SKEWNESS

Definition:

Skewness is the degree of the asymmetry, or departure from symmetry of a distribution.

1. Karl Pearson's coefficient of Skewness (Sk_p)

(i) $$Sk_p = \frac{\text{Mean} - \text{Mode}}{\text{Standard Deviation}}$$

$$Sk_p = \frac{\overline{X} - Z}{\sigma}$$

(ii) When mode is ill-defined, the following formula can be used:

$$Sk_p = \frac{3(\text{Mean} - \text{Median})}{\text{Standard Deviation}}$$

$$Sk_p = \frac{3(\overline{X} - M)}{\sigma}$$

2. Bowley's coefficient of Skewness (Sk_B)

$$Sk_B = \frac{Q_3 + Q_1 - 2M}{Q_3 - Q_1}$$

$$Sk_B = \frac{Q_3 + Q_1 - 2M}{Q_3 - Q_1}$$

Karl Pearson's coefficient of Skewness (Sk_p)

1. From the marks secured by 120 students in Section A and B of a class, the

Following measures are obtained:

Section A: $\overline{X}$ = 46.83; S.D = 14.8; Mode = 51.67

Section B: $\overline{X}$ = 47.83; S.D = 14.8; Mode = 47.07

Determine which distribution of marks is more skewed.

Solution: Karl Pearson's coefficient of Skewness

For Section A: $Sk_p = \frac{\overline{X} - Z}{\sigma} = \frac{46.83 - 51.67}{14.8} = \frac{-4.84}{14.8} = -0.3270$

For Section B: $Sk_p = \frac{\overline{X} - Z}{\sigma} = \frac{47.83 - 47.07}{14.8} = \frac{0.76}{14.8} = 0.05135$

Marks of Section A is more Skewed. But marks of Section A is negatively Skewed. Marks of Section B are Positively skewed.

2. From a moderately skewed distribution of retail prices for men's shoes, it is found that the mean price is Rs. 20 and the median price is Rs. 17. If the coefficient of variation is 20%, find the Pearsonian coefficient of skewness of the distribution.

Solution: Given: C.V. = 20, $\overline{X}$ = 20, M = 17

$$\text{C. V.} = \frac{\sigma}{\overline{X}} X 100$$

$$20 = \frac{\sigma}{20} X 100 = 20 \text{ X } 20 /100 = 400/100 = 4$$

$$\sigma = 4$$

$$Sk_p = \frac{3(\overline{X} - M)}{\sigma} = \frac{3(20 - 17)}{4} = 9/4 = 2.25$$

3. Calculate Karl Pearson's coefficient of Skewness for the following data.

X	X^2
25	625
15	225
23	529
40	1600
27	729
25	625
23	529
25	625
20	400
$\sum X$ =223	$\sum X^2$ = 5887

$$\overline{X} = \frac{\sum X}{N} = \frac{223}{9} = 24.78$$

$$\sigma = \sqrt{\frac{\sum X^2}{N} - \left[\frac{\sum X}{N}\right]^2} = \sqrt{\frac{5887}{9} - (24.78)^2}$$

$$= \sqrt{654.1111 - 614.0484} = \sqrt{40.06} = 6.33$$

Z = 25

$$Sk_p = \frac{\overline{X} - Z}{\sigma} = \frac{24.78 - 25}{6.33} = \frac{-0.22}{6.33} = -0.0348$$

4. Calculate Karl Pearson's coefficient of Skewness for the following data.

Wage per Item	Number of items			

Rs.(x)	f	fx	x^2	fx^2
12	10	120	144	1440
15	25	375	225	5625
20	40	800	400	16000
25	70	1750	625	43750
30	32	960	900	28800
40	13	520	1600	20800
50	10	500	2500	25000
	$\sum f = 200$	$\sum fx = 5025$		$\sum fX^2 = 141415$

$$\overline{X} = \frac{\sum fX}{\sum f} = \frac{5025}{200} = 25.13$$

$$\sigma = \sqrt{\frac{\sum fX^2}{\sum f} - \left[\frac{\sum fX}{\sum f}\right]^2} = \sqrt{\frac{141415}{200} - (25.13)^2} = \sqrt{707.075 - 631.5169} = \sqrt{75.5581} = 8.69$$

Greatest frequency = 70, Z = 25

$$Sk_p = \frac{\overline{X} - Z}{\sigma} = \frac{25.13 - 25}{8.69} = 0.13/8.69 = 0.0149$$

5. Calculate Karl Pearson's coefficient of Skewness for the following data.

Profit (Rs.Lakhs)	No of Companies f	m	fm	m^2	fm^2
10-20	18	15	270	225	4050
20-30	$20 = f_0$	25	500	625	12500
30-40	$30 = f_1$	35	1050	1225	36750
40-50	$22 = f_2$	45	990	2025	44550
50-60	10	55	550	3025	30250
	$\sum f = 100$		$\sum fm = 3360$		$\sum fm^2 = 128100$

$$\overline{X} = \frac{\sum fm}{\sum f} = 3360/100 = 33.6$$

$$\sigma = \sqrt{\frac{\sum fm^2}{\sum f} - \left[\frac{\sum fm}{\sum f}\right]^2} = \sqrt{\frac{128100}{100} - (33.6)^2} = \sqrt{1281 - 1128.96} = \sqrt{152.04} = 12.33$$

