Unit-III Statistical techniques

Moments :- The moments of the distribution are the arithmetic mean of the various powers of the deviations of the object from the some given no.

Moment about mean (Central moment) :-

1. For an individual series

If $x_1, x_2, x_3 ----- x_n$ are the values of the variable then, rth moment μ_r about mean $\bar{x}$ is defined a

$$\mu_r = \frac{\sum(x_i-\bar{x})^r}{n} \quad \text{where } r = 0,1,2,3 --- n.$$

2. For frequency series :-

$$\mu_r = \frac{\sum f_i(x_i-\bar{x})^r}{\sum f_i}, \quad r = 0,1,2,3 ---- n$$

same formula is used for class distribution series.

$$\boxed{\mu_0 = 1}$$

Date
20-03

$$\mu_1 = \frac{\sum_{i=1}^{n} f_i(x_i-\bar{x})}{\sum_{i=1}^{n} f_i} = \frac{\sum_{i=1}^{n} f_i x_i}{\sum_{i=1}^{n} f_i} - \frac{\bar{x}\sum_{i=1}^{n} f_i}{\sum_{i=1}^{n} f_i}$$

$$\mu_1 = \bar{x} - \bar{x}(1) = 0$$

$$\mu_2 = \frac{\sum_{i=1}^{n} f_i(x_i-\bar{x})^2}{\sum_{i=1}^{n} f_i} = (S.D)^2 = \text{variance}$$

Moment about any arbitrary no. :-

The rth moment about any arbitrary no. 'a' is denoted by μ_r' & is defined as

$$\mu_r' = \frac{\sum_{i=1}^{n} f_i (x_i - a)^r}{\sum_{i=1}^{n} f_i} \quad , r = 0, 1, 2 \ldots n$$

$\boxed{\mu_0' = 1}$, $\boxed{\mu_1' = \bar{x} - a}$

Moment about origin :

The rth moment about origin is denoted by μ_r'

$$\mu_r' = \frac{\sum_{i=1}^{n} f_i x_i^r}{\sum_{i=1}^{n} f_i} \quad , r = 0, 1, 2, 3 \ldots$$

$\boxed{\mu_0' = 1}$, $\boxed{\mu_1' = \bar{x}}$

Relation between μ_r & μ_r' :-

$$\mu_r = {}^rc_0 \mu_r' (\mu_1')^0 - {}^rc_1 \mu_{r-1}' (\mu_1')^1 + {}^rc_2 \mu_{r-2}' (\mu_1')^2 + {}^rc_3 \mu_{r-3}' (\mu_1')^3 + \ldots$$

Que:- The first 4 moments of a distribution about the value 4 are the variables are $-1.5, 17, -30, 108$. Find the moment about mean & about the origin & about the point $x = 2$

Sol:- $\mu_1' = -1.5$, $\mu_2' = 17$, $\mu_3' = -30$, $\mu_4' = 108$

$\mu_1 = 0$

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

$$\mu_3 = 3c_0 \mu_3' (\mu_1')^0 - 3c_1 \mu_2' (\mu_1')^1 + 3c_2 \mu_1' (\mu_1')^2 - 3c_3 \mu_0' (\mu_1')^3$$

$$\mu_3 = 1 \times (-30) - 3 \times 17(-1.5) + 3 \times (-1.5)^3$$
$$- 1 \times 1 \times (-1.5)^3$$

$$\mu_3 = 39.75$$

$$\mu_4 = {}^4C_0 \mu_4' (\mu_1')^0 - {}^4C_1 \mu_3' (\mu_1')^1 + {}^4C_2 \mu_2' (\mu_1')^2$$
$$- {}^4C_3 \mu_1' (\mu_1')^3 + {}^4C_4 \mu_0' (\mu_1')^4$$

$$= 1 \times 108 - 4 \times (-30)(-1.5) + 6 \times 17(-1.5)^2$$
$$- 4 \times (-1.5)^4 + 1 \times 1(-1.5)^4$$

$$\mu_4 = 142.3125$$

moment about origin

$$\mu_1' = \bar{x} - a \quad \Rightarrow -1.5 = \bar{x} - 4 \quad \Rightarrow \bar{x} = 2.5$$

first moment about origin, $\mu_1' = \bar{x} = 2.5$

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

or $14.75 = \mu_2' - (2.5)^2$

$$\mu_2' = 14.75 + 6.25 = 21$$

Date
21-03

Que:- Calculate the variance and the third central moment from the following data

x_i	0	1	2	3	4	5	6	7	8
f_i	1	9	26	59	72	52	29	7	1

Sol:-

x_i	f_i	$d = x_i - a$ where $a = 4$	$f_i d$	$f_i d^2$	$f_i d^3$
0	1	-4	-4	16	-64
1	9	-3	-27	81	-243
2	26	-2	-52	104	-208
3	59	-1	-59	59	-59
4	72	0	0	0	0
5	52	1	52	52	52
6	29	2	58	116	232
7	7	3	21	63	189
8	1	4	4	16	64
	256		-7	507	-37

$$\mu_1' = \frac{\sum_{i=1}^{n} f_i(x_i - a)}{\sum_{i=1}^{n} f_i} = \frac{-7}{256} = -0.027$$

$$\mu_2' = \frac{\sum_{i=1}^{n} f_i(x_i - a)^2}{\sum_{i=1}^{n} f_i} = \frac{507}{256} = 1.9805$$

$$\mu_3' = \frac{\sum_{i=1}^{n} f_i(x_i - a)^3}{\sum_{i=1}^{n} f_i} = \frac{-37}{256} = -0.1445$$

Now, $\mu_2 = \mu_2' - (\mu_1')^2 = 1.9805 - (-0.02734)^2$

$\mu_2 = 0.9797$

Hence, variance $= \mu_2 = 0.9797$

Now, $\mu_3 = {}^3C_0 \mu_3' - {}^3C_1 \mu_2' \mu_1' + {}^3C_2 \mu_1' \mu_1'^2 - {}^3C_3 \mu_0' \mu_1$

$\mu_3 = (-0.1445) - 3(1.9805)(-0.02734) + 3(-0.02734)$

$- (-0.02734)^3$

$\mu_3 = 0.017899$ = third central moment

Que:- From the following data calculate the first four moment about mean

Class:	0-10	10-20	20-30	30-40	40-50	50-60	60-
frequency:	4	8	11	15	11	7	4

Sol:-

class	f_i	x_i	$d = x_i - a$	$f_i d$	$f_i d^2$	$f_i d^3$	$f_i d^4$
0-10	4	5	-30	-120	3600	-108000	3240000
10-20	8	15	-20	-160	3200	-64000	1280000
20-30	11	25	-10	-110	1100	-11000	110000
30-40	15	35	0	0	0	0	0
40-50	11	45	10	110	1100	11000	110000
50-60	7	55	20	140	2800	56000	1120000
60-70	4	65	30	120	3600	108000	3240000
	60			-20	~~69400~~ 15400	-8000	~~11080000~~ 9100000

$$\mu_1' = \frac{\sum_{i=1}^{n} f_i(x_i - a)}{\sum_{i=1}^{n} f_i} = \frac{-20}{60} = -0.333$$

