Regression ①

line of regression of y on x is

$$y = a + bx$$

line of regression of x on y is

$$x = a + by$$

eqn of line of regression of y on x is

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

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

$b_{yx} = r\dfrac{\sigma_y}{\sigma_x}$ = regression coefficient of y on x

eqn of line of regression of x on y is

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

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

$b_{xy} = r\dfrac{\sigma_x}{\sigma_y}$ = regression coefficient of x on y

① line of regression of y on x

$$y = a + bx$$

$$\Sigma y = a\cancel{\Sigma x}\, + b\Sigma x$$

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

Q. line of regression of x on y

$x = a + by$

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

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

are normal eqns

Ex: 1 Find the equations of the lines of regression

x :	5	6	7	8	9
y :	2	4	5	6	8

also find r

Soln.

Sr No.	x	x^2	y	y^2	xy
1	5	25	2	4	10
2	6	36	4	16	24
3	7	49	5	25	35
4	8	64	6	36	48
5	9	81	8	64	72
N=5	$\Sigma x = 35$	$\Sigma x^2 = 255$	$\Sigma y = 25$	$\Sigma y^2 = 145$	Σxy 189

line of regression of y on x is

$y = a + bx$

$\Sigma y = aN + b\Sigma x$ $\therefore$ $25 = 5a + 35b$ — (1)

$\Sigma xy = a\Sigma x + b\Sigma x^2$ $189 = 35a + 255b$ — (2)

$a = -4.8, \ b = 1.4$

$y = -4.8 + 1.4x$

line of regression of x on y (3)

$x = a + by$

$\Sigma x = aN + b\Sigma y$ $\therefore$ $35 = 5a + 25b$ — (3)

$\Sigma xy = a\Sigma y + b\Sigma y^2$ $189 = 25a + 145b$ — (4)

$a = 2.2$, $b = 0.56$

$x = 2.2 + 0.56y$

$b_{yx} = 1.4$ $b_{xy} = 0.56$

$r = \sqrt{b_{yx} b_{xy}} = \sqrt{(1.4)(0.56)} = 0.88$

(i) $b_{yx} = \frac{\Sigma xy}{\Sigma x^2}$ $b_{xy} = \frac{\Sigma xy}{\Sigma y^2}$

$x = X - \bar{X}$, $y = Y - \bar{Y}$

(ii) $dx = x - A$, $dy = Y - B$

$b_{yx} = \frac{\Sigma dxdy - \frac{\Sigma dx \Sigma dy}{N}}{\Sigma d^2x - \frac{(\Sigma dx)^2}{N}}$

$b_{xy} = \frac{\Sigma dxdy - \frac{\Sigma dx \cdot \Sigma dy}{N}}{\Sigma d^2y - \frac{(\Sigma dy)^2}{N}}$

(iii) x, y actual value

$b_{yx} = \frac{\Sigma XY - \frac{\Sigma X \cdot \Sigma Y}{N}}{\Sigma X^2 - \frac{(\Sigma X)^2}{N}}$

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

$b_{xy} = \frac{\Sigma XY - \frac{\Sigma X \Sigma Y}{N}}{\Sigma Y^2 - \frac{(\Sigma Y)^2}{N}}$

property

(i) $b_{yx}\, b_{xy} = r^2$

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

both b_{yx}, b_{xy} are positive or both are negative

(ii) $b_{yx} \leq \frac{1}{b_{xy}}$

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

(iv) ~~b_{xy}~~ If $r = \pm 1$

$b_{xy} = \frac{1}{b_{yx}}$

EX: State true or false with reasoning

$2x + y = 3$ and $x = 2y + 3$ cannot be the lines of regression

Soln The line of regression of $2x + y = 3$ is

$y = -2x + 3$, $b_{yx} = -2$

The line of regression of $x = 2y + 3$ is

$b_{xy} = 2$

$b_{yx} = -2$, $b_{xy} = 2$ are opposite signs

& both are greater than 1

again line of regression of x on y of $2x + y = 3$

is $x = -\frac{y}{2} + \frac{3}{2}$ $b_{xy} = -\frac{1}{2}$

The line of regression of y on x or

$x \pm y \pm 3$ is $y = +\frac{1}{2}x \mp \frac{3}{2} = \frac{1}{2}x - \frac{3}{2}$

