5. Regression Analysis

Introduction

- Regression is the average linear relationship between two or more variables.
- The word regression implies “estimation of prediction”. In other words through regression equations we can quantify the relationship between two variables and we can predict the average value of one variable corresponding to a specific value of the other.
- It establishes a functional relationship between two variables.
- Regression equation enables us to find the nature and the extent of relationship between two variables. Correlation can measure only the degree of association between the two variables whereas regression quantifies such relationship.
- The two variables are dependent and independent variable. Thus, we try to estimate the average value of dependent variable, for a specified value of independent variable using regression analysis.
- If there are two variables, then the independent variable is called the “Regressor” or “Explaining Variable” and the dependent variable is called the “Regressed” or “Explained Variable”.
- Regression analysis is an **absolute measure** showing a change in the value of y or x for a corresponding change in the value of x or y whereas correlation coefficient is a **relative measure** of linear relationship between x and y.
- This average linear relationship between two variables is expressed by means of two straight line equation known as regression lines or regression equations.
- If there are two variables x and y we can have the following two types of regression lines,
  i. Regression equation of y on x (y dependent, x independent)
  ii. Regression equation of x on y (x dependent, y independent)

**REGRESSION LINES**

<table>
<thead>
<tr>
<th>Regression equation of y on x:</th>
<th>Regression equation of x on y:</th>
</tr>
</thead>
<tbody>
<tr>
<td>((Y - \bar{Y}) = b_{yx} (X - \bar{X}))</td>
<td>((X - \bar{X}) = b_{xy} (Y - \bar{Y}))</td>
</tr>
<tr>
<td>(b_{yx}) stands for regression coefficient of y on x</td>
<td>(b_{xy}) stands for regression coefficient of x on y</td>
</tr>
<tr>
<td>Here y depends on x</td>
<td>Here x depends on y</td>
</tr>
<tr>
<td>Here y is a dependent/explained and x is an independent variable</td>
<td>Here x is a dependent/explained and y is an independent variable</td>
</tr>
<tr>
<td>This equation will be of the form y = ax + b</td>
<td>This equation will be of the form x = by + a</td>
</tr>
<tr>
<td>This equation is used to estimate the value of y given the value of x</td>
<td>This equation is used to estimate the value of x given the value of y</td>
</tr>
<tr>
<td>The slope of this equation is (b_{yx})</td>
<td>The slope of this equation is (b_{xy})</td>
</tr>
<tr>
<td>The regression line of y on x is the straight line on the scatter diagram for which the sum of squares of vertical distances of the points is minimum.</td>
<td>The regression line of x on y is derived by the minimization of horizontal distance in the scatter diagram using method of least square.</td>
</tr>
<tr>
<td>The principle which is applied for deriving the two lines of regression is known as “Method of Least Squares”.</td>
<td>The principle which is applied for deriving the two lines of regression is known as “Method of Least Squares”.</td>
</tr>
</tbody>
</table>
### Regression Coefficient of y on x ($b_{yx}$):

1. **Using covariance:**
   
   $$b_{yx} = \frac{\text{Cov}(x, y)}{\sigma_x^2}$$

2. **Without any deviations (Directly from x and Y values):**
   
   $$b_{yx} = \frac{\sum xy - \frac{1}{n} \sum x \sum y}{\sum x^2 - \left(\frac{\sum x}{n}\right)^2}$$

3. **When deviations are taken from actual mean i.e., $\bar{x}$ and $\bar{y}$ such that $u = x - \bar{x}$ and $v = y - \bar{y}$:**
   
   $$b_{yx} = \frac{\sum (x - \bar{x})(y - \bar{y})}{\sum (x - \bar{x})^2} = \frac{\sum uv}{\sum u^2}$$

4. **When deviations are taken from assumed mean say A & B for x and y, $u = x - A$, $v = y - B$:**
   
   $$b_{yx} = \frac{\sum uv - \frac{1}{n} \sum u \sum v}{\sum u^2 - \left(\frac{\sum u}{n}\right)^2}$$

