## Mean Deviation:

Mean deviation of a set of observations of a series is the arithmetic mean (AM) of all the deviations, (without their algebraic signs), taken from its central value [mean (or) median (or) mode].

(OR)

Mean deviation is the average of the modulus of the deviations of the observations in the series taken from mean or median or mode.

## Formula:

### I. Raw data:

$$\text{Mean deviation about mean} = \frac{\Sigma |x_i - \bar{x}|}{n}$$

$$\text{Mean deviation about median} = \frac{\Sigma |x_i - \text{median}|}{n}$$

$$\text{Mean deviation about mode} = \frac{\Sigma |x_i - \text{mode}|}{n}$$

### II. For frequency distribution

$$\text{Mean deviation about mean} = \frac{\Sigma f_i |x_i - \bar{x}|}{N}$$

$$\text{Mean deviation about median} = \frac{\Sigma f_i |x_i - \text{median}|}{N}$$

$$\text{Mean deviation about mode} = \frac{\Sigma f_i |x_i - \text{mode}|}{N}.$$

where $N = \Sigma f_i$

## Co efficient of Mean deviation

$$\text{Co eff of Mean deviation} = \frac{\text{Mean deviation about } A}{A}$$

where $A$ = Mean (or) Median (or) Mode.

Scanned by CamScanner

## Coefficient of Mean Variation

$$\text{Coeff of Mean Variation} = \frac{\text{mean deviation about } A}{A} \times 100$$

(or)

$$= \text{Coeff of mean deviation} \times 100.$$

## Problems:

1. The marks obtained by 10 students in an examinations were as follows

70, 65, 68, 70, 75, 73, 80, 70, 83, 86.

Find the mean deviation about the mean and co-eff of ~~variatiation~~. Variation.

**Soln:**

$$\text{Mean deviation about mean} = \frac{\Sigma |x_i - \bar{x}|}{n}.$$

$$\text{where } \bar{x} = \frac{70+65+68+70+75+73+80+70+83+86}{10}$$

$$= \frac{740}{10} = 74.$$

$$\text{Mean deviation} = \frac{\Sigma |x_i - \bar{x}|}{n}$$

$$= \frac{4+9+6+4+1+1+6+4+9+12}{10}$$

$$= \frac{56}{10} = 5.6.$$

$$\text{Coeff of Variation} = \frac{\text{Mean deviation}}{\text{mean}} \times 100$$

$$= \frac{5.6}{74} \times 100 = 7.57\%.$$

Scanned by CamScanner

2. Find the mean deviation from mean and median for the following data.

| marks | 20 | 18 | 16 | 14 | 12 | 10 | 8 | 6 |
|---|---|---|---|---|---|---|---|---|
| No of students | 2 | 4 | 9 | 18 | 27 | 25 | 14 | 1 |

Also calculate coeff of variation.

Soln:

i) Mean deviation about mean $= \dfrac{\Sigma f|x_i - \bar{x}|}{N}$, $N = \Sigma f$

$\bar{x} = \dfrac{\Sigma fx}{\Sigma f}$.

| $x$ | $f$ | $c.f$ | $xf$ | $\lvert x-\bar{x}\rvert$ | $f\lvert x-\bar{x}\rvert$ |
|---|---|---|---|---|---|
| 6 | 1 | 1 | 6 | 6 | 16 |
| 8 | 14 | 15 | 112 | 4 | 24 |
| 10 | 25 | 40 | 250 | 2 | 36 |
| 12 | 27 | 67 | 324 | 0 | 36 |
| 14 | 18 | 85 | 252 | 2 | 0 |
| 16 | 9 | 94 | 144 | 4 | 50 |
| 18 | 4 | 98 | 72 | 6 | 56 |
| 20 | 2 | 100 | 40 | 8 | 6 |
| $\Sigma$ | 100 | 1200 | | | 224 |

$\bar{x} = \dfrac{1200}{100} = 12$

Mean deviation about mean $= \dfrac{\Sigma f|x_i - \bar{x}|}{\Sigma f} = \dfrac{224}{100} = 22.4$

