Introduction to Stochastic Processes

Thiyanga S. Talagala

Table of contents

Preface

This is a Quarto book.

To learn more about Quarto books visit https://quarto.org/docs/books.

1 Introduction

1.1 What is a Stochastic Process?

First, let's see what does "stochastic" mean? The meaning of "stochastic" is **random**. The term "process" refers to a mathematical or statistical model that describes the evolution of a random variable over time. In the study of a stochastic process, we examine a collection of random variables indexed by a certain parameter, typically time, representing the evolution of a system over a series of discrete or continuous instances.

For example, suppose we monitor the weather condition every hour in a day sunny, rainy, and cloudy. Then you are essentially observing a stochastic process. This process describes how the weather condition evolves over time.

Can we describe this situation using a single random variable? No, we cannot. We need a sequence of random variables index by time as follows:

X_0 - weather condition from 00:00 to 01:00

X_1 - weather condition from 01:00 to 02:00

.

.

.

X_{23} - weather condition from 23:00 - 00:00

The above scenario can be framed as a stochastic process. Here's how it relates to the concept of a stochastic process:

Time index: The time index is the hour of the day as 0, 1, 2, 3, ... 23. Each hour is a specific point in time.

Random variable: The weather conditions at each hour can be viewed as random variables. These random variables can be take different values such as sunny, rainy and cloudy. The weather conditions sunny,rainy and cloudy are called *states* (see Section for more information).

In many real life situations, observations are made over a period of time. Stochastic processes are used to model and analyze such time-dependent random phenomena, allowing you

to study the probabilistic behavior and make predictions about future states based on past observations. When dealing with stochastic processes, we can address various probabilistic questions, including but not limited to:

1. **Conditional probabilities:** For instance, given that the weather has been cloudy for the first five hours of the day, you can use the stochastic process to estimate the likelihood of it remaining cloudy or changing to a different condition in the next hour.

2. **Time to an event:** For example, you can estimate how long it will take for the weather to change from cloudy to sunny.

3. **Transition probabilities:** For instance, you can determine the likelihood of going from a rainy day to a sunny day or vice versa.

4. **Frequency of Events:** You can examine the frequency of specific events occurring within a given time frame.

These are just a few examples of the probabilistic questions that can be addressed using stochastic processes. The specific questions you can answer will depend on the nature of the process and the data available for analysis.

1.2 Definition of a stochastic process

Definition 1

A stochastic process is a collection of random variables $\{X_t, t \in T\}$ or $X(t), t \in T$ where T is an index set. That is for each $t \in T$, the random variable X_t (or $X(t)$) is a random variable.

1.3 Parameter space

In definition 1, the index set T is called the parameter space. It is usually interpreted as a time variable, telling us when the process is measured. The parameter space can be discrete or continuous.

1.3.1 Discrete-parameter process

When T is a countable set, the process is said to be a discrete-parameter process. A discrete-parameter stochastic process is defined as follows:

$$\{X_t : t \in T\}$$

Example: Number of Customers arriving each hour to a particular super market (Discrete Parameter Space)

In this scenario, you are interested in the number of customers arriving during each discrete time interval, typically on an hourly basis.

1.3.2 Continuous-parameter space

When T is an interval of the real line, the process is said to be a continuous-parameter process. A continuous-parameter stochastic process is defined as follows:

$$\{X(t) : t \in T\}$$

Example 1: Number of Customers Arriving from 8 AM to 10 PM (Continuous Parameter Space):

In this scenario, you are interested in the total number of customers arriving over a continuous time period, specifically from 8 AM to 10 PM.

1.4 State space

The set of possible values of an individual random variable X_t or $X(t)$ of a stochastic process is called the state space. The state space can be discrete or continuous.

1.5 Random Variable in Probability Theory vs Stochastic Theory

1.5.1 Probability theory

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. A measurable mapping $X : \Omega \to \mathbb{R}$ is called a random variable. The $X(\omega)$ for $\omega \in \Omega$ iscalled a realization of X.

Example:

1.5.2 Stochastic theory

Suppose that $(\Omega, \mathcal{F}, \mathbb{P})$ is a probability space, the function $X : T \times \Omega \to \mathbb{R}$.

- We will always assume that the cardinality of T is infinite, either countable or uncountable.

If $T = \mathbb{Z}^+$ then we called X a discrete time stochastic process.

If $I = [0, \infty)$, then X is said to be a continuous time stochastic processes.

1.6 Sample path (trajectory) of a stochastic process

The function $t \to X_t(\omega)$ $(t \to X(t)(\omega))$ is called a sample path of the stochastic process. For each

Example:

1.7 Types of Stochastic Processes

Depending on the parameter space and state space we can define four type of stochastic processes.

1. **Discrete-Parameter Discrete-State Space Stochastic Processes:**

- Parameter Space: Discrete

- State Space: Discrete

- Examples: Assessment of crop condition during routine field inspections in agriculture is categorized as: healthy, pest-infested, diseased, damaged. These field inspections are typically conducted at regular intervals, such as once a week

2. **Continuous-Parameter Discrete-State Space Stochastic Processes:**

- Parameter Space: Continuous

- State Space: Discrete

- Examples:

3. **Discrete-Parameter Continous-State Space Stochastic Processes:**

- Parameter Space: Discrete

- State Space: Continuous

- Examples:

4. **Continuous-Parameter Continuous-State Space Stochastic Processes:**

- Parameter Space: Continuous

- State Space: Continuous

- Examples:

