Unit-5

(least)^2 method

Curve Fitting :-

Straight line equation :- (using least squar method)

$$y = a + bx \rightarrow (1)$$

Normal eqn of Straight line eqn are

apply Σ

$$\Sigma y = Na + b\Sigma x \rightarrow (2)$$

multiply x & apply Σ

$$\Sigma xy = a\Sigma x + b\Sigma x^2 \rightarrow (3).$$

ex: i) Fit the straight line eqn to the following table by using least squar method

x	1	2	3	4	5
y	14	27	40	55	68

Sol: We know $y = a + bx \rightarrow (1)$

Normal eqn,

$$\Sigma y = Na + b\Sigma x \rightarrow (2)$$

$$\Sigma xy = a\Sigma x + b\Sigma x^2 \rightarrow (3)$$

x	y	x^2	xy
1	14	1	14
2	27	4	54
3	46	9	120
4	55	16	220
5	68	25	340
15	204	55	748

$N = 5$, $\Sigma x = 15$

$\Sigma y = 204$, $\Sigma x^2 = 55$

$\Sigma xy = 748$

sub in (2) & (3)

$$5a + 15b = 204 \rightarrow (4)$$

$$15a + 55b = 748 \rightarrow (5)$$

$a = 0$, $b = \frac{68}{5} = 13.6$

Sub a & b in ①

$$y = a + bx$$

$$y = 0 + (13.6)x$$

$$y = (13.6)x$$

ii) Fit straight line eq^n to the full data

x	0	2	5	7
y	-1	5	12	20

We know $y = a + bx$

Normal eq $\sum y = Na + b\sum x$

$$\sum xy = a\sum x + b\sum x^2$$

x	y	x^2	xy
0	-1	0	1
2	5	4	10
5	12	25	60
7	20	49	140
14	36	78	211

$N = 4$, $\Sigma x = 14$, $\Sigma y = 36$

$\Sigma x^2 = 78$, $\Sigma xy = 210$

$$\Rightarrow \quad 36 = 4a + b14$$
$$210 = a14 + b78$$

$$a = -1.1379$$
$$b = 2.8965$$

$$\Rightarrow \quad y = -1.1379 + x(2.8965)$$

is required SLE

iii)

x	1	3	5	7	9
y	1.5	2.8	4.0	4.7	6.0

x	y	x^2	xy
1	1.5	1	1.5
3	2.8	9	8.4
5	4.0	25	20.0
7	4.7	49	32.9
9	6.0	81	54.0
25	19.0	165	116.8

$\Sigma x = 25$, $\Sigma y = 19.0$

$\Sigma x^2 = 165$, $\Sigma xy = 116.8$

$y = a + bx$.

$\Sigma y = Na + b\Sigma x$ | $\Sigma xy = a\Sigma x + b\Sigma x^2$

$19 = 5(a) + b(25)$ | $116.8 = a(25) + b(165)$

$a = 1.075$, $b = 0.545$

$\Rightarrow y = 1.075 + x(0.545)$ //

is the SLE

iv) H.W

x	6	7	7	8	8	8	9	9	10
y	5	5	4	5	4	3	4	3	3

$\Sigma x^2 = 588$ ~~$\Sigma y^2 = 286$~~ $\Sigma xy = 286$

$\Sigma x = 72$ $\Sigma y = 36$.

$36 = 9(a) + b(72)$

$286 = 72(a) + b(588)$

$a = 5.33$ $b = -0.166$

$y = 5.33 - x(0.166)$ //

Parabola eqⁿ :

$y = a + bx + cx^2$ is parabolic eqⁿ

Normal eqⁿ of the parabola eqⁿ are

$$\Sigma y = Na + b\Sigma x + c\Sigma x^2$$

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

$$\Sigma x^2 y = a\Sigma x^2 + b\Sigma x^3 + c\Sigma x^4$$

eg i) Fit a parabola eqⁿ $y = a + bx + cx^2$ for following data

x	1	2	3	4	5
y	10	12	8	10	14

≥ we k=

$$y = a + bx + cx^2$$

Normal eqⁿ

$$\Sigma y = Na + b\Sigma x + c\Sigma x^2$$

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

$$\Sigma x^2 y = a\Sigma x^2 + b\Sigma x^3 + c\Sigma x^4.$$

x	y	xy	x^2y	x^2	x^3	x^4
1	10	10	10	1	1	1
2	12	24	48	4	8	16
3	8	24	72	9	27	81
4	10	40	160	16	64	256
5	14	70	350	25	125	625
15	54	168	640	55	225	979

