Chapter - 05
Regression Analysis

Meaning of regression, regression lines, regression equation and estimation calculation of regression equation when regression co-efficients are given (simultaneous equation method excluded) Problems.

MEANING:-

When it is known that two variables are closely related or correlated, the estimation of one variable, given the value of other variable can be made.

"It is a statistical tool with the help of which we are in a position, to estimate the unknown value of one variable, from known values of another variable is known as Regression".

The dictionary meaning of the term regression is 'the act of returning back' or stepping back'. Literally regression is opposite to the word 'Regression'.

In statistics regression is used to estimate the probable trend in a related variable when the value of other related variable is given, taking into the account, the past trend.

The term regression was first used by British Biometrician SIR FRANCIS GALTON in 19th century.

REGRESSION ANALYSIS:-

It is a technique which measures the average relationship between two or more variables and helps in estimating the value of dependent variable given the value of independent variable.

DEPENDENT VARIABLE:-

It is one whose value is influenced by other variables. In other words the variable which is to estimated/predicted is called dependent variable.

INDEPENDENT VARIABLE:-

It is one which influences the value of the other related variable. In other words the variables which is used for estimation or prediction is called Independent variable.

REGRESSION LINES / ESTIMATING LINES:-

The line of regression is the straight line which gives the best estimate of one variables for any given value of the other variables. It is a graphic technique to show the functional relationship between two variables.

Ex:- X is dependent variable and Y is independent variable. It shows the average relationship between the two variables X and Y. (and vice-versa).

In case of two variable X and Y as stated earlier we shall have two regression lines they are:-

a) The regression line of X on Y
b) The regression line of Y on X.

REGRESSION EQUATION:-

Regression equations are algebraic expression of the regression lines, for each of the regression line has regression equations.

a) The regression equation of Y on X.
b) The regression equation of X on Y.

a) THE REGRESSION EQUATION OF Y ON X :-

$$Y_c = a + bx$$

$$\Sigma Y = Na + b\Sigma X$$

$$\Sigma XY = a\Sigma X + b\Sigma X^2$$

b) THE REGRESSION EQUATION OF X ON Y :-

$$X_c = a + bY$$

$$\Sigma X = Na + b\Sigma Y$$

$$\Sigma XY = a\Sigma Y + b\Sigma Y^2$$

REGRESSION EQUATIONS THROUGH CORRELATION CO-EFFICIENT:-

A regression co-efficient represents the increment in the value of dependent value for unit change in the value of independent variable. In other words, Regression co-efficient represents the rate of change of dependent variable with respect to the rate of change of independent variable.

In two regression equations, $Y_c = a + b\Sigma$ and $x_c = a + bY$ is regarded as regression co-efficient. As there are two regression lines, two regression equations, there are two regression co-efficient.

a) The regression co-efficient of Y on X

$$b_{yx} = \frac{cov(x, y)}{v(y)} = r\frac{\sigma_y}{\sigma_x}$$

Date
Page

b) The regression co-efficient of X on Y

$$b_{xy} = \frac{cov(x,y)}{v(x)} = r\frac{\sigma_x}{\sigma_y}$$

WHEN DEVIATIONS ARE TAKEN FROM ACTUAL MEAN:-

a) The regression coefficient of Y on X under actual mean.

$$b_{yx} = \frac{\Sigma xy}{\Sigma x^2} = r\frac{\sigma_y}{\sigma_x}.$$

Thus the regression eqⁿ of Y on X can be written as $(Y - \bar{Y}) = b_{yx}(X - \bar{X})$.

b) The regression coefficient of X on Y under actual mean.

$$b_{xy} = \frac{\Sigma xy}{\Sigma y^2} = r\frac{\sigma_x}{\sigma_y}.$$

Regression eqⁿ of Y on X $= (Y - \bar{Y}) = b_{yx}(X - \bar{X})$.

