CHAPTER

9

MEASURES OF DISPERSION

OBJECTIVES

After studying the material in this chapter, you should be able to :

- *Understand the meaning of the term dispersion.*
- *Know the requisites of an ideal measure of dispersion.*
- *Differentiate between relative and absolute measures of dispersion.*
- *Identify different types of measures of dispersion.*
- *Calculate various measures of dispersion, such as the range, quartile deviation, mean deviation, standard deviation and variance.*
- *Know the merits and demerits of different measures of dispersion.*
- *Appreciate properties of standard deviation.*
- *Understand the concept of coefficient of variation.*

9.1 INTRODUCTION

The various measures of central tendency discussed in the previous chapter do not by themselves give an adequate description of the data. Regardless of the measure that may be used, it does not tell us whether the data are grouped closely together or spread over a large range of values. It is quite possible to have two or more sets of observations with the same mean but still differ considerably in the variability of their measurements about the average. For example, consider the following three sets of observations, each containing 9 items :

											Total	Mean
Set A	:	10	10	10	10	10	10	10	10	10	90	10
Set B	:	6	7	8	9	10	11	12	13	14	90	10
Set C	:	2	4	6	8	10	12	14	16	18	90	10

All the three sets of data have the same mean but otherwise they are quite different. Each set has a sum of 90, but in the first set this sum is dispersed among all the values. In fact, none of the observations in set *A* deviates from the mean and we say that there is no dispersion at all in set *A*. However, comparing the observations in sets *B* and *C*, it is quite obvious that set *B* is more uniform than set *C*. We say that the variability or the dispersion of the observations from the average is less for set *B* than for set *C*. Thus we need some measure of dispersion of a set of numbers. The dispersion of a data measures the degree to which numerical data tend to spread about an average value.

The measures of central tendency are, therefore, inadequate to describe the data completely. They must be supported and supplemented by some other measures.

9.2 DISPERSION

In the following we shall give some important definitions of dispersion.

1. *Dispersion is the measure of the variations of the items.* **—A.L. Bowley**
2. *Dispersion is the measure of the extent to which the individual items vary.* **—L.R. Connor**
3. *The degree to which numerical data tend to spread about an average value is called the variation or dispersion of the data.* **— Spiegel**
4. *Dispersion or spread is the degree of the scatter or variation of the variables about a central value.* **—Brooks and Dick**
5. *The term dispersion is used to indicate the facts that within a given group, the items differ from one another in size or in other words, there is a lack of uniformity in their sizes.* **—W.I. King**

It is clear from these definitions that dispersion (also called scatter, spread or variation) measures the extent to which the individual items vary from a central value.

9.3 MEASURES OF DISPERSION

The degree to which numerical data tend to spread about an average value is called the **variation** or **dispersion** of the data. There are various measures of dispersion, the most important being the *range*, the *quartile deviation*, the *mean deviation* and the *standard deviation*. The measures of dispersion can be classified as follows:

(i) *Absolute measures of dispersion.*

(ii) *Relative measures of dispersion.*

Absolute Measures of Dispersion. The measures of dispersion which are expressed in the same statistical unit in which the original data are given are termed as **absolute measures of dispersion.** They have a serious drawback because the dispersion obtained from these measures cannot be used to compare the variability of two distributions expressed in different units.

Relative Measures of Dispersion. These measures are pure numbers independent of the units of measurement and can be used to compare the variability of two distributions expressed in different units.

9.4 PROPERTIES OF A GOOD MEASURE OF DISPERSION

A good measure of dispersion should possess the following properties:

1. *It should be simple to understand.*
2. *It should be easy to calculate.*
3. *It should be rigidly defined.*
4. *It should be based on all the observations.*
5. *It should be amenable to further mathematical treatment.*
6. *It should have sampling stability.*
7. *It should not be unduly affected by extreme observations.*

9.5 RANGE

Definition. The **range** of a set of data is defined as the difference between the largest and the smallest value in the set. Symbolically,

$$\text{Range} = L - S$$

where L = largest value and S = smallest value.

