

CORRELATION & REGRESSION

Correlation:

Correlation analysis is a statistical technique used to measure the degree and direction of relationship between the variables.

if $y = f(x)$ $x =$ independent $y =$ dependent.

Types of correlation: (i) (Positive and Negative correlation)

x	y	Types of correlation
Increase	Increase	Positive
Decrease	Decrease	Positive
Increase	Decrease	Negative
Decrease	Increase	Negative
Increase	No change	No correlation
Decrease	No change	No correlation

- (ii) **Simple correlation:-** When only two variables are studied, it is a case of simple correlation.
- (iii) **Multiple correlation:-** When three or more variables are studied, it is a case of multiple correlation.
- (iv) **Partial multiple correlation:-** In partial multiple correlation one studies three or more variables but considers only two variables to be influencing each other and effect of other variables being held constant.
- (v) **Total multiple correlation:-** In total multiple correlation one studies three or more variables without including the effect of any variable held as constant.
- (vi) **Linear Correlation:-** If the amount of change in one variable bears a constant ratio to the amount of change in the other variable, then correlation is said to be linear.
- (vii) **Non Linear (Curvilinear) Correlation:-** If the amount of change in one variable does not bear a constant ratio to the amount of change in the other variable, then correlation is said to be non-linear. If such variables are plotted on a graph, the point would fall on a curve and not on a straight line.

Univariate Data:- When there is only one variable.

Bivariate frequency distribution:- When two or more variables are considered at the same time of point, known as bivariate frequency distribution.

Bivariate frequency table:-

Age in Years(x)/weight y in Kg.	40-45	45-50	50-55	55-60	60-65
45-50	2	5	8	3	0
50-55	1	3	6	10	2
55-60	0	2	5	12	1

Note: We can obtain two types of univariate distributions which are known as

- (i) Marginal distribution
- (ii) Conditional distribution

Marginal Frequency Distribution for x.

x	40-45	45-50	50-55	55-60	60-65
y	3	10	19	25	3

(Marginal Frequency Distribution for y.

x	40-45	50-55	55-60
y	18	22	20

(Conditional frequency distribution for age x when weight y is between 55-60.

x	40-45	45-50	50-55	55-60	60-65
y	0	2	5	12	1

Note: If there are m classification for x and n classification for y, there would be altogether (m + n) conditional distribution.

Multi Variate: When more than two variable at the some point of time.

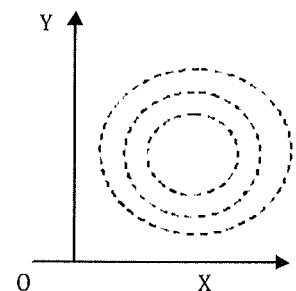
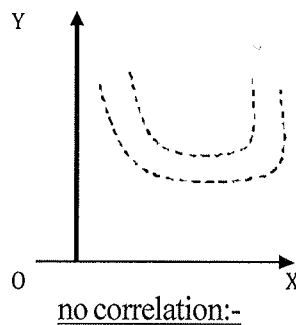
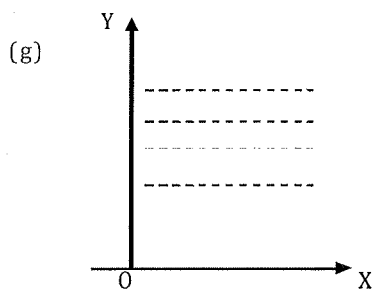
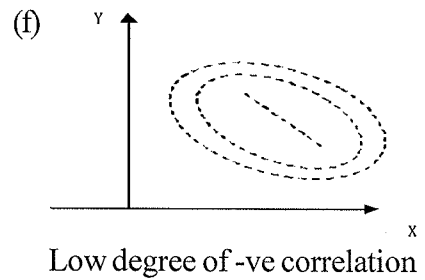
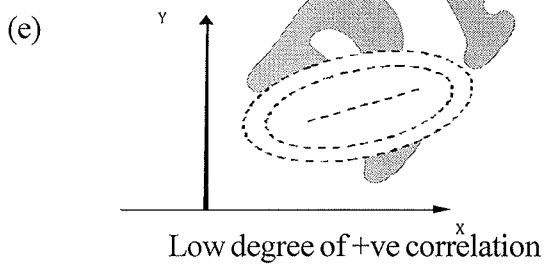
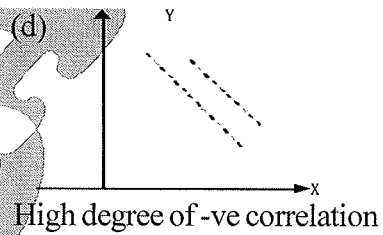
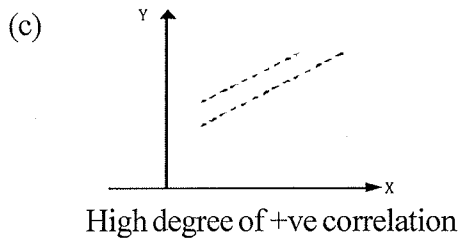
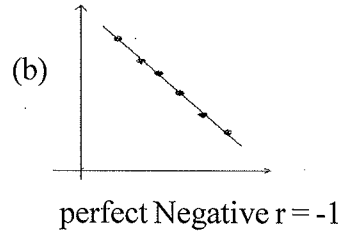
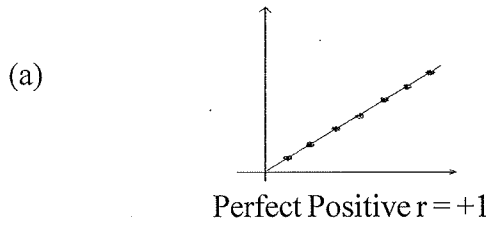
Covariance:-

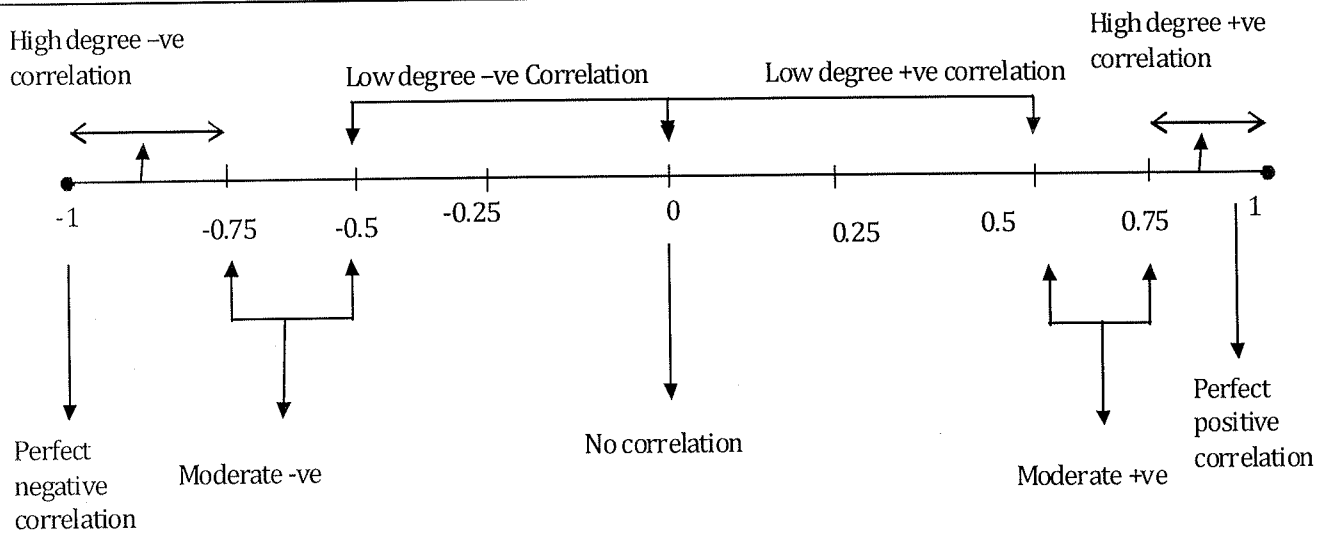
- (i)
$$\text{cov}(x, y) = \frac{\sum(x - \bar{x})(y - \bar{y})}{n}$$
- (ii)
$$\text{cov}(x, y) = \frac{\sum xy}{n} - \frac{\sum x}{n} \cdot \frac{\sum y}{n}$$
- (iii)
$$\text{cov}(x, y) = E(x, y) - E(x)E(y).$$

Measures of Correlation

- (i) Scatter diagram
- (ii) Karl Pearson's product moment correlation coefficient
- (iii) Spearman's rank correlation coefficient
- (iv) Coefficient of concurrent deviations.

(I) Scatter Diagram method:





II Karl Pearson's Product Moment Correlation Coefficient:-

Best method for finding correlation between two variables provided the relationship between the two variables is linear. Correlation coefficient = r or $\rho(x, y)$.

1. $r = \frac{\text{cov}(x, y)}{S_x S_y} = \frac{\Sigma(x - \bar{x})(y - \bar{y})}{NS_x S_y}$ $S_x = \text{S.D. of } x$ $S_y = \text{S.D. of } y$. $N = \text{No of obs pairs.}$

2. $r = \frac{N\Sigma xy - \Sigma x \Sigma y}{\sqrt{N\Sigma x^2 - (\Sigma x)^2} \sqrt{N\Sigma y^2 - (\Sigma y)^2}}$

3. $r = \frac{N\Sigma dx dy - \Sigma dx \Sigma dy}{\sqrt{N\Sigma dx^2 - (\Sigma dx)^2} \sqrt{N\Sigma dy^2 - (\Sigma dy)^2}}$ Where $dx = X - A$, $A = \text{Assumed mean of } x \text{-series.}$
 $dy = Y - B$, $B = \text{Assumed mean of } y \text{-series.}$

4. $r = \frac{\Sigma xy}{\sqrt{\Sigma x^2 \Sigma y^2}}$ where $x = x - \bar{x}$: $y = y - \bar{y}$ or $r = \frac{\Sigma xy}{NS_x S_y}$

Properties of correlation coefficient:-

- (i) The coefficient of correlation is a unit free measure.
- (ii) The coefficient of correlation remains invariant under a change of origin and/or scale of the variables under consideration depending on the sign of scale factors.

$$u = \frac{x - a}{b} \quad v = \frac{y - c}{d}$$

$$r_{xy} = \frac{bd}{|b||d|} r_{uv}$$

- (iii) The coefficient of correlation always lies between -1 and 1, including both the limiting values.
 $-1 \leq r \leq 1$

Probable Error in Correlation:-

$$P.E. = 0.6745 \times \frac{(1 - r^2)}{\sqrt{n}} = \frac{2}{3} \times \frac{1 - r^2}{\sqrt{n}}$$

n = No of pairs of observation; r = coefficient of correlation

Properties of Probable Error:-

- (i) If $r < 6 \times P.E.$ then it is not significant
- (ii) If $r \geq 6 \times P.E.$ then it is significant.
- (iii) Limit of correlation coefficient $P = r \pm P.E.$
 P = Correlation coefficient of population.
- (iv) Probable error is never negative.