$\Sigma x = 15$, $\Sigma y = 54$, $\Sigma xy = 168$, $\Sigma x^2y = 640$

$\Sigma x^2 = 55$, $\Sigma x^3 = 225$, $\Sigma x^4 = 979$

$N = 5$

$\Rightarrow \quad 54 = 5a + 15b + 55c$

$\Rightarrow \quad 168 = 15a + 55b + 225c$

$\Rightarrow \quad 640 = 55a + 225b + 979c.$

$a = 14, \quad b = -3.6857$

$c = 0.7142$

$\Rightarrow \quad y = 14 + (-3.6857)x + 0.7142x^2$ //

H.W

i) →

x	0	1	2
y	1	6	17

$\Sigma x = 3$ $\Sigma x^2 = 4$ $\Sigma xy = 40$

$\Sigma y = 24$ $\Sigma x^3 = 9$ $\Sigma x^2 y =$

$\Sigma x^4 = 17$

ii) →

x	1	5	7	9	12
y	10	15	12	15	21

Exponential function :

$$y = a e^{bx} \rightarrow ①$$

$$\log y = \log (a e^{bx})$$

$$\log y = \log a + \log e^{bx}.$$

$$\log y = \log a + bx$$

$$Y = A + BX \rightarrow ②$$

where $Y = \log y$

$A = \log a$

$B = b$

$X = x$

Normal eqn of 2 are

$$\Sigma Y = NA + B \Sigma x$$

$$\Sigma XY = A \Sigma X + B \Sigma X^2$$

i) Fit the curve $y = ae^{bx}$ from the following

x	1	5	7	9	12
$y = F(x)$	10	15	12	15	21

$\Rightarrow$ Given $y = ae^{bx} \longrightarrow (1)$

Taking log on B.S

$\log y = \log a + \log_e e^{bx}$

$\log y = \log a + bx.$

$Y = A + BX$ where $Y = \log y$

$A = \log a$

$B = b$

$X = x$

The normal eqn,

$\Rightarrow \Sigma Y = NA + B\Sigma x$

$\Sigma XY = A\Sigma X + B\Sigma X^2$

$x = X$	y	$\log y = Y$	XY	X^2
1	10	2.302	2.302	1
5	15	2.708	13.54	25
7	12	2.484	17.388	49
9	15	2.708	24.372	81
12	21	3.044	36.528	144
34		13.246	94.184	300

Here $N = 5$, $\Sigma x = 34$, $\Sigma Y = 13.246$

$\Sigma XY = 94.184$, $\Sigma x^2 = 300$

Sub value in 2 & 3

$$5A + 34B = 13.246$$

$$34A + 300B = 94.184$$

soln

$$A = 2.243$$

$$b = B = 0.059$$

$$\log a = A$$

$$a = e^A$$

$a = e^{2.243}$

$a = 9.421$

$\Rightarrow \quad y = 9.421\, e^{0.059x}$

ii) Fit the curve $y = ae^{bx}$ from the Follow

x	2	3	4	5	6
y	8.3	15.4	33.1	65.2	127.4

Sol

$y = ae^{bx}$

$\log y = \log a + \log_e e^{bx}$

$= \log a + bx$

$Y = A + BX$

$Y = \log y$

$X = x$

$B = b$

$X = x.$

Normal eq

$\Sigma Y = NA + B\Sigma X$

$\Sigma XY = A\Sigma X + B\Sigma X^2$

$X = x$	y	$\ln y = Y$	XY	X^2
2	8.3	2.116	4.232	4
3	15.4	2.734	8.202	9
4	33.1	3.499	13.996	16
5	65.2	4.177	20.885	25
6	127.4	4.847	29.082	36
20		17.373	76.397	90

$N = 5$; $\Sigma X = 20$, $\Sigma Y = 17.373$

$\Sigma XY = 76.397$ $\quad \Sigma X^2 = 90$

$$\Rightarrow \quad 5A + 20B = 17.373$$

$$20A + 90B = 76.397$$

$$A = 0.7126$$

$$b = B = 0.6905$$

$$a = e^A = 2.0392$$

$$\therefore \quad y = (2.0392)\, e^{0.6905 x}$$

HW
EF

i)

x	0	1	2	3	4	5	6	7	8
y	20	30	52	77	135	211	326	550	1052

ii)

x	77	100	185	239	285	-
y	2.4	3.4	7.0	11.1	19.6	

Correlation :- Correlation is a staticial analysis which measures an analysis the degree or extent to which two variables fluctuate with reference to each other

Types of Correlation :- It is classified into many types

- Positive & negative Correlation
- Simple & multiple "
- Partial & total "
- linear & non linear "

Method of Studying Correlation

There are two different method to find the relation b/n variables

i) Graphic method

ii) Mathamatical "

Graphic Method :-

i) Scatter diagram

ii) Simple graph

Mathamatical methods.

i) Karl pearson's co-efficient of correlation (KPC)

ii) Spearman's rank co-efficient of correlation (SRC)

iii) Co-efficient of concurrent diviation.

thod

iv) Method of least squares

→ Correlation coefficient denoted by "r"

