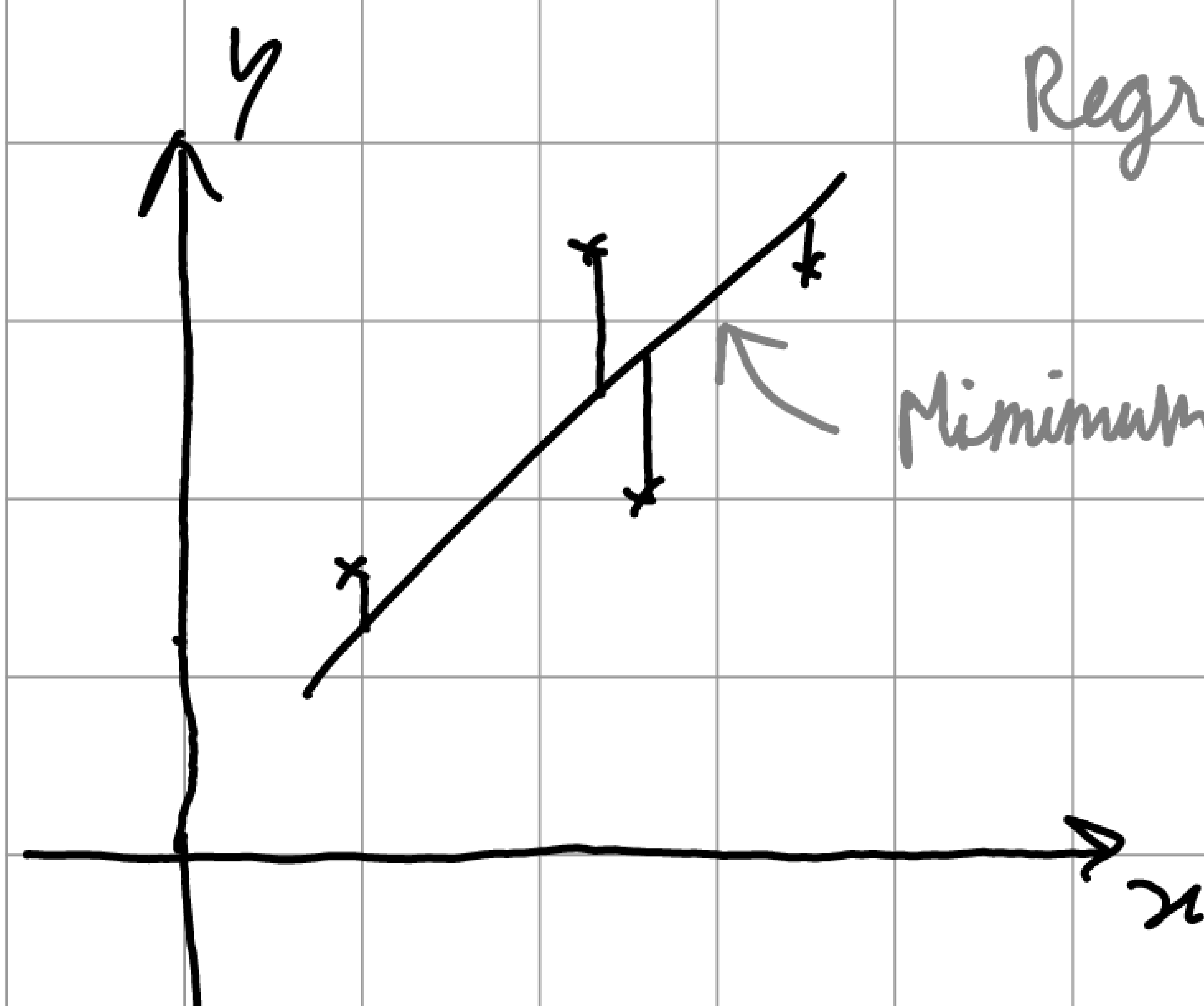
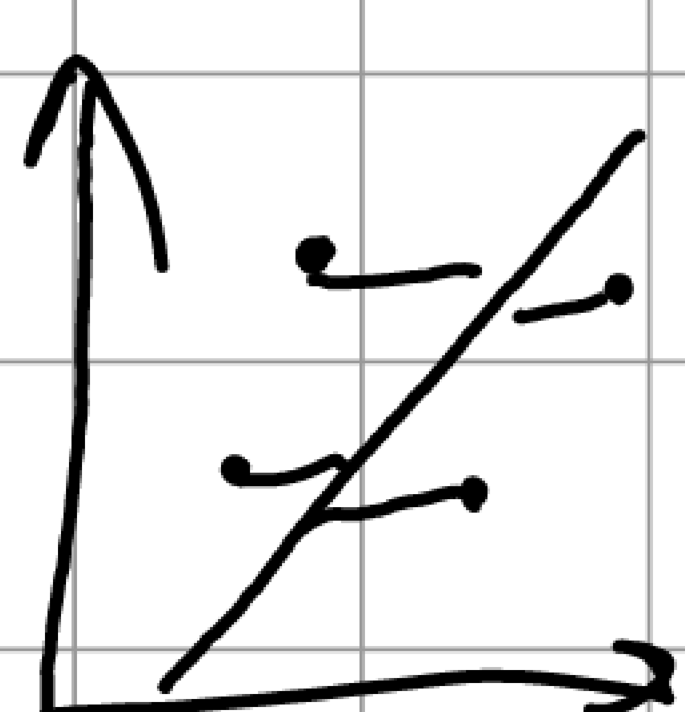
Regression lines