Standard Error:- $S.E = \frac{1 - r^2}{\sqrt{n}}$

Coefficient of determination:- It is the square of the coefficient of correlation i.e. r^2 where r is the coefficient of correlation.

Coefficient of determination = $\frac{\text{Explained variance}}{\text{Total Variance}}$

Coefficient of Non-determination = $\frac{\text{Unexplained variance}}{\text{Total Variance}}$

i.e. $K^2 = 1 - r^2$ Where K^2 = coefficient of non-determination.

Coefficient of Alienation:- $\sqrt{1 - r^2}$

Limit of Standard Error = $r \pm 3 S.E$

$r = \frac{LSE + USE}{2}$ LSE = Lower Standard error; USE = Upper Standard error.

III Spearman's Rank Correlation Coefficient:-

It is also known as Ranking method, Rank difference method or qualitative or attribute method of correlation.

$$r_R = 1 - \frac{6\sum d_i^2}{n(n^2-1)} ; d_i = x_i - y_i \text{ represent the difference in ranks; } n = \text{No. of individuals.}$$

In case of tied Ranks:- $r_R = 1 - \frac{6 \left[\sum d_i^2 + \frac{\sum(t_j^3 - t_j)}{12} \right]}{n(n^2-1)}$ Where t_j represents the j th tie length

(i) $-1 \leq r_R \leq 1$ (ii) Suitable for qualitative data.

IV Concurrent Deviation method:-

$$r_c = \pm \sqrt{\pm \frac{(2c - m)}{m}}$$

C = No of concurrent deviation (No. of +ve sign in the product of deviation column)
 m = total No. of pairs of deviation.

If $(2c - m) > 0$ we take the +ve sign both inside and outside the radical sign.

$(2c - m) < 0$ we take -ve sign both inside and outside the radical sign.

- Note:**
- (i) $-1 \leq r_c \leq 1$
 - (ii) Suitable for large n.
 - (iii) If all concurrent deviation are +ve then perfect the correlation i.e. if $c = n$ then $r = 1$.
 - (iv) If all concurrent deviations are -ve then perfect -ve correlation i.e. $c = 0$ then $r = -1$.
 - (v) No. of +ve concurrent deviation = No. of -ve concurrent deviation then $r = 0$.
 - (vi) If No. of +ve concurrent deviation > No. of -ve concurrent deviation then r will be +ve otherwise -ve

Spurious correlation: (Non-Sense Correlation):- It is the correlation between two variables having no casual relationship.

$$\text{If } u = \frac{x - a}{h} \qquad v = \frac{y - b}{k}$$

$$\text{Cov}(x, y) = h.k. \text{cov}(u, v).$$

- (i) Covariance has no range whereas correlation coefficient has a range $-1 \leq r \leq 1$
- (ii) The correlation and covariance for any two independent variable is always zero.

Some Examples of positive correlation are:

- (i) Age of applicants for life insurance and the premium of insurance.
- (ii) Production of pig iron and soot content in Durgapur.
- (iii) The ages of husbands and wives.
- (iv) Years of education and income.
- (v) Amount of rainfall and yield of crop.
- (vi) Age and income of employed persons.
- (vii) Speed of an automobile and the distance required to stop the car often applying brakes.
- (viii) Sale of cold drinks and day temperature.

Some Examples of Negative correlation are:-

- (i) Unemployment index and purchasing power of the common man.
- (ii) Prices and Demand for goods under normal times.
- (iii) 'Insurance companies' profits and no. of claims they have to pay.
- (iv) Sale of woollen garments and day temperature.
- (v) Production and price per unit.

Regression Analysis:-

By regression, we mean average relationship between two or more variables. One of these variables, is called the dependent or explained variable and other variable independent or explaining variable. If the explaining variables are two or more then two it will be called multiple regression analysis.

Difference between correlation and regression:-

Correlation	Regression
1. Correlation measures degree and direction of relationship between the variables.	1. Regression measures the nature and extent of average relationship between two or more variables in terms of original units of data.
2. It is relative measure	2. It is absolute measure
3. It is independent of change of both origin and scale	3. It is independent of origin and not scale.
4. It is independent of units of measurement.	4. It is not independent of units of measurement.
5. Any is symmetric i.e. $r_{xy} = r_{yx}$. r_{xy}, r_{yx} are correlation coefficient.	5. b_{xy} is not symmetric $b_{xy} \neq b_{yx}$. b_{xy} and b_{yx} are regression coefficient.

Regression Lines:-

(i) Regression Line of X on Y.

$X = a + b Y$; X = Dependent variable Y = Independent Variable.

Note: This line gives the probable value of X for any given value of Y.

Another way:

$$X - \bar{X} = b_{xy} (Y - \bar{Y}); X - \bar{X} = r \cdot \frac{\sigma_x}{\sigma_y} (Y - \bar{Y}) \text{ where } r = \text{coefficient of correlation between } x \text{ and } y$$

(ii) Regression Line of y on x.

$Y = a + b X$; X = Independent variable Y = dependent variable

Note: This line gives the probable value of y for any given value of x.

Another way:-

$$Y - \bar{y} = b_{yx} (X - \bar{X}); Y - \bar{y} = r \cdot \frac{\sigma_y}{\sigma_x} (X - \bar{X}) \text{ where } r = \text{coefficient of correlation between } X \text{ and } Y.$$

Properties of Linear regression:-

- (i) $b_{yx} \cdot b_{xy} = r^2$ i.e. $r = \sqrt{b_{yx} \cdot b_{xy}}$
- (ii) r, b_{yx} and b_{xy} all have the same sign.
- (iii) If the correlation coefficient r is zero, the regression coefficient b_{yx} and b_{xy} are also zero.
- (iv) Regression lines always intersect at their means.
- (v) The angle between the two regression lines depends on the correlation coefficient (r).

Value of r.	Angle between Regression lines
(a) if $r = 0$	Regression lines are perpendicular to each other.
(b) if $r = +1$ or -1	regression line coincide (Become identical)

Note:- As value of r increases numerically from 0 to 1, the angle between regression equation decreases from 90° to 0° .

i.e. the farther the two regression lines are from each other, the lesser is the degree of correlation.

- (vi) The value of x or y can be estimated from linear equation if $r \neq 0$.

Properties of Regression coefficients:-

- (i) Both regression coefficient have the same signs, i.e. either they will be positive or negative.
- (ii) If one of the regression coefficients is greater than unity, the other must be less than unity.
- (iii) A.M. $> r$ i.e. Arithmetic mean of regression coefficient is greater than correlation coefficient.
- (iv) r is G.M. of regression coefficient.

$$\text{i.e. } r = \pm \sqrt{b_{xy} \cdot b_{yx}}$$

- (v) If regression coefficient have +ve sign, r will also be positive and if regression coefficient have a -ve sign, r will also be negative.
- (vi) If $b_{xy} = b_{yx}$ then $r = b_{xy} = b_{yx}$.
- (vii) If $\sigma_x = \sigma_y$ then $r = b_{xy} = b_{yx}$.
- (viii) If $u = \frac{x-a}{p}$ and $v = \frac{y-b}{q}$ then $b_{yx} = \frac{q}{p} \times b_{uv}$; and $b_{xy} = \frac{p}{q} \times b_{uv}$
- (ix) For regression Line Y on X

$\Sigma Y - Y_c = 0$ and $\Sigma (Y - Y_c)^2$ is least where Y is observed value and Y_c is estimated value. $Y - Y_c$

Y_c is known as error or residue or vertical deviation & it may be positive or negative or 0.

- (x) For regression Line X on Y.

$\Sigma (X - X_c) = 0$ and $\Sigma (X - X_c)^2 = \text{Least}$. $X - X_c = \text{Error or Residue Horizontal deviation}$

Error may be positive or Negative or Zero

Exercise - I

1. If $\text{cov}(x, y) = 15$, what restrictions should be put for the standard deviations of x and y ? [SM]
 - (a) No restriction
 - (b) The product of the standard deviations should be more than 15
 - (c) The product of the standard deviations should be less than 15
 - (d) The sum of the standard deviation should be less than 15

2. What is the coefficient of correlation from the following data? [SM]

x:	1	2	3	4	5
y:	8	6	7	5	5

 - (a) 0.75
 - (b) -0.75
 - (c) -0.85
 - (d) 0.82

3. If for two variable x and y , the covariance, variance of x and variance of y are 40, 16 and 256 respectively, what is the value of the correlation coefficient? [SM]
 - (a) 0.01
 - (b) 0.625
 - (c) 0.4
 - (d) 0.5

4. From the following data [SM]

x:	2	3	5	4	7
y:	4	6	7	8	10

Two coefficient of correlation was found to be 0.93. What is the correlation between u and v as given below?

u:	-3	-2	0	-1	2
v:	-4	-2	-1	0	2

 - (a) -0.93
 - (b) 0.93
 - (c) 0.57
 - (d) -0.57

5. Referring to the data presented in Q.No. 4, what would be the correlation between u and v ?

u	10	15	25	20	35
v:	-24	-36	-42	-48	-60 [SM]

 - (a) -0.6
 - (b) 0.6
 - (c) -0.93
 - (d) 0.93

6. If the coefficient of correlation between two variables is -0.9, then the coefficient of determination is [SM]
 - (a) 0.9
 - (b) 0.81
 - (c) 0.1
 - (d) 0.19

7. If the coefficient of correlation between two variables is 0.7 then the percentage of variation unaccounted for is [SM]
 - (a) 70%
 - (b) 30%
 - (c) 51%
 - (d) 49%

8. If the covariance between two variables is 20 and the variance of one of the variables is 16, what would be the variance of the other variable?
 - (a) More than 100
 - (b) More than 10
 - (c) Less than 10
 - (d) More than 1.25

9. If $r = 0.6$ then the coefficient of non-determination is
 - (a) 0.4
 - (b) -0.6
 - (c) 0.36
 - (d) 0.64

10. If $u + 5x = 6$ and $3y - 7v = 20$ and the correlation coefficient between x and y is 0.58 then what would be the correlation coefficient between u and v ?
 - (a) 0.58
 - (b) -0.58
 - (c) -0.84
 - (d) 0.84

11. From the data given below, calculate coefficient of correlation and interpret it:

	X	Y
Number of items	8	8
Mean	68	69
Sum of squares of deviations from mean	36	44
Sum of products of deviations of X and Y-series =	24	

(a) $r = 0.603$, PE = 0.15 (b) $r = 0.703$, PE = 0.15 (c) $r = 0.803$, PE = 0.17 (d) $r = 0.903$, PE = 0.19

12. No. of pairs of observation of X and Y-series = 1000

Standard deviation of X-series = 4.5

Standard deviation of Y-series = 3.6

Summation of product of corresponding deviations from respective means of X and Y-series = 4800.

