3.4 MEASURES OF DISPERSION

3.4.1 Concept of Dispersion:-

3.4.1.1 Dispersion

The averages are representatives of a frequency distribution. But they fail to give a complete picture of the distribution. They do not tell anything about the scatterness of observations within the distribution.

The word dispersion means deviation or difference. In statistics refers to deviation of the values of a variable from their central value. Measures of dispersion indicate the extent to which individual observations vary from their averages i.e. mean, median or mode. It shows the spread of items of a series from their central value. This is known as variation or dispersion.

3.4.1.2 Definitions:

"Dispersion is the measure of variation of the variables about a central value"

OR

"Dispersion is a measure of the extent to which the individual observations vary"

OR

"The scatterness or variation of observations from their average are called the dispersion"

3.4.2 Requisites of a good average (or) characteristics for ideal measures of dispersion:

(i) It should be rigidly defined.

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

(iii) It should be based on all the observations of the series.

(iv) It should be used for further algebraic treatments.

(v) It should not be affected much by the sampling fluctuations.

(vi) It should not be affected by the extreme observations in the series.

3.4.2.1 Objectives or uses of dispersion:

(i) Measures of dispersion tell us whether an average is a true representative of the series or not.

(ii) The extent of variability between two or more series can be compared with the help of the measures of dispersion.

(iii) It is used to determine the degree of uniformity, reliability and consistency amongst two or more sets of data.

(iv) Measures of dispersion are used in the statistical measures like correlation, regression etc. for further analysis.

3.4.2.2 Absolute and relative measures of dispersion:

The dispersion of a series may be measured either absolutely or relatively. **If the dispersion is expressed in terms of the original units of the series, it is called absolute measure of dispersion.** The disadvantage of absolute measure of dispersion is that it is not suitable for comparative study of the characteristics of two or more series.

For example the income of workers may be in rupees, while their heights may be in inches. Thus a comparison to measure their variations cannot be made as both are in different units.

So for comparison point of view it is necessary to calculate the relative measures of dispersion which are expressed as percentage form (i.e. unit less number). These types of expressions are called coefficients of dispersion or coefficient of variation.

3.4.2.3 Measures of Dispersion:

The following are the important measures of dispersion.

[1] Range

[2] Quartile deviation or Semi inter-quartile range.

[3] Mean absolute deviation or Mean deviation

[4] Variance

[5] Standard deviation

3.4.3 Range:

The difference between the values of the greatest observation and the smallest observation in a distribution is called the range. It is denoted by R. It is given by

Range = Maximum observation -Minimum observation

OR

It is the simplest measure of dispersion. Range is defined as the difference between the two extreme values of the series i.e. it is the difference between the largest and smallest value of the series.

Symbolically

R = L – S, this is the absolute measure of dispersion.

Where,

R = Absolute value of range

L= largest value of the series

S= Smallest value of the series.

3.4.3.1 Coefficient of range:

The relative measure of range is known as Coefficient of range which can be calculated by the following formula.

$$\text{Coefficient of Range} = \frac{\text{L-S}}{\text{L+S}}$$

The Co-efficient of range is otherwise known as “the ratio of range” or “the coefficient of scatter”.

3.4.3.2 In case of grouped frequency distribution:

Range has nothing to do with the frequencies of a series. Grouped frequency distribution (i.e. continuous series), range is calculated by taking the difference between the upper limit of the largest class and the lower limit of the lowest class. However, range can also be calculated taking the difference between the midpoints of the largest and smallest classes. In practice both the methods are used. But results in both the cases are not the same.

3.4.3.3 Merits and Demerits of Range

Merits

[i] It is simple to understand.

[ii] It is easy to calculate.

[iii] In certain types of problems like quality control, weather forecasts, share price analysis etc. Range is most widely used.

Demerits

[i] It is very much affected by the extreme items.

[ii] It is based on only two extreme observations.

[iii] It cannot be calculated from open-end class intervals.

[iv] It is not suitable for mathematical treatment.

[v] It is a very rarely used measure.

3.4.4 Quartile deviation or Semi inter-quartile range.

This is the measure of dispersion. Quartile deviation is based on central 50% of the items. The quartile deviation or semi inter quartile range is defined as half the difference between the third quartile (Q_3) and first quartile (Q_1). Symbolically,

$$Q.D. = \frac{Q_3 - Q_1}{2} \quad \text{(Absolute measure)}$$

Where,

Q_1 and Q_3 are the first and third quartiles respectively and $Q_3 - Q_1$ is called 'Inter quartile range'.

3.4.4.1 Coefficient of Quartile deviation:

The relative measure of quartile deviation is known as coefficient of quartile deviation and is defined as

Coeficientf of Q.D.= $\frac{Q_3 - Q_1}{Q_3 + Q_1}$

Merits:

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

(ii) It is rigidly defined

(iii) It is not much affected by the extreme values of the series

(iv) It is easy to calculate in case of open-end series.

Demerits:

(i) It is not capable for further algebraic treatments.

(ii) It is too much affected by sampling fluctuations.

(iii) It is not based on all the observations of the series.

(iv) It does not show the scatter ness around any average.

3.4.5 Mean absolute deviation:

While calculating range and quartile deviation all the values does not take into account. For the purpose of calculation only two values are considered. That is why they are called positional average. To overcome these things, mean deviation and standard deviation have been developed.

Definition: The mean deviation or mean absolute deviation is defined as the arithmetic mean of the deviations of the items from any measure of central tendency i.e. mean, median, mode etc. ignoring $\pm$(algebraic) signs. Theoretically it is advantageous to take deviations of items from median. This is so because the sum of the deviations of items from median is minimum when plus and minus signs are ignored but in practice mean is frequently used for calculation of mean deviation.

3.4.5.1 Calculation of mean deviation- Individual series:

If $x_1, x_2, x_3, ..., x_n$ are 'n' individual observed values of a variable X with any central value A, then the mean deviation about any central value is defined as

$$\text{M.D.about } A = \frac{1}{n}\sum_{i=1}^{n}|x_i - A|$$

Where A is any central value i.e. mean or median or mode etc.

3.4.5.2 Calculation of mean deviation -Discrete series (or) ungrouped frequency distribution:

If $x_1, x_2, x_3, ..., x_n$ are 'n' individual observed values of a variable X occurred with frequencies $f_1, f_2, f_3, ..., f_n$ and with any central value A, then mean deviation about any central value is defined as

$$\text{M.D.about } A = \frac{1}{N}\sum_{i=1}^{n} f_i |x_i - A|$$

3.4.5.3 Continuous series (or) Grouped frequency distribution:

The calculation of mean deviation in case of grouped frequency distribution is same as the calculation of mean deviation in case of discrete frequency distribution. The only difference is to use mid points of the classes in place of values of the variable. Symbolically,

$$\text{M.D.about } A = \frac{1}{N}\sum_{i=1}^{n} f_i |m_i - A|$$

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

3.4.5.4 Coefficient of mean deviation:

The relative measure of mean deviation is known as 'coefficient of mean deviation' and is defined as

$$\text{Coefficient of M.D.} = \frac{\text{M.D.}}{A},$$

Where, A is central value from which the deviations are taken.

3.4.6 Variance:

Definitions:

The variance is defined as the arithmetic mean of the squares of the deviations of the observations from their arithmetic mean. It is denoted by σ^2 OR

Variance: The Square of the standard deviation is called variance and is denoted by σ^2.

3.4.7 Standard deviation:

The standard deviation is defined as the square root of the arithmetic mean of the squares of the deviations of the observations from their arithmetic mean. It is denoted by σ.

