

June 13, 2024 02:30 pm - 05:30 pm 1T01875 - T.E. Computer Science & Engineering
(Artificial Intelligence & Machine Learning) (Choice Based) (R-2019 'C' Scheme) SEMESTER - V /
48895 - Department Optional Course - 1: Statistics for Artificial Intelligence & Data Science
QP CODE: 10056538

Duration: 3hrs

[Max Marks:80]

N.B. 1. Question No. 1 is compulsory.

2. Attempt any three questions out of remaining five.

3. All questions carry equal marks

4. Assume Suitable data, if required and state it clearly.

- 1 Attempt any four: 20
- (a) Define Confidence Interval?
 - (b) In a certain property investment company with an international presence, workers have a mean hourly wage of \$12 with a population standard deviation of \$3. Given a sample size of 30, estimate and interpret the SE of the sample mean.
 - (c) What is hypothesis testing? Explain type I and type II errors?
 - (d) What do you mean by correlation and regression? Explain with example.
 - (e) What is analysis of variance? Explain its usage.
- 2 (a) X is a normally distributed variable with mean $\mu = 30$ and standard deviation $\sigma = 4$. Find 10
- a) $P(x < 40)$
 - b) $P(x > 21)$
 - c) $P(30 < x < 35)$
- (b) . Some vehicles pass through a junction on a busy road at an average rate of 300 10
per hour.
- a. Find out the probability that none passes in a given minute.
 - b. What is the expected number of passing in two minutes?
 - c. Find the probability that this expected number found above actually pass through in a given two-minute period.
- 3 (a) For a certain type of computers, the length of time between charges of the 10
battery is normally distributed with a mean of 50 hours and a standard deviation of 15 hours. John owns one of these computers and wants to know the probability that the length of time will be between 50 and 70 hours.
- (b) The average score on a test is 80 with a standard deviation of 10. With a new 10
teaching curriculum introduced it is believed that this score will change. On random testing, the score of 38 students, the mean was found to be 88. With a 0.05 significance level, is there any evidence to support this claim?
- 4 a) Explain QQ plots in detail. Show how scatterplots explores relationships 10
between variables.

- b) Given four samples A, B, C, D. Solve using one-way ANOVA to identify any difference between samples. 10

Observation	A	B	C	D
1	8	12	18	13
2	10	11	12	9
3	12	9	16	12
4	8	14	6	16
5	7	4	8	15

- 5 a) What is F-Test? If the F statistic as 2.38 and the degrees of freedom obtained by him were 8 and 3. Find out the F value from the F Table and determine whether we can reject the null hypothesis at 5% level of significance (one-tailed test). 10
- b) Find the simple linear regression equation that fits the given data and coefficient of determination: 10

X	Y
2	69
9	98
5	82
5	77
3	71
7	84

- 6 a) Explain Binomial distribution in detail. 10
 Bottles of water have a label stating that the volume is 12 oz. A consumer group suspects the bottles are under-filled and plans to conduct a test. What would a Type I error in this situation mean?
- b) Write short notes on (any two) 10
1. Chi-square distribution.
 2. Weibull distribution.
 3. Stem & Leaf Plot
 4. Box Plot