In case of a grouped frequency distribution, range is defined as the difference between the upper limit of the highest class and the lower limit of the smallest class.

Coefficient of Range. The range is an absolute measure of dispersion and is expressed in the unit of measurement of values of a distribution. Hence it cannot be used to compare two distributions expressed in different units. To overcome this difficulty we need a relative measure which is independent of the units of measurement. This relative measure, called the **coefficient of range,** is defined as follows:

$$\text{Coefficient of range} = \frac{\text{Range}}{\text{Sum of the largest and the lowest values}} = \frac{L-S}{L+S}.$$

Example 1. Find the range and the coefficient of range for the following observations:

65, 70, 82, 59, 81, 76, 57, 60, 55 and 50

Solution.

$$\text{Range} = L - S$$

Here, $L = 82$ and $S = 50$

$\therefore$

$$\text{Range} = 82 - 50 = 32$$

$$\text{Coefficient of range} = \frac{L-S}{L+S} = \frac{82-50}{82+50} = \frac{32}{132} = 0.24.$$

Example 2. Calculate range and coefficient of range from the following data:

Marks :	10 – 20	20 – 30	30 – 40	40 – 50	50 – 60	60 – 70	70 – 80	80 – 90
No. of Students :	2	6	12	18	25	20	10	7

Solution. Range = $L - S = 90 - 10 = 80$

and Coefficient of range = $\frac{L-S}{L+S} = \frac{90-10}{90+10} = \frac{80}{100} = 0.8.$

Note. It should be noted that in the calculation of 'Range' only the values of the variable are taken into account and the frequencies are completely ignored.

9.6 MERITS AND DEMERITS OF RANGE

Merits. The range possesses the following merits:

1. *It is simple to understand and easy to calculate.*
2. *It requires minimum time to calculate the value of range.*

Demerits. The range has the following drawbacks:

1. *It is not based on all the observations.*
2. *Range is a poor measure of variation. It considers only the extreme values and tells us nothing about the distribution of numbers in between. Consider, for example, the following two sets of data, both with a range of 12:*

A :	8	9	10	11	13	14	15	17	20
B :	8	12	12	12	13	13	13	14	20

In set A the mean and median are both 13, but the numbers vary over the entire interval from 8 to 20. In set B the mean and median are also 13, but most of the values are closer to the center of the data.

3. *It is very much affected by fluctuations of sampling. Its value varies widely from sample to sample.*
4. *It cannot be calculated for grouped frequency distribution with open-end classes.*
5. *It is not suitable for further mathematical treatment.*

Uses of Range. Although the range is a poor measure of variation, it does have some useful applications in the following areas:

1. ***Industry.*** It is used in industry where the range for measurements on items coming off an assembly line might be specified in advance. As long as all measurements fall within the specified range, the process is said to be in control.
2. ***Weather Forecasts.*** It is used by the meteorological department for weather forecasts. The range in determining the difference between the minimum temperature and the maximum temperature is of great importance to the general public because they come to know the limits within which the temperature is likely to vary on a particular day.
3. ***Stock Exchange.*** It is useful in studying the fluctuations in the share prices.
4. ***Day-to-day Living.*** The range is a most widely used measure of dispersion in

our day-today life. For example, questions such as "What are the monthly wages of workers in the factory?" "What is the daily sales in a departmental store?" "How much do you travel in a month by your own car" are all answered in the form of range.

9.7 QUARTILE DEVIATION OR SEMI INTER-QUARTILE RANGE

Definition. **Inter-quartile range** is an absolute measure of dispersion defined by the formula:

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

where Q_1 and Q_3 are the first (or lower) and the third (or upper) quartiles respectively.

As inter-quartile range is based only on Q_1 and Q_3, it measures the length of the interval that contains 50% of the data.

Definition. **Quartile deviation,** also called **semi inter-quartile range,** is an absolute measure of dispersion defined by the formula:

$$\text{Quartile Deviation } (Q.D.) = \frac{Q_3 - Q_1}{2}$$

Quartile deviation gives the average amount by which the two quartiles differ from the median. In a symmetrical distribution, the two quartiles Q_1 and Q_3 are equidistant from the median, *i.e.*, median $- Q_1 = Q_3 -$ median and as such median $\pm$ $Q.D.$ covers exactly 50% of the data.