KPC of correlation :-

$$r = \frac{\Sigma XY}{\sqrt{\Sigma X^2 \Sigma Y^2}}$$

where $X = x - \bar{x}$; $\bar{x}$ is mean of x series

$Y = y - \bar{y}$; $\bar{y}$ is " " y-series

$\Rightarrow \bar{x} = \frac{\Sigma x}{N}$; $\bar{y} = \frac{\Sigma y}{N}$

Note: Correlation co-efficient lies b/n −1 & +1

ex: i) Find if their is any significant correlation b/n the heigh & weight given below:

Height	57	59	62	63	64	65	55	58	57
weight	113	117	126	126	130	129	111	116	112

w.k.

$$Y = \frac{\Sigma XY}{\sqrt{\Sigma X^2 \Sigma Y^2}}$$

$\bar{x} = 60$

$\bar{y} = 120$

x	y	$X = x - \bar{x}$	$Y = y - \bar{y}$	X^2	Y^2	XY
57	113	-3	-7	9	49	21
59	117	-1	-3	1	9	3
62	126	2	6	4	36	12
63	126	3	6	9	36	18
64	130	4	10	16	100	40
65	129	5	9	25	81	45
55	111	-5	-9	25	81	45
58	116	-2	-4	4	16	8
57	112	-3	-8	9	64	24
540	1080			102	472	216

$\bar{x} = \frac{540}{9} = 60$ $\bar{y} = \frac{1080}{9} = 120$

$$r = \frac{216}{\sqrt{102 \times 472}}$$

$$= 0.98$$

ii) Psychological tests of intaligence & engineering ability ~~work~~ were applying to 10 students. Here is a record of ungrouped data showing intelligence ratio.

IR	105	104	102	101	100	99	98	96	93	92
ER	101	103	100	98	95	96	104	92	97	94

x	y	X	Y	X^2	Y^2	XY
105	101	6	3	36	9	18
104	103	5	5	25	25	25
102	100	3	2	9	4	6
101	98	2	0	4	0	0
100	95	1	-3	1	9	-3
99	96	0	-2	0	4	0
98	104	-1	+6	1	36	-6
96	92	-3	-6	9	36	18
93	97	-6	-1	36	1	6
92	94	-7	-4	49	16	28
990	980			170	140	92

99 98

$$r = \frac{\Sigma XY}{\sqrt{\Sigma X^2 . \Sigma Y^2}}$$

$$= \frac{\cancel{170}\ 92}{\sqrt{140 . \cancel{82}\ 170}}$$

$$= 0.59$$

$\therefore$ co-relation co-efficient 0.59

H.W Find KPC co-relation from the data

i)

wages	100	101	102	102	100	99	97	98	96
cost of living	98	99	99	97	95	92	95	94	90

ii)

x	12	9	8	10	11	13	7
y	14	8	6	9	11	12	3

II) When deviations are taken from an assumed mean.

$$r = \frac{\Sigma XY - \frac{\Sigma X \cdot \Sigma Y}{N}}{\sqrt{\left(\Sigma X^2 - \frac{(\Sigma X)^2}{N}\right)\left(\Sigma Y^2 - \frac{(\Sigma Y)^2}{N}\right)}}$$

ex: i) Calculate KPC for the following data

$\frac{356}{10} = 35.6$

x	28	41	40	38	35	33	40	32	36	33
y	23	24	33	34	30	26	28	31	36	38

what inference will you draw from the estimat

$\Sigma x = 356$ $\qquad$ $\Sigma y = 303$

$\bar{x} = \frac{\Sigma x}{N} = \frac{356}{10}$ $\qquad$ $\bar{y} = \frac{\Sigma y}{N} = \frac{303}{10}$

$= 35.6$ $\qquad$ $= 30.3$

x	y	X	Y	XY	X^2	Y^2
28	23	-7	-7	49	49	49
41	24	6	-6	-36	36	36
40	33	5	3	15	25	9
38	34	3	4	12	9	16
35	30	0	0	0	0	0
33	26	-2	-4	8	4	16
40	28	5	-2	-10	25	4
32	31	-3	1	-3	9	1
36	36	1	6	6	1	36
33	38	-2	8	-16	4	36
356	30.3	6	3	19	162	203

$\Sigma X = 6$, $\Sigma Y = 3$, $\Sigma X^2 = 162$, $\Sigma Y^2 = 203$

$\Sigma XY = 25$

$$r = \frac{\Sigma XY - \frac{\Sigma X \Sigma Y}{N}}{\sqrt{\left(\Sigma x^2 - \frac{(\Sigma x)^2}{N}\right)\left(\Sigma y^2 - \frac{(\Sigma Y)^2}{N}\right)}}$$

$$= \frac{25 - \frac{6.3}{10}}{\sqrt{\left(162 - \frac{36}{16}\right)\left(203 - \frac{9}{10}\right)}}$$

$$= \underline{\underline{0.12}}.$$

ii) Find a sutable correlation from the following data.

