STA169-2081 ✡

Tribhuvan University
Institute of Science and Technology
2081
✡

Bachelor Level / First Year/ Second Semester/ Science
Computer Science and Information Technology **(STA. 169)**
(Statistics I)
NEW COURSE

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

Candidates are required to give their answers in their own words as for as practicable.
The figures in the margin indicate full marks.

Section A

Long answer questions.
Attempt any **TWO** questions. **(2×10=20)**

1. What do you mean by dispersion? Distinguish between absolute and relative measures of dispersion. Ten typists applying for a job in a company made the following scores in a typing speed test.

Applicant	1	2	3	4	5	6	7	8	9	10
Speed (words/min)	55	56	80	71	87	62	76	90	73	68

Compute mean, median, range and standard deviation of above data.

2. Define positive and negative correlation. What are the required assumptions for correlation analysis? A data analytic company wants to find the relation between traffic in website (X) per day and server downtime (Y) in minutes per day. The collected data are:

X	9	10	12	9	10	13	13	19
Y	26	38	27	45	55	80	84	100

(i) Find the correlation coefficient between x and y. interpret the value.

(ii) Find the regression equation of y on x. Estimate the value of y when x= 16.

(iii) Interpret the value of y -intercept and slope of the line.

3. What is the significance of normal probability distribution? What are the main properties of normal curve? The time it takes a cell to divide (called mitosis) is normally distributed with an average time of one hour and standard deviation of 4 minutes. What is the probability that a cell divides in less than 55 minutes? (ii) What is the time that it takes approximately 99% of all the cells to complete mitosis?

Section B

Short answer questions.
Attempt any **EIGHT** questions. **(8×5=40)**

4. Define statistics. Discuss in brief importance and limitations of statistics.

5. What are the measures of central tendency? The following are the marks obtained by 95 students in Statistics.

Marks	20-30	30-40	40-50	50-60	60-70	70-80
No. of students	10	17	19	26	15	8

Find (i) the modal marks (ii) the lowest marks of the top 10% student.

6. A factory produce two types of electric lamps A and B, in an experiment relating their life, the following results were obtained:

Length of life in1000 hours	4-6	6-8	8-10	10-12
Number of lamps A	7	9	11	3
Number of lamps B	5	10	13	2

State (i) which model has a longer average life? (ii) Which model is more consistent in terms of life?

7. The following scores represent the final examination score for an elementary statistics course: 45, 65, 23, 32, 57, 74, 60, 30, 17, 10, 80 and 50. (i) Compute five number summary (ii) Construct a box and whisker plot and interpret the result.

8. Explain the mathematical approach of probability. A list of important customers contains 25 names. Among them 18 persons have their accounts in good standing while 7 are delinquent. Three persons will be selected at random from this list and status of their accounts checked. Calculate the probability that (i) all accounts are delinquent (ii) one account is delinquent and others are in good standing.

9. State Baye's theorem. A problem on multiple-choice quiz is answered correctly with probability 0.90 if a student is prepared. An unprepared student guesses between 4 possible answers, so that the probability of choosing the right answer is 1/4. Eighty percent of students prepare for the quiz. If a student X gives correct answer to this problem, what is the probability that he did not prepare for the quiz?

10. Of the items produced by a company, 3% are defective. A sample of 25 items is selected for inspection. Use both binomial and Poisson distribution to answer the question: what is the probability that exactly two defectives are found? Compare the result.

11. Suppose X and Y are independent random variables having the joint probability distribution:

f(x, y)		X	
		2	4
y	1	0.10	0.20
	3	0.15	0.30
	5	0.20	0.05

Find (i) marginal probability distribution of X and Y. (ii) E(X) (iii) E(Y) (iv) E (2X+3Y) (v) E (XY).

12. A private owned business operates both a drive - in facility and walk – in facility. On a randomly selected day, let X and Y respectively, are the proportions of the time that the drive- in and the walk –in facilities are in use, and suppose that the joint density function of these random variables is:

$f(x, y) = 0.4\,(2x+3y), \quad 0< x< 1, \quad 0 < y <1$

$= 0, \quad$ elsewhere.

Find $P\,(0< x < 0.5, \quad 0.25 < y < 0.5)$

13. Write short note on any two: (i) Primary data and secondary data. (ii) Sampling error and non sampling error (iii) Interval scale and ratio scale of measurements.