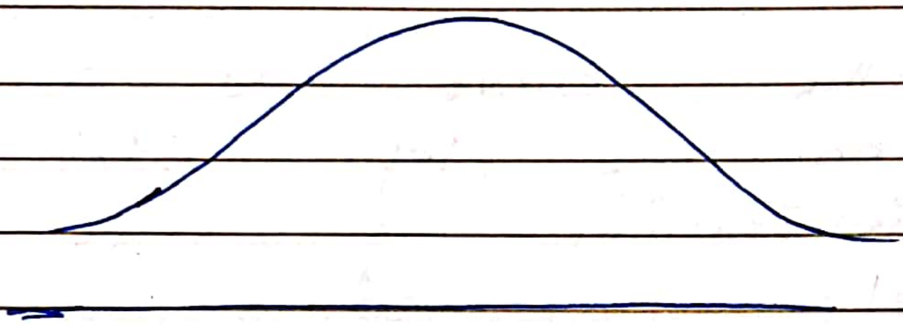
Shape Measures

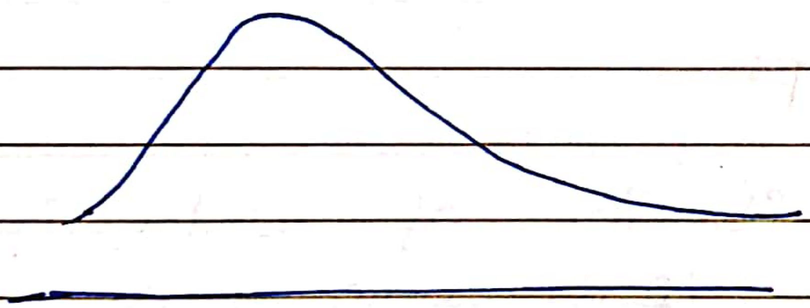
- Skewness
  - Mean, Median & Mode Method.
  - Moment Method.
- kurtosis
  - Moment Method.

Skewness:

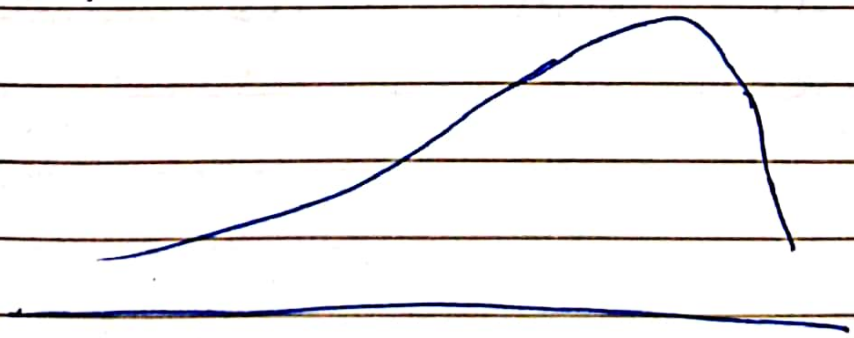
If the distribution of data is not symmetrical then it will be skewed. So, the lack of symmetry ~~is~~ in a dataset is called skewness.



Symmetrical distribution of data.



Positively skewed distribution.



Negatively skewed distribution

⇒ Positive Skewness:

If the curve has a longer tail on the right side then the distribution is called positively skewed.

⇒ Negative Skewness:

If the curve has a longer tail on the left side then the distribution is called negatively skewed.

⇒ Symmetry:

A distribution of data is called symmetrical if its curve has same shape on both sides from central line.

⟹ Mean, Median & Mode Method:

→ If Mean = Median = Mode then it is symmetrical distribution.

→ If Mean > Median > Mode then it is positively skewed distribution.

→ If Mean < Median < Mode then it is negatively skewed distribution.

→ Example :

For any distribution the mean is 45, the median is 30 & the mode is zero. is the distribution symmetrical, positively skewed or negatively skewed?

Mean = 45
Median = 30
Mode = 0

Here

Mean > Median > Mode.

45 > 30 > 0

So, the distribution is posively skewed.

⇒ Example:

For any dataset the mean is 2600, Median is 2800 & the Mode is 3000. Find shape of the dataset.

Mean = 2600

Median = 2800

Mode = 3000

Here

$2600 < 2800 < 3000$

Mean < Median < Mode.

So, dataset is Negatively skewed.

→ Example:

For weekly income data set, the mean is \$700, Median is \$500 & the Mode is \$300. Is the dist. symmetrical, positively skewed or negatively skewed?