1.8 Stochastic proceses vs Time series

Figure Figure **??** shows a weekly dengue cases in Sri Lanka from 2006 - Week 52 to 2023 - Week 8. The data are available in the denguedatahub package in R (**dengudatahub?**). The first few rows of the dataset is shown below.

year	week	start.date	end.date	district	cases
2006	52	12/23/2006	12/29/2006	Colombo	71
2007	1	12/30/2006	1/5/2007	Colombo	40
2007	2	1/6/2007	1/12/2007	Colombo	43
2007	3	1/13/2007	1/19/2007	Colombo	38
2007	4	1/20/2007	1/26/2007	Colombo	52
2007	5	1/27/2007	2/2/2007	Colombo	69

Let's define the set of random variables as follows:

X_0 - Cases of dengue during the fifty second week of 2006

X_1 - Cases of dengue during the first week of 2007

X_2 - Cases of dengue during the second week of 2007

.

.

.

Do you consider Figure **??** as a visual representation of the above **Stochastic Process**?

The answer is "No". Figure **??** is not a visual representation of the stochastic process. It is a realization of the above stochastic process. According to the data set in the fifty second week of 2006, there were 76 cases were reported. This is the observed value of the random variable X_0. Similarly, the observed value of the random variable X_1 is 40. Hence, Figure **??** represents the set of observed values of the the stochastic process. This is called a time series. "In other words, time series is a realization of a stochastic process. When we say a time series is a
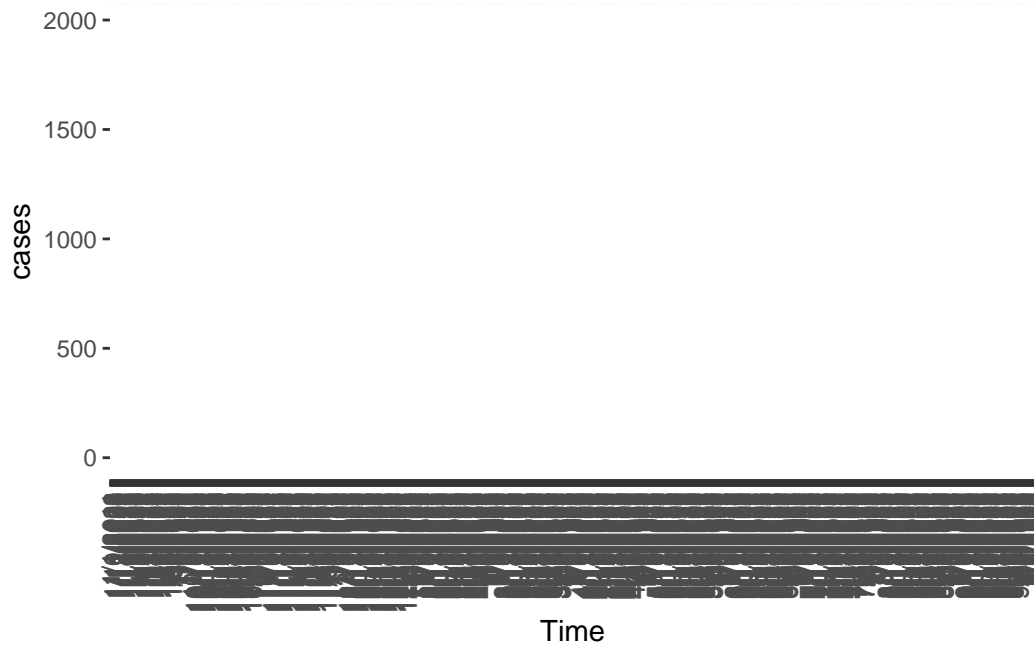
Figure 1.1: Weekly dengue cases in Sri Lanka from 2007 to 2023 August

realization, we mean that it represents a specific outcome or path or trajectory of a stochastic process. A realization is essentially a particular observed sequence of values that the process can take. Therefore, when we say a time series is a realization of a stochastic process, we are highlighting that the observed sequence of data points in a time series is **one possible outcome** of a random process that unfolds over time. The stochastic nature implies that, even though the underlying process has certain statistical properties, the specific values observed at any given point in time are not predetermined and can exhibit variability.

1.9 Stochastic process vs a deterministic process

Consider the following data generating process and its visual representation. Let's define the set of random variables as follows:

X_0 - value at time $t = 1$

X_1 - value at time $t = 2$

X_2 - value at time $t = 3$

.

.

9

.

