# Chapter 5: Measures of Dispersions

## Part-2

## For Grouped data

# Mean Deviation (M.D.):

For grouped data:

Sample Mean Deviation, $M.D.(\bar{x}) = \dfrac{\sum_{i=1}^{n} f_i \left| x_i - \bar{x} \right|}{n-1}$ ;

Population Mean Deviation, $M.D.(\bar{x}) = \dfrac{\sum_{i=1}^{n} f_i \left| x_i - \bar{x} \right|}{n}$

Where $x_1, x_2, \ldots, x_n$ be the mid-values and $f_1, f_2, \ldots, f_n$ be their corresponding frequencies.

**Example 3.** (a) *Find the Mean Deviation from the Mean for the following data :*

| Class Interval : | 0–10 | 10–20 | 20–30 | 30–40 | 40–50 | 50–60 | 60–70 |
|---|---|---|---|---|---|---|---|
| Frequency : | 8 | 12 | 10 | 8 | 3 | 2 | 7 |

(b) *Also find the mean deviation about median.*

(c) *Compare the results obtained in (a) and (b).*

**Solution.**

CALCULATIONS FOR M.D. ABOUT MEAN AND MEDIAN

| *Class Interval* | *Mid-value* ($X$) | *Frequency* ($f$) | *Less than c.f.* | $fX$ | $\lvert X-\bar{X}\rvert = \lvert X-29\rvert$ | $f\lvert X-\bar{X}\rvert$ | $\lvert X-Md\rvert = \lvert X-22\rvert$ | $f\lvert X-Md\rvert$ |
|---|---|---|---|---|---|---|---|---|
| 0—10 | 5 | 8 | 8 | 40 | 24 | 192 | 17 | 136 |
| 10—20 | 15 | 12 | 20 | 180 | 14 | 168 | 7 | 84 |
| 20—30 | 25 | 10 | 30 | 250 | 4 | 40 | 3 | 30 |
| 30—40 | 35 | 8 | 38 | 280 | 6 | 48 | 13 | 104 |
| 40—50 | 45 | 3 | 41 | 135 | 16 | 48 | 23 | 69 |
| 50—60 | 55 | 2 | 43 | 110 | 26 | 52 | 33 | 66 |
| 60—70 | 65 | 7 | 50 | 455 | 36 | 252 | 43 | 301 |
| Total | | $N = 50$ | | $\sum fX = 1,450$ | | $\sum f\lvert X-\bar{X}\rvert = 800$ | | $\sum f\lvert X-Md\rvert = 790$ |

($a$). Mean $(\overline{X}) = \frac{1}{N}\sum fX = \frac{1,450}{50} = 29.$

Mean Deviation about mean $= \frac{1}{N}\sum f \mid X - \overline{X} \mid = \frac{800}{50} = 16.$

($b$) $(N/2) = (50/2) = 25$. The c.f. just greate than 25 is 30. Hence, the corresponding class 20—30 is the median class. Using the median formula, we get

$$Md = l + \frac{h}{f}\left(\frac{N}{2} - C\right) = 20 + \frac{10}{25}(25 - 20) = 20 + 2 = 22$$

$\therefore$ Mean Deviation about median $= \frac{1}{N}\sum f \mid X - Md \mid = \frac{790}{50} = 15{\cdot}8$

($c$) From ($a$) and ($b$), we observe that :

M.D. about median < M.D. about mean.

In fact, we have the following general result :

"Mean deviation is least when taken about median."

# Variance

For grouped data:

$$\text{Sample Variance, } \sigma^2 = \frac{\sum_{i=1}^{n} f_i (x_i - \bar{x})^2}{n-1}; \quad or, \quad \sigma^2 = \frac{\sum_{i=1}^{n} f_i x_i^2}{n-1} - \bar{x}^2;$$

$$\text{Population Variance, } \sigma^2 = \frac{\sum_{i=1}^{n} f_i (x_i - \bar{x})^2}{n}; \quad or, \quad \sigma^2 = \frac{\sum_{i=1}^{n} f_i x_i^2}{n} - \bar{x}^2$$

Where $x_1, x_2, \ldots, x_n$ be the mid-values and $f_1, f_2, \ldots, f_n$ be their corresponding frequencies.

