Tribhuwan University
Institute of Science and Technology
2078

Bachelor Level / second-semester / Science Full marks: 60 **Computer Science and Information Technology(STA164)** Pass marks: 24 (Statistics I) 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

Long answer questions:

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

1. What are different methods of measuring dispersion? Sample of polythene bags from two manufactures, A, B are tested by a prospective buyer for bursting pressure and the results are as follows.

Bursting pressure		5-10	10-15	15-20	20-25	25-30	30-35
Number of bags manufactured by	A	2	9	29	54	11	5
	B	9	11	18	32	27	13

Which set of bags has more uniform pressure? If prices are the same, which manufacturer's bags would be preferred by the buyer? Use appropriate statistical tools.

2. Write the properties of correlation coefficient. The time it takes to transmit a file always depends on the file size. Suppose you transmitted 30 files, with the average size of 126K bytes and the standard deviation of 35 Kbytes. The average transmitted time was 0.04 seconds with the standard deviation 0.01 seconds. The correlation coefficient between the time and size was 0.86. Based on these data, fit a linear regression model and predict the time it will take to transmit a 400K byte file.

3. (a) What do you understand about the Poisson distribution? What are its main features?

(b) What do you mean by joint probability distribution function? Write down its properties.

Group B

Short answer questions:

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

4. If 50 images of your website, 10 have black and white images, and their average scanned image occupies 2.5 megabytes of memory. The total image occupied by the entire work is 281 megabytes. Find the average occupied megabytes of those color images.

5. Calculate Q_1, D_7, and P_{58} from the following data and interpret the results.

Weight	0-10	10-15	20-25	25-30	30-35	35-40	40-45	45-50	50-60
No. of person	4	8	30	15	13	6	4	4	1

6. The following joint probability data apply to fatigue tests to be run on bronze strips. X represents failure (in 10^5) when alternate strips are bent at a high level of deflection. Y represents the same at a lower deflection level.

X \ Y	20	30	40	50
4	0.01	0.03	0.05	0.02
5	0.03	0.1	0.08	0.04
6	0.02	0.08	0.12	0.11
7	0.02	0.04	0.07	0.18

a) Find the marginal probability distribution for X and for Y.

b) Determine the conditional probability distribution of Y given X=5.

c) Are X and Y independent?

7. Fit a binomial distribution to the following data.

X	0	1	2	3	4	5	6
f	5	8	15	14	10	6	2

8. If two random variables have the joint probability density function

$$f(x,y) = \begin{cases} K(2x+3y), for\ 0 \le x \le 1, & 0 \le y \le 1 \\ 0, & otherwise \end{cases}$$

Find (i) constant k (ii) Conditional probability density function of X given Y (iil) identify whether X and Y are independent. 9.

Compute first four moments about arbitrary point 4 from following distribution and describe the characteristics of data.

X	2	3	4	5	6
f	1	3	7	2	1

10. The lifetime of a certain electronic component is a normal random variate with the exception of 5000 hours and a standard deviation of 100 hours. Compute the probabilities under the following conditions:

a) Lifetime of components is less than 4000 hours

b) Lifetime of components between 3000 to 6500 hours

c) Lifetime of components more than 6000

11. Calculate Spearman's rank correlation coefficient for the following ranks given by three judges in a music contest.

1st Judge	2	1	4	6	5	8	9	10	7	3
2nd Judge	4	3	2	5	1	6	8	9	10	7
3rd Judge	5	8	4	7	10	2	1	6	9	3

Indicate which pair of judges has the nearest approach to music.

12. What do you mean by sampling? Explain the difference between stratified sampling and cluster sampling.

13. State with suitable examples the role played by computer technology in applied statistics and the role of statistics in information technology.