Mean = \$700

Median = \$500

Mode = \$300

Here

Mean > Median > Mode

\$700 > \$500 > \$300

So it is positively skewed distribution

⇒ Example:

The weekly sales from a sample of Hi-Tec electronic supply stores were organized into frequency distribution. The mean of weekly sales was computed to be \$105,900, the median \$105,000 & the mode \$104,500.

Mean = \$105,900

Median = \$105,000

Mode = \$104,500

Here

Mean > Median > Mode

\$105,900 > \$105,000 > \$104,500

So it is positively skewed.

⟹ Example :
if a dataset have ;
Mean = 20

Median = 20

Mode = 20

Here
Mean = Median = Mode

So
it is symmetricaly distributted dataset.

→ Example :

If a distribution has mean 1403, median 1460 & Mode 1487. What can you say about the skewness.

Mean = 1403
Median = 1460
Mode = 1487

The distribution is negatively skewed because,

Mean < Median < Mode

1403 < 1460 < 1487.

⇒ Example :

What can you say about skewness if ,

Mean = 140
Median = 145
Mode = 148.5.

The distribution is negatively skewed because,

Mean < Median < Mode.

140 < 145 < 148.5.

⟹ Example:

The mean, median & mode of the weekly income of women workers from a locality are 3133.33, 3020.01, 2804.35 respectively. Calculate skewness of the dist.

Mean = 3133.33

Median = 3020.01

Mode = 2804.35

The distribution is positively skewed because;

Mean > Median > Mode.

3133.33 > 3020.01 > 2804.35.

⇒ Moments:

→ The moments are defined as the average of different powers of deviations of observations from their mean.

⇒ Merits of Moments:

1. Moments tell us about the mean, variance, skewness, kurtosis of dataset.

2. From moments, we can find shape of any dataset.

Moments

- Moments about mean (central Moments)
- Non-central Moments.

⟹ Central Moments:

$$\mu_r = \frac{\Sigma(x-\mu)^r}{N}$$

First Four moments about mean | central moments are;

$$\mu_1 = \frac{\Sigma(x-\mu)^1}{N}$$

$$\mu_2 = \frac{\Sigma(x-\mu)^2}{N}$$

$$\mu_3 = \frac{\Sigma(x-\mu)^3}{N}$$

$$\mu_4 = \frac{\Sigma(x-\mu)^4}{N}.$$

→ Skewness (Moment Method):

The numerical measure that is used to know about the symmetry & skewness of dataset is $\sqrt{\beta_1}$

$$\sqrt{\beta_1} = \frac{\mu_3}{\sqrt{\mu_2^3}}$$

If $\sqrt{\beta_1} = 0$ → Symmetrical

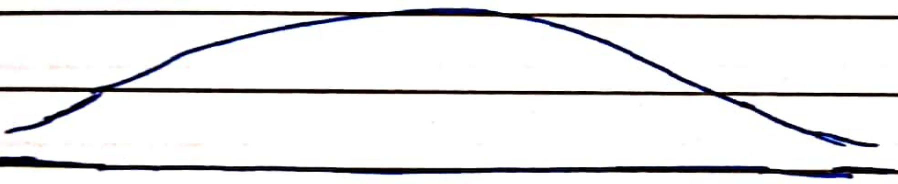
If $\sqrt{\beta_1} > 0$ → Positively Skewed

If $\sqrt{\beta_1} < 0$ → Negatively Skewed.

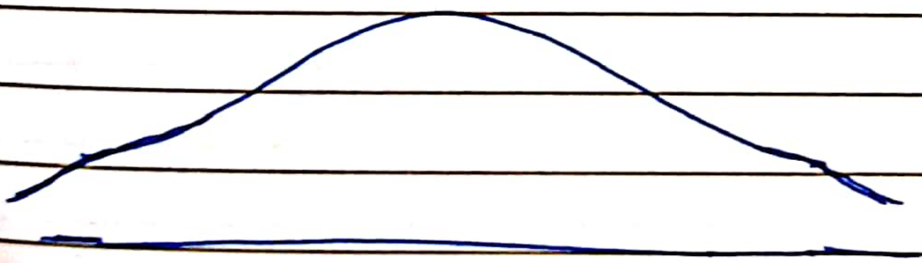
→ Kurtosis :

Kurtosis is defined as the degree of peakness or flatness.

