20BDS0405

① Find the coefficient of correlation between industrial production and export using following data:-

| Production(X) | 55 | 56 | 58 | 59 | 60 | 60 | 62 |
|---|---|---|---|---|---|---|---|
| Export(Y) | 35 | 38 | 37 | 39 | 44 | 43 | 44 |

↳ Sol^n:-

| X | Y | $X^2$ | $Y^2$ | XY |
|---|---|---|---|---|
| 55 | 35 | 3025 | 1225 | 1925 |
| 56 | 38 | 3136 | 1444 | 2128 |
| 58 | 37 | 3364 | 1369 | 2146 |
| 59 | 39 | 3481 | 1521 | 2301 |
| 60 | 44 | 3600 | 1936 | 2640 |
| 60 | 43 | 3600 | 1849 | 2580 |
| 62 | 44 | 3844 | 1936 | 2728 |
| $\Sigma X=$ 410 | $\Sigma Y=$ 280 | $\Sigma X^2=$ 24050 | $\Sigma Y^2=$ 11280 | $\Sigma XY=$ 16448 |

Here, N=7,

$E(X) = \bar{X} = \frac{\Sigma X}{N} = \frac{410}{7} = 58.57$

$E(Y) = \bar{Y} = \frac{\Sigma Y}{N} = \frac{280}{7} = 40$

$E(X^2) = \frac{\Sigma X^2}{N} = \frac{24050}{7} = 3435.71$

$E(Y^2) = \frac{\Sigma Y^2}{N} = \frac{11280}{7} = 1611.43$

$E(XY) = \frac{\Sigma XY}{N} = \frac{16448}{7} = 2349.7$

$Cov(X,Y) = E(XY) - E(X)\cdot E(Y) = 2349.7 - 58.57\times 40 = 6.914$

$\sigma_X = \sqrt{EX^2 - (EX)^2} = \sqrt{3435.71 - 58.57^2} = 2.29$

$\sigma_Y = \sqrt{EY^2 - (EY)^2} = \sqrt{1611.43 - 40^2} = 3.38$

Coefficient of correlation $(r) = \frac{Cov(X,Y)}{\sigma_X \cdot \sigma_Y}$

$= \frac{6.914}{2.29\times 3.38} = 0.893$

$\boxed{r = 0.893}$

20BDS0405

② Calculate the coefficient of correlation and obtain the least square regression line of y on x for the following data :-

| X | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|---|---|---|---|---|---|---|---|---|---|
| Y | 9 | 8 | 10 | 12 | 11 | 13 | 14 | 16 | 15 |

Also, estimate the value of y corresponding to the average to x = 6.2.

| X | Y | $x^2$ | $y^2$ | XY |
|---|---|---|---|---|
| 1 | 9 | 1 | 81 | 9 |
| 2 | 8 | 4 | 64 | 16 |
| 3 | 10 | 9 | 100 | 30 |
| 4 | 12 | 16 | 144 | 48 |
| 5 | 11 | 25 | 121 | 55 |
| 6 | 13 | 36 | 169 | 78 |
| 7 | 14 | 49 | 196 | 98 |
| 8 | 16 | 64 | 256 | 128 |
| 9 | 15 | 81 | 225 | 135 |
| $\Sigma x =$ 45 | $\Sigma y =$ 108 | $\Sigma x^2 =$ 285 | $\Sigma y^2 =$ 1356 | $\Sigma xy =$ 597 |