Do you consider Figure **??** as a realization of a stochastic process?

```r
t <- 1:100
xt <- sin(2*pi*t)
df <- data.frame(xt=xt, t=t)
ggplot(data=df, aes(x=t, y=xt)) +
  geom_point() +
  geom_line()
```
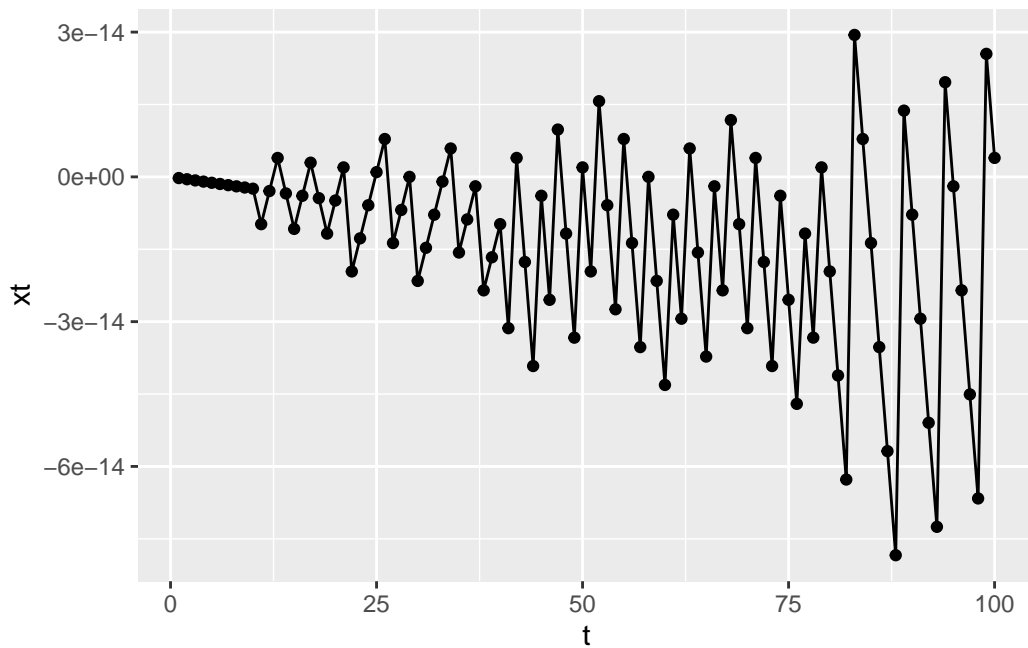

Figure 1.2: Visual representation of data

The answer is "No". The reason is when you observe the above R-code you can see the values are generated based on the function $\sin(2\pi t)$. Hence, there is no randomness associated with he random variables. This type of process is called a deterministic process.

1.10 Stationary condition of a stochastic process

1.11 Independent increments condition of a stochastic process

1.12 Stationary increments condition of a stochastic process

1.13 Stochastic processes vs probaility calculation in a single random variable

1.14 Applications of stochastic processes

2 Discrete Parameter Markov Chains

2.1 Introduction

A discrete parameter Markov chain process is a modeling approach used to represent systems that evolve over time in **discrete steps**, where the **future state depends only on the current state**. This is very useful in analyzing distinct states and transitions. Discrete Parameter Markov Chains is also known as "Discrete-time Markov Chains".

Definition

Let $\{X_n; n = 0, 1, 2, ...\}$ be a stochastic process that takes on a finite or countable number of possible values. If $X_n = i$, then the process is said to be in state i at time n.

The discrete-parameter, discrete state stochastic process $\{X_n; n = 0, 1, 2, ...\}$ is called a **discrete-parameter Markov chain** if for all states $i_0, i_1, ... i_{n-1}, i, j$ and all $n \geq 0$,

$$P(X_{n+1} = j | X_n = i, X_{n-1} = i_{n-1}, ..., X_1 = i_1, X_0 = i_0) = P(X_{n+1} = j | X_n = i).$$

This means, a stochastic process is a Markov chain if the probability of moving to the next state depends only on the current state and not on the sequence of events that preceded it.

2.2 One-step transition probabilities

We have a set of states, $S = \{i_0, i_1, i_2, ... i_{n-1}, i, j\}$. The process starts in one of these states and moves successively from one state to another. Each move is called a **step**.

$$P_{nij} = P(X_{n+1} = j | X_n = i, X_{n-1} = i_{n-1}, ..., X_1 = i_1, X_0 = i_0) = P(X_{n+1} = j | X_n = i)$$

If the chain is currently in state i, then it moves to state j at the next step with a probability denoted by p_{nij}, and this probability does not depend upon which states the chain was in before the current state.

The probabilities p_{nij} are called one-step transition probabilities.

2.3 Time Homogeneous Discrete-Parameter Markov Chain

If the conditional probability $P(X_{n+1} = j|X_n = i)$ does not depend on n, then the process is known as time homogeneous Markov chain process or stationary Markov chain process. Then we can write the conditional probability p_{nij} as $P_{i,j}$. Moreover, when there is no risk of confusion, we can write $P_{i,j}$ simply as P_{ij}.

2.3.1 Intuition behind time-homogenous Markov chain process

Example 1:

Suppose we observe the condition of crop's soil moisture on every morning. We record the conditions as dry, normal and wet. Let's denote the states as follows:

- state 0: dry

- state 1: normal

- state 2: wet

Let's assume that the soil moisture condition on a given day depends only on the moisture condition of the previous day. Furthermore, in this case conditional probability $P(X_{n+1} = j|X_n = i)$ actually depend on n. The probability of moving wet to wet $P(X_{n+1} = 2|X_n = 2)$ is not same for the whole year. This probability is small during the dry season and very high during the rainy season. If you are living in a country with four seasons, then this probability will vary according to the seasons: winter, summer, spring, and autumn. Hence, this is a non-stationary discrete-parameter discrete-state space Markov chain process. We can use a stationary Markov chain only for a short period of time.

In this book we only consider time-homogeneous Markov chain processes.

Example 2

Let's consider modeling the mood of a person as a time-homogeneous Markov chain with three states:

- state 0: happy

- state 1: neutral

- state 2: sad

Here is the time homogeneous transition probability matrix

$$P = \begin{bmatrix} p_{00} & p_{01} & p_{02} & \cdots \\ p_{10} & p_{11} & p_{12} & \cdots \\ . & . & . & \cdots \\ p_{i0} & p_{i1} & p_{i2} & \cdots \\ . & . & . & \cdots \end{bmatrix}$$

2.4 Exercise

1. Suppose there are two types of crops (Crop A and Crop B) planted in different fields, and farmers perform certain operations that may affect the composition of crops in each field. After each agricultural season, a random selection of crops from one field is transferred to the other. Let X_t denote the number of Crop A plants in the first field after the tth season. What are the parameter space and state space for this agricultural scenario, and can $\{X_t\}$ be considered a Markov chain given certain conditions on the planting and transfer processes?