29 232

Fertilizer	15	18	20	24	30	35	40	50
produce	85	93	95	105	120	130	150	160

938

117.25

x	y	X	Y	X^2	Y^2	XY
15	85	-14	-32	196	1024	448
18	93	-11	-24	121	576	264
20	95	-9	-22	81	484	198
24	105	-5	-12	25	144	60
30	120	+1	+3	1	9	3
35	130	+6	+13	36	169	78
40	150	+11	+33	121	1089	363
50	160	+21	+43	441	1849	903
232	938	0	2	1022	5344	2317

$\bar{x} = 29$ / $\bar{y} = 117.25$

$$r = \frac{\Sigma XY - \frac{\Sigma X \Sigma Y}{N}}{\sqrt{\left(\Sigma X^2 - \frac{(\Sigma X)^2}{N}\right)\left(\Sigma Y^2 - \frac{(\Sigma Y)^2}{N}\right)}}$$

$$= \frac{2317 - 0}{\sqrt{(1022) \times \left(5344 - \frac{4}{8}\right)}}$$

$$r = 0.99$$

H.W Calculate KPC for the following data

i)

x	38	45	46	38	35	38	46	32	36	38
y	28	34	38	34	36					36

ii)

Height of F	65	66	67	67	68	69	71	73
Heigh of S	67	68	64	68	72	70	69	70

III Rank Co-relation co-efficient. (ρ)

$$\rho = 1 - \frac{6 \Sigma D^2}{N(N^2-1)}$$

where ρ is rank co-efficient of co-relation

D^2 sum of the squares of differences of two ranks

N is the number of ~~operation~~ observation

eg i) Following are the ranks obtained by 10 student in 2 students. S & M. to what extent the knowledge of the students in 2 sub is related

---

Statics	1	2	3	4	5	6	7	8	9	10
Maths	2	4	1	5	3	9	7	10	6	8

Statics x	Maths y	$D = x - y$	D^2
1	2	−1	1
2	4	−2	4
3	1	2	4
4	5	−1	1
5	3	2	4
6	9	−3	9
7	7	0	0
8	10	−2	4
9	6	−3	9
10	8	2	4
55	55		40

$$\rho = 1 - \frac{6\Sigma D^2}{N(N^2-1} = 1 - \frac{4\cancel{0}(6)}{1\cancel{0}(99)}$$

$$\rho = 0.75$$

ii) A random sample of 5 college students is selected their grades in math & static, are found to be

	1	2	3	4	5
Maths	85	60	73	40	90
static	93	75	65	50	80

Maths	Stat	D	D^2
85	93	-8	64
60	75	-15	225
73	65	8	64
40	50	-10	100
90	80	10	100
		-15	553

$$\rho = 1 - \frac{6(553)}{5(24)}$$

$= -26.65$ X

Math x	Stat y	Rank in x	Rank in y	$D = x - y$	D^2
85	93	2	1	1	1
60	75	4	3	1	1
73	65	3	4	-1	1
40	50	5	5	0	0
90	80	1	2	-1	1
					4

$$\rho = 1 - \frac{6 \Sigma D^2}{N(N^2-1)}$$

$$= 0.8$$

H.W i) RCC 16 students

x	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
y	1	10	3	4	5	7	2	6	8	11	15	9	14	12	16	13

calculate the RCC

IV ~~Equal~~ or repeated ranks :-

$$\rho = 1 - 6\left\{\frac{\Sigma D^2 + \frac{1}{12}(m^3 - m) + \frac{1}{12}(m^3 - m) + \ldots}{N(N^2 - 1)}\right\}$$

m = the no. of items whose rank are common

i) From the following data calculate the RCC after making adjust for the tied rank.

x	48	33	40	9	16	16	65	24	16	57
y	13	13	24	6	15	4	20	9	6	19

x	y	rank in x	rank in y	D x - y	D^2
48	13	3	5.5	-2.5	6.25
33	13	5	5.5	-0.5	0.25
40	24	4	1	-3	9
9	6	~~8~~ 10	8.5	1.5	2.25
16	15	8	4	4	16
16	4	8	10	-2	4
65	20	1	2	-1	1
24	9	6	7	-1	1
16	6	8	8.5	-0.5	0.25
57	19	2	3	-1	1
					41

$N = 10, \ \Sigma D^2 = 41$

$$\rho = 1 - 6 \left\{ \frac{\Sigma D^2 + \frac{1}{12}(m^3 - m) + \frac{1}{12}(m^3 - m) \ldots}{N(N^2 - 1)} \right.$$

116 - repeated 3 times in X - series $\Rightarrow m = 3$

113 " 2 times in Y " $\Rightarrow m = 2$

6 " 2 time " " " $\Rightarrow m = 2$

$$= 1 - 6 \left\{ \frac{41 + \frac{1}{2}(27 - 3) + \frac{1}{2}(8 - 2) + \frac{1}{2}(8 - 2)}{10(99)} \right\}$$

