Q A computer while calculating correlation coeff. betn 2 variables X & Y from 25 pairs of observations obtained the following results:

$n = 25$, $\Sigma X = 125$, $\Sigma X^2 = 650$, $\Sigma Y = 100$, $\Sigma Y^2 = 460$, $\Sigma XY = 508$

It was however, later discovered at the time of checking that he had copied down two pairs as

X	6	14
Y	7	6

while the correct values were

X	8	6
Y	12	8

Obtain the correct value of correlation coeff.

Soln

Corrected $\Sigma X = 125 - 6 - 8 + 8 + 6 = 125$, Corrected $\Sigma Y = 100 - 14 - 6 + 12 + 8 = 100$.

" $\Sigma X^2 = 650 - 6^2 - 8^2 + 8^2 + 6^2 = 650$, " $\Sigma Y^2 = 460 - 14^2 - 6^2 + 12^2 + 8^2 = 436$.

" $\Sigma XY = 508 - 6\times14 - 8\times6 + 8\times12 + 6\times8 = 520$.

$\bar{X} = \frac{1}{n}\Sigma X = 5$, $\bar{Y} = \frac{1}{n}\Sigma Y = 4$, $Cov(X,Y) = \frac{1}{n}\Sigma XY - \bar{X}\bar{Y} = \frac{520}{25} - 5\times4 = \frac{4}{5}$.

$\sigma_X^2 = \frac{1}{n}\Sigma X^2 - \bar{X}^2 = \frac{650}{25} - 5^2 = 1$, $\sigma_Y^2 = \frac{1}{n}\Sigma Y^2 - \bar{Y}^2 = \frac{1}{25}\times436 - 16 = \frac{36}{25}$.

$\therefore$ Corrected $r(X,Y) = \frac{Cov(X,Y)}{\sigma_X\sigma_Y} = \frac{4/5}{1\times\frac{6}{5}} = \frac{2}{3} = {\cdot}67$.

Q In a partially destroyed laboratory, record of an analysis of correlation data, the following results only are legible:

Var. $X = 9$, Regression eqs: $8X - 10Y + 66 = 0$, $40X - 18Y = 214$.

What are: (i) the mean values X and Y, (ii) the correlation coeff. betn X and Y, and (iii) the standard deviation of Y?

Soln (i) Since both lines of regression pass through the pt. $(\bar{X}, \bar{Y})$, we have $8\bar{X} - 10\bar{Y} + 66 = 0$ & $40\bar{X} - 18\bar{Y} = 214$. Solving, we get $\bar{X} = 13$, $\bar{Y} = 17$.

(ii) Let $8X - 10Y + 66 = 0$ and $40X - 18Y = 214$ be the lines of regression of Y on X and X on Y resp. These eqs. can be put in the form:

$Y = \frac{8}{10}X + \frac{66}{10}$ & $X = \frac{18}{40}Y + \frac{214}{40}$.

$\therefore b_{YX}$ = Regression coeff. of Y on X $= \frac{8}{10} = \frac{4}{5}$.

& b_{XY} = " " X on Y $= \frac{18}{40} = \frac{9}{20}$.

Hence, $r^2 = b_{YX}.b_{XY} = \frac{4}{5}\times\frac{9}{20} = \frac{9}{25}$ $\therefore r = \pm\frac{3}{5} = \pm0{\cdot}6$.

But since both the regression coefficients are +ve, we take $\boxed{r = +0.6}$.

(iii) We have $b_{YX} = r.\frac{\sigma_Y}{\sigma_X} \Rightarrow \frac{4}{5} = \frac{3}{5}\times\frac{\sigma_Y}{3}$. Hence $\boxed{\sigma_Y = 4}$

Scanned with Cam