2.5 One-step transition probability matrix

Let P denote the matrix of one-step transition probabilities P_{ij}, so that

$$P = \begin{bmatrix} p_{00} & p_{01} & p_{02} & \cdots \\ p_{10} & p_{11} & p_{12} & \cdots \\ . & . & . & \cdots \\ p_{i0} & p_{i1} & p_{i2} & \cdots \\ . & . & . & \cdots \end{bmatrix}$$

Since probabilities are nonnegative and since the process must make a transition into some state, we have

$p_{ij} \geq 0$, for $i, j \geq 0$, $\sum_{j=0}^{\infty} p_{ij} = 1$, for $i = 0, 1, ...$

2.6 Higher (n-step) transition probabilities

P_{ij} - One step transition probabilities

P_{ij}^n - n - step transition probabilities

Probability that a process in state i will be in state j after n additional transitions. That is,

$$P_{ij}^n = P(X_{n+k} = j | X_k = i), \ n \geq 0, \ i, j \geq 0.$$

2.7 Chapman-Kolmogorov equations

$$P_{ij}^{n+m} = \sum_{k=0}^{\infty} P_{ik}^n P_{kj}^m \text{ for all n, m} \geq 0, \text{ all i, j,}$$

where, $P_{ik}^n P_{kj}^m$ represents the probability that starting in i the process will go to state j in $n + m$ with an intermediate stop in state k after n steps.

This can be used to compute n-step transition probabilities

2.8 Exercise

1. Show that $P_{ij}^{n+m} = \sum_{k=0}^{\infty} P_{ik}^n P_{kj}^m$ for all n, m ≥ 0, all i, j.

2.9 n - step transition matrix

The n-step transition matrix is

$$\mathbf{P}^{(n)} = \begin{bmatrix} P_{00}^{(n)} & P_{01}^{(n)} & P_{02}^{(n)} & \cdots \\ P_{10}^{(n)} & P_{11}^{(n)} & P_{12}^{(n)} & \cdots \\ . & . & . & \cdots \\ . & . & . & \cdots \\ . & . & . & \cdots \end{bmatrix}$$

The Chapman-Kolmogrov equations imply

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

In particular,

$$\mathbf{P}^{(2)} = \mathbf{P}^{(1)} \mathbf{P}^{(1)} = \mathbf{PP} = \mathbf{P}^2.$$

By induction,

$$\mathbf{P}^{(n)} = \mathbf{P}^{(n-1+1)} = \mathbf{P}^{n-1} \mathbf{P} = \mathbf{P}^n.$$

2.10 n - step transition matrix

Proposition

$$P^{(n)} = P^n = P \times P \times P \times ... \times P, \ n \geq 1.$$

That is, $P^{(n)}$ is equal to P multiplied by itself n times.

2.11 Types of States

Definition: If $P_{ij}^{(n)} > 0$ for some $n \geq 0$, state j is **accessible** from i.

Notation: $i \rightarrow j$.

Definition: If $i \rightarrow j$ and $j \rightarrow i$, then i and j **communicate**.

Notation: $i \leftrightarrow j$.

2.12 Theorem:

Communication is an equivalence relation:

   (i) $i \leftrightarrow i$ for all i (reflexive).

  (ii) $i \leftrightarrow j$ implies $j \leftrightarrow i$ (symmetric).

 (iii) $i \leftrightarrow j$ and $j \leftrightarrow k$ imply $i \leftrightarrow k$ (transitive).

2.13 In-class: Proof

   (i) $i \leftrightarrow i$ for all i (reflexive).

2.14 In-class: Proof

(ii) $i \leftrightarrow j$ implies $j \leftrightarrow i$ (symmetric).

2.15 In-class: Proof

(iii) $i \leftrightarrow j$ and $j \leftrightarrow k$ imply $i \leftrightarrow k$ (transitive).

2.16 In-class: Proof

2.17 Note:

- Two states that communicate are said to be in the same **class**.

- The concept of communication divides the state space up into a number of separate classes.

In-class: demonstration

2.18 Theorem

Definition: An equivalence class consists of all states that communicate with each other.

Remark: Easy to see that two equivalence classes are disjoint.

Example: The following P has equivalence classes $\{0, 1\}$ and $\{2, 3\}$

$$\mathbf{P} = \begin{bmatrix} 0.5 & 0.5 & 0 & 0 \\ 0.5 & 0.5 & 0 & 0 \\ 0 & 0 & 0.75 & 0.25 \\ 0 & 0 & 0.25 & 0.75 \end{bmatrix}$$

What about this?

$$\mathbf{P} = \begin{bmatrix} 0.5 & 0.5 & 0 & 0 \\ 0.5 & 0.3 & 0.2 & 0 \\ 0 & 0 & 0.75 & 0.25 \\ 0 & 0 & 0.25 & 0.75 \end{bmatrix}$$

2.19 Irreducible

Definition: A MC is irreducible if there is only one equivalence class (i.e., if all states communicate with each other).

What about these?

