BHARATHIDASAN UNIVERSITY

Tiruchirappalli- 620024

Tamil Nadu, India.

Programme: M.Sc. Statistics

Course Title: Stochastic Processes and Time Series Analysis

Course Code: 23ST02DEC

Unit-I

Stochastic Processes

Dr. T. Jai Sankar
Associate Professor and Head
Department of Statistics

Ms. I. Angel Agnes Mary
Guest Faculty
Department of Statistics

<ins>UNIT - I</ins>

STOCHASTIC PROCESS

Stochastic:

The word stochastic comes from the Greek word *stokhazesthai* meaning to aim or guess. The definition of stochastic is random or involving chance or probability. When the order of events is randomly determined and it cannot be predicted what will come first, second or third, this is an example of when the order of events would be described as stochastic.

Stochastic Model:

A stochastic model represents a situation where uncertainty is present. In other words, it's a model for a process that has some kind of randomness. In the real word, uncertainty is a part of everyday life, so a stochastic model could literally represent *anything*. The opposite is a deterministic model, which predicts outcomes with 100% certainty. Deterministic models always have a set of equations that describe the system inputs and outputs exactly. On the other hand, stochastic models will likely produce different results every time the model is run.

All stochastic models have the following in common:

- They reflect all aspects of the problem being studied,
- Probabilities are assigned to events within the model,
- Those probabilities can be used to make predictions or supply other relevant information about the process.

Steps for Building a Stochastic Model

The basic steps to build a stochastic model are:

- Create the sample space (Ω) — a list of all possible outcomes,
- Assign probabilities to sample space elements,
- Identify the events of interest,
- Calculate the probabilities for the events of interest.

A very simple example of this process in action: You are rolling a die in a casino. If you roll a six or a one, you win $10. The steps would be:

1. The sample space includes all possibilities for die roll outcomes: $\Omega = \{1,2,3,4,5,6\}$.
2. The probability for any number being rolled is 1/6.
3. The event of interest is "roll a 6 or roll a 1".
4. The probability for "roll a 6 or 1" is 1/6 + 1/6 = 2/6 = 1/3.

Differences between Behaviors of Deterministic and Stochastic Models

Deterministic Model	Stochastic Model
Variables are functions of time only	Variables depend on time and probability
All mechanisms are described precisely	Process could have two sources of variability (demographic and environmental)
Captures only mean characteristics of process	Captures variations from mean behavior
Deterministic path provides expected value	Stochasticity leads to variances and covariances
Trajectory is fixed between simulations	Variability between simulations
For given parameter set, one simulation is sufficient	Needs many simulations
Behavior entirely governed by parameters	Allows "chance" to play a role
If we knew perfectly the present state, we could predict future states accurately	If we knew perfectly the present state, model assigns only a probability distribution to future states
Perfect reproducibility	Each realization is different
Deterministic dynamics, in general, have equilibrium behavior	Stochasticity can excite system to sustained oscillations (resonance) or can drive system to extinction
Mathematically easier	Often harder to analyze mathematically

Random Experiment

A random experiment is a physical situation whose outcome cannot be predicted until it is observed.

Sample Space

A sample space, Ω, is a set of possible outcomes of a random experiment.

Example

Random experiment: Toss a coin once.

Sample space: Ω = {head, tail}

Random Variable

A random variable, X, is defined as a function from the sample space to the real numbers: $X : \Omega \to R$. That is, a random variable assigns a real number to every possible outcome of a random experiment.

Example

Random experiment: Toss a coin once.

Sample space: Ω = {head, tail}.

An example of a random variable: $X : \Omega \to R$ maps "head" → 1, "tail" → 0.

Essential point

A random variable is a way of producing random real numbers.

Stochastic Process or Random Process

A ***stochastic process*** or ***random process*** is a collection of random variables, representing the evolution of some system of random values over time.

A family (or collection) of random variables that are indexed by a parameter, such as time, is called a stochastic process.

Definition

A stochastic process is a family of random variables, $\{X(t) : t \in T\}$, where t usually denotes time. That is, at every time t in the index set T may discrete ($T = \{0, 1, 2, 3, 4, ...\}$) or continuous ($T = [0, \infty]$).

Example of Stochastic Process

Suppose we toss a six-sided dies several times, and is interested in the number which appears at the n^{th} toss. Let X(1) denote the number which appears at the first toss, X(2), the number which appears at the second toss, and so on. It can be seen that, we can describe the outcome of this experiment by defining a family of random variables $\{X(t); t \in T\}$, where $T = \{1, 2, 3, . . .\}$ and X(t) is the number which appears at the n^{th} toss of the die. This family or collection of random variables, indexed by the parameter n, is an example of a stochastic process. In this example, t is called the index parameter of the stochastic process, while T is called the index set of the stochastic process.

Discrete-time process

$\{X(t) : t \in T\}$ is a discrete-time process if the set T is finite or countable. This is generally means $T = \{0, 1, 2, 3, . . .\}$

Thus a discrete-time process is $\{X(0), X(1), X(2), X(3), . . .\}$: a random number associated with every time 0, 1, 2, 3, . . .

Continuous-time process

$\{X(t) : t \in T\}$ is a continuous-time process if T is not finite or countable. This is generally means $T = [0, \infty)$, or $T = [0, K]$ for some K.

Thus a continuous-time process $\{X(t) : t \in T\}$ has a random number X(t) associated with every instant in time.

(Note that X(t) need not change at every instant in time, but it is allowed to change at any time; i.e. not just at t = 0, 1, 2, . . . , like a discrete-time process.)

State Space

The state space, S, is the set of real values that X(t) can take. Every X(t) takes a value in R, but S will often be a smaller set: $S \subseteq R$. For example, if X(t) is the outcome of a coin tossed at time t, then the state space is S = {0, 1}. The state space S is discrete if it is finite or countable. Otherwise it is continuous.

Types of stochastic Processes

There are four types of stochastic processes.

- Discrete or Continuous state processes
- Markov Processes
- Birth-death Processes
- Poisson Processes

Classification of Stochastic process:

Stochastic process is classified in four categories on the basis of state space and time space:

- Stochastic Processes with Discrete Parameter and State Spaces
- Stochastic Processes with Continuous Parameter and Discrete State Space
- Stochastic Processes with Discrete Parameter and Continuous State Space
- Stochastic Processes with Continuous Parameter and State Spaces

Stochastic Processes with Discrete Parameter and State Spaces

If both t and X_t belongs to N, the set of natural numbers, then we have models like markov chain. For Example, if X_t means the fit (0 or 1) in position t of a sequence of transmitted bits, then X_t can be modelled as markov chain with two states. This leads to the error correction algorithm in data transmission,

A Brand-Switching Model for Consumer Behavior

Before introducing a new brand of coffee, a manufacturer wants to study consumer behavior relative to the brands already available in the market. Suppose there are three brands already available in the market. Suppose there are three brands on sale, say A, B, C. The consumers either buy the same brand for a few months or change their brands every now and then. There is also a strong possibility that when a superior brand is introduced, come of the old brands will be left with only a few customers. Sample surveys are used to gauge consumer behavior.

In such a survey, conducted over a period of time, suppose the estimates obtained for the consumer brand-switching behavior are as follows: Out of those why buy A is one month, during the next months 60% buy A again, 30% switch to brand B and 10% switch to brand C. For brands B and C these figures are, B to A 50%, B to B 30%, B to C 20%, C to A 40%, C to B 40%, C to C 20%.