$$E(x) = \bar{x} = \frac{\Sigma x}{N} = \frac{45}{9} = 5$$

$$E(y) = \bar{y} = \frac{\Sigma y}{N} = \frac{108}{9} = 12$$

$$E(x^2) = \frac{\Sigma x^2}{N} = \frac{285}{9} = 31.67$$

$$E(y^2) = \frac{\Sigma y^2}{N} = \frac{1356}{9} = 150.67$$

$$E(xy) = \frac{\Sigma xy}{N} = \frac{597}{9} = 66.33$$

Correlation Coefficient ($r$) $= \frac{Cov(X,Y)}{\sigma_x \cdot \sigma_y}$

$$= \frac{E(XY) - E(X) \cdot E(Y)}{\sqrt{Ex^2 - (Ex)^2}\sqrt{Ey^2 - (Ey)^2}}$$

$$= \frac{66.33 - 5 \times 12}{\sqrt{31.67 - 5^2}\sqrt{150.67 - 12^2}}$$

$$= 0.949$$

Regression Coefficient of Y on X is :-

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

$$= r \cdot \frac{\sqrt{Ey^2 - (Ey)^2}}{\sqrt{Ex^2 - (Ex)^2}}$$

$$= 0.949 \times \frac{\sqrt{150.67 - 12^2}}{\sqrt{31.67 - 25}}$$

$$= 0.949$$

Then, Regression Eqn of Y on X is:-

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

$$\Rightarrow Y - 12 = 0.949(X - 5)$$

$$\Rightarrow Y = 0.949x + 12 - 0.949 \times 5$$

$$\Rightarrow \boxed{Y = 0.949x + 7.255}$$

When $x = 6.2$,

$$Y = 0.949 \times 6.2 + 7.255$$

$$\Rightarrow \boxed{Y = 13.139}$$

20BDS0405

② Calculate the coefficient of correlation and obtain the least square regression line of y on x for the following data :-

| X | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|---|---|---|---|---|---|---|---|---|---|
| Y | 9 | 8 | 10 | 12 | 11 | 13 | 14 | 16 | 15 |

Also, estimate the value of y corresponding to the average to x = 6.2.

| X | Y | $X^2$ | $Y^2$ | XY |
|---|---|---|---|---|
| 1 | 9 | 1 | 81 | 9 |
| 2 | 8 | 4 | 64 | 16 |
| 3 | 10 | 9 | 100 | 30 |
| 4 | 12 | 16 | 144 | 48 |
| 5 | 11 | 25 | 121 | 55 |
| 6 | 13 | 36 | 169 | 78 |
| 7 | 14 | 49 | 196 | 98 |
| 8 | 16 | 64 | 256 | 128 |
| 9 | 15 | 81 | 225 | 135 |
| $\Sigma X = 45$ | $\Sigma Y = 108$ | $\Sigma X^2 = 285$ | $\Sigma Y^2 = 1356$ | $\Sigma XY = 597$ |

$$E(X) = \bar{X} = \frac{\Sigma X}{N} = \frac{45}{9} = 5$$

$$E(Y) = \bar{Y} = \frac{\Sigma Y}{N} = \frac{108}{9} = 12$$

$$E(X^2) = \frac{\Sigma X^2}{N} = \frac{285}{9} = 31.67$$

$$E(Y^2) = \frac{\Sigma Y^2}{N} = \frac{1356}{9} = 150.67$$

$$E(XY) = \frac{\Sigma XY}{N} = \frac{597}{9} = 66.33$$

Correlation Coefficient $(r) = \frac{Cov(X,Y)}{\sigma_x \times \sigma_y}$

$$= \frac{E(XY) - E(X).E(Y)}{\sqrt{E_X^2 - (E_X)^2}\sqrt{E_Y^2 - (E_Y)^2}}$$

$$= \frac{66.33 - 5 \times 12}{\sqrt{31.67 - 5^2}\sqrt{150.67 - 12^2}}$$

$$= 0.949$$

Regression Coefficient of Y on X is :-

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

$$= r . \frac{\sqrt{E_{Y^2} - (E_Y)^2}}{\sqrt{E_{X^2} - (E_X)^2}}$$

$$= 0.949 \times \frac{\sqrt{150.67 - 12^2}}{\sqrt{31.67 - 25}}$$

$$= 0.949$$

Then, Regression Eqn of Y on X is :-

$$Y - \bar{Y} = b_{yx}(X - \bar{X})$$

$$\Rightarrow Y - 12 = 0.949(X - 5)$$

$$\Rightarrow Y = 0.949X + 12 - 0.949 \times 5$$

$$\Rightarrow \boxed{Y = 0.949x + 7.255}$$

When x = 6.2,

$$Y = 0.949 \times 6.2 + 7.255$$

$$\Rightarrow \boxed{Y = 13.139}$$

(3). The two lines of regression are $4x-5y+33=0$ and $20x-9y-107=0$. The variance of X is 25. Find (i) the mean values of X and Y
(ii) the variance of Y.
(iii) the Correlation Coefficient betn X & Y.