Calculate coefficient of correlation between X and Y

(a) $r = 0.289$ (b) $r = 0.296$ (c) $r = 0.278$ (d) $r = 0.267$

13. If the covariance between X and Y-series is +6 and variance of X and Y-series are respectively 6 and 6.8, find the coefficient of correlation between them.

(a) $r = 0.246$ (b) $r = 0.757$ (c) $r = 0.939$ (d) $r = 0.838$

14. Find the coefficient of correlation between X and Y if covariance between X and Y variables is 10 and the variance of X and Y are respectively 16 and 9.

(a) $r = 0.36$ (b) $r = 0.55$ (c) $r = 0.74$ (d) $r = 0.83$

15. The coefficient correlation between two variables X and Y is -.75 and their covariance is -15. If variance of X is 25, find the second moment about mean of Y-series.

(a) 16 (b) 17 (c) 18 (d) 19

16. The coefficient correlation between two variables X and Y is +.8 and their covariance is +40. If the variance of X-series is 16, find the standard deviation of Y-series.

(a) 15.2 (b) 17.3 (c) 22.4 (d) 12.5

17. Find P.E if $r = +.8$ and $N = 25$.

(a) PE = 0.1487 (b) PE = 0.0486 (c) PE = 0.2488 (d) PE = 0.3489

18. Find r if P.E. of 10 observation is .07.
- (a) 0.75 (b) 0.64 (c) 0.82 (d) 0.73
19. Find out the limit if $r = .6$ and $N = 16$.
- (a) 0.492 to 0.708 (b) 0.595 to 0.605 (c) 0.395 to 0.606 (d) 0.195 to 0.509
20. For what value of N , the coefficient of correlation equal to $+0.6$ will be significant.
- (a) $N = 16$ (b) $N = 17$ (c) $N = 18$ (d) $N = 19$
21. While computing rank correlation coefficient between profit and investment for the last 6 years of a company the difference in rank for a year was taken 3 instead of 4. What is the rectified rank correlation coefficient if it is known that the original value of rank correlation coefficient was 0.4 ?[SM]
- (a) 0.3 (b) 0.2 (c) 0.25 (d) 0.28
22. Eight contestants in a musical contest were ranked by two judges A and B in the following manner:
- | Serial Number of the contestants: | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|-----------------------------------|---|---|---|---|---|---|---|---|
| Rank by Judge A: | 7 | 6 | 2 | 4 | 5 | 3 | 1 | 8 |
| Rank by Judge B: | 5 | 4 | 6 | 3 | 8 | 2 | 1 | 7 |
- The rank correlation coefficient is [SM]
- (a) 0.65 (b) 0.63 (c) 0.60 (d) 0.57
23. Following are the marks of 10 students in Botany and Zoology:
- | Serial No.: | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|-------------------|----|----|----|----|----|----|----|----|----|----|
| Marks in Botany: | 58 | 43 | 50 | 19 | 28 | 24 | 77 | 34 | 29 | 75 |
| Marks in Zoology: | 62 | 63 | 79 | 56 | 65 | 54 | 70 | 59 | 55 | 69 |
- The coefficient of rank correlation between marks in Botany and Zoology is [SM]
- (a) 0.65 (b) 0.70 (c) 0.72 (d) 0.75
24. What is the value of Rank correlation coefficient between the following marks in Physics and Chemistry:
- | Roll No.: | 1 | 2 | 3 | 4 | 5 | 6 |
|---------------------|----|----|----|----|----|--------|
| Marks in Physics: | 25 | 30 | 46 | 30 | 55 | 80 |
| Marks in Chemistry: | 30 | 25 | 50 | 40 | 50 | 78[SM] |
- (a) 0.782 (b) 0.85 (c) 0.89 (d) 0.857
25. For 10 pairs of observations, No. of concurrent deviations was found to be 4. What is the value of the coefficient of concurrent deviation ?[SM]
- (a) $\sqrt{0.2}$ (b) $-\sqrt{0.2}$ (c) $1/3$ (d) $-1/3$
26. The coefficient of concurrent deviation for p pairs of observations was found to be $1/\sqrt{3}$. If the number of concurrent deviations was found to be 6, then the value of p is . [SM]
- (a) 10 (b) 9 (c) 8 (d) none of these

27. What is the coefficient of concurrent deviations for the following data :

Supply:	68	43	38	78	66	83	38	23	83	63	53
Demand:	65	60	55	61	35	75	45	40	85	80	85[SM]
(a) 0.82			(b) 0.85			(c) 0.89			(d) -0.81		

28. What is the coefficient of concurrent deviations for the following data :

Year:	1996	1997	1998	1999	2000	2001	2002	2003
Price:	35	38	40	33	45	48	49	52
Demand:	36	35	31	36	30	29	27	24 [SM]
(a) -1			(b) 0.43			(c) 0.5		(d) $\sqrt{2}$

PARAS

Exercise - II

1. Given the following data :[SM]

Variable :	x	y
Mean :	80	98
Variance:	4	9

Coefficient of correlation = 0.6

What is the most likely value of y when x = 90 ?

- (a) 90 (b) 103 (c) 104 (d) 107

2. Given below the information about the capital employed and profit earned by a company over the last twenty five years :

	Mean	SD
Capital employed (000 Rs)	62	5
Profit earned (000 Rs)	25	6

Correlation coefficient between capital employed and profit = 0.92. The sum of the Regression coefficient for the above data would be :[SM]

- (a) 1.871 (b) 2.358 (c) 1.968 (d) 2.346

3. In case the correlation coefficient between two variables is 1 , the relationship between the two variables would be[SM]

- (a) $y = a + bx$ (b) $y = a + bx, b > 0$
 (c) $y = a + bx, b < 0$ (d) $y = a + bx$, both a and b being positive.

4. If the relationship between two variables x and y in given by $2x + 3y + 4 = 0$, then the value of the correlation coefficient between x and y is [SM]

- (a) 0 (b) 1 (c) -1 (d) negative

5. You are given below the following information about advertisement and sales.

	Adv. Expenditure (X) (Rs. crores)	Sales (Y) (Rs. crores)
Mean	20	120
S.D.	5	25

Correlation coefficient, $r = + 0.8$

- (i) Calculate the two regression equations.
 (ii) Find the likely sales when advertisement expenditure is Rs. 25 crores.
 (iii) What should be the advertisement budget if the company wants to attain sales target of Rs. 150 crores ?

- (a) $x = 0.8 + 0.16y, y = 40 + 4x, 140cr., 24.8cr.$ (b) $x = 0.9 + 0.76y, y = 54 + 4x, 240cr., 35.8cr.$
 (c) $x = 0.6 + 0.74y, y = 64 + 4x, 340cr., 14.8cr.$ (d) $x = 0.7 + 0.75y, y = 63 + 4x, 341cr., 15.8cr.$

6. Given :
- | | X-series | Y-series |
|--------------------|----------|----------|
| Mean | 6 | 13 |
| Standard Deviation | 2.45 | 2.61 |
- Sum of the product of deviation from Mean of X and Y-series is 30 and number of items is 5.
- (i) Obtain the two regression equations.
 (ii) Estimate the value of X, when Y = 10
- (a) $r = 0.938, X_c = -5.44 + 0.88y, Y = 0.999x + 7$ (b) $r = 0.838, X_c = 5.44 - 0.87y, Y = x + 8$
 (c) $r = 0.738, X_c = -4.44 + 0.86y, Y = 0.9x + 7$ (d) $r = 0.638, X_c = -3.44 + 0.85y, Y = x$
7. Given :
- | | X-series | Y-series |
|--------------------|----------|----------|
| Mean | 18 | 100 |
| Standard Deviation | 14 | 20 |
- Coefficient of correlation between X and Y-series = +.8
- (A) Find the most probable value of Y if X is 70 and most probable value of X if Y is 90.
- (a) $Y = 359.44, X = 14.2$ (b) $Y = 259.44, X = 16.3$
 (c) $Y = 459.44, X = 15.1$ (d) $Y = 159.44, X = 12.4$
8. If the line $Y = 13 - 3X / 2$ is the regression equation of y on x then b_{yx} is
- (a) $\frac{2}{3}$ (b) $-\frac{2}{3}$ (c) $\frac{3}{2}$ (d) $-\frac{3}{2}$
9. The line $X = 31/6 - Y/6$ is the regression equation of
- (a) Y on X (b) X on Y (c) both (d) we can not say
10. If x and y satisfy the relationship $y = -5 + 7x$, the value of r is
- (a) 0 (b) -1 (c) +1 (d) none
11. 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 ?[SM]
- (a) 0.45 (b) 0.9375 (c) 0.6 (d) none of these
12. 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 ?[SM]
- (a) 0.5 (b) $-1/\sqrt{2}$ (c) -0.5 (d) none of these
13. 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 ?[SM]
- (a) $2/\sqrt{3/2}$ (b) $16/3$ (c) $4/3$ (d) 5
14. 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 ?[SM]
- (a) 1 (b) -1 (c) 7 (d) none of these

