1128
B.Sc. (Hons.) Bio-Informatics
Third Semester
BIN-3004: Statistical Methods

Time allowed: 3 Hours   **Max. Marks: 60**

NOTE: *Attempt* <u>*five*</u> *questions in all, including Question No. I which is compulsory and selecting two questions from each Unit. Simple calculator is allowed. Statistical tables may be provided on demand.*

x-x-x

Question(1)

(a) Where to use Spearman's rank correlation? Explain with example..

(b) If the correlation coefficient between random variables x and y is positive then show that $E(xy) > E(x)E(y)$.

(c) Discuss the use of regression lines.

(d) What factors are responsible for determining the number of replications.

(e) What assumptions are generally made for a non parametric test?

(f) What are the advantages and disadvantages of completely randomized design. $(6 \times 2 = 12)$

UNIT I

Question (2):

(a) What is Spearman's rank correlation coefficient? Prove Spearman's rank correlation coefficient and establish its limits.

(b) Two office in charges rank ten subordinates according to their suitability for promotion as follows:

Subordinate	A	B	C	D	E	F	G	H	I	J
In Charge 1	1.5	1.5	3	4	6	6	6	8	9.5	9.5
In charge 2	1	2	4	4	4	6	7	8	9	10

Find the Spearman's rank correlation coefficient.   (6,6)

Question (3):

(a) What are regression lines? Why is it necessary to consider two lines of regression? In case the two lines are identical, prove that the correlation coefficient is $+1$ *or* -1. If the two variables are independent then show that the two regression lines are perpendicular.

(b) From the following data:

Sales	91	97	108	121	67	124	51	73	111	57
Purchase	71	75	69	97	70	91	39	61	80	47

Obtain the two regression equations.   (6,6)

Question (4):

(a) Explain briefly how the width of a confidence interval decreases with decrease in the confidence level. Give an example.

(b) (i) How one can compute sample size in case of proportions?

P.T.o.

(ii) A company has installed a new machine that makes a part that is used in clocks. The company wants to estimate the proportion of these parts produced by this machine that are defective. The company manager wants this estimate to be within .02 of the population proportion for a 95% confidence level. What is the most conservative estimate of the sample size that will limit the maximum error to within .02 of the population proportion?   (6,6)

UNIT II

Question (5):
(a) Explain the terms:
(i) Level of Significance   (ii) Critical values
(iii) Acceptance and rejection Regions   (iv) One-tailed and two-tailed tests.
(b)(i) Explain the difference between the observed frequencies and expected frequencies for a goodness of fit test.
(ii) The following table lists the age distribution for a sample of 100 persons arrested for drunk driving.

Age	16 - 25	26 - 35	36 - 45	46 - 55	56 and older
Arrests	32	25	19	16	8

Using the 1% significance level, can we reject the null hypothesis that the proportion of people arrested for drunk driving is the same for all age groups?   (6,6)

Question (6):
(a) Briefly explain when one-way ANOVA procedure is used to make a test of hypothesis.
(b) Fifteen fourth grade students were randomly assigned to three groups in order to experiment with three different methods of teaching arithmetic. At the end of the semester, the same test was given to all the 15 students. The table gives the scores of students in the three groups.

Method I	Method II	Method III
48	55	84
73	85	68
51	70	95
65	69	74
87	90	67

Calculate the value of the test statistics F.   (6,6)

Question (7):
(a) Explain the analysis of Completely randomized design and discuss its uses.
(b) Discuss (i) Sign test and (ii) Wilcoxon signed-rank test.   (6,6)

x- x-x