# Standard Deviation

**For grouped data:**

Sample Standard Deviation, $S.D. = \sigma = \sqrt{\dfrac{\sum_{i=1}^{n} f_i (x_i - \bar{x})^2}{n-1}}$; $\quad or, \quad \sigma = \sqrt{\dfrac{\sum_{i=1}^{n} f_i x_i^2}{n-1} - \bar{x}^2}$;

Population Standard Deviation, $\sigma^2 = \dfrac{\sum_{i=1}^{n} f_i (x_i - \bar{x})^2}{n}$; $\quad or, \quad \sigma^2 = \dfrac{\sum_{i=1}^{n} f_i x_i^2}{n} - \bar{x}^2$

Where $x_1, x_2, \ldots, x_n$ be the mid-values and $f_1, f_2, \ldots, f_n$ be their corresponding frequencies.

## 3. Different Formulae for Variance for Raw Data

If $x_1, x_2, \ldots, x_n$ are the $n$ observations, then

$$\sigma_x^2 = \frac{1}{n}\sum (X - \bar{X})^2 \qquad \text{............(i)}$$

$$\Rightarrow \qquad \sigma_x^2 = \frac{1}{n}\sum X^2 - \bar{X}^2 = \frac{1}{n}\sum X^2 - \left(\frac{1}{n}\sum X\right)^2 \quad \text{[Taking } f = 1 \text{ and } N = n \text{ in (i)]} \qquad \text{............(ii)}$$

**4.** If we are given $\bar{X}$ and $\sigma_x^2$, then we can obtain the values of $\sum X$ and $\sum X^2$ as discussed below.

From **(ii)**, we have

$$\bar{X} = \frac{1}{n}\sum X \quad \Rightarrow \quad \sum X = n\bar{X} \qquad \text{..........(iii)}$$

and

$$\sigma_x^2 = \frac{1}{n}\sum X^2 - \bar{X}^2 \quad \Rightarrow \quad \sum X^2 = n(\sigma_x^2 + \bar{X}^2) \qquad \text{..........(iv)}$$

Formulae **(iii)** and **(iv)** are very useful when we are given the values of the mean and standard deviation (or variance) and later on it is found that one or more of the observations are wrong and it is required to compute the mean and variance after replacing the wrong values by correct values or after deleting the wrong values.

**Example:** Compute the mean deviation, variance, and standard deviation of the given data:

| Class Interval | 48.5 – 53.5 | 53.5 - 58.5 | 58.5 - 63.5 | 63.5-68.5 | 68.5-73.5 | 73.5-78.5 | 78.5-83.5 | 83.5-88.5 | 88.5-93.5 | 93.5-98.5 |
|---|---|---|---|---|---|---|---|---|---|---|
| **Frequency** | 2 | 2 | 3 | 5 | 5 | 5 | 5 | 7 | 10 | 6 |

**Solution:** For solving the given problem first we construct the below table.

| Class Interval | Mid-Point ($x_i$ ) | Frequency ($f_i$) | ($f_i x_i$) | $\bar{x}$ | $\lvert x_i - \bar{x} \rvert$ | $\lvert x_i - \bar{x} \rvert^2$ | $f_i \lvert x_i - \bar{x} \rvert$ | $f_i \lvert x_i - \bar{x} \rvert^2$ |
|---|---|---|---|---|---|---|---|---|
| 48.5-53.5 | 51 | 2 | 102 | | 28.5 | 812.25 | 56.2 | 1624.5 |
| 53.5-58.5 | 56 | 2 | 112 | | 23.1 | 533.61 | 46.2 | 1067.22 |
| 58.5-63.5 | 61 | 3 | 183 | | 18.1 | 327.61 | 54.3 | 982.83 |
| 63.5-68.5 | 66 | 5 | 330 | | 13.1 | 171.61 | 65.5 | 858.05 |
| 68.5-73.5 | 71 | 5 | 335 | | 8.1 | 65.61 | 40.5 | 328.05 |
| 73.5-78.5 | 76 | 5 | 380 | $\frac{\sum f_i x_i}{n}$ =79.1 | 3.1 | 9.61 | 15.5 | 48.05 |
| 78.5-83.5 | 81 | 5 | 405 | | 1.9 | 3.61 | 9.5 | 18.05 |
| 83.5-88.5 | 86 | 7 | 602 | | 6.9 | 47.61 | 48.5 | 333.27 |
| 88.5-93.5 | 91 | 10 | 910 | | 11.9 | 141.61 | 119 | 1416.1 |
| 93.5-98.5 | 96 | 6 | 576 | | 16.9 | 285.61 | 101.4 | 1713.66 |
| **Total** | | $\sum f_i = 50 = n$ | **3955** | | $\sum \lvert x_i - \bar{x} \rvert = 131.6$ | $\sum \lvert x_i - \bar{x} \rvert^2 = 2398.74$ | $\sum f_i \lvert x_i - \bar{x} \rvert = 556.4$ | $\sum f_i \lvert x_i - \bar{x} \rvert^2 = 8389.78$ |

