Unit-4: Markov Chain

Concept of Markov Chain:

Important classes of stochastic processes are Markov chains and Markov processes. A Markov chain is a discrete-time process for which the future behavior (if given the past and the present) is only depends on the present and not on the past. A Markov process is the continuous-time version of a Markov chain. Many queuing models are in fact Markov processes.

A Markov chain is named after Russian mathematician Andrey Markov. It is a mathematical system that undergoes transitions from one state to another, between a finite or countable number of possible states. It is a random process characterized as memory-less it means the next state depends only on the current state and not on the sequence of events that preceded it. This specific kind of "memory-less-ness" is called the Markov property.

The probability that the process moves from any given state to any other particular state is always same regardless (without considering) of the history of the process.

A Markov chain consists of state and transition probabilities. Each transition probabilities are the probability of moving from one state to another in one step. The transition probabilities are independent of the past and depend only on the two states involved (present and next). The matrix of transition call probabilities are called transition matrix. Markov modeling is an extremely important to the field of modeling and analysis of telecommunication networks.

Use of Markov Chain:

Markov chain is mostly used

- To predict the behavior of customer in terms of brand loyalty and switching pattern
- To predict the state of machine (working or not working) used in manufacture a product.
- To predict weather of future either rain or not rain

Key features of Markov chain:

A sequence of trail of an experiment is a Markov chain if:

1. Finite numbers of states
2. The outcome of each experiment is one of a set of discrete state
3. Only one state is possible at a time
4. The outcome of the experiment depends only on the present state and not on the past state.
5. The transition probability remains constant from one to the next.

Transition Probability in Markov Chain:

The probability of moving from one state to another state or remaining in the same state during a single time period is called transition probability.

Transition probability is mathematically expressed as:

P_{ij} = P (Next State S_j at t = 1 | Initial State S_i at t = 0)

Where,

P_{ij} = Probability of initial and next state

t = time

Transition Probability Matrix (TPM):

The matrix which is used to predict the movement of system from one state to another state is called TPM.

$$P = \begin{array}{cc} & \text{Next state}(j)\,[n=1] \\ & \begin{array}{ccccc} S_1 & S_2 & S_3 & \cdots & S_n \end{array} \\ \begin{array}{cc} & S_1 \\ \text{Initial state}(i) & S_2 \\ [n=0] & S_3 \\ & \vdots \\ & S_n \end{array} & \begin{bmatrix} P_{11} & P_{12} & P_{13} & \cdots & P_{1n} \\ P_{21} & P_{22} & P_{23} & & P_{2n} \\ P_{31} & P_{32} & P_{33} & & P_{3n} \\ \vdots & & & & \\ P_{n1} & P_{n2} & P_{n3} & & P_{nn} \end{bmatrix} \end{array}$$

$P_{11} = P(S_1 \text{ at } t=1 \mid S_1 \text{ at } t=0)$

$P_{12} = P(S_2 \text{ at } t=1 \mid S_1 \text{ at } t=0)$

$P_{21} = P(S_1 \text{ at } t=1 \mid S_2 \text{ at } t=0)$

Similarly

$P_{nn} = P(S_n \text{ at } t=1 \mid S_n \text{ at } t=0)$

Here, we are moving on step at a time, so this is called ***1-Step Transition Probability.***

$R_1 = R_0 P$

Where, R_0 is the initial round of probability. R_1 is round 1 probability. P is the TPM.

2-Step Transition Probability:

In this transition we move 2-Step forward.

$$
P^2 = \begin{array}{c} \\ \text{Initial state(i) [n=0]} \end{array}
\begin{array}{c|cccc|}
 & \multicolumn{4}{c}{\text{Next state(j) [n=1]}} \\
 & S_1 & S_2 & S_3 & S_n \\
S_1 & P_{11}^2 & P_{12}^2 & P_{13}^2 & P_{1n}^2 \\
S_2 & P_{21}^2 & P_{22}^2 & P_{23}^2 & P_{2n}^2 \\
S_3 & P_{31}^2 & P_{32}^2 & P_{33}^2 & P_{3n}^2 \\
\vdots & \vdots & & & \\
S_n & P_{n1}^2 & P_{n2}^2 & P_{n3}^2 & P_{nn}^2
\end{array}
$$

$P^2_{11} = P(S_1 \text{ at } t=2 \mid S_2 \text{ at } t=0)$

$R_2 = R_1 P$

Or,

$R_2 = R_0 P^2$

N-Step Transition Probability:

In this transition we move N-Step forward.

