U.S.N.										

B.M.S. College of Engineering, Bengaluru-560019

Autonomous Institute Affiliated to VTU

September / October 2023 Supplementary Examinations

Programme: B E **Semester: IV**
Branch: Artificial Intelligence and Machine Learning **Duration: 3 hrs.**
Course Code: 20AM4PCPSM **Max Marks: 100**
Course: Probability and Statistics for Machine Learning **Date: 13.09.2023**

Instructions: 1. Answer any FIVE full questions, choosing one full question from each unit.
2. Missing data, if any, may be suitably assumed.

UNIT - I

1 a) Detail on the process of computing probabilities of events along with necessary relational representatives. **10**

b) Among 10 laptop computers, five are good and five are defective. Unaware of this, a customer buys 6 laptops. **10**
i) What is the probability of exactly 2 defective laptops among them?
ii) Given that atleast 2 purchased laptops are defective, what is the probability that exactly 4 are defective?

UNIT - II

2 a) There are 4 keys in a bunch. The house owner tries with each key to unlock the door until he succeeds. Let X be the number of wrong keys he tries before finding the right one. Apply suitable discrete random distribution and identify the values of E(X) and Var(X). **10**

b) Outline the difference between the Binomial and Poisson distributions with respect to their Expectation and Variance inferred through their probability mass functions (pmf)s. **10**

OR

3 a) Let X and Y be continuous joint random variables with joint pdf **10**

$f_{X,Y}(x,y) = \{cx+1$ if $x,y>=0$, $x+y=1$ and 0 otherwise$\}$

i) Find the value of constant c

ii) Find the marginal pdfs $f_X(x)$ and $f_Y(y)$

b) The baby yawn time is recorded between 0 to 23 seconds observed over certain period of time. Apply suitable distribution and compute its mean and standard deviation. Do justify the reason for the distribution selected for computation. **10**

UNIT - III

4 a) Using Simple Descriptive statistics, compute **10**

i) Mean, median, mode, the quartiles Q1, Q2, Q3 and interquartile range of the following data: 19, 29, 24, 22, 21, 31, 35, 17. Plot the graph to represent the quartiles identified through the mathematical relations.

ii) Relate the computed Quartiles to Percentiles numerically.

b) Illustrate the difference between sampling and non-sampling errors with related examples in real time. **5**

c) By applying the concepts of mean and median, Is it possible to interpret the nature of data distribution through its shape? How does mean and median vary as the shape changes? **5**

UNIT - IV

5 a) Installation of certain hardware takes random time with standard deviation of 5 minutes. **10**

i) A computer technician installs this hardware on 64 different computers with the average installation time of 42 minutes. Construct 95% confidence interval for the population - mean installation time.

ii) Suppose that the population mean installation time is 40 minutes and a technician installs the hardware on your PC. What is the probability that the installation time will be within the interval computed in **i)**.

b) Let X1, X2, …, Xn ~ N(θ, σ^2). Compute the Method of Moments estimator for the parameter vector (θ, σ^2). **10**

UNIT - V

6 a) Derive the estimates of confidence interval for the mean of responses. **10**

b) Elaborate the Logistic Regression Model working on the concept of Gradient Descent in Detail. **10**

OR

7 a) Using the given x and y values: **10**

(x)	0	1	2	3	4
(y)	2	3	5	4	6

i) Determine the coefficients of the least square regression line

ii) y = a x + b.

iii) Estimate the value of y when x=10.

b) Summarize multivariate ANOVA estimates in detail. **10**

****