↳ Soln

20BDS0405

$4x-5y+33=0$ ———— eqn (i).
$20x-9y-107=0$ ———— eqn (ii).

- eqn (ii) $-5\times$ eqn (i) gives,
$20x-9y-107-20x+25y-165=0$
$\Rightarrow 16y = 272$
$\Rightarrow y = 17$
Similarly, $x = 13$. Hence, the mean is $(13,17) = (\bar{x},\bar{y})$.

From inspection, we can observe that the first eqn is regression equation of Y on X and second one is X on Y.
From eqn (i) and (ii),
Slope of eqn (i) = regression coefficient of Y on X ($b_{yx}$)
$= \frac{-4}{-5}$
$= 4/5$
Similarly,
Slope of eqn (ii) = regression coefficient of X on Y ($b_{xy}$)
$= 1/\left(\frac{-20}{-9}\right)$
$= 9/20$

Since $b_{xy} = \frac{9}{20}$, $b_{yx} = \frac{4}{5}$.
Correlation Coefficient of X and Y is:- $r = \sqrt{b_{xy}\cdot b_{yx}}$
$\Rightarrow r = \sqrt{\frac{4}{5}\times\frac{9}{20}} = \frac{3}{5} = 0.6$

Thus, $b_{xy} = r.\frac{\sigma_x}{\sigma_y}$
$\Rightarrow \sigma_y = \frac{r.\sigma_x}{b_{xy}} = \frac{0.6\times\sqrt{25}}{0.45} = \frac{0.6\times 5}{0.45}$
$\Rightarrow \sigma_y = \frac{20}{3} = 6.67$
Variance of $Y = \sigma_y^2 = \left(\frac{20}{3}\right)^2 = \frac{400}{9} = 44.44$

Hence, Ans:-
$(\bar{x},\bar{y}) = (13,17)$
$\sigma_y^2 = 44.44$
$r = 0.6$

4). Find the coefficient of correlation for the following table:-

20BDS0405

| X | 10 | 14 | 18 | 22 | 26 | 30 |
|---|---|---|---|---|---|---|
| Y | 18 | 12 | 24 | 6 | 30 | 36 |

| $X$ | $Y$ | $X^2$ | $Y^2$ | $XY$ |
|---|---|---|---|---|
| 10 | 18 | 100 | 324 | 180 |
| 14 | 12 | 196 | 144 | 168 |
| 18 | 24 | 324 | 576 | 432 |
| 22 | 6 | 484 | 36 | 132 |
| 26 | 30 | 676 | 900 | 780 |
| 30 | 36 | 900 | 1296 | 1080 |
| $\Sigma X=$ 120 | $\Sigma Y=$ 126 | $\Sigma X^2=$ 2680 | $\Sigma Y^2=$ 3276 | $\Sigma XY=$ 2772. |

$N=6$.

Here, $E(x)=\bar{X}=\frac{\Sigma X}{N}=\frac{120}{6}=20$

$E(Y)=\bar{Y}=\frac{\Sigma Y}{N}=\frac{126}{6}=21$

$E(X^2)=\frac{\Sigma X^2}{N}=\frac{2680}{6}=446.67$

$E(Y^2)=\frac{\Sigma Y^2}{N}=\frac{3276}{6}=546$ , $E(XY)=\frac{\Sigma XY}{N}=462$

Now, standard deviation of X and Y are:-

$$\sigma_x=\sqrt{\frac{\Sigma X^2}{N}-\left(\frac{\Sigma X}{N}\right)^2}=\sqrt{446.67-20^2}$$
$$=6.83$$

Variance$(\sigma_x^2)=46.67$

$$\sigma_y=\sqrt{\frac{\Sigma Y^2}{N}-\left(\frac{\Sigma Y}{N}\right)^2}=\sqrt{546-21^2}$$
$$=10.247$$

Variance$(\sigma_y^2)=105$.