Quartile deviation, being an absolute measure of dispersion, cannot be used to compare two distributions expressed in different units. For comparative study of variability of two distributions, a relative measure known as *coefficient of quartile deviation* is defined.

Definition. The ***coefficient of quartile deviation*** is a relative measure of dispersion defined by

$$\text{Coefficient of Quartile Deviation} = \frac{\frac{Q_3 - Q_1}{2}}{\frac{Q_3 + Q_1}{2}} = \frac{Q_3 - Q_1}{Q_3 + Q_1}$$

Coefficient of quartile deviation is a pure number independent of the units of measurement and can be used to compare two distributions expressed in different units.

Example 3. Calculate quartile deviation from the following distribution :

X	:	5 – 7	8 – 10	11 – 13	14 – 16	17 – 19
Frequency	:	14	24	38	20	4

[*Delhi Univ. B.Com. (P) 2001*]

Solution.

Calculations For Quartiles

Class Boundary	Frequency (*f*)	Less than c.f.
4.5 – 7.5	14	14
7.5 – 10.5	24	38
10.5 – 13.5	38	76
13.5 – 16.5	20	96
16.5 – 19.5	4	100
	$N = \sum f = 100$	

Computation of Q_1 **:** We have $\frac{N}{4} = \frac{100}{4} = 25$; and the c.f. just greater than or equal to 25 is 38. The class corresponding to this c.f. is 7.5 – 10.5. Thus Q_1– class is 7.5 – 10.5.

$$\therefore \quad Q_1 = l + \frac{\frac{N}{4} - C}{f} \times h = 7.5 + \frac{25 - 14}{24} \times 3 = 7.5 + 1.375 = 8.875$$

Computation of Q_3 **:** $\frac{3N}{4} = 75$; and the c.f. just greater than or equal to 75 is 76. The class corresponding to this c.f. is 10.5 – 13.5. Thus Q_3 – class is 10.5 – 13.5.

$$\therefore \quad Q_3 = l + \frac{\frac{3N}{4} - C}{f} \times h = 10.5 + \frac{75 - 38}{38} \times 3 = 10.5 + 2.92 = 13.42$$

Computation of Quartile Deviation:

$$Q.D. = \frac{Q_3 - Q_1}{2} = \frac{13.42 - 8.875}{2} = 2.27 \text{ (app.)}$$

Example 4. Find the value of third quartile if the values of first quartile and quartile deviation are 90 and 20 respectively. ***[Delhi Univ. B.Com. (P) 2013]***

Solution. Quartile deviation is given by

$$Q.D. = \frac{Q_3 - Q_1}{2}$$

Substituting $Q_1 = 90$ and $Q.D. = 20$ in the above formula, we obtain $20 = \frac{Q_3 - 90}{2}$

$$\Rightarrow \quad Q_3 - 90 = 20 \times 2 = 40 \quad \Rightarrow \quad Q_3 = 90 + 40 = 130.$$

Example 5. If the first quartile is 48 and quartile deviation is 6, find the median (assuming the distribution to be symmetrical). ***[Delhi Univ. B.Com. (P) 1996]***

Solution. Quartile deviation is given by

$$Q.D. = \frac{Q_3 - Q_1}{2}$$

Substituting $Q_1 = 48$ and $Q.D. = 6$ in the above formula, we obtain $6 = \frac{Q_3 - 48}{2}$

$$\Rightarrow \quad Q_3 - 48 = 12 \quad \Rightarrow \quad Q_3 = 60$$

For a symmetrical distribution, the first and the third quartiles are equidistant from the median. That is,

$$Q_3 - Md = Md - Q_1$$

or,

$$2\ Md = Q_1 + Q_3 \Rightarrow Md = \frac{Q_1 + Q_3}{2}$$

$\therefore$

$$Md = \frac{48+60}{2} = \frac{108}{2} = 54.$$

Example 6. Find the inter-quartile range and the coefficient of quartile deviation from the following data:

Marks less than :	10	20	30	40	50	60	70	80
No. of Students :	4	16	40	76	96	112	120	125

Solution. We shall first convert the given cumulative frequency distribution into an ordinary frequency distribution.