If we are interested in the number of people who buy a certain brand of coffee, then that number could be represented as a stochastic process. The behavior of the consumer can also be considered a stochastic process that can enter three different states A, B, C. Some of the questions that arise are: What is the expected number of months that a consumer stays with one specific brand? What are the mean and variance of the number using a particular brand after a certain number of months? Which is the product preferred most by the customers in the long run? Suppose, for instance, that consumer preferences are observed on a monthly basis. Then we have a discrete-time, discrete-state stochastic process.

Stochastic Processes with Continuous Parameter and Discrete State Space

If the index space I is a finite or infinite interval the sample paths $\{X_t(\omega)\}_{t \in I}$ may have the probability that it is bounded integrals, continuous, differentiable, whether it has a limit at ∞, their probability distribution is an interval.

For example, the number of students waiting for a bus at any time of day in this case the parameter space is continuous.

Consider the size of a population at a given time there again a continuous-time, discrete state stochastic process as the population is finite.

Stochastic Processes with Discrete Parameter and Continuous State Space

If the index set of the process is N and range is R, the sample sequences of the process $\{X_i\}_{i \in N}$, where sample sequence $\{X_i(\omega)\}_{i \in N}$ their arises whether the probability that each sample sequence is bounded, monotonic, limit approaches ∞ and also the series obtained from a sample sequence from f(i) converges and to know the probability distribution of the sum. There application includes markov chain monte carlo.

For example, consider the values of the Dow-Jones Index at the end of the n^{th} week. Then we have a discrete-time stochastic process with the continuous state space $(0,\infty)$.

Jobs of varied length come to a computing center from various sources. The number of jobs arriving, as well as their length, can be said to follow certain distributions. Under these conditions the number of jobs waiting at any time and the time a job has to spend in the system can be represented by stochastic processes.

Under a strictly first-come, first-served policy, there is a good chance of a long job delaying a much more important shorter fob over a long period of time. For the efficient operation of the system, in addition to minimizing the number of jobs waiting and the total delay, it may be necessary to adopt a different service policy.

A round-robin policy in which the service is performed on a single job only for a certain length of time, say 3 or 5 sec, and those jobs that need more service are put back in the queue, is one of the common practices adopted under these conditions. Consider accumulated workload observed at specified points in time.

Stochastic Processes with Continuous Parameter and State Spaces

The problem was concerned with a particle floating on a liquid surface receiving kicks from the molecules of the liquid. The particle then viewed as being subjects are small and close to whether is treated as being continuous and since the particle is constrained to the surface of the liquid by surface tension is at each point is time a vector parallel to the surface. This random force is described by two components stochastic processes to real valued random variables are associated to each point of the two random variables being R, giving the X and Y components of the force. For example, consider waiting time of an arriving job until it gets into service, with the arriving time of the job is the parameter.

Properties of stochastic processes

The classical types of stochastic processes, characterized by different dependence relationship among X(t).

(i) Process with independent increments

Given a stochastic process {X(t)}, if the random variables $X(t_2) - X(t_1)$, $X(t_3) - X(t_2)$, … , $X(t_n) – X(t_{n-1})$, are independent for all choices of $t_1, t_2, \ldots, t_n$, satisfying $t_1 < t_2 < \ldots < t_n$, then we say that {X (t)} is a stochastic process with independent increments.

(ii) Process with the Markovian property

The basic property of a Markov chain is that only the most recent point in the trajectory affects what happens next. This is called the Markov Property. It means that X_{t+1} depends upon X_t, but it does not depend upon $X_{t-1}, \ldots, X_1, X_0$.

Definition

A stochastic process $\{X_t\}$ is a Markov process if

$$P(X_{t+1} = k_{t+1}|X_t = k_t, X_{t-1} = k_{t-1}, \cdots, X_1 = k_1) = P((X_{t+1} = k_{t+1}|X_t = k_t)$$

If a stochastic process $\{X_t\}$ has the Markovian property, then given the present state X_t, the past states $X_1, X_2, \ldots, X_{t-1}$ is not needed to predict the future state X_{t+1}.

(iii) Process with stationary increments

A stochastic process $\{X(t);\ t \in T\}$ is said to have stationary increments if the distribution of the increment, $X(t_1 + h) - X(t_1)$, depends only on the length h of the interval and not on the time t.

For a process with stationary increments, the distribution of $X(t_1 + h) - X(t_1)$ is the same as the distribution of $X(t_2 + h) - X(t_2)$, no matter the values of t_1, t_2 *and* h. If $\{X_t\}$ is a stochastic process with stationary increments, then the distribution of $X(t)$ is the same for each t. This also means that the particular times at which we examine the process is irrelevant.

Markov Model:

In probability theory, a Markov model is a stochastic model used to model randomly changing systems where it is assumed that future states depend only on the present state and not on the sequence of events that preceded it (that is, it assumes the Markov property).

Generally, this assumption enables reasoning and computation with the model that would otherwise be intractable.

Examples:

- Snake & ladder game
- Weather system

Assumptions for Markov model:

- ❖ A fixed set of states and fixed transition probabilities, and the possibility of getting from any state to another through a series of transitions.
- ❖ A Markov process converges to a unique distribution over states. This means that what happens in the long run won't depend on where the process started or on what happened along the way.
- ❖ What happens in the long run will be completely determined by the transition probabilities – the likelihoods of moving between the various states.

Types of Markov models and when to use which model

	System state is fully observable	**System state is partially observable**
System is autonomous	Markov Chain	Hidden Markov Model
System is controlled	Markov Decision Process	Partially observable Markov decision process

State Space and Time Space

		State Space	
		Discrete	**Continuous**
Time Space	**Discrete**	Markov Chain	A Markov chain on a measurable state space
	Continuous	Markov Process	Continuous stochastic process

Markov Chain

A Markov chain is a stochastic model that uses mathematics to predict the probability of a sequence of events occurring based on the most recent event. A common example of a Markov chain in action is the way Google predicts the next word in your sentence based on your previous entry within Gmail.

A Markov chain is a stochastic model created by Andrey Markov that outlines the probability associated with a sequence of events occurring based on the state in the previous event. It's a very common and easy to understand model that's frequently used in industries that deal with sequential data such as finance. Even Google's page rank algorithm, which determines what links to show first in its search engine, is a type of Markov chain. Through mathematics, this model uses our observations to predict an approximation of future events.

Definition 1

The discrete parameter Markov process $\{X_t;\ t \in T\}$ is known as Markov Chain with state space either discrete or continuous.

Consider a simple coin tossing experiment repeated for a number of times (costively), two possible outcomes for each trial are 'Head' and 'Tail'. Assume that Head occurs with probability p and that Tail occurs with probability q, so that $p + q = 1$.

Let us denote the outcomes of the n^{th} toss of the unbiased coin by X_t. Then

$$X_t = \begin{cases} 1 & if\ head\ occurs \\ 0 & if\ tail\ occurs \end{cases} \quad for\ n = 1, 2, \ldots \ldots$$

That is $P\{X_t = 1\} = p, and\ P_r\, h\{X_t = 0\} = q$. Hence the sequence of random variables, $X_1, X_2, \ldots, X_{t-1.}$ Can be written as $\{X_t\text{: } t \geq 1\}$, which is a Markov chain.

