CORRELATION

→ Correlation:- Statistical tecniques used to measure the degree & direction of the relation b/w two variables.

- positive correlation : $x\uparrow$ $y\uparrow$
- negative correlation: $x\downarrow$ $y\downarrow$ (Inverse relation)

→ Linear correlation.

- measure of linear correlation
  1. Graphical method (scatterness diagram method)
  2. non-graphical
     a. Non karl pearson's co-efficient correlation
     b. spearman's Rank correlation
     c. concurrent deviations methods.

1. Graphical method.

- If the points plotted are concentrated from lower left corner to upper right corner. it is positive correlation
  - $0 < r < 1$ → positive correlation
  - $r = 1$ → perfect positive correlation
- If the points plotted are concentrated from upper left to lower right corner, it is negative correlation
  - $-1 < r < 0$ → negative correlation
  - -1 → perfect negative -"-
- If the points are scattered without any pattern the variables are un-correlated
  - $r = 0$ → (uncorrelated, no correlation)

2. non-graphical

a. karl pearson's method.

Co-varionce = cov (x,y)

$$\text{cov}(x,y) = \frac{\Sigma(x_i-\bar{x}).(y-\bar{y})}{N} = \boxed{\frac{\text{cov}(x,y)}{\sigma_x \sigma_y}}$$

- cov(x,y) can be any real number.
- change in origin – NO
- change in scale – Yes

$$\sigma_x = \sqrt{\frac{\Sigma(x-\bar{x})^2}{N}} \qquad \sigma_y = \sqrt{\frac{\Sigma(y-\bar{y})^2}{N}}$$

$r = \dfrac{\text{cov}(x,y)}{\sigma_x \sigma_y}$ # $\dfrac{\Sigma(x-\bar{x}).(y-\bar{y})}{N\sigma_x \sigma_y}$

$\dfrac{\Sigma(x-\bar{x})(y-\bar{y})}{\sqrt{\Sigma(x-\bar{x})^2 . \Sigma(y-\bar{y})^2}}$

#

$$r = \frac{\Sigma xy - \frac{\Sigma x \times \Sigma y}{N}}{\sqrt{\Sigma x^2 - \frac{(\Sigma x)^2}{N}} \times \sqrt{\Sigma y^2 - \frac{(\Sigma y)^2}{N}}}$$

$U = x_i - A$ & $V = y - A$

$$r = \frac{\Sigma UV - \frac{\Sigma U \times \Sigma V}{N}}{\sqrt{\Sigma U^2 - \frac{(\Sigma V)^2}{N}} \sqrt{\Sigma V^2 - \frac{(\Sigma V)^2}{N}}}$$

karl pearson co-efficient of correlation.

- change in origin → No change
- change of scale → No change.

→ [Spearman's Rank correlation/spearman method]

→ used for relation B/w Qualitative character

→ level of agreement b/w two judges.

$$r = 1 - \frac{6\Sigma D^2}{N^3 - N}$$

→ If some elements repeat.

$$r = 1 - \frac{6[\Sigma D^2 + \frac{1}{12}[m_1^3 - m_1] + \frac{1}{12}[m_2^3 - m_2] + \frac{1}{2}[m_3^3 - m_3]}{N^3 - N}$$

$\left.\begin{matrix} M_1 \\ M_2 \\ M_3 \end{matrix}\right\}$ = frequency of repeating numbers.

→ [Concurrent deviation method]

$\left.\begin{matrix} x\uparrow\ y\uparrow \rightarrow \text{concurrent} \\ x\downarrow\ y\downarrow \rightarrow \text{concurrent} \end{matrix}\right\}$ $x\uparrow$ $y\downarrow$ → non concurrent

$$r = \pm\sqrt{\pm\left(\frac{2c-n}{n}\right)}$$

c = Total concurrent deviation

n = Total no of pairs

→ Co-efficient of determination

- The coefficient of determination is used to expalin the relationship between an independent and dependent variable.
- measures the amount of change in dependent variable due to change in independent variable.

formula

$$\text{COD} = r^2 = \frac{\text{Explained variance}}{\text{Total variance}}$$

→ Co-efficient of non determination

$= 1 - r^2$

CORRELATION

→ Correlation :- Statistical tecniques used to measure the degree & direction of the relation b/w two variables.

