REGRESSION

What is Regression?

- ❖ Let us take the example of Investment and return
- ❖ Test Score and CA foundation Clear

X Y X

Y is dependendent on X

Y is dependendent on X

Mathematical relation → equation

A mathematical equation that allows us to predict value of one variable from known values of one or more variables is called a regression equation.

X on Y → X is dependendent on Y

WHAT IS REGRESSION?

X	Y
2	5
5	8
8	12
10	14

$y = 2 + 3x$

Simple Regression

only two variable

Linear Equation

y on x: $y = a + bx$

x on y: $x = c + dy$

DEPENDENT VARIABLE

The variable whose value is to be predicted is called the dependent variable or explained variable.

INDEPENDENT VARIABLE

The variables which are used to predict the values of a dependent variable are called independent variables or explanatory variables

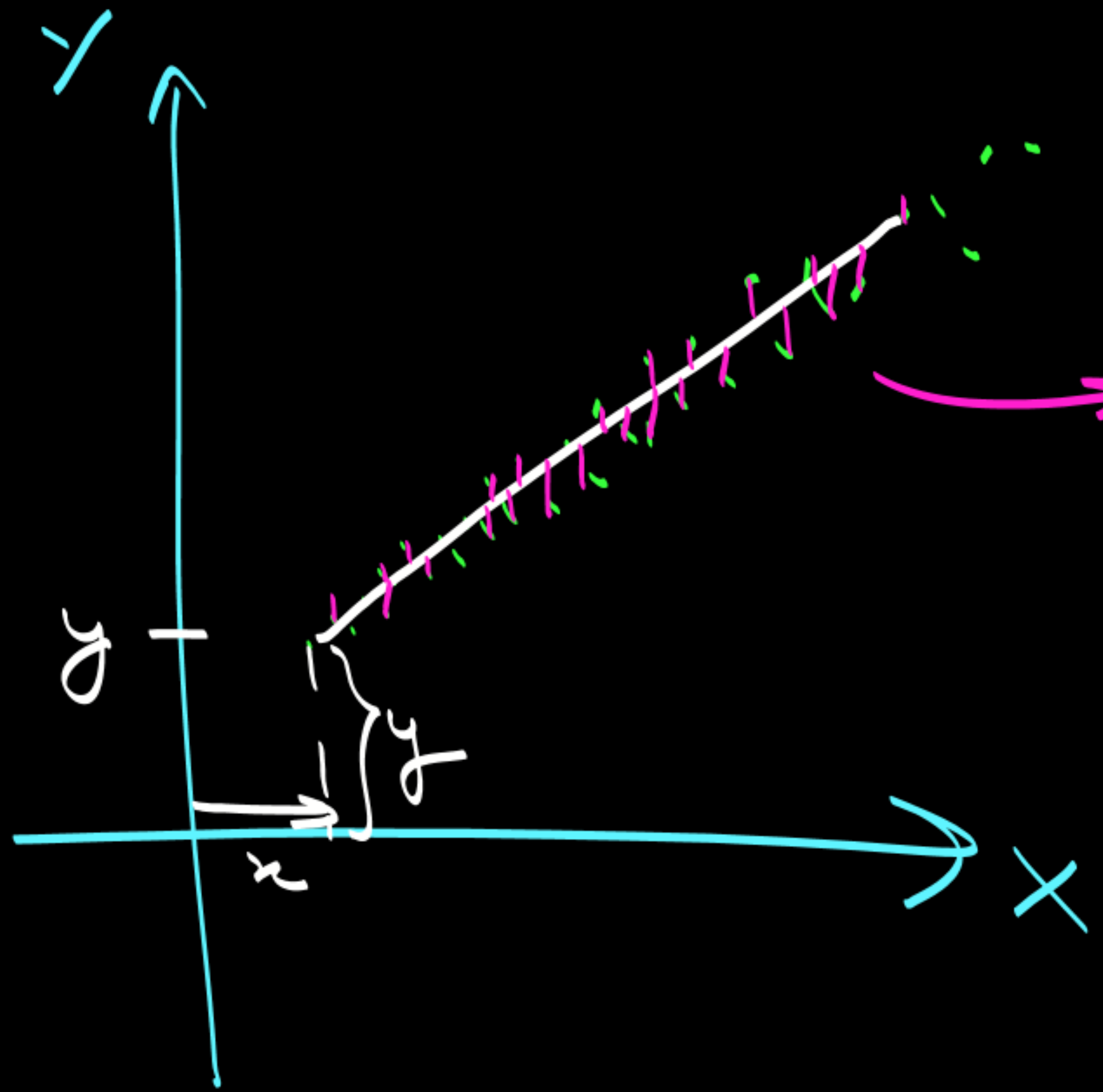
SIMPLE REGRESSION ANALYSIS AND SIMPLE LINEAR ANALYSIS

The regression analysis confined to the study of only two variables, a dependent variable and an independent variable, is called simple regression analysis.

When the relationship between the dependent variable and the independent variable is linear, the technique for prediction is called simple linear regression.