WHEN DEVIATIONS ARE TAKEN FROM ASSUMED MEAN:-

a) The regression co-efficient of Y on X under assumed mean.

$$b_{yx} = \frac{N\Sigma d_x d_y - (\Sigma d_x)(\Sigma d_y)}{N\Sigma d_x^2 - (\Sigma d_x)^2}$$

b) The regression coefficient on X on Y under assumed mean.

$$b_{xy} = \frac{N\Sigma d_x d_y - (\Sigma d_x)(\Sigma d_y)}{N\Sigma d_y^2 - (\Sigma d_y)^2}$$

Kavitha

PROBLEMS ON REGRESSION UNDER ACTUAL MEAN:-

=> Regression equation of y on x

$= (y - \bar{y}) = byx (x - \bar{x})$

$byx = \frac{\Sigma xy}{\Sigma x^2} = r \frac{\sigma y}{\sigma x}$

=> Regression equation of x on y

$= (x - \bar{x}) = bxy (y - \bar{y})$

$bxy = \frac{\Sigma xy}{\Sigma y^2} = r \frac{\sigma x}{\sigma y}$

PROBLEMS :-

1) From the following data obtain the regression equation x on y and the regression equation y on x

x	y	x	x^2	y	y^2	xy
6	9	0	0	1	1	0
4	8	-2	4	0	0	0
8	7	2	4	-1	1	-2
10	5	4	16	-3	9	-12
2	11	-4	16	3	9	-12
30	40		40		20	-26

$\bar{x} = \frac{\Sigma x}{N} = \frac{0}{5} = 6$

$$\bar{Y} = \frac{\Sigma y}{N} = \frac{40}{5} = 8$$

Regression equation Y on X :-

$$(Y - \bar{Y}) = byx\,(X - \bar{X})$$

$$(Y - 8) = byx\,(X - 6)$$

$$(Y - 8) = \frac{\Sigma xy}{\Sigma x^2}\,(X - 6)$$

$$Y - 8 = \frac{-26}{40}\,(X - 6)$$

$$Y - 8 = -0.65\,(X - 6)$$

$$Y - 8 = -0.65X + 3.9$$

$$Y = -0.65X + 3.9 + 8$$

$$\boxed{Y = -0.65X + 11.9}$$

Regression equation X on Y :-

$$(X - \bar{X}) = bxy\,(Y - \bar{Y})$$

$$(X - 6) = bxy\,(Y - 8)$$

$$X - 6 = \frac{-26}{20}\,(Y - 8)$$

$$X - 6 = -1.3\,(Y - 8)$$

$$X - 6 = -1.3Y + 10.4$$

$$X = -1.3Y + 10.4 + 6$$

$$X = -1.3Y + 16.4$$

(OR)

$$\boxed{X = 16.4 - 1.3Y}$$

26. Formulate both the regression lines from the following data. Predict y when $x = 50$ and x when $y = 25$.

x	y	x	x^2	y	y^2	xy
40	30	1	1	-4	16	-4
32	35	-7	49	1	1	-7
38	40	-1	1	6	36	-6
42	36	3	9	2	4	6
36	28	-3	9	-6	36	18
46	35	7	49	1	1	7
234	204		118		94	14

$$\bar{x} = \frac{\Sigma x}{N} = \frac{234}{6} = 39$$

$$\bar{y} = \frac{\Sigma y}{N} = \frac{204}{6} = 34$$

Regression equation y on x:-

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

$$(y - 34) = b_{yx}(x - 39)$$

$$(y - 34) = \frac{14}{118}(x - 39)$$

$$(y - 34) = 0.1186(x - 39)$$

$$y - 34 = 0.1186x - 4.6254$$

$$y = 0.1186x - 4.6254 + 34$$

$$y = 0.1186x + 29.3746$$

Date
Page

when $x = 50$

$y = 0.1186(50) + 29.3746$

$y = 5.93 + 29.3746$

$y = 35.3046$

Regression equation X on Y :-

$(x - \bar{x}) = bxy\,(y - \bar{y})$

$(x - 39) = bxy\,(y - 34)$

$x - 39 = \frac{14}{94}\,(y - 34)$

$x - 39 = 0.1489\,(y - 34)$

$x - 39 = 0.1489y - 5.0626$

$x = 0.1489y - 5.0626 + 39$

$x = 0.1489y + 33.9374$

when $y = 25$

$x = 0.1489(25) + 33.9374$

$x = 3.7225 + 33.9374$

$x = 37.6599$

34 The following data relate to the ages of Husband and wives.

Age of Husband (yr)	25	28	30	32	35
Age of wives (yr)	20	26	29	30	25
	36	38	39	42	55
	18	26	35	35	46.

