MANIPAL
ACADEMY of HIGHER EDUCATION
(Institution of Eminence Deemed to be University)

MDS5204/ Segment 1

INTRODUCTION TO STOCHASTIC PROCESSES

Table of Contents

Introduction

The motivation behind stochastic processes is to study the dynamics and a well-defined theory which is needed to study the characteristics of realistic models. The deterministic models do not involve the concept of uncertainty, so the probability models are more realistic than deterministic models.

Learning Objectives

At the end of this topic, you will be able to:

- Explain the deterministic and stochastic model
- Define sigma-algebra and the probability measure
- Explain stochastic processes and sample path
- Define state space and parameter space

1. Deterministic Model

A deterministic model is a mathematical or computational representation of a system or process in which the outcome is fully determined (completely certain) by the initial conditions and the underlying rules or laws that govern the system. A deterministic model allows you to calculate a future event exactly without the involvement of randomness. If the outcome of an experiment is certain (always the same), it is a deterministic experiment**.**

Some examples of deterministic model are as follows:

- A simple physics model that describes the motion of an object under known forces.
- A computer simulation of a manufacturing process that follows a fixed set of rules.
- A mathematical model of a machine learning algorithm that predicts outcomes based on input data.

Hence, in a deterministic model everything is 'certain' as it works under fixed initial conditions which results in fixed outcomes and zero involvement of randomness.

2. Stochastic Model

Stochastic model is the model that includes random variables.

Some examples of Stochastic model are as follows:

- Stock market model: It takes random fluctuations into account in stock prices.
- Weather forecasting model: It includes random variations in atmospheric conditions.
- Population growth model: It includes random fluctuations in birth and death rates in the population.

Let's consider an example of bacterial growth. In the deterministic growth model, number of bacteria presents after t minutes can be determined. The deterministic growth model is to assert the population of bacteria's growth at a fixed rate, like 20% per minute with zero randomness involved. The deterministic model does not address the uncertainty present in the reproduction rate of individual organisms.

The stochastic process captures the uncertainty and makes the model more realistic than the deterministic model as it considers the reproduction rate of individual organism. In many biological processes, the exponential distribution is a common choice for modelling the time of birth and death. A simple stochastic growth model assumes that the time taken by individual bacterium to reproduce is an independent exponential random variable—for example, say, with a rate parameter of 0.20.

Since randomness can be included in the model, by applying the stochastic processes or the stochastic model, the dynamics of the system can be studied in a better way.

It helps to answer the following questions:

1. What is the average number of bacteria present at time *t* such as after 10 min, 1 hour, or a day?
2. What is the probability that the number of bacteria will exceed some threshold after *t* minutes?
3. What is the distribution of the time taken by the bacteria to double in size?
4. What is the likelihood that the population goes extinct, or reaches a long-term equilibrium?

Therefore, in all cases probability is used for the quantification of uncertainty in the system which makes the stochastic model much better than the deterministic model.

3. Prerequisites for Stochastic Process

There are some prerequisites that are required to be known, before implementing a stochastic process. They are:

- **Experiment**: It is a measurement procedure or observation process.
- **Outcome**: It is the result of an experiment, where if one outcome occurs, then no other outcome can occur at the same time. For example, when a coin is tossed, obtaining a head or tail is an outcome.
- **Random Experiment**: It is a well-defined procedure that produces an observable outcome that could not be perfectly predicted in advance.
- **Event**: It is a single outcome or a collection of outcomes of an experiment.

- **Sample Space ($\boldsymbol{\Omega}$)**: It is a set of possible outcomes of a random experiment. For example, tossing a coin is a random experiment and getting head or tail is the outcome. So, the collection of these outcomes will become sample space, represented as,

$$Random\ experiment\text{: Toss a coin once}$$
$$Sample\ Space\text{: } \boldsymbol{\Omega} = \{\textbf{head},\ \textbf{tail}\}$$

- **Random Variable**: A random variable (X) is defined as a function from the sample space to the real numbers that is "$\mathrm{X}: \Omega \rightarrow R$" where X is a random variable, Ω is a sample space and a domain and R is the set of real numbers and co-domain. So, every element of this sample space will be associated to some element of this set of real numbers. A random variable assigns a real number to every possible outcome of a random experiment. In the given random experiment, tossing a coin and the collection of outcomes of getting a head and tail will be the sample space. X is a random variable of getting a head or tail. As shown below, the head is associated with a real number 1 and the tail is associated with 0.

