H/W

Chapter-10
Correlation and Regression

Page No.
Date: / /

Important formulas:

γ = Karl Pearson's correlation coefficient

1. $$\gamma = \frac{Cov(x, y)}{\sqrt{Var(x)} \cdot \sqrt{Var(y)}}$$

Where,

$$Cov(X, Y) = \frac{1}{n} \Sigma (x - \bar{x})(y - \bar{y})$$

2. $$\gamma = \frac{n \Sigma XY - \Sigma X \cdot \Sigma Y}{\sqrt{n \Sigma X^2 - (\Sigma X)^2} \cdot \sqrt{n \Sigma Y^2 - (\Sigma Y)^2}}$$

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

4. $$\gamma = \frac{\Sigma xy}{n \sigma_x \cdot \sigma_y}$$

5. $$\gamma = \frac{\Sigma xy}{\sqrt{\Sigma x^2} \cdot \sqrt{\Sigma y^2}}$$

Exercise - 10.1

1. Find the correlation coefficient between the two variables under the ~~foll~~ following conditions.

a. Cov (X, Y) = covariance of the variables X and Y = 18

Variance of X ~~and~~ = 16 and the variance of Y = 81.

Solution:

We know,

$$\gamma = \frac{Cov(X,Y)}{\sqrt{Var(X)} \cdot \sqrt{Var(Y)}}$$

or, $$\gamma = \frac{18}{\sqrt{16} \cdot \sqrt{81}}$$

$$\therefore \gamma = 0.5$$

b. $Cov(X,Y) = -16.5$, $Var(X) = 2.89$ and $Var(Y) = 100$

Solution:

We know,

$$\gamma = \frac{Cov(X,Y)}{\sqrt{Var(X)} \cdot \sqrt{Var(Y)}}$$

$$= \frac{-16.5}{\sqrt{2.89} \cdot \sqrt{100}}$$

$$= -0.97$$

c. $\Sigma(X-\bar{X})^2 = 40$, $\Sigma(Y-\bar{Y})^2 = 63$ and $\Sigma(X-\bar{X}).(Y-\bar{Y}) = 35$.

Solution:

We know,

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

$$= \frac{35}{\sqrt{40} \cdot \sqrt{63}}$$

$$= 0.697$$

Date: / /

d. $n = 15$, $\sigma_x = 3.2$, $\sigma_y = 3.4$ and $\Sigma(x-\bar{x})(Y-\bar{Y})$ $= 122$.

Solution:

We know,

$$r = \frac{\Sigma(x-\bar{x})(y-\bar{y})}{n\,\sigma_x \cdot \sigma_y}$$

$$= \frac{122}{15 \times 3.2 \times 3.4}$$

$$= 0.75$$

e. $n = 10$, $\Sigma X = 60$, $\Sigma Y = 60$, $\Sigma x^2 = 400$, $\Sigma Y^2 = 580$ and $\Sigma XY = 415$

Solution:

We know,

$$r = \frac{n\Sigma XY - \Sigma X \cdot \Sigma Y}{\sqrt{n\Sigma x^2 - (\Sigma x)^2} \cdot \sqrt{n\Sigma y^2 - (\Sigma y)^2}}$$

$$= \frac{415 - 60 \times 60}{\sqrt{10(400) - (60)^2} \cdot \sqrt{10(580) - (60)^2}}$$

$$= 0.59$$

f. $n = 10$, $\bar{x} = 5$, $\bar{Y} = 3$, $\Sigma x^2 = 290$, $\Sigma Y^2 = 300$, $\Sigma XY = 115$

Solution.

We know,

$$r = \frac{\Sigma XY - n\bar{x}\bar{y}}{\sqrt{\Sigma x^2 - n\bar{x}^2} \cdot \sqrt{\Sigma y^2 - n\bar{y}^2}}$$

$$= \frac{115 - 10 \times 5 \times 3}{\sqrt{290 - 10 \times 5^2} \cdot \sqrt{300 - 10 \times 3^2}}$$

$$= -0.38$$

9.