3.4.7.1 Formulae for variance and Standard deviation - Individual series:

If $x_1, x_2, x_3, ..., x_n$ are 'n' individual observed values of a variable X, with mean $\bar{x}$ then the variance is defined as

$$\sigma^2 = \frac{1}{n}\sum_{i=1}^{n}\left(x_i - \bar{x}\right)^2$$

it can also be written as

$$\sigma^2 = \frac{1}{n}\sum_{i=1}^{n}\left(x_i - \bar{x}\right)^2$$

$$\sigma^2 = \frac{1}{n}\sum_{i=1}^{n}x_i^2 - \bar{x}^2$$

Standard deviation is defined as

$$(\sigma) = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(x_i - \bar{x}\right)^2}$$

Where, $\bar{x}$ is mean of the series.

The formula can also extended as below

$$(\sigma)=\sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(x_i-\bar{x}\right)^2}$$

$$(\sigma)=\sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i^2-2x_i\bar{x}+\bar{x}^2)}$$

$$(\sigma)=\sqrt{\frac{1}{n}\sum_{i=1}^{n}x_i^2-2\bar{x}\frac{1}{n}\sum_{i=1}^{n}x_i+\frac{1}{n}\sum_{i=1}^{n}\bar{x}^2}$$

$$(\sigma)=\sqrt{\frac{1}{n}\sum_{i=1}^{n}x_i^2-2\bar{x}^2+\bar{x}^2}$$

$$(\sigma)=\sqrt{\frac{1}{n}\sum_{i=1}^{n}x_i^2-\bar{x}^2}$$

$$\therefore(\sigma)=\sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(x_i-\bar{x}\right)^2}=\sqrt{\frac{1}{n}\sum_{i=1}^{n}x_i^2-\bar{x}^2}$$

3.4.7.2 Steps for computation of standard deviation for Individual data.

(i) Find the mean of the given observation.

(ii) Find the deviation of each observation from the mean i.e. subtract the mean from each observation.

(iii) Square the deviation obtained in step 2.

(iv) Find the sum of square in step 3. This gives $\sum_{i=1}^{n}d_i^2$

(v) Divide $\sum_{i=1}^{n}d_i^2$ by the number of observations n

(vi) Find square root of the quantity in step 5.

(vii) The number so obtained is nothing but standard deviation.

3.4.7.3 Calculation of Standard deviation -Discrete series:

Direct method:

If $x_1,x_2,x_3,....,x_n$ are 'n' individual observed values of a variable X occurred with frequencies $f_1,f_2,f_3,....,f_n$ and with mean $\bar{x}$, then variance and Standard deviation is defined as

$$\text{Variance} = (\sigma^2) = \frac{1}{\sum_{i=1}^{n}f_i}\sum_{i=1}^{n}f_i\left(x_i-\bar{x}\right)^2=\frac{1}{N}\sum_{i=1}^{n}f_i\left(x_i-\bar{x}\right)^2$$

$$\text{Standard deviation}=(\sigma)=\sqrt{\frac{1}{\sum_{i=1}^{n} f_i}\sum_{i=1}^{n} f_i\left(x_i-\bar{x}\right)^2}=\sqrt{\frac{1}{N}\sum_{i=1}^{n} f_i\left(x_i-\bar{x}\right)^2}$$

Where,

$$\bar{x}=\frac{\sum f_i x_i}{\sum f_i}=\frac{\sum f_i x_i}{N}$$

This is the mean of the series.

The above formula can be extended to

$$\text{variance}=\sigma^2=\frac{\sum_{i=1}^{n} f_i\left(x_i-\bar{x}\right)^2}{\sum_{i=1}^{n} f_i}=\frac{\sum_{i=1}^{n} f_i x_i^{\,2}}{\sum_{i=1}^{n} f_i}-\bar{x}^2=\frac{\sum_{i=1}^{n} f_i x_i^{\,2}}{N}-\bar{x}^2$$

$$\text{Standard deviation}=\sigma=\sqrt{\frac{\sum_{i=1}^{n} f_i\left(x_i-\bar{x}\right)^2}{\sum_{i=1}^{n} f_i}}=\sqrt{\frac{\sum_{i=1}^{n} f_i x_i^{\,2}}{\sum_{i=1}^{n} f_i}-\bar{x}^2}=\sqrt{\frac{\sum_{i=1}^{n} f_i x_i^{\,2}}{N}-\bar{x}^2}$$

Where,

$$\bar{x}=\frac{\sum f_i x_i}{\sum f_i}=\frac{\sum f_i x_i}{N}$$

Short cut method:

In this method the formula for variance and standard deviation is given by

$$\text{Var}(x)=\sigma^2=\frac{\sum_{i=1}^{n} f_i d_i^{\,2}}{\sum_{i=1}^{n} f_i}-\left(\frac{\sum_{i=1}^{n} f_i d_i}{\sum_{i=1}^{n} f_i}\right)^2=\frac{\sum_{i=1}^{n} f_i d_i^{\,2}}{N}-\left(\frac{\sum_{i=1}^{n} f_i d_i}{N}\right)^2$$

$$\text{S.D.}(x)=\sigma=\sqrt{\frac{\sum_{i=1}^{n} f_i d_i^{\,2}}{\sum_{i=1}^{n} f_i}-\left(\frac{\sum_{i=1}^{n} f_i d_i}{\sum_{i=1}^{n} f_i}\right)^2}=\sqrt{\frac{\sum_{i=1}^{n} f_i d_i^{\,2}}{N}-\left(\frac{\sum_{i=1}^{n} f_i d_i}{N}\right)^2}$$

Where,

$d_i = x_i - A$; A = assumed mean; $\sum_{i=1}^{n} f_i = N$

3.4.7.4 Steps for computation of standard deviation for ungrouped data.

(i) Find the mean of the given observation.

(ii) Find the deviation of each observation from the mean i.e. subtract the mean from each observation.

(iii) Square the deviation obtained in step 2.

(iv) Find the sum of square in step 3. This gives $\sum_{i=1}^{n} d_i^2$

(v) Multiply frequencies and deviations then find $\sum_{i=1}^{n} f_i d_i$

(vi) Multiply frequencies and squared deviations then find $\sum_{i=1}^{n} f_i d_i^2$

(vii) Find square root of the quantity in step 5.

(viii) The number so obtained is nothing but standard deviation.

3.4.7.5 Continuous series (or) Grouped frequency distribution:

The calculation of variance and standard deviation in case of grouped frequency distribution is.

Step deviation method (or) change of origin and scale:

In this method the formula for variance and standard deviation is given by

$$\mathrm{Var}(x) = \sigma^2 = h^2 \times \left[\frac{\sum_{i=1}^{n} f_i u_i^2}{\sum_{i=1}^{n} f_i} - \left(\frac{\sum_{i=1}^{n} f_i u_i}{\sum_{i=1}^{n} f_i} \right)^2 \right] = h^2 \times \left[\frac{\sum_{i=1}^{n} f_i u_i^2}{N} - \left(\frac{\sum_{i=1}^{n} f_i u_i}{N} \right)^2 \right]$$

$$\mathrm{S.D.}(x) = \sigma = h \times \sqrt{\frac{\sum_{i=1}^{n} f_i u_i^2}{\sum_{i=1}^{n} f_i} - \left(\frac{\sum_{i=1}^{n} f_i u_i}{\sum_{i=1}^{n} f_i} \right)^2} = h \times \sqrt{\frac{\sum_{i=1}^{n} f_i u_i^2}{N} - \left(\frac{\sum_{i=1}^{n} f_i u_i}{N} \right)^2}$$

Where,

$d_i = \frac{x_i - A}{h}$; A = assumed mean; h = class interval

3.4.7.6 Steps for computation of standard deviation for grouped data.

(i) Find the mean of the given observation.

