
Correlation:
Ht
Wt
Page-2
(A) Spearman's Rank (R)
Non-Repeated Rank
Repeated Rank
$R = 1 - \frac{6 \Sigma d^2}{n^3 - n}$
$R = 1 - \frac{6\left[\Sigma d^2 + \frac{1}{12}(m_1^3 - m_1) + \frac{1}{12}(m_2^3 - m_2) \ldots\right]}{n^3 - n}$
m = Rank कितनी बार Repeat हुआ है ?
$d = R_1 - R_2$ OR $R_2 - R_1$

Ex 1. Find Spearman's Rank coeff of correlation

X :	10	13	12	15	17
Y :	5	6	3	2	1

Sol:

Sl.No	X	R_1	Y	R_2	$d = R_1 - R_2$	d^2
1	10	5	5	2	3	9
2	13	3	6	1	2	4
3	12	4	3	3	1	1
4	15	2	2	4	-2	4
n=5	17	1	1	5	-4	16
						$\Sigma d^2 = 34$

~~X 1,2,3,4,5~~ ~~Y 1,2,3,4,5~~

$$R = 1 - \frac{6\Sigma d^2}{n^3 - n}$$

$$= 1 - \frac{6 \times 34}{5^3 - 5}$$

$$= 1 - \frac{6 \times 34}{120}$$

$$R = -0.7 \quad (-1 \le R \le 1)$$

Ex 2. Find Spearman's Rank coeff of correlation

X : 32 55 49 60 43 37 43 49 10 20

Y : 40 30 70 20 30 50 72 60 45 25

Sl NO	X	R_1	Y	R_2	$d = R_1 - R_2$	d^2
1	32	8	40	6	2	4
2	55	2	30	7.5	-5.5	30.25
3	49	3.5	70	2	1.5	2.25
4	60	1	20	10	-9	81
5	43	5.5	30	7.5	-2	4
6	37	7	50	4	3	9
7	43	5.5	72	1	4.5	20.25
8	49	3.5	60	3	0.5	0.25
9	10	10	45	5	5	25
n=10	20	9	25	9	0	0

$\Sigma d^2 = 176$

3 3 4 4 5 5 7 8 9 10

$\frac{3+4}{2} = 3.5$ 5.5

$m_1 = 2, m_2 = 2, m_3 = 2$

$$R = 1 - \frac{6\left[\Sigma d^2 + \frac{1}{12}(m_1^3 - m_1) + \frac{1}{12}(m_2^3 - m_2) + \frac{1}{12}(m_3^3 - m_3)\right]}{n^3 - n}$$

$$= 1 - \frac{6\left[176 + \frac{1}{12}(6) + \frac{1}{12}(6) + \frac{1}{12}(6)\right]}{10^3 - 10}$$