Example 1

$$\mathbf{P} = \begin{bmatrix} 0.5 & 0.5 & 0 & 0 \\ 0.5 & 0.5 & 0 & 0 \\ 0 & 0 & 0.75 & 0.25 \\ 0 & 0 & 0.25 & 0.75 \end{bmatrix}$$

Example 2

$$\mathbf{P} = \begin{bmatrix} 0.5 & 0.5 & 0 & 0 \\ 0.5 & 0.3 & 0.2 & 0 \\ 0 & 0 & 0.75 & 0.25 \\ 0 & 0 & 0.25 & 0.75 \end{bmatrix}$$

Example 3

$$\mathbf{P} = \begin{bmatrix} 0.5 & 0.5 \\ 0.25 & 0.75 \end{bmatrix}$$

Example 4

$$\mathbf{P} = \begin{bmatrix} 0.25 & 0 & 0.75 \\ 1 & 0 & 0 \\ 0 & 0.5 & 0.5 \end{bmatrix}$$

2.20 Identify the equivalence classes

Consider a Markov chain with a state space $S = \{0, 1, 2, 3, 4\}$ and having the following one-step transition probability matrix.

$$\mathbf{P} = \begin{bmatrix} 0.4 & 0.2 & 0 & 0.4 & 0 \\ 0.2 & 0.4 & 0.1 & 0.3 & 0 \\ 0.1 & 0.2 & 0.5 & 0.1 & 0.1 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

2.21 Problems [1]

Example 4.10

Example 4.11

Example 4.12

2.22 Theorem

The relation of communication partitions the state space into mutually exclusive and exhaustive classes. (The states in a given class communicate with each other. But states in different classes do not communicate with each other.)

2.23 Definition

Let f_i denote the probability that, starting in state i, the process will ever re-enters state i, i.e,

$$f_i = P(X_n = i \text{ for some } n \geq 1 | X_0 = i)$$

Example 1

Consider the Markov chain consisting of the states 0, 1, 2, 3 with the transition probability matrix,

$$\mathbf{P} = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

---

[1]Introduction to Probability Models, Sheldon M. Ross

Find f_0, f_1, f_2, f_3.

2.24 Recurrent and transient states

Let f_i be the probability that, starting in state i, the process will ever re-enter state i. State i is said to be recurrent if $f_i = 1$ and transient if $f_i < 1$.

Example 1

Consider the Markov chain consisting of the states 0,1,2 with the transition probability matrix

$$\mathbf{P} = \begin{bmatrix} 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & 0 \end{bmatrix}$$

Determine which states are transient and which are recurrent.

Example 2

Consider the Markov chain consisting of the states 0, 1, 2, 3 with the transition probability matrix

$$\mathbf{P} = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

Determine which states are transient and which are recurrent.

Example 3

Consider the Markov chain consisting of the states 0, 1, 2, 3, 4 with the transition probability matrix

$$\mathbf{P} = \begin{bmatrix} \frac{1}{2} & 0 & \frac{1}{2} & 0 & 0 \\ \frac{1}{4} & \frac{1}{2} & \frac{1}{4} & 0 & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & 0 & 0 & \frac{1}{2} & \frac{1}{2} \end{bmatrix}$$

Determine which states are transient and which are recurrent.

2.25 Theorem

if state i is recurrent then, starting in state i, the process will re-enter state i again and again and again—in fact, infinitely often.

2.26 Theorem

For any state i, let f_i denote the probability that, starting in state i, the process will ever re-enter state i. If state i is transient then, starting in state i, the number of time periods that the process will be in state i has a geometric distribution with finite mean $\frac{1}{1-f_i}$.

Proof: In-class

2.27 Theorem

State i is

$$\text{recurrent if } \sum_{n=1}^{\infty} P_{ii}^n = \infty,$$

$$\text{transient if } \sum_{n=1}^{\infty} P_{ii}^n < \infty,$$

Proof: In-class

2.28 Corollary 1

If state i is recurrent, and state i communicates with state j $(i \leftrightarrow j)$, then state j is recurrent.

Proof: In-class

2.29 Corollary 2

In a Markov Chain with a finite number of states not all of the states can be transient (There should be at least one recurrent state).

Proof: In-class

2.30 Corollary 3

If one state in an equivalent class is transient, then all other states in that class are also transient.

Proof: In-class

2.31 Corollary 4

Not all states in a finite Markov chain can be transient. This leads to the conclusion that **all states of a finite irreducible Markov chain are recurrent**.

3 Practical

3.1 Package

```
#pak::pak("thiyangt/stochastic")
library(stochastic)
```

3.1.1 n-step transition probability matrix

```
library(stochastic)
x <- c(0.2,0.8,0.4,0.6)
nstepmat(x, 2, 3)
```

```
$one_step
     [,1] [,2]
[1,]  0.2  0.8
[2,]  0.4  0.6

$n_step
       [,1]  [,2]
[1,] 0.328 0.672
[2,] 0.336 0.664
```

```
nstepmat(x, 2, 13)
```

```
$one_step
     [,1] [,2]
[1,]  0.2  0.8
[2,]  0.4  0.6

$n_step
        [,1]      [,2]
```

```
[1,]  0.3333333  0.6666667
[2,]  0.3333333  0.6666667
```

3.1.2 Compute stationary probabilities

```
mat <- matrix(c(0.5, 0.5, 0.7, 0.3), byrow=TRUE, ncol=2)
stationary_prob(onestep=mat)
```

```
[1]  0.5833333  0.4166667
```

