CORRELATION AND REGRESSION ASSIGNMENT

1. If *a*, *b*, *h*, *k* are constants, while *U* and *V* are $U = \frac{X-a}{h}$, $V = \frac{Y-b}{k}$, then

(a) *Cov* (*X*, *Y*) = *Cov* (*U*, *V*) (b) *Cov* (*X*, *Y*) = *hk Cov* (*U*, *V*)

(c) *Cov* (*X*, *Y*) = *ab Cov* (*U*, *V*) (d) *Cov* (*U*, *V*) = *hk Cov* (*X*, *Y*)

2. Let *X*, *Y* be two variables with correlation coefficient $\rho(X, Y)$ and variables *U*, *V* be related to *X*, *Y* by the relation *U* = 2*X*, *V* = 3*Y*, then $\rho(U, V)$ is equal to

(a) $\rho(X, Y)$ (b) $6\rho(X, Y)$ (c) $\sqrt{6}\rho(X, Y)$ (d) $\frac{3}{2}\rho(X, Y)$

3. If *X* and *Y* are two uncorrelated variables and if $u = X + Y$, $v = X - Y$, then *r*(*u*, *v*) is equal to

(a) $\frac{\sigma_x^2 + \sigma_y^2}{\sigma_x^2 - \sigma_y^2}$ (b) $\frac{\sigma_x^2 - \sigma_y^2}{\sigma_x^2 + \sigma_y^2}$ (c) $\frac{\sigma_x^2 + \sigma_y^2}{\sigma_x \sigma_y}$ (d) None of these

4. If $\bar{x} = \bar{y} = 0, \sum x_i y_i = 12, \sigma_x = 2, \sigma_y = 3$ and *n* = 10, then the coefficient of correlation is

(a) 0.4 (b) 0.3 (c) 0.2 (d) 0.1

5. Let *X* and *Y* be two variables with the same variance and *U* and *V* be two variables such that *U* = *X* + *Y*, *V* = *X* – *Y*. Then *Cov* (*U*, *V*) is equal to

(a) *Cov* (*X*, *Y*) (b) 0 (c) 1 (d) – 1

6. For the following data

	x	*y*
Mean	65	67
Standard deviation	5.0	2.5
Correlation coefficient	0.8	

Then the equation of line of regression of *y* on *x* is

(a) $y - 67 = \frac{2}{5}(x - 65)$ (b) $y - 67 = \frac{1}{5}(x - 65)$ (c) $x - 65 = \frac{2}{5}(y - 67)$ (d) $x - 65 = \frac{1}{5}(y - 67)$

7. If the lines of regression of *y* on *x* and that of *x* on *y* are $y = kx + 4$ and $x = 4y + 5$ respectively, then

(a) $k \le 0$ (b) $k \ge 0$ (c) $0 \le k \le \frac{1}{4}$ (d) $0 \le k \le 1$

8. From the following observations $\{(x, y)\} = \{(1,7),(4,5),(7,2),(10,6),(13,5)\}$. The line of regression of *y* on *x* is

(a) $7x + 30y - 187 = 0$ (b) $7x - 30y - 187 = 0$ (c) $7x - 30y + 187 = 0$ (d) None of these

9. If the variance of $x = 9$ and regression equations are $4x - 5y + 33 = 0$ and $20x - 9y - 10 = 0$, then the coefficient of correlation between *x* and *y* and the variance of *y* respectively are

(a) 0.6; 16 (b) 0.16; 16 (c) 0.3; 4 (d) 0.6; 4

10. If the two lines of regression are $x + 4y = 3$ and $3x + y = 15$, then value of *x* for $y = 3$ is

(a) 4 (b) –9 (c) – 4 (d) None of these

11. Which of the following two sets of regression lines are the true representative of the information from the bivariate population

I. $x + 4y = 15$ and $y + 3x = 12, \bar{x} = 3, \bar{y} = 3$ II. $3x + 4y = 9$ and $4x + y = 1, \bar{x} = -\frac{5}{10}, \bar{y} = \frac{30}{13}$

(a) Both I and II (b) II only (c) I only (d) None of these

12. Out of the two lines of regression given by $x + 2y = 4$ and $2x + 3y - 5 = 0$, the regression line of *x* on *y* is

(a) $x + 2y = 4$ (b) $2x + 3y - 5 = 0$

(c) The given lines cannot be the regression lines (d) $x + 2y = 0$

13. Regression of savings (*S*) of a family on income *Y* may be expressed as $S = a + \frac{Y}{m}$, where *a* and *m* are constants. In a random sample of 100 families the variance of savings is one-quarter of the variance of incomes and the correlation coefficient is found to be 0.4. The value of *m* is

(a) 2 (b) 5 (c) 8 (d) None of these

14. If there exists a linear statistical relationship between two variables *x* and *y*, then the regression coefficient of *y* on *x* is

(a) $\frac{cor(x, y)}{\sigma_x . \sigma_y}$ (b) $\frac{cor(x, y)}{\sigma_y^2}$ (c) $\frac{cor(x, y)}{\sigma_x^2}$

GRAVITY CLASSES

(d) $\frac{cor(x,y)}{\sigma_x}$, where σ_x, σ_y are standard deviations of x and y respectively.