5. **Using ‘r’**
   
   $$b_{yx} = r \cdot \frac{\sigma_y}{\sigma_x}$$

   $$\sigma_x = S.D(x), \sigma_y = S.D(y)$$

   and $r = \text{Correlation co-efficient between x and y}$

### Regression Coefficient of x on y ($b_{xy}$):

1. **Using covariance:**
   
   $$b_{xy} = \frac{\text{Cov}(x, y)}{\sigma_y^2}$$

2. **Without any deviations (Directly from x and Y values):**
   
   $$b_{xy} = \frac{\sum xy - \frac{1}{n} \sum x \sum y}{\sum y^2 - \left(\frac{\sum y}{n}\right)^2}$$

3. **When deviations are taken from actual mean i.e., $\bar{x}$ and $\bar{y}$ such that $u = x - \bar{x}$ and $v = y - \bar{y}$:**
   
   $$b_{xy} = \frac{\sum (x - \bar{x})(y - \bar{y})}{\sum (y - \bar{y})^2} = \frac{\sum uv}{\sum v^2}$$

4. **When deviations are taken from assumed mean say A & B for x and y, $u = x - A$, $v = y - B$:**
   
   $$b_{xy} = \frac{\sum uv - \frac{1}{n} \sum u \sum v}{\sum v^2 - \left(\frac{\sum v}{n}\right)^2}$$

5. **Using ‘r’**
   
   $$b_{xy} = r \cdot \frac{\sigma_x}{\sigma_y}$$

   $$\sigma_x = S.D(x), \sigma_y = S.D(y)$$

   and $r = \text{Correlation co-efficient between x and y}$

**Properties of regression coefficients**

1. $b_{yx} = \text{slope of the regression line of y on x which measures the change in variable y for a unit change in variable x.}$
2. $b_{xy} = \text{slope of the regression line of x on y which measures the change in variable x for a unit change in variable y.}$
3. Correlation coefficient is symmetric i.e., $r_{yx} = r_{xy}$ but regression coefficients are not symmetric $b_{yx} \neq b_{xy}$.
4. When $r = 0$, both the regression coefficients are 0.
5. Both the regression coefficients will have same sign.

6. Correlation coefficient is the geometric mean between regression coefficients i.e.,
\[ r = \pm \sqrt{b_{yx} \cdot b_{xy}} \]

7. Sign analogy of \( b_{yx}, b_{xy} \) and \( r \)

<table>
<thead>
<tr>
<th>( b_{yx} )</th>
<th>( b_{xy} )</th>
<th>( r )</th>
</tr>
</thead>
<tbody>
<tr>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

Note:
When \( b_{yx} \) and \( b_{xy} \) are of opposite signs, data are inconsistent, \( r \) is imaginary.

8. Regression coefficients are independent of the Change of Origin but they are dependent on Change of Scale. If \( u = \frac{x - a}{c} \) and \( v = \frac{y - b}{d} \) then

i. \( b_{yx} = b_{vu} \cdot \frac{d}{c} \)

ii. \( b_{xy} = b_{uv} \cdot \frac{c}{d} \)

9. There is no specific range within which two regression coefficients will lie but their values should be such that the square root of the product of two regression coefficients must lie between -1 and +1 (both inclusive). Thus, if one of the regression coefficient, is greater than unity then the other must be less than unity.

Properties of regression lines:
- Two regression lines always intersect at their mean or average values \((\bar{x}, \bar{y})\). In other words if we solve two regression equations we get the average values of \( x \) and \( y \).
- When \( r = 0 \), then
  i. \( b_{yx} = b_{xy} = 0 \)
  ii. The two regression lines thus reduces to; \( y = \bar{y} \) and \( x = \bar{x} \)
  iii. Nothing can be predicted from the two regression lines since, the variables become independent.
  iv. The angle between the two regression lines becomes 90° i.e., they are perpendicular to each other.
- When \( r = \pm 1 \), then
  i. The two regression lines become identical i.e., they coincide.
  ii. \( b_{yx} = \frac{1}{b_{xy}} \)
  iii. Perfect linear co-relationship is observed and the angle between the two regression lines becomes 0°.
  iv. For a particular value of \( x \) we shall obtain a specific value of \( y \).
• As the angle between two regression lines numerically decreases from 90° to 0°, the correlation increases from 0 to 1 and the two regression lines comes closer to each other.