$= 0.73$

iii)

x	68	64	75	50	64	80	75	40	55	64
y	62	58	68	45	81	60	68	48	50	70

x	y	Rank of x	Rank of y	D	D^2
68	62	4	5	-1	1
64	58	~~5.5~~ → 6	7	-1	1
75	68	2.5	3.5	-1	1
50	45	9	10	-1	1
64	81	~~5.5~~ → 6	1	5	25
80	60	1	6	-5	25
75	65	2.5	3.5	-1	1
40	48	10	9	1	1
55	50	8	8	0	0
64	70	~~5.5~~ → 6	2	4	16
					72

123

12

64 - repeated 3 times in X series $m = 3$

75 - repeated 2 times in X series $\Rightarrow m = 2$

68 - " " " " $m = 2$

$N = 10$ $\quad \Sigma D^2 = 72$

$$\rho = 1 - \frac{6\left\{ \sum D^2 + \frac{1}{12}(m^3 - m) + \frac{1}{12}(m^3 - m) + \cdots \right\}}{N(N^2 - 1)}$$

$$= 1 - \frac{6\left\{ 72 + \frac{1}{12}(27 - 3) + \frac{1}{12}(8 - 2) + \frac{1}{12}(8 - 2) \right\}}{(10).(99)}$$

$$= 0.54$$

iii) A sample of 12 fathers & their elder sons gave the following data about their elder sons calculate RCC

father	65	63	67	64	68	62	70	66	68	67	69	71
son	68	66	68	65	69	66	68	65	71	67	68	70

x	y	R in x	R in y	$D = x - y$	D^2
65	68	9	5.5		
63	66	11	9.5		
67	68	6.5	5.5		
64	65	10	11		
68	69	4.5	3		
62	66	12	9.5		
70	68	2	5.5		
66	65	8	1		
68	71	4.5	8		
67	67	6.5	5.5		
69	68	3	2		
71	70	1			

9/2/19

Regression :-

If the 2 regression lines of y on x and x on y respectively $a_1x + b_1y + c_1 = 0$

$a_2x + b_2y + c_2 = 0$

then the correlation coefficient $r = \sqrt{\frac{a_1}{a_2} \frac{b_2}{b_1}}$

Ex: 1) The lines of regression of y on $\bar{x}$ and $\bar{x}$ on y are $4x - 5y + 33 = 0$

$20x - 9y - 107 = 0$ respectively Calculate $\bar{x}$, $\bar{y}$ and the correlation coefficient of 'y on x' and

Given regression lines are

$4x - 5y + 33 = 0$

$20x - 9y - 107 = 0$

We know that, the two regression lines

$a_1 x + b_1 y + c_1 = 0$

$a_2 x + b_2 y + c_2 = 0$

$\bar{x} = 13$
$\bar{y} = 17$

then $r = \sqrt{\frac{a_1}{a_2} \frac{b_2}{b_1}}$

Here $a_1 = 4$, $b_1 = -5$, $c_1 = 33$, $a_2 = 20$, $b_2 = -9$, $c_2 = -107$

$r = \sqrt{\frac{4}{20} \times \left(\frac{-9}{-5}\right)} = \sqrt{\frac{36}{100}} = \frac{6}{10} = 6 = 0.6$

Correlation coefficient $r = 0.6$

$4x - 5y = -33 \Rightarrow 20x - 25y = -165$

$20x - 9y = 107$ $\quad -20x \mp 9y = -107$

$-\cancel{8y} = -2\cancel{72}$

$-16y = 272 = 17$

$\Rightarrow$ $\Rightarrow \cancel{y = 8}$

$4x - 5(8) = -33$

$\cancel{4x - 40 = 33} \Rightarrow 4x = 7 \Rightarrow x = \frac{7}{4} \Rightarrow 1.75$

$4x - 85 = 33$

~~Mean of x $\bar{x} = 1.75$~~ $\quad x = \frac{52}{4} = x = 13$

~~Mean of y $\bar{y} = 8$.~~ $\quad \bar{x} = 13$

$\bar{y} = 17$

The lines of regression in bivariate distribution are

$x + 9y = 7$ | Find correlation coefficient and $\bar{x}$ & $\bar{y}$

$y + 4x = \frac{49}{3}$

A) Given, 2 regression lines are

$x + 9y = 7$

$y + 4x = \frac{49}{3}$

we know that $a_1 x + b_1 y + c_1 = 0$

$a_2 x + b_2 y + c_2 = 0$

$a_1 = 1,\ b_1 = 9,\ c_1 = -7,\ a_2 = 4,\ b_2 = 1,\ c_2 = -\frac{49}{3}$

$r = \sqrt{\frac{1}{4} \cdot \frac{9}{1}} \Rightarrow r = \sqrt{\frac{9}{4}} = \frac{3}{2} \Rightarrow r = 1.5$

$x + 9y = 7 \Rightarrow 4x + 36y = 28$

$4x + y = \frac{49}{3}$ $\quad$ $4x + y = \frac{49}{3}$

$35y = 28 - \frac{49}{3}$

$y = \left(\frac{28 - 49/3}{35}\right)$

$x = 4,\ y = \frac{1}{3}$

... of y ...