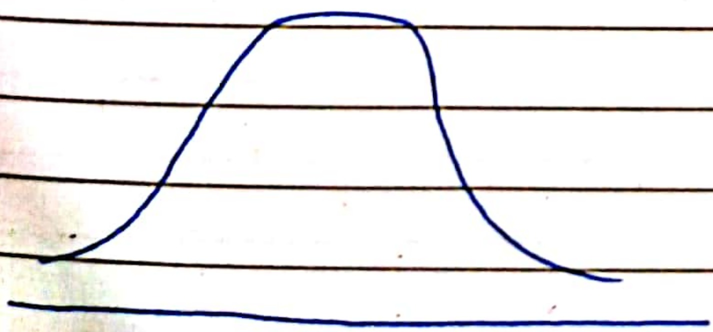
It indicates the length of the peakedness of symmetrical distribution. Symmetrical distribution may be platykurtic, mesokurtic or leptokurtic i.e,



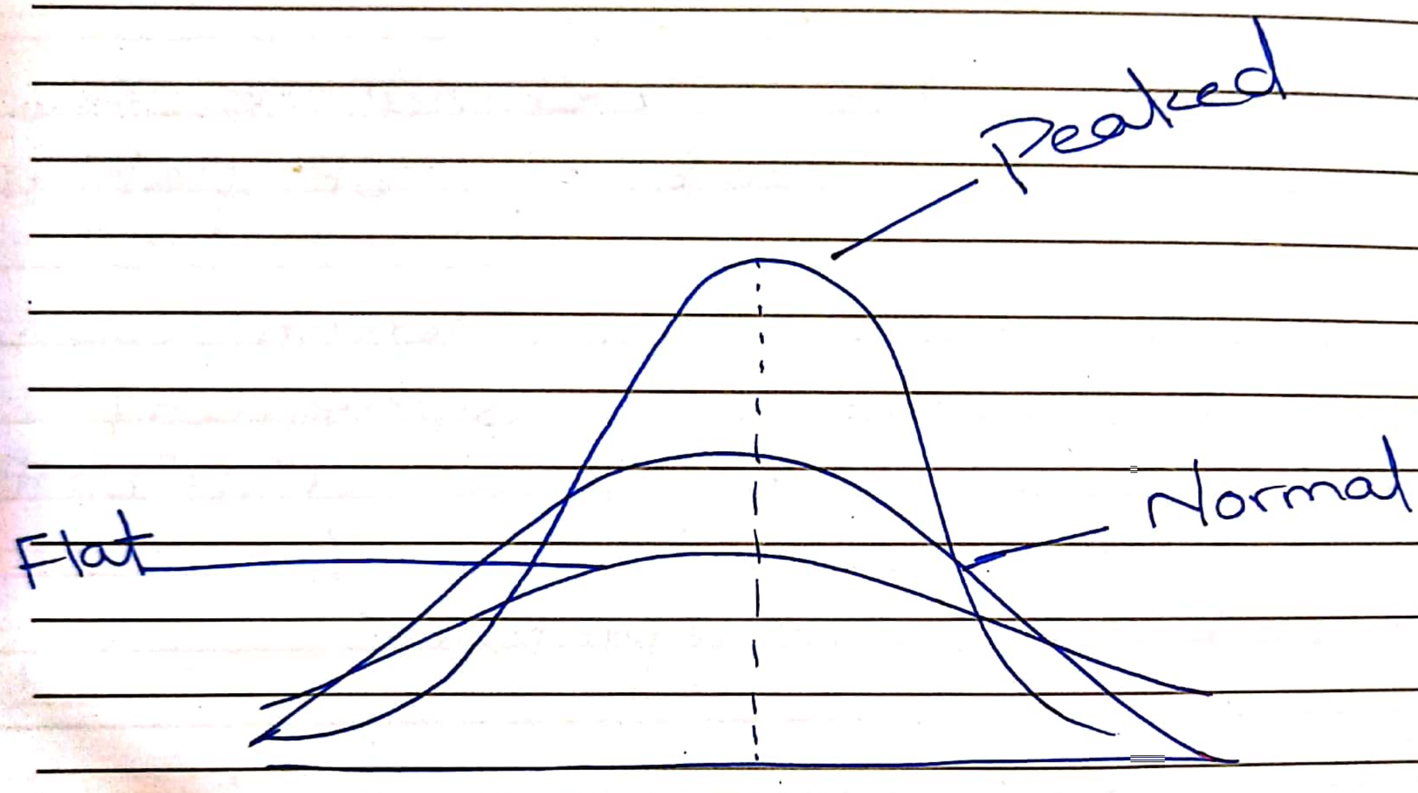
Platykurtic (flat)



Mesokurtic (Normal)



Leptokurtic (Peaked)


Peaked
Normal
Flat

→ Kurtosis (Moment Method):

