MOMENTS (μ)

Moments about Mean

If $x_1, x_2, \ldots, x_n$ are values of a variable x with corresponding frequencies $f_1, f_2, \ldots, f_n$ respectively then r^{th} moment about mean $\bar{x}$ is denoted by μ_r and defined as

Ungrouped Data: $$\mu_r = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^r \quad ; \quad r = 0, 1, 2, \ldots$$

Grouped Data: $$\mu_r = \frac{1}{N}\sum_{i=1}^{n} f_i (x_i - \bar{x})^r \quad ; \quad r = 0, 1, 2, \ldots$$

$$N = \sum_{i=1}^{n} f_i$$

* The second moment about the mean coincides with the variance.

i.e. μ_2 = Variance

* Moments are used to describe the the various characteristics of frequency distribution

Moments about arbitrary point

If $x_1, x_2, \dots, x_n$ are the values of a variable x with the corresponding frequencies $f_1, f_2, \dots, f_n$ the rth moment about any point A is denoted by μ'_r and is defined as

Ungrouped Data :

$$\mu'_r = \frac{1}{n} \sum_{i=1}^{n} (x_i - A)^r \quad ; \quad r = 0, 1, 2, \dots$$

Grouped Data :

$$\mu'_r = \frac{1}{N} \sum_{i=1}^{n} f_i (x_i - A)^r \quad ; \quad r = 0, 1, 2, \dots$$

where $N = \sum_{i=1}^{n} f_i$

Relation betn μ_r & μ'_r

$$\mu_r = \mu'_r - {}^{r}c_1 \mu'_{r-1} \mu'_1 + {}^{r}c_2 \mu'_{r-2} \mu_1'^{2} + \dots + (-1)^r \mu_1'^{r}$$

Important

$$\boxed{\begin{aligned} \mu_1 &= 0 \\ \mu_2 &= \mu'_2 - \mu_1'^{2} \\ \mu_3 &= \mu'_3 - 3\mu'_2 \mu'_1 + 2(\mu'_1)^3 \\ \mu_4 &= \mu'_4 - 4\mu'_3 \mu'_1 + 6\mu'_2 \mu_1'^{2} - 3(\mu'_1)^4 \end{aligned}}$$

Moments about the origin

If $x_1, x_2, \ldots, x_n$ are the values of the variable x with the corresponding frequencies $f_1, f_2, \ldots, f_n$ respectively. the rth moment about the origin is denoted by v_r and is defined as

$$v_r = \frac{1}{N} \sum_{i=1}^{n} f_i x_i^r \quad ; \quad r = 0, 1, 2, \ldots$$

$$\text{where} \quad N = \sum_{i=1}^{n} f_i$$

Relation between μ_r and v_r

$$v_r = \mu_r + {}^rc_1 \mu_{r-1} \bar{x} + {}^rc_2 \mu_{r-2} \bar{x}^2 + \ldots + \bar{x}^r$$

$$\boxed{\begin{aligned} v_1 &= \bar{x} = A + \mu_1' \quad ; \ A = \text{assumed mean} \\ v_2 &= \mu_2 + \bar{x}^2 \\ v_3 &= \mu_3 + 3v_2\bar{x} - 2\bar{x}^3 \\ v_4 &= \mu_4 + 4v_1v_3 - 6v_1^2v_2 + 3v_1^4 \end{aligned}}$$

* Sheppard's Correction for Moments

$$\mu_2 \text{ (corrected)} = \mu_2 - \frac{h^2}{12}$$

$$\mu_4 \text{ (corrected)} = \mu_4 - \frac{h^2 \mu_2}{2} + \frac{7h^4}{240}$$

where h is width of the class interval

Note: μ_1 & μ_3 require no correction.

* These corrections are required for grouped data as in grouped data we assume that observations are concentrated at the mid-point of the class intervals.