$$\mu_2' = \frac{15400}{60} = 256.666$$

$$\mu_3' = \frac{-8000}{60} = -133.33$$

$$\mu_4' = \frac{9100000}{60} = 151666.66$$

$\mu_1 = 0$

$\mu_2 = \mu_2' - \mu_1'^2 = 256.66 - (-0.333)^2 = 256.56$

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

$= (-133.33) - 3(256.66)(-0.333) + 2(-0.333)^3$

$\mu_3 = 120.70$

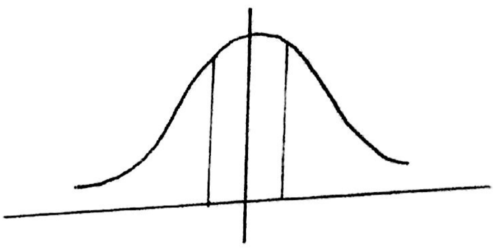
$\mu_4 = \mu_4' - 4\mu_3'\mu_1' + 6\mu_2'\mu_1'^2 - 4\mu_1'\mu_1'^3 + \mu_0'\mu_1^4$

$= 151666.67 - 4(-133.33)(-0.333) + 6(256.66)(-0.33)^2 - 3(-0.333)^4$

$\mu_4 = 151659.8022$

Date 3-03

Skewness :- For a symmetrical distribution, the frequency are symmetrically distributed about the mean i.e variates equidistant from the mean have equal frequencies.



Mean, median, mode coincides

fig. Symmetrical distribution

If the curve of distribution is not symmetrical it may be tail on either side of the distribution. Skewness is defined as the lack of symmetry in a frequency distribution.

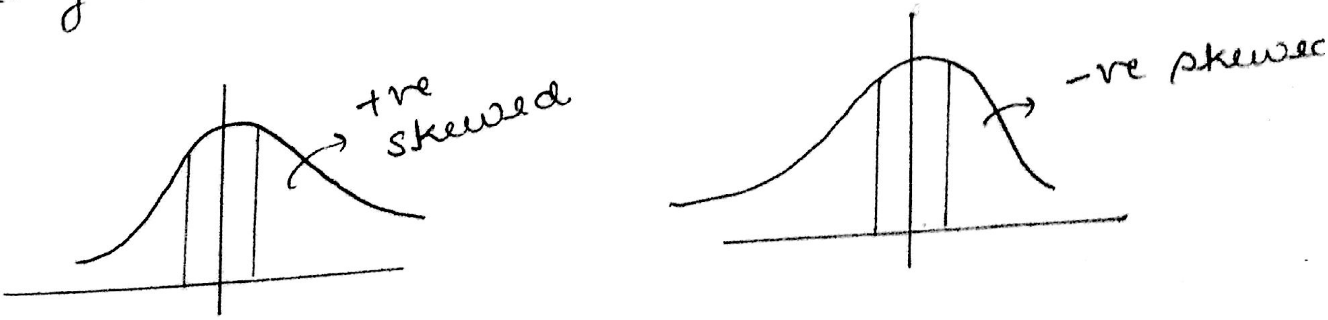


fig. unsymmetrical distribution

Test of Skewness :-

1. If arithmetic mean = median = mode, then there is no skewness.

2. If A.M is less than mode, then there tail is on left side i.e the distribution is -ve & is called negatively skewed.

If A.M is greater than mode, the tail is on right side i.e the distribution is +ve & is called positively skewed

Method of measuring skewness :-

1. Karl Pearson's method
2. Method of moments

1. Method of moments :-

$$\text{moment coefficient of skewness} = \frac{\mu_3}{\sqrt{\mu_2^3}}$$

Que:- The first three moment of a distribution of 0, 15, -31. Find the moment coefficient of skewness

Sol :- $\mu_1 = 0, \quad \mu_2 = 15, \quad \mu_3 = -31$

$$\text{coefficient of skewness} = \frac{-31}{\sqrt{(15)^3}} = -2.066$$

Que:- The first four moment of a distribution about the value 5 of the variable are 2, 20, 40, 50. Calculate the moment coefficient of skewness.

Sol :- $\mu_1' = 2, \quad \mu_2' = 20, \quad \mu_3' = 40, \quad \mu_4' = 50$

2. Karl Pearson's coefficient of skewness

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

Que:- Karl Pearson's coefficient of skewness of a distribution is 0.32. Its standard deviation is 6.5 & mean is 29.6. find mode of distri.

Sol:- $Sk_p = 0.32$, $S.D = 6.5$, Mean $= 29.6$

$$Sk_p = \frac{\text{Mean} - \text{mode}}{S.D} \Rightarrow 0.32 = \frac{29.6 - \text{mode}}{6.5}$$

mode = 27.52

Que:- For a skewed data, the A.M is 100, the variance is 35 & Karl pearson's coefficient of skewness is 0.2. Find mode & median.

Sol: $\bar{x} = 100$, variance $= 35$, $Sk_p = 0.2$

$$Sk_p = \frac{\bar{x} - \text{mode}}{\sqrt{\text{variance}}} \Rightarrow 0.2 = \frac{100 - \text{mode}}{\sqrt{35}}$$

$\therefore$ mode = 98.816

Mode = 3 median - 2 mean

$\Rightarrow 98.816 = 3(\text{median}) - 2 \times 100$

$\therefore$ median = 99.6

Que:- The sum of 20 observation is 300 & sum of their square is 5000. The median is 15. Find the karl pearson's coefficient of skewness.

Sol:- $\sum x_i = 300$, $\sum x_i^2 = 5000$, $n = 20$

median = 15.

$$\text{Mean} = \frac{\sum x_i}{n} = \frac{300}{20} = 15$$

$$S.D = \sqrt{\frac{\sum x_i^2}{n} - (\bar{x})^2} = \sqrt{250 - 225} = 5$$

Scanned by CamScanner

$$Skp = \frac{3(\text{mean} - \text{median})}{S.D} = \frac{3(15-15)}{5} = 0.$$

Date 24-03

Que:- Find the coefficient of skewness by Karl pearson's method for the following data.

value :	6	12	18	24	30	36	42
frequency :	4	7	9	18	15	10	3

Sol:-

x_i	f_i	$x_i f_i$	$f_i x_i^2$
6	4	24	144
12	7	84	1008
18	9	162	2916
24	18	432	10368
30	15	450	13500
36	10	360	12960
42	3	126	5292
	66	1638	46188

$$\bar{x} = \frac{\sum x_i f_i}{\sum f_i} = \frac{1638}{66} = 24.82$$

$$S.D = \sqrt{\frac{\sum f_i x_i^2}{n} - (\bar{x})^2}$$

$$= \sqrt{\frac{46188}{66} - (24.82)^2}$$

$$S.D = 9.15$$

$$Skp = \frac{\text{median} - \text{mode}}{S.D} = \frac{24.82 - 24}{9.15} = 0.089$$

