

**INSTITUTE OF AERONAUTICAL ENGINEERING****(Autonomous)**

Dundigal, Hyderabad - 500 043

B TECH V SEMESTER CIE-II EXAMINATIONS, FEBRUARY- 2024**Regulation: UG20****STATISTICAL FOUNDATIONS OF DATA SCIENCE****CSE (AI&ML)****Time: 2 Hours****Max Marks: 20****Answer any FOUR questions****All parts of the question must be answered in one place only**

1. (a) Discuss set theory with its operations and properties. Outline about the term collectively exhaustive event. [BL: Understand| CO: 4|Marks: 2]
(b) A jar has 1000 coins, of which 999 are fair and 1 is double headed. Pick a coin at random, and toss it 10 times. Given that you see 10 heads, what is the probability that the next toss of that coin is also a head? [BL: Apply | CO: 4|Marks: 3]
2. (a) How important is the standard deviation to understanding relationships between data points? Why do you think students find it so hard to understand this concept? What could students do to better understand the standard deviation? [BL: Understand| CO: 5|Marks: 2]
(b) A telecom service provider claims that individual customers pay on an average 400 rs. per month with standard deviation of 25 rs. A random sample of 50 customers bills during a given month is taken with a mean of 250 and standard deviation of 15. What to say with respect to the claim made by the service provider? [BL: Apply | CO: 5|Marks: 3]
3. (a) Give the details on statistical analysis of data and explain properties and scales of measurement. [BL: Understand| CO: 5|Marks: 2]
(b) The following data represents shear strength (X) of the test spot weld 392, 376, 401, 367, 389, 362, 409, 415, 358, 375.
i) Assuming that X is normally distributed, estimate the true average shear strength and standard deviation of shear strength using the method of maximum likelihood.
ii) Obtain the MLE of $P(X \leq 400)$. [BL: Apply | CO: 5|Marks: 3]
4. (a) Interpret the effective data presentation techniques. Differentiate between correlation and causation. [BL: Understand| CO: 6|Marks: 2]
(b) Mention the critical elements of visualization that we can use to illustrate data stories. Describe line graph and box & whisker plot in data visualization. [BL: Apply | CO: 2|Marks: 3]
5. (a) Summarize about the graphs and statistics used in data analysis. Describe Simpson's paradox and how does it pertain to confounding? [BL: Understand| CO: 6|Marks: 2]
(b) Identify six factors that an entrepreneur should consider to make audio visual communication effective [BL: Apply | CO: 6|Marks: 3]