• Angle between two regression lines; if A is the angle between two regression lines then tan A = \[ \pm \frac{1 - r^2}{r} \left( \frac{\sigma_x \cdot \sigma_y}{\sigma_x^2 + \sigma_y^2} \right) \]

Miscellaneous Properties:
• In regression analysis, the difference between the Observed value and the Estimated value is known as Residue or Error.
• Proportion of Total Variance explained by regression analysis is \( r^2 \).
• Proportion of Total Unexplained Variance is \( 1 - r^2 \).
• Standard error of estimate of \( x(S_{xy}) \) is given by \( S_{xy} = \sqrt{\frac{\sum (x - \bar{x})^2}{N}} \) or \( \sigma_x \sqrt{1 - r^2} \)
• Standard error of estimate of \( y(S_{yx}) \) is given by \( S_{yx} = \sqrt{\frac{\sum (y - \bar{y})^2}{N}} \) or \( \sigma_y \sqrt{1 - r^2} \)
• When \( r^2 = 1 \), then;
  i. \( \frac{\text{Explained variance}}{\text{Total variance}} = 1 \)
  ii. Explained variance = Total Variance
  iii. The whole of the total variance is explained by regression.
  iv. The unexplained variation is zero
  v. All the points on the scatter diagram will lie on the regression line
  vi. There is a perfect linear dependence between the variables
  vii. The two regression lines coincide
  viii. For a given value of one variable, we have a fixed value of the other variable

Regression Lines
1. From the following data, find the regression equation of \( X \) on \( Y \):

<table>
<thead>
<tr>
<th>X</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Y</td>
<td>2</td>
<td>3</td>
<td>5</td>
<td>4</td>
<td>6</td>
</tr>
</tbody>
</table>

a) \( X = 0.9Y + 0.6 \)
b) \( X = 0.9Y - 0.6 \)
c) \( X = 0.9Y + 1.3 \)
d) \( X = 0.9Y - 1.3 \)

2. Taking data from the previous question, find the regression equation of \( Y \) on \( X \).

a) \( Y = 0.9X - 1.3 \)
b) \( Y = 0.9X - 0.6 \)
c) \( Y = 0.9X + 1.3 \)
d) \( Y = 0.9X + 0.6 \)
3. The information given below relates to the advertisement expenses and sales revenue of a company.

<table>
<thead>
<tr>
<th>Adv Exp Rs. Lakh</th>
<th>Sales Rs Lakh</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>20</td>
</tr>
<tr>
<td>S.D.</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>100</td>
</tr>
<tr>
<td></td>
<td>12</td>
</tr>
</tbody>
</table>

\[ r = +0.8 \]

What should be the advertisement expenditure, if the company wants to attain a sales target of Rs. 120 lakh?

a) Rs. 20,00,000  
b) Rs. 24,00,000  
c) Rs. 18,00,000  
d) None of the above

4. Determine the regression line of Y on X from the data given below:

<table>
<thead>
<tr>
<th>X</th>
<th>Y</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>6</td>
<td>4</td>
</tr>
<tr>
<td>8</td>
<td>5</td>
</tr>
<tr>
<td>9</td>
<td>7</td>
</tr>
<tr>
<td>11</td>
<td>8</td>
</tr>
<tr>
<td>14</td>
<td>9</td>
</tr>
</tbody>
</table>

a) \[ Y = 0.46X - 0.52 \]  
b) \[ Y = 0.64X + 0.52 \]  
c) \[ Y = 0.46X + 0.52 \]  
d) \[ Y = 0.64X - 0.25 \]

5. Given the following data, find the marks in statistics for a student who gets 90 marks in accountancy:

<table>
<thead>
<tr>
<th>Statistics</th>
<th>Accountancy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average Marks</td>
<td>67</td>
</tr>
<tr>
<td>S.D. of Marks</td>
<td>8</td>
</tr>
</tbody>
</table>

Correlation Coefficient \[ r = +0.75 \]

6. In the estimate of regression equation of two variable X and Y, the following results were obtained: Mean(X) = 90, Mean(Y) = 70, \( n = 10 \), \( \sum x^2 = 6360 \), \( \sum y^2 = 2860 \), \( \sum xy = 3900 \) (x and y are the deviations taken from respective means). Obtain the regression equation of Y on X.