3.1.3 Simulate a Markov Chain process

```
init <- c(0.1, 0.9)
mat <- matrix(c(0.5, 0.5, 0.7, 0.3), byrow=TRUE, ncol=2)
simmarkov(init, mat, 100, c("Rainy", "Sunny"))
```

```
  [1] "Sunny" "Rainy" "Rainy" "Sunny" "Sunny" "Rainy" "Sunny" "Rainy" "Sunny"
 [10] "Sunny" "Rainy" "Sunny" "Sunny" "Rainy" "Sunny" "Rainy" "Sunny" "Rainy"
 [19] "Sunny" "Sunny" "Rainy" "Rainy" "Sunny" "Rainy" "Sunny" "Rainy" "Sunny"
 [28] "Rainy" "Sunny" "Rainy" "Sunny" "Rainy" "Rainy" "Sunny" "Sunny" "Rainy"
 [37] "Sunny" "Rainy" "Sunny" "Sunny" "Rainy" "Sunny" "Rainy" "Rainy" "Sunny"
 [46] "Rainy" "Sunny" "Sunny" "Rainy" "Rainy" "Sunny" "Sunny" "Rainy" "Sunny"
 [55] "Rainy" "Sunny" "Rainy" "Sunny" "Rainy" "Rainy" "Rainy" "Rainy" "Rainy"
 [64] "Sunny" "Rainy" "Sunny" "Rainy" "Rainy" "Sunny" "Rainy" "Rainy" "Sunny"
 [73] "Rainy" "Sunny" "Sunny" "Rainy" "Sunny" "Sunny" "Rainy" "Rainy" "Rainy"
 [82] "Sunny" "Rainy" "Sunny" "Rainy" "Rainy" "Sunny" "Rainy" "Sunny" "Rainy"
 [91] "Rainy" "Sunny" "Rainy" "Sunny" "Rainy" "Sunny" "Rainy" "Sunny" "Rainy"
[100] "Rainy" "Rainy"
```

3.2 In-class illustration

We are given sequences of emotional states for five individuals (person1–person5). Each sequence represents how a person's mood changed over time.

```
person1 <- c("happy", "calm", "neutral", "sad", "angry", "stressed", "happy")
person2 <- c("happy", "neutral", "calm", "sad", "angry", "happy", "stressed")
person3 <- c("calm", "happy", "neutral", "sad", "angry", "stressed", "happy")
person4 <- c("happy", "calm", "sad", "neutral", "angry", "stressed", "happy")
person5 <- c("happy", "calm", "sad", "angry", "neutral", "happy", "stressed")
person6 <- c("happy", "calm", "neutral", "sad")
```

We want to use these sequences to:

```
Estimate a transition probability matrix

Predict the next emotional state for person6

Visualise the Markov chain
```

3.3 Question 1

An insurance company offers four levels of discounts to policyholders based on their claim history:

Level	Discount
1	0%
2	20%
3	30%
4	40%

The following are the rules for shifting between the four levels:

- A policyholder at Level 1 moves to Level 2 with probability 0.5, stays at Level 1 with probability 0.5.

- A policyholder at Level 2 moves to Level 1 with probability 0.2, moves to Level 3 with probability 0.3, and stays at Level 2 with probability 0.5.

- A policyholder at Level 3 moves to Level 2 with probability 0.3, moves to Level 4 with probability 0.4, and stays at Level 3 with probability 0.3.

- A policyholder at Level 4 moves to Level 3 with probability 0.2, and stays at Level 4 with probability 0.8.

Tasks

1. Represent the system as a **Markov chain** using a transition probability matrix in R.

2. Write an R function to **simulate the Markov chain** for 50 time periods starting from Level 1.

3. Compute the **stationary distribution** of the Markov chain.

4. Plot the **simulated discount levels over time**.

5. Compute the **limiting (long-run) probabilities** of being in each discount level.

3.4 Question 2

Generating and Simulating a Lung Cancer Metastasis Markov Chain

Read "A Stochastic Markov Chain Model to Describe Lung Cancer Growth and Metastasis" by Paul K. Newton et al. (2012). Based on the model presented in the paper:

1. Identify the key anatomical sites relevant for lung cancer metastasis and represent them as states in a discrete-time Markov chain.

2. Using the descriptions and data from the paper, construct a transition probability matrix that represents the likelihood of metastasis from one site to another in one time step.

3. Write an R function to simulate the evolution of a single patient's metastasis over n time steps, starting from the primary lung tumor.

4. Use your function to simulate multiple patients (e.g., 1000) over a fixed number of steps (e.g., 20) and summarize the proportion of patients in each site at the final step.

5. Discuss how your simulation results relate to the biological progression described in the paper.

4 Birth Process, Death Process and Birth-Death Process.

- A birth-death process is a continuous-time Markov chain used to describe the evolution of the system by counting the number of individuals in the system over time.

- In this dynamic system, each individual has the potential to either give birth to a new individual or undergo a death event.

- The rates of these birth and death events at any given time depends upon the current number of existing individuals in the system.

4.1 Applications of Birth-Death Process

- Genetics, epidemiology: Study of infectious diseases, birth-death processes can be employed to model the spread of infections within a population

- Ecology: model the birth and death rates of different species in a habitat

- Epidemiology: model number of infected individuals

4.2 Definition: Transition Probability

Let $N(t)$ be the number of individuals at arbitrary time t for $t \geq 0$ and $\{N(t) : t \geq 0\}$ be the sequence of random variables which define a birth-death Markov chain with a birth rate λ_i and death rate μ_i for state i. Assume that an arbitrary community with r individuals, thus $N(0) = r$ for some $r > 0$.