Definition 2

Let $\{X_0, X_1, X_2, \ldots\}$ be a sequence of discrete random variables. Then $\{X_0, X_1, X_2, \ldots\}$ is a Markov chain if it satisfies the Markov property:

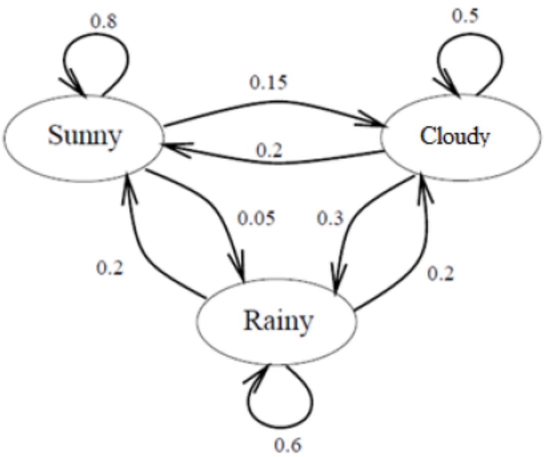
$$P(X_{t+1} = s \mid X_t = s_t, \ldots, X_0 = s_0) = P(X_{t+1} = s \mid X_t = s_t),$$

for all $t = 1, 2, 3, \ldots$ and for all states $s_0, s_1, \ldots, s_t$, s and $t \geq 0$ is called a Markov Chain.

Markov Chain: Weather Example

- Design a Markov Chain to predict the weather of tomorrow using previous information of the past days.
- Our model has only 3 states: $S=\{S_1,S_2,S_3\}$, and the name of each state is $S_1=Sunny$, $S_2=Rainy$, $S_3=Cloudy$.
- To establish the transition probabilities relationship between states we will need to collect data.

- Assume the data produces the following transition probabilities:



$$P(Sunny|Sunny) = 0.8$$
$$P(Rainy|Sunny) = 0.05$$
$$P(Cloudy|Sunny) = 0.15$$
} 1

$$P(Sunny|Rainy) = 0.2$$
$$P(Rainy|Rainy) = 0.6$$
$$P(Cloudyy|Rainy) = 0.2$$
} 1

$$P(Sunny|Cloudy) = 0.2$$
$$P(Rainy|Cloudy) = 0.3$$
$$P(Cloudy|Cloudy) = 0.5$$
} 1

- Let's say we have a sequence: Sunny, Rainy, Cloudy, Cloudy, Sunny, Sunny, Sunny, Rainy, ….; so, in a day we can be in any of the three states.

- We can use the following state sequence notation: q_1, q_2, q_3, q_4, q_5,….., where $qi \epsilon \{Sunny, Rainy, Cloudy\}$.

- In order to compute the probability of tomorrow's weather we can use the Markov property:

$$P(q_1, \ldots, q_n) = \prod_{i=1}^{n} P(q_i|q_{i-1})$$

Exercise 1: Given that today is Sunny, what's the probability that tomorrow is Sunny and the next day Rainy?

$$P(q_2, q_3|q_1) = P(q_2|q_1)P(q_3|q_1, q_2)$$

$$= P(q_2|q_1)\ P(q_3|q_2)$$
$$= P(Sunny|Sunny)\ P(Rainy|Sunny)$$
$$= (0.8)(0.05)$$
$$= 0.04$$

Exercise 2: Assume that yesterday's weather was Rainy, and today is Cloudy, what is the probability that tomorrow will be Sunny?

$$P(q_3|q_1, q_2) = P(q_3|q_2)$$

$$= P(Sunny|Cloudy)$$

$$= 0.2$$

Markov process

A ‘continuous time’ stochastic process that fulfils the Markov property is called a Markov process.

Markov process is to identify the probability of transitioning from one state to another. One of the primary appeals to Markov is that the future state of a stochastic variable is only dependent on its present state. An informal definition of a stochastic variable is described as a variable whose values depend on the outcomes of random occurrences. Markov process is a stochastic process which has memoryless characteristics.

Definition of Markov process

A stochastic process $\{X(t),\ t \in T\}$ is said to be Markov if the future of the process is independent of its past, conditioned on the present value of the process. That is, for any choice of sampling instances $t_1 < t_2 < \ldots < t_k$,

$$\Pr[X(t_k) = x_k \mid X(t_1) = x_1, \cdots X(t_{k-1}) = x_{k-1}) = \Pr[X(t_k) = x_k \mid X(t_{k-1}) - x_{k-1})$$

the process $\{X(t),\ t \in T\}$ is a Markov process.

Definition of Process Diagram:

The process diagram of a Markov chain is a directed graph describing the Markov process. Each node represents a state from the state space. The edges are labeled by the probabilities of going from one state to the other states. Edges with zero transition probability are usually discarded.

Create a Markov Chain Model

A Markov chain model is dependent on two key pieces of information:

- Transition matrix
- Initial state vector

Transition Matrix

Denoted as “P,” This N×N matrix represents the probability distribution of the state’s transitions. The sum of probabilities in each row of the matrix will be one, implying that this is a stochastic matrix.

Initial State Vector

Denoted as “S,” this N×1 vector represents the probability distribution of starting at each of the N possible states. Every element in the vector represents the probability of beginning at that state.

Transition probabilities:

We have a set of states, S = $\{s_1; s_2; \ldots ; s_r\}$. The process starts in one of these states and moves successively from one state to another. Each move is called a step. If the chain is currently in state s_i, then it moves to state s_j at the next step with a probability denoted by p_{ij} , and this probability does not depend upon which states the chain was in before the current state.

The probabilities p_{ij} are called *transition probabilities.* The process can remain in the state it is in, and this occurs with probability p_{ii}. An initial probability distribution, defined on *S*, specifies the starting state.

Transition Probability Table

$$\mathbf{P} = \begin{bmatrix} P_{11} & P_{12} & P_{13} \\ P_{21} & P_{22} & P_{23} \\ P_{31} & P_{32} & P_{33} \end{bmatrix}$$

$$P_{ij} \geq 0,\ i = 1,\ldots,n;\ j = 1,\ldots,n\ \ and\ \ \sum_{j=1}^{n} P_{ij} = 1$$

$$\begin{matrix} P_{11} = 0.7 & P_{12} = 0.2 & P_{13} = 0.1 \\ P_{21} = 0. & P_{22} = 0.6 & P_{23} = 0.4 \\ P_{31} = 0.3 & P_{32} = 0.5 & P_{33} = 0.2 \end{matrix}$$

Example:

Cheezit, a lazy hamster, only knows three places in its cage: (a) the pine wood shaving that offers him bedding where it sleeps, (b) the feeding trough that supplies him with food, and (c) the wheel where it makes some exercise.

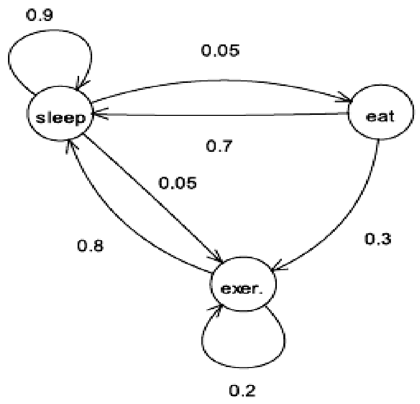
After every minute, the hamster either gets to some other activity, or keeps on doing what he's just been doing. Referring to Cheezit as a process without memory is not exaggerated at all:

- When the hamster sleeps, there are 9 chances out of 10 that it won't wake up the next minute.
- When it wakes up, there is 1 chance out of 2 that it eats and 1 chance out of 2 that it does some exercise.
- The hamster's meal only lasts for one minute, after which it does something else.
- After eating, there are 3 chances out of 10 that the hamster goes into its wheel, but most notably, there are 7 chances out of 10 that it goes back to sleep.
- Running in the wheel is tiring: there is an 80% chance that the hamster gets tired and goes back to sleep. Otherwise, it keeps running, ignoring fatigue.

