Ch= 17 correlation and Regression

* Correlation

→ Karl Pearson's Method :

correlation coefficient = r

(i) $r_{xy} = \dfrac{\text{covariance}}{(\text{sd of } x)(\text{sd of } y)} = \dfrac{\text{cov}(x,y)}{s_x \cdot s_y}$

where, $\text{cov}(x,y) = \dfrac{\Sigma(x-\bar{x})(y-\bar{y})}{n}$

$$\boxed{= \dfrac{\Sigma xy - n\bar{x}\bar{y}}{n}}$$

$\therefore s_x = \sqrt{\dfrac{\Sigma(x-\bar{x})^2}{n}}$ and $s_y = \sqrt{\dfrac{\Sigma(y-\bar{y})^2}{n}}$

(ii) $r_{xy} = \dfrac{\Sigma(x-\bar{x})(y-\bar{y})}{\sqrt{\Sigma(x-\bar{x})^2} \cdot \sqrt{\Sigma(y-\bar{y})^2}}$

(iii) $r_{xy} = \dfrac{n\Sigma xy - (\Sigma x)(\Sigma y)}{\sqrt{n\Sigma x^2 - (\Sigma x)^2} \cdot \sqrt{n\Sigma y^2 - (\Sigma y)^2}}$

mean of x = $\bar{x} = \dfrac{\Sigma x_i}{n}$

mean of y = $\bar{y} = \dfrac{\Sigma y_i}{n}$

(iv) $r = \dfrac{n\Sigma uv - (\Sigma u)(\Sigma v)}{\sqrt{n\Sigma u^2 - (\Sigma u)^2} \cdot \sqrt{n\Sigma v^2 - (\Sigma v)^2}}$

where, $u = X - A$ or $\dfrac{x - A}{cx}$,

$v = y - B$ or $\dfrac{y - B}{cy}$

(v) $r_{\bar{x}\bar{y}} = \dfrac{\Sigma (x-\bar{x})(y-\bar{y})}{n \cdot Sx \cdot Sy}$

(vi) $r_{\bar{x}\bar{y}} = \dfrac{\Sigma xy - n\bar{x}\bar{y}}{n \cdot Sx \cdot Sy}$

(Specially for short sums)

→ Spearman's Rank Correlation Method.

(i) $r = 1 - \dfrac{6\Sigma d^2}{n(n^2-1)}$ (when the obs. are not repeated)

(ii) $r = 1 - \dfrac{6[\Sigma d^2 + CF]}{n(n^2-1)}$ (when some of the obs. repeated)

where, d = rank of x - Rank of y

$\boxed{d = R_x - R_y.}$

CF = Correlation factor $= \Sigma \left(\dfrac{m^3 - m}{12} \right)$

m = No. of times obs. is repeated

Date :

* Regression

→ Equation of regression Line

$$\hat{y} = a + bx$$

where, $b = b_{yx}$ = Regression Coefficient

(i) $b_{yx} = \dfrac{\Sigma (x-\bar{x})(y-\bar{y})}{\Sigma (x-\bar{x})^2}$

(ii) $b_{yx} = \dfrac{n\Sigma xy - (\Sigma x)(\Sigma y)}{n\Sigma x^2 - (\Sigma x)^2}$

(iii) $b = b_{yx} = \dfrac{n\Sigma uv - (\Sigma u)(\Sigma v)}{n\Sigma u^2 - (\Sigma u)^2}$.

Here, $u = X - A$ and $v = y - B$

(iv) $b = \dfrac{n\Sigma uv - (\Sigma u)(\Sigma v)}{n\Sigma u^2 - (\Sigma u)^2} \times \dfrac{Cy}{Cx}$

Here, $u = \dfrac{X-A}{Cx}$ & $v = \dfrac{Y-B}{Cy}$

(v) $b_{yx} = r \cdot \dfrac{Sy}{Sx}$

(vi) $b_{yx} = \dfrac{Cov(x,y)}{Sx^2} = \dfrac{r \cdot Sx \cdot Sy}{Sx^2}$

(vii) $a = \bar{y} - b\bar{x}$

(viii) Coefficient of Determination (R^2)

$$R^2 = [r(y, \hat{y})]^2 = [r(x,y)]^2 = r^2$$

=) in <u>Illustration</u>

Coefficient of Determination

$$R^2 = \left[\frac{\text{cov}(x,y)}{S_x \cdot S_y}\right]^2.$$

error find karwa mate

$$e = y - \hat{y}$$

→ Coefficient of concurrent Deviations

refer = Example 17.15 (Pg - 17.24)

formula = $r_c = \pm \sqrt{\pm \frac{(2c-m)}{m}}$

where, m = no. of pairs of deviation

c = No. of positive signs in product of Deviation column

= No. of concurrent deviation.

(ix) Coefficient of non-determination

$= (1 - r^2)$