Kurtosis :-

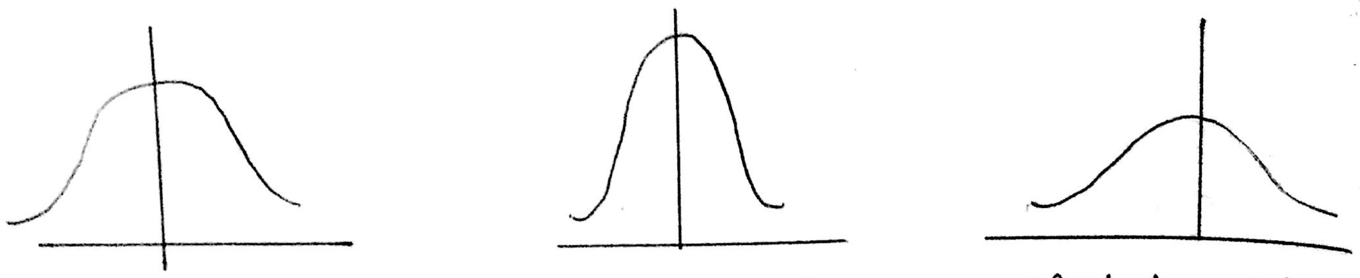
Given two frequency distribution which have the same variability as measured by the S.D, they may be relatively more or less flat than the normal curve.

A frequency curve may be symmetrical but it may not be equally flat with the normal curve. The relative flatness of the top is called Kurtosis & is measured by β_2.

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

Normal curve or mesokurtic :-

A curve which are neither flat nor sharpening are called normal curve



<u>Leptokurtic curve</u>:-

A curve which are more sharpening than the normal curve are called leptocurtic curve

<u>Platykurtic curve</u>:-

A curve which are flatter than normal curve are called platykurtic curve

(i) $\beta_2 > 3$, frequency distri is leptokurtic curve

(ii) $\beta_2 = 3$,, ,, ,, mesokurtic curve

(iii) $\beta_2 < 3$,, ,, ,, platykurtic curve

<u>Que</u>:- The first four moment of a frequency distri a 0, 100, -7, 35000. Discuss the kurtosis of the distribution.

<u>Sol</u>:- $\mu_1 = 0$, $\mu_2 = 100$, $\mu_3 = -7$, $\mu_4 = 35000$

$$\beta_2 = \frac{\mu_4}{\mu_2^2} = \frac{35000}{(100)^2} = \frac{35\cancel{000}}{10\cancel{000}} = 3.5$$

$\therefore$ $\beta_2 > 3 \Rightarrow$ leptokurtic curve

Ques: The first four moment of a distribution about $x=4$ are 1, 4, 20, 45. obtain the various characteristic of distribution on the basis of given information. comment on the nature of distribution.

Sol:

$$\mu_2 = (S.D)^2 = (5)^2 = 25$$

Solution of quest^n on next page

(i) $\beta_2 = \frac{\mu_4}{\mu_2^2} > 3 \Rightarrow \frac{\mu_4}{(25)^2} > 3$

$\Rightarrow \mu_4 > 3 \times (25)^2$ or $\mu_4 > 1875$

(ii) $\beta_2 = \frac{\mu_4}{\mu_2^2} = 3 \Rightarrow \mu_4 = 3 \times (25)^2 = 1875$

(iii) $\beta_2 = \frac{\mu_4}{\mu_2^2} < 3 \Rightarrow \mu_4 < 3 \times (25)^2$

$\Rightarrow \mu_4 < 1875.$

Date 25-03

Curve fitting :-

$(x_1, y_1)\ (x_2, y_2)\ (x_3, y_3) \cdots\cdots (x_m, y_m)$

1. Method of least square :-

$y = a_0 + a_1 x + a_2 x^2 + a_3 x^3 + \cdots a_n x^n$ (expected value)

$y_1' = a_0 + a_1 x_1 + a_2 x_1^2 + a_3 x_1^3 + \cdots a_n x_1^n$

$y_2' = a_0 + a_1 x_2 + a_2 x_2^2 + a_3 x_2^3 + \cdots a_n x_2^n$

$\vdots$

$y_m' = a_0 + a_1 x_m + a_2 x_m^2 + a_3 x_m^3 + \cdots a_n x_m^n$

$(y_1 - y_1'), (y_2 - y_2'), (y_3 - y_3') \cdots\cdots (y_m - y_m')$

these are called residuals.

in general, $(y_i - y_i')$

$$V = \sum_{i=1}^{m} (y_i - y_i')^2$$

or $$V = \sum_{i=1}^{m} \left(y_i - a_0 - a_1 x_i - a_2 x_i^2 - a_3 x_i^3 \cdots a_n x_i^n\right)^2$$

$$\frac{\partial V}{\partial a_0} = 0, \quad \frac{\partial V}{\partial a_1} = 0, \quad \frac{\partial V}{\partial a_2} = 0, \cdots \frac{\partial V}{\partial a_n} = 0$$

Scanned by CamScanner

Que:- The S.D of a symmetric distribution is 5. What must be the value of fourth moment about mean in order that distribution are
(i) leptokurtic (ii) Mesokurtic (iv) Platykurtic

$$\sum_{i=1}^{m} 2(y_i - a_0 - a_1 x_i - a_2 x_i^2 - \cdots - a_n x_i^n)(-1) = 0$$

$$\sum_{i=1}^{m} 2(y_i - a_0 - a_1 x_i - a_2 x_i^2 - \cdots - a_n x_i^n)(-x_i) = 0$$

these equation are normal eqn.

Fit a straight line :-

$(x_1, y_1), (x_2, y_2)\ (x_3, y_3) \cdots (x_m, y_m)$

$y = a + bx$, $y_1' = a + bx_1$, $y_2' = a + bx_2$

$y_3' = a + bx_3 \cdots y_m' = a + bx_m$

$$V = \sum_{i=1}^{m} (y_i - y_i')^2 = \sum_{i=1}^{m} (y_i - a - bx_i)^2$$

$$\frac{\partial v}{\partial a} = 0 , \quad \frac{\partial v}{\partial b} = 0$$

$$\Rightarrow \left.\begin{array}{l} \sum_{i=1}^{m} 2(y_i - a - bx_i)(-1) = 0 \quad -① \\ \sum_{i=1}^{m} 2(y_i - a - bx_i)(-x_i) = 0 \quad -② \end{array}\right\} \text{normal eq}^n$$

from (1)

$$\sum_{i=1}^{m} (y_i - a - bx_i) = 0$$

or $$\sum_{i=1}^{m} y_i - a\sum_{i=1}^{m} 1 - b\sum_{i=1}^{m} x_i = 0$$

or $$Y - ma - bX = 0 \Rightarrow \boxed{Y = ma + bX}$$

Scanned by CamScanner

Similarly from (2)

$$\sum_{i=1}^{m}(-x_i y_i) + a\sum_{i=1}^{m} x_i + b\sum_{i=1}^{m} x_i^2 = 0$$

or $-XY + aX + bX^2 = 0$

or $\boxed{XY = aX + bX^2}$

also, $Y = ma + bX$

$$\left.\begin{aligned} \Sigma Y &= ma + b\Sigma X \\ \Sigma XY &= a\Sigma X + b\Sigma X^2 \end{aligned}\right\} \text{normal equ.}$$

$y = a + bx + cx^2$

$xy = ax + bx^2 + cx^3$

$x^2y = ax^2 + bx^3 + cx^4$

$y = \frac{a}{x} + bx$

$\frac{y}{x} = \frac{a}{x^2} + b$

$xy = a + bx^2$

Que:- Using the method of a least square fit a straight line from the following data

x	0	2	4	5	6
y	5.012	10	15	21	30.