	Series X	Series Y
No. of pair of observation	10	10
Standard deviation	2.05	2.41
Sum of square of ~~derivati~~ deviations from their resp. means:	42	58

Sum of products of deviations of X and Y from their respective means = 36.

Solutions:

Here, $n = 10$

$\sigma_x = 2.05$

$\sigma_y = 2.41$

$\Sigma(x-\bar{x})^2 = 42$

$\Sigma(y-\bar{y})^2 = 58$

$\Sigma(x-\bar{x})(y-\bar{y}) = 36$

We know,

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

$$= \frac{36}{\sqrt{42} \cdot \sqrt{58}}$$

$$= 0.73$$

2.a. Calculate Karl Pearson's correlation coefficient between the two variables height (in cm) and weight (in kg) from the data given below.

Height:	160	162	165	161	162
Weight:	63	62	64	60	61

Solution:

X	Y	X^2	Y^2	XY
160	63	25600	3969	10080
162	62	26244	3844	1044
165	64	27225	4096	10560
161	60	25921	3600	9660
162	61	26244	3721	9882
$\Sigma X = 810$	$\Sigma Y = 310$	$\Sigma X^2 = 131234$	$\Sigma Y^2 = 19230$	$\Sigma XY = 50226$

Now,

$$r = \frac{n\Sigma XY - \Sigma X . \Sigma Y}{\sqrt{n\Sigma X^2 - (\Sigma X)^2} \cdot \sqrt{n\Sigma Y^2 - (\Sigma Y)^2}}$$

$$= \frac{5 \times 50226 - 810 \times 310}{\sqrt{5 \times 131234 - (810)^2} \cdot \sqrt{5 \times 19230 - (310)^2}}$$

$$= 0.51$$

b. Find the correlation coefficient between the two Variables X and Y from the following data.

X :	5	7	1	3	4
Y :	2	3	4	5	6

Solution:

X	Y	X^2	Y^2	XY
5	2	25	4	10
7	3	49	9	21
1	4	1	16	4
3	5	9	25	15
4	6	16	36	24
$\Sigma X = 20$	$\Sigma Y = 20$	$\Sigma X^2 = 100$	$\Sigma Y^2 = 90$	$\Sigma XY = 74$

Now,

$$r = \frac{n\Sigma XY - \Sigma X . \Sigma Y}{\sqrt{n\Sigma X^2 - (\Sigma X)^2} \cdot \sqrt{n\Sigma Y^2 - (\Sigma Y)^2}}$$

$$= \frac{5 \times 74 - 20 \times 20}{\sqrt{5 \times 100 - 400} \cdot \sqrt{5 \times 90 - 400}}$$

$= -0.42$

c. Calculate Karl Pearson's Correlation coefficient between the sales and expenses in thousand rupees of 5 firms.

Sales :	43	41	36	34	50
Expenses :	12	24	15	21	19

Solution:

Sales (X)	Expenses (Y)	XY	X^2	Y^2
43	12	516	1849	144
41	24	984	1681	576
36	15	540	1296	225
34	21	714	1156	441
50	19	950	2500	361
$\Sigma X = 204$	$\Sigma Y = 91$	$\Sigma XY = 3704$	$\Sigma X^2 = 8482$	$\Sigma Y^2 = 1747$

Now,

$$r = \frac{n\Sigma XY - \Sigma X.\Sigma Y}{\sqrt{n\Sigma X^2 - (\Sigma X)^2} \cdot \sqrt{n\Sigma Y^2 - (\Sigma Y)^2}}$$

$$= \frac{5 \times 3704 - (204 \times 91)}{\sqrt{5 \times 8482 - (204)^2} \cdot \sqrt{5 \times 1747 - (91)^2}}$$

$$= -0.07$$

d. Determine the degree of relationship between the ages of the husbands and their wives from the following data

