# UNIT 4 : MEASURES OF DISPERSION

## UNIT STRUCTURE


4.1 Learning Objectives

4.2 Introduction

4.3 Concept of Dispersion

- 4.3.1 Range
- 4.3.2 Quartile Deviation
- 4.3.3 Mean Deviation
- 4.3.4 Standard Deviation
- 4.3.5 Coefficient of Variation

4.4 Let Us Sum Up

4.5 Further Readings

4.6 Answers To Check Your Progress

4.7 Model Questions


## 4.1 LEARNING OBJECTIVES

After going through this unit, you will able to learn

- concept of dispersion
- range
- quartile deviation
- mean deviation
- standard deviation
- coefficient of variation.

## 4.2 INTRODUCTION

Dispersion means ‘the degree of scatter of data’. Measures of dispersion provide ‘certain specific information concerning the scatter or dispersion of values in a distribution’. These measures viz. range, quartile deviation, mean deviation, standard deviation will be discussed in this unit.

## 4.3 CONCEPT OF DISPERSION

Let us consider the following two sets of observations.

A. 25, 26, 27, 28, 29

B. 3, 7, 13, 50, 60

In each of A, and B, there are 5 observations. The mean in each case is 27. But the observations in A are more concentrated than observations in B. Though the mean is the same, the two sets are differently constituted. The greatest variation from the mean in A is 4; whereas in B it is 35. Thus we see that the measures of central tendency do not give a clear idea of a distribution. These measues must be supplemented by other measures. Dispersion is one of such measures.

### 4.3.1 Range

Range is the difference between the maximum and the minimum values of the observations. Thus,

Range= M–m, where M is the maximum value and m is the minimum value.

**Example 1:** Find the range of the following observations.

23, 16, 14, 29, 5.6, 27, 13, 59

**Solution :** We arrange the data in ascending order.

13, 14, 16, 23, 27, 29, 56, 59

Here maximum value = 59

Minimum value = 13

$\therefore$ Range = 59 – 13 = 46

**Note :** Though range is can be calculated easily, it is not a reliable measure. It is not based on the entire set of data. It depends upon the two extremes. In the above example, the two extremes are 13, and 59. Range is based upon these two.

**Coefficient of range :** Coefficient of range is defined as follows.

Coefficient of range $\frac{\quad}{\quad}$, where M is the maximum

value, and m is the minimum value.

**Example 2:** Find the coefficient of range of the observations given in Example 1.

**Solution :** Here M = 59, M = 13

$\therefore$ Coefficient of range =

## 4.3.2 Quartile Deviation

The difference between the third quartile and the first quartile is called interquartile range, and half the interquartile range is called quartile deviation (QD). Thus quartile deviation Q is given by

,

Where $Q_3$ is the third quartile, and $Q_1$ is the first quartile.

Coefficient of quartile deviation is defined as follows.

Coefficient of quartile deviation $= \dfrac{Q_3 - Q_1}{Q_3 + Q_1}$

$QD = Q = \dfrac{59-13}{59+13} = \dfrac{46}{72} = \dfrac{23}{36}$

**Example 3:** Find the quartile deviation for the following data

| Class-interval (Marks) | 0-10 | 10-20 | 20-30 | 30-40 | 40-50 | 50-60 |
|---|---|---|---|---|---|---|
| No. of students | 4 | 6 | 20 | 10 | 7 | 3 |

**Solution :** We know that $Q_1 = L + \dfrac{\frac{N}{4} - c}{f} \times h$,

Where L is the lower limit of the first quartile class,

N is the total number of frequencies,

C is the cumulative frequency of the class just above the quartile class,

f is the frequency of the quartile class,

h is the class interval.

Now,

| Class interval | Number of students (f) | c.f. |
|---|---|---|
| 0-10 | 4 | 4 |
| 10-20 | 6 | 10 |
| 20-30 | 20 | 30 |
| 40-50 | 7 | 47 |
| 50-60 | 3 | 50 |

$\therefore$ N = 50,

12.5 is contained the minimum c.f. 30.

$\therefore$ 1st quartile class : 20 – 30

L = 20, c = 10, f = 20, h = 10

$$= 20 + \frac{2.5}{2} = 20 + 1.25 = 21.25$$

Again, $Q_3 = L + \frac{\frac{3N}{4} - c}{f} \times h$

$$\frac{3N}{4} = \frac{3 \times 50}{4} = 37.5$$

$\therefore$ 3rd quartile class : 30 – 40

L = 30, c = 30, f = 10 h = 10

= 30 + 7.5 = 37.5

$\therefore$ Quartile deviation =

$$= \frac{16.25}{2}$$

= 8.125

**Note :** Quartile deviation is not affected by the extreme values.



## CHECK YOUR PROGRESS 1

**Q 1:** Write down one advantage and disadvantage of range.

**Q 2:** Find the range and coefficient of range of the following observations

43, 13, 26, 11, 39, 56, 7, 98, 96, 34, 65, 42 .

**Q 3:** Is quartile deviation affected by the extreme value ?

**Q 4:** What is two range of the following observations ?

6, 6, 6, 6, 6, 6, 6, 6, 6.

