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

CLASS: IMSC | **SEMESTER: V**
BRANCH: CQEDS | **SESSION: MO/2024**

SUBJECT: ED303 MULTIVARIATE DATA ANALYSIS

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 handbook/Graph paper etc. to be supplied to the candidates in the examination hall.

			CO	BL
Q.1(a)	Let X be a 3-dimensional random vector with mean and covariance as follows. $S = \begin{bmatrix} 3 & -3/2 & 0 \\ -3/2 & 1 & 1/2 \\ 0 & 1/2 & 1 \end{bmatrix}$, $\bar{X} = \begin{bmatrix} 2 \\ 10 \\ 5 \end{bmatrix}$, $a = \begin{bmatrix} 2 \\ 2 \\ -1 \end{bmatrix}$, and $b = \begin{bmatrix} 1 \\ -1 \\ 3 \end{bmatrix}$ Find, (a) E(a'X) and E(b'X). (b) Var(a'X) and Var(b'X). (c) Cov(a'X, b'X).	[5]	1	2
Q.1(b)	If X is distributed as $N_p(\mu, \Sigma)$, what can you say about the distribution of any linear combination of variables, say, a'X, where a is a real vector of size p x 1? Support your answer with a proper argument and obtain the distribution of a'X.	[5]	2	4
Q.2(a)	Find the maximum likelihood estimators of 2x1 mean vector μ and variance-covariance matrix Σ based on the random sample $X = \begin{bmatrix} 3 & 6 \\ 4 & 4 \\ 5 & 7 \\ 4 & 7 \end{bmatrix}$.	[5]	2	3
Q.2(b)	Explain the MANOVA model for comparing the g population mean vectors with the MANOVA table and write the expression for Wilk's lambda test statistics.	[5]	3	3
Q.3(a)	Let X be normally distributed as $N_p(\mu, \Sigma)$. Derive the likelihood ratio test statistics to test the hypothesis about the mean vector $H_0: \mu = \mu_0$ vs $H_1: \mu \neq \mu_0$.	[5]	3	3
Q.3(b)	Evaluate the value of T^2 statistics and test the hypothesis $H_0: \mu = \begin{bmatrix} 7 \\ 11 \end{bmatrix}$ vs $H_1: \mu \neq \begin{bmatrix} 7 \\ 11 \end{bmatrix}$ using the sample data $X = \begin{bmatrix} 2 & 12 \\ 8 & 9 \\ 6 & 9 \\ 8 & 10 \end{bmatrix}$, at 5% level of significance. Use $F_{2,2} = 19.00$.	[5]	3	3

PTO

Q.4(a)	Explain the confidence region and simultaneous confidence interval. Write the expression of the exact simultaneous and Bonferroni confidence interval.	[5]	4	3
Q.4(b)	Let Π_1 and Π_2 be two exponential populations with different scale parameters λ_1 and λ_2 , (q_1, q_2) and (C(1|2), C(2|1)) are the prior probabilities and cost of misclassifications. Derive the classification region to classify a new observation, say Z, into one of the two populations Π_1 or Π_2.	[5]	4	3
Q.5	$$\Sigma = \begin{bmatrix} 5 & 2 \\ 2 & 2 \end{bmatrix}$$ a) Determine the population principal components X_1 and X_2. b) Compute the proportion of the total population variance explained by both components. Compute $\rho_{X1,Z1}$, $\rho_{X1,Z2}$, and $\rho_{X2,Z1}$. Also, interpret the value of the correlation coefficient.	[10]	4	4

::::::20/11/2024::::::M