6 Dispersion

Introduction

The various measures of central tendency that we studied in the last chapter describe the characteristic of the entire data by means of a single value that is the central value of the distribution. But with these measures alone we cannot form a clear idea about the distribution. Consider the following series :

A	28	29	30	30	32
B	30	30	30	30	30
C	1	2	30	30	87

Compute the mean, median and mode of the two data sets. As you see from your results, the two datasets have the same mean, median and mode, all are equal to 30. The three data sets also contain same number of elements. But the three sets are different. What is the main difference among them?

Three data sets have same average, but they differ in deviation of observations. ie, Data set C is more scattered than in data sets A and B. The values in C are more spread out. ie, They lie farther away from their mean than in the case of A and B. Thus, The degree to which numerical data tend to spread about an average value is called the dispersion, or variation, of the data.

Dispersion

> Dispersion is the degree of scatter or variation of the variable about a central value.

There are sevaral measures of variability or dispersion, they are

1. Range
2. Quartile Deviation
3. Mean Deviation

4. Standard Deviation

Properties of Measure of Dispersion

A good measure of dispersion should possess the following properties

- It should be rigidly defined.
- It should be based on all observations.
- It should be simple to understand and easy to calculate.
- It should be amicable to further algebraic treatment.
- It should not be unduly affected by extreme values.

Importance of Measuring Dispersion

- It tells about the variability of a data.
- It enables us to compare two or more distributions.
- It is of great importance in advanced statistical analysis.

6.1 Range

Range(R) is the difference between the highest value (H) and lowest value (L) in a set of data. In symbols $R = H - L$

- Higher value of range implies higher dispersion and vice versa.

Merits

- It is the simplest measure of dispersion.
- It is simple to understand and easy to calculate.

Demerits

- It gives importance to the two extreme values only,and therefore it may be unduly influenced by extreme values.
- It is not a reliable measure of dispersion on many occasions.

Illustration 6.1

Given below the data of price (Rupees) of 1gm gold in the first week of August 2013.

Monday	*Tuesday*	*Wednesday*	*Thursday*	*Friday*	*Saturday*
2530	*2500*	*2430*	*2560*	*2680*	*2600*

Calculate range of price in the week

Solution.

$$\text{Range} = \text{highest value} - \text{lowest value}$$
$$= 2680 - 2430$$
$$= 250 \text{ rupees}$$

Activity

Look at the series 10,30,50,100

- Calculate range if 30 is replaced by 20.
- Calculate range if 100 is not present in it.

6.2 Quartile Deviation (QD)

One drawback of range is that it depends only on extreme values and we cannot find range in open end classes. As a modification the difference between 3^{rd} quartile and 1^{rd} quartile is called inter quartile range and half of the difference between 3^{rd} quartile and 1^{st} quartile is called Quartile Deviation (QD) . Quartile deviation is also called semi-inter quartile range.

$$\text{Inter quartile range} = Q_3 - Q_1$$

$$\text{Quartile deviation} = \frac{Q_3 - Q_1}{2}$$

Where $Q_3 = 3^{rd}$ quartile, $Q_1 = 1^{st}$ quartile

Quartile Deviation for a Raw Data

To find QD for a raw data consisting of n values ,first arrange the data in ascending order of magnitude, then find Q_1 and Q_3 as explained in Chapter 5 ,Section 6

$$QD = \frac{Q_3 - Q_1}{2}$$

Illustration 6.2

A statistical data was collected from 11 school children on the number of hours they spend watching television in one week. The data is given below 3, 8.5, 12, 9, 16.5, 9, 14, 20, 18, 19, 20. Find quartile deviation

Solution.