**Q 5:** Find the quartile deviation for the following data.

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

### 4.3.3 Mean Deviation

Mean deviation is the mean of the numerical deviations from the mean, median or mode.

**Mean deviation for ungrouped data**

Let $x_1, x_2, x_3, \ldots\ldots x_n$ be n observations

Let $\bar{x}$ be the mean, and M be the median of the distribution.

(A) Mean deviation about the mean, denoted by MD ($\bar{x}$), is given by

$$MD(\bar{x}) = \frac{\sum_{i=1}^{n} |x_i - \bar{x}|}{n}$$

(B) Mean deviation about the mode, denoted by MD(M), is given by

$$MD(M) = \frac{\sum_{i=1}^{n} |x_i - M|}{n}$$

**Example 4:** Find the mean deviation (a) about the mean, (b) about the median for the following observations

4, 5, 7, 8, 9, 13, 16, 18

**Solution :** (a) $\bar{x} = \frac{4+5+7+8+9+13+16+18}{8} = \frac{80}{8} = 10$

$\therefore$ the values of are –6, –5, –3, –2, –1, 3, 6, 8.

$\therefore$ the values of are 6, 5, 3, 2, 1, 3, 6, 8.

$$\therefore MD(\bar{x}) = \frac{\sum_{i=1}^{8} |x_i - \bar{x}|}{8}$$

(b) The observations are in ascending order :

4, 5, 7, 8, 9, 13, 16, 18.

There are 8 observations

Median $M = \frac{1}{2}\left(\frac{8}{2}\text{th observation} + \left(\frac{8}{2}+1\right)\text{th observation}\right)$

$= \frac{1}{2}$(4th observation + 5th observation)

$$= \frac{1}{2} \times (8+9) = \frac{17}{2} = 8.5$$

$\therefore$ the values of ($x_i$–M) are

–4.5, –3.5, –1.5, –0.5, 0.5, 4.5, 7.5, 9.5

the values of ($x_i$–M) are

4.5, 3.5, 1.5, 0.5, 0.5, 4.5, 7.5, 9.5

$$= \frac{4.5+3.5+1.5+0.5+0.5+4.5+7.5+9.5}{8}$$

$$= \frac{32}{2} = 4$$

**Mean deviation for discrete frequency distribution :**

Let $x_1, x_2, x_3, \ldots\ldots, x_n$ be n observations occurring with frequencies $f_1, f_2, f_3, \ldots\ldots, f_n$ respectively.

(A) Mean deviation about the mean, denoted by $MD(\bar{x})$, is given by

$$MD(\bar{x}) = \frac{\sum_{i=1}^{n} f_i |x_i - \bar{x}|}{\sum_{i=1}^{n} f_i}, \quad \text{Where}$$

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

(B) Mean deviation about the mode, denoted by MD (M) is given by

$$MD(M) = \frac{\sum_{i=1}^{8} f_i \, | x_i - M |}{N}$$

**Example 5:** Find the mean deviation (a) about the mean (b) about the median for the following distribution

| $x_i$ | 6 | 12 | 18 | 24 | 30 | 86 |
|---|---|---|---|---|---|---|
| $f_i$ | 5 | 4 | 11 | 6 | 4 | 6 |

**Solution :**

(a)

| $x_i$ | $f_i$ | $f_i x_i$ | $x_i - \bar{x}$ | $\lvert x_i - \bar{x} \rvert$ | $f_i \lvert x_i - \bar{x} \rvert$ |
|---|---|---|---|---|---|
| 6 | 5 | 30 | –15 | 15 | 75 |
| 12 | 4 | 48 | –9 | 9 | 36 |
| 18 | 11 | 198 | –3 | 3 | 33 |
| 24 | 6 | 144 | 3 | 3 | 18 |
| 30 | 4 | 120 | - | 9 | 36 |
| 36 | 6 | 216 | 15 | 15 | 90 |

$$N = \sum_{i=1}^{6} f_i = 36, \quad \sum_{i=1}^{6} f_i x_i = 756$$

$$\therefore \bar{x} = \frac{\Sigma f_i x_i}{N} = \frac{756}{36} = 21$$

$$\therefore MD(\bar{x}) = \frac{\Sigma f_i \, | x_i - \bar{x} |}{N} = \frac{288}{36} = 8$$

(B)

| $x_i$ | $f_i$ | cf | $x_i - M$ | $\lvert x_i - M \rvert$ | $f_i [x_i - M]$ |
|---|---|---|---|---|---|
| 6 | 5 | 5 | –12 | 12 | 60 |
| 12 | 4 | 9 | –6 | 6 | 24 |
| 18 | 11 | 20 | 0 | 0 | 0 |
| 24 | 6 | 26 | 6 | 6 | 36 |
| 30 | 4 | 30 | 12 | 12 | 48 |
| 36 | 6 | 35 | 18 | 18 | 104 |

N = 36, an even number

$\therefore$ Median M = [ th observation + $\left(\frac{36}{2}+1\right)$th observation ]