Covariance of X and Y is:- $E(XY)-E(X).E(Y)$
$$=462-20\times21$$
$$=42.$$

Coefficient of correlation $r=\frac{Cov.(X,Y)}{\sqrt{Var(X)}.\sqrt{Var(Y)}}=\frac{42}{6.83\times10.247}$
$$=0.6001$$

20BDS0406

(5) Obtain the line of regression from the following data. Also, test the validity of linear regression model.

| X: | 25 | 28 | 35 | 32 | 31 | 36 | 29 | 38 | 34 | 32 |
|---|---|---|---|---|---|---|---|---|---|---|
| Y: | 43 | 46 | 49 | 41 | 36 | 32 | 31 | 30 | 33 | 39 |

| X | Y | $X^2$ | $Y^2$ | XY |
|---|---|---|---|---|
| 25 | 43 | 625 | 1849 | 1075 |
| 28 | 46 | 784 | 2116 | 1288 |
| 35 | 49 | 1225 | 2401 | 1715 |
| 32 | 41 | 1024 | 1681 | 1312 |
| 31 | 36 | 961 | 1296 | 1116 |
| 36 | 32 | 1296 | 1024 | 1152 |
| 29 | 31 | 841 | 961 | 899 |
| 38 | 30 | 1444 | 900 | 1140 |
| 34 | 33 | 1156 | 1089 | 1122 |
| 32 | 39 | 1024 | 1521 | 1248 |
| $\Sigma X=$ 320 | $\Sigma Y=$ 380 | $\Sigma X^2=$ 10380 | $\Sigma Y^2=$ 14838 | $\Sigma XY=$ 12067 |

Here,

$E(XY)=\frac{\Sigma XY}{N}=\frac{12067}{10}=1206.7$

$E(X)=\bar{X}=\frac{\Sigma X}{N}=\frac{320}{10}=32$

$E(Y)=\bar{Y}=\frac{\Sigma Y}{N}=\frac{380}{10}=38$

~~$\sigma_x = \bar{Y}$~~

$E(X^2)=\frac{\Sigma X^2}{N}=\frac{10380}{10}=1038$

$E(Y^2)=\frac{\Sigma Y^2}{N}=\frac{14838}{10}=1483.8$

$\sigma_x=\sqrt{E(X^2)-(E(X))^2}$

$=\sqrt{1038-32^2}$

$=3.742$

$\sigma_y=\sqrt{E(Y^2)-(E(Y))^2}$

$=\sqrt{1483.8-38^2}$

$=6.309$

Now, Coefficient of correlation of X and Y is:-

$r=\frac{Cov(X,Y)}{\sigma_x \sigma_y}=\frac{E(XY)-E(X).E(Y)}{\sigma_x.\sigma_y}$

$=\frac{1206.7-32\times38}{3.742\times6.309}$

$=-0.3939$

Coefficient of regression of X on Y is:- $b_{xy}=r\frac{\sigma_x}{\sigma_y}=-0.3939\times\frac{3.742}{6.309}$

$=-0.234$

Regression Eq^n of X on Y is:- $(X-\bar{X})=b_{xy}(Y-\bar{Y})$

$\Rightarrow X-32=-0.234(Y-38)$

$\Rightarrow \boxed{X=-0.234Y+40.892}$

Coefficient of regression of Y on X is:- $b_{yx}=r\frac{\sigma_y}{\sigma_x}=-0.3939\times\frac{6.309}{3.742}$

$=-0.664$

Regression Eq^n of Y on X is:- $Y-\bar{Y}=b_{yx}(X-\bar{X})$

$\Rightarrow (Y-38)=-0.664(X-32)$

$\Rightarrow \boxed{Y=-0.664X+59.248}$

For verification, we calculate $b_{xy}$ and $b_{yx}$ from another method and show that it is equal.

$$b_{xy} = \frac{n\sum xy - \sum x \cdot \sum y}{n\sum y^2 - (\sum y)^2} = \frac{10 \times 12067 - 320 \times 380}{10 \times 14838 - 380^2} = \frac{-930}{3980} = -0.2337.$$

$$b_{yx} = \frac{n\sum xy - \sum x \cdot \sum y}{n\sum x^2 - (\sum x)^2} = \frac{10 \times 12067 - 320 \times 380}{10 \times 10380 - 320^2} = \frac{-930}{1400} = -0.664$$

The values of regression coefficients $b_{xy}$ and $b_{yx}$ obtained from this method is same as obtained earlier.