Find KPCC and most likely age of husband when wife's age is 25 years.

solut

X	Y	x	x^2	y	y^2	xy
25	20	-11	121	-9	81	99
28	26	-8	64	-3	9	24
30	29	-6	36	0	0	0
32	30	-4	16	1	1	-4
35	25	-1	1	-4	16	4
36	18	0	0	-11	121	0
38	26	2	4	-3	9	-6
39	35	3	9	6	36	18
42	35	6	36	6	36	36
55	46	19	361	17	289	323
360	290		648		598	494

$$\bar{X} = \frac{360}{10} = \underline{\underline{36}}$$

$$\bar{Y} = \frac{290}{10} = \underline{\underline{29}}$$

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

$$= \frac{494}{\sqrt{648 \times 598}}$$

$$= \frac{494}{\sqrt{387504}}$$

$$= \frac{494}{622.4981} = \underline{\underline{0.7935}}$$

$\therefore$ Their exist a high degree positive correlation.

$\Rightarrow$ Regression eqn on x on y :-

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

$$(x-36) = b_{xy}(y-29)$$

$$x-36 = \frac{494}{598}(y-29)$$

$$x-36 = 0.8260(y-29)$$

$$x-36 = 0.8260y - 23.954$$

$$x = 0.8260y - 23.954 + 36$$

$$x = 0.8260y + 12.046$$

when wife $(y) = 25$

$$x = 0.8260(25) + 12.046$$

$$x = 20.65 + 12.046$$

$$x = 32.696$$

$\Rightarrow$ Regression equation on y on x :-

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

$$(y-29) = b_{yx}(x-36)$$

$$(y-29) = \frac{494}{648}(x-36)$$

$$y-29 = 0.7623(x-36)$$

$$y-29 = 0.7623x - 27.4428$$

$$y = 0.7623x - 27.4428 + 29$$

$$y = 0.7623x + 1.5572$$

4) obtain the two regression equations from the following data when the deviations are obtained from mean.

$N = 20$, $\bar{x} = 4$, $\bar{y} = 2$, $\Sigma x^2 = 1680$,
$\Sigma y^2 = 320$, $\Sigma xy = 480$.

solut

→ Regression equation on x on y:-

$$(x - \bar{x}) = b_{xy}(y - \bar{y})$$
$$(x - 4) = b_{xy}(y - 2)$$
$$x - 4 = \frac{480}{320}(y - 2)$$
$$x - 4 = 1.5(y - 2)$$
$$x - 4 = 1.5y - 3$$
$$x = 1.5y - 3 + 4$$
$$x = 1.5y + 1$$

→ Regression equation y on x:-

$$(y - \bar{y}) = b_{yx}(x - \bar{x})$$
$$(y - 2) = b_{yx}(x - 4)$$
$$y - 2 = \frac{480}{1680}(x - 4)$$
$$y - 2 = 0.2857(x - 4)$$
$$y - 2 = 0.2857x - 1.1428$$
$$y = 0.2857x - 1.1428 + 2$$
$$y = 0.2857x + 0.8572$$

PROBLEMS ON REGRESSION EQUATION ON ASSUMED MEAN:-

$\Rightarrow$ Regression of Y on X

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

$$b_{yx} = \frac{N\Sigma dxdy - \Sigma dx \Sigma dy}{N\Sigma dx^2 - (\Sigma dx)^2}$$

$\Rightarrow$ Regression of X on Y = $(X-\bar{X}) = b_{xy}(Y-\bar{Y})$

$$b_{xy} = \frac{N\Sigma dxdy - \Sigma dx \Sigma dy}{N\Sigma dy^2 - (\Sigma dy)^2}$$

5) The following table shows the age (x) and blood pressure (y) of 6 persons.

X	Y
52	62
45	53
36	51
72	75
65	79
47	60

Obtain two regression equations. Also find the expected blood pressure of person who is 50 years old.

Kavitha

Soln :- Assumed mean method

x	y	dx	dx^2	dy	dy^2	dxdy
52	62	16	256	11	121	176
45	53	9	81	2	4	18
36	51	0	0	0	0	0
72	75	36	1296	24	576	864
65	79	29	841	28	784	812
47	60	11	121	9	81	99
317	380	101	2595	74	1566	1969

calculation of regression co-efficient

$$b_{yx} = \frac{N\Sigma dxdy - \Sigma dx \Sigma dy}{N\Sigma dx^2 - (\Sigma dx)^2}$$