If let say y depends on x, then equation will be :-

Method of Least Square



$\sum (y_i - y) \rightarrow$ sum min

$y = a + bx$

dependent

independent

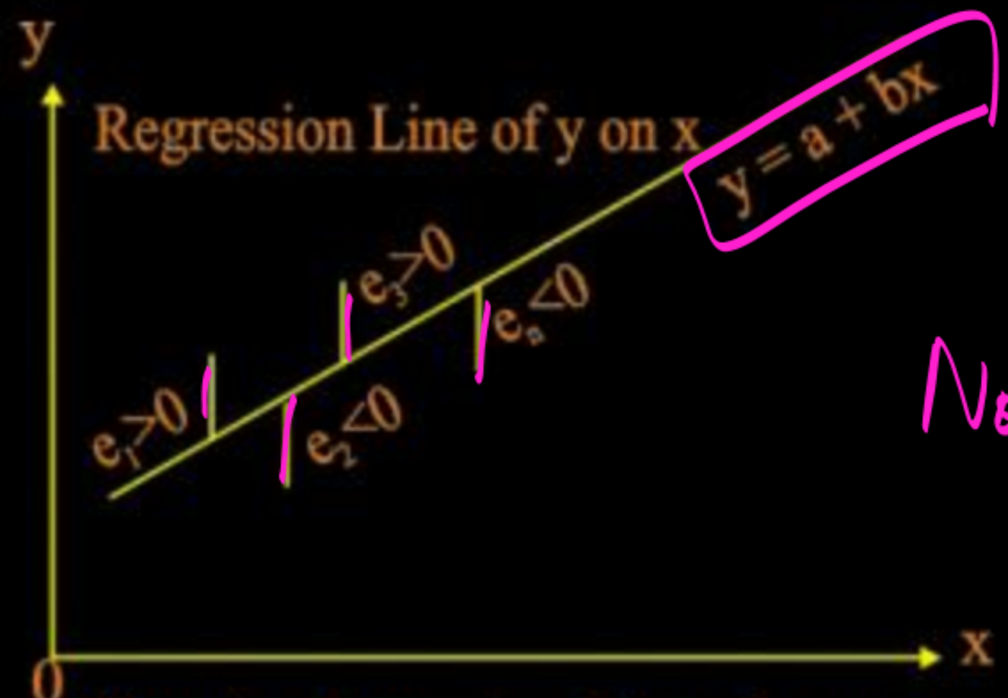
given

METHOD OF LEAST SQUARE – REGRESSION

If a line of best fit approximating the given data has the equation then the method of least squares requires that we must determine constants a and b so as to minimize

$n \to$ no. of pairs (x, y)

X	Y
2	3
4	5
8	6



Showing Regression Line of y on x

$\sum |y_i - y| \to$ minimum

Normal eqn $\to$ $\sum y = n.a + b\sum x$

$\sum xy = a\sum x + b\sum x^2$

❖ These equations known as the normal equation for estimating a and b, are given by

METHOD OF LEAST SQUARE – REGRESSION

y on x $\quad y = a + bx$

- Solving equation simultaneously for a and b we obtain

$$\sum x \left(\sum y = an + b\sum x \right) - (1)$$

$$n \left(\sum xy = a\sum x + b\sum x^2 \right) - (2)$$

$$\Rightarrow \quad b = \frac{n\sum xy - \sum x \sum y}{n\sum x^2 - (\sum x)^2} = \frac{Cov(x,y)}{\sigma_x^2}$$

$(2) - (1)$

In eq 1

$$\frac{\sum y}{n} = \frac{a \cdot \cancel{n}}{\cancel{n}} + \frac{b\sum x}{n}$$

$$\Rightarrow \bar{y} = a + b\bar{x}$$

$$\Rightarrow a = \bar{y} - b\bar{x}$$

$$r = \frac{Cov(x,y)}{\sigma_x \cdot \sigma_y}$$

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

* y on x $\Rightarrow$ $y = a_{yx} + b_{yx}x$ — (1)

constant of y on x

coeff of y on x

$$\Rightarrow \boxed{b_{yx} = \frac{n\Sigma xy - \Sigma x \cdot \Sigma y}{n\Sigma x^2 - (\Sigma x)^2} = \frac{Cov(x,y)}{\sigma_x^2}}$$

$$\Rightarrow \boxed{a_{yx} = \bar{y} - b_{yx}\bar{x}} \text{ — (2)}$$

$\Rightarrow$ Put a_{yx} from (2) in (1)

$$\Rightarrow y = \bar{y} - b_{yx}\bar{x} + b_{yx}x$$

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

* x on y :-

$x = c + d\,y$

$c \to a_{xy}$, $d \to b_{xy}$

$\Rightarrow x = a_{xy} + b_{xy}\,y$

$$\Rightarrow \boxed{b_{xy} = \frac{n\Sigma xy - \Sigma x \Sigma y}{n\Sigma y^2 - (\Sigma y)^2} = \frac{Cov(x,y)}{\sigma_y^2}}$$

$$\Rightarrow \boxed{a_{xy} = \bar{x} - b_{xy}\bar{y}}$$

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

QUESTION

regression eq^n → y on x
&

Calculate the regression coefficients from the following information:

$\sum X = 50$, $\sum Y = 30$, $\sum XY = 1000$, $\sum X^2 = 3000$, $\sum Y^2 = 1800$, n=10

y on x:

$$\bar{y} = \frac{\Sigma y}{n} = \frac{30}{10} = 3, \quad \bar{x} = \frac{50}{10} = 5$$

$$b_{yx} = \frac{n\Sigma xy - \Sigma x \Sigma y}{n\Sigma x^2 - (\Sigma x)^2} = \frac{10(1000) - 50 \times 30}{10(3000) - (50)^2} = 0.309$$

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

$$\Rightarrow y - 3 = 0.309(x - 5)$$

QUESTION

x on y

Calculate the regression coefficients from the following information:

$\sum X = 50$, $\sum Y = 30$, $\sum XY = 1000$, $\sum X^2 = 3000$, $\sum Y^2 = 1800$, n=10

$\bar{x} = 5$, $\bar{y} = 3$

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

where $b_{xy} = \dfrac{n\sum xy - \sum x \sum y}{n\sum y^2 - (\sum y)^2} = 0.497$

$x - 5 = 0.497(y - 3)$

QUESTION

Following table gives the age of cars of a certain make and annual maintenance costs. Obtain the regression equation for costs related to age:

Age of cars in years: x	2	4	6	8
Maintenance cost (in hundred) y	10	20	25	30

Y on X

$$b_{yx} = \frac{n\Sigma XY - \Sigma X \Sigma Y}{n\Sigma X^2 - (\Sigma X)^2}$$

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

$n = 4$

X	Y	XY	X^2
2	10	$2\times10=$	$2\times=$
4	20	$4\times20=$	$4\times=$
6	25	$6\times25=$	$6\times=$
8	30	$8\times30=$	$8\times=$
$\Sigma X=20$	$\Sigma Y=85$	$\Sigma XY=490$	$\Sigma X^2=120$

GT GT

$\bar{x} = \frac{20}{4} = 5$, $\bar{y} = \frac{85}{4} = 21.25$

$$b_{yx} = \frac{4\times490 - 20\times85}{4\times120 - (20)^2}$$

$$b_{yx} = \frac{1960 - 1700}{80} = 3.25$$

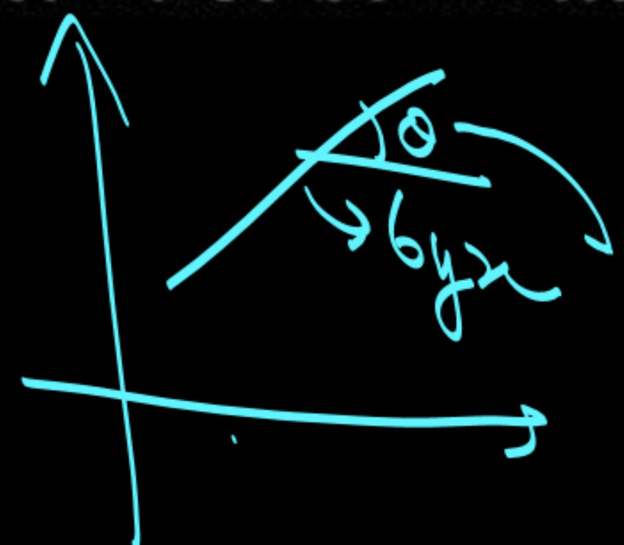
$$\Rightarrow y - 21.25 = 3.25(x - 5)$$

METHOD OF LEAST SQUARE – REGRESSION

- The constant b is called the regression coefficient of Y on X is denoted by b_{yx}.
- It measures the change in Y corresponding to a unit change in X.
- Thus, b_{yx} represent the slope of the line of regression of Y on X
- The equation of the line of regression of Y on X can also be written as

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

Slope



METHOD OF LEAST SQUARE – REGRESSION

- On the other hand, if we wish to estimate a value of X for a given value of Y, we have to obtain regression line of X on Y:

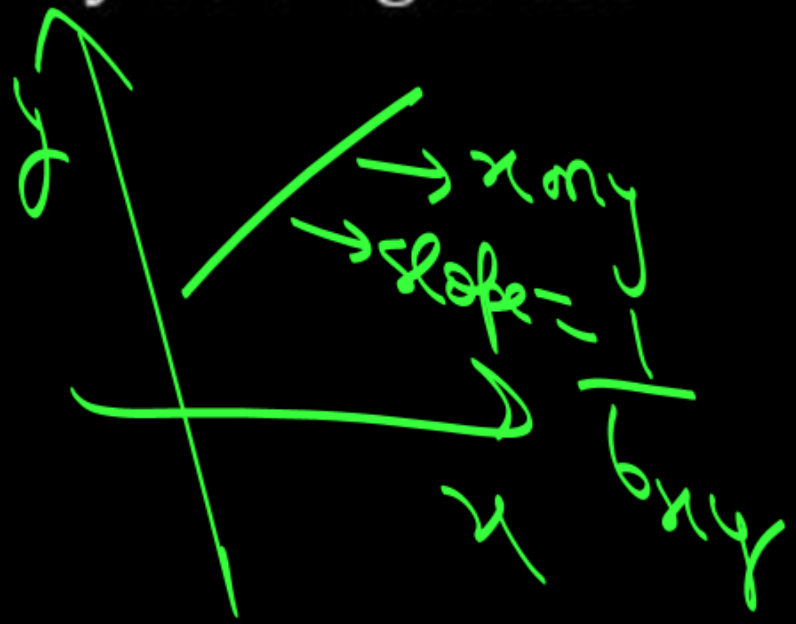
$$x = a_{xy} + b_{xy}\, y$$

$$b_{xy} = \frac{n\sum xy - \sum x \sum y}{n\sum y^2 - (\sum y)^2} \quad \& \quad a_{yx} = \bar{x} - b_{xy}\bar{y}$$

- The two normal equation for estimating c and d are given by Solving these normal equations simultaneously for c and d, we obtain

$$\boxed{x - \bar{x} = b_{xy}(y - \bar{y})} \rightarrow \text{slope} = \frac{1}{b_{xy}}$$

$$\frac{1}{b_{xy}}(x - \bar{x}) = y - \bar{y} \Rightarrow y = \frac{1}{b_{xy}}(x - \bar{x}) + \bar{y}$$

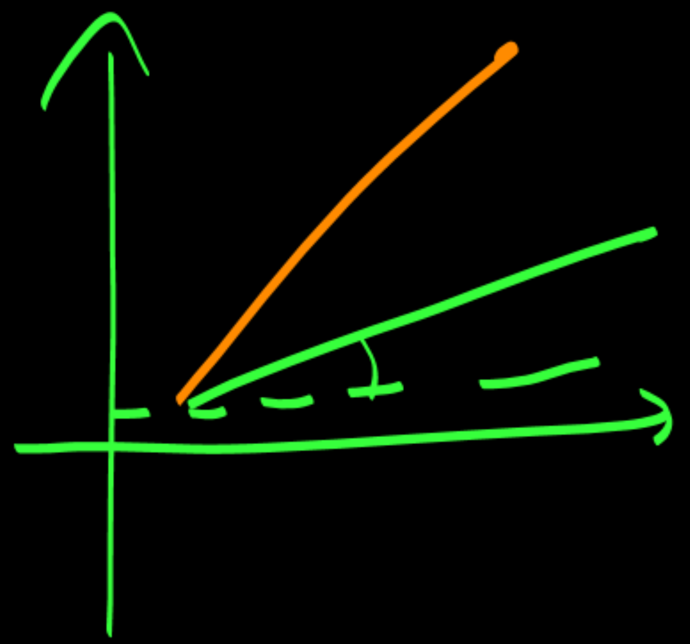


METHOD OF LEAST SQUARE – REGRESSION

- The constant d is called the regression coefficient of X on Y and is denoted by $b_{xy.}$ it measure the change in X corresponding to a unit change in Y.