Arrange the data in ascending order of magnitude as follows

3, 8.5, 9, 9, 12, 14, 16.5, 18, 19, 20, 20

$$\begin{aligned} Q_1 &= \text{value of } \left(\frac{n+1}{4}\right)^{\text{th}} \text{ item} \\ &= \text{value of} \left(\frac{11+1}{4}\right)^{\text{th}} \text{ item} \\ &= 3^{\text{rd}}\text{item} \\ &= 9 \\ Q_3 &= \text{value of } \left[3\left(\frac{n+1}{4}\right)\right]^{\text{th}} \text{ item} \\ &= \text{value of } 9^{\text{th}}\text{item.} \\ &= 19 \end{aligned}$$

$$QD = \frac{19-9}{2}$$
$$= 5$$

Quartile Deviation for a Discrete Frequency Distribution

Consider a discrete frequency distribution consisting of N values. To find QD first arrange the values in the ascending order of magnitude, then find Q_1 and Q_3 for a discrete frequency distribution as explained in chapter 5, section 9. Then $QD = \dfrac{Q_3 - Q_1}{2}$.

Illustration 6.3

Calculate Q D from the following data

Score	*5*	*10*	*15*	*20*	*25*	*30*
Number of students	*4*	*7*	*15*	*8*	*7*	*2*

Solution.

Less than cumulative frequency table is given below

Scores (x)	Number ofstudents(f)	Less than cum. frequency
5	4	4
10	7	11
15	15	26
20	8	34
25	7	41
30	2	43
	$N = 43$	

$$
\begin{aligned}
Q_1 &= \text{Value of}\left(\frac{N+1}{4}\right)^{\text{th}} \text{ item} \\
&= \text{Value of}\left(\frac{43+1}{4}\right)^{\text{th}} \text{ item} \\
&= \text{Value of } 11^{\text{th}} \text{ item} \\
&= 10 \\
Q_3 &= \text{Value of}\left(\frac{3(N+1)}{4}\right)^{\text{th}} \text{ item} \\
&= \text{Value of}\left(\frac{3(43+1)}{4}\right)^{\text{th}} \text{ item} \\
&= \text{Value of } 33^{\text{rd}} \text{ item} \\
&= 20
\end{aligned}
$$

$$
\begin{aligned}
QD &= \frac{Q_3 - Q_1}{2} \\
&= \frac{20 - 10}{2} \\
&= 5
\end{aligned}
$$

Quartile Deviation for a Continuous Frequency Distribution

You have already learnt to find Q_3 and Q_1 in a continuous frequency distribution in Chapter 5, Section 6

$QD = \dfrac{Q_3 - Q_1}{2}$

Illustration 6.4

A survey of domestic consumption of electricity in a colony gave the following distribution of units consumed

units	*Below 100*	*100-200*	*200-300*	*300-400*	*400-500*	*500-600*	*600-700*	*above 700*
No.of consumers	*20*	*21*	*30*	*46*	*20*	*25*	*16*	*10*

Find quartile deviation.

Prepare less than cumulative frequency table as explained in unit 3

No.of units	No.of consumers(f)	Less than frequency
Below 100	20	20
100-200	21	41
200-300	30	71
300-400	46	117
400-500	20	137
500-600	25	162
600-700	16	178
700& above	10	188
	$N = 188$	

$$Q_1 = l_1 + \left(\frac{\frac{N}{4} - m_1}{f_1}\right) \times C_1$$

$$\frac{N}{4} = \frac{188}{4}$$

$$= 47$$

The value 47 lies in 200-300 class

Therefore Q_1 lies in 200-300 class

$$Q_1 = 200 + \frac{47 - 41}{30} \times 100$$

$$= 200 + 20$$

$$= 220$$

$$Similarly, Q_3 = l_3 + \left(\frac{\frac{3N}{4} - m_3}{f3}\right) \times C_3$$

$$3 \times \frac{N}{4} = 141$$

The value 141 lies in 500-600 class

$$Q_3 = 500 + \left[\frac{141 - 137}{25}\right] \times 100$$
$$= 500 + 16$$
$$= 516$$

$$QD = \frac{516 - 220}{2}$$
$$= 148$$

Merits

- It can be calculated for open end classes.
- It is not unduly affected by extreme values.