Sol:- let us assume, $y = a + bx$. — (1)

by method of least square normal eqn are

$\Sigma Y = 5a + b\Sigma X$ — (2)

$\Sigma XY = a\Sigma X + b\Sigma X^2$ — (3)

x	y	xy	x^2
0	5.012	0	0
2	10	20	4
4	15	60	16
5	21	105	25
6	30	180	36
$\Sigma x = 17$	$\Sigma Y = 81.012$	$\Sigma xy = 365$	$\Sigma x^2 = 81$

$81.012 = 5a + 17b$

$365 = 17a + 81b$

$a = 3.077$

$b = 3.86$

$\therefore\ y = 3.077 + 3.86x.$

Date 27-03

$$y = ae^{bx}$$

$$\log_e y = \log_e a + bx \log_e e$$

$$Y = A + bx$$

$$\Sigma Y = nA + b\Sigma X$$

$$\Sigma XY = A\Sigma x + b\Sigma x^2$$

$$Y = \log_e y, \quad A = \log_e a,$$

Que:- Using least square method to fit a curve of the form $y = ae^{bx}$ from the following data

x	1	1.2	1.4	1.6
y	40.17	73.196	133.372	243.02

Sol:-

let $y = ae^{bx}$ —— (1)

$$\log_e y = \log_e a + bx$$

$$Y = A + bx$$

normal equations are

$$\Sigma Y = nA + b\Sigma x$$

$$\Sigma xY = A\Sigma x + b\Sigma x^2$$

x	y	Y	xY	x^2
1	40.17	3.693	3.693	1
1.2	73.196	4.293	5.1516	1.44
1.4	133.372	4.893	6.8502	1.96
1.6	243.02	5.493	8.7888	2.56

$\Sigma x = 5.2$ $\quad \Sigma Y = 18.372$ $\quad \Sigma xY = 24.48$ $\quad 6.96 = \Sigma x^2$

$$\Rightarrow \quad 18.372 = 4A + 5.2b$$

$$24.48 = 5.2A + 6.96b$$

$$\Rightarrow \quad A = 0.716, \quad b = 2.98$$

$$\log_e a = 0.716 \quad \Rightarrow \quad a = e^{0.716} = 2.046$$

Date 30-03

Correlation :-

Covariance :- Let the corresponding value of two variables X & Y are $(x_1, y_1), (x_2, y_2), (x_3, y_3), \ldots (x_n, y_n)$.

Then, covariance b/w X & Y is denoted by $cov(X,Y)$

$$cov(X,Y) = \frac{(x_1-\bar{x})(y_1-\bar{y}) + (x_2-\bar{x})(y_2-\bar{y}) + \cdots + (x_n-\bar{x})(y_n-\bar{y})}{n}$$

or $$Cov(X,Y) = \frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})$$

where $\bar{x}$ & $\bar{y}$ are the mean of x_i & y_i.

→ In a bivariat distribution, if the change in one variable affect a change in the other variable, the variable are called correlated.

Method of determining Correlation

1. **Karl Pearson's coefficient of correlation**

Correlation between two variables X & Y denoted by 'r' & is defined as

$$r = \frac{\text{covariance of }(X,Y)}{\sqrt{\text{variance of }X}\sqrt{\text{variance of }Y}}$$

or $$r = \frac{1/n\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})}{\sqrt{\frac{1}{n}\Sigma(x_i-\bar{x})^2}\sqrt{\frac{1}{n}\Sigma(y_i-\bar{y})^2}}$$

or $$r = \frac{\Sigma(x_i-\bar{x})(y_i-\bar{y})}{\sqrt{\Sigma(x_i-\bar{x})^2}\sqrt{\Sigma(y_i-\bar{y})^2}}$$

Que: Find the correlation coefficient b/w x & y for the following data

x:	21	23	30	54	57	58	72	78	87	90
y:	60	71	72	83	110	84	100	92	113	135

Sol:

x_i	y_i	$x_i-\bar{x}$	$y_i-\bar{y}$	$(x_i-\bar{x})^2$	$(y_i-\bar{y})^2$	$(x_i$ $(y_i$
21	60	-36	-32	1296	1024	115
23	71	-34	-21	1156	441	71
30	72	-27	-20	729	400	54
54	83	-3	-9	9	81	27
57	110	0	18	0	324	0
58	84	1	-8	1	64	-8
72	100	15	8	225	64	120
78	92	21	0	441	0	0
87	113	30	21	900	441	630
90	135	33	43	1089	1849	141
570	920	0	-8	5846	4688	459

$\bar{x} = \dfrac{570}{10} = 57$, $\bar{y} = \dfrac{920}{10} = 92$

correlation coefficient

$$r = \frac{\Sigma(x_i-\bar{x})(y_i-\bar{y})}{\sqrt{\Sigma(x_i-\bar{x})^2}\sqrt{\Sigma(y_i-\bar{y})^2}} = \frac{4594}{\sqrt{5846}\sqrt{4688}}$$

$$r = \frac{4596}{76.459 \times 68.468} = 0.87754$$

Scanned by CamScanner

Date 31-03

Spearman's rank correlation coefficient

$$r = 1 - \frac{6\Sigma d^2}{n(n^2-1)}$$ where $d = r_1 - r_2$

1. For equal roots

$$r = 1 - \frac{6\left[\Sigma d^2 + \frac{1}{12}(m_1^3 - m_1) + \frac{1}{12}(m_2^3 - m_2) + \cdots\right]}{n(n^2-1)}$$

where, $m_1, m_2, m_3 \cdots$ are number of rank repeatition

Que:- Compute Spearman's rank correlation coefficient 'r' for the following data.

Person	:	A	B	C	D	E	F	G	H	I	J
Rank in statistic	:	9	10	6	5	7	2	4	8	1	3
Rank in income	:	1	2	3	4	5	6	7	8	9	10

Sol:-

Person	Rank in St. (r_1)	Rank in income (r_2)	$d = r_1 - r_2$	d^2
A	9	1	8	64
B	10	2	8	64
C	6	3	3	9
D	5	4	1	1
E	7	5	2	4
F	2	6	-4	16
G	4	7	-3	9
H	8	8	0	0
I	1	9	-8	64
J	3	10	-7	49

$\Sigma d^2 = 280$

$$r = 1 - \frac{6 \times 280}{10(10^2-1)} = 1 - \frac{6 \times 280}{10 \times 99} = -0.697$$

Scanned by CamScanner

Que:- Obtain the rank correlation coefficient for the following data

x	y	Rank(x) (r_1)	Rank(y) (r_2)	d ($r_1 - r_2$)	d^2
68	62	4.	5	-1	1
64	58	6	7	-1	1
75.	68	2.5	3.5	-1	1
50	45	9	10	-1	1
64	81	6	1	5	25
80	60	1	6	-5	25
75	68	2.5	3.5	-1	1
40	48	10	9	1	1
55	50	8	8	0	0
64	70.	6	2	4	16

$\Sigma d^2 = 72.$

for equal roots

$$r = 1 - \frac{6\left(72 + \frac{1}{12}(2^3-2) + \frac{1}{12}(3^3-3) + \frac{1}{12}(2^2-2)\right)}{10 \times 99}$$

$$r = 1 - \frac{6 \times 75}{10 \times 99} = 0.545$$

Regression :- If the scatter diagram indicates some relationship between two variables x & y, then the dots in the scatter diagram will be concentrated around the curve. This curve is called as curve of regression.