If two variables are corelated we can get value of other variable if value of one variable is given



Regression line of y on x

Minimum distance lines



Reg of x on y

Regression line of y on x is given by

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

y on x

$$b_{yx} = r \frac{\sigma_y}{\sigma_x}$$

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

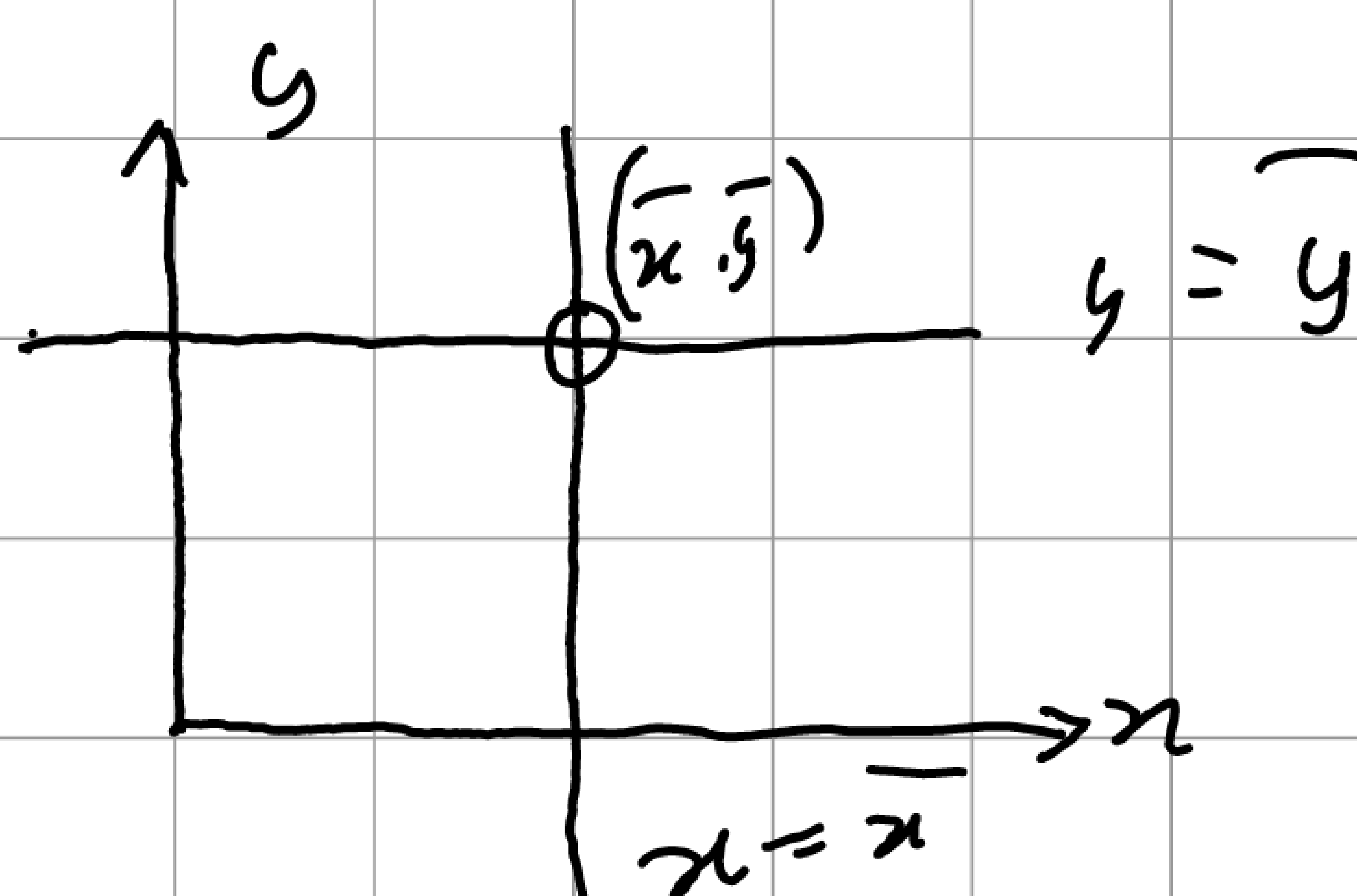
$$b_{xy} = r \frac{\sigma_x}{\sigma_y}$$