Hence we get, Sample mean deviation $M.D.(\bar{x}) = \frac{\sum_{i=1}^{n} f_i |x_i - \bar{x}|}{n-1} = \frac{556.4}{49} = 11.355$

Sample Variance, $\sigma^2 = \frac{\sum_{i=1}^{n} f_i (x_i - \bar{x})^2}{n-1} = \frac{8389.78}{49} = 171.22$

Sample Standard deviation, $S.D. = \sigma = \sqrt{\frac{\sum_{i=1}^{n} f_i (x_i - \bar{x})^2}{n-1}} = \sqrt{\frac{8389.78}{49}} = \sqrt{171.22} = 13.085$

**Example:** The lengths of 32 leaves were measures correct to nearest mm. Find the range, mean deviation, variance, and standard deviation of the given data.

| **Length** | 20-22 | 23-25 | 26-28 | 29-31 | 32-34 |
|---|---|---|---|---|---|
| **Frequency** | 3 | 6 | 12 | 9 | 2 |

**Solution:** For solving the given problem first we construct the below table.

| Length | **Mid-Point** $(x_i)$ | **Frequency** $(f_i)$ | $(f_i x_i)$ | $\bar{x}$ | $\lvert x_i-\bar{x}\rvert$ | $\lvert x_i-\bar{x}\rvert^2$ | $f_i\lvert x_i-\bar{x}\rvert$ | $f_i\lvert x_i-\bar{x}\rvert^2$ |
|---|---|---|---|---|---|---|---|---|
| 20-22 | 21 | 3 | 63 | $\frac{\sum f_i x_i}{n} = 27.09$ | 6.09 | 37.0881 | 18.27 | 111.2643 |
| 23-25 | 24 | 6 | 144 | | 3.09 | 9.5481 | 18.54 | 57.2886 |
| 26-28 | 27 | 12 | 324 | | 0.09 | 0.0081 | 1.08 | 0.0972 |
| 29-31 | 30 | 9 | 270 | | 2.97 | 8.8209 | 26.73 | 79.3881 |
| 32-34 | 33 | 2 | 66 | | 5.91 | 34.9281 | 11.82 | 69.8562 |
| **Total** | | $\sum f_i = 32 = n$ | **867** | | $\sum\lvert x_i-\bar{x}\rvert = 18.15$ | $\sum\lvert x_i-\bar{x}\rvert^2 = 90.3933$ | $\sum f_i\lvert x_i-\bar{x}\rvert = 76.44$ | $\sum f_i\lvert x_i-\bar{x}\rvert^2 = 317.8944$ |

In a frequency distribution, the range is calculated by taking the difference between the lower limit of the lower class and the upper limit of the highest class.

$$\text{Range} = L - S = 34 - 20 = 14$$

Now, Sample mean deviation $M.D.(\bar{x}) = \dfrac{\sum_{i=1}^{n} f_i |x_i - \bar{x}|}{n-1} = \dfrac{76.44}{32-1} = 2.4658$

Sample Variance, $\sigma^2 = \dfrac{\sum_{i=1}^{n} f_i (x_i - \bar{x})^2}{n-1} = \dfrac{317.8944}{31} = 10.2547$

