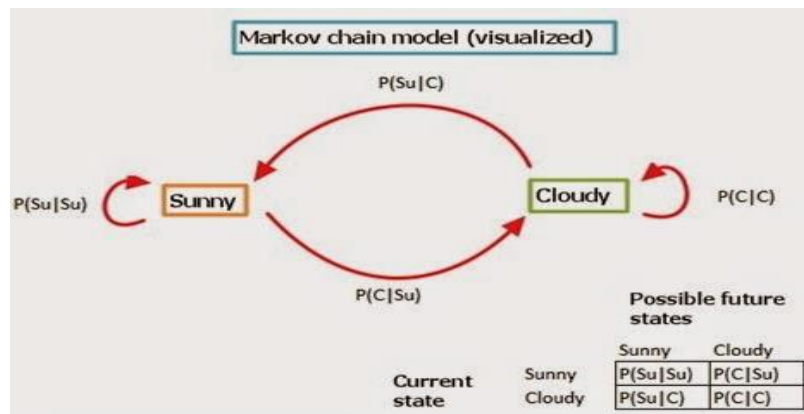


MAT 3103: Computational Statistics and Probability
Chapter 7: Stochastic Process



Stochastic Process:

A random variable $X(t)$ is said to be a stochastic process if X takes different values at different points of time t . Here, $X(t)$ indicates the states of the process.

Why we need to study Stochastic Process:

"Stochastic" is a Greek word which means "random" or "chance". Stochastic process deals with models which involve uncertainties or randomness. Uncertainty, complexity and dynamism have been continuing challenges to our understanding and control of our physical environment. Every day we encounter signals which cannot be modeled exactly by an analytic expression or in a deterministic way. Examples of such signals are ordinary speech waveforms, seismological signals, biological signals, temperature histories, communication signals etc. In manufacturing domain no machine is totally reliable. Every machine fails at some random time. Thus, in a typical manufacturing system which involves a large number of machines, the total number of machines at any time cannot be determined in a deterministic way. In a market driven economy, the stock market is volatile, the interest rates fluctuate in a random fashion. One can give any number of examples from our daily life events where uncertainty prevails in an essential way. This gives us the realization that many real-life phenomena require the analysis of a system in a probabilistic setting rather than in a deterministic setting. Thus, stochastic models are becoming increasingly important for understanding or making performance evaluation of complex systems in a broad spectrum of fields.

Reaching AIUB from your home: let us say there are three different routes to do this, all of them equal in distance/traffic congestion/safety. And also let us say these routes intersect at various points before you reach AIUB. Now, as soon as you get out of your home, you have three choices [A, B, C]. You may pick any of them with probability $1/3$. Say you pick route A. And at each intersection, you have more possibilities: stay on route A, switch to route B or switch to route C. You may pick any of the above options according to your whim. So, what happened here? The final route sequence you will take to reach AIUB (home \rightarrow A, A, B, C, A, A, C \rightarrow AIUB) is not "deterministic". In this simplistic example, your end point (AIUB) is deterministic, but your route sequence is a stochastic process.

Stochastic processes have played a significant role in various **engineering disciplines** like power systems, signal processing, manufacturing systems, automotive technology, semiconductor manufacturing, communication networks, robotics, wireless networks etc. In communication networks, unpredictability and randomness arise for several reasons. For one, connections come in and leave randomly. At the time of connection establishment, circuit switched networks typically need to find sufficient resources for the connection in the absence of which, a connection is refused. For packet switched networks such as the internet, even though the above constraint is not there, however, each packet is routed individually by the switches based on current information available and can take a different path. Also, packets can be dropped if sufficient resources are not available. In most networks, the packets that individual connections pump into the network are of varying sizes (one packet may have a different size than the other in each connection) and thus each packet holds the network for a varying (random) amount of time. Moreover, the links can fail randomly and so reliability of the medium is also an issue. From the perspective of the user, the total time needed to transmit, say a file, should be the least possible. Whereas processing, transmission and propagation delays of packets are not significant in general, queuing delays are. The latter depend on the size of packets and also the number of connections operational at a given time. Stochastic processes are thus crucially used in the design, analysis and control of networks. Control in networks can be broadly classified under four heads - admission control, routing, flow and congestion control, and resource allocation. Designing good control strategies requires a good knowledge of stochastic control, parameter estimation, simulation-based optimization, queuing theory, learning theory etc., for all of which stochastic processes form the key ingredient.

