KARL PEARSON'S COEFFICIENT OF C

A numerical measure of the degree of Correlation, denoted by r or r_{xy} between two Variables x & y was introduced by the Karl Pearson

It is called the Coefficient of Correlation and is defined by

$$r = \frac{\sum (x_i - M_x)(y_i - M_y)}{\sqrt{\left[\left\{\sum (x_i - M_x)^2\right\} \cdot \left\{\sum (y_i - M_y)^2\right\}\right]}} \rightarrow ①$$

$$r = \frac{\sum (x_i - M_x)(y_i - M_y)}{n \sigma_x \sigma_y} \rightarrow ②$$

where M_x is Mean of x, M_y is the mean of y, σ_x & σ_y are Standard deviation of x & y

here Summation extends over i from 1 to n

The formula ② can also be stated as

$$r = \frac{Cov(x, y)}{\sqrt{(Var x)(Var y)}}$$

$$r = \frac{\mathrm{Cov}(x,y)}{\left[(\mathrm{Var}\,x)(\mathrm{Var}\,y)\right]^{1/2}}$$

$$r = \frac{\frac{1}{n}\sum_{i=1}^{n}(x_i - M_x)(y_i - M_y)}{\left[\left\{\frac{1}{n}\sum x_i^2 - M_x^2\right\}\left\{\frac{1}{n}\sum y_i^2 - M_y^2\right\}\right]^{1/2}}$$

Since

$$\sum (x_i - M_x)(y_i - M_y) = \sum x_i y_i - M_x \sum y_i - M_y \sum x_i + n M_x M_y$$

$$= \sum x_i y_i - \cancel{M_x . n M_y} - \cancel{M_y . n M_x} + \cancel{n M_x} M_y$$

$\therefore$ Mean of y

Mean $= \frac{y_1 + y_2 + y_3 \dots y_n}{n}$

$M_y = \frac{\sum y_i}{n}$

So

$$= \sum x_i y_i + n M_x M_y$$

$$\mathrm{Cov}(x,y) = \frac{1}{n}\sum_{i=1}^{n}(x_i - M_x)(y_i - M_y)$$

$$= \frac{1}{N}\left[\sum x_i y_i + n M_x M_y\right]$$

$$= \sum x_i y_i + n M_x M_y$$

i.e. $\sigma_x = \sqrt{\frac{1}{N}\sum x_i^2 - M_x^2}$ and $\sigma_y = \sqrt{\frac{1}{N}\sum y_i^2 - M_y^2}$

The coefficient of correlation r has no units and is merely a number.

→ Find the Coefficient of Correlation
[betw]een x & y from the following

x : 1 2 3 4 5 6 7 8 9

y : 9 8 10 12 11 13 14 16 15

Solution →

$M_x = \frac{\Sigma x_i}{n} = \frac{45}{9} = 5$ $\quad M_y = \frac{\Sigma y_i}{n} = \frac{108}{9} = 12$

x	$x - M_x$	$(x-M_x)^2$	y	$(y-M_y)$	$(y-M_y)^2$	$(x-M_x)(y-M_y)$
1	-4	16	9	-3	9	12
2	-3	9	8	-4	16	12
3	-2	4	10	-2	4	4
4	-1	1	12	0	0	0
5	0	0	11	-1	1	0
6	1	1	13	1	1	1
7	2	4	14	2	4	4
8	3	9	16	4	16	12
9	4	16	15	3	9	12
$\Sigma x_i = 45$		$\Sigma(x_i-M_x)^2 = 60$	$\Sigma y_i = 108$		$\Sigma(y_i-M_y)^2 = 60$	$\Sigma(x_i-M_x)(y_i-M_y) = 57$

Hence the Coefficient of Correlation

$$r = \frac{\Sigma(x_i-M_x)(y_i-M_y)}{\sqrt{\Sigma(x_i-M_x)^2 \Sigma(y_i-M_y)^2}} = \frac{57}{\sqrt{60 \times 60}} = \frac{57}{60}$$

$$= 0.95$$

Ans.