Sample Standard deviation, $S.D. = \sigma = \sqrt{\dfrac{\sum_{i=1}^{n} f_i (x_i - \bar{x})^2}{n-1}} = \sqrt{\dfrac{317.8944}{31}} = \sqrt{10.2547} = 3.2023$

**Example 6·12.** *Calculate the mean and standard deviation from the following data :*

| *Value* | : | *90—99* | *80—89* | *70—79* | *60—69* | *50—59* | *40—49* | *30—39* |
|---|---|---|---|---|---|---|---|---|
| *Frequency* | : | *2* | *12* | *22* | *20* | *14* | *4* | *1* |

**Solution.**

CALCULATIONS FOR MEAN AND S.D.

| *Class* | *Mid-value* $(x)$ | *Frequency* $(f)$ | $d = \frac{x - 64{\cdot}5}{10}$ | $fd$ | $fd^2$ |
|---|---|---|---|---|---|
| 90—99 | 94·5 | 2 | 3 | 6 | 18 |
| 80—89 | 84·5 | 12 | 2 | 24 | 48 |
| 70—79 | 74·5 | 22 | 1 | 22 | 22 |
| 60—69 | 64·5 | 20 | 0 | 0 | 0 |
| 50—59 | 54·5 | 14 | –1 | –14 | 14 |
| 40—49 | 44·5 | 4 | –2 | – 8 | 16 |
| 30—39 | 34·5 | 1 | –3 | –3 | 9 |
| Total | | $N = 75$ | | $\sum fd = 27$ | $\sum fd^2 = 127$ |

$$\text{Mean} = A + \frac{h \sum fd}{N} = 64{\cdot}5 + \frac{10 \times 27}{75} = 64{\cdot}5 + 3{\cdot}6 = 68{\cdot}1$$

$$\text{S.D.} = h \cdot \sqrt{\frac{\sum fd^2}{N} - \left(\frac{\sum fd}{N}\right)^2} = 10 \times \sqrt{\frac{127}{75} - \left(\frac{27}{75}\right)^2} = 10 \times \sqrt{1{\cdot}6933 - 0{\cdot}1296}$$

$$= 10 \times \sqrt{1{\cdot}5637} = 10 \times 1{\cdot}2505 = 12{\cdot}505.$$

**Example 6·13.** *The arithmetic mean and the standard deviation of a set of 9 items are 43 and 5 respectively. If an item of value 63 is added to the set, find the mean and standard deviation of 10 items.*

**Solution.** We are given : $n = 9,\ \bar{x} = 43$ and $\sigma = 5$.

$$\bar{x} = \frac{\sum x}{n} \quad \Rightarrow \quad \sum x = n\bar{x} = 9 \times 43 = 387 \qquad \ldots(i)$$

Also
$$\sigma^2 = \frac{\sum x^2}{n} - (\bar{x})^2 \quad \Rightarrow \quad \sum x^2 = n(\sigma^2 + \bar{x}^2)$$

$$\therefore \quad \sum x^2 = 9(25 + 43^2) = 9(25 + 1849) = 9 \times 1874 = 16866 \qquad \ldots(ii)$$

If a new item 63 is added, then number of items becomes 10.

$$\text{New } (\sum x) = \sum x + 63 = 387 + 63 = 450 \qquad \text{[From } (i)]$$

$$\therefore \quad \text{New mean} = \frac{450}{10} = 45$$

$$\text{New } (\sum x^2) = \sum x^2 + 63^2 = 16866 + 3969 = 20835$$

$$\text{New } s.d. = \sqrt{\frac{\text{New } (\sum x^2)}{10} - (\text{New mean})^2} = \sqrt{\frac{20835}{10} - (45)^2}$$

$$= \sqrt{2083{\cdot}5 - 2025} = \sqrt{58{\cdot}5} = 7{\cdot}65.$$

**Example 6.15.** *The variance of a series of numbers 2, 3, 11 and x is $12\frac{1}{4}$. Find the value of x.*

**Solution.** We are given $n = 4$.

| $X$ | 2 | 3 | 11 | $x$ | $\sum X = 16 + x$ |
|---|---|---|---|---|---|
| $X^2$ | 4 | 9 | 121 | $x^2$ | $\sum X^2 = x^2 + 134$ |