Notes

① If $r = 0$

$y = \bar{y}$

$x = \bar{x}$



② If $r = \pm 1$ Both lines are coincident

$$b_{yx} = \frac{\sum xy - N\bar{x}\bar{y}}{\sum x^2 - n\bar{x}^2}$$

Properties i) $r = \pm\sqrt{b_{yx}\, b_{xy}}$

ii) $b_{xy} > 1$ $b_{yx} < 1$

iii) $\frac{b_{xy} + b_{yx}}{2} \geq r$

iv) b_{xy} & b_{yx} independant of change of origin but not change of scale

Angle between two regression lines is

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

① Panel of 2 judges a & b graded the performance by awarding marks as follows

	1	2	3	4	5	6	7	8
a	36	32	34	31	32	32	34	38
b	35	33	31	30	34	32	36	–

judge b could not attend 8^{th} performance
If b was present how many marks he would have given to 8^{th} performance

$\longrightarrow$

To find y given x $\quad b_{yx}$

$$b_{yx} = \frac{\Sigma yx - \bar{x}\bar{y}N}{\Sigma x^2 - N\bar{x}^2}$$

x	y	u ($x-30$)	v ($y-30$)	u^2	v^2	uv
36	35	6	5	36	25	30
32	33	2	3	4	9	6
34	31	4	1	16	1	4
31	30	1	0	1	0	0
32	34	2	4	4	16	8
32	32	2	2	4	4	4
34	36	4	6	16	36	24
	Σ	21	21	81	91	76

$$\bar{u} = \frac{21}{7} = 3$$

$$\bar{v} = \frac{21}{7} = 3$$

$$\frac{\Sigma uv - N\bar{u}\bar{v}}{\Sigma u^2 - N\bar{u}^2}$$

$$= \frac{76 - 7 \times 3 \times 3}{81 - 7 \times 3 \times 3}$$

$$= \frac{13}{18} = 0.72$$

$$v - \bar{v} = b_{vu}(u - \bar{u})$$

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

$$x - 33 = 0.72(38 - 33)$$

$$\therefore x \sim 36$$

$N = 12$

$\Sigma x = 120$

$\Sigma y = 432$

$\Sigma xy = 4992$

$\Sigma x^2 = 1392$

$\Sigma y^2 = 18252$

find b_{xy}

$\bar{x} = 10$

$\bar{y} = 36$

$$b_{xy} = \frac{\Sigma xy - n\bar{x}\bar{y}}{\Sigma y^2 - n\bar{y}^2} = \frac{4992 - 10 \times 36 \times 12}{18252 - 12 \times 36 \times 36}$$