Process Diagrams:

Process diagrams offer a natural way of graphically representing Markov processes – similar to the state diagrams of finite automata.

For instance, the previous example with our hamster in a cage can be represented with the process diagram.



Process diagram of a Markov process.

Transition probabilities

$$P = \begin{array}{c} \\ S \\ E \\ Ex \end{array}\begin{array}{c} \begin{array}{ccc} S & E & Ex \end{array} \\ \begin{bmatrix} 0.9 & 0.05 & 0.05 \\ 0.7 & 0 & 0.3 \\ 0.8 & 0 & 0.2 \end{bmatrix} \end{array}$$

P(sleep | sleep)=0.9	P(sleep | eat)=0.7	P(sleep | exercise)=0.8
P(eat | sleep)=0.05	P(eat | eat)=0	P(eat | exercise)=0
P(exercise | sleep)=0.05	P(exercise | eat)=0.3	P(exercise | exercise)=0.2

Example: Markov Chain

Suppose that on any given sunny day, the next day's weather has a 60% chance of being sunny, a 30% chance of being cloudy, and a 10% chance of being rainy. On any given cloudy day, the next day's weather has a 40% chance of being sunny, a 30% chance of being cloudy, and a 30% chance of being rainy. Lastly, on any given rainy day, there is a 20% chance of being sunny, a 50% chance of being cloudy, and a 30% chance of being rainy.

(i) To find the transition probability matrix and also find the initial state probability matrix.

(ii) Draw the process diagram.

(iii) To determine the probabilities of the weather conditions after 5 days from today.

Procedure

- To calculate the probabilities of weather condition and from the transition probability matrix.
- To draw the process diagram for given probabilities.
- To estimate the probabilities of the weather conditions after 5 days.

$$S_1 = S_0 \times P$$

$$P^2 = P \times P$$

$$S_2 = S_0 \times P^2$$

$$P^3 = P^2 \times P$$

$$S_3 = S_0 \times P^3$$

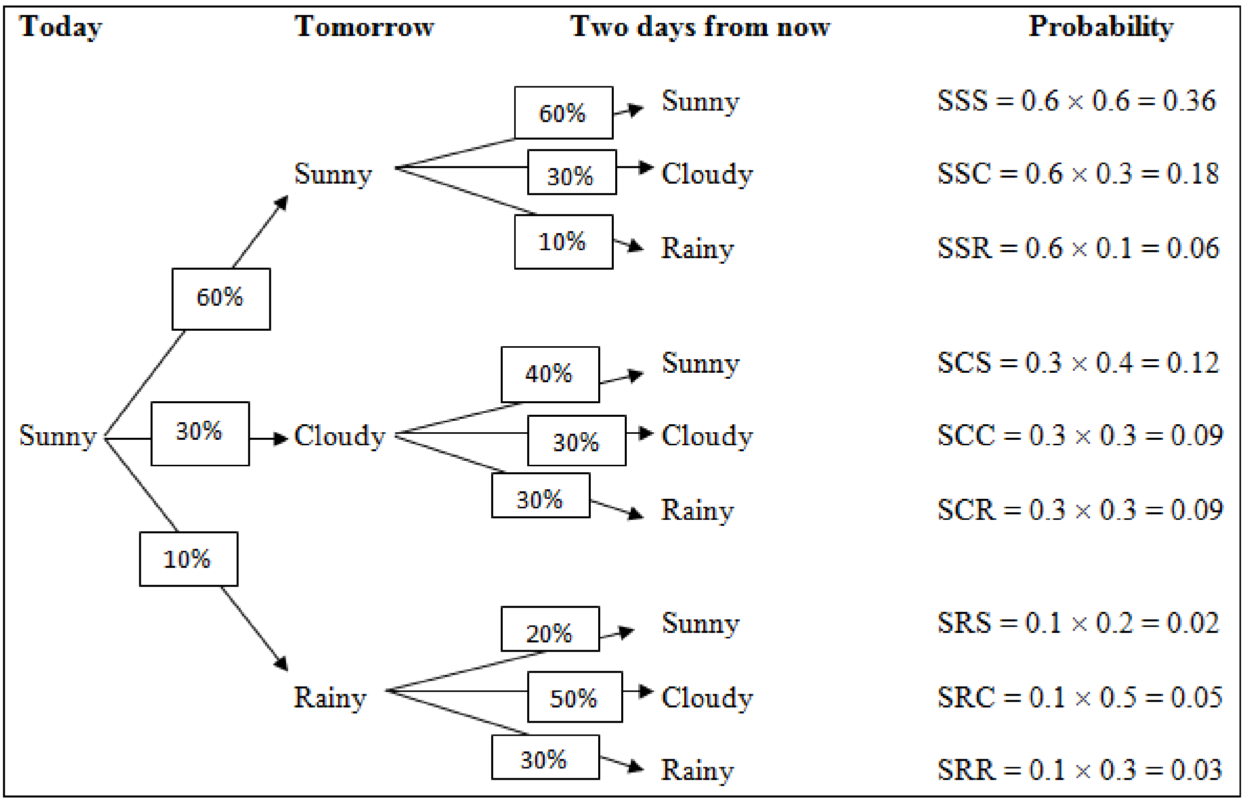
$$P^4 = P^3 \times P$$

$$S_4 = S_0 \times P^4$$

$$P^5 = P^4 \times P$$

$$S_5 = S_0 \times P^5$$

Calculation



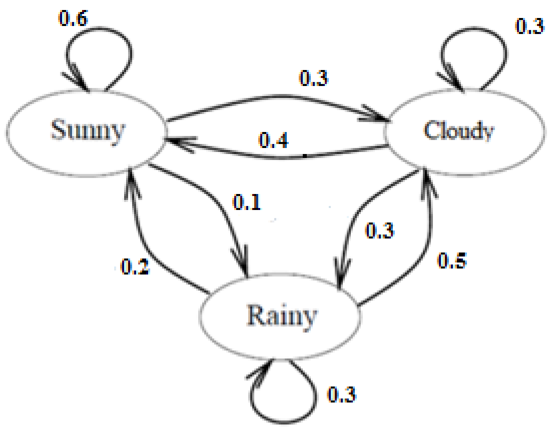
(i) The transition Probability matrix is,

$$P = \begin{bmatrix} 0.6 & 0.3 & 0.1 \\ 0.4 & 0.3 & 0.3 \\ 0.2 & 0.5 & 0.3 \end{bmatrix}$$

The initial state probability matrix is,

$$S_0 = [1 \quad 0 \quad 0]$$

(ii) The process diagram is



(iii) To calculate the probabilities of the weather conditions after 5 days from today

Here the initial vector is $S_0 = [1 \quad 0 \quad 0]$ because it was sunny on Sunday.

