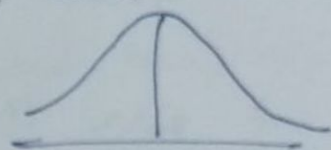


Skewness

(4)

means symmetry or lack of symmetry of data. It is observed from frequency curve.

Draw a line at the value of mode then if we find the curve on either side of the line same equal, then data is symmetric.

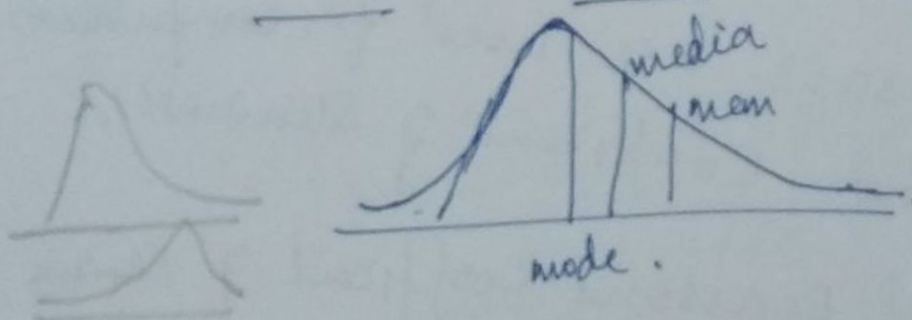


mean = median
= mode.

Positive skewness :-

when tail of the curve of the frequency distribution nose on right and

mean > median > mode.



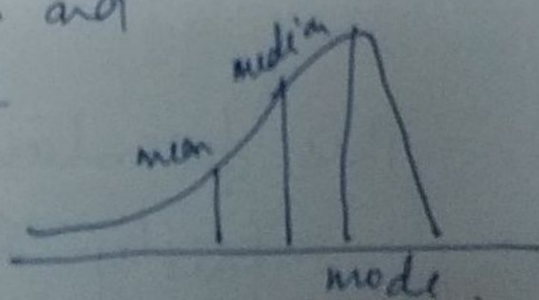
(Right skewed)

Negative skewness

when tail of the curve of frequency distribution elongates more on the left and

mean < median < mode

(left skewed)



Examples

measure of skewness can be both absolute and relative.

Since in symmetrical distribution, mean, median and mode are identical,

→ more the mean moves away from mode, the larger the asymmetry or skewness.

→ An absolute measure cannot be used for comparison because same amount of skewness has different meanings in distribution

$$Sk = \text{mean} - \text{median}$$

$$Sk = \text{mean} - \text{mode}$$

Relative measures are used for comparison and are called coefficient of skewness.

① Based on Karl Pearson's coefficient of skewness

$$L_k = \frac{\text{Mean} - \text{mode}}{\sigma}$$

This formula is used if data is fairly symmetric.

But if the given data is more skewed, ⁽⁵⁾

$$\text{Then } S_k = \frac{3(\text{mean} - \text{median})}{\sigma}$$

$$\text{and mode} = 3\text{median} - 2\text{mean}$$

Ex: For a distribution, Pearson's coefficient of Skewness is 0.64, SD is 13 and mean = 59.2. Find mode and median.

Sol $S_k = 0.64, \sigma = 13, \text{mean} = 59.2$

$$S_k = \frac{\text{mean} - \text{mode}}{\sigma}$$

$$0.64 = \frac{59.2 - \text{mode}}{13}$$

$$\text{mode} = -13(0.64) + 59.2$$

$$\boxed{\text{mode} = 50.88}$$

$$\text{mode} = 3\text{median} - 2\text{mean}$$

$$50.88 = 3\text{median} - 2(59.2)$$

$$\text{median} = \frac{50.88 + 118.4}{3} = \frac{169.28}{3}$$

$$\boxed{\text{median} = 56.42}$$

Ex

$$\text{Coff. of Skewness} = 1.28.$$

$$\text{mean} = 164$$

$$\text{mode} = 100, \text{ find S.D.}$$

$$S_k = \frac{\text{mean} - \text{mode}}{\sigma}$$

$$\sigma = \frac{\text{mean} - \text{mode}}{S_k} = \frac{164 - 100}{1.28} = \underline{\underline{50}}$$

Bowley's Coefficient of Skewness
It is based on quantiles.

$$S_k = \frac{(Q_3 - Q_2) - (Q_2 - Q_1)}{(Q_3 - Q_1)}$$

$$S_k \Rightarrow \frac{Q_3 - 2Q_2 + Q_1}{(Q_3 - Q_1)}$$

Ex: For Bowley's coefficient of skewness is 1.2.
If the sum of 1st & 3rd quartile is 200
median = 76. find third quartile.

$$Q_1 + Q_3 = 200, \quad Q_2 = 76, \quad S_k = 1.2.$$

Sol

$$1.2 = \frac{200 - 2 \times 76}{Q_3 - Q_1} \Rightarrow$$

$$Q_3 - Q_1 = \frac{48}{1.2} = 40 \quad \text{--- (1)}$$

$$Q_1 + Q_3 = 200$$

$$Q_1 = 200 - Q_3 \quad \text{--- (2)}$$

Solving (1) & (2)

$$\underline{\underline{Q_3 = 120}}, \quad \underline{\underline{Q_1 = 80.}}$$

(6)

The following table gives distribution of weight of 100 new born babies. calculate Pearson coefficient of Skewness.

weight	3-5	5-7	7-9	9-11	11-13
No. of babies	10	30	28	18	14

Sol

Class Interval	mid value 100 f	f x.	xf
3-5	10	4	40
5-7	30	6	180
7-9	28	8	224
9-11	18	10	180
11-13	14	12	168

$$\bar{x} = \frac{\sum f_i x_i}{N} = \frac{792}{100} = 7.92.$$

Kurtosis

is a statistical measure that defines how heavily the tails of a distribution differ from those of a normal distribution.