Demerits

- It is not based on all observations.
- It is not capable of further algebraic treatments.

Know your progress

1. The following data represents the profit(in lakhs) of 50 business men in a large city. Find QD of the data

Profit	20-29	30-39	40-49	50-59	60-69
No.of Business man	8	12	20	7	3

2. The heights of students (in inches) is given below. Calculate QD of the data
   Heights: 55, 54, 57, 67, 60, 61, 58, 63

3. Prices of shares of a company were as under from Monday to Saturday for 40 weeks. Find QD of shares

Day	Mon	Tue	Wed	Thu	Fri	Sat
Price	150	200	190	210	230	180
No.weeks	5	5	8	10	5	7

Activity

Collect a data on height of students in your class and find quartile deviation and range. Interpret your results.

6.3 Mean Deviation (MD)

Even though quartile deviation is an improvement over range, it depends only on two values Q_3 and Q_1. As a solution we try to introduce another measure of dispersion which depends on all values in a series. Mean deviation is the arithmetic mean of absolute deviation of the observations from an assumed average.

Mean Deviation for a Raw Data

Let us consider a raw data of 'n observations. mean deviation is calculated as follows

Mean deviation= $\dfrac{\sum|x-A|}{n}$, where A is any average

Compute an average (A) which is required to calculate mean deviation (mean or median or mode)

Find the absolute deviation of the observations from the average to each value. It is denoted by $|x-A|$

The arithmetic mean of these absolute deviations is mean deviation

Illustration 6.5

Eleven students were selected and asked how many hours each of them studied the day before the final examination in statistics. Their answers were recorded here 8, 11, 5, 4, 5, 0, 2, 6, 9, 3, 2. Calculate Mean Deviation about mean

Solution.

$$\text{Mean deviation about mean}=\frac{\sum|x-\text{Mean}|}{n}$$

Duration of study(x)	$|x-\text{Mean}|$
8	$|8-5|=3$
11	$|11-5|=6$
5	$|5-5|=0$
4	$|4-5|=1$
5	$|5-5|=0$
0	$|0-5|=5$
2	$|2-5|=3$
6	$|6-5|=1$
9	$|9-5|=4$
3	$|3-5|=2$
2	$|2-5|=3$
$\sum x=55$	$\sum|x-\text{mean}|=28$

$$\begin{aligned}\text{Mean} &= \frac{\sum x}{n}\\ &= \frac{55}{11}\\ &= 5\\ \text{Mean deviation} &= \frac{\sum|x-\text{Mean}|}{n}\\ &= \frac{28}{11}=2.54.\end{aligned}$$

Mean Deviation for a Discrete Frequency Distribution

For a discrete frequency distribution consisting of N observations, the procedure and formulae is as follows

Mean deviation= $\frac{\sum f|x-A|}{N}$ where A is any average

Step 1

Find the average (A) which is required to calculate mean deviation.

Step 2

Find the absolute deviations of the observations from the average to each value.ie $|x-A|$.

Step 3

Multiply $|x-A|$ by their frequency f. Thus we get $f|x-A|$ to each observations.

Step 4

The arithmetic mean of these values is the mean deviation.

Illustration 6.6

The following table shows the number of books read by students in a literature class consisting of 28 students,in a month.

No. of Books	*0*	*1*	*2*	*3*	*4*
No.of students	*2*	*6*	*12*	*5*	*3*

Calculate mean deviation about mode of number of books read

Solution. Mean deviation about mode = $\frac{\sum f|x-\text{Mode}|}{N}$

mode=2 (the value which has maximum frequency).