15. The two lines of regression are given by $8x + 10y = 25$ and $16x + 5y = 12$ respectively. If the variance of x is 25, what is the standard deviation of y ? [SM]
(a) 16 (b) 8 (c) 64 (d) 4
16. The difference between the observed value and the estimated value in regression analysis is known as
(a) Error (b) Residue (c) Deviation (d) (a) or (b)
17. For certain data $Y = 1.3 X$ and $X = 0.7 Y$ are the two regression equations. Compute the coefficient of correlation between X and Y .
(a) $r = .852$ (b) $r = .753$ (c) $r = .954$ (d) $r = .655$
18. For certain data, $4 Y = 9 X + 15$ and $25 X = 6 Y + 7$ are two regression equations. Find the mean values and coefficient of correlation.
(a) $\bar{X} = 1.57, \bar{Y} = 3.53, r = 0.431$ (b) $\bar{X} = 2.57, \bar{Y} = 9.53, r = 0.734$
(c) $\bar{X} = 4.57, \bar{Y} = 5.53, r = 0.532$ (d) $\bar{X} = 3.57, \bar{Y} = 8.53, r = 0.633$
19. The two regression equations between X and Y are $12 X - 15 Y + 99 = 0$ and $60 X - 27 Y = 321$. The standard deviation of X is 6.
Calculate :
(a) the mean value of X and Y ,
(b) the coefficient of correlation between X and Y , and
(c) standard deviation of Y .
(a) $\bar{X} = 13, \bar{Y} = 17, r = 0.6, \sigma_y = 8$ (b) $\bar{X} = 11, \bar{Y} = 20, r = 0.3, \sigma_y = 8$
(c) $\bar{X} = 15, \bar{Y} = 14, r = 0.7, \sigma_y = 8$ (d) $\bar{X} = 15, \bar{Y} = 19, r = 0.1, \sigma_y = 8$
20. For certain data of X and Y series, the equations of regression of y on X and X on Y are $6 Y = 5 X + 90$ and $15 X = 8 Y + 130$ respectively. The standard deviation of X is 4. Find out \bar{X}, \bar{Y}, r and σ_y .
(a) $\bar{X} = 30, \bar{Y} = 40, r = 0.667, \sigma_y = 5$ (b) $\bar{X} = 30, \bar{Y} = 42, r = 0.567, \sigma_y = 3$
(c) $\bar{X} = 30, \bar{Y} = 43, r = 0.467, \sigma_y = 2$ (d) $\bar{X} = 30, \bar{Y} = 43, r = 0.367, \sigma_y = 1$

Exercise - III

1. **Bivariate Data are the data collected for[SM]**
 (a) Two variables (b) More than two variables
 (c) Two variables at the same point of time (d) Two variables at different points of time
2. **For a bivariate frequency table having (p + q) classification the total number of cells is [SM]**
 (a) p (b) p + q (c) q (d) pq
3. **Some of cell the frequencies in a bivariate frequency table may be[SM]**
 (a) Negative (b) Zero (c) a or b (d) none of these
4. **For a $p \times q$ bivariate frequency table , the maximum number of marginal distribution is [SM]**
 (a) p (b) p + q (c) 1 (d) 2
5. **For a $p \times q$ classification of bivariate data , the maximum number of conditional distributions is[SM]**
 (a) p (b) p + q (c) pq (d) p or q
6. **Correlation analysis aims at [SM]**
 (a) Predicting one variable for a given value of the other variable
 (b) Establishing relation between two variables
 (c) Measuring the extent of relation between two variables
 (d) Both (b) and (c)
7. **When high values of one variable are associated with high values of the other & low values of one variable are associated with low values of another, then they are said to be[SM]**
 (a) positively correlated (b) directly correlated (c) both (d) none
8. **If high values of one tend to low values of the other, they are said to be[SM]**
 (a) negatively correlated (b) inversely correlated (c) both (d) none
9. **Correlation coefficient between two variables is a measure of their linear relationship .[SM]**
 (a) true (b) false (c) both (d) none
10. **Correlation coefficient is dependent of the choice of both origin & the scale of observations.[SM]**
 (a) True (b) false (c) both (d) none
11. **Correlation coefficient is a pure number.[SM]**
 (a) true (b) false (c) both (d) none
12. **Correlation coefficient is _____ of the units of measurement.[SM]**
 (a) dependent (b) independent (c) both (d) none
13. **The value of correlation coefficient lies between[SM]**
 (a) -1 and +1 (b) -1 and 0
 (c) 0 and 1 Inclusive of these two values (d) none.
14. **Correlation coefficient can be found out by[SM]**
 (a) Scatter Diagram (b) Rank Method (c) both (d) none.

15. Covariance measures _____ variations of two variables.[SM]
(a) joint (b) single (c) both (d) none
16. In calculating the Karl Pearson's coefficient of correlation it is necessary that the data should be of numerical measurements. The statement is[SM]
(a) valid (b) not valid (c) both (d) none
17. Rank correlation coefficient lies between[SM]
(a) 0 to 1 (b) -1 to +1 inclusive of these value
(c) -1 to 0 (d) both
18. A coefficient near +1 indicates tendency for the larger values of one variable to be associated with the larger values of the other.[SM]
(a) true (b) false (c) both (d) none
19. In rank correlation coefficient the association need not be linear.[SM]
(a) true (b) false (c) both (d) none
20. In rank correlation coefficient only an ascending/descending relationship is required.[SM]
(a) false (b) true (c) both (d) none
21. Great advantage of _____ is that it can be used to rank attributes which can not be expressed by way of numerical value .[SM]
(a) concurrent correlation (b) regression (c) rank correlation (d) none
22. The sum of the difference of rank is.[SM]
(a) 1 (b) -1 (c) 0 (d) none.
23. Correlation methods are used to study the relationship between two time series of data which are recorded annually, monthly, weekly, daily and so on.[SM]
(a) True (b) false (c) both (d) none
24. Age of Applicants for life insurance and the premium of insurance - correlation is[SM]
(a) positive (b) negative (c) zero (d) none
25. "Unemployment index and the purchasing power of the common man" _____ Correlation is[SM]
(a) positive (b) negative (c) zero (d) none
26. Production of pig iron and soot content in Durgapur - Correlations are [SM]
(a) positive (b) negative (c) zero (d) none
27. "Demand for goods and their prices under normal times" _____ Correlation is[SM]
(a) positive (b) negative (c) zero (d) none
28. _____ is a relative measure of association between two or more variables.[SM]
(a) Coefficient of correlation (b) Coefficient of regression (c) both (d) none
29. The square of coefficient of correlation 'r' is called the coefficient of[SM]
(a) determination (b) regression (c) both (d) none

30. Simple correlation is called[SM]
(a) linear correlation (b) nonlinear correlation (c) both (d) none
31. A scatter diagram indicates the type of correlation between two variables.[SM]
(a) true (b) false (c) both (d) none
32. If the pattern of points (or dots) on the scatter diagram shows a linear path diagonally across the graph paper from the bottom left- hand corner to the top right, correlation will be[SM]
(a) negative (b) zero (c) positive (d) none
33. The correlation coefficient being +1 if the slope of the straight line in a scatter diagram is[SM]
(a) positive (b) negative (c) zero (d) none
34. The correlation coefficient being -1 if the slope of the straight line in a scatter diagram is[SM]
(a) positive (b) negative (c) zero (d) none
35. The more scattered the points are around a straight line in a scattered diagram the _____ is the correlation coefficient.[SM]
(a) zero (b) more (c) less (d) none
36. If the values of y are not affected by changes in the values of x, the variables are said to be[SM]
(a) correlated (b) uncorrelated (c) both (d) zero
37. If the amount of change in one variable tends to bear a constant ratio to the amount of change in the other variable, then correlation is said to be[SM]
(a) non linear (b) linear (c) both (d) none
38. Variance may be positive, negative or zero.[SM]
(a) true (b) false (c) both (d) none
39. Covariance may be positive, negative or zero.[SM]
(a) true (b) false (c) both (d) none
40. Correlation coefficient between x and y = correlation coefficient between u and v[SM]
(a) true (b) false (c) both (d) none
41. In case ' The ages of husbands and wives' _____ correlation is[SM]
(a) positive (b) negative (c) zero (d) none
42. In case 'Insurance companies' profits and the no of claims they have to pay "[SM]
(a) positive correlation (b) negative correlation (c) no correlation (d) none
43. In case 'Years of education and income' [SM]
(a) positive correlation (b) negative correlation (c) no correlation (d) none
44. In case 'Amount of rainfall and yield of crop' [SM]
(a) positive correlation (b) negative correlation (c) no correlation (d) none
45. For calculation of correlation coefficient, a change of origin is[SM]
(a) not possible (b) possible (c) both (d) none

46. In case of employed persons 'Age and income' correlation is[SM]
(a) positive (b) negative (c) zero (d) none
47. In case 'Speed of an automobile and the distance required to stop the car often applying brakes' - correlation is[SM]
(a) positive (b) negative (c) zero (d) none
48. In case 'Sale of woolen garments and day temperature' correlation is[SM]
(a) positive (b) negative (c) zero (d) none
49. In case 'Sale of cold drinks and day temperature' _____ correlation is
(a) positive (b) negative (c) zero (d) none
50. In case of 'Production and price per unit' - correlation is[SM]
(a) positive (b) negative (c) zero (d) none
51. Co-variance measures the joint variations of two variables.[SM]
(a) true (b) false (c) both (d) none
52. The minimum value of correlation coefficient is[SM]
(a) 0 (b) -2 (c) 1 (d) -1
53. The maximum value of correlation coefficient is[SM]
(a) 0 (b) 2 (c) 1 (d) -1
54. In Method of Concurrent Deviations, only the directions of change (Positive direction / Negative direction) in the variables are taken into account for calculation of[SM]
(a) coefficient of S.D (b) coefficient of regression. (c) coefficient of correlation (d) none
55. A coefficient near +1 indicates tendency for the larger values of one variable to be associated with the larger values of the other.
(a) True (b) False (c) Both (a) and (b) (d) None of these
56. If the value of correlation coefficient is positive, then the points in a scatter diagram tend to cluster:
(a) From lower left corner to upper right corner.
(b) From lower right corner to upper left corner.
(c) From lower left corner to lower right corner.
(d) From lower right corner to upper right corner.
57. Product moment correlation coefficient may be defined as the ratio of:
(a) The covariance between the variables to the product of the variance of them.
(b) The product of standard deviations of the two variables to the covariance between them.
(c) The covariance between the variables to the product of their standard deviations.
(d) Either (b) or (c).

