Exam Date & Time: 18-Jul-2022 (02:00 PM - 05:00 PM)

MANIPAL ACADEMY OF HIGHER EDUCATION

IV Semester MAKEUP Examination
Engineering Mathematics IV (MAT 2213)

MATHEMATICAL FOUNDATION FOR DATA SCIENCE-II [MAT 2213]

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

Descriptive Questions

Answer all the questions. Section Duration: 180 mins

1) A) Suppose that there are five types of breakfast cereal, which we call A, B, C, D and E. Customers tend to stick to the same brand. Those who choose type A, choose it the next time around with probability 0.8; those who choose type B, choose it the next time with probability 0.9. The probabilities for types C, D and E are given by 0.7, 0.8 and 0.6 respectively. When customers change brand, they choose one of the other four, equally probably. Construct the transition probability matrix after representing the situation as a Markov chain. (3)

B) For a Markov chain with state space S={0, 1, 2, 3} having transition probability matrix given below, identify the transient and recurrent states, if any, with the help of a transition probability diagram (3)

$$\begin{bmatrix} 0 & 0 & 0.5 & 0.5 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}$$

C) For a Markov chain with state space S={1, 2, 3, 4, 5} having transition probability matrix given below, draw all sample paths of length four that begin in state 1. What is the probability of being in each of the states 1 through 5 after four steps beginning in state 1? (4)

$$P = \begin{bmatrix} 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0.8 & 0.2 & 0 & 0 \\ 0.4 & 0 & 0.6 & 0 & 0 \end{bmatrix}$$

2) A) Let $Y = (Y_1 \quad Y_2 \quad Y_3)^T$ be a random vector with mean vector $\mu = (1 \quad -1 \quad 3)^T$ and covariance matrix $\Sigma = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 2 & 3 \\ 0 & 3 & 10 \end{bmatrix}$. Let $Z = 2Y_1 - 3Y_2 + Y_3$. Find $E(Z)$ and $Var(Z)$ (3)

B) Given the following information regarding the variables, determine the communalities and the specific variances corresponding to each of them. (3)

Variables	V1	V2	V3	V4	V5
Loadings lambda_1	0.969	0.519	0.785	0.971	0.704
Loadings lambda_2	-0.231	0.807	-0.587	-0.210	0.667

C) Consider a sample of size 20 and we have two variables namely X_1 and X_2 with sample means 71.45 and 164.7. Let the sample originate from $N_2(\mu, \Sigma)$ where $\Sigma = \begin{bmatrix} 20 & 100 \\ 100 & 1000 \end{bmatrix}$. Test the hypothesis $H_0: \mu = (70 \quad 170)^T$. (4)

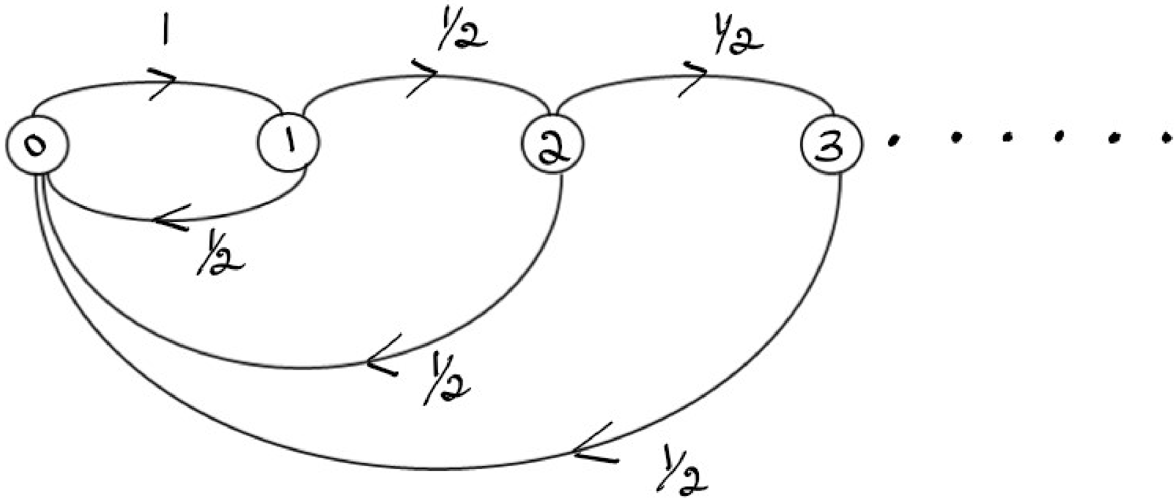
The cutoff value at 5% level of significance is 5.99.

3) A) Let X have a bivariate normal distribution $N(O, \Sigma)$ with $\Sigma = \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}, \rho > 0$. Determine the eigen values and the corresponding eigen vectors of Σ and using this information, obtain the first and the (5)

second principal components of $X.$ Also, find the variances of the two principal components.