Line of regression :- Let n order pair $(x_1, y_1), (x_2, y_2)$ (x_3, y_3) ----- (x_n, y_n). Let a no. of pairs are two correlated variables.

Suppose we have to find out the unknown values y for a known values of x, then we have line of regression of y on x. i.e $y = a + bx$. Here y is dependent variable & x is independent variable.

If we have to find unknown values of x for a known values of y, then we have line of regressi

Scanned by CamScanner

of x on y i.e $x = a + by$ where x is dependent variable & y is independent variable. Thus, we have two line of regression.

Equation to the line of regression

let $y = a + bx$ —— ①

$\Sigma y = na + b\Sigma x$ —— ②

$\Sigma xy = a\Sigma x + b\Sigma x^2$ —— ③

dividing eq. (2) by n

$$\frac{\Sigma y}{n} = a + b\frac{\Sigma x}{n} \Rightarrow \bar{y} = a + b\bar{x} \quad (\bar{x}, \bar{y})$$

from (3),

$$\Sigma(x-\bar{x})(y-\bar{y}) = a\Sigma(x-\bar{x}) + b\Sigma(x-\bar{x})^2$$

or $\Sigma(x-\bar{x})(y-\bar{y}) = b\Sigma(x-\bar{x})^2$

$$\therefore \quad b = \frac{\Sigma(x-\bar{x})(y-\bar{y})}{\Sigma(x-\bar{x})^2} \quad —— ④$$

correlation coefficient

$$r = \frac{\Sigma(x-\bar{x})(y-\bar{y})}{\sqrt{\Sigma(x-\bar{x})^2}\sqrt{\Sigma(y-\bar{y})^2}} \quad — ⑤$$

Date
01-04

$$r = \frac{\Sigma(x-\bar{x})(y-\bar{y})}{n\sqrt{\frac{\Sigma(x-\bar{x})^2}{n}}\sqrt{\frac{\Sigma(y-\bar{y})^2}{n}}} = \frac{\Sigma(x-\bar{x})(y-\bar{y})}{n\,\sigma_x\,\sigma_y}$$

or $\Sigma(x-\bar{x})(y-\bar{y}) = r\,n\,\sigma_x\,\sigma_y$

putting this value in eq. (4) we get

$$b = \frac{r\,n\,\sigma_x\,\sigma_y}{\Sigma(x-\bar{x})^2} = \frac{r\,\sigma_x\,\sigma_y}{\frac{\Sigma(x-\bar{x})^2}{n}}$$

$$b = \frac{r\,\sigma_x\,\sigma_y}{\sigma_x^2} = \frac{r\,\sigma_y}{\sigma_x}$$

$\therefore$ $\boxed{b = \frac{r\sigma_y}{\sigma_x}}$ → slope of line of regression

the line of regression passes through the point $(\bar{x}, \bar{y})$ & having slope $b = \frac{r\sigma_y}{\sigma_x}$

Hence, the equation of line of regression is

$$y - \bar{y} = b(x - \bar{x})$$

or $\boxed{y - \bar{y} = \frac{r\sigma_y}{\sigma_x}(x - \bar{x})}$ which are called regression line of y on x.

Similarly, regression line of x on y.

$$x - \bar{x} = \frac{r\sigma_x}{\sigma_y}(y - \bar{y})$$

$b_{yx} = \frac{r\sigma_y}{\sigma_x}$. $b_{xy} = \frac{r\sigma_x}{\sigma_y}$ are called coefficient of regression

or $\boxed{b_{yx} \cdot b_{xy} = r^2}$

Que:- If θ be the acute angle between the two regression line in the case of two variables x & y. Show that $\tan\theta = \left(\frac{1-r^2}{r}\right)\frac{\sigma_x\sigma_y}{\sigma_x^2+\sigma_y^2}$

where, r, σ_x, σ_y have their usual meanings. Explain the significance where $r = 0$ & $r = \pm 1$

Sol:- Let the two regression lines be

$$y - \bar{y} = \frac{r\sigma_y}{\sigma_x}(x - \bar{x}) \quad \text{---} (1)$$

$$\& \quad x - \bar{x} = \frac{r\sigma_x}{\sigma_y}(y - \bar{y}) \quad \text{---} (2)$$

from (1) $y = \frac{r\sigma_y}{\sigma_x}(x - \bar{x}) + \bar{y}$

Let m_1 be the slope of 1st & m_2 be the slope of 2nd.

$\Rightarrow$ $m_1 = \frac{r\sigma_y}{\sigma_x}$, $m_2 = \frac{r\sigma_x}{\sigma_y} = \frac{1}{r}\frac{\sigma_y}{\sigma_x}$

we know that

$$\tan\theta = \frac{|m_1 - m_2|}{1+m_1 \cdot m_2} = \frac{\left|\frac{r\sigma_y}{\sigma_x} - \frac{1}{r}\frac{\sigma_y}{\sigma_x}\right|}{1+\left(\frac{r\sigma_y}{\sigma_x} \times \frac{1}{r}\frac{\sigma_y}{\sigma_x}\right)}$$

$$\tan\theta = \frac{\left|\left(\frac{r^2-1}{r}\right)\frac{\sigma_y}{\sigma_x}\right|}{1+\frac{\sigma_y^2}{\sigma_x^2}} = \frac{\left|\left(\frac{r^2-1}{r}\right)\frac{\sigma_y}{\sigma_x}\right|}{\frac{\sigma_x^2+\sigma_y^2}{\sigma_x^2}}$$

or $$\tan\theta = \frac{\frac{1-r^2}{r} \times \sigma_x^2 \times \frac{\sigma_y}{\sigma_x}}{\sigma_x^2+\sigma_y^2} = \frac{1-r^2}{r}\frac{\sigma_x \cdot \sigma_y}{\sigma_x^2+\sigma_y^2}$$

$$\therefore \tan\theta = \left(\frac{1-r^2}{r}\right)\frac{\sigma_x\sigma_y}{\sigma_x^2+\sigma_y^2} \quad \text{Hence proved} \quad (3)$$

(i) if $r=0$, there is no relationship b/w the variables and they are independent

On putting $r=0$ in the eq. (3) we get $\tan\theta = \infty$

or $\theta = 90°$ & so, line (1) & (2) are perpendicular.