$\bar{x} = \frac{\Sigma x}{N} = \frac{317}{6} = 52.83$

$\bar{y} = \frac{\Sigma y}{N} = \frac{380}{6}$

$= 63.333$

$$= \frac{6(1969) - (101)(74)}{6(2595) - (101)^2}$$

$$= \frac{11814 - 7474}{15570 - 10201}$$

$$= \frac{4340}{5369}$$

$$= 0.8083$$

$$b_{xy} = \frac{N\Sigma dx\,dy - \Sigma dx\,\Sigma dy}{N\Sigma dy^2 - (\Sigma dy)^2}$$

$$= \frac{6(1969) - (101)(74)}{6(1566) - (74)^2}$$

$$= \frac{11814 - 7474}{9396 - 5476}$$

$$= \frac{4340}{3920}$$

$$= 1.1071$$

Calculation of regression equations.

=> Regression of Y on X.

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

$$(y - 63.333) = 0.8083\,(x - 52.83)$$

$$y - 63.33 = 0.8083x - 42.7024$$

$$y = 0.8083x - 42.7024 + 63.33$$

$$y = 0.8083x + 20.6276$$

=> Regression of X on Y

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

$$(x - 52.83) = 1.1071\,(y - 63.33)$$

$$x - 52.83 = 1.1071y - 70.1126$$

$$x = 1.1071y - 70.1126 + 52.83$$

$$x = 1.1071y - 17.2826$$

Prediction

when Age = 50 BP = ?

$Y = 0.8083x + 20.6276$

$Y = 0.8083(50) + 20.6276$

$Y = 40.415 + 20.6276$

$Y = 61.0426$

6) Given the bivariate data

x	1	2	3	5	1	1	3	7
y	6	0	0	1	1	2	8	1

a) Fit a regression line y on x and hence predict y if x = 5

b) Fit a regression line x on y and hence predict x if y = 25

soln ⇒ Assumed mean method.

x	y	dx	dx^2	dy	dy^2	dxdy
1	6	0	0	5	25	0
2	0	1	1	-1	1	-1
3	0	2	4	-1	1	-2
5	1	4	16	0	0	0
1	1	0	0	0	0	0
1	2	0	0	1	1	0
3	5	2	4	4	16	8
7	1	6	36	0	0	0
23	16	15	61	8	44	5

$$\bar{x} = \frac{\Sigma x}{N} = \frac{23}{8} = 2.875$$

$$\bar{y} = \frac{\Sigma y}{N} = \frac{16}{8} = 2$$

Calculation of regression co-efficient

$$b_{yx} = \frac{N\Sigma dxdy - \Sigma dx \Sigma dy}{N\Sigma dx^2 - (\Sigma dx)^2}$$

$$= \frac{8(5) - (15)(8)}{8(61) - (15)^2}$$

$$= \frac{40 - 120}{488 - 225}$$

$$= \frac{-80}{263}$$

$$= -0.3041$$

$$b_{xy} = \frac{N\Sigma dxdy - \Sigma dx \Sigma dy}{N\Sigma dx^2 - (\Sigma dx)^2}$$

$$= \frac{8(5) - 15(8)}{8(44) - (8)^2}$$

$$= \frac{40 - 120}{352 - 64} = \frac{-80}{288} = -0.2777$$

Calculation of regression equations

=> Regression of Y on X :-

$(Y-\bar{Y}) = byx\,(x-\bar{x})$

$(Y-2) = -0.3041\,(x-2.875)$

$(Y-2) = -0.3041x + 0.8742$

$Y = -0.3041x + 0.8742 + 2$

$Y = -0.3041x + 2.8742$

when $x = 5$

$Y = -0.3041(5) + 2.8742$

$Y = -1.5205 + 2.8742$

$Y = 1.3537$

=> Regression of x on Y

$(x-\bar{x}) = bxy\,(Y-\bar{Y})$

$(x-2.875) = -0.2777\,(Y-2)$

$x - 2.875 = -0.2777Y + 0.5554$

$x = -0.2777Y + 0.5554 + 2.875$

$x = -0.2777Y + 3.4304$

when $Y = 2.5$

$x = -0.2777(2.5) + 3.4304$

$x = -0.69425 + 3.4304$

$x = 2.73615$

Q6 Calculate 'r' if $bxy = 0.8$ and $byx = 0.96$

Solut

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

$= \sqrt{0.8 \times 0.96}$

Date
Page

$r = \sqrt{0.768}$

$r = 0.8763$

∴ Their exist a high degree positive correlation.

