

First Mid Term Examination, ME-781, September 15, 2022

Name:

Roll No:

Total Time 2 hours; Total Marks 110

Open notes (self-hand-written) examination.

Make any necessary assumption required to solve the questions.

- 3+7** 1. The price of a flight ticket between two cities is a function of the distance between the two cities and whether the source and destination cities are metros or not. Let the price of the flight ticket be denoted by Y , the distance between two cities be denoted by d , and the class of the source and the destination cities be C_s and C_d , respectively. The value of d could be any positive number (including zero), and the class of the cities could be only Metro (depicted by M) and non-Metro (depicted by N).

Then construct an expression (showing various parameters) for the following 2 models for the prediction of the price of the flight ticket Y :

- Linear model in d , C_s , and C_d , and
 - Non-linear model (in d , C_s , and C_d) with all possible quadratic terms
- 10** 2. Which data scale do the following variables belong to, and what all statistics can be used on them (provide a brief reason):
- Name of places
 - Cell number of a person
 - Age of students in a school
 - Ranking of a student in his/her class
 - Rate of change of fever of a hospitalized patient (in deg C per minute)
- 10** 3. Let set $A=\{0,2,3\}$ and set $B=[-1,1]$. Then draw the following sets:
- $C = A \times B$
 - $D = A \oplus b(0,1)$
 - Smallest convex set which contains C
 - Smallest convex set, which contains D
- 10** 4. A fair dice has numberings from 5 to 10, which forms the sample space. Random events are elements of the σ -algebra of sample space.
- Write the smallest σ -algebra for the above probability space.
 - Write the smallest σ -algebra, which contains an event that the number on the dice is a prime number.
 - Calculate the probability of each of the random events contained in the above two σ -algebras.
- 10** 5. Show whether the following are dissimilarity measures (in a 2D space) or not:
- $\sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}$
 - $\sqrt{|(x_1 - x_2) \cdot (y_1 - y_2)|}$
 - $\sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2} + \sqrt{|(x_1 - x_2) \cdot (y_1 - y_2)|}$
 - $(|x_1| - |x_2|) \cdot (|y_1| - |y_2|)$
- 10** 6. In a p -dimensional space, all the data points are only on the positive half side for each of the P -dimensions.
- If we are measuring the dissimilarity of each data point with the origin, then is it possible to determine the Euclidean norm and Frobenius (or Hilbert-Schmidt) norm using Lebesgue (or Minkowski) norm?
- Support your answer with the necessary derivation.

- 10 7. Statistically, the probability of rain on September 15 in Mumbai is 60%. Many students forget their umbrellas, and it is found that 40% of the students get their shoes wet when it rains. However, 5% of the students get their shoes wet while drinking water from the cooler. If a student comes to the class with wet shoes (on September 15), then what is the probability that it is not raining?
- 10 8. In a class of 100 students, each student is given a separate coin to perform a certain experiment. The students are told that their coin is either a fair coin or a coin with the same image on both sides. The students have to determine which of the two cases is true for their coin by observing the outcomes of six coin tosses. However, they are not allowed to flip the coin and see both sides. If all the students were given a fair coin and they were asked to use a p-value of 2.5%, then what is the expected number of students who would conclude that it is a coin with the same image on both sides?
- 10 9. The average age of students appearing for class 10th board exam is 14 years. The standard deviation of the age of these students is one year. If randomly 30 students appearing for the 10th board exam are chosen, what would be the expected value of the sample mean, sample variance, and expected value of the variance of the sample mean?
- 10 10. Let x be a uniform random variable between 0 and 3. Then, derive an expression for the probability density function of $9-x^2$.
- 10 11. For a linear regression with three features (X_1 , X_2 , and X_3), express the matrix form on Hessian of RSS ($\nabla^2 \text{RSS}(\beta)$).