15. If $ax + by + c = 0$ is a line of regression of y on x and $a_1x + b_1y + c_1 = 0$ that of x on y, then

(a) $a_1b \le ab_1$ (b) $aa_1 = bb_1$ (c) $ab_1 \le a_1b$ (d) None of these

16. Least square lines of regression give best possible estimates, when $\rho(X, Y)$ is

(a) <1 (b) > –1 (c) –1 or 1 (d) None of these

17. Which of the following statement is correct

(a) Correlation coefficient is the arithmetic mean of the regression coefficient

(b) Correlation coefficient is the geometric mean of the regression coefficient

(c) Correlation coefficient is the harmonic mean of the regression coefficient

(d) None of these

18. The relationship between the correlation coefficient r and the regression coefficients b_{xy} and b_{yx} is

(a) $r = \frac{1}{2}(b_{xy} + b_{yx})$ (b) $r = \sqrt{b_{xy}.b_{yx}}$ (c) $r = (b_{xy}b_{yx})^2$ (d) $r = b_{xy} + b_{yx}$

19. If the coefficient of correlation is positive, then the regression coefficients

(a) Both are positive (b) Both are negative (c) One is positive and another is negative

(d) None of these

20. If b_{yx} and b_{xy} are both positive (where b_{yx} and b_{xy} are regression coefficients), then

(a) $\frac{1}{b_{yx}} + \frac{1}{b_{xy}} < \frac{2}{r}$ (b) $\frac{1}{b_{yx}} + \frac{1}{b_{xy}} > \frac{2}{r}$ (c) $\frac{1}{b_{yx}} + \frac{1}{b_{xy}} < \frac{r}{2}$ (d) None of these

21. If x_1 and x_2 are regression coefficients and r is the coefficient of correlation, then

(a) $x_1 - x_2 > r$ (b) $x_1 + x_2 < r$ (c) $x_1 + x_2 \ge 2r$ (d) None of these

22. If one regression coefficient be unity, then the other will be

(a) Greater than unity (b) Greater than or equal to unity (c) Less than or equal to unity (d) None of these

23. If one regression coefficient be less than unity, then the other will be

(a) Less than unity (b) Equal to unity (c) Greater than unity (d) All of the above

24. If regression coefficient of y on x is 2, then the regression coefficient of x on y is

(a) 2 (b) $\frac{1}{2}$ (c) $\le \frac{1}{2}$ (d) None of these

25. The lines of regression of x on y estimates

(a) x for a given value of y (b) y for a given value of x (c) x from y and y from x (d) None of these

26. The statistical method which helps us to estimate or predict the unknown value of one variable from the known value of the related variable is called

(a) Correlation (b) Scatter diagram (c) Regression (d) Dispersion

27. The coefficient of correlation between two variables x and y is 0.8 while regression coefficient of y on x is 0.2. Then the regression coefficient of x on y is

(a) –3.2 (b) 3.2 (c) 4 (d) 0.16

28. If the lines of regression coincide, then the value of correlation coefficient is

(a) 0 (b) 1 (c) 0.5 (d) 0.33

29. Two lines of regression are $3x + 4y - 7 = 0$ and $4x + y - 5 = 0$. Then correlation coefficient between x and y is

(a) $\frac{\sqrt{3}}{4}$ (b) $-\frac{\sqrt{3}}{4}$ (c) $\frac{3}{16}$ (d) $-\frac{3}{16}$

30. If the two lines of regression are $4x + 3y + 7 = 0$ and $3x + 4y + 8 = 0$, then the means of x and y are

(a) $-\frac{4}{7}, -\frac{11}{7}$ (b) $-\frac{4}{7}, \frac{11}{7}$ (c) $\frac{4}{7}, -\frac{11}{7}$ (d) 4, 7

31. The two regression lines for a bivariate data are $x + y + 50 = 0$ and $2x + 3y + K = 0$. If $\bar{x} = 0$, then $\bar{y}$ is

(a) 50 (b) $K - 100$ (c) – 50 (d) $50 + K$

32. The two regression lines are $2x - 9y + 6 = 0$ and $x - 2y + 1 = 0$. What is the correlation coefficient between x and y

(a) $-\frac{2}{3}$ (b) $\frac{2}{3}$ (c) $\frac{4}{9}$ (d) None of these

33. If the two regression coefficient between x and y are 0.8 and 0.2, then the coefficient of correlation between them is

(a) 0.4 (b) 0.6 (c) 0.3 (d) 0.5

34. The two lines of regression are given by $3x + 2y = 26$ and $6x + y = 31$. The coefficient of correlation between *x* and *y* is