Another area where stochastic processes have applications is in the area of **neuroscience**. Analysis of EEG signals from the brain makes intensive use of stochastic processes. Among the awesome repertoire of tasks that the human brain can accomplish, one of the more fascinating ones is how the electrical activity of millions of brain cells (neurons) is translated into precise sequences of movements. One of the greatest challenges in applied neuroscience is to build prosthetic limbs controlled by neural signals from the brain. The ultimate goal is to provide paralytic patients and amputees with the means to move and communicate by controlling the prosthetic device using brain activity. Scientists and engineers are slowly getting closer to building such devices thanks to studies revealing a strong connection between the activity of neurons in the brain's cerebral cortex and the movements of limbs. To realize the above goal of building prosthetic limbs, one tool which plays a critical role is the theory of stochastic processes.

Examples of Stochastic Process:

- ⇒ Number of signals sent from a station by time t .
- ⇒ Number of e-mails received by a cyber cafe by time t .
- ⇒ Number of customers entered to a market by time t .
- ⇒ Number of vehicles entered to a CNG filling station by time t .

Counting Process:

A stochastic process $\{N(t), t \geq 0\}$ is said to be a counting process if $N(t)$ represents the total number of events that have occurred up to time t .

Under what conditions a stochastic process becomes counting process?

- ⇒ $N(t) \geq 0$.
- ⇒ $N(t)$ is an integer.
- ⇒ If $s \leq t$ then $N(s) \leq N(t)$.
- ⇒ If $s < t$, then $N(t) - N(s)$ is the number of events occurred during the interval $[s, t]$.

Examples of Counting Process:

- ⇒ Total number of signals sent from a station up to time t .
- ⇒ Total number of e-mails received by a cyber cafe up to time t .

- ⇒ Total number of customers entered to a market by time t .
- ⇒ Total number of vehicles entered to a CNG filling station by time t .

Markov Process:

A stochastic process in state i moves to state j with associated probabilities p_{ij} such that $p_{ij} \geq 0$ and $\sum_{j=0}^{\infty} p_{ij} = 1$. Such a stochastic process is called Markov process. The parametric space of the above discussed Markov process is discrete and hence this process is called **Markov Chain**.

Understanding Markov Chain with real life applications:

As an example of Markov chain, we can explain the students' admission and completion of semester in **AIUB**. The students get admission in semester 1. After one semester, some of them are promoted to semester 2, that is, they move from state 1 to state 2. This movement is associated with probability. Again, when they are promoted to semester 2, there are some students in semester 2 who move to semester 3 with probability of promotion. This process of admission and movement of students from one semester to another continues, moving from state i to state j with associated probability p_{ij} . These probabilities of movement can be shown in matrix form as $P = \begin{bmatrix} p_{00} & p_{01} \\ p_{10} & p_{11} \end{bmatrix}$. This P is called one-step transition probability matrix (TPM). The transition of the process can occur after k steps, where $(p_{ij})^{(k)}$ is the transition probability from i to j in k steps.

Dean of **FST** in **AIUB** states his intention to give a promotion by the end of next semester to the employee P. And P forwards the same message to Q, who further gives the news to R and so on. In such scenario there is a probability p that a person will replace the answer from yes to no while conveying the message to the next person and a probability y that the person will change the answer from no to yes.