Marks	*f*	*c.f.*
0 – 10	4	4
10 – 20	12	16
20 – 30	24	40
30 – 40	36	76
40 – 50	20	96
50 – 60	16	112
60 – 70	8	120
70 – 80	5	125
	$N = 125$	

Computation of Q_1: $\frac{N}{4} = \frac{125}{4} = 31.25$; the c.f. just greater than or equal to 31.25 is 40. Therefore Q_1 lies in the class 20 – 30 and is given by

$$Q_1 = l + \frac{\frac{N}{4} - C}{f} \times h = 20 + \frac{31.25 - 16}{24} \times 10 = 26.35.$$

Computation of Q_3: $\frac{3N}{4} = 93.75$; the *c.f.* just greater than or equal to 93.75 is 96. Therefore Q_3 lies in the class 40 – 50 and is given by

$$Q_3 = l + \frac{\frac{3N}{4} - C}{f} \times h = 40 + \frac{93.75 - 76}{20} \times 10 = 48.88$$

$\therefore$ Inter-quartile range $= Q_3 - Q_1 = 48.88 - 26.35 = 22.53$

and Coefficient of $Q.D. = \frac{Q_3 - Q_1}{Q_3 + Q_1} = \frac{48.88 - 26.35}{48.88 + 26.35} = \frac{22.53}{75.23} = 0.30$ (app.)

9.8 MERITS AND DEMERITS OF QUARTILE DEVIATION

Merits. 1. *It is simple to understand and easy to calculate.*

2. *It is superior to range as a measure of dispersion.*

3. *It can be computed from the distribution having open-end classes. In fact, it is the only measure of dispersion which can be obtained while dealing with a distribution having open-end classes.*

4. *Quartile deviation is not affected at all by extreme observations as it ignores 25% of the data from the beginning of the distribution and another 25% of the data from the top end.*

5. *Quartile deviation is useful specially when it is desired to study variability in the central half part of the data.*

Demerits. 1. *Quartile deviation is not based on all the observations. In fact, it ignores 25% of the data at the lower end and 25% of the data at the upper end. Hence it cannot be considered as a good measure of dispersion.*

2. *Quartile deviation is not suitable for further mathematical treatment.*

3. *It is affected considerably by sampling fluctuations.*

EXERCISE 9.1

Theory Questions

1. (*a*) What is meant by dispersion? **[*Delhi Univ. B.Com. (P)* 1995, 2004]**

   (*b*) List the various methods of measuring dispersion. **[*Delhi Univ. B.Com. (P)* 1993, 2005, 2006]**

2. What do you mean by relative measure of dispersion? **[*Delhi Univ. B.Com. (P)* 2003]**

3. Distinguish clearly between absolute and relative measures of dispersion. **[*Delhi Univ. B.Com. (P)* 2008, 2012]**

4. Explain briefly (*i*) Range, and (*ii*) Semi-interquartile range. **[*Delhi Univ. B.Com. (P)* 1980]**

5. What is coefficient of quartile deviation? State the formula. **[*Delhi Univ. B.Com. (P)* 1994]**

9.9 MEAN DEVIATION OR AVERAGE DEVIATION

As discussed earlier, the two measures of dispersion, *viz.*, range and quartile deviation are not based on all the observations. Moreover, they do not show any scatterness around an average and thus completely ignore the composition of the series. However, if we wish to measure variation in the sense of showing the scatterness around an average, we must include the deviations of each and every item from an average. *Mean deviation* or the *average deviation* helps us in achieving this goal. As the name suggests, this measure of dispersion is obtained on taking the average (arithmetic mean) of the deviations of the given values from a measure of central tendency. According to **Clark** and **Schkade:**

"Average deviation is the average amount of scatter of the items in a distribution from either the mean or the median, ignoring the signs of the deviations. The average that is

taken of the scatter is an arithmetic mean which accounts for the fact that this measure is often called the mean deviation".