No. of books	$\|x - \text{mode}\|$	f	$f\|x - \text{mode}\|$
0	$\|0-2\| = 2$	2	4
1	$\|1-2\| = 1$	6	6
2	$\|2-2\| = 0$	12	0
3	$\|3-2\| = 1$	5	5
4	$\|4-2\| = 2$	3	6
		$N = 28$	$\sum f\|x - \text{mode}\| = 21$

$$\text{Mean deviation about mode} = \frac{21}{28}$$
$$= 0.75$$

Mean Deviation for a Continuous Frequency Distribution

To find mean deviation, convert the continuous frequency distribution into discrete frequency distribution by taking middle values of each classes. The procedure and formulae are same in a discrete frequency distribution

Illustration 6.7

Calculate mean deviation from median of the following data

Height in cm	*100-120*	*120-140*	*140-160*	*160-180*	*180-200*
No.of students	*4*	*6*	*10*	*8*	*5*

Solution. Mean deviation about median= $\dfrac{\sum f|x - \text{median}|}{N}$

Find middle values of each class taken then prepare a table as given below

Class	Mid. value(x)	Freq. (f)	Cum. Freq.	$\|x - \text{median}\|$	$f\|x - \text{median}\|$
100-120	110	4	4	$\|110-153\| = 43$	172
120-140	130	6	10	$\|130-153\| = 23$	138
140-160	150	10	20	$\|150-153\| = 3$	30
160-180	170	8	28	$\|170-153\| = 17$	136
180-200	190	5	33	$\|190-153\| = 37$	185
Total		33			661

$$N = 33\,,\ \sum f|x - \text{median}| = 661$$

$$\frac{N}{2} = \frac{33}{2} = 16.5$$

The class having cum.frequency 16.5 is 140-160

Therefore 140-160 is median class

$$\begin{aligned}\text{median} &= l + \frac{\frac{N}{2} - m}{f} \times c \\ &= 140 + \frac{16.5 - 10}{10} \times 20 \\ &= 153\end{aligned}$$

$$\begin{aligned}\text{Mean deviation about median} &= \frac{\sum f|x - \text{median}|}{N} \\ &= \frac{661}{33} \\ &= 20.03\end{aligned}$$

Merits

- It is based on all observations.
- It can be calculated from any value.
- It is not much affected by extreme values.

Demerits

- Ignoring signs of deviations may create artificiality.
- Computation is not much easier.

Know your progress

1. The following table shows the frequency distribution of grade points

Grade point	8	6	4	2
No.of students	4	20	5	1

Find mean deviation of the grade points about mode.

2. Find mean deviation about mean of the data which relate to the sales of 100 companies.

Sales(Rs.thousand)	40-50	50-60	60-70	70-80	80-90	90-100
No.of days	10	15	25	30	12	8

Activity

Look at the data 5, 8, 3, 2, 12

- Calculate mean deviation about mean
- Calculate mean deviation about median
- Calculate mean deviation about mode
- Try yourself to find which of them is minimum

6.4 Standard Deviation(SD)

In order to have a more meaningful measure to compute the variability of a data we use SD. The concept of SD was introduced by Karl Pearson. A relatively small standard deviation indicates high degree of uniformity in the data with not much variation of individual values from their mean. In particular, if all the observations are equal the SD is equal to zero.

Standard deviation (SD) is defined as the positive square root of the mean of squares of deviations from the arithmetic mean. It is denoted by Greek letter σ (sigma).It cannot be negative. It is the best measure of dispersion.

Properties

1. The minimum value of SD is zero.
2. The sum of squared deviation is minimum when taken about mean.

Standard deviation for a raw data

Let us consider a raw data of n observations. SD can be calculated as follows

$$SD, \sigma = \sqrt{\frac{\sum(x-\overline{x})^2}{n}}$$

For computation purpose we can use another formula derived from the above. The procedure is discussed below

Step 1

Calculate arithmetic mean $\overline{x}$ of the given data

Step 2

Calculate x^2 for each observation.