Marketers use Markov Chain to predict brand switching behavior within their customers. Let us take the case of Detergent Brands. Some consumers might be using "**Tide**", Some would be using "**Surf Excel**", Others would be using something else. In this case a Brand Marketer would be interested in knowing what is the probability of a consumer to switch to using "Surf Excel" next month in case he is using "Tide" this month. He would be interested in knowing what is the probability of a customer to continue using "Tide" next month in case he is using "Tide" in the

current month, etc. Using Markov chain helps them get the probability of each of the cases thus helping them predict brand switching behavior.

Markov chain can also be explained with the example of a **frog** in a pond **jumping** from lily pad to lily pad with the relative transition probabilities. Lily pads in the pond represent the finite states in the Markov chain and the probability is the odds of frog changing the lily pads.

Google Page Rank: The entire web can be thought of as a Markov model, where every web page can be a state and the links or references between these pages can be thought of as, transitions with probabilities. So basically, irrespective of which web page you start surfing on, the chance of getting to a certain web page, say, X is a fixed probability.

Typing Word Prediction: Markov chains are known to be used for predicting upcoming words. They can also be used in auto-completion and suggestions.

Subreddit Simulation: Surely, you've come across Reddit and had an interaction on one of their threads or subreddits. Reddit uses a subreddit simulator that consumes a huge amount of data containing all the comments and discussions held across their groups. By making use of Markov chains, the simulator produces word-to-word probabilities, to create comments and topics.

Text generator: Markov chains are most commonly used to generate dummy texts or produce large essays and compile speeches. It is also used in the name generators that you see on the web.

Poisson process:

A counting process $\{N(t), t \geq 0\}$ is said to be a Poisson process if event $N(t)$ occurs at Poisson rate λ and follows:

$$\Rightarrow N(0) = 0$$

$$\Rightarrow N(t) \text{ increases independently.}$$

$$\Rightarrow P[N(t)] = \frac{e^{-\lambda t} (\lambda t)^n}{n!} ; n = 0, 1, 2, \dots, \infty.$$

Distribution of inter arrival time in case of Poisson Process:

Let T_1 be the time for the occurrence of first event, i. e. within time T_1 the first event will occur. Now, if $n > 1$, T_n is the elapsed time between the occurrence of $(n-1)^{th}$ and n^{th} event. Occurring after some time t . we need $P(T > t)$. But

$$P[N(0) = 0] = \frac{e^{-\lambda t} (\lambda t)^0}{0!} = e^{-\lambda t}$$

$$P(T > t) = e^{-\lambda t}, \quad P(T < t) = 1 - e^{-\lambda t}, \quad P(t_1 < T < t_2) = e^{-\lambda t_1} - e^{-\lambda t_2}$$

The probability function of T is as $f(T) = \lambda e^{-\lambda t}$

$$\text{So, } \mu'_1 = \frac{1}{\lambda}, \quad \mu_2 = \frac{1}{\lambda^2}, \quad \beta_1 = 4, \quad \beta_2 = 9$$

The total time until the occurrence of the n^{th} event is $S_n = T_1 + T_2 + \dots + T_n$

We need expected total time until the occurrence of n^{th} event

$$E(S_n) = E(T_1 + T_2 + \dots + T_n) = E(T_1) + E(T_2) + \dots + E(T_n) = \frac{n}{\lambda}$$

Example 7.1: Let us consider a two- state Markov process, where

State 0: today is a sunny day, and then tomorrow will be a sunny day with probability 0.6.

State 1: today is not a sunny day, tomorrow will be a sunny day with probability 0.5.

Find the probability that sunny day will be continued for four days starting from today.

