Co-rrelation:

In a bivariate distribution, if the change in one-variable affects a change in the other variable, the variables are said to be co-rrelated.

Eg:→ Pressure and volume. These two parameters are inversely proportional ie the change in one parameter affects the other. Therefore Pressure and volume are co-rrelated.

If the two variables deviate in the same direction ie, if the increase (or decrease) in one results in a corresponding increase (or decrease) in the other, then the corelation is said to be positive.

Eg:→ (i) Height and weight. ~~Tall people~~

(ii) Income and expenditure etc.

If the two variables deviate in opposite direction ie if the increase (or decrease) in one results in a corresponding decrease (or increase) in the other, correlation is said to be inverse ~~ore~~ or negative.

Eg:→ Pressure and volume ~~of~~ ~~go~~

Karl Pearson's co-efficient of ~~co~~ correlation

Correlation co-efficient between two variables x and y, usually denoted by $r(x,y)$ or r_{xy} and is defined as,

$$r_{xy} = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum (x_i - \bar{x})^2 \sum (y_i - \bar{y})^2}}$$

$$= \frac{\frac{1}{n} \sum (x_i - \bar{x})(y_i - \bar{y})}{\sigma_x \sigma_y}.$$

where $\sigma_x = \sqrt{\frac{1}{n} \sum (x_i - \bar{x})^2}$, $\sigma_y = \sqrt{\frac{1}{n} \sum (y_i - \bar{y})^2}$

(Remember this formula)

If the values of x and y are very large, we can use the change of origin and scale like,

$$u = \frac{x-a}{h}, \quad v = \frac{y-b}{k},$$

a, b, h, k are suitable constant,

then $r_{xy} = r_{uv}$.

Different form of correlation co-efficient :

We know that $r_{xy} = \dfrac{\frac{1}{n}\sum (x_i - \bar{x})(y_i - \bar{y})}{\sigma_x \sigma_y}.$

Now, $\frac{1}{n}\sum (x_i - \bar{x})(y_i - \bar{y}) = \frac{1}{n}\sum (x_i y_i - x_i \bar{y} - y_i \bar{x} + \bar{x}\bar{y})$

$$= \frac{1}{n}\sum x_i y_i - \bar{y}\cdot\frac{1}{n}\sum x_i - \bar{x}\,\frac{1}{n}\sum y_i + \frac{\bar{x}\bar{y}}{n}\sum 1$$

$\because$ $\bar{x}$, $\bar{y}$ are means of x and y and as numbers, so we can put $\bar{x}$, $\bar{y}$ outside the $\sum$.

$$= \frac{1}{n}\sum x_i y_i - \bar{y}\bar{x} - \bar{x}\bar{y} + \bar{x}\bar{y}$$

$$= \frac{1}{n}\sum x_i y_i - \bar{x}\bar{y} \qquad \left[\because \frac{1}{n}\sum x_i = \bar{x},\ \frac{1}{n}\sum y_i = \bar{y}\right.$$

Now, $\sigma_x^2 = \frac{1}{n}\sum'(x_i-\bar{x})^2$

$$= \frac{1}{n}\sum'(x_i^2 - 2x_i\bar{x} + \bar{x}^2)$$

$$= \frac{1}{n}\sum' x_i^2 - 2\bar{x}\frac{1}{n}\sum' x_i + \bar{x}^2 \frac{1}{n}\sum' 1.$$

$$= \frac{1}{n}\sum' x_i^2 - 2\bar{x}\cdot\bar{x} + \bar{x}^2 \cdot \frac{1}{n}\cdot n.$$

$$= \frac{1}{n}\sum' x_i^2 - 2\bar{x}^2 + \bar{x}^2$$

$$= \frac{1}{n}\sum' x_i^2 - \bar{x}^2$$

lly, $\sigma_y^2 = \frac{1}{n}\sum' y_i^2 - \bar{y}^2$

$$\therefore r_{xy} = \frac{\frac{1}{n}\sum'(x_i-\bar{x})(y_i-\bar{y})}{\sigma_x\sigma_y}$$

$$= \frac{\frac{1}{n}\sum' x_i y_i - \bar{x}\,\bar{y}}{\sqrt{\left(\frac{1}{n}\sum' x_i^2 - \bar{x}^2\right)\left(\frac{1}{n}\sum' y_i^2 - \bar{y}^2\right)}}$$

$$= \frac{n\sum' x_i y_i - \sum' x_i \cdot \sum' y_i}{\sqrt{n\sum' x_i^2 - (\sum' x_i)^2}\sqrt{n\sum' y_i^2 - (\sum' y_i)^2}}$$

In solving problem, we can use transformation $u = \frac{x-a}{h}$, $v = \frac{y-b}{k}$.

Then the formula became,

$$\boxed{r_{xy} = r_{uv} = \frac{n\sum' uv - \sum' u \sum v}{\sqrt{n\sum' u^2 - (\sum' u)^2}\sqrt{n\sum' v^2 - (\sum' v)^2}}}$$

Remember this formular for problem,

here, $u = \frac{x-a}{h}$, $v = \frac{y-b}{k}$.