Hence, the line of regression is valid and linear model of regression is valid as well.

20BDS0405

6). The ranks of some 7 students in Mathematics and Physics are as follows. Calculate the rank correlation coefficient for proficiency in mathematics and physics.

| Rank in Maths X | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|---|
| Rank in Maths Y | 4 | 3 | 1 | 2 | 6 | 5 | 7 |

| $R_x$ | $R_y$ | $d = R_x - R_y$ | $d^2$ |
|---|---|---|---|
| 1 | 4 | -3 | 9 |
| 2 | 3 | -1 | 1 |
| 3 | 1 | 2 | 4 |
| 4 | 2 | 2 | 4 |
| 5 | 6 | -1 | 1 |
| 6 | 5 | 1 | 1 |
| 7 | 7 | 0 | 0 |

$\Sigma d^2 = 20$

Here, $n = 7$

Using formula of rank correlation coefficient,

$$\rho = 1 - \frac{6 \Sigma d^2}{n(n^2-1)}$$

$$= 1 - \frac{6 \times 20}{7(48)}$$

$$= 0.643$$

20BDS0405

7. A sample of 12 fathers and their eldest sons have the following data about their height in inches. Calculate the rank correlation coefficient.

| Fathers (x) | 65 | 63 | 67 | 64 | 68 | 62 | 70 | 66 | 68 | 67 | 69 | 71 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Sons (Y) | 68 | 66 | 68 | 65 | 69 | 66 | 68 | 65 | 71 | 67 | 68 | 70 |

| X | Y | Rx | Ry | d=(Rx-Ry) | $d^2$ |
|---|---|---|---|---|---|
| 65 | 68 | 9 | 5.5 | 3.5 | 12.25 |
| 63 | 66 | 11 | 9.5 | 1.5 | 2.25 |
| 67 | 68 | 6.5 | 5.5 | 1 | 1 |
| 64 | 65 | 10 | 11.5 | -1.5 | 2.25 |
| 68 | 69 | 4.5 | 3 | 1.5 | 2.25 |
| 62 | 66 | 12 | 9.5 | 2.5 | 6.25 |
| 70 | 68 | 2 | 5.5 | -3.5 | 12.25 |
| 66 | 65 | 8 | 11.5 | -3.5 | 12.25 |
| 68 | 71 | 4.5 | 1 | 3.5 | 12.25 |
| 67 | 67 | 6.5 | 8 | -1.5 | 2.25 |
| 69 | 68 | 3 | 5.5 | -2.5 | 6.25 |
| 71 | 70 | 1 | 2 | -1 | 1 |

$\Sigma d^2 = 72.5$

$N = 12$

Here, Repeated data are:-

For X,

67 (2 times) and 68 (2 times).

For Y,

68 (4 times) and 66 (2 times) and 65 (2 times).

$\therefore S_1 = 2, S_2 = 2, S_3 = 4, S_4 = 2, S_5 = 2.$

Then,

$$Adj\cdot(\Sigma d^2) = \Sigma d^2 + \frac{S_1^3 - S_1}{12} + \frac{S_2^3 - S_2}{12} + \frac{S_3^3 - S_3}{12} + \frac{S_4^3 - S_4}{12} + \frac{S_5^3 - S_5}{12}$$

$$= 72.5 + \frac{8-2}{12} + \frac{8-2}{12} + \frac{64-4}{12} + \frac{8-2}{12} + \frac{8-2}{12}$$

$$= 79.5$$

Then, Rank Correlation Coefficient $(\rho) = 1 - \frac{6\Sigma d^2}{n(n^2-1)}$

$$= 1 - \frac{6 \times 79.5}{12(144-1)} \boxed{= 0.722}$$

|20BDS0405|

(8) From the data relating to the yield of dry bark ($X_1$), height ($X_2$) and girth ($X_3$) for 18 Cinchona plants the following simple correlation coefficients were obtained:-

$r_{12} = 0.77$, $r_{13} = 0.73$, $r_{23} = 0.52$.

