Statistics

Definition of Statistics –
Statistics is a branch of mathematics dealing with data collection, organization, analysis, interpretation and presentation.

Ungrouped (Raw) data – Ungrouped data is the data which we first gather from an experiment or study.
e.g. List of numbers.

Grouped data – Grouped data is data that is bundled together in categories.
e.g. Time taken by group of students to answer question
Discrete Variate – If a variate takes integral values or values at interval then it is called discrete variate.
e. g. 1) Number of runs scored by a cricketer
2) Number of telephone calls on today

Continuous data – Continuous data is quantitative data and it is measurable. It has an infinite number of possible values within. a selected range.
 e.g. temperature range.

Continuous Variate - If a variate takes every value within the specified range, then it is called continuous variate.
e.g. 1) The temperature of a patient
2) The salary of an employee

Dispersion - Dispersion is the degree of variation of the variable about a central value.
We study following measures of dispersion.
1) Range   2) Mean Deviation (M.D.)   3) Standard Deviation (S.D.)   4) Variance
Let us consider above measures in detail

1) Range
The range is the simplest measure of dispersion. It is defined as the difference between the value of largest observation and the value of smallest observation included in the distribution.

Note – 1) For ungrouped data,
$$\text{Range} = \text{Largest Value} - \text{Smallest Value}$$
$$\therefore R = L - S$$
where R = Range, L = Largest Value, S = Smallest Value

2) For grouped frequency distribution,
$$\text{Range} = \text{Upper limit of highest class} - \text{Lower limit of lower class}$$
$$\therefore R = U_L - L_S$$
where R = Range, U_L = Upper limit of the highest class interval,
L_S = Lower limit of the lowest class interval

3) Coefficient of $\text{Range} = \dfrac{L-S}{L+S}$

Examples

1) A batsman's score in four test matches against a certain country was
61,2,61,42,59,78,13,221
Find the range and coefficient of range of his score. [R= 219, Coeff. = 0.9820]
2) Find the range and coefficient of range from the following data:
50,90,120, 40,180,200,80 [B.T.E.2013,2016] [R = 160, Coe. = 0.67]

3) Find range and coefficient of range of the following distribution: **[B.T.E.2015]**

x_i	10	20	30	40	50
f_i	7	5	3	2	1

[Range = 40, Coeff. = 2/3 = 0.67]

4) Find range and coefficient of range for the following distribution:

Marks	0-10	10-20	20-30	30-40	40-50
No. of students	8	12	10	15	5

[R =50 , Coeff. =1]

5) From the following data, calculate range and coefficient of range: **[BTE2016]**

Marks	10-19	20-29	30-39	40-49	50-59	60-69
No. of students	6	10	16	14	8	4

[R = 60, Coeff. = 0.76]

Examples for Tutorial

Find the range and coefficient of range from the following:
1) 5, 25, 65, 55, 35, 45, 15. **[B.T.E.2017]** [Range = 60, Coefficient of range = 60/70 = 0.857]
2) Weight in kg.: 70,75,69,80,85,83,65,89,73,84,90 [Range = 25] **[B.T.E.2015]**
3) 27,30,35,36,38,40,43 [Range = 16, Coefficient of Range = 0.23]
4) 800, 725, 750, 900, 925, 910, 1000, 790, 870, 920. **[B.T.E.2016]**
5) 45,42,39,40,48,41,45,44. **[BTE 2017] [9]**
6)

Weights	115-125	125-135	135-145	145-155	155-165
Frequency	4	5	6	3	2

7)

Marks	0-10	10-20	20-30	30-40	40-50	50-60
No. of students	10	20	30	50	40	30

[Answer: Range = 50, Coefficient of Range = 0.833]
8)

Marks	20-29	30-39	40-49	50-59	60-69	70-79	80-89	90-99
No. of students	10	15	16	20	21	22	9	8

[Range = 80, Coefficient of range =0.672]   **[BTE2014, 2017]**

Mean Deviation (M.D.)
Mean deviation is defined as the arithmetic mean of all the absolute deviations from any one of its averages. Generally we take mean or median.
In this topic, we study mean deviation from mean.
Note -
Mean Deviation of ungrouped/raw data
1) If $x_1, x_2, x_3, ---- x_n$ are n observations whose mean is $\bar{x}$, then the mean deviation is given by,

$$\text{Mean Deviation} = \frac{\sum |x_i - \bar{x}|}{n} = \frac{\sum |d_i|}{n}$$