8) Calculate 'r' if $bxy = 0.36$ and $byx = -1.38$

Soln

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

$= \sqrt{-1.38 \times 0.36}$

$= \sqrt{-0.4968}$

$= 0.7048$

NOTE:- Since regression co-efficient are r are of different singus are cannot be computed.

9) Calculate 'r' and two regression co-efficients when $r = 0.9$, $\sigma x = 10$, $\sigma y = 1.5$

Soln Regression equation of y on x

$(y - \bar{y}) = byx\,(x - \bar{x})$

$byx = \frac{\Sigma xy}{\Sigma x^2}$

$= r \times \frac{\sigma y}{\sigma x}$

Date
Page

$= 0.9 \times \frac{1.5}{10}$

$= \frac{1.35}{10}$

$= 0.135$

Regression equation of x on y

$(x - \bar{x}) = bxy\ (y - \bar{y})$

$bxy = \frac{\Sigma xy}{\Sigma y^2} = r \times \frac{\sigma_x}{\sigma_y}$

$= 0.9 \times \frac{10}{1.5}$

$= \frac{9}{1.5}$

$= 6$

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

$= \sqrt{0.135 \times 6}$

$= \sqrt{0.81}$

$= 0.9$

Kavitha

10) If one of the regression co-efficients is 1.5 and 'r' is 0.55. Find the value of the other regression co-efficient.

solu:- $b_{yx} = 1.5$ $r = 0.55$

$$r = \sqrt{b_{yx} \times b_{xy}}$$

$$0.55 = \sqrt{1.5 \times b_{xy}}$$

SBS

$$(0.55)^2 = \left(\sqrt{1.5 \times b_{xy}}\right)^2$$

$$(0.55)^2 = 1.5 \times b_{xy}$$

$$0.3025 = 1.5 \times b_{xy}$$

$$\frac{0.3025}{1.5} = b_{xy}$$

$$b_{xy} = 0.2016$$

11) Find 'r' when regression co-efficients are -0.6 and -1.4.

solu:

$$r = \sqrt{b_{yx} \times b_{xy}}$$

$$= \sqrt{-0.6 \times -1.4}$$

$$= \sqrt{0.84}$$

$$= 0.9165$$

NOTE:- If both the regression co-efficients are negative then r should be negative. Hence r is equal to -0.91. There exist a high degree negative correlation.

124 Findout σ_y and r, when $\sigma_x = 3$, $x = 0.85y$, $Y = 0.89X$.

Solut Regression Equation of x on Y = $X = 0.85Y$ where $b_{xy} = 0.85$

Regression equation of Y on X = $Y = 0.89X$, where $b_{yx} = 0.89$.

$$r = \sqrt{b_{yx} \times b_{xy}}$$

$$= \sqrt{0.89 \times 0.85}$$

$$= \sqrt{0.7565}$$

$$= 0.8697$$

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

$$0.85 = 0.8697 \times \frac{3}{\sigma_y}$$

$$0.85 = \frac{2.6091}{\sigma_y} \Rightarrow \sigma_y = \frac{2.6091}{0.85}$$

$$\sigma_y = 3.069$$

Kavitha

14 MARKS PROBLEMS:-

1) From the following data,
a) The two regression coefficients
b) The two regression equations
c) The co-efficient of correlation between the marks in statistics and mathematics
d) The most likely marks in mathematics when marks in statistics are 30.

Marks in stats	25	28	35	32	31
Marks in Maths	43	46	49	41	36
	36	29	38	34	32
	32	31	30	33	39

Soln ⇒ Actual mean method.

X	Y	x	x^2	y	y^2	xy
25	43	-7	49	5	25	-35
28	46	-4	16	8	64	-32
35	49	3	9	11	121	33
32	41	0	0	3	9	0
31	36	-1	1	-2	4	2
36	32	4	16	-6	36	-24
29	31	-3	9	-7	49	21
38	30	6	36	-8	64	-48
34	33	2	4	-5	25	-10
32	39	0	0	1	1	0
320	380		140		398	-93

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

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

(a) calculation of regression co-efficient :-

$$b_{yx} = \frac{\Sigma xy}{\Sigma x^2} = \frac{-93}{140} = -0.6642$$

$$b_{xy} = \frac{\Sigma xy}{\Sigma y^2} = \frac{-93}{398} = -0.2336$$

(b) calculation of regression equations:-

⇒ Regression equation of x on y :-

$$(x - \bar{x}) = b_{xy}(y - \bar{y})$$
$$(x - 32) = -0.2336 (y - 38)$$
$$x - 32 = -0.2336y + 8.8768$$
$$x = -0.2336y + 8.8768 + 32$$
$$x = -0.2336y + 40.8768$$