B) If the data matrix $X = \begin{bmatrix} 4 & 5 & 10 \\ 5 & 4 & 12 \\ 6 & 6 & 15 \end{bmatrix}$, find the covariance matrix S. (2)

C) Consider the Markov chain with state space $S = \{0,\ 1,\ 2,\ 3,\ \ldots\ldots\}$ and the state transition diagram as shown below. Find the stationary distribution, if exists.

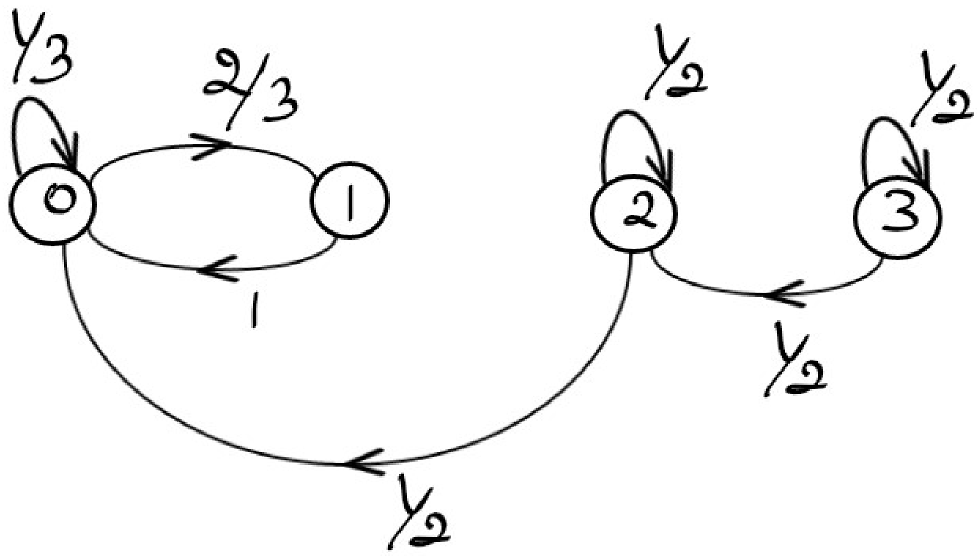


(3)

.

4) The state transition diagram for a Markov chain with state space $S = \{0, 1, 2, 3\}$ is given below. Classify the states. Identify the recurrent and transient states. Are there any positive recurrent states? Find out the periodicity of recurrent states, if exist.

A)



(3)

.

B) Let $\{X_n\ , n = 0, 1, 2, \cdots\}$ be a Markov chain with state space $S = \{0, 1\}$ and one-step transition probability matrix $P = \begin{pmatrix} 0.1 & 0.9 \\ 0.3 & 0.7 \end{pmatrix}$. Let the initial distribution be $P(0) = (0.5, 0.5)$. Find the value of the probability $P(X_3 = 0)$. (3)

C) Consider a Gambler's ruin problem with $p = 0.4$ and $n = 6$, given the following probabilities, $p_{12} = p_{23} = p_{34} = p_{45} = p_{56} = 0.4$, and $p_{10} = p_{21} = p_{32} = p_{43} = p_{54} = 0.6$. Starting in state 3, determine (4)

a. the expected amount of time spent in state 3 and the expected amount of time spent in state 2

b. the probability of ever visiting state 4.

5) Let $X = \begin{bmatrix} X_1 \\ X_2 \end{bmatrix}$ be a bivariate normal random vector and let $X \sim N_2\left(\begin{bmatrix} 5 \\ 10 \end{bmatrix}, \begin{bmatrix} 16 & 12 \\ 12 & 36 \end{bmatrix}\right)$.

A) (a) Find the distribution of $X_1 + X_2$ and $X_1 - X_2$. (3)

(b) Are $X_1 + X_2$ and $X_1 - X_2$ independent? Justify your answer.

B) Let $Y = \begin{bmatrix} Y_1 \\ Y_2 \\ Y_3 \end{bmatrix} \sim N_3(\mu, \Sigma)$ where $\mu = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$ and $\Sigma = \begin{bmatrix} 22 & 4 & -2 \\ 4 & 3 & -2 \\ -2 & -2 & -2 \end{bmatrix}$. Find the conditional distribution of Y_1 given that $Y_2 = 0$ and $Y_3 = 1$. (3)

C) Consider the hypothetical distances between pairs of five objects as follows: (4)

Objects	1	2	3	4	5
1	0				
2	9	0			
3	3	7	0		
4	6	5	9	0	
5	11	10	2	8	0

Using complete linkage clustering identify the clusters and find the final distance matrix. Draw the dendrogram for this complete linkage clustering.

-----End-----