Age of the husband (X):	23	22	24	23	26	27
Age of the wives (Y):	20	18	20	21	21	22

Solution:

X	Y	XY	X^2	Y^2
23	20	460	529	400
22	18	396	484	324
24	20	480	576	400
23	21	483	529	441
26	21	546	676	441
27	22	594	729	484
$\Sigma X = 145$	$\Sigma Y = 122$	$\Sigma XY = 2959$	$\Sigma X^2 = 3523$	$\Sigma Y^2 = 2490$

We know,

$$r = \frac{n\Sigma XY - \Sigma X \cdot \Sigma Y}{\sqrt{n\Sigma X^2 - (\Sigma X)^2} \cdot \sqrt{n\Sigma Y^2 - (\Sigma Y)^2}}$$

$$= \frac{6 \times 2959 - 145 \times 122}{\sqrt{6 \times 3523 - (145)^2} \cdot \sqrt{6 \times 2490 - (122)^2}}$$

$$= 0.80$$

3. From the following table calculate Karl Pearson's correlation coefficient between the two variables.

X :	6	2	10	4	8
Y :	9	11	?	8	7

Arithmetic means of X and Y are 6 and 8 respectively.

Solution:

X	Y	$x = X - \bar{X}$	$y = Y - \bar{Y}$	x^2	y^2	xy
6	9	0	1	0	1	0
2	11	-4	3	16	9	-12
10	z=5	4	-3	16	9	-12
4	8	-2	0	4	0	0
8	7	2	-1	4	1	-2
$\Sigma X = 30$	$\Sigma Y = 40$			40	20	-26

We know,

$\bar{X} = 6$, $\bar{Y} = 8$

We also know,

$$Y = \frac{\Sigma y}{n} = \frac{9 + 11 + z + 8 + 7}{5}$$

$$\therefore z = 5$$

Now,

$$r = \frac{\Sigma xy}{\sqrt{\Sigma x^2} \cdot \sqrt{\Sigma y^2}}$$

$$= \frac{-26}{\sqrt{40} \cdot \sqrt{20}}$$

$$= -0.92$$

4. In order to find the correlation coefficient between the two variables x and y from 12 pair of observation, the following calculations were made.
$\Sigma X = 30$, $\Sigma Y = 5$, $\Sigma X^2 = 670$, $\Sigma Y^2 = 285$, $\Sigma XY = 334$

On subsequent verification, it was found that the pair $(X = 11, Y = 4)$ was copied wrong; the correct value being $(X = 10, Y = 14)$. Find the correct value of correlation coefficient.

Solution:

Given, $\Sigma X = 30$, $\Sigma Y = 5$, $\Sigma X^2 = 670$, $\Sigma Y^2 = 285$
$\Sigma XY = 334$

If $X = 11$ and $Y = 4$ was copied wrong then the corrected value will be;

Corrected $\Sigma X = 30 - 11 + 10 = 29$

Corrected $\Sigma Y = 5 - 4 + 14 = 15$

Corrected $\Sigma X^2 = 670 - 11^2 + 10^2 = 649$

Corrected $\Sigma Y^2 = 285 - 4^2 + 14^2 = 465$

Corrected $\Sigma XY = 334 - 11 \times 4 + 10 \times 14 = 430$

Now,

$$\text{Corrected, } r = \frac{n \cdot \text{Corr.} \Sigma XY - \text{Corr.} \Sigma X \cdot \text{Corr.} \Sigma Y}{\sqrt{n \cdot \text{corr.} \Sigma X^2 - (\text{Corr.} \Sigma X)^2} \cdot \sqrt{n \cdot \text{Corr.} \Sigma Y^2 - (\text{Corr.} \Sigma Y)^2}}$$

Page No.

Date: / /

$$= \frac{12 \times 430 - 29 \times 15}{\sqrt{12 \times 649 - (29)^2} \cdot \sqrt{12 \times 465 - (15)^2}}$$

$$= 0.775$$