Day 1:

As Monday is first day away from Sunday so there have to determine the first step of the transition probability matrix.

$$P = \begin{bmatrix} 0.6 & 0.3 & 0.1 \\ 0.4 & 0.3 & 0.3 \\ 0.2 & 0.5 & 0.3 \end{bmatrix}$$

$$S_1 = S_0 \times P$$

$$S_1 = [1 \quad 0 \quad 0] \begin{bmatrix} 0.6 & 0.3 & 0.1 \\ 0.4 & 0.3 & 0.3 \\ 0.2 & 0.5 & 0.3 \end{bmatrix}$$

$$S_1 = [0.6 \quad 0.3 \quad 0.1]$$

Thus, the probability that if it is sunny on Sunday, it will be sunning next day is 0.6, clouding next is 0.3 and raining next day is 0.1.

Day 2:

As Tuesday is Second day away from Sunday so there have to determine the Second step of the transition probability matrix.

$$P^2 = P \times P$$

$$P^2 = \begin{bmatrix} 0.6 & 0.3 & 0.1 \\ 0.4 & 0.3 & 0.3 \\ 0.2 & 0.5 & 0.3 \end{bmatrix} \times \begin{bmatrix} 0.6 & 0.3 & 0.1 \\ 0.4 & 0.3 & 0.3 \\ 0.2 & 0.5 & 0.3 \end{bmatrix}$$

$$P^2 = \begin{bmatrix} 0.50 & 0.32 & 0.18 \\ 0.42 & 0.36 & 0.22 \\ 0.38 & 0.36 & 0.26 \end{bmatrix}$$

$$S_2 = S_0 \times P^2$$

$$S_2 = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0.50 & 0.32 & 0.18 \\ 0.42 & 0.36 & 0.22 \\ 0.38 & 0.36 & 0.26 \end{bmatrix}$$

$$S_2 = \begin{bmatrix} 0.5 & 0.32 & 0.18 \end{bmatrix}$$

Thus, the probability that if it is sunny on Sunday, it will be sunning next Tuesday is 0.5, clouding next Tuesday is 0.32 and raining next Tuesday is 0.18.

Day 3:

As Wednesday is Third day away from Sunday so there have to determine the third step of the transition probability matrix.

$$P^3 = P^2 \times P$$

$$P^3 = \begin{bmatrix} 0.50 & 0.32 & 0.18 \\ 0.42 & 0.36 & 0.22 \\ 0.38 & 0.36 & 0.26 \end{bmatrix} \times \begin{bmatrix} 0.6 & 0.3 & 0.1 \\ 0.4 & 0.3 & 0.3 \\ 0.2 & 0.5 & 0.3 \end{bmatrix}$$

$$P^3 = \begin{bmatrix} 0.464 & 0.336 & 0.200 \\ 0.440 & 0.344 & 0.216 \\ 0.424 & 0.352 & 0.224 \end{bmatrix}$$

$$S_3 = S_0 \times P^3$$

$$S_3 = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0.464 & 0.336 & 0.200 \\ 0.440 & 0.344 & 0.216 \\ 0.424 & 0.352 & 0.224 \end{bmatrix}$$

$$S_3 = \begin{bmatrix} 0.464 & 0.336 & 0.2 \end{bmatrix}$$

Thus, the probability that if it is sunny on Sunday, it will be sunning next Wednesday is 0.464, clouding next Wednesday is 0.336 and raining next Wednesday is 0.2.

Day 4:

As Thursday is four days away from Sunday so there have to determine the 4th step of the transition probability matrix.

$$P^4 = P^3 \times P$$

$$P^4 = \begin{bmatrix} 0.464 & 0.336 & 0.200 \\ 0.440 & 0.344 & 0.216 \\ 0.424 & 0.352 & 0.224 \end{bmatrix} \times \begin{bmatrix} 0.6 & 0.3 & 0.1 \\ 0.4 & 0.3 & 0.3 \\ 0.2 & 0.5 & 0.3 \end{bmatrix}$$

$$P^4 = \begin{bmatrix} 0.4528 & 0.3400 & 0.2072 \\ 0.4448 & 0.3432 & 0.2120 \\ 0.4400 & 0.3448 & 0.2152 \end{bmatrix}$$

$$S_4 = S_0 \times P^4$$

$$S_4 = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0.4528 & 0.3400 & 0.2072 \\ 0.4448 & 0.3432 & 0.2120 \\ 0.4400 & 0.3448 & 0.2152 \end{bmatrix}$$

$$S_4 = \begin{bmatrix} 0.4528 & 0.34 & 0.2072 \end{bmatrix}$$

Thus, the probability that if it is sunny on Sunday, it will be sunning next Thursday is 0.4528, clouding next Thursday is 0.34 and raining next Thursday is 0.2072.

Day 5:

As Friday is fifth days away from Sunday so there have to determine the fifth step of the transition probability matrix.

$$P^5 = P^4 \times P$$

$$P^5 = \begin{bmatrix} 0.4528 & 0.3400 & 0.2072 \\ 0.4448 & 0.3432 & 0.2120 \\ 0.4400 & 0.3448 & 0.2152 \end{bmatrix} \times \begin{bmatrix} 0.6 & 0.3 & 0.1 \\ 0.4 & 0.3 & 0.3 \\ 0.2 & 0.5 & 0.3 \end{bmatrix}$$

$$P^5 = \begin{bmatrix} 0.449 & 0.341 & 0.209 \\ 0.447 & 0.342 & 0.211 \\ 0.445 & 0.343 & 0.212 \end{bmatrix}$$

$$S_5 = S_0 \times P^5$$

$$S_5 = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0.449 & 0.341 & 0.209 \\ 0.447 & 0.342 & 0.211 \\ 0.445 & 0.343 & 0.212 \end{bmatrix}$$

$$S_5 = \begin{bmatrix} 0.449 & 0.341 & 0.209 \end{bmatrix}$$

Thus, the probability that if it is sunny on Sunday, it will be sunning next Friday is 0.449, clouding next Friday is 0.341 and raining next Friday is 0.209.

Result

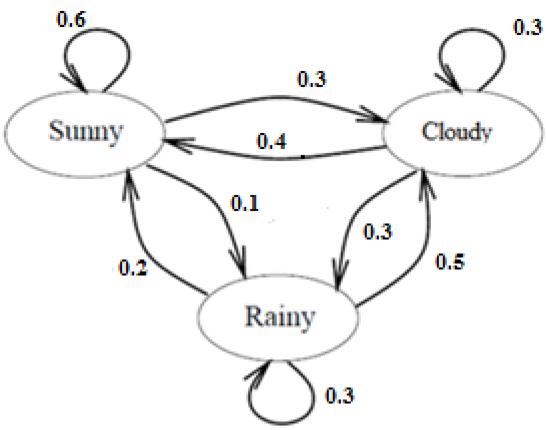
(i) The transition Probability matrix is,

$$P = \begin{bmatrix} 0.6 & 0.3 & 0.1 \\ 0.4 & 0.3 & 0.3 \\ 0.2 & 0.5 & 0.3 \end{bmatrix}$$

The initial state probability matrix is,

$$S_0 = [1 \quad 0 \quad 0]$$

(ii) The process diagram is



(iii) To calculate the probabilities of the weather conditions after 5 days from today

Day 1:

$$P = \begin{bmatrix} 0.6 & 0.3 & 0.1 \\ 0.4 & 0.3 & 0.3 \\ 0.2 & 0.5 & 0.3 \end{bmatrix}$$

$$S_1 = [0.6 \quad 0.3 \quad 0.1]$$

Thus, the probability that if it is sunny on Sunday, it will be sunning next day is 0.6, clouding next is 0.3 and raining next day is 0.1.