Computation of Mean Deviation – Individual Observations

For a given set of n observations $X_1, X_2, \ldots, X_n$, the mean deviation *(M.D.)* about an average, say A, is given by

$$\text{Mean Deviation (about an average } A) = \frac{\sum |X - A|}{n} = \frac{\sum |D|}{n},$$

where $|D| = |X - A|$ (read as mod $(X - A)$ is the modulus value (or absolute value) of the deviation of X from A, ignoring $\pm$ signs.

Remark 1. Usually, we obtain mean deviation about any one of the three averages Mean (M), Median (Md) or Mode (Mo). But since mode is generally ill-defined, in practice, mean deviation is computed about mean or median. However, there is an advantage of computing mean deviation about median because *the sum of the deviations of items from median is least when signs are ignored.* Nevertheless, in practice, the mean is more frequently used in computing the average deviation and this is the reason why it is more commonly referred to as mean deviation.

Procedure for Computing the Mean Deviation

We now outline the procedure for computing the mean deviation:

Step 1. *Calculate the average A about which mean deviation is to be computed, by the methods discussed earlier.*

Step 2. *Find the deviation of each observation X from A and denote it by D. That is, find $D = X - A$.*

Step 3. *Find the absolute value of the deviation of each observation from A ignoring $\pm$ signs and denote it by $|D|$.*

Step 4. *Find the sum of all absolute deviations obtained in Step 3 to get $\sum |D|$.*

Step 5. *Divide the sum obtained in Step 4 by the number of observations to get the required mean deviation about the average A.*

Computation of Mean Deviation – Discrete Series

In case of discrete series where the variable X takes the values $X_1, X_2, \ldots, X_n$ with respective frequencies $f_1, f_2, \ldots, f_n$, the mean deviation about an average A is given by

$$\text{Mean Deviation (about an average } A) = \frac{\sum f |X - A|}{N} = \frac{\sum f |D|}{N},$$

where $N = \sum f$ is the total frequency and $D = X - A$

Procedure for Computing the Mean Deviation

Step 1. *Calculate the average A about which mean deviation is to be computed.*

Step 2. *Take the deviation of each observation from A and denote it by D. That is, find $D = X - A$.*

Step 3. *Find the absolute value of the deviation of each observation from A ignoring $\pm$ signs and denote it by $|D|$.*

Step 4. *Multiply each absolute deviation $|D|$ by the corresponding frequency f to get $f|D|$.*

Step 5. *Add all the products obtained in Step 4 to get $\sum f|D|$.*

Step 6. *Divide the sum obtained in Step 5 by N, the total frequency, to get the required mean deviation.*

Computation of Mean Deviation – Continuous Series

The computation of the mean deviation in the case of continuous series is exactly the same as discussed above for discrete series. The only difference is that here we have to obtain the class marks (or mid-values) of the various classes and take absolute deviations of these values from the average A. Thus, if $X_1, X_2, \ldots, X_n$ are the class marks (or mid-values) of a set of grouped data with corresponding class frequencies $f_1, f_2, \ldots, f_n$, then the mean deviation about an average A is given by

$$\text{Mean Deviation (about an average } A) = \frac{\sum f|X-A|}{N} = \frac{\sum f|D|}{N},$$

where $N = \sum f$ is the total frequency

Coefficient of Mean Deviation

The measures of mean deviation as defined above are absolute measures depending on the units of measurement. The relative measure corresponding to the mean deviation, called the **coefficient of mean deviation,** is given by

$$\text{Coefficient of } M.D. = \frac{\text{Mean Deviation}}{\text{Average about which it is calculated}}$$

$$\therefore \quad \text{Coefficient of } M.D. \text{ about mean} = \frac{M.D}{\text{Mean}}$$

and $$\text{Coefficient of } M.D. \text{ about median} = \frac{M.D}{\text{Median}}$$

Coefficient of mean deviation is a pure number independent of the units of measurement and can be used to compare two distributions expressed in different units.