11) ~~3x + 2y =~~ Two random variables have the regression lines with equations $3x + 2y = 26$, $6x + y = 31$. Find the correlation coefficient, mean of x & mean of y

A) Given 2 regression lines are

$$3x + 2y = 26$$
$$6x + y = 31.$$

~~Here $a_1 = 3$, $b_1 = 2$~~

We know that $a_1 x + b_1 y + c_1 = 0$

$a_2 x + b_2 y + c_2 = 0$ then

the correlation coefficient $r = \sqrt{\frac{a_1}{a_2} \frac{b_2}{b_1}}$

$a_1 = 3$, $b_1 = 2$, $c_1 = -26$

$a_2 = 6$, $b_2 = 1$, $c_2 = -31$.

$r = \sqrt{\frac{3}{6} \frac{1}{2}} \Rightarrow r = \sqrt{\frac{1}{4}} \Rightarrow r = \frac{1}{2} = 0.5$

$x = 4$
$y = 7$

$① \times 2 \Rightarrow \quad 6x + 4y = 52$

$② \times 1 \quad \underline{-6x \pm y = -31}$

$3y = 21 \Rightarrow y = 7$

∴ Mean of $\bar{x} = 4$

Mean of $\bar{y} = 7$.

① $3x + 2(7) = 26$

$3x = 26 - 14 \Rightarrow 3x = 12$
$x = 4$

Angle b/w two regression lines

— Let the line of regression x on y & y on x are given by $x - \bar{x} = r \frac{\sigma x}{\sigma y}(y - \bar{y})$

Slope of the line $m_1 = \frac{1}{r} \frac{\sigma y}{\sigma x}$

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

Slope of the line $m_2 = r \frac{\sigma y}{\sigma x}$.

we have $\tan\theta = \frac{m_1 - m_2}{1 + m_1 m_2}$ //

$$\tan\theta = \frac{\frac{1}{r}\frac{\sigma y}{\sigma x} - r\frac{\sigma y}{\sigma x}}{1 + \frac{1}{\cancel{r}}\frac{\sigma y}{\sigma x}\frac{\cancel{r}}{1}\frac{\sigma y}{\sigma x}}.$$

$$= \frac{\frac{1 - r^2}{r}\left(\frac{\sigma y}{\sigma x}\right)}{1 + \left(\frac{\sigma y}{\sigma x}\right)^2}$$

$$\tan\theta = \left(\frac{1-r^2}{r}\right)\left(\frac{\sigma x \cdot \sigma y}{\sigma x^2 + \sigma y^2}\right)$$

⇒ i) if θ is angle b/w two regulation lines & standard deviation y is twice the standard deviation of x & r = 0.25 Find tan θ.

⇒ Given

$$\sigma y = 2\sigma x$$

$$r = 0.25$$

$$\frac{1-r^2}{r} = 3.75$$

we know that ~~standa~~

$$\tan\theta = \left(\frac{1-r^2}{r}\right)\left(\frac{\sigma x \cdot \sigma y}{\sigma x^2 + \sigma y^2}\right)$$

$$= \left(\frac{1-0.25^2}{0.25}\right)\left(\frac{\sigma x \cdot 2\sigma x}{\sigma x^2 + 4\sigma x^2}\right)$$

$$= 3.75\left(\frac{2}{5}\right) = 1.5$$

$$\tan\theta = 1.5$$

$$\theta = 56.30^\circ$$

ii) $\sigma x = \sigma y = \sigma$. ~~tan~~ the angle b/n regression lin is $\tan^{-1}\left(\frac{4}{3}\right)$ find r

$$\tan\theta = \frac{4}{3}$$

$$\sigma x = \sigma y = \sigma$$

$$\Rightarrow \quad \tan\theta = \left(\frac{1-r^2}{r}\right)\left(\frac{\sigma x \cdot \sigma y}{\sigma x^2 + \sigma y^2}\right)$$

$$\frac{4}{3} = \left(\frac{1-r^2}{r}\right)\left(\frac{\sigma^2}{\sigma^2+\sigma^2}\right)$$

$$\frac{4}{3} = \frac{1-r^2}{r}\left(\frac{1}{2}\right)$$

$$\Rightarrow \quad \frac{8}{3} = \frac{1-r^2}{r} \quad \Rightarrow \quad 8r = 3 - 3r^2$$

$$\Rightarrow \quad 3r^2 + 8r - 3$$

$$r = \frac{1}{3} = 0.33$$

$$r = -3 \quad \times$$

$$\therefore r = 0.33$$

iii) the tangent of the angle b/n two regerant ulm is 0.6

$$\sigma x = \frac{1}{2} \sigma y$$

$$\tan^{-1}(0.6) = \cancel{0.01}\ 30.96$$

$$\sigma x = \frac{1}{2} \sigma y. \quad ; \quad \sigma y = 2\sigma x.$$

$$\tan\theta = \left(\frac{1-r^2}{r}\right)\left(\frac{\sigma x \cdot \sigma y}{\sigma x^2 + \sigma y^2}\right)$$