Day 2:

$$P^2 = \begin{bmatrix} 0.50 & 0.32 & 0.18 \\ 0.42 & 0.36 & 0.22 \\ 0.38 & 0.36 & 0.26 \end{bmatrix}$$

$$S_2 = [0.5 \quad 0.32 \quad 0.18]$$

Thus, the probability that if it is sunny on Sunday, it will be sunning next Tuesday is 0.5, clouding next Tuesday is 0.32 and raining next Tuesday is 0.18.

Day 3:

$$P^3 = \begin{bmatrix} 0.464 & 0.336 & 0.200 \\ 0.440 & 0.344 & 0.216 \\ 0.424 & 0.352 & 0.224 \end{bmatrix}$$

$$S_3 = [0.464 \quad 0.336 \quad 0.2]$$

Thus, the probability that if it is sunny on Sunday, it will be sunning next Wednesday is 0.464, clouding next Wednesday is 0.336 and raining next Wednesday is 0.2.

Day 4:

$$P^4 = \begin{bmatrix} 0.4628 & 0.3400 & 0.2072 \\ 0.4448 & 0.3432 & 0.2120 \\ 0.4400 & 0.3448 & 0.2152 \end{bmatrix}$$

$$S_4 = [0.4528 \quad 0.34 \quad 0.2072]$$

Thus, the probability that if it is sunny on Sunday, it will be sunning next Thursday is 0.4528, clouding next Thursday is 0.34 and raining next Thursday is 0.2072.

Day 5:

$$P^5 = \begin{bmatrix} 0.449 & 0.341 & 0.209 \\ 0.447 & 0.322 & 0.211 \\ 0.445 & 0.343 & 0.212 \end{bmatrix}$$

$$S_5 = [0.449 \quad 0.341 \quad 0.209]$$

Thus, the probability that if it is sunny on Sunday, it will be sunning next Friday is 0.449, clouding next Friday is 0.341 and raining next Friday is 0.209.

The n-step transition probability matrix

We define $P_{ij}^{(n)}$ as the probability that the chain is in state E_j after n steps given that the chain started in state E_i. The first step transition probabilities $P_{ij}^{(1)} = P_{ij}$ are simply the elements of the transition matrix T. We intend to find a formula for $P_{ij}^{(n)}$.

Now, by definition,

$$P_{ij}^{(n)} = P(X_n = j | X_0 = i)$$

And also

$$P_{ij}^{(n)} = \sum_{k=1}^{m} P(X_n = j, X_{n-1} = k | X_0 = i)$$

for $n \geq 2$, since the chain must have passed through one of all the m possible states at step $n - 1$.

For any three events A, B, and C, we have available the identity

$$P(A \cap B|C) = P(A|B \cap C) \times P(B|C)$$

Interpreting A as $X_n = j$, B as $X_{n-1} = k$, and C as $X_0 = i$, it follows that

$$\text{in General n : } P_{ij}^{(n)} = P(A \cap B \mid C) = P(X_n = j, X_{n-1} = k \mid X_0 = i);$$

$$= \sum_{k=1}^{m} P(X_n = j \mid X_{n-1} = k, X_0 = i) \times P(X_{n-1} = k \mid X_0 = i)$$

$$= \sum_{k=1}^{m} P(X_n = j \mid X_{n-1} = k) \times P(X_{n-1} = k \mid X_0 = i)$$

$$= \sum_{k=1}^{m} P_{jk}^{(1)} P_{ik}^{(n-1)}$$

Using the Markov property again. These are known as the Chapman–Kolmogorov equations. Putting n successively equal to *2, 3, . . .*, we find that the matrices with these elements are, using the product rule for matrices,

$$\left[P_{ij}^{(2)}\right] = \left[\sum_{k=1}^{m} P_{ik}^{(1)} P_{jk}^{(1)}\right] = T^2$$

$$\left[P_{ij}^{(3)}\right] = \left[\sum_{k=1}^{m} P_{ik}^{(2)} P_{jk}^{(1)}\right] = T^2 T = T^3$$

Since $P_{ik}^{(2)}$ are the elements of T^2, and so on. Generalising this rule,

$$\left[P_{ij}^{(n)}\right] = T^n$$

is proven by induction to n.

Calculation Procedure for n-state Transition probability matrix:

Rows indicate the current state and column indicate the transition. For example, given the current state of A, the probability of going to the next state A is s. Given the current state A', the probability of going from this state to A is r. Notice that the rows sum to 1. We will call this matrix P.

$$P = \begin{matrix} A \\ A' \end{matrix} \begin{bmatrix} s & 1-s \\ r & 1-r \end{bmatrix}$$

Initial State distribution matrix:

- This is the initial probabilities of being in state A as well as not A, A'. Notice again that the row probabilities sum to one, as they should.

$$S_0 = [t \quad (1-t)]$$

First and second state matrices:

- If we multiply the Initial state matrix by the transition matrix, we obtain the first state matrix.

$$S_1 = S_0P$$

- If the first state matrix is multiplied by the transition matrix we obtain the second state matrix:

$$S_2 = S_1P = S_0P \,.\, P = S_0P^2$$

n^{th} – State matrix: If this process is repeated, we will obtain the following expression: The entry in the i^{th} row and j^{th} column indicates the probability of the system moving from the i^{th} state to the j^{th} state in n observations or trials.

$$S_n = S_{n-1}P = S_0P^n$$

Example: n^{th} State Transition Probability Matrix

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 of 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 of the next year.

(i) Find the transition matrix.

(ii) If 90% of the drivers in the community are the low-risk category in this year, what is the probability that a driver chosen at random from the community will be the low-risk category in the next 6 years?

Procedure

- To calculate the transition probability matrix

$$P = \begin{bmatrix} L & (1-L) \\ L' & (1-L') \end{bmatrix}$$

- To calculate the initial state distribution matrix

$$S_0 = [t \quad (1-t)]$$

- To calculate the low-risk category in the next 6 years

Year 1: $S_1 = S_0 \times P$

Year 2: $S_2 = S_1 \times P$

Year 3: $S_3 = S_2 \times P$

Year 4: $S_4 = S_3 \times P$

Year 5: $S_5 = S_4 \times P$

Year 6: $S_6 = S_5 \times P$

Calculation

- The transition matrix, P

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

- Initial State distribution matrix, S_0

$$S_0 = \begin{array}{c} \begin{array}{cc} \text{L} & \text{L'} \end{array} \\ \begin{bmatrix} 0.90 & 0.10 \end{bmatrix} \end{array}$$

- First year Probability, S_1

$$S_1 = S_0 P$$

$$S_1 = \begin{bmatrix} 0.90 & 0.10 \end{bmatrix} \begin{bmatrix} 0.98 & 0.02 \\ 0.78 & 0.22 \end{bmatrix}$$

$$S_1 = \begin{array}{c} \begin{array}{cc} \text{L} & \text{L'} \end{array} \\ \begin{bmatrix} 0.96 & 0.04 \end{bmatrix} \end{array}$$

- Second year Probability, S_2

$$S_2 = S_1 \times P = S_0 \times P \,.\, P = S_0 P^2$$

$$S_2 = \begin{bmatrix} 0.96 & 0.04 \end{bmatrix} \begin{bmatrix} 0.98 & 0.02 \\ 0.78 & 0.22 \end{bmatrix}$$

$$S_2 = \begin{array}{c} \begin{array}{cc} \text{L} & \text{L'} \end{array} \\ \begin{bmatrix} 0.972 & 0.028 \end{bmatrix} \end{array}$$