Example 7. Find the mean deviation about the median for the following data:

56 46 79 26 85 39 65 99 29 72

Find also the coefficient of mean deviation.

Solution. Arranging the data in ascending order of magnitude, we get

26 29 39 46 56 65 72 79 85 99

Here,

$$n = 10 \qquad \therefore \quad \frac{n+1}{2} = \frac{10+1}{2} = 5.5$$

$\therefore$

$$\begin{aligned}\text{Median} &= \text{size of } \left(\frac{n+1}{2}\right)\text{th item} = \text{size of (5.5)th item}\\ &= \text{5th item} + 0.5\,(\text{6th item} - \text{5th item})\\ &= 56 + 0.5\,(65-56) = 56 + 4.5 = 60.5.\end{aligned}$$

Calculation for Mean Deviation

X	*Deviation from Median* $D = X - Md$	*Absolute Deviation from Median* $\|D\|$
26	–34.5	34.5
29	–31.5	31.5
39	–21.5	21.5
46	–14.5	14.5
56	–4.5	4.5
65	4.5	4.5
72	11.5	11.5
79	18.5	18.5
85	24.5	24.5
99	38.5	38.5
$n = 10$		$\sum\|D\| = 204$

$$\therefore \quad M.D. \text{ (about median)} = \frac{\sum|D|}{n} = \frac{204}{10} = 20.4.$$

$$\text{Coefficient of } M.D. = \frac{\text{Mean Deviation}}{\text{Median}} = \frac{20.4}{60.5} = 0.34.$$

Example 8. Calculate mean deviation about the mean for the following data:

X	:	10	11	12	13	14	Total
f	:	3	12	18	12	3	48

Solution. **Calculations for Mean Deviation**

X	f	fX	$D = X - \bar{X}$	$\|D\|$	$f\|D\|$
10	3	30	– 2	2	6
11	12	132	– 1	1	12
12	18	216	0	0	0
13	12	156	1	1	12
14	3	42	2	2	6
	$N = \sum f = 48$	$\sum fX = 576$			$\sum f\|D\| = 36$

$$\text{Mean} : \bar{X} = \frac{\sum fX}{\sum f} = \frac{576}{48} = 12$$

$$\text{Mean Deviation about Mean} = \frac{\sum f|D|}{N} = \frac{36}{48} = 0.75.$$

Example 9. Calculate the mean deviation from the median for the following data:

Marks :	10 – 20	20 – 30	30 – 40	40 – 50	50 – 60	60 – 70	70 – 80	80 – 90
No. of Students :	2	6	12	18	25	20	10	7

Also calculate the coefficient of mean deviation from median.

[*Delhi Univ. B.Com. (P) 1982*]

Solution. **Calculation for Mean Deviation**

Marks	Mid-value X	No. of Students f	Less than cf	D	|D|	f|D|
10 – 20	15	2	2	–39.8	39.8	79.6
20 – 30	25	6	8	–29.8	29.8	178.8
30 – 40	35	12	20	–19.8	19.8	237.6
40 – 50	45	18	38	–9.8	9.8	176.4
50 – 60	55	25	63	0.2	0.2	5.0
60 – 70	65	20	83	10.2	10.2	204.0
70 – 80	75	10	93	20.2	20.2	202.0
80 – 90	85	7	100	30.2	30.2	211.4
		$N = \sum f = 100$				$\sum f\|D\| = 1294.8$

Computation of Median : We have $\frac{N}{2} = 50$. The c.f. just greater than or equal to 50 is 63 and the corresponding class interval is 50-60. Thus the median class is 50 - 60.

$$\therefore \text{ Median } (Md) = l + \frac{\frac{N}{2} - C}{f} \times h = 50 + \frac{50 - 38}{25} \times 10 = 50 + \frac{120}{25} = 50 + 4.8 = 54.8.$$

