All India Shri Shivaji Memorial Society's

Institute of Information Technology

Course material for Engineering Mathematics-III Branch: Computer/IT

Topics: Statistics, Correlation and Regression

Q. Obtain arithmetic mean for the following data.

x	0	1	2	3	4	5	6	7	8	9	10
f	4	5	12	12	13	16	15	13	12	5	6

Q. Obtain arithmetic mean for the following data.

Marks obtained	0-10	10-20	20-30	30-40	40-50	50-60	60-70	70-80	80-90	90-100
Number of students	8	20	14	16	20	25	13	10	5	2

Q. Arithmetic mean of four numbers is 16. One item 20 is replaced by 24, what is the arithmetic mean?

Q. Obtain median for the following data.

x	1	3	5	7	9	11	13	15	17
f	3	6	8	12	16	16	15	10	5

Q. Obtain median for the following data.

Wages in Rupees	10-20	20-30	30-40	40-50	50-60
Number of labourers	5	8	13	10	8

Q. Find mode of the following distribution.

Marks obtained	0-10	10-20	20-30	30-40	40-50	50-60	60-70
Number of students	4	7	8	12	25	18	10

Q. Calculate arithmetic mean and mean deviation for the following data.

x	1	2	3	4	5	6
f	3	4	8	6	4	2

Q. Calculate standard deviation for the following data.

Marks obtained	0-10	10-20	20-30	30-40	40-50	50-60
Number of students	5	9	15	12	10	3

Q. Standard deviation of three numbers 9, 10, 11 is a) 2/3 b) 1/3 c)$\sqrt{2/3}$ d) $\sqrt{1/3}$

Q. Standard deviation of four numbers 9, 11, 13, 15 is a) 2 b) 4 c)$\sqrt{6}$ d) $\sqrt{5}$

Q. If $\sum x = 235,\ \sum x^2 = 6750,\ n = 10$ then standard deviation of the distribution is

a) 11.08 b) 13.08 c) 8.08 d) 7.6

Q. Coefficient of the variation of the data 1, 3, 5, 7, 9 is a) 54.23 b)56.57 c) 55.41 d)60.19

Q. The standard deviation and arithmetic mean of the distribution are 12 and 45.5 respectively. Coefficient of the variation of the data is a) 26.37 b)32.43 c) 12.11 d) 22.15

Q. The standard deviation and arithmetic mean of three distributions X, Y, Z is as below.

Distribution	standard deviation	arithmetic mean
X	18.0	5.4
Y	22.5	4.5
Z	24.0	6.0

The more stable distribution is a) X b) Y c) Z d) X and Z

Prepared by: Department of Mathematics, IoIT

Q. The standard deviation and arithmetic mean of three distributions X, Y, Z is as below.

Distribution	standard deviation	arithmetic mean
X	50	24.43
Y	46	25.495
Z	40	27

The more stable distribution is a) X b) Y c) Z d) Y and Z

Q. The standard deviation and arithmetic mean of three distributions X, Y, Z is as below.

Distribution	standard deviation	arithmetic mean
X	532	11
Y	831	9
Z	650	10

The more stable distribution is a) X b) Y c) Z d) Y and Z

Q. Fluctuations in the aggregate marks obtained by two groups of students are given below. Find which of the two shows greater variability.

Group A	518	519	530	530	544	542	518	550	527	527	531	550	550	529	528
Group B	825	830	830	819	814	814	844	842	842	826	832	835	835	840	840

Q. Calculate first four central moments about the mean for the following data.

x	2.0	2.5	3.0	3.5	4.0	4.5	5.0
f	4	36	60	90	70	40	10

Q. Calculate first four central moments for the following data.

Marks obtained	0-10	10-20	20-30	30-40	40-50
Number of students	6	26	47	15	6

Q. Calculate first four central moments for the following data. Hence find β_1 and β_2. (nov-17)

x	0	1	2	3	4	5	6
f	5	15	17	25	19	14	5

Q The first four moments of a distribution about 4 of the variable are -1.5, 17, -30 and 108. Find central moments and β_1 and β_2. Also comment on skewness and kurtosis of the distribution.(may-17)

Q. The first four moments of a distribution about 30.2 of the variable are 0.255, 6.222, 30.211 and 400.25. Find central moments and β_1 and β_2. Also comment on skewness and kurtosis of the distribution.

