Synthesis Lectures on
Mathematics & Statistics

SYNTHESIS
COLLECTION OF TECHNOLOGY

Liliana Blanco-Castañeda
Viswanathan Arunachalam

Applied Stochastic Modeling

Springer

Synthesis Lectures on Mathematics & Statistics

This series includes titles in applied mathematics and statistics for cross-disciplinary STEM professionals, educators, researchers, and students. The series focuses on new and traditional techniques to develop mathematical knowledge and skills, an understanding of core mathematical reasoning, and the ability to utilize data in specific applications.

Liliana Blanco-Castañeda ·
Viswanathan Arunachalam

Applied Stochastic Modeling

Springer

Liliana Blanco-Castañeda
Universidad Nacional de Colombia
Bogotá, Colombia

Viswanathan Arunachalam
Universidad Nacional de Colombia
Bogotá, Colombia

ISSN 1938-1743 ISSN 1938-1751 (electronic)
Synthesis Lectures on Mathematics & Statistics
ISBN 978-3-031-31281-6 ISBN 978-3-031-31282-3 (eBook)
https://doi.org/10.1007/978-3-031-31282-3

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Preface

This book is designed as a reference text for students and researchers who need to consult stochastic models in their professional work and are unfamiliar with the mathematical and statistical theory required to understand these methods. The most relevant concepts and results for the development of the examples are presented, omitting rigorous mathematical proofs and giving only the guidelines of those fundamental in constructing the models.

The book is divided into five chapters. Chapter 1 presents a compilation of the discrete-time Markov chain and the most relevant concepts and theorems. Chapter 2 deals with the Poisson process and some important properties. Chapter 3 is devoted to studying the continuous-time Markov chain and its applications, with particular emphasis on the birth and death process. Chapter 4 deals with Branching processes, particularly the Galton-Watson process with two types of individuals and is presented as an example to describe the evolution of the SARS-CoV2 virus in Bogota. Chapter 5 presents the theory corresponding to a hidden Markov model developed for the behavior of the horizontal displacements of the behaviors of two animals from their observed trajectories in order to identify hidden behavioral states and determine the preferences of habitat.

We thank our students for their assistance in typesetting and preparing programming codes, which has served as the platform for this project. We thank our Editors, Ms. Susanne Steitz-Filler and Ms. Melanie Rowen, Springer Nature, for their advice and technical support.

Bogotá, Colombia
February 2023

Liliana Blanco-Castañeda
Viswanathan Arunachalam

Contents

1 Discrete-Time Markov Chain

Markov chains are named after the Russian mathematician Andrei Andreyevich Markov (1856–1922) who introduced them in his work "Extension of the law of large numbers to dependent quantities", published in 1906, in which he developed the concept of the law of large numbers and the central limit theorem for sequences of dependent random variables [1]. As a disciple of the Russian mathematician Patnufy Chebyschev (1821–1894), he made great contributions to probability theory, number theory, and analysis. He worked as a professor at the University of Saint Petersburg since 1886, from where he retired in 1905, although he continued teaching until the end of his life.

Markov developed his theory of chains from a completely theoretical point of view, he also applied these ideas to chains of two states, vowels, and consonants, in some literary texts of the Russian poet Aleksandr Pushkin (1799–1837). Markov analyzed the sequences of vowels and consonants in Pushkin's verse work "Eugene Onegin", concluding that the letters are not arranged independently in the poem but that the placement of each letter depends on the previous letter.

Markov lived through a period of great political activity in Russia and became actively involved. In the year 1902, the Russian novelist, Maxim Gorky was elected to the Russian Academy of Sciences in 1902, but the direct order of the Tsar canceled his election. Markov protested and refused the honors he was awarded the following year. Later, when the interior ministry ordered university professors to report any anti-government activity by their students, he objected, claiming that he was a professor of probability and not a policeman [2]. Currently, Markov chains are used to find the author of a text [3] and in web search systems such as Google [4].