58. **The coefficient of correlation between two variables**
 (a) Is expressed as the product of units of the two variables.
 (b) Can have any unit.
 (c) Is a unit free measure.
 (d) None of these
59. **In case the correlation coefficient between two variables is 1, then the relationship between the two variables would be:**
 (a) $y = a + bx, b < 0$ (b) $y = a + bx, b > 0$
 (c) $y = a + bx$ (d) $y = a + bx$, both a and b being positive.
60. **For finding correlation between two attributes, we consider:**
 (a) Scatter diagram (b) Pearson's correlation coefficient
 (c) Spearman's rank correlation coefficient (d) Coefficient of concurrent deviation.
61. **If the values of y are not affected by changes in the values of x, the variables are said to be:**
 (a) correlated (b) Uncorrelated (c) Both (a) and (b) (d) zero.
62. **Whatever may be the value of r, positive or negative, its square will be:**
 (a) Negative only (b) Positive only (c) zero only (d) none of these
63. **The covariance between two variables is:**
 (a) Strictly negative (b) Strictly positive (c) always 0 (d) Either positive or negative or zero
64. **A scatter diagram indicates the type of correlation between two variables.**
 (a) True (b) False (c) Both (a) and (b) (d) None of these
65. **Scatter diagram helps us to:**
 (a) Find the direction of correlation between two variables.
 (b) Obtain the mathematical relationship between two variables.
 (c) Compute the extent of correlation between two variables.
 (d) Both (a) and (c).
66. **If the plotted points in a scatter diagram lie from upper left to lower right, then the correlation is:**
 (a) zero (b) Positive (c) negative (d) none fo these
67. **If the pattern of points (or dots) on the scatter diagram shows a linear path diagonally across the graph paper from the bottom left-hand corner to the top right, correlation will be:**
 (a) negative (b) zero (c) positive (d) none of these
68. **If the plotted points in a scatter diagram are evenly distributed, then the correlation is:**
 (a) zero (b) positive (c) negative (d) (a) or (b)
69. **If all the plotted points in a scatter diagram lie on a single line, then the correlation is:**
 (a) perfect positive (b) Perfect negative (c) both (a) and (b) (d) Either (a) or (b)

70. **When $r = 1$, all the points in a scatter diagram would lie:**
(a) on a straight line directed from lower left to upper right.
(b) on a straight line.
(c) on a straight line directed from upper left to lower right.
(d) both (a) and (b)
71. **The correlation coefficient being -1 if the slope of the straight line in a scatter diagram is:**
(a) positive (b) negative (c) zero (d) none of these
72. **The correlation coefficient being $+1$ if the slope of the straight line in a scatter diagram is:**
(a) Positive (b) negative (c) zero (d) none of these
73. **Maximum value of rank correlation coefficient is:**
(a) 0 (b) $+1$ (c) -1 (d) None of these
74. **In rank correlation the association need not be linear.**
(a) False (b) True (c) Both (a) and (b) (d) none of these
75. **When $r = 0$, then $\text{cov}(x,y) =$**
(a) 0 (b) -1 (c) $+1$ (d) none of these
76. **Correlation coefficient is.....of the units of measurement.**
(a) independent (b) dependent (c) Both (a) and (b) (d) none of these
77. **For covariance the number of variate values in the two given distributions should be:**
(a) unequal (b) equal
(c) any number in one and any number in the other. (d) none of these
78. **What is spurious correlation ?[SM]**
(a) It is a bad relation between two variables
(b) It is very low correlation between two variables
(c) It is the correlation between two variables having no causal relation
(d) It is a negative correlation
79. **Pearson's correlation coefficient is used for finding[SM]**
(a) Correlation for any type of relation (b) Correlation for linear relation only
(c) Correlation for curvilinear relation only (d) Both (b) and (c)
80. **Product moment correlation coefficient is considered for[SM]**
(a) Finding the nature of correlation (b) Finding the amount of correlation
(c) Both (a) and (b) (d) Either (a) and (b)
81. **..... is concerned with the measurement of the "strength of association" between variables.**
(a) correlation (b) regression (c) both (d) none
82. **Karl Pearson's coefficient is defined from**
(a) ungrouped data (Linear) (b) grouped data (Non-Linear) (c) both (d) none

83. The correlation between shoe-size and intelligence is [SM]
(a) Zero (b) Positive (c) Negative (d) None of these
84. If there is a perfect disagreement between the marks in Geography and Statistics, then what would be the value of rank correlation coefficient ?[SM]
(a) Any value (b) Only 1 (c) Only -1 (d) (b) or (c)
85. A relationship $r^2 = 1 - \frac{500}{300}$ is not possible[SM]
(a) true (b) false (c) both (d) none
86. For finding the degree of agreement about beauty between two Judges in a Beauty Contest, we use
(a) Scatter diagram (b) Coefficient of rank correlation
(c) Coefficient of correlation (d) Coefficient of concurrent deviation
87. Rank correlation coefficient lies between
(a) 0 to 1 (b) -1 to +1 inclusive of these values
(c) -1 to 0 (d) both
88. When we are not concerned with the magnitude of the two variables under discussion we consider:
(a) Rank correlation coefficient (b) Product moment correlation coefficient
(c) Coefficient of concurrent deviation (d) (a) or (b) but not (c).
89. What is the quickest method to find correlation between two variables ?
(a) Scatter diagram (b) Method of concurrent deviation
(c) Method of rank correlation (d) Method of product moment correlation
90. Scatter diagram is considered for measuring[SM]
(a) Linear relationship between two variables (b) Curvilinear relationship between two variables
(c) Neither (a) nor (b) (d) Both (a) and (b)

Exercise - IV

1. **Regression analysis is concerned with[SM]**
 - (a) Establishing a mathematical relationship between two variables
 - (b) Measuring the extent of association between two variables
 - (c) Predicting the value of the dependent variable for a given value of the independent variable
 - (d) Both (a) and (c)

2. _____ gives the mathematical relationship of the variables.[SM]
 - (a) correlation
 - (b) regression
 - (c) both
 - (d) none

3. **The lines of regression passes through the points, bearing no. of points on both sides.[SM]**
 - (a) equal
 - (b) unequal
 - (c) zero
 - (d) none

4. **Under Algebraic Method we get ----- linear equations .[SM]**
 - (a) one
 - (b) two
 - (c) three
 - (d) none

5. **In linear equations $Y = a + bX$ and $X = a + bY$ 'a' is the[SM]**
 - (a) intercept of the line
 - (b) slope
 - (c) both
 - (d) none

6. **In linear equations $Y = a + bX$ and $X = a + bY$ 'b' is the[SM]**
 - (a) intercept of the line
 - (b) slope of the line
 - (c) both
 - (d) none

7. **The regression equations $Y = a + bX$ and $X = a + bY$ are based on the method of[SM]**
 - (a) greatest squares
 - (b) least squares
 - (c) both
 - (d) none

8. **The line $Y = a + bX$ represents the regression equation of[SM]**
 - (a) Y on X
 - (b) X on Y
 - (c) both
 - (d) none

9. **The line $X = a + bY$ represents the regression equation of[SM]**
 - (a) Y on X
 - (b) X on Y
 - (c) both
 - (d) none

10. **Two regression lines always intersect at the means.[SM]**
 - (a) true
 - (b) false
 - (c) both
 - (d) none

11. **r, b_{xy}, b_{yx} all have _____ sign.[SM]**
 - (a) different
 - (b) same
 - (c) both
 - (d) none

12. **The regression coefficients are zero if r is equal to[SM]**
 - (a) 2
 - (b) -1
 - (c) 1
 - (d) 0

13. **The regression lines are identical if r is equal to[SM]**
 - (a) +1
 - (b) -1
 - (c) +2
 - (d) 0

14. The regression lines are perpendicular to each other if r is equal to [SM]
 (a) 0 (b) +1 (c) -1 (d) +1
15. Feature of Least Square regression lines are :- The sum of the deviations at the Y's or the X's from their regression lines are zero. [SM]
 (a) true (b) false (c) both (d) none
16. The coefficient of determination is defined by the formula [SM]
 (a) $r^2 = 1 - \frac{\text{unexplained variance}}{\text{total variance}}$ (b) $r^2 = \frac{\text{explained variance}}{\text{total variance}}$ (c) both (d) none
17. The relation $r_{xy} = \text{cov}(x,y) / \sigma_x \sigma_y$ is [SM]
 (a) true (b) false (c) both (d) none
18. Two regression lines coincide when
 (a) $r = 0$ (b) $r = 2$ (c) $r = +1$ (d) none
19. Neither y nor x can be estimated by a linear function of the other variable when r is equal to [SM]
 (a) +1 (b) -1 (c) 0 (d) none
20. When $r = 0$ then $\text{cov}(x,y)$ is equal to [SM]
 (a) +1 (b) -1 (c) 0 (d) none
21. b_{xy} is called regression coefficient of [SM]
 (a) x on y (b) y on x (c) both (d) none
22. b_{yx} is called regression coefficient of [SM]
 (a) x on y (b) y on x (c) both (d) none
23. The slopes of the regression line of y on x is [SM]
 (a) b_{yx} (b) b_{xy} (c) b_{xx} (d) b_{yy}
24. The slopes of the regression line of x on y is [SM]
 (a) b_{yx} (b) b_{xy} (c) $1/b_{xy}$ (d) $1/b_{yx}$
25. The angle between the regression lines depends on [SM]
 (a) correlation coefficient (b) regression coefficient (c) both (d) none
26. If x and y satisfy the relationship $y = -5 + 7x$, the value of r is [SM]
 (a) 0 (b) -1 (c) +1 (d) none
27. If b_{yx} and b_{xy} are negative, r is [SM]
 (a) positive (b) negative (c) zero (d) none
28. Correlation coefficient r lie between the regression coefficients b_{yx} and b_{xy} [SM]
 (a) true (b) false (c) both (d) none

29. Since the correlation coefficient r cannot be greater than 1 numerically, the product of the regression must[SM]
 (a) not exceed 1 (b) exceed 1 (c) be zero (d) none
30. The correlation coefficient r is the _____ of the two regression coefficients b_{yx} and b_{xy} [SM]
 (a) A.M (b) G.M (c) H.M (d) none
31. Which is true?[SM]
 (a) $b_{yx} = r \frac{\sigma_x}{\sigma_y}$ (b) $b_{yx} = r \frac{\sigma_y}{\sigma_x}$ (c) $b_{yx} = r \frac{\sigma_{xy}}{\sigma_x}$ (d) $b_{yx} = r \frac{\sigma_{yy}}{\sigma_x}$
32. The correlation coefficient lies between[SM]
 (a) -1 and +1 inclusive of these two value (b) 0 and + 1 (c) -1 and (d) none
33. r_{12} is the correlation coefficient between[SM]
 (a) x_1 and x_2 (b) x_2 and x_1 (c) x_1 and x_3 (d) x_2 and x_3
34. r_{12} is the same as r_{21} [SM]
 (a) true (b) false (c) both (d) none
35. When $r = 0$, the regression coefficients are[SM]
 (a) 0 (b) 1 (c) -1 (d) none
36. If these are two variables x and y , then the number of regression equations could be:
 (a) 2 (b) 1 (c) Any number (d) 3
37. The method applied for deriving the regression equations is known as:
 (a) Concurrent deviation (b) Least squares (c) Product moment (d) Normal equation.
38. What are the limits of the two regression coefficients?
 (a) Must be positive (b) No limit.
 (c) One positive and the other negative. (d) Both positive or both negative
39. Two regression lines coincide when:
 (a) $r = 0$ (b) $r = 2$ (c) $r = + 1$ (d) None of these
40. The two lines of regression become identical when:
 (a) $r = 1$ (b) $r = - 1$ (c) $r = 0$ (d) (a) or (b)
41. The difference between the observed value and the estimated value in regression analysis is known as:
 (a) residue (b) error (c) deviation (d) Either (a) or (b)
42. The regression lines are perpendicular to each other if r is equal to:
 (a) 0 (b) +1 (c) -1 (d) 1
43. Two regression lines always intersect at the means.
 (a) True (b) False (c) Both (a) and (b) (d) None of these