$$= 1 - \frac{6[176 + 0.5 + 0.5 + 0.5]}{990}$$

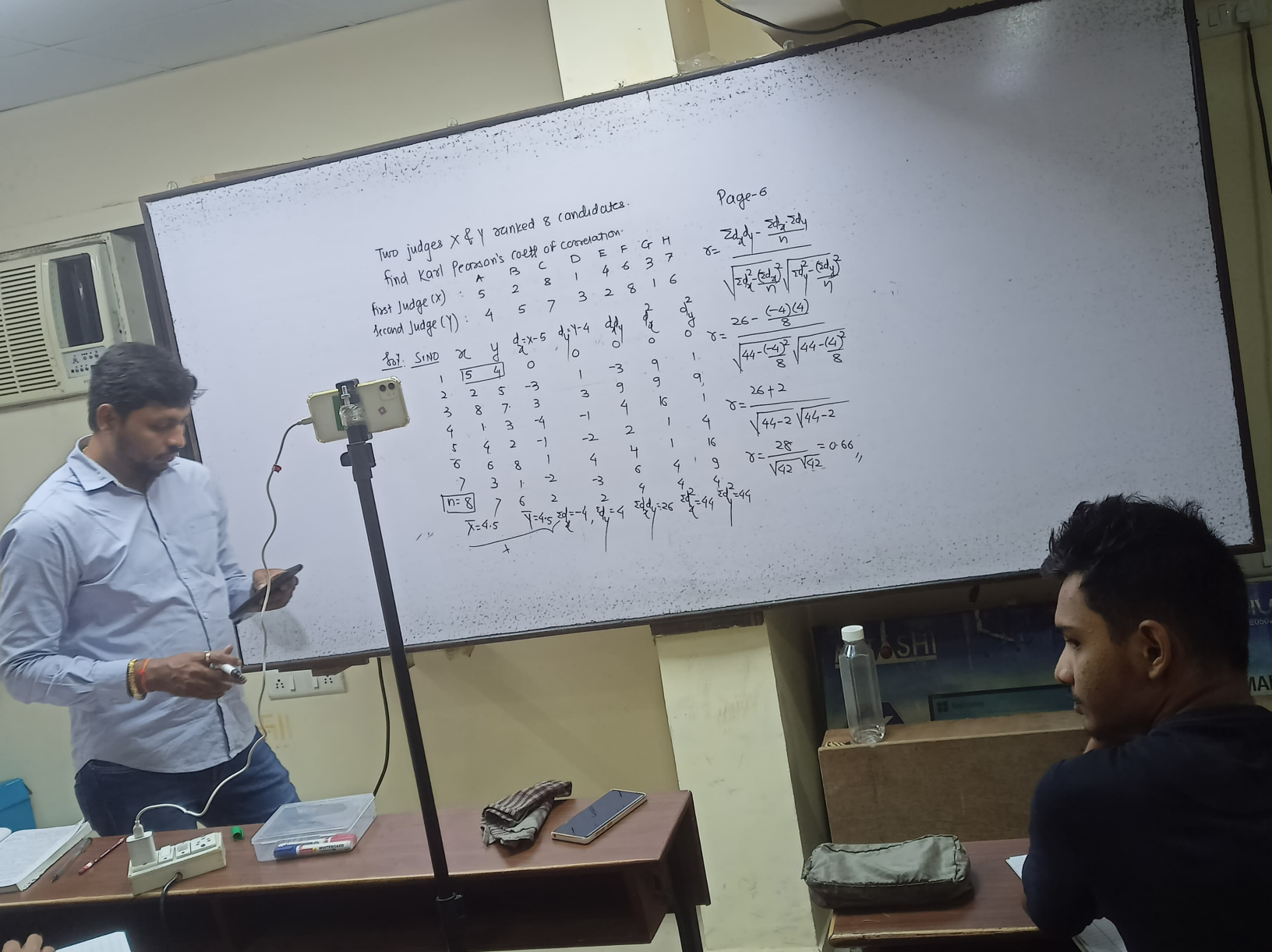
$R = -0.25$ $\quad -1 \le R \le 1$

(B) Karl pearson's coefficient of correlation.

$$r = \frac{\Sigma d_x d_y - \frac{\Sigma d_x \cdot \Sigma d_y}{n}}{\sqrt{\Sigma d_x^2 - \frac{(\Sigma d_x)^2}{n}} \sqrt{\Sigma d_y^2 - \frac{(\Sigma d_y)^2}{n}}} \quad \text{(Table)}$$

$$= \frac{\Sigma xy - \frac{\Sigma x \cdot \Sigma y}{n}}{\sqrt{\Sigma x^2 - \frac{(\Sigma x)^2}{n}} \sqrt{\Sigma y^2 - \frac{(\Sigma y)^2}{n}}} \quad \left(\Sigma x, \Sigma y, \Sigma xy \right)$$

$$= \frac{\frac{\Sigma (x-\bar{x})(y-\bar{y})}{n}}{\sqrt{\frac{\Sigma (x-\bar{x})^2}{n}} \sqrt{\frac{\Sigma (y-\bar{y})^2}{n}}} \quad \left(\Sigma (x-\bar{x}), \Sigma (y-\bar{y}) \right)$$


Page-6
Two judges X & Y ranked 8 candidates.
Find Karl Pearson's coeff of correlation.
A B C D E F G H
first Judge (x) : 5 2 8 1 4 6 3 7
Second Judge (Y) : 4 5 7 3 2 8 1 6
Sol: SIND
n = 8
X̄ = 4.5
Ȳ = 4.5
Σd_x = -4, Σd_y = 4
Σd_x d_y = 26
Σd_x² = 44
Σd_y² = 44
r = 26 / √(44-2)√(44-2)
r = 28/42 = 0.66