L. Blanco-Castañeda and V. Arunachalam, *Applied Stochastic Modeling*, Synthesis Lectures on Mathematics & Statistics, https://doi.org/10.1007/978-3-031-31282-3_1

1.1 Introduction to Stochastic Processes

Definition 1.1 A *stochastic process* is a family or a collection of random variables $X = \{X_t, t \in T\}$ defined on a common probability space $(\Omega, \Im, P)$ with taking values in a measurable space $(S, \mathcal{S})$, called the *state space*. The set of parameters T is called the *parameter space* of the stochastic process, which is usually a subset of $\mathbb{R}$.

The mapping defined for each fixed $\omega \in \Omega$, the function $t \to X_t(\omega), t \in \mathbb{R}$, is called *sample path* or a *realization* of the stochastic process X. The process path associated with ω provides a mathematical model for a random experiment whose outcome can be observed continuously in time.

The set of possible values of the indexing parameter which can be either discrete or continuous. For our convenience, when the indexing parameter is discrete, the family is represented by $\{X_n, n = 0, 1, 2 \dots\}$. In case of continuous time both $\{X_t, t \in T\}$ and $\{X(t), t \in T\}$ are used. If the state space and the parameter space of a stochastic process are discrete, then the process is called stochastic sequence, and often referred as a chain.

Stochastic processes can be classified, in general, into the following four types of processes:

1. Discrete time, discrete state space (DTDS).
2. Discrete time, continuous state space (DTCS).
3. Continuous time, discrete state space (CTDS).
4. Continuous time, continuous state space (CTCS).

Definition 1.2 Let $\{X_t; t \in T\}$ be a stochastic process and $\{t_1, t_2, \dots, t_n\} \subset T$ where $t_1 < t_2 < \dots < t_n$. The function

$$F_{t_1 \dots t_n}(x_1, \dots, x_n) := P(X_{t_1} \le x_1, \dots, X_{t_n} \le x_n)$$

is called *the finite-dimensional distribution of the process*.

Definition 1.3 If, for all $t_0, t_1, t_2, \dots, t_n$ such that $t_0 < t_1 < t_2 < \dots < t_n$, the random variables $X_{t_0}, X_{t_1} - X_{t_0}, X_{t_2} - X_{t_1}, \dots, X_{t_n} - X_{t_{n-1}}$ are independent, then the process $\{X_t; t \in T\}$ is said to be a process with independent increments.

Definition 1.4 A stochastic process $\{X_t; t \in T\}$ is said to have *stationary increments* if $X_{t_2+s} - X_{t_1+s}$ has the same distribution as $X_{t_2} - X_{t_1}$ for all choices of $t_1, t_2 \in T$ and $s > 0$.

Definition 1.5 If for all $t_1, t_2, \dots, t_n$ the joint distributions of the vector random variables

$$(X(t_1), X(t_2), \dots, X(t_n)) \text{ and } (X(t_1 + h), X(t_2 + h) \dots, X(t_n + h))$$

are the same for all $h > 0$, then the stochastic process $\{X_t; t \in T\}$ is said to be a stationary stochastic process of order n (or simply a stationary process). The stochastic process $\{X_t; t \in T\}$ is said to be a strong stationary stochastic process or strictly stationary process if the above property is satisfied for all n.

Definition 1.6 A stochastic process $\{X_t; t \in T\}$ is called a second-order process if $E\left(X_t^2\right) < \infty$ for all $t \in T$.