Co-eff of variation $= \dfrac{MD}{\bar{x}} \times 100 = \dfrac{22.4}{12} \times 100 = 18.67\%$

Scanned by CamScanner

ii) Mean deviation about median $= \frac{\Sigma f |x_i - \text{median}|}{\Sigma f}$, $N = \Sigma f$

Median = Average of $\left(\frac{N}{2}\right)^{th}$ and $\left(\frac{N}{2}+1\right)^{th}$ item

$$= \frac{\left(\frac{100}{2}\right)^{th} + \left(\frac{100}{2}+1\right)^{th} \text{term}}{2} = \frac{12+12}{2}$$

$= 12.$

mean deviation about median $= \frac{\Sigma f |x_i - 12|}{\Sigma f}$

$$= \frac{224}{100} = 2.24$$

Coefficient of variation $= \frac{MD}{\text{median}} \times 100 = \frac{2.24}{12} \times 100 = 18.67\,\%$

3. Calculate the mean deviation from the mean for the following data.

| class interval | 0 – 4 | 4 – 8 | 8 – 12 | 12 – 16 | 16 – 20 |
|---|---|---|---|---|---|
| Frequency | 4 | 6 | 8 | 5 | 2 |

Soln. : Assuming that the frequencies in each class are centered as its mid value.

| Class interval | $x$ | $f$ | $xf$ | $\lvert x-\bar{x}\rvert$ | $f\lvert x-\bar{x}\rvert$ |
|---|---|---|---|---|---|
| 0 – 4 | 2 | 4 | 8 | 7.2 | 28.8 |
| 4 – 8 | 6 | 6 | 16 | 3.2 | 19.2 |
| 8 – 12 | 10 | 8 | 80 | 0.8 | 6.4 |
| 12 – 16 | 14 | 5 | 70 | 4.8 | 24.0 |
| 16 – 20 | 18 | 2 | 36 | 8.8 | 17.6. |
| | $\Sigma$ | 25 | 230 | | 96. |

$\bar{x} = \frac{\Sigma f x}{\Sigma f} = \frac{230}{25} = 9.2$, Mean deviation $= \frac{\Sigma f |x-\bar{x}|}{\Sigma f} = \frac{96}{25} = 3.84$

Scanned by CamScanner

## Standard deviation [or Root mean square deviation].

Standard deviation is the positive square root of the average of squared deviations taken from arithmetic mean.

It is denoted by $\sigma$ (or) S.D. (or) s.d.

### Formula:

#### I - Raw data:

**Direct Method:**

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

**Short cut Method:**

$$\sigma = \sqrt{\frac{\Sigma d^2}{n} - \left(\frac{\Sigma d}{n}\right)^2}$$

, where $d = x - A$

and $A$ = Assumed mean.

$n$ = total number of observations

#### II. Discrete frequency distribution

**Direct Method:**

$$\sigma = \sqrt{\frac{\Sigma f (x-\bar{x})^2}{N}}, \quad N = \Sigma f$$

**Short cut Method:**

$$\sigma = \sqrt{\frac{\Sigma f d^2}{N} - \left(\frac{\Sigma f d}{N}\right)^2}, \quad N = \Sigma f$$

**Step deviation Method:**

In this method we divide the deviations by a common class interval and use the following formula for computing standard deviation:

Scanned by CamScanner

$$\sigma = \sqrt{\frac{\Sigma fd^2}{N} - \left(\frac{\Sigma fd}{N}\right)^2} \times i$$

where $i$ = common class interval

$d = \frac{x - A}{i}$ , $A$ = assumed mean.

## III - continuous.

The standard deviation of a continuous series can be calculated by any one of the methods discussed for discrete frequency distribution. However, in practice only step deviation method is mostly used.

$$\sigma = \sqrt{\frac{\Sigma fd^2}{N} - \left(\frac{\Sigma fd}{N}\right)^2} \times i$$

where $i$ = class interval (or the common factor in case the class intervals are unequal).