$b_{yx} = \frac{1}{2}$

$\therefore b_{xy} = \frac{1}{2},\ b_{yx} = \frac{1}{2}$

Ex (2) State true or false with justification

If two lines of regression are $x+3y-5=0$ and $4x+3y-8=0$ then correlation coefficients $+0.5$

Soln The line $x+3y-5=0$ be the line of regression of x on y, is $x = -3y+5$

$b_{xy} = -3$ The line $4x+3y-8=0$ be the line of regression of y on x is

$3y = -4x+8 \therefore y = -\frac{4}{3}x + \frac{8}{3}$ $b_{yx} = -\frac{4}{3}$

$r = \pm\sqrt{b_{yx}\, b_{xy}} = \sqrt{(-3)(-\frac{4}{3})} = \sqrt{4} = 2$

r cannot be greater than 1

our supposition is wrong

$\therefore$ The line $x+3y-5=0$ be the line of regression of y on x is $y = -\frac{1}{3} + \frac{5}{3}$

$b_{yx} = -\frac{1}{3}$

The line $4x+3y-8=0$ of regression of x on y

is $4x = -\frac{3}{4}y + 2$ $b_{xy} = -\frac{3}{4}$

$b_{yx} = -\frac{1}{3}$ $b_{xy} = -\frac{3}{4}$

$r = \sqrt{b_{yx} b_{xy}} = \sqrt{-\frac{1}{3}\left(-\frac{3}{4}\right)} = \frac{1}{2}$

$= 0.5$

statements are

angle betn the lines of regression

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

Ex: Find the coefficient of regression and hence the each lines of regression

x:	78	36	98	25	75	82	90	62	65	39
Y:	84	51	91		60	68	62	86	58	53 47

Find Y when x = 57, Find x when y = 90

Sol^n

Sr No	X	$x = X - \bar{X}$ $= X - 65$	x^2	Y	$y = Y - \bar{Y}$ $= Y - 66$	y^2	xy
1	78	13	169	84	18	324	234
2	36	-29	841	51	-15	225	435
3	98	33	1089	91	25	625	825
4	25	-40	1600	60	-6	36	240
5	75	10	100	68	2	4	20
6	82	17	289	62	-4	16	-68
7	90	25	625	86	20	400	500
8	62	-3	9	58	-8	64	24
7	65	0	0	53	-13	169	0
10	39	-26	676	47	-19	361	494
$N = 10$	$\Sigma X = 650$		$\Sigma x^2 = 5398$	$\Sigma Y = 660$	Σ	$\Sigma y^2 = 2224$	$\Sigma xy = 2704$

$\bar{X} = \frac{\Sigma X}{N} = \frac{650}{10} = 65$

$\bar{Y} = \frac{\Sigma Y}{N} = \frac{660}{10} = 66$

$b_{yx} = \frac{\Sigma xy}{\Sigma x^2} = \frac{2704}{5398} = 0.5009$

$b_{xy} = \frac{\Sigma xy}{\Sigma y^2} = \frac{2204}{\cancel{660}2224} = 1.215$

The line of regression of y on x is

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

$y - 66 = 0.5009(x - 65)$

$y - 66 = 0.5(x - 65)$

The eqn of line of regression of x on y

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

$x-65 = 1.215(y-66)$

$x = 65 + 1.215y - (1.215)(66)$

$y = 90$

$x = 94$

$x = 50$

$y = 66 + 0.5x - (0.5)(65)$

$\underline{y = 58}$

Ex: A chemical engineer is ~~Thesting~~ investigating the effect of process of operating temperature x on product yield y

x : 100 110 120 130 140 150 160 170 180 190

y : 45 51 54 61 66 70 74 78, 85, 89.

The eqn of line of product yield on basic temperature

find r =