Q. Calculate the first four moments about the mean of the following data

x	0	1	2	3	4	5	6	7	8
$f(x)$	1	8	28	56	70	56	28	8	1

Soln: let us calculate moment about $x = 4$

$$\mu_r' = \frac{1}{N} \Sigma f_i (x_i - 4)^r = \frac{1}{N} \Sigma f_i d_i^r \quad \text{where } d_i = x_i - 4$$

x	f	$d = x-4$	fd	fd^2	fd^3	fd^4
0	1	-4	-4	16	-64	256
1	8	-3	-24	72	-216	648
2	28	-2	-56	112	-224	448
3	56	-1	-56	56	-56	56
4	70	0	0	0	0	0
5	56	1	56	56	56	56
6	28	2	56	112	224	448
7	8	3	24	72	216	648
8	1	4	4	16	64	256

$\Sigma f = 256$ $\quad \Sigma d = 0$ $\quad \Sigma fd = 0$ $\quad \Sigma fd^2 = 512$ $\quad \Sigma fd^3 = 0$ $\quad \Sigma fd^4 = 2816$

$$N = \Sigma f = 256$$

$$\mu_1' = \frac{1}{N} \Sigma fd = 0$$

$$\mu_2' = \frac{1}{N} \Sigma fd^2 = \frac{512}{256} = 2$$

$$\mu_3' = \frac{1}{N} \Sigma fd^3 = 0$$

$$\mu_4' = \frac{1}{N} \Sigma fd^4 = \frac{2816}{256} = 11$$

Moments about mean,

$$\mu_1 = 0$$

$$\mu_2 = \mu_2' - \mu_1'^2 = 2$$

$$\mu_3 = \mu_3' - 3\mu_2'\mu_1' + 2(\mu_1')^3 = 0$$

$$\mu_4 = \mu_4' - 4\mu_3'\mu_1' + 6\mu_2'\mu_1'^2 - 3\mu_1'^4 = 11$$

Q. From the following data, calculate moments about

(i) assumed mean 25 arbitrary

(ii) actual mean

(iii) moments about zero origin

Note : Effect of change of orgin and scale on moments

Assume $d_i = \dfrac{x_i - A}{h}$ $\Rightarrow x_i = A + d_i h$

$\mu_r' = h^r \cdot \frac{1}{N}\sum_i f_i d_i$ and $\mu_r = h^r \cdot \frac{1}{N}\sum f_i (d_i - \bar{d})^r$

where $\bar{x} = A + h\bar{d}$

Variable	Mid point (x)	Freq (f)	$d = \frac{x-25}{10}$	fd	fd^2	fd^3	fd^4
0–10	5	1	−2	−2	4	−8	16
10–20	15	3	−1	−3	3	−3	3
20–30	25	4	0	0	0	0	0
30–40	35	2	1	2	2	2	2
		$N=\Sigma f=10$		$\Sigma fd=-3$	$\Sigma fd^2=9$	$\Sigma fd^3=-9$	$\Sigma fd^4=21$

$\boxed{h = 10}$

① Moments about assumed mean 25,

$$\mu_1' = h \times \frac{1}{N} \Sigma fd = 10 \times \frac{-3}{10} = -3$$

$$\mu_2' = h^2 \times \frac{1}{N} \Sigma fd^2 = (10)^2 \times \frac{9}{10} = 90$$

$$\mu_3' = h^3 \times \frac{1}{N} \Sigma fd^3 = (10)^3 \times \frac{(-9)}{100} = -900$$

$$\mu_4' = h^4 \times \frac{1}{N} \Sigma fd^4 = (10)^4 \times \frac{21}{10} = 21000$$