7. Given that the means of X and Y are 65 and 67 respectively. Their standard deviations are 2.5 and 3.5 respectively, and the coefficient of correlation between them is 0.8. Obtain the best estimate of X, when \( Y = 70 \).

a) 56.61  
b) 66.71  
c) 79.81  
d) 70.71

8. With \( b_{xy} = 0.5 \), \( r = 0.8 \) and variance of y = 16, standard deviation of x equals to:

a) 6.4  
b) 2.5  
c) 10.0  
d) 26.5
9. For the regression line of y on x, $2x + 3y + 50 = 0$, find the value of $b_{yx}$.
   a) $\frac{2}{3}$
   b) $-\frac{2}{3}$
   c) $-\frac{3}{2}$
   d) None of the above

10. Regression equation of Y on X is $8X – 10Y + 66 = 0$ and $SD(x) = 3$, find the value of Cov (x, y).
    a) 11.25
    b) 7.2
    c) 2.4
    d) None of the above

11. If Mean of x = 10, Mean of y = 50, $SD(x) = 3$, $SD(y) = 15$, $r = 0.9$, then find the estimated value of x for corresponding y = 100.
    a) 18
    b) 19
    c) 20
    d) 21

12. For 100 students of a class, the regression equation of marks statistics (X) on the marks in Economics(Y) is $3Y-5X+180=0$. The mean marks in Economics is 50 and variance of marks in statistics is 4/9 of the variance of the marks in Economics. The mean marks in statistics is & Correlation coefficient is
    a) 56, $r = 0.7$
    b) 55, $r = 0.8$
    c) 66, $r = 0.9$
    d) 65, $r = 0.7$

13. While calculating the coefficient of correlation between two variables X and Y the following results were obtained:
The number of observation N=25
   $\sum X = 125, \sum Y = 100, \sum X^2 = 650, \sum Y^2 = 460, \sum XY = 508$, It was however, later discovered at the time of checking that two pairs of observations (X,Y) were copied (6,14) and (8,6) while the correct values (8,12) and (6,8) respectively. The correct value of byx is:
    a) 0.6
    b) 0.7
    c) 0.8
    d) 0.65

14. Obtain the estimate of ‘m’ if regression equation of saving (x) on income (y) is expressed as $x= a+y/m$ where ‘m’ are constants and it is known that the coefficient of correlation between x and y based on a sample of 100 families is 0.4 and the variance of saving is one quarter the variance of incomes.
    a) 5
    b) 4
    c) 3
    d) 2
15. Find the regression line of profits on output from the following data using:

<table>
<thead>
<tr>
<th>Output (1000 units)</th>
<th>5</th>
<th>7</th>
<th>9</th>
<th>11</th>
<th>13</th>
<th>15</th>
</tr>
</thead>
<tbody>
<tr>
<td>Profits per unit (Rs.)</td>
<td>1.7</td>
<td>2.4</td>
<td>2.8</td>
<td>3.4</td>
<td>3.7</td>
<td>4.4</td>
</tr>
</tbody>
</table>

a) $y=0.257x+0.497$
b) $y=0.255x+0.487$
c) $y=0.277x+0.447$
d) None

**Properties of Regression Coefficients**

16. If $b_{yx} = 0.8$, $b_{xy} = 0.46$, then the value of $r$ is:
   a) 0.60
   b) 0.61
   c) 0.51
   d) None of the above

17. If $r = 0.4$, $b_{xy} = 0.2$, then find the value of $b_{yx}$.
   a) $-0.8$
   b) 0.8
   c) 0.2
   d) None of the above

18. Regression Coefficient of $y$ on $x = 0.8$. Regression coefficient of $x$ on $y = 0.2$ coefficient of correlation = -0.4. Given data is:
   a) Accurate
   b) Inaccurate
   c) True
   d) None