The stationary transition probability is defined as follows:

$$P_{ij}(t) = P\{N(s+t) = j | N(s) = i\} \text{ for all } s \text{and } t > 0$$

4.3 Instantaneous birth rate and instantaneous death rate

Consider a population of $N(t)$ individuals. Suppose in next time interval $(t, t+h)$ probability of population increase of 1 (called a birth) is $\lambda_i h + o(h)$ and probability of decrease of 1 (death) is $\mu_i h + o(h)$. Here λ_i is called the instantaneous birth rate and μ_i is called instantaneous death rate.

Further, we have

$$
q_{ij} = \begin{cases} 0, & \text{if } |i-j| \geq 1 \\ \lambda_i, & \text{if } j = i+1 \\ \mu_i, & \text{if } j = i-1 \end{cases}
$$

4.4 Postulates

Let

$$
P_{ij}(h) = P(N(t+h) = j | N(t) = i) \text{ for all } t \geq 0,
$$

Further, $P_{ij}(t)$ satisfy,

1. $P_{i,i+1}(h) = P(N(t+h) = i+1 | N(t) = i) = \lambda_i h + o(h), i \geq 0$

2. $P_{i,i-1}(h) = P(N(t+h) = i-1 | N(t) = i) = \mu_i h + o(h), i \geq 1$

3. $P_{i,i}(h) = P(N(t+h) = i | N(t) = i) = 1 - (\lambda_i + \mu_i)h + o(h), i \geq 0$

4. $P(N(t+h) = i+m | N(t) = i) = o(h), |m| > 1$

5. $\mu_0 = 0, \lambda_0 > 0, \mu_i, \lambda_i > 0, i = 1, 2, 3, , , ,$

4.5 Exercise

Show that the birth process, death process and birth-death process are Markov processes

4.6 Special cases

1. When all $\mu_i = 0$, called a pure birth process.

2. When all $\lambda_i = 0$ called a pure death process

3. When $\lambda_n = n\lambda$ and $\mu_n = n\lambda$ is a linear birth and death process.

4. When $\lambda_n = \lambda$ and $\mu_n = 0$ is called a Poisson process.

4.7 Exercise

Derive the distribution of length of stay (elapsed time between two consecutive occurrences) for a birth, death and birth-death process.

> Help 1

Important facts about the exponential distribution

Fact 1:...................................

Fact 2:...................................

Fact 3:...................................

Fact 4:...................................

> Help 2

Distribution of length of stay - Continuous parameter Markov chain processes

Suppose that a continuous-time time-homogeneous Markov chain $\{N(t) : t \geq 0\}$ enters state i at some time, (say at time $s \geq 0$)) and let T_i denote the amount of time that the process stays in state i. Then, for any $u > 0$

$$P[T_i \geq u] = P[N(t) = i; s < t < s + u | N(s) = i)]$$

for any $s \geq 0$.

4.8 Instanttaneous Probability, Transition Probability, Probability Mass Function

- In-class discussion

5 Applications

5.1 Computer Science

5.1.1 PageRank Algorithm

- *Eirinaki, M., Vazirgiannis, M., & Kapogiannis, D. (2005, November). Web path recommendations based on page ranking and markov models. In Proceedings of the 7th annual ACM international workshop on Web information and data management (pp. 2-9).*

  Extracted "Markov models have been widely used for modelling **users' navigational behaviour** in the Web graph, using the transitional probabilities between web pages, as recorded in the web logs. **The recorded users' navigation is used to extract popular web paths and predict current users' next steps**. Such purely usage-based probabilistic models, however, present certain shortcomings. Since the prediction of users' navigational behaviour is based solely on the usage data, structural properties of the Web graph are ignored."

5.1.2 Machine Learning

- Hidden Markov Models (HMMs) are widely used in speech recognition and natural language processing.

5.1.3 Monte Carlo Markov Chain (MCMC) Applications

Randomized Algorithms – Monte Carlo Markov Chain (MCMC) methods are used in simulations and Bayesian inference.

Markov Chain Monte Carlo (MCMC) is a class of algorithms used to sample from complex probability distributions when direct sampling is difficult. It combines Markov chains with Monte Carlo methods to generate samples that approximate a target distribution.

How MCMC Works

Markov Chain Property

The process moves from one state to another based on a transition rule, where the next state depends only on the current state (not past states).

Monte Carlo Sampling

Random samples are generated to approximate integrals, expectations, or probabilities in Bayesian inference and statistical modeling.

Convergence to Target Distribution

Over many iterations, the chain reaches a steady-state (stationary distribution) that represents the desired probability distribution.

Popular MCMC Algorithms

1. Metropolis-Hastings Algorithm
2. Gibbs Sampling
3. Hamiltonian Monte Carlo (HMC)

Applications of MCMC

Bayesian Statistics → Posterior inference in Bayesian models.

Machine Learning → Parameter estimation in probabilistic models.

Physics & Chemistry → Simulating molecular dynamics and statistical mechanics.

Finance → Risk assessment and portfolio optimization.

Genetics & Epidemiology → Evolutionary models and disease spread prediction

5.2 Biology

5.2.1 Population Genetics – Evolutionary changes in gene frequencies are modeled with Markov chains.

5.2.2 Ecological Modelling

Balzter, H. (2000). Markov chain models for vegetation dynamics. Ecological modelling, 126(2-3), 139-154.

Extracted "The aim of this paper is to assess the applicability of Markov chain models to **predict vegetation changes** using several different data sets, both from the scientific literature and own observations. It is anticipated that the results and the insights in model behaviour will also be useful for large-scale landscape ecosystem models by using Markov chains as submodels."

