BIRLA INSTITUTE OF TECHNOLOGY, MESRA, RANCHI
(END SEMESTER EXAMINATION)

CLASS: IMSC **SEMESTER : VIII**
BRANCH: QEDS **SESSION : SP/2025**

SUBJECT: ED419 RESAMPLING METHODS AND STATISTICAL COMPUTATION

TIME: 3 HOURS **FULL MARKS: 50**

INSTRUCTIONS:
1. The question paper contains 5 questions each of 10 marks and total 50 marks.
2. Attempt all questions.
3. The missing data, if any, may be assumed suitably.
***4*. Before attempting the question paper, be sure that you have got the correct question paper.**
5. Tables/Data hand book/Graph paper etc. to be supplied to the candidates in the examination hall.

			CO	BL
Q.1(a)	Use K-fold cross validation to test the hypothesis that the drug significantly reduces recovery time compared to the placebo. Calculate the p-value. $\boldsymbol{\alpha = 0.05}$.	[5]	CO1	3,4

Patient	Group	Recovery Time
1	Drug	8
2	Drug	7
3	Drug	9
4	Drug	6
5	Drug	7
6	Drug	8
7	Drug	9
8	Drug	10
9	Drug	7
10	Drug	6
11	Placebo	12
12	Placebo	14
13	Placebo	13
14	Placebo	15
15	Placebo	14
16	Placebo	13
17	Placebo	12
18	Placebo	14
19	Placebo	13
20	Placebo	15

Q.1(b)	Write three advantages and three disadvantages of using Jackknife method.	[3]	CO2	2
Q.1(c)	Given the data X = [12,15, 14, 10, 13, 17, 12, 14, 15, 16]. Calculate 98% bootstrap confidence interval of the median. (Do it for 10 iterations).	[2]	CO1	2,3
Q.2(a)	Define influence function? Derive the influence function of the median.	[5]	CO4	1,2
Q.2(b)	Derive Jackknife variance estimate of an estimator. Use this result to find the Jackknife standard error estimate of $\bar{x}^2$ of the following data [3, 9, 10, 7, 4]	[5]	CO2	2,3
Q.3(a)	Write the use of bootstrap technique in linear regression? Write the major differences between the residual bootstrap and the paired bootstrap method.	[4]	CO3	1,2
Q.3(b)	Explain multivariate Delta method in vector form. If $\sqrt{n}(X_n - \mu_0) \to N(0, V)$ where $\mu_0 = (1,2), X_n = (x_{n1}, x_{n2})$ and $V = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$. Then find the asymptotic distribution of $f(X_n) = \begin{pmatrix} e^{2x_{n1}+x_{n2}^3} \\ e^{x_{n1}^2+x_{n2}^3} \end{pmatrix}$.	[6]	CO4	3
Q.4(a)	Define projection depth. Prove that the projection depth is affine invariant.	[5]	CO4	1,2
Q.4(b)	Given $X = \{(1,1), (2,3), (4,2), (5,5), (6,1)\}$. Compute the spatial depth of $x = (3,2)$. Find the spatial median for the given data.	[5]	CO4	2

PTO

Q.5(a) The compressive strength of an alloy fastener used in aircraft construction is being studied. Ten loads were selected over the range 2500-3500 psi and a number of fasteners were tested at those loads. The numbers of fasteners failing at each load were recorded. The data is as follows. Fit a logistic regression model to the data. [5] CO5 2,3

Load (psi), x	Sample Size, n	Number Failing, r
2500	50	10
2700	70	17
2900	100	30
3100	60	21
3300	40	18
3500	85	43

Q.5(b) The joint distribution of (x, θ, n) is $\left(n_{C_x}\right)\theta^x(1-\theta)^{n-x}\frac{\Gamma(a+b)}{\Gamma a \Gamma b}\theta^{a-1}(1-\theta)^{b-1}e^{-\lambda}\frac{\lambda^n}{n!}$ $x = 0,1,\dots,n; 0 < \theta < 1, n = 0, 1, \dots$. Find the conditional distribution of $n|\ \theta, x$. [2] CO5 3,4

Q.5(c) Consider the dataset as follows. For constructing a tree-based method consider the splits of the data with respect to the size ≤ 900 and distance to center ≤ 4. Find the impurity functions with respect to these two splits. Among them find the best split. [3] CO5 2,3

Size (sqft)	Distance to Center (kms)	Rent ($1000s)
600	15	1.2
800	10	1.6
1000	5	2.0
1200	3	2.5
1500	2	3.0

:::::29/04/2025 E:::::