(ii) if $r = \pm 1$, on putting these values of r in eq. (3) we get $\tan\theta = 0 \Rightarrow \theta = 0°$ i.e line (1) & (2) are coincide. The correlation b/w the variables is perfect.

Que:- If the coefficient of correlation b/w two variables x & y is 0.5 and the acute angle b/w their line of regression is $\tan^{-1}(3/5)$. Show that $\sigma_x = \frac{1}{2}\sigma_y$

Sol:-

given $r = 0.5$

$\theta = \tan^{-1}(3/5) \Rightarrow \tan\theta = 3/5.$

$$\Rightarrow \left(\frac{1-r^2}{r}\right)\frac{\sigma_x\sigma_y}{\sigma_x^2+\sigma_y^2} = \frac{3}{5}$$

or $$\left(\frac{1-1/4}{1/2}\right)\frac{\sigma_x\sigma_y}{\sigma_x^2+\sigma_y^2} = \frac{3}{5}$$

$$\Rightarrow \quad \frac{\cancel{8}}{2}\frac{\sigma_x \sigma_y}{\sigma_x^2+\sigma_y^2} = \frac{\cancel{8}}{5} \quad \Rightarrow \quad \frac{\sigma_x \sigma_y}{\sigma_x^2+\sigma_y^2} = \frac{2}{5}$$

or $\quad 5\sigma_x \sigma_y = 2(\sigma_x^2+\sigma_y^2)$

$\Rightarrow \quad 2\sigma_x^2 + 2\sigma_y^2 - 5\sigma_x\sigma_y = 0$

or $\quad 2\sigma_x^2 - 4\sigma_x\sigma_y + 2\sigma_y^2 - \sigma_x\sigma_y = 0$

or $\quad 2\sigma_x(\sigma_x - 2\sigma_y) + \sigma_y(2\sigma_y - \sigma_x) = 0$

$\Rightarrow \quad (\sigma_x - 2\sigma_y)[2\sigma_x - \sigma_y] = 0$

$\sigma_x = 2\sigma_y$, $\quad \sigma_x = \frac{1}{2}\sigma_y$. Hence proved

Moment Generating function

Date 03-04

Moment generating function is on indirect method for computing moment. This method depends on the finding of the moment generating fn

1. In case of continuous variable x, it is defined as

$$M(t) = \int_a^b e^{tx} \cdot f(x)\,dx$$

where, integral is a function of parameter t. and the limit a, b can be taken as $-\infty$ 4 ∞

let us see how $M(t)$ generate moment

$$M(t) = \int_a^b \left(1 + tx + \frac{(tx)^2}{2!} + \frac{(tx)^3}{3!} + \cdots \frac{(tx)^r}{r!}\right) f(x)\,dx$$

or $$M(t) = \int_a^b f(x)\,dx + t\int_a^b x f(x)\,dx + \frac{t^2}{2!}\int_a^b x^2 f(x)\,dx + \cdots \frac{t^r}{r!}\int_a^b x^r f(x)\,dx.$$

or $$M(t) = \mu_0' + t\mu_1' + \frac{t^2}{2!}\mu_2' + \frac{t^3}{3!}\mu_3' + \cdots \frac{t^r}{r!}\mu_r'$$

2. In case of discrete distribution, the moment generating function

$$M(t) = \sum_{x=1}^{\infty} e^{tx} f(x)$$

Que:- Find the moment generating function for triangular distri defined by,

$$f(x) = \begin{cases} x & ; \ 0 \le x \le 1 \\ 2-x & ; \ 1 \le x \le 2 \end{cases}$$

Sol:

$$M(t) = \int_0^2 e^{tx} \cdot f(x)\,dx$$

$$M(t) = \int_0^1 e^{tx} \cdot x \, dx + \int_1^2 e^{tx}(2-x)\, dx$$

$$= \left[x\left(\frac{e^{tx}}{t}\right) - (1)\left(\frac{e^{tx}}{t^2}\right) \right]_0^1 + \left[(2-x)\left(\frac{e^{tx}}{t}\right) - (-1)\left(\frac{e^{tx}}{t^2}\right.\right.$$

$$\text{or } M(t) = \left[\cancel{\frac{e^t}{t}} - \cancel{\frac{e^t}{t^2}} + \frac{1}{t^2} \right] + \left[+ \frac{e^{2t}}{t^2} - \cancel{\frac{e^t}{t}} + \cancel{\frac{e^t}{t^2}} \right]$$

$$= \frac{1}{t^2} - \cancel{\frac{e^{2t}}{t^2}} = \frac{1}{t^2}\cancel{\left(1 - e^{2t}\right)} = \frac{(1-e^t)^2}{t^2}$$

Que:- find M...

Binomial Probability Distribution :-

Let there be n independent trials in an experiment. Let a random variable x denotes the no. of success in the n independent trial.

Let p is the probability of success & q be the probability of failure in a single trial.

$$\therefore \quad p+q = 1.$$

Let the trial be independent & p is constant for every trial. Let us find the probability or 'r' success in 'n' trial

$$P(X=r) = {}^{n}C_{r}\, p^{r} q^{n-r}, \quad \text{where } r = 0, 1, 2, 3 \ldots n$$

Recurrence formula for binomial distribution.

$$P(X=r) = {}^{n}C_{r}\, p^{r} q^{n-r}$$

$$P(X=r+1) = {}^{n}C_{r+1}\, p^{r+1} q^{n-(r+1)}$$

Scanned by CamScanner

$$\frac{P(X=r+1)}{P(X=r)} = \frac{{}^nC_{r+1}\, p^{r+1} \cdot q^{n-(r+1)}}{{}^nC_r\, p^r \cdot q^{n-r}} = \frac{(n-r)}{(r+1)}\frac{p}{q}$$

$$\boxed{P(X=r+1) = \left(\frac{n-r}{r+1}\right)\frac{p}{q}\, P(X=r)}$$

Mean & variance of discrete binomial distri.

We know that the moment generating fn. about the origin is given as

$$M(t) = \sum e^{tx} P(X=x)$$

$$= \sum_{x=0}^{n} e^{tx}\, {}^nC_x\, p^x q^{n-x}$$

$$\text{or } M(t) = \sum_{x=0}^{n} {}^nC_x (pe^t)^x q^{n-x}$$

$$= (q+pe^t)^n$$

$$\text{Mean} = \mu_1' = \left[\frac{\partial}{\partial t} M(t)\right]_{t=0} = \left[\frac{\partial}{\partial t}(q+pe^t)^n\right]_{t=0}$$

$$= \left[n(q+pe^t)^{n-1} \cdot pe^t\right]_{t=0}$$

$$\boxed{\text{Mean} = np}$$

$$\mu_2' = \left[\frac{\partial^2}{\partial t^2} M(t)\right]_{t=0} = np\left[\frac{\partial}{\partial t}(q+pe^t)^{n-1} \cdot e^t\right]_{t=0}$$

$$= np\left[(n-1)(q+pe^t)^{n-2}\, pe^{2t} + (q+pe^t)^{n-1} e^t\right]_{t=0}$$

$$= np[(n-1)p+1] = np[np-p+1]$$

$$= np[np+(1-p)] = np[np+q]$$

$$= n^2p^2 + npq$$

Scanned by CamScanner

$\mu_2 = \mu_2' - \mu_1'^2 = n^2p^2 + npq - n^2p^2$

$\Rightarrow \boxed{\mu_2 = npq = \text{variance}}$

Date 07.04

Que:- A die is tossed thrice. A success is getting 1 or 6 on a toss. Find the mean & variance of the no. of success.