- Third year Probability, S_3

$$S_3 = S_2 \times P$$

$$S_3 = \begin{bmatrix} 0.972 & 0.028 \end{bmatrix} \begin{bmatrix} 0.98 & 0.02 \\ 0.78 & 0.22 \end{bmatrix}$$

$$S_3 = \begin{bmatrix} 0.9744 & 0.0256 \end{bmatrix}$$

- Fourth year Probability, S_4

$$S_4 = S_3 \times P$$

$$S_4 = [0.9744 \quad 0.0256]\begin{bmatrix} 0.98 & 0.02 \\ 0.78 & 0.22 \end{bmatrix}$$

$$\begin{matrix} & L & L' \\ S_4 = & [0.97488 & 0.02512] \end{matrix}$$

- Fifth year Probability, S_5

$$S_5 = S_4 P$$

$$S_5 = [0.97488 \quad 0.02512]\begin{bmatrix} 0.98 & 0.02 \\ 0.78 & 0.22 \end{bmatrix}$$

$$\begin{matrix} & L & L' \\ S_5 = & [0.97498 & 0.02502] \end{matrix}$$

- Sixth year Probability, S_6

Use the formula $S_n = S_0 P^n$ to find the 6th state matrix

$$S_6 = S_5 \times P$$

$$S_6 = [0.97488 \quad 0.02502]\begin{bmatrix} 0.98 & 0.02 \\ 0.78 & 0.22 \end{bmatrix}$$

$$\begin{matrix} & L & L' \\ S_6 = & [0.974996 & 0.025004] \end{matrix}$$

After Six states, the percentage of low-risk drivers has increased to 0.97499

Result

- The transition matrix, P

$$P = \begin{matrix} L \\ L' \end{matrix}\begin{bmatrix} 0.98 & 0.02 \\ 0.78 & 0.22 \end{bmatrix}$$

- Initial State distribution matrix, S_0

$$S_0 = [0.90 \quad 0.10]$$

- To calculate the low-risk category in the next 6 years

Year 1: $S_1 = [0.96 \quad 0.04]$

Year 2: $S_2 = [0.972 \quad 0.028]$

Year 3: $S_3 = [0.9744 \quad 0.0256]$

$$\text{Year 4: } S_4 = [0.97488 \quad 0.02512]$$

$$\text{Year 5: } S_5 = [0.97498 \quad 0.02502]$$

$$\text{Year 6: } S_6 = [0.974996 \quad 0.025004]$$

After Six states, the percentage of low-risk drivers has increased to 0.97499

Chapman – Kolmogorov equation

Considered unit-step or one-step transition probabilities, the probability of X_n given X_{n-1}, i. e. the probability of the outcome at the n^{th} step or trial given the outcome at the previous step; p_{jk} gives the probability of unit-step transition from the state j at a trial to the state k at the next following trial. The m-step transition probability is denoted by

$$\Pr\{X_{m+n} = k|X_n = j\} = p_{jk}^{(m)}$$

$p_{jk}^{(m)}$ gives the probability that from the state j at n^{th} trial, the k is reached at $(m + n)^{th}$ trial in m steps, i. e. the probability of transition from the state j to the state k in exactly m steps. The number n does not occur in the r. h. s. of the relation and the chain is homogeneous. The one-step transition probabilities $p_{jk}^{(1)}$ are denoted by p_{jk} for simplicity. Consider

$$\mathrm{p}_{\mathrm{jk}}^{(2)} = \Pr\{X_{n+2} = k|X_n = j\}$$

The state k can be reached from the state j in two steps through some intermediate state r. Consider a fixed value of r; we have

$$pr\{X_{n+2} = k, X_{n+1} = r|X_n = j\} = pr\{X_{n+2} = k, X_{n+1} = r|X_n = j\} \Pr\{X_{n+1} = r|X_n = j\}$$

$$= p_{rk}^{(1)} p_{jr}^{(1)} = p_{rk} p_{jr}$$

Since these intermediate state r can assume values r = 1, 2, …., we have

$$\mathrm{p}_{\mathrm{jk}}^{(2)} = \Pr\{X_{n+2} = k|X_n = j\} = pr\{X_{n+2} = k, X_{n+1} = r|X_n = j\}$$

$$= \sum_r \mathrm{p}_{\mathrm{jr}} \mathrm{p}_{\mathrm{jk}}$$

(summing over for all intermediate states).

By induction, we have

$$p_{jk}^{(m+1)} = pr\{X_{n+m+1} = k|X_n = j\}$$

$$= \sum_r \Pr\{X_{n+m+1} = k|X_{n+m} = r\} \; Pr\{X_{n+m} = r|X_n = j\}$$

$$= \sum_r \Pr\{P_{jr}^{(m)}\}$$

Similarly, we get

$$p_{jk}^{(m+1)} = \sum_r \mathrm{p_{jr}} \mathrm{p_{jr}^{(m)}}$$

In general, we have

$$p_{jk}^{(m+n)} = \sum_r \mathrm{p_{rk}^{(n)}} \mathrm{p_{jr}^{(m)}} = \sum_r \mathrm{p_{jr}^{(n)}} \mathrm{p_{rk}^{(m)}}$$

This equation is a special case of Chapman-Kolmogorov equation, which is satisfied by the transition probabilities of a Markov chain.

From the above argument, we get

$$p_{jk}^{(m+n)} \geq \sum_r \mathrm{p_{rk}^{(n)}} \mathrm{p_{jr}^{(m)}}, \text{ for any r.}$$

Let P = (P_{jk}) denote the transition matrix of the unit-step transition and $P^{(m)} = P_{jk}^{(m)}$ denote the m-step transition matrix. For m = 2, we have the matrix $P^{(2)}$ whose elements are given by. It follow that the elements of $P^{(2)}$ are the elements of the matrix obtained by multiplying the matrix P by itself, i. e.

$$P^{(2)} = P \times P = P^2$$

Similarly,

$$P^{(m+1)} = P^{(m)} \times P$$

And

$$P^{(m+n)} = P^{(m)} \times P^{(n)}$$

It should be noted that there exist non -Markov chain whose transition probabilities satisfy Chapman – Kolmogorov equation.

