BACKGROUND

Markov chains are an important mathematical tool in stochastic processes. The underlying idea is the Markov Property, in order words, that some predictions about stochastic processes can be simplified by viewing the future as independent of the past, given the present state of the process. This is used to simplify predictions about the future state of a stochastic process.

Andrei Markov was a Russian mathematician who lived between 1856 and 1922. He was a poorly performing student and the only subject he didn' t have difficulties in was mathematics. He later studied mathematics at the university of Petersburg and was lectured by Pafnuty Chebyshev, known for his work in probability theory. Markov' s first scientific areas were in number theory, convergent series and approximation theory. His most famous studies were with Markov chains, hence the name and his first paper on the subject was published in 1906. He was also very interested in poetry and the first application he found of Markov chains was in fact in a linguistic analysis of Pusjkins work *Eugene Onegin.*

INTRODUCTION TO MARKOV CHAINS

Markov chains are a fundamental part of stochastic processes. They are used widely in many different disciplines. A Markov chain is a stochastic process that satisfies the Markov property, which means that the past and future are independent when the present is known. This means that if one knows the current state of the process, then no additional information of its past states is required to make the best possible prediction of its future. This simplicity allows for great reduction of the number of parameters when studying such a process.

In mathematical terms, the definition can be expressed as follows:

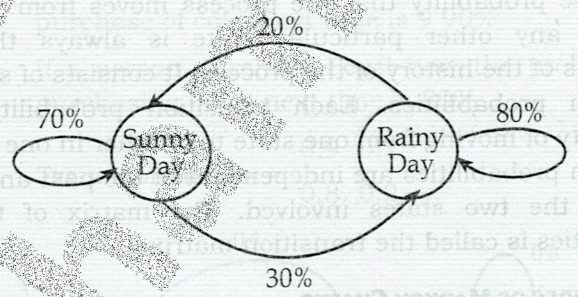
A stochastic process $X = \{X_n, n \in N\}$ in a countable space S is a discrete-time Markov chain if:

For all $n \geq 0$, $X_n \in S$

For all $n \geq 1$ and for all $i_0, \ldots i_{n-1}, i_n \in S$, we have :

$$P\{X_n = i_n \mid X_{n-1} = i_{n-1}, \ldots, X_0 = i_0\} = P\{X_n = i_n \mid X_{n-1} = i_{n-1}\}$$

Markov chains are used to compute the probabilities of events occurring by viewing them as states transitioning into other states, or transitioning into the same state as before. We can take weather as an example: If we arbitrarily pick probabilities, a prediction regarding the weather can be the following: If it is a sunny day, there is a 30% probability that the next day will be a rainy day, and a 20% probability that if it is a rainy day, the day after will be a sunny day. If it is a sunny day, there is therefore a 70% chance that the next day will be another sunny day, and if today is a rainy day, there is 80% chance that the next day will be a rainy day as well. This can be summarized in a transition diagram, where all of the possible transitions of states are described:



To approach this mathematically one views today as the current state, S_0, which is a $1 \times m$ vector. The elements of this vector will be the current state of the process. In our weather example, we define $S = [Sunny\ Rainy]$. Where S is called our state space, in which all the elements are all the possible states that the process can attain. If, for example, today is a sunny day, then the so vector will be $S_0 = [1\ 0]$, because there is 100% chance of a sunny day and zero chance of it being a rainy day. To get to the next state, the transition probability matrix is required, which is just the state transition probabilities summarized in a matrix. In this case it will be as follows:

$$P = \begin{array}{c|cc} & S & R \\ \hline S & 0.7 & 0.3 \\ R & 0.2 & 0.8 \end{array} \quad \text{or more generally in this case:}$$

$$P = \begin{array}{c|cc} & S & R \\ \hline S & \alpha & 1-\alpha \\ R & \beta & 1-\beta \end{array}$$

To get to the next state, S_1, you simply calculate the matrix product $S_1 = S_0P$. Since calculations for successive states of S is only of the type $S_n = S_{n-1}P$, the general formula for computing the probability of a process ending up in a certain state is $S_n = S_0P_n$.This allows for great simplicity when calculating the probabilities far into the future. For example, if today is a sunny day then the state vector 120 days from now, S^{120} , is S^{120} = [0.4 0.6] . [3]

So, A Markov chain is a mathematical model for a process which moves step by step through various states. In a Markov chain, the probability that the process moves from any given state to any other particular state is always the same, regardless of the history of the process. It consists of states and transition probabilities. Each transition probability is 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. The matrix of transition probabilities is called the transition matrix.

KEY FEATURES OF MARKOV CHAINS

A sequence of trials of an experiment is a Markov chain if

- the outcome of each experiment is one of a set of discrete states.
- the outcome of an experiment depends only on the present state, and not on any past states.
- the transition probabilities remain constant from one transition to the next.