19. If $b_{xy} = -1.2$ and $b_{yx} = -0.3$, then the coefficient of correlation between $x$ and $y$ is:
   a) $-0.698$
   b) $-0.36$
   c) $-0.51$
   d) $-0.6$

20. If the regression coefficient of $y$ on $x$ is $4/3$, then the regression coefficient of $x$ on $y$ is:
   a) More than 1
   b) Less than 1
   c) Less than zero
   d) None of the above

21. Given $b_{xy} = 1.36$, $b_{yx} = 0.613$, then the value of coefficient of determination is:
   a) 0.734
   b) 0.634
   c) 0.534
   d) 0.834
22. Given $b_{xy} = 0.756$, $b_{yx} = 0.659$, then the value of coefficient of non-determination is given by:
   a) 0.402
   b) 0.502
   c) 0.602
   d) 0.702

23. Find the mean of $x$ and $y$, if the two regression lines are $3x – y – 5 = 0$ and $2x – y – 4 = 0$.
   a) 1, -2
   b) –1, 2
   c) 2, -1
   d) –2, -1

24. Given $x = 4y + 5$ and $y = kx + 9$ are the lines of regression of $x$ on $y$ and $y$ on $x$ respective.
   If $k = 1/16$, the the value of $r$ is:
   a) 0.4
   b) 0.5
   c) 0.6
   d) 0.9

25. A relationship $r^2 = 1 – 5/14$, is it possible?
   a) Possible
   b) Impossible
   c) May be possible
   d) None of the above

26. If the Coefficient of non-determination is 0.502, then the Coefficient of alienation is:
   a) 0.71
   b) 0.61
   c) 0.51
   d) 0.81

27. Given $\sigma_x = 4$, $r=1$, then the standard error of estimate of $X$ on $Y$ is:
   a) 1
   b) 0
   c) 2
   d) 3

28. Given $b_{yx} = 3$, $\sigma_y = 12$ and $r=1$ then the standard error of estimate of $X$ on $Y$ is:
   a) 4
   b) 3
   c) 2
   d) 0
29. If \( u = 2x + 5 \), \( v = -3y - 6 \), and the regression coefficient of \( y \) on \( x \) is 2.4, the regression coefficient of \( v \) on \( u \) is:
   a) 1.6
   b) – 3.6
   c) 3.6
   d) – 1.6

30. If \( u = 2x + 5 \), \( v = -3y + 1 \), and the regression coefficient of \( y \) on \( x \) is – 1.2, the regression coefficient of \( v \) on \( u \) is:
   a) 1.8
   b) – 1.8
   c) 3.26
   d) 0.8

31. If \( 4u = 2x + 7 \) and \( 6v = 2y - 15 \), and the regression coefficient of \( y \) on \( x \) is 3, then the regression coefficient of \( v \) on \( u \) is:
   a) 3
   b) –2
   c) 4.5
   d) 2

32. Out of the two regression lines \( x + 2y = 4 \) and \( 2x + 3y - 5 = 0 \), which is the regression line of \( x \) on \( y \)?
   a) \( 2x + 3y - 5 = 0 \)
   b) \( x + 2y = 4 \)
   c) Both a) and b) above
   d) Neither a) nor b) above

33. In a partially destroyed record, the following data are available: variance of \( x = 9 \) and the regression lines are \( 8x - 10y + 66 = 0 \) and \( 40x - 18y = 214 \), find the coefficient of correlation between \( x \) and \( y \).
   a) – 0.6
   b) 0.36
   c) – 0.36
   d) + 0.6

34. Taking data from the previous question, what would be the standard deviation of \( y \)?
   a) 9
   b) 4
   c) 6
   d) 5
From the following regression equations answer the questions that follows:
3x – 2y – 10 = 0
24x – 25y + 145 = 0

35. What is the ratio of mean value of x and y?
   a) 4 : 5
   b) 5 : 4
   c) 20: 17
   d) 19:25

36. What is the ratio of standard deviation of x and y?
   a) 6 : 5
   b) 7 : 8
   c) 1 : 1
   d) 5 : 6

37. What is the value of correlation coefficient between x and y?
   a) – 0.8
   b) – 0.56
   c) + 0.8
   d) None of the above

Theoretical Aspects

Introduction

38. The word regression is used to denote ________ of the average value of one variable for
    a specified value of the other variable.
   a) Estimation
   b) Prediction
   c) Either a) or b) above
   d) None of the above

39. Regression methods are meant to determine:
   a) The nature of relationship between the variables.
   b) The functional relationship between the two variables.
   c) Both a) and b) above
   d) Neither a) nor b) above.

