23

1M4A24231 (Pages : 2) Reg. No:.....

Name: .....

FAROOK COLLEGE (AUTONOMOUS), KOZHIKODE

Fourth Semester M.Sc Statistics Degree Examination, April 2024

MST4C13 – Multivariate Analysis

(2022 Admission onwards)

Time: 3 hours Max. Weightage : 30

(Use of Scientific Calculator is permitted)

PART A

Answer any four questions. Weightage 2 for each question

1. Define a multivariate data with an illustrative example.
2. If $X \sim N_p(\mu, \Sigma)$, obtain the distribution of a subvector of X.
3. Show that Wishart distribution is a multivariate analogue of Chi-square distribution.
4. Define (i) partial and multiple correlation ii) partial regression coefficients.
5. Define Mahalanobis D^2 statistic. How is it related with Hotellings T^2 statistic? Mention any one use of the D^2 statistic.
6. Give a brief idea of classification problem. How the Fishers discriminant function is useful in classification problems?
7. What is meant by factor analysis? Mentioning its uses.

(4x2=8 Weightage)

PART B

Answer any four questions. Weightage 3 for each question

8. Obtain the MLE's of μ and Σ in $N_p(\mu, \Sigma)$. Show that they are independent.
9. If $X = \begin{pmatrix} X^{(1)} \\ X^{(2)} \end{pmatrix}$, where $X \sim N_p(\mu, \Sigma)$, then obtain the distribution of $X^{(1)} + LX^{(2)}$, where L is a matrix of constants with appropriate order.
10. If y_1, y_2, y_3 are i.i.d. standard normal variables verify whether quadratic form $3{y_1}^2 + 2{y_2}^2 + 2{y_3}^2 + 4\,y_1y_2 + 2y_2y_3$ is independent of the linear form $y_1 + y_2 + y_3$.

11. Define Wishart distribution. Find its characteristic function. Establish its additive property.
12. Explain the test for equality of two multivariate normal populations.
13. Derive the test for independence of sub vectors of a multivariate normal random vector.
14. Derive the variances of the first k principal components. How do you interpret the result obtained through principal component analysis?

(4x3=12 Weightage)

PART C

Answer Any 2 questions. Weightage 5 for each question

15. Define a multivariate normal distribution of X of order 'p'. Find its characteristic function. Obtain the distribution of the linear combination of the components of X.
16. (a) Define generalized variance and derive its distribution .

    (b) Define a matrix variate gamma distribution of one parameter. Obtain its characteristic function.
17. (a) Explain the multivariate Fisher-Behren problem.

    (b) Explain sphericity test.
18. (a) Describe the procedure of classification into one of two known multivariate normal populations with the same known dispersion matrix.

    (b)Explain the basic concepts and applications of cluster analysis.

(2x5=10 Weightage)

22

1M4A24232 (Pages : 1) Reg. No:.....

Name: .....

FAROOK COLLEGE (AUTONOMOUS), KOZHIKODE

Fourth Semester M.Sc Statistics Degree Examination, April 2024

MST4E23 – Applied Algorithms and Analysis of Multi Type and Big Data

(2022 Admission onwards)

Time: 3 hours Max. Weightage : 30

Part A

Answer any four questions. Each questions carries 2 weightage.

1. Define separating hyperplane.
2. Define structured data and provide an example.
3. What is perceptual map.
4. How the One-Versus-One and One-Versus-All strategies are applied in Support Vector Machines for multiclass classification.
5. How is Similarities Data different from Preferences Data?
6. What is the significance of having a well-defined analytics life cycle in managing big data projects
7. Define aggregate and disaggregate analysis in the context of Multi-Dimensional Scaling.

Part B

Answer any four questions. Each questions carries 3 weightage.

8. Explain EM algorithm with two component Gaussian mixture model.
9. Elaborate on the kernel method in Support Vector Machines and how it allows the algorithm to handle non-linear decision boundaries.
10. Define metadata and its role in managing Big Data.
11. Provide examples of unstructured data and discuss the challenges associated with analyzing such data.
12. Differentiate between descriptive, predictive, and prescriptive analytics.
13. In what scenarios would you prefer using metric methods over non-metric methods.
14. How can MDS results inform decision-making in a business or marketing context?

Part C

Answer any two questions. Each questions carries 5 weightage.

15. Elaborate on EM as a Maximization–Maximization Procedure.
16. Provide a detailed comparison between Support Vector Classifiers and Maximal margin classifiers in the machine learning classifications.
17. Outline the key phases of the Big Data Analytics Life Cycle and the significance of each phase.
18. Elaborate on the decision framework for perceptual mapping