- Clearly ,$1/b_{xy}$ represent the slope of the regression line of X on Y.

$y = a + bx$



Slope

y on $x \Rightarrow y = a_{yx} + b_{yx} x$

x on $y \Rightarrow x = a_{xy} + b_{xy} y$

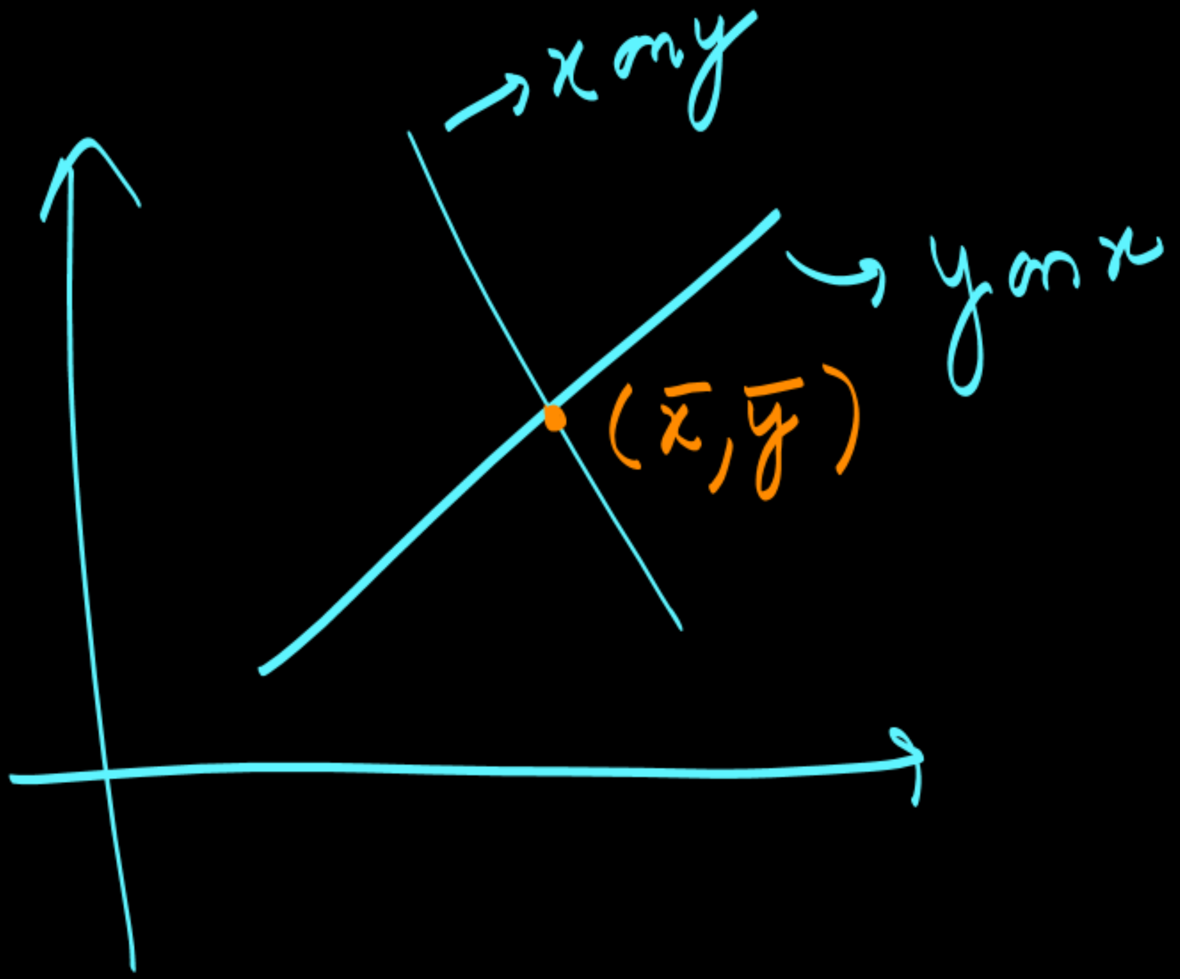
$\Rightarrow b_{xy} y = x - a_{xy}$

$\Rightarrow y = \frac{1}{b_{xy}}(x - a_{xy})$

Slope

METHOD OF LEAST SQUARE – REGRESSION

Line of regression of X and Y passes the point $(\bar{X}, \bar{Y})$ and hence the equation of the line of regression of X on Y can be written as



QUESTION

In the estimation of regression equation of two variables X and Y, the following results were obtained:

$\sum X = 900, \sum Y = 700, \sum X^2 = 6360, \sum Y^2 = 2860 \quad \sum XY = 3900, n = 10$

Obtain two regression equations

$\bar{x} = \frac{900}{10} = 90$

$\bar{y} = \frac{700}{10} = 70$

y on x

$b_{yx} = \frac{10(3900) - 900 \times 700}{10(6360) - (900)^2} = 0.791$

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

$y - 70 = 0.791(x - 90)$

x on y

$b_{xy} = \frac{10(3900) - 900 \times 700}{10(2860) - (700)^2} = 1.288$

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

$x - 90 = 1.288(y - 70)$

QUESTION

Given $\bar{x} = 50$, $\bar{y} = 20$, $\sigma_x = 20$, $\sigma_y = 30$ and cov (X, Y) = -100 find:

1. Correlation coefficient
2. Both the regression coefficients

1) $r = \dfrac{Cov(X,Y)}{\sigma_x \cdot \sigma_y} = \dfrac{-100}{20 \times 30} = \dfrac{-1}{6}$

2) $b_{yx} = \dfrac{Cov(x,y)}{\sigma_x^2} = \dfrac{-100}{(20)^2} = \dfrac{-100}{400} = -\dfrac{1}{4}$

3) $b_{xy} = \dfrac{Cov(x,y)}{\sigma_y^2} = \dfrac{-100}{(30)^2} = \dfrac{-100}{900} = \dfrac{-1}{9}$

* $b_{yx} = \dfrac{Cov(X,Y)}{\sigma_x \cdot \sigma_x} \times \dfrac{\sigma_y}{\sigma_y} = r\,\dfrac{\sigma_y}{\sigma_x}$

* $b_{xy} = \dfrac{Cov(x,y)}{\sigma_y \cdot \sigma_y} \cdot \dfrac{\sigma_x}{\sigma_x} = r\,\dfrac{\sigma_x}{\sigma_y}$

* $b_{yx} \cdot b_{xy} = r\dfrac{\cancel{\sigma_y}}{\cancel{\sigma_x}} \cdot r\dfrac{\cancel{\sigma_x}}{\cancel{\sigma_y}}$

$\Rightarrow \boxed{b_{yx} \cdot b_{xy} = r^2}$

$\therefore g \Rightarrow -\dfrac{1}{4} \times -\dfrac{1}{9} = \left(-\dfrac{1}{6}\right)^2 \Rightarrow \dfrac{1}{36} = \dfrac{1}{36}$

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

$$y - \bar{y} = r\frac{\sigma_y}{\sigma_x}(x - \bar{x})$$

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

$$x - \bar{x} = r\frac{\sigma_x}{\sigma_y}(y - \bar{y})$$

QUESTION

(15 min) Ques + Break

The following data relate to the mean and SD of the prices of two shares in a stock Exchange:

Share :	Mean (in Rs.)	SD (in Rs.)
x Company A	$\bar{x}$ 44	σ_x 5.6
y Company B	$\bar{y}$ 58	σ_y 6.3

Coefficient of correlation between the share prices = 0.48 $= r$

Find the most likely price of share A corresponding to a price of Rs.60 of share B and also the most likely price of share B for a price of Rs. 50 of share A.

1) $x = ?$ $y = 60$

x on $y \Rightarrow x - \bar{x} = r \frac{\sigma_x}{\sigma_y}(y - \bar{y})$

$\Rightarrow x - 44 = 0.48 \left(\frac{5.6}{6.3}\right)(60 - 58) \Rightarrow x = 44.85$

2) $x = 50$ $y = ?$

y on $x \Rightarrow y - \bar{y} = r \frac{\sigma_y}{\sigma_x}(x - \bar{x})$

$y - 58 = 0.48 \times \frac{6.3}{5.6}(50 - 44) \Rightarrow y = 61.24$

PROPERTIES OF REGRESSION COEFFICIENTS

Property 1. The coefficients of correlation and two regression coefficients has the same signs.

b_{yx}, b_{xy} & r

b_{yx}	b_{xy}	r
+ve	+ve	+ve
−ve	−ve	−ve

$-1 \quad r \quad 1$

Property 2. The coefficients of correlation are the geometric mean between the regression coefficients.

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

$$\Rightarrow 0 \le b_{yx} \times b_{xy} \le 1$$

$-1 \le r \le 1 \rightarrow r^2 \rightarrow (-1 \text{ to } 0)^2$

$r \rightarrow 0$

$\Rightarrow 0 \le r^2 \le 1 \qquad r \rightarrow 0 \text{ to } 1$

QUESTION

If r = -0.6 and one regression coefficient is-0.8, find the other regression coefficient.

b_{xy}

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

$$\Rightarrow -0.8 \times b_{yx} = (-0.6)^2$$

$$\Rightarrow b_{yx} = \frac{0.36}{-0.8} = -0.45$$

PROPERTIES OF REGRESSION COEFFICIENTS

Property 3. If one of the regression coefficients is greater than unity, the other must be less than the unity. < 1

$b_{xy} > 1$

$b_{yx} \cdot b_{xy} \leq 1$

$\Rightarrow b_{yx} \leq \frac{1}{b_{xy}}$

Let's say

$b_{xy} = 2$

$b_{yx} \times 2 \leq 1$

$\rightarrow b_{yx} \leq \frac{1}{2}$

PROPERTIES OF REGRESSION COEFFICIENTS

Property 4. The two lines of regression intersect at the point $(\bar{X}, \bar{Y})$ where x and y are the variables under consideration.

Property 5. The regression coefficients are independent of change of origin but not for scale.

$$\left.\begin{array}{l} u = a + bx \\ v = c + dy \end{array}\right\} \quad \begin{array}{l} b_{vu} = \frac{d}{b} b_{yx} \\ b_{uv} = \frac{b}{d} b_{xy} \end{array}$$

QUESTION

The regression coefficients of regression equation of X on Y is 2.4 and the same for regression equation of Y on X is 0.8 are the regression coefficients consistent?

$$b_{yx} \cdot b_{xy} \leq 1$$

$$\Rightarrow 2.4 \times 0.8 = 1.92 \nleq 1$$

as product of reg. coeff is not less than 1.

They are not consistent.