Step 3

Calculate $\frac{\sum x^2}{n}$, where 'n' is no. of values

Step 4

Calculate $\sigma = \sqrt{\frac{\sum x^2}{n} - \overline{x}^2}$

$$SD, \sigma = \sqrt{\frac{\sum x^2}{n} - \overline{x}^2}$$

Illustration 6.8

A survey was conducted for the number of road accidents in a major city during 11 successive weeks. The results are given below 8 ,6 ,3 ,0 ,5 ,9 ,2 ,1 ,3 ,5 ,2 . Clculate SD of road accidents

Solution.

$$SD = \sqrt{\frac{\sum x^2}{n} - \overline{x}^2}$$

x	8	6	3	0	5	9	2	1	3	5	2	$\sum x = 44$
x^2	64	36	9	0	25	81	4	1	9	25	4	$\sum x^2 = 258$

$$\begin{aligned} \overline{x} &= \frac{44}{11} \\ &= 4 \\ \sigma &= \sqrt{\frac{258}{11} - 4^2} \\ &= \sqrt{23.45 - 16} \\ &= 2.73 \end{aligned}$$

Standard Deviation for a Discrete Frequency Distribution

For a discrete frequency distribution consisting of N $(\sum f)$ observations, the procedure and formulae is discussed below

Step 1

Calculate the arithmetic mean of the data, $\overline{x} = \dfrac{\sum fx}{N}$

Step 2

Calculate x^2 and fx^2 for each value

Step 3

Calculate $\sum fx^2$

Step 4

Calculate SD $= \sqrt{\dfrac{\sum fx^2}{N} - (\overline{x})^2}$

$$SD, \sigma = \sqrt{\frac{\sum fx^2}{N} - (\overline{x})^2}$$

Illustration 6.9

25 students were given an arithmetic test. The time in minute to complete the *test is as follows*

Time in minutes	1	2	3	4	5
No.of students	4	3	10	5	3

Calculate SD of their time to complete the test

Solution.

Prepare a table given below

x	f	x^2	fx	fx^2
1	4	1	4	4
2	3	4	6	12
3	10	9	30	90
4	5	16	20	80
5	3	25	15	75
	$N = 25$		$\sum fx = 75$	$\sum fx^2 = 261$

$$\overline{x} = \frac{\sum fx}{N}$$
$$= \frac{75}{25}$$
$$= 3$$

$$\sigma = \sqrt{\frac{\sum fx^2}{N} - \overline{x}^2} = \sqrt{\frac{261}{25} - 3^2}$$
$$= \sqrt{10.44 - 9}$$
$$= \sqrt{1.44}$$
$$= 1.2$$

Standard Deviation for a Continuous Frequency Distribution

To find SD, convert the continuous frequency distribution into discrete frequency distribution by taking mid values of each class. The procedure and formulae are same as in the case of a discrete frequency distribution

Illustration 6.10

A study of 100 engineering companies gives the following information

Profit(in crore)	0-10	10-20	20-30	30-40	40-50	50-60
No.of companies	8	12	20	30	20	10

Find SD of profit earned.

Solution.

Computation table

x	f	x^2	fx	fx^2
5	8	25	40	200
15	12	225	180	2700
25	20	625	500	12500
35	30	1225	1050	36750
45	20	2025	900	40500
55	10	3025	550	30250
	$N = 100$		$\sum fx = 3220$	$\sum fx^2 = 122900$

$$\begin{aligned}
\overline{x} &= \frac{\sum fx}{N} \\
&= \frac{3220}{100} \\
&= 32.2 \\
\sigma &= \sqrt{\frac{\sum fx^2}{N} - \overline{x}^2} \\
&= \sqrt{\frac{122900}{100} - (32.2)^2} \\
&= \sqrt{192.16} \\
&= 13.86
\end{aligned}$$

Merits

- It is rigidly defined and based on all observations.
- It is capable of further algebraic treatments.
- It is less affected by the fluctuations of sampling than other measures of dispersion.