$d = \frac{m - A}{i}$ , $m$ is mid value
$A$ is Assumed mean.

## Coeff of standard deviation

$$\text{Co-eff of SD} = \frac{\sigma}{\bar{x}}$$

## Coeff of variations

$$\text{Co-eff of Variations (C.V)} = \frac{\sigma}{\bar{x}} \times 100$$

(or)

$$= \text{coeff of SD} \times 100.$$

Scanned by CamScanner

Problems.

1. Find the standard deviation of 16, 13, 17, 22.

Soln:

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

| $x$ | $x-\bar{x}$ | $(x-\bar{x})^2$ |
|---|---|---|
| 16 | $-1$ | 1 |
| 13 | $-4$ | 16 |
| 17 | 0 | 0 |
| 22 | 5 | 25 |
| | $\Sigma$ | 42 |

$$\therefore SD = \sigma = \sqrt{\frac{\Sigma(x-\bar{x})^2}{n}}$$

$$= \sqrt{\frac{42}{4}} = 3.24$$

2. Find the mean and s.d. of the following data, 48, 43, 65, 57, 31, 60, 37, 48, 59, 78.

Soln: Let $A = 50$

| $x$ | $d = x - A$ | $d^2$ |
|---|---|---|
| 48 | $-2$ | 4 |
| 43 | $-7$ | 49 |
| 65 | 15 | 225 |
| 57 | 7 | 49 |
| 31 | $-19$ | 361 |
| 60 | 10 | 100 |
| 37 | $-13$ | 169 |
| 48 | $-2$ | 4 |
| 59 | 9 | 81 |
| 78 | 28 | 784 |
| $\Sigma$ | 26 | 1826 |

Scanned by CamScanner

$$\text{Mean} = \bar{x} = A + \frac{\Sigma d}{n}$$

$$= 50 + \frac{26}{10}$$

$$= 52.6$$

$$SD = \sqrt{\frac{\Sigma d^2}{n} - \left(\frac{\Sigma d}{n}\right)^2}$$

$$= \sqrt{\frac{1826}{10} - \left(\frac{26}{10}\right)^2}$$

$$= 13.26$$

**3.** Find mean and SD from the following data.

| Size of the item | 10 | 11 | 12 | 13 | 14 | 15 | 16 |
|---|---|---|---|---|---|---|---|
| frequency | 2 | 7 | 11 | 15 | 10 | 4 | 1 |

Also find the co-eff of variation

Soln:

| $x$ | $f$ | $d = x - A$ $= x - 13$ | $fd$ | $d^2$ | $fd^2$ |
|---|---|---|---|---|---|
| 10 | 2 | −3 | −6 | 9 | 18 |
| 11 | 7 | −2 | −14 | 4 | 28 |
| 12 | 11 | −1 | −11 | 1 | 11 |
| 13 | 15 | 0 | 0 | 0 | 0 |
| 14 | 10 | 1 | 10 | 1 | 10 |
| 15 | 4 | 2 | 8 | 4 | 16 |
| 16 | 1 | 3 | 3 | 9 | 9 |
| | 50 | | −10 | | 92 |

$$\text{Mean} = \bar{X} = A + \frac{\Sigma fd}{\Sigma f} = 13 + \frac{(-10)}{50} = 12.8$$

Scanned by CamScanner

$$SD = \sigma = \sqrt{\frac{\Sigma f d^2}{N} - \left(\frac{\Sigma f d}{N}\right)^2}$$

$$= \sqrt{\frac{92}{50} - \left(\frac{-10}{50}\right)^2}$$

$$= 1.342.$$

$$\text{Co-eff of variation} = \frac{\sigma}{\bar{x}} \times 100 = \frac{1.342}{12.8} \times 100$$

$$= 10.4.$$

4. Daily Sales (recorded in rupees) of a retail shop are given below

| Daily sales in Rs (X-midpoint of interval) | 102 | 106 | 110 | 114 | 118 | 122 | 126 |
|---|---|---|---|---|---|---|---|
| No of days (frequency-f) | 3 | 9 | 25 | 35 | 17 | 10 | 1 |

Calulate the mean and the standard deviation of the above data, Also calulate the co-eff of variation.

