, 0.009

(iii) 5,10,200,12375,2575

[28] Find the G.M following data.

Marks	10	20	30	40	50	60
No. of Students	12	15	25	10	6	2

[29] Find the average rate of growth of population which in the first decade has increased of 20%, in the second decade by 30% and in third by 45%.

[30] If n_1=2, n_2 = 3 and n_3=5 and GM_1=8, GM_2=10 and GM_3=15. find the combined geometric mean of all the observations.

[31] The G.M of 10 items of a series was 16.2. It was later found that one of the items was wrongly taken as 12.9 instead of 21.9. Calculate the correct G.M.

[32] If A.M and G.M of two numbers are 12.5 and 10 respectively. Find those numbers.

[33] If A.M and G.M of two numbers are 10 and 8 respectively. Find those numbers.

[34] Find the H.M of 3,4,12

[35] Find H.M of $\frac{1}{5}, \frac{1}{6}, \frac{1}{7}, \frac{1}{8}, \frac{1}{9}$

[36] A person travelled 20 km at 5kmph and again 24 km at 4kmph. Find his average speed.

[37] Find the average speed of a motor car which covers as follows. First 10 km at speed of 40 kmph. Second 20 km at a speed of 60 kmph. Third 30 km at a speed of 80 kmph.

[38] Calculate median, quartiles, 5th decile and 45th percentile of the following

(i) 391, 591, 407, 384, 1490, 2488, 672, 522, 753, 777

(ii) 31, 28, 49, 57, 31, 56, 27, 49

(iii) 16, 14, 11,11, 13, 10, 10, 9, 7, 7, 4, 3, 2, 1

[39] Find D_6 , P_{65} for the following data

10, 20, 25, 30, 35, 40, 50, 55, 60

[40] Find three quartiles, D_2 , P_5, P_{90} from the following data.

Marks	10	20	30	40	50	60
No. of Students	12	15	25	10	6	2

[41] Find three quartiles, D_2, P_5, P_{90} of following data.

CI	10-20	20-40	40-70	70-120	120-200

Frequency	4	10	26	8	2

[42] Find the modal marks of the following data.

Marks	0-9	10-19	20-29	30-39	40-49	50-59	60-69	70-79	80-89	90-99
No of students	6	29	87	181	247	263	133	43	9	2

[43] Mean of the 10 observations is 20. If each observation is increased by 5 what is the mean of the resultant series?

[44] Mean of the 5 observations is 10. If each observation is doubled then what is the mean of the new series.

[45] How mean, median is affected if every value of the variable is increased by 2 and multiplied by 5?

[46] The GM and HM of two observations are respectively 18 and 10.8. Find the observations.

[47] The arithmetic mean of 10 observations is 72.5 and the arithmetic mean of 9 observations is 63.2, find the value of 10^{th} observation.