$$Random\ Experiment \text{ : Toss a coin once}$$
$$Sample\text{: } \Omega = \{\text{head}, \text{tail}\}$$
$$An\ example\ of\ random\ variable\text{: } X\text{: } \Omega \rightarrow \mathrm{R} \text{ maps "head"} \rightarrow 1,\ \text{"tail"} \rightarrow 0$$

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

- **Sigma-Algebra**: Collection of all subjects of sample space (Ω) to which we are willing to assign probabilities-these subsets are called events. Suppose, Sigma-Algebra is denoted with (β), then β be a set whose elements are subsets of sample space (Ω) then β is a sigma-algebra if it satisfies the following axioms:
    1. The sample space (Ω) that belongs to the β that is a Sigma algebra ($\Omega \in \beta$).
    2. If A lies in β $(\mathrm{A} \in \beta)$, then A^{C} should also be a member of β $(\mathrm{A}^{\mathrm{C}} \in \beta)$.
    3. If A_{n} is a countable collection of elements of β, and A is equals to union of A_{n} $(\mathrm{A} = \cup_{\mathrm{n}=1}^{\infty} \mathrm{A}_{\mathrm{n}})$, then A will belong to β $(\mathrm{A} \in \beta)$.

- **Probability Measure**: It is a function that associates probability to each of the events belonging to the Sigma-algebra β. Let P be a function that associates a real number to each element of the Sigma algebra β, then P is a probability measure if it satisfies the following axioms.

1. If A belongs to β ($A \in \beta$), then the probability of A should always be greater or equal to 0 ($P(A) \geq 0$).
2. The probability of the sample space that is a certain event is always equal to one ($P(\Omega) = 1$).
3. If $\{A_n\}$ is a countable collection of disjoint elements of β (that is $A_j \cap A_K = \emptyset$ if $j \neq k$), then the probability of union of A_n is equal to the summation of the probability of A_n ($P(\bigcup_{n=1}^{\infty} A_n) = \sum_{n=1}^{\infty} P(A_n)$).

A probabilistic experiment consists in extracting a ball from a box containing two balls, one red (R) and one blue (B). So, the sample space (Ω) will consist of red ball and blue ball. (σ) Sigma algebra of events, that is, β, will consist of the null set (Ø), the sample space (Ω), extracting a red ball $\{R\}$ and extracting a blue ball $\{B\}$. Sample space (Ω) means either a blue or red ball extracted.

$$\Omega = \{R, B\},$$

$$\sigma - \text{algebra of events i.e.}$$

$$\beta = \{\emptyset, \Omega, \{R\}, \{B\}\}$$

The probability measure (P) on the Sigma algebra of even β is,

$$P(\beta) = \begin{cases} 0 \; if \; \beta = \emptyset \\ {}^{1}/_{2} \; if \; \beta = \{R\} \\ {}^{1}/_{2} \; if \; \beta = \{B\} \\ 1 \; if \; \beta = \Omega \end{cases}$$

- **Probability Space**: A probability space is a triplet (Ω, β, P) where Ω is sample space, β is a sigma-algebra of events and P is a probability measure on β.

4. Stochastic Processes

A stochastic process is a family of indexed random variable $\{X(t,w)\}; t \in T; w \in \Omega\}$ defined on a probability space (Ω, β, P) where T is an arbitrary set. This set, $\{X(t,w)\}; t \in T; w \in \Omega\}$ is known as stochastic process. It consists of two arguments, one is (t) and another one is (w), where (w) belongs to the sample space (Ω) and (t) belongs to the index parameter (T).

It also consists of the sample space (Ω), the Sigma algebra (β) and the probability measure (P). Here, T may be time, length, or distance and can be an arbitrary set.

There are many ways of visualising a stochastic process, such as

- For each choice of $t \in T, X(t,w)$ is a random variable.
- For each choice of $\omega \in \Omega,\ X(t,w)$ is a function of t.
- For each choice of ω and $t, X(t,w)$ is a number.
- In general, it is an ensemble (family) of function $X(t,w)$ having two arguments where, t and w can take different possible values.