5.2.3 Epidemiology – Predicting disease spread using Markov models in SIR models.

Teaching in the Time of the COVID-19 Pandemic: dynamic Markov models in epidemiology

- Susceptible Infected Recovered (SIR)
- Susceptible Exposed Infected Recovered (SEIR)

5.2.3.1 Dynamic Markov Models (DMMs)

DMMs refer to Markov models where transition probabilities change over time, making them more flexible than standard homogeneous Markov chains. These models are useful in scenarios where the system evolves dynamically, such as in financial markets, weather forecasting, and biological processes.

Another application of DMM

Guo, X., Liu, Y., Tan, K., Mao, W., Jin, M., & Lu, H. (2021). Dynamic Markov Model: Password Guessing Using Probability Adjustment Method. Applied Sciences, 11(10), 4607. https://doi.org/10.3390/app11104607

Extracted

"Password attack technology is mainly used for recovering passwords…"

5.2.3.2 Susceptible-Infected-Recovered (SIR) models

These model assumes people can be in 3 states:

1 Susceptible: people have no immunity to the disease. These people can become infected by coming into contact with an infected person.

2 Infected: people have the disease and they can spread it.

3 Recovered people survive the disease and have immunity.

- These are states as we saw before but in these models they are referred as **compartments**. For this reason models like these are called compartmental disease models. However, there is not much agreement on names, but you can think of these models as dynamic Markov models that come in two flavors, deterministic and stochastic.

- They are "dynamic" because people move from one compartment (state) to the other at different rates over time. In other words, the transition probabilities are not fixed, they change over time.

- The transition probabilities are themselves a function of other parameters in the model (which we will cover soon).

The SIR model description is extracted from here.

5.3 Economics

5.3.1 Consumer Behavior Modeling – Predicting how consumers switch between different brands.

- *Sharma, V., & Sonwalkar, J. (2016). Predicting the Switching Intention of Cell-Phone Brands: Application of Markov Chain Models. Journal of Management Research and Analysis, 3(2), 95-100.*

consumer research to identify the perceived value and performing a detailed cost analysis[13][25][15][26]. Studies also found that overall dissatisfaction considered to be the major reasons for switching behaviour[27][9][28] in any industry.

Rationally, it is very difficult for a manager to analyse such erratic consumer behaviour, as it varies industry to industry and it depends upon many different variable, which is almost impossible to control by any organization. Furthermore, determining and correcting for the cause of this behaviour is essential. This article presents evidence gleaned from the laboratory that such behaviour is symptomatic of the lack of salience, which can be corrected with a simple instructional variation.

Objective
1. Mobile Brand switching behaviour of the youths

was Indore city and data was collected through convenient random sampling procedure. Around 500 questionnaires were circulated among the youth, out of which 320 were finalized for the final analysis. The questionnaire had only two questions, the first was which brand of mobile hand set they are presently using and which brand of mobile they may buy in near future or will stick to the same mobile brand. The major brands which were considered for the study were Apple, Blackberry, Nokia, Samsung, Sony and any other.

Objective 1
Brand switching behaviour of the youth for mobile they are using?
In order to find weather there is any brand switching or not in the market a chi-square test was applied.

Table 1: Brand of first Handset * Which Brand of Mobile Handset you will buy in future

Which Brand of Mobile Handset you are using		Which Brand of Mobile Handset you will buy in future												
		Nokia/Microsoft	Samsung	Apple	ZTE	LG	Huawei	TCL	Sony	Lenovo	Micromax	Xiomi	Others	Total
	Nokia/Microsoft	10	0	2	0	0	0	0	1	0	7	2	0	22
	Samsung	8	32	10	0	0	0	4	2	0	4	1	2	63
	Apple	4	2	44	0	0	2	2	0	0	0	0	0	54
	ZTE	0	8	2	12	0	0	2	0	0	3	1	2	30
	LG	4	0	12	0	3	0	4	0	0	1	1	2	27
	Huawei	0	0	6	4	0	9	2	0	0	2	2	2	27
	TCL	1	2	2	0	0	3	2	0	0	2	1	2	15
	Sony	2	0	0	0	0	0	0	8	0	2	0	0	12
	Lenovo	0	0	2	2	0	0	0	0	6	4	0	1	15
	Micromax	3	1	2	0	0	0	0	0	1	13	0	0	20
	Xiomi	2	0	2	0	0	0	0	0	1	5	8	2	20
	Others	3	0	0	0	0	0	0	0	0	0	0	12	15
	Total	37	45	84	18	3	12	16	13	8	43	16	25	320

Source: As per the data collected by the researcher

33

The above image is extracted from the paper "Sharma, V., & Sonwalkar, J. (2016). Predicting the Switching Intention of Cell-Phone Brands: Application of Markov Chain Models. Journal of Management Research and Analysis, 3(2), 95-100."

5.3.2 Macroeconomic Modeling – Analyzing business cycles and economic growth using Markov processes.

5.3.3 Income Distribution – Modeling income mobility across different economic classes.

5.4 Finance

5.4.1 Stock Price Modeling – The Markov property is fundamental in models like the Black-Scholes model.

5.4.2 Credit Ratings – Credit rating transitions of companies are often modeled using Markov chains.

5.4.3 Risk Management – Estimating probabilities of financial crises based on economic indicators.

5.5 Insurance

5.5.1 Actuarial Science – Predicting claim occurrences and policyholder behavior.

5.5.2 Risk Analysis – Assessing life insurance risks with Markovian mortality models.

5.5.3 Policyholder Behavior – Modeling lapse rates and claims under different conditions.

6 Summary

In summary, this book has no content whatsoever.