(ii) Find the deviation of each observation from the mean i.e. subtract the mean from each observation.

(iii) Square the deviation obtained in step 2.

(iv) Find the sum of square in step 3. This gives $\sum_{i=1}^{n} d_i^{\,2}$

(v) Multiply frequencies and deviations then find $\sum_{i=1}^{n} f_i d_i$

(vi) Multiply frequencies and squared deviations then find $\sum_{i=1}^{n} f_i d_i^{\,2}$

(vii) Find square root of the quantity in step 5.

The number so obtained is nothing but standard deviation.

3.4.7.7 Merits and Demerits of Standard Deviation

Merits

(i) It is rigidly defined and its value is always definite and based on all the observations and the actual signs of deviations are used.

(ii) As it is based on arithmetic mean, it has all the merits of arithmetic mean.

(iii) It is the most important and widely used measure of dispersion.

(iv) It is possible for further algebraic treatment.

(v) It is less affected by the fluctuations of sampling and hence stable.

(vi) It is the basis for measuring the coefficient of correlation and sampling.

Demerits

(i) It is not easy to understand and it is difficult to calculate.

(ii) It gives more weight to extreme values because the values are squared up.

(iii) As it is an absolute measure of variability, it cannot be used for the purpose of comparison.

3.4.7.8 Combined variance and combined Standard deviation

Suppose there are two groups. First is of size n_1 with arithmetic mean $\bar{x}_1$ and variance σ_1^2 similarly for second is of size n_2 with arithmetic mean $\bar{x}_2$ and variance σ_2^2. Then the variance and standard deviation of combined group of size $n_1 + n_2$ is given by

$$\text{combined variance} = \frac{n_1\sigma_1^2 + n_2\sigma_2^2 + n_1 d_1^2 + n_2 d_2^2}{n_1 + n_2}$$

$$\text{combined variance} = \frac{n_1\left(\sigma_1^2 + d_1^2\right) + n_2\left(\sigma_2^2 + d_2^2\right)}{n_1 + n_2}$$

$$\text{combined Standard deviation} = \sqrt{\left[\frac{n_1\sigma_1^2 + n_2\sigma_2^2 + n_1 d_1^2 + n_2 d_2^2}{n_1 + n_2}\right]}$$

$$\text{combined Standard deviation} = \sqrt{\left[\frac{n_1\left(\sigma_1^2 + d_1^2\right) + n_2\left(\sigma_2^2 + d_2^2\right)}{n_1 + n_2}\right]}$$

$$\bar{x}_C = \frac{n_1\bar{x}_1 + n_2\bar{x}_2}{n_1 + n_2} \quad ;d_1 = \bar{x}_C - \bar{x}_1 \quad \& \quad d_2 = \bar{x}_C - \bar{x}_2$$

3.4.8 Measures of dispersion for comparison (relative)

There are two types of relative dispersion i.e. Coefficient of range and coefficient of variation.

3.4.8.1 Coefficient of range is defined as

$$\text{Coefficent of Range} = \frac{\text{Maximum observation - Minimum observation}}{\text{Maximum observation + Minimum observation}}$$

$$\text{Coefficent of Range} = \frac{X_{max} - X_{min}}{X_{max} + X_{min}}$$

3.4.8.2 Coefficient of variation:

The standard deviation of heights of plants cannot be compared with the standard deviation of weights of the grains, as both are expressed in different units, i.e. heights in centimetre and weights in kilograms. Therefore the standard deviation must be converted into a relative measure of dispersion for the purpose of comparison. The relative measure of standard deviation is called coefficient of standard deviation and is given by

Coefficient of S.D = $\frac{\sigma}{\bar{x}}$

Coefficient of variation is defined as

$$C.V. = \frac{\sigma}{\bar{x}} \times 100$$

The series or groups of data for which the C.V. is greater it indicate that the group is more variable, less stable, less uniform, less consistent or less homogeneous.

If the value of C.V. is less, it indicates that the group is less variable or more stable or more uniform or more consistent or more homogeneous.

Note:

[1] Coefficient of variation is always expressed in percentage.

[2] Standard deviation gives total variation in mean while coefficient of variation gives percentage variation.

[3] Coefficient of variation is used for comparing variation between two different data sets.

[4] It has no units.

Numerical Example:

[1] Calculate variance and standard deviation for the following data:

2, 4, 6, 8, 10, 12, 14

Solution:

X	2	4	6	8	10	12	14	56
X^2	4	16	36	64	100	144	196	560

$$\text{Variance } = \sigma^2 = \frac{1}{n}\sum_{i=1}^{n} x_i^{\,2} - \bar{x}^2$$

$$\text{Standard deviation } = \sigma = \sqrt{\frac{1}{n}\sum_{i=1}^{n} x_i^{\,2} - \bar{x}^2} \quad ; \quad \bar{x} = \frac{\sum x_i}{n}$$

Now,

$$\bar{x} = \frac{\sum x_i}{n}$$

$$\bar{x} = \frac{56}{7}$$

$$\bar{x} = 8$$

$$\text{Variance } = \sigma^2 = \frac{1}{7}560 - 8^2$$

$$\text{Variance } = \sigma^2 = 80 - 64$$

$$\text{Variance } = \sigma^2 = 16$$

Similarly,

$$\text{Standard deviation } = \sigma = \sqrt{\frac{1}{n}\sum_{i=1}^{n} x_i^{\,2} - \bar{x}^2}$$

$$\text{Standard deviation } = \sigma = \sqrt{\text{Var}(x)}$$

$$\text{Standard deviation } = \sigma = \sqrt{16}$$

$$\text{Standard deviation } = \sigma = 4$$

[2] Calculate variance and standard deviation for the following data:

102, 104, 106, 108, 110, 112, 114

Solution:

X	102	104	106	108	110	112	114	56

$d_i = x_i - A$ $d_i = x_i - 108$	-6	-4	-2	0	2	4	6	0
d_i^2	36	16	4	00	4	16	36	112

$$\text{Variance } = \sigma^2 = \frac{1}{n}\sum_{i=1}^{n} d_i^2 - \left(\bar{d}\right)^2$$

$$\text{Standard deviation } = \sigma = \sqrt{\frac{1}{n}\sum_{i=1}^{n} d_i^2 - \left(\bar{d}\right)^2} \quad ; \qquad \left(\bar{d}\right) = \frac{\sum d_i}{n}$$

Where, $d_i = x_i - A$; A = assumed mean = 108

Now,

$$\bar{d} = \frac{\sum d_i}{n} = \bar{d} = \frac{0}{7} = \bar{d} = 0$$

$$\text{Variance } = \sigma^2 = \frac{1}{7}112 - 0^2$$

$$\text{Variance } = \sigma^2 = 16 - 0$$

$$\text{Variance } = \sigma^2 = 16$$

Similarly,

$$\text{Standard deviation } = \sigma = \sqrt{\text{Var}(x)}$$

$$\text{Standard deviation } = \sigma = \sqrt{16}$$

$$\text{Standard deviation } = \sigma = 4$$

Remarks: Above two examples, we conclude that add or subtract a constant number in each observation, then the variance and standard deviation will not be changed. This is known as change of origin property.