Soln: Let $A = 114$

$$\therefore d = \frac{x - A}{i}$$

$$= \frac{x - 114}{4} \quad , i = 4 \quad \text{(since mid value is given)}$$

Scanned by CamScanner

| No of days $x$ | $f$ | $d = \frac{x-114}{4}$ | $fd$ | $d^2$ | $fd^2$ |
|---|---|---|---|---|---|
| 102 | 3 | -3 | -9 | 9 | 27 |
| 106 | 9 | -2 | -18 | 4 | 36 |
| 110 | 25 | -1 | -25 | 1 | 25 |
| 114 | 35 | 0 | 0 | 0 | 0 |
| 118 | 17 | 1 | 17 | 1 | 17 |
| 122 | 10 | 2 | 20 | 4 | 40 |
| 126 | 1 | 3 | 3 | 9 | 9 |
| $\Sigma$ | 100 | | -12 | | 154 |

$$\bar{x} = A + \frac{\Sigma fd}{N} \times i$$

$$= 114 + \frac{-12}{100} \times 4$$

$$= 113.52$$

$$\sigma = \sqrt{\frac{\Sigma fd^2}{N} - \left(\frac{fd}{N}\right)^2} \times i, \quad N = \Sigma f$$

$$= \sqrt{\frac{154}{100} - \left(\frac{-12}{100}\right)^2} \times 4$$

$$= 4.94$$

Co eff of variation : $C.V = \frac{\sigma}{\bar{x}} \times 100 = \frac{4.94}{113.52} \times 100$

$$= 4.35 \%$$

Scanned by CamScanner

5. Find the standard deviation for the following distribution.

| marks | 10-20 | 20-30 | 30-40 | 40-50 | 50-60 | 60-70 | 70-80 |
|---|---|---|---|---|---|---|---|
| No of students | 5 | 12 | 15 | 20 | 10 | 4 | 2 |

Soln: let $A = 45$.

| Marks | f | midvalue (m) | $d = \frac{m-45}{10}$ | fd | $fd^2$ |
|---|---|---|---|---|---|
| 10-20 | 5 | 15 | -3 | -15 | 45 |
| 20-30 | 12 | 25 | -2 | -24 | 48 |
| 30-40 | 15 | 35 | -1 | -15 | 15 |
| 40-50 | 20 | 45 | 0 | 0 | 0 |
| 50-60 | 10 | 55 | 1 | 10 | 10 |
| 60-70 | 4 | 65 | 2 | 8 | 16 |
| 70-80 | 2 | 75 | 3 | 6 | 18 |
| | | | | -30 | 152 |
| $\Sigma$ | 68 | | | | |

$$\sigma = \sqrt{\frac{\Sigma fd^2}{N} - \left(\frac{\Sigma fd}{N}\right)^2} \times i \quad , N = \Sigma f$$

$$= \sqrt{\frac{152}{68} - \left(\frac{-30}{68}\right)^2} \times 10$$

$$= 14.3$$

Scanned by CamScanner

## Karl Pearson's co-eff of correlation

Karl Pearson [1857-1936] was a great statistician. He gave the following mathematical formula for measuring the magnitude of linear co-eff between two variables.

If X and Y are two variables, then the correlation co-eff $\rho(X, Y)$ between them is given by

$$r_{xy} = r = \rho(X, Y) = \rho = \frac{cov[X, Y]}{\sigma_X \sigma_Y}$$

$$= \frac{cov[X, Y]}{\sqrt{Var X}\sqrt{Var Y}}$$

where $cov[X, Y] = E[XY] - E[X]E[Y]$

$$= \frac{\Sigma xy}{n} - \left(\frac{\Sigma x}{n}\right)\left(\frac{\Sigma y}{n}\right)$$

$$\sigma_X^2 = Var[X] = \frac{\Sigma x^2}{n} - \left(\frac{\Sigma x}{n}\right)^2$$

$$\sigma_Y^2 = Var[Y] = \frac{\Sigma y^2}{n} - \left(\frac{\Sigma y}{n}\right)^2$$