QUESTION

If the relationship between two variables x and u is u + 3x = 10 and between two other variables y and v is 2y + 5v = 25, and the regression coefficient of y on x is known as 0.80, what would be the regression coefficient of v on u?

$$u+3x=10 \longrightarrow u=10-3x$$

$$5v+2y=25 \Rightarrow v=\frac{25}{5}-\frac{2}{5}y$$

$$b_{yx}=0.8$$

$$b_{vu}=\frac{\left(-\frac{2}{5}\right)}{-3}\,b_{yx}=+\frac{2}{5}\times\frac{1}{+3}\times 0.8=\underline{0.10667}$$

QUESTION

For some bivariate data, the following results were obtained:
Mean value of variable X = 53.2 and Y =39.5 regression coefficient of Y on X = -1.5 and of X on Y = -0.38 what should be the most likely value of X when Y = 50? Also find the coefficients of correlation between two variables.

$\bar{X} = 53.2$

$\bar{Y} = 39.5$

$b_{yx} = -1.5$

$b_{xy} = -0.38$

$X = ?$
$Y = 50$ → X on Y → $x - \bar{x} = b_{xy}(y - \bar{y})$

$\rightarrow x - 53.2 = -0.38(50 - 39.5)$

$\rightarrow x = 53.2 - 3.99$

$\rightarrow x = 49.21$

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

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

$= -\sqrt{-1.5 \times -0.38} = -0.755$

QUESTION

x on $y \rightarrow$ $\Sigma x = na + b\Sigma y$

$\Sigma xy = a\Sigma y + b\Sigma y^2$

Following are the two normal equations obtained for deriving the regression line of y and x:

Q16, Set (B), (ICAI)

$5a + 10b = 40$

$10a + 25b = 95$

y on x $\begin{cases} \Sigma y = na + b\Sigma x \\ \Sigma xy = a\Sigma x + b\Sigma x^2 \end{cases}$

The regression line of y on x is given by

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

(A) $2x + \cancel{3}y = 5$

(B) $2y + 3x = 5$

(C) $y = 2 + 3x$

(D) $y = 3 \cancel{+} 5x$

$n = 5$, $\Sigma x = 10$, $\Sigma y = 40$, $\Sigma xy = 95$, $\Sigma x^2 = 25$

$\bar{y} = \frac{40}{5} = 8$, $\bar{x} = \frac{10}{5} = 2$

$b_{yx} = \frac{n\Sigma xy - \Sigma x \Sigma y}{n\Sigma x^2 - (\Sigma x)^2} = \frac{5 \times 95 - 10 \times 40}{5 \times 25 - 10^2} = \frac{75}{25} = 3$

y on $x \rightarrow$ $y - 8 = 3(x - 2)$

$\Rightarrow y - 8 = 3x - 6$

$\Rightarrow y = 3x + 2$

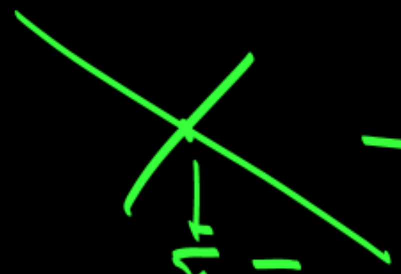
QUESTION

If the regression line of y on x and of x on y are given by $2x + 3y = -1$ and $5x + 6y = -1$ then the arithmetic means of x and y are given by

Q17, Set (B), (ICAI)

- **A** (1, –1)
- **B** (–1, 1)
- **C** (–1, –1)
- **D** (2, 3)

$\bar{x}, \bar{y}$



$\bar{x}, \bar{y}$ → Simultaneous Sol.

Simultaneous

$5 \times \quad 2\bar{x} + 3\bar{y} = -1 \Rightarrow 10\bar{x} + 15\bar{y} = -5$

$2 \times \quad 5\bar{x} + 6\bar{y} = -1 \Rightarrow 10\bar{x} + 12\bar{y} = -2$

$\quad - \quad - \quad +$

$3\bar{y} = -3$

$\Rightarrow \bar{y} = -1$

$\Rightarrow 2\bar{x} + 3(-1) = -1$

$\Rightarrow 2\bar{x} = -1 + 3$

$\Rightarrow 2\bar{x} = 2 \Rightarrow \bar{x} = 1$

QUESTION

Given the regression equations as $3x + y = 13$ and $2x + 5y = 20$, which one is the regression equation of y on x?

Q18, Set (B), (ICAI)

(A) 1^{st} equation

(B) 2^{nd} equation

(C) Both (a) and (b

(D) None of these

$y = a_{yx} + b_{yx}x$ $\quad$ $x = a_{xy} + b_{xy}y$

$3x+y=13 \mid 2x+5y=20$

Assume ~~y on x~~ x on y $\quad$ ~~x on y~~ y on x

$y = 13-3x$ $\quad$ $2x = \frac{20}{2} - \frac{5}{2}y$

$b_{yx} = -3$ $\quad$ $b_{xy} = -\frac{5}{2}$

Check $b_{yx} \times b_{xy} \le 1 \rightarrow -3 \times -\frac{5}{2} \not\le 1 \Rightarrow \frac{15}{2} \not\le 1$ ← Our assumption is wrong

$3x+1y=13 \mid 2x+5y=20$

$\frac{x}{y}$ $\quad$ $\frac{3}{1} > \frac{2}{5}$

x on y $\quad$ y on x

QUESTION

Given the following equations: $2x-3y = 10$ and $3x + 4y = 15$, which one is the regression x on y ?

Q19, Set (B), (ICAI)

$\frac{x}{y}$ $\quad \frac{2}{3} < \frac{3}{4}$

y on x $\quad$ x on y

A. 1^{st} equation

B. 2^{nd} equation

C. both the equations

D. none of these

QUESTION

If $u = 2x + 5$ and $v = -3y-6$ and regression coefficient of y on x is 2.4, what is the regression coefficient of v on u?