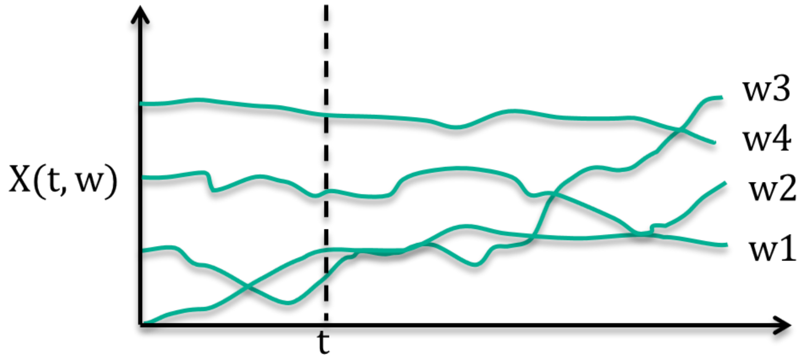

Figure 1: Stochastic Processes

The set of functions $(\mathrm{X}(\mathrm{t}, w_1),\ \mathrm{X}(\mathrm{t}, w_2), \mathrm{X}(\mathrm{t}, w_3), \dots, \mathrm{X}(\mathrm{t}, w_n)$ corresponding to the N outcomes of an experiments is called an ensemble. Each member $\mathrm{X}(\mathrm{t}, w_i)$ is called a sample function of the stochastic process.

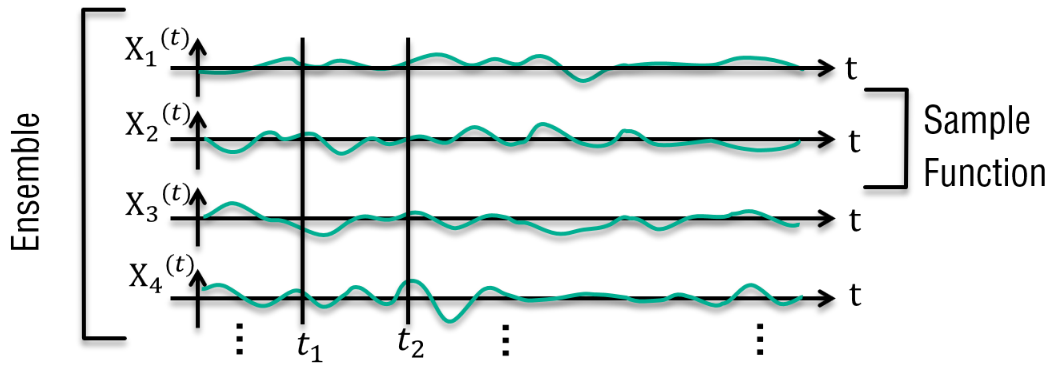

Figure 2: 1st Observation of Stochastic Process

In this observation of Stochastic Process, the X_4 represents $\mathrm{X}(t, w_4)$, X_3 represents $\mathrm{X}(t, w_3)$, X_2 represents $\mathrm{X}(\mathrm{t}, w_2)$, and X_1 represents $\mathrm{X}(\mathrm{t}, w_1)$. All the above paths are known as Sample paths or Sample function of a stochastic process. The values of $\mathrm{X}(\mathrm{t})$ at a particular time t_1 define a random variable $\mathrm{X}(\mathrm{t}_1)$ or represented by X_1.

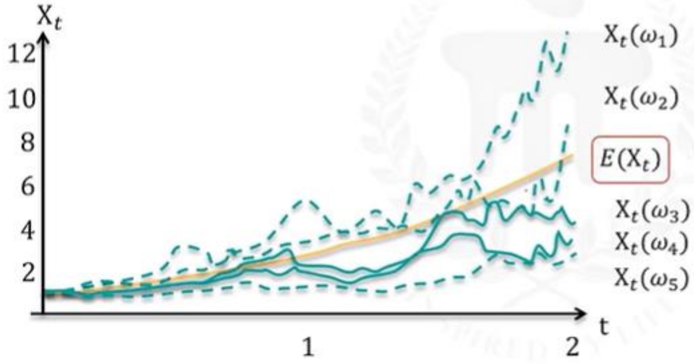

Figure 3: Observation of Stochastic Process

The graph shows five sample paths of a stochastic process $\{X_t\}$ where $t \geq 0$. So, $X_t(w_1)$, $X_t(w_2)$…, these are nothing but X_1, X_2, X_3, X_4, X_5, all are different sample paths or sample functions of the stochastic process X_t. The average or the expected value $E\{X_t\}$ of the process is also drawn (in yellow), and that it is not a sample path; it is a deterministic function of t.