44. **The regression line of y on x is derived by**
 (a) The minimisation of vertical distances in the diagram.
 (b) The minimisation of horizontal distances in the scatter diagram.
 (c) both (a) and (b)
 (d) None of these
45. **The line of regression x on y is given by:**
 (a) $x - \bar{x} = r(\sigma_x / \sigma_y)(y - \bar{y})$
 (b) $y - \bar{y} = r(\sigma_x / \sigma_y)(x - \bar{x})$
 (c) $(x - \bar{x}) = r(\sigma_y / \sigma_x)(y - \bar{y})$
 (d) None
46. **The line of regression of y on x is given by:**
 (a) $(x - \bar{x}) = r(\sigma_y / \sigma_x)(y - \bar{y})$
 (b) $(y - \bar{y}) = r(\sigma_y / \sigma_x)(x - \bar{x})$
 (c) $(y - \bar{y}) = r(\sigma_x / \sigma_y)(x - \bar{x})$
 (d) None of these
47. **Coefficient of correlation is:**
 (a) G.M. of the coefficient of regression.
 (b) A.M. of the coefficient of regression.
 (c) H.M. of the coefficient of regression.
 (d) Product of G.M. and A.M. of the regression coefficients.
48. **If the two lines of regression for a bivariate data coincide, then:**
 (a) The two variates are independent.
 (b) There is a perfect correlation between the two variates.
 (c) There is not a perfect positive correlation between the two variables.
 (d) None of these
49. **The statistical method, which helps us to estimate or predict the unknown value of one variable from the known value of the related variables, is called:**
 (a) Correlation (b) Scatter diagram (c) Regression (d) Dispersion.
50. **Which of the following statements is not true**
 (a) The regression lines are drawn by applying the principal of least squares.
 (b) The correlation : $r = b_{xy} \times b_{yx}$
 (c) b_{xy} and b_{yx} must be both either positive or negative.
 (d) The sign of r is the same as that of b_{xy} or b_{yx} .
51. **Slope of regression line of Y on X is:**
 (a) $\frac{1}{b_{yx}}$ (b) b_{yx} (c) $\frac{Cov(X,Y)}{\sigma_x}$ (d) None of these
52. **If the angle between two lines of regression is 90° , then it represents**
 (a) perfect positive correlation (b) Perfect negative correlation
 (c) no linear correlation (d) None of these

53. Since Blood Pressure of a person depends on age , we need consider
- (a) The regression equation of Blood Presure on age
 - (b) The regression equation of age on Blood Presure
 - (c) Both (a) and (b)
 - (d) Either (a) and (b)

Question Bank

1. **The coefficient of correlation r between x and y when: (N-06)**
 $Cov(x, y) = -16.5, Var(x) = 2.89, Var(y) = 100$ is:
 (a) -0.97 (b) 0.97 (c) 0.89 (d) -0.89

2. **Take 200 and 150 respectively as the assumed mean for x and y series of 11 values, then $dx = x - 200, dy = Y - 150, \sum dx = 13, \sum dx^2 = 2667, \sum dy = 42, \sum dy^2 = 6964, \sum dx dy = 3943.$ The value of r is: (N-06)**
 (a) 0.77 (b) 0.9 (c) 0.92 (d) 0.82

3. **The coefficient of correlation between x and y is 0.6 U and V are two variables defined as $U = \frac{x-3}{2}, V = \frac{y-2}{3}$ then the coefficient of correlation between U, V and is: (M-07)**
 (a) 0.6 (b) 0.4 (c) 0.8 (d) 1

4. **If the covariance between two variables is 20 and the variance of one of the variables is 16, what would be the variance of the other variable? (A-07)**
 (a) More than 10 (b) More than 100 (c) More than 1.25 (d) Less than 10

5. **Assume 69 and 112 as the mean values for X and Y respectively. (N-07)**
 $\sum dx = 47, \sum dx^2 = 1475, \sum dy = 108, \sum dy^2 = 3468, \sum dx dy = 2116$ and $N = 8$
Where $dx = X - 69, dy = Y - 112.$ Then the value of r is;
 (a) 0.95 (b) 0.65 (c) 0.75 (d) 0.85

6. **The coefficient of correlation between x and y series from the following data: (J-08)**

	X Series	Y Series
No. of pairs of observations	15	15
Arithmetic Mean	25	18
Standard Deviation	3.01	3.03
Sum of the squares of deviation from mean	136	138

Sum of the product of the deviations of x and y series from their respective means = 122, is:
 (a) 0.89 (b) 0.99 (c) 0.69 (d) 0.91

7. The coefficient of correlation is significant if: (D-08)
(a) $r > 5 P. E$ (b) $r \leq 6 P.E$ (c) $r \geq 6 P. E$ (d) $r = 6 P. E$
8. Correlation coefficient between X and Y will be negative when:- (D-09)
(a) X and Y are decreasing (b) X is increasing, Y is decreasing
(c) X and Y are increasing (d) None of these
9. If 'r' is the simple correlation coefficient, the quantity r^2 is known as: (J-10)
(a) Coefficient of determination (b) Coefficient of Non-determination
(c) Coefficient of alienation (d) None of the above
10. If the correlation coefficient between x and y is r, then between $U = \frac{x-5}{10}$, $V = \frac{y-7}{2}$ is (J-10)
(a) r (b) -r (c) $(r-5)/2$ (d) $(r-7)/10$
11. If the sum of the product of deviations of x and y series from their means is zero, then the coefficient of correlation will be (D-10)
(a) 1 (b) -1 (c) 0 (d) None of these
12. The covariance between two variables X and Y is 8.4 and their variances are 25 and 36 respectively. Calculate Karl Pearson's coefficient of correlation between them. (J-11)
(a) 0.82 (b) 0.28 (c) 0.01 (d) 0.09
13. If the covariance between variables X and Y is 25 and variance of X and Y are respectively 36 and 25, then the coefficient of correlation is (J-12)
(a) 0.409 (b) 0.416 (c) 0.833 (d) 0.0277
14. If the coefficient of correlation between x and y series is -0.38. The linear relations between x & u and y & v are $3x+5u=3$ and $-8y-7v=44$ respectively. Then the coefficient of correlation between u & v is: (D-12)
(a) 0.38 (b) -0.38 (c) 0.40 (d) None
15. The coefficient of correlation between two variables x and y is 0.28. There covariance is 7.6. If the variance of x is 9, then the standard deviation of y is: (J-13)
(a) 8.048 (b) 9.048 (c) 10.048 (d) 11.048
16. Price and Demand is example of (D-13)
(a) No correlation (b) Positive correlation (c) Negative correlation (d) none of these

17. Determine the coefficient of correlation between x and y series: (D-13)
 Sum of product deviation of x and y series from mean = 122

	X-Series	Y-Series
Number of items	15	15
Arithmetic mean	25	18
Sum of Square of deviation of mean	136	138

- (a) -0.89 (b) 0.89 (c) 0.69 (d) -0.69
18. If the value of correlation coefficient between x & y is 1. then the value of correlation coefficient between $x - 2$ and $\frac{-y}{2} + 1$ is (D-14)
 (a) 1 (b) -1 (c) -1/2 (d) 1/2
19. If the sum of squares of the rank difference in mathematics and physics marks of 10 students is 22, then the coefficient of rank correlation is: (F-07)
 (a) 0.267 (b) 0.867 (c) 0.92 (d) None
20. For the following data, the coefficient of rank correlation is: (M-07)
- | | | | | | |
|-------------------|---|---|---|---|---|
| Rank in Botany: | 1 | 2 | 3 | 4 | 5 |
| Rank in Chemistry | 2 | 3 | 1 | 5 | 4 |
- (a) 0.93 (b) 0.4 (c) 0.6 (d) None
21. In rank correlation, the association need not be linear (Mathematical): (N-07)
 (a) True (b) False (c) Partly True (d) Partly False
22. The coefficient of rank correlation of marks obtained by 10 students, in English and Economics was found to be 0.5. It was later discovered that the difference in ranks in the two subjects obtained by one students was wrongly taken as 3 instead of 7. The correct coefficient of rank correlation is: (F-08)
 (a) 0.32 (b) 0.26 (c) 0.49 (d) 0.93
23. If the sum of square of differences of rank is 50 and number of items is 8 then what is the value of rank correlation coefficient. (D-08)
 (a) 0.59 (b) 0.40 (c) 0.36 (d) 0.63
24. Ranks of two _____ characteristics by two judges are in reverse order then find the value of Spearman rank correlation co-efficient. (J-09)
 (a) -1 (b) 0 (c) 1 (d) 0.75

25. The ranks of five participants given by two judges are

(D-10)

Participants

	A	B	C	D	E
Judge 1	1	2	3	4	5
Judge 2	5	4	3	2	1

Rank correlation coefficient between ranks will be :

- (a) 1 (b) 0 (c) -1 (d) $\frac{1}{2}$

26. Three competitors in a contest are ranked by two judges in the order 1,2,3 and 2,3,1 respectively. Calculate the Spearman's rank correlation coefficient. (J-11)

- (a) -0.5 (b) -0.8 (c) 0.5 (d) 0.8

27. When each individual gets the exactly opposite rank by the two judges, then the rank correlation will be _____. (J-11)

- (a) 0 (b) -1 (c) +1 (d) $\frac{1}{2}$.