$$0.6 = \left(\frac{1-r^2}{r}\right)\left(\frac{\sigma x \cdot 2\sigma x}{\sigma x^2 + 4\sigma y^2}\right)$$

$$\frac{5}{2} \times 0.6 = \frac{1-r^2}{r}$$

$$1.5 = \frac{1-r^2}{r}$$

$$r^2 + 1.5r - 1 = 0$$

$$r = \frac{1}{2} = 0.5 \,//$$

$$r = -2 \ \times$$

Note the Co-relation of $r = \dfrac{cov(x,y)}{\sigma x . \sigma y}$

where

$$cov(x,y) = \frac{\Sigma(x-\bar{x})(y-\bar{y})}{N}$$

$$\sigma x = \sqrt{\frac{\Sigma(x-\bar{x})^2}{N}} \; ; \; \sigma y = \sqrt{\frac{\Sigma(y-\bar{y})^2}{N}}$$

The regression line on x on y is

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

The regression line on y on x is

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

Q i) Calculate the co-efficient of co-relation & obtain the least square regression lines for the following data

x	1	2	3	4	5	6	7	8	9
y	9	8	10	12	11	13	14	16	15

x
1
2
3
4
5
6
7
8
9
45

$\bar{x}$ =
$\bar{y}$
σ
σ
co
r =

x	y	$x-\bar{x}$	$y-\bar{y}$	$(x-\bar{x})^2$	$(y-\bar{y})^2$	$(x-\bar{x})(y-\bar{y})$
1	9	-4	-3	16	9	12
2	8	-3	-4	9	16	12
3	10	-2	-2	4	4	4
4	12	-1	0	1	0	0
5	11	0	-1	0	1	0
6	13	1	1	1	1	1
7	14	2	2	4	4	4
8	16	3	4	9	16	12
9	15	4	3	16	9	12
45	108			60	60	56

$\bar{x} = 5$ $\quad N = 9$ $\quad \Sigma(x-\bar{x})^2 = 60$ $\quad \Sigma(x-\bar{x})(y-\bar{y}) = 56$

$\bar{y} = 12$ $\quad \Sigma(y-\bar{y})^2 = 60$

$$\sigma x = \sqrt{\frac{\Sigma(x-\bar{x})^2}{N}} = \sqrt{\frac{(60)^2}{9}} = \cancel{20}\ 2.581$$

$$\sigma y = \sqrt{\frac{\Sigma(y-\bar{y})^2}{N}} = \sqrt{\frac{(60)^2}{9}} = \cancel{20}\ 2.581$$

$$\operatorname{cov}(x,y) = \frac{\Sigma(x-\bar{x})(y-\bar{y})}{N} = \frac{56}{9} = 6.33$$

$$r = \frac{6.33}{\cancel{20 \times 20}\ 2.581 \times 2.581} = 0.95$$

case (i) x on y

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

$$(x - 5) = 0.95\left(\frac{\cancel{2.581}}{\cancel{2.581}}\right)(y - 12)$$

$$\Rightarrow \quad x - 5 = 0.95y - 11.4$$

$$\Rightarrow \quad x - 0.95y + 6.4 = 0$$

case (ii) y on x

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

$$(y - 12) = 0.95\left(\frac{\cancel{2.581}}{\cancel{2.581}}\right)(x - 5)$$

$$y - 12 = 0.95x - 4.75$$

$$0.95x - y + 7.25 = 0$$

→ Find the co-relation co-efficient b/n x & y for the following data & also find two regration lines

1	2	3	4	5	6	7	8	9	10
10	12	16	28	25	36	41	49	40	50

x	y	$x-\bar{x}$	$y-\bar{y}$	$(x-\bar{x})^2$	$(y-\bar{y})^2$	$(x-\bar{x})(y-\bar{y})$
1	10	−4	−20	16	400	80
2	12	−3	−18	9	324	54
3	16	−2	−14	4	196	28
4	28	−1	−2	~~2~~ 1	4	2
5	25	0	−5	0	25	0
6	36	1	+6	1	36	6
7	41	2	+11	4	121	22
8	49	3	+19	9	361	57
9	40	4	+10	16	100	40
10	50	5	+20	25	400	100
55	307			85	1967	389