Next state(j) [n=1]

$$P^n = \begin{array}{c c} & \begin{array}{ccccc} S_1 & S_2 & S_3 & \cdots & S_n \end{array} \\ \begin{array}{c} \text{Initial state(i) [n=0]} \end{array} \begin{array}{c} S_1 \\ S_2 \\ S_3 \\ \vdots \\ S_n \end{array} & \begin{bmatrix} P^n_{11} & P^n_{12} & P^n_{13} & \cdots & P^n_{1n} \\ P^n_{21} & P^n_{22} & P^n_{23} & \cdots & P^n_{2n} \\ P^n_{31} & P^n_{32} & P^n_{33} & \cdots & P^n_{3n} \\ \vdots & & & & \\ P^n_{n1} & P^n_{n2} & P^n_{n3} & \cdots & P^n_{nn} \end{bmatrix} \end{array}$$

$P^n_{11} = P(S_1 \text{ at } t=n \mid S_2 \text{ at } t=0)$

$R_n = R_{n-1} P$

Or,

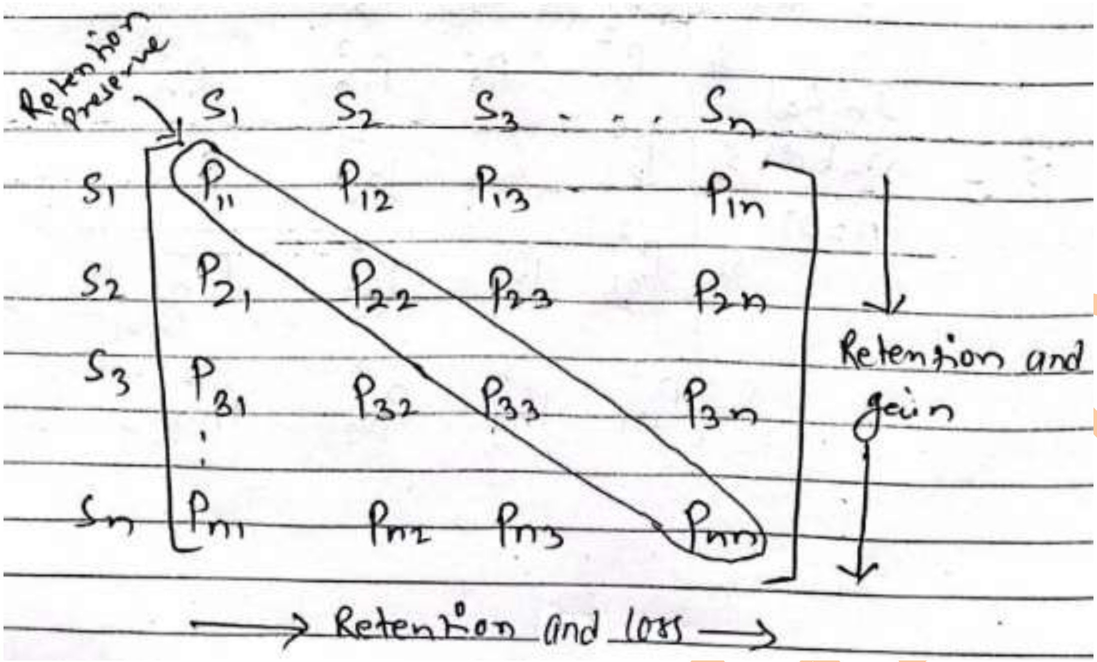
$R_n = R_0 P^n$

Assumption of TPM:

1. Sum of row element of TPM = 1
2. Each element of TPM is probability, hence all the value should be non-negative and range from 0 to 1 i.e. ($0 <= P_{ij} <= 1$)
3. This is a square matrix because row shows initial state and col shows next state

Retention Loss and Retention Gain:

- If you move row wise Retention will be loss
- If you move col wise Retention will be gain
- If you move diagonally Retention will be preserved



Example 1:

In a certain market, only two brands of cold drinks Coke and Pepsi are sold. Consider a customer last purchased Coke, there is 80% chance that he would buy the same brand in the next purchase. While if a customer purchases Pepsi, there is 90% chance that his next purchase would be Pepsi. Using this information

a. Develop Transaction Probability Matrix (TPM)
b. Interpret the state transition matrix in terms of Retention, loss and gain
c. Draw Transition Diagram

Solution:

Here given that,

If Customer purchase Coke now, then

Probability of buy Coke in next purchase = 0.8

Probability of buy Pepsi in next purchase = 0.2

If Customer purchase Pepsi now, then

Probability of buy Pepsi in next purchase = 0.9

Probability of buy Coke in next purchase = 0.1

a) TPM:

Next Purchase [n=1]

$$
P = \begin{array}{c|cc}
 & \text{Coke} & \text{Pepsi} \\
\hline
\text{Initial Purchase (n=0) Coke} & 0.8 & 0.2 \\
\text{Pepsi} & 0.1 & 0.9
\end{array}
\qquad
P = \begin{bmatrix} 0.8 & 0.2 \\ 0.1 & 0.9 \end{bmatrix}
$$

b) Retention, loss and gain

$P_{11} = P_{\text{Coke Coke}}$ = P(Coke at t=1 | Coke at t = 0)

= 0.8

Hence, there is 80% chance to retention of Coke.

$P_{12} = P_{\text{Coke Pepse}}$ = P(Pepsi at t=1 | Coke at t = 0)

= 0.2

Hence, there is 20% chance to loss of Coke. That means 20% chance to gain of Pepsi.

$P_{21} = P_{\text{Pepsi Coke}}$ = P(Coke at t=1 | Pepsi at t = 0)

= 0.1

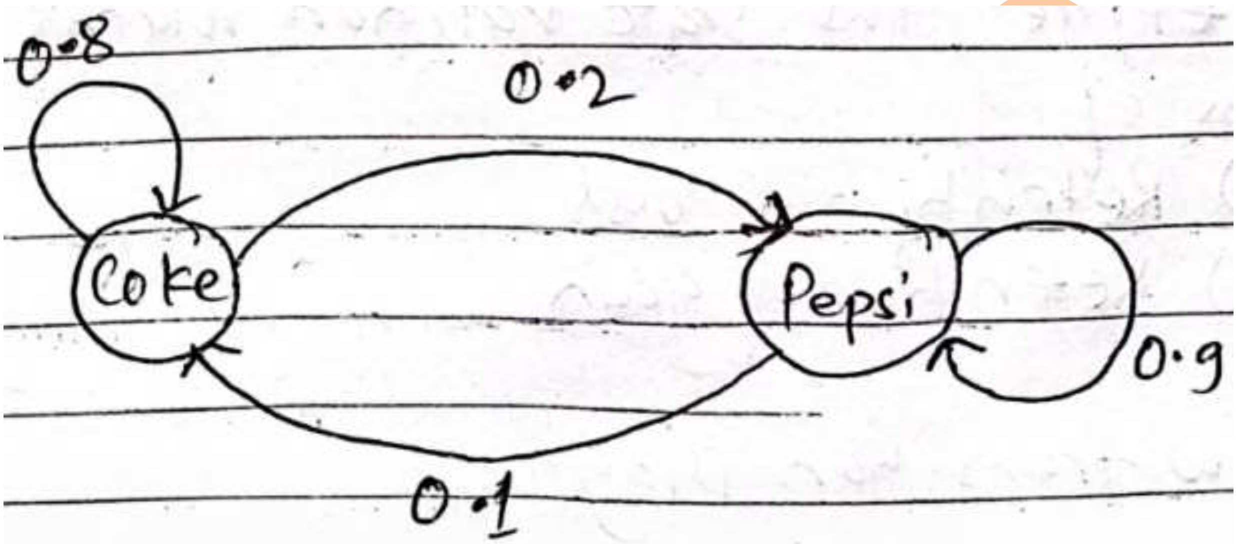
Hence, there is 10% chance to loss of Pepsi. That means 10% chance to gain of Coke.

$P_{22} = P_{Pepsi\ Pepsi} = P(\text{Pepsi at } t=1 \mid \text{Pepsi at } t = 0)$

$= 0.9$

Hence, there is 90% chance to retention of Pepsi.

c. **Transition Diagram**



Here we can move from Coke to Pepsi and Pepsi to Coke. Hence this is called non-absorbing state.

Note: If we cannot move in all the States then this is called absorbing state.