Demerits

- It is difficult to calculate.
- It is impossible to find it in open end classes.

Know your progress

1. Calculate SD of daily income of 10 persons in rupees given below 227, 235, 255, 269, 292, 299, 312, 321, 333, 348
2. The following table gives the distribution of expenditure of 100 families in a village. Calculate SD

Income	0-1000	1000-2000	2000-3000	3000-4000	4000-5000	5000-6000
No.of families	18	26	30	12	10	4

Activity

If a student scored equal scores in 6 subjects in an examination, calculate SD and interpret your result.

Activity

Observe the data 40 ,42 ,38 ,44 ,46 ,48 ,50

- Compute SD of the data.
- Add 3 to each value then find new SD.

- Subtract 3 to each value then find new SD.
- Multiply 3 by each value then find new SD.
- Divide 3 by each value then find new SD.
- Interpret your findings

Variance

The term variance was introduced by R.A.Fisher. It is defined as the square of SD

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

$$Or, \text{SD} = \sqrt{\text{variance}} = \sigma$$

It has many applications in advanced statistical analysis

6.5 Relative measures of dispersion

The measures of dispersion so far discussed are called as absolute measures of dispersion (Range, QD, MD, SD). Because they measures the variability or deviation of values in a data and expressed in the original unit of that data. If the units are different, absolute measures of dispersion is not suitable for a direct comparison of two or more series because it may result an incorrect conclusion regarding variation of the data. For these reasons, a measure of dispersion which is independent of unit of measurement is suggested. Such a measure is called relative measure of dispersion. It is the ratio of absolute measures of dispersion to an average with which the measure of dispersion is computed

Features of relative measure:

- It is a ratio.
- It is a pure number.
- It is free from the unit of measurement of observations.

- It is used for comparing two or more sets of data.

The two important relative measures of dispersion we are going to discuss are

1. Coefficient of variation
2. Coefficient of quartile deviation

Coefficient of Variation (CV)

It is the most commonly used measure to compare the consistency or stability between two or more sets of data. Coefficient of variation is defined as SD divided by its arithmetic mean expressed in percentage

$$Coefficient\ of\ variation,\ CV = \frac{\sigma}{\overline{x}} \times 100$$

Suppose, between two groups we have to find the consistent group. First find the CV for each group. The group with less CV is considered to be the consistent group. It is free from the unit of original measurements.

Illustration 6.11

In two factories A & B located in same industrial area, the weekly wage(in Rupees) and standard deviations are as follows

Factory	*Average*($\overline{x}$)	$SD(\sigma)$	*Number of workers*
A	*500*	*5*	*476*
B	*600*	*4*	*524*

1. *Which factory pays larger amount as weekly wages?*
2. *Which factory has greater consistency in individual wages?*

Solution. Here $n_1 = 476, \overline{x}_1 = 500\ \sigma_1 = 5, n_2 = 524, \overline{x}_2 = 600, \sigma_2 = 4$

1.

$$\text{Total wages paid by factory A} = 500 \times 476$$
$$= 238000$$
$$\text{Total wages paid by factory B} = 600 \times 524$$
$$= 314400$$

2.

$$\text{CV of factory A} = \frac{\sigma_1}{\bar{x}_1} \times 100$$
$$= \frac{5}{500} \times 100$$
$$= 1$$
$$\text{CV of factory B} = \frac{\sigma_2}{\bar{x}_2} \times 100$$
$$= \frac{4}{600} \times 100$$
$$= 0.67$$

Since total wages of factory B is greater than factory A, factory B pays larger amount as weekly wages. Since CV is greater for A compared to B, factory A has greater variability.

Know your progress

Prices of a particular commodity in 5 years at two different cities in Kerala and Tamilnadu are as follows

Tamilnadu(RS)	20	22	19	22	23
Kerala(RS)	18	12	10	20	15

Which state has stable price?justify your answer

Activity

Form two groups in your class, collect the grades scored by statistics examination and then find which of the two groups is more stable

.