Regression

(A) Regression coefficient (b)

- y on x
- x on y

$$b_{yx} = r\frac{\sigma_y}{\sigma_x} = \frac{\frac{\Sigma d_x d_y - \frac{\Sigma d_x \cdot \Sigma d_y}{n}}{n}}{\Sigma d_x^2 - \frac{(\Sigma d_x)^2}{n}}$$

$$b_{xy} = r\frac{\sigma_x}{\sigma_y} = \frac{\Sigma d_x d_y - \frac{\Sigma d_x \Sigma d_y}{n}}{\Sigma d_y^2 - \frac{(\Sigma d_y)^2}{n}}$$

(B) Lines of regression

- y on x
- x on y

$y - \bar{y} = b_{yx}(x - \bar{x})$ $\quad x - \bar{x} = b_{xy}(y - \bar{y})$

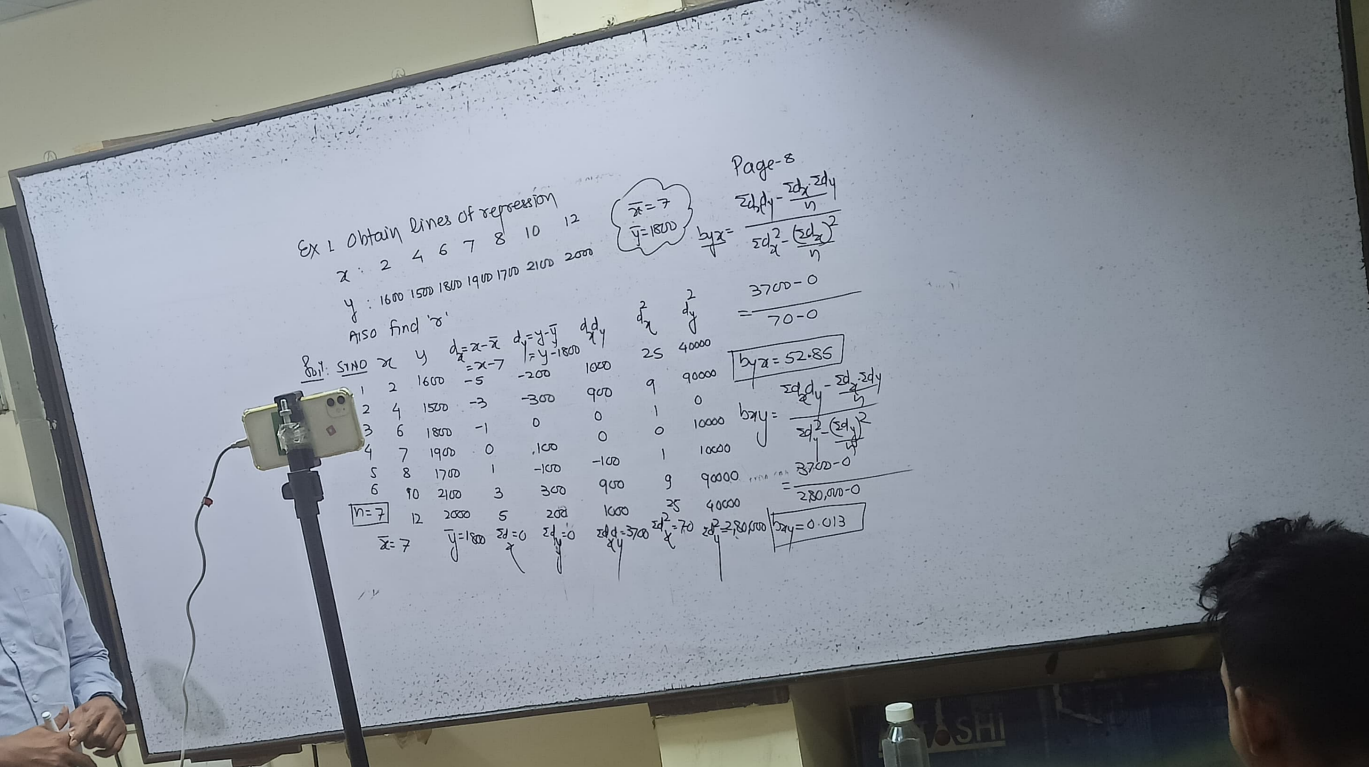
(C) Angle between lines

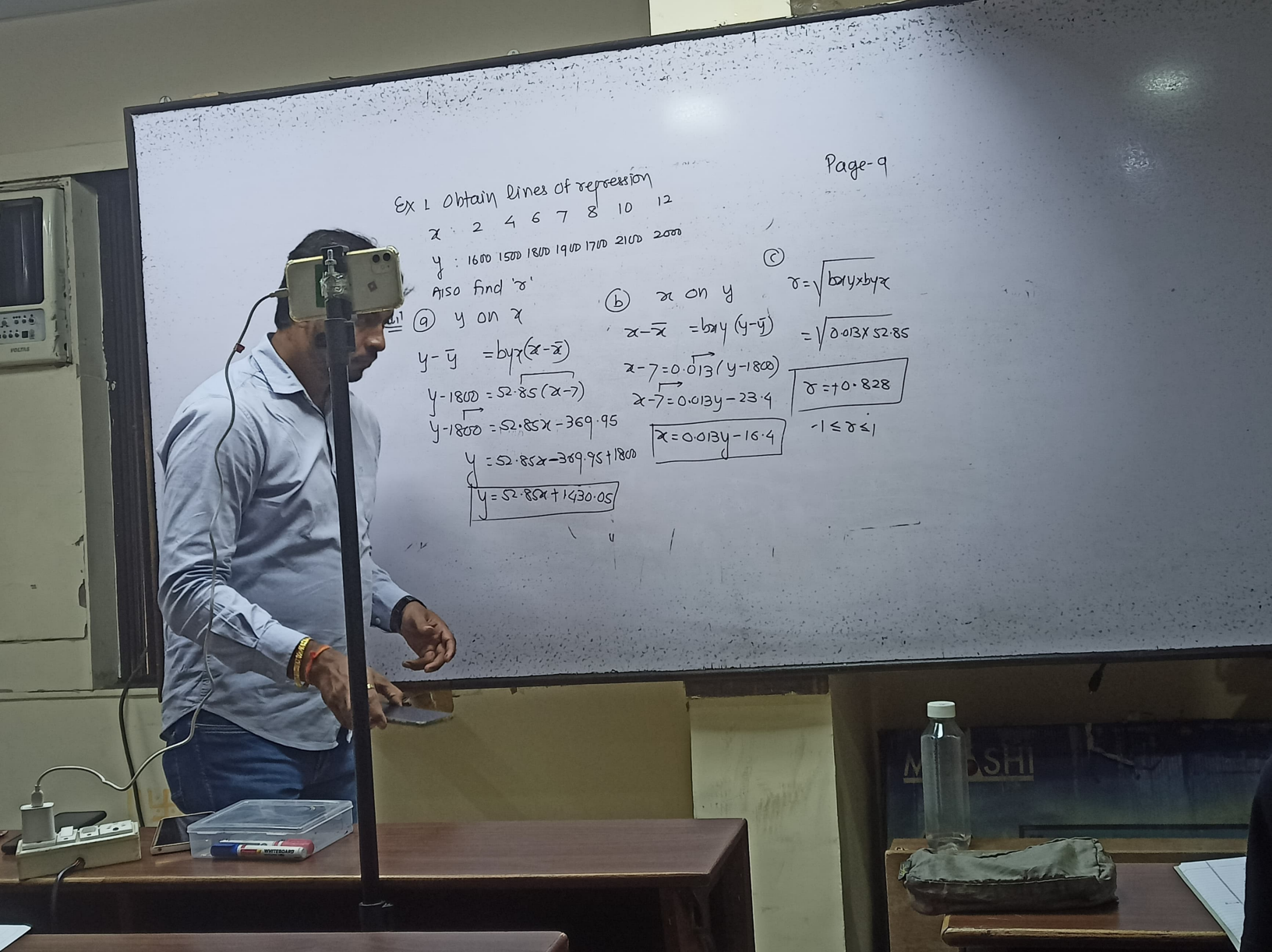
$$\tan\theta = \left(\frac{1-r^2}{r}\right)\left(\frac{\sigma_x \cdot \sigma_y}{\sigma_x^2 + \sigma_y^2}\right)$$