Example 1.1 Let Z_1 and Z_2 be independent normally distributed random variables, each having mean 0 and variance σ^2. Let $\lambda \in \mathbb{R}$ and

$$X_t = Z_1 \cos(\lambda t) + Z_2 \sin(\lambda t), \quad t \in \mathbb{R}.$$

$\{X_t; t \in T\}$ is a second-order stationary process. ▲

Definition 1.7 A second-order stochastic process $\{X_t; t \in T\}$ is called *covariance stationary* or *weakly stationary* if its mean function $m(t) = E[X_t]$ is independent of t and its covariance function $Cov(X_s, X_t)$ depends only on the difference $| t - s |$ for all $s, t \in T$. That is:

$$Cov(X_s, X_t) = f(| t - s |) .$$

Definition 1.8 A stochastic process that is not stationary in any sense is called an evolutionary stochastic process.

Definition 1.9 A stochastic process $\{X_t; t \in T\}$ is a *Gaussian process* if the random vectors $(X(t_1), X(t_2), \ldots, X(t_n))$ have a joint Normal distribution for all $(t_1, t_2, \ldots, t_n)$ and $t_1 < t_2 < \cdots < t_n$.

Definition 1.10 Let $\{X_t; t \geq 0\}$ be a stochastic process defined over a probability space $(\Omega, \Im, P)$ and with state space $(\mathbb{R}, \mathcal{B})$. We say that the stochastic process $\{X_t; t \geq 0\}$ is called a Markov process if for any $0 \leq t_1 < t_2 < \cdots < t_n$ and for any states $B, B_1, B_2, \ldots, B_{n-1} \in \mathcal{B}$:

$$P\left(X_{t_n} \in B \mid X_{t_1} \in B_1, \ldots, X_{t_{n-1}} \in B_{n-1}\right) = P\left(X_{t_n} \in B \mid X_{t_{n-1}} \in B_{n-1}\right) . \tag{1.1}$$

The above condition (1.1) is called the Markov property, and has the following implications: Any stochastic process with independent increments is a Markov process. Also, the Markov process is such that, given the value of X_s, for $t > s$, the distribution of X_t does not depend on the values of X_u, for $u < s$.

1.2 Discrete-Time Markov Chain

The Markov chain is defined as a sequence of random variables taking a finite or countable set of values and characterized by the Markov property. This section discusses the most important properties of the discrete-time Markov chain (for more details see [5, 6]).

Definition 1.11 The stochastic process $\{X_n; n \in \mathbb{N}\}$ with $n = 0, 1, \ldots$ is called a *discrete-time Markov chain* if for all for all $t_0 < t_1 < \cdots < t_{n+1}$ with $t_i \in T$ and $i, j, i_0, i_1, \ldots, i_{n-2} \in$ S We have

$$P\left(X_n = j \mid X_{n-1} = i, X_{n-2} = i_{n-2}, \ldots, X_0 = i_0\right) = P\left(X_n = j \mid X_{n-1} = i\right) \tag{1.2}$$

with

$$P(X_0 = i_0, \ldots, X_{n-1} = i) > 0.$$

Here the future state $X_n = j$ of the Markov chain depends only on the present state $X_{n-1} = i$, but not on the past "$X_{n-2}, X_{n-3}, \ldots, X_0$".

Let $\{X_n; n \in \mathbb{N}\}$ be a discrete-time Markov chain. If $X_0 = i_0$, then the chain is said to have started in the state i_0. If $X_n = i_n$ then the chain is said to be at time n in state i_n. The sequence of states $i_0, i_1, \ldots, i_n$ is said to be the complete history of the chain up to the time n, if $X_0 = i_0,\ X_1 = i_1, \ldots, X_n = i_n$.

Theorem 1.1 *The stochastic process $\{X_n; n \in \mathbb{N}\}$ with set of states S is a Markov chain, if and only if, for any finite sequence of natural numbers $0 \leq n_0 < n_1 < \cdots < n_k$ and for any choice $i_{n_0}, i_{n_1}, \ldots, i_{n_k} \in S$ it is satisfied that:*

$$P\left(X_{n_k+m} = j \mid X_{n_k} = i_{n_k}, \ldots, X_{n_0} = i_{n_0}\right) = P\left(X_{n_k+m} = j \mid X_{n_k} = i_{n_k}\right) \tag{1.3}$$

for any integer $m > 1$.

