Exam Date & Time: 09-Jun-2022 (02:00 PM - 05:00 PM)

MANIPAL INSTITUTE OF TECHNOLOGY
MANIPAL
(A constituent unit of MAHE, Manipal)

IV Semester BTECH End Semester Examination< DATA SCIENCE & ENGINEERING />

MATHEMATICAL FOUNDATION FOR DATA SCIENCE-II [MAT 2213]

Marks: 50 **Duration: 180 mins.**

Descriptive Questions

Answer all the questions.

1)

A) A prisoner in a prison is put in the following situation. A regular deck of 52 cards is placed in front of him. He must choose cards one at a time to determine their color. Once chosen, the card is replaced in the deck and the deck is shuffled. If the prisoner happens to select three consecutive red cards, he is executed. If he happens to select six cards before three consecutive red cards appear, he is granted freedom. Construct the transition probability matrix after representing the prisoner's situation as a Markov chain. (3)

B) For a Markov chain with state space S={a, b, c, d, e} having transition probability matrix given below, identify the closed sets and based on this information, state whether the chain is irreducible or not. (3)

$$P = \begin{bmatrix} 1/2 & 0 & 1/2 & 0 & 0 \\ 0 & 1/4 & 0 & 3/4 & 0 \\ 0 & 0 & 1/3 & 0 & 2/3 \\ 1/4 & 1/2 & 0 & 1/4 & 0 \\ 1/3 & 0 & 1/3 & 0 & 1/3 \end{bmatrix}$$

C) For a Markov chain with state space S={1, 2, 3, 4, 5, 6, 7, 8} having transition probability matrix given below, identify the transient, recurrent and absorbing states, if any. If there exists recurrent states, classify them into null recurrent and non-null states. Also determine the periodicity of the positively recurrent states. (4)

$$P = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.5 & 0 & 0.5 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0.3 & 0.7 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0.2 & 0.4 & 0 & 0 & 0.1 & 0 & 0.1 & 0.2 \\ 0 & 0 & 0.3 & 0 & 0 & 0.4 & 0 & 0.3 \end{bmatrix}$$

2)

A) A student's study habits are as follows. If he studies one night, he is 70% sure not to study the next night. On the other hand, if he does not study one night, he is 60% sure not to study the next night. Write the transition probability matrix. Are the states communicating? Are there any absorbing barriers? In the long run, how often does he study? (3)

B) Explain the four types of stochastic processes giving one example each, clearly mentioning the state space and the index set. (3)

C) Given $Y = (Y_1 \quad Y_2 \quad Y_3 \quad Y_4)^T \sim N_4(\mu, \Sigma)$ with $\mu = (1 \quad 2 \quad 3 \quad 4)^T$ and $\Sigma = \begin{bmatrix} 4 & 2 & -1 & 2 \\ 2 & 6 & 3 & -2 \\ -1 & 3 & 5 & -4 \\ 2 & -2 & -4 & 4 \end{bmatrix}$, find $E(Y_1 | Y_2\, Y_3\, Y_4)$ and $Var(Y_1 | Y_2\, Y_3\, Y_4)$. (4)

3)

A) It was required to test five varieties of crop at 6 different locations. The error sum of squares matrix was obtained to be $\begin{bmatrix} 3279 & 802 \\ 802 & 4017 \end{bmatrix}$ and the total sum of squares and products matrix was $\begin{bmatrix} 6067 & 3352 \\ 3352 & 6880 \end{bmatrix}$. Using the information provided, clearly state the hypotheses and obtain the test statistic to test for the similarities in the crop varieties at 5% level given the cutoff (table value) as 0.468.What is your inference? (2)

B) Given the distance matrix between pairs of five objects {1, 2, 3, 4, 5}, obtain the clusters and the final distance matrix D using single linkage. (3)

$$D = \begin{bmatrix} 0 & 9 & 3 & 6 & 11 \\ 9 & 0 & 7 & 5 & 10 \\ 3 & 7 & 0 & 9 & 2 \\ 6 & 5 & 9 & 0 & 8 \\ 11 & 10 & 2 & 8 & 0 \end{bmatrix}$$

C) In an orthogonal exploratory factor model (assume $m = 1$ factor model), suppose the manifest variables X_1, X_2 and X_3 have the sample covariance matrix $S = \begin{bmatrix} 100 & 40 & 0 \\ 40 & 60 & 0 \\ 0 & 0 & 30 \end{bmatrix}$. The estimated Eigen values of the covariance matrix S are $\hat{Q}_1 = 124.72, \hat{Q}_2 = 35.28, \hat{Q}_3 = 30$ (5)

a. Find the estimated Eigen vector corresponding to the largest Eigen value

b. Find the estimated loading matrix $\widehat{\Lambda}$

c. Find the communality for the variable X_2

4)

A) Let $\{X_n, n = 0, 1, 2, \dots\}$ be a Markov chain with state space $\{0, 1, 2\}$, the initial probability vector $P(0) = (\frac{1}{4}, \frac{1}{2}, \frac{1}{4})$ and the one – step transition probability matrix (3)

$$P = \begin{bmatrix} \frac{1}{4} & \frac{3}{4} & 0 \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ 0 & \frac{1}{4} & \frac{3}{4} \end{bmatrix}.$$

a. Find $P(X_0 = 0, X_1 = 1, X_2 = 1)$

b. Find $P(X_7 = 0|X_5 = 0)$

B) Consider the pdf $f(x_1,x_2) = \begin{cases} \frac{x_1+3x_2}{2} \ ; & 0 \le x_1,x_2 \le 1 \\ 0 \ ; & otherwise \end{cases}$

a. Find the mean vector μ (3)

b. Find the marginal density of X_2

C) Let $X = \begin{bmatrix} X_1 \\ X_2 \end{bmatrix}$ be the data matrix and $S = \begin{bmatrix} 14 & -11 \\ -11 & 23 \end{bmatrix}$ be the sample covariance matrix with sample size $n = 50$.

a. Find the first principal component Z_1 (4)

b. Find the variance explained by the first principal component

5)

A) For the data matrix $X = \begin{bmatrix} 10 & 100 \\ 12 & 110 \\ 11 & 105 \end{bmatrix}$ calculate the covariance matrix S. Using S obtain the correlation matrix R. (3)

B) Consider the Markov chain with state space $S = \{0,1,2,\dots\}$ and transition probabilities $P_{0,1} = 1 \ , \quad P_{i,i+1} = P_{i,0} = \frac{1}{2}$ for all $i \ge 1$. Draw the state transition diagram and find the stationary distribution. (3)

C) Let $X = \begin{bmatrix} X_1 \\ X_2 \\ X_3 \end{bmatrix}$ be a random vector with $X \sim N_3\left(\begin{bmatrix} -3 \\ 1 \\ 4 \end{bmatrix}, \begin{bmatrix} 1 & -2 & 0 \\ -2 & 5 & 0 \\ 0 & 0 & 2 \end{bmatrix}\right)$.

(a) Check whether (X_1,X_2) and X_3 independent or not. Justify your answer (4)

(b) Find the distribution of $\frac{X_1+X_2}{2}$ and X_3.

-----End-----