- positive correlation : $x\uparrow$ $y\uparrow$
- negative correlation: $x\downarrow$ $y\downarrow$ (Inverse relation)

→ Linear correlation.

- measure of linear correlation
  1. Graphical method (scatterness diagram method)
  2. non-graphical
     a. non Karl pearson's co-efficient correlation
     b. spearman's Rank correlation
     c. concurrent deviations methods.

1. Graphical method.

- If the points plotted are concentrated from lower left corner to upper right corner. it is positive correlation
  - $0 < r < 1$ → positive correlation
  - $r = 1$ → perfect positive correlation
- If the points plotted are concentrated from upper left to lower right corner, it is negative correlation
  - $-1 < r < 0$ → negative correlation
  - -1 → perfect negative —"—
- If the points are scattered without any pattern the variables are un-correlated
  - $r = 0$ → (uncorrelated, no correlation)

2. non-graphical

a. Karl pearson's method.

co-variance = cov (x,y)

$\boxed{cov(x,y)}$

Karl pearson co-efficient of correlation.

- Change in origin → No change
- Change of scale → No change

→ [Spearman's Rank correlation/spearman method]

→ used for relation B/w qualitative character
→ level of agreement b/w two judges.

$$r = 1 - \frac{6\Sigma D^2}{N^3 - N}$$

→ If some elements repeat.

$$r = 1 - \frac{6\left[\Sigma D^2 + \frac{1}{12}[m_1^3 - m_1] + \frac{1}{12}[m_2^3 - m_2] + \frac{1}{2}[m_3^3 - m_3]\right]}{N^3 - N}$$

$\left.\begin{matrix} M_1 \\ M_2 \\ M_3 \end{matrix}\right\}$ = frequency of repeating numbers.

→ [Concurrent deviation method]

$\left.\begin{matrix} x\uparrow\ y\uparrow \to \text{concurrent} \\ x\downarrow\ y\downarrow \to \text{concurrent} \end{matrix}\right\}$ $x\uparrow$ $y\downarrow$ → non concurrent

$$r = \pm\sqrt{\pm\left(\frac{2c - n}{n}\right)}$$

c = Total concurrent deviation
n = Total no of pairs

→ Co-efficient of determination

- The coefficient of determination is used to explain the relationship between an independent and dependent variable.
- measures the amount of change in dependent variable due to change in independent

> $r = 0 \rightarrow$ (uncorrelated, no correlation)

2. non-graphical

a. Karl Pearson's method.

co-variance = cov (x,y)

$$cov(x,y) = \frac{\Sigma(x_i - \bar{x}).(y-\bar{y})}{N} = \boxed{\frac{cov(x,y)}{\sigma x \, \sigma y}}$$

> cov(x,y) can be any real number.

> change in origin - NO

change in scale - Yes

$$\sigma x = \sqrt{\frac{\Sigma(x-\bar{x})^2}{N}} \qquad \sigma y = \sqrt{\frac{\Sigma(y-\bar{y})^2}{N}}$$

$$\# \; r = \frac{cov(x,y)}{\sigma x \, \sigma y} \quad \# \quad \frac{\Sigma(x-\bar{x}).(y-\bar{y})}{N \sigma x \, \sigma y}$$

$$= \frac{\Sigma(x-\bar{x})(y-\bar{y})}{\sqrt{\Sigma(x-\bar{x})^2 . \Sigma(y-\bar{y})^2}}$$

$$r = \frac{\Sigma xy - \frac{\Sigma x \times \Sigma y}{N}}{\sqrt{\Sigma x^2 - \frac{(\Sigma x)^2}{N}} \times \sqrt{\Sigma y^2 - \frac{(\Sigma y)^2}{N}}}$$

$U = x_i - A$ & $V = y - A$

$$= \frac{\Sigma UV - \frac{\Sigma U \times \Sigma V}{N}}{\sqrt{\Sigma U^2 - \frac{(\Sigma V)^2}{N}} \sqrt{\Sigma V^2 - \frac{(\Sigma V)^2}{N}}}$$

- The coefficient of determination is used to explain the relationship between an independent and dependent variable.
- measures the amount of change in dependent variable due to change in independent variable.

formula

$$COD = r^2 = \frac{\text{Explained variance}}{\text{Total variance}}.$$

→ co-efficient of non determination

$= 1 - r^2$