PSOT (sem-4)

classmate
Date
Page

Mod 2

→ Correlation

Mean $\bar{x} = \frac{\Sigma x_i}{n}$

Variance $\sigma_x^2 = \frac{\Sigma (x-\bar{x})^2}{n}$

S.D. $\sigma_x = \sqrt{\frac{\Sigma (x-\bar{x})^2}{n}}$

covariance $\text{cov}(x,y) = \frac{\Sigma xy}{n} - \left(\frac{\Sigma x}{n}\right)\left(\frac{\Sigma y}{n}\right)$

* KP coefficient of correlation (r)

① $r = \frac{\text{cov}(x,y)}{\sigma_x \ \sigma_y}$

② If $\bar{x}$ & $\bar{y}$ integer

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

④ If $\bar{x}$ & $\bar{y}$ decimal/fraction

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

⑤ Similarly, $r = \frac{\Sigma xy - n\bar{x}\bar{y}}{\sqrt{\Sigma x^2 - n\bar{x}^2}\sqrt{\Sigma y^2 - n\bar{y}^2}}$

Change of scale & origin allowed
$-1 \leq r \leq 1$

classmate
Date
Page

* Spearman's coefficient of Rank correlation (R)

$R_1, R_2 \rightarrow$ Ranks of x & y

$d = R_1 - R_2$

$$R = 1 - \frac{6 \Sigma d^2}{n^3 - n} \qquad \text{check } \Sigma d = 0$$

Cond^n of equal Ranks

Values in sample are repeated

$$R = 1 - \frac{6\left[\Sigma d^2 + \frac{1}{12}(m_1^3 - m_1) + \frac{1}{12}(m_2^3 - m_2) + \dots\right]}{n^3 - n}$$

m_1 = How many times 1st repeating value is repeated

m_2 = How many times 2nd repeating value is repeated

→ Steps :

To find 'r'

1. Make table x, y, xy, x^2, y^2 Σ / –
2. Identify available values
3. Use formulas out of the 4.

Sometimes if change of origin (– ~~smallest + biggest / 2~~ $\bar{x}$ ~~or~~ $\bar{y}$ (lowest + Highest)/2)
& assign them as u, v & again steps 1, 2 & 3

To find 'R'

1. Make table by giving Ranks to values
   x, R_1 y, R_2 $d = R_1 - R_2$ d^2
   $\Sigma d = 0$
2. Apply formula

Sometimes condition of eq ranks middle of two ranks for both & different formula.

Regression

→ Regression line of y on x

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

Regression coefficient of y on x

$$b_{yx} = r \frac{\sigma_y}{\sigma_x}$$

$$= \frac{\Sigma xy - n\bar{x}\bar{y}}{\Sigma x^2 - n(\bar{x})^2}$$

→ Regression line of x on y

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

Regression coefficient of x on y

$$b_{xy} = r \frac{\sigma_x}{\sigma_y}$$

$$= \frac{\Sigma xy - n\bar{x}\bar{y}}{\Sigma y^2 - n(\bar{y})^2}$$

Also, Rember

$y = mx + c$

$(\bar{x}, \bar{y}) \rightarrow$ pt of intsec of R lines.

Properties

① $r = \sqrt{b_{xy} \times b_{yx}}$

$r^2 = b_{xy} \times b_{yx}$

② $\frac{b_{xy} + b_{yx}}{2} > r$

③ b_{xy}, b_{yx} & r → have the same sign

④ If θ → acute angle b/w lines of regression

$$\tan\theta = \left(\frac{1 - r^2}{r}\right)\left(\frac{\sigma_x \sigma_y}{\sigma_x^2 + \sigma_y^2}\right)$$

PAGE No.
DATE / /

* Properties of Regression coefficients

I : The correlation coefficient is geometric mean b/w regression coefficients

$$r = \sqrt{bxy \times byx}$$

II : If one of the regression coefficients is greater than 1, then the other one has to be less than 1. $\left[r^2 = \underset{\ngtr 1}{bxy \times byx} \right]$

III : $\dfrac{bxy + byx}{2} > r$

Absolute sense

IV : The regression coefficients are independent of change of origin but not of change of scale.

V : The regression coefficients & the correlation coefficient have same sign.
[bxy, byx & r]

VI Angle
If θ is the acute angle b/w lines of regression, then $\tan\theta$ is given by

$$\tan\theta = \left(\frac{1 - r^2}{r}\right)\left(\frac{\sigma x \; \sigma y}{\sigma x^2 + \sigma y^2}\right)$$

Teacher's Signature:..................