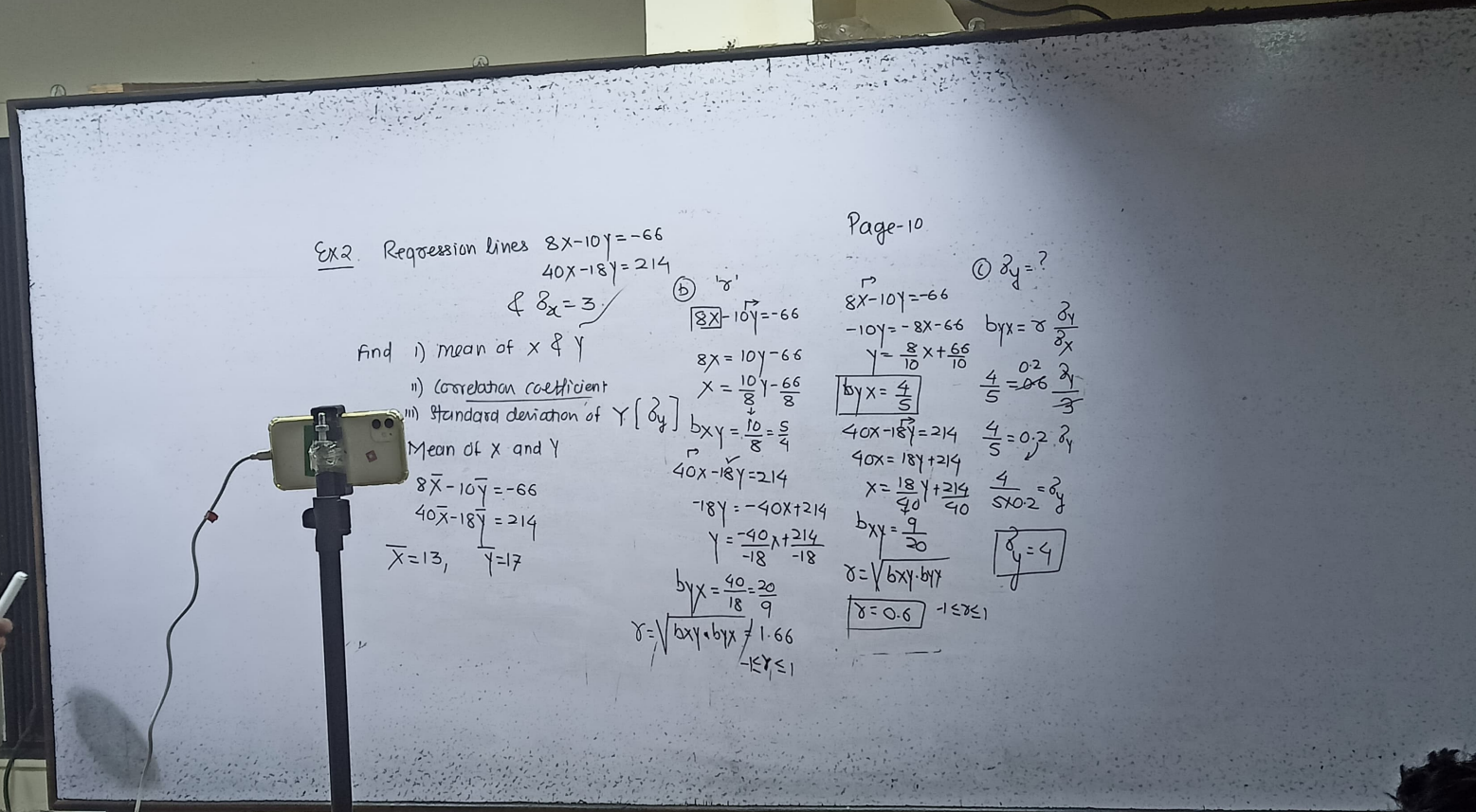
(D) Note

i) $r = \sqrt{b_{xy} * b_{yx}}$

ii) Sign of 'r' is same as sign of b_{xy} & b_{yx}


Page-8
Ex 1: Obtain lines of regression
x : 2 4 6 7 8 10 12
y : 1600 1500 1800 1900 1700 2100 2000
Also find 'r'
$\bar{x}=7$
$\bar{y}=1800$
$b_{yx}=\frac{\Sigma d_x d_y-\frac{\Sigma d_x \Sigma d_y}{n}}{\Sigma d_x^2-\frac{(\Sigma d_x)^2}{n}}$
$=\frac{3700-0}{70-0}$
$b_{yx}=52.85$
$b_{xy}=\frac{\Sigma d_x d_y-\frac{\Sigma d_x \Sigma d_y}{n}}{\Sigma d_y^2-\frac{(\Sigma d_y)^2}{n}}$
$=\frac{3700-0}{280,000-0}$
$b_{xy}=0.013$
By: S.NO x y $d_x=x-\bar{x}=x-7$ $d_y=y-\bar{y}=y-1800$ $d_x d_y$ d_x^2 d_y^2
1 2 1600 -5 -200 1000 25 40000
2 4 1500 -3 -300 900 9 90000
3 6 1800 -1 0 0 1 0
4 7 1900 0 100 0 0 10000
5 8 1700 1 -100 -100 1 10000
6 10 2100 3 300 900 9 90000
n=7 12 2000 5 200 1000 25 40000
$\bar{x}=7$ $\bar{y}=1800$ $\Sigma d_x=0$ $\Sigma d_y=0$ $\Sigma d_x d_y=3700$ $\Sigma d_x^2=70$ $\Sigma d_y^2=280,000$


Page-9
Ex 1 Obtain lines of regression
x : 2 4 6 7 8 10 12
y : 1600 1500 1800 1900 1700 2100 2000
Also find 'r'
(a) y on x
y - ȳ = byx(x - x̄)
y - 1800 = 52.85(x - 7)
y - 1800 = 52.85x - 369.95
y = 52.85x - 369.95 + 1800
y = 52.85x + 1430.05
