Exam.Code:0041
Sub. Code: 1004

2122

B.Sc. (Hons.) Bio-Informatics

Third Semester

BIN-3004: Statistical Methods

Time allowed: 3 Hours Max. Marks: 60

NOTE: *Attempt five questions in all, including Question No. I which is compulsory and selecting two questions from each Unit.*

x-x-x

I. Attempt the following:

a) Coefficient of correlation between X and Y is 0.3. Their covariance is 9. The variance of X is 16. Find the standard deviation of Y series. **(2)**

b) Define distribution free test. **(1)**

c) Differentiate between critical region and acceptance region. **(2)**

d) Define the following terms used in design and analysis of experiments:

i. Treatment
ii. Experimental unit
iii. Yield
iv. Block **(4)**

e) A statistical measure of the frequency with which two features are found in association with each other is known as _____. **(1)**

f) What is contingency table? **(1)**

g) Define Yates correction. **(1)**

UNIT - I

II. **(a)** Differentiate between Karl Pearson's Correlation and Spearman's Correlation? Give examples also. Does the correlation coefficient imply 'causation' between two variables?

(b) A computer while calculating the correlation coefficient between the variable X and Y obtained the following results:

$N=30 ; \sum X = 120, \sum X^2 = 600, \sum Y = 90, \sum Y^2 = 250, \sum XY = 335$

It was, however, later discovered at the time of checking that it had copied down two pairs of observations as: (X, Y) : (8, 10) (12, 7)

While the correct values were: (X, Y) : (8, 12) (10, 8)

Obtain the correct value of the correlation coefficient between X and Y. **(6,6)**

III. **(a)** Define the following:

(i) Lines of Regression

(ii) Regression Coefficients

(b) Why do we have, in general, two regression lines? Write the properties of regression coefficients.

Contd.....P/2

(c) Out of the two lines of regression given by: X+2Y-5=0 and 2X+3Y-8=0, which one is the regression line of X on Y? **(4, 4, 4)**

IV. **(a)** Write a short note on determination of sample size for estimating means.
(b) How do you find the sampling distribution of a sample proportion?
(c) What is the standard error of the sampling distribution of the sample proportion?
(d) How do we construct the confidence interval for the estimation of population mean when standard deviation is (i) known and (ii) unknown? **(3, 3, 2, 4)**

UNIT - II

V. **(a)** What is testing of hypothesis? How the decision is taken in testing of hypothesis?
(b) Explain the following terms:
i. Type of Hypothesis
ii. Level of Significance
iii. Contingency table
iv. Type of Errors **(4, 8)**

VI. Three hundred apples were distributed among 10 persons. The distributions were as follows:

Number of persons:	0	1	2	3	4	5	6	7	8	9
Apples distributed:	28	29	33	31	26	35	32	30	31	25

Apply χ^2-test to test the hypothesis that the apples were distributed in equal numbers. For 9 degrees of freedom and 5% level of significance the table of χ^2 is 16.92. **(12)**

VII. **(a)** Derive the sign test, stating clearly the assumptions made.
(b) An examination designed to measure the basic I.Q. was given. Random samples were taken of 20 boys and 20 girls joining as management trainees in a company. The scores obtained by them out of 50 are given below:

Boys	28	30	32	34	46	45	39	15	33	23	27	29	36	48	47	45	26	28	47	42
Girls	35	37	31	33	41	44	38	16	42	46	13	34	12	10	40	41	42	26	13	15

By applying U-test determine whether there is a significant difference in the average I.Q. of boys and girls (use .05 level of significance). **(4, 8)**

X-X-X