= th observation +19th observation

$$=\frac{1}{2}(18+18)=18$$

$$\therefore\ MD(M)=\frac{\sum_{i=1}^{6} f_i\,|x_i-M|}{N}=\frac{272}{36}=7.5$$

**Mean deviation for continuous frequency distribution :**

(A) Mean deviation about the mean, denoted by MD (), is given by

$MD(\bar{x})=\frac{\Sigma f_i\,|x_i-\bar{x}|}{N}$, where $N=\Sigma f_i$

(B) Mean deviation about the median, denoted by MD (M) is given by

$$MD(M)=\frac{\Sigma f_i\,|x_i-M|}{N}$$

**Example 6:** Find the mean deviation about the mean for the following distribution.

| Class-interval (Marks) | 0-10 | 10-20 | 20-30 | 30-40 | 40-50 | 50-60 |
|---|---|---|---|---|---|---|
| Frequency (Number of students) | 4 | 8 | 14 | 16 | 4 | 2 |

**Solution :**

| Class-interval | Frequency ($f_i$) | Mid point ($x_i$) | $f_ix_i$ | $x_i-\bar{x}$ | $\lvert x_i-\bar{x}\rvert$ | $f_i\lvert x_i-\bar{x}\rvert$ |
|---|---|---|---|---|---|---|
| 0–10 | 6 | 5 | 30 | –22 | 22 | 132 |
| 10–20 | 8 | 15 | 120 | –12 | 12 | 96 |
| 20–30 | 14 | 25 | 350 | –2 | 2 | 28 |
| 30–40 | 16 | 35 | 560 | 8 | 8 | 128 |
| 40–50 | 4 | 45 | 180 | 18 | 18 | 72 |
| 50–60 | 2 | 55 | 110 | 28 | 28 | 56 |

$N=\Sigma f_i=50,\quad \Sigma f_ix_i=1350$

$$\therefore\ x = \frac{\Sigma f_i x_i}{N} = \frac{1350}{50} = 27$$

$$MD(\bar{x}) = \frac{\Sigma f_i \mid x_i - \bar{x} \mid}{N}$$

$$= \frac{512}{50} = 10.24$$

**Example 7:** Find the mean deviation about the median for the following distribution

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

**Solution :**

| Class-interval | frequency ($f_i$) | Midpoint ($x_i$) | c.f. | $\lvert x_i - M \rvert$ | $\lvert f_i - M \rvert$ |
|---|---|---|---|---|---|
| 0–10 | 5 | 5 | 5 | 20 | 100 |
| 10–20 | 10 | 15 | 15 | 10 | 100 |
| 20–30 | 20 | 25 | 35 | 0 | 0 |
| 30–40 | 5 | 35 | 40 | 10 | 50 |
| 40–50 | 10 | 45 | 50 | 20 | 200 |

$$M = L + \frac{\frac{N}{2} - c}{f} \times h$$

Here $N = \Sigma f_i = 50$ $\quad \therefore\ \frac{N}{2} = \frac{50}{2} = 25$

The cumulative frequency just greater than 25 is 35, and the corresponding class-interval is 20–30.

$\therefore$ Median class is 20–30

Now, median

Here L = 20, $\frac{N}{2} = 25$ c = 15, f = 20, h = 10

$$\therefore M = 20 + \frac{25-15}{20} \times 10 \quad = 20 + \frac{10}{20} \times 10 = 20 + 5 = 25$$

$$\therefore MD(M) = \frac{\Sigma f_i \mid x_i - M \mid}{N} = \frac{450}{50} = 9$$



## CHECK YOUR PROGRESS 2

**Q 1:** Find the mean deviation about the mean for the following observations.

1, 2, 3, 4, 5, 6, 7

**Q 2:** Find the mean deviation about the median for the following observations

2, 3, 4, 7, 8, 9, 10, 11, 14

**Q 3:** Find the mean deviation (a) about the mean, (b) about the median for the following distribution.

| $x_i$ | 5 | 7 | 9 | 10 | 12 | 15 |
|---|---|---|---|---|---|---|
| $f_i$ | 8 | 6 | 2 | 2 | 2 | 6 |

**Q 4:** The mean of 2, 4, 6, k, 8 and 15 is 7. Find the mean deviation (a) about the mean, (b) about the median.

**Q 5:** Find the mean deviation about the mean for the following distribution.

| Class-interval | 0–10 | 10–20 | 20–30 | 30–40 | 40–50 |
|---|---|---|---|---|---|
| frequency | 5 | 10 | 20 | 5 | 10 |

**Q 6:** Find the mean deviation about the median for the following distribution.

| Class-interval | 0–10 | 10–20 | 20–30 | 30–40 | 40–50 |
|---|---|---|---|---|---|
| frequency | 5 | 8 | 15 | 16 | 6 |

### 4.3.4 Standard Deviation

**(A) Standard deviation for ungrouped data :**

Let $x_1, x_2, x_3, \ldots, x_n$ be n observations. Let $\bar{x}$ be the mean. The standard deviation (SD), denoted by $\sigma$ (Greek letter small 'sigma'), is defined by