[3] Calculate variance and standard deviation for the following data:

x_i	2	4	6	8	10	12	14
f_i	3	6	8	10	7	5	2

Solution:

x_i	f_i	x_i^2	$f_i x_i^2$	$f_i x_i$
2	2	4	8	4
4	4	16	64	16
6	5	36	180	30
8	7	64	448	56
10	5	100	500	50
12	3	144	432	36
14	2	196	392	28
56	28	560	2024	220

We know,

$$\text{variance} = \sigma^2 = \frac{\sum_{i=1}^{n} f_i\left(x_i-\bar{x}\right)^2}{\sum_{i=1}^{n} f_i} = \frac{\sum_{i=1}^{n} f_i x_i^{\,2}}{\sum_{i=1}^{n} f_i} - \left(\frac{\sum_{i=1}^{n} f_i x_i}{\sum_{i=1}^{n} f_i}\right)^2$$

$$\text{variance} = \sigma^2 = \frac{\sum_{i=1}^{n} f_i x_i^{\,2}}{\sum_{i=1}^{n} f_i} - \left(\frac{\sum_{i=1}^{n} f_i x_i}{\sum_{i=1}^{n} f_i}\right)^2$$

$$\text{variance} = \sigma^2 = \frac{2024}{28} - \left(\frac{220}{28}\right)^2$$

$$\text{variance} = \sigma^2 = 72.28-(7.85)^2$$

$$\text{variance} = \sigma^2 = 72.28-61.62$$

$$\text{variance} = \sigma^2 = 10.66$$

Similarly,

$$\text{Standard deviation} = \sigma = \sqrt{\frac{\sum_{i=1}^{n} f_i \left(x_i - \bar{x}\right)^2}{\sum_{i=1}^{n} f_i}} = \sqrt{\frac{\sum_{i=1}^{n} f_i x_i^{\,2}}{\sum_{i=1}^{n} f_i} - \left(\frac{\sum_{i=1}^{n} f_i x_i}{\sum_{i=1}^{n} f_i}\right)^2}$$

$$\text{Standard deviation} = \sigma = \sqrt{Var(x)}$$

$$\text{Standard deviation} = \sigma = \sqrt{10.66}$$

$$\text{Standard deviation} = \sigma = 3.26$$

[4] Calculate variance and standard deviation by short-cut method for the following data:

x_i	102	104	106	108	110	112	114
f_i	3	6	8	10	7	5	2

Solution:

x_i	f_i	$d_i = x_i - A$ $d_i = x_i - 108$	d_i^2	$f_i d_i$	$f_i d_i^2$
102	2	-6	36	-12	72
104	4	-4	16	-16	64
106	5	-2	4	-10	20
108	7	0	0	00	00
110	5	2	4	10	20
112	3	4	16	12	48
114	2	6	36	12	72
756	28	0	112	-4	296

We know,

$$Var = \sigma^2 = \frac{\sum_{i=1}^{n} f_i d_i^{\ 2}}{\sum_{i=1}^{n} f_i} - \left(\frac{\sum_{i=1}^{n} f_i d_i}{\sum_{i=1}^{n} f_i} \right)^2$$

$$Var = \sigma^2 = \frac{296}{28} - \left(\frac{-4}{28} \right)^2$$

$$Var = \sigma^2 = 10.57 - \left(-0.14\right)^2$$

$$Var = \sigma^2 = 10.57 - 0.0196$$

$$Var = \sigma^2 = 10.55$$

$$S.D = \sigma = \sqrt{\frac{\sum_{i=1}^{n} f_i d_i^{\ 2}}{\sum_{i=1}^{n} f_i} - \left(\frac{\sum_{i=1}^{n} f_i d_i}{\sum_{i=1}^{n} f_i} \right)^2}$$

$$S.D. = (\sigma) = 3.24$$

Where,

$d_i = x_i - A$; A = assumed mean; $\bar{d} = \frac{\sum f_i d_i}{N}$

$$Var(\sigma^2) = 3.24$$

[5] Calculate variance and standard deviation by step-deviation method for the following data:

Marks obtained	0-10	10-20	20-30	30-40	40-50	50-60	60-70
No. of students	3	7	11	15	12	81	4

Solution:

Marks obtained	f_i	x_i	$d_i = \frac{x_i - A}{h}$ $d_i = \frac{x_i - 35}{10}$	d_i^2	$f_i d_i$	$f_i d_i^2$
0-10	3	5	-3	9	-9	27
10-20	7	15	-2	4	-14	28
20-30	11	25	-1	1	-11	11
30-40	15	35	0	0	00	00

40-50	12	45	1	1	12	12
50-60	8	55	2	4	16	32
60-70	4	65	3	9	12	36
	60				6	146

We know,

$$\text{Var} = \sigma^2 = h^2 \times \left[\frac{\sum_{i=1}^{n} f_i d_i^2}{\sum_{i=1}^{n} f_i} - \left(\frac{\sum_{i=1}^{n} f_i d_i}{\sum_{i=1}^{n} f_i} \right)^2 \right]$$

Where, $d_i = \frac{x_i - A}{h}$; A = assumed mean; h = class interval

$$\text{Var} = \sigma^2 = 10^2 \times \left[\frac{146}{60} - \left(\frac{6}{60} \right)^2 \right]$$

$$\text{Var} = \sigma^2 = 100 \times \left[2.43 - (0.1)^2 \right]$$

$$\text{Var} = \sigma^2 = 100 \times [2.43 - 0.01]$$

$$\text{Var} = \sigma^2 = 100 \times [2.42]$$

$$\text{Var} = \sigma^2 = 242$$

$$\text{S.D}(\sigma) = h \times \sqrt{\frac{\sum_{i=1}^{n} f_i d_i^2}{\sum_{i=1}^{n} f_i} - \left(\frac{\sum_{i=1}^{n} f_i d_i}{\sum_{i=1}^{n} f_i} \right)^2}$$

$$\text{S.D}(\sigma) = \sqrt{\text{Var}(x)}$$

$$\text{S.D}(\sigma) = \sqrt{242}$$

$$\text{S.D}(\sigma) = 15.55$$

[6] Calculate variance and standard deviation by step-deviation method for the following data:

Weekly wages in Rs.	0-40	40-80	80-120	120-160	160-200	200-240	240-280	280-320

No. of students	1	7	12	25	30	15	6	4

Solution:

Marks obtained	f_i	x_i	$d_i=\frac{x_i-A}{h}$ $d_i=\frac{x_i-180}{40}$	d_i^2	f_id_i	$f_id_i^2$
0-40	1	20	-4	16	-4	16
40-80	7	60	-3	09	-21	63
80-120	12	100	-2	04	-24	48
120-160	25	140	-1	01	-25	25
160-200	30	180	0	00	00	00
200-240	15	220	1	01	15	15
240-280	6	260	2	04	12	24
280-320	4	300	3	09	12	36
Total	100		-4		-35	227

We know,

$$Var=\sigma^2=h^2\times\left[\frac{\sum_{i=1}^{n}f_id_i^2}{\sum_{i=1}^{n}f_i}-\left(\frac{\sum_{i=1}^{n}f_id_i}{\sum_{i=1}^{n}f_i}\right)^2\right]$$

Where,

$d_i=\frac{x_i-A}{h}$; A = assumed mean; h = class interval

$$Var=\sigma^2=40^2\times\left[\frac{227}{100}-\left(\frac{-35}{100}\right)^2\right]$$

$$Var=\sigma^2=1600\times\left[2.27-(-0.35)^2\right]$$

$$Var=\sigma^2=1600\times\left[2.27-0.1225\right]$$

$$Var=\sigma^2=1600\times\left[2.1475\right]$$

$$Var=\sigma^2=3436$$

$$S.D = \sigma = h \times \sqrt{\frac{\sum_{i=1}^{n} f_i d_i^{\,2}}{\sum_{i=1}^{n} f_i} - \left(\frac{\sum_{i=1}^{n} f_i d_i}{\sum_{i=1}^{n} f_i}\right)^2}$$

$$S.D = \sigma = \sqrt{Var(x)}$$

$$S.D = \sigma = \sqrt{3436}$$

$$S.D = \sigma = 58.62$$

[7] The Frequency distributions of seed yield of 50 seasamum plants are given below. Find the standard deviation.