(a) $-\frac{1}{3}$ (b) $\frac{1}{3}$ (c) $-\frac{1}{2}$ (d) $\frac{1}{2}$

35. If the lines of regression be $x - y = 0$ and $4x - y - 3 = 0$ and $\sigma_x^2 = 1$, then the coefficient of correlation is

(a) – 0.5 (b) 0.5 (c) 1.0 (d) – 1.0

36. A student obtained two regression lines as $L_1 \equiv x - 5y + 7 = 0$ and $L_2 \equiv 3x + y - 8 = 0$. Then the regression line of *y* on *x* is

(a) L_1 (b) L_2 (c) Neither of the two (d) $x - 5y = 0$

37. If b_{yx} and b_{xy} are regression coefficients of *y* on *x* and *x* on *y* respectively, then which of the following statement is true

(a) $b_{xy} = 1.5, b_{yx} = 1.4$ (b) $b_{xy} = 1.5, b_{yx} = 0.9$ (c) $b_{xy} = 1.5, b_{yx} = 0.8$ (d) $b_{xy} = 1.5, b_{yx} = 0.6$

38. Angle between two lines of regression is given by

(a) $\tan^{-1}\left(\frac{b_{yx} - \frac{1}{b_{xy}}}{1 + \frac{b_{xy}}{b_{yx}}}\right)$ (b) $\tan^{-1}\left(\frac{b_{yx} - b_{xy} - 1}{b_{yx} + b_{xy}}\right)$ (c) $\tan^{-1}\left(\frac{b_{xy} - \frac{1}{b_{yx}}}{1 + \frac{b_{xy}}{b_{yx}}}\right)$ (d) $\tan^{-1}\left(\frac{b_{yx} - b_{xy}}{1 + b_{yx}.b_{xy}}\right)$

39. If acute angle between the two regression lines is θ, then

(a) $\sin\theta \geq 1 - r^2$ (b) $\tan\theta \geq 1 - r^2$ (c) $\sin\theta \leq 1 - r^2$ (d) $\tan\theta \leq 1 - r^2$

40. If the angle between the two lines of regression is 90°, then it represents

(a) Perfect correlation (b) Perfect negative correlation (c) No linear correlation (d) None of these

41. If $2x + y = 7$ and $x + 2y = 7$ are the two regression lines respectively, then the correlation co-efficient between *x* and *y* is

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

42. For a perfect correlation between the variables *x* and *y*, the line of regression is $ax + by + c = 0$ where $a, b, c > 0$; then $\rho(x, y) =$

(a) 0 (b) –1 (c) 1 (d) None of these

43. If two random variables *X* and *Y* of a bivariate distribution are connected by the relationship $3x + 2y = 4$, then correlation coefficient r_{xy} equals

(a) 1 (b) –1 (c) 2/3 (d) –2/3

44. Two variables *x* and *y* are related by the linear equation $ax + by + c = 0$ .The coefficient of correlation between the two is +1,if

(a) *a* is positive (b) *b* is positive (c) *a* and *b* both are positive (d) *a* and *b* are of opposite sign

45. If the two lines of regression are $5x + 3y = 55$ and $7x + y = 45$, then the correlation coefficient between *x* and *y* is

(a) +1 (b) –1 (c) $-\sqrt{\frac{5}{21}}$ (d) $-\sqrt{\frac{21}{5}}$

46. The error of prediction of *x* from the required line of regression is given by,
(where ρ is the co-efficient of correlation)

(a) $\sigma_x(1 - \rho^2)$ (b) $n\,\sigma_x^2(1 - \rho^2)$ (c) $\sigma_x^2(1 - \rho^2)$ (d) $n\,\sigma_y^2(1 - \rho^2)$

47. Probable error of *r* is

(a) $0.6745\left(\frac{1 - r^2}{\sqrt{n}}\right)$ (b) $0.6754\left(\frac{1 + r^2}{\sqrt{n}}\right)$ (c) $0.6547\left(\frac{1 - r^2}{n}\right)$ (d) $0.6754\left(\frac{1 - r^2}{n}\right)$

48. If the two variables *x* and *y* of a bivariate distribution have a perfect correlation, they may be connected by

(a) $xy = 1$ (b) $\frac{a}{x} + \frac{b}{y} = 1$ (c) $\frac{x}{a} + \frac{y}{b} = 1$ (d) None of these

49. If *x* and *y* are related as $y - 4x = 3$, then the nature of correlation between *x* and *y* is

(a) Perfect positive (b) Perfect negative (c) No correlation (d) None of these

50. If $\sum x = 15$, $\sum y = 36$, $\sum xy = 110$, $n = 5$ then $Cov(x, y)$ equals

(a) $\frac{1}{5}$ (b) $\frac{-1}{5}$ (c) $\frac{2}{5}$ (d) $-\frac{2}{5}$