The square of the standard deviation is called variance. Thus,

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

Formula for a can be simplified as follows

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

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

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

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

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

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

Thus, $SD = \sigma = \sqrt{\frac{\Sigma\bar{x}_i^{\,2}}{n} - \bar{x}^2}$

$$\text{Variance} = \sigma^2 = \frac{\Sigma\bar{x}_i^{\,2}}{n} - \bar{x}^2$$

**Example 8:** Find the SD of the first five natural numbers.

**Solution :** The first five natural numbers are 1, 2, 3, 4, 5.

$$\bar{x} = \frac{1+2+3+4+5}{5} = \frac{15}{5} = 3$$

$$\therefore\ \bar{x}^2 = 3^2 = 9$$

$$\frac{\Sigma x_i^{\,2}}{n} = \frac{1^2+2^2+3^2+4^2+5^2}{5} = \frac{55}{5} = 11$$

$$\therefore\ \frac{\Sigma x_i^{\,2}}{n} - \bar{x}^2 = 11 - 9 = 2$$

$$\therefore SD = \sigma = \sqrt{2}$$

**Example 9:** The standard deviation of the observations a, 3, 6, 11 is 3.5. Find a.

**Solution :** We have $\bar{x} = \frac{a+3+6+11}{4} = \frac{a+20}{4}$

$$\frac{\Sigma x_i^{\,2}}{n} = \frac{a^2 + 3^2 + 6^2 + 11^2}{4} = \frac{a^2 + 9 + 36 + 121}{4} = \frac{a^2 + 166}{4}$$

Now $\sigma^2 = \frac{\Sigma x_i^{\,2}}{n} - \bar{x}^2$

or, $(3.5)^2 = \frac{a^2 + 166}{4} - \frac{\left(a^2 + 20\right)^2}{4}$

or, $\frac{49}{4} = \frac{4(a^2 + 166) - (a^2 + 40a + 400)}{16}$

or, $49 = \frac{3a^2 - 40a + 264}{4}$

or, $3a^2 - 40a + 264 = 196$

or, $3a^2 - 40a + 68 = 0$

or, $3a^2 - 6a - 34a + 68 = 0$

or, $3a(a-2) - 34(a - 2) = 0$

or, $(a - 2)(3a - 34) = 0$

$\therefore$ $a - 2 = 0$ or $3a - 34 = 0$

$a = 2$ or

**Two properties**

**Property :** If each of the observations $x_1, x_2, x_3, \ldots, x_n$ is increased by a non-zero constant k, the standard deviation as well as the variance remains the same.

Proof: Let the observations be $x_1, x_2, \ldots, x_n$.

$$\therefore\ \bar{x} = \frac{x_1 + x_2 + \ldots + x_n}{n}$$

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

Now the new observations are

$y_1 = x_1 + k, y_2 = x_2 + k, \ldots, y_n = x_n + k$

$\therefore$ new mean

$$= \frac{(x_1 + n_2 + \ldots + x_n) + (k + k + \ldots + k)}{n}$$

$$= \frac{x_1 + n_2 + \ldots + x_n}{n} + \frac{nk}{n}$$

$$= \bar{x} + k$$

Now, new

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

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

= old SD.

Thus, the SD remains the same.

**Property :** If each of the observations $x_1, x_2, \ldots, x_n$ is multiplied by a non-zero constant k, the standard deviation is also multiplied by the same constant k.

**Proof:** Let the observations be $x_1, x_2, \ldots, x_n$.

$$\therefore \bar{x} = \frac{x_1 + x_2 + \ldots + x_n}{n}$$

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

Now, the new observations are

$y_1 = kx_1, y_2 = kx_2, \ldots, y_n = kx_n$

$\therefore$ new mean

$$= \frac{k(x_1 + x_2 + \ldots + x_n)}{n}$$

$$= k\,\bar{x}$$

Now, new

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

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

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

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

= k (old SD)

**Note :** If each of the observations $x_1, x_2, \ldots, x_n$ is multiplied by a non-zero constant, the variance is multiplied by $k^2$.

**(B) Standard deviation for discrete frequency distribution**

Let $x_1, x_2, \ldots, x_n$ be n observations occurring with frequencies $f_1, f_2, \ldots, f_n$ respectively.

The SD is defined by

$$SD = \sigma = \frac{1}{n} \sum_{i=1}^{n} f_i (x_i - \bar{x})^2, \qquad \text{where } N = \Sigma f$$

**Example 10:** FInd the standard deviation for the following distribution

| $x_i$ | 5 | 15 | 25 | 35 | 45 |
|---|---|---|---|---|---|
| $f_i$ | 5 | 4 | 6 | 3 | 2 |

**Solution :**

| $x_i$ | $f_i$ | $f_i x_i$ | | $(x_i - \bar{x})^2$ | $f_i (x_i - \bar{x})^2$ |
|---|---|---|---|---|---|
| 5 | 5 | 25 | –16.5 | 272.25 | 136.25 |
| 15 | 4 | 60 | –6.5 | 42.25 | 169 |
| 25 | 6 | 150 | 3.5 | 12.25 | 73.5 |
| 35 | 3 | 105 | 13.5 | 182.25 | 546.75 |
| 45 | 2 | 90 | 23.5 | 552.25 | 1104.5 |

Now, $N = \Sigma f_i = 20, \ \Sigma f_i x_i = 430$

$$\therefore \bar{x} = \frac{430}{20} = 21.5$$

$$\therefore \Sigma f_i (x_i - \bar{x})^2 = 3255$$

$$\therefore SD = \sqrt{\frac{3255}{20}} = \sqrt{162.75}$$

**(C) Statdard deviation for continuous frequency distribution.**

The standard deviation for for a continuous frequency distribution is given by

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

where $N = \sum_{i=1}^{n} f_i$, $x_i$ is the midpoint of the i th class.