Q20, Set (B), (ICAI)

(A) 3.6

(B) –3.6

(C) 2.4

(D) –2.4

$u = 2x + 5$ $\quad$ $b_{yx} = 2.4$

$v = -3y - 6$

$b_{vu} = \frac{-3}{2} \times b_{yx} = \frac{-3}{2} \times 2.4 = \underline{-3.6}$

QUESTION

If 4y – 5x = 15 is the regression line of y on x and the coefficient of correlation between x and y is 0.75, what is the value of the regression coefficient of x on y?

Q21, Set (B), (ICAI)

(A) 0.45

(B) 0.9375

(C) 0.6

(D) None of these

y on x $\quad 4y - 5x = 15$

$$y = \frac{15}{4} + \frac{5}{4}x$$

$$b_{yx} = \frac{5}{4}, \quad r = 0.75$$

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

$$\Rightarrow (0.75)^2 = \frac{5}{4} \times b_{xy}$$

$$\Rightarrow b_{xy} = \frac{0.5625}{1.25} = 0.45$$

QUESTION

If the regression line of y on x and that of x on y are given by y = –2x + 3 and 8x = –y + 3 respectively, what is the coefficient of correlation between x and y?

Q22, Set (B), (ICAI)

(A) 0.5

(B) $-1/\sqrt{2}$

(C) –0.5

(D) None of these

$-\frac{y}{8} = -\frac{1}{8} \times y$

y on x

$\boxed{y = -2x + 3}$

x on y

$8x = -y + 3$

$\to x = -\frac{y}{8} + \frac{3}{8}$

x

y

$\frac{2}{1} < \frac{8}{1}$

y on x

$\to$ x on y

$b_{yx} = -2$

$b_{xy} = -\frac{1}{8}$

$r^2 = -2 \times -\frac{1}{8}$

$\Rightarrow r^2 = \frac{1}{4}$

$\Rightarrow r = \sqrt{\frac{1}{4}} = -\frac{1}{2} = -0.5$

QUESTION

If the regression coefficient of y on x, the coefficient of correlation between x and y and variance of y are $-3/4$, $\frac{\sqrt{3}}{2}$ and 4 respectively, what is the variance of x?

Q23, Set (B), (ICAI)

(A) $2/\sqrt{3/2}$

(B) $16/3$

(C) $4/3$

(D) 4

b_{yx} $\quad r \quad \sigma_y^2$

$b_{yx} = -\frac{3}{4}$, $r = \frac{\sqrt{3}}{2}$ $\qquad \sigma_y^2 = 4$, $\sigma_x^2 = ?$

$\sigma_y = 2$

$\Rightarrow r\frac{\sigma_y}{\sigma_x} = -\frac{3}{4}$

$\Rightarrow \frac{\sqrt{3}}{\cancel{2}} \times \frac{\cancel{2}}{\sigma_x} = -\frac{3}{4}$

$\Rightarrow (\sigma_x)^2 = \left(\sqrt{3} \times \frac{4}{-3}\right)^2 \rightarrow \sigma_x^2 = (\sqrt{3})^2 \times \frac{4^2}{3^2} = 3 \times \frac{16}{9_3} = \frac{16}{3}$

QUESTION

If $y = 3x + 4$ is the regression line of y on x and the arithmetic mean of x is -1, what is the arithmetic mean of y?

Q24, Set (B), (ICAI)

(A) 1

(B) –1

(C) 7

(D) None of these

$y = 3x + 4$

$\bar{x} = -1$

$\bar{y}$

$\rightarrow \bar{y} = 3\bar{x} + 4$

$= 3(-1) + 4$

$= 1$

QUESTION

If there are two variables x and y, then the number of regression equations could be

Q32, Set (A), (ICAI)

A. 1

B. 2

C. Any number

D. 3

QUESTION

Since Blood Pressure of a person depends on age, we need to consider

Q33, Set (A), (ICAI)

A. The regression equation of Blood Pressure on age

B. The regression equation of age on Blood Pressure

C. Both (a) and (b)

D. Either (a) or (b).

QUESTION

The method applied for deriving the regression equations is known as

Q34, Set (A), (ICAI)

(A) Least squares

(B) Concurrent deviation

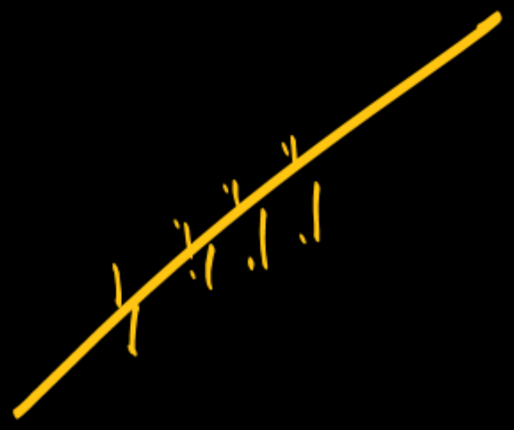
(C) Product moment

(D) Normal equation

QUESTION

The difference between the observed value and the estimated value in regression analysis is known as

Q35, Set (A), (ICAI)



(A) Error

(B) Residue

(C) Deviation

(D) (A) or (B).

QUESTION

The errors in case of regression equations are

Q36, Set (A), (ICAI)

A. Positive

B. Negative

C. Zero

D. All these

QUESTION

The regression line of y on x is derived by

Q37, Set (A), (ICAI)

A. The minmisation of vertical distance in the catter diagram

B. The minimisation of horizontal distances in the scatter diagram

x on y

C. Both (A) and (B)

D. (A) or (B).

QUESTION

The two lines of regression become identical when

Q38, Set (A), (ICAI)

A $r = 1$

B $r = -1$

C $r = 0$

D (A) or (B).

$r = 1$

$b_{yx} \times b_{xy} = 1$

$\Rightarrow \boxed{b_{yx} = \frac{1}{b_{xy}}}$

$r = -1$

$b_{yx} \times b_{xy} = (-1)^2 = 1$

$b_{yx} \times b_{xy} = 1$

$\Rightarrow b_{yx} = \frac{1}{b_{xy}}$

QUESTION

What are the limits of the two regression coefficients?