(ii) Moments about mean,

$$\mu_1 = 0$$

$$\mu_2 = \mu_2' - \mu_1'^2 = 90 - (-3)^2 = 81$$

$$\mu_3 = \mu_3' - 3\mu_2' \mu_1' + 2\mu_1'^3$$

$$= -900 - 3((-3)(900)) + 2(-3)^3$$

$$= -144$$

$$\mu_4 = \mu_4' - 4\mu_3' \mu_1' + 6\mu_2' \mu_1'^2 - 3(\mu_1')^4$$

$$= 2100 - 4(-900)(-3) + 6(90)(-3)^2 - 3(-3)^4$$

$$= 14817$$

(iii) Moments about zero i.e. origin

$$v_1 = \bar{x} = A + \mu_1' = 25 - 3 = 22$$

$$v_2 = \mu_2 + (v_1)^2 = 81 + (22)^2 = 565$$

$$v_3 = \mu_3 + 3v_1 v_2 - 2v_1^3 = -144 + 3(22)(565) - 2(22)^3 = 15650$$

$$v_4 = \mu_4 + 4v_1 v_3 - 6v_1^2 v_2 + 3v_1^4 = \text{calculate ?}$$

Q. The first four moments of a distribution about $x = 4$ are 1, 4, 10 and 45. Determine the values of the mean and variance

Soln: Given,

$$A = 4$$

$$\mu_1' = 1$$

$$\mu_2' = 4$$

$$\mu_3' = 10$$

$$\mu_4' = 45$$

We know that,

$$\bar{x} = A + \mu_1'$$

$$\Rightarrow \bar{x} = 4 + 1 = 5$$

$$\Rightarrow \text{Mean} = 5$$

Again,

$$\text{Variance} = \mu_2 = \mu_2' - (\mu_1')^2 = 4 - (1)^2 = 3$$

9. The first four moments about the mean of a frequency distribution with class size 3 are $\mu_1 = 0$, $\mu_2 = 8.7579$, $\mu_3 = -2.924$, $\mu_4 = 281.453$. Use Sheppard's formula to find the correct moments.

Soln: Given,

$$\mu_1 = 0$$

$$\mu_2 = 8.7579$$

$$\mu_3 = -2.924$$

$$\mu_4 = 281.453$$

$$h = 3$$

Using Sheppard's Correction

$$\mu_1 (\text{corrected}) = \mu_1 = 0$$

$$\mu_2 (\text{corrected}) = \mu_2 - \frac{h^2}{12} = 8.7579 - \frac{3^2}{12}$$

$$= 8.0079$$

$\mu_3 \text{(corrected)} = \mu_3 = -2.924$

$$\mu_4 \text{(corrected)} = \mu_4 - \frac{h^2 \mu_2}{2} + \frac{7h^4}{240}$$

$$= 281.453 - \frac{(3)^2 \times 8.0079}{2} + \frac{7(3)^4}{240}$$

$$= 244.40495$$

Q1. The first four moments of a distribution about the value 4 of the variable are $-15, 17, -30$ and 108. Find the moments about the mean.

Q2. If the first four moments of a distribution about the value 5 of the variable are $-4, 22, -117$ and 560. Find the moments about the origin.

Q. Find the first four moments from the following data

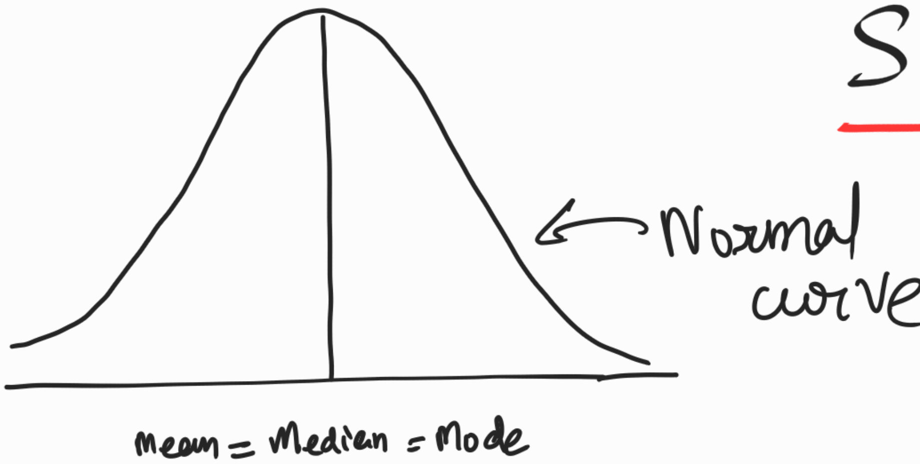
(1)