Seed yield in gms (x)	3.5-4.5	4.5-5.5	5.5-6.5	6.5-7.5	7.5-8.5
No. of plants (f)	4	6	15	15	10

Solution:

Marks obtained	f_i	x_i	$d_i = \frac{x_i - A}{h}$ $d_i = \frac{x_i - 6}{1}$	d_i^2	$f_i d_i$	$f_i d_i^2$
3.5-4.5	4	4	-2	4	-8	16
4.5-5.5	6	5	-1	1	-6	06
5.5-6.5	15	6	0	0	0	00
6.5-7.5	15	7	1	1	15	15
7.5-8.5	10	8	2	4	20	40
Total	50		0		21	77

We know,

$$S.D = \sigma = h \times \sqrt{\left[\frac{\sum_{i=1}^{n} f_i d_i^{\,2}}{\sum_{i=1}^{n} f_i} - \left(\frac{\sum_{i=1}^{n} f_i d_i}{\sum_{i=1}^{n} f_i}\right)^2\right]}$$

Where, $d_i = \frac{x_i - A}{h}$; A = assumed mean; h = class interval

$$S.D = \sigma = 1 \times \sqrt{\left[\frac{77}{50} - \left(\frac{21}{50}\right)^2\right]}$$

$$S.D = \sigma = 1 \times \sqrt{\left[1.54 - (0.42)^2\right]}$$

$$S.D = \sigma = 1 \times \sqrt{[1.54 - 0.1764]}$$

$$S.D = \sigma = 1 \times \sqrt{[1.3636]}$$

$$S.D = \sigma = 1 \times 1.167$$

$$S.D = \sigma = 1.167$$

[8] Calculate coefficient of variation for the following data:

Weekly wages in Rs.	0-40	40-80	80-120	120-160	160-200	200-240	240-280	280-320
No. of students	1	7	12	25	30	15	6	4

Solution:

Marks obtained	f_i	x_i	$d_i = \frac{x_i - A}{h}$ $d_i = \frac{x_i - 180}{40}$	d_i^2	$f_i d_i$	$f_i d_i^2$
0-40	1	20	-4	16	-4	16
40-80	7	60	-3	09	-21	63
80-120	12	100	-2	04	-24	48
120-160	25	140	-1	01	-25	25
160-200	30	180	0	00	00	00
200-240	15	220	1	01	15	15
240-280	6	260	2	04	12	24
280-320	4	300	3	09	12	36
Total	100		-4		-35	227

We know,

$$C.V. = \frac{\sigma}{\overline{x}} \times 100$$

$$S.D = \sigma = h \times \sqrt{\frac{\sum_{i=1}^{n} f_i d_i^2}{\sum_{i=1}^{n} f_i} - \left(\frac{\sum_{i=1}^{n} f_i d_i}{\sum_{i=1}^{n} f_i}\right)^2}$$

Where, $d_i = \frac{x_i - A}{h}$; A = assumed mean; h = class interval

$$\overline{d} = \frac{\overline{x} - A}{h} \Rightarrow h\overline{d} = \overline{x} - A \Rightarrow \overline{x} = A + h\overline{d}\ ; \qquad \overline{d} = \frac{1}{N}\sum_{i=1}^{n} f_i d_i$$

$$\overline{x} = A + h\overline{d}$$

$$\overline{x} = A + \frac{h}{N}\sum_{i=1}^{n} f_i d_i$$

$$\overline{x} = 180 + \frac{40}{100}(-35)$$

$$\overline{x} = 180 - \frac{1400}{100}$$

$$\overline{x} = 180 - 14$$

$$\overline{x} = 166$$

$$S.D = \sigma = 40 \times \sqrt{\left[\frac{227}{100} - \left(\frac{-35}{100}\right)^2\right]}$$

$$S.D = \sigma = 40 \times \sqrt{\left[2.27 - (-0.35)^2\right]}$$

$$S.D = \sigma = 40 \times \sqrt{[2.27 - 0.1225]}$$

$$S.D = \sigma = 40 \times \sqrt{[2.1475]}$$

$$S.D = \sigma = 40 \times 1.4654$$

$$S.D = \sigma = 58.62$$

$$C.V. = \frac{\sigma}{\bar{x}} \times 100$$

$$C.V. = \frac{58.62}{166} \times 100$$

$$C.V. = \frac{5862}{166}$$

$$C.V. = 35.31\%$$

[9] Consider the measurement on, yield and plant height of a paddy variety. The mean and standard deviation for yield are 50 kg and 10 kg respectively. The mean and standard deviation for plant height are 55 cm and 5 cm respectively.

Solution: we have given,

Mean for yield = 50 kg; S.D. for yield = 10 kg

Mean for plant = 55 cm; S.D. for yield = 5 cm

We know,

$$C.V. = \frac{\sigma}{\bar{x}} \times 100$$

Here, the measurements for yield and plant height are in different units. Hence, the variability's can be compared only by using coefficient of variation.

For yield,

$$C.V. = \frac{\sigma}{\bar{x}} \times 100$$

$$C.V. = \frac{10}{50} \times 100$$

$$C.V. = 20\%$$

For plant height,

$$C.V. = \frac{\sigma}{\bar{x}} \times 100$$

$$C.V. = \frac{5}{55} \times 100$$

$$C.V. = \frac{500}{55}$$

$$C.V. = 9.09\%$$

The yield is subject to more variation than the plant height.

[10] Calculate coefficient of variation for the following data:

No. of lectures attended	5-10	10-15	15-20	20-25	25-30	30-35	35-40
No. of students	6	8	16	28	32	25	5

Solution:

No. of lectures attended	No. of students (f_i)	x_i	$d_i = \frac{x_i - A}{h}$ $d_i = \frac{x_i - 22.5}{5}$	d_i^2	$f_i d_i$	$f_i d_i^2$
05-10	6	7.5	-3	9	-18	54
10-15	8	12.5	-2	4	-16	32
15-20	16	17.5	-1	1	-16	16
20-25	28	22.5	0	0	0	00
25-30	32	27.5	1	1	32	32
30-35	25	32.5	2	4	50	100
35-40	5	37.5	3	9	15	45
Total	100		0	28	47	279

We know,

$$C.V. = \frac{\sigma}{\bar{x}} \times 100$$

$$S.D = \sigma = h \times \sqrt{\frac{\sum_{i=1}^{n} f_i d_i^2}{\sum_{i=1}^{n} f_i} - \left(\frac{\sum_{i=1}^{n} f_i d_i}{\sum_{i=1}^{n} f_i}\right)^2}$$

