Tribhuwan University
Institute of Science and Technology
2078

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. There are three brands of computers namely Dell, Lenovo and HP. The following are the lifetime of 15 computers in years.

Serial number	Computer brand	Life time in years
1	Dell	15
2	Lenovo	10
3	HP	9
4	Dell	12
5	Lenovo	6
6	HP	7
7	Dell	4
8	Lenovo	8
9	HP	13
10	Dell	11
11	HP	5
12	Lenovo	7
13	Dell	3
14	HP	5
15	Lenovo	4

Apply appropriate statistical test to identify whether the average life time (in years) is significantly different across three brands of computer at 5% level of significance. You can again tabulate the data initially in the required format for statistical analysis.

2. Explain the sampling distribution of mean with reference to some numerical example. Illustrate the practical implications of Central Limit Theorem (CLT) in inferential statistics?

3. A study was conducted among IT officers working in different IT Centers in Kathmandu valley, one of the objectives of the study was to quantify the effect of age and working hour per day on Computer Vision Syndrome (CVS). The CVS was measured in a continuum measurement scale varying from 0 to 50. Few parts of the data were taken randomly from the surveyed data and provided in the following table for statistical analysis.

Respondents' ID	001	007	125	231	99	299	145
Scales of CVS	6	7	5	11	3	29.0	28
Age of respondents (in year)	24	26	30	41	47	50	52
Working hour (per day)	4	5	6	8	3	6	7

Recognize which one is dependent variable? Assuming that the relationship between CVS, age and working hour is linear. Fit a multiple linear regression model to address the objective of the study and interpret the model appropriately.

Group B

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

4. The following are the details of working hours in class room per week of male and female faculty working in the area of Computer Science and Information Technology of Tribhuvan University.

	Male faculty	Female faculty
Sample size	60	30
Average working hour per week	12	9
Standard deviation of working hours per week	4	3

Apply independent t-test to examine whether the average working hour in class room per week is significant different between male and female faculty, at 1% level of significance. State also null and alternative hypothesis appropriately.

5. A survey was conducted among 70 students studying B.Sc. CSIT in some colleges randomly. Among them 50 students secured more than 80% marks in Statistics. Compute 99% and 95% confidence interval for the population proportion of students who secured more than 80% marks in the subject Statistics, and comment on the results.

6. In location 1, there are 250 corona positive cases out of 460 persons and in location 2, 250 positive cases reported out of 650 persons. Can it be concluded that the proportion of corona positive cases is higher in location 1 compared to location 2? Test at 10% level of significance.

7. Previous literature has reported that the average age of B.Sc. CSIT enrolling students in Tribhuvan University is 22 years. A researcher has doubts on this information and he feels that the average age to be less than 22 years. In order to examine this, the following sample data were collected randomly from the enrolling students of CSIT.

Age in years	20	19	22	23	19	20	21	20	19	20

Set up null and alternative hypothesis and test whether the researcher doubt will be justified. Use 5% level of significance. Assume that the parent population from which samples are drawn, is normally distributed.

8. Apply Mann-Whitney U test for examining the following knowledge score on IT among two groups of IT workers at 5% level of significance?

Group A :	5	8	2	7	6
Group B:	9	12	4	6	

9. A survey was conducted to see the association between job opportunity status (Yes vs. No) of IT workers and gender. The survey has reported the following details.

Gender	Job opportunity status of IT workers	
	Yes	No
Male	70	35
Female	40	90

Do the information provide sufficient evidence to conclude that gender is associated with job opportunity status of IT workers? Use Chi-square test at 5% level of significance.

10. State mathematical model for Statistical analysis of m x m LSD for one observation per experimental unit. Also prepare a dummy ANOVA table for this.

11. Define Markov chain and introduce its basic notations. Also explain the characteristics of a Markov chain.

12. Write short notes on the following:

I. Rationale of using non-parametric statistical test

II. Estimation of minimum sample size for the given proportion