Example: Chapman-Kolmogorov equation

Consider a Markov chain with the following transition probability matrix.

$$P = \begin{bmatrix} 0.1 & 0.2 & 0.7 \\ 0.2 & 0.2 & 0.6 \\ 0.6 & 0.1 & 0.3 \end{bmatrix}$$

and initial distribution $P^{(0)} = [0.7 \quad 0.2 \quad 0.1]$. Determine the conditional probabilities,

(i) $P[X_3 = 2|X_1 = 1]$, (ii) $P[X_3 = 2|X_0 = 1]$, (iii) $P[X_4 = 2|X_0 = 1]$, (iv) $P[X_3 = 1|X_0 = 1]$ and (v) $P[X_4 = 1|X_0 = 1]$

Procedure

- To calculate the first state probability

$$P^{(1)} = P^{(0)} \times P$$

- To calculate the conditional probabilities

$$P[X_3 = 2|X_1 = 1] = P_{12}^{(2)} P_1^{(1)}$$

$$P[X_3 = 2|X_0 = 1] = P_{12}^{(3)} P_1^{(0)}$$

$$P[X_4 = 2|X_0 = 1] = P_{12}^{(4)} P_1^{(0)}$$

$$P[X_3 = 1|X_0 = 1] = P_{11}^{(3)} P_1^{(0)}$$

$$P[X_4 = 1|X_0 = 1] = P_{11}^{(4)} P_1^{(0)}$$

Calculation

Given transition probability matrix of $\{X_n\}$ is

$$P = \begin{bmatrix} 0.1 & 0.2 & 0.7 \\ 0.2 & 0.2 & 0.6 \\ 0.6 & 0.1 & 0.3 \end{bmatrix}$$

By conditional probability

$$P(X_0 = 1) = 0.7 = P_1^{(0)}$$

$$P(X_0 = 2) = 0.2 = P_2^{(0)}$$

$$P(X_0 = 3) = 0.1 = P_3^{(0)}$$

The first state probability

$$P^{(1)} = P^{(0)} \times P$$

$$P^{(1)} = [0.7 \quad 0.2 \quad 0.1] \times \begin{bmatrix} 0.1 & 0.2 & 0.7 \\ 0.2 & 0.2 & 0.6 \\ 0.6 & 0.1 & 0.3 \end{bmatrix}$$

$$P^{(1)} = [0.17 \quad 0.19 \quad 0.64]$$

By conditional probability

$$P(X_1 = 1) = 0.17 = P_1^{(1)}$$

$$P(X_1 = 2) = 0.19 = P_2^{(1)}$$

$$P(X_1 = 3) = 0.64 = P_3^{(1)}$$

To calculate two step TPM

$$P^2 = \begin{bmatrix} 0.1 & 0.2 & 0.7 \\ 0.2 & 0.2 & 0.6 \\ 0.6 & 0.1 & 0.3 \end{bmatrix} \times \begin{bmatrix} 0.1 & 0.2 & 0.7 \\ 0.2 & 0.2 & 0.6 \\ 0.6 & 0.1 & 0.3 \end{bmatrix}$$

$$P^2 = \begin{bmatrix} 0.47 & 0.13 & 0.4 \\ 0.42 & 0.14 & 0.44 \\ 0.26 & 0.17 & 0.57 \end{bmatrix}$$

To calculate three step TPM

$$P^3 = \begin{bmatrix} 0.47 & 0.13 & 0.4 \\ 0.42 & 0.14 & 0.44 \\ 0.26 & 0.17 & 0.57 \end{bmatrix} \times \begin{bmatrix} 0.1 & 0.2 & 0.7 \\ 0.2 & 0.2 & 0.6 \\ 0.6 & 0.1 & 0.3 \end{bmatrix}$$

$$P^3 = \begin{bmatrix} 0.313 & 0.160 & 0.527 \\ 0.334 & 0.156 & 0.51 \\ 0.402 & 0.143 & 0.455 \end{bmatrix}$$

To calculate four step TPM

$$P^4 = \begin{bmatrix} 0.313 & 0.160 & 0.527 \\ 0.334 & 0.156 & 0.51 \\ 0.402 & 0.143 & 0.455 \end{bmatrix} \times \begin{bmatrix} 0.1 & 0.2 & 0.7 \\ 0.2 & 0.2 & 0.6 \\ 0.6 & 0.1 & 0.3 \end{bmatrix}$$

$$P^4 = \begin{bmatrix} 0.3795 & 0.1473 & 0.4732 \\ 0.3706 & 0.1490 & 0.4804 \\ 0.3418 & 0.1545 & 0.5037 \end{bmatrix}$$

The conditional probabilities is,

(i) $P[X_3 = 2|X_1 = 1] = P_{12}^{(2)} P_1^{(1)}$

$$= 0.13 \times 0.17$$

$$P[X_3 = 2|X_1 = 1] = 0.0221$$

(ii) $P[X_3 = 2|X_0 = 1] = P_{12}^{(3)} P_1^{(0)}$

$$= 0.16 \times 0.7$$

$$P[X_3 = 2|X_1 = 1] = 0.112$$

(iii) $P[X_4 = 2|X_0 = 1] = P_{12}^{(4)} P_1^{(0)}$

$$= 0.1473 \times 0.7$$

$$P[X_3 = 2|X_1 = 1] = 0.10311$$

(iv) $P[X_3 = 1|X_0 = 1] = P_{11}^{(3)} P_1^{(0)}$

$$= 0.313 \times 0.7$$

$$P[X_3 = 2|X_1 = 1] = 0.2191$$

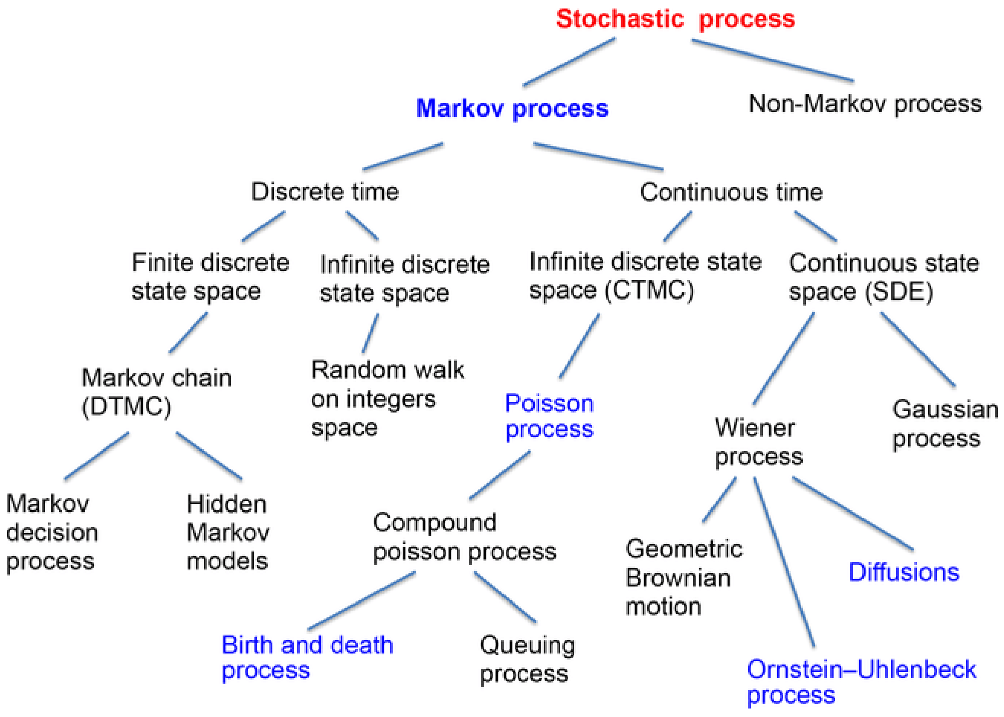
(v) $P[X_4 = 1|X_0 = 1] = P_{11}^{(4)} P_1^{(0)}$

$$= 0.3795 \times 0.7$$

$$P[X_3 = 2|X_1 = 1] = 0.2657$$

Result

- $P[X_3 = 2|X_1 = 1]$ =0.0221
- $P[X_3 = 2|X_0 = 1]$ =0.112
- $P[X_4 = 2|X_0 = 1]$ =0.10311
- $P[X_3 = 1|X_0 = 1]$ =0.2191
- $P[X_4 = 1|X_0 = 1]$ =0.2657



Skeleton of Stochastic Process