(b) x on y
x - x̄ = bxy(y - ȳ)
x - 7 = 0.013(y - 1800)
x - 7 = 0.013y - 23.4
x = 0.013y - 16.4
(c)
r = √(bxy × byx)
= √(0.013 × 52.85)
r = +0.828
-1 ≤ r ≤ 1


Ex2 Regression lines $8X-10Y=-66$
$40X-18Y=214$
& $\sigma_x=3$
Find i) Mean of X & Y
ii) Correlation coefficient
iii) Standard deviation of Y [σ_y]
Mean of X and Y
$8\bar{X}-10\bar{Y}=-66$
$40\bar{X}-18\bar{Y}=214$
$\bar{X}=13$, $\bar{Y}=17$
(b)
$8X-10Y=-66$
$8X=10Y-66$
$X=\frac{10}{8}Y-\frac{66}{8}$
$b_{XY}=\frac{10}{8}=\frac{5}{4}$
$40X-18Y=214$
$-18Y=-40X+214$
$Y=\frac{-40}{-18}X+\frac{214}{-18}$
$b_{YX}=\frac{40}{18}=\frac{20}{9}$
$r=\sqrt{b_{XY}\cdot b_{YX}}=1.66$
$-1\leq r\leq 1$
Page-10
① $\sigma_y=?$
$8X-10Y=-66$
$-10Y=-8X-66$
$Y=\frac{8}{10}X+\frac{66}{10}$
$b_{YX}=\frac{4}{5}$
$40X-18Y=214$
$40X=18Y+214$
$X=\frac{18}{40}Y+\frac{214}{40}$
$b_{XY}=\frac{9}{20}$
$r=\sqrt{b_{XY}\cdot b_{YX}}$
$r=0.6$ $-1\leq r\leq 1$
$b_{YX}=r\frac{\sigma_y}{\sigma_x}$
$\frac{4}{5}=0.6\frac{\sigma_y}{3}$
$\frac{4}{5}=0.2\,\sigma_y$
$\frac{4}{5\times 0.2}=\sigma_y$
$\sigma_y=4$

If the tangent of angle made by line of regression of y on x is 0.6 and $\sigma_x = \frac{1}{2}\sigma_y$. find correlation coefficient. $\sigma_y = 2\sigma_x$

