Another Equation for Correlation Coefficient

$$r = \frac{n\sum XY - (\sum X)(\sum Y)}{\sqrt{n\sum X^2 - (\sum X)^2} \times \sqrt{n\sum Y^2 - (\sum Y)^2}}$$

where n is the number of observations.

Example: Find Correlation Coefficient.

X : 1 2 3 4 5 6 7

Y : 6 8 11 9 12 10 14

Sol:

X	Y	X^2	Y^2	XY
1	6	1	36	6
2	8	4	64	16
3	11	9	121	33
4	9	16	81	36
5	12	25	144	60
6	10	36	100	60
7	14	49	196	98
Total: 28	70	140	742	309

Correlation Coeffient $r = \frac{n\sum XY - (\sum X)(\sum Y)}{\sqrt{n\sum X^2 - (\sum X)^2} \times \sqrt{n\sum Y^2 - (\sum Y)^2}}$

$$= \frac{7 \times 309 - 28 \times 70}{\sqrt{7 \times 140 - (28)^2} \times \sqrt{7 \times 742 - (70)^2}}$$

$$= \frac{203}{14 \times 17.15} = \underline{\underline{0.845}}.$$

Regression

Regression equation of y on x is given as:

$$y - \bar{y} = \frac{Cov(x,y)}{\sigma_x^2}(x - \bar{x})$$

or,

$$y - \bar{y} = b_{yx}(x - \bar{x})$$, where $\bar{x}$ is the mean

where,

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

Regression equation of x on y is given as:

$$x - \bar{x} = \frac{Cov(x,y)}{\sigma_y^2}(y - \bar{y})$$, where $\bar{y}$ is the mean

or,

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

where,

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

* Form the two regression Equations:

X:	1	2	3	4	5	6	7
Y:	2	4	7	6	5	6	5

Ans.

X	Y	X^2	Y^2	XY
1	2	1	4	2
2	4	4	16	8
3	7	9	49	21
4	6	16	36	24
5	5	25	25	25
6	6	36	36	36
7	5	49	25	35
$\Sigma X = 28$	$\Sigma Y = 35$	$\Sigma X^2 = 140$	$\Sigma Y^2 = 191$	$\Sigma XY = 151$

Now,

$$\bar{x} = \frac{1+2+3+4+5+6+7}{7} = \frac{28}{7} = \underline{\underline{4}}$$

$$\bar{y} = \frac{2+4+7+6+5+6+5}{7} = \frac{35}{7} = \underline{\underline{5}}$$

Regression Equation of y on x is given by

$$y - \bar{y} = b_{yx}(x - \bar{x}) \quad \text{--- (1)}$$

where, $b_{yx} = \dfrac{n \Sigma XY - \Sigma X \Sigma Y}{n \Sigma X^2 - (\Sigma X)^2}$

Now,

$$b_{yx} = \frac{7\times 151 - 980}{7\times 140 - (28)^2}$$

$$= \frac{77}{196} = \underline{\underline{0.392}}$$

Substituting the value of b_{yx} in (1)

we have,

$$y - \bar{y} = 0.392(x - \bar{x})$$

Substituting the values for $\bar{y}$ & $\bar{x}$,

$$y - 5 = 0.392(x - 4)$$

$$y - 5 = 0.392x - 1.568$$

$$y = 0.392x + 3.432$$

$\therefore$ The regression Equation of y on x is :

$$\underline{\underline{y = 0.392x + 3.432}}$$

Now, Regression Equation of x on y is given by

$$x - \bar{x} = b_{xy}(y - \bar{y}) \quad \text{—— (2)}$$

where,

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

Now,

$$b_{xy} = \frac{7\times151 - 980}{7\times191 - (35)^2}$$

$$= \frac{77}{112} = \underline{\underline{0.687}}$$

Substituting in (1) we have,

$x - \bar{x} = 0.687\ (y - \bar{y})$

Substituting the values of $\bar{x}$ and $\bar{y}$,

$x - 4 = 0.687\ (y - 5)$

$x - 4 = 0.687y - 3.435$

$x = 0.687y + 0.565$

∴ The regression Equation of x on Y is

$$\underline{\underline{x = 0.687y + 0.565}}$$

* Find the two Regression equations for the following data:

x:	2	3	4	5	6
y:	3	5	4	8	9.

* a) Find the correlation Coefficient b/w x & y

b) Form the two regression equations.

x:	12	20	15	22	18	24	20	12	15	22
y:	30	35	28	36	29	39	30	25	30	38