Date: 30th March,2021

VIRTUAL COACHING CLASSES
ORGANISED BY BOS (ACADEMIC), ICAI

FOUNDATION LEVEL
PAPER 3: BUSINESS MATHEMATICS AND LOGICAL REASONING & STATISTICS

Faculty: CA Rashmi Lonikar, M.Sc., FCA, DISA

REGRESSION ANALYSIS

Regression analysis is a set of statistical processes for estimating the relationships between a dependent variable and one or more independent variables .

Regression analysis is a way of mathematically sorting out which of those variables does indeed have an impact.

- It answers the questions:
- Which factors matter most?
  - Which can we ignore?
    - How do those factors interact with each other?

You have your **dependent variable** — the main factor that you're trying to understand or predict. And then you have your **independent variables** — the factors you suspect have an impact on your dependent variable.

Helps model the **future** relationship between them

Suppose you're a sales manager trying to predict next month's numbers.Let's say you find out the average monthly rainfall for the past three years as well. Then you plot all of that information on a chart that looks like this:

- Equation of y on x   y = a + bx where

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

$$b = \frac{n\sum xy - (\sum x)(\sum y)}{n(\sum x^2) - (\sum x)^2}$$

$$a = \frac{\sum y - b\sum x}{n}$$

Regression analysis includes several variations, such as linear, multiple linear, and nonlinear. The most common models are simple linear and multiple linear.

Regression line of y on x is $y = a + bx$

A, b, are regression parameters

Here $b = b_{yx}$

$$\Sigma y_i = na + b\Sigma x_i$$

$$\Sigma x_i y_i = a\Sigma x_i + b\Sigma x_i^2$$

Solving there two equations

$$b = \frac{Cov(x,y)}{S_x^{\ 2}}$$

$$= \frac{r.S_x.S_y}{S_x^2}$$

Standard form and normal equations - regression equation

$$\Sigma y_i = na + b\Sigma x_i$$

$$\Sigma x_i y_i = a\Sigma x_i + b\Sigma x_i^2$$

Solving there two equations

$$b = \frac{Cov(x,y)}{S_x^2}$$

$$= \frac{r.S_x.S_y}{S_x^2}$$

$$\frac{(y-\bar{y})}{S_y} = \frac{r(x-\bar{x})}{S_x} \quad \ldots$$

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

Regression Equation of y on x:

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

Regression Equation of x on y:

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

$$b_{yx} = r\,(\sigma y / \sigma x)$$

$$b_{xy} = r\,(\sigma x / \sigma y)$$

$$b_{yx} = \frac{Cov(x,y)}{S_x^{\,2}} = \frac{r.S_y}{S_x} \ldots\ldots\ldots$$

$$b\hat{} = b_{xy} = \frac{cov(x,y)}{S_y^2} = \frac{r.S_x}{S_y} \ldots$$

Example 16.15: Find the two regression equations from the following data:

x:	2	4	5	5	8	10
y:	6	7	9	10	12	12

Hence estimate y when x is 13 and estimate also x when y is 15.

Find Summation x, summation y, summation xy, summation x^2 and summation $y^{2.}$

Find b and a you get equation

- .

$$\text{cov}(x, y) = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{n} = \frac{\sum x_i y_i}{n} - \bar{x}\,\bar{y}$$

$$S_x = \sqrt{\frac{\sum(x_i - \bar{x})^2}{n}} = \sqrt{\frac{\sum x_i^2}{n} - \bar{x}^2}$$

$$S_y = \sqrt{\frac{\sum(y_i - \bar{y})^2}{n}} = \sqrt{\frac{\sum y_i^2}{n} - \bar{y}^2}$$

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

$$b = \frac{\text{Cov}(x, y)}{S_x^{\,2}} = \frac{r.S_x.S_y}{S_x^2}$$

Q2COMPREHENSIVE:

The time x in years that an employee spent at a company and the employee’s hourly pay, y, for 5 employees are listed in the table below. Calculate and interpret the correlation coefficient r. Find the equation of the regression line. Use the equations to predict the hourly pay rate of an employee who has worked for 20 years.