**Another form of formula:**

$$\rho(X, Y) = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\Sigma(x_i - \bar{x})^2}\sqrt{\Sigma(y_i - \bar{y})^2}}$$

Scanned by CamScanner

## Rank Correlation:

If $(x_i, y_i)$, $i = 1, 2, \ldots, n$ be the ranks of the individuals in two characteristics A and B respectively, then the rank correlation co-efficients is given by.

$$r = 1 - \frac{6}{n(n^2-1)} \sum_{i=1}^{n} d_i^2$$

where $d_i = x_i - y_i$, difference between ranks.

$n$ = no. of items

This formula is called spearman's formula for the rank correlation co-eff.

## Repeated Rank correlation:

$$r[x, y] = 1 - \frac{6}{n(n^2-1)} \left[ \sum_{i=1}^{n} d_i^2 + \frac{m(m^2-1)}{12} + \frac{m(m^2-1)}{12} + \cdots \right]$$

where $m$ is the number of times an item (Rank) is repeated.

Scanned by CamScanner

① Calculate the co-eff of correlation between X and Y for the following data.

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| X : | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
| Y : | 9 | 8 | 10 | 12 | 11 | 13 | 14 | 16 | 15 |

Soln:

Method-1:

| | X | Y | $X^2$ | $Y^2$ | XY |
|---|---|---|---|---|---|
| | 1 | 9 | 1 | 81 | 9 |
| | 2 | 8 | 4 | 64 | 16 |
| | 3 | 10 | 9 | 100 | 30 |
| | 4 | 12 | 16 | 144 | 48 |
| | 5 | 11 | 25 | 121 | 55 |
| | 6 | 13 | 36 | 169 | 78 |
| | 7 | 14 | 49 | 196 | 96 |
| | 8 | 16 | 64 | 256 | 128 |
| | 9 | 15 | 81 | 225 | 135 |
| $\Sigma$ | 45 | 108 | 285 | 1356 | 597 |

$$Cov[X,Y] = \frac{\Sigma XY}{n} - \frac{\Sigma X}{n}\frac{\Sigma Y}{n} = \frac{597}{9} - \left(\frac{45}{9}\right)\left(\frac{108}{9}\right) = 6.3333$$

$$\sigma_X^2 = \frac{\Sigma X^2}{n} - \left(\frac{\Sigma X}{n}\right)^2 = \frac{285}{9} - \left(\frac{45}{9}\right)^2 = 6.6667$$

$$\sigma_X = 2.5820$$

$$\sigma_Y^2 = \frac{\Sigma Y^2}{n} - \left(\frac{\Sigma Y}{n}\right)^2 = \frac{1356}{9} - \left(\frac{108}{9}\right)^2 = 6.6667$$

$\sigma_Y = 2.5820$

Scanned by CamScanner

$$\rho = \frac{Cov[X,Y]}{\sigma_X \ \sigma_Y}$$

$$= \frac{6.3333}{(2.5820)(2.5820)} = 0.95$$

Method-II

$$\bar{X} = \frac{\Sigma X}{n} = \frac{45}{9} = 5, \quad \bar{Y} = \frac{\Sigma Y}{n} = \frac{108}{9} = 12$$

| $X$ | $Y$ | $X-\bar{X}$ | $Y-\bar{Y}$ | $(X-\bar{X})^2$ | $(Y-\bar{Y})^2$ | $(X-\bar{X})(Y-\bar{Y})$ |
|---|---|---|---|---|---|---|
| 1 | 9 | -4 | -3 | 16 | 9 | 12 |
| 2 | 8 | -3 | -4 | 9 | 16 | 12 |
| 3 | 10 | -2 | -2 | 4 | 4 | 4 |
| 4 | 12 | -1 | 0 | 1 | 0 | 0 |
| 5 | 11 | 0 | -1 | 0 | 1 | 0 |
| 6 | 13 | 1 | 1 | 1 | 1 | 1 |
| 7 | 14 | 2 | 2 | 4 | 4 | 4 |
| 8 | 16 | 3 | 4 | 9 | 16 | 12 |
| 9 | 15 | 4 | 3 | 16 | 9 | 12 |
| $\Sigma$ | | 0 | 0 | 60 | 60 | 57 |