40. Regression analysis is a mathematical measure of the __________ relationship between
    two or more variables in term of the original units of the data.
   a) Continuous
   b) Direct
   c) Average
   d) Both b) and c) above.
41. The two types of variables in regression analysis are:
   a) Direct & Indirect
   b) Dependent & Independent
   c) Discrete & Continuous
   d) None of the above

42. The dependent variable in the regression analysis is one:
   a) Which influences the value of the independent variable.
   b) Whose value is to be predicted.
   c) Which can choose its value independently.
   d) None of the above.

43. If the curve plotted on a Scatter Diagram is a straight line, it is called the:
   a) Line of correlation
   b) Line of scatter diagram
   c) Line of regression
   d) Both a) and c) above

44. The estimation in regression analysis is done by means of suitable equations, derived
   on the basis of available bivariate data. Such an equation is known as:
   a) Recursive Equations
   b) Recurring Equations
   c) Regression Equations
   d) Both a) and c) above

45. The geometrical representation of regression equations is known as:
   a) Straight Line
   b) Regression Line
   c) Regression Curve
   d) Scatter Curve

46. The line of regression is:
   a) The line which gives the best estimate to the value of one variable for any specified
      value of the other variable.
   b) The line which gives the best estimate to the value of all variables for any arbitrary
      value of a constant variable.
   c) The line showing the nature of relationship between two or more variables.
   d) None of the above.

47. For two variables, the number of regression lines would be:
   a) 1
   b) 3
   c) 2
   d) Greater than 3
48. For the relationship between two variables $X$ and $Y$, we have which of the followings lines of regression?
   a) $Y$ on $X$
   b) $X$ on $Y$
   c) Both a) and b) above
   d) Only b) above

49. Which of the following regression equations is used to estimate $Y$, when the value of $X$ is known?
   a) $X$ on $Y$
   b) $Y$ on $X$
   c) Either of the two can be used
   d) Neither of the two can be used

50. Since Yield of a crop depends upon amount of rainfall, we need to consider:
   a) The regression equation of yield on rainfall
   b) The regression equation of rainfall on yield
   c) Any one of a) or b) above can be considered
   d) Neither of a) or b) can be considered

51. The line $Y = a + bX$, represents the regression equation of:
   a) $X$ on $Y$
   b) $Y$ on $X$
   c) Both a) and b) above
   d) Neither a) nor b) above

52. The method applied for deriving the regression equations is:
   a) Fitting of Normal Equations
   b) Rank Correlation Method
   c) Least Square Method
   d) Product Moment Method

53. In a regression equation $Y = a + bX$, what is “$b$”?
   a) The intercept which the line cuts in the $Y$-axis.
   b) The intercept which the line cuts in the $X$-axis.
   c) The slope of the line.
   d) Both a) and c) above.

54. Regression coefficient of $X$ on $Y$ is denoted by:
   a) $b_{xy}$
   b) $b_{yx}$
   c) $a$
   d) $a_{xy}$
55. The slope of the regression line of x on y is given by:
   a) $b_{xy}$
   b) $b_{yx}$
   c) $1/b_{xy}$
   d) $1/b_{yx}$

56. The regression line of x on y is derived by:
   a) The minimization of vertical distance in the scatter diagram.
   b) The minimization of horizontal distance in the scatter diagram.
   c) Either a) or b) above.
   d) Both a) and b) above.

57. Regression line of y on x is:
   a) That straight line on the scatter diagram for which the sum of squares of vertical
distances of the points is minimum.
   b) That straight line on the radar diagram for which the sum of squares of horizontal
distances of the points is maximum.
   c) That straight line on the scatter diagram for which the sum of squares of vertical
distances of the points is maximum.
   d) None of the above.