Solution: The transition probability matrix: $P = \begin{bmatrix} p_{00} & p_{01} \\ p_{10} & p_{11} \end{bmatrix} = \begin{bmatrix} 0.6 & 0.4 \\ 0.5 & 0.5 \end{bmatrix}$. We need P_{00}^4 in P^4 .

$$P^2 = \begin{bmatrix} 0.6 & 0.4 \\ 0.5 & 0.5 \end{bmatrix} \begin{bmatrix} 0.6 & 0.4 \\ 0.5 & 0.5 \end{bmatrix} = \begin{bmatrix} 0.56 & 0.44 \\ 0.55 & 0.45 \end{bmatrix}$$

$$P^4 = \begin{bmatrix} 0.56 & 0.44 \\ 0.55 & 0.45 \end{bmatrix} \begin{bmatrix} 0.56 & 0.44 \\ 0.55 & 0.45 \end{bmatrix} = \begin{bmatrix} 0.5556 & 0.4444 \\ 0.5555 & 0.4445 \end{bmatrix}$$

The required probability is 0.5556.

Example 7.2: A communication system sends the digits 0 and 1 as a mode of communication.

The digit sent from the center reaches to the destination after several steps as it is sent. A digit 0 sent from the center reaches to the next step as zero with probability 0.4. Find the probability that the digit 0 sent from the center will reach to the destination at 5th stage.

Solution: The transition probability matrix: $P = \begin{bmatrix} p_{00} & p_{01} \\ p_{10} & p_{11} \end{bmatrix} = \begin{bmatrix} 0.4 & 0.6 \\ 0.6 & 0.4 \end{bmatrix}$. To reach the destination at 5th step, it needs further (5-1) = 4 steps. Thus, we need P_{00}^4 in P^4 .

$$P^2 = \begin{bmatrix} 0.4 & 0.6 \\ 0.6 & 0.4 \end{bmatrix} \begin{bmatrix} 0.4 & 0.6 \\ 0.6 & 0.4 \end{bmatrix} = \begin{bmatrix} 0.52 & 0.48 \\ 0.48 & 0.52 \end{bmatrix}$$

$$P^4 = \begin{bmatrix} 0.52 & 0.48 \\ 0.48 & 0.52 \end{bmatrix} \begin{bmatrix} 0.52 & 0.48 \\ 0.48 & 0.52 \end{bmatrix} = \begin{bmatrix} 0.5008 & 0.4992 \\ 0.4992 & 0.5008 \end{bmatrix}$$

The required probability is 0.5008.

Example 7.3: Signals sent from a station reach its goal at Poisson rate $\lambda = 5$ per hour. Find the probability that the elapsed time between the entrance of 5th and 6th signal is (i) more than 1 hour, (ii) less than 0.5 hour, (iii) 0.5 to 1 hour. Find the expected time until the 9th signal reaches its goal.

Solution: Let T be the elapsed time between the entrance of (n-1)th and nth signal and S_n be the waiting time until the nth signal reaches to the destination.

i. $P(T > 1) = e^{-\lambda t} = e^{-5 \times 1} = 0.00674$

ii. $P\left(T < \frac{1}{2}\right) = 1 - e^{-\lambda t} = 1 - e^{-5 \times \frac{1}{2}} = 0.91792$

iii. $P(0.5 < T < 1) = e^{-\lambda t_1} - e^{-\lambda t_2} = e^{-5 \times 0.5} - e^{-5 \times 1} = 0.08208 - 0.00674 = 0.07534$

$$E(S_n) = \frac{n}{\lambda} = \frac{9}{5} = 1.8 \text{ hour}$$

Example 7.4: Emails enter to a cybercafé after several stages. An email sent from a cybercafé reaches to the next stage with probability 0.7. Find the probability that an email sent from cyber cafe reaches to another cybercafé to the address at 4th stage.