$$r(x,y) = \rho(X,Y) = \frac{\Sigma(x-\bar{x})(y-\bar{y})}{\sqrt{\Sigma(x-\bar{x})^2}\sqrt{\Sigma(y-\bar{y})^2}}$$

$$= \frac{57}{\sqrt{60}\sqrt{60}} = 0.95.$$

Scanned by CamScanner

2. Ten student's got the following percentage of marks in economics and statistics. Calculate the rank correlation co-eff

| Economics | 78 | 36 | 98 | 25 | 75 | 82 | 90 | 6 | 65 | 39 |
|---|---|---|---|---|---|---|---|---|---|---|
| Statistics | 84 | 51 | 91 | 60 | 68 | 62 | 86 | 58 | 53 | 47 |

Soln:

| X | Y | Rank of X $x_i$ | Rank of Y $y_i$ | $d_i = x_i - y_i$ | $d_i^2$ |
|---|---|---|---|---|---|
| 78 | 84 | 4 | 3 | 1 | 1 |
| 36 | 51 | 8 | 9 | -1 | 1 |
| 98 | 91 | 1 | 1 | 0 | 0 |
| 25 | 60 | 9 | 6 | 3 | 9 |
| 75 | 68 | 5 | 4 | 1 | 1 |
| 82 | 62 | 3 | 5 | -2 | 4 |
| 90 | 86 | 2 | 2 | 0 | 0 |
| 6 | 58 | 10 | 7 | 3 | 9 |
| 65 | 53 | 6 | 8 | -2 | 4 |
| 39 | 47 | 7 | 10 | -3 | 9 |
| | | | | $\Sigma$ | 38 |

$$r = 1 - \frac{6}{n(n^2-1)} \Sigma d_i^2$$

$$= 1 - \frac{6}{10(10^2-1)}(38)$$

$$= 0.769$$

Scanned by CamScanner

3. Find the rank correlation co-eff between X and Y for the given data

| X | 90 | 82 | 82 | 82 | 81 | 71 | 63 | 63 | 49 | 38 |
|---|---|---|---|---|---|---|---|---|---|---|
| Y | 75 | 72 | 71 | 71 | 71 | 71 | 50 | 40 | 32 | 32 |

Soln:

| X | Y | Rank of X $x_i$ | Rank of Y $y_i$ | $d_i = x_i - y_i$ | $d_i^2$ |
|---|---|---|---|---|---|
| 90 | 75 | 1 | 1 | 0 | 0 |
| 82 | 72 | 3 | 2 | 1 | 1 |
| 82 | 71 | 3 | 4.5 | -1.5 | 2.25 |
| 82 | 71 | 3 | 4.5 | -1.5 | 2.25 |
| 81 | 71 | 5 | 4.5 | 0.5 | 0.25 |
| 71 | 71 | 6 | 4.5 | 1.5 | 2.25 |
| 63 | 50 | 7.5 | 7 | 0.5 | 0.25 |
| 63 | 40 | 7.5 | 8 | -0.5 | 0.25 |
| 49 | 32 | 9 | 9.5 | -0.5 | 0.25 |
| 38 | 32 | 10 | 9.5 | 0.5 | 0.25 |
| | | | | $\Sigma$ | 9 |

$$r(x,y) = 1 - \frac{6}{n(n^2-1)}\left[\Sigma d_i^2 + \frac{m(m^2-1)}{12} + \frac{m(m^2-1)}{12} + \cdots\right]$$

$$= 1 - \frac{6}{10(10^2-1)}\left[9 + \frac{3(3^2-1)}{12} + \frac{2(2^2-1)}{12} + \frac{4(4^2-1)}{12} + \frac{2(2^2-1)}{12}\right]$$

$$= 0.8969$$

Scanned by CamScanner

# Measures of Skewness, Kurtosis and Moments

## Skewness:

Averages determine the central point of distribution, but they give no information about the shape of the frequency curve.