The measure of kurtosis based on moments is β_2

$$\beta_2 = \frac{\mu_4}{\mu_2^2}$$

For $\beta_2 = 3$ ⟶ Mesokurtic

For $\beta_2 < 3$ ⟶ Platykurtic

For $\beta_2 > 3$ ⟶ Leptokurtic

⟹ Example:

calculate first four moments about mean for the following set of marks obtained in the examination

45, 32, 37, 46, 39, 36, 41, 48 & 36.
Also calculate $\sqrt{\beta_1}$

$$\mu = \frac{\Sigma X}{N}$$

$$= \frac{45+32+37+46+39+36+41+48+36}{9}$$

$$\boxed{\mu = 40}$$

we know that:

$$\mu_1 = \frac{\Sigma (x-\mu)}{N}$$

$$\mu_2 = \frac{\Sigma (x-\mu)^2}{N}$$

$$\mu_3 = \frac{\Sigma (X-\mu)^3}{N}$$

$$\mu_4 = \frac{\Sigma (X-\mu)^4}{N}$$

X	X − M	$(X-M)^2$	$(X-M)^3$	$(X-M)^4$
45	5	25	125	625
32	−8	64	−512	4096
37	−3	9	−27	81
46	6	36	216	1296
39	−1	1	−1	1
36	−4	16	−64	256
41	1	1	1	1
48	8	64	512	4096
36	−4	16	−64	256
	0	232	186	10708

$$\mu_1 = \frac{0}{9} = \boxed{0}$$

$$\mu_2 = \frac{232}{9} = \boxed{25.78}$$

$$\mu_3 = \frac{186}{9} = \boxed{20.67}$$

$$\mu_4 = \frac{10708}{9} = \boxed{1189.78}$$

$$\sqrt{\beta_1} = \frac{\mu_3}{\sqrt{\mu_2^3}}$$

$$= \frac{20.67}{\sqrt{(25.78)^3}}$$

$$= \frac{20.67}{\sqrt{17133.60}}$$

$$= \frac{20.67}{130.8954}$$

$$\sqrt{\beta_1} = 0.16 > 0$$

Positively Skewed.

→ Example :

First four central moments are

0, 43.7220, 43.7407, 4363.0296

Find $\sqrt{\beta_1}$.

Here;

$\mu_1 = 0$

$\mu_2 = 43.7220$

$\mu_3 = 43.7407$

$\mu_4 = 4363.0296$.

So;

$$\sqrt{\beta_1} = \frac{\mu_3}{\sqrt{\mu_2^3}}$$

Scanned with CamScanner

$$\sqrt{\beta_1} = \frac{43.7407}{\sqrt{(43.7220)^3}}$$

$$= \frac{43.7407}{\sqrt{83,579.556}}$$

$$= \frac{43.7407}{289.1013}$$

$$\sqrt{\beta_1} = 0.1513$$

As $\sqrt{\beta_1} > 0$

So, the given distribution is positively skewed.

⟹ Example:

First four central moments are;

0 , 43.4988 , 17.3354 , 4131.1478

Find $\sqrt{\beta_1}$:

Here;

$\mu_1 = 0$

$\mu_2 = 43.4988$

$\mu_3 = 17.3354$

$\mu_4 = 4131.1478$

We know that;

$$\sqrt{\beta_1} = \frac{\mu_3}{\sqrt{\mu_2^3}}$$

$$\sqrt{\beta_1} = \frac{17.3354}{\sqrt{(43.4988)^3}}$$

$$= \frac{17.3354}{\sqrt{82,306.0631}}$$

$$= \frac{17.3354}{286.89}$$

$$\sqrt{\beta_1} = 0.0604 > 0$$

Positively Skewed.

⟹ Example:

First four moments about mean are 0, 11, 49, 192.

Find $\sqrt{\beta_1}$

$\mu_1 = 0$

$\mu_2 = 11$

$\mu_3 = 49$

$\mu_4 = 192.$

$$\sqrt{\beta_1} = \frac{\mu_3}{\sqrt{\mu_2^3}} = \frac{49}{\sqrt{(11)^3}}$$

$\sqrt{\beta_1} = 1.34$

$\sqrt{\beta_1} > 0$

Positively Skewed.