Sol: $\tan\theta = 0.6$ & $\boxed{\sigma_y = 2\sigma_x}$

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

$$0.6 = \left(\frac{1-r^2}{r}\right)\left[\frac{\sigma_x \cdot 2\sigma_x}{\sigma_x^2 + 4\sigma_x^2}\right]$$

$$0.6 = \left(\frac{1-r^2}{r}\right)\left[\frac{2\sigma_x^2}{5\sigma_x^2}\right]$$

$$0.6 = \left(\frac{1-r^2}{r}\right)\left(\frac{2}{5}\right)$$

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

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

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

$$r = 0.5 \text{ & } r \neq -2$$

Curve fitting

$\Sigma 1 = n$

Straight line

$y = a + bx$

$\Sigma y = \Sigma a + \Sigma bx$

$\Sigma y = a\Sigma 1 + b\Sigma x$

$\boxed{\Sigma y = an + b\Sigma x}$

$xy = ax + bx^2$

$\boxed{\Sigma xy = a\Sigma x + b\Sigma x^2}$

Parabola

$y = a + bx + cx^2$

$\boxed{\Sigma y = an + b\Sigma x + c\Sigma x^2}$

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

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

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

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

$\Sigma a = a\Sigma 1$

$= an$

Ex 1. Fit a straight line to the following data.

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

Sol: a) Equation

$y = a + bx$

$\Sigma y = an + b\Sigma x$ — ①

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

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

b) Table

S.NO	x	y	xy	x^2
1	3	8	24	9
2	4	6	24	16
3	5	5	25	25
4	6	7	42	36
5	7	6	42	49
$n=6$	8	4	32	64

$\Sigma x = 33$, $\Sigma y = 36$, $\Sigma xy = 189$, $\Sigma x^2 = 199$

c) Calculation

$\Sigma y = an + b\Sigma x$

$36 = 6a + 33b$

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

$189 = 33a + 199b$

$a = 8.82$ & $b = -0.51$

$\because y = a + bx$

$y = 8.82 - 0.51x$

Ex·2.- Fit a second degree curve.

x	0	1	2	3	4
y	1	1·8	2·3	5·5	6·3

Sol:

$\Sigma y = an + b\Sigma x + c\Sigma x^2$ —①

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

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

ⓑ Table

SlNO	x	y	x^2	xy	x^3	x^2y	x^4
1	0	1	0	0	0	0	0
2	1	1·8	1	1·8	1	1·8	1
3	2	2·3	4	4·6	8	9·2	16
4	3	5·5	9	16·5	27	49·5	81
n=5	4	6·3	16	25·2	64	100·8	256

$\Sigma x = 10$ $\Sigma y = 16·9$ $\Sigma x^2 = 30$ $\Sigma xy = 48·1$ $\Sigma x^3 = 100$ $\Sigma x^2y = 161·3$ $\Sigma x^4 = 354$

ⓒ Calculation.

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

$16·9 = 5a + 10b + 30c$ $48·1 = 10a + 30b + 100c$ $161·3 = 30a + 100b + 354c$

$a = 0·91$

$b = 0·65$

$c = 0·19$

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

$y = 0·91 + 0·65x + 0·19x^2$

Ex·2:- Fit a second degree curve.

x	0	1	2	3	4
y	1	1·8	2·3	5·5	6·3

Sol: a) Equation

$y = a + bx + cx^2$

$\Sigma y = an + b\Sigma x + c\Sigma x^2$ —①

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

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

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

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

b) Table

Sl NO	x	y	x^2	xy	x^3	x^2y	x^4
1	0	1	0	0	0	0	0
2	1	1·8	1	1·8	1	1·8	1
3	2	2·3	4	4·6	8	9·2	16
4	3	5·5	9	16·5	27	49·5	81
$n=5$	4	6·3	16	25·2	64	100·8	256

$\Sigma x = 10$ $\Sigma y = 16·9$ $\Sigma x^2 = 30$ $\Sigma xy = 48·1$ $\Sigma x^3 = 100$ $\Sigma x^2y = 161·3$ $\Sigma x^4 = 354$

c) Calculation.

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

$16·9 = 5a + 10b + 30c$ $48·1 = 10a + 30b + 100c$ $161·3 = 30a + 100b + 354c$

$a = 0·91$

$b = 0·65$

$c = 0·19$

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

$y = 0·91 + 0·65x + 0·19x^2$