**Properties:**

58. When the slope of two regression lines are equal:
   a) The lines are perpendicular to each other.
   b) The lines will coincide.
   c) The lines will be parallel to each other.
   d) None of the above.

59. If $r = +1$, the two lines of regression become:
   a) Perpendicular to each other.
   b) Identical
   c) Parallel to each other.
   d) Either a) or c) above.

60. The regression coefficients are zero if:
   a) $r = 1$
   b) $r = 2$
   c) $r = 0$
   d) $r = -1$

61. Correlation coefficient is the ________ of the two regression coefficients.
   a) Harmonic Mean
   b) Geometric Mean
   c) Arithmetic Mean
   d) Both b) and c) above
62. The sign analogy of correlation coefficients and two regression coefficients is:
   a) +, +, +
   b) -, -, -
   c) +, +, +
   d) Both b) and c) above

63. The point of intersection of two regression lines is:
   a) (x, y)
   b) (\(\bar{x}, \bar{y}\))
   c) (\(\sigma_x, \sigma_y\))
   d) None

64. As “r” increases numerically from 0 to 1, the angle between the regression lines:
   a) Increases from 0° to 90°
   b) Diminishes from 90° to 0°
   c) Increases from 0° to 180°
   d) Both a) and c) above

65. When r = 0, the regression lines are:
   a) Parallel to each other
   b) Perpendicular to each other
   c) Coincides
   d) Either a) or b) above

66. If correlation coefficient between two variables is zero, which of the following is true?
   a) Both regression coefficients are greater than one.
   b) Both regression coefficients are negative.
   c) One of the regression coefficient is zero.
   d) Both regression coefficients are zero.

67. In regression analysis the difference between the observed value and the estimated value is known as:
   a) Residue
   b) Deviation
   c) Error
   d) Either a) or c) above

68. Which of the following(s) is/are TRUE regarding regression coefficient?
   a) If \(b_{xy} > 0\), then \(r < 0\)
   b) If \(b_{xy} < 0\), then \(r > 0\)
   c) If the variable X and Y are independent, the regression coefficient is zero.
   d) The range of regression coefficient is –1 to +1.
69. Which of the following statement/s is/are FALSE regarding the regression coefficient?
   
a) If one of the regression coefficient is greater than unity the other one is less than unity.
   
b) The product of two regression coefficient is equal to the square of the correlation coefficient between the two variables.
   
c) The regression coefficient lies between – infinity to + infinity.
   
d) None of the above is FALSE.

70. The regression coefficients remain unaffected due to a:
   
a) Change of scale
   
b) Change of origin
   
c) Both a) and b) above
   
d) None of the above

<p>| | | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>38</td>
<td>c</td>
<td>46</td>
<td>a</td>
<td>54</td>
<td>a</td>
<td>62</td>
</tr>
<tr>
<td>39</td>
<td>c</td>
<td>47</td>
<td>c</td>
<td>55</td>
<td>a</td>
<td>63</td>
</tr>
<tr>
<td>40</td>
<td>c</td>
<td>48</td>
<td>c</td>
<td>56</td>
<td>b</td>
<td>64</td>
</tr>
<tr>
<td>41</td>
<td>b</td>
<td>49</td>
<td>b</td>
<td>57</td>
<td>a</td>
<td>65</td>
</tr>
<tr>
<td>42</td>
<td>b</td>
<td>50</td>
<td>a</td>
<td>58</td>
<td>b</td>
<td>66</td>
</tr>
<tr>
<td>43</td>
<td>c</td>
<td>51</td>
<td>b</td>
<td>59</td>
<td>b</td>
<td>67</td>
</tr>
<tr>
<td>44</td>
<td>c</td>
<td>52</td>
<td>c</td>
<td>60</td>
<td>c</td>
<td>68</td>
</tr>
<tr>
<td>45</td>
<td>b</td>
<td>53</td>
<td>c</td>
<td>61</td>
<td>b</td>
<td>69</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>70</td>
</tr>
</tbody>
</table>