Computation of Mean Deviation :

$$M.D. = \frac{\sum f|D|}{N} = \frac{1294.8}{100} = 12.948.$$

$\therefore$ Coefficient of *M.D.* from median

$$= \frac{M.D. \text{ (about median)}}{\text{Median}} = \frac{12.948}{54.8} = 0.24 \text{ (app.).}$$

Example 10. Calculate the mean deviation from the mean for the following data:

Marks :	0 – 10	10 – 20	20 – 30	30 – 40	40 – 50	50 – 60	60 – 70
No. of Students :	6	5	8	15	7	6	3

Solution. **Calculation of Mean Deviation From Mean**

Marks	Mid-value X	No. of Students f	fX	$D = X - \bar{X}$	| D |	f | D |
0 – 10	5	6	30	–28.4	28.4	170.4
10 – 20	15	5	75	–18.4	18.4	92.0
20 – 30	25	8	200	–8.4	8.4	67.2
30 – 40	35	15	525	1.6	1.6	24.0
40 – 50	45	7	315	11.6	11.6	81.2
50 – 60	55	6	330	21.6	21.6	129.6
60 – 70	65	3	195	31.6	31.6	94.8
		N = 50	$\sum fX = 1670$			$\sum f\|D\| = 659.2$

Computation of Mean: $\bar{X} = \frac{\sum fX}{N} = \frac{1670}{50} = 33.4$

Computation of Mean Deviation about Mean:

$$M.D. = \frac{\sum f|D|}{N} = \frac{659.2}{50} = 13.18 \text{ (app.).}$$

Example 11. Calculate the mean deviation and its coefficient:

Marks	: 21 – 25	26 – 30	31 – 35	36 – 40	41 – 45	46 – 50	51 – 55
No. of Students	: 5	15	28	42	15	12	3

[Delhi Univ. B.Com. (P) 2004]

Solution. The mean deviation gives the best results when deviations are taken from median. Since nothing is specified in the question, we shall take deviations from median.

Calculation of Mean Deviation

Marks	*Mid-Value* X	*No. of Students* f	*c.f.*	$D = X - Md$	$\vert D \vert$	$f\vert D\vert$
20.5 – 25.5	23	5	5	–13.93	13.93	69.65
25.5 – 30.5	28	15	20	– 8.93	8.93	133.95
30.5 – 35.5	33	28	48	–3.93	3.93	110.04
35.5 – 40.5	38	42	90	1.07	1.07	44.94
40.5 – 45.5	43	15	105	6.07	6.07	91.05
45.5 – 50.5	48	12	117	11.07	11.07	132.84
50.5 – 55.5	53	3	120	16.07	16.07	48.21
		$N = \sum f = 120$				$\sum f\vert D\vert = 630.08$

Computation of Median : $\frac{N}{2} = \frac{120}{2} = 60$; the *c.f.* just greater than or equal to 60 is 90.

Thus median lies in the class 35.5 - 40.5 and is given by

$$\text{Median } (Md) = l + \frac{\frac{N}{2} - C}{f} \times h = 35.5 + \frac{60 - 48}{42} \times 5 = 35.5 + 1.43 = 36.93$$

$$\therefore \quad M.D. \text{ (about Median)} = \frac{\sum f|D|}{N} = \frac{630.08}{120} = 5.26$$

$$\text{Coefficient of Mean Deviation (about median)} = \frac{M.D.}{\text{Median}} = 0.142.$$

9.10 MERITS AND DEMERITS OF MEAN DEVIATION

Merits. 1. *It is easy to understand and simple to calculate.*

2. *It is based on each and every item of the data.*

3. *It is rigidly defined.*

4. *As compared to standard deviation, it is less affected by extreme observations.*
5. *Since deviations are taken from a central value, comparison about formation of different distributions can easily be made.*

Demerits. 1. *The major drawback of mean deviation is that algebraic signs are ignored while taking the deviations of the items.*

2. *It is not suitable for further mathematical treatment.*
3. *It cannot be computed for distribution with open-end classes.*
4. *It is rarely used in sociological studies.*

EXERCISE 9.2

Theory Questions

1. What do you understand by mean deviation?
2. What is coefficient of mean deviation? State the formula.