**Example 11:** Find the standard deviation for the following distribution.

| Class-interval | 0–10 | 10–20 | 20–30 | 30–40 | 40–50 |
|---|---|---|---|---|---|
| Frequency | 5 | 10 | 20 | 5 | 10 |

**Solution :**

| Class-interval | frequency | Midpoint | $f_i x_i$ | $x_i - \bar{x}$ | $(x_i - \bar{x})^2$ | $f_i (x_i - \bar{x})^2$ |
|---|---|---|---|---|---|---|
| | $f_i$ | $x_i$ | | | | |
| 0–10 | 5 | 5 | 25 | –21 | 441 | 2205 |
| 10–20 | 10 | 15 | 150 | –11 | 121 | 1210 |
| 20–30 | 20 | 25 | 500 | –1 | 1 | 20 |
| 30–40 | 5 | 35 | 175 | 9 | 81 | 405 |
| 40–50 | 10 | 45 | 450 | 19 | 361 | 3610 |

Now, $N = \Sigma f_i = 50, \ \Sigma f_i x_i = 1300$

$$\therefore \bar{x} = \frac{\Sigma f_i x_i}{N} = \frac{1300}{50} = 26$$

Now, $\Sigma f_i (x_i - \bar{x})^2 = 7450$

$$\therefore SD = \sigma = \sqrt{\frac{1}{N}\Sigma f_i (x_i - \bar{x})^2} = \sqrt{\frac{7450}{50}} = \sqrt{149}$$

**Note :** Standard deviation explains the average amount of variation on either side of the mean. It is considered to be the best measure of dispersion. It is widely used.

There is an empirical (based on experience) relation between quartile deviation, mean deviation, and standard deviation. It is given by

6 QD = 5 MD = 4 SD



## CHECK YOUR PROGRESS 3

**Q 1:** The standard deviation of 5 observations is 13.25. If each of the observations is increased by 3, what is the new standard deviation?

**Q 2:** The standard deviation of 7 observations is 15.75. If each of the observations is multiplied by 2, what is the new standard deviation?

**Q 3:** The standard deviation of 5 observations is 27.45. What is the (a) quartile deviation, (b) mean deviation?

**Q 4:** Find the standard deviation of the following observations 13, 16, 17, 22

**Q 5:** Find the SD of the first n natural numbers.

**Q 6:** The mean and variance of the observations 1, 2, a, 6 and b are 4.4, and 8.24 respectively. Find the values of a and b.

**Q 7:** Find the SD for the following distribution

| $x_i$ | 6 | 10 | 14 | 18 | 20 |
|---|---|---|---|---|---|
| $f_i$ | 2 | 4 | 7 | 12 | 5 |

**Q 8:** Find the SD for the following distribution

| Class-interval | 0–10 | 10–20 | 20–30 | 30–40 | 40–50 |
|---|---|---|---|---|---|
| Frequency | 2 | 4 | 7 | 12 | 5 |

### 8.3.5 Coefficient of Variation

Let the mean and the standard deviation of a distribution be $\bar{x}$ and $\sigma$ respectively. The coefficient of variation (CV) is defined by

This measure was introduced by statistician karl pearson.

**Note :** 1. CV is a unitless measure of dispersion.

2. CV is the best measure to compare the variability of two distributions.

3. A distribution with less CV is considered more consistent (less

variable).

4. A distribution with more CV is considered less consistent (more variable).

**Example 12:** The mean and standard deviation of heights of adults (age 30 years), and children (age 10 months) are given below -

| | Adults | Children |
|---|---|---|
| Mean | 160 cm | 60 cm |
| SD | 10 cm | 5 cm |

Show that there is a greater variation in the heights of children than that of adults.

**Solution :**

CV (adults)

CV (children)

Now CV (children) > CV (adults)

$\frac{\sigma}{\bar{x}} \times 100 = \frac{10}{160} \times 100 = 6.25$ $\frac{\sigma}{\bar{x}} \times 100 = \frac{5}{60} \times 100 = 8.33$

there is greater variation in the heights of children than that of adults.



**CHECK YOUR PROGRESS 4**

**Q 1:** The mean blood pressure of 25 boys is 120 mm Hg, SD is 10 mm Hg. The mean height and SD of the same set of boys is 160 cm and 5 cm respectively. Show that blood pressure is more variable than height.



## 4.4 LET US SUM UP

- Measures of dispersion provide 'certain specific information concerning the scatter of dispersion of values in a distribution'.
- Range, Quartile deviation, Mean Deviation, Standard Deviation are measures of dispersion.
- Range is the difference between the maximum and the minimum

values of the observations.