Find the partial correlation coefficient coefficient $r_{12.3}$ and Multiple correlation coefficient $R_{1.23}$.

↳ Sol:

Here, It is given that,

$r_{12} = 0.77$

$r_{13} = 0.73$

$r_{23} = 0.52$.

Then, Partial correlation Coefficient $(r_{12.3}) = \dfrac{r_{12} - r_{13} \cdot r_{23}}{\sqrt{1-r_{13}^2}\sqrt{1-r_{23}^2}}$

$= \dfrac{0.77 - 0.73 \times 0.52}{\sqrt{1-0.73^2} \cdot \sqrt{1-0.52^2}}$

$= 0.6687$

Now,

Multiple Correlation Coefficient $(R_{1.23}) = \sqrt{\dfrac{r_{12}^2 + r_{13}^2 - 2r_{12} \cdot r_{13} \cdot r_{23}}{1 - r_{23}^2}}$

$= \sqrt{\dfrac{0.77^2 + 0.73^2 - 2 \times 0.77 \times 0.73 \times 0.52}{1 - 0.52^2}}$

$= 0.861$

208DS0405

(9) The sale of a product in lakhs of rupees (Y) is expected to be influenced by two variables namely advertising expenditure $X_1$ (in 1000 Rs) and the number of sales persons ($X_2$) in a region. Following is the sample data for 8 states of the region.
Develop a multiple regression model for the data:-

| Area | Y | $X_1$ | $X_2$ |
|---|---|---|---|
| 1 | 110 | 30 | 11 |
| 2 | 80 | 40 | 10 |
| 3 | 70 | 20 | 7 |
| 4 | 120 | 50 | 15 |
| 5 | 150 | 60 | 19 |
| 6 | 90 | 40 | 12 |
| 7 | 70 | 20 | 8 |
| 8 | 120 | 60 | 14 |

Sol<sup></sup>

The regression model for above data will be in the form

$$Y = b_0 + b_1X_1 + b_2X_2 \quad \text{—— eqn ①}$$

The equation of normals are:-

$$\Sigma Y = nb_0 + b_1\Sigma X_1 + b_2\Sigma X_2 \quad \text{—— eqn ②}$$

$$\Sigma YX_1 = b_0\Sigma X_1 + b_1\Sigma X_1^2 + b_2\Sigma X_1X_2 \quad \text{—— eqn ③}$$

$$\Sigma YX_2 = b_0\Sigma X_2 + b_1\Sigma X_1X_2 + b_2\Sigma X_2^2 \quad \text{—— eqn ④}$$

Hence, the data analysis table is:-

| Y | $X_1$ | $X_2$ | $X_1^2$ | $X_2^2$ | $YX_1$ | $YX_2$ | $X_1X_2$ |
|---|---|---|---|---|---|---|---|
| 110 | 30 | 11 | 900 | 121 | 3300 | 1210 | 330 |
| 80 | 40 | 10 | 1600 | 100 | 3200 | 800 | 400 |
| 70 | 20 | 7 | 400 | 49 | 1400 | 490 | 140 |
| 120 | 50 | 15 | 2500 | 225 | 6000 | 1800 | 750 |
| 150 | 60 | 19 | 3600 | 361 | 9000 | 2850 | 1140 |
| 90 | 40 | 12 | 1600 | 144 | 3600 | 1080 | 480 |
| 70 | 20 | 8 | 400 | 64 | 1400 | 560 | 160 |
| 120 | 60 | 14 | 3600 | 196 | 7200 | 1680 | 840 |
| $\Sigma Y =$ 810 | $\Sigma X_1 =$ 320 | $\Sigma X_2 =$ 96 | $\Sigma X_1^2 =$ 14600 | $\Sigma X_2^2 =$ 1260 | $\Sigma YX_1 =$ 35100 | $\Sigma YX_2 =$ 10470 | $\Sigma X_1X_2 =$ 4240 |