Where, $d_i = \frac{x_i - A}{h}$; A = assumed mean; h = class interval

$\bar{d} = \frac{\bar{x} - A}{h} \Rightarrow h\bar{d} = \bar{x} - A \Rightarrow \bar{x} = A + h\bar{d}$; $\bar{d} = \frac{1}{N}\sum_{i=1}^{n} f_i d_i$

$$\bar{x} = A + h\bar{d}$$

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

$$\bar{x} = 22.5 + \frac{5}{120}(47)$$

$$\bar{x} = 22.5 + \frac{235}{120}$$

$$\bar{x} = 22.5 + 1.96$$

$$\bar{x} = 24.46$$

$$S.D = \sigma = h \times \sqrt{\frac{\sum_{i=1}^{n} f_i d_i^2}{\sum_{i=1}^{n} f_i} - \left(\frac{\sum_{i=1}^{n} f_i d_i}{\sum_{i=1}^{n} f_i}\right)^2}$$

$$S.D = \sigma = 5 \times \sqrt{\left[\frac{279}{120} - \left(\frac{47}{120}\right)^2\right]}$$

$$S.D = \sigma = 5 \times \sqrt{\left[2.325 - (0.392)^2\right]}$$

$$S.D = \sigma = 5 \times \sqrt{[2.325 - 0.154]}$$

$$S.D = \sigma = 5 \times \sqrt{[2.17]}$$

$$S.D = \sigma = 5 \times 1.47$$

$$S.D = \sigma = 7.35$$

$$C.V. = \frac{\sigma}{\bar{x}} \times 100$$

$$C.V. = \frac{7.35}{24.46} \times 100$$

$$C.V. = \frac{735}{24.46}$$

$$C.V. = 30.05\%$$

[11] 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. Also, variances are 25 and 16 respectively. Obtain the combined variance and standard deviation.

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

$$\text{combined variance} = \frac{n_1\sigma_1^2 + n_2\sigma_2^2 + n_1 d_1^2 + n_2 d_2^2}{n_1 + n_2}$$

$$\text{combined variance} = \frac{n_1\left(\sigma_1^2 + d_1^2\right) + n_2\left(\sigma_2^2 + d_1^2\right)}{n_1 + n_2}$$

$$\text{combined Standard deviation} = \sqrt{\left[\frac{n_1\sigma_1^2 + n_2\sigma_2^2 + n_1 d_1^2 + n_2 d_2^2}{n_1 + n_2}\right]}$$

$$\text{combined Standard deviation} = \sqrt{\left[\frac{n_1\left(\sigma_1^2 + d_1^2\right) + n_2\left(\sigma_2^2 + d_2^2\right)}{n_1 + n_2}\right]}$$

$$\bar{x}c = \frac{n_1\bar{x}_1 + n_2\bar{x}_2}{n_1 + n_2} \qquad ; d_1 = \bar{x}_c - \overline{x_1} \qquad \& \, d_2 = \bar{x}_c - \overline{x_2}$$

$$\overline{x_c} = 4000, \qquad \bar{x}_1 = 4200, \qquad \bar{x}_2 = 3200$$

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

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

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

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

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

$$-200\ n_1 = -800\ n_2$$

$$200\ n_1 = 800\ n_2$$

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

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

male and female employees in the company are 4 & 1

$$d_1 = \bar{x}_C - \bar{x}_1$$

$$d_1 = 4000 - 4200$$

$$d_1 = -200$$

$$d_1^2 = 40000$$

&

$$d_2 = \bar{x}_C - \bar{x}_2$$

$$d_2 = 4000 - 3200$$

$$d_2 = 800$$

$$d_2^2 = 6400$$

$$\text{combined variance} = \frac{4(25+40000)+1(16+6400)}{4+1}$$

$$\text{combined variance} = \frac{4(40025)+1(6416)}{5}$$

$$\text{combined variance} = \frac{160100+6416}{5}$$

$$\text{combined variance} = \frac{166516}{5}$$

combined variance = 33303.2

$$\text{combined Standard deviation} = \sqrt{\left[\frac{n_1\left(\sigma_1^2+d_1^2\right)+n_2\left(\sigma_2^2+d_2^2\right)}{n_1+n_2}\right]}$$

$$\text{combined Standard deviation} = \sqrt{33303.2}$$

combined Standard deviation = 182.49

[12] The means of two samples of sizes 50 and 100 with their means are 70 and 40 with their variances are 100 and 25 respectively. Obtain the combined mean and combined standard deviation.

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

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

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

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

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

$\bar{x}_c = 50$

$d_1 = \bar{x}_c - \bar{x}_1$

$d_1 = 50-70$

$d_1 = -20$

$d_1^2 = 400$

&

$d_2 = \bar{x}_c - \bar{x}_2$

$d_2 = 50-40$

$d_2 = 10$

$d_2^2 = 100$

$$\text{combined Standard deviation} = \sqrt{\left[\frac{n_1\left(\sigma_1^2+d_1^2\right)+n_2\left(\sigma_2^2+d_2^2\right)}{n_1+n_2}\right]}$$

$$\text{combined Standard deviation} = \sqrt{\left[\frac{50(100+400)+100(25+100)}{50+100}\right]}$$

$$\text{combined Standard deviation} = \sqrt{\left[\frac{50(500)+100(125)}{50+100}\right]}$$

$$\text{combined Standard deviation} = \sqrt{\left[\frac{25000+12500}{150}\right]}$$

$$\text{combined Standard deviation} = \sqrt{\left[\frac{37500}{150}\right]}$$

$$\text{combined Standard deviation} = \sqrt{250}$$

$$\text{combined Standard deviation} = 15.81$$

[A] Fill in the blanks

[1] Less the coefficient of variation …….consistent is the series

Answer:- More

[2] Coefficient of variation is usually expressed in……..

Answer:- Percentage

[3] If quartile deviation of certain items is 20 and mean is 50 the coefficient of variation is …… percent.

Answer:- 60

[4] The sum of squares of deviation taken from mean 40 for 9 sample observation is 288 the coefficient of variation is …..

Answer:- 14.125

[5] Out of all measures of dispersion the easiest one to calculate is

Answer:-Range.

[6] Formula for range of a set of values $x_1, x_2, x_3, \cdots, x_n$ is .

Answer:- $R=X_{max}-X_{min}$

[7] Which of the measures of dispersion get a negative value.

Answer:-Range.

[8] The mean and standard deviation of a set of values of a normal distribution are 66 and 4 respectively. Then the mean in which almost 95% value is

Answer:-58 to 74.

[9] Which measure of dispersion are most affected by Extreme values

Answer:-Range

[10] Range of a set of values is 65 and maximum value in the series 83 the minimum values of the series is

Answer:-18

[11] If the minimum value in set is 9 and its range is 57 the maximum value of a set is

Answer:-66

[12] The range of the set of values 15, 12, 17, 6, 9, 18, 21 is

Answer:-15.

[13] Coefficient of quartile deviation is given by the formula

Answer:- Coefficient of Q. D. = $\frac{(Q_3-Q_1)}{(Q_3+Q_1)}$

[14] Quartile deviation or semi interquartile deviation is given by the formula

Answer:- Q. D. = $\frac{(Q_3-Q_1)}{2}$

[15] The empirical relation between quartile deviation and standard deviation in normal distribution

Answer:-3 Q.D. = 2 S.D.

[16] The empirical relation between quartile deviation and mean deviation from

Answer:-6 Q.D. = 5 M.D.

[17] An empirical relation between standard deviation and mean deviation about mean and quartile deviation is

Answer:-4 S.D. = 5 M.D. = 6 Q.D.

[18] An empirical relation between range and quartile deviation about mean is

Answer:-R = 9 Q.D.

[19] Which measure of dispersion ensures lowest degree reliability

Answer:-Quartile deviation

[20] If the standard deviation of a distribution is 15 the quartile deviation of the distribution is

Answer:-10

[21] If the first 25% observation of a series are 20 and last 25% observation of a series are 50 or more the quartile deviation is

Answer:-15

[22] Which measures of dispersion is least affected by Extreme values

Answer:-Quartile deviation

[23] Quartile deviation is equal to

Answer:-Half of the interquartile range.