Sol: probability of getting 1 or 6 on a toss

$= p = 2/6 = 1/3$

$\therefore q = 1-p = 1-1/3 = 2/3$

also, $n = 3$

mean $= np = 3 \times 1/3 = 1$

variance $= npq = 3 \times 1/3 \times 2/3 = 2/3$

Que:- If 10% of the bolts produced to m/c are defective find the probability that out of 10 bolts taken at random (i) 1 (ii) none (iii) at most 2 bolts will be defective.

Sol: probability of success (defective) $= \frac{10}{100} = 1/10$

q (effective) $= 1 - 1/10 = 9/10$

$n = 10$

(i) $P(x=1) = {}^{10}C_1 p^1 q^9 = {}^{10}C_1 \left(\frac{1}{10}\right)^1 \left(\frac{9}{10}\right)^9$

$= .3874$

(ii) $P(x=0) = {}^{10}C_0 \left(\frac{1}{10}\right)^0 \left(\frac{9}{10}\right)^9 = .3486$

(iii) $P(x \leq 2) = P(x=0) + P(x=1) + P(x=2)$

$= 0.3486 + 0.3874 + {}^{10}C_2 (1/10)^2 (9/10)^8$

$= 0.9297.$

Poisson distribution :-

If the parameters n & p of a binomial distribution are known, we can find the distribution. But in situation where n is very large & p is very small, application of binomial distri is very complicated. Then, we consider as $n \to \infty$, $p \to 0$ such that np always remains finite. we get the poisson approximⁿ to the binomial distribution. Now, for poisson distri.

$$P(X=r) = \frac{e^{-\lambda}\lambda^r}{\lfloor r}, \quad \text{where. } r = 0, 1, 2, 3 \dots \infty$$

where λ is a finite no. i.e $\boxed{\lambda = np}$

The sum of the probabilities $P(r)$ for $r = 0, 1, 2 \dots \infty$ is 1.

$$P(r) = P(0) + P(2) + P(2) + P(3) \dots$$

$$= \frac{e^{-\lambda}}{\lfloor 0} + \frac{e^{-\lambda}.\lambda}{\lfloor 1} + \frac{e^{-\lambda}.\lambda^2}{\lfloor 2} + \dots$$

$$= e^{-\lambda}\left[1 + \frac{\lambda}{\lfloor 1} + \frac{\lambda^2}{\lfloor 2} + \dots\right]$$

$$= e^{-\lambda} \times e^{\lambda} = 1.$$

Recurrence formula for P.D :-

$$P(X=r) = \frac{e^{-\lambda}.\lambda^r}{\lfloor r}$$

$$P(X=r+1) = \frac{e^{-\lambda}.\lambda^{r+1}}{\lfloor r+1}$$

$$P(X=r+1) = \frac{\lambda}{r+1} P(X=r), \quad r = 0, 1, 2, 3 \dots$$

Date
08-04

Mean & Variance of P.D

Mean, $\mu = \sum_{x=0}^{\infty} x P(X=x) = \sum_{x=0}^{\infty} x \frac{e^{-\lambda} \cdot \lambda^x}{\lfloor x}$

$\mu = e^{-\lambda} \sum_{x=0}^{\infty} \frac{x \cdot \lambda^x}{\lfloor x} = e^{-\lambda} \sum_{x=1}^{\infty} \frac{\lambda^x}{\lfloor x-1}$

$= e^{-\lambda} \left[\frac{\lambda}{\lfloor 0} + \frac{\lambda^2}{\lfloor 1} + \frac{\lambda^3}{\lfloor 2} + \cdots \right]$

or $\mu = e^{-\lambda} \cdot \lambda \left[1 + \frac{\lambda}{\lfloor 1} + \frac{\lambda^2}{\lfloor 2} + \cdots \right]$

$= \lambda \cdot e^{-\lambda} \cdot e^{\lambda} = \lambda$

$\therefore \mu_1' = \lambda$

Variance $= \sigma^2 = \sum_{x=0}^{\infty} x^2 P(X=x) - \mu_1'^2$

or $= \sum_{x=0}^{\infty} x^2 \frac{e^{-\lambda} \cdot \lambda^x}{\lfloor x} - \lambda^2$

$= e^{-\lambda} \left[\frac{1^2 \lambda}{\lfloor 1} + \frac{2^2 \lambda^2}{\lfloor 2} + \frac{3^2 \lambda^3}{\lfloor 3} + \cdots \infty \right] - \lambda^2$

$= e^{-\lambda} \cdot \lambda \left[1 + \frac{2\lambda}{\lfloor 1} + \frac{3\lambda^2}{\lfloor 2} + \cdots \right] - \lambda^2$

$= e^{-\lambda} \cdot \lambda \left[1 + \frac{(\lambda + \lambda)}{\lfloor 1} + \frac{(\lambda^2 + 2\lambda^2)}{\lfloor 2} + \frac{(\lambda^3 + 3\lambda^3)}{\lfloor 3} \right]$

Variance $= e^{-\lambda} \cdot \lambda \left[1 + \frac{\lambda}{\lfloor 1} + \frac{\lambda^2}{\lfloor 2} + \cdots \infty + \frac{\lambda}{\lfloor 1} + \frac{2\lambda^2}{\lfloor 2} + \frac{3\lambda^3}{\lfloor 3} + \cdots \right] - \lambda^2$

$= e^{-\lambda} \cdot \lambda \left[e^{\lambda} + \lambda \left\{ 1 + \frac{\lambda}{\lfloor 1} + \frac{\lambda^2}{\lfloor 2} + \cdots \right\} \right] - \lambda^2$

$= e^{-\lambda} \cdot \lambda \left[e^{\lambda} + \lambda e^{\lambda} \right] - \lambda^2 = \lambda + \cancel{\lambda^2} - \cancel{\lambda^2}$

$\boxed{\sigma^2 = \lambda}$

Que: If there are three misprints in a book of 1000 page. Find the probability that a given page will contain (1) no misprint (2) more than 2 misprints

Sol: $n = 1000$

$P(\text{misprint / page}) = p = 3/1000$

$\lambda = np = 1000 \times \frac{3}{1000} = 3$

$$P(x=0) = \frac{e^{-\lambda} \cdot \lambda^{r}}{\lfloor r} = e^{-3} \cdot 1 = 0.0497$$

for a single page, $n=1$, $p = 3/1000$

$\lambda = 3/1000$

$$P(x=0) = e^{-3/1000} = 0.997$$

for more than 2 misprints

$$P(x>2) = P(x=3) = \frac{e^{-3/1000} \cdot (3/1000)^3}{3!}$$

$$= 4.48 \times 10^{-9}$$

Que:- Six coins are tossed for 6400 times. Using the P.D find the approximate probability of getting 6 heads x times.