28. For some bivariate data, the following results were obtained for the two variables x and y: $\bar{x} = 53.2$, $\bar{y} = 27.9$, $b_{yx} = -1.5$, $b_{xy} = -0.2$. The most probable value of y when x = 60 is: (N-06)

- (a) 15.6 (b) 13.4 (c) 19.7 (d) 17.7

29. The following data is given, based on 450 students for marks in Statistics and Economics at a certain examination: (M-07)

Mean marks in Statistics = 40

Mean marks in Economics = 48

S.D. of marks (Statistics) = 12

Variance of marks (Economics) = 256

Sum of the products of deviations of marks from their respective mean = 42075

The average marks in economics of candidates who obtained 50 marks in Statistics is:

- (a) 45 (b) 54.5 (c) 54 (d) 47.5

30. The method applied of deriving regression equations is known as:

(F-08)

- (a) Concurrent deviation (b) Product moment (c) Least Squares (d) Normal equation

31. Given: $\bar{x} = 16, \sigma_x = 4.8$

$\bar{y} = 20, \sigma_y = 9.6$

The coefficient of correlation between x and y is 0.6. What will be the regression coefficient of 'x' on 'y'?

- (a) 0.03 (b) 0.3 (c) 0.2 (d) 0.05

32. If Y is dependent variable and X is Independent variable and the S.D of X and Y are 5 and 8 respectively and Co-efficient of co-relation between X and Y is 0.8. Find the Regression coefficient of Y on X : (D-11)

- (a) 0.78 (b) 1.28 (c) 6.8 (d) 0.32

33. Two random variables have the regression lines $3x + 2y = 26$ and $6x + y = 31$. The coefficient of correlation between x and y is : (F-07)

- (a) -0.25 (b) 0.5 (c) -0.5 (d) 0.25

34. The lines of regression are as follows; $5x - 145 = -10y$; $14y - 208 = -8x$ (N-07)

The mean value (\bar{x}, \bar{y}) is:

- (a) (12, 5) (b) (5,7) (c) (7,12) (d) (5,12)

35. If the lines of regression in a bivariate distribution are given by $x+2y = 5$ and $2x +3y = 8$, then the coefficient of correlation is: (J-08)

- (a) 0.866 (b) -0.666 (c) 0.667 (d) -0.866

36. If the correlation coefficient between two variables is 1, then the two lines of regressions are: (J-08)

- (a) Parallel (b) At right angles (c) Coincident (d) None of these

37. Given the regression equations as $3x + y = 13$ and $2x + 5y = 20$. Find regression equation of y on x. (D-08)

- (a) $3x + y = 13$ (b) $2x + y = 20$ (c) $3x + 5y = 13$ (d) $2x + 5y = 20$

38. The two regression equations are: (J-09)

$2x + 3y + 18 = 0$

$X + 2y - 25 = 0$

Find the value of y if x = 9

- (a) -8 (b) 8 (c) -12 (d) 0

39. The correlation coefficient between x and y is $-\frac{1}{2}$. The value of $b_{xy} = -1/8$ Find b_{yx} . (J-09)
- (a) -2 (b) -4 (c) 0 (d) 2
40. Which of the following regression equations represent regression line of Y on X : (D-09)
- $7x + 2y + 15 = 0$, $2x + 5y + 10 = 0$
- (a) $7x + 2y + 15 = 0$ (b) $2x + 5y + 10 = 0$ (c) Both (a) and (b) (d) None of these
41. The two regression lines are $7x - 3y - 18 = 0$ and $4x - y - 11 = 0$ Find the values of b_{yx} and b_{xy} (D-09)
- (a) $7/3$, $1/4$ (b) $-7/3$, $-1/4$ (c) $-3/7$, $-1/4$ (d) None of these
42. _____ of the regression coefficients is greater than the correlation coefficient (J-10)
- (a) Combined mean (b) Harmonic mean (c) Geometric mean (d) Arithmetic mean
43. If two lines of regression are $x + 2y - 5 = 0$ and $2x + 3y - 8 = 0$. (J-10)
- The regression line of y on x is
- (a) $x + 2y - 5 = 0$ (b) $2x + 3y - 8 = 0$ (c) Any of the two lines (d) None of the two lines
44. Regression coefficient are _____. (D-10)
- (a) dependent of change of origin and of scale
- (b) Independent of both change of origin and of scale.
- (c) Dependent of change of origin but not of scale.
- (d) Independent of change of origin but not of scale
45. If the two lines of regression are $x + 2y - 5 = 0$ and $2x + 3y - 8 = 0$ (D-10)
- The regression line of y on x is
- (a) $x + 2y - 5 = 0$ (b) $2x + 3y - 8 = 0$ (c) Any of the two line (d) None of the two lines.
46. For a bivariate data, two lines of regression are $40x - 18y = 214$ and $8x - 10y + 66 = 0$, then find the values of \bar{x} and \bar{y} (J-11)
- (a) 17 and 13 (b) 13 and 17 (c) 13 and -17 (d) -13 and 17

47. Out of the following which one affects the regression co-efficient. (D-11)
- (a) Change of Origin Only (b) Change of scale Only
 (c) Change of scale & origin both (d) Neither Change of origin nor change of scale
48. For a bivariate data, the lines of regression of Y on X, and of X on Y are respectively $2.5Y - X = 35$ and $10X - Y = 70$, then the Correlation coefficient r is equal to: (D-11)
- (a) 0.2 (b) - 0.2 (c) 0.5 (d) - 0.5
49. If one of regression coefficient is _____ unity, the other must be _____ unity. (D-11)
- (a) more than, more then (b) Less than, Less then (c) more than, less than (d) Positive, Negative
50. If the regression lines are $8x - 10y + 66 = 0$ and $40x - 18y = 214$, the correlation coefficient between 'X' and 'y' is: (J-12)
- (a) 1 (b) 0.6 (c) -0.6 (d) -1
51. The coefficient of correlation between two variables x and y is the simple _____ of the two regression coefficients. (J-12)
- (a) Arithmetic Mean (b) Geometric Mean (c) Harmonic Mean (d) None of the above.
52. If 2 variables are uncorrelated, their regression lines are: (J-12)
- (a) Parallel (b) Perpendicular (c) Coincident (d) Inclined at 45 degrees.
53. If \bar{x}, \bar{y} denote the arithmetic means, σ_x, σ_y denotes the standard deviations. b_{xy}, b_{yx} denote the regression coefficients of the variables 'x' and 'y' respectively, then the point of intersection of regression lines x on y & y on x is _____. (J-12)
- (a) (\bar{x}, \bar{y}) (b) (σ_x, σ_y) (c) (b_{xy}, b_{yx}) (d) (σ_x^2, σ_y^2)
54. If $y = 18x + 5$ is the regression line of x on y then the value of b_{xy} is _____. (D-12)
- (a) $\frac{5}{18}$ (b) 18 (c) 5 (d) $\frac{1}{18}$
55. For certain x and y series which are correlated, the two lines of regression are $5x - 6y + 9 = 0$, $15x - 8y + 130 = 0$. The correlation coefficient is: (D-12)
- (a) $\frac{4}{5}$ (b) $\frac{3}{4}$ (c) $\frac{2}{3}$ (d) $\frac{1}{2}$
56. If the regression equations are $8x - 3y + 50 = 0$ and $14x - 7y - 60 = 0$ and standard deviation of y is 1. The coefficient of correlation is = (J-13)
- (a) 2 (b) 1 (c) 0.87 (d) -0.87

57. Two variables x and y are related according to $4x+3y=7$. Then x and y are : (J-13)
 (a) Positively correlated (b) Negatively correlated (c) Correlation is zero (d) None of these
58. If mean of X and Y variables is 20 and 40 respectively and the regression coefficient Y on X is 1.608 then the regression line of Y on X is : (D-13)
 (a) $Y = 1.608 X + 7.84$ (b) $Y = 1.56 X + 4.84$ (c) $Y = 1.608 X + 4.84$ (d) $Y = 1.56 X + 7.84$
59. When the value of correlation coefficient is +1 or -1, then the two regression lines will..... (D-13)
 (a) have 30° angle between them (b) have 45° angle between them
 (c) coincide (d) be perpendicular to each other.
60. Two regression lines for a bivariate data are : $2x - 5y + 6 = 0$ and $5x - 4y + 3 = 0$. Then the coefficient of correlation should be: (J-14)
 (a) $\frac{-2\sqrt{2}}{5}$ (b) $\frac{2}{5}$ (c) $\frac{+2\sqrt{2}}{5}$ (d) $\frac{\sqrt{2}}{5}$
61. If the means of the two variables ' x ' and ' y ' are 3 and 1 respectively. Then the equation of two regression lines are _____. (J-14)
 (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) $5x + 7y + 22 = 0, 6x + 2y + 20 = 0$
62. The equation of two lines of regression for ' x ' and ' y ' and $64x = 24 + 45y$ are $5x = 22 + y$ then the value of regression coefficient of ' y ' on ' x ' will be _____. (J-14)
 (a) 5 (b) $1/5$ (c) $64/45$ (d) $45/64$
63. If the correlation coefficient between two variables is zero, then the lines of regression are: (D-14)
 (a) Parallel (b) Perpendicular (c) Coincide (d) None of these
64. The equations of two regression lines are $x + y = 6$ and $x + 2y = 10$, then the value of correlation coefficient between x and y is: (D-14)
 (a) $-1/2$ (b) $+1/2$ (c) $-1/\sqrt{2}$ (d) $+1/\sqrt{2}$
65. In case 'Insurance Companies' Profits and the no. of claims they have to pay" : [D-15]
 (a) Positive correlation (b) Negative correlation (c) No correlation (d) None of these
66. If the coefficient of correlation between X and Y variables is +0.90 then what will be the coefficient of determination? [J-16]
 (a) 0.30 (b) 0.81 (c) 0.94 (d) None of these
67. If $r = 0.6$, then the coefficient of determination is. [J-16]
 (a) 0.4 (b) -0.6 (c) 0.36 (d) 0.64

