Tribhuwan University
Institute of Science and Technology
2077

Bachelor Level / Third Semester / Science — Full marks: 60
Computer Science and Information Technology(STA210) — Pass marks: 24
((TU CSIT) Statistics II) — Time: 3 hours
Candidates are required to give their answers in their own words as far as practicable.
The figures in the margin indicate full marks.

Group A

Attempt any Two questions: (2 x 10 = 20)

1. Describe the concept of sampling distribution of mean with reference to the population data (20, 21, 22 & 23) of size 4. In order to explain this, perform simple random sampling with replacement taking all possible samples with sample size n =2. While describing the sampling distribution following issues will be covered.

I. Population mean & population variance, and its distribution

II. Sample mean & sample variance, and its distribution

III. Comparison of population mean and sample mean; population variance and sample variance; population distribution and sampling distribution based on the given data.

IV. Standard error of mean

V. Final comments based on your result.

2. It was reported somewhere that children whenever plays the game in computer, they used the computer very roughly which may reduce the lifetime of computer. The random access memory (RAM) of computer also plays a crucial role on the lifetime of a computer. A researcher wanted to examine how the lifetime of a personal computer which is used by children is affected by the time (in hours) spends by the children per day to play games and the available random access memory (RAM) measured in megabytes (MB) of a used computer. The data is provided in the following table.

Lifetime (years)	5	1	7	2	3	4	6
Play time (hours)/day	2	8	1	5	6	3	2
RAM (in MB)	8	2	6	3	2	4	7

Identify which one is dependent variable? Solve this problem using multiple linear regression model and provide problem specific interpretations based on the regression model developed.

3. Explain the fundamental concepts of Latin Square Design (LSD) with its necessary conditions. Perform the analysis of variance from the following data and make final comments based on the analysis.

A(5)	B(10)	C(15)
C(20)	A(15)	B(10)
B(20)	C(5)	A(10)

Group B

Attempt any Eight questions: (8 x 5 = 40)

4. A dealer of a DELL company located at New Road claimed that the average lifetime of a multimedia projector produced by Dell Company is greater than 60,000 hours with standard deviation of 6000 hours. In order to test his claim, sample of 100 DELL projectors are taken and the average life time was monitored and it was found to be 55,000 hours. Test the claim of the dealer at 5% level of significance.

5. Based on he following information, perform the following :

I. Test whether two mean are significantly different (α = 5%) using independent test.

II. Compute 95% confidence interval estimation for the difference of mean.

III. Show the linkage between testing of hypothesis and confidence interval estimation in this problem,

	Group I	Group II
Sample mean	10	15
Sample standard deviation	3	5
Sample size	49	64

6. A study of 1000 computer engineers conducted by their professional organization reported that 300 stated that their firms' greatest concern was to uplift the professional quality of work. In order to conduct a follow up study to estimate the population proportion of computer engineers to fulfill their greatest concern within ±0.01 with 99% confidence interval, how many computer engineers would be required to be surveyed?

7. A survey was conducted to see the association between hacking status of the email and the type of e-mail account. The survey has reported the following cross tabulation.

Type of e-mail account	Hacking status	
	Yes	No
Yahoo	60	15
Gmail	20	120

Do the information provide sufficient evidence to conduct that the type email account and the hacking status is associated? Use Chi-square test at 1% level of significance.

8. A machine produces metal rods used in an automobile suspension system. A random sample of 6 rods is selected and diameter is measured. The measuring data (in millimeters) are as follows. Assuming that the samples drawn from the normally distributed population.

8.24	8.26	8.20	8.28	8.21	8.23

Find 95% two sided confidence interval on the mean rod diameter, and interpret the result with reference to the given problem.

9. Use Mann-Whitney U test to assess whether the following satisfaction score based on the performance of two different special types of gadgets at 5% level of significance.

Gadget A	50	40	30	20
Gadget B	40	30	10	40

10. Define Markov chain and describe its characteristics.

11. Every day is generally considered as either sunny or rainy. A sunny day is followed by another sunny day with probability 0.8 where as a rainy day is followed by a sunny day with probability 0.4. Suppose it rains on Monday. Make forecasts for Tuesday and Wednesday.

12. Write short notes on the following:

I. Test of equality of two variances

II. Adjusted R^2