⇒ Regression equation y on x :-

$$(y - \bar{y}) = b_{yx}(x - \bar{x})$$
$$y - 38 = -0.6642 (x - 32)$$
$$y - 38 = -0.6642x + 21.2544$$
$$y = -0.6642x + 21.2544 + 38$$
$$y = -0.6642x + 59.2544$$

(c) Calculation of co-efficient correlation

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

$$= \frac{-93}{\sqrt{140 \times 398}}$$

$$= \frac{-93}{\sqrt{55720}}$$

$$= \frac{-93}{236.0508}$$

$$= -0.3939$$

OR

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

$$= \sqrt{-0.2336 \times -0.6642}$$

$$= \sqrt{0.1551}$$

$$= 0.39$$

∴ The regression co-efficient are negative hence correlation co-efficient should be negative = $r = -0.39$.

∴ Their exist moderate degree negative correlation

(d) $x = 30 \quad Y = ?$

$Y = -0.6642X + 59.2544$
$Y = -0.6642(30) + 59.2544$
$Y = -19.926 + 59.2544$
$Y = 39.3284$

Q. From the following data you are required to:-

(a) Form two regression equation
(b) Estimate the value of x when $Y = 10$ and the value of Y when $x = 15$
(c) Find correlation co-efficient through regression coefficients

Particulars	x	Y
No. of pairs of observations (N)	5	5
Mean ($\bar{x}$, $\bar{Y}$)	6	8
Sum of Squares of deviation ($\Sigma x^2, \Sigma y^2$)	40	20
Sum of the product of deviation	-26	

soln
(a) Calculation of regression equation
⇒ Regression equation of x on Y:-

$(x - \bar{x}) = bxy\,(Y - \bar{Y})$

$(x - 6) = \frac{-26}{20}(Y - 8)$

$x - 6 = -1.3\,(Y - 8)$

$x - 6 = -1.3y + 10.4$

$x = -1.3y + 10.4 + 6$

$x = -1.3y + 16.4$

⟹ Regression equation of y on x:-

$(y - \bar{y}) = byx(x - \bar{x})$

$(y - 8) = \frac{-26}{40}(x - 6)$

$y - 8 = -0.65(x - 6)$

$y - 8 = -0.65x + 3.9$

$y = -0.65x + 3.9 + 8$

$y = -0.65x + 11.9$

(b) when $y = 10$

$x = -1.3y + 16.4$

$x = -1.3(10) + 16.4$

$x = -13 + 16.4$

$x = 3.4$

when $x = 15$

$y = -0.65x + 11.9$

$y = -0.65(15) + 11.9$

$y = -9.75 + 11.9$

$y = 2.15$

(c) Calculation of correlation co-efficient through regression co-efficients

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

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

$$= \sqrt{-1.3 \times -0.65}$$

$$= \sqrt{0.845}$$

$$= 0.9192$$

36 The data about the sales and advertisement expenditure of a firm are given below:-

Particulars	Sales (in crore (₹))	Advertisement Exp's (in crore ₹)
Mean $(\bar{x}, \bar{y})$	40	6
Standard deviation (σ_x, σ_y)	10	1.5
Co-efficient of correlation (r)	0.9	

(i) Estimate the likely sales for proposed advertisement expenses of ₹10 crores.

(ii) What should be the advertisement expenses if the firm proposes a sales target of ₹60 crores?

Solut Regression co-efficient

$$b_{xy} = r\frac{\sigma_x}{\sigma_y} = 0.9 \times \frac{10}{1.5} = 6$$

$$b_{yx} = r\frac{\sigma_y}{\sigma_x} = 0.9 \times \frac{1.5}{10} = 0.135$$

→ Regression equation of x on y:-

$$(x-\bar{x}) = b_{xy}(y-\bar{y})$$
$$(x-40) = 6(y-6)$$
$$x-40 = 6y-36$$
$$x = 6y-36+40$$
$$x = 6y+4$$

when $y = 10$

$$x = 6(10)+4$$
$$x = 60+4$$
$$x = 64$$

b) Regression equation y on x:-

$$(y-\bar{y}) = b_{yx}(x-\bar{x})$$
$$(y-6) = 0.135(x-40)$$
$$y-6 = 0.135x-5.4$$
$$y = 0.135x-5.4+6$$
$$y = 0.135x+0.6$$

when $x = 60$

$$y = 0.135(60)+0.6$$
$$y = 8.1+0.6$$
$$y = 8.7$$

Q) The following results of capital employed and profit earned by a firm in 10 successive years are calculated.