120BDS0405

Here, $n=8$

$\Sigma X_1 = 320$, $\Sigma X_2 = 96$

$\Sigma X_1^2 = \cancel{810}14600$, $\Sigma X_2^2 = 1260$

$\Sigma Y = 810$

$\Sigma X_1 X_2 = 4240$, $\Sigma Y X_1 = 35100$ $\Sigma Y X_2 = 10470$

Thus, normal equations become:-

$$810 = 8b_0 + 320b_1 + 96b_2 \text{ ——— eq}^n \text{ (5)}$$

$$35100 = 320b_0 + 14600b_1 + 4240b_2 \text{ ——— eq}^n \text{ (6)}$$

$$10470 = 96b_0 + 4240b_1 + 1260b_2 \text{ ——— eq}^n \text{ (7)}$$

Solving equation (5) and (6),

Doing (6) − 40×(5), we get,

$$35100 - 40\times 810 = 320b_0 - 8\times 40 b_0 + 14600b_1 - 320\times 40 b_1 + 4240b_2 - 96\times 40 b_2$$

$$\Rightarrow 2700 = 1800b_1 + 400b_2$$

$$\Rightarrow 27 = 18b_1 + 4b_2$$

$$\Rightarrow b_1 = \frac{27 - 4b_2}{18} \text{ ——— eq}^n \text{ (8)}$$

similarly, Doing eq$^n$ (7) − 12×(5), we get,

$$10470 - 12\times 810 = 96b_0 - 8\times 12 b_0 + 4240b_1 - 12\times 320 b_1 + 1260b_2 - 96\times 12 b_2$$

$$\Rightarrow 750 = 400b_1 + 108b_2$$

$$\Rightarrow 750 = \frac{\cancel{400}^{200}}{\cancel{18}_{9}}(27 - 4b_2) + 108b_2$$

$$\Rightarrow 6750 = 5400 - 800b_2 + 972b_2$$

$$\Rightarrow 1350 = 172b_2$$

$$\Rightarrow b_2 = \frac{675}{86} = 7.849$$

Similarly, from eq$^n$ (8),

$$b_1 = \frac{27 - 4\times 7.849}{18} = -0.244$$

Substituting $b_1, b_2$ in eq$^n$ (5), we get,

$$810 = 8b_0 + 320\times(-0.244) + 96\times 7.849$$

$$\Rightarrow \cancel{810}\, 8b_0 = 134.576$$

$$\Rightarrow b_0 = 16.822$$

Thus, the required regression model of multivariables is:-

$$\boxed{Y = 16.822 - 0.244X_1 + 7.849X_2}$$

10. From the following data, obtain $(R_{1\cdot23})^2$. | 20BDS0405

| $X_1$ | 65 | 72 | 54 | 68 | 55 | 59 | 78 | 58 | 57 | 51 |
|---|---|---|---|---|---|---|---|---|---|---|
| $X_2$ | 56 | 58 | 48 | 61 | 50 | 51 | 55 | 48 | 52 | 42 |
| $X_3$ | 9 | 11 | 8 | 13 | 10 | 8 | 11 | 10 | 11 | 7 |

Sol.

Here, n = 10 (no. of observations).

| $X_1$ | $X_2$ | $X_3$ | $X_1X_2$ | $X_2X_3$ | $X_1X_3$ | $X_1^2$ | $X_2^2$ | $X_3^2$ |
|---|---|---|---|---|---|---|---|---|
| 65 | 56 | 9 | 3640 | 504 | 585 | 4225 | 3136 | 81 |
| 72 | 58 | 11 | 4176 | 638 | 792 | 5184 | 3364 | 121 |
| 54 | 48 | 8 | 2592 | 384 | 432 | 2916 | 2304 | 64 |
| 68 | 61 | 13 | 4148 | 793 | 884 | 4624 | 3721 | 169 |
| 55 | 50 | 10 | 2750 | 500 | 550 | 3025 | 2500 | 100 |
| 59 | 51 | 8 | 3009 | 408 | 472 | 3481 | 2601 | 64 |
| 78 | 55 | 11 | 4290 | 605 | 858 | 6084 | 3025 | 121 |
| 58 | 48 | 10 | 2784 | 480 | 580 | 3364 | 2304 | 100 |
| 57 | 52 | 11 | 2964 | 572 | 627 | 3249 | 2704 | 121 |
| 51 | 42 | 7 | 2142 | 294 | 357 | 2601 | 1764 | 49 |
| $\Sigma X_1 =$ 617 | $\Sigma X_2 =$ 463 | $\Sigma X_3 =$ 98 | $\Sigma X_1X_2 =$ 32495 | $\Sigma X_2X_3 =$ 5178 | $\Sigma X_3X_1 =$ 6137 | $\Sigma X_1^2 =$ 38753 | $\Sigma X_2^2 =$ 27423 | $\Sigma X_3^2 =$ 990 |