Measures of dispersion give some ideas of the spread of a variable about its average. Both these measures do not study whether a distribution is symmetrical or not.

Skewness is a measure to study this aspect of a statistical distribution.

If a distribution is not symmetrical, we say that it is skewed.

In a Symmetrical distribution, the values of the variable equi-distant from their mean have equal frequencies.

In a perfectly symmetrical distribution mean, median and mode coincide. If this is not the case, the distribution is said to be skewed.

Scanned by CamScanner

When the distribution is skewed to the right, mean is greater than the mode. When the distribution is skewed to the left, mean is less than the mode.

The measure of skewness based on these is due to pearson and is called pearson's coefficient of skewness.

$$\text{Pearson's co-eff of skewness} = \frac{\text{Mean} - \text{Mode}}{SD}.$$

When mode is not well defined,

$$\text{Pearson's co-eff of skewness} = \frac{3(\text{mean} - \text{Median})}{SD}$$

Also in a symmetrical distribution, the quartiles $Q_1$ and $Q_3$ are equidistant from their median. If it is not the case, then we say that the distribution is skewed.

The measure of skewness based on this is due to Bowley.

$$\text{Bowley's co-eff of skewness} = \frac{Q_3 + Q_1 - 2M}{Q_3 - Q_1}$$

Note:

When the distribution is skewed to the right mean > mode. Therefore, from pearson's formula we see that the co-eff of skewness is positive. $||^{ly}$, when the distribution is skewed to the left, the skewness is negative.

Scanned by CamScanner

1. Calculate Pearson's measure of skewness for the following data

| Size | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 |
|---|---|---|---|---|---|---|---|---|
| frequency | 2 | 11 | 36 | 64 | 39 | 30 | 22 | 2 |

Soln:

| $x$ | $f$ | $d$ | $fd$ | $fd^2$ |
|---|---|---|---|---|
| 7 | 2 | -3 | -6 | 18 |
| 8 | 11 | -2 | -22 | 44 |
| 9 | 36 | -1 | -36 | 36 |
| 10 | 64 | 0 | 0 | 0 |
| 11 | 39 | 1 | 39 | 39 |
| 12 | 30 | 2 | 60 | 120 |
| 13 | 22 | 3 | 66 | 198 |
| 14 | 2 | 4 | 8 | 32. |
| | 206 | | 109 | 487 |

$\because d = x - 10.$

This is a discrete dat.

Maximum frequency corresponds to $x = 10$.

$\therefore$ Mode $= 10$.

Let $A = 10$, $d = x - 10$

$$\text{mean} = A + \frac{\Sigma fd}{\Sigma f}$$

$$= 10 + \frac{109}{206}$$

$$= 10.53$$

[$\because$ Also by direct method $\text{mean} = \frac{\Sigma fx}{\Sigma f} = \frac{2169}{206} = 10.53$

$$SD = \sigma = \sqrt{\frac{\Sigma fd^2}{N} - \left(\frac{\Sigma fd}{N}\right)^2}$$

$$= \sqrt{\frac{487}{206} - \left(\frac{109}{206}\right)^2} = 1.44$$

Scanned by CamScanner

Pearson's coeff of skewness $= \frac{\text{Mean} - \text{Mode}}{SD}$

$= \frac{10.53 - 10}{1.44}$

$= 0.3680$

2. Calulate pearson measure of skewness for the distribution.

| class interval | 0-9 | 10-19 | 20-29 | 30-39 | 40-49 |
|---|---|---|---|---|---|
| frequency | 8 | 15 | 23 | 16 | 9 |