Example 2:

Q) The school of International studies for population found out by its survey that the mobility of a population of a state to the village, town, and city is in the following percentage:

$$
\begin{array}{ll}
 & \text{To} \\
 & \begin{array}{ccc} \text{Village} & \text{Town} & \text{City} \end{array} \\
\text{From} \begin{array}{r} \text{Village} \\ \text{Town} \\ \text{City} \end{array} & \begin{bmatrix} 50\% & 30\% & 20\% \\ 10\% & 70\% & 20\% \\ 10\% & 40\% & 50\% \end{bmatrix}
\end{array}
$$

a) Interpret the state rasition matrix in terms of

i) Retention & loss

ii) Retention & gain

b) Draw transition Diagram

Solution:

a) Retention, loss and gain

P_{11} = 50 % retention to village

P_{12} = 30 % loss to village and 30% gain to town

P_{13} = 20 % loss to village and 20% gain to city

P_{21} = 10 % gain to village and 10% loss to town

P_{22} = 70 % retention to town

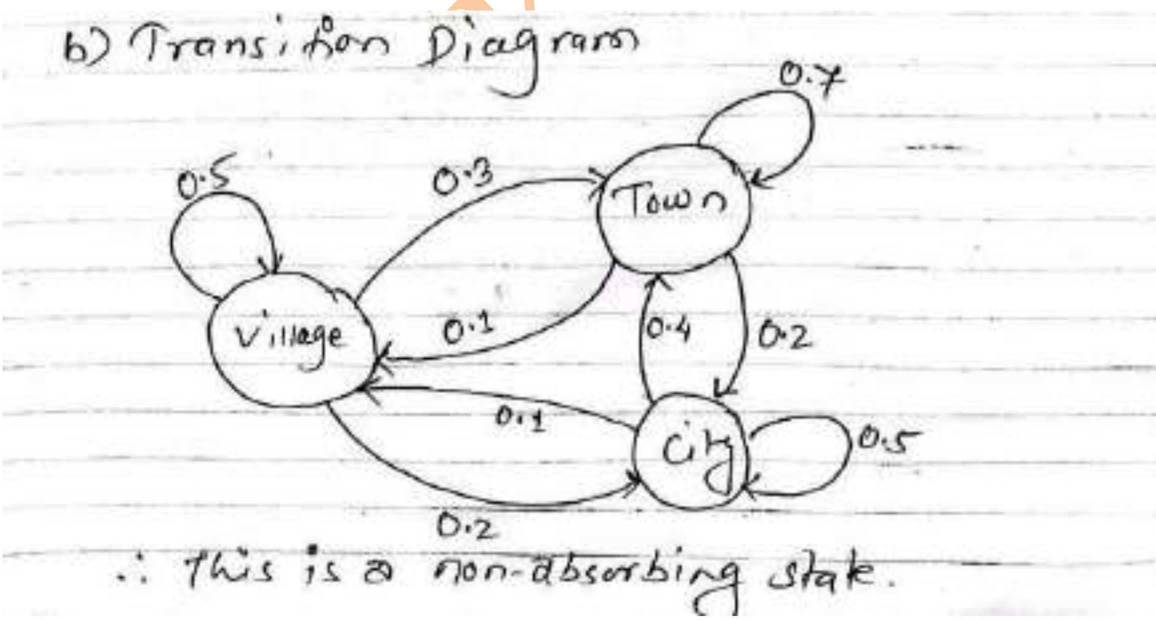
P_{23} = 20 % loss to town and gain to city

P_{31} = 10 % loss to city and gain to village

P_{32} = 40 % loss to city and gain to town

P_{33} = 50 % retention to city

b) Transition Diagram



Example 3:

There is a probability of weather as:

Rainy Today => 40% Rainy Tomorrow

=> 60% Not Rainy Tomorrow

Not Rain Today => 20% Rain Tomorrow

=> 80% Not Rain Tomorrow

a. What will be probability if today is raining then not rain the day after tomorrow?

b. What will be the probability if today is raining then rain day after tomorrow?

c. What will be the probability if today is not raining then not raining after 2 days?

d. What will be the probability if today is not raining then raining after 2 days?

e. Draw the Markov Chain Diagram

Solution:

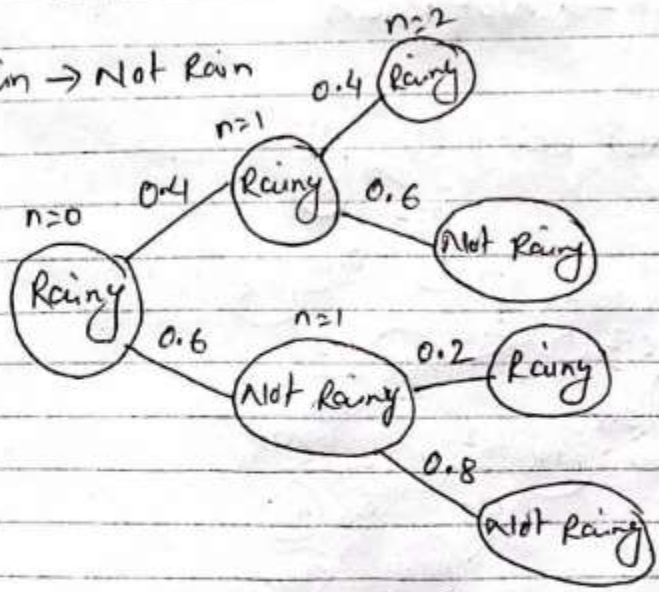
TPM of given information:

Next stage(j) [n=1]

$$P = \begin{array}{cc} \text{Initial stage(i)} \\ n=0 \end{array} \begin{array}{c} \\ \text{Raing} \\ \text{Not Rainy} \end{array} \begin{array}{c} \begin{array}{cc} \text{Rainy} & \text{Not Rainy} \end{array} \\ \begin{bmatrix} 0.4 & 0.6 \\ 0.2 & 0.8 \end{bmatrix} \end{array}$$

a) Here, we have to calculate probability of not rain day after tomorrow i.e. n=2 (not rainy) by considering today is rain i.e. n=0 (rainy)

a) Rain → Not Rain



$$P_{12} = (0.4 \times 0.6) + (0.6 \times 0.8)$$

$$= 0.24 + 0.48$$

$$= 0.72$$

∴ The probability of not raining day after tomorrow is 0.72.

Alternative

Matrix method

a) $R = [R \quad NR]$

$R_0 = [1 \quad 0]$ today is raining & 0% not rainy

$R_1 = R_0 P$

$= [1 \quad 0] \begin{bmatrix} 0.4 & 0.6 \\ 0.2 & 0.8 \end{bmatrix}$

$= [0.4 \times 1 + 0.2 \times 0 \qquad 0.6 \times 1 + 0.8 \times 0]$

$R_1 = [0.4 \quad 0.6]$

$R_2 = R_1 . P$

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

$= \begin{bmatrix} (0.4 \times 0.4)+(0.6 \times 0.2) & (0.4 \times 0.6)+(0.6 \times 0.8) \end{bmatrix}$

$= \begin{bmatrix} 0.16 + 0.12 & 0.24 + 0.48 \end{bmatrix}$

$= \begin{bmatrix} 0.28 & 0.72 \end{bmatrix}$

$\therefore$ The probability of not rain day after tomorrow is 0.72.

The probability of rain day after tomorrow is 0.28.

b) Here we have to calculate probability of rainy day after tomorrow i.e. n=2(rainy) by considering today is rainy i.e. n=0 (rainy)

b) Rain → Rain

$P_{11} = (0.4 \times 0.4) + (0.6 \times 0.2)$

$= 0.16 + 0.12$

$= 0.28$

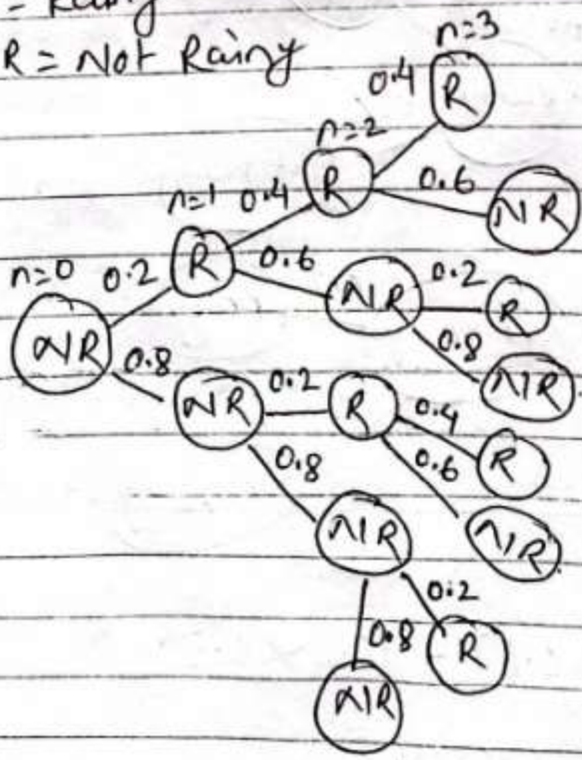
$\therefore$ The probability of raining day after tomorrow is 0.28.

