

Tribhuvan University
Institute of Science and Technology
2082
☆

Bachelor Level / Second Year/ Third Semester/ Science
Computer Science and Information Technology (STA 215)
(Statistics II)
(NEW COURSE)

Full Marks: 60
Pass Marks: 24
Time: 3 hours

Candidates are required to give their answers in their own words as far as practicable. All notations have the usual meanings. The figures in the margin indicate full marks.

Group A

Attempt any TWO questions.

(2×10=20)

1. The following sample data were collected to determine the relationship between two processing variables and the current gain of a transistor in an integrated circuit.

Diffusion time in hours (x_1)	1.5	1.2	0.5	2.7	1.6	0.3
Sheet resistance ohm-cm (x_2)	65	140	70	92	123	105
Current gain (y)	5.2	9.8	7.5	10.8	12.6	9.1

- (i) Fit a regression equation and use this equation to estimate the expected current gain when the diffusion time is 1.8 hours and the sheet resistance is 90 ohm –cm.
- (ii) Compute the coefficient of multiple determination and interpret its value.
2. Company A recently developed an epoxy painting process for corrosion protection on exhaust components. Mr.Y, the owner, wishes to determine whether the lengths of life for the paint are equal for three different conditions: saltwater, fresh-water without weeds, and freshwater with a heavy concentration of weeds. Accelerated-life tests were conducted in the laboratory, and the number of hours the paint lasted before peeling was recorded.

Saltwater	167.4	189.6	177.2	168.4	180.3
Freshwater without weeds	160.5	177.6	185.3	170.9	176.6
Fresh water with weeds	180.7	164.4	172.9	169.2	170.7

Use the Kruskal-Wallis test at 0.01 level of significance to determine whether the lasting quality of the paint is same for the three water conditions.

3. Differentiate between Z -test at t –test. Industrial wastes, sewage dumped into our rivers, streams absorb oxygen and thereby reduce the amount of dissolved oxygen available for fish and other forms of aquatic life. One state agency requires a minimum of 8 parts per million (ppm) of dissolved oxygen in order for the oxygen content to be sufficient to support aquatic life. Eight water specimens taken from a river at a specification location during the low- water season gave the readings 4.5, 5.1, 4.8, 4.9, 5.1, 4.9, 5.0 and 4.7 ppm of dissolved oxygen. Do the data provide sufficient evidence to indicate that the dissolved oxygen content is less than 5ppm? Use 5% level of significance.

Group B

Attempt any EIGHT questions.

(8×5=40)

4. A population consists of the four number 4, 7, 10 and 11. (i) Write down all possible samples of size 2 which can be drawn without replacement from this population. (ii) Find the mean of sampling distribution of means and show that it is equal to the population mean. (iii) Find the variance of sampling distribution of means. (iv) Find the standard error of sample mean.

5. Diet X runs a number of weight reduction centers within a large town in the north east of Nepal. From the historical data it was found that the weight of the participants is normally distributed with mean 76 kg and standard deviation 5.2 kg. Calculate the probability that the average sample weight is greater than 78 when 25 participants are randomly selected for the sample.
6. Define point and interval estimation. A sample poll of 1000 voters chosen at random from all voters in a given district indicated that 58% of them were in favor of a particular candidate. (i) Compute standard error of sample proportion of voters who are in favor of the candidate. (ii) Find 99% confidence limits for the proportion of all the voters in favor this candidate.
7. In study of 200 subjects with acute ankle sprains were tested to compare two treatment methods, one (called Treatment) with oxygen at 1.2 atmosphere absolute pressure and the other (called Control) with air at 1.9 atmosphere absolute pressure in hyperbaric chamber. The subjects were divided equally in a double-blind manner between the two methods of treatment. A seven points Ankle Function Score (AFS) was tested to evaluate the effectiveness of each method and higher AFS indicates higher level of recovery. At the end, the following information was gathered.

	Control	Treatment
Number of subjects	100	100
Mean AFS	5.4	6.5
Standard deviation of AFS	0.8	0.6

Using 5% level of significance, test whether treatment is better than control.

8. What is meant by a Randomized block design? What are the assumptions made in the analysis a randomized block design?
9. The following ANOVA summary table was obtained from a multiple regression model with two independent variables.

Source of Variation	Sum of square	Degree of freedom	Mean sum of square	F -value
Regression	60	?	?	?
Error	?	?	?	
Total	190	20		

i) Fill in the blanks.

ii) Test the significance of the overall regression model at 5% level of significance.

10. A study was conducted to determine the effect of early child care on infant-mother attachment patterns. In the study, 90 infants were classified as either "secure" or "anxious" using the Ainsworth strange situation paradigm. In addition, the infants were classified according to the average number of hours per week that they spent in child care. The data were presented in the table.

	Low (0-5hours)	Moderate (6-19hours)	High (20-54)
Secure	21	35	5
Anxious	10	11	8

Do the data provide sufficient evidence to indicate that there is a difference in attachment pattern for the infants depending on the amount of the time spent in child care? Use 5% level of significance.

11. Define stochastic process. In a town each day is either sunny or rainy. A sunny day is followed by another sunny day with probability 0.8, whereas a rainy is followed by sunny day is with probability 0.4. If it rains on Sunday, make forecast for Monday and Tuesday.

12. Write short notes on:

i) Parametric and non parametric test.

ii) Type I and type II error in testing of hypothesis.