X	**5**	**3**	**4**	**10**	**15**
Y	25	20	21	35	38

R = 0.97 Y = 16.11 + 1.58X, ANS 47.71

18.17/pg 18.32

Means of share prices of companies A and B are 44 and 58 and standard deviation is 5.60 and 6.30 respectively. The coefficient of correlation is 0.48. Find share price of A when share price of B is Rs 60 and share price of B and share price of A is Rs 50.

Hint : Use r and standard deviations to find b

Byx = 0.54, a = 34.24, y = 34.24+ 0.54x, when x = 50, y = 61.24

Bxy = 0.43 a = 19.25 x = 19.25+0.43y when y = 60, x = 44.85

Example 18.20: For the variables x and y, the regression equations are given as $7x - 3y - 18 = 0$ and $4x - y - 11 = 0$

(i) Find the arithmetic means of x and y.

(ii) Identify the regression equation of y on x.

(iii) Compute the correlation coefficient between x and y.

(iv) Given the variance of x is 9, find the SD of y.

Using the regression equations, find the means and coefficient of correlation r

2Y – X- 50 = 0

3Y – 2X -10 = 0

R2 = 0.75 mean x = 130 mean y = 90

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

for means solve simultaneous equations

PROPERTIES OF REGRESSION

(i) **The regression coefficients remain unchanged due to a shift of origin but change due to a shift of scale.**

This property states that if the original pair of variables is (x, y) and if they are changed to the pair (u, v) where

$$u = \frac{x-a}{p} \text{ and } v = \frac{y-c}{q}$$

$$b_{yx} = \frac{q}{p} \times b_{vu} \quad \text{.......................... (18.28)}$$

$$\text{and } bxy = \frac{p}{q} \times b_{uv} \quad \text{........................ (18.29)}$$

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

(iii) The coefficient of correlation between two variables x and y in the simple geometric mean of the two regression coefficients. The sign of the correlation coefficient would be the common sign of the two regression coefficients.

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

Example 18.19: 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 = X-A/P   V = Y-C/Q

P = 1/3 Q = -5/2 BUV =8/75

1. For bi-variate data, $\bar{x}$ = 53.2 $\bar{y}$ = 27.9b_{yx}= -1.5 and b_{xy}= -0.2. The most probable value of y when x = 60 is ……….

Ans: 17.7

2. The regression lines 3x + y = 13 and 2x + 5y = 20. Find regression equation line of y on x.

Ans: 2x + 5y = 20

3. Two random variables have the regression lines 3x + 2y = 26 and 6x + y = 31. The coefficient of correlation is ……

Ans: -0.5

4. The coefficient of correlation between x and y is -1/2. The value of b_{xy}= -1/8. Find byx

Ans: -2

5. Given $\bar{x}$ =16 $\bar{y}$ = 20s_x= 4.8 and s_y= 9.6, R = 0.6. The regression coefficient of x on y is…..

Ans: 0.3

6.The standard deviation of x and y are 5 and 8 respectively and coefficient of correlation is 0.8. The regression coefficient of y on x is…..

Ans: 1.28

---

7. If mean of x and y variables is 20 and 40 respectively. The regression coefficient of y on x is 1.608. The regression line of y on x is

(a) $y = 1.608x + 7.84$ (b) $y = 1.56x + 4.84$
© $y = 1.608x + 4.84$ (d) $y = 1.56x + 7.84$

Ans: a

8. If the mean of x and y variables is 3 and 1 respectively. The regression lines are

(a) $5x + 7y - 22 = 0$ & $6x + 2y - 20 = 0$ (b) $5x + 7y - 22 = 0$ & $6x + 2y + 20 = 0$

(c) $5x + 7y + 22 = 0$ & $6x + 2y - 20 = 0$ (d) $5x + 7y + 22 = 0$ & $6x + 2y + 20 = 0$

Ans: a

9. The regression line of x on y is 3x +2y = 100 and then the value of bxy is ….

Ans: -2/3

---

10. The regression line of y on x is given by y = x + 2 and the Karl Pearson's coefficient of correlation is 0.5 then find $\frac{Sy^2}{Sx^2}$

THANK YOU