- The difference between the third quartile and the first quartile is called interquartile range, and half the interquartile range is called quartile deviation. Quartile deviation is not affected by the extreme values.
- Mean deviation is the mean of the numerical deviations from the mean, median, or mode.
- The standard deviation (SD) of the observations $x_1, x_2, \ldots, x_n$; denoted by a (Greek letter small 'sigma'); is defined by

- SD of observations $x_1, x_2, \ldots, x_n$ occurring with frequencies $f_1, f_2, \ldots, f_n$ respectively is defined by

$$SD = \sigma = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2}\ , \qquad \text{where } N = \sum_{i=1}^{n} f_i$$

- Square of SD is called variance.
- Coefficient of variation (CV) is defined by $CV = \frac{\sigma}{\bar{x}} \times 100, \quad \bar{x} \neq 0$

  where a is the SD, and the mean of the distribution.
- CV is a unitless measure of dispersion.
- A distribution with less CV is considered more consistent.
- A distribution with more CV is considered less consistent.



## 4.5 FURTHER READINGS


1) Business Mathematics, Qazi Zameeruddin, V K Khanna, and S K Bhambri, Vikas Publishing House Pvt Ltd, Second Reprint, 1995.
2) Agarwal, D. R. (2006). Business Statistics. Delhi, Vrinda Publications.
3) Gupta S.C. (1994). Fundamentals of statistics. New Delhi, Himalayan Publishing House.
4) Rajagopalan, S. P. & Sattanthan R. (2009).Business Statistics and Operations Research. New Delhi, Tata McGraw-Hill
5) Sharma, J. K. (2007). Business Statistics. New Delhi. Pearson

Education Ltd.

6) Verma, A. P. (2007). Business Statistics. Guwahat, Asian Books Private Limited.



## 4.6 ANSWERS TO CHECK YOUR PROGRESS

**Check Your Progress 1**

**Ans to Q No 1: Advantage :** Range can be calculated easily.

**Disadvantage :** It is not a reliable measure. It is not based on the entire set of data. It depends upon the two extremes.

**Ans to Q No 2:** We arrange the data in ascending order. 7, 11, 13, 26, 34, 39, 42, 56, 65, 96, 98.

Here, maximum value = 98

minimum value = 7

$\therefore$ Range = 98 – 7 = 91

$$Q_1 = L + \frac{\frac{N}{4} - c}{f} \times h \qquad \frac{N}{4} = \frac{50}{4} = 12.5$$

**Ans to Q No 3:** Quartile deviation is not affected by the extreme values.

**Ans to Q No 4:** In this case, all the observations are identical.

Maximum value = Minumum value = 6.

Range = 6 – 6 = 0

**Ans to Q No 5:**

| Class interval | Number of students (f) | c.f. |
|---|---|---|
| 0-10 | 5 | 5 |
| 10-20 | 8 | 13 |
| 20-30 | 15 | 28 |
| 30-40 | 16 | 44 |
| 40-50 | 6 | 50 |

Now, N = 50,

$\therefore$ 1st Quartile class : 10–20

L = 10, c = 5, f = 8, h = 10

Now

$$= 10 + \frac{12.5 - 5}{8} \times 10$$

= 10 + 9.375 = 19.375

Now, $\frac{3N}{4} = \frac{3 \times 50}{4} = 37.5$

$\therefore$ 3rd Quartile class : 30–40

L = 30, c = 28, f = 16, h = 10

$$= 30 + \frac{37.5 - 28}{16} \times 10$$

= 30 + 5.9375 = 35.9375

$\therefore$ Quartile deviation

$$= \frac{35.9375 - 19.375}{2}$$

$$= \frac{16.5625}{2} = 8.28125$$

**Check Your Progress 2**

**Ans to Q No 1:** $\bar{x} = \frac{1+2+3+4+5+6+7}{7} = \frac{28}{7} = 4$

The values of $(x_i - \bar{x})$ are

–3, –2, –1, 0, 1, 2, 3

The values of $|x_i - \bar{x}|$ are

3, 2, 1, 0, 1, 2, 3

$$\therefore MD(\bar{x}) = \frac{\sum_{i=1}^{7} |x_i - \bar{x}|}{7} = \frac{3+2+1+0+1+2+3}{7}$$

$$= \frac{12}{7}$$

**Ans to Q No 2:** The observations are in ascending order

2, 3, 4, 7, 8, 9, 10, 11, 14

There are 9 observations

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

= 5th observation = 8

The values of $(x_i - M)$ are –6, –5, –4, –1, 0, 1, 2, 3, 6

The values of $|x_i - M|$ are 6, 5, 4, 1, 0, 1, 2, 3, 6

$$\therefore MD(M) = \frac{\Sigma |x_i - M|}{9}$$

$$= \frac{6+5+4+1+0+1+2+3+6}{9}$$

$$= \frac{28}{9}$$

**Ans to Q No 3:** We have

(a)

| $x_i$ | $f_i$ | $f_i x_i$ | $x_i - \bar{x}$ | $\lvert x_i - \bar{x} \rvert$ | $f_i \lvert x_i - \bar{x} \rvert$ |
|---|---|---|---|---|---|
| 5 | 8 | 40 | –4 | 4 | 32 |
| 7 | 6 | 42 | –2 | 2 | 12 |
| 9 | 2 | 18 | 0 | 0 | 0 |
| 10 | 2 | 20 | 1 | 1 | 2 |
| 12 | 2 | 24 | 3 | 3 | 6 |
| 15 | 6 | 90 | 6 | 6 | 36 |