$$\sigma_X^2 = \frac{\sum X^2}{n} - \left(\frac{\sum X}{n}\right)^2 = \frac{x^2 + 134}{4} - \left(\frac{16 + x}{4}\right)^2$$

$$= \frac{1}{16}\left[4\,(x^2 + 134) - (16 + x)^2\right] = \frac{1}{16}\left[4x^2 + 536 - (256 + x^2 + 32x)\right]$$

$$\Rightarrow \quad \sigma_X^2 = \frac{1}{16}\left[3x^2 - 32x + 280\right] = 12\frac{1}{4} = \frac{49}{4} \text{ (Given).}$$

$$\Rightarrow \quad 3x^2 - 32x + 280 = 49 \times 4 = 196 \quad \Rightarrow \quad 3x^2 - 32x + 84 = 0$$

$$\therefore \quad x = \frac{32 \pm \sqrt{(-32)^2 - 4 \times 3 \times 84}}{2 \times 3} = \frac{32 \pm \sqrt{1024 - 1008}}{6} = \frac{32 \pm 4}{6}$$

$$\Rightarrow \quad x = \left(\frac{32 + 4}{6} \text{ or } \frac{32 - 4}{6}\right) = \left(6 \text{ or } \frac{14}{3}\right).$$

**Example 6·16.** *The mean of 5 observations is 4·4 and the variance is 8·24. If three of the five observations are 1, 2 and 6, find the values of the other two.*

**Solution.** We are given $n = 5$, $\bar{x} = 4{\cdot}4$ and $\sigma^2 = 8{\cdot}24$

We have : $\sum x = n\bar{x} = 5 \times 4{\cdot}4 = 22$

and $\sum x^2 = n(\sigma^2 + \bar{x}^2) = 5(8{\cdot}24 + 19{\cdot}36) = 5 \times 27{\cdot}60 = 138$

Three observations are 1, 2 and 6. Let the two unknown observations be $x_1$ and $x_2$. Then

$$\sum x = 1 + 2 + 6 + x_1 + x_2 = 22 \quad \Rightarrow \quad x_1 + x_2 = 22 - 9 = 13 \qquad \ldots(*)$$

$$\sum x^2 = 1^2 + 2^2 + 6^2 + x_1^2 + x_2^2 = 138 \quad \Rightarrow \quad x_1^2 + x_2^2 = 138 - 41 = 97 \qquad \ldots(**)$$

Substituting the value of $x_2$ from (*) in (**) we get

$$x_1^2 + (13 - x_1)^2 = 97$$

$$\Rightarrow \quad x_1^2 + [13^2 + x_1^2 - 2 \times 13 \times x_1] = 97 \qquad [\because (a-b)^2 = a^2 + b^2 - 2ab]$$

$$\Rightarrow \quad x_1^2 + (169 + x_1^2 - 26x_1) = 97$$

$$\Rightarrow \quad 2x_1^2 - 26x_1 + 72 = 0$$

Solving as a quadratic equation in $x_1$, we get

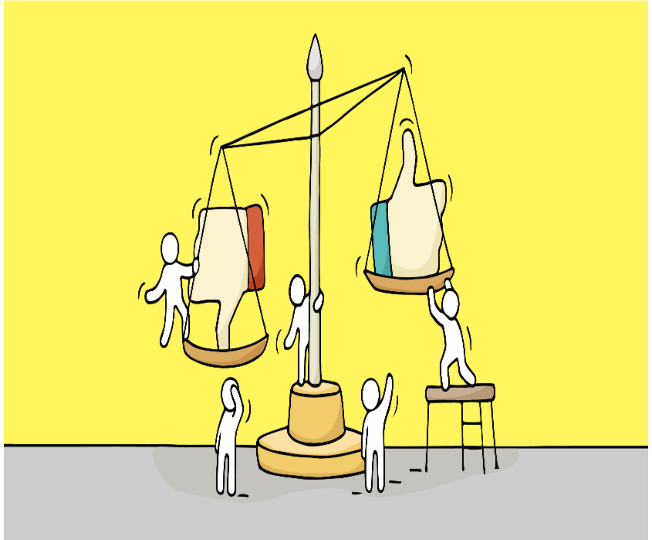
$$x_1 = \frac{26 \pm \sqrt{(26)^2 - 4 \times 2 \times 72}}{2 \times 2} = \frac{26 \pm \sqrt{676 - 576}}{4} = \frac{26 \pm 10}{4}$$