Definition 1.12 Let $\{X_n; n \in \mathbb{N}\}$ be a Markov chain with discrete-time parameter. The probabilities

$$p_{ij} := P\left(X_{n+1} = j \mid X_n = i\right) \tag{1.4}$$

with $i, j \in S$ are called *transition probabilities*. The matrix form of the transition probability is written as

$$\mathbf{P} = (p_{ij}) = \begin{pmatrix} p_{00} & p_{01} & p_{02} & \cdots \\ p_{10} & p_{11} & p_{12} & \cdots \\ p_{20} & p_{21} & p_{22} & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

is called the *transition probabilities matrix* or *stochastic matrix*, and satisfies the following:

$$\begin{aligned} p_{ij} &\geq 0 \text{ for all } i, j \in S \\ \sum_j p_{ij} &= 1 \text{ for all } i \in S\,. \end{aligned}$$

Remark 1.1 • A Markov chain $\{X_n; n \geq 0\}$ is called *homogeneous* if the transition probabilities do not depend on time-step n. That, is for $n \in \mathbb{N}$

$$p_{ij} := P\left(X_1 = j \mid X_0 = i\right) = P\left(X_{n+1} = j \mid X_n = i\right).$$

- The transition probabilities with the initial distribution $\pi_i^{(0)} := P(X_0 = i)$ completely determines the Markov chain. That is, if $\{X_n, n = 0, 1, 2, \dots\}$ is a Markov chain, then for all n and $i_0, \dots, i_n$ the set of states the satisfies the following:

$$P\left(X_0 = i_0, \dots, X_n = i_n\right) = \pi_i^{(0)} P\left(X_1 = i_1 \mid X_0 = i_0\right) \dots P\left(X_n = i_n \mid X_{n-1} = i_{n-1}\right).$$

Example 1.2 Suppose a random experiment is performed where there are only two possible outcomes success or failure, with a probability of success $0 < p < 1$ and probability of failure $q := 1 - p$. Let X_n be the random variable denoting the number of successes in n repetitions of the experiment. The random variable X_n has a binomial distribution of parameters n and p and the sequence $\{X_n; n \geq 1\}$ is a Markov chain with state set $S = \{0, 1, 2, \dots\}$ and transition matrix

$$\mathbf{P} = \left(p_{i,j}\right)_{i,j \in S}$$

with

$$p_{ij} = \begin{cases} p \text{ if } j = i + 1 \\ q \text{ if } \quad j = i \\ 0 \quad \text{otherwise} \end{cases}$$

Example 1.3 (*Random walk*) Let $(Y_n)_{n \geq 1}$ be a sequence of independent and equally distributed variables with values in $\mathbb{Z}$. The process $\{X_n; n \geq 0\}$ defined by:

$$\begin{aligned} X_0 &:= 0 \\ X_n &:= \sum_{k=1}^{n} Y_k \end{aligned}$$

is a Markov chain with state set $\mathbb{Z}$ and matrix of transition $\mathbf{P} = \left(p_{i,j}\right)_{i,j \in \mathbb{Z}}$ where $p_{i,j} = P\left(Y_1 = j - i\right)$.

Example 1.4 Suppose we have two players A and B at the beginning of the game, player A has a capital of $x \in \mathbb{Z}^+$ dollar and player B a capital of $y \in \mathbb{Z}^+$ dollar. Say $a := x + y$. In each round of the game, either A wins B a dollar with probability p or B wins A a dollar with probability q being $p + q = 1$. The game continues until one of the two players loses his capital, that is, until $X_n = 0$ or $X_n = a$.