Here, $E(X_1) = \frac{\Sigma X_1}{N} = \frac{617}{10} = 61\cdot7$

$E(X_2) = \frac{\Sigma X_2}{N} = \frac{463}{10} = 46\cdot3$

$E(X_3) = \frac{\Sigma X_3}{N} = \frac{98}{10} = 9\cdot8$

$E(X_1X_2) = \frac{\Sigma X_1X_2}{N} = \frac{32495}{10} = 3249\cdot5$

$E(X_2X_3) = \frac{\Sigma X_2X_3}{N} = \frac{5178}{10} = 517\cdot8$

$E(X_1X_3) = \frac{\Sigma X_1X_3}{N} = \frac{6137}{10} = 613\cdot7$

$E(X_1^2) = \frac{\Sigma X_1^2}{N} = \frac{38753}{10} = 3875\cdot3$

$E(X_2^2) = \frac{\Sigma X_2^2}{N} = \frac{27423}{10} = 2742\cdot3$

$E(X_3^2) = \frac{\Sigma X_3^2}{N} = \frac{990}{10} = 99$

20BDS0405

Now,

$$Cov(X_1,X_2) = E(X_1X_2) - E(X_1)\cdot E(X_2)$$
$$= 3249.5 - 61.7 \times 46.3$$
$$= 392.79$$

$$Cov(X_2X_3) = E(X_2X_3) - E(X_2)\cdot E(X_3) = 517.8 - 46.3 \times 9.8 = 64.06$$

$$Cov(X_1X_3) = E(X_1X_3) - E(X_1)\cdot E(X_3) = 613.7 - 61.7 \times 9.8 = 9.04$$

$$\sigma_{x_1} = \sqrt{Ex_1^2 - (E(X_1))^2} = \sqrt{3875.3 - 61.7^2} = \sqrt{68.41} = 8.27$$

$$\sigma_{x_2} = \sqrt{Ex_2^2 - (E(X_2))^2} = \sqrt{2742.3 - 46.3^2} = \sqrt{598.61} = 24.466$$

$$\sigma_{x_3} = \sqrt{E(X_3^2) - (E(X_3))^2} = \sqrt{99 - 9.8^2} = \sqrt{2.96} = 1.72$$

$$r_{12} = \frac{Cov(X_1,X_2)}{\sigma_{x_1} \times \sigma_{x_2}} = \frac{392.79}{8.27 \times 24.466} = 1.94$$

$$r_{23} = \frac{Cov(X_2,X_3)}{\sigma_{x_2}\cdot\sigma_{x_3}} = \frac{64.06}{24.466 \times 1.72} = 1.522$$

$$r_{13} = \frac{Cov(X_1X_3)}{\sigma_{x_1}\cdot\sigma_{x_3}} = \frac{9.04}{8.27 \times 1.72} = 0.635$$

~~Hence~~, Multiple correlation coefficient $(R_{1.23}) = ?$

$$R_{1.23} = \sqrt{\left(\frac{r_{12}^2 + r_{13}^2 - 2r_{12}\cdot r_{23}\cdot r_{13}}{1 - r_{23}^2}\right)}$$

$$= \sqrt{\frac{0.8^2 + 0.64^2 - 2 \times 0.8 \times 0.79 \times 0.64}{1 - 0.79^2}}$$

$$= \sqrt{0.64}$$

$$\Rightarrow R_{1.23}^2 = (\sqrt{0.64})^2 = 0.64$$

$$\boxed{\therefore R_{1.23} = 0.64}$$

Scanned with CamScanner