$\frac{1}{2}$

$N = \Sigma f_i = 26, \ \Sigma f_i x_i = 234$ $\therefore \bar{x} = \frac{\Sigma f_i x_i}{N} = \frac{234}{26} = 9$

$$\therefore MD(\bar{x}) = \frac{\Sigma f_i |x_i - \bar{x}|}{N} = \frac{88}{26} = 3.38$$

(b)

| $x_i$ | $f_i$ | cf | $x_i - M$ | $\lvert x_i - M \rvert$ | $f_i \lvert x_i - M \rvert$ |
|---|---|---|---|---|---|
| 5 | 8 | 8 | –2 | 2 | 16 |
| 7 | 6 | 14 | 0 | 0 | 0 |
| 9 | 2 | 16 | 2 | 2 | 4 |
| 10 | 2 | 18 | 3 | 3 | 6 |
| 12 | 2 | 20 | 5 | 5 | 10 |
| 15 | 6 | 26 | 8 | 8 | 48 |

N = 26, an even number

$\therefore$ Median = $\left[ \frac{26}{2} \text{th observation} + \left(\frac{26}{2}+1\right) \text{th observation} \right]$

$= \frac{1}{2} \left[ 13\text{th observation} + 14\text{th observation} \right]$

$$= \frac{1}{2}(7+7) = 7$$

$$MD(M) = \frac{\Sigma f_i \mid x_i - M \mid}{N} = \frac{84}{20} = 3.23$$

**Ans to Q No 4:** We have

$$7 = \frac{2+4+6+k+8+15}{6}$$

or, $7 = \frac{35+k}{6}$ or 42 = 35 + k, or k = 7

$\therefore$ the observations are 2, 4, 6, 7, 8, 15.

(a) We have

$\therefore$ the values of are

–5, –3, –1, 0, 1, 8

$\therefore$ the values of $<|x_i - \bar{x}|$ are

5, 3, 1, 0, 1, 8

(b) The observations are in ascending order :

2, 4, 6, 7, 8, 15

There are 6 observations.

Median M = $\frac{1}{2}\left[\ \frac{6}{2}\text{ th observation} + \left(\frac{6}{2}+1\right)\text{th observation}\ \right]$

$= \frac{1}{2}$ (3rd observation +4th observation)

$$= \frac{1}{2} \times (6+7) = \frac{13}{2} = 6.5$$

The values of $(x_i - M)$ are

–4.5, –2.5, –0.5, 0.5, 1.5, 8.5

$\therefore$ the values of $|x_i - M|$ are

4.5, 2.5, 0.5, 0.5, 1.5, 8.5

$$= \frac{18}{6} = 3$$

**Ans to Q No 5:**

| Class-interval | Frequency $f_i$ | Mid point $x_i$ | $f_ix_i$ | $x_i - \bar{x}$ | $\lvert x_i - \bar{x} \rvert$ | $f_i \lvert x_i - \bar{x} \rvert$ |
|---|---|---|---|---|---|---|
| 0–10 | 5 | 5 | 25 | –21 | 21 | 105 |
| 10–20 | 10 | 15 | 150 | –11 | 11 | 110 |
| 20–30 | 20 | 25 | 500 | –1 | 1 | 20 |
| 30–40 | 5 | 35 | 175 | 9 | 9 | 45 |
| 40–50 | 10 | 45 | 450 | 19 | 19 | 190 |

$$N = \sum_{i=1}^{S} f_i = 50$$

$$\sum_{i=1}^{S} f_i x_i = 1{,}300$$

$$\therefore \bar{x} = \frac{\sum_{i=1}^{S} f_i x_i}{N} = \frac{1300}{50} = 26$$

$$M = L + \frac{\frac{N}{2} - c}{f} \times h \qquad \therefore MD(\bar{x}) = \frac{\sum f_i \lvert x_i - \bar{x} \rvert}{N}$$

$$= \frac{470}{50} = 9.4$$

**Ans to Q No 6:**

| Class-interval | Frequency ($f_i$) | Mid point ($x_i$) | cf | $\lvert x_i - M \rvert$ | $f_i \lvert x_i - M \rvert$ |
|---|---|---|---|---|---|
| 0–10 | 6 | 5 | 5 | 23 | 115 |
| 10–20 | 8 | 15 | 13 | 13 | 104 |
| 20–30 | 15 | 25 | 28 | 3 | 45 |
| 30–40 | 16 | 35 | 44 | 7 | 112 |
| 40–50 | 6 | 45 | 50 | 17 | 102 |

Here $N = \Sigma f_i = 50$ $\qquad \therefore \frac{N}{2} = \frac{50}{2} = 25$

The cumulative frequency just greater than 25 is 28, and the corresponding class-interval is 20–30. $\therefore$ Median class is 20–30.