Soln:

| class | Mid 'x' | f | d | fd | $fd^2$ |
|---|---|---|---|---|---|
| -0.5-9.5 | 4.5 | 8 | -2 | -16 | 32 |
| 9.5-19.5 | 14.5 | 15 | -1 | -15 | 15 |
| 19.5-29.5 | 24.5 | 23 | 0 | 0 | 0 |
| 29.5-39.5 | 34.5 | 16 | 1 | 16 | 16 |
| 39.5-49.5 | 44.5 | 9 | 2 | 18 | 36 |
| | | 71 | | 3 | 99 |

$\left[\because d = \frac{x - 24.5}{10}\right]$

$\text{Mean} = A + \frac{\Sigma fd}{N} \times i$

$= 24.5 + \frac{3}{71} \times 10$

$= 24.9225$

$\text{Mode} = l + \frac{f_m - f_1}{2f_m - f_2 - f_1} \times i$

$= 19.5 + \frac{23 - 15}{2(23) - 15 - 16} \times 10 = 21$

Scanned by CamScanner

$$\sigma = \sqrt{\frac{\Sigma fd^2}{N} - \left(\frac{\Sigma fd}{N}\right)^2} \times i \qquad [\because N = \Sigma f]$$

$$= \sqrt{\left(\frac{99}{71}\right) - \left(\frac{3}{71}\right)^2} \times 10$$

$$= 11.8008.$$

Pearson's co eff of skewness $= \dfrac{\text{Mean} - \text{Mode}}{\sigma}$

$$= \frac{24.9225 - 21}{11.8008}$$

$$= 0.332$$

(3) Calculate the Bowley's measure of skewness for the following data.

| Payment of Commission | 100-120 | 120-140 | 140-160 | 160-180 | 180-200 | 200-220 | 220-240 | 240-260 | 260-280 | 280-300 |
|---|---|---|---|---|---|---|---|---|---|---|
| No. of salesmen | 4 | 10 | 16 | 29 | 52 | 80 | 42 | 23 | 17 | 7. |

Soln:

| class | f | cf |
|---|---|---|
| 100 – 120 | 4 | 4 |
| 120 – 140 | 10 | 14 |
| 140 – 160 | 16 | 30 |
| 160 – 180 | 29 | 59 |
| 180 – 200 | 52 | 111 |
| 200 – 220 | 80 | 191 |
| 220 – 240 | 42 | 233 |
| 240 – 260 | 23 | 256 |
| 260 – 280 | 17 | 273 |
| 280 – 300 | 7 | 280 |

Scanned by CamScanner

To find $Q_1$: $\frac{iN}{4}$, where $N = 54$.

$\frac{N}{4} = \frac{280}{4} = 70.$

$Q_1$ lies in the group $\cancel{180}$-200, $L = 180$, $C = 59$, $h = 20$, $f = 52$

$\therefore Q_1 = L + \left(\frac{\frac{N}{4} - C}{f}\right) \times h = 180 + \frac{70 - 59}{52} \times 20 = 184.23$

To find $Q_2$ (median): $\frac{iN}{4}$

$\frac{N}{2} = \frac{280}{4} = 140$

$Q_2$ lies in the group 200-220, $L = 200$, $C = 111$, $h = 20$, $f = 80$

$\therefore Q_2 = L + \left(\frac{\frac{N}{2} - C}{f}\right) \times h = 200 + \frac{140 - 111}{80} \times 20 = 207.25$

To find $Q_3$: $\frac{iN}{4}$

$\frac{3N}{4} = \frac{3 \times 280}{4} = 210$

$Q_3$ lies in the group 220-240; $L = 220$, $C = 191$, $h = 20$; $f = 42$

$\therefore Q_3 = L + \left(\frac{\frac{3N}{4} - C}{f}\right) \times h = 220 + \frac{210 - 191}{42} \times 20 = 229.05$.

Bowley's co-eff of skewness $= \frac{Q_3 + Q_1 - 2M}{Q_3 - Q_1}$

$= \frac{229.05 + 184.23 - (2 \times 207.25)}{229.05 - 184.23}$

$= -0.03.$

Note: There is a very small negative co-eff of skewness $\therefore$ the distribution is almost symmetrical.

Scanned by CamScanner