Solution: The transition probability matrix: $P = \begin{bmatrix} p & 1-p \\ 1-p & p \end{bmatrix} = \begin{bmatrix} 0.7 & 0.3 \\ 0.3 & 0.7 \end{bmatrix}$. It will reach to the destination at (4-1) = 3 steps. Thus, we need P_{00}^3 in P^3 .

$$P^2 = \begin{bmatrix} 0.7 & 0.3 \\ 0.3 & 0.7 \end{bmatrix} \begin{bmatrix} 0.7 & 0.3 \\ 0.3 & 0.7 \end{bmatrix} = \begin{bmatrix} 0.58 & 0.42 \\ 0.42 & 0.58 \end{bmatrix}$$

$$P^3 = \begin{bmatrix} 0.58 & 0.42 \\ 0.42 & 0.58 \end{bmatrix} \begin{bmatrix} 0.7 & 0.3 \\ 0.3 & 0.7 \end{bmatrix} = \begin{bmatrix} 0.532 & 0.468 \\ 0.468 & 0.532 \end{bmatrix}$$

The required probability is 0.532.

- 7.4.** Customers enter to a mobile repairing shop from 9 – 10 and from 10 – 11 with probability 0.6. Customers also enter to the shop from 10 – 11 but not from 9 – 10 with probability 0.8. Find the probability that the customers will enter to the repairing shop up to 2 PM starting from 9 AM.
- 7.5.** Customers enter to a mobile repairing shop from 9 – 10 and from 10 – 11 with probability 0.6. Customers also enter to the shop from 10 – 11 but not from 9 – 10 with probability 0.4. Find the probability that the customers will enter to the repairing shop up to 1 PM starting from 9 AM.
- 7.6.** E. mails enter to a cybercafé at a Poisson rate $\lambda = 2$ per minute. Find the probability that the elapsed time between the entrance of 10th and 11th mail is (i) more than 1 minute, (ii) less than 2 minutes, (iii) between 1 to 2 minutes.

- 7.7.** The customer enters to a bank at Poisson rate $\lambda = 25$ per hour. Find the probability that the elapsed time between the entrance of 10th and 11th customer is (i) less than 2minutes, (ii) more than 5 minutes, (iii) 3 to 5 minutes.
- 7.8.** E-mails enter to a cybercafé at a Poisson rate $\lambda = 2$ per minute. Find the expected time until the entrance of 10th mail.
- 7.9.** The customer enters to a bank at Poisson rate $\lambda = 25$ per hour. Find the expected time until the entrance of 50th customer.

Sample MCQ

1. Customers enter to a cyber cafe from 9 – 10 am and from 10 – 11 am with probability 0.6. Customers also enter to the shop from 10 – 11 am but not from 9 – 10am with probability 0.8. Find the probability that the customers will enter to the cyber cafe up to 12 pm starting from 9 am.

- a) 0.3847 b) 0.6163 c) 0.3360 **d) 0.6640**

2. E-mails enter to a server after several stages . An e.mail sent from a server reaches to the next stage with probability 0.4 . Find the probability that an e.mail sent from server reaches to another server to the address at 4th stage.

- a) 0.1843 **b) 0.4960** c) 0.9266 d) 0.5040

3. Telephone calls enter to a hospital at a Poisson rate = 2 per minute. Find the probability that the elapsed time between the entrance of 11th and 12th call is more than 1 minute.

- a) 0.13534** b) 0.98168 c) 0.11702 d) 0.0634

4. Telephone calls enter to a hospital at a Poisson rate = 2 per minute. Find the probability that the elapsed time between the entrance of 11th and 12th call is less than 2.

- a) 0.13534 **b) 0.98168** c) 0.11702 d) 0.0634

5. Telephone calls enter to a hospital at a Poisson rate = 2 per minute. Find the probability that the elapsed time between the entrance of 11th and 12th call is between 1 to 2 minutes.

- a) 0.13534 b) 0.98168 **c) 0.11702** d) 0.0634

6. Telephone calls enter to a hospital at a Poisson rate = 3 per minute. Find the expected time until the entrance of 15th call. 2 points

- a) 10 minutes b) 15 minutes c) 45 minutes **d) 5 minutes**