Now, median

Here L = 20, $\frac{N}{2} = 25$, c = 13, f = 15, h = 10

$$\therefore\ M = 20 + \frac{25-13}{15}\times 10 \quad = 20 + \frac{12}{15}\times 10 \quad = 20 + 8 = 28$$

$$\therefore MD(M) = \frac{\Sigma f_i \mid x_i - M \mid}{N} = \frac{478}{50} = 9.56$$

**Check Your Progress 3**

**Ans to Q No 1:** We know that if each of the observations is increased by a constant, the standard deviation remains the same.

$\therefore$ the new standard deviation is also 13.25.

**Ans to Q No 2:** The new standard deviation = 5 x 15.75 = 78.75

**Ans to Q No 3:** We know that the quartile deviation, mean deviation and standard deviation are connected by 6 QD = 5 MD = 450

Here SD = 27.45

6 QD = 4 x 27.45 or, QD =

**Ans to Q No 4:** We know that $\sigma^2 = \frac{\Sigma x_i^{\ 2}}{N} - \bar{x}^2$

Now, $\bar{x} = \frac{13+16+17+22}{4} = \frac{68}{4} = 17$

$$\therefore\ \sigma^2 = \frac{13^2+16^2+17^2+22^2}{4} - 17^2$$

$$= \frac{169+256+289+484}{4} - 289$$

$$= \frac{1198}{4} - 289 = 299.5 - 289 = 10.5$$

$$\therefore\ SD = \sigma = \sqrt{10.5}$$

**Ans to Q No 5:** We have $\sigma^2 = \frac{\Sigma x_i^{\ 2}}{N} - \bar{x}^2$

Now $\bar{x} = \frac{1+2+3+\ldots..+n}{n} = \frac{n(n+1)}{2n} = \frac{n+1}{2}$

$$\frac{\Sigma x_i^{\ 2}}{n} = \frac{1^2+2^2+3^2+\ldots..+n^2}{n}$$

$$= \frac{n(n+1)(2n+1)}{n} = \frac{(n+1)(2n+1)}{6}$$

$$\therefore\ \sigma^2 = \frac{(n+1)(2n+1)}{6} - \frac{(n+1)^2}{4}$$

$$= \frac{2(n+1)(2n+1) - 3(n+1)^2}{12}$$

$$= \frac{(n+1)[2(2n+1) - 3(n+1)]}{12}$$

$$= \frac{(n+1)(4n+2-3n-3)}{12}$$

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

$$\therefore SD = \sigma = \sqrt{\frac{n^2-1}{12}}$$

**Ans to Q No 6:** We have $\frac{1+2+a+6+b}{5} = 4.4$

or, 9 + a + b = 4.5 x 5 = 22

$\therefore$ a + b = 22 – 9 = 13 (1)

Again,

or, 1 + 4 + $a^2$ + 36 + $b^2$ – 19.36 x 5 = 8.24 x 5

$\frac{1^2+2^2+a^2+6^2+b^2}{5} - (4.4)^2 = 8.24$

or, 41 + $a^2$ + $b^2$ – 96.8 = 41.2

or, $a^2 + b^2$ = 97 (2)

From (1), b = 13 – a

$\therefore$ from (2), $a^2 + (13 - a)^2 = 97$

or, $a^2 + 169 - 26a + a^2 = 97$

or, $2a^2 - 26a + 72 = 0$

or, $a^2 - 13a + 36 = 0$

or, $a^2 - 4a - 9a + 36 = 0$

or, a(a–4) – 9(a–4) = 0

or, (a–4) (a–9) = 0

a = 4, or a = 9

If a = 4, b = 13 – 4 = 9

If a = 9, b = 13 – 9 = 4

the required observations are 4, 9.

**Ans to Q No 7:** Try yourself.

**Ans to Q No 8:** Try yourself.

**Check Your Progress 4**

**Ans to Q No 1:** CV (blood pressure)

(approx)

$$\text{CV (height)} = \frac{\sigma}{\bar{x}} \times 100 \quad = 3.125$$

now CV (blood pressure) > CV (height)

$\therefore$ blood pressure is more variable than height.



## 4.7 MODEL QUESTIONS

**Q 1:** In a class test 5 students obtain the following marks out of 20 in a particular subject

| A | B | C | D | E |
|---|---|---|---|---|
| 10 | 12 | 14 | 12 | 18 |

Calculate the standard deviation of these marks.

**Q 2:** Calculate the mean deviation about (a) mean, and (b) median from the following data.

| Class-interval | 20–30 | 30–40 | 40–50 | 50–60 | 60–70 |
|---|---|---|---|---|---|
| frequency | 120 | 201 | 150 | 75 | 25 |

**Q 3:** The mean and variance of 7 observations are 8 and 16 respectively. If 5 of these observations are 6, 8, 10, 12, 14; find the other two observations.

**Q 4:** The scores of two crickters A and B in five matches are given below.

| | | | | | |
|---|---|---|---|---|---|
| A : | 58 | 59 | 60 | 54 | 65 |
| B : | 82 | 12 | 44 | 08 | 64 |

Which crickter may be considered as more consistent?

**Q 5:** Coefficients of variation of two sets of observations are 58 and 69. Their standard deviations are 21.2 and 15.6 respectively. FInd their arithmetic means.

*** ***** ***