2) Mean Deviation of discrete frequency distribution

$$\text{Mean Deviation} = \frac{\sum f_i |x_i - \bar{x}|}{\sum f_i} = \frac{\sum f_i |d_i|}{N}$$

where $N = \sum f_i$

3) Mean Deviation of a grouped data
Mean deviation of a grouped data is given by,

$$\text{Mean Deviation} = \frac{\sum f_i |x_i - \bar{x}|}{\sum f_i} = \frac{\sum f_i |d_i|}{N}$$

where $N = \sum f_i$

where x_i = Mid-value or class mark

Examples

1) Find the mean deviation of the data 12,6,7,3,15,10,18,5. [M.D. = 4.25]
2) Calculate mean deviation about mean of the following distribution.**[B.T.E. 2015]**

x_i	3	4	5	6	7	8
f_i	4	9	10	8	6	3

[Mean = 5.3, M.D. = 1.195]

3) Calculate mean deviation from mean for the following data:

Age in years	0-10	10-20	20-30	30-40	40-50	50-60	60-70
No. of persons	4	6	10	20	10	6	4

[Mean = 35, M.D. = 11.33]

Examples for Tutorial

1) Find the mean deviation of the data 2,3,4,7,6 [Ans. M.D. = 1.5]

2) A bowler threw maiden overs in a cricket test series as
                    3,3,4.2.5,5, 7,7,8,3,0,5
Calculate mean deviation. [Mean Deviation = 1.83]
3) The weights of a sample of crates containing books for the book store are (in kgs.)
103, 97, 101, 106, 103

4) Calculate mean deviation about mean of the following distribution. **[B.T.E. 2017]**

Marks	3	4	5	6	7	8
No. of students	1	3	7	5	2	2

[M.D. = 1.05]

5) Calculate mean deviation about mean of the following distribution.**[B.T.E. 2017]**

x_i	10	11	12	13	14
f_i	3	12	18	12	3

[Mean = 12, M.D. = 0.75]

6) Find mean deviation for the following data:

Age in years	20-22	22-24	24-26	26-28	28-30	30-32	32-34
No. of employees	70	90	110	140	130	80	80

[Mean = 27.0857 years, M.D. = 2.96 years]

7) Find mean deviation for the following data:

Class-interval	0-10	10-20	20-30	30-40	40-50
Frequency	15	20	40	20	5

[M.D.= 8.6]

8) Find the mean deviation from mean of the following distribution:
[B.T.E.2013,2014]

Marks	0-10	10-20	20-30	30-40	40-50
No. of students	5	8	15	16	6

[M.D. = 9.44]

9) Calculate mean deviation for the following data: **[B.T.E.2014]**

Expenditure (in Rs.)	40-59	60-79	80-99	100-119	120-139
No. of families	50	300	500	200	60

[M.D. = 13.50]

10) Find the mean deviation from median of the following distribution: **[B.T.E.2016]**

Weight in grams	10-15	15-20	20-25	25-30	30-35	35-40	40-45
No. of items	7	12	16	25	19	15	6

[$M.D. = 6.5$]

Standard Deviation (S.D.)

The standard deviation is defined as the square root of the mean of the squares of the deviations from mean.

Standard deviation is mean of the mean It is denoted by the Greek letter σ (read as sigma). It is useful in comparing sets of data which may have the same mean but a different range.

Note

1) The standard deviation for raw data is given by,

$$\sigma = \sqrt{\frac{\sum(x_i - \bar{x})^2}{n}} = \sqrt{\frac{\sum x_i^2}{n} - \bar{x}^2}.$$

where n is number of observations in the given data.

2) The standard deviation for frequency distribution is given by,

$$\sigma = \sqrt{\frac{\sum f_i(x_i - \bar{x})^2}{\sum f_i}} = \sqrt{\frac{\sum f_i x_i^2}{N} - \bar{x}^2} \quad \text{where } \bar{x} = \frac{\sum f_i x_i}{\sum f_i} \text{ and } N = \sum f_i$$

The standard deviation is always measured from arithmetic mean.

Variance

Variance is the arithmetic mean of the squares of deviations of all observations of the distribution from arithmetic mean. In other words, the square of the standard deviation is called variance. It is denoted by the symbol σ^2.

Note -

1) The variance for raw data is given by,

$$\therefore \text{Variance} = \frac{\sum(x_i - \bar{x})^2}{n} = \frac{\sum x_i^2}{n} - \bar{x}^2.$$

2) The variance for frequency distribution is given by,

$$\text{Variance} = \frac{\sum f_i(x_i - \bar{x})^2}{\sum f_i} = \frac{\sum f_i x_i^2}{N} - \bar{x}^2, \quad \text{where } N = \sum f_i.$$

Coefficient of Standard Deviation

The rate of change of standard deviation with respect to mean is called coefficient of standard deviation.