Q. The first four moments of a distribution about 5 of the variable are 2, 20, 40 and 50. Find central moments and β_1 and β_2. Also comment on skewness and kurtosis of the distribution.

Q. The first four central moments of a distribution are 0, 2.5, 0.7 and 18.75. Find central moments and β_1 and β_2. Also comment on skewness and kurtosis of the distribution.

Q. If $\beta_2 = 2.5$, then distribution is

a) Leptokurtic b) Mesokurtic c) Platykurtic d) none of these.

Q. If $\beta_2 = 3.9$ then distribution is

a) Leptokurtic b) Mesokurtic c) Platykurtic d) none of these.

Q. The first moment of the distribution about the value 5 is 2. Arithmetic mean of the distribution is

a) 15 b) 17 c) 18 d)16

Prepared by: Department of Mathematics, IoIT

Q. The first and second moments of the distribution about the value 3 are 2 and 20. Second moment about the mean is a) 12 b) 14 c) 20 d)16

Q. The first second and third moments of the distribution about the value 5 are 2 and 20 and 40. Third moment about the mean is a) 64 b) -64 c) 32 d)-32

Q. The first and second moments of the distribution about value 2 are 1 and 16. Variance of the distribution is a) 12 b) 3 c) 15 d) 17

Q. Calculate Karl Pearson's coefficient of correlation for the following data. (Dec-12)

x	23	28	42	17	26	35	29	37	16	46
y	25	22	38	21	27	39	24	32	18	44

Q. Find correlation coefficient between x and y, given that

$n = 25, \sum x = 75, \ \sum y = 100, \ \sum x^2 = 250, \sum y^2 = 500, \sum xy = 325.$

Q. Calculate coefficient of correlation for the following data.

X	10	14	18	22	22	30
Y	18	12	24	6	30	36

Q. Find correlation coefficient between x and y, given that

$n = 100, \sigma_x = 10, \sigma_y$=12, $\bar{x} = 62, \bar{y} = 53, \sum(x - \bar{x})(y - \bar{y}) = 8000$

Q. Find the line regression for the following data.

x	10	14	19	26	30	34	39
y	12	16	18	26	29	35	38

Q. Find the line regression for the following data.

x	2	3	5	7	9	10	12	15
y	2	5	8	10	12	14	15	16

Find estimate of (i) y when x=6 and (ii) x when y=20

Q. By the method of least squares, find the straight line that best fits the following data. (may-17)

X	1	2	3	4	5
Y	14	27	40	55	68

Q. By the method of least squares, find the straight line that best fits the following data. (nov-17)

X	0	5	10	15	20	25
Y	12	15	17	22	24	30

Q. Following is the information given.

	Variable x	Variable y
Arithmetic mean	10	90
Standard Deviation	3	12

Coefficient of correlation between x and y is 0.8. Find the lines of regression.

Q. Following is the information given.

	Variable x	Variable y
Arithmetic mean	30	500
Standard Deviation	3	100

Coefficient of correlation between x and y is 0.8. Find the lines of regression.

Prepared by: Department of Mathematics, IoIT

Q. The correlation coefficient between x and y is 0.6. If $\sigma_x = 1.5, \sigma_y = 2.00, \bar{x} = 10, \ \bar{y} = 20$ then find lines of regression.

Q. The correlation coefficient between x and y is 0.7111. If $\sigma_x = 4,\ \sigma_y = 1.8, \bar{x} = 5,\ \bar{y} = 4$ then find lines of regression.

Q. Following is the information given.

	Variable x	Variable y
Arithmetic mean	8.2	12.4
Standard Deviation	6.2	20

Coefficient of correlation between x and y is 0.9. Find the linear regression estimate of x given that y=10.

Q. If the two lines of regression are $9x + y - \lambda = 0,\ 4x + y = \mu\ and$ means of and y are 2 and -3 respectively, find values of λ and μ. Also find coefficient of correlation between x and y.

Q. The regression equations are 8x-10y+66=0, 40x-18y=214. The value of variance of x is 9. Find (1) the mean value of x and y (2) correlation coefficient (3) standard deviation of y.