$\Rightarrow 5.5 = \bar{x}$ $\quad \bar{y} = 30.7$

$\bar{x} = 5 \quad \bar{y} = 30$

$$\text{cov}(x,y) = \frac{\Sigma(x-\bar{x})(y-\bar{y})}{N}$$

$$= \frac{389}{10} = 38.9$$

$$r = \frac{38.9}{\sigma x \cdot \sigma y} \qquad \sigma x = \sqrt{\frac{\Sigma (x-\bar{x})^2}{N}}$$

$$= \frac{38.9}{2.9154 \times 14.0249} \qquad = \sqrt{\frac{85}{10}}$$

$$= 0.905 \qquad = 2.9154$$

i) x on y

$$\sigma y = \sqrt{\frac{\Sigma (y-\bar{y})^2}{N}}$$

$$(x - \bar{x}) = r \frac{\sigma x}{\sigma y} (y - \bar{y}) \qquad = \sqrt{\frac{1967}{10}}$$

$$(x-5) = 0.95\left(\frac{2.9154}{14.0249}\right)(y-30) \qquad = 14.0249$$

$$(x-5) = 0.19(y-30)$$

$$x - 5 = 0.19y - 5.7$$

$$x - 0.19y + 0.7 = 0$$

ii) y on x.

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

$$(y-30) = 4.57(x-5)$$

$$(y-30) = 4.57x - 22.85$$

$$4.57x - y + 7.15 = 0$$

iii) Find r & 2 regration line

x	1	3	4	5	7	8	10
y	2	6	8	10	14	16	20

x	y	$x-\bar{x}$	$y-\bar{y}$	$(x-\bar{x})^2$	$(y-\bar{y})^2$	$(x-\bar{x})(y-\bar{y})$
1	2	$-4 \cancel{2}$	$\cancel{5}\ -8$	16	64	32
3	6	$-2 \cancel{0}$	$\cancel{1}\ -4$	4	16	8
4	8	$-1 \cancel{7}$	$\cancel{7}\ -2$	1	4	2
5	10	$0 \cancel{2}$	$\cancel{2}\ 0$	0	0	0
7	14	$2 \cancel{4}$	$\cancel{7}\ 4$	4	16	8
8	16	$3 \cancel{5}$	$\cancel{4}\ 6$	9	36	18
10	20	$5 \cancel{7}$	$\cancel{13}\ 10$	25	100	50
38	76			59	236	118

$\bar{x} = \cancel{5.4} = \boxed{5}$

$\cancel{\bar{y}\ 7.6 = 7}$

$\bar{y} = 10.8 = 16$

$$\sigma x = \sqrt{\frac{\Sigma(x-\bar{x})^2}{N}} = 2.90$$

$$\sigma y = \sqrt{\frac{\Sigma(y-\bar{y})^2}{N}} = 5.80$$

$$\operatorname{cov}(x,y) = \frac{\Sigma(x-\bar{x})(y-\bar{y})}{N} = 16.85$$

$$r = \frac{16.8\cancel{5}}{2.90 \times 5.80} = 0.99$$

i) x on y

$$(x-5) = 0.99\left(\frac{2.90}{5.80}\right)(y-10)$$

$$(x-5) = 0.495(y-10)$$

$$(x-5) = 0.495y - 4.95$$

$$x - 0.495y - 0.05 = 0$$

ii) y on x

$$(y-10) = 0.99\left(\frac{5.80}{2.90}\right)(x-5)$$

$$(y-10) = 1.98(x-5)$$

$$y - 10 = 1.98x - 9.9$$

$$1.98x - y + 0.1 = 0$$

iv) find co-relation b/n x & y

x	55	56	58	59	60	60	62
y	35	38	38	39	44	43	45

x	y	$x-\bar{x}$	$y-\bar{y}$	$(x-\bar{x})^2$	$(y-\bar{y})^2$	$(x-\bar{x})(y-\bar{y})$
55	35	-3	-5	9	25	15
56	38	-2	-2	4	4	4
58	38	0	-2	0	4	0
59	39	1	-1	1	1	-1
60	44	2	4	4	16	8
60	43	2	3	4	9	6
62	45	4	5	16	25	20
410	282			38	84	52

$$\bar{x} = \frac{410}{7} = 58.57 = 58$$

$$\bar{y} = \frac{282}{7} = 40.28 = 40$$

$$\sigma x = \sqrt{\frac{38}{7}} = 2.32 \qquad \sigma y = \sqrt{\frac{84}{7}} = 3.46$$

$$\text{cov}(xy) = \frac{\Sigma (x-\bar{x})(y-\bar{y})}{7} = \frac{52}{7} = 7.42$$

$$r = \frac{7.42}{2.32 \times 3.46} = 0.924$$

i) x on y

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

$$(x - 58) = 0.92 \left(\frac{2.32 \; \cancel{3.46}}{3.46} \right) y - 40)$$

$$x - 58 = 0.616 (y - 40)$$

$$x - 58 = y - 24.67$$

$$x - \overset{0.616}{y} - 33.33 = 0$$

ii) y on x

$$\cancel{x} \; (y - 40) = 0.92 \left(\frac{3.46}{2.32} \right) (x - 58)$$

$$(y - 40) = 1.372 (x - 58)$$

$$y - 40 = \overset{1.372}{x} - 79.58$$

$$0 = \overset{1.372}{x} - y - 39.58$$