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

Coefficient of variation for comparing two sets of observations:

When two sets of observations are expressed in different units, then we calculate coefficient of variation by using percentage method.

The coefficient of variation is a relative measure and is defined by,

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

Note – When we want to compare the consistency (variability) of different groups, we compare their coefficient of variance.

1) A group of data having higher coefficient of variance is less consistent i.e. more variable.

2) A group of data having lower coefficient of variance is more consistent i.e. less variable.

The set of observations which has a smaller coefficient of variation is said to be more consistent i.e. it has smaller variability otherwise it has greater variability.

Examples: 1) A bowler threw maiden overs in a cricket test series as

$$3,3,4.2.5,5, 7,7,8,3,0,5$$

Calculate standard deviation and variance.

2) Calculate standard deviation for the following table: **[B.T.E.2013,2015]**

Weekly Expenditure	5	10	15	20	25
No. of Students	6	16	28	38	46

3) The frequency distribution of the cost of production of steel castings in rupees per 10 kg. for different foudaries are given below:

Rs.	5.5-6.5	6.5-7.5	7.5-8.5	8.5-9.5	9.5-10.5	10.5-11.5
No.	3	5	20	16	10	6

Find the standard deviation of the distribution. **[B.T.E.'90]**

4) The outcome of the data giving medical report of 100 students admitted to the polytechnic courses results into

Set I	Set II
Chest Measurement (in cm.)	Weight (in kg.)
$\bar{x} = 83.4$	$\bar{x} = 51.85$
$\sigma = 6.7$	$\sigma = 7.45$

Which of the two sets is more consistent?

5) If mean is 82.5, standard deviation is 7.2, find coefficient of variance. **[B.T.E.2015]** [C.V.= 8.73]

Examples for Tutorial

1)The weights of 100 students are given by the following distribution : **[B.T.E.2016]**

Weight above or equal to	36	41	46	51	56	61	66	71
No. of students	100	96	79	56	28	11	5	2

Calculate (i) Mean (ii) Variance of the data using Step Deviation Method . No student has weight above 75 kg. [Mean = 51.85, Variance = 55.35]

2)The following data gives the percentage defective observed on a radio production line:

$$12,17,30,6,8,10,11,13,7,4,2 \text{ and } 9.$$

Find range and standard deviation.
[Range = 28,   Standard Deviation = 6.99]

3) Calculate mean and standard deviation of the following frequency distribution:**[B.T.E.2016]**

Class interval	0-10	10-20	20-30	30-40	40-50
Frequency	14	23	27	21	15

4) Find the standard deviation of the following data: :**[B.T.E.2014]**

Class interval	0-10	10-20	20-30	30-40	40-50
Frequency	3	5	8	3	1

5) Find the standard deviation of the following: **[B.T.E.2015]**

Class interval	0-20	20-40	40-60	60-80	80-100
Frequency	20	130	220	70	60

$[S.D.(\sigma) = 20.38]$

6) Find the standard deviation of the following: **[B.T.E.2017]**

Class interval	0-5	5-10	10-15	15-20	20-25	25-30	30-35	35-40
Frequency	3	5	9	15	20	16	10	2

$[\sigma = 8.177]$

7) Find variance and coefficient of variance for the following data: **[B.T.E.2015]**

Class-interval	55-65	65-75	75-85	85-95	95-105	105-115	115-125
No. of workers	10	12	15	20	14	7	2

8) Find variance and coefficient of variance of the following data: **[B.T.E.2014]**

Class-interval	0-10	10-20	20-30	30-40	40-50
Frequencies	14	23	27	21	15

9) Find variance and coefficient of variance for the following distribution: **[B.T.E.2014]**

Class-interval	20-25	25-30	30-35	35-40	40-45	45-50	50-55	55-60	60-65
Frequencies	25	30	50	90	75	60	35	25	15

11) Find variance from the following series.**[B.T.E.2015]**

Age Under	10	20	30	40	50	60	70	80
No. of persons	15	30	53	75	100	110	115	125

$[S.D.(\sigma) = 19.75, \ Variance(\sigma^2) = 390.0625]$

12) Calculate the standard deviation and coefficient of variation for the following data giving average annual rainfall in cm. of 100 places. **[B.T.E.2016]**

Rainfall	70-80	80-90	90-100	100-110	110-120	120-130	130-140	140-150
Number of places	6	7	12	19	21	18	11	6

[Answer: S.D. = 18.25, C.V. = 16.3]

13) Coefficient of variation of a distribution is 75% and standard deviation is 24. What is its mean? [Mean = 32]

14) Coefficient of variation of certain distribution is 5 and mean is 60. Find the standard deviation? [S.D. = 3]