$= 0.259$

$$b_{yx} = \frac{4992 - 10 \times 36 \times 12}{1392 - 10 \times 10 \times 12} = 0.259$$

③ Given info of marks of 60 students

	x Maths	y English
Mean	80	50
S.D.	13	10

$r = 0.4$

i) Estimate Mark of student in Maths who scored 60 marks in English & Marks of student in Eng who scored 70 in Maths

Given $\bar{x} = 80$

$\bar{y} = 50$

$\sigma_x = 15$

$\sigma_y = 10$

$r = 0.4$

$$b_{xy} = r\,\frac{\sigma_x}{\sigma_y} = 0.4 \times \frac{15}{10} = 0.6$$

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

$$\therefore\ x - 80 = 0.6\,(60 - 50)$$

$$x = 86$$

$$b_{yx} = r\,\frac{\sigma_y}{\sigma_x} = 0.4 \times \frac{10}{15} = ☺$$

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

$$y = ☺\,(70 - 80) + 50$$

$$= 47.36$$

(4) Given $6y = 5x + 90$ (1)

$15x = 8y + 130$ (2)

$(\sigma x)^2 = 16$

(a) find $\bar{x}$ & $\bar{y}$

(b) r

(c) $(\sigma_y)^2$

$\bar{x}$ & $\bar{y}$ are intersection of regression lines

Solve (1) & (2), $\bar{x} = 30$

$\bar{y} = 40$

Let line (1) be bxy

$x = \frac{6}{5} y - 90 \quad b_{xy} = 6/5$

$y = \frac{15}{8} x - 130 \quad b_{yx} = 15/8$

$r = \sqrt{b_{xy} b_{yx}} = 1 \ldots$

$r > 1$ Not Possible

$\therefore$ line ① is b_{yx}

$y = \frac{5}{6}x + \frac{90}{6}$ $\qquad b_{xy} = \frac{5}{6}$

$x = \frac{8}{15}y + \frac{130}{5}$ $\qquad b_{yx} = \frac{8}{15}$

$r = \sqrt{b_{xy} b_{yx}} = 0.6$

$b_{yx} = \frac{5}{6}$ $\qquad \frac{5}{6} = r\frac{\sigma_y}{\sigma_x}$

$= 0.67 \times \frac{\sigma_y}{3}$

$\therefore \sigma_y = 5$ $\qquad \therefore \sigma_y^2 = 25$

④ $\sigma = \sigma_y = \sigma_x$ and angle between lines is $\tan^{-1}(3)$ find the coeficiant of correlation

$\theta = \tan^{-1}(3)$ - given

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

$$3 = \frac{(1-r^2)}{r} \frac{\sigma^2}{2\sigma^2}$$

$$r^2 + 6r - 1 = 0$$

$$\therefore r = -3 \pm \sqrt{10}$$

$$= -6\ldots \quad \text{or} \quad 0.16$$

$1 < r < -1 \quad \therefore r = 0.16$

⑤ Means of x & y are $\bar{x} = 5$, $\bar{y} = 10$. Line of reg. of y on x is parallel to line $20y = 9x + 40$. Find y for $x = 30$

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

is parallel to

$$y = \frac{9}{20}x + c.$$

$$\therefore b_{yx} = \frac{9}{20}$$

$$\therefore (y - 10) = \frac{9}{20}(30 - 5)$$

$$\therefore y = \frac{9}{20} \times 25 + 10$$

$$\therefore y = 21.25$$

$$b_{xy} = r\frac{\sigma_x}{\sigma_y} = \frac{x-\bar{x}}{y-\bar{y}}$$

$$\sigma_x = \sqrt{\Sigma x^2 + N\bar{x}^2}$$

$$b_{xy} = \frac{\Sigma xy + N\bar{x}\bar{y}}{\cancel{\sigma_x}\,\sigma_y}\,\frac{\cancel{\sigma_x}}{\sigma_y}$$

$$= \frac{\Sigma xy + N\bar{x}\bar{y}}{\Sigma y + N\bar{y}^2}$$