Q. The regression equations are 3x+2y-26=0, 6x+y-31=0. Find (1) the mean value of x and y (2) correlation coefficient. (nov 17)

Q. If $\sum xy = 2800, \bar{x} = 16, \bar{y} = 16, n = 10$, variance of x is 36 and that of y is 25 then r(x,y) is equal to

a) 0.95 b) 0.73 c) 0.8 d) 0.65

Q. If $\sum xy = 190, \bar{x} = 4, \bar{y} = 4, n = 10$, variance of x is 1.732 and that of y is 2 then r(x,y) is equal to

a) 0.9128 b) 0.8660 c) 0.7548 d) 0.5324

Q. If $cov(x, y) = 25.8, \sigma_x = 6, \sigma_y = 5$ then r(x,y) is a) 0.5 b) 0.75 c) 0.91 d) 0.86

Q. If covariance between x and y is 10, variance of x and y are 16 and 9 respectively, then r(x,y) is

a) 0.8333 b) 0.6333 c) 0.527 d) 0.745

Q. If the two regression coefficients are 0.16 and 4 then correlation coefficient is

a) 0.08 b) -0.8 c) 0.8 d)0.64

Q. If the two regression coefficients are -8/15and -5/6 then correlation coefficient is

a) 0.5 b) -0.667 c) 0.537 d)-1.5

Q. cov(x,y) for $\sum xy = 2638, \bar{x} = 14, \bar{y} = 17, n = 10$ is a) 24.2 b) 25.8 c) 23.9 d) 20.5

Q. cov(x,y) for $\sum xy = 1242, \bar{x} = -5.1, \bar{y} = -10, n = 10$ is a) 67.4 b) 83.9 c) 58.5 d) 73.2

Q. cov(x,y) for $\sum (x + y)^2 = 10623, \sum x^2 = 2219,\ \sum y^2 = 3056\ \bar{x} = 14.7, \bar{y} = 17, n = 10$ is

a) 1.39 b) 13.9 c) 139 d) -13.9

Q. r(x,y) for $n = 10, \sum x = 140, \sum y = 150, \sum x^2 = 1980, \sum y^2 = 2465, \sum xy = 2160$ is

a) 0.753 b) 0.4325 c) 0.556 d)0.9013

Q. r(x,y) for $\bar{x} = 5.5,\ \bar{y} = 4, \sum x^2 = 385, \sum y^2 = 192, \sum (x + y)^2 = 947$ is

a) -0.924 b) -0.681 c) -0.542 d)-0.813

Prepared by: Department of Mathematics, IoIT

Q. Number of observations for $r = 0.5, \sum xy\ = 350, \sigma_x = 1, \sigma_y = 4, \bar{x} = 3, \bar{y} = 4$ is a) 25 b)5 c)15 d) 20

Q. Number of observations for $r = 0.022, \sum xy = 33799, \sigma_x = 4.5, \sigma_y = 64.605, \bar{x} = 68, \bar{y} = 62.125$ is

a) 10 b)5 c)7 d) 8

Q. Coefficient of correlation between x and y is 0.8 and their covariance is 20 and variance of x is 16. Standard deviation of y is a)8.25 b) 6.25 c)6.75 d) 7.5

Q. For a given set of bivariate data $\bar{x} = 53.2, \bar{y} = 27.9, \sigma_x = 11, \sigma_y = 8, r = 0.66$, using line of regression x on y, the most probable value of x when y=75 is a) 29.143 b) 24.325 c) 31.453 d) 26.925

Q. For a given set of bivariate data $\bar{x} = 36, \bar{y} = 85$, regression coefficient of y on x is -1.5, using line of regression yon x, the most probable value of y when x=60 is a) 15.7 b) 13.7 c) 17.7 d) 21.7

Q. For a given set of bivariate data $\bar{x} = 2, \bar{y} = -3$, regression coefficient of x on y is -0.11, using line of regression x on y, the most probable value of x when y=10 is a) 0.77 b) 1.77 c) 0.87 d) 0.57

Q. Line of regression y on x is 8x+2y=26, line of regression of x on y is 6x+y=31. Variance of x is 25. Then standard deviation of y is a)-15 b) 15 c)1.5 d)-1.5

Q. Line of regression y on x is 8x-10y+66=0, line of regression of x on y is 40x-18y-214=0. Variance of y is 16. Then standard deviation of x is a)3 b) 2c)6 d)7

Q. Line of regression y on x is 8x-10y+66=0, line of regression of x on y is 40x-18y-214=0. r(x,y) is

a)0.6 b) 0.5 c) 0.75 d)0.45

Q. The regression lines are 9x+y=15, 4x+y=5. r(x,y) is a)0.444 b)-0.11 c)0.663 d)0.7