Let $X_n :=$ "capital of A after the nth game round." The sequence $(X_n)_{n\in\mathbb{N}}$ is a Markov chain with set of states $S = \{0, 1, 2, \ldots, a\}$ and transition matrix is given by

$$\mathbf{P} = \begin{pmatrix} 1 & 0 & 0 & 0 & \cdots & 0 \\ q & 0 & p & 0 & \cdots & 0 \\ 0 & q & 0 & p & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ldots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \ldots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 1 \end{pmatrix}$$

Definition 1.13 Let $\{X_n; n \in \mathbb{N}\}$ be a Markov chain. The transition probability in m steps, $p_{ij}^{(m)}$, is the probability that from the state i, the state reached at state j at the mth step and defined as

$$p_{ij}^{(m)} = P\left(X_m = j \mid X_0 = i\right). \tag{1.5}$$

The $p_{ij}^{(m)}$ is stationary, if and only if, for all $n \in \mathbb{N}$.

$$p_{ij}^{(m)} = P\left(X_{n+m} = j \mid X_n = i\right) = P\left(X_m = j \mid X_0 = i\right) \tag{1.6}$$

A Markov chain whose transition probabilities in m steps are all stationary is called a homogeneous Markov chain. The transition matrix for $m-$ transition probabilities is written as

$$\mathbf{p}^m = \left(p_{ij}^m\right)_{i,j\in S} \tag{1.7}$$

Homogeneous Markov chains can be represented by a network in which the vertices indicate the states of the chain, and the arcs indicate the transitions between one state and another. For example, if $\{X_n; n \in \mathbb{N}\}$ is a Markov chain with set of states $S = \{0, 1, 2, 3\}$ with transition probability matrix

$$\mathbf{P} = \begin{pmatrix} \frac{1}{5} & \frac{1}{5} & 0 & \frac{3}{5} \\ 0 & \frac{1}{3} & \frac{2}{3} & 0 \\ \frac{1}{2} & 0 & 0 & \frac{1}{2} \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \end{pmatrix}.$$

The graphical representation of the state transition is shown in Fig. 1.1.

The following Chapman-Kolmogorov equation gives a method of computing $n-$step transition probabilities.

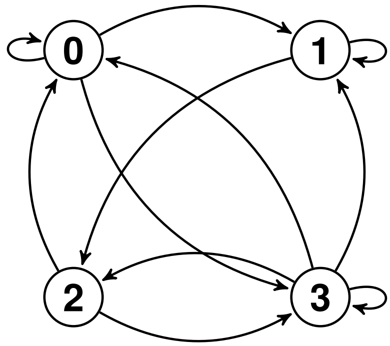

Fig. 1.1 State transition diagram

Proposition 1.1 *If $\{X_n; n \in \mathbb{N}\}$ is a homogeneous Markov chain and if $k < m < n$ then for states $h, i, j \in S$, we have*

$$p_{hj}^{n} = \sum_{i \in S} p_{ij}^{n-m} p_{hi}^{m}. \tag{1.8}$$

Remark 1.2 The above proposition which states that the transition matrix in m steps is the mth power of the transition matrix. That is,

$$\mathbf{P}^{(n)} = \mathbf{P}^{n} \tag{1.9}$$

Example 1.5 A Markov chain $\{X_n; n \geq 1\}$ with set of states $S = \{0, 1\}$ and transition matrix

$$\mathbf{P} = \begin{pmatrix} 1-a & a \\ b & 1-b \end{pmatrix}$$

where a and b are real numbers with $0 < a < 1$ and $0 < b < 1$.

The eigenvalues of the matrix $\mathbf{P}$ are $\lambda_1 = 1$ and $\lambda_2 = 1 - a - b$ and the corresponding eigenvectors are

$$\upsilon_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \quad \text{and} \quad \upsilon_2 = \begin{pmatrix} -a \\ b \end{pmatrix}$$

Then

$$\mathbf{P} = ADA^{-1}$$

with

$$A = \begin{pmatrix} 1 & -a \\ 1 & b \end{pmatrix}$$

$$D = \begin{pmatrix} 1 & 0 \\ 0 & 1-a-b \end{pmatrix}$$

Since