Marks	0-10	10-20	20-30	30-40	40-50	50-60	60-70
No. of students	8	12	20	30	15	10	5

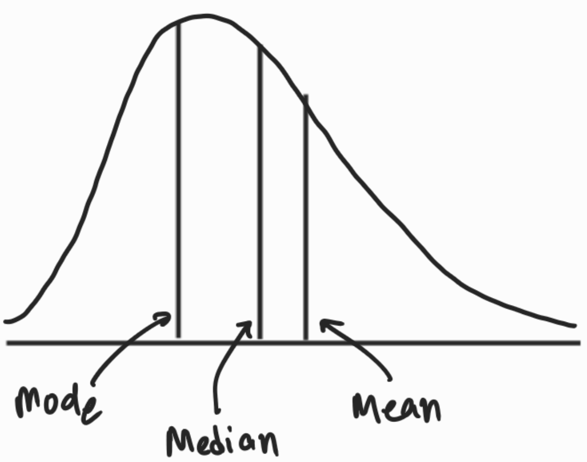
(2)

Mid-value	5	10	15	20	25	30	35
Frequency	8	15	20	32	23	17	5

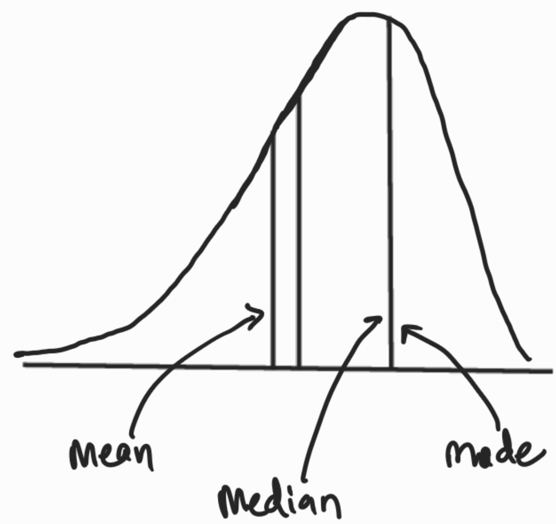
Skewness



Symmetrical Distribution



Positively skewed Distribution



Negatively skewed distribution

Defⁿ: Skewness of a distribution can be defined as the tendency of a distribution to depart from symmetry.

Measures

* Karl Pearson's method

$S.D = \sqrt{var(X)}$

$var(X) = \mu_2$

$= E(X^2) - [E(X)]^2$

Karl Pearson's coefficient of skewness

$$* \; S_K = \frac{\text{Mean} - \text{Mode}}{S.D.}$$

$$= \frac{3(\text{Mean} - \text{Median})}{S.D.}$$

(using mode = 3 median − 2 mean)

* value of Karl Pearson's coefficient lies between -3 & 3.

Method of Moments

$$\beta_1 = \frac{\mu_3^2}{\mu_2^3} = \frac{(\mu_3)^2}{(\mu_2)^3}$$

$\leftarrow$ (3rd moment about mean)^2

Relative measure of skewness

Q. Karl Pearson's coefficient of skewness of a distribution is 0.32, its standard deviation is 6.5 and mean is 29.6. Find the mode of the distribution.

$$\text{Coefficient of skewness} = \frac{\text{Mean} - \text{Mode}}{\text{S.D.}}$$

$$\Rightarrow 0.32 = \frac{29.6 - \text{Mode}}{6.5}$$

$$\Rightarrow \text{Mode} = 29.6 - (0.32 \times 6.5)$$

$$= 27.52$$

Q. Compute the coefficient of skewness from the following data

25, 15, 23, 40, 27, 25, 23, 25, 20

Soln: Mean, $E[x] = \frac{25+15+23+40+27+25+23+25+20}{9}$

$$= \frac{223}{9}$$

$$= 24.78$$

Mode $= 25$ (data having highest freq)