Coefficient of Quartile Deviation

Quartile deviation is an absolute measure of dispersion .The relative measure corresponding to quartile deviation is coefficient of quartile deviation.

$$\text{Coefficient of Quartile Deviation} = \frac{Q_3 - Q_1}{Q_3 + Q_1}$$

It can be used to compare the degree of variation in different situations

6.6 Covariance

It is a measure of strength of linear relationship between two variables. Covariance indicates whether the variables are positively related or negatively related in a bivariate distribution.

$COV(x,y) = \dfrac{\sum(x-\overline{x})(y-\overline{y})}{n}$, where '$n$ is the number of observations in a data.

For computation purpose we can use another formulae

$$COV(x,y) = \frac{\sum xy}{n} - \overline{x} \times \overline{y}$$

If $COV(x,y)$ is positive the variables move in same direction. If it is negative they move in opposite direction .If two variables are unrelated, then $COV(x,y)$ should be zero.

Illustration 6.12

The following data pertains to the length of service(in years) and the annual income in thousands for a sample of 8 employees in an industry. Find covariance.

Length of service(x)	6	8	9	10	11	13	15	16
Annual income (y)	14	17	15	18	16	22	25	25

Solution. $COV(x,y) = \dfrac{\sum xy}{n} - \overline{x} \times \overline{y}$

x	y	xy
6	14	84
8	17	136
9	15	135
10	18	180
11	16	176
13	22	286
15	25	375
16	25	400
$\overline{x} = 11$	$\overline{y} = 19$	$\sum xy = 1772$

$$\text{Cov(x,y)} = \frac{1772}{8} - 11 \times 19$$
$$= 12.5$$

Let us sum up

In this chapter we have discussed the concept of dispersion and some absolute and relative measures of dispersion . The absolute measures of dispersion helps us to calculate the deviation of values among themselves or deviation of values from an average. Among this, Range and quartile deviation attempt to measure deviation of values among themselves. But mean deviation and standard deviation attempt to measure variation of values from their average. The relative measures of dispersion are free from the unit of measurement of observations and are used to compare two sets of data. A measure of relationship between two variables is also introduced using covariance.

Learning outcomes

After transaction of this unit, the learner:-

- recognises the importance of measuring dispersion.
- explains and evaluates the measures of dispersion-Range, QD,MD and SD.
- distinguishes absolute and relative measures of dispersion.

Evaluation Items

1. If a constant is subtracted from each observation then variance ________
2. If each observation is divided by 10 then SD of the new observations is ________
3. If first 25% of values in a data is 20 or less and last 25% is 50 or more then QD is ________
4. If the lowest value of a set is 9 and its range is 57 then the highest value is ________
5. If the CV of a distribution is 50 and its SD 20,the arithmetic mean shall be ________
6. The SD of five observations 5 ,5 ,5 ,5, 5 is ________
7. The mean of squared deviation from mean is called ________
8. Indicate whether the following statements are true or false
   (a) Range is the best measure of dispersion
   (b) QD is more suitable in the case of open-end distributions
   (c) Absolute measure of dispersion can be used for the purpose of comparison
   (d) Mean deviation is least when deviations are taken from median
   (e) SD can never be negative

9. Distinguish between absolute and relative measures of dispersion

10. Relative measures of dispersion are used for comparison , why?

11. What do you understand by dispersion of a set of values? How do you measure dispersion?

12. What are the different absolute measures of dispersion?

13. Why SD is called the best measures of dispersion?

14. Suppose each measurement in a distribution is multiplied by 2 what happens to the following?

    (a) Mean of the distribution.

    (b) Variance of the distribution.

    (c) SD of the distribution.