[24] Which measure of dispersion can be calculated in case of open end intervals

Answer:-Quartile deviation

[25] If the first quartile $Q_1 = 30$ and third quartile $Q_3 = 60$ then the quartile deviation is.

Answer:-15

[26] Mean deviation is minimum when deviations are taken from

Answer:- median

[27] The relationship between mean deviation and standard deviation is

Answer:- S.M.D. = 4 S.D.

[28] Empirical relation between range and mean deviation is

Answer:-2R = 15 M.D.

[29] If the quartile deviation of a series is 60 the mean deviation of this series is

Answer:-72

[30] The measures of dispersion which ignores signs of the deviation from a central value is.

Answer:-mean deviation

[31] For a set of values mean deviation is always less than

Answer:- Standard deviation .

[B] Choose the correct alternative from the following.

[1] The difference between the largest and the smallest data values is the

(a) variance (b) interquartile range

(c) range **(**d) coefficient of variation

Answer:- (c) range

[2] A researcher has collected the sample data. The mean of the sample is 5. 3, 5, 12, 3, 2. The range is

(a) 1 (b) 2 (c) 10 **(**d) 12

Answer:- (**c) 10**

[3] Which of the following is not a measure of dispersion?

(a) the range (b) the quartiles

(c) the standard deviation (d) the variance

Answer:- (**b) the quartiles**

[4] 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

[5] The standard deviation of a set of 50 observation is 8. If each observation is multiplied by 2, then the value of S.D. will be

(a) 4 (b) 8

(c) 16 (d) None

Answer:- (c) 16

[6] Which of the following measures of dispersion is affected most by extreme values of observations in a data set ?

(a) Range (b) Quartile Deviation

(c) Mean Deviation (d) Standard Deviation

Answer:- (**d) Standard Deviation**

[7] Which of the following is a relative measure of dispersion ?

(a) Standard Deviation (b) Variance

(c) Coefficient of Variation (d) All

Answer:- (c) Coefficient of Variation

[8] If mean and coefficient of variation of a set of data is 10 and 5, then S.D. is

(a) 10 (b) 50 (c) 5 (d) None

Answer:- (**b) 50**

[9] The hourly wages of a sample of 130 system analysts are given below. mean = 60; range = 20; mode = 73; variance = 324; median = 74. The coefficient of variation equals.

(a). 0.30% (b) 30% (c) 5.4% (d) 54%

Answer:- (**b) 30%**

[10] The descriptive measure of dispersion that is based on the concept of a deviation about the mean is

(a) the range

(b) the interquartile range

(c) the absolute value of the range

(d) the standard deviation

Answer:- (d) **the standard deviation**

[11] The numerical value of the standard deviation can never be

(a) larger than the variance (b) zero

(c) negative (d) smaller than the variance

Answer:- (c) negative

[12] The variance can never be

(a) zero

(b) larger than the standard deviation

(c) negative

(d) smaller than the standard deviation

Answer:- (c) negative

[13] The sum of deviations of the individual data elements from their mean is

(a) always greater than zero

(b) always less than zero

(c) Depending on the data elements

(d) always equal to zero

Answer:- (d) always equal to zero

[14] The square of standard deviation is known as

(a) mean (b) standard deviation

(c) variance (d) none of these

Answer:- (**c) variance**

[15] Standard deviation is

(a) relative measure (b) absolute measure

(c) both (d) none of these

Answer:- (**b) absolute measure**

[16] Standard deviation is always taken from

(a) median (b) mode

(c) mean (d) none of these

Answer:- (c) mean

[17] The standard deviation of 5,5,5,5,5,5,5 will be

(a) 1 (b) 0

(c) 5 (d) none of these

Answer:- (b) 0

[18] Coefficient of variation is

(a) relative measure (b) absolute measure

(c) both (a) & (b) (d) none of these

Answer:- (a) relative measure

[19] Out of will measures of dispersion, the easiest one to calculate is

(a) Standard deviation (b) Range

(c) variance (d) quartile deviation.

Answer:- (b) Range

[20] Formula for range (R) of a set of values $x_1, x_2, x_3, \cdots, x_n$

(a) $R = X_{max} - X_{min}$ (b) $R = |X_{max} - X_{min}|$

(c) $R = X_{min} - X_{max}$ (d) NOTA

Answer:- (a) $R = X_{max} - X_{min}$

[21] The empirical relation between range and standard deviation is

(a) R= 3 S.D (b) R= 2 S.D

(c) R= 6 S.D (d) R= 4 S.D

Answer:- (c) R= 6 S.D

[22] The mean and standard deviation of a set of values from a normal distribution are 66 and 4 respectively. The range in which almost 95 percent values lie in

(a) 62 to 70 (b) 62 to 74

(c) 58 to 74 (d) 66 to 74

Answer:-(c) 58 to 74

[23] Range of a set of values is 65 and maximum value in the series is 83. The minimum value of the series is

(a) 74 (b) a

(c) 18 (d) NOTA

Answer:- (c) 18

[24] If the minimum value in a set is 9 and it range is 57, the maximum value of the set is

(a) 33 (b) 48

(c) 66 (d) NOTA

Answer:-(c) 66

[25] The range of value for the frequency distribution given in question 81 is

(a) 2 (b) 10

(c) 6 (d) 8

Answer:- (d) 8

[26] The range of set of values 15,12,27,6,9,18,21, is

(a) 21 (b) 4.5

(c) 0.64 (d) 3

Answer:- (a) 21

[27] Coefficient of quartile deviation is given by the formula

(a) Coefficient of Q. D. = $\frac{(Q_3+Q_1)}{(Q_3-Q_1)}$

(b) Coefficient of Q. D. = $\frac{(Q_3-Q_1)}{(Q_3-Q_1)}$

(c) Coefficient of Q. D. = $\frac{(Q_3+Q_1)}{(Q_1-Q_3)}$

(d) Coefficient of Q. D. = $\frac{(Q_3-Q_1)}{(Q_3+Q_1)}$

Answer:- (d)

[28] The empirical relationship between Q.D and standard deviation in normal distribution is

(a) 3Q.D =2 S.D (b) 4Q.D =3 S.D

(c) 6Q.D= 5 S.D (d) 5Q.D=4 S.D

Answer:- (a) 3Q.D =2 S.D

[29] The empirical relationship between Q.D and M.D form mean is

(a) 3 Q.D =5M.D (b) 6 Q.D = 3 M.D

(c) 5 Q.D = 6 M.D (d) 6 Q.D = 5 M.D

Answer:- (d) 6 Q.D = 5 M.D

[30] An empirical relationship between range and Q.D about mean is

(a) R = 4Q.D (b) R = 9 Q.D

(c) R = 6 Q.D (d) none of the above

Answer:- (b) R = 9 Q.D

[31] Which measure of dispersion ensures lowest degree of reliability?

(a) range (b) M.D

(c) Q.D. (d) S.D

Answer:- (c) Q.D.

[32] If the S.D of a distribution is 15, the Q.D of the distribution is

(a) 15.0 (b) 12.5

(c) 10.00 (d) NOTA

Answer:- (c) 10.00

[33] If the first 25 percent observations of series are 20 or less and last 25 percent observation of a series are 50 or more, the Q.D is ; 15

(a) 25 (b) 35

(c) 15 (d) 30

Answer:- (c) 15

[34] Quartile Deviation is equal to

(a) interquartile range

(b) double the interquartile range

(c) half of the interquartile range

(d) none of the above

Answer:- (c) half of the interquartile range

[35] If the first quartile Q_1 =20 and third quartile Q_3 =50, the quartile deviation is

(a) 35 (b) 15

(c) 2.5 (d) 0.8

Answer:- (b) 15

[36] Mean Deviation is minimum when deviation are taken from

(a) mean (b) median

(c) mode (d) zero

Answer:- (b) median

[37] The relationship between mean deviation and standard deviation is :

(a) 3 S.D =5M.D. (b) 6 S.D = 3 M.D.