→ Example:

First three moments about mean are; 0, 2.78, −1.03 what can you say about skewness

$$\mu_1 = 0$$

$$\mu_2 = 2.78$$

$$\mu_3 = -1.03$$

$$\sqrt{\beta_1} = \frac{\mu_3}{\sqrt{\mu_2^3}}$$

$$= \frac{-1.03}{\sqrt{(2.78)^3}}$$

$$\sqrt{\beta_1} = -0.22 < 0$$

Negatively Skewed

⟹ Example:

First three moments about mean are 0, 3 & 0. What can you say about skewness. Is the distribution symmetrical, positively skewed or negatively skewed?

$\mu_1 = 0$

$\mu_2 = 3$

$\mu_3 = 0$

$$\sqrt{\beta_1} = \frac{\mu_3}{\sqrt{\mu_2^3}}$$

$$\sqrt{\beta_1} = \frac{0}{\sqrt{3^3}} = 0$$

The distribution is symmetrical.

⇒ Example:

First four moments about mean are 0, 47, −105, 5621.

Find β_2.

$\mu_1 = 0$

$\mu_2 = 47$

$\mu_3 = -105$

$\mu_4 = 5621$

$$\beta_2 = \frac{\mu_4}{\mu_2^2}$$

$$= \frac{5621}{(47)^2}$$

$\beta_2 = 2.54 < 3$

platy-kurtic.

→ Example:

First four moments about mean are, 0, 45.81, 43.74 & 4917.37. Find β_2.

$\mu_1 = 0$

$\mu_2 = 45.81$

$\mu_3 = 43.74$

$\mu_4 = 4917.37$

$$\beta_2 = \frac{\mu_4}{(\mu_2)^2}$$

$$= \frac{4917.37}{(45.81)^2} = 2.34 < 3$$

platykurtic.

→ Example:

For a dataset the second and fourth moments about mean are 2.6364 & 28.30256. Find β_2.

$\mu_2 = 2.6364$

$\mu_4 = 28.30256.$

$$\beta_2 = \frac{\mu_4}{\mu_2^2}$$

$$= \frac{28.30256}{(2.6364)^2}$$

$\beta_2 = 4.07 \quad > 3$

Leptokurtic

⇒ First four moments about mean are 0, 73.91, -21.546 & 12110.94. Find out whether the distribution is leptokurtic or platykurtic.

$\mu_1 = 0$

$\mu_2 = 73.91$

$\mu_3 = -21.546$

$\mu_4 = 12110.94$

$$\beta_2 = \frac{\mu_4}{\mu_2^2} = \frac{12110.94}{(73.91)^2}$$

$$\boxed{\beta_2 = 2.22}$$

Since $\beta_2 < 3$

The distribution is platykurtic.

→ Example:

second moment about mean of two datasets are 9 & 16 while fourth moment about mean are 230 & 780 respectively. Find out which of the distribution is

i) platykurtic.

ii) leptokurtic.

Dataset 1	Dataset 2
$\mu_2 = 9$	$\mu_2 = 16$
$\mu_4 = 230$	$\mu_4 = 780$
$\beta_2 = \frac{\mu_4}{\mu_2^2}$	$\beta_2 = \frac{\mu_4}{\mu_2^2}$

Scanned with CamScanner

$$\beta_2 = \frac{230}{(9)^2}$$

$$\beta_2 = 2.84$$

Since;

$\beta_2 < 3$

Therefore the dist is platykurtic

$$\beta_2 = \frac{780}{(16)^2}$$

$$\beta_2 = 3.05$$

Since,

$\beta_2 > 3$

Therefore the dist. is Leptokurtic

⟹ Example :

The second moment about mean of two datasets are 13.76 and 63.0 while the fourth moments about the mean are 528.06 & 9500 respectively. which of the dist. is

a) Platykurtic.

b) Leptokurtic.

c) Mesokurtic.

Dataset 1 :

$\mu_2 = 13.76$

$\mu_4 = 528.06$

$$\beta_2 = \frac{\mu_4}{\mu_2^2}$$

$$\beta_2 = \frac{528.06}{(13.76)^2}$$

$$\beta_2 = 2.79$$

$$\beta_2 < 3$$

Platykurtic.

Dataset 2 :

$$\mu_2 = 63.0$$

$$\mu_4 = 9500$$

$$\beta_2 = \frac{\mu_4}{\mu_2^2}$$

$$= \frac{9500}{(63)^2}$$

$$\beta_2 = 2.39 < 3$$

Platykurtic.

→ Example:

For a dataset

$\mu_2 = 9$ &

$\mu_4 = 243$.

Find β_2:

$$\beta_2 = \frac{\mu_4}{\mu_2^2}$$

$$= \frac{243}{(9)^2}$$

$$= \frac{243}{81}$$

$\beta_2 = 3 \Rightarrow$ Mesokurtic