$$E[x^2] = \frac{(25)^2+(15)^2+(23)^2+(40)^2+(27)^2+(25)^2+(23)^2+(25)^2+(20)^2}{9}$$

$$= \frac{625+225+529+1600+729+625+529+625+400}{9}$$

$$= \frac{5887}{9}$$

$$= 654.11$$

S.D., $\sigma = \sqrt{E[x^2]-(E[x])^2} = \sqrt{654.11 - 614.05}$

$$= \sqrt{40.06}$$

$$= 6.33$$

$$\text{Coefficient of Skewness} = \frac{\text{Mean} - \text{Mode}}{\sigma}$$

$$= \frac{24.78 - 25}{6.33}$$

$$= -0.03$$

Q. For the following distribution, find the first four moments about the mean. Also find the value of β_1. Is it symmetrical distribution?

Note: If $\beta_1 = 0$ then distribution is symmetrical

Q. Calculate the Karl Pearson's coefficient of skewness from the data given below

Income	400-500	500-600	600-700	700-800	800-900
Expenditure	8	16	20	17	3

KURTOSIS

Kurtosis measures the extent to which the curve is more peaked or more flat-topped than the normal curve.

If the curve concentrates
- too much on centre → leptokurtic
- less on centre → Platykurtic

If the curve is normal curve then Kurtosis is equal to zero and known as Mesokurtic.

Kurtosis is measured by the coefficient β_2 given by

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

where $\mu_4 \rightarrow 4^{th}$ moment

$\mu_2 \rightarrow 2^{nd}$ moment

* Greater the value of β_2, the more peaked the dist.

Fisher's measure of Kurtosis is expressed in terms of

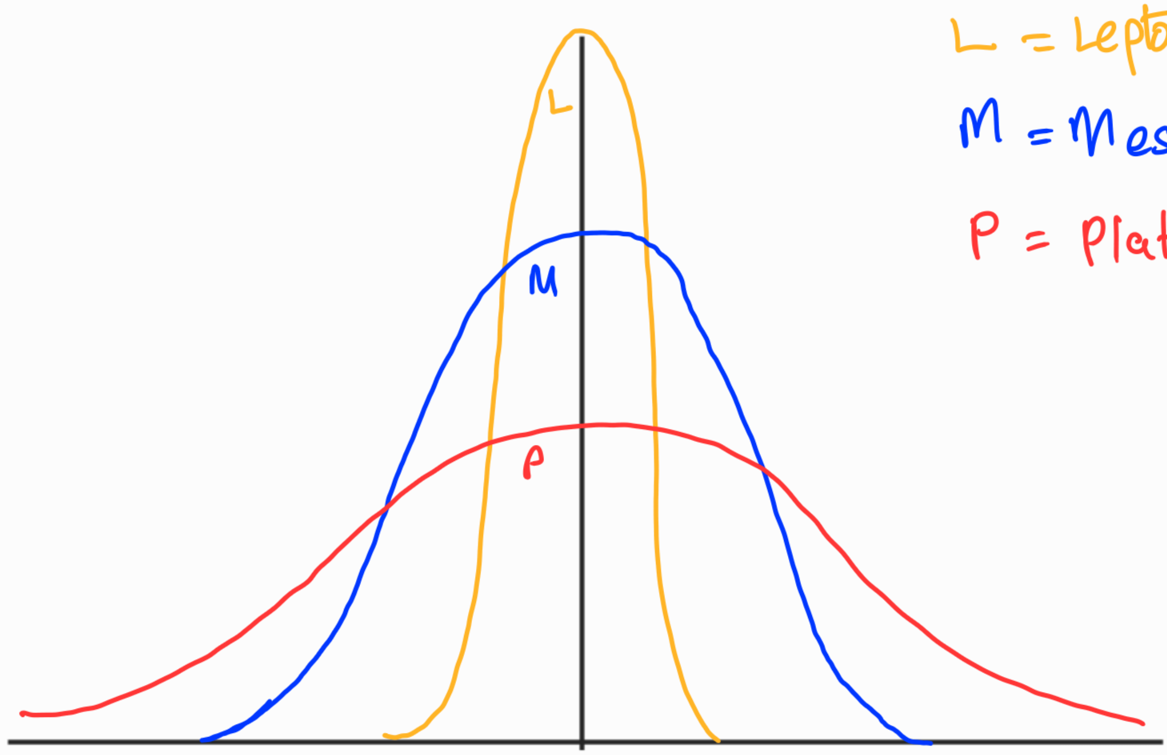
$$\boxed{\begin{aligned} \gamma_2 &= \beta_2 - 3 \\ \gamma_1 &= +\sqrt{\beta_1} \end{aligned}}$$