Kurtosis is used to measure financial risk.

→ A large kurtosis is associated with a high level of risk for an investment

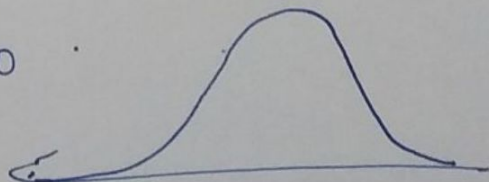
because it indicates that there are high probabilities of extremely large & extremely small returns.

→ A small kurtosis has a moderate level of risk because the probabilities of extreme returns are relatively low.

Kurtosis of normal distribution is 3.

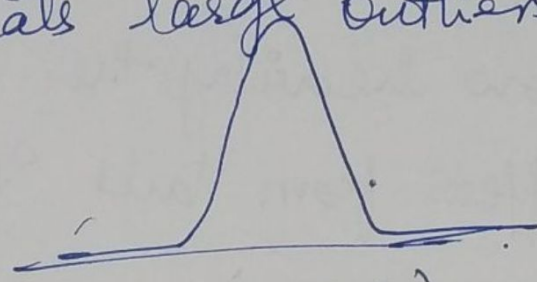
Types of Kurtosis

1) Mesokurtic — Data that shows excess kurtosis is zero.



2) Leptokurtic - shows heavy tails on either side - has positive excess kurtosis. - indicates large outliers.

(has more returns).
Investment returns are riskier



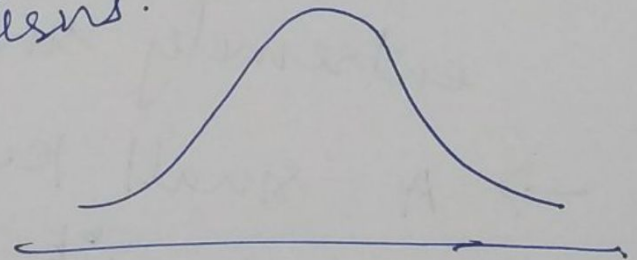
(kurtosis is > 3)

3) platykurtic - It has distribution with flat tails.

The flat tails indicate small outliers in the distribution.

- small portion of investment would have extreme returns.

(kurtosis < 3)



A standard normal distribution has kurtosis of 3 and is recognised as mesokurtic.

$$\rightarrow \text{kurtosis} = n \times \frac{\sum (x_i - \bar{x})^4}{\left[\sum (x_i - \bar{x})^2 \right]^2} = \frac{\text{Fourth moment}}{(\text{Sec. moment})^2}$$

$$\text{Excess kurtosis} = \text{kurtosis} - 3.$$

find the composition of kurtosis
if lept, platy or mesokurtic.

26, 12, 16, 56, 112, 24.

x_i	$x_i - \bar{x}$	$(x - \bar{x})^2$	$(x - \bar{x})^4$
26	26 - 41 = -15	225	50625
12	-29	841	707281
16	-25	625	390625
56	+15	225	50625
112	71	5041	25411681
24	-17	289	83521
$\bar{x} = 41$		7246	4449059.67 26694358

second moment $= (x - \bar{x})^2 = 7246 / 6 = 1207$
 $(x - \bar{x})^4 = 52504516 / 6 = 8750752.67$

$k = \frac{6 \times 52504516}{26694358 \times n} = 6 \times 0.5084$

fourth moment $= 26694358 / 6 = 4449059.67$

kurtosis $= \frac{\text{fourth moment}}{(\text{second moment})^2} = \frac{4449059.67}{(1207)^2} = 3.05$

Since kurtosis of distribution is more than 3, it means Leptokurtic distribution.

Ex. 42, 20, 38, 78, 54, 26

$$\text{kurtosis} = 2.35.$$

Since kurtosis of distribution is < 3 , it is platykurtic distribution.

$$\text{Second moment} = \sum (x - \bar{x})^2$$

$$\text{Fourth moment} = \sum (x - \bar{x})^4$$

Moments about mean

- 1) First calculate mean of the values
 - 2) Subtract this mean from each value
 - 3) Raise their power to 5^{th} , for each of differences.
 - 4) Add the numbers together after step 3
 - 5) Finally divide the sum by no. of values.
- $$\text{Moment } M_5 = \frac{(x_1 - \bar{x})^5 + (x_2 - \bar{x})^5 + (x_3 - \bar{x})^5 + \dots + (x_n - \bar{x})^5}{n}.$$