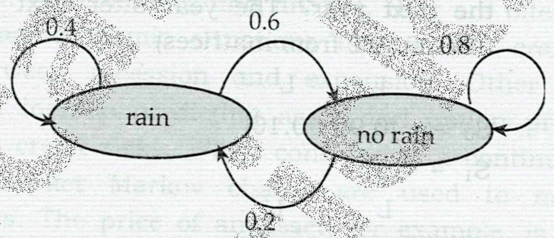
PROCESS EXAMPLES

1. **Weather**

   a. If it rains today there is 40% probability of raining tomorrow

   b. If it does not rains today there is 20% probability of raining tomorrow

   $$\text{Transition Matrix} = \begin{matrix} 0.4 & 0.6 \\ 0.2 & 0.8 \end{matrix}$$

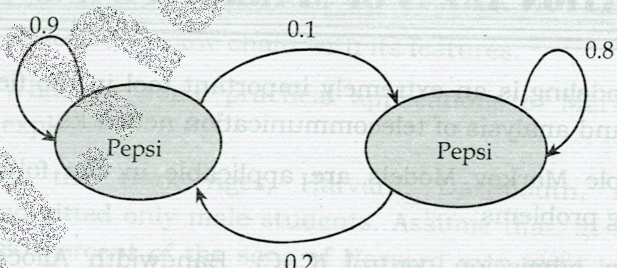


2. **Coke and Pepsi Example**

   a. If a person purchase coke now the probability of purchase of coke next time is 9 90%

   b. If a person purchase Pepsi now the probability of purchasing Pepsi next time is 80%

   $$\text{Transition Matrix} = \begin{matrix} 0.9 & 0.1 \\ 0.8 & 0.2 \end{matrix}$$



3. An insurance company classifies drivers as low-risk if they are accident free for one year. Past records indicate that 98% of the drivers in the low-risk category (L) will remain in that category the next year, and 78% of the

drivers who are not in the low risk category (L') one year will be in the low-risk category the next year.

a. Find the transition matrix, P

$$P = \begin{array}{c|cc} & L & L' \\ \hline L & 0.98 & 0.02 \\ L' & 0.78 & 0.22 \end{array}$$

b. If 90% of the drivers in the community are in the low-risk category this year, what is the probability that a driver chosen at random from the community will be in the low-risk category the next year? The year after next ? (answer 0.96, 0.972 from matrices)

$$\begin{array}{ccc} & L & L' \\ S_0 = & [0.90 & 0.10] \end{array}$$

$S_1 = S_0P$

$$\begin{array}{ccc} & L & L' \\ S_1 = & [0.96 & 0.04] \end{array}$$

$S_2 = S_1$

$= S_0P.P$

$= S_0P_L^2$

$$\begin{array}{ccc} & L & L' \\ S_2 = & [0.972 & 0.028] \end{array}$$

APPLICATION AREAS OF MARKOV CHAINS

Markov Modeling is an extremely important tool in the field of modeling and analysis of telecommunication networks.

For example Markov Models are applicable in the following networking problems:

Connection admission control (CAC), Bandwidth Allocation, Routing, Queuing and scheduling

Markov chains are used to analyze trends and predict the future such as weather, stock market, genetics, product success, etc.

Markov models have also been used in internet application to analyze web navigation behavior of users. A user's web link transition on a particular website can be modeled using first- or second order. Markov models and can be used to make predictions regarding future navigation and to personalize the web page for an individual user.

Since Markov chains can be designed to model many real-world processes, they are used in a wide variety of situations. These fields range from the mapping of animal life populations to search-engine algorithms, music composition and speech recognition. In economics and finance, they are often used to predict macroeconomic situations like market crashes and cycles between recession and expansion. Other areas of application include predicting asset and option prices, and calculating credit risks. When considering a continuous-time financial market Markov chains are used to model the randomness. The price of an asset, for example, is set by a random factor - a stochastic discount factor - which is defined using a Markov chain.

DISCUSSION EXERCISE

1. Describe Markov chain with its features.
2. What are some practical applications of Markov chain explain?
3. In the Dark Ages, Harvard, Dartmouth, and Yale admitted only male students. Assume that, at that time, 80 percent of the sons of Harvard men went to Harvard and the rest went to Yale, 40 percent of the sons of Yale men went to Yale, and the rest split evenly between Harvard and Dartmouth; and of the sons of Dartmouth men, 70 percent went to Dartmouth, 20 percent to

Harvard, and 10 percent to Yale. (i) Find the probability that the grandson of a man from Harvard went to Harvard. (ii) Modify the above by assuming that the son of a Harvard man always went to Harvard. Again, find the probability that the grandson of a man from Harvard went to Harvard.

4. Assume that a man's profession can be classified as professional, skilled laborer, or unskilled laborer. Assume that, of the sons of professional men, 80 percent are professional, 10 percent are skilled laborers, and 10 percent are unskilled laborers. In the case of sons of skilled laborers, 60 percent are skilled laborers, 20 percent are professional, and 20 percent are unskilled. Finally, in the case of unskilled laborers, 50 percent of the sons are unskilled laborers, and 25 percent each are in the other two categories. Assume that every man has at least one son, and form a Markov chain by following the profession of a randomly chosen son of a given family through several generations. Set up the matrix of transition probabilities. Find the probability that a randomly chosen grandson of an unskilled laborer is a professional man.

❑❑❑