68. In a beauty contest there were 10 competitors. Rank of these candidates are assigned by two judges A and B. The sum of squares of differences of ranks is 44. The value of rank correlation is : [D-16]
- (a) 0.70 (b) 0.73 (c) 0.80 (d) 0.60
69. When $r = 1$, all the points in a scatter diagram would lie: [J-15]
- (a) On a straight line directed from lower left to upper right.
 (b) On a straight line.
 (c) On a straight line directed from upper left to lower right.
 (d) Both (a) and (b)
70. The two regression lines are [J-15]
- $16x - 20y + 132 = 0$ and $80x - 30y - 428 = 0$, the value of correlation coefficient is
- (a) 0.6 (b) -0.6 (c) 0.54 (d) 0.45
71. Two regression equations are as follows: [J-16]
- Regression equation of x on y : $5x - y = 22$
 Regression equation of y on x : $64x - 45y = 24$
 What will be the mean of x and y?
- (a) $\bar{x} = 8, \bar{y} = 6$ (b) $\bar{x} = 6, \bar{y} = 6$ (c) $\bar{x} = 6, \bar{y} = 8$ (d) $\bar{x} = 8, \bar{y} = 8$
72. The two lines of regression become identical when [J-16]
- (a) $r = 1$ (b) $r = -1$ (c) $r = 0$ (d) (a) or (b)
73. The two regression lines passing through [D-16]
- (a) Represent means (b) Represent S.Ds (c) (a) and (b) (d) None of these
74. Out of the following the one which effects the regression coefficient is [D-16]
- (a) Change of origin only (b) Change of scale only
 (c) Change of scale and origin both (d) Neither change in origin nor change of scale
75. Correlation between temperature and power consumption is [J-17]
- (a) Positive (b) Negative (c) Zero (d) None
76. Coefficient of correlation between X & Y is 0.6. If both X and Y are multiplied by -1. Then resultant coefficient of correlation is [J-17]
- (a) 0.6 (b) -0.6 (c) 1/0.6 (d) None

77. If the two regression lines are $x + y = 1$ and $x - y = 1$ then \bar{x} and \bar{y} are [J-17]
 (a) 1, 0 (b) 0, 1 (c) 1, 1 (d) None
78. If $r = 0.6$ then the coefficient of non-determination is: [D-17]
 (a) 0.4 (b) -0.6 (c) 0.36 (d) 0.64
79. If there is a constant increase in the series then the obtained graph is: [D-17]
 (a) Convex (b) Concave (c) Parabola (d) Straight Line from left to right
80. The correlation coefficient is the _____ of the two regression coefficients b_{yx} and b_{xy} : [D-17]
 (a) AM (b) GM (c) HM (d) None of these
81. Regression coefficients are independent of _____: [D-17]
 (a) Change of origin (b) Change of scale (c) Both (a) and (b) (d) None of these
82. $5y = 9x - 22$ & $20x = 9y + 350$ are two regression lines. Find the correlation coefficient between x & y : [D-17]
 (a) 0.9 (b) 0.1 (c) -0.9 (d) -0.1
83. If $r = 0.58$, correlation coefficient of $u = -5x + 3$ and $v = y + 2$ is _____: (J-18)
 (a) 0.58 (b) -0.58 (c) 0.62 (d) None the above
84. If the sum of squares of deviations of ranks of 8 students is 50 then the rank correlation coefficient is _____: (J-18)
 (a) 0.40 (b) 0.45 (c) 0.5 (d) 0.8
85. Regression lines are parallel then $r =$: (J-18)
 (a) ± 1 (b) $-1/2$ (c) 0 (d) None the above
86. The two line of regression intersect at the point (N-18)
 (a) Mean (b) Mode (c) Median (d) None of these
87. If the two lines of regression are $x + 2y - 5 = 0$ and $2x + 3y - 8 = 0$, then the regression line of y on x is (N-18)
 (a) $x + 2y - 5 = 0$ (b) $2x + 3y - 8 = 0$ (c) $x + 2y = 0$ (d) $2x + 3y = 0$
88. If the two regression lines are $3X = Y$ and $8Y = 6X$, then the value of correlation coefficient is (N-18)
 (a) 0.5 (b) -0.5 (c) 0.75 (d) -0.80
89. The regression coefficient is independent of the change of (N-18)
 (a) Scale (b) Origin (c) Scale and Origin both (d) None of these
90. If the correlation coefficient between the variables X and Y is 0.5, then the correlation coefficient between the variables $2x-4$ and $3-2y$ is (N-18)
 (a) 1 (b) 0.5 (c) -0.5 (d) 0

91. Given the following series: [J-19]

X	10	13	12	15	8	15
Y	12	16	18	16	7	18

The rank correlation coefficient r =

(a) $1 - \frac{6\sum d^3 + \sum_{i=1}^2 m_i(m_i^3 - 1)}{12m(n^2 - 1)}$

(b) $1 - \frac{6 \left[\sum d_i^2 + \sum_{i=1}^3 \frac{m_i(m_i^2 - 1)}{12} \right]}{n(n^2 - 1)}$

(c) $1 - \frac{6\sum d^2 + \sum_{i=1}^2 m_i(m_i^2 - 1)}{n(n^2 - 1)}$

(d) $1 - 6\sum d^2 + \sum_{i=1}^3 \frac{m_i(m_i^2 - 1)}{12n(n^2 - 1)}$

92. Find the probable error if $r = \frac{2}{\sqrt{10}}$ and $n = 36$. [J-19]

- (a) 0.6745 (b) 0.06745 (c) 0.5287 (d) None

93. If the regression line of y on X is given by $Y = x + 2$ and Karl Pearson's coefficient of correlation is 0.5

then $\frac{\sigma_y^2}{\sigma_x^2} = \underline{\hspace{2cm}}$ [J-19]

- (a) 3 (b) 2 (c) 4 (d) None

94. A.M. of regression coefficient is [J-19]

- (a) Equal to r (b) Greater than or equal to r (c) Half of r (d) None of these

95. If the equation of the two regression line are $2x - 3y = 0$ and $4y - 5x = 8$ then the correlation coefficient between x and y is equal to [N-19]

- (a) $\sqrt{\frac{15}{8}}$ (b) $\sqrt{\frac{8}{15}}$ (c) $\sqrt{\frac{6}{15}}$ (d) $\sqrt{\frac{1}{15}}$

96. Find the correlation coefficient [N-19]

X 5 4 3 2 1

Y 1 2 3 4 5

- (a) 1 (b) -1 (c) 0 (d) None of these

97. If scatter diagram from a line move from lower to left to upper right corner then the correlation is.

[N-19]

- (a) Perfect Positive (b) Perfect negative
(c) Simple Positive (d) No correlation

98. Consider to regression line $3x + 2y = 26$, $6x + y = 31$ find correlation coefficient between x and y [N-19]

- (a) 0.5 (b) -0.5 (c) 0.25 (d) -0.25

99. If correlation coefficient between x and y is 0.5 then Find the correlation coefficient between $2x-3$ and $3-5y$ is [N-19]
- (a) 0.5 (b) -0.5 (c) 2.5 (d) -2.5
100. Scatter diagram does not help us to [N-20]
- (a) Identify whether variables are correlated or not
(b) find the type of correlation
(c) determine the linear or nonlinear type
(d) find the numerical value of the correlation coefficient
101. Which of the following is spurious correlation ? [N-20]
- (a) Negative correlation
(b) Correlation between 2 variables having no causal relation
(c) Bad relation between 2 variables
(d) Very low correlation between 2 variables
102. The _____ are used when one wants to visually examine the relationship between two variables. [N-20]
- (a) Bar graphs (b) pie charts (c) Line charts (d) Scatter plots
103. The covariance between two variable in [N-20]
- (a) Strictly positive (b) Strictly negative
(c) Always zero (d) Either positive or negative or zero
104. A partially legible working sheet for the calculation of rank correlation coefficient revealed that the coefficient of Rank correlation is $1/3$ and the sum of squared differences of ranks is 80. What is the number of observations ? [N-20]
- (a) 9 (b) 8 (c) 7 (d) 6

[CH.-18] [Correlation & Regression]

EXERCISE - I

1	b	7	c	13	c	19	a	25	d
2	c	8	a	14	d	20	d	26	a
3	b	9	d	15	a	21	b	27	c
4	b	10	b	16	d	22	d	28	a
5	c	11	a	17	b	23	c		
6	b	12	b	18	c	24	d		

EXERCISE - II

1	d	5	a	9	b	13	b	17	c
2	a	6	a	10	c	14	a	18	b
3	b	7	d	11	a	15	b	19	a
4	c	8	d	12	c	16	d	20	a

EXERCISE - III

1	c	23	a	45	b	67	c	89	b
2	d	24	a	46	a	68	a	90	d
3	b	25	b	47	a	69	d		
4	d	26	a	48	b	70	a		
5	b	27	b	49	a	71	b		
6	d	28	a	50	b	72	a		
7	c	29	a	51	a	73	b		
8	c	30	a	52	d	74	b		
9	a	31	a	53	c	75	a		
10	b	32	c	54	c	76	a		
11	a	33	a	55	a	77	b		
12	b	34	b	56	a	78	c		
13	a	35	c	57	c	79	b		
14	c	36	b	58	c	80	c		
15	a	37	b	59	b	81	a		
16	a	38	b	60	c	82	a		
17	b	39	a	61	b	83	a		
18	a	40	a	62	b	84	c		
19	a	41	a	63	d	85	a		
20	b	42	b	64	a	86	b		
21	c	43	a	65	a	87	b		
22	c	44	a	66	c	88	c		

EXERCISE-IV

1	d	12	d	23	a	34	a	45	a
2	b	13	a	24	b	35	a	46	b
3	d	14	a	25	a	36	a	47	a
4	b	15	a	26	c	37	b	48	b
5	a	16	c	27	b	38	d	49	c
6	b	17	a	28	a	39	c	50	b
7	b	18	c	29	a	40	d	51	b
8	a	19	c	30	b	41	d	52	c
9	b	20	c	31	b	42	a	53	a
10	a	21	a	32	a	43	a		
11	b	22	b	33	a	44	a		

QUESTION BANK

1	a	22	b	43	a	64	c	85	a
2	c	23	b	44	d	65	b	86	a
3	a	24	a	45	a	66	b	87	a
4	b	25	c	46	b	67	c	88	a
5	a	26	a	47	b	68	b	89	b
6	a	27	b	48	a	69	a	90	c
7	c	28	d	49	c	70	c	91	b
8	b	29	b	50	b	71	c	92	b
9	a	30	c	51	b	72	d	93	c
10	a	31	b	52	b	73	a	94	b
11	c	32	b	53	a	74	b	95	b
12	b	33	c	54	d	75	a	96	b
13	c	34	d	55	c	76	a	97	a
14	b	35	d	56	c	77	d	98	b
15	b	36	c	57	b	78	d	99	b
16	c	37	d	58	a	79	d	100	d
17	b	38	b	59	c	80	b	101	b
18	b	39	a	60	c	81	a	102	d
19	b	40	b	61	a	82	a	103	d
20	c	41	a	62	c	83	b	104	a
21	a	42	d	63	b	84	a		