Sol:- $n = 6400$, $r = x$

$$p = (1/2)(1/2)(1/2)(1/2)(1/2)(1/2) = \frac{1}{2^6}$$

$$\lambda = 6400 \times \left(\frac{1}{2^6}\right) = 100$$

$$P(x=x) = \frac{e^{-100}(100)^{x}}{\lfloor x}$$

Normal distribution

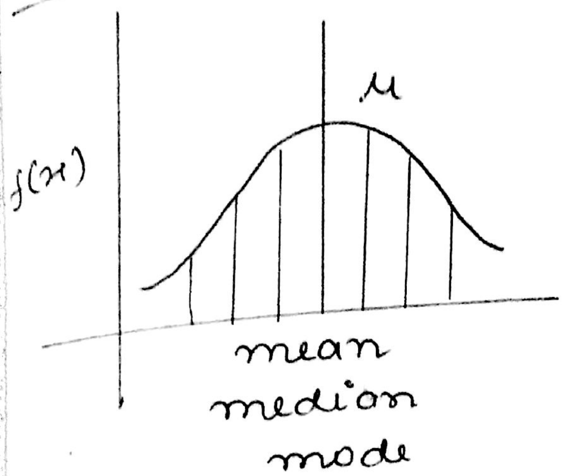
Normal distribution is a continuous distribution. It is derived as the limiting form of binomial distribution. For large value of n & p, the probability of success is closed to 1/2 & the probability density fn.

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

where, μ = mean, $\sigma \rightarrow$ standard deviation

x is called the normal variate & $f(x)$ is called probability density fn. of the N.D.

Normal distribution curve :-



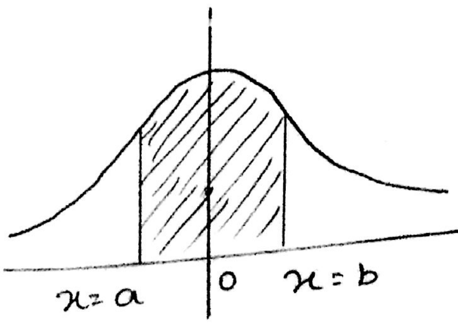
Properties of N.D :-

(i) $f(x) \geq 0$

(ii) $\int_{-\infty}^{\infty} f(x)\,dx = 1$

(total area under the normal curve above the x-axis is 1)

or area under normal probability curve



$$A = \int_a^b f(x)\,dx \rightarrow P(a \leq x \leq b)$$

$$A = \int_a^{\infty} f(x)\,dx \rightarrow P(x \geq a)$$

$$A = \int_{-\infty}^{b} f(x)\,dx \rightarrow P(x \leq b)$$

Standard form of the normal distribution

Let x be a normal variable with mean μ & S.D σ then, the random variable $z = \frac{x-\mu}{\sigma}$ has a N.D with mean 0 & S.D 1.

The random variable z is called the standard normal variable.

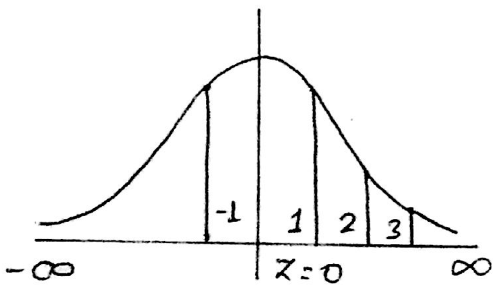
The probability density fn of the standard normal variable z is given by

$$f(z) = \frac{1}{\sqrt{2\pi}} e^{-z^2/2}, \quad -\infty < z < \infty$$

& the corresponding distribution fn is given by

$$P(a \le z \le b) = \int_a^b f(z)\,dz$$

Area under normal standard distribution



$$A = \int_a^b f(z)dz = P(a \le z \le b)$$

$$P(-1 < z < 1) = 0.68$$

$$P(-2 < z < 2) = 0.95$$

Que:- If x is normally distributed with mean 2 and variance 1. then find $P(|x-2|<1)$.

Sol:- $\mu = 2$, $\sigma = 1$, $z = \frac{x-\mu}{\sigma} = \frac{x-2}{1}$

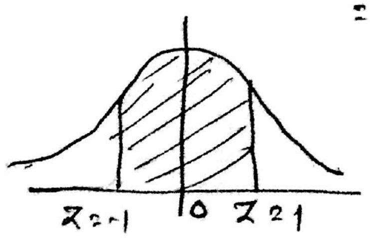
$$P(|x-2|<1) = P(|z|<1)$$

$$= P(-1<z<1)$$

$$= P(-1<z<0) + P(0<z<1)$$

$$= 2P(0<z<1)$$

$$= 2(0.3413) = 0.6826$$



Q. If x is a normal variate with mean 30 & S.D = 5 find the probability that

(i) $26 \le x \le 40$ (ii) $x \ge 45$ (iii) $|x-30| > 5$

Sol.

$\mu = 30, \ \sigma = 5$

$$z = \frac{x-\mu}{5} = \frac{x-30}{5}$$

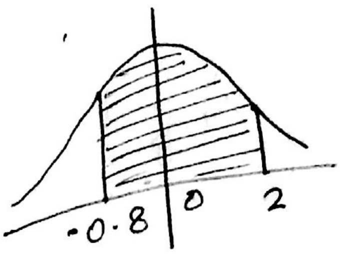
(i) $$P(26 \le x \le 40) = P\left(\frac{26-30}{5} \le \frac{x-30}{5} \le \frac{40-30}{5}\right)$$

or $P(-0.8 \le z < 2)$

$$= P(-0.8 \le z \le 0) + P(0 \le z \le 2)$$

$$= P(0 \le z \le 0.8) + P(0 \le z \le 2)$$

$$= 0.2881 + 0.4772 = 0.7653$$



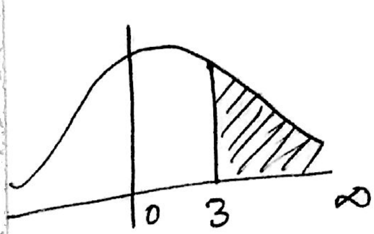
(ii) $$P(x \ge 45) = P\left(\frac{x-30}{5} \ge \frac{45-30}{5}\right)$$

$$= P(z \ge 3)$$

$$= P(0 \le z \le \infty) - P(0 \le z \le 3)$$

$$= 0.5 - 0.4987$$

$$= 0.0013$$



(iii) $$P(|x-30| > 5) = P\left(\left|\frac{x-30}{5}\right| > \frac{5}{5}\right)$$

$$= P\left(|z| > \frac{5}{5}\right) = P(|z| > 1)$$

$$= P(-\infty < z < \infty) - P(-1 \le z \le 1)$$

$$= 1 - 2P(0 < z < 1)$$

$$= 1 - 2 \times 0.3413 = 1 - 0.6826$$

$$= 0.3284 = 0.3174$$