Q39, Set (A), (ICAI)

(A) No limit

(B) Must be positive

(C) One positive and the other negative

(D) Product of the regression coefficient must be numerically less than unity. ✓

$b_{yx} \times b_{xy} \leq 1$

QUESTION

The regression coefficients remain unchanged due to a

Q40, Set (A), (ICAI)

A. Shift of origin

B. Shift of scale

C. Both (a) and (b)

D. (a) or (b).

QUESTION

If the coefficient of correlation between two variables is -0.9, then the coefficient of determination is

Q41, Set (A), (ICAI)

(A) 0.9

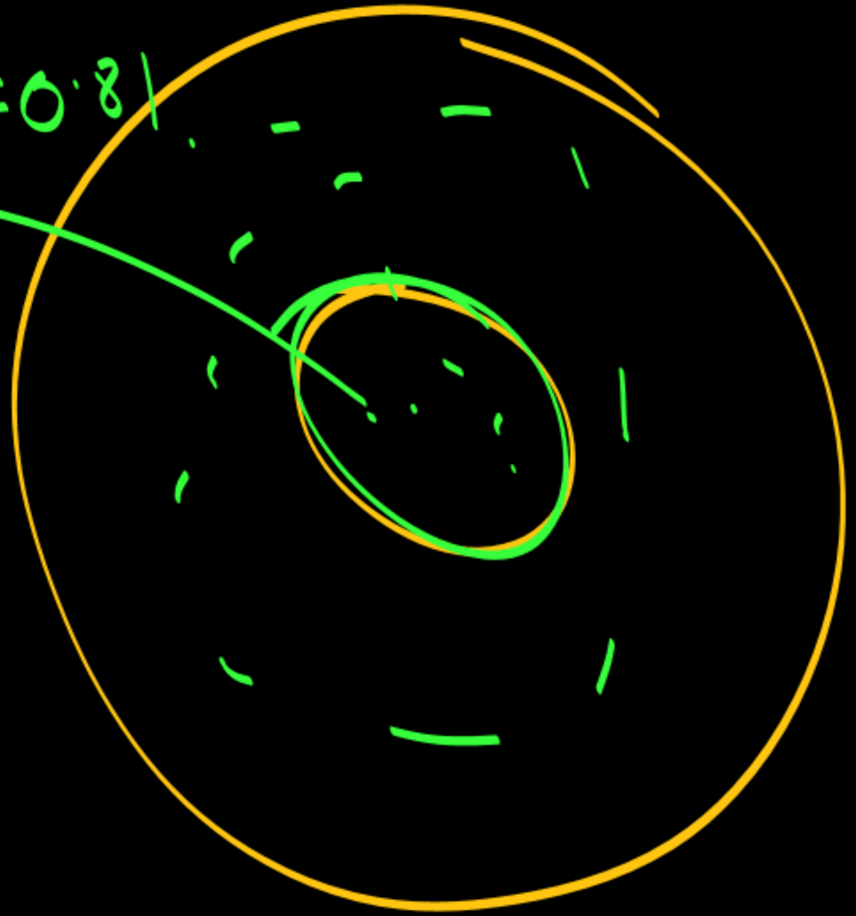
(B) 0.81

(C) 0.1

(D) 0.19.

Coeff of det. $= r^2 = (-0.9)^2 = 0.81$

Coeff of non-det. $= 1 - r^2$



QUESTION

If the coefficient of correlation between two variables is 0.7 then the percentage of variation unaccounted for is

Q42, Set (A), (ICAI)

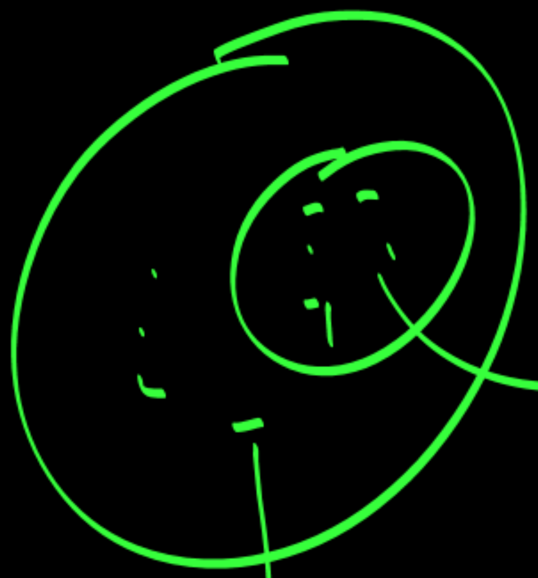
- **A** 70%
- **B** 30%
- **C** 51%
- **D** 49%

Coeff of nondet $= 1-(0.7)^2$

$= 1-0.49 = 0.51$

% area unacc. $= 0.51 \times 100\%$

$= 51\%$



QUESTION

Q4, Set (C), (ICAI)

The following results relate to bivariate data on (x, y):
$\sum xy = 414$, $\sum x = 120$, $\sum y = 90$, $\sum x^2 = 600$, $\sum y^2 = 300$ $n = 30$ Later or, it was known that two pairs of observations (12, 11) and (6, 8) were wrongly taken, the correct pairs of observations being (10, 9) and (8, 10). The corrected value of the correlation coefficient is

Record as odd

20min

(A) 0.752

(B) 0.768

(C) 0.846

(D) 0.953

QUESTION

rahulbhutani.ka@gmail.com

Coefficient of correlation between x and y for 20 items is 0.4. The AM's and SD's of x and y are known to be 12 and 15 and 3 and 4 respectively. Later on, it was found that the pair (20, 15) was wrongly taken as (15, 20). Find the correct value of the correlation coefficient.

$\rightarrow$ Sameer

Rec.

20min.

THANK YOU