Particulars	Mean	SD
capital employed (₹ in '000)	55	28.7
profit earned (₹ in '000)	13	8.5
co-efficient of correlation	0.96	

(i) obtain the two regression equations.
(ii) Estimate the amount of profit to be earned if capital employed is ₹ 50000/-.
(iii) Estimate the amount of capital to be employed if profit earned is ₹20000/-.

solut → Regression co-efficient

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

$$= 0.96 \times \frac{8.5}{28.7}$$

$$= 0.2843$$

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

$$= 0.96 \times \frac{28.7}{8.5}$$

$$= 3.2414$$

(a) calculation of regression equation

⟹ Regression equation y on x

$$(y-\bar{y}) = b_{yx}(x-\bar{x})$$
$$(y-13) = 0.2843(x-55)$$
$$y-13 = 0.2843x - 15.6365$$
$$y = 0.2843x - 15.6365 + 13$$
$$y = 0.2843x - 2.6365$$

⟹ Regression equation x on y

$$(x-\bar{x}) = b_{xy}(y-\bar{y})$$
$$(x-55) = 3.2414(y-13)$$
$$(x-55) = 3.2414y - 42.1382$$
$$x = 3.2414y - 42.1382 + 55$$
$$x = 3.2414y + 12.8618$$

(b) when $x = 50$

$$y = 0.2843x - 2.6365$$
$$y = 0.2843(50) - 2.6365$$
$$y = 14.215 - 2.6365$$
$$y = 11.5785$$

(c) when $y = 20$

$x = 3.2414y + 12.8618$

$x = 3.2414(20) + 12.8618$

$x = 64.828 + 12.8618$

$x = 77.6898$

∴ Amount of profit is 4.5785.
Amount of capital employed is 77.6898.

8) To study the relationship between expenditure on accommodation (x) and expenditure on food and entertainment on food and entertainment (y) an enquiry into 50 families gave the following results.

$\Sigma X = 8500$, $\Sigma Y = 9600$ $\sigma_x = 60$, $\sigma_y = 20$
$r = 0.6$.

Soln: Calculation of Mean

$$\bar{X} = \frac{\Sigma X}{N} = \frac{8500}{50} = 170$$

$$\bar{Y} = \frac{\Sigma Y}{N} = \frac{9600}{50} = 192$$

Calculation of regression co-efficients

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

$$= 0.6 \times \frac{60}{20}$$

$$= 1.8$$

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

$$= 0.6 \times \frac{20}{60}$$

$$= 0.2$$

Calculation of regression equation.

Regression equation of Y on X :-

$$(Y - \bar{Y}) = b_{yx}(X - \bar{X})$$
$$(Y - 192) = 0.2(X - 170)$$
$$Y - 192 = 0.2X - 34$$
$$Y = 0.2X - 34 + 192$$
$$Y = 0.2X + 158$$

Regression equation of x on y:-

$$(x - \bar{x}) = bxy(y - \bar{y})$$
$$x - 170 = 1.8(y - 192)$$
$$x - 170 = 1.8y - 345.6$$
$$x = 1.8y - 345.6 + 170$$
$$x = 1.8y - 175.6$$

All the Best

Scanned with OKEN Scanner

ASSIGNMENT PROBLEMS:-

1. You are given

Particulars	X	Y
Mean $(\bar{x}, \bar{y})$	18	100
Standard deviation $(\sigma x, \sigma y)$	14	20
co-efficient correlation (r)	0.8	

Findout most probable value of Y if X is 70 and that of X if Y is 90.

Solu:- Regression co-efficient

$$bxy = r\frac{\sigma x}{\sigma y} = 0.8 \times \frac{14}{20}$$

$$= 0.56$$

$$byx = r\frac{\sigma y}{\sigma x} = 0.8 \times \frac{20}{14}$$

$$= 1.1428$$

Calculation of Regression equation.