A stochastic process $X = (X_t | t \in I)$ is a collection of random variables X_t. In this, w is suppressed and $X(t)$ or X_t is used to denote a random process. Taking values in some (real-valued) set $S, X_t(\omega) \in S$, and indexed by a real-valued (time) parameter $t \in I$. Here, I is referred to as the parameter space and S is referred to as the sample space. Stochastic processes are also called as random processes (or just processes).

5. State Space and Parameter Space

The index set $I \subset \mathbb{R}$ over which random variable is indexed or as discussed earlier that is $X_t, t \in T$ is called the parameter space. The parameter space can be represented with both T and I. Previously, it has been represented with T and now it is being represented with I, as it belongs to the set of real numbers. The value set $S \subset \mathbb{R}$, is also a subset of real number, called as the state space of the process. The values added to the collection of random variables is known as the state space. For example, while tossing a coin, the values taken as random variables of head or tail, then the collection of these values will be the state space. Similarly, while rolling a dice, the collection of outcomes as 1, 2, 3, 4, 5 and 6, will also become the state space. Sometimes notation X_t is used to refer to the whole stochastic process (instead of a single random variable).

MANIPAL ACADEMY of HIGHER EDUCATION
(Institution of Eminence Deemed to be University)

6. Examples of Stochastic Process

Consider a simple experiment like throwing true dice.

- **Example 1**

  Suppose that X_n is the outcome of the nth throw, $\mathrm{n} \geq 1$. Since it's a throw, it will always be greater than equals to one. Here, $\{X_n, n \geq 1\}$ is a family of random variables such that for a distinct value of n $(1,2 \ldots)$, one gets a distinct random variable X_n; $\{X_n, n \geq 1\}$. It constitutes a stochastic process, known as Bernoulli process.

- **Example 2**

  Suppose that X_n is the number of sixes in the first n throws. For a distinct value of $n = 1,2, \ldots$, we get a distinct binomial variable X_n; $\{X_n, n \geq 1\}$ which gives a family of random variables, and this is also a stochastic process.

- **Example 3**

  Suppose that X_n is the maximum number shown in the first n throws. Here, $\{X_n, n \geq 1\}$ constitutes a stochastic process. For example, suppose "n" is the number of times a dice is thrown then if $n = 1$, and if maximum value is 6, then $X_1 = 6$. Similarly, if $n = 2$ and then $X_2 = 4$ and if $n = 3$ and then $X_3 = 5$. Here, X will become X_1, X_2, X_3. So, these will become the different values of random variables. Here again, $\{X_n, n \geq 1\}$ which gives a family of random variables is a stochastic process.

- **Example 4**

  Consider that there are r cells and an infinitely large number of identical balls that are thrown at random, one by one, into the cells. Assume that the ball being thrown is equally likely to go into any of the cells. Suppose that $\mathrm{X_n}$ is the number of occupied cells after n throws. Then $\{\mathrm{X_n}, \mathrm{n} \geq 1\}$ constitutes a stochastic process.

- **Example 5**

  Consider a random event occurring in time, such as the number of telephone calls received at a switchboard. Suppose that $\mathrm{X(t)}$ is the random variable which represents the number of incoming calls in an interval $(0, \mathrm{t})$ of duration t units. Here, the random variable is the number of incoming calls. The number of calls within a fixed interval of specified duration, say, one unit of time, is a random variable $X(1)$. Here, $X(1)$ means how many calls are received in one unit of time. The family $\{X(t), t \in T\}$ constitutes a stochastic process where $T = [0, \infty)$ is a continuous variable.

7. Summary

In this topic, we discussed:

- Deterministic model and Stochastic model
- Stochastic processes and Sample path
- State space and Parameter space