(c) 5 S.D = 6 M.D. (d) 4 S.D = 5 M.D.

Answer:- (d) 4 S.D = 5 M.D.

[38] An empirical relation between range and mean deviation is

(a) 2R = 15 M.D. (b) 4R = 10 M.D.

(c) 5R = 6 M.D. (d) none of the above

Answer:- (a) 2R = 15 M.D.

[39] If the Q.D. of a series 60, the M.D. of the series is

(a) 72 (b) 48

(c) 50 (d) 75

Answer:- (a) 72

[40] For a set of values : M.D. is always less than S.D about median is

(a) 1.43 (b) 1.00

(c) 2.43 (d) 6

Answer:- (a) 1.43

[41] The correct relation between variance and S.D of a variable X is

(a) $\text{S.D.}=\left[\text{var}(x)\right]^{2}$ (b) $\text{S.D.}=\left[\text{var}(x)\right]^{\frac{1}{2}}$

(c) $\text{S.D.}=\left[\text{var}(x)\right]$ (d) none of the above

Answer:- (b)

[42] If a constant value 5 is subtracted from each observation of a set, the variance is

(a) reduced by 5 (b) reduced by 25

(c) unaltered (d) increased by 25

Answer:- (c) unaltered

[43] If the values of a set are measured, in cm, the unit of variance will be

(a) no unit (b) cm

(c) cm^2 (d) cm^3

Answer:- (c) cm^2

[44] The variance of first n natural number is

(a) $\frac{(n^2-1)}{12}$ (b) $\frac{(n+1)^2}{12}$

(c) $\frac{(n^2+1)}{12}$ (d) none of the above

Answer:- (a)

[45] If a random variable X has mean 3 and S.D 5, then variance of a variable Y = 2x-5 is

(a) 45 (b) 100

(c) 15 (d) 40

Answer:- (b) 100

[46] All values in a sample are same, and then the variance is

(a) 0 (b) one

(c) not calculable (d) all of the above

Answer:- (a) 0

[47] Which one of the given measures of dispersions is considers best?

(a) S.D. (b) range

(c) variance (d) coefficient of variation

Answer:- (a) S.D.

[48] If each observation of a set is divided by 10, the S.D of the new observation is.

(a) $1/10^{th}$ of S.D of original obs.

(b) $1/100^{th}$ of S.D of original obs.

(c) not changed

(d) 10 times of S.D of original obs.

Answer:- (a) $1/10^{th}$ of S.D of original obs.

[49] If the M.D. of a distribution is 20.20, the S.D of a distribution is

(a) 15.15 (b) 25.25

(c) 30.30 (d) NOTA

Answer:- (b) 25.25

[50] If the mean and variance of A and B are as, X_A =15.0, X_B=20.0 and σ_A^2 =25 and σ_B^2 =16 which of the two series is more consistent.

(a) series A

(b) series B

(c) series A and B equally consistent

(d) none of the above.

Answer:- (b) series B

[51] Any discrete set of values, the correct relation between M.D. and S.D.

(a) M.D > S.D **(b) M.D < S.D**

(c) M.D ≤S.D (d) M.D ≥ S.D

Answer:- (b) M.D < S.D

[52] Which one property out of the following does not hold good in case of standard deviation?

(a) It is distorted by extreme values.

(b) It is not very sensitive to sampling fluctuation as compared to other measures.

(c) It is a unitless measure of dispersion.

(d) It is must used measure of dispersion.

Answer:-(c) It is a unitless measure of dispersion.

[53] The relation between variance and standard deviation is

(a) variance is square root of standard deviation.

(b) standard deviation is the square of variance .

(c) variance is equal to standard deviation.

(d) square of standard deviation is equal to variance.

Answer:- (d) square of standard deviation is equal to variance.

[54] The S.D. of a set of value will be:

(a) positive when the values are positive.

(b) positive when the values are negative.

(c) always positive.

(d) all of the above.

Answer:-(c) always positive.

[55] Formula for coefficient of variation is

(a) $\text{C.V.}=\frac{\text{S.D.}}{\text{Mean}}\times 100$

(b) $\text{C.V.}=\frac{\text{Mean}}{\text{S.D.}}\times 100$

(c) $\text{C.V.}=\frac{\text{S.D.}}{\text{Mean}}$

(d) NOTA

Answer:- (a)

[56] Average wages of workers of a factory are Rs. 550.00 is per month and S.D of wages is 110. The coefficient of variation is

(a) C.V = 30%　　(b) C.V = 15%

(c) C.V = 500%　　(d) C.V = 20%

Answer:- (d) C.V = 20%

[57] The mean and standard deviation of a set of values are 25 and 5, respectively. If a constant of value 5 is added to each value, the coefficient of variation of the new set of value is

(a) 250%　　(b) 600%

(c) 20%　　(d) 16.6%

Answer:- (d) 16.6%

[58] If each values of a series is divided by 5, its coefficient of variation is reduced by

(a) 0%　　(b) 5%

(c) 10%　　(d) 20%

Answer:- (a) 0%

[59] If each value of a series is multiplied by 10, the coefficient of variation will be increased by

(a) 5%　　(b) 10%

(c) 15%　　(d) 0%

Answer:- (d) 0%

[60] If a constant value 10 is a subtracted from each value of a series, the coefficient of variation will be: increased in comparison to original value

(a) decreased in comparison to original value.

(b) increased in comparison to original value.

(c) same as original value.

(d) none of the above.

Answer:- (**b) increased in comparison to original value.**

[61] If each value of a series is multiplied by constant "C", the C.V. as compared to original value each

(a) increased　　(b) decreased

(c) unaltered. (d) none of the above

Answer:-(c) unaltered.

[62] If each value of a set is divided by constant 'd' , the C.V. will

(a) same as original value.

(b) less than original value

(c) more than original value.

(d) none of above .

Answer:- (a) same as original value.

EXCERCISE:

[1] Define dispersion.

[2] Define C.V. What are its uses?

[3] What are the differences between absolute measure and relative measure of dispersion?

[4] Calculate standard deviation for the following data.

Class	5-15	20-40	40-60	60-80	80-100
Frequency	5	15	32	40	11

[5] Calculate standard deviation for the following data.

Class	0-20	20-40	40-60	60-80	80-100
Frequency	5	12	32	40	11

[6] Calculate coefficient of variation for the following data:

10, 16, 30, 32, 12

[7] The monthly wages paid to the workers in two firms A and B gives the following result:

	Firm A	Company B
No. of workers	400	500
Mean wages Rs.	1500	1600
Standard deviation of wages Rs.	30	40

Which firm is more consistent in paying wages?

[8] Information of shares of two companies is given below.

	Company A	Company B
Mean Price Rs.	27	25
Standard deviation of prices Rs.	20	12

Which company's share pieces are more stable?

[9] Calculate Range and coefficient of Range for the following frequency distribution:

Electricity Consumption	100-2000	200-300	300-400
No. of families	25	35	40

[10] Calculate Range and coefficient of Range for the following frequency distribution:

Roll No.	1	2	3	4	5	6	7
Marks of Students	5	15	25	35	45	55	20