$D_1 = f_1 - f_0 = 30-20 = 10$: $D_2 = f_1 - f_2 = 30-22 = 8$: L = 30: i=10

$$Z = L + \left[\frac{D_1}{D_1 + D_2}\right] i = 30 + \left[\frac{10}{10+8}\right] 10 = 30 + \left[\frac{10}{18}\right] 10 = 30+5.56 = 35.56$$

$$Sk_p = \frac{\overline{X} - Z}{\sigma} = \frac{33.6 - 35.56}{12.33} = -1.96/12.33 = -0.1590$$

6. Calculate Karl Pearson's coefficient of Skewness for the following data.

Weight (lbs)	No of Students f	m	fm	m^2	fm^2	cf
90-100	4	95	380	9025	36100	4
100-110	2	105	210	11025	22050	6
110-120	18	115	2070	13225	238050	24
120-130	22	125	2750	15625	343750	46
130-140	21	135	2835	18225	382725	67
140-150	19	145	2755	21025	399475	86
150-160	10	155	1550	24025	240250	96
160-170	3	165	495	27225	81675	99
170-180	2	175	350	30625	61250	101
	$\sum f = 101$		$\sum fm =$ 13395		$\sum fm^2 =$ 1805325	

$\overline{X} = \frac{\sum fm}{\sum f}$ = 13395/101 = 132.62

$\sigma = \sqrt{\frac{\sum fm^2}{\sum f} - \left[\frac{\sum fm}{\sum f}\right]^2} = = \sqrt{\frac{1805325}{101} - (132.62)^2} = \sqrt{17874.51 - 17588.06}$

$= \sqrt{286.45} = 16.9$

$\frac{\sum f}{2}$ = 101/2 = 50.5, Median Class = 130-140,L = 130,p.c.f =46, f = 21,i = 10

$M = L + \left[\frac{\sum f/2 - p.c.f}{f}\right]i = 130 + \left[\frac{50.5 - 46}{21}\right]10 = 130 + \left[\frac{4.5}{21}\right]10 =$ 130+2.14 = 132.14

$Sk_p = \frac{3(\overline{X} - M)}{\sigma} = \frac{3(132.62 - 132.14)}{16.9} = \frac{3(0.48)}{16.9}$ = 1.44/16.9 = 0.0852

BOWLEY'S COEFFICIENT OF SKEWNESS

7. Compare the Skewness of A and B

	Q_1	M	Q_3
Series A	40	60	80
Series B	62.85	65.25	72.15

Series A

$Sk_B = \frac{Q_3 + Q_1 - 2M}{Q_3 - Q_1} = \frac{80 + 40 - 2(60)}{80 - 40} = \frac{120 - 120}{40} = 0$

Series B

$Sk_B = \frac{Q_3 + Q_1 - 2M}{Q_3 - Q_1} = \frac{72.15 + 62.85 - 2(65.25)}{72.15 - 62.85} = \frac{135 - 130.5}{9.3}$ = 4.5/9.3 = 0.4839

In series A there is no skewness, In Series B there is moderate positive skewness.

8.. Calculate Bowley's coefficient of Skewness.

No of child per family	No of Families	

x	f	cf
0	7	7
1	10	17
2	16	33
3	25	58
4	18	76
5	11	87
6	8	95
	$\sum f = 95$	

Solution:

Position of $Q_1 = \frac{\sum f + 1}{4} = 95+1/4 = 96/4 = 24$

$Q_1 = 2$

Position of $Q_3 = 3(\frac{\sum f + 1}{4}) = 3(24) = 72$

$Q_3 = 4$

Position $M = \frac{\sum f + 1}{2} = 95+1/2 = 96/2 = 48$

$M = 3$

$$Sk_B = \frac{Q_3 + Q_1 - 2M}{Q_3 - Q_1} = \frac{4 + 2 - 2(3)}{4 - 2} = \frac{6 - 6}{2} = 0$$

9. Calculate Bowley's coefficient of Skewness.

Weekly Wages (Rs.)	No of Workers f	cf
Below 200	10	10
200-250	25	35
250-300	145	180
300-350	220	400

350-400	70	470
400 & above	30	500
	$\sum f = 500$	

$\frac{\sum f}{4} = \frac{500}{4} = 125$, Q_1 Class = 250-300, $L_1 = 250$, p.c.$f_1 = 35$, $f_1 = 145$, $i_1 = 50$

$$Q_1 = L_1 + \left[\frac{\sum f/4 - p.c.f_1}{f_1}\right] i_1 = 250 + \left[\frac{125-35}{145}\right]50 = 250 + \left[\frac{90}{145}\right]50$$

$Q_1 = 250 + 31.03 = Rs.281.03$

$3\left(\frac{\sum f}{4}\right) = 3(125) = 375$, Q_3 Class = 300-350, $L_3 = 300$, p.c.$f_3 = 180$, $f_3 = 220$, $i_3 = 50$

$$Q_3 = L_3 + \left[\frac{3(\sum f/4) - p.c.f_3}{f_3}\right] i_3 = 300 + \left[\frac{375-180}{220}\right]50 = 300 + \left[\frac{195}{220}\right]50$$

$Q_3 = 300 + 44.32 = Rs.\ 344.32$

Median:

$\frac{\sum f}{2} = 500/2 = 250$, Median Class = 300-350, L = 300, p.c.f = 180,

f = 220, i = 50

$$M = L + \left[\frac{\sum f/2 - p.c.f}{f}\right] i = 300 + \left[\frac{250-180}{220}\right]50 = 300 + \left[\frac{70}{220}\right]50 = 300 + 15.91$$

M = Rs. 315.91

$$Sk_B = \frac{Q_3 + Q_1 - 2M}{Q_3 - Q_1} = -0.1022$$