$$= \frac{26 + 10}{4} \text{ or } \frac{26 - 10}{4} = 9 \text{ or } 4$$

Substituting in (*), we get $x_1 = 9 \Rightarrow x_2 = 13 - 9 = 4$ or $x_1 = 4 \Rightarrow x_2 = 13 - 4 = 9$

Hence, the other two numbers are 4 and 9.

# Merits and limitations of Standard Deviation



**Merits of Standard Deviation:**
Among all measures of dispersion Standard Deviation is considered superior because it possesses almost all the requisite characteristics of a good measure of dispersion.

1) It is based on all the observations of the data set.
3) It is amenable to further mathematical calculation.

**Limitations:**
1) It is more affected by extreme items.

# Exercises

**7.** 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 |

**Ans.** Mean = 33·4 ; M.D. about mean = 13·184.

**8.** (*a*) Mean deviation may be calculated from the arithmetic mean or the median or the mode ? Which of these three measures is the minimum ?

(*b*) Find out mean deviation and its coefficient from median from the following series :

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| *Size of items* : | 4 | 6 | 8 | 10 | 12 | 14 | 16 |
| *Frequency* : | 2 | 1 | 3 | 6 | 4 | 3 | 1 |

**Ans.** 2·4 ; 0·24

**9.** Calculate the mean deviation about the mean for the following data :

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| $x$ : | 5 | 15 | 25 | 35 | 45 | 55 | 65 |
| $f$ : | 8 | 12 | 10 | 8 | 3 | 2 | 7 |

Also find the M.D. about median and comment on the results obtained in (*a*) and (*b*).

**Ans.** Mean = 29; M.D. about mean = 16. ; Median = 22 ; M.D. about median = 15·8.

**10.** Compute the mean deviation from the median and from mean for the following distribution of the scores of 50 college students.

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| *Score* | : | 140—150 | 150—160 | 160—170 | 170—180 | 180—190 | 190—200 |
| *Frequency* | : | 4 | 6 | 10 | 10 | 9 | 3 |

**Ans.** 10·24 ; 10·56.

**11.** Calculate Mean Deviation from Median from the following data :

| | | | | | | |
|---|---|---|---|---|---|---|
| *Wages in Rs. (Mid-value)* | : | 125 | 175 | 225 | 275 | 325 |
| *No. of persons* | : | 3 | 8 | 21 | 6 | 2 |

**Ans.** Median = 221·43 ; M.D. (Median) = 31·607.

**14.** (*a*) *Find the mean and standard deviation of the first n natural numbers.*

(*b*) *Hence deduce the mean and s.d. of the numbers 1, 2, 3, 4, 5, 6, 7, 8, 9, 10.*

**15.** Find out the mean and standard deviation of the following data.

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| *Age under (years)* | : | 10 | 20 | 30 | 40 | 50 | 60 | 70 | 80 |
| *No. of persons dying* | : | 15 | 30 | 53 | 75 | 100 | 110 | 115 | 125 |

**18.** The standard deviation calculated from a set of 32 observations is 5. If the sum of the observations is 80, what is the sum of the squares of these observations ?

**Ans.** $\sum X^2 = 1000$.

**19.** The mean of 200 items is 48 and their standard deviation is 3. Find the sum and sum of squares of all items.

**Ans.** 9,600 ; 4,62,600.

**20.** Given : No. of observations ($N$) = 100; Arithmetic average ($\bar{X}$) = 2 ; Standard deviation ($s_x$) = 4

find $\sum X$ and $\sum X^2$.

**Ans.** $\sum X = 200$, $\sum X^2 = 2000$.

**21.** The mean of 5 observations is 3 and variance is 2. If three of the five observations are 1, 3, 5, find the other two.

**Ans.** 2, 4.


Thank you