c) Here, we have to calculate probability of not rainy after 2 days i.e. n=3 (not rainy) by considering today is not rainy i.e. n=0(not rainy)

c) Not Rain → Not Rain

R = Rainy

NR = Not Rainy



$$P_{22} = (0.2 \times 0.4 \times 0.6) + (0.2 \times 0.6 \times 0.8) + (0.8 \times 0.2 \times 0.6) + (0.8 \times 0.8 \times 0.8)$$

$$= 0.048 + 0.096 + 0.096 + 0.512$$

$$= 0.75$$

Hence, the probability of not rainy after 2 days is 0.75

a) Here, we have to calculate probability of rainy after 2 days i.e. n=3 (rainy) by considering today is not rainy i.e. n=0 (not rainy)

d) Not Rain → Rain

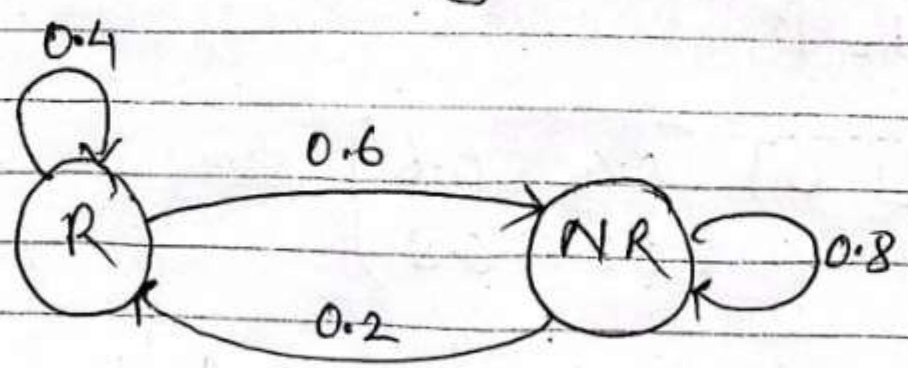
$$P_{21} = (0.2\times0.4\times0.4)+(0.2\times0.6\times0.2)+(0.8\times0.2\times0.4)+(0.8\times0.8\times0.2)$$

$$= 0.032 + 0.024 + 0.064 + 0.128$$

$$= 0.25$$

∴ The probability of rain after 2 days is 0.25

e) Markov chain Diagram



Class Work:

Suppose 60% of all people ride bike and 40% car. What fraction of people will be ridding bike after 3 years from now? [Hint: find probability of riding bike at n=4]

Applications of Markov Chain or Process:

Physics:

Markovian systems appear extensively in thermodynamics and statistical mechanics, whenever probabilities are used to represent unknown or un-modeled details of the system, if it can be assumed that the dynamics are time-invariant, and that no relevant history need be considered which is not already included in the state description.

Queuing theory:

Markov chains are the basis for the analytical treatment of queues (queuing theory). Agner Krarup Erlang initiated the subject in 1917. This makes them critical for optimizing the performance of telecommunications networks, where messages must often compete for limited resources (such as bandwidth).

Internet applications:

The Page Rank of a webpage as used by Google is defined by a Markov chain. It is the probability to be at page i in the stationary distribution on the following Markov chain on all (known) web pages

Statistics:

Markov chain methods have also become very important for generating sequences of random numbers to accurately reflect very complicated desired probability distributions, via a process called Markov chain Monte Carlo (MCMC) And many more.

Assignment:

1. Define and describe Markov chain in detail with the help of suitable examples. Also describe at least three of application of Markov chain.
2. Define a Markov chains and its application
3. What are the key features of Markov chains

End of Unit-4