$\gamma_2 = 0$ i.e. $\beta_2 = 3$ ← normal curve | mesokurtic

$\gamma_2 < 0$ ie $\beta_2 < 3$ ← Platykurtic

$\gamma_2 > 0$ i.e. $\beta_2 > 3$ ← Leptokurtic

$\gamma_1 = 0$ ← curve is symmetrical



L = Leptokurtic
M = Mesokurtic
P = Platykurtic

Q. The first four moments of a distribution about the value 4 of the variable are -1.5, 17, -30 and 108. Discuss the Kurtosis of the distribution.

$$\mu_1' = -1.5$$

$$\mu_2' = 17$$

$$\mu_3' = -30$$

$$\mu_4' = 108$$

Now,

$$\mu_2 = \mu_2' - (\mu_1')^2 = 17 - (-1.5)^2 = 14.75$$

$$\mu_4 = \mu_4' - 4\mu_1'\mu_3' + 6\mu_2'\mu_1'^2 - 3(\mu_1')^4$$

$$= 108 - 4(-1.5)(-30) + 6(17)(-1.5)^2 - 3(-1.5)^4$$

$$= 142.31$$

$$\therefore \beta_2 = \frac{\mu_4}{\mu_2^2} = \frac{142.31}{(14.75)^2} = 0.654 < 3$$

$\Rightarrow$ Distribution is platykurtic.

Q. The following data are given to an economist for the purpose of economic analysis. The data refers to the length of a certain type of batteries

$n = 100$, $\sum fd = 50$, $\sum fd^2 = 1970$, $\sum fd^3 = 2948$

$\sum fd^4 = 86752$ in which $d = x - 48$. Do you think that this distribution is platykurtic?

For platykurtic, $\beta_2 < 3$

Now,

$$\mu_1' = \frac{\sum fd}{N} = \frac{50}{100} = 0.5$$

$$\mu_2' = \frac{\sum fd^2}{N} = \frac{1970}{100} = 19.70$$

$$\mu_3' = \frac{\sum fd^3}{N} = \frac{2948}{100} = 29.48$$

$$\mu_4' = \frac{\sum f d^4}{N} = \frac{86752}{100} = 867.52$$

Again,

$$\mu_2 = \mu_2' - (\mu_1')^2 = 19.70 - (0.5)^2 = 19.45$$

$$\mu_3 = \mu_3' - 3\mu_1'\mu_2' + 2\mu_1'^3 = 29.48 - 3(0.5)(29.48) + 2(0.5)^3$$

$$= -14.49$$

$$\mu_4 = \mu_4' - 4\mu_1'\mu_3' + 6\mu_2'\mu_1'^2 - 3(\mu_1')^4$$

$$= 867.52 - 4(0.5)(29.48) + 6(19.70)(0.5)^2 - 3(0.5)^4$$

$$= 837.92$$

Now,

$$\beta_2 = \frac{\mu_4}{\mu_2^2} = \frac{837.92}{(19.45)^2} = 2.21$$

$\therefore$ $\beta_2 < 3$

$\therefore$ Distribution is platykurtic.

Q. Discuss the kurtosis of the following distribution

Marks	0-10	10-20	20-30	30-40	40-50	50-60	60-70
No. of students	8	12	20	30	15	10	5

Q. The first four moments of a distribution about the value 5 of the variable are 2, 20, 40 and 50. Obtain as far as possible the various characteristics of this distribution on the basis of the information given.