⇒ Regresssion equation of Y on X:-

$$(Y-\bar{Y}) = byx\,(X-\bar{X})$$

$$(Y-100) = 1.1428\,(X-18)$$

$$Y-100 = 1.1428X - 20.5704$$

$$Y = 1.1428X - 20.5704 + 100$$

$$\boxed{Y = 1.1428X + 79.4296}$$

⇒ Regression equation of X on Y:-

$$(X-\bar{X}) = bxy\,(Y-\bar{Y})$$

$$(X-18) = 0.56(Y-100)$$

$$X-18 = 0.56Y - 56$$

$x = 0.56y - 56 + 18$

$\boxed{x = 0.56y - 38}$

$\Rightarrow$ When $x = 70$

$y = 1.1428x + 79.4296$

$y = 1.1428(70) + 79.4296$

$y = 79.996 + 79.4296$

$\boxed{y = 159.4256}$

When $y = 90$

$x = 0.56y - 38$

$x = 0.56(90) - 38$

$x = 50.4 - 38$

$\boxed{x = 12.4}$

$\Rightarrow$ You are given

Particulars	x	y
Mean $(\bar{x}, \bar{y})$	18	100
variance (σ_x^2, σ_y^2)	14	20
co-efficient of correlation (r)	0.8	

compute the regression line x on y and calculate x when y is 88

Soln:- variance :- $\sigma_x = \sqrt{14} = 3.7416$

(SD) $\sigma_y = \sqrt{20} = 4.4721$

calculation of regression co-efficient

$$b_{yx} = r\frac{\sigma_y}{\sigma_x} = 0.8 \times \frac{4.4721}{3.7416}$$

$$= 0.9561$$

$$b_{xy} = r\frac{\sigma_x}{\sigma_y} = \frac{0.8 \times 3.7416}{4.4721}$$

$$= 0.6693$$

calculation of regression equations.

⇒ Regression equation of x on y :-

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

$$(x - 18) = 0.6693(y - 100)$$

$$x - 18 = 0.6693y - 66.93$$

$$x = 0.6693y - 66.93 + 18$$

$$\boxed{x = 0.6693y - 48.93}$$

⇒ Regression equation of y on x :-

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

$$(y - 100) = 0.9561(x - 18)$$

$$y - 100 = 0.9561x - 17.2098$$

$$y = 0.9561x - 17.2098 + 100$$

$$\boxed{y = 0.9561x + 82.7902}$$

Prediction:-

when $y = 88$

$$x = 0.6693y - 48.93$$

$$x = 0.6693(88) - 48.93$$

$$x = 58.8984 - 48.93$$

$$\boxed{x = 9.9684}$$

Scanned with OKEN Scanner

3) The following data relate to marks obtained by 250 student in Accountancy and Statistics in B.Com Examination of BCU.

Subject	Mean	Variance
Accountancy (x)	48	16
Statistics (y)	55	25
co-efficient of correlation	0.8	

(i) Find two regression equations
(ii) Estimate the marks obtained by a student in statistics who secured 50 marks in accountancy.
(iii) Estimate the marks obtained in accountancy when marks in statistics are 65.

Solu:- $SD = \sigma x = \sqrt{16} = 4$

$\sigma y = \sqrt{25} = 5$

calculation of regression co-efficient.

$$bxy = r \frac{\sigma x}{\sigma y} = 0.8 \times \frac{4}{5} = 0.64$$

$$byx = r \frac{\sigma y}{\sigma x} = 0.8 \times \frac{5}{4} = 1$$

(i) calculation of regression equations
⇒ Regression equation of Y on X :-

$(Y - \bar{Y}) = byx (X - \bar{X})$
$(Y - 55) = 1(X - 48)$
$Y - 55 = X - 48$

$$y = x - 48 + 55$$
$$\boxed{y = x + 7}$$

=> Regression equation of x on y:-

$$(x - \bar{x}) = b_{xy}(y - \bar{y})$$
$$(x - 48) = 0.64(y - 55)$$
$$x - 48 = 0.64y - 35.2$$
$$x = 0.64y - 35.2 + 48$$
$$\boxed{x = 0.64y + 12.8}$$

(ii) when $x = 50$

$$y = x + 7$$
$$y = 50 + 7$$
$$\boxed{y = 57}$$

(iii) when $y = 65$

$$x = 0.64y + 12.8$$
$$x = 0.64(65) + 12.8$$
$$x = 41.6 + 12.8$$
$$\boxed{x = 54.4}$$