15. For a data consisting of 9 observations, the sum is 9 values is 360. And the sum of squares of deviation taken from mean is 288. Find

    (a) standard deviation

    (b) $C.V$.

16. The number accidents occurred due to careless driving on a busy road during 7 days in week is reported below

No.of accidents	7	10	4	7	9	3	2

    Determine mean deviation about mean.

17. Following are the responses from 52 students to the question about how much money they spend everyday

Money spent	No.students
5	2
10	5
15	10
20	7
25	8
30	5
35	6
40	4
45	5

Find quartile deviation.

18. For a newly created post the manager interviewed the candidates in 5 days. The number of candidates in each day were 16,19,15,15,&14 respectively . Find variance

19. A factory produces two types of electric lamps A & B. In an experiment relating to their life the following results were obtained.

Length of life (Hrs)	Number of lamps (A)	Number of lamps (B)
500-700	5	4
700-900	11	30
900-1100	26	12
1100-1300	10	8
1300-1500	8	6

(a) Which bulb A or B has more average life.

(b) Which is more consistent.

20. Annual tax paid by certain employees is given below

Tax Paid in thousand	5-10	10-15	15-20	20-25	25-30	30-35	35-40
No. of Emplyees	18	30	46	28	20	12	6

Find semi inter quartile range.

21. For 108 randomly selected higher secondary students the following I Q distribution was obtained

I Q	90-98	99-107	108-116	117-125	126-134
No. students	6	10	25	15	4

Find standard deviation of IQ.

22. Prices of a particular commodity in 5 years in two cities are given below

Price in city A	20	22	19	23	16
Price in city B	10	20	18	12	15

Which city has stable prices using coefficient of quartile deviation.

23. The number of motor cycle accident cases reported in a hospital is as follows

Age	0-10	10-20	20-30	30-40	40-50	50-60
No. of accidents	15	49	37	20	6	1

Find mean deviation about median.

24. The following income distribution was obtained from 1000 persons

Income more than	1000	900	800	700	600	500	400	300	200
No.of persons	0	50	110	200	400	650	825	950	1000

Find quartile deviation.

25. Nine students of B.Com class of a college have obtained the following marks in statistics out of 100. Calculate standard deviation and variance

Sl no	1	2	3	4	5	6	7	8	9
Marks in Statistics	5	10	20	25	40	42	45	48	70

26. The age group and the number of diabetic patients treated in a hospital is listed below

Age	20	30	40	50	60	70	80
No.of .patients	3	10	30	40	35	10	7

Determine mean deviation about median.

27. The number of television sets sold in a week from a home appliance show room is given below

Number of sets	5	6	7	8	4	3	1

Find standard deviation and variance.

28. In a village ,25% of the persons earned more than Rs.10000/- whereas 75% earned only more than Rs.5000/-. Calculate relative and absolute values of dispersion.

29. The following table relates to advertisement expenditure(y) and sales (x) of a company for a period of 10 years. Calculate $COV(x,y)$

Sales(x)	50	5	50	40	30	20	20	15	10
Adv.Exp.(y)	700	200	650	500	450	400	300	250	210

Answers

1. does not change 2. 1/10th of SD of original obs 3. 15 4. 66 5. 40 6. 0 7. SD 8. (a) false (b)true (c) false (d)true (e) true 14. (A) .2 times (B) 4 times (C) .2 times 15. SD=5.66,CV=6.28 16. MD about mean= 2.57 17. QD=10 18. var=2.96 19. (A) bulb A has more average life (B) bulb A is more consistent 20. semi inter quartile range=5.48 21. SD=9.36 22. coe:QD of cityA=0.125,coe: QD of cityB=0.267 therefore cityA has more stable price 23. MD about median=8.906 24. QD=137.5 25. SD=19.46, variance=378.69 26. MD about median=10 27. SD=